notes

USING DATA-CENTRIC METHODS TO TEACH INTRODUCTORY STATISTICS

Editors:

Donna LaLonde
American Statistical Association

Deirdre Longacher Smeltzer
Mathematical Association of America

Associate Editors:

Michael Brilleslyper
Florida Polytechnic University

Jenna Carpenter
Campbell University

Sarah Holsted
Broad-based Knowledge, LLC

Kathryn Kozak
Coconino Community College

Ambika Silva
College of the Canyons

MAA

NOTES / VOLUME 99

Using Data-Centric Methods to Teach Introductory Statistics

StatPREP is an initiative of the

Mathematical Association of America (MAA),

the American Statistical Association (ASA),

and the
American Mathemattical Association of Two-Year Colleges (AMATYC).

This work is supported by the National Science Foundation (grant DUE-1626337). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Electronic ISBN 978-1-61444-335-3

Published in the United States of America

Using Data-Centric Methods to Teach Introductory Statistics

Editors

Donna LaLonde
American Statistical Association

Deirdre Longacher Smeltzer
Mathematical Association of America

Associate Editors

Michael Brilleslyper
Florida Polytechnic University

Jenna Carpenter
Campbell University

Sarah Holsted
Broad-based Knowledge, LLC

Kathryn Kozak
Coconino Community College

Ambika Silva
College of the Canyons

Published and Distributed by

The Mathematical Association of America

The MAA Notes Series, started in 1982, addresses a broad range of topics and themes of interest to all who are involved with undergraduate mathematics. The volumes in this series are readable, informative, and useful, and help the mathematical community keep up with developments of importance to mathematics.

MAA Notes

14. Mathematical Writing, by *Donald E. Knuth, Tracy Larrabee, and Paul M. Roberts.*
16. Using Writing to Teach Mathematics, *Andrew Sterrett,* Editor.
17. Priming the Calculus Pump: Innovations and Resources, Committee on Calculus Reform and the First Two Years, a subcommittee of the Committee on the Undergraduate Program in Mathematics, *Thomas W. Tucker,* Editor.
18. Models for Undergraduate Research in Mathematics, *Lester Senechal,* Editor.
19. Visualization in Teaching and Learning Mathematics, Committee on Computers in Mathematics Education, *Steve Cunningham and Walter S. Zimmermann,* Editors.
20. The Laboratory Approach to Teaching Calculus, *L. Carl Leinbach et al.,* Editors.
21. Perspectives on Contemporary Statistics, *David C. Hoaglin and David S. Moore,* Editors.
22. Heeding the Call for Change: Suggestions for Curricular Action, *Lynn A. Steen,* Editor.
24. Symbolic Computation in Undergraduate Mathematics Education, *Zaven A. Karian,* Editor.
25. The Concept of Function: Aspects of Epistemology and Pedagogy, *Guershon Harel and Ed Dubinsky,* Editors.
26. Statistics for the Twenty-First Century, *Florence and Sheldon Gordon,* Editors.
27. Resources for Calculus Collection, Volume 1: Learning by Discovery: A Lab Manual for Calculus, *Anita E. Solow,* Editor.
28. Resources for Calculus Collection, Volume 2: Calculus Problems for a New Century, *Robert Fraga,* Editor.
29. Resources for Calculus Collection, Volume 3: Applications of Calculus, *Philip Straffin,* Editor.
30. Resources for Calculus Collection, Volume 4: Problems for Student Investigation, *Michael B. Jackson and John R. Ramsay,* Editors.
31. Resources for Calculus Collection, Volume 5: Readings for Calculus, *Underwood Dudley,* Editor.
32. Essays in Humanistic Mathematics, *Alvin White,* Editor.
33. Research Issues in Undergraduate Mathematics Learning: Preliminary Analyses and Results, *James J. Kaput and Ed Dubinsky,* Editors.
34. In Eves' Circles, *Joby Milo Anthony,* Editor.
35. You're the Professor, What Next? Ideas and Resources for Preparing College Teachers, The Committee on Preparation for College Teaching, *Bettye Anne Case,* Editor.
36. Preparing for a New Calculus: Conference Proceedings, *Anita E. Solow,* Editor.
37. A Practical Guide to Cooperative Learning in Collegiate Mathematics, *Nancy L. Hagelgans, Barbara E. Reynolds, SDS, Keith Schwingendorf, Draga Vidakovic, Ed Dubinsky, Mazen Shahin, G. Joseph Wimbish, Jr.*
38. Models That Work: Case Studies in Effective Undergraduate Mathematics Programs, *Alan C. Tucker,* Editor.
39. Calculus: The Dynamics of Change, CUPM Subcommittee on Calculus Reform and the First Two Years, *A. Wayne Roberts,* Editor.

40. Vita Mathematica: Historical Research and Integration with Teaching, *Ronald Calinger,* Editor.
41. Geometry Turned On: Dynamic Software in Learning, Teaching, and Research, *James R. King and Doris Schattschneider,* Editors.
42. Resources for Teaching Linear Algebra, *David Carlson, Charles R. Johnson, David C. Lay, A. Duane Porter, Ann E. Watkins, William Watkins,* Editors.
43. Student Assessment in Calculus: A Report of the NSF Working Group on Assessment in Calculus, *Alan Schoenfeld,* Editor.
44. Readings in Cooperative Learning for Undergraduate Mathematics, *Ed Dubinsky, David Mathews, and Barbara E. Reynolds,* Editors.
45. Confronting the Core Curriculum: Considering Change in the Undergraduate Mathematics Major, *John A. Dossey,* Editor.
46. Women in Mathematics: Scaling the Heights, *Deborah Nolan,* Editor.
47. Exemplary Programs in Introductory College Mathematics: Innovative Programs Using Technology, *Susan Lenker,* Editor.
48. Writing in the Teaching and Learning of Mathematics, *John Meier and Thomas Rishel.*
49. Assessment Practices in Undergraduate Mathematics, *Bonnie Gold,* Editor.
50. Revolutions in Differential Equations: Exploring ODEs with Modern Technology, *Michael J. Kallaher,* Editor.
51. Using History to Teach Mathematics: An International Perspective, *Victor J. Katz,* Editor.
52. Teaching Statistics: Resources for Undergraduate Instructors, *Thomas L. Moore,* Editor.
53. Geometry at Work: Papers in Applied Geometry, *Catherine A. Gorini,* Editor.
54. Teaching First: A Guide for New Mathematicians, *Thomas W. Rishel*.
55. Cooperative Learning in Undergraduate Mathematics: Issues That Matter and Strategies That Work, *Elizabeth C. Rogers, Barbara E. Reynolds, Neil A. Davidson, and Anthony D. Thomas,* Editors.
56. Changing Calculus: A Report on Evaluation Efforts and National Impact from 1988 to 1998, *Susan L. Ganter*.
57. Learning to Teach and Teaching to Learn Mathematics: Resources for Professional Development, *Matthew Delong and Dale Winter.*
58. Fractals, Graphics, and Mathematics Education, Benoit Mandelbrot and Michael Frame, Editors.
59. Linear Algebra Gems: Assets for Undergraduate Mathematics, *David Carlson, Charles R. Johnson, David C. Lay, and A. Duane Porter,* Editors.
60. Innovations in Teaching Abstract Algebra, *Allen C. Hibbard and Ellen J. Maycock,* Editors.
61. Changing Core Mathematics, *Chris Arney and Donald Small,* Editors.
62. Achieving Quantitative Literacy: An Urgent Challenge for Higher Education, *Lynn Arthur Steen.*
64. Leading the Mathematical Sciences Department: A Resource for Chairs, *Tina H. Straley, Marcia P. Sward, and Jon W. Scott,* Editors.
65. Innovations in Teaching Statistics, *Joan B. Garfield,* Editor.
66. Mathematics in Service to the Community: Concepts and models for service-learning in the mathematical sciences, *Charles R. Hadlock,* Editor.
67. Innovative Approaches to Undergraduate Mathematics Courses Beyond Calculus, *Richard J. Maher,* Editor.
68. From Calculus to Computers: Using the last 200 years of mathematics history in the classroom, Amy Shell-Gellasch and Dick Jardine, Editors.
69. A Fresh Start for Collegiate Mathematics: Rethinking the Courses below Calculus, *Nancy Baxter Hastings,* Editor.
70. Current Practices in Quantitative Literacy, *Rick Gillman*, Editor.
71. War Stories from Applied Math: Undergraduate Consultancy Projects, *Robert Fraga*, Editor.
72. Hands On History: A Resource for Teaching Mathematics, *Amy Shell-Gellasch*, Editor.
73. Making the Connection: Research and Teaching in Undergraduate Mathematics Education, *Marilyn P. Carlson and Chris Rasmussen*, Editors.
74. Resources for Teaching Discrete Mathematics: Classroom Projects, History Modules, and Articles, *Brian Hopkins*, Editor.
75. The Moore Method: A Pathway to Learner-Centered Instruction, *Charles A. Coppin, W. Ted Mahavier, E. Lee May, and G. Edgar Parker*.
76. The Beauty of Fractals: Six Different Views, *Denny Gulick and Jon Scott*, Editors.

77. Mathematical Time Capsules: Historical Modules for the Mathematics Classroom, *Dick Jardine and Amy Shell-Gellasch*, Editors.
78. Recent Developments on Introducing a Historical Dimension in Mathematics Education, *Victor J. Katz and Costas Tzanakis*, Editors.
79. Teaching Mathematics with Classroom Voting: With and Without Clickers, *Kelly Cline and Holly Zullo*, Editors.
80. Resources for Preparing Middle School Mathematics Teachers, *Cheryl Beaver, Laurie Burton, Maria Fung, and Klay Kruczek*, Editors.
81. Undergraduate Mathematics for the Life Sciences: Models, Processes, and Directions, *Glenn Ledder, Jenna P. Carpenter, and Timothy D. Comar*, Editors.
82. Applications of Mathematics in Economics, *Warren Page*, Editor.
83. Doing the Scholarship of Teaching and Learning in Mathematics, *Jacqueline M. Dewar and Curtis D. Bennett*, Editors.
84. Insights and Recommendations from the MAA National Study of College Calculus, *David Bressoud, Vilma Mesa, and Chris Rasmussen,* Editors.
85. Beyond Lecture: Resources and Pedagogical Techniques for Enhancing the Teaching of Proof-Writing Across the Curriculum, *Rachel Schwell, Aliza Steurer and Jennifer F. Vasquez,* Editors.
86. Using the Philosophy of Mathematics in Teaching Undergraduate Mathematics, *Bonnie Gold, Carl E. Behrens, and Roger A. Simons,* Editors.
87. The Courses of History: Ideas for Developing a History of Mathematics Course, *Amy Shell-Gellasch and Dick Jardine,* Editors.
88. Shifting Contexts, Stable Core: Advancing Quantitative Literacy in Higher Education, *Luke Tunstall, Gizem Karaali, and Victor Piercey,* Editors.
89. MAA Instructional Practices Guide, *Martha L. Abell, Linda Braddy, Doug Ensley, Lewis Ludwig, Hortensia Soto,* Project Leadership Team.
90. What Could They Possibly Be Thinking!?! Understanding your college math students, *Dave Kung and Natasha Speer*.
91. Mathematical Themes in a First-Year Seminar, *Jennifer Schaefer, Jennifer Bowen, Mark Kozek, and Pamela Pierce,* Editors.
92. Addressing Challenges to the Precalculus to Calculus II Sequence through Case Studies: Report based on the National Science Foundation Funded Project Precalculus through Calculus II, *Estrella Johnson, Naneh Apkarian, Kristen Vroom, Antonio Martinez, Chris Rasmussen, and David Bressoud*, Editors.
93. Engaging Students in Introductory Mathematics Courses through Interdisciplinary Partnerships: The SUMMIT-P Model *Susan L. Ganter, Debra Bourdeau, Victor Piercey, and Afroditi V. Filippas,* Editors.
94. Expanding Undergraduate Research in Mathematics: Making UR more inclusive, *Michael Dorff, Jan Rychtář, Dewey Taylor,* Editors.
95. Sharing and Storing Knowledge about Teaching Undergraduate Mathematics: An Introduction to a Written Genre for Sharing Lesson-specific Instructional Knowledge, *Douglas Lyman Corey and Steven R. Jones*, Editors.
96. Justice Through the Lens of Calculus: Framing New Possibilities for Diversity, Equity, and Inclusion, *Matthew Voigt, Jess Ellis Hagman, Jessica Gehrtz, Brea Ratliff, Nathan Alexander, and Rachel Levy*, Editors.
97. Equitable and Engaging Mathematics Teaching: A Guide to Disrupting Hierarchies in the Classroom, *Daniel Reinholz*
98. An Aspirational Approach to the Mathematical Preparation of Teachers, *James A. M. Álvarez, Elizabeth G. Arnold, and Elizabeth A. Burroughs*, Editors.
99. Using Data-Centric Methods to Teach Introductory Statistics, *Donna LaLonde and Deirdre Longacher Smeltzer,* Editors.

MAA Service Center
P.O. Box 91112
Washington, DC 20090-1112
1-800-331-1MAA FAX: 1-301-206-9789

Acknowledgments

The work and outcomes of the Professional Development Emphasizing Data-Centered Resources and Pedagogies for Instructors of Undergraduate Introductory Statistics (StatPREP, NSF Award #1626337) project involved a network of mathematicians, statisticians, mathematics and statistics education professionals, Introductory Statistics instructors, and MAA staff members.

We gratefully acknowledge the contributions of all of the many StatPREP workshop participants whose efforts were the foundation for this volume. Their thoughtful engagement during workshops and webinars was instrumental in shaping the ideas presented here. Each participant brought a unique perspective to the discussions, enriching our understanding of the challenges and opportunities in statistics education. The collaborative spirit of the StatPREP participants created an environment that pushed the boundaries of traditional approaches and fostered innovative thinking.

Likewise, we extend our deepest gratitude to the instructors and hub leaders who guided these workshops with exceptional skill and insight. Their commitment to fostering meaningful dialogue and their ability to connect theory with practice ensured the success of this project. Their tireless efforts, both during and after the workshops, are reflected in the pages of this volume.

We are also grateful for the support of the broader statistics education community whose efforts fostered a culture of inclusion and support ensuring many opportunities for meaningful professional development. Finally, we are thankful for support from the American Statistical Association, the American Mathematical Association of Two- Year Colleges, the Mathematical Association of America, and the National Science Foundation; we particularly acknowledge the dedicated staff of the Mathematical Association of America whose time and expertise were instrumental in ensuring all the success of the StatPREP activities.

Contents

Introduction: StatPREP Notes

Michael Brilleslyper, *Florida Polytechnic University*

StatPREP was born in 2016 and represents a practice-focused trajectory within the statistics reform efforts. Combined with the rapid development of data science and the need for a data literate workforce, StatPREP provided a path for instructors to bring forth a modern approach to introductory statistics. The program was both a professional development opportunity for instructors and a means of creating an accessible approach for students to work with real data and to experience the power of computing and visualization as the central tenets of teaching statistics.

Indeed, the StatPREP mission can be summed up in a single sentence: *To empower statistics instructors to create data-driven learning opportunities so that students can gain the skills and knowledge needed to work and live in a data-centric world.*

This MAA Notes volume is the StatPREP story. But more than that, it is a collection of writings about teaching statistics differently than how most of us learned it. StatPREP methodology and tools provide a replacement for many traditional topics, not an addition to existing paradigms. The goal is not to add more content, but rather to bring forth statistical methods and concepts differently. Students need to interact with real data in a way that directly impacts their learning. Our world demands a workforce that is data-literate, computationally savvy, and more.

The StatPREP story is also about the several hundred statistics instructors who participated in summer StatPREP workshops and who were willing volunteers in testing and adopting materials for teaching statistics. Their stories of change are often inspiring, and we highlight some of those here. Rather than simply providing tools for change, we make the case for change. What started out as a grant idea intended to reach two-year college faculty grew to be a community of educators, across institutions of all types, focused on bringing data-centric statistics education into the mainstream. The background, the philosophy, the tools, the opportunities and obstacles, the personal stories, and the vision for the future are collected in this volume, written by a group of dedicated educators from among the StatPREP community.

This volume is divided into three parts:

I. The StatPREP Approach tells the story of StatPREP's origins, lays the foundation for the StatPREP philosophy, explains the role of community, discusses tools and resources, and tells a personal account of embracing change to inspire others to do the same.

II. StatPREP in the Classroom describes how a traditional statistics classroom can be transformed into one that emphasizes the centrality of data and computing, and how to use StatPREP products such as the Little Apps in the development of data-centric lessons.

III. StatPREP as a Catalyst for Statistics and Data Science Education provides information that underpins much of the first two parts, including an overview of data considerations, a brief review of the literature on statistics education reform, and a chapter from the evaluation team describing the impact of StatPREP within the landscape of professional development initiatives. Some of this assessment material may be of particular interest to administrators or for those planning to seek funding for their own projects.

Many recent projects in statistics education reform trace their origins and motivations to the 2016 "Guidelines in Assessment and Instruction in Statistics Education (GAISE) College Report" [67]. The 2016 report updates the original GAISE College Report from 2005 and highlights the following six key principles around which statistics education should rally.

1. Teach statistical thinking.

    (a) Teach statistics as an investigative process of problem-solving and decision making.
    (b) Give students experience with multivariable thinking.

2. Focus on conceptual understanding.
3. Integrate real data with a context and purpose.
4. Foster active learning.
5. Use technology to explore concepts and analyze data.
6. Use assessments to improve and evaluate student learning

The influence of this report is far-reaching and, indeed, many authors of this volume make reference to the 2016 GAISE College Report. Because the report both informed the StatPREP work and was important to the professional development of individual authors, we have chosen to leave some of these redundancies in place. In Part 2 of this volume we will provide a succinct overview of the report highlighting the alignment of StatPREP goals with the 2016 recommendations. We encourage the reader to study the original report in its entirety. StatPREP fully embraced the GAISE College Report, including sessions about the report at the workshops, and directly putting into action recommendations 1 through 5.

Other themes make frequent appearances in many of the articles contained herein. These include the use of real data, a strong case for the use of computers and professional statistical software, a deliberate move towards re-sampling and variation as a way to instill conceptual understanding versus an algebraic or formulaic approach. Different authors view these ideas slightly differently, and we have made no attempt to agree on a single set of definitions or ideas. Indeed, we see considerable value in an exposure to an array of interpretations and approaches. StatPREP is not a uniquely prescriptive way to teach statistics. Rather, StatPREP is a community of educators, a set of tools, and a philosophy that puts data at the center of student learning.

This volume likely has something for everyone who teaches or will teach statistics. It is for educators looking to modernize their approach or who wish to incorporate the GAISE College Report principles into their teaching. It can be viewed as a prologue to data science, one of the fastest growing careers and fields of study. It is for faculty eager to include computing in their instruction but unsure how to proceed. Faculty from both two- and four-year colleges have participated in StatPREP, and have been part of the StatPREP leadership. The program reached out to instructors across the spectrum of higher education to include part-time and adjunct faculty (groups for whom professional development opportunities are often hard to come by). But mostly, this Notes volume is for anyone interested in the future of statistics education.

How should you read this book? Of course you are free to flip to any page and dig in. However, we believe setting the stage for change helps change happen. Each of the three parts begins with an Overview, and these will be helpful in orienting you to the content and the goals provided by the various authors. We particularly recommend reading all the material in Part I: StatPREP Project Background and Context. Part II StatPREP in the Classroom digs into demonstrating how StatPREP resources can be used to teach statistics, while Part III StatPREP as a Catalyst for Statistics and Data Science Education provides both context and a look at what the future may hold. This volume's companion online Library has a wealth of resources.

This Notes volume is the culmination of a seven-year program. The leadership team envisioned this book to capture the history of StatPREP and to extend its reach to a much broader community of educators. The array of willing writers that stepped forward to share their expertise and their experiences is a testament to the power of community and the dedication of faculty committed to improving statistics education on a National level. We hope this volume inspires you to approach teaching statistics differently. Wholesale change is not required. Make small changes, try something new, reach out to the community, and, most of all, have fun putting data and computing at the center of student learning.

Part I

The StatPREP Approach

Part I Overview

Michael Brilleslyper, *Florida Polytechnic University*

Part I, The StatPREP Approach, is intended to provide context and foundation for readers who may be new to the StatPREP project. Chapter 1 opens with the background and history of StatPREP, which highlights the importance of offering professional development opportunities that are timely and relevant to the needs within higher education. Chapter 2 follows with a philosophical essay in which Danny Kaplan, the principal architect behind many of StatPREP's most innovative initiatives, provides a powerful argument for placing meaningful interaction with data at the center of statistics instruction. Kaplan makes a case for computing over algebra, and provides several other thought-provoking commentaries on the current state of statistics education.

We then shift gears and provide a personal account of change. The author of Chapter 3, Jennifer Ward (an early StatPREP workshop participant), recounts her own experiences of having a strong and supportive community that helped her bring change to her own teaching. The change that Ward, and many others, incorporated was only possible because tools were developed to facilitate that change.

A centerpiece of the StatPREP philosophy is the use of computing to understand data. Kaplan developed six applications (called Little Apps) that provide a web-based, low barrier entry point for exploring real data with powerful, professional computing software. In Chapter 4, Kaplan describes the uniform design and use of the Little Apps. Much more detailed accounts of the Little Apps and how to effectively use them are given in Part II, StatPREP in the Classroom.

Central to much of modern statistical analysis is the use of modern programming languages such as R and Python. Indeed, data science programs require students to master one or both languages as part of their programs. Bringing high-level programming languages into the introductory statistics course may seem unattainable, but many instructors have developed pathways that allow students to interact with languages such as R, while learning foundational principles of statistics. In Chapter 5, Amelia McNamara describes one approach to using R in introductory courses. McNamara gives an overview of what is possible and provides suggestions on how to "gently" begin the process of using a modern programming language in your teaching.

Over the course of StatPREP's seven-year run, a wealth of material was developed: the Little Apps, R-modules, Webinars, classroom activities, and more. Tracking down useful and tested resources through general internet searches can be daunting and unproductive. Thus, we provide a Library of carefully curated StatPREP resources. These online collections allow you to further explore how you can incorporate portions of StatPREP in your own teaching.

We close Part I with another personal story of change. In Chapter 6, Megan Briet-Goodwin describes how making small changes in her teaching led to dramatic changes in her approach to statistics and to how she interacted with her students. Her experience is not unique, and it shows what is possible when people are willing to take small risks.

1

StatPREP Project Background

Sarah Holsted, *Broad-based Knowledge, LLC*

Flora McMartin, *Broad-based Knowledge, LLC*

1.1 Introduction

The 2016 report *A Common Vision for Undergraduate Mathematical Sciences Programs in 2025* included a challenge for the mathematics and statistics community to close the ever-widening gap between traditional higher education introductory statistics courses and the data-driven workplace that college graduates enter [169]. Specifically, A Common Vision states, "The material should be motivated by a variety of examples and real data sets, including data collected by students." Concurrently, in 2016 the National Science Foundation (NSF) issued a request for proposals to the Improving Undergraduate STEM Education (IUSE) program that would "improve the preparation of undergraduate students so they can succeed as productive members of the future STEM workforce, regardless of career path, and be engaged as members of a STEM-literate society."

Three professional associations—the American Mathematical Association of Two-Year Colleges (AMATYC), the American Statistical Association (ASA), and the Mathematical Association of America (MAA)—came together to address this priority through the StatPREP project (DUE-1626337). Originally called the Statistics Professional Enhancement Program for a Data-Centric World, StatPREP had the goal of fostering the widespread use of data-centered methods and pedagogies in introductory statistics courses to help prepare students to meet the demands of a data-driven workplace.

To accomplish this goal, the StatPREP project leadership team created a professional development program for mathematics faculty to learn to teach modern methods of data analytics in introductory statistics courses. The target audience for the workshops were higher education mathematics faculty, with an emphasis on faculty members teaching at two-year institutions.

The intended outcomes for the StatPREP project were guided by three strategies:

1. To provide faculty members with a jump start to using modern content and instructional practices, linking them to familiar topics in their current courses. This approach did not require faculty members to change their textbook or syllabus, an approach that would be constrained by departmental coordination of their courses and by time limitations induced by heavy workloads.

2. To shift faculty members' beliefs about the need to modernize statistics education in addition to their beliefs about teaching and learning.

3. To expose faculty members to a new way of thinking about data, how students should interact with it, and what it means to think statistically as opposed to thinking mathematically about statistics ([4]; [9]).

1.2 Building on the Literature and Prior Work

The StatPREP professional development model was designed around four principles grounded in research on professional development programs and the change and innovation process in higher education [74].

1. Workshops must be complemented by ongoing support as new ideas are implemented. Rutz, et al, reported that while participating faculty can learn core ideas and strategies in workshops, ongoing learning and support coupled with the workshops is key to changing teaching practice [166].

2. Participant self-reflection is critical. Bouwma-Gearhart asserted that science faculty members must identify the areas where they want to improve their teaching practice as a necessary first step in the change process [21]. Change is less about the thing being changed (e.g., curriculum) and more about changing beliefs about teaching and learning ([42]; [52]; [105]; [111]; [182]; [192]; and [121]). An effective approach to creating a less stressful learning environment in situations that require faculty members to question their approaches is to engage a facilitator external to the institution ([177]; [52]; [182]; and [64]).

3. Building community is valuable. For mathematics faculty involved in a seminar focused on teaching concepts of proof, Blanton and Stylianou [19] built a community of practice grounded in discussions of mathematical content to highlight issues of practice ([114]; [203]; [204]). Working collaboratively, building partnerships, and creating networks among participants are key to establishing support and buy-in for change ([52]; [105]; [107]; [106]; [146]; [16]; [64]; and [182]).

4. Change takes time. Duration, experiential learning opportunities, and peer interactions all contribute to meaningful faculty development interventions ([183]; [18]; and [69]). Change takes time; plan for the long-term ([177]; [16]; and [107]).

The StatPREP professional development model was also based directly on the decade-long MAA Professional Enhancement Program (PREP), which adhered to the four principles for professional development outlined above, and internal reports, which include information from the external evaluator's reports showing it was an effective mechanism for developing and supporting participants [58]. The lessons learned that most informed the StatPREP model include

1. travel and lodging costs to attend the workshop were prohibitive for many faculty, and

2. attending workshops lasting more than two days were not feasible for many faculty members.

Consequently, StatPREP aimed to emulate effective elements of PREP through the creation of regional hubs and a national support network, while minimizing the barriers to participation by (1) basing the program in urban areas to reduce the cost of attendance, and (2) limiting the workshop duration to 1.5 days.

Additionally, the StatPREP model drew on lessons learned from the Project MOSAIC workshops (NSF DUE-0920350) and built on the foundation laid by the TANGO project (NSF DUE-1432251). TANGO has established regional hubs and created a cadre of faculty who have been mentored in modern statistics practices.

1.3 Phase I – 2017–2018: Initial Implementation

From 2017-2018, the StatPREP professional development model was designed around the process of building regional hubs. Each regional hub would have a Hub Leader, who would host multi-day professional development workshops during the summer. The workshops were held in the same locations two years in a row. After the workshops, the Hub Leaders would grow regional communities of practice around mathematics faculty teaching introductory statistics by providing ongoing support for workshop participants and their colleagues through meet-ups or via an email listserv. The workshop content incorporated recommendations from the 2016 GAISE College Report and the workshop resources would be available online [67].

Recognizing the significant challenges facing instructors, particularly the heavy workload of instructors of two-year institutions, the sometimes-rigid curriculums found in institutions, and the lack of access to or knowledge about computer technology available for teaching statistics, the StatPREP leadership team sought to introduce and link workshop attendees to modern content and practices familiar to their current courses and practices. They offered workshops with not only new information, but also new resources, such as Little Apps, and new approaches, such as the use of the open-source and free software, R / Posit. The aim was to help workshop participants use small steps to integrate these resources and approaches in some way into one or two existing class sessions. Their belief was that once started, workshop attendees would experience positive changes in students learning statistics, which would further encourage attendees to seek out and implement data-centric statistical teaching methods.

1.4 Phase II – 2018–2019: Assessment and Revision

In Fall 2018, the StatPREP project leadership team convened a strategic planning meeting to assess the design of the StatPREP professional development model and workshop content based on the leadership team's observations and experiences, informal feedback, and formal evaluation activities. The discussion resulted in the primary observation that: *the current program is not yet adequately equipping participants to make substantive changes in the ways they teach statistics*. Table 1.1 summarizes the outcomes of the meeting in a series of observations and changes proposed for subsequent implementation of the professional development model. After the strategic planning meeting and through 2019, the project leadership team implemented the proposed changes, including teaching a revised workshop curriculum during the summer 2019 workshops.

1.5 Phase III – 2020–2022: Pandemic Pivot

Due to the COVID-19 pandemic, StatPREP workshops were not held in summer 2020. In 2021, still responding to the pandemic, the StatPREP professional development model was adapted to be delivered entirely online, and the workshop content was adapted and expanded. Table 1.2 compares the elements of the face-to-face and online StatPREP professional development model. In 2022, with COVID uncertainties still looming, the StatPREP team opted to offer a face-to-face workshop at just one site (May 2022), with no online workshops being held in summer 2022.

Based on data provided by project participants before they attended a workshop, the StatPREP project reached its intended audience. Over the duration of the StatPREP project, approximately 737 people attended StatPREP summer workshops in-person and online, and 49% of responding participants were from two-year colleges.

Since its inception in 2016, the StatPREP project has featured significant engagement with both two-year and four-year Introductory Statistics instructors. This has occurred through in-person and online summer workshops, in-semester webinars, cultivation of an online discussion platform, a monthly e-newsletter, and development of Little Apps, a set of interactive web-based tools for data-centric teaching and learning of statistical concepts. Through the creation of this MAA Notes volume and a companion online Library, many of the resources created during this project will remain accessible after the project ends.

Observation: Workshop Content	**Proposed Changes—Implemented Summer 2019**
The focus of StatPREP needs to shift towards using data-centric methods and StatPREP software (web-based applications that allow experimentation and visualization using real data, and cloud-based Posit) to teach standard topics and concepts differently, as opposed to using a data science approach to certain concepts, to replace commonly taught topics. This shift will drive changes to development of tools and workshop organization.	Beginning with summer 2019, StatPREP workshops will be focused around using data and software to provide an alternate or additional means of teaching the following six topics: · Exploratory data analysis (EDA). Focus on mean, median, and mode. Using a data driven approach, we emphasize EDA is about summarizing and simplifying data. · Normal distributions (to include some probability). Using a data-driven focus we emphasize what it means for something to be "rare" or "unusual." · Hypothesis testing. Here, the StatPREP applets allows for experimentation with different samples and different sample sizes. The emphasis will be that sample size matters and to explore "significance vs. substance." · Confidence intervals. Using data, we approach this topic as a way to "compare with precision." · Sampling. This fundamental idea is best explored using large data sets. We emphasize that "samples vary." · Regression. The StatPREP approach allows multiple models and easy experimentation. The emphasis is that "functions capture patterns."
Observation: Project Administration	**Proposed Changes—Implemented Fall 2018-Fall 2019**
Project Team roles and responsibilities need to be clarified. — Hub leaders are unclear on their roles. — Hub leaders need to be more involved and need to be given more guidance on how to engage their hubs.	The leadership team will more clearly delineate individual responsibilities to ensure ownership and follow-through. — Beginning in summer 2019, pre-workshop activities and the workshop agenda will change to reflect the following: · Workshop leaders will engage hub leaders prior to the workshop to make sure they are fully vested in the goals and content of the workshop and not merely the logistics of planning it. · Organize participants into groups of 4 prior to the workshop. Have them sit in these groups and collaborate during the workshop. · De-brief and review pre-workshop reading and activities. · Make greater utilization of hub leaders in actually leading workshop activities. Ensure hub leaders are seen as a resource.
Observation: Project Communications	**Proposed Changes—Implemented Fall 2018-Fall 2022**
The StatPREP leadership team needs a clear way to disseminate information to hub leaders and workshop participants, such as announcing upcoming webinars and also a way to encourage regular communication among participants — The StatPREP website needs to be cleaned up, simplified and kept current.	The StatPREP project leadership team will create a series of newsletters ("StatPREP: News You Can Use") and plan to do five per year. Newsletters are planned for an annual cycle with distribution in September, November, February, April, and late summer. — Work will be done on the statprep.org project website to tag specific resources to the six main topics described above.

Table 1.1: Revising the StatPREP Professional Development Model

Element	**Face-to-Face (2017-2019; 2022)**	**Online (2021)**
Time	Summer	Summer
Location	2- or 4-year campus in AMATYC region	Online
Workshop Duration	Twelve hours over the course of two days	Six hours over two days
Workshop Distribution	Offered at four different locations	Entire series of four sessions held once in June and once in July
Participants	StatPREP project team; Hub Leaders	StatPREP project team; Hub Leaders
Content	Several topics covered at each workshop (e.g., Little Apps, Posit, orientation to StatPREP resources, finding real data)	One topic (Little Apps, Introduction to Posit, Advanced Posit) was covered in-depth at each workshop session. One panel presentation "Data Science at Two-Year Colleges"
Participant Compensation	Travel and lodging	None

Table 1.2: Comparing the Elements of the StatPREP Professional Development Model

2

Philosophy of StatPREP

Daniel Kaplan, *Macalester College*

The key idea behind the StatPREP Project was that introductory statistics courses ought to be centered on the interpretation and extraction of meaning from data. While this statement may seem obvious, many faculty members teaching introductory statistics have little or no background in statistics, data, or science, a situation recognized in "Qualifications for teaching an introductory statistics course," published jointly in February 2014 by the American Statistical Association and the Mathematical Association of America. The report highlighted that many instructors may see "data" as just another topic to be added to the traditional probability-based and heavily formulaic statistics course.

A widely respected and influential statistics educator, George Cobb, entitled a 2014 paper, "Mere renovation is too little too late: we need to rethink our undergraduate curriculum from the ground up" [44]. The StatPREP Project provided an important vehicle for such rethinking of the statistics curriculum, which I summarize below as five principles.

Of course, StatPREP is not the only locus of rethinking introductory statistics. For useful examples, look at textbooks such as the Lock family's Unlocking the Power of Data or the Tintle group's Introduction to Statistical Investigations . Especially valuable are the online, freely available, open-source textbooks such as Introduction to Modern Statistics or Statistical Inference via Data Science. StatPREP materials, however, are compatible with any textbook and can help bring a new way of thinking into your class ([117]; [189]; [102]).

2.1 Five Principles

Most current statistics textbooks make claims of using real data in examples and exercises. However, many of these examples utilize small data sets that afford a specific context in which to discuss summary statistics or to provide a concrete situation in which to apply a statistical test using by-hand calculations.

These sources of data are "real" in the sense that they were collected from observations by some persons, for some purpose, and then published. Perhaps a more meaningful definition of "real" should go beyond this, requiring that the data provide a genuine illustration of the proper organization of data, the effective display of patterns, the playing out of an important class of experimental or sampling design, the opportunity for open-ended exploration, generation of hypotheses about relationships among variables, choices of appropriate analysis technique, enhancement of student motivation, and so on. "Real data" should support a meaningful lesson about extracting information from data, not just number fodder for calculations. (The StatPREP newsletter in 2018 provided a concise, pragmatic definition of "real data.")

Within this overarching general theme of using real data that supports lessons, I enumerate five specific principles that are embodied in StatPREP. These five principles formed the core themes for workshops and webinars, and are deeply embedded in the development of StatPREP tools and tutorials for teaching statistics.

1. Data graphics can be powerful.

2. Skill in properly organizing data is important.
3. A modern approach to working with data should be employed.
4. Statistics should be empowering.
5. Statistics is primarily a computational discipline, not a mathematical one.

2.2 Data Graphics can be Powerful

Introductory textbooks generally have lots of graphics, but the large majority of them are about statistical theory rather than revealing patterns in data. Such theory-based graphics are easily spotted because one or both axes are not scaled in units of any data variable.

The StatPREP Little Apps illustrate two good rules for data graphics:

1. Each of the axes forming the graphics frame should refer to a data variable. (The “grammar of graphics” usefully generalizes “axis” to “aesthetics,” which might include color, facets, and so on. [206])
2. Each point drawn in the graphics frame should stand for variables from an individual row of a data frame.

Of course, these two rules describe a particular type of data graphic: the point plot (More commonly referred to as ”scatterplot”).

Many textbook graphics do not use both axes to represent data variables. The histogram provides an illustration: one of the axes presents the result of a frequency calculation on the variable depicted in the other axis. This is wasteful of a scarce resource: there are only two axes that most people can perceive readily.

Judicious use of axis labeling will select a variable identified as “explanatory” by the analysis for the horizontal axis and a “response” variable of interest for the vertical axis. Such axes can present either quantitative or categorical variables. Figures 2.1 and 2.2 illustrate graphics in which one of the two axis variables is categorical and the other numerical.

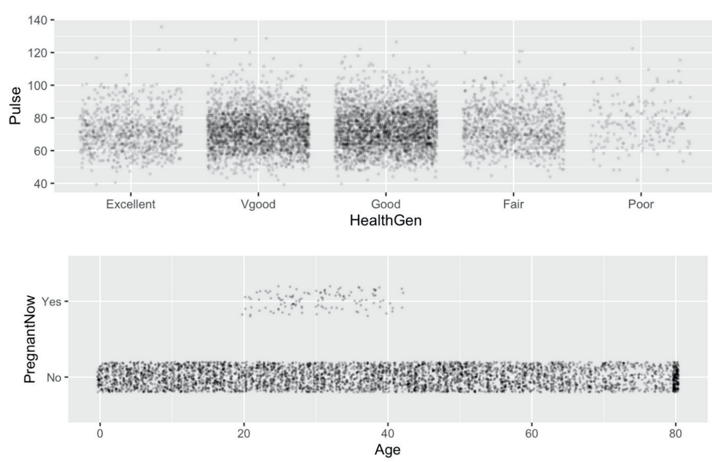


Figure 2.1: Graphics can be powerful.

The experienced instructor will note that many textbook graphics are about the distribution of a single variable. Such single-variable displays can be kept within the general data graphics framework: a response variable and explanatory variables, with the single variable of interest in the position of the response variable. The graph appearance will be that of one column in Figure 2.1 (top), which could be labeled “All” or simply 1. Arranging things in this manner lets students use the same (vertical) habits to read data graphics for a single variable as when there is a non-trivial explanatory variable.

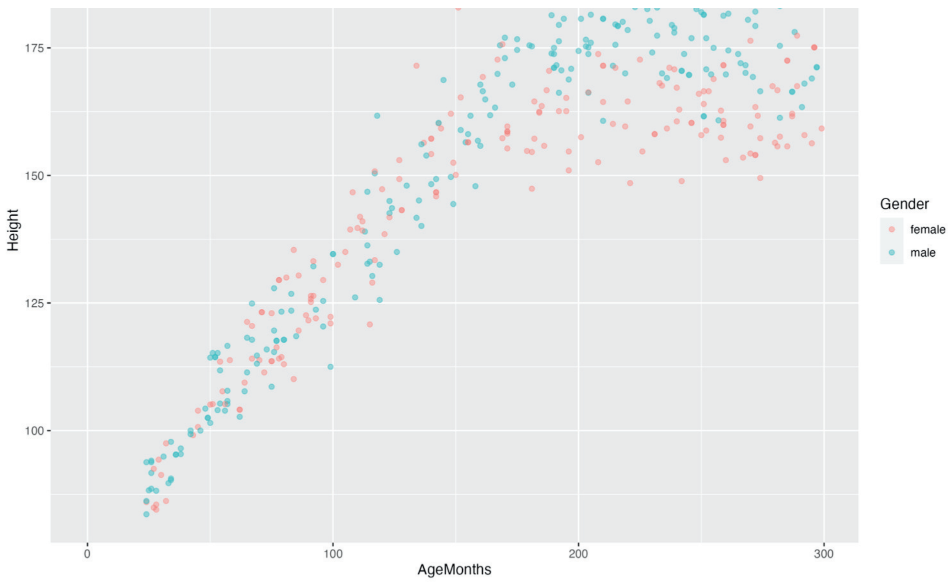


Figure 2.2: Graphics can be powerful.

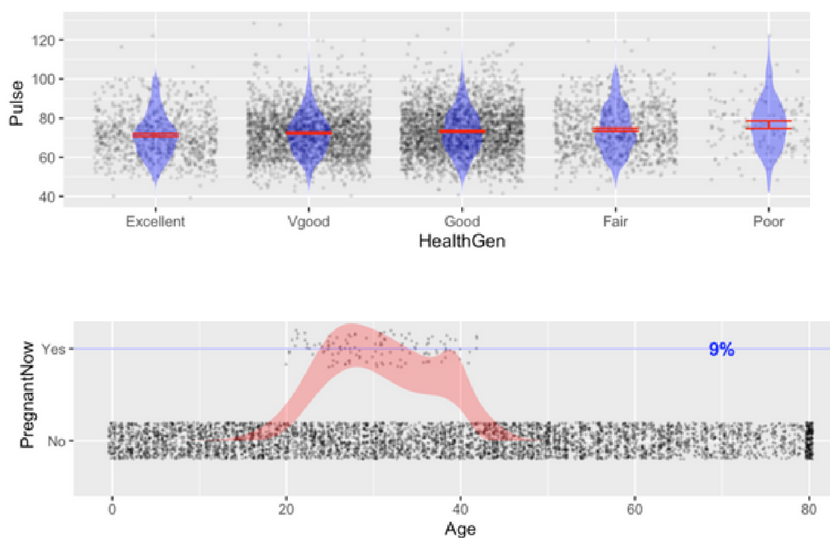


Figure 2.3: Examples of Annotations on Top of Data

The graphics so strongly featured in traditional texts——confidence intervals, boxplots, histograms, and the like——are most effective when placed as annotations on top of data. For most purposes, only three different types of annotation are needed: error bars, density display, confidence or prediction band for numerical explanatory variables. Figures 2.3 and 2.4 show such annotations in the contexts seen in Figures 2.1 and 2.2.

Annotations that might be represented by a point estimate or regression line are instead presented with confidence bands. One of the goals of Introductory Statistics is to encourage students to think in terms of uncertainty and precision. Shouldn't we be demonstrating these concepts from the very beginning?

A fundamental reason to show statistics as annotations layered on data is clear communication. For instance, the top panel of Figure 2.3 shows that the mean pulse rate tends to rise as self-assessed health gets worse. The confidence bands are narrow, as expected with large n. The p-value from an ANOVA test (not shown) is $p < 0.0000000001$. This abstract measure suggests a very "significant" relationship between pulse rate and self-assessed health. But the data shows that the pulse variation from individual to individual covers the same range for all classes of health; the relationship between pulse and health is insubstantial. A graph without the data (such as Figure 2.5) gives a completely

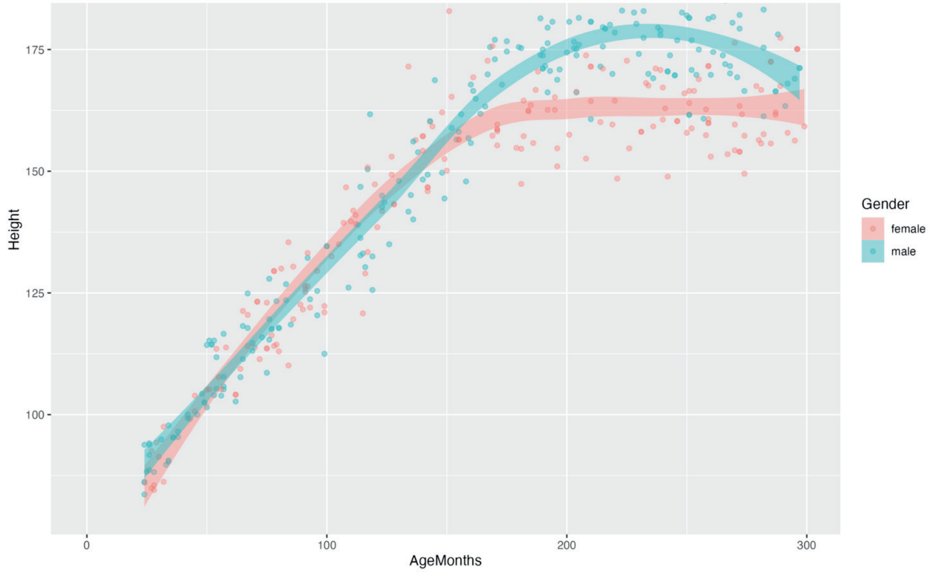


Figure 2.4: Additional Example of Annotations on Top of Data

different impression.

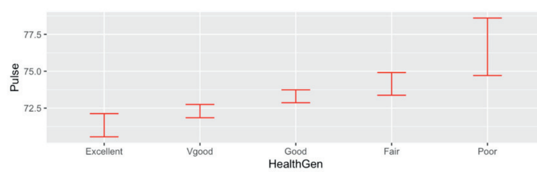


Figure 2.5: Graph without Data

2.3 Developing Skills in Organizing Data

A data-centric statistics course ought to take all aspects of data seriously: collection, storage, wrangling, display. There are simple conventions for organizing data that are within the grasp of any student. For an introductory course the "data frame" provides an appropriate organization: each column is a variable, each row is a "unit of observation."

Here are two examples of data-organization failures.

Consider this typographically compact bad presentation of numbers:

Sex						
Female	6.2	3.1	4.9	3.7	3.2	4.3
Male	2.1	4.8	5.3	4.9		

Basic questions to ask: Is the row a unit of observation? Are the columns variables? In the original textbook source, each number refers to a different person.

Our second example shows that data conventions may be violated even while using the row/column case/variable organization. The context is body temperature.

Subject	Age	Sec	8AM day 1	12AM day 1	8AM day 2	12AM day 2
1	22	M	98.0	98.0	98.0	98.6
2	23	M	97.0	97.6	97.5	—
3	22	M	98.6	98.8	97.8	98.6

There are six columns. Are there six variables? Or are there four: age, sex, temperature, date-time? Organizing the records as four variables provides useful flexibility. For instance, suppose sometimes the temperature was taken at 11AM, or covered a span of more than two days. Data in the format shown would be difficult to handle on a computer. Unless the point is to teach data cleaning, it's much better to use examples that are already tidy. Statistical Inference via Data Science is particularly good at instilling tidy principles by example, and showing straightforward but powerful data wrangling techniques that provide the flexibility needed for constructing data graphics. Instructors might reasonably differ on how much data wrangling to include in a course, but the only excuse for ill-organized data is to show techniques to organize it properly.

2.4 Employ a Modern Approach

Generations of students have relied on algebra to convey principles of statistical inference, though with mixed results at best. Simulation (e.g. resampling, bootstrapping, permutation) provides a much cleaner demonstration of the ideas of sampling variation and even the Null hypothesis. The power of combining simulation with graphical display is well illustrated by the work of Wild and Pfannkupf, Towards more accessible conceptions of statistical inference, which brings sophisticated statistical thinking to the primary and secondary school level [208].

Similarly, traditional textbooks recapitulate the early history of statistical methods in the late 19$^{\text{th}}$ and early 20$^{\text{th}}$ centuries. Marquee topics are one- and two-sample t-tests, one- and two-sample p-tests, and simple regression. A large amount of attention is devoted to the problems introduced by tiny data, that is, data with only a handful of units of observation.

Surveys of statistical techniques that are actually found in research literature point to the importance of techniques, such as logistic regression, that provide an interpretable way to include one or more covariates in the description of relationships between variables, and to quantify relationships with an effect size [97]. Many research findings turn on the choice of covariates, often reported in the press with the phrase "after adjusting for." Without an appropriate choice of covariates, a p-value (or an effect size) has no good meaning and loses relevance to the real world.

2.5 Statistics Should be Empowering

Traditionally, statistics instruction has embraced the notion of statistics as the gatekeeper to publication. This approach emphasizes that researchers can fool themselves by, for example, reading meaning into "insignificant" differences or by accepting observational data (that is, data not from a controlled experiment) as indicating causal connections. Certainly, avoiding fooling yourself with data is an important component of statistical methods.

It is also important to see the ways in which statistical methods can be empowering. Examples include untangling covariates with multivariable modeling, responsible thinking about correlation and causality using directed acyclic graphs (DAGs), rich data graphics that reveal patterns without overstating them, and the ability to cross-index data from different sources. Another example is the method of case-control studies, hardly mentioned in traditional statistical texts, but invaluable in settings like public health where the incidence of a condition is very low. A "simple random sample," the idealized data-collection method of introductory statistics texts, is effectively useless in such situations.

2.6 Statistics is a Computational Discipline

Needless to say, the storage, re-organization (wrangling), graphical display, and modeling of data are properly accomplished using a computer. Similarly, an entirely satisfactory basis for statistical inference (except for tiny data!) involves random re-selection or shuffling of data frames or individual variables. That is, simple computer operations are a solid basis for statistics.

To demonstrate that statistics is not fundamentally mathematical, consider, for instance, how statistics might operate if the central limit theorem were not a subject of mathematical proof. Statistical practice would go on unchanged. Simulations would easily demonstrate the ubiquity of the bell-shaped distribution of summary statistics.

Similarly, there is no algebraic mathematical basis for effective graphics; it is a matter of psychophysics. Database organization, although often theorized as a "relational algebra," can be used effectively by people who have none of the skills they might once have encountered in high-school algebra. Calculus techniques such as symbolic integration can be used for certain simple distributions, but these techniques are not applicable to many of the distribution curves that feature in introductory statistics.

When we envision statistics as a computational discipline, priorities for statistical curricular space immediately shift to the provision of computational skills. This will be a large, long-term undertaking that will be disruptive to established mathematics curricula and will meet considerable resistance. As a profession, we have taken only a dozen years to fully embrace the possibilities created by the computer in the form of the smartphone or the web search. But a dozen years is nothing when it comes to the formation of mathematics curricula. A data-centric statistics curriculum will only emerge to the extent that instructors develop computational skills and confidence.

2.7 Conclusion

The last 15-plus years of development in cutting-edge statistics education has demonstrated that it is entirely feasible to teach introductory statistics in a data-centric way. The enthusiasm for data skills is demonstrated by the widespread embrace of the paradigm of "data science." The societal need for data skills is demonstrated by the heavy demand for them in the workplace.

I am optimistic that data-centric statistics will continue to diffuse slowly through the academy. It will be driven by the creativity of innovative textbook authors as seen in the open-source textbooks mentioned in the introduction. It will be supported by the many young researchers who depend on the computational engines such as R that undergird much of their work. These future faculty will be well placed to use and promote a data-centric approach to statistics.

Chapter 2 Appendix: Real Data

In 2018, the "News from Danny" column in the StatPREP newsletter discussed the possible definition of "real data." We include the content of that column, verbatim, noting that any definition of "real data" is highly subjective and the benchmark identified in the 2018 discussion may be more accurately thought of as a general guideline.

> At October's StatPREP meeting [in 2018] at the Mathematical Association of America's DC headquarters the new MAA deputy executive director, Rachel Levy asked a simple question: What's real data?
>
> A core recommendation of the American Statistical Association's GAISE College Report is to "use real data" when teaching statistics. Prof. Levy wasn't looking to prompt a philosophical discussion of the nature of reality, but to define a benchmark. If a widely lauded, consensus report from the world's leading organization of statisticians calls for every introductory course to use real data, we need a way for instructors to know, for sure, whether they are in compliance. And so we discussed what is "real" when it comes to teaching statistics using real data. Our conclusion:
>
> *Data is real when it has at least 1000 rows, at least 5 variables, and was not initially collected with a primary purpose of teaching statistics.*
>
> How did we come up with this definition? In part, we looked at the examples of "real data" in the GAISE College Report, for instance a dataset on housing with 2930 rows and 80 variables, or a dataset on 53,940 diamonds with 10 variables. But mainly, we looked at the reasons motivating the recommendation to teach with real data: which practices are encouraged and which discouraged. These are: teach statistics as an investigative process, foster active learning, give students experience with multivariate thinking, use technology but focus on concepts.
>
> Why 1000 rows? Working with data on this scale requires using appropriate technology, the sort used in the data workplace. Graphics with 1000 points can be rich enough to see relationships, even when there are multiple variables. And with 1000 rows, a central concept in statistical reasoning, sampling variation, can be shown directly using random selection.
>
> Why 5 variables? "Multivariate" is at least three and there are three basic roles played by variables in data analysis: response variable, explanatory variable(s), covariate(s). But we need more than three because both categorical and quantitative variables can star in any of these roles. And we need room for students to explore actively which can be as simple as letting them choose which variables to relate to which.

Of course we understand that there is no hard statistical boundary between $n = 999$ and $n = 1000$, just as there is no hard boundary at $p = 0.05$.

Now that you have precise criteria for the “real” in “use real data,” our next task will be to define “use.”

3

The StatPREP Community

Jennifer Ward, *Clark College*

Developing community is central to the StatPREP program. In fact, two of the project goals are "establishing regional communities of practice to support instructors who teach introductory statistics" and "establishing a national online support network comprising instructors who teach introductory statistics and statistics education experts." Community takes on many forms within the StatPREP program, from interpersonal relationships to nationwide networks. The participants have diverse backgrounds and teaching experiences, so I am but one voice in a chorus of many who can attest to the benefits, personally and professionally, of StatPREP. This chapter will be my story of the communities that StatPREP nurtured. I will share how a serendipitous meeting over lunch gave me a new perspective of what's possible in statistics education when you have the opportunity, such as with the StatPREP conference, to meet people generous with their time and just as enthusiastic to "talk shop" as you are.

3.1 A Fortuitous Beginning

In early 2018, a statistics colleague of mine at Clark College forwarded an email to our department. The email began with the line "*Whether you are teaching statistics for the first time or the fortieth time, the MAA StatPREP program is for you!*" Having taught statistics for over 15 years and serving as an adjunct statistics instructor juggling work at two community colleges, I definitely fell into the latter category. The email continued: "*Learn to use data-centered methods and pedagogies in your introductory statistics course. Collaborate with your peers and interact with national experts in statistics education.*"

The idea of putting data at the center of my statistics courses intrigued me. I decided to apply to StatPREP out of professional curiosity and an interest in aligning my teaching with national recommendations.

First, the American Statistical Association recommends a data-centered approach to teaching statistics, as spelled out in Guideline #3 of the 2016 Guidelines for Assessment and Instruction in Statistics Education (College GAISE, [67]): "Integrate real data with a context and purpose." I reflected upon my teaching: do I authentically include real data in my class? What can I learn at this workshop that can help realign my teaching practices with Guideline #3? Will I find sources of interesting data that my students feel is real and not contrived?

Second, StatPREP prioritized access for participants by locating workshops in colleges near a major airport in cities geographically dispersed across the United States. In my case, a StatPREP hub was going to be held two hours from my house and participants would be given a nominal stipend to cover expenses (lodging, fuel, etc). When the participants are local to each other, there's also a higher chance of running into a fellow StatPREP attendee at, say, your local American Mathematical Association of Two-Year Colleges (AMATYC) conference. I attended StatPREP before the COVID-19 pandemic, so developing a community was more dependent on in-person events.

Third, StatPREP carves out time to learn, reflect, and make over parts of my statistics class. The opportunity to connect with fellow community college statistics teachers was intentional and central to the StatPREP grant.

Lastly and personally, the original AMATYC email invitation to apply StatPREP created an opportunity for me to find a passion. For too long, I was too busy teaching at multiple schools to take the time to figure out what I was authentically interested in. While the opportunities to be involved on committees may be plentiful, they may not be fulfilling to one's soul. With StatPREP, the right opportunity had presented itself and I had to take it.

3.2 Workshop Connections

A colleague at Clark College who would be teaching statistics for the first time later that year, replied to the email announcement to ask if anyone else was interested in attending the Seattle Hub of StatPREP. My excitement for StatPREP grew knowing that I didn't have to attend a conference alone. Within a few months, we were both accepted to StatPREP.

The first sense of community I experienced from StatPREP was developing a better working relationship with one of my Clark colleagues. We carpooled from Vancouver, WA, to Highline College, two hours each way. The car ride gave us time to get excited about StatPREP and also time to review what we'd learned after the workshop ended. During the workshop, it helped to see a familiar face in the halls between sessions. After the conference, we were able to support each other because of our common interest in using real data when teaching students. The difference in our experience with teaching statistics was mutually beneficial; I was a resource for my colleague in addressing her questions about familiar content and her questions prompted me to think critically about my answers to ensure I wasn't passing along incorrect information or forgetting important details. These conversations continue today but online. The StatPREP online discussion is like a virtual car ride with colleagues. Someone asks for advice or shares a resource, and the "passengers" share their perspectives.

At the StatPREP workshop, breakfast and lunch were provided each day which was another avenue that built community among its participants. In my second year at the StatPREP hub, I sat down at a lunch table with a number of people I didn't know yet, and easily struck up a conversation with the man sitting next to me. We discussed the morning's sessions and teaching statistics with the other people now sitting at the table. I shared other ideas and professional development experiences I'd had, including a pipe dream of writing something that I could publish. I told my lunch acquaintance about my recent professional development experience called Teaching Squares. He casually replied that I should write about Teaching Squares in JSE (The Journal of Statistics Education, now called The Journal of Statistics and Data Science Education, JSDSE). The confidence of his tone—-as if this task was fully attainable—-was shocking and yet oddly inspiring. Why NOT share this unique experience in Teaching Squares though writing a paper? I left lunch with so many new ideas and thoughts based on our table conversation. Only after lunch was I able to put a name to the face of my mealtime companion: Nick Horton, a StatPREP hub leader, who would go on to be a Vice President of the American Statistical Association.

3.3 Leveraging Community Resources

StatPREP offered my first foray into meeting and learning from influential people in the statistics education world by giving all participants the opportunity to interact with professionals who are doing great work. Funding for community college faculty to travel to expensive and distant conferences is often scarce. StatPREP did us all a service by giving us access to professors we wouldn't ordinarily run into. Through the StatPREP community, I have built relationships with many of the leading voices in statistics education—people whom I have now met and even been on committees with.

After the StatPREP conferences ended, former and future attendees were able to stay connected with an online discussion board (aka the virtual car ride), hosted by the MAA. The MAA Connect StatPREP Hub Community discussion board supports the project's goals of "establishing a national online support network comprising instructors who teach introductory statistics and statistics education experts." Not only were project leaders such as Danny Kaplan, Kate Kozak, and Ambika Silva just a discussion post away, over 200 StatPREP attendees from all over the country have a centralized place to ask questions, give their own input, share teaching resources, stay informed about upcoming events, and more. I personally used the discussion board to share activities that I had created using Little Apps.

One very useful aspect of StatPREP, especially for faculty newer to teaching statistics, is how willing instructors are to share what they created for the classes. If you check out the Library on MAA Connect StatPREP community, you'll find recordings of webinars, sources for data sets, example assessments, class activities, and more.

The MAA and StatPREP regularly presented webinars from 2017-2022 as a way of engaging and connecting with statistics instructors who were not part of a StatPREP cohort. The webinars were designed to reconnect StatPREP participants, share teaching resources with non-StatPREP educators, and highlight some of the useful technology resources created for the StatPREP project: Little Apps. I was invited to provide a webinar sharing my experience using the Little Apps in the online classroom.

I like to incorporate the Little Apps at points in the term where the visual demonstration can say more than words on the paper. I use the Little App named Regression to quickly show students the connection between a correlation and the scatterplot of data. By resampling, students can see that sample statistics vary based on the sample. There are data sets preloaded into the Little Apps, such as the NHANES datasets [30] or Lock5 data [117], but you can also upload your own data into the Little Apps. If you want your students to analyze data that's more "true to life" and see themselves as a "real" statistician, a calculator is not the right tool for the job. The Little Apps make it easy to explore multivariate relationships with the click of a button.

An explicit goal of StatPREP is to provide easily accessible professional development opportunities for two-year college faculty, such as myself. After I attended a StatPREP workshop in 2018, I was "bitten by the conference bug." I soon found myself signing up for any conference I could get my community college employers to pay for: the Women in Statistics and Data Science (WSDS) conference in Ohio, the virtual Electronic Conference on Teaching Statistics (eCOTS)/US Conference on Teaching Statistics (USCOTS), and the virtual Joint Statistical Meetings (JSM). Did I want to meet other statistics teachers I could bounce ideas off of? Yes. Was I the awkward one at WSDS who knew no one and talked to strangers just so I wouldn't be standing alone? You bet. My awkwardness eventually turned into bravery. At each conference I met a few more people, recognized a few more names, and subsequently joined a strong community of statistics educators who are supportive and inspiring. Had it not been for the easy access and opportunity to attend a StatPREP conference close to where I lived, I'm not sure I would be here writing these very words.

To bring this story full circle, look what a forwarded email, from a colleague to his department, has led to! My colleague opened up his community to me, and he may not even realize it. If you belong to a professional organization, like MAA, AMATYC, or the ASA, consider the positive impact you may have on your colleagues when you connect with your community. I know that I will continue to be a part of this community by developing my mentoring network, asking for help teaching R, and sharing what I create with colleagues. Our StatPREP community is powerful.

4

StatPREP Tools: The Little Apps

Daniel Kaplan, *Macalaster College*

4.1 Introduction

The instructor setting out to make changes in the way she teaches statistics is likely to encounter an evangelist: a person who has seen the statistical light and wants to bring others to see the way of statistics in that same light. For many of the earliest StatPREP participants, my evangelical message was, “Learn statistical computing. Teach with computing, not algebra.”

Many StatPREP instructors do not have extensive experience with computer languages or with the pedagogy of computing, let alone the technical skills to provide computing facilities to their students. Preaching computing to such instructors understandably does not help them better accomplish their job.

The situation reminded me of a story from the first century BC. A merchant—that is, someone who like today’s statistics instructors is already busy with the demands of the moment—nonetheless wanted to live an ethical life to the best of his ability. The merchant did not have time to study volumes of scripture, let alone the even more voluminous commentary on scripture. He sought instruction and was prescribed a course of intensive study. To make clear that his resources and time for study were limited, he re-framed his request: “Teach me as I stand on one foot.” The sages receiving this request thought it absurd and sent the merchant on his way without any instruction.

One sage, Hillel (c. 100 BC to 10 AD), understood why the request to condense all of ethics into one, short maxim that could easily be applied in everyday life was reasonable. Hillel replied to the merchant, “What you do not want someone to do to you, do not do to him or her.”

Inspired by Hillel, I committed myself to construct a stand-on-one-foot simplification for statistical computing. I wanted to keep the essentials of statistical thinking and put them within easy reach of even the busiest of instructors. In today’s world, the web app is the favored form of making computation accessible. Like Hillel’s maxim summarizing ethics, web apps can be carried along with you as you go about your business. A student needs only a phone with a web browser.

The StatPREP Little Apps present the essential elements of computational statistical thought:

- Since statistics is about data, every Little App is centered on an engine for displaying data graphically along with annotations showing appropriate statistical summaries.
- Since data are rich and various, the Little App graphical engine can be easily applied to any of hundreds of data sets, many containing twenty or more variables.
- Since statistical inference is important, every Little App provides immediate access to the essence of inference.

Figure 4.1: The central graphical display.

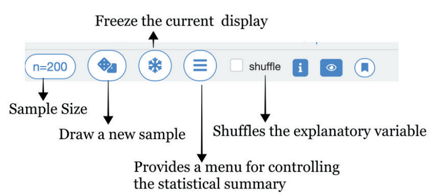


Figure 4.2: Control bar

4.2 Anatomy of a Little App

Before explaining why the Little Apps are designed as they are, we will show you the design, using the t-test Little App as an example. In a printed report such as this, displays are necessarily static. You may want to open your web browser to be able to interact with the Little App. The link for the t-test app is https://maa-statprep.shinyapps.io/Little_App_T/.

Each Little App is built around a graphical display of data and, layered on the data, a graphical statistical summary. Figure 4.1 is an example of the graphical display. Each Little App provides four different views, represented by the choices at the top of the window.

- The “Data” view allows for selection of which data set and which variables within that data set to display.
- The “Graph” view shows the data graphic in a full-window mode, which makes it easiest to discuss the data with your class.
- The “Compare” view shows two graphical displays, side-by-side, making it easy to compare.
- The “Stats” view gives a more-or-less traditional numerical statistical summary. Each of the four views includes a control bar, as shown in Figure 4.2.

Pressing a control button or checking “shuffle” immediately updates the current display. The “frozen” display remains as it was, although it can be set to the current display by pressing the Freeze button.

4.3 Little Apps and Statistical Inference

The features of the Little App graphical interface listed above, which are common to all the Little Apps, are oriented to helping students assimilate ideas that are central to statistical inference. Textbooks typically break down inference into separate topics: difference in proportions, difference in means, slope of a simple regression line, and so on. But all these topics are built on a small set of common ideas.

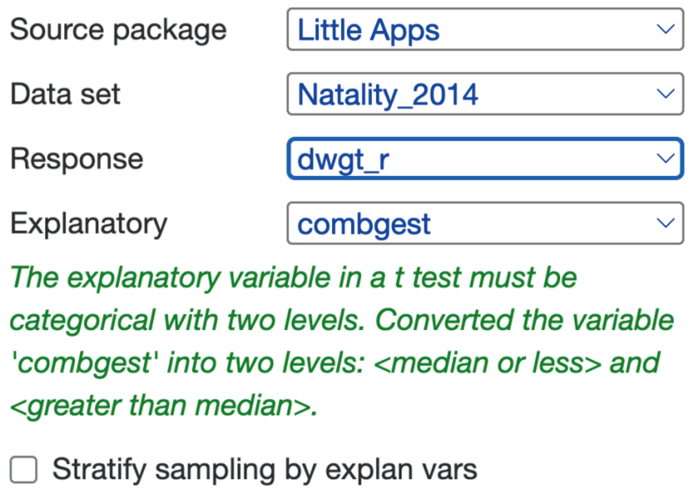


Figure 4.3: The data and variable selectors are much the same in every Little App. This is the one for the t-test app.

There is an extensive theoretical framework for statistical inference, dating from the early decades of the 20th century and featuring concepts with which every statistics instructor is familiar, such as population, sample, parameter, statistic, sampling distribution, sampling distribution under the Null.

Our goal in introductory statistics is not the exposition of a theory, but the assimilation by students of the whys and wherefores of established statistical techniques. We can facilitate the achievement of this goal by organizing inference concepts differently.

A case in point comes from the sampling distribution. As the proponents of the bootstrap and permutation approach to inference have shown us, a student can assimilate the idea of sampling variation by direct experience: drawing a new sample and observing that a summarizing statistic changes. Visually, change is most easily perceived when there is discernible movement involved.

To demonstrate sampling variation, therefore, all the Little Apps contain a button that takes a new sample and immediately displays in the central graphic the change, both in the data and in any summary statistics being displayed. Similar instantaneous change is invoked by the shuffle control. Whenever a new sample is generated with shuffling turned on, the explanatory variables are shuffled and the graphical display updated. An effective teaching opportunity can be made by setting up the Little App by displaying a response and explanatory variable that are strongly correlated. Then toggle the shuffle control on and off. The student can quickly see that shuffling/permuting the explanatory variable leads to a discernibly different display; the world of the *Null hypothesis* is different from the world in which the data were collected.

Another key idea for students to assimilate is that the amount of sampling variation depends on the *size of the sample*. This idea can be expressed in terms of the sampling distribution by noting that the standard error of a statistic is proportional to $1/\sqrt{n}$. Or, it can be shown much more directly by changing the sample size with the $n = 200$ control and then repeatedly generating new samples.

To summarize, the Little Apps controls and graphics provide a way to create student intuition about inference by side-stepping the abstraction of a sampling distribution in favor of an in-the-moment comparison of successive instances of sampling.

4.4 Little App Data

The Data view allows the user to select a data set and variables of interest. The controls appear as in Figure 4.3. To organize the description, we will follow from top to bottom the selectors and messages in Figure 4.3.

Note: No R programming or commands are needed to use the Little Apps. Nonetheless, the capabilities of the Little Apps are built on top of the features of the R ecosystem. In order to make a large number of datasets available, the Little Apps exploit the "package" system in R.

R allows any contributor to distribute added functionality and data though the package system. Many R packages—there are more than 10,000 available through the official channel and an uncounted number available by other channels——are designed specifically to provide ready access to data. Think of the choices for packages available under "source package" as akin to the various sections of a public library: non-fiction, self-help, home improvement, and so on. You go to the library section of interest and then have access to any of the books shelved in that section.

For simplicity, the Little Apps provide access to just a handful of packages that have been curated by the authors of some of the popular introductory statistics textbooks. Some are open-access: the Introduction to Modern Statistics, Statistics using Technology, and the Project MOSAIC datasets associated with Statistical Modeling: A Fresh Approach. Others are from commercial publishers, such as The Lock family's Statistics: Unlocking the Power of Data.

Data set: In a Little App, each of the packages available under the "source package" drop-down corresponds to a *data frame*. A data frame is a row-and-column oriented table. Each column is a variable that has a name assigned by the author of the package. The "data set" selector lists all of the data frames available in the selected package.

Response variable: Every app requires that a response variable be specified. The variables available in the selected data set are listed in the "response" selector.

Explanatory variable(s): Almost always, there is another variable by which the app user seeks to "explain" or "account for" the response variable. This will be familiar to most instructors as the setting for "simple regression": the response variable is y and the explanatory variable is x. (In multiple regression, there are multiple explanatory variables. Some Little Apps have a selector for this.)

For those familiar with the settings described by "two-sample" tests, take note that this terminology is replaced by the more general "response vs explanatory" framework. For a t-test, for example, the explanatory variable describes which entry in the response variable corresponds to each of the two groups. (In the t-test App, "turning off" by changing the explanatory variable to "Select explanatory var:" results in the explanatory variable switching to "one-sample t-test" mode.)

Message (in green type): Sometimes the selected explanatory or response variables are not directly suitable for the statistical method being used in the App. For instance, the t-test involves a quantitative response variable and a categorical explanatory variable (with two levels: the two groups being compared). Any variable that does not meet the requirements is coerced into one that does. In Figure 4.3, the message is explaining that the quantitative explanatory variable has been converted into a two-level categorical variable.[1] The opposite sort of conversion is also available: a categorical variable can be converted into a 0-1 encoding.

In the t-test App, a categorical response variable is converted into a 0-1 numerical representation, with 1 representing the most heavily populated categorical level. This may be unfamiliar, but it is meaningful. Effectively, the t-test App can handle situations that traditionally would be called "differences in proportions." After all, the mean of a 0-1 variable is the proportion of 1s.

Stratify sampling: When the explanatory variable is categorical, stratifying the sample will attempt to draw equal numbers of specimens from each of the levels. (It can't always succeed.)

Data display: At the bottom of the data view, the data in the current sample are displayed. (For large samples, only a subset is shown.)

The multiple text-book-oriented packages listed in the Little Apps were not designed specifically for the Little Apps. For instance, some data sets may have only one variable, rendering them unsuitable for, say, the regression Little App.

As a rule, there is documentation available for each data set. The quality of the documentation is often excellent, but sometimes not, depending on the priorities of the author of the package holding the data set.

[1] As the name implies, a two-level categorical variable is one that classifies observations as belonging to one of two categories.

4.5 Exploiting Multiple Variables in Multiple Data Sets

Instructors will often want to select a data set and response/explanatory variables that illustrate a particularly statistical topic of interest. Directing students to the example is a matter of specifying the source package, data set, and response/explanatory variables. This is a useful pedagogy for emphasizing the framing of a research hypothesis.

The ability to rapidly change the choice of dataset or of response and explanatory variables supports an exploratory pedagogy of trial and error. For instance, a lesson on correlation coefficients might involve directing students to find a data set and variables that produce a correlation greater than 0.5 or a negative correlation.

Had Hillel been a statistics instructor in the 21st century, perhaps he would have provided the sage advice and methods needed to bring a deep understanding of statistical inference to the students of the world. In his absence, we are forced to rely on technology and the accumulated experience of a small, but growing group of educators. The Little Apps are an attempt to consolidate deep ideas about variability and data into a digestible form that empowers students to learn important concepts. It is not necessary to stand on one foot while using the Little Apps, but it might just amuse your students.

5

Moving to R: A Gentle Introduction

Amelia McNamara, *University of St. Thomas*

5.1 Introduction

Imagine a familiar scenario from many companies: your boss asks you to compile a report about the products the company produces, including summary statistics on the most popular products, a plot of popularity over time, and a statistical model of what attributes lead to the most sales. You open the data as an Excel spreadsheet and produce the relevant statistical products (graphs, models, summary statistics), then paste them into a Word document to write the report. So far, so good.

You send it to your boss, who is thrilled with the report, but wants it updated to remove one category of products. You need to go back to the spreadsheet, filter the sheet to remove those products, and then update each table and graph in your Word document. This is starting to get more tedious.

The next year, your boss asks you to produce an analogous report with the most recent 12 months of data. You open your year-old spreadsheet and remind yourself what you did, trying to replicate that in the new spreadsheet, producing the same sorts of tables and graphs and pasting them into a new Word doc. It's hard to remember what you did, and you end up duplicating a lot of work.

Instead of this paradigm, imagine an alternate scenario: this time, you use the programming language R to do the analysis. Instead of clicking and dragging in Excel to generate a plot, you write a line of code. To produce the report, you click a button that runs all the code you've written and creates a PDF you can send to your boss. When they ask you to remove one category, you go back to the code, change one line, and re-run the document. Every statistic, graph, and model automatically updates and the report looks just as nice. The next year when you need to make the new report, you re-open your code file, change the input data, and re-run the PDF.

That's the argument for using a programming language as a professional, and we believe those skills will be relevant for students in a variety of future careers.

There are teaching-related benefits, too. When teaching a GUI tool (like Excel, Minitab, etc), you need to explain each of the steps to take for a particular method. If you are demoing live and a student misses a step, you have to go back and show them again. And if they are working on homework or a project and run into an issue, you end up debugging from screenshots sent to your email. This can be painful.

Teaching programming can be intimidating (and yes, sometimes painful!) but it has many benefits. One is that when you teach a programming language, the last few commands you have run will be visible on the screen, so even if you are demoing live, students can explicitly see the past steps you took. This reduces the amount of times you have to go back and re-do an entire process. You can provide code files up front to students, for them to reference and modify as they work. And if students run into issues when they are working independently, they email you chunks of code and error messages, which allow you to reproduce issues and pinpoint errors much more easily from afar.

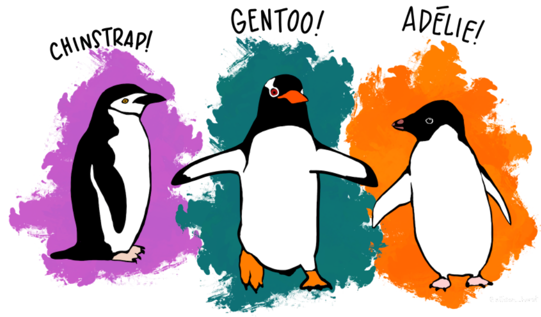


Figure 5.1: Illustration of the three penguin species represented in the Palmer penguins dataset. Artwork by @allison_horst.

5.2 Day One of R

Obviously, writing a full report with R isn't going to happen in the first day (or maybe even the first semester!) of teaching and learning R. But, gentle introduction to the language can set the stage for more advanced work later on. Instead of focusing on that future stage, let's reset the clock to day one.

On the first day students begin working with code, we recommend "starting with cake," showing students something cool they can accomplish on that first day. In many courses, this means that the first R content you teach will be focused on data visualization, a topic many students find interesting, and which produces satisfying results within the first hour or so [200].

For example, on day one you could introduce the Palmer penguins dataset [94]. This dataset is available through R, and contains information collected by researchers about penguins on islands in the Palmer Archipelago. Students enjoy this data because it is genuine research data, and it is also accompanied by cute penguin illustrations by one of the scientists (e.g., Figure 5.1).

In their first hour working with R, students could load the data,

```
library(mosaic)
library(palmerpenguins)
data(penguins)
```

Scroll through it in Posit, as in Table 5.2, and then begin making data visualizations

```
gf_histogram(f̃lipper_length_mm, data = penguin)
```

to achieve Figure 5.3.

If you want to keep day one easy, you could show just those four lines of code. We'll talk through them each in turn.

R relies on the use of functions, which are pieces of code followed by parentheses. You can think of them like math functions (e.g. $f(x)$, "f of x"), and in fact, the parentheses are typically pronounced as "of," just as would be the case in a math formula. Inside the parentheses are the "arguments"—that is, whatever is being fed into the function.

The library() function is used to load additional packages. R Packages include R functions, the documentation that describes how to use them, and sample data. The creators of the mosaic package have all been involved in StatPREP to one degree or another, so it can be a good package to use for a gentle introduction to R [150]. By running

`library(mosaic)` you are loading the package in for use.

The palmerpenguins pacakage contains the Palmer penguins data. By running

`library(palmerpenguins)` the package is loaded, meaning the data is available to your R session, but not loaded in yet. To load the data, you run

```
data(penguins)
```

Those first three lines are setup code, and your students will become extremely familiar with the library() and data() functions as they use R.

The more interesting line of code is

`gf_histogram(f̃lipper_length_mm, data = penguin)` which actually makes the plot. The letters 'gf' stand for 'graphics formula,' and are at the beginning of every data visualization function from mosaic. We might read that line of code as "gf histogram of flipper length mm, data equal to penguins." The first argument for a gf_ function is the penguins

Filter

	species	island	bill_length_mm	bill_depth_mm	flipper_length_mm	body_mass_g	sex	year
1	Adelie	Torgersen	39.1	18.7	181	3750	male	2007
2	Adelie	Torgersen	39.5	17.4	186	3800	female	2007
3	Adelie	Torgersen	40.3	18.0	195	3250	female	2007
4	Adelie	Torgersen	NA	NA	NA	NA	NA	2007
5	Adelie	Torgersen	36.7	19.3	193	3450	female	2007
6	Adelie	Torgersen	39.3	20.6	190	3650	male	2007
7	Adelie	Torgersen	38.9	17.8	181	3625	female	2007
8	Adelie	Torgersen	39.2	19.6	195	4675	male	2007
9	Adelie	Torgersen	34.1	18.1	193	3475	NA	2007
10	Adelie	Torgersen	42.0	20.2	190	4250	NA	2007
11	Adelie	Torgersen	37.8	17.1	186	3300	NA	2007
12	Adelie	Torgersen	37.8	17.3	180	3700	NA	2007

Figure 5.2: The Palmer penguins data previewed in Posit.

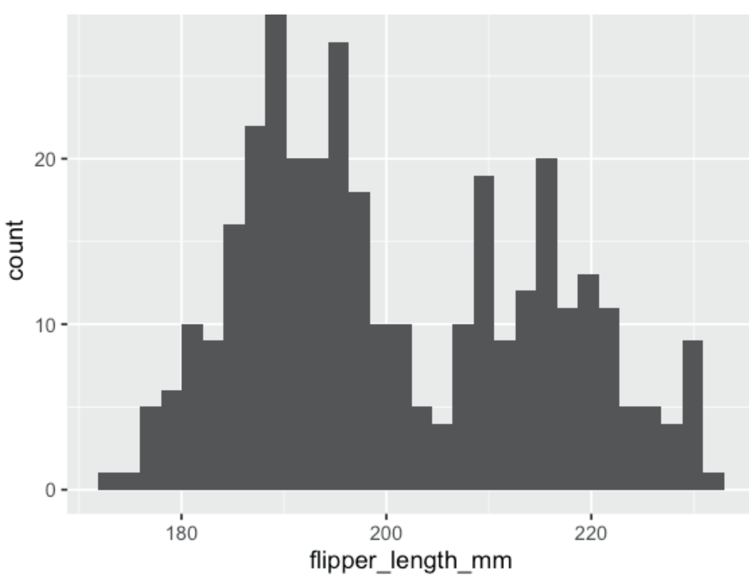


Figure 5.3: A histogram of flipper lengths of the Palmer penguins, achievable in the first hour students are working with code. This figure can kick off a discussion of center, shape, and spread.

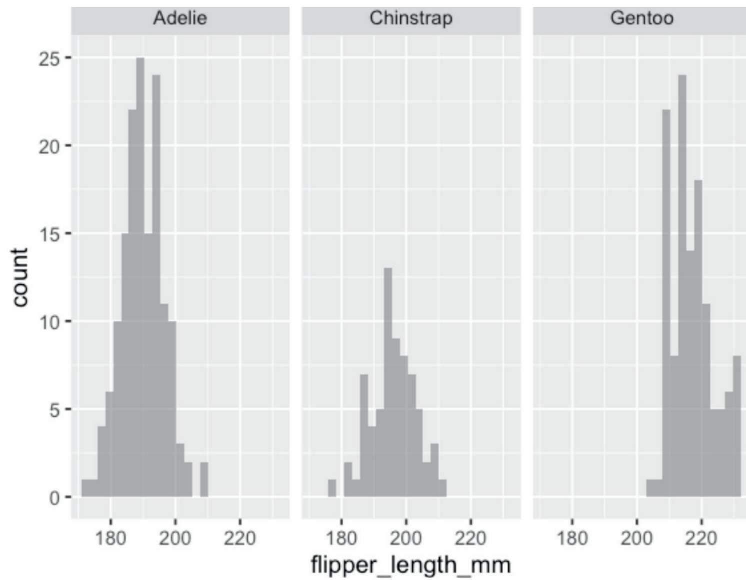


Figure 5.4: A multivariate histogram of flipper length, faceted by species.

variable(s) you want to visualize, followed by a comma and data = followed by the name of the dataset. In this case, our dataset is called penguins and contains the variable flipper_length_mm. (If you want to remind yourself of the other variables in the data, check out Table 5.2.)

With those four lines of code and the resulting plot, you could kick off a discussion with your students about the center, shape, and spread of the distribution. You could ask students things like "what do you notice?" and "what do you wonder?" (These questions are inspired by the American Statistical Association and New York Times collaboration, "What's Going on in this Graph?", [75].) If you have time and would like to go further, modifications to the plot could be made in order to explore what makes the plot appear bimodal. For example, you could "facet" the graph by species as seen in Figure 5.4. This view makes it clear that the three species have very different flipper lengths.

5.3 What is R, and Why Teach It?

Hopefully I have convinced you that you don't need to overwhelm your students with code on the first day they see programming. But, I haven't really explained what R is. At its core, R is a programming language designed for doing statistics and data analysis. It is open source, so no cost and you can do what you want with it [153]. It was developed in the 1990s as an open-source alternative to the S programming language.

There are many other tools you could use for analysis and teaching. For example, if you wanted to use a Graphical User Interface (GUI), you could use Stata, Minitab, JMP, or many other software packages. GUI tools typically have the advantage of a very low barrier to entry. They can be less intimidating to students than a programming language, and include menus and wizards to guide the analyst through a particular type of analysis. But, they usually cost money, are difficult to debug (especially from afar), and many GUIs designed to teach introductory statistics can't be used outside the classroom setting.

In contrast, programming languages are typically free, easy to debug (even over email), and are used by professionals in the fields of statistics and data science. For all of these reasons, we strongly believe in teaching a programming language to statistics students, even those at the introductory level. Although it is slightly more challenging to get started with programming, students appreciate using an authentic tool in the classroom.

The three most common programming languages for working with data are R, Python, and Julia. We believe that R is the best language for teaching and learning statistics. The main reason for this is that package designers have spent years developing R packages to ease teaching and learning.

When you download R, it comes with a few "base" packages, which allow you to perform standard statistical tests and create graphics. However, the power of R really comes from its add-on packages. Anyone can create an R package to extend the language, implement new statistical methods, or customize graphics to a particular data type.

For example, if you want to do spatial statistics, you may want the sp or sf packages. If you are interested in highly customizable data visualizations, you may want ggplot2. For classification and regression models, car would be the place to start. These add-on packages are hosted in the Comprehensive R Archive Network (CRAN), and they can be installed using a single line of R code. As of this writing, there are over 18,000 packages on CRAN. Luckily for you, it will be sufficient to utilize just a handful of packages as you get started. For a semester-long introductory statistics course, you can use fewer than five packages.

For example, a class using what is commonly called the Lock5 book written by Robin Lock, Patti Frazer Lock, Kari Lock Morgan, Eric F. Lock, and Dennis F. Lock [117] would probably want

library(Lock5Data) which contains all the datasets described in the book, and

library(mosaic) to provide consistent functions for summary statistics and graphics. However, there are other options for packages depending on the content and goals of your course.

5.3.1 R syntax

Every programming language uses a very specific syntax, which is particular to the language. Syntax governs which pieces of code will be correctly interpreted by the language, and which will produce an error. Human language has syntax as well, the rules that govern which sentences are understandable or make sense. However, when a child or language learner makes a syntax error in human language, we can typically understand what they are trying to say.

Programming languages do not have this flexibility of understanding. A missed parenthesis or a capital letter where a lowercase one was expected will result in code not running. This can be frustrating for students, because identifying syntax errors requires a lot of attention to detail, and understanding of which details matter.

An additional complication specific to R is that the add-on packages for R can use different syntactic structure. There is an aphorism about Python that "There should be one– and preferably only one –obvious way to do [things]" [143]. This is not true for R. In R, there are often at least three equally valid ways to "say" the same thing. These three syntaxes are base, formula, and tidyverse [122]. The base syntax is used by the base R language, and is characterized by the use of dollar signs and square brackets, as seen in **code chunk 1**.[1] The formula syntax uses the tilde to separate response and explanatory variable(s), as seen in **code chunk 2**. The tidyverse syntax uses a data-first approach, and the pipe to move data between steps as seen in **code chunk 3**.

code chunk 1. Making a histogram of the flipper length, then taking the mean of length, using the base R syntax.

```
hist(penguins$flipper_length_mm)
hist(penguins$flipper_length_mm, na.rm = TRUE)
```

code chunk 2. Making a histogram of the flipper length, then taking the mean of length, using the formula R syntax.

```
gf_histogram(~flipper_length_mm, data = penguins)
mean(~flipper_length_mm, data = penguins, na.rm = TRUE)
```

code chunk 3. Making a histogram of the flipper length, then taking the mean of length, using the tidyverse R syntax.

```
ggplot(penguins) +
    geom_histogram(aes(x = flipper_length_mm))
penguins | >
    drop_na(flipper_length_mm)| >
    summarize(mean(flipper_length_mm))
```

While we suggest picking one syntax and using it consistently throughout your course, it is important to be aware of the other syntaxes, primarily so as to reduce confusion when searching online; if you have only been shown the formula syntax but your search returns tidyverse syntax, it will be baffling and intimidating. I have created a "cheatsheet" that shows the same basic tasks done in all three syntaxes, so students can learn the "analogies" between them [122].

[1] These "code chunks" would be called "Listings" in R terminology.

5.4 Best Practices for Teaching R

5.4.1 Explicitly teach R

Research in computer science education suggests that explicit, direct programming education helps students learn. We believe that if you are going to have students use R in your course, you must incorporate teaching R into the class. There are many strategies for teaching R, including live-coding or providing example code.

The global coding education nonprofit, The Carpentries, recommends live coding. Live coding is when you begin with a blank screen, which your students can see, and generate all code "live." Live coding has many benefits. Among them, it slows the instructor down, because they need to type all the code they expect their students to be following along with. It also provides many teachable moments when the instructor makes a typo or other error and can demonstrate how to recover from errors, a crucial element of coding. If you are interested in instituting live coding in your class, The Carpentries has resources for how to institute it well. One recommendation is to come prepared with notes, either on paper or on a second screen, and try to stick to your "script." This may feel forced or unnatural at first, but the structure allows for a lot of learning to take place.

Another option is to give students example code, and have them read through the code and try to decipher it. If this is incorporated into class, it can be a productive strategy. If you are going to do this, we recommend reading code out loud.

In our StatPREP workshops, we often split the difference between providing example code and live-coding by preparing documents we distributed to participants at the beginning of a session. These documents are typically a mixture of code and explanatory text, with questions and blanks to be filled in. The blanks are filled in by workshop participants, either as they work together in small groups, or as they watch the instructor live-code. The documents provide a welcome safety blanket for both the instructor and the participants.

No matter which strategy you choose, we recommend explicit instruction. Like learning a second human language as an adult, learning a programming language is not something you can just "pick up." In order to learn it, you need instruction, explanation, structure, and practice.

5.4.2 Start with cake

As we discussed in the introduction, the idea of starting with cake is to begin with something interesting and achievable, so students quickly experience success. This idea comes from Mine Çetinkaya-Rundel's Data Science in a Box curriculum [31]. Many instructors use data visualization as their "cake," but depending on your student population, there could be other good introductory topics. Many people believe self-motivated learning based on a project of interest to the learner is the best way to go. The main point is, don't feel like you need to save the good stuff as a payoff for the end of the semester. Learning programming is tough, and it needs to feel worthwhile right off the bat to keep learners engaged.

Starting with cake is the opposite approach from how R was historically taught. Many R users were taught by instructors who believed it was important to understand the "fundamentals" of R before going on, and would teach the documentation page for R from start to finish [152]. One of the first topics in this approach is the concept of a vector, followed by object classes, arrays, and matrices. None of these topics are necessary for an introductory course in statistics or data science.

5.4.3 Be consistent

We have discussed the variety of valid R syntaxes above. There are arguments to be made for teaching one syntax over another. The authors of the mosaic package argue for the use of the formula syntax, and many educators argue for the use of tidyverse ([150]; [35]).

The research evaluating which syntax is better on an empirical basis is limited. So far there is little difference in error rates, although there may be some task-specific differences [157]. Each syntax has its benefits and drawbacks for teaching [123]. For example, formula syntax allows for an extremely consistent syntax throughout an entire semester-long introductory statistics course. The same class in tidyverse syntax involves a few more functions, and the syntax is not quite as consistent. But, the tidyverse syntax supports data wrangling, which the formula syntax does not. So for introductory data science classes, or introductory statistics classes in which many students go on to a second data or statistics course, tidyverse may be the better choice.

Most educators agree that the base R syntax is not the best one for novices, although there may be situations where it was appropriate.

No matter which syntax you choose, it is important to be as consistent as possible. The materials from McNamara are available online, and provide a full semester's introductory statistics course in each of the formula and tidyverse syntaxes, as a starting point [123]. We recommend being intentional about choosing syntax a priori, and developing as many instructional materials as possible before the course begins.

As an example of what we mean by consistency, see the following formula syntax code:

```
gf_point(bill_length_mm~flipper_length_mm, data = penguins)
cor(bill_length_mm~flipper_length_mm, data = penguins, use = "complete.obs")
lm(bill_length_mm~flipper_length_mm, data = penguins)
```

This code works with two quantitative variables: bill length and flipper length. It creates a data visualization of bill length and flipper length (a scatterplot), computes a summary statistic of the same two variables (using the "cor" function—for "correlation"), and then models the relationship (using the "lm" function—for "linear regression model"). For each of these tasks, the specification of the variables inside the parentheses is almost identical.

Similarly, consider

```
gf_boxplt(body_mass_g~species, data = penguins)
mean(body_mass_g~species, data = penguins, na.rm = TRUE)
aov(body_mass_g~species, data = penguins)
```

This code works with a quantitative variable and a categorical variable: body mass and penguin species. It creates a data visualization of the two variables (side by side boxplots), finds summary statistics (means for each group), and then models the relationship (using ANOVA). Again, the content inside the parenthesis is highly consistent, and the code looks similar to the code above.

The choice of R package also varies based on syntax choice. If you plan to teach in formula syntax, the main R package to learn is mosaic, authored by Randy Pruim, Danny Kaplan, and Nick Horton. If you want to teach tidyverse syntax, explore the tidyverse package, as well as infer.

5.4.4 Start small

There is no need to jump fully into the ecosystem of R, or to draw it through your entire course the first time you teach it. The little apps are designed to allow for some material to be taught without code, and to make the transition to code easier.

R really shines when it comes to data visualization, and that is the "cake" that typically gets students excited (see Subsection 5.4.2). So a first iteration of using R might be just doing a unit on data visualization in R.

5.4.5 Use Posit

R is a programming language, and can be used at the command line, or in an extremely basic graphical user interface (GUI) that comes with the installation of the language. However, it is much easier to use and teach within a supportive Integrated Development Environment (IDE). There are several choices for IDE, but we suggest Posit (Posit Team 2014). A screenshot of Posit is shown in Figure 5.5.

Posit is used by many professional R programmers, and the majority of R educators. It provides supportive features like tab-completion (so you don't have to type all code completely out), a data environment visualizer (so you can remember what datasets you have loaded), a dedicated plot window (to allow you to look back and forth between code and plots), and much more.

There are a number of ways to use Posit. The cheapest way is to use the Posit Desktop software, which students install on their own computers. The desktop application is free to download, and available for Mac, Windows, and Linux. The drawback to this approach is you will find yourself acting as tech services for your class of students. Most of the kinks have been ironed out of Posit over the years, but there are always one or two students with unique computer setups that require more specialized know-how in order to make it work.

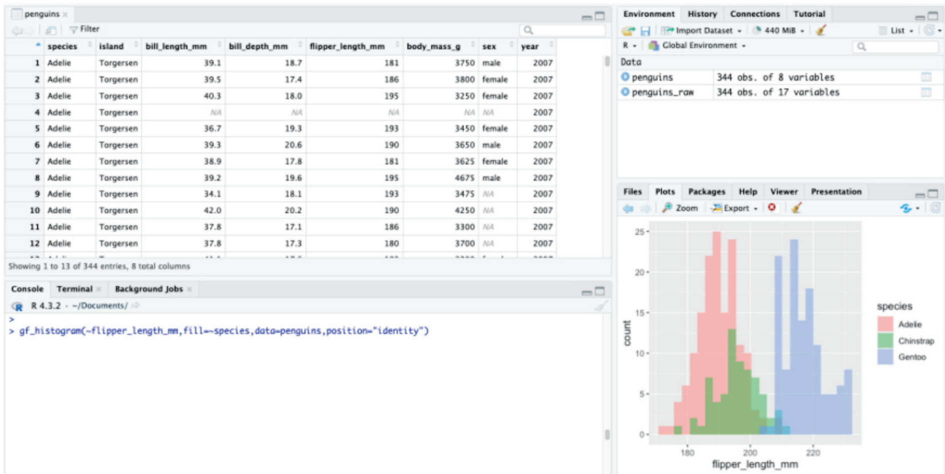


Figure 5.5: Screenshot of Posit, showing data preview pane, Environment, Console, and plot pane.

Alternatively, you could have students use Posit in their browser. In this scenario, students do not need to install anything onto their computer, and you can ensure everything is set up before the start of the term (you can even preload course documents!). But, browser-access versions of Posit cost money. If your institution has information technology support and server space, they might be willing to install Posit Workbench on a server for you. Server costs will need to be paid, but a few institutions have been willing to foot that bill. If you don't have support staff to manage a server for you, Posit Cloud could be a good option. Posit Cloud is hosted by the company Posit, so you don't have to manage the server yourself. But, it costs money once you go over a certain number of computer hours. You could have students pay for Posit Cloud or arrange to have your institution pay the costs (bulk discounts are offered).

5.4.6 Teach reproducible workflows

One benefit of teaching and using a programming language is that it supports reproducible workflows. If your students are going on to become scientists, they will certainly use the skills of reproducible research in their careers. But no matter their career path, having a good workflow is a skill that will serve them well.

Foundational computer scientist Don Knuth pioneered the idea of "literate programming," mixing human-readable text with code [109]. In the R world, this has been implemented in the document format RMarkdown, or its next-generation counterpart, Quarto ([164]; [5]). Throughout this section we will use RMarkdown as the terminology, because existing StatPREP materials have been prepared as RMarkdown, but almost everything will apply to Quarto as well.

RMarkdown documents allow you to write paragraphs of explanatory text (for example, describing statistical concepts, or explaining code), mixed with code "chunks." Text can be formatted using Markdown syntax (using asterisks to make words bold, and pound signs to indicate section headers). An RMarkdown file is plain text, so it can be opened in any text editor, but it works best when opened in Posit. Posit provides additional supportive features for authoring and editing RMarkdown documents.

Once you have a document with text and code you are happy with, you can "knit" or "render" it to an output format of your choosing. Common output formats are HTML, PDF, and Word documents. When Posit renders the document, it formats the text according to Markdown formatting, runs the R code in the code chunks, and inserts any output (graphs, summary statistics, model tables, etc) into the document. The result is a polished document that can be submitted as a homework assignment or final project. See Figure 5.6 to view the difference between an RMarkdown file and the knitted HTML that can be produced from that file.

Educators have been using RMarkdown in classes for years, including very introductory courses [14]. Students love RMarkdown, because the results look professional, and seem to come together "by magic." Teaching RMarkdown

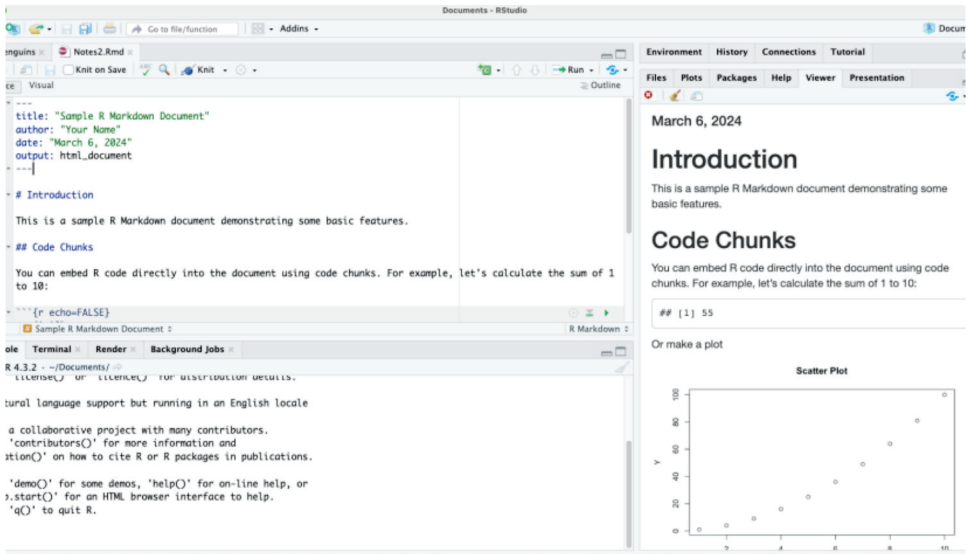


Figure 5.6: Screenshot of Posit showing an RMarkdown document (top left) with text mixed with code chunks, and a knitted HTML document (right) showing the formatted output, including plots.

provides benefits for the instructor as well because it ensures that all code submitted in a rendered assignment works. If Posit encounters an error when rendering the document, it will cause error and the student will have to debug their code. This reduces the errors in submitted work to conceptual errors, rather than mixing both coding and conceptual errors.

5.5 Challenges

As we have mentioned many times, learning a programming language is challenging, and R is no different. Some common challenges are addressed here.

5.5.1 Non-R-related

A number of the challenges that come with teaching R are actually not R-related, but more related to computer skills more generally. As The Verge reports, “Students who grew up with search engines might change STEM education forever” [41]. Computer programming relies on knowing where files “are.” That is, where they reside on your computer as it relates to the file directory structure. If you want to load an Excel spreadsheet into R, you need to know if that document is stored in the Downloads folder, the Documents folder, or some sub-sub folder buried deep in your personal organizational system.

Because Gen Z has grown up with search engines like Google, cloud storage systems like Google Drive, and computer search algorithms like Spotlight and Windows Search, they are often not familiar with where files “live.” It can be challenging to convince them to adapt an organizational structure for their files, even a folder for a class.

Using a version of Posit that students log into through their browser (like Posit Cloud) helps a bit, but students still typically need to retrieve files from your Learning Management System (LMS) and save them into their Posit Cloud account.

5.5.2 R-related

Challenges related to R typically relate to syntax and the opacity of some R errors. Like all programming languages, R does not understand commands that are even a single keystroke off. The difference between year and Year will throw an error, for example. Other common confusions are between the tilde (used in formula syntax) and a single dash or “minus sign.” Looking at $\sim$ and – on a screen, it can be hard to tell the difference. Similarly, R makes a distinction between a backtick (‘) and a single quotation mark (’). The list goes on.

Students can get frustrated if they cannot identify the source of an error, and it takes practice to see them. Error messages often do not help. For example, when students mix up the tilde and minus sign, they will get an error stating that one of their variables is not found, even though the variable name is spelled correctly.

One particularly ironic R error is **Error in plot.new() : figure margins too large**, which means that the plot window is too small to display the plot; that is, the figure is too large to be shown there.

Students default to using search engines to find answers, but it can be difficult to find answers to R-related questions online. One reason for this difficulty is the proliferation of syntaxes. If students search "how to make a scatterplot in R," they are likely to find results showing them how to accomplish that task in syntaxes they have never seen before. One solution for this problem is to encourage students to include the name of a package in their search query, like "how to make a scatterplot in R mosaic" or "how to make a scatterplot in R tidyverse."

5.5.3 Overcoming challenges

We do not believe any of these challenges are insurmountable. Utilizing StatPREP resources will help avoid some issues, and others are inherent to programming and working with computers. When I teach R, one of the first things I assure students is "it will get easier with time." I haven't found a way to circumvent the frustration of making syntax errors, or not understanding where files are. But since my assignments all start off the same way (open a template RMarkdown document and add code), have consistent syntax (so the functions become familiar over time), and are submitted the same way (knitting an RMarkdown document to HTML and uploading to our Learning Management System), the same errors tend to pop up repeatedly for the first few weeks. This means students get practice overcoming the same challenge repeatedly, and pretty quickly learn how to avoid it. "Oh, if I move my RMarkdown file to my class folder right away, I don't get that error!" This sort of trial-and-error learning and resilience in the face of obstacles are valuable experiences for students.

5.6 StatPREP and R

StatPREP has been focused on R since the beginning. We have developed resources for teaching R, or teaching with R (the little apps are written in R, even the ones in which you don't see any code). We have led repeated professional development workshops for faculty who wish to deepen their skills in working with data and teaching modern statistics courses. While this chapter has given you a way to start out on day one, this is not a solid enough knowledge base for teaching a full semester course. This is where the StatPREP resources come in.

5.6.1 Professional development

Overall, the StatPREP team practices what we preach in our own work and our professional development workshops. We use reproducible workflows in all our work, so workshop materials are developed in RMarkdown, and little apps are written in R. We teach within Posit. We provide explicit instruction in R, and try to "start with cake," by beginning with data visualization or something else we believe will motivate our participants.

The main place where we have deviated from our own recommendations are

1. We don't start small.
2. We're not consistent.

This is mostly due to the nature of professional development. If we have two days to teach 30 people the R required for a statistics course, we can't move slowly or start small. If anything, our workshops are a firehose of information. We spend an hour on something that would be spread out over days or weeks in a class.

When you look at materials provided by StatPREP, you will see an overwhelming consistency with the formula syntax. This is no coincidence. The authors of the mosaic and ggformula packages have all been involved in StatPREP to one degree or another.

However, participants in StatPREP professional development workshop were provided resources in both the formula and tidyverse syntaxes. This is because we believe it is important for an instructor to be somewhat familiar with a wider variety of commands.

5.7 Resources

I encourage you to explore the resources made available by the StatPREP team. One useful resource is the set of textbook companions. Two of the textbooks for which there are companion materials are open source and freely available online. This is in line with the ethos of R, which is open source software. Those books are

- OpenIntro Statistics by David Diez, Mine Çetinkaya-Rundel, and Christopher Barr [32]
- Statistics Using Technology by Kathryn Kozak, a StatPREP co-PI [112]

In addition, there are other resources available online that would allow you to dig more deeply into R. For example, the authors of the OpenIntro book have created lab materials to accompany the book, using both base R and tidyverse syntaxes [32].

If you use the Lock5 textbook, I have written lab materials to accompany that book, using both formula and tidyverse syntaxes [124]. These are just a few examples of R materials. Because R is open source software, many instructors choose to make their teaching materials freely available on the web.

In fact, this is another benefit to R as a programming language: the community. R is known for having an open, welcoming community. For example, the group R-ladies Global seeks to support women and minority genders in using R [168]. They host meetups in cities across the globe, likely including one not too far from you. There are also general R user groups that meet up in person and online communities like the R for Data Science Slack ([103]; [155]). While many programmers use the online site StackOverflow to ask and answer questions, R users realized the environment there was not welcoming for newcomers, so they created a separate website where even the most basic questions are welcomed: the Posit Community [147]. Finally, the StatPREP team is all very friendly and welcomes questions. So if you're stuck on something that is preventing you from using R or the StatPREP resources, I would recommend reaching out.

6

Small Changes: What *Can* We Do?

Megan Breit-Goodwin, *Anoka-Ramsey Community College*

6.1 Introduction

I was scared when I began teaching statistics. My teacher preparation and the entirety of my teaching career had lived inside the algebra through calculus pathway. Statistics was outside my area of expertise and comfort. Not only was I new to teaching statistics, I was relatively new to statistics itself! I had completed one statistics course as an undergraduate student, and my graduate program included two applied statistics courses. This paled in comparison to the volume of mathematics courses I completed as a student. The content outlined for my assigned introductory statistics course included material I was not confident that I understood, let alone felt prepared to teach.

In this chapter, I will share my own personal transformation in the teaching and learning of introductory statistics. My hope is that by sharing my journey, I will encourage others along their own paths, whether they are taking the first steps in a new direction, or thinking about deepening their current practices in new ways. A distinctive feature of this Notes volume is the inclusion of classroom activities that can be tried with students. Making small changes, often trying just one new activity at a time, helped me make profound shifts in my teaching and transformed introductory statistics for me and my students.

I began teaching statistics with the tools and strategies I used as an algebra teacher. The course was formula-heavy, computation-driven (by hand or by calculator), with content largely delayed until after the computations were complete. This approach seemed consistent with the learning outcomes articulated in the course outline.

In the first semester I taught introductory statistics, I learned a lot about the mathematics behind the statistics, but it didn't go well for my students. Assessments showed that my students were somewhat proficient in the computations I presented, but students did not demonstrate understanding nor meaningful interpretation of what they computed. For example, many students demonstrated proficiency in constructing a confidence interval of a mean, but struggled to interpret a confidence interval to make an inference about questions relating to the context of the data. I was noticing that students were approaching the content of the course as something that was solely procedural, and they were not engaging in either the contexts or the statistical thinking that were central to the course. This was a problem, and a challenge for me as their teacher.

I needed to learn new ways of working with data and teaching the content of the introductory statistics course. I wanted to explore what was possible for the course, and create a better experience for my students and myself. I was ready to leverage the energy of a beginner in order to grow.

6.2 First Steps Toward Change

My first approach to changing the way I taught statistics was to weave in more of the mathematics behind the computations students were using. This followed my own introduction to statistics as a student which was focused on the

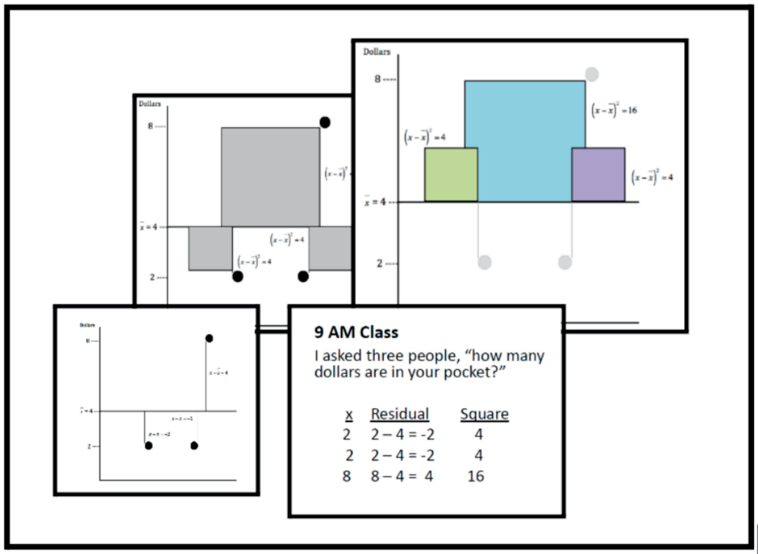


Figure 6.1: Lecture Slides

beautiful mathematics behind the methods. I tried to illustrate the mathematics through examples that were simple enough that students could “do” the math, tracing computations through tables and equations. I spent hours creating lecture slides that illustrated the mathematics using simple and contrived contexts such as counting dollars in a pocket, or the number of chocolate chips inside of a cookie.

These changes helped me learn more about the statistics, but it wasn’t met with student enthusiasm. I asked for student input on how the course was going. There were many neutral responses, but there was also significant and direct feedback. One student wrote, “Stop talking about the cookies, nobody cares about the cookies.” Students told me in their feedback that this attempt at mathematically-rich “bite size” (excuse the pun) examples was not working.

Next, I tried bringing in real studies from the news and from scientific and academic papers across different disciplines. A few students appeared to find this interesting, but again, it didn’t have the impact I was hoping for. My students continued to speak honestly when asked for feedback. There was an encouraging note from a student that said, “You are at your best when you are talking about the ‘real stuff,’ and we learn best when you are at your best.” That feedback gave me pause. It directly spoke to the connection between my own experience in the course, and my students’ experience. Students could see my own enthusiasm for statistics, and this was something to build upon.

I was engaging with the “real stuff” in my teaching and that had started to cause a shift in what my students were experiencing. However, I wasn’t actually engaging students in the “real stuff.” Instead I was introducing rules and tools that “other people” used. Understanding and interpreting outcomes and data is important, but what was missing was actually working with data in ways that would open opportunities for my students to do statistics.

I wanted to make changes in my teaching that would have the impacts that I wanted to see in student learning and outcomes. I needed to start by learning more myself, so that I could articulate and understand those changes and the impacts.

6.3 My Own Learning

The next phase of my learning began by familiarizing myself with the *College Guidelines for Assessment and Instruction in Statistics Education* (GAISE 2016) report.

The recommendations read like a list of things I was not yet doing as a statistics teacher. It was clear to me that my class, and my teaching, were not aligned to the 2016 GAISE College Report standards. That was a problem. I needed to better understand what these recommendations meant, and how to enact them in my classroom. I was particularly drawn to the idea of teaching statistical thinking, which is defined as “the type of thinking that statisticians use when approaching or solving statistical problems” (GAISE, p. 1) Crucially, integrating statistical thinking did not have to come at the expense of content; rather, this approach offers a way to authentically engage students in learning the content of the course.

I needed support to take the steps to make changes to my teaching. This is why I chose to participate in the Saint Paul, MN, StatPREP hub in 2017. I wanted to learn, and I knew StatPREP was an opportunity for me to learn in collaboration with other instructors that were teaching at colleges near me.

6.4 The StatPREP Experience

I entered my StatPREP experience with a lot of enthusiasm, and a bit of skepticism. The StatPREP project was aligned with the College GAISE standards and would help me develop the practices I knew were missing in my teaching. However, I was worried that it might not work the way it was intended in my classes and with my students. I had doubts about my own ability to teach data centered statistics. Did I know enough about working with data to teach this way? Could I manage the technology? What if students wanted to discuss the relevance of the context? Was I capable of giving up this control of my classroom? It felt like I was taking a risk, and that was scary for me.

The first StatPREP workshop allowed me to be a student again, which was exactly what I needed to become comfortable in trying new teaching methods. It was fun to play with data, create activities, and see the teaching modeled. There were challenging moments, frustrating moments, and exciting moments. The energy created through collaboration with other faculty while we explored data sets and thought about what kind of questions emerged from the data confirmed this was the direction I would take my teaching. We learned new ways to teach statistics through data that could motivate course content, lessons, or even whole units of study. We spent time in small groups working with data sets to develop activities that could be used in our courses. It was in the time spent just exploring the data and brainstorming ideas with my colleagues that made me realize that pushing my students to compute or construct things prior to exploration of the data was taking away the space that was needed for students to truly engage with the context and grow in their statistical thinking. The workshop motivated me and gave me enough confidence to take the next steps and begin making changes. I was equipped with tools and resources that I could use to try just one new thing at a time. StatPREP gave me a community of educators who were making changes along with me, and this helped me feel more confident as I started to make plans for the next semester's statistics course.

6.5 Small Changes and a Moment of Panic

Knowing my propensity to dive in and overwhelm myself, I intended to make only small changes in my teaching the semester after my first StatPREP workshop. I committed to making changes to the first unit of the course, which focuses on an introduction to data and descriptive statistics. This is the unit I was most comfortable teaching; I wanted to start the semester strong, and I needed to ensure I had the energy to make changes.

I knew it was important for me to limit the scope of the technology I was going to learn to teach with, so I decided to leverage a pedagogical statistical software that was included with the text. I selected a single data set—namely, a subset of 2009-2010 National Health and Nutrition Examination Survey (NHANES) data—and then spent a lot of time preparing to teach with it.

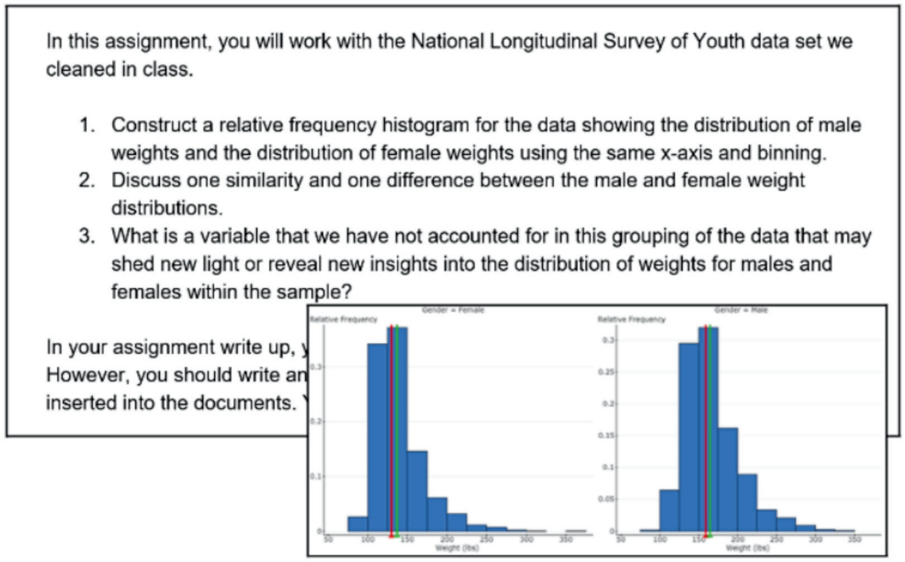
I was nervous for the first day with my students using the data set, which caused me to over prepare. I knew how to use the software to do anything I could imagine students would want to do with the data. I practiced, practiced again, and practiced again. I was ready.

Then came the first day in class.

The anticipated challenges of getting everyone onto a computer, into the program, and connected with the data set went okay. After I led students through the creation of their first histogram and boxplot, I asked them to take a few minutes to play with the data, encouraging them to make pictures of the distributions. Then, the inevitable happened. A student asked how we would look at just the 18-22 year olds in the data set.

I froze. I didn't know how to do this. I hadn't practiced this. The sweat began to pool, and my heart was racing. I was unprepared, this wasn't possible, and I was ready to go back to the way things were. It was likely only 10 seconds that I stood frozen at the front of the class, but it felt like 10 minutes. But then a different student said, "Oh! Let's try this..." I stepped aside and handed my computer over to a student to present.

I literally got out of my own way, and my students' way. I was no longer in the role of "the expert"; my students were now leading the way. The thing I feared most happened on day one, and we all made it through. It was good.



In this assignment, you will work with the National Longitudinal Survey of Youth data set we cleaned in class.

1. Construct a relative frequency histogram for the data showing the distribution of male weights and the distribution of female weights using the same x-axis and binning.
2. Discuss one similarity and one difference between the male and female weight distributions.
3. What is a variable that we have not accounted for in this grouping of the data that may shed new light or reveal new insights into the distribution of weights for males and females within the sample?

In your assignment write up, y
However, you should write an
inserted into the documents. Y

Figure 6.2: Assignments and Histograms

6.6 Opening Opportunities for Statistical Thinking

Students completed an assignment the second week of that semester in which they discussed similarities and differences between male and female weight distributions as observed in relative frequency histograms constructed from the NHANES data set. Student work on the assignment demonstrated proficiency in describing a similarity of shape between the distributions with language like "the graphs are skewed right." While this is accurate and an important observation, it fell short of a rich and robust discussion of the data that was represented. I wanted my students to connect with the units of analysis (teenagers), and the variable measured (weight) to structure their discussion and descriptions.

I adjusted the instructions to create space for students to work with the same data set, but look at the distributions of female and male heights. Students were tasked to collaboratively develop descriptions that addressed the similarities and differences between the distributions. The discussion during the activity was rich, supported by their use of technology and flexibly looking at graphical representations of the distributions. I did more listening than speaking, often just prompting students to "tell me more" or asking, "what does that mean?" Together, the students were able to create a rich comparison between the two distributions, noting important differences between the groups and using language that engaged the context of teenagers and heights. Students began to question if the differences they were observing were really meaningful differences. When students worked with another pair of distributions on their next assignment, they demonstrated a notable improvement in their explanation.

The experiences from that first unit created momentum and increased my confidence to make more changes. It was clear to me that my students' experiences in the course were changing, and so was my experience as a teacher. I was beginning to move away from teaching statistics, and beginning to teach students.

6.7 Small Changes that Transformed the Course

I continued making changes one lesson, one assignment, and sometimes one question at a time that first semester. I tried new things, with new confidence, even when I wasn't sure how it was going to look in action. I was transparent with my students that I was learning to see and teach statistics differently, and they were generous with me when class sometimes went in unexpected directions. Technology became an asset, and my fears about using technology in the classroom dissipated. I trusted that my students and I could figure out how to use the technology (usually in the moment). When things didn't work out perfectly, we became comfortable with that being part of the process and the messiness of what we were doing together.

Data-centered lessons allowed my students to enter course content through exploration, and the questions that emerged through this exploration were rich. We spent time building the background context of the data we were working with, and discussing the measures or concepts we were computing or examining. This wasn't something we rushed through. Connecting to the realities, the humanity, and the stories of what we measured and explored was central to our statistical thinking. For example, when working with the NHANES data set, students discussed limitations of the exploration due to what wasn't included. Conversations about how race and ethnicity is and is not measured, and how age and generation intersects with measures of race in ways that may impact health measures. The implications of this complexity for medical practitioners in their interactions with patients also became a topic of interest for my students, many of whom where going into allied health fields such as nursing. My students and I were becoming comfortable in these spaces together.

It became clear that if I was to make changes to the teaching and learning in the class, I also needed to change the way I was measuring student learning, understanding, and practices. The prior homework and testing practices I used were not aligned to what I now valued and the experiences that students were having in class. I could no longer ask them a list of de-contextualized questions designed to capture a computation around a specific concept or skill. I brought in more data sets, and students engaged with them across several weeks of assignments and assessments. This allowed me to assess the ways students interacted with data and move that forward in their thinking, to iteratively ask and answer new questions. This approach had the added benefit of challenging students to think across contexts while measuring their understanding of the course learning outcomes. Changing assessment methods allowed me to deepen the ways I was listening to and understanding my students' statistical thinking, and helped me become more flexible in my teaching.

The semester I taught statistics following my first StatPREP experience was the first time I utilized an inquiry-oriented classroom. It was also the first time I felt shared ownership of the teaching and learning in the class between myself and my students. For statistical thinking to be central to the curriculum and teaching and learning in the course, my students needed to know that they belonged in that space and could take ownership of the space. If I could wish one experience on all educators it is this: a few moments in class where students are completely absorbed in the material, listening to the ways they bring each other deeper into the content.

6.8 Reflection Motivates the First Step

Making changes to my statistics teaching was scary at first. What made it possible for me to take the first steps towards change was reflection on the teaching and learning that were happening in my classes. It was important for me to slow down and think about what was going well, and what wasn't going well. This helped me identify an opportunity for my own development as a teacher, and encouraged me to take action.

I like to start small, because small changes have always had profound impacts on me and my students. This might look like trying one of the ideas or activities shared in this book. I personally found the StatPREP Little Apps to be fun and welcoming resources to use with students. The first time I used a Little App activity, I introduced it by saying, "We are going to just try something today. I'm new to this, so I don't know what is going to happen, but let's try it together!" When you try something new, think about what surprised you about your students' experience. What surprised you about your own experience? What do you want to learn more about?

Before I made changes to my statistics teaching, lesson planning was not fun. There was so much content to cover, and the disconnect between what I was preparing and what my students were (and were not) learning was frustrating. When I began making small changes, lesson planning was sometimes scary for me. I was not secure in my ability to teach in these new ways, and I tried to prepare for all possible outcomes to what may unfold in the classroom. However, the ways these new approaches actually unfolded in my classroom—and the way the classroom environment shifted—helped to increase my own confidence to try new things. I now have implemented StatPREP- and College GAISE-informed methods across all learning units in my introductory statistics class. I did this by taking small steps in my teaching in collaboration with my students, and with the support and encouragement of other educators.

Lesson planning is now fun. One night I was up much too late, playing with a data set that I wanted to bring into class. My husband saw what I was doing and asked, "Are you working, or are you playing?" The question captures how teaching statistics feels to me: it is playful work. Now, my lesson planning centers on identifying a few key ideas for a class session, and how to create openness and space for students to generate the questions and guide the inquiry that will bring us to those ideas. That is where the learning happens: my students learning, and my learning.

I am not done making changes to my teaching. Teaching statistics continues to challenge me to be a student of my students. I continue to make changes one step at a time, and look forward to growing that way. This text has great examples of the impacts of R in the classroom, including bite-sized things I look forward to trying with my students.

Making small changes, and being open with my students about my own learning, helped me become more comfortable taking risks and engaging with my students' statistical thinking. Giving myself the freedom to make small changes prevented me from becoming overwhelmed. The excitement, challenge, and outcomes of each change motivated me to take even more steps as we progressed. The dynamics of an inquiry-oriented classroom require vulnerability and authenticity from students and myself. This was something I was scared of, but didn't know I feared until I began this journey. And this is a joyful outcome to have realized. It allows me to be more present, more responsive, and open to the adventure of what unfolds in my classroom. I know how I will continue to learn and grow: one small change at a time.

Part II

StatPREP in the Classroom

Part II Overview

Kathryn Kozak, *Coconino Community College*
Ambika Silva, *College of the Canyons*

The StatPREP project focused on bringing a modern, data-centric approach to the introductory statistics classroom. The chapters in Part II StatPREP in the Classroom highlight some of the tangible products that were developed through this project. By focusing on the Little Apps and related classroom activities, these chapters demonstrate a modern approach to some Introductory Statistics topics. The StatPREP approach emphasized conceptual understanding over the traditional computational approach.

In 2005 the American Statistical Association (ASA) endorsed the Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report [67]. The six recommendations put forward in the report stood the test of time and were reaffirmed in the revised 2016 GAISE College Report. To reflect the changing landscape of the practice of statistics, and looking to the future of statistics education, the 2016 revision simplified the language of some of the recommendations and added two new emphases. Factors that influenced the direction of the report included the increase in the number of students studying statistics, the new emphasis on statistical concepts and methods in K–12 state standards, the increase in available data and technology, and the emerging discipline of data science. The revisions were motivated to first focus on what we teach and then on how to teach introductory statistics courses.

The revised recommendations are

1. Teach statistical thinking.
   (a) Teach statistics as an investigative process of problem-solving and decision-making.
   (b) Give students experience with multivariable thinking.
2. Focus on conceptual understanding.
3. Integrate real data with a context and purpose.
4. Foster active learning.
5. Use technology to explore concepts and analyze data.
6. Use assessments to improve and evaluate student learning.

The 2016 GAISE College Report recognized there was no single introductory statistics course; courses are designed to fulfill the needs of many different student audiences and as such have different prerequisites. The recommendations were developed to be applicable to the variety of introductory courses allowing that the specifics of implementation would vary but guide the instructional practice. The overarching goal was to ensure students are engaged in active learning, develop the ability to think statistically and use the appropriate technology to explore concepts and analyze data. It is worth noting that the practice of statistics and data science continues to evolve and work is underway on a revision of the 2016 GAISE College Report. This revision is anticipated to support existing recommendations.

As acknowledged in the preceding paragraph, there is much variation in introductory statistics courses. Faculty use a diverse range of techniques for teaching undergraduate statistics courses: some use tables and calculators while others are using statistical software provided by a textbook publisher or computer languages like Python or R. To encourage a more modern approach, one of the focuses was on statistical computing using R. These initial efforts were supported by the creation of tutorials in R. When reflecting on the first year of StatPREP, the leadership team realized that scaffolding was needed to support instructors who were new to statistical computing. In response, the StatPREP team created tools and activities to provide a gentle introduction to the data-centric approach. These tools, called Little Apps, facilitate a data-centric approach to teaching Introductory Statistics without the cognitive load of learning R. To bring active learning into the classroom, activities were created that utilize and complement the Little Apps tools. Little Apps activities enabled instructors to develop lessons that answered the question, “What can I do now?” supporting small changes that allow students to explore statistics concepts.

The chapters of Part II can be considered a “how to” manual for making small changes to an introductory statistics course. Chapter 7, written by Kelly McConville, provides context for the recommendation, “give students experience with multivariable thinking.” This is followed by chapters showcasing the Little Apps that naturally lend themselves to exploration with several explanatory variables.

While all six Little Apps provide a comprehensive toolbox for teaching statistics, we focus on three: Point and Density, Confidence and T, and Regression Models. Authors Carol Howald, Dustin Silva, and Helen Burn provide explanations and examples of how to use these Little Apps to teach these critical concepts. The specific lesson ideas include detailed activities that are carefully explained. The Little Apps are immediately accessible to students and instructors via any web browser, tablet, or smartphone. They allow a student to use interesting data, with user-controlled professional graphical displays at the center of the lesson.

In addition to the Little App activities described in this Part, there are other activities that use these Little Apps. These activities are available in the companion Library in the MAA Connect StatPREP Community. Also available are Center and Spread, Resampling, and Stratification and Confounding Little Apps. Although they do not have associated activities, they are very useful in class discussions and student exploration.

In the final chapter of Part II, Joe Roith and Maria Tackett address how probability fits within the framework of a modern approach to teaching statistics. They discuss obstacles and challenges in teaching probability and suggest ways to use modern technology as a vehicle for emphasizing concepts over algebraic formulas.

After reading this Part, the answer to the question “What can I do now?” will be clear. Small changes can make a big difference and may lead to larger institutional changes both locally and globally.

Using the Little Apps

Before you read the chapters in this Part, we suggest you take time to become familiar with the design features of the Little Apps. Figure II.1 shows the common features. As was discussed in Part I, each of the Little Apps focuses on

Figure II.1: Overview of features common to all Little Apps.

a specific statistical methodology or technique, but all Little Apps share common design features. We have created a short introductory video to help you get started.

Using the Little Apps provides you and your students easy access to interesting data sets, and the ability to select variables from the data set to encourage statistical thinking. These are shown in the figure as dropdown options - Data set, Response, and Explanatory.

The top menu has entries Data, Graph, Compare, and Stats, which allow the user to navigate to different views. For example, Stats displays a numerical statistical summary of the data. The information icon (a white letter i in a blue box), which is a part of the sub menu, provides access to a pop-up window that describes all the features as well as the icons. Since these features are consistent across the Little Apps, students can easily transfer their knowledge and quickly be able to utilize a new Little App.

The other options available are selecting the sample size (n), generating a new random sample (image of two dice), the ability to freeze the current sample so it can used with compare and displayed in the Stats window (image of a snowflake), controls specific the particular methodology or technique (image of three parallel lines), the information icon (the letter i), the codebook which provides a description of the variables in the data set (image on an eye), the option to bookmark and/or download a sample of the data (image is a bookmark).

You will read more about Points and Densities, Confidence and T, and Regression Modeling, in the chapters that make up Part II. The other Little Apps are Center and Spread, Resampling, and Stratification and Confounding.

7

Multivariate Thinking

Kelly McConville, *Harvard University*

7.1 Introduction

An overarching goal of introductory statistics is to teach students how to extract knowledge from data. These data rarely represent a single variable acting independently, so teaching students to learn from data requires teaching them how to simultaneously consider many variables and their various associations. In essence, extracting knowledge from data requires multivariate thinking.

Let's consider a hypothetical. Suppose a company asks one of our introductory statistics students to help them answer the question, "Do the men at this company make more money than the women?" The student's first inclination might be to find the difference in average incomes between the men and the women and to see how far that number is from zero. This, along with faceted histograms of the income distributions by gender, is a sensible first step, but should their analyses stop, only ever focusing on these two variables? How can we train our students to broaden their analyses so that they also consider and account for the impacts of additional variables, such as experience, education, or job type? A multivariate approach would provide a richer and more definitive answer to the pay equity question posed.

Loosely speaking, we are exhibiting multivariate thinking if we consider more than one variable when answering a question with data. This means that bivariate analyses, including inference for a difference in means, a difference in proportions, or a correlation, are multivariate. However, as suggested in the previous paragraph, we'd recommend also moving beyond the bivariate and providing students with opportunities to grapple with three or more variables at once. Give your students practice asking:

- What variables relate to my data question? Which of those do I have access to?
- How should the variables be incorporated into my analyses?
- How does accounting for different variables and their relationships impact my conclusions?

In this chapter, we provide ideas and examples on how to infuse multivariate thinking throughout the introductory statistics course. We focus our examples on connecting multivariate thinking to different aspects of the data analysis process. Through repeated exposure and practice, our students can learn to harness the richness of their data to answer real, relevant questions about our complex world.

7.2 Where to Begin

Statistical paradoxes are a great way to pique student interest in the course material and increase their engagement ([128]; [180]). Simpson's Paradox serves as the prime example for the importance of multivariate thinking; the paradox represents the situation where the relationship between two variables *flips* when we account for a third variable.

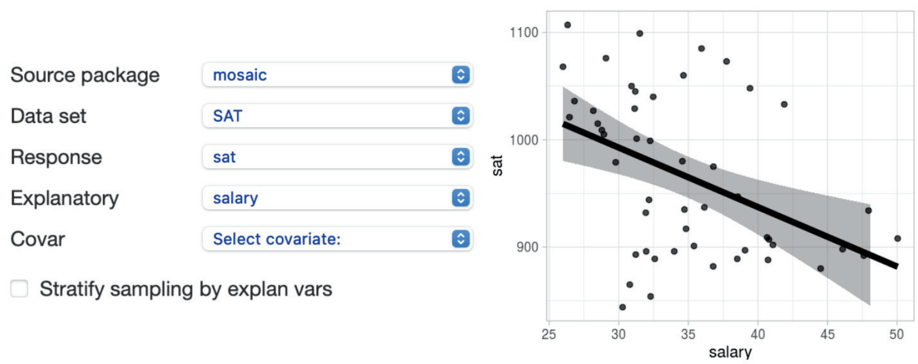


Figure 7.1: Stratification Little App explores the relationship between statewide average SAT scores and average teacher salaries.

Indeed the knowledge we extract from the data leads to a very different conclusion once we factor in additional information! In our own teaching, we often have our students wrestle with a real-world example of Simpson's Paradox on the first day of the term, which gets them practicing multivariate thinking right from the start.

Luckily the world is rich in interesting examples of Simpson's Paradox, and even better, many statistics and data science educators have written about how to bring these examples into the classroom ([210]; [172] ; [188]). If you'd like to present a social justice example, Jeff Witmer explores Stand Your Ground data where white defendants are more likely to be convicted than racial and ethnic minority defendants until you control for the race of the victim [210]. Then, regardless of the race of the victim, the racial and ethnic minority defendants are more likely to be convicted. Simpson's Paradox Worksheet offers an extended exercise for students. For another timely example, von Kugelgen, et al., found Simpson's Paradox in data on COVID-19 fatalities early on in the COVID-19 pandemic [199]. Comparing Italy and China, early data exhibited a higher case fatality rate overall in Italy but a lower rate within every age group. We would note though that with this and other pandemic examples, it is important to take a "trauma-informed" approach where you reduce the chance that students who have been directly impacted by the pandemic are re-traumatized by the discussion [187].

The Stratification Little App can also be used to illustrate Simpson's Paradox as it includes an example given in the 2016 GAISE College Report [67]. The College GAISE example explores the relationship between statewide average student SAT scores and average teacher salaries in the mid-1990s. Using the Stratification Little App and the options showcased in Figure 7.1, we obtain a scatterplot that shows a negative linear relationship between teacher salary and SAT scores. As the average teacher salary increases, the average SAT score for the state decreases linearly. At this point, we recommend asking the students:

- Should we conclude that decreasing teacher pay will lead to higher SAT scores?
- What other factors might influence or relate to SAT scores?

As 2016 GAISE College Report points out, student participation in the SAT varies greatly by state. In states with low participation, it is generally only the students considering out-of-state colleges who take the exam and these tend to be Midwestern states with lower teacher salaries. In Figure 7.2, we incorporate student participation by breaking the data into 4 groups or strata based on the quartiles of the variable measuring the fraction of all eligible students taking the SAT. And now we see Simpson's Paradox in action! After controlling for student participation, average teacher salary and average SAT score now exhibit either no relationship or a positive linear relationship. The 2016 GAISE College Report report suggests that "simple methods such as stratification can allow students to think beyond

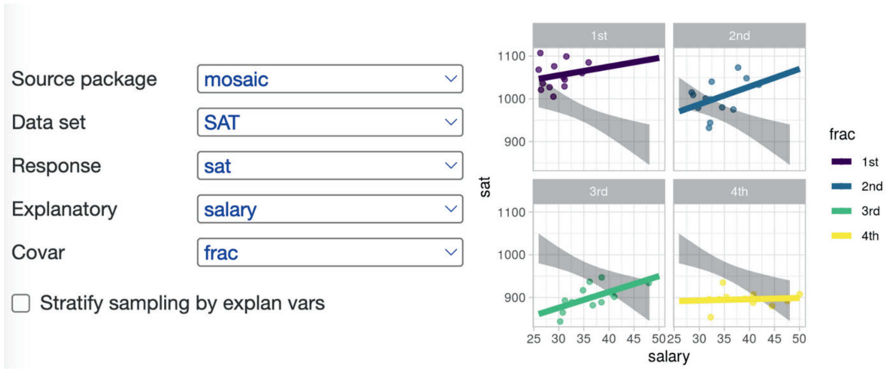


Figure 7.2: After incorporating student participation into the graph of teacher salaries and SAT scores, we see no relationship or a positive linear relationship across the four levels of participation.

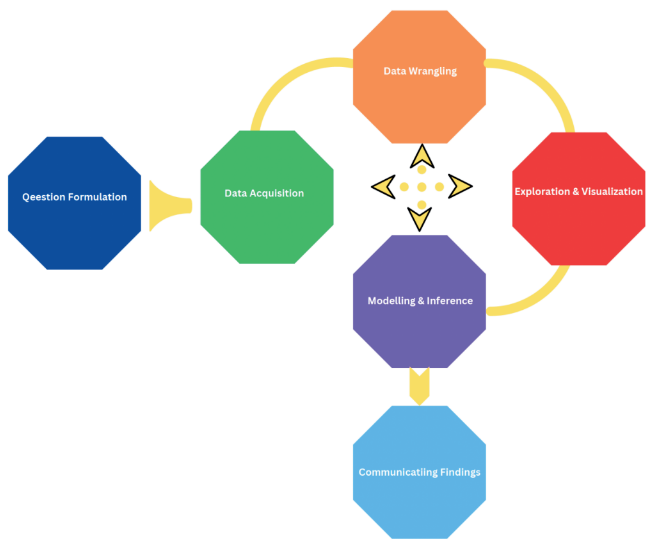


Figure 7.3: A diagram of the data analysis process.

two dimensions and reveal effects of confounding variables. Introducing this thought process early on helps students easily transition to analyses involving multiple explanatory variables."

7.3 Mutivariate Thinking Throughout the Data Analysis Process

In this section, we discuss ways to infuse multivariate thinking into an introductory statistics course by focusing on concrete examples related to steps of the data analysis process (as illustrated in Figure 7.3). By connecting multivariate thinking to several course topics, students get repeated exposure and hopefully become more comfortable disentangling the relationships in their data.

7.3.1 Data Acquisition

How we collect data impacts what we can conclude: to which population we can generalize our results, and whether we can claim a causal link between variables. Also, when teaching about data collection, we often stress the impact of two forms of randomness, random sampling and random assignment. Selecting a random sample from the population ensures that the sample is representative, while the random assignment of the treatment variable to the study participants allows causal links to be studied between the treatment and response variables. Both random sampling and random assignment can be enhanced by multivariate thinking through the concepts of stratification and blocking.

Stratified sampling, a type of random sampling, occurs when the population is broken into subgroups called strata and a random sample is taken within each stratum. A benefit of this approach is that the researchers control the sample sizes for each stratum and so can increase the precision of estimates in the strata. Strata are chosen to align with the study objectives and multivariate thinking helps the researcher decide which variables should be used to construct the strata. The 2011 - 2012 National Health and Nutrition Examination Survey was stratified by geography and population demographics to achieve large enough sample sizes for Hispanic people, non-Hispanic Black and Asian people, older adults, and low-income people [30]. For another example, Houghton, et al., stratified by sex and age in their study of Vitamin D levels in Thai schoolchildren so that they could explore if the relationship between age and Vitamin D levels varies by sex [98].

Just as we might stratify before randomly sampling, in experiments we might consider blocking before randomly assigning. With blocking, the subjects are first placed into blocks based on additional variables, and then within each block, the subjects are randomly assigned to the treatment groups. For example, if we are looking at the effectiveness of a vaccine or a placebo for a particular virus but worry that its effectiveness may be lower for smokers, we could first block by smoking status and then within each block, we would randomly assign half of the subjects to get the vaccine and half to get the placebo. Picking effective blocking variables requires thinking of potentially confounding variables, which is an exercise in multivariate thinking!

Beyond just how we collect data, what we collect data on also greatly influences our conclusions. For example, in the hypothetical pay discrimination study, we can only control for the variables to which we have access. In introductory statistics courses, most of the data students are exposed to were collected by someone else, possibly to answer a different question than the one the class is considering. This is why instead of titling the section "Data Collection," we used the more passive term "acquisition," signaling that this included both data students directly collected and found data. When working with found data, it is still important for students to think through how the data were collected but also to consider what other variables it would be useful to have. And, in a course that covers computing, this could mean learning to join several found datasets together to increase the information known about the sampled units. For example, if we want to understand what impacts the total daily check-outs for a community bike share program, we may want to merge these data with daily weather data or a dataset of holidays, as the variables in these other datasets may help explain usage patterns.

By teaching students to think carefully and expansively about both how they acquire their data and what data they collect, they will be able to ask and answer richer questions later in the data analysis process.

7.3.2 Data Wrangling

Data wrangling, which may also go by the names "data cleaning," "data transforming," and "data munging," is a very important step of the data analysis process that is sometimes overlooked in introductory statistics courses. It represents the process of transforming the raw data into a format ready for analysis. Without this skill, students will only be able to visualize, summarize, and model squeaky-clean datasets, which are common in textbooks but rare in the real world.

Data wrangling requires the user to make several subjective choices with the data. To name a few, these include determining which categories to collapse together for a categorical variable with many, sparsely populated categories, if and how the data should be subsetted down to a smaller set of more relevant observations, and how to handle item-level missing values in the dataset. These choices can have profound implications further downstream in the analysis, especially when estimating the relationships between variables.

To make this point more concrete, let's focus on the handling of missing values in a dataset. Computationally, missing values are often a source of frustration because one "NA" can cause code to error out (not execute). So, when we encounter an "NA," what should we do with it and why is this discussion important to a chapter on "multivariate thinking"? In other words, how does our handling of an "NA" for one variable, have any impact on the other variables in our dataset?

One of the datasets contained in the Little Apps comes from the National Health and Nutrition Examination Survey (NHANES), which is collected by the US National Center for Health Statistics Centers for Disease Control and Prevention (CDC) [30]. There is also a package called "NHANES" that was written for the statistical software package R ([149]; [153]). The "NHANES" R package contains a similar dataset and focuses on data from 2009-2012, which have been resampled to mimic a simple random sample of "the non-institutionalized civilian resident population of the United States" [149].

Let's suppose we want to estimate the average height and the correlation between the weight and height of the US population. In the NHANES dataset, there is a Weight (in kg) variable but seemingly two potential height variables: Height (Standing height in cm) and Length (Recumbent length in cm). Some of the variables in the NHANES dataset have missing values and so let's consider a few common strategies for handling NAs.

- Complete cases approach: Remove any row with a missing value so that you have a dataset with no missing values.
- Keep all approach: Keep all rows and only remove a case with a missing value when performing a specific analysis operation.
- Subset of complete cases approach: Remove any row with a missing value for the subset of variables (Weight, Height, Length) used directly in the analysis.

	Complete Cases	**Keep All**	**Subset**
Average Height (in cm)	NA	161.9	95.0
Average Length (in cm)	NA	85.0	96.0
Correlation of Height and Weight	NA	0.7	0.7
Average Age (in years)	NA	36.7	2.5

Table 7.1: Table of the mean values of variables Height, Length, Age, and the correlation between Height and Weight under the three strategies for handling missing values.

Table 7.1 summarizes the mean value of variables Height, Length, Age, and the correlation between Height and Weight under the three strategies for handling missing values. All of the estimates are NA for the complete cases approach because each row of the dataset had at least one missing value and so every row was removed under this strategy. Under the other two strategies we get wildly different means for Height and Length which can be explained when we consider another variable Age. It turns out the Length variable was only measured on participants who were three years old or younger. This means the subset approach dropped anyone older than three from the dataset. Interestingly, although the estimated correlation coefficient of weight and height is roughly the same for two of the strategies, we see in Figure 7.4 that only one of the datasets actually exhibits a linear trend! By coloring the point based on age, we see both the commonalities of the datasets and how age impacts the relationship between height and weight.

7.3.3 Data Visualization

If the world is now awash in data, it is also awash in beautiful, but complicated, visualizations of that data. Visit the website of any major news source and you will find articles containing engaging, interactive displays of the data. These graphs often include three or more variables and visualize complex relationships. By giving students practice interpreting similarly difficult graphs in introductory statistics, we are helping prepare them to be effective citizen statisticians.

The "What's Going on in This Graph?" series, a collaboration with the American Statistical Association and the New York Times Learning Network, provides not only a large array of graphs to explore but also a structure for that exploration. For each graph covered in the series, they ask the reader to consider the following four questions and one task: "What do you notice? What do you wonder? How does this relate to you and your community? What's going on in this graph? Create a catchy headline that captures the graph's main idea." For example, Figure 7.5 from the article

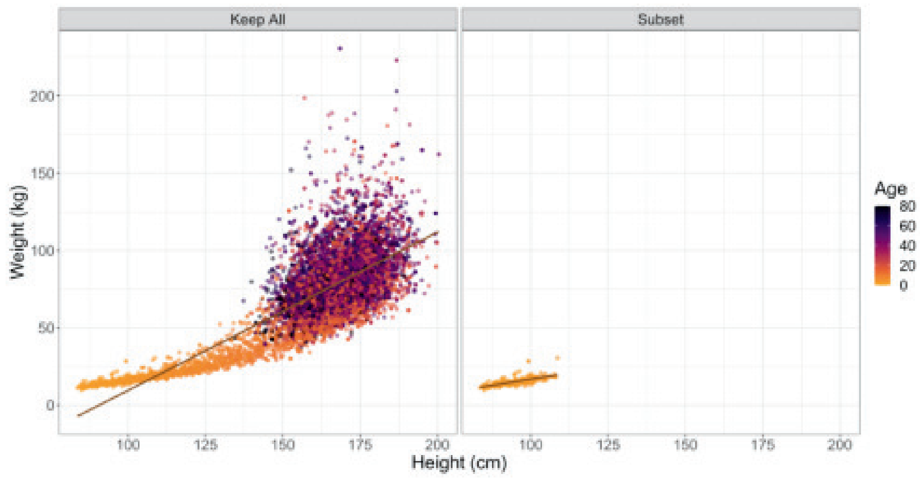


Figure 7.4: A comparison of the Keep All and Subset of Complete Cases approaches.

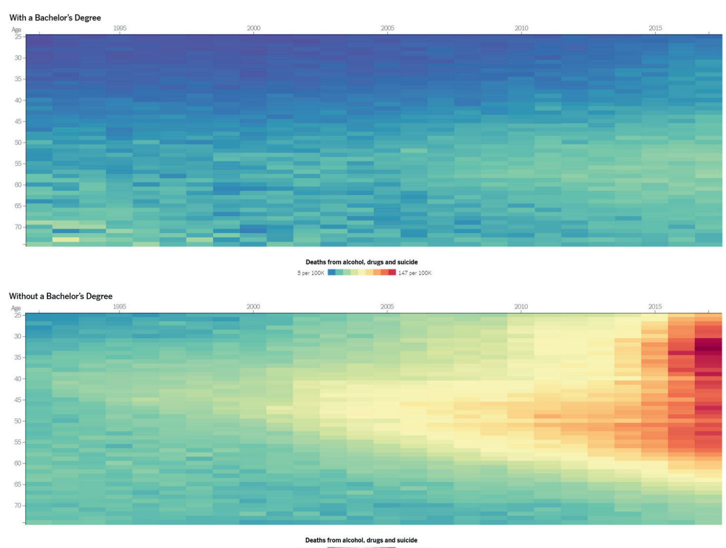


Figure 7.5: Figure from the article ”Teach About Inequality with These 28 New York Times Graphs.”

“Teach About Inequality With These 28 New York Times Graphs” considers the relationship between education level and rates of death of non-Hispanic white people from alcohol, drugs, and suicide using a heatmap [75]. To tell a fuller story, the graph incorporates two more variables into the mix: age and year. After asking your students to answer the questions posed by the series editor, you could consider asking them to explain the interconnected story of age, year, and education on death rates by alcohol, drugs, and suicide. For those without a college degree, we see that over the past three decades, the rates of death have greatly increased, especially for those under 50, whereas the death rates for those with a degree have only increased slightly and remain fairly low for younger people. Also, ask your students to explain what would be lost if we removed one of the variables from the visualization or what might be gained if we incorporated another variable.

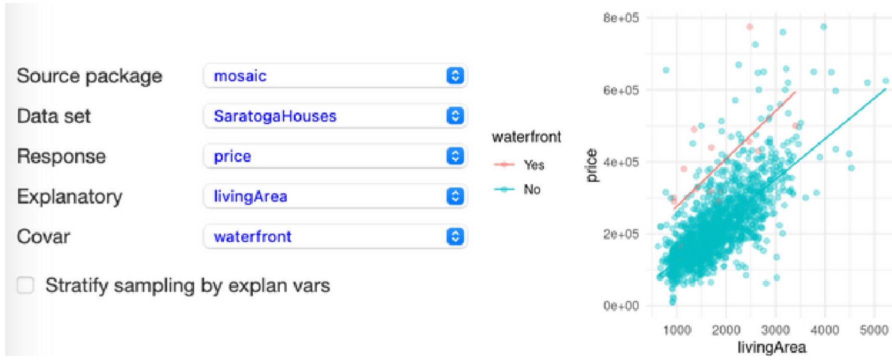


Figure 7.6: Using the Regression Little App to predict the price of a house using the size of the living area and whether or not the house is on the waterfront.

7.3.4 Modeling

The statistical models of today are often dizzying displays of multivariate thinking. Sometimes hundreds of variables are employed to build the most predictive model possible and multiple courses are needed to prepare students to deploy the wide range of available methods, including linear models, hierarchical models, machine learning models, etc. So what place does modeling have in the introductory statistics course? College GAISE argues that students should be exposed to multiple linear regression models with a focus more on how the variables interact and less on the mechanics and math. Here we provide an example based on the ASA's Stats 101 toolkit using data collected on house prices in Saratoga Springs, New York in 2006.

Let's explore the question, "Does the relationship between the size and price of a house vary by whether or not the house is on the waterfront?" Using the Regression Little App, we can explore this question visually as shown in Figure 7.6. Notice the two parallel estimated regression lines: one for houses on the waterfront and one for houses not on the waterfront. This means that the relationship between price and size does not vary by whether or not the house is on the waterfront. In other words, controlling for the size of the house, whether it is small, medium, or large, the difference in price between the waterfront and non-waterfront property stays roughly the same. To estimate that difference, we go to the "Stats" tab of the Little App after first setting the model parameters using the "App controls" button as shown in Figure 7.7. The third row of the regression table states that the estimate for the variable "waterfrontNo" is -165000. This means that for a given house size, a non-waterfront house will cost about \$165,000 less than a waterfront house.

In contrast, the relationship between the size and price of a house differs by whether or not the house has central air conditioning, as seen in Figure 7.8. The slope of the relationship is steeper for houses with central air than for houses without. We can find the estimated slopes for the two lines from the regression table, shown in Figure 7.9.[1] For houses with central air, we expect the price of the house to increase by \$132 for every 1 square foot increase in the living area, whereas, for those without central air, we expect the price of the house to only increase by $\$132 - \$44.6 = \$87.4$ dollars for every 1 square foot increase in the living area. "How much is a Fireplace Worth?" explores another regression example with these data to see how the price of a fireplace changes depending on what other variables are included in the model [53]. All of these examples get students to see how conclusions change depending on the variables we include in our analyses–a key aspect of multivariate thinking. And, by leading with graphs, students can develop intuition before diving into the model equation and interpretations. For more examples of teaching modeling in introductory statistics, see [144] and [211].

[1] Note that the Regression Little App by default fits the model with the "Include interaction" box checked under "App Controls."

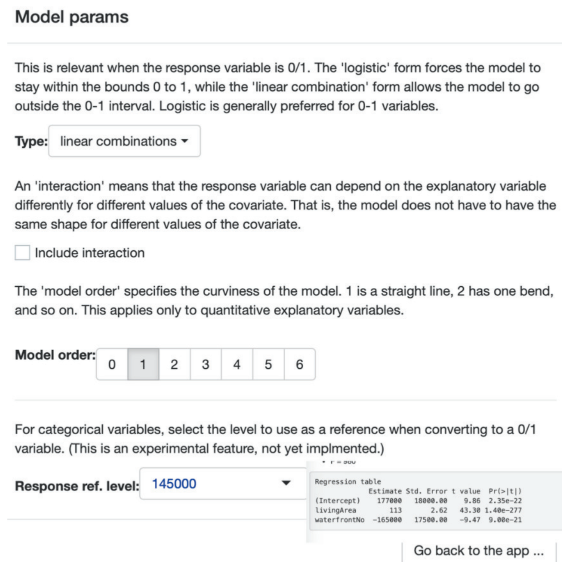


Figure 7.7: Using the Regression Little App to estimate the difference in prices between waterfront and non-waterfront houses after controlling for the size of the living area.

Figure 7.8: Using the Regression Little App to predict the price of a house using the size of the living area and whether or not the house has central air.

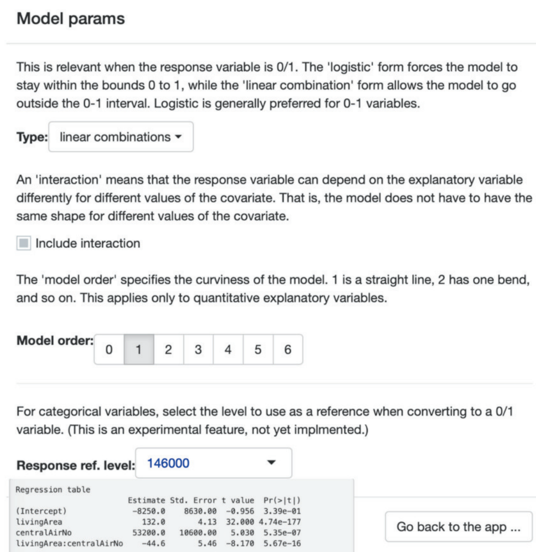


Figure 7.9: Using the Regression Little App to fit the regression model for price of a house using the size of the living area and whether or not the house has central air, and an interaction term.

7.4 Conclusions: Multiple Experiences with Multivariate Thinking

We know that many introductory statistics courses are already bursting at the seams with topics to cover, so it is a difficult sell to suggest adding new content. Luckily, adding multivariate thinking doesn't require adding a new module as it can be infused into pre-existing materials. When teaching about data acquisition, consider featuring examples of stratification or blocking and asking students which additional data might have enhanced the analyses. When students are wrangling data, have them explore how their transformation choices impact other relationships in their data. When visualizing data, make sure students deconstruct and construct graphs of several variables and see how the addition or subtraction of a variable impacts the story. When covering models, have students explore how the conclusions change when adding a new variable into the mix, and along the way make sure students practice communicating their findings, both in oral and written form. As the GAISE College Report states, ”We must prepare our students to answer challenging questions that require them to investigate and explore relationships among many variables. Doing so will help them to appreciate the value of statistical thinking and methods.”

8

Points and Density

Carol Howald, *Howard Community College*

8.1 Introduction

Introductory Statistics presents students with basic statistical terminology and concepts including the vocabulary of study design and sampling techniques, the use of descriptive statistics for investigating center and spread and the population parameters they estimate, the role of sample variation, and representations of data that support inferences about a population. These topics have been chosen so that students can read and understand research within the disciplines related to their program of study. This course can also provide an excellent environment to develop professional skills, such as teamwork, written and oral communication, and problem clarification and resolution in the context of data analysis. In addition, these topics should help students become more informed citizens in our data-driven society, an important goal often present in a college's general education model.

As detailed in the Part II Overview, the American Statistical Association updated its GAISE College Report in 2016 [67]. The issues that prompted this work include the increasing rate that college students were studying statistics, with many of them taking courses at community colleges where faculty may lack formal training in statistics. It is important to recognize both the commonalities and differences between mathematical reasoning and statistical reasoning. Mathematicians are often well-versed in the theoretical development of statistical measures. They are often less familiar with the inference processes and the technology that allows one to investigate real-life data sets and explain the strengths and weaknesses of a model when making inferences from a probability sample. In addition, more and better technology options for visualizing data are constantly becoming available. These tools change how professional fields analyze and present findings.

The GAISE College Report identifies six recommendations for teaching introductory statistics courses and suggest topics that may be omitted from introductory statistics courses to create room for these recommendations: 1) reduce the emphasis on probability to basic probability about random variables with the binomial as one special case; 2) reduce time spent constructing statistical plots by hand given the wide-spread availability of technology and instead focus on developing a deeper understanding of what we can tell about the data from the plots; 3) reduce coverage of basic statistics that are now commonly found in the K– 12 curriculum such as pie charts, histograms, calculating means and medians; and 4) eliminate the use of statistical tables given these skills no longer reflect modern statistical practice with the ability to look these up on easily accessible apps and programs such as Excel.

8.2 Data Exploration in an Introduction to Statistics Classroom

As we consider data exploration class activities, work tasks, and assessments it is important to evaluate statistics students' incoming skills, programs of study, and career plans. With curriculum recommendations being implemented in our nation's high schools based on documents such as the National Council of Teachers of Mathematics (NCTM)

Principles and Standard for School Mathematics and the Common Core State Standards Initiative, some students arrive in our classrooms with rich vocabularies and experience summarizing, representing, and interpreting categorical and quantitative data [129]. Other students arrive in our classrooms who experienced mathematics curriculums prior to the implementation of such recommendations. By incorporating tasks that students can access from multiple levels of understanding and providing students with the opportunity to discuss their findings with others, we support all students as they transition from using informal language to describe patterns in data to a more nuanced interpretation incorporating the language of statistics.

We can create an inclusive and empathic classroom culture by considering their programs of study and career plans. For many students, mathematics courses have represented an alienating experience where they struggle to see the relevance in their lives and future plans. By getting to know our students, we can incorporate statistical investigations and data sets that pique their interest and demonstrate the relevance of the course. By valuing their interests and needs, we have a context for instruction that holds their interests and supports the development of a strong conceptual understanding of key contexts.

As we examine several data exploration activities developed during the StatPREP project, we will highlight how they address the national recommendations for developing statistical reasoning (such as using large data sets and experiencing sample variation), how students can access the tasks from different levels of understanding, and how they can contribute to creating an inclusive classroom culture. Many of the tasks were developed with the National Health and Nutrition Examination Survey (NHANES) as the context [30]. All the activities can be adapted for use with other large data sets that are relevant to the students served in a particular introductory statistics class. These activities incorporate the use of the Little Apps developed through the StatPREP project. Finding no-cost resources that can serve many purposes for our students also provides accessibility for students with financial limitations.

Before considering the Points and Density Little App activities, it is useful to emphasize one of the significant strengths of the Little Apps: they allow students to visualize and explore interesting data. One way to support student learning and to allow them to make meaningful conclusions is to use jittering. To demonstrate, consider the scatterplot shown in Figure 8.1, which was created using R and a 300-entry data set that provides data on the meetings chaired by individuals who are members of a hypothetical team and the length of the meetings. In Figure 8.2, the data has been plotted using the jitter function in R. Essentially, jittering adds noise to the plot to make it easier to interpret. Of course, over-jittering could distort the data (as in Figure 8.3) and should be avoided.

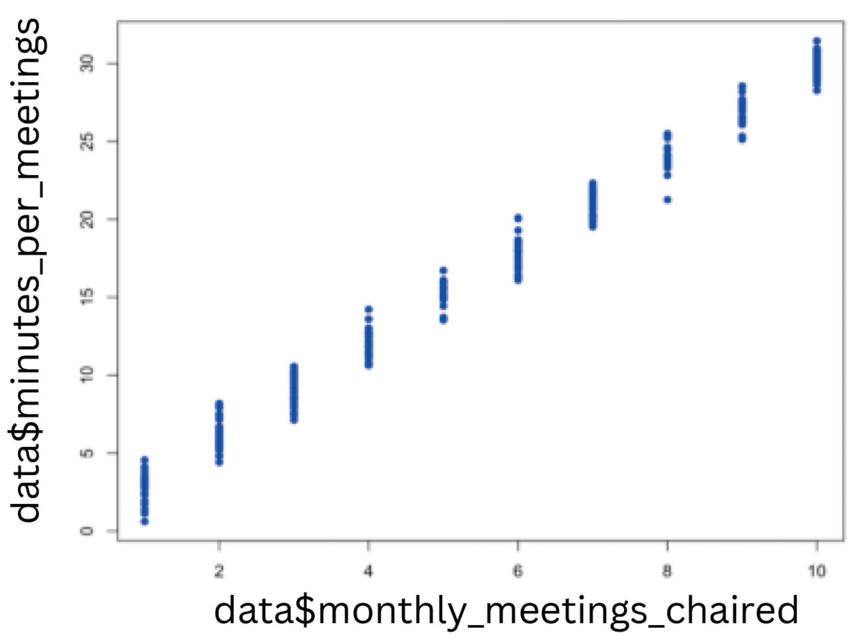


Figure 8.1: Scatterplot of Meetings Chaired Data.

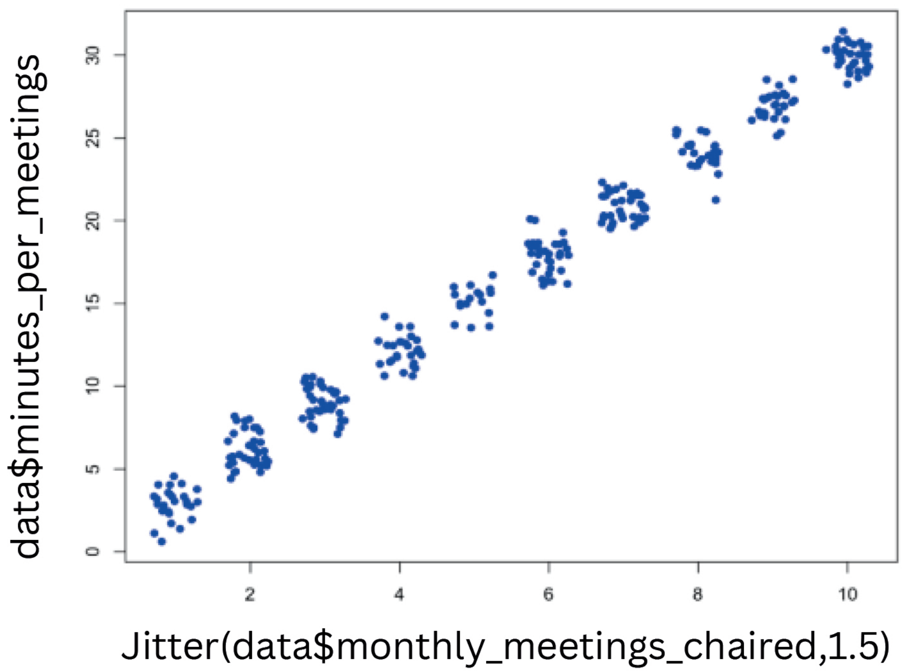


Figure 8.2: Jittered Meetings Chaired Data

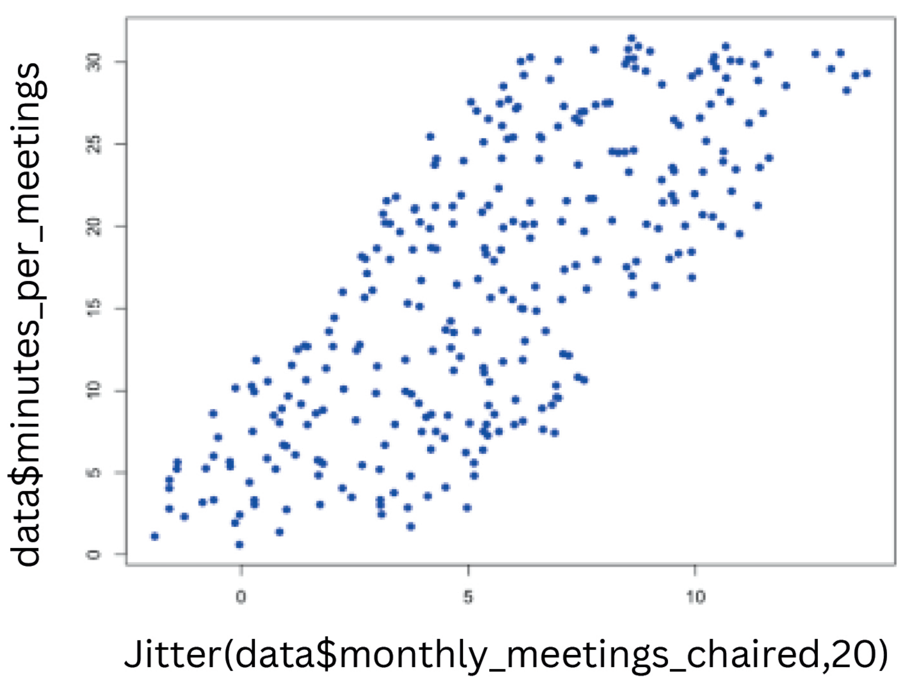


Figure 8.3: Over-jittered Meetings Chaired Data

8.3 Classroom Introduction to the Little App: Points and Density

StatPREP's Little Apps are interactive web tools with large data sets available for exploring statistical concepts. It is helpful to introduce students to the Little Apps with simple activities centered on concept and vocabulary development early in the semester. When we begin looking at more complex data representations with large data sets, the students are already comfortable with variable types, the purpose of a data codebook, and the shared features they will experience across the six Little Apps. This chapter will focus on the Points and Density Little App.

8.3.1 Laying the Groundwork

Early objectives in introductory statistics courses include recognizing different types of statistical studies, variables, and sampling techniques. The following activity provides students with the opportunity to apply some of the language developed and lays the groundwork for introducing the Points and Density Little App to students in a meaningful context. This activity is built around the NHANES study because many of my students are interested in health-related careers but a similar task could be built around disciplines of interest to your students.

Activity: Studies, variables, and sampling techniques

Instructions: Take 2 minutes to read the following study description. In groups of 3 discuss and answer the following questions.

The National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The survey is unique in that it combines interviews and physical examinations. All participants visit the physician. Dietary interviews and body measurements are included for everyone. All but the very young have a blood sample taken and will have a dental screening.

The survey examines a nationally representative sample of about 5,000 persons each year. These persons are located in counties across the country, 15 of which are visited each year.

The results of NHANES benefit people in the United States in important ways. Facts about the distribution of health problems and risk factors in the population give researchers important clues to the causes of disease. Information collected from the current survey is compared with information collected in previous surveys. This allows health planners to detect the extent various health problems and risk factors have changed in the U.S. population over time. By identifying the healthcare needs of the population, government agencies, and private sector organizations can establish policies and plan research, education, and health promotion programs that help improve present health status and will prevent future health problems. (Reference: National Center for Health Statistics NHANES data)

1. Does the NHANES program of studies represent experimental studies or observational studies? Explain.
2. Does the summary suggest a simple random sample of people living in the US? Why or why not?
3. Predict 3 categorical variables and 3 quantitative variables that you believe might be collected from each participant during the study. For categorical variables, indicate possible categories of participant responses. For quantitative variables, indicate possible units of measure.

As the groups work, instructors have the opportunity to make a formative assessment of student understanding of the terms and concepts involved. Select two or three groups to share their responses to questions one and two, choosing groups that provide a variety of perspectives and language. Encourage the class to compare and contrast the arguments provided and come to a consensus.

By asking small groups of students to predict and discuss possible variables of interest in question three, we give them the opportunity to brainstorm multiple variables that might be related to health issues, giving us insight into what variables students may be familiar with or interested in. Enhancing students' ability to recognize and understand types of variables will support their success in selecting appropriate data representations, analysis tools, and appropriate statistical tests. Having them also brainstorm possible categorical responses or quantitative units allows instructors to evaluate student understanding of these variable types. Encouraging students to make predictions about variables or outcomes is an excellent way to get them invested in exploration.

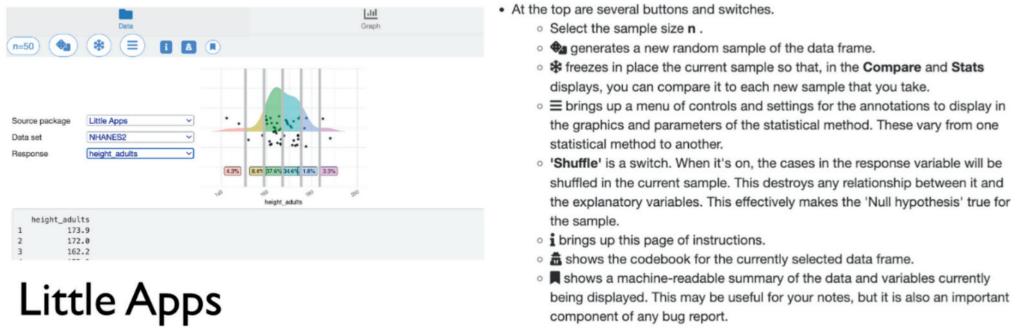


Figure 8.4: Exploring the Little Apps.

8.3.2 Initial Exploration of the Little App

The Points and Density Little App provides a smoothed density graph for a selected variable with the data points overlaid. Instead of a dot plot being used for the points, which would be cumbersome with a large sample, the points are "jittered" around the horizontal axes. (The points height are randomly assigned within a small vertical window.) The density of these jittered points helps students make sense of the meaning of the smoothed density graph.

To use this or any of the other Little Apps, it is helpful to have a general understanding of the tool. Figure 8.4 provides an overview.

The following activity encourages students to explore the app while consolidating their understanding of different variable types. The activity allows students to self-assess whether they used the terms categorical and quantitative accurately. In this task, students are only asked to access the codebook and to generate a sample for a few variables. As the class shares their results, the instructor can call up individual variables that were identified, indicate how an individual data point is represented (the jittered dot), the distribution shape (the density of the jittered dots) and how the smoothed density curve relates to that distribution.

Activity: Variables and codebooks

Visit the Points and Density Little App. This tool contains data from the NHANES studies. Previously you predicted a number of variables that might be of interest to the researchers. Although this tool does not contain all the collected data, it does contain many of the variables used in the studies.

1. Many studies use a codebook to document the variables in the study, the units of measure used for quantitative variables (for example, are heights recorded in inches or cm?), and the categories used for qualitative variables. Hover over each icon near the top of the page to find the **codebook** icon. Click on the codebook and list any of the variables you predicted in the previous activity that were included in the study.

2. Identify and list below 3 different variables in the codebook that you think are qualitative and 3 different variables that you think are quantitative.

Qualitative	Quantitative
(a)	(a)
(b)	(b)
(c)	(c)

3. Close the codebook by clicking dismiss at the bottom. For each of the variables you listed above, find the variable under the drop-down menu for the response variable to generate a sample of 50 responses. Use the data provided to confirm your decision of Qualitative or Quantitative. For each variable, share one data point, illustrating a category for a qualitative variable or a measurement for a quantitative variable.

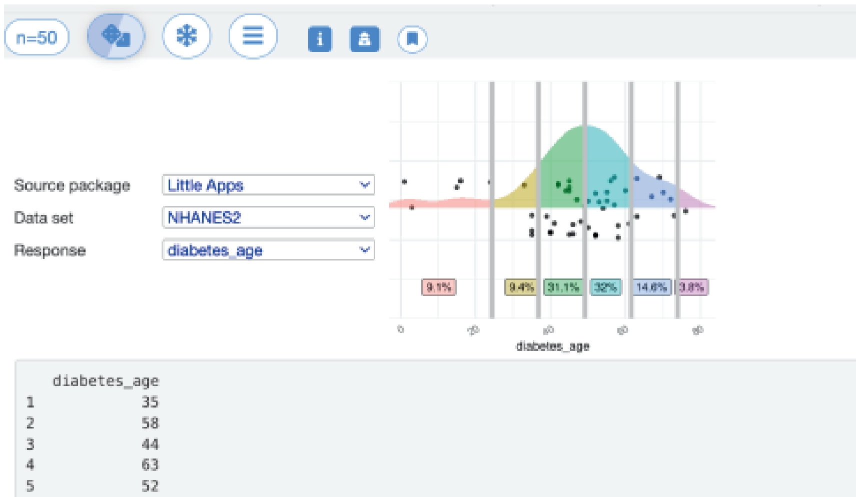


Figure 8.5: Variables and codebooks

As an example of number 3 in the above activity, a student might indicate they found the variable diabetes_age, defined as the age of study participant when first told they had diabetes, to be a discrete quantitative variable with age 37 as one reported outcome (see Figure 8.5). Useful questions to ask include ‘Which black dot represents this outcome?’, ‘Do you think this person is above or below the median age when a person is informed they have diabetes (are more than half the dots above or below 37)?’ and ‘What is the relationship between the clusters of dots and the shape of the curve behind them?’.

When the instructor pulls up a sample, students will often note that it is different from their results, giving an early opportunity to consider sampling variation using statistical concepts previously addressed. Is the new sample similar in range, shape, and mean (by default, the center vertical gray line in the graph)? If we increase the sample size in the app, do the density plots of our individual samples look more or less similar in shape to each other? This provides a good transition to introducing concepts of skew and symmetry and their relation to measures of center.

8.3.3 Distribution shape and its relation to measures of center

In the next activity, students use their familiarity with the data set to make some predictions about the shape of the distribution for a variable of interest and the relationship between the mean and median of the data. Although the activity below uses the NHANES data set again, one could use any large data set of interest for a similar experience. To upload their own data, use the data icon to change the source package to UPLOAD. The app will prompt you to find a .cvs or .rda file to upload.

Initially, it is good practice in graph reading and interpretation to let students work from the definition of median to evaluate the relationship between mean and median. As a student gains confidence with these skills, they can demonstrate using the control button (icon with 3 horizontal lines) to toggle between marking the mean or median as the center of the graph.

Activity: Distribution shapes and measures of center

Visit the Points and Density Little App. In the NHANES data set, set the sample size to 100. Choose a quantitative variable that you believe has a skewed distribution. Remember to use the codebook if you need to decipher what the name of a variable represents.

From the graph, indicate the direction of the skew. You can click on the graph icon at the top to access a larger view. Investigate whether the mean is higher or lower than the median for the sample. Recall the center grey line will represent the mean of the data by default. The median represents the value where half the data points will be above this value and half are below. Predict if the mean will be higher or lower than the median for the sample and estimate the value of each measure of center from the graph.

1. Quantitative Variable name:
2. Direction of skew:
3. Which center is estimated to be higher, Mean or Median? Explain.
4. Resample 3 times by pressing the dice icon. Does this relationship seem to hold over several samples?

As the students share some of their results, they generalize about the impact of a heavy skew or extreme outliers on the value of the mean relative to the median. This also provides an important experience to establish how resampling can be useful in identifying meaningful differences and similarities among samples. At any time students can freeze a sample for comparison using the snowflake icon near the top of the app. This sample will then be available under the compare icon. Each new sample will appear adjacent for easy viewing. With these types of experiences with sample variation and identifying consistencies across samples, students are much more prepared to understand inference concepts such as the 95% confidence interval. These features of the Points and Density App are also useful when developing other concepts underlying inference concepts, such as standard deviation and z-scores.

8.3.4 Common and rare: Using standard deviation as a ruler for expressing frequency

Building meanings for terms such as common and rare, and using statistical concepts such as standard deviation and z-scores are other important introductory statistics objectives for data exploration. We want to continue to incorporate how variability within a sample distribution should be taken into consideration. We want students to understand what a common or rare sample result means in the context of a study.

It is important for students to realize researchers bring bias from their personal experiences to any statistical investigation. We need visualization tools and interpretation skills to move past our own experiences of what is "typical." To begin, have students get their personal conceptions of a variable on the table. I often ask students to discuss in small groups a range of women's ages that they think of as a "common" age to give birth and an "unusual" age to give birth before looking at any data (and stipulate no quick googling!). Each group reports a range on the board after three minutes. Typically there will be a lot of variations in their responses and some discussion about how old their own mothers were when the student was born or how old the student was if they have children. This provides some motivation to establish a need for a definition of common and rare and to look at data for U.S. women to take into consideration more than our own experiences.

As a class, we might decide that the youngest and oldest 5% of women are unusually low or high and the middle 50% is very common. Using the Natality_2014 data set from the CDC in the Points and Density App allows us to estimate these ages. In the Data tab, switch the Data set to Natality_2014 and select *mager*, the age of the mother when she gave birth (see the codebook for details), as the response variable. Once a sample is selected, under the graph tab, the gray bar in the center represents the mean age of the mother at birth and each of the other bars can be dragged to capture the designated percentages of the data points (see Figure 8.6). If one hovers the cursor over a bar, the z-score will be displayed above the graph. This provides a way to introduce z-scores conceptually with a large data set before focusing on the mechanics of the computation. Students are already familiar with standard deviation as a measure of spread (the typical distance a data point is from the mean). The mean age at birth is 27 years of age and the top 5% of the ages has a cut-off value of 39 years of age with a z-score of 1.7. Alternatively, we might have defined ages above 2 standard deviations above the mean as rare and moved the bar to find the corresponding age and percentage of the data points.

The following task was designed with health science students in mind. The data set and terms involved could be varied for other student populations. The objective of the activity is to have students investigate typical and atypical pulse rates through data exploration, demonstrating their understanding of the concepts of mean, standard deviation, z-score, and percentiles. Students are asked to find their own pulse rate. This fairly quick act hooks their attention without asking them to share numbers that can be more stress-inducing like height or weight. Some of the students will know how to accurately measure pulse rate depending on their program progression; others are interested in learning and practicing. I have them collect this data in class and use it with the following at-home assessment.

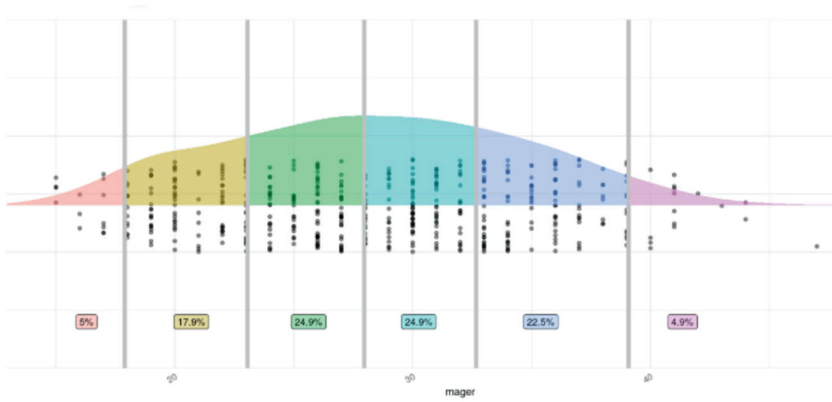


Figure 8.6: Common and Rare Example.

Assessment: Investigating Pulse Rate

1. What is your current 60-second pulse rate? How does it compare to the U.S. adult population?

   Visit the Points and Density Little App to use the NHANES data set to determine and interpret the z-score and percentile for your current pulse rate.

2. *Tachycardia* is the medical term for a heart rate over 100 beats per minute. Use your cursor to move the rare marker to represent Tachycardia. Describe in terms of approximate standard deviations from the mean. Is this a common event for a US adult? Explain and include a screenshot of the distribution supporting your work.

3. *Bradycardia* is the medical term for a heart rate that is less than 60 beats per minute. Use your cursor to move the lower rare marker to represent Bradycardia. Describe in terms of approximate standard deviations from the mean. Is this a rare event for a U.S. adult? Explain and include a screenshot of the distribution supporting your work.

Using summative assessment tasks where large data sets and an appropriate technology tool are required convey to our students the importance of ‘doing’ statistics as it is experienced in their discipline. Often, simple data sets and calculators are used to simplify the summative assessment process but students often equate what is on the test with what to value the most.

In addition to accurately applying concepts and interpreting results, we need our students to pose and investigate good questions. By allowing students to choose a source package, data set, and variable to consider, they have the opportunity to practice posing meaningful questions. Using the codebook to clarify the variables available in a data set, they can form and investigate a question of interest around defining common and rare events. Students can be asked to give an overview of the distribution using the appropriate statistical language and to define common and rare outcomes based on percentiles and z-scores, with a justification of why knowledge of these outcomes is relevant in our society. Even within the NHANES data set, there is a wide scope of variables available beyond typical health markers like pulse rate, including income level, drinking levels, age at onset of events such as sexual activity and marijuana use, and time spent sleeping, watching TV, or using computer time.

One unexpected perk of these investigations early in the course is the multivariable questions that arise out of the initial investigation of a single variable within the Points and Density App. Is the level of drinking representing the highest 10% of drinkers the same for males and females? Is the range of household income within 2 standard deviations of the mean the same for people with only a high school diploma as it is for those with a college degree? Their questions can help instructors build interesting and relevant examples and tasks as they move through the rest of the course using data in which their students now have a vested interest.

8.4 Conclusions

By addressing the national recommendations for developing statistical reasoning (such as using large data sets, embracing technologically driven representation, and exposing my students regularly to sample variation), my classroom has become more centered on real data sets that bring both complexity and excitement to my teaching. These data exploration investigations, tasks, and assessments have increased my students' ability to read and understand research within the disciplines related to their program of study. Regular group work around data exploration has given my students a contextualized way to transition from the informal language they initially use to describe patterns in data to a more nuanced interpretation incorporating the language of statistics.

Since replacing small class surveys with the use of Little Apps and the data exploration activities developed during the StatPREP project, my course is more robust and relevant to my students. The technology provides inexpensive ways to represent data in ways they will see in professional journals and media outlets and the data we explore represents our world more accurately. Finding no-cost resources that can serve many purposes provides accessibility for our students with financial limitations.

The ability of the technology to let students experience sample variation regularly across many concepts has built my students' confidence in the statistical process. Previously, some of my students had true difficulty understanding how a single sample could provide insight. They embraced the chance of a sampling error and ignored the much more likely chance of a representative sample. With quick access to repeated samples within the app and the clear visualization of the distribution, students were more confident that they could make claims about a population with the analysis of a carefully collected sample.

9

Confidence and T

Dustin Silva, *College of the Canyons*

9.1 Introduction

The purpose of this chapter is to provide introductory statistics instructors with accessible starting points for implementing practices that foster student engagement in learning about confidence intervals. We start by explaining why we focus on confidence intervals rather than hypothesis testing. Next, we explore StatPREP's Confidence and T Little App going over the mechanics of the technology and providing some insights on using the Little App to further develop a student's conceptual understanding. After that, we overview various activities that are ready to use in an introductory statistics classroom. Assessment suggestions as well as other classroom-relevant topics are also introduced. We end with a summary and key takeaways.

StatPREP's goals, including using technology, focusing on conceptual understanding, and incorporating active learning, provide the framework for our discussion. Using technology in the classroom supports active learning and helps increase student understanding. A tool such as the Confidence and T Little App can increase student engagement and give more time to focus on concepts rather than formulas.

Why Confidence Intervals? When discussing topics about inference, the more useful approaches to interpreting study results such as confidence intervals tend to be overshadowed by a strong emphasis on hypothesis testing [68]. A confidence interval is a range of values that will contain the population parameter for a specified proportion of trials. Confidence intervals provide all the information that a test of statistical significance provides and more, but since confidence intervals avoid the term "significance," they avoid the misleading interpretation of that word. Moreover, aspects of confidence intervals are seen in the media, so this is a concept students will likely encounter outside of the classroom.

9.2 Confidence Intervals in an Introductory Statistics Course

This section discusses ways to talk about confidence intervals in an introductory statistics course by focusing on how to introduce the topic in a way that will pique student interest in the topic, and addresses some of the challenges encountered in teaching this topic. Focusing on the StatPREP principles of using technology to visualize data to help develop conceptual understanding over computations aids students in recognizing the value of this important topic in science and society.

9.2.1 Where to Begin: Confidence Intervals

The way a topic is introduced can promote some understanding of it before getting into the finer details. When introducing confidence intervals, starting with a question like "how long did it take you to get to school or class?" can be helpful. A student may respond that it takes 30 minutes for them to get to school. I follow up with something like, "Exactly 30 minutes? Is it 30 minutes every time, or would you say 30 minutes is about the average time it takes you to get to school?" and students tend to say "it's an average." I follow up by asking if they can give me a range of how many minutes it typically takes. They may respond that 25 to 35 minutes is the range of time it takes which leads us into using an interval to estimate an average time. I'm able to build on this familiar situation and refer back to it as we move forward. Establishing this context serves as a great precursor for a discussion about defining the population and what inferences can be made.

9.2.2 Instructional Challenges: Confidence Intervals

There are challenges, as with any topic, in teaching confidence intervals. Students in Introductory statistics courses typically have diverse backgrounds and varied mathematical preparation, and some topics tend to be easier for students to pick up on than others. Here are some examples that students may struggle with on the subject of confidence intervals:

- Identifying the type of variable, categorical or quantitative, and whether to use a proportion or mean.
- Interpreting and understanding a confidence interval.
- Understanding what it means to be 95% confident.
- Getting caught up on formulas.
- Missing the overarching meaning of a confidence interval.

How do we, or should we, teach the material? There may not be an approach that works for everyone; it may depend on the goals for your course and whether your students are more likely to be producers or consumers of statistics. Assessing this may help guide the instructor in how they present and teach the material so that it meets the course learning outcomes. This can also make it more relevant to students and create more interest and understanding.

9.2.3 Bridging Gaps: What does it mean to be 95% confident?

In my experience, one of the most common roadblocks to understanding confidence intervals is the confidence level. What does it mean to be 95% confident? Students may not understand the distinction between confidence and probability. A population mean is either in the confidence interval or it is not in the interval. Using applets to produce displays of sample intervals and seeing which ones do or do not contain the population is a great way to show this.

A story can help with understanding. Imagine your doctor advises you to have a medical procedure that has a 90% success rate. What does this mean? It means that 90% of the procedures are successful and 10% are not. So if you have the procedure, would you say there is a 90% chance it was successful? Most students would say that seems correct. However, once the procedure has been done, it either worked or it didn't; speaking of "the chance it works" is no longer meaningful. Another way to think about this scenario is to represent all the procedures with slips of paper and randomly picking one. There is a 90% chance of picking one that was successful, but once you pick one then the chance is gone for that particular procedure (case/sample). It either was or was not successful.

9.3 Little App: Confidence and T

9.3.1 Access the Little App "Confidence & T"

To access the Little App "Confidence & T," you may go to this browser-based website: Little_app_T If you want a refresher on how functionality of the Little Apps, review this short video.

We are able to customize confidence interval(s) by showing the confidence interval (clicking on "Show conf interval" as shown by an example below in Figures 9.1 and 9.2.

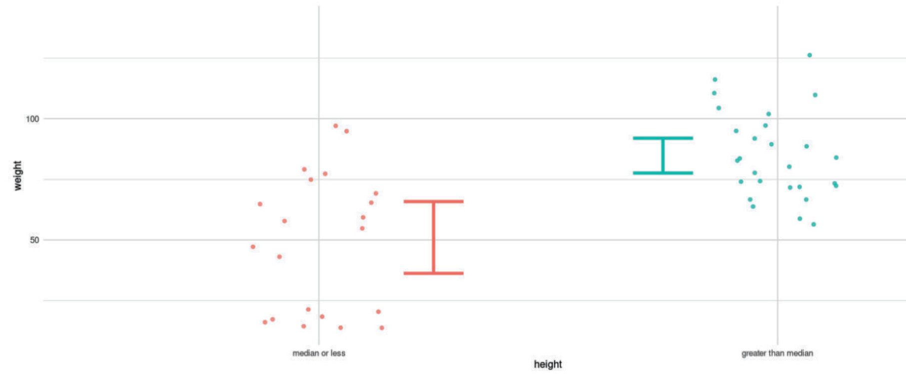


Figure 9.1: Confidence and T Little App Example

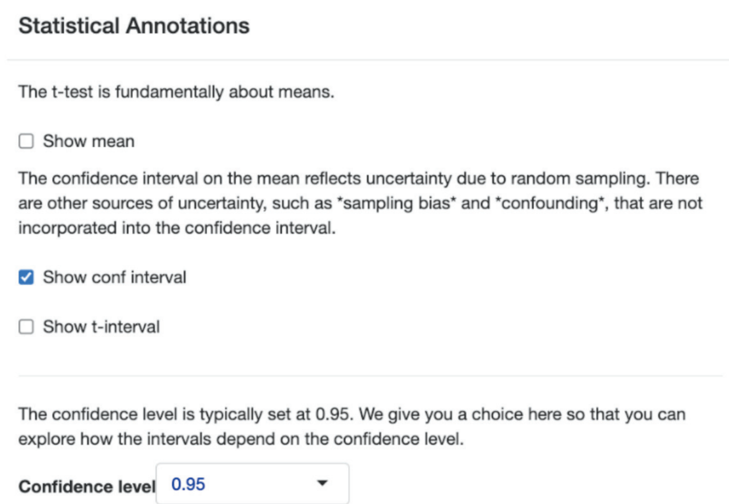


Figure 9.2: Confidence and T Little App Example

Teacher Takeaway: The confidence interval can be estimated based on graphs, using the visualizations to find the upper and lower bounds of the respective intervals, as well as the mean or margin of error of each. With these visualizations, we can also see whether the confidence intervals overlap or not, and what it tells us. Using the graphic above, we can see that while there is overlap in the data, the confidence intervals do not overlap, which suggests there is a statistically significant difference between the two groups and provides evidence that the average weight of those with a median or lower height differs from the average weight of those who are taller. Visually, it suggests taller people have a heavier average weight than the average weight of the rest. Students can estimate the lower limits and upper limits, or the margin of error. After constructing the confidence interval, click on the lower bound of one of the confidence intervals and drag the cursor to the upper bound. Then let go of the mouse to see an approximate length of the confidence interval, which can be divided by two to get the margin of error.

By also clicking on "Show t-interval," we see the appearance of a confidence interval in the middle of the aggregated data, as shown in Figures 9.3 and 9.4.

The "Show t-interval" gives the difference between the two samples. Instead of showing the difference in means, the t-interval in the Little App is centered on the sample mean of the first group. This allows the reader to see how much more or less the mean of the first group is from the second group. The idea behind this is to allow the reader to discuss how much more or less the first mean could be from the second mean rather than discussing how different the two means are. Figure 9.5 has another example of the t-interval. This example displays the sample income of different home ownership types. In this figure, the t-interval doesn't contain the sample mean of the second group, and it is

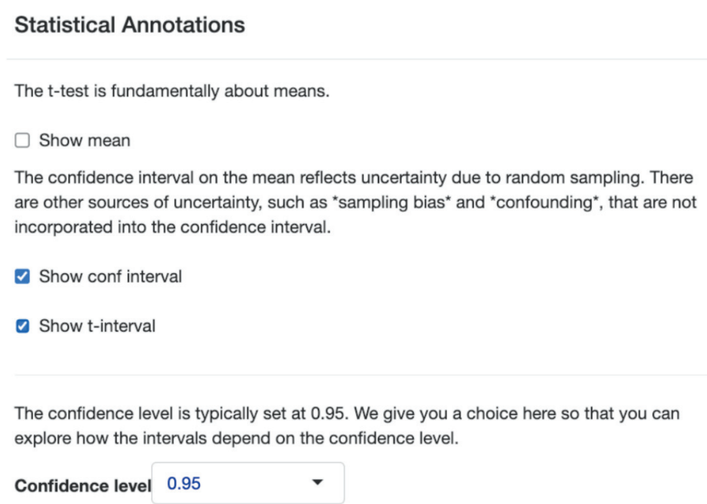


Figure 9.3: Confidence and T Little App Example, with ”Show t-interval” Clicked

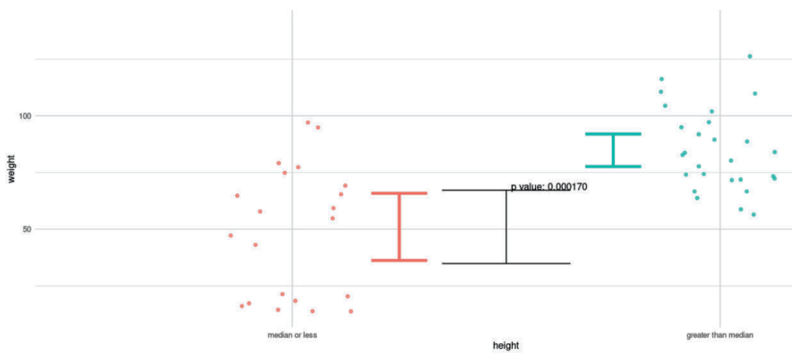


Figure 9.4: Confidence and T Little App Example, with ”Show t-interval” Clicked

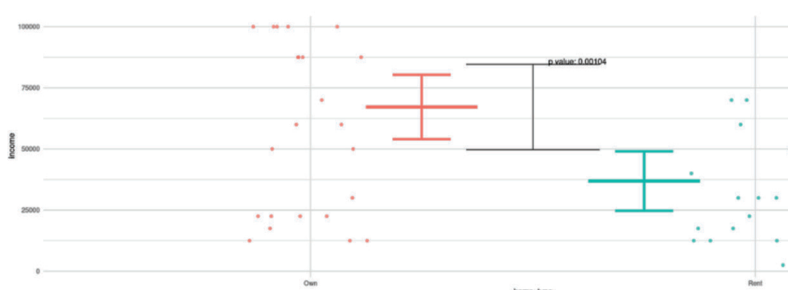


Figure 9.5

evident that the first group has a sample mean that is always more than the second group. Clicking on the Stats tab reveals that the t-interval for this sample is CI = 12850 to 47730 at a 95% level, which means that the mean income level of people who own a home is anywhere from $12,850 to $47,730 more than the mean income level of people who rent a home.

9.4 Activities Using Confidence and T Little App

As we explore this Confidence and T Little App, we will explore concepts around confidence intervals using ready-to-use activities that can be modified as instructors see fit for their courses. Students learn best when they are engaged in their learning [62]. We will now focus on classroom activities that give students the opportunity to be actively engaged in their own learning, and give faculty resources that may immediately be implemented in their Introductory Statistics courses. After each activity, discussions about the activity and this topic will be addressed.

Productive Struggle: Productive struggle in a mathematics class can be beneficial for student learning. Students' productive struggle in mathematics is described as the rigorous learning which can effectively promote grit and creative problem-solving [161]. We utilize the word struggle to convey the intent for students to exert effort to make sense of mathematics and to understand topics that are not immediately apparent [57]. Once students have an understanding they may have to clarify questions with discussions. It is important to follow up on their questions after activities such as these. Questioning techniques cannot meaningfully foster student engagement if students are not given sufficient time to think about the questions posed and to respond [1]. When students productively struggle, they are able to come up with solutions to problems that are within reach even when the concepts may be initially difficult to comprehend, whereas unproductive struggles do not give students useful knowledge.

9.4.1 Confidence and T Classroom Activity #1: What is a confidence interval?

In this first example of how to utilize the Confidence and T Little App, we learn about confidence intervals, their legitimacy, and how a confidence interval changes by both confidence level and sample size.

Typically in an introductory statistics course, we start with data analysis by finding sample statistics such as the mean or proportion. We would at that time use these statistics as our best guess at estimating the population parameter. When introducing confidence intervals, we can describe them as another type of estimate, but instead of being just one number, the interval is a range of values calculated from a given set of sample data and is an interval based on a sample statistic. The confidence interval found is likely to include an unknown population parameter.

We are highlighting a part of the activity below. The complete activity can be found at StatPREP Little Apps: What is a confidence interval?

Activity 1 Learning Objectives

1. Identify the characteristics of a confidence interval estimate.
2. Demonstrate how the confidence level affects the width of a confidence interval.
3. Demonstrate how the sample size affects the width of a confidence interval.

Activity 1 Highlight. The highlight shown in 9.6 is taken from Activity #1, which is accessible via the link found in the preceding paragraph. This example highlights students taking a sample and creating confidence intervals with

5. In order for confidence intervals to behave in the right way, that is, to cover the population parameter at a frequency specified by the confidence level, valid confidence intervals tend to change systematically with the sample size and with the confidence level.

   i. Keeping one sample on the display by pressing the freeze the current plot icon (snowflake icon) change the confidence level. Then click the Compare button in the top tool bar. Now chance the confidence level and compare the length of the interval. State if the interval length became wider or narrower for each increae in confidence level.

Confidence level	Interval length
50%	
80%	
95%	
99%	
99.99%	

   ii. Come up with a general statement about how the length of the confidence interval depends on the confidence level.

Your answer:

Figure 9.6: Confidence and T Little App Activity 1

different confidence levels. Students then use the lengths of each confidence interval and accompanying level to draw a general conclusion about the relationship between the length of an interval and confidence level.

Activity 1 Discussion Questions. This activity provokes a lot of questions. Here are some questions you can consider asking your class following the complete activity.

1. Why would we obtain different values of a statistic for different samples drawn from the same population?
2. Where is the mean located relative to the lower and upper bounds of the confidence interval?
3. If sample size stays the same then what other factors could affect the distance between the lower and upper bounds?
4. What does changing the sample size do to the confidence interval?

9.4.2 Confidence and T Classroom Activity #2: Comparing Two Confidence Intervals

This second example utilizes the Confidence and T app to support a discussion on sampling variability. This topic in an introductory statistics course is typically done before confidence intervals are introduced. The notion of sampling variability can be introduced on the very first day and highlighted throughout the course. For inference, we can use this Little App to explore if there's sufficient evidence that the two means are different. This brings in the notion of hypothesis testing as well as reinforces the idea around sampling variability. Typically we would use a two-sample t-test for such problems, but it is nice for students to build up intuition first and then formalize inference later in the course with the help of other statistical software.

We are highlighting a part of the activity below. The complete activity can be found at StatPREP Little Apps: Comparing two confidence intervals. After this activity, discussions about the activity and this topic will be addressed.

Activity 2 Learning Objectives

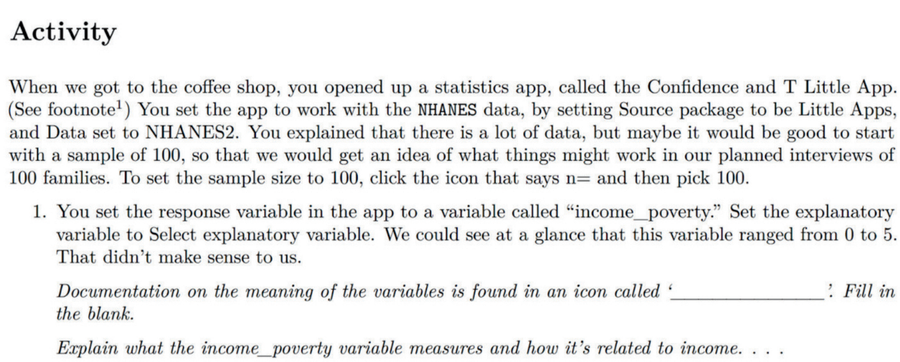
1. Understand the meaning of the variables in data sets (Codebook).
2. Use a jitter plot to describe patterns in data, and compare among groups.
3. Understand the random nature of sampling.
4. Informally evaluate if there's good evidence that two means are different.

Activity 2 Highlight.

The highlight shown in 9.7 is taken from Activity #2, illustrates how an activity can be given to students with outlined steps for students to do as well as prompts to respond to along the way. Students not only become more comfortable navigating the applet but are analyzing real data as they are learning.

Activity 2 Discussion Questions. This activity provokes a lot of questions as we introduce the notion of sample variability to students. Here are some questions you can consider asking your class following the complete activity.

1. Why would we obtain different values of a statistic for different samples drawn from the same population?



Activity

When we got to the coffee shop, you opened up a statistics app, called the Confidence and T Little App. (See footnote[1]) You set the app to work with the `NHANES` data, by setting Source package to be Little Apps, and Data set to NHANES2. You explained that there is a lot of data, but maybe it would be good to start with a sample of 100, so that we would get an idea of what things might work in our planned interviews of 100 families. To set the sample size to 100, click the icon that says n= and then pick 100.

1. You set the response variable in the app to a variable called "income_poverty." Set the explanatory variable to Select explanatory variable. We could see at a glance that this variable ranged from 0 to 5. That didn't make sense to us.

   Documentation on the meaning of the variables is found in an icon called '_______________'. Fill in the blank.

   Explain what the income_poverty variable measures and how it's related to income. . . .

Figure 9.7: Confidence and T Little App Activity 2

3. Explore how sample size effects the length of the confidence intervals and whether they overlap. (Make sure the "Shuffle" box is *unchecked*, so you are looking at the actual data rather than a simulation.)
   - Set the sample size to n = 20. Measure the length of the confidence intervals using the app's measuring stick (click on the top of the confidence interval and then drag to the bottom of the interval). Do this for several new samples to get an idea of the *typical* length of the confidence intervals for a sample size n = 20. Write down your result.
   - Set the sample size to n = 200, ten times larger than the previous step. Again, measure the typical length of the confidence intervals over several new samples. Write down your result.
   - Set the sample size to n = 2000, ten times larger than the previous step and one-hundred times larger than when n = 20. Again, measure the typical length of the confidence intervals over several new samples.
   - Compare the lengths of the confidence intervals that you measures for n = 20, 200, 2000. Describe the pattern:

Figure 9.8: Confidence and T Little App Activity 3

2. Do you think that you can say much about an individual person's level for the poverty variable based on their home ownership?
3. What does the income_poverty variable measure and how does it relate to income?
4. Is there a consistent pattern or is the mean for "Own" sometimes above and sometimes below the mean for "Rent"?

9.4.3 Confidence and T Classroom Activity #3: Comparing Two Groups

In the third example, we consider the "Comparing Two Groups" Activity. In this activity, students take the data previously used and then layer confidence intervals onto the data. With this new information, students analyze two sets of data to see if there is a significant difference in means and relate this to whether the confidence interval overlaps or not. StatPREP Little Apps: Comparing two groups. After this activity, discussions about the activity and this topic will be addressed.

Activity 3 Learning Objectives

1. Analyze confidence intervals for two populations that overlap or do not overlap.
2. Understand how sample size affects the length of the confidence intervals and whether they overlap.

Activity 3 Highlight.

The highlight shown in 9.8 is taken from Activity #3, which can be found in the StatPREP Community library. In this example, students take a sample of various sizes and create a confidence interval based on the sample. Students then use the lengths of each confidence interval and accompanying sample size to draw a general conclusion about the relationship between the length of an interval and sample size.

Activity 3 Discussion Questions. Here are some questions you can consider asking your class following the complete activity.

1. Do larger sample sizes lead to larger confidence intervals?
2. What's the ratio of the length of the $n = 2000$ confidence interval to the $n = 20$ confidence interval? Can you see a simple relationship between this ratio and the ratio in a sample size of $2000/20 = 100$?
3. Test out your theory for the relationship between sample size and confidence interval length using the lengths of the confidence interval for $n = 2000$ compared to $n = 200$. Does your theory work?

9.5 Assessment Using Confidence and T Little App

Assessment plays a vital role in the process of learning and motivation for students. There are various types of assessment.

- Formative assessment is a process used during instruction and allows instructors to give feedback and suggest needed adjustments in order to increase students' understanding of intended learning outcomes. Formative assessments can be done informally during activities such as the examples in this Notes volume.

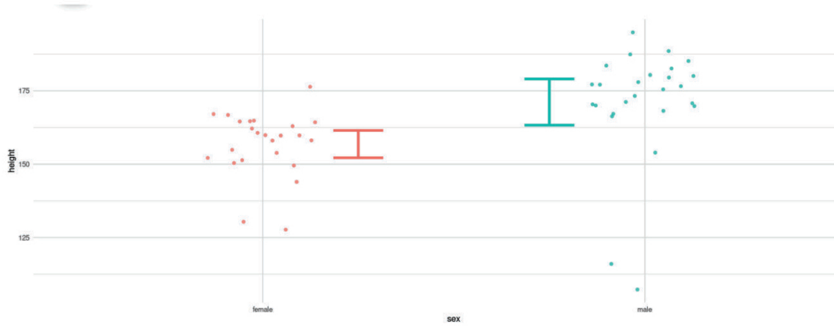


Figure 9.9: NHANES Data

- Summative assessment is the process of evaluating or certifying learning at the end of a specific period of instruction. Summative assessments are typically higher-stakes assignments such as exams or projects.

The StatPREP Little App "Confidence and T" can be used for both summative and formative assessments. For written assignments regardless of the type, students are encouraged to mimic what is done in class. One way to extend Little App activities is to use the download image feature and ask questions about the graphic related to the concepts learned during lectures.

Example 9.1. Confidence and T Assessment Question for "What is a confidence interval?"
Given two samples of size fifty of National Health and Nutrition Evaluation Survey data as shown in Figure 9.9, answer the following questions.

1. Do males or females have more variability in their heights? Explain using the data provided.
2. Is there a significant difference in heights between males and females? Explain using the data provided.
3. Which confidence interval has a larger margin or error and explain why based on the data? You do not need to calculate anything for this question.

Sample answers:

1. Do males or females have more variability in their heights? Explain using the data provided.
   (a) Males have more variability as their confidence interval is longer in width, or we can say they have a larger margin of error.
2. Is there a significant difference in heights between males and females? Explain using the data provided.
   (a) There is a significant difference since there is no overlap between the individual confidence intervals.
3. Which confidence interval has a larger margin or error and explain why based on the data? You do not need to calculate anything for this question.
   (a) The males have a larger margin of error as the distance from the middle of the confidence interval (the sample mean) to either of the limits (upper or lower) is larger than for females.

Example 9.2. Confidence and T Assessment Question for "Comparing Two Groups" Activity
Given four random samples showing the home type and income poverty level as shown in Figure 9.10, answer the following questions.

1. Who has a larger mean income poverty number in each of the four samples?
2. Are the four random samples showing evidence that those that own have a higher income typically than those that rent? Why or why not?

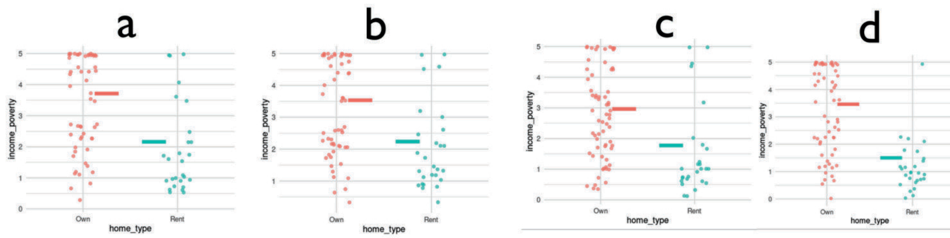


Figure 9.10: Home Type and Income Poverty Level

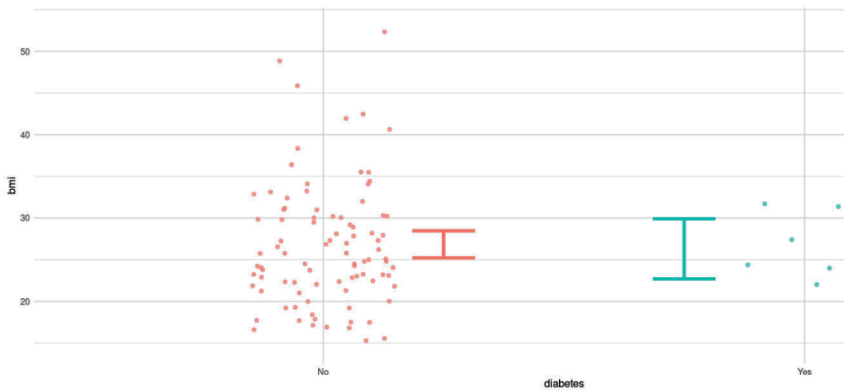


Figure 9.11: ”Common and Rare” Assessment Activity

Example 9.3. Confidence and T Assessment Question for “Common and Rare” Activity

Given two samples from the National Health and Nutrition Evaluation Survey data each of size one hundred as shown in Figure 9.11, answer the following question.

1. What does the overlap tell us about the BMI between diabetic versus non-diabetics?

Sample answer: If an overlap exists between the two confidence intervals, then we state that there’s not a statistically significant difference between the two population means. However, if there is no overlap we can state that it suggests that there is a statistically significant difference between the two population means.

9.6 Conclusion

There are some overarching takeaways:

- Being able to easily download graphics can be very helpful to instructors and students. For example, if an instructor or students have limited access to technology, the instructor may choose to download pictures in advance and provide print-outs for students. This allows students to take notes or use them in activities. Being able to download graphics also is useful for creating assessments.
- Using “All” is a great way to discuss a “population” with students. A discussion of the population and the parameter in context may be required. This may serve as a good reminder that we usually do not know what the measure of a population will be, which is why we take samples to make an inference about the population. Here, we have access to the population and can use that to see how our random samples relate to the population.
- Being able to take random samples quickly and show their distributions graphically is a great way to illustrate how samples vary from sample to sample. Adjusting the sample size and repeating the sampling process can also help show how sample size and variability relate. This is ideally done in such a way that each student can see this as they click on their own new random samples. Alternatively, the instructor can model this if they are not in a classroom equipped for students to have access to individual computers.

The purpose of this chapter was to give faculty teaching Introductory Statistics courses accessible starting points for implementing practices that allow for more student engagement in learning about confidence intervals. Using technology is good in the classroom as active learning helps student understanding. Using technology such as the Confidence and T Little App can increase student engagement and give more time to focus on concepts rather than formulas. Our goal was to share classroom-relevant topics that could be easily introduced.

10

Regression

Helen Burn, *Highline College*

10.1 Introduction

In an introductory statistics course, regression modeling generally refers to linear regression—one of the oldest and most widely used statistical techniques for describing and modeling correlations between a quantitative response variable and one or more explanatory variables. Simple linear regression is often covered in introductory statistics courses in which the response variable, y, is modeled as a straight-line function of the explanatory variable $x : y = f(x) = a + bx$. The linear growth rate, b, and intercept, a, are often utilized as a straightforward, albeit limited, way of describing the connection between the response and explanatory variables. However, simple linear regression is often deemed too simple to accurately describe modern research which involves large, real data sets and seldom consists of just a single explanatory variable. Therefore, at the introductory level, having students identify and address confounding variables, as well as using technology to explore the insertion of a covariate in simple linear regression, can introduce a more all-encompassing technique: multiple regression [25].

10.2 Topics Related to Linear Regression in a Traditional Introductory Statistics Course

Learning outcomes around simple linear regression (hereafter referred to as linear regression) begin with interpreting scatterplots before introducing concepts such as determining the suitability of a linear model for a data set and using the regression model for prediction. My own analysis of course outcomes in introductory statistics courses in Washington state showed a common basic set of topics but with wide variability in including analysis of residuals and conducting hypothesis testing under the null hypothesis that the linear growth rate or the y-intercept of the regression model are possibly zero [24]. Thus, this chapter focuses less on residual analysis and hypothesis testing.

The overall choice of topics is influenced by contextual factors including prerequisite course requirements, the length of the course (semester versus quarter), the existence of statewide learning outcomes, and the use of statistical software. Topics covered in a typical introductory statistics course include:

1. Create visualizations of bivariate data through scatter plots, using technology where appropriate.
2. Relate the location of each point in a scatter plot to the corresponding data.
3. Sensibly choose which variable should be the response and which is the explanatory variable (Note: Other nomenclature for explanatory/response includes predictor/response or independent/dependent variables.)
4. Determine whether a straight-line model is appropriate for describing a relationship between a pair of variables.

5. Use technology to find a linear regression model $y = f(x) = a + bx$, correlation coefficient r, and coefficient of determination R^2, for a pair of variables.

6. Evaluate the regression model by interpreting the correlation coefficient r, coefficient of determination R^2, and growth rate b of the regression model.

7. Use the regression equation for prediction.

8. Identify the residual of a point (difference between predicted value and observed value) given the location of the point and the regression function and recognize what residuals from a regression model indicate.

9. Conduct hypothesis testing under the null hypothesis that the regression model slope and intercept are possibly both 0.

StatPREP adheres to the 2016 GAISE College Report and recommends using technology to visualize data and to promote conceptual understanding over computation [67]. In the context of linear regression, this includes being able to interpret model statistics, such as the correlation coefficient, r, in terms of the direction (positive, negative, or null) and strength of the relationship between a pair of variables, as well as the coefficient of determination, R^2, which expresses the proportion variance in the response variable predicted by the explanatory variable. Students should also have an understanding of the growth rate, b, of the regression model in terms of the relationship between incremental change in input variables and the corresponding incremental change in output. Finally, the topic of regression typically includes using the regression model to predict a value of y given the value of x.

It is worth considering that presenting simple linear regression as a way to predict a value of y given the value of x can be somewhat misleading, as a proper statistical prediction should not be in the form of a single number. Instead, a meaningful prediction is that the output y for any given x is predicted to have the form of a normal distribution with mean $a + bx$ and a standard deviation corresponding roughly to the standard deviation of the residuals of the y-values from the corresponding model value. It is also noteworthy that mathematicians generally describe a and b in a formula like $a + bx$ as "coefficients" or "parameters." In statistics, the meaning of "parameter" is different (referring to a population) and the values of a and b generated by regression are "statistics" (being derived from a sample drawn from the population). Throughout this chapter, r, R^2, and b are referred to as model statistics.

10.3 Regression Modeling Little App Overview

The Regression Modeling Little App was designed to give instructors and students the option to do the following.

- Work with real-life, large data sets and upload data sets;
- Create scatter plots to visually recognize settings and variables for which regression is an appropriate technique;
- Select and explore different explanatory and response variables and sample sizes;
- Visualize the correlation coefficient r
- Add a covariate to the linear regression to promote multivariable thinking;
- Resample from the data set to demonstrate sample variability; and
- Compute regression model statistics (r, R^2, b), model residuals, and test statistics and p-values for hypothesis testing.

Upon opening the Regression Modeling Little App, users will be presented with the StatPREP landing page for the app shown in Figure 10.1.

Figure 10.1 shows that I am in the "Data" screen and am accessing the data set "NHANES2" (National Health and Nutrition Examination Survey) from the Centers for Disease Control and Prevention (2022). The NHANES2 data appeals to my pre-nursing majors who are the majority student audience for my introductory statistics course. As indicated in Figure 10.1, the response variable is set to "systolic" and the explanatory variable is set to "diastolic," referring to systolic and diastolic blood pressure respectively, as noted in the Codebook that is always available in the app. The upper left-hand corner of Figure 10.1 shows the sample size is currently set to 50, and a portion of the data is

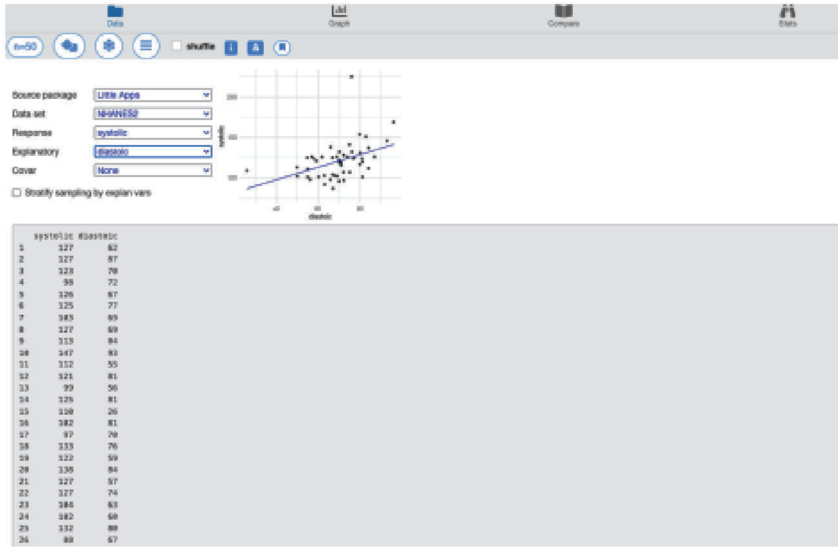


Figure 10.1: Landing page for Regression Modeling Little App.

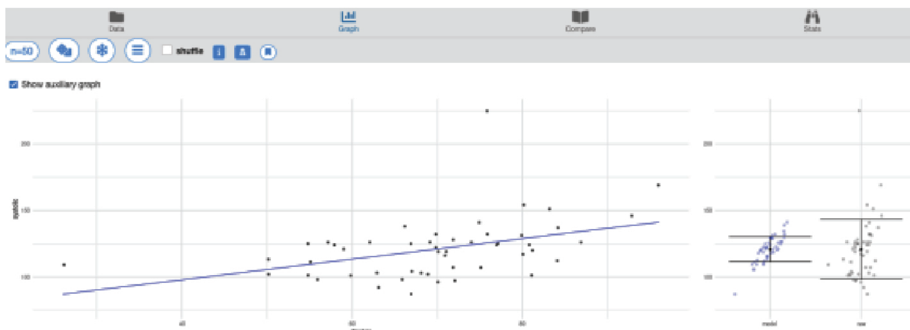


Figure 10.2: The "Graph" option enables the user to enlarge the scatter plot and shows an auxiliary graph that represents a visualization of the correlation coefficient r.

visible in the figure. However, the entire sample is available to be copied and pasted into any preferred spreadsheet or data analysis software. The Little App allows you to change the sample size, and to resample by clicking on the image of the dice. Figure 10.2 shows an enlarged version of the scatter plot, which can be achieved by selecting "Graph" from the top menu bar.

Figure 10.2 also includes a unique feature of the Little App that provides a visualization of the correlation coefficient r, shown as the ratio of the length of the model value interval (left-most I bar) to the length of the raw values interval (right-most I-bar) (see [25]). In addition, it is important to note that the correlation coefficient r is typically utilized as a means to quantify the strength of the relationship between two quantitative variables. However, it is also applicable in situations where one or both of the variables are dichotomous (e.g., yes/no or win/lose or A/B). Many variables in the NHANES2 data are dichotomous, and thus regression modeling techniques can be utilized.

Figure 10.3 shows a useful Little App feature that allows users to compare two scatter plots in a side-by-side format. After selecting the snowflake button from the menu, the first sample is frozen. Then a second sample can be drawn and the "Compare" feature selected. The scatter plots reveal a similar trend, although with many variables the trend is likely to change with subsequent resampling. As illustrated, the plots also display an auxiliary plot of the correlation coefficient r, aiding students in drawing conclusions about the relationship between the variables.

Figure 10.4 shows the comprehensive model statistics provided by the Regression Modeling Little App when the user selects the "Stats" button at the top-right of the menu. It reveals many more metrics than those usually taught in typical introductory statistics courses which are usually focused on R^2 (R-Sq) and the values shown in the regression table.

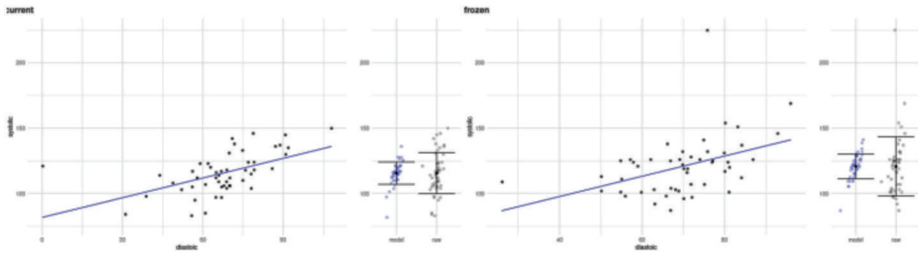


Figure 10.3: Comparing two samples side-by-side using “Freeze” and “Compare” to illustrate sample variability.

current

- n = 50
- dflex = 1
- var_raw = 247.9
- var_model = 72.84

thus ...

- R-sq = 0.294
- F = 20

```
Regression table
             Estimate Std. Error t value Pr(>|t|)
(Intercept)    81.800      7.810   10.50 5.45e-14
diastoic        0.502      0.112    4.47 4.79e-05
```

```
Analysis of Variance Table
Response: systolic
          Df Sum Sq Mean Sq F value    Pr(>F)
diastoic   1 3569.2  3569.2  19.971 4.789e-05 ***
Residuals 48 8578.6   178.7
---
```

frozen

- n = 50
- dflex = 1
- var_raw = 514.6
- var_model = 90.05

thus ...

- R-sq = 0.175
- F = 10

```
Regression table
             Estimate Std. Error t value Pr(>|t|)
(Intercept)    66.700     17.200    3.88 0.000321
diastoic        0.776      0.243    3.19 0.002500
```

```
Analysis of Variance Table
Response: systolic
          Df  Sum Sq Mean Sq F value   Pr(>F)
diastoic   1  4412.4  4412.4  10.181 0.002502 **
Residuals 48 20802.1   433.4
---
```

Figure 10.4: Regression model statistics are provided through the Stat feature of the Regression Modeling Little App.

Figure 10.5 shows the option to add a covariate to the regression to support students’ multivariable thinking. The data is stratified by sex (female/male), and adding covariates is accomplished on the Data page by adding a variable in the “Covar” field (see Figure 10.1). Although the NHANES limits “sex” to a dichotomous variable indicating biological sex, I recommend discussion with students about more current approaches to collecting data on sex and gender. As can be seen in Figure 10.5, the linear trend line is shown for females (red) and males (blue).

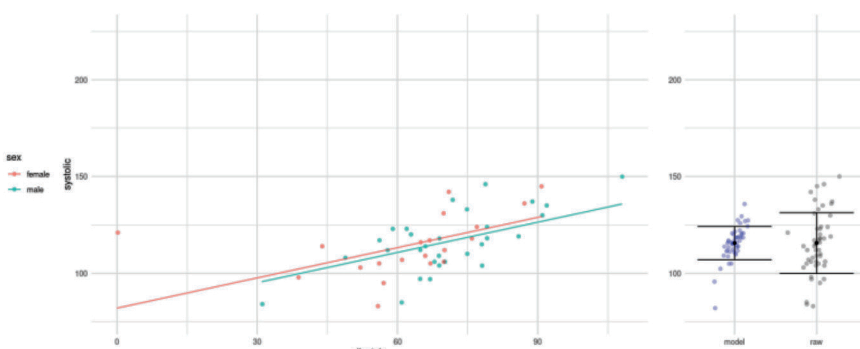


Figure 10.5: The covariate “sex” (female/male) is added to the regression on the Data page under Covar.

10.4 Teaching Opportunities and Challenges

10.4.1 Concepts Students Struggle With When Learning Linear Regression

When learning linear regression, students often have difficulty understanding which variable should be the explanatory and which variable should be the response variable, and reasoning simultaneously about model statistics such as r, R^2, and b as they assess the suitability of the linear model for a given data set. It is noteworthy that real-life correlations based on a single explanatory variable can tend to be weak or moderate, which is an additional fact that students should consider when using the regression model for prediction. Furthermore, despite students' exposure to linear equations and functions in their prior mathematics courses, students have difficulty interpreting the growth rate, b, of the regression model in terms of the relationship between incremental change in x and the corresponding incremental change in y with appropriate units of measure; and translating a difference in the input to the corresponding difference in the output. Finally, students face additional challenges in courses where residual analysis and hypothesis testing are included. If our instructional goal is to increase students' capacity to reason about data, the challenges identified here need to be addressed.

10.4.2 How We Can Bridge Gaps in Student Knowledge

The challenges we identified can be addressed by adopting several recommendations from the StatPREP project team. The first recommendation is to modify the way the topic of linear regression is typically motivated. Specifically, the approach in traditional statistics textbooks is to begin by selecting two variables, x and y, and ask if there is a bivariate relationship between the two variables. In contrast, an approach that is effective and more authentic to the way linear regression is used in practice is to select a specific response variable of interest, y, and ask: What predicts y? The regression Little App accommodates this approach by placing the response variable field before the explanatory variable (Figure 10.1). This approach prompts important discussion about which variable should be the explanatory and which variable should be the response variable.

A second recommendation is to invest more effort from the start in providing an overview of the concepts around linear regression using real data and the Graph and Data features available in the Little App. This includes selecting variables, resampling, examining scatter plots and auxiliary graphs, and the model statistics. An authentic and holistic dive into the topic should expose students to the various concepts, tools, and necessary vocabulary to assess the suitability of a linear regression model. It is recommended that teachers provide students with practice on the individual components, which will further aid in their overall understanding of the linear regression concept. Additionally, teachers should ensure that students are aware of the meaning of superscripts (e.g., R^2) and subscripts (e.g., x_i). Those who utilize Greek symbols in teaching regression should ensure that students have a knowledge of their Roman equivalents (e.g., β and b). The StatPREP Little Apps use the following symbols.

- R^2 to represent the coefficient of determination
- r to represent the correlation coefficient
- S_x and S_y for the standard deviations of the x and y variables, respectively
- b to represent the linear growth rate

Finally, additional attention should be given by instructors to nomenclature and how mathematicians and statisticians use the terms parameter and coefficient differently. As mentioned previously, mathematicians generally describe a and b in a formula like $a + bx$ as "coefficients," but a "coefficient" in a formula like $a + bx$ is not particularly similar to a "correlation coefficient r". The correlation coefficient r ($r = bS_y/S_x$) is a result that combines three pieces of information: the growth rate, b, and the standard deviations, S_x and S_y.

10.5 Activities Using the Regression Little App

This section presents brief descriptions of three activities for teaching concepts around linear regression. The StatPREP Hub Community Library contains full activities around regression for you to print and use.

Activity 1

Activity 1 can occur early in the treatment of linear regression and is intended to motivate the overall topic and introduce students to scatter plots and vocabulary associated with linear regression.

Goals

- Motivate the concept of linear regression
- Create scatter plots for bivariate data

- Introduce vocabulary and concepts associated with simple linear regression
- Interpret the correlation coefficient and coefficient of determination

Instructions

Introduce linear regression to students by asking questions such as, *What predicts a person's Body Mass Index (BMI)*? Possible predictors that can be explored in the NHANES2 dataset include weight, height, blood pressure, and hours slept per night. As you explore possible explanatory variables, you could lead a class discussion by demonstrating the following activities.

- Resample from the data set to help students visually recognize statistical concepts and variables for which regression is an appropriate technique and also reinforce the concept of sample variability.
- Demonstrate how to use the "Stats" button to compute the regression model statistics (r and R^2) and interpret them in terms of pos/neg/null and the strength of the linear correlation.
- Ask the question, *What real-world phenomenon accounts for the correlation? What are examples of lurking or confounding variables?*
- Include a covariate (e.g., sex) to encourage multivariable thinking.

This activity represents a shift away from the traditional approach of asking whether there is an association between bivariate data. The shift opens opportunities for students to explore data to find possible predictors, or explanatory variables, for Body Mass Index (BMI). Students get more engaged after discovering that BMI is a controversial statistic that was first developed by Adolphe Quetelet (1796 – 1874) and continues to be a meaningful measure despite criticisms ([193]; [43]).

Activity 2

The purpose of this activity is to provide students with an opportunity to actively explore data and improve their skills in assessing and evaluating the suitability of a linear model for a given data set. Students should evaluate the model and draw conclusions using the statistics r, R^2, and b. The activity could be modified to be suitable for courses that include analysis of residuals and hypothesis testing.

Goals

- Use technology to find a linear regression model and model statistics (r, R^2, and b).
- Guide students to interpret the results of the model statistics.
- Use the regression equation for prediction.

Instructions

Select a dataset in the Regression Modeling Little App and a response variable of interest. Find two explanatory variables that might predict the response variable. For example, students might work with the Natality_2014 data set and choose to investigate Mother's age at birth or length of gestation as explanatory variables for baby's birth weight. Based on resampling and model statistics, determine which explanatory variable is a better predictor of the response variable. Write the linear regression equation using the stronger correlation and use the equation to make one prediction.

In the activity, the regression equation should be used for prediction, and students should be able to explain the direction (positive, negative, or null) and strength of the relationship between a pair of variables, and the proportion of variance in the response variable predicted by the explanatory variable. Additionally, students should be able to explain the growth rate of the regression model by describing the change in x in relation to the corresponding change in y.

Students respond well to exploratory activities such as this although they may need scaffolding around how to interpret the model statistics. I provide students with templates for interpreting each statistic which they can use to write their own interpretations. Students may gravitate towards qualitative ordinal variables in the data set, so be prepared for that. Overall, an activity of this type is key to helping students make sense of how to use the different model statistics to draw conclusions about the relationship between the explanatory and response variables.

Example Template: *The r value is ______, suggesting the strength of the model is weak, moderate, strong (circle one).*

The explanatory variable ______ explains ______ percent of the variation in the response variable ______.

As the explanatory variable ______ increases by _________ the response variable ______ will go up-or-down (circle one) by _________.

In filling in the two blanks following the word "by," make sure to give the units of the variables. You can find the units by looking at the codebook.

Activity 3

This activity focuses specifically on interpreting the growth rate, b, of the regression model. As mentioned previously, students are often challenged when asked to apply their prior learning of linear growth to linear regression in introductory statistics courses. The growth rate poses additional challenges compared to other model statistics because units of measure are involved. Ultimately, however, interpreting b is essential to interpreting the relationship and it serves as an effect size of the explanatory variable on the response variable. In the activity below, students are provided with a template to interpreting b to scaffold their learning.

Goal

- Interpret the growth rate, b, of the regression in terms of the relationship between incremental change in x and the corresponding incremental change in y and translate a difference in the input to the corresponding difference in the output.

Instructions

1. Choose a response variable of interest.
2. Find an explanatory variable where the regression line slopes up.
3. Find an explanatory variable where the regression line slopes down.
4. For each variable, find the numerical value of the growth rate, b, of the line. Then summarize the relationship in the way shown in the template below.

Example Template: *As the explanatory variable ______ increases by _________, the response variable ______ will go up-or-down (circle one) by _________.*

In filling in the blank following the word "by," make sure to give the units of the variables. You can find the units by looking at the codebook.

Templates such as the ones above have increased my students' ability to interpret the model growth rate, b, and have resulted in an overall improvement in assessment scores in this traditionally challenging topic.

10.6 Assessment

The activities below offer ideas for formative and summative assessments.

1. Provide each group of students with three to five scatterplots created on the Little App so the groups can discuss which have a positive, negative, or no association.
2. Transform the Little App output into a card sort in which students match scatterplots with model statistics.
3. Provide two explanatory variables for a response variable and have students compute the model statistics. They should then determine which is the better predictor and explain their conclusion based on the data, including doing one prediction with the model.
4. Have students use the Little App to generate computer output including model statistics from a given explanatory and response variable. They should then summarize the results and assess the suitability of the linear model for the two variables.
5. Provide students two variables to explore and ask them to decide which should be the response variable and which should be the explanatory variable. Using the Regression Modeling Little App, they should then compute the model statistics (r, R^2, and b) and interpret the model statistics using the templates or model sentences developed in class.

10.7 Conclusion

This chapter showed how the topic of linear regression in an introductory statistics course can be elevated and enriched through the use of the StatPREP Regression Little App. The chapter not only provided insight into how modern data-science approaches to the topic differ from traditional approaches, but also offered ideas for classroom activities and assessments that can improve student learning of linear regression.

11

Probability and Probability Distributions

Joe Roith, *St. Olaf College*
Maria Tackett, *Duke University*

11.1 Introduction

Recent innovations in Introductory Statistics courses recognize the need to prepare students for the modern world—a world in which they interact with data and statistics on a daily basis. With this new data landscape in mind, there has been a move to focus more on teaching students methods for developing statistical literacy and conceptual understanding rather than formulas and procedures. The 2016 GAISE College Report [67] highlights this focus in their first two recommendations for introductory statistics courses:

1. Teach statistical thinking.
2. Focus on conceptual understanding.

Similarly, the 2014 Curriculum Guidelines for Undergraduate Programs in Statistical Science state that undergraduate statistics programs "should emphasize concepts and approaches for working with complex data" [9].

Recent textbooks such as *Introduction to Modern Statistics* by Mine Çetinkaya-Rundel and Johanna Hardin [34], *Introduction to Statistical Investigations* by Nathan Tintle et al. [189], and *Statistics: Unlocking the Power of Data* by Robin Lock, et al. [117], follow this concept-driven approach as they use simulation-based methods and present few equations to introduce and explain many of the concepts traditionally taught in introductory statistics courses.

While this emphasis on conceptual understanding applies to all topics in introductory statistics, this chapter focuses specifically on how it can influence the way probability and probability distributions are taught in these courses. Under this framework, the learning outcomes for probability and probability distributions are multifold and are motivated by the goal to prepare students to more effectively understand statistical inference. The extent of coverage of these topics varies by the institution; however, based on our experience and those of colleagues at multiple institutions, we have found the following to represent typical learning outcomes for probability and probability distributions in introductory courses.

Students are able to

- *understand and apply probability rules to answer questions about data;*
- *understand general probability concepts, such as intersections, unions, and the complement rule;*
- *define the appropriate conditional probability calculation needed to answer a specific research question, calculate the probability, and interpret the result;*
- *explain the difference between continuous and discrete distributions and identify when each might be appropriate to model a given phenomenon;*

- *identify and be comfortable calculating probabilities from probability distributions needed for inference taught in the introductory courses: Normal and t-distributions (and also χ^2 and F distributions if needed in the course);*
- *use technology (whatever technology is used in the course) to calculate probabilities and probability distributions.*

In this chapter, we outline some of the primary challenges and opportunities for teaching probability and probability distributions. We discuss common issues such as determining the amount of time to dedicate to such topics, strategies for focusing on conceptual understanding over formulas and equations, and how to teach these topics using technology that can be used to analyze large data sets. We also address how this framework can be implemented and adapted, even in courses under strictly defined institution and/or state-wide curriculum guidelines. Since the topics of probability and probability distributions are not directly addressed by the StatPREP LittleApps, we provide examples demonstrating these ideas.

11.2 Challenges and alternative approaches

11.2.1 Challenge: Time restrictions

Introductory statistics (and data science) courses serve broad and diverse student populations who are interested in a range of academic disciplines with a variety of motivations for taking the course. These courses, then, are designed to prepare this diverse student population for a wide range of purposes, from engaging with statistics as a citizen in the modern data-driven world to applying statistical methods to conduct research in their field to pursuing further study in statistics in more advanced courses. Given the different student motivations for taking the course, along with required content from the institution, state, or other administrative authority, it is no surprise that the curriculum for the introductory course is packed full of content that at times can feel rushed in a typical 11–16 week term.

In a standard introductory statistics course, an in-depth understanding of probability and probability distributions is often not a learning outcome students are expected to achieve by the end of the term. It is usually much more important that students have a rich understanding of topics such as exploratory data analysis using visualizations and summary statistics, conducting statistical inference to draw conclusions about a population, developing statistical literacy, and being "critical consumers of statistically-based results" as described in the 2016 GAISE College Report. Therefore, the amount of class time and the depth to which these topics are taught and assessed need to reflect this balance in the desired learning outcomes.

Part of balancing content to reflect the desired learning outcomes is the consideration of students' cognitive load. *Cognitive load* is described in Teaching Tech Together by Greg Wilson as "the mental effort needed to solve a problem" [209]. Wilson explains the three components of cognitive load theory: (1) what students have to keep in mind to learn new material (intrinsic load); (2) mental effort to link new information to old (germane load); and (3) anything that distracts from learning (extraneous). Intrinsic and germane load are required for students to actually learn and not just memorize new material, so the instruction must focus students' mental efforts on intrinsic and germane load (e.g., learning, connecting, and applying new concepts) while minimizing the mental effort on extraneous load (e.g., memorizing too many terms and formulas) to more effectively teach this content.

Students are able to retain only a finite amount of information, and there are limits to the amount of content students can feasibly learn in a single course, especially when students are taking multiple courses in a given term. It is not beneficial to students to try to cover as much content as possible in an introductory course, so it is imperative to be mindful of how much we expect students to learn, and hence the amount of content we teach.

Taking all these factors into account leads to the first challenge of teaching probability and probability distributions in introductory courses: the time restriction does not allow for teaching these topics in full. Many mathematics and statistics curricula at the university level offer entire courses on probability and probability distributions, so introductory courses should streamline the content to what is most important for this student population to learn.

11.2.2 Alternative approach for teaching probability

Probability is often introduced as a set of rules and equations rather than methods that can be used to understand random phenomena and statistical inference. There is an opportunity to focus on a more conceptual understanding of probability and ultimately help students better understand statistical inference. Keeping this in mind, we discuss some strategies for streamlining probability content to cultivate an understanding of basic probability rules and conditional probability.

Use data and contexts that are relatable to students. Many students are using probabilistic thinking in their everyday lives (sometimes unknowingly), so using relatable examples can help them draw from their intuition as a starting point for understanding probability.

Use words in place of symbols when possible. The overuse of symbols adds to the students' extraneous cognitive load, as it adds to the amount of new information they need to understand without necessarily deepening their understanding of the key ideas. We encourage instructors to expose students to these mathematical symbols, so students at least recognize them in subsequent statistics classes, but without an expectation that students are able to fluently read and produce such symbols. Table 11.1 shows how common probability topics can be introduced using only mathematical symbols, a combination of words and symbols, and words only.

Concept	**Avoid**	**Better**	**Best**
Intersection	$P(A \cap B)$	$P(A$ and $B)$	Probability of A and B
Union	$P(A \cup B)$	$P(A$ or $B)$	Probability of A or B
Complement	$P(A^c)$	$P(A$ *complement*$)$	Probability of A complement (not A)
Conditional	$P(A\|B)$	$P(A$ given $B)$	Probability of A given B

Table 11.1: Use words rather than symbols, when possible

Motivate probability with relevant questions. Introduce probability using a compelling analysis question and explain how probability will provide a way to quantify and describe the random phenomena that students will eventually use to draw conclusions about the analysis questions. Additionally, end probability exercises with probing questions that encourage students to think ahead toward statistical inference. For example, we could ask: "Are attitudes about climate change different among current generations?" Students may have preconceived ideas about this question prompting a discussion that they can investigate with data. We will revisit this in the next section with an example involving survey data about attitudes on climate change.

Focus on Conditional Probability. Students need a basic understanding of conditional probability, so they can more fully understand the definition of p-values (the probability of obtaining the observed result or a more extreme one given the null hypothesis is true) later on as they learn statistical inference. Similar to introducing probability rules, conditional probability is often introduced by presenting the formula

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

and having students plug in the values for $P(A \cap B)$ and $P(B)$ to calculate an answer for $P(A|B)$. While computing the conditional probability using the provided formula and value may seem straightforward for students, they often have difficulty describing what the formula and respective calculations mean. One common way to encourage conceptual understanding is to use a Venn diagram to provide students a visualization of what "A given B" looks like; however, using Venn diagrams in conjunction with formulas still requires an additional cognitive step to connect the visual understanding from the Venn diagram to the components of the formula.

When the primary goal is conceptual understanding, students can be exposed to the formula for conditional probability but engage with the calculations in a more intuitive form, such as using a two-way table. Using a two-way table like the one shown in Table 11.2, students can draw upon their intuitive understanding as they can visualize the values being calculated in the conditional probability. Thus, students can focus more on the conceptual understanding of the calculation steps rather than dedicating mental load to explaining formulas and matching formulas to diagrams. The visual benefits of the two-way table can also help students describe the differences between marginal, joint, and conditional probabilities, i.e., $P(A)$ versus $P(A \cap B)$ versus $P(A|B)$.

Below is an example activity demonstrating how a two-way table can be used to teach basic probability rules and conditional probabilities, incorporating the strategies mentioned above. The scenario and table are from Exercise 2 in Chapter 18 of Introduction to Modern Statistics. The data are originally from a 2021 study by the Pew Research Center, in the article "Key findings: How Americans' attitudes about climate change differ by generation, party, and other factors" [66]. The exercise solutions are available in the supplemental document Teaching Probability Example.

Generation	Took Action	Didn't Take Action	Total
Gen Z	292	620	912
Millennial	885	2,275	3,160
Gen X	809	2,709	3,518
Boomer & older	1,276	4,798	6,074
Total	3,262	10,402	13,664

Table 11.2: U.S. Adults, personal responses to climate change by generation

Example 11.1. Over 13,000 U.S. adults were asked whether or not they have taken personal action to address climate change. The responses and generation of survey respondents are shown in Table 11.2. The goal of this analysis is to determine if generation and action on climate change are independent, i.e., there is no association between generation and action on climate change.

1. Before looking at the data, do you think generation and action on climate change are independent? Why or why not?
2. What is the probability a randomly selected respondent is part of Gen Z?
3. What is the probability a randomly selected respondent said they have taken personal action on climate change?
4. What is the probability a randomly selected respondent is part of Gen Z **and** said they have taken action on climate change?
5. What is the probability a randomly selected respondent said they've taken action on climate change **given** they are part of Gen Z?
6. What is the probability a randomly selected respondent is part of Gen Z **given** they said they have taken personal action on climate change?
7. Do the data suggest generation and action on climate change are independent? Explain your response.

11.2.3 Alternative approach for teaching probability distributions

Instructors may feel that to understand inferential procedures, students must understand the workings of the theoretical distribution upon which each test is based. This is typically not the case, as an understanding of the general structure of inferential procedures, such as the test statistic and p-value can be applied without knowledge of probability distributions. The 2016 GAISE College Report recommends less emphasis on probability theory and we recommend extending that to distributions as well, with the exception of special cases, like the binomial distribution, when required.

This provides an opportunity we rarely see in our statistics courses: the chance to simplify or even eliminate content. Juggling all of the recommendations makes it difficult to find room to include topics like randomization tests and simulation-based inference. But, as Utts suggests in *The Many Facets of Statistics Education: 175 Years of Common Themes*, students often are taught too much hypothesis testing without contextualizing the important ideas of practical significance, power, and effect size [196]. We believe the same is true for probability and probability distributions.

There is an opportunity here to decrease the emphasis or eliminate the teaching of theoretical distributions in our introductory classes. When we think about what a typical student needs to know from our course, it is not the parameters of the F-distribution, or even of the Normal distribution (in the theoretical context). By reducing this content, we create time to work on data literacy skills such as causal inference, effect size and power implications on significance, and interpretation of results.

In *Challenges and Opportunities for Statistics and Statistical Education: Looking Back, Looking Forward* Horton recommends "Increasingly, students need experience analyzing larger, real-world datasets and to be aware of what techniques do and do not scale well" [95]. For example, rather than having students practice "reading a Normal distribution" in an abstract way (P(Z $>$ 1.5)), we should allow them to explore a real variable such as the age of mother at birth. This data resembles a Normal distribution, and probabilities can be calculated in the context of a real-world application. The symbolic representation of probability with context, such as $P(\text{age} > 32)$ is a better connection between data and theory, but using language like "what is the probability that a mother will be over the age of 32" can force students to think critically about what is really being asked.

Exploring sampling distributions through simulations or null distributions through randomization can provide a natural approach to understanding probability's role in statistical inference. A more detailed description of the randomization and the shuffling process can be seen in Sections 11.1-11.2 of Introduction to Modern Statistics. As a simple example, randomly shuffling treatment and control assignments in an experiment can show students the null distribution of samples assuming there is no relationship between variables (see Example 11.2). In an activity such as this, students are able to see the true treatment provided to individual cases in the data (Treatment vs. Control). Then, with the help of statistical software or by providing the simulations, the treatments are randomly rearranged among the measurements. This process mimics the null hypothesis that both treatments provide the same effect by shuffling the values as if they all came from the same population. By calculating the difference of the means between these shuffled groups and repeating the process many times, students actively "build" the null distribution. Then, by calculating the proportion of simulated null samples that recreate the observed difference between control and treatment groups, we can illustrate p-values in a more intuitive way. After several examples like this, a natural observation is the consistent shape of these null distributions from completely different examples. This provides an organic segue into discussions about probability distributions such as the normal and t-distribution. Examining examples of these distributions makes it simple to attach a name to them and even discuss some general properties, such as parameters for the center and spread. Once this level of understanding has been reached, there is no benefit to increasing the repertoire of distributions to the students' toolbox; they can understand the role of the test statistic in a chi-square test of independence or F-test without knowing the theoretical distributions at all. This approach reduces cognitive load and actually emphasizes the connection between inferential procedures rather than promoting them as disjoint sets of concepts. It also aligns with the idea Gould lays out in *Statistics and the Modern Student* of teaching by defining the data, creating critically thinking "citizen statisticians," and teaching with technology [77].

11.3 Additional challenges

11.3.1 Effectively teaching p-values

Students will almost certainly encounter p-values in their subsequent courses, research, work, and daily lives. Even within the introductory course, many curricula include topics such as χ^2 tests that rely on p-values. As such, it is important that students have an understanding of what p-values mean, how to effectively use them to draw statistical conclusions, and the limitations of reliance on this single statistic to draw scientific conclusions. This section will focus on effectively teaching p-values and how it relates to teaching probability and probability distributions. Later chapters in this volume will provide a more in-depth discussion on effectively teaching statistical inference as a whole.

The same "focus on conceptual understanding" approach for teaching probability and probability distributions can be applied to teaching p-values as well. At this point in the curriculum, students have gained some experience with probabilities and using simulation to understand probability distributions. In each of these topics, the emphasis was on a conceptual understanding of the content, so this will provide the foundation for students to gain an understanding of calculating, interpreting, and using p-values.

As with probability rules, p-values are often introduced to students in terms of a formula such as p-value $= P(T \geq t | H_0 \textit{ true})$, which leads to more formulas to calculate test statistics, such as $T = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}$. Though most students will be able to identify each part of the formula, it requires a high cognitive load for those learning p-values for the first time to combine the components of the formula to construct a definition of the p-value. As with teaching probability, one way to reduce the cognitive load is to use words rather than formulas when describing what the p-value is and how to calculate it. For example, the p-value could instead be introduced as "given the null hypothesis is true, the p-value is the probability of observing a statistic [mean, proportion, slope, etc.] at least as extreme as the one observed in the data." While this can still feel like a fairly complex definition, it removes the extra step of translating the formula into words and eventually translating it into calculations.

Additionally, students can use simulation to generate the null distribution and thus have a visual of the distribution and what they are calculating to obtain the p-value. Using the simulated null distribution, students are able to calculate the p-value as the number of simulations that resulted in a statistic "at least as extreme" as the one observed in the data.

At this point, they have learned how to calculate probabilities from a simulated distribution, so more attention can be paid to what is meant by "at least as extreme" in the context of the hypothesis test being conducted (For example, see Example using StatKey). As with probability rules, we still encourage instructors to show students the formulas, so they have at least been exposed to them if they encounter the formulas in future work.

Below are portions from Exercise 8 in Chapter 11 of *Introduction to Modern Statistics* (with some adaptations) illustrating the simulation-based inference, the emphasis on the conceptual understanding of p-values, and how they are applied in practice.

Example 11.2. The Stanford University Heart Transplant Study was conducted to determine whether an experimental heart transplant program increased lifespan. Each patient entering the program was designated an official heart transplant candidate, meaning that they were gravely ill and would most likely benefit from a new heart. Some patients got a transplant and some did not. The variable *transplant* indicates which group the patients were in; patients in the treatment group got a transplant and those in the control group did not. Of the 34 patients in the control group, 30 died. Of the 69 people in the treatment group, 45 died. Another variable called *survived* was used to indicate whether the patient was alive at the end of the study. Figure 11.1 depicts this information.

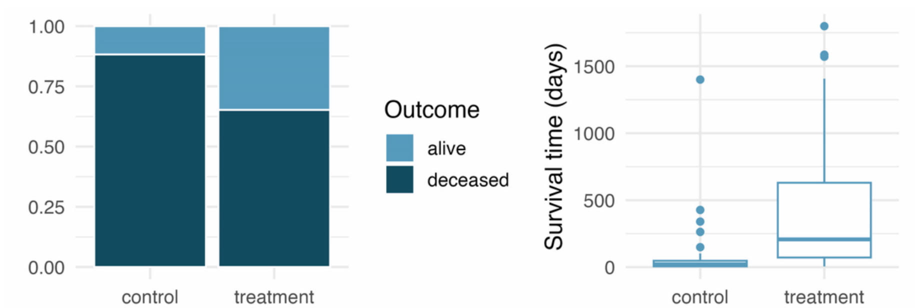


Figure 11.1: Stanford University Heart Transplant Study Data

1. Does the stacked bar plot in Figure 11.1indicate that survival is independent of whether the patient got a transplant? Explain your reasoning.

2. What do the box plots in Figure 11.1 suggest about the efficacy (effectiveness) of the heart transplant treatment?

3. What proportion of patients in the treatment group and what proportion of patients in the control group died?

4. One approach for investigating whether the treatment is effective is to use a randomization technique. Considering the patients who received transplants and those who did not, what are the claims being tested? Design a process that does not require statistical software to test using a randomization approach.

5. What do the simulation results shown in Figure 11.2 suggest about the effectiveness of the transplant program? Use the distribution to calculate a p-value, then use the p-value to determine and support your response.

The goal of this exercise is for students to use exploratory data analysis and hypothesis testing via randomization to draw conclusions about the effectiveness of the transplant program. At this point, students have had some exposure to randomization when they learned probability distributions, as described in the previous section. The exercise starts with questions 1 and 2 by having students use visualizations to develop initial insights about whether survival is independent of treatment. From the visualization in question 1, students can observe that there is a higher proportion of patients in the treatment group who survived compared to the control group. From this observation, students may determine that survival and treatment do not appear to be independent. If students are unsure how to use their observations from the plot to make an assessment about independence, the instructor can help guide them by asking the probing question, "Does knowing a patient is in the treatment or control group tell you something about the probability they survived?"

Question 2 asks students to make an initial assessment about the program's effectiveness using new information about the survival time. From the box plots, students can observe that outside of a few outlier observations in the control group, patients in the treatment group generally lived much longer than those in the control group.

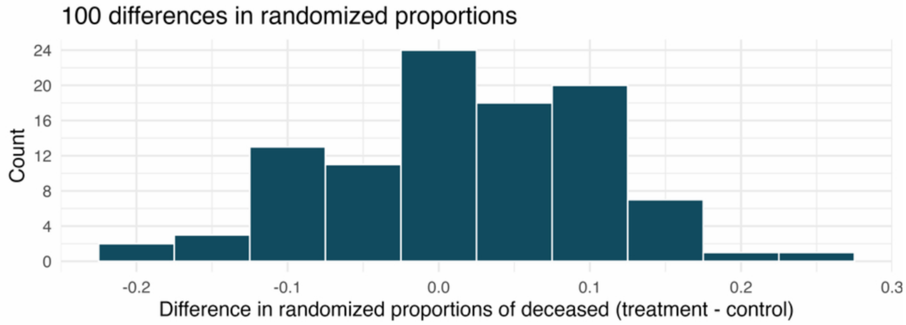


Figure 11.2: Heart Transplant Study Simulation Results

Once students have developed some initial intuition about the effectiveness of the treatment, they begin calculations in question 3 using the observed data. From the exercise description, they calculate that the proportion of patients who died in the treatment group is $\frac{45}{69} = 0.652$ and the proportion in the control group is $\frac{30}{34} = 0.882$. At this point, the instructor may also ask students to calculate the difference in proportions (*treatment* - *control*) $= 0.652 - 0.882 = -0.23$, as they will need this value in question 5 to calculate the p-value.

Now that students have examined the observed data, they conduct a hypothesis test in questions 4 and 5. Question 4 asks students to specify the null and alternative hypotheses. Students can be prompted to think of the null hypothesis as the "status quo" or "there is nothing going on" hypothesis. In this scenario, it means there is no difference in survival outcomes between the treatment and control groups, and thus survival is independent of whether a patient received a transplant. The alternative hypothesis, then, is that survival is not independent of whether a patient received a transplant.

Students draw their final conclusions based on the null distribution produced by the randomization test shown in question 5. To calculate the p-value, instructors can start by having students observe that the statistic used for this test is the difference in proportions and think about what the center of the null distribution—i.e. the difference in the proportion of patients who died between treatment and control—is expected to be based on the null hypothesis from question 4. The center of the distribution is expected to be 0, and students will observe this consistent with the graph shown in part 5. Next, students can mark where the observed difference in proportions (-0.23) lies on the histogram of the null distribution.

Now students are ready to estimate the number of times one would expect to get a difference in proportions of -0.23 or better (less than -0.23) under the null distribution. Though this value can't be calculated exactly from the graph, students can estimate the p-value will be less than or equal to $\frac{2}{1000} = 0.002$. This is a very small p-value. Students then assess whether the small p-value indicates the observed data are consistent with the null distribution and thus the null hypothesis. The lack of consistency will prompt students to conclude that the data are more in favor of the alternative hypothesis, and there is evidence that survival and treatment are not independent. Furthermore, using the observations from questions 1 and 5, they can conclude that there is evidence that the transplant program (treatment) is effective, as the visualizations showed lower death rates and longer survival times among patients in the program.

As the example illustrates, this approach to teaching p-values in a way motivated by relevant questions and drawing on insights from initial data analysis encourages students to think more critically about p-values rather than memorizing a set of formulas. Additionally, because students have an understanding that is built on both intuition and previous skills learned in the course, they have more opportunity to think critically about the practical implications of their results rather than ending at the point of a result being "statistically significant" or not. As indicated by the multiple articles, statements, and reports about the misuse of p-values in scientific research, this ability to think critically about the results and their potential implications is more important than being able to execute calculations from a series of equations. See the supplemental document Notes on p-values for more information on the conversation about p-values in the scientific and statistics communities.

11.4 Preparing students for the next course

There may be a fear that by removing some of the in-depth explorations of probability, logic, probability distributions, and similar topics, we will be doing a disservice to those students who will move on to an intermediate course in statistics or probability. The concern is that those students will not receive the depth of knowledge in these topics if the introductory course focuses more on breadth. Students who are interested in the theoretical details of these topics might be bored or disengaged. Even students who do not pursue additional statistics courses may encounter the technical terminology of probability and distributions later in their undergraduate studies or careers. If these are left out of an introductory statistics course, these students will be at a disadvantage when encountering these concepts in later courses.

However, we argue that for students who are inclined to further their statistical education, many of the probability terms and concepts traditionally taught in an introductory statistics course can more easily be understood by motivated students in an intermediate statistics course. A simple primer of the theoretical topics can be supplemental materials for these students, rather than required materials for all students in an introductory course. In *What Educated Citizens Should Know About Statistics and Probability*, Utts points out that "many who take our large introductory statistics courses will never actually do statistical analysis of their own" [197]. And for those who will encounter these terms later in their careers, we question whether they have the motivation to retain these abstract concepts beyond the introductory class.

Rather than viewing the removal of such topics as a disservice to the few students who may choose to pursue further statistics courses, we see this as an opportunity to engage and inspire the many more students who suffer anxiety when confronted with mathematical formulas, theorems, and abstract logic. The chance to show probability and distributions in an applicable way will be a service to the majority of students in an introductory statistics course. It rightly places the emphasis on exposing students to the concepts and ideas they will most likely encounter in the format with which they are most likely to interact. Exposure to greater breadth will draw students in, rather than turn them away with technical depth.

11.5 Current Landscape: External curriculum guidelines

We acknowledge that many states and university systems require institutions to cover specific topics in an introductory statistics course in order for credits to transfer to the larger institution. This often includes several probability and distribution concepts that we recommend de-emphasizing. We consider the guidelines for two states, Maryland and California. In the University system of Maryland, the general statistics studies course description begins with:

A first college-level statistics course provides students with a basic understanding of statistics and prepares them to solve problems that involve collecting and analyzing meaningful data. This includes the study of measures of central tendency, measures of variation, graphical representation of data, least squares regression, correlation, ***probability distributions*** *[emphasis added], sampling techniques, parameter estimation, and hypothesis testing. Technology and statistical literacy will be integrated throughout the course.*"

The document goes on to list "perform elementary probability calculations and solve problems by applying appropriate standard probability distributions, including discrete, binomial, uniform and normal distributions" as an intended learning outcome.

In the California State University system, the C-ID Descriptor Documents are "posted for general use after a statewide review by the discipline faculty who teach these courses in California's public post-secondary institutions. The descriptors are offered to encourage wider articulation and to expand the lower division curricular offerings and thereby increase the variety—and ensure the rigor—of our many Community College courses." For statistics courses:

"The use of ***probability techniques*** *[emphasis added], hypothesis testing, and predictive techniques to facilitate decision-making. Topics include descriptive statistics;* ***probability*** *[emphasis added] and sampling distributions; statistical inference; correlation and linear regression; analysis of variance, chi-square and t-tests; and application of technology for statistical analysis including the interpretation of the relevance of the statistical findings. Applications using data from a broad range of disciplines.*"

The document goes on to list binomial distributions, normal distributions, as well as discrete distributions, and expected values as content for an introductory statistics course.

Each state and system has its own set of guidelines. This makes it challenging to create a consistent and pedagogically sound approach to teaching statistics at the introductory level. But we believe that our suggestions can be implemented thoughtfully and responsibly to meet such guidelines and still place the right emphasis on teaching a data-centric statistics course.

The opportunities presented here guide instructors in carefully planning how they will incorporate these topics into their courses in a meaningful way. The historic temptation, usually spurred on by outdated textbooks, is to create separate units or modules in a course that address probability and distributions. As stated above, this provides a disjoint experience for students. They may begin the semester interacting with real data, summarizing and visualizing data sets that have a context to which they can relate—but before continuing with inference and modeling of this real data, they spend a week or more learning about abstract probability rules and theoretical distributions. The understandable result is that students often find it difficult to connect the two worlds of theory and application. At the very least, when required to include these topics, instructors should focus on the applications of such concepts as they relate to the major themes of the introductory class.

Most guidelines include a statement on using technology as a tool for teaching. We recommend leveraging this for probability as well as for exploratory and inferential techniques. Teaching probability and conditional probability in the context of p-values can be a useful application while also fulfilling requirements. Teaching normal and binomial distributions after exposing students to sampling distributions or null distributions provides a clear link between major concepts of variability and how we use theoretical distributions as a model for what we often see occurring naturally. If we take an approach that continues the focus on application. The theoretical connections may be pointed out after students have demonstrated an understanding of the importance and use of modeling and inference.

11.6 Looking Ahead

It should be noted there are ongoing discussions in many states around changes to curriculum guidelines for statistics. One example where these conversations have resulted in recent updates is the state of Arizona which uses the following language:

"The idea of the curriculum is to not be really specific in some areas. An example is with descriptive statistics, the course must have measures of central tendency, but the types of measures are not specified. This is similar in other areas, such as probability and what types of inferential statistics tests are conducted. ***Probability should be taught but deemphasised*** *[emphasis added]. The other suggestion is that calculations should be conducted using statistical software and not using calculators and definitely not by hand. z and t tables should not be used any more or at least limited use."*

It is encouraging to see this type of change in the language of recommended best practices for teaching introductory statistics. In most cases, these are simply guidelines that are reviewed occasionally by the state or university system. Syllabi and teaching artifacts may be reviewed to determine whether a course change still aligns with the curriculum guidelines. We believe that even with more restricted guidelines, there is room to teach probability and probability distributions in a thoughtful and applicable way while still meeting standards.

11.7 Conclusion

In this chapter we presented some of the challenges and opportunities for teaching probability and probability distributions in a way that emphasizes conceptual understanding instead of reliance on theory and formulas. We presented strategies and example exercises that encourage students to build upon their intuitive probabilistic thinking as they think critically about how probability and probability distributions can be used to derive insights and begin to draw conclusions about data. The approach presented in this chapter is largely motivated by the goal of helping students get a better grasp of statistical inference without overburdening their cognitive load. This provides a complement to the data analysis process introduced in Chapter 7.

As introductory statistics classes continue to evolve to best serve the modern student, state and institution-wide curricula are adapting this approach, with some de-emphasizing probability in the introductory course. There are also increasingly more textbooks, along with interactive web-based tools such as the Little Apps available for instructors as they facilitate this new approach to learning probability and probability distributions in their courses.

Chapter 11 Appendix: Note on p-values

There is an ongoing discussion about the use of p-values and the determination of conclusions as "significant" or "not significant" in scientific research. In 2015 American Statistical Association convened a committee of 26 statistical experts to address the use (and misuse) of p-values in research and develop a statement on p-values and statistical significance. The 2016 article "Statisticians Found One Thing They Can Agree On: It's Time To Stop Misusing P-Values" on the data journalism website FiveThirtyEight highlights some of the primary issues identified by committee members with the approach to statistical inference in scientific research [10]. These issues include the lack of ability to reproduce results that were deemed "statistically significant" in the previous studies, the reliance on p-values as a replacement for scientific reasoning, the inability to understand effect size from the p-value, and the use of p-values without adjusting for factors that impact the results such as biased data and multiple testing, among others. This is not to mention the major disadvantage to the use of p-values in current research - the common misunderstanding of what the p-value actually means. As stated in the article, a "common misconception among nonstatisticians is that p-values can tell you the probability that a result occurred by chance".

There has been some movement away from p-values, such as the ban on p-values by the journal Basic and Applied Social Psychology, as highlighted in the 2015 Nature article "Psychology Journal Bans P-Values" [213], but they are still widely used in academia and industry. Additionally, while there is general agreement among statisticians about the problems with over reliance on p-values, there is not widespread consensus on an alternative approach to statistical inference. Much of this debate in the statistics community was presented in the 2019 American Statistician special issue "Statistical Inference in the 21st Century: A World Beyond $p < 0.05$" [201].

Part III

StatPREP as a Catalyst for Statistics and Data Education

Part III Overview

Jenna Carpenter, *Campbell University*

This third part of the Notes volume contains a variety of articles that support and underpin the first two parts. While Part I StatPREP Project Background and Context describes the origins and philosophy of StatPREP, faculty and administrators interested in implementing the types of changes for which we advocate might find it useful to read more of the research undergirding our approaches. From our seven years of work on the StatPREP Project, we encountered some topics repeatedly, like questions and issues around use of real data, when considering how to reform the way introductory statistics is taught across a wide range of institutions, programs, settings, and student groups. And when the COVID pandemic necessitated a pivot to doing our work online in 2020 and 2021, we saw an opportunity to leverage and build on StatPREP's data-centric approach to introductory statistics for the creation of programs in data science. Certainly the interest in and proliferation of initiatives in data science have grown much faster that we could have imagined back in 2016.

The Abbreviated Literature Review of Chapter 12 contains a great review of the research that supports the StatPREP philosophy and the 2016 GAISE College Report recommendations that are the motivation for our work. To read more about what we learned from the StatPREP grant itself, check out Chapter 13: Unpacking the Impact of the StatPREP Project (or, for a deep dive, check out the formal assessment and evaluation report in the Library.

This volume argues that introductory statistics courses need to be data-centric and use authentic data. Chapter 14 The Data Story explores that further, describing how to provide authentic data-rich learning experiences, challenges that instructors are likely to encounter with this approach, and strategies for managing those issues.

For those readers interested in expanding the ideas here beyond introductory statistics to create courses, concentrations, certificates, minors, and majors in data science, check out Chapter 15 Undergraduate Data Science. This chapter offers perspectives on the definition of data science, as well as recommendations for implementing Data Science programs at both two-year colleges and four-year colleges or universities. These articles contain additional references to existing programs, along with different ways to focus and structure such efforts.

12

Making the Case: An Abbreviated Literature Review

Anna Bargagliotti, *Loyola Marymount University*

Susan Peters, *University of Louisville*

12.1 Introduction

StatPREP is an attempt to embody many of the recommendations made in the literature concerning how students should learn and experience introductory statistics. Here we present a brief literature review that outlines the evidence behind the StatPREP approach in its effort to promote teaching with real data and using computing to teach introductory statistics, along with resources for further reading.

StatPREP is a NSF-funded project that developed a series of professional development workshops for college instructors to help them teach with real data. The goal of StatPREP and the developed materials is to change how introductory statistics is taught by emphasizing real data wrangling and visualization techniques. StatPREP focuses on using data that resonates with students in order to develop both student interest and understanding in place of using rote, formulaic-based approaches with small teaching data sets in uninteresting contexts. This approach to teaching introductory statistics emphasizes teaching with real data, teaching with technology, active learning, and teaching to develop students' statistical habits of mind and is supported by considerable research in statistics education.

There is no question that the demand for people educated in statistics and data science has grown tremendously over the past decade. Jobs related to statistics continue to flourish with growth expected to be 35% between 2023 and 2033 according to the Bureau of Labor Statistics [194]. Key attributes to be successful in such jobs are good computing, analytic, and statistical skills; good communication skills; ability to work with real data; verbal and visual storytelling ability with data; and ability to work as part of a team [50]. As a result, data education entails ensuring that students not only have sound computing and statistical skills, but also have good communication skills and the ability to be team players ([92]; [215]). As noted by Horton and Hardin, "the idea that an undergraduate statistics [major] develops general problem-solving skills to use data to make sense of the world is powerful" [82]. This is what offerings in statistics should strive to achieve—nimble computing data problem solvers ([137]; [138]).

In 2016, the American Statistical Association (ASA) released the revised GAISE College Report [67]. As detailed in the Part II Overview, the report calls for introductory statistics to be taught using six key principles. The StatPREP philosophy and approach align with these recommendations. More specifically, three important, fundamental, and particularly timely themes central to the StatPREP approach are that (1) students need to employ technology, (2) students need to explore real data sets, and (3) students need to engage in active learning. Statistics should be guided and taught through the statistical investigative process of formulating a question, collecting and considering appropriate data to answer that question, choosing the appropriate analysis technique to answer that question, and interpreting the results to answer the question [67]. The main goal of StatPREP is to help instructors teach with data using modern pedagogy that emphasizes the use of computing to wrangle data and visualize data.

12.2 Teaching with Real Data

Teaching statistics using real data has been a longstanding recommendation by the statistics education community (see, for example, calls for using real data in the classrooms from [22], [54], [56], [71], [93], [100], [136], [165], and [176]). Evidence exists that the use of real data helps to promote understanding and engagement. For example, Neumann, Hood, and Neumann noted that the use of real data in an introductory statistics classroom led to increased motivation for learning by students [135]. In science classrooms, Schultheis and Kjelvik found that use of Data Nuggets classroom activities that focus on authentic data for K–16 students promoted students' ability to engage with the scientific process [175]. Makar and Confrey worked with pre-service teachers and also found that working with authentic data on equity-related issues increased student engagement [118]. Overall, the consensus is that using real-life data that is relevant and interesting to students helps to motivate them and to deepen their statistical reasoning ([45]; [72]; [170]; [176]). Meaningful data sets can be an effective vehicle for teaching statistics by enabling students to develop analytical skills through realistic research situations.

Several important reports also have stated the need for students to work with real data. The MAA Committee on the Undergraduate Program in Mathematics (CUPM) Curriculum Guide report (2015) recommends that all mathematical sciences major programs include concepts and methods from data analysis and computing using authentic data sets. As noted above, the 2016 GAISE College Report also includes working with real data as one of the necessary six components for structuring a statistics course. In addition, recommendations of the ASA for undergraduate programs in statistical science state that programs should "emphasize concepts and approaches for working with complex data and provide experiences in designing studies and analyzing real data (defined as data that have been collected to solve an authentic and relevant problem)" [9].

In addition to increasing student motivation for learning, the use of real data can also promote multivariate thinking. Real data tend to be complex, messy, and have multiple variables. Having students develop multivariate thinking was one of the main additions to the modernized GAISE College Report. The report notes that multivariate thinking should be promoted from as early as kindergarten [12]. Horton also maintains that introductory statistics courses at the university level should have students engage in multivariate thinking and multivariate methods [95]. As data are growing both in size and speed, it is imperative to have students wrangle and visualize complex data sets. Specific to visualizations, Nolan and Perrett advocate for the use of more visualization in introductory courses [139]. This is a natural extension from using real and complex data because data of this type often necessitates the use of visualizations to investigate it.

12.3 Teaching with Technology

The use of technology and computing in statistics classrooms has exploded as student-friendly statistical software and computing power has become more accessible. In addition to advocating for the use of real data, ASA's Guidelines for Undergraduate Programs in Statistical Science called for increased focus on teaching computing as part of undergraduate statistics education. Since working with data necessitates computational skills, there is widespread consensus that computing and technology should be interwoven with statistics education at all levels (e.g., [59]; [137]). In addition, with more accessibility to large data sets and access through technology, ethical considerations about data gain importance. Unfortunately, even courses designed specifically for students to work with real data often fail to consider data ethics. For example, Lyford and Mei examined available course descriptions and syllabi of data science courses offered in the United States and found that only 30% and 41% of introductory and non-introductory courses mention ethics in their course descriptions, and only 9% and 40% of the respective courses mention ethics in their syllabi [125].

The push to bring more computing into the classroom is common in the literature (see, for example, [36]; [63]). Bringing more computing and technology into the classroom can facilitate the employment of new teaching approaches and pedagogy. For example, Chance and colleagues point out the power of technology for enabling students to actively explore and visualize data to see statistical concepts, and the authors specifically point to the use of applets for students to dynamically interact with data [37]. Others point to the power of using apps for instructors to teach specific content in a manner that may not be possible using standard software packages [55]. A common use of apps and applets in statistics education is simulation to develop conceptual understanding of concepts such as chance, randomness, and inference (e.g, [37]).

With the emergence of introductory textbooks that build and teach basic inference through simulation and randomization, a proposed pedagogical approach for introductory statistics is to have students carry out simulations to gain insight into the seemingly difficult and abstract concepts such as sampling distributions, p-values, and hypothesis testing (e.g., [39]; [117]; [189]). To date, some research indicates that teaching introductory statistics with a heavy use of simulations and randomization techniques increases student understanding ([40]; [110]; [126]; [127]; [145]; [190]). Other research, however, points to potential pitfalls of teaching introductory statistics in this manner ([89]; [91]; [202]).

In general, there is no consensus on the optimal way to teach undergraduate statistics courses; however, there is an abundance of literature that supports the use of technology and computing in the classroom for data analysis. For example, Hardin and colleagues narrate the use of computing in the statistical curriculum over the past 10 years and note that computational reasoning for data analysis should be central to the statistics classroom [82]. The Journal of Statistics and Data Science Education put out a special issue focused on computing in the statistics curriculum. The collection of articles in the special issue all call for increased use of computing and suggest methods, pedagogy, and projects to achieve such a goal (e.g., [23]; [33]; [80]; [96]; [173]).

The use of computing and technology has become more accessible for students with the design of specific packages in statistical software (e.g., R's Mosaic package) and point-and-click student-friendly software (e.g., StatCrunch, JMP). Pruim, Kaplan, and Horton state that as students increasingly use real complex data in the classroom, having accessible ways for them to engage in data management, exploration of data, visualization and modeling is necessary [150]. Moreover, several studies show that using computing to visualize data and engage students in data analysis allows students to carry out the statistical investigative process and develop statistical thinking ([2]; [95]; [120]; [200]).

12.4 Teaching to Develop Statistical Habits of Mind through Active Learning

An important goal for statistics education is to develop students' statistical habits of mind. Burrill and Biehler posited several important statistical habits of mind that include using real data; building intuitions by, for example, using simulations; beginning data exploration with a graph; exploring alternative representations of data; investigating and exploring data before introducing formulas; and engaging in doing statistics [26]. Although not linked explicitly to aspects of the statistical investigative process by the authors, the activity needed to develop these habits of mind aligns with engaging in the entire statistical investigation process, as recommended in recent reports ([13]; [61]; [67]).

Lee and Tran explicitly align overlapping statistical habits of mind with different aspects of the investigative process, many of which students can develop through active learning from teaching with real data and technology [116]. Active learning has long been promoted in the statistics education literature (e.g., [45]; [70]). Strong cases have been made in the literature that active learning increases student performance in STEM disciplines (e.g., [29]; [62]; [185]; [113]; [104]; [38]). Aligned with this literature base, StatPREP's goals for statistics instructors to incorporate statistical analysis software and computing technology, complex data, open-ended investigations, and analytical thinking into their existing courses have the benefit of focusing teaching on developing students' statistical habits of mind through active learning [116]. For example, teaching with real, multivariate data can promote focus on developing habits of mind related to posing questions by bringing context to the fore and encouraging students' consideration of contextually-based questions and ways to explain and control variability to answer those questions. Teaching with real data also requires consideration of context when interpreting results, specifically when making claims about the results and arguments to support the claims based on the data. The latter also necessitates clear communication to effectively answer a posed investigative question [12]. Garfield argues that statisticians must change their culture within the discipline and put more focus on how to communicate statistical contributions in a manner that is accessible to others outside of the discipline, and Radke-Sharp argues that interpretation and communication are equally as important to analysis and computation ([70]; [156]). Emphasizing argumentation in written and oral form within the statistics classroom can help develop students' communication skills.

Teaching with technology can promote students' strategic and appropriate use of tools to transnumerate data by creating multiple data representations and computing statistics and coordinating the information to better understand the data and to reason about distribution as an aggregate ([207]; [81]). As data sets have become more complex and as technology has been infused more frequently in the statistics classroom, visualizations have also become an important tool for communicating data. (See, for example, Franconeri et al., who provide research-based recommendations for features that data visualizations should have in order to be effective in communicating ideas, [60]. Borner et al. describe a data visualization literacy framework to show how to manipulate characteristics of data visualizations (e.g., color, symbols, shape) to better communicate through data [20]. Presenting data visualizations to support arguments as well as writing clear interpretations and summaries are important skills to be developed within the statistics education classroom.

12.5 Extended Professional Development for Instructors

An important part of the StatPREP success in supporting instructors in modernizing the manner in which they teach statistics has been the design of the professional learning program for these instructors. Considerable consensus exists on characteristics of high-quality professional learning for K–12 teachers [87]. ”High quality” professional learning experiences focus on content ([46]; [49]; [76]; [179]), are of sustained duration ([46]; [49]; [76]; [179]), and offer teachers opportunities for active learning of both content ([69]; [162]) and pedagogy [69]; [73]; [86]; [115]). Additional research suggests successful professional learning programs provide opportunities for teachers to interact with other teachers in a collaborative environment to form communities of practice [171] and support networks ([84]; [167]).

The latter is supported by research in higher education as well, which provides evidence that professional learning communities help motivate and support change in teaching practice [83]. For example, Hilliard describes the structure of successful learning communities in higher education and notes that such communities can serve as change agents for improving student learning and success [90]. Stoll and colleagues provide an important paper reviewing the literature to date. In their paper, they note that professional learning communities in education have the potential for capacity building and helping develop teaching practices that can promote student success [184]. Vescio, Ross, and Adams illustrate that professional learning communities have a positive impact on teaching as well as increased student learning [198].

StatPREP participants came together for summer workshops with the shared common goal of changing their statistical teaching practices; inevitably, communities of like-minded instructors were formed. While the StatPREP workshops have ended, the community of practice remains. The field-tested StatPREP materials are available in the Library that serves as a companion to this Notes volume, and StatPREP enthusiasts can be found on the MAA Connect StatPREP community who are eager to continue the conversation about a data-centric approach to teaching introductory statistics.

13

Unpacking the Impact of the StatPREP Project

Sarah Holsted, *Broad-based Knowledge, LLC*

Flora McMartin, *Broad-based Knowledge, LLC*

13.1 Introduction

The purpose of this chapter is to tell the story of the StatPREP project using data that was collected and reported by the StatPREP project evaluation team over the duration of the StatPREP project (2016-2022). Broad-based Knowledge, LLC (BbK[1]), designed, developed, and conducted the StatPREP evaluation. With hindsight, it is easy to identify a number of choices and activities that impacted the project's trajectory. However, we will focus on impacts that we believe will endure beyond the StatPREP project and that can inform future efforts to change the way introductory statistics is taught.

13.2 Evaluation Analysis & Limitations

The data used in this chapter is drawn from surveys administered to workshop participants before and after each workshop (pre-/post-surveys) and surveys administered to participants one-to-three years after attending a workshop (post-post survey). Please see the final project evaluation report in the Notes volume library for details on methods, analysis, and results.

In 2022, the StatPREP evaluation team conducted a summative evaluation of the StatPREP Project. The evaluation focused on the impact of the StatPREP summer workshops. The primary project evaluation question was: "What methods of professional development best support modernization of the statistics curriculum and pedagogies in two-year institutions?"

Our analysis focused on three areas of interest:

- The impact of <u>workshop mode</u> (in-person or online);
- The <u>type of institution</u> where an attendee taught (2-year or 4-year institution); and,
- The <u>level of teaching experience</u> (less than six years and six years or greater).

The analysis of <u>level of teaching experience</u> led to few meaningful results, so these were not included in the final evaluation report. However, analysis did show there were impacts associated with <u>workshop mode</u> and <u>type of institution</u>.

Limitations to keep in mind:

[1] BbK aims to assist faculty members in their evaluation of innovations in higher education, especially in the areas of science, technology, engineering, and mathematics education. The BbK evaluation team was led by Dr. Flora McMartin from Richmond, CA, with significant contributions from Sarah Holsted and Raleigh Watts.

- The post-post survey was conducted less than one year after the 2021 online workshops. Therefore, attendees at the 2021 workshops may have had less time to implement changes than attendees at previous workshops (2017-2019, in-person).
- The quasi-experimental comparison of in-person and online professional development models in this report was made possible only because of the pandemic.

13.3 From Professional Development Model to Workshop Mode

In this section, we briefly describe the development of the StatPREP professional development model and its implementation and then report the results of the summative evaluation analysis related to the impact of workshop mode.

A synthesis of lessons learned from previous professional development efforts informed the design of the StatPREP professional development model in 2016. The goal of the StatPREP professional development program was for mathematics faculty to learn to teach modern methods of data analytics in introductory statistics courses. Please see Chapter 1 StatPREP Project Background for a detailed description on the evolution of the StatPREP Professional development model.

- Giersch & McMartin identified four principles for professional development, which were grounded in research on professional development programs and the change and innovation process in higher education [74].
  - Workshops must be complemented by ongoing support as new ideas are implemented.
  - Participant self-reflection is critical.
  - Building community is valuable.
  - Change takes time.
- The TANGO project (NSF DUE-1432251), established regional hubs and created a cadre of faculty who were mentored in modern statistics practices.
- From the decade-long MAA PREP program, lessons emerged about workshop duration and defraying travel and lodging costs to ensure participation.

As a result of these lessons, the StatPREP professional development model elements consisted of: creating regional hubs and a national support network; hosting workshops in urban areas to reduce the cost of attendance; and, limiting the duration of each workshop to 1.5 days.

From 2017-2019 and in 2022, 11 StatPREP workshops (1.5 days each, in-person) were held in seven locations around the U.S. (See Figure 13.1.)

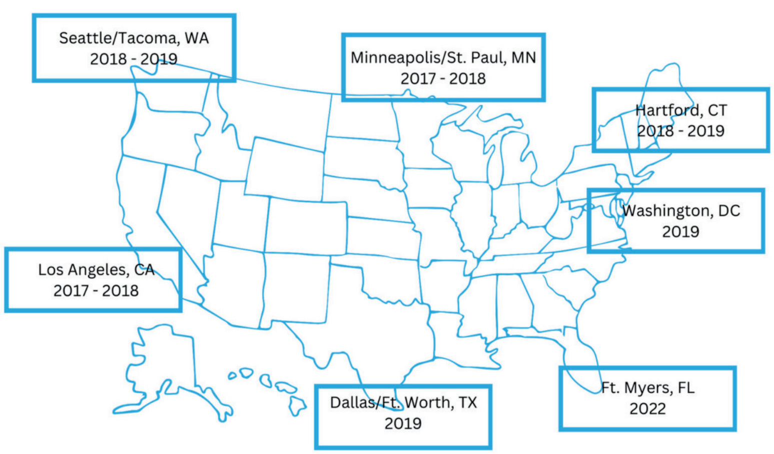


Figure 13.1: Geographic Distribution of StatPREP workshops.

Due to the pandemic, in-person StatPREP workshops were not held in summer 2020. The StatPREP professional development model was adapted to be delivered entirely online. The workshop content was modified and expanded

Element	In-Person (2017-2019; 2022)	Online (2021)
Time	Summer	Summer
Location	2- or 4-year campus in AMATYC regions	Online
Workshop Duration	12 hours over the course of 1.5 days	Six hours over two days
Workshop Distribution	Offered at seven different locations	Entire series of four sessions held once in June and once in July
Participants	Limited geographically by driving distance and room size (30-45 people)	Unlimited; national
Presenters	StatPREP project team; Hub Leaders	StatPREP project team; Hub Leaders
Content	Several topics covered at each workshop (e.g., Little Apps, RStudio, orientation to StatPREP resources, finding real data)	One topic covered in-depth at each workshop session (e.g., Little Apps, Introduction to RStudio, Advanced RStudio) and one panel on "Data Science at Two-Year Colleges"
Participant Compensation	Travel and lodging	None

Table 13.1: Comparison of the Elements of the StatPREP Professional Development Model

into a four-part series, and the series was offered twice during the summer. Table 13.1 compares the elements of the two modes for delivering StatPREP professional development workshops.

Based on data provided by project participants before they attended a workshop, the StatPREP project reached its intended audience.

- Over the duration of the StatPREP project, approximately 737 people attended StatPREP summer workshops in-person and online (Figures 13.2 and 13.3).
- 49% of responding participants were from two-year colleges, 75% were faculty members, and 85% were teaching full-time (Figures 13.4 and 13.5).

Figure 13.2: Mode of Participation

The unplanned deployment of online workshops in 2021 allowed the StatPREP evaluation team the opportunity to compare the impact of the original, in-person model of professional development workshops with the online professional development workshops. The results of the final evaluation report strongly suggest that:

1. The in-person workshop model was more effective in promoting change among attendees than the online workshop.
2. The in-person workshop model, combined with a target audience of 2-Year faculty members, was the most effective of all the workshop offerings.
3. Attending the in-person workshop may result in long-term changes.

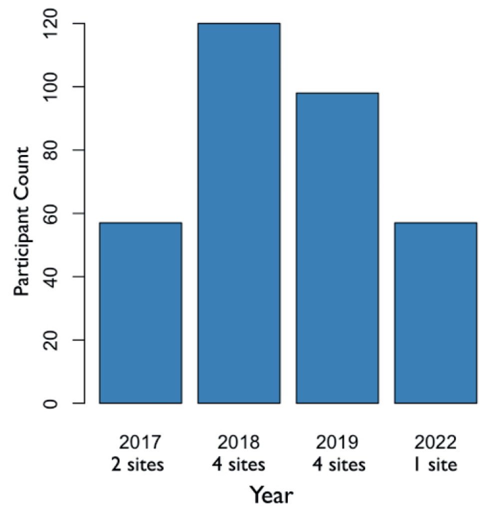


Figure 13.3: Participant Numbers, by Year

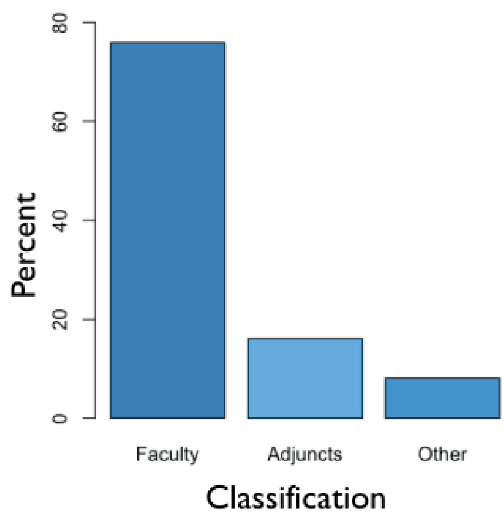


Figure 13.4: Participants' Position Classification

13.4 From College GAISE Recommendations to StatPREP Workshop Impacts

One of the goals of the StatPREP Project was to "increase faculty members' capacity to enact curricular change by incorporating statistical analysis software and computing technology, complex data, open-ended investigations, and statistical thinking into their existing courses." This goal grew from the recognition that on many campuses introductory statistics is taught by mathematics instructors who need professional development in order to address the recommendations made by the 2016 GAISE College Report [67].

In this section, we focus on the impacts on participants' changes in teaching practices and beliefs about teaching and learning statistics that are associated with the type of institution where participants taught (2-year and 4-year).

In our analysis of pre-survey and post-post survey results, we were guided by Cohen and Wolf's methods to measure the magnitude of difference between the two groups (participants from 2-year and 4-year institutions) through the use of effect sizes ([47]; [212]). We then used those effect sizes to determine if an intervention had an "educational impact" or a "practical impact."

The evaluation team determined that effect sizes that fell between 0.20 and 0.49 were likely to result in an educational impact. This means that while the effect size was not large, change did occur as a result of the intervention (i.e., attending a workshop) and was potentially beneficial to the participant.

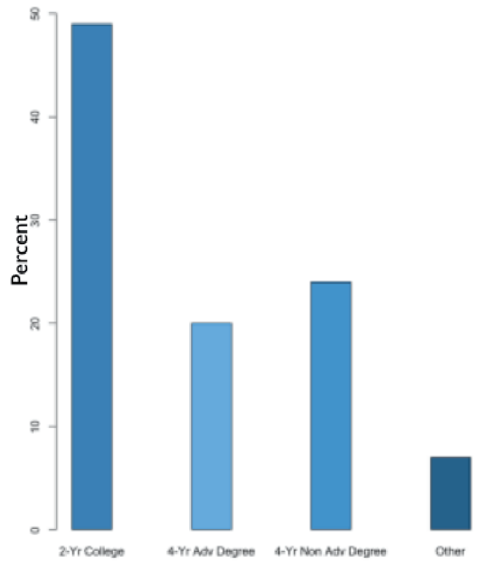


Figure 13.5: StatPREP Participants' Type of Institution

In contrast, effect sizes that were over 0.50 were likely to have a practical impact. This means that the impact of the intervention was large enough to be meaningful to participants. A large effect size with practical impacts also means that this information is "actionable" and we recommend that these interventions should be incorporated into future professional development interventions.

To what extent did respondents change their teaching practice?

The StatPREP Project used an incremental approach to making change. The idea was that, once started, workshop participants would experience positive changes in their students' learning statistics, which would further encourage faculty members to seek out and implement data-centric statistical teaching methods.

The StatPREP workshops, follow-on webinars, and online resources did provide participants with an opportunity to make changes in their teaching with respect to: the technology tools they used for teaching (Figure 13.6); their teaching strategies (e.g., moving away from all-lecture to use of active learning strategies); and, the content they used for teaching (e.g., introducing students to real data) (Figure 13.7).

To what extent did StatPREP workshop participants change their beliefs about teaching statistics?

StatPREP's aim was to encourage mathematics instructors to use modern statistical content and data-centric instructional practices required mathematics instructors to change how they teach as well as what they believe about teaching and student learning. As research across STEM education has shown, change is less about the thing being changed (i.e., teaching practice, curriculum) and more about changing instructors' beliefs about teaching and student learning

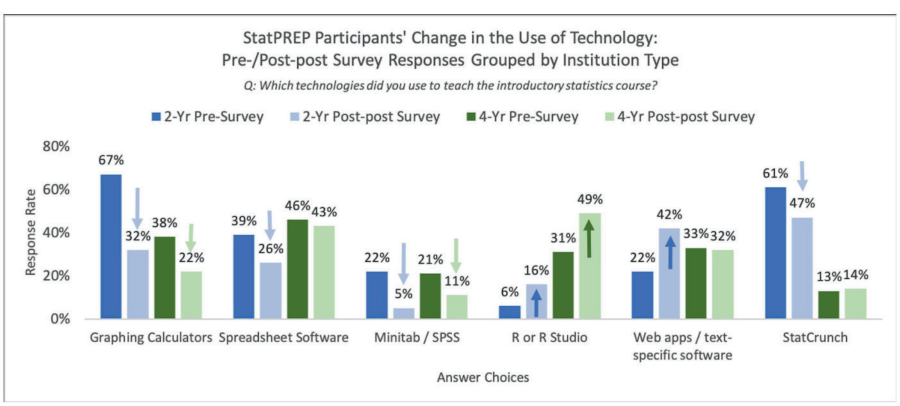


Figure 13.6: Participants' Use of Technology

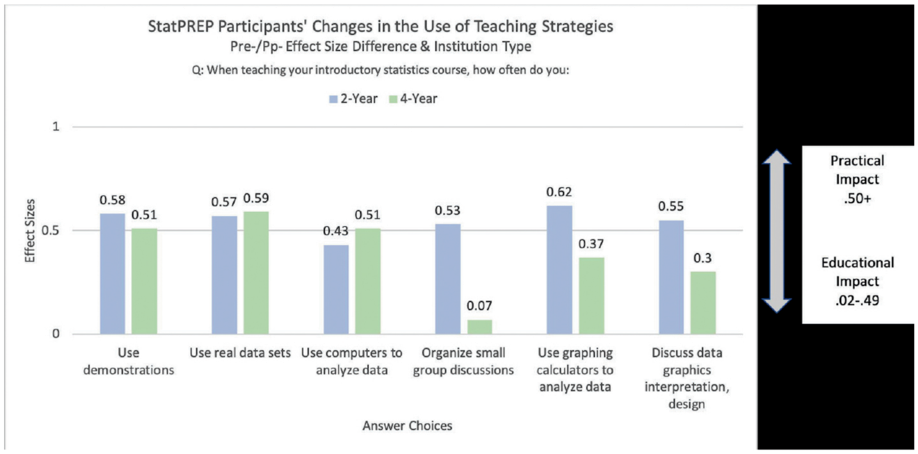


Figure 13.7: Participants' Teaching Strategies

generally ([192]; [148]) and with relation to the College GAISE recommendations in particular [214].

The incremental approach favored by the StatPREP Project follows the recommendations of Prosser and Trigwell, who suggest beliefs about teaching and student learning are key to adoption of innovations [148]. Research conducted by McKenna and Yalvac regarding adoption builds on this, suggesting that sustained reflections on the part of the innovation adopter is necessary to make change "stick" [121]. Wenger et. al and Lave and Wenger argue that learning is social and that communities of practice and social networks are necessary for people to learn ([204]; [114]). King claims that (positive) personal experiences with transformation learning are necessary to be open to innovation and adoption [108].

StatPREP's in-person workshop design built on the research listed above by providing workshop participants with opportunities to try new teaching resources during the workshop and to reflect on how they might use them in their own teaching. Organizing around the workshops at the regional level provided a means for local communities of practice around StatPREP to emerge, and online webinars supported a more national community of practice as did the StatPREP newsletter and website. Our analysis showed that the online workshop provided these same opportunities but less so and with a diffuse emphasis.

- Recognizing that changing beliefs is very difficult to accomplish in one workshop, the impact of attending a StatPREP workshop ranged from *no impact* to mainly an *educational impact* (Figure 13.8).
- Beliefs changed markedly depending on mode of workshop and type of workshop attendee.
- Analysis by mode of workshop and target audience for the nine indicators related to beliefs about data-centric teaching showed more change in more indicators for attendees of an online workshop and attendees from 4-year institutions. However, there is also a noticeable lack of overlap between the actual effect sizes between these groups, which may suggest a different set of priorities between the groups. Possible explanations could be that the different modes stressed content differently, and / or that participants attended workshops for different reasons.

13.5 Discussion and Conclusion

A recurring theme in this Notes volume are the recommendations for teaching statistics put forward in the 2016 GAISE College Report. When thinking about the StatPREP story, it is useful to remember that the GAISE College Report was released in 2016, the same year that the proposal to fund the StatPREP project was submitted to the National Science Foundation Improving Undergraduate STEM Education (IUSE) program. At that time, using data to teach statistics and evaluating the pedagogy and impact of teaching statistics was already happening [214]. However, as we write this chapter in 2023, data science and data visualization have "gone mainstream." The need for using modern methods to teach statistics is just as urgent as it was in 2016.

The selected results from the summative evaluation that are reported in this chapter suggest that StatPREP products and aspects of the professional development model might be successfully adapted or adopted by future professional

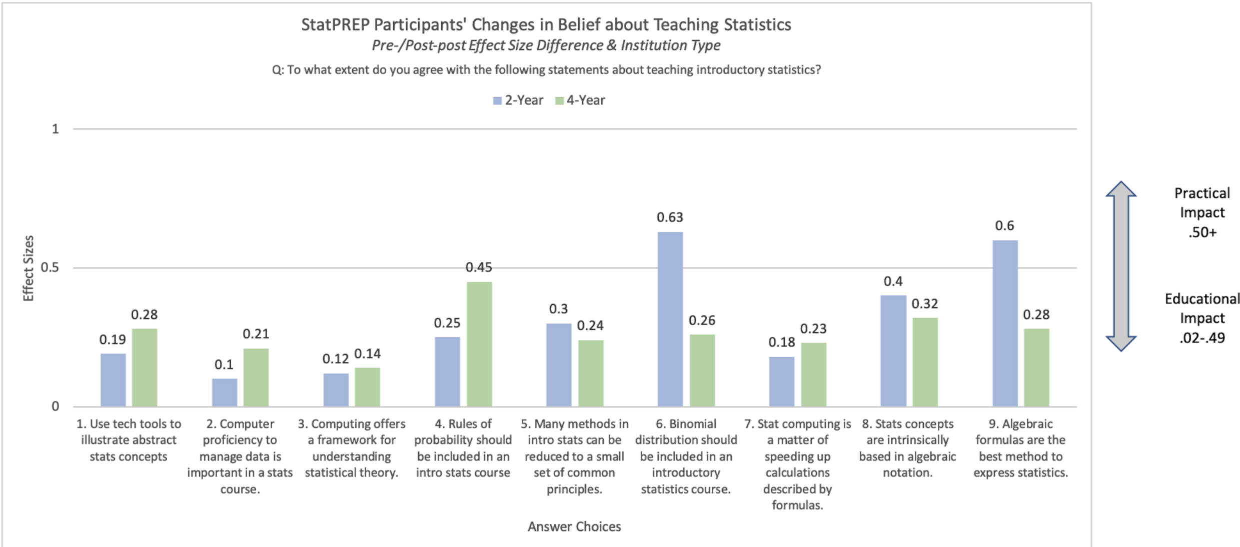


Figure 13.8: Participants' Beliefs about Teaching Strategies

development programs. The possibilities are framed as questions below.

What kind of impact is expected from an intervention: educational or practical?

Stakeholders use impact to measure the success of professional development programs. Understanding the costs or benefits of various program design elements can directly affect the type of impact a professional development program might have. The StatPREP workshops resulted in change for all attendees mainly at the educational impact level with some attendees experiencing change at a practical impact level.

Looking to the future, specifically with regard to elements of the StatPREP professional development model:

- Is it better to implement online workshops with broader, more diverse audiences (i.e., both 2-year and 4-year faculty members) that result in *educational impacts*?
- Is it more effective to use in-person workshops with the more targeted audience of 2-year faculty members that result in *practical impacts*?
- Is it better to combine activities to accomplish both?

What audience might be served best with this professional development program?

Appropriately targeting the audience for a particular professional development program can lead to more powerful impacts from an intervention (i.e., workshops). StatPREP's primary audience was 2-year mathematics faculty members who teach introductory statistics courses. Over time, however, about 50% of the attendees were 4-year faculty members. For many workshop outcomes, the 2-year faculty members involved in StatPREP experienced practical impacts while their 4-year colleagues did not. And, 2-year attendees experienced more changes to their teaching; they used significantly more computer technology; and, they used real data and data-centric teaching strategies.

Looking to the future, specifically with regard to the target audience for StatPREP professional development model:

- Is it better to target 2-year faculty members only, thereby having more and more powerful impacts?
- Is it better to continue to address a broader audience resulting in less powerful impacts?

In short: These questions are relevant beyond the StatPREP project and future efforts to effect change in teaching introductory statistics courses. Delivering and engaging with online-only professional development during the pandemic had an impact, which the StatPREP evaluation team had an opportunity to study. But, the overarching question remains: is there sufficient value in an online-only professional development model to implement it at scale going forward, knowing that in-person workshops are more effective but also knowing that online workshops, while popular, result in limited impact? Future professional development programs will need to find a balance that best meets their program needs and budgets.

14

The Data Story

Anna Bargagliotti, *Loyola Marymount University*
Donna LaLonde, *American Statistical Association*

14.1 Introduction

"Teaching statistics and data science generally requires lots of data—small data, big data, toy data, simulated data, and anything in between." (Dogucu and Çetinkaya-Rundel, 2022). This quote captures both the theme of this chapter and the StatPREP vision. A data-centric approach to teaching is the thread that brings together all the content of this volume. It is appropriate that data are the "star of the show." We will review the case for the importance of data, how to provide authentic learning experiences that engage students in exploring issues of data quality, how to find data sets that support the learning goals of your curriculum, how to navigate the challenges of using data that deal with social issues, and how to guide the ethical development of your students. Space does not allow for comprehensive coverage of these important topics in this chapter. Our goal is to begin a discussion, which we hope continues in the StatPREP Hub community. Readers are encouraged to join this community on MAA Connect.

14.2 The Case for a Data-Centric Approach

The Pareto principle, or the 80/20 rule, has been used in many contexts including to describe the work of data scientists. Applied to data science, the principle is framed as 80% of the work done by data scientists is "Data Gathering, Preparation, and Exploration" (Donoho, 2017) and 20% is dedicated to analysis. More importantly from a pedagogical perspective is the link between the statistical investigative question, the data collected or considered, the analysis, and the interpretation of the results. The process is not linear and will often require many iterations of each of the phases. The guidance from the 2016 GAISE College Report frames this from the student perspective: "Students should not leave their introductory statistics course with the mistaken impression that statistics consists of an unrelated collection of formulas and methods. Rather, students should understand that statistics is a problem-solving and decision making process that is fundamental to scientific inquiry and essential for making sound decisions" [67].

Before we consider how to provide students with authentic experiences considering the question of data quality, we should be explicit about our definition of data. Data can be numbers, counts, and measurements, but also images, video, sounds, or words, so we need to grapple with the fast-changing nature of data and the tools required for analysis to ensure that our curriculum is relevant for students. As statistics educators, we juggle the conundrum of too many topics and too little time. Our primary learning outcomes will guide the types of data that we are able to explore with our students; however, being cognizant of the types of data will make us more likely to recognize opportunities to introduce explorations that involve many data types.

For example, we could combine an activity designed to build community with the exploration of image data. This activity is an extension based on a lesson plan, Using Photographs as Data Sources to Tell Stories About Our Favorite

Outdoor Spaces, originally developed by Leticia Perez and published in the online, open-source Statistics Teacher Journal. An instructor could utilize this introductory activity during the first class meeting for the course by sharing a picture of their favorite outdoor space. In a writing assignment to be completed for the next class period, the students could be asked to share an image of their favorite outdoor space and respond to the prompt, "How do our favorite outdoor spaces compare?" The goal of this activity is to begin acclimating the student to a broader understanding of types of data and to introduce the concept of telling a story with data. It may also provide a "teachable moment" to address the issue of reproducibility: Will the image used be available for future researchers and how will this be ensured?

14.3 Messy Data and Data Quality

> "Tidy datasets are all alike, but every messy dataset is messy in its own way."
>
> —Hadley Wickham

This quote is the epigraph that introduces chapter 6 of the book *R for Data Science* [205]. Because every messy dataset is messy in its "own way," students need to encounter many messy datasets. Although each dataset will be messy in its own way, with experience students will be able to deal with unique dataset characteristics and make progress toward answering their statistical investigative question. In the introduction to the *R for Data Science* books, the authors proudly claim that a focus on small, in-memory datasets works well to provide the necessary foundation to be able to tackle big data. (Wickham and Grolemund) [205].

The statistical problem-solving process starts with formulating a question followed by collecting or considering the data. The first important data quality question that must be answered is, does the data support the investigation? Other important questions include: Is the meaning of the variables clear? Are you confident in the integrity of the data collection method? Are there missing values? The answers to these questions inform next steps in the statistical problem-solving process.

Consider this dataset, how much money do you make? — Ask a Manager, from the blog "Ask a Manager." Readers are asked to provide salary information by filling out a form anonymously. Any successful lesson must capture the students' attention, and how much money a person in a particular job makes is likely to be a topic that satisfies this criterion. The responses are available at Ask A Manager Salary Survey 2021 (Responses). We use this data set to highlight some of the important data quality and data cleaning issues that are a part of the data science lifecycle. The most obvious issue is that the data is collected anonymously via an online form, so there is no mechanism to validate the responses. Several of the questions allow for "free form" text entry, which are difficult to code. The question "What Industry do you work in?" allows for multiple responses, and each individual completing the form will use their own context to interpret the choices. This is authentic secondary data and the context is accessible, thus providing an opportunity for exploring important data quality issues with students.

Issues of data quality are often discussed in the context of reproducibility and replicability. The National Academies Reproducibility and Replicability in Science report (2019) defines **reproducibility** as "computational reproducibility—obtaining consistent computational results using the same input data, computational steps, methods, code, and conditions of analysis." The report also defines **replicability** to mean "obtaining consistent results across studies aimed at answering the same scientific question, each of which has obtained its own data." It is beyond the scope of this section to explore the tools that are used to support a reproducible workflow. The article An Invitation to Teaching Reproducible Research: Lessons from a Symposium in the Journal of Statistics and Data Science Education is a good starting point for exploring this topic [11].

14.4 Data Detectives

Finding data that fits your pedagogical needs often takes you on a hero's journey from your "safe" place to an unknown world. In such a journey, the stage of "meeting the mentor" is critical for cultivating the hero's willingness to plunge into the unfamiliar world. Fortunately, the StatPREP community, and the larger statistics education community, is filled with mentors who have collected data resources. The previous sections of this volume provide excellent examples of data sets and supporting analysis tools. For example, data from the National Health and Nutrition Evaluation Survey (NHANES) is curated and provided with a codebook that describes the variables [30]. Using one of the Little Apps lessons to present a topic from a data-centric perspective is a small change one can make immediately and with confidence that there are mentors available to support the change.

When you are ready for bigger steps, there are excellent online resources available to help navigate both the pedagogical and technological issues. For instance, the Data and Story Library supports searching for data sets by statistical method. Many colleagues make their course websites available; the Intro to Data Science course designed by Ben Baumer and the Statistical Learning course designed by Lucy D'Agostino McGowan are two examples. Many excellent open source textbooks favor a data-centric approach and thus provide good sources of data; consider, for example, Introduction to Modern Statistics by Mine Çetinkaya-Rundel and Johanna Hardin [34].

The Census at School project demonstrates an approach that was designed to engage middle and high school students in contributing their data to an analysis of a larger dataset. In this project, a class completes a survey and the system allows the class to compare their data with a random sample of data from other students. One of the strengths of this approach is that it gives students experience with data collection and with a messy, more complex dataset.

This, in turn, provides an opportunity to discuss the data collection instrument and best practices around the collection of data. The Census at Schools' existing instrument treats gender as dichotomous which does not reflect current best practice. Students will have implicit and explicit biases, and it is important to frame this discussion as one of professional responsibility, recognizing the impact of statistical practice on society, groups, and individuals. Although designed for pre-college students, this resource has been used by many college-level introductory statistics instructors.

14.5 Cleaning and Wrangling

Chapter 5 of this book provides a gentle introduction to R, which includes a description of the tidyverse. Regardless of whether or not you use R in your teaching, it is important to introduce students to the concepts of data cleaning and data wrangling. In the R community, creating "tidy" data involves importing the raw data set and "making it tidy" so each column is a variable and each observation is a row. The next step is to "transform" the data in order to be able to focus on specific observations and use existing variables to create new variables, e.g., a rate. In the tidyverse, these steps are called data wrangling.

Throughout this volume and in the broader literature, data science is acknowledged as an emerging discipline. One ramification of being early in the development of the discipline is that terms may be used inconsistently from one source to another. Data cleaning, data transformation, data wrangling, and data munging are examples of such terms. These terms are used as synonyms by some and as distinct actions by others. What is critical is to understand the activities that are associated with the initial steps of the data science life cycle. One definition states that "data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset." "Data transformation (or wrangling) is the process of converting data from one format or structure into another" [186].

There is an expression, "garbage in garbage out (GIGO)," which offers an appropriate warning to explore the concepts of data cleaning and data wrangling. Table 14.1 displays a random sample from the Census at School data.

There is missing data as well as "?" and "idk" responses. Before this dataset can be used in the statistical problem-solving process, the data needs to be "cleaned" by removing these entries. It is worth noting that there are more sophisticated statistical techniques for dealing with missing values than mere deletion of observations (e.g., imputation).

Consider another sample of Census and School data, shown in Table 14.2.

Visually reviewing the entries, we surmise that 6 and 5'5 are likely not in the appropriate units. If the dataset were too large to be assessed visually, we could utilize technology to accomplish our initial exploration.

Importing the dataset into the Common Online Data Analysis Platform - CODAP allows us to see the two entries that require additional investigation. Discussing how to proceed with the analysis provides an opportunity to cultivate a rich class discussion about the impact of data quality on the investigation and the conclusions. Students should consider what changes could be made in data collection to mitigate this issue. As a part of communicating their results, students should be explicit about what decisions were made.

Armspan_cm
182.8
182
?
155.6
175
166
Idk

Table 14.1: A sample of Census at School data

Height_cm	Footlength_cm
?	?
152.4	22
158	23
160	22
165	26
177	9
167	88
6	11
165.5	24
142.2	20.32
179	26
5'5	

Table 14.2: Another sample of Census at School data

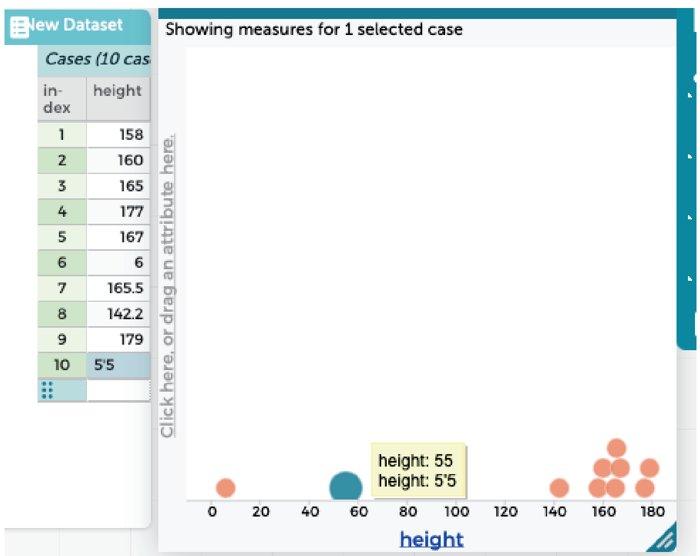


Figure 14.1: Graphing Census Data

14.6 Ethics Using Data

As a deluge of data are collected and data are more and more readily available, ethical considerations for how data are obtained, data privacy implications, potential biases present in data, and the possibility of targeted manipulation of consumers are all important ethical considerations. Baumer, et al., make the case for featuring ethical consideration prominently in all statistics and data science programs at the undergraduate level [15]. They propose course-level as well as departmental-level recommendations for institutions to implement in order to infuse ethics discussions in undergraduate programs. Utts also highlights the importance of making ethics part data of science programs and work discussions [195]. For example, she suggests numerous statistical practices addressing data ethics that instructors can bring into their classrooms (e.g., ethical reporting, ethical consent, societal vs. individual societal benefits).

To underscore the implications that ethical considerations can have for our society, we can examine recent high profile ethical conundrums. For example, the 2016 presidential election illustrated how Facebook data were used without consent to target individuals for voting [51]. The publication of FitBit data by the company helped reveal locations of US Army bases abroad which put military personnel in danger (U.S. soldiers are revealing sensitive and dangerous information by jogging - The Washington Post [178]). The recognition that data are being collected on each of us individually and without our explicit consent often being repurposed for profits in marketing and even political weaponization should be prominent when introducing students to working with data. Data ethics can be a difficult topic to understand and grasp, particularly because of the speed at which data collection mechanisms change.

Within the statistical problem-solving process, ethical considerations of data should be explored when considering whether data are appropriate for use to answer statistical questions. The American Statistical Association has developed Ethical Guidelines for Statistical Practice, which outline the statistical practitioner's responsibilities to all stakeholders and provide guidance on maintaining the integrity of data and methods. This includes a requirement to "understand the provenance of the data—including origins, revisions, and any restrictions on usage—and fitness for use prior to conducting statistical practices." Class projects that utilize secondary data sources provide an excellent opportunity to help students put the guidelines into practice.

In the article Ten simple rules for integrating ethics into statistics and data science instruction, Rochelle Tractenberg offers this advice: "Aim to equip, not solely instruct." Build in opportunities for students to practice using the guidelines as they complete investigations. It takes practice to focus not on finding the "right answer" but on using the guidelines to support ethical decision-making [191].

14.7 Data Considerations

This chapter has laid out the importance of data in society today, how to explore data quality and adjust for quality concerns, and how to find and access data. Once data have been identified or collected, then data wrangling can be used to curate the data for use in answering particular statistical questions of interest. Data ethics are especially important when using secondary data, yet also relevant when working with primary data. The manner in which data were collected impacts what can be inferred and concluded from the data. Overall, the data story takes us through the relevance of data and how we navigate the use of data with our students. Careful discussion within the classroom about data conception is invaluable and aligned with the StatPREP vision.

15

Undergraduate Data Science

Rachel Saidi, *Montgomery College*
Rebecca Sharples, *Purdue University*
Mark Daniel Ward, *Purdue University*

15.1 What is Data Science?

Data science is a new and evolving interdisciplinary field. A data scientist is someone who makes value out of data. Data scientists utilize scientific and statistical methods and algorithms to analyze and interpret data, which allows them to provide insights related to data-driven issues that their organization is experiencing. Data science requires computational, statistical, and mathematical thinking as well as skills in communication and a deep understanding of ethical implications surrounding the data lifecycle.

Data science may often be confused with computer programming or computer science. While the two disciplines have some overlap in that computer programming has a variety of applications and uses some of the same tools and software as data science, a student in computer programming uses coding techniques to develop applications, but they might never get experience handling a data set larger than what can be processed on their local machine. With data science, students explore, transform, and analyze multifaceted and complex datasets to communicate insights to various audiences. They learn how to use tools like R and Python (and others) in the full data cycle, from data collection and data wrangling to creation and delivery of data-driven visualizations and applications. Students analyze and manipulate datasets of various sizes, including some data sets that are so large that they need to be stored and processed in a high performance computing environment or a cloud-based computing environment. The students' local machines are only used to connect to the computational environment; the computational work is performed remotely.

The StatPREP methodology emphasizes incorporating these elements of computational statistics and data wrangling into the introductory curriculum. StatPREP seeks to provide students with a more realistic experiential learning environment, preparing them to address real-world data driven issues. An excellent resource that describes the recent understanding of data science pedagogy and curriculum is the Two-Year College Data Science Summits [78].

Students may not be aware that companies of all sizes and from all industries are hiring people with the full range of data science and analytical skills, from employees with basic data literacy to full data scientists. Data scientists who choose to work at small or midsize companies have the opportunity to make a tremendous impact on their teams at an early point in their careers. Beyond data science skills, other skills increasingly in demand include knowledge of enterprise systems and architecture, data engineering, data-driven software, modeling, machine learning, natural language processing, user interface / user experience, and data visualization. The prevalence of data science adoption throughout a corporation can give data scientists and statisticians an ever-broader opportunity to collaborate and influence the trajectory of data science use in practice. On the other hand, this presents a new challenge to universities, to continuously update their curricula and to stay in touch with current trends in the data sciences.

As societies move to datafication, competencies in data literacy and data acumen are increasingly important for all students, in every area of study. Increasing the number of students enrolled in data science and data analytics courses and programs will transform how communities approach complex, multifaceted problems and decisions with respect to public policy, public awareness, and individual autonomy. Industries previously run based on small disaggregated sets of data will see a transformation in the way they do business as their data competent employees transform their decision models.

15.2 How has Data Science and StatPREP Influenced Pedagogy?

In an increasingly data-driven world, data science, a new type of professional literacy, continues to gain traction in a wide range of industries. The rise of data science has also caused faculty who offer introductory data science courses to embed data science innovations into their programs, even during the first-year statistics curriculum. StatPREP takes this trend one step further when providing training to faculty and students, and encourages creativity, curiosity, training in scientific methods and specific technical toolsets, an understanding of the ethics surrounding the data lifecycle, and skills-development in communication and storytelling, mathematics, and statistics. Our ever-evolving, data-driven world demands not only an increasing number of data scientists and data analysts, but also data competent professionals in all fields.

Data science can be considered to be a two-year "infrastructure," serving as a key aspect of decision making in any organization. For example, Rawlings-Goss, et al., discusses the pervasive impact of data science across all industries [158]. In addition, data science builds on statistics, computer science, applied mathematics, but also requires domain knowledge, in order to achieve meaningful outcomes. Parker, Burgess, and Bourne state, "data science transcends traditional disciplinary boundaries to discover new insights not owned by any one existing discipline, driven by endless streams of digital data with the promise of translation to societal benefit" [142].

The U.S. Bureau of Labor Statistics (BLS) projects strong growth in the data science field. The BLS predicts that data science jobs within the mathematical sciences will increase by about 28% through 2026, which is approximately 11.5 million new jobs in the field [160]. A defining characteristic of the 21st century is the imperative need for data collection and data-informed decision-making. Accordingly, organizations are hiring employees with the technical skills needed to create high quality data science applications, environments, and products. Organizations also continue to value soft-skills that are critical to effectively communicate data-driven results and to consider the ethical implications of data analysis. The StatPREP methodology combines both the technical and professional development curriculum to help students become well-rounded researchers with the ability not only to analyze the data but to also interpret and communicate the findings and business impact.

15.3 The Two-Year College Perspective on Data Science

Two-year colleges provide an opportunity for students to try new subjects, change careers, upskill, or begin exploring higher education at affordable rates. Many students might begin their exploration of data science by taking a course at a local two-year college. There is much excitement and optimism at two-year colleges to help build data science learning and career pathways for these students. As the field of data science and data analytics continues to grow, and the need for more data scientists expands, two-year colleges play an important role in providing students with necessary skills to transfer to four-year schools or enter careers in data science. Within the last decade, data science programs have been increasingly appearing in different formats at two-year colleges across the country. The myriad reasons why students are choosing to explore data science are as varied as the diverse populations of students joining the programs.

Two-year colleges cater to students of all ages, from dual-enrolled high school students to retired seniors. Although a possibly over-used descriptor, data science is truly an interdisciplinary subject that spans all domains. In a single two-year college data science class, one might find a historian with a Ph.D., a corporate manager who wants to apply data analytics trends, a dual-enrolled high school computer science student, a math major, a business major, and a general studies student who wants to know more about data science. They come to the class with diverse abilities in mathematics, statistics, coding, and communications. Faculty need to be adept at teaching students with these varied abilities and perspectives, and willing to learn new technologies as they emerge.

Faculty also need to be able to leverage the varied perspectives in a way that attacks data-driven problems from different angles. For instructors who are accustomed to teaching relatively homogeneous baseline skills in a course like calculus, it is an entirely different experience to teach an introductory data science class. Faculty must accommodate the wide range of students' skills. Many data science programs have project-based assessments. Faculty in such courses must be willing to help students to find appropriate datasets to explore, to understand the variable definitions, and to navigate wrangling and cleaning, analyzing, and revealing insights. Without a singular "answer key" for a traditional quiz or test, helping students work on all aspects of the projects as well as grading them is labor intensive. Faculty need to navigate such challenges for a data science program to flourish. Faculty also face an additional set of potentially daunting issues to create and build a new data science program at their institution.

Two-year colleges have a unique position in higher education, welcoming students from all backgrounds and abilities, addressing workforce needs, satisfying requirements of administrators, programs, curriculum committees, boards of trustees, articulation agreements, and much more. Because data science is new and ever-evolving, meeting these requirements is a moving target. Embedding all necessary topics and skills into course sequences that meet course prerequisites is a puzzle that necessitates negotiations and good will with other departments both within and outside of the two-year school. Such course development and program development also requires awareness of four-year data science programs, where students may transfer after finishing their two year degree. "How to fit it all in" is a phrase commonly repeated amongst data science coordinators at two-year colleges, who face the challenge of supporting the multitude of stakeholders, including:

- high school students interested in starting their higher education path,
- non-traditional working professional students desiring to upskill,
- students interested in transferring to an undergraduate or graduate four-year institution,
- businesses, corporations, industries, and government entities needing more competent and skilled employees,
- administrators from within and outside of the organization with competing interests and perspectives,
- faculty who are qualified (or seek to become qualified) to teach data science classes,
- students entering the program with varied levels of mathematics, statistics, communication, and coding abilities.

"Data science" is an umbrella term. The educational requirements for data science positions vary, depending on the student's career path. Many schools have chosen to offer stackable credential options, from badges to certificates to degrees. The educational requirement for an entry-level data scientist or data analyst position is typically an associate's degree or certificate, while other mid- to high-level positions such as business or data analyst, senior data scientist, mathematician, or statistician require a bachelor's or even a master's degree. Completion of a data science associate of science degree provides the minimum college credits and curricula required to apply to many entry-level positions, such as data entry, database management, data science, and data analytics. An associate's degree (as opposed to a certificate) also provides greater flexibility for students planning to transfer to a four-year institution.

15.4 The Four-Year College Perspective on Data Science

Data science is an ever-evolving field at four-year colleges and universities, and manifests itself through various paths and degree programs. Although data science degrees are becoming increasingly prevalent, other established programs are beginning to include data science as a minor or concentration within their majors, or as supplementary certificate programs. For example, data science is often coupled with Computer Science, Statistics, Actuarial Science, and/or Mathematics degree programs. Sometimes the programs are dual-degree with Data Science as the second degree. Other programs offer Data Science as a concentration within the primary degree program. Some institutions offer a Data Science certificate that supplements existing degree programs not affiliated with a Data Science major/minor, but allows the students to learn how data analytics or data science skills and data competencies are applied in a particular field. For example, a university may have a Data Science major offered by its College of Science, College of Engineering, or even College of Data Science, as well as a broader Data Science certificate that any student in the university can pursue.

Some undergraduate institutions are implementing active learning models or experiential learning models, in which the university partners with local companies to provide applied data science opportunities for their students. This approach aligns with the StatPREP methodology and seeks to ensure that students have the preparation to work with real data that might be messy but more realistic once they enter the workforce. When applying an experiential learning pedagogy, an interdisciplinary team of students may work on a data-driven project with mentors from industry, creating an environment for tackling real-world problems from many different perspectives simultaneously, increasing the likelihood of success. Such experiences also serve to broaden a faculty member's understanding of how data science is used in practice by businesses, corporations, industries, and national laboratories. Industries need to be able to process increasingly large sets of data (think: petabytes and exabytes of data!). Having the ability to recruit and retain employees with these skills can propel these organizations to the next level of success.

Data science program coordinators must continually address the question of "what is next" for graduates of their programs, in terms of internship and career opportunities. Partnerships with local industries and organizations may facilitate these experiences. Local industries may not understand how students with experience in data science can help them. They often do not know the full extent of students' competencies in constantly evolving technologies and methodologies. When universities work with local industries to place students in internships and other experiential learning settings, these mysteries are solved. Institutions that develop projects and relationships with local businesses also have the opportunity to improve the local economy by creating opportunities for students that would not have existed otherwise. One of the benefits for students at four-year institutions, especially those students who work with the same partnerships for multiple years, is that they become a part of an organic recruiting pipeline with the companies they work with. An example of this is Purdue University's Data Mine Corporate Partners program. If students can get face-to-face time with company mentors on a weekly basis throughout the entirety of the school year, they are able to develop deep and meaningful relationships. Students participating in these types of Corporate Partners projects can reap the benefits of developing existing rich connections with hiring managers and trusted colleagues, when the companies are recruiting and the students are looking for future job opportunities. Universities are also seeing an increasing demand from graduate students who are already in the workforce wanting to deepen their data analytics and data science skills. Mid-career employees can also benefit from working directly with undergraduate and graduate students on data science projects arising in industry.

Experiential learning programs often do not include formalized lectures, but instead provide hands-on projects with extremely large data sets. Such programs enable students to learn how to apply a variety of data science tools and methodologies. Data Science programs and certificates have risen from a need for data competence in a world saturated with complex information and the need to be able to process that information in a way that is logical and functional. Students who may not have (otherwise) chosen to learn data science competencies can find an increased sense of self-efficacy and belonging, by participating in an active learning environment. Such learning programs can, for instance, introduce students from agriculture, business schools, engineering, liberal arts, medical and nursing programs, technical programs, veterinary science, etc., to data science skills that they would not have learned otherwise. Nursing students can use neural networks to predict cardiac arrests. Physics students can utilize large-scale data analytics and data science techniques to search for dark matter.

Four-year college students have the opportunity to go beyond introductory Python, R, and Structured Query Language (SQL) coursework. One helpful resource for universities planning to implement a data science program of any kind is configuration and access to a supercomputing cluster in which hundreds or thousands of jobs can be batched in parallel for large-scale data analysis. Some four-year colleges have made investments in cloud-based computing environments, such as Amazon Web Services, Azure, or Google Cloud. The National Science Foundation's Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, serves thousands of researchers across the US. The NSF has built some of the largest supercomputers in the United States. Researchers and their students can request computing allocations on supercomputing clusters maintained by the ACCESS program. JupyterLab is a common environment for students to access such computing resources. The benefit for faculty who are teaching Introductory Statistics courses is that they do not need to deal with students having separate resources on their individual computers, and their university does not need to provide a computing lab. Students can typically log in to one of the clusters that ACCESS maintains, and run JupyterLab directly on their computer, without needing to install any tools or libraries on their own computer.

The American Statistical Association (ASA) is partnering with The Data Mine at Purdue on a National Science

Foundation grant that enables 300 students from Minority Serving Institutions (MSIs) to work on projects with Corporate Partners and research faculty. This National Data Mine Network (NDMN) gives students a collaborative learning environment, including learning materials, access to a high performance computing cluster, and research projects with mentors from industry and from research institutions. This enables these students to use real-world data from companies and to participate in larger research teams.

15.5 How Data Science and Statistics is Changing Recruiting and Retention at All Levels

A new data science certificate, degree, or program at any two or four-year institution will require an ongoing advertising campaign (both internal and external to the organization) to let students know about its existence. Such advertisements should be directed to college students currently taking introductory statistics classes, students collegewide, people in the two-year, and high school students. It would also be beneficial to begin outreach for data science earlier in the lives of students, particularly girls. Studies have shown that girls are likely to conform to gender norms within their careers, including not feeling empowered or smart enough to explore STEM related fields [119].

As schools determine what type of program(s) to offer students (non-credit courses, certificates, or degrees), these decisions drive the kinds of mathematics and statistics skills they require. Many have chosen to start with low barriers to entry in terms of math and statistics requirements. Often, an introduction to statistics course may be the only math/statistics prerequisite for students to begin taking data classes. This introductory level may be sufficient for a student to earn a certificate. On the other hand, an associate of arts degree may require pre-calculus or probability skills; an associate of science might require students to take lower-division calculus and linear algebra courses. Where the data science program resides often impacts those types of decisions and often determines the answer to the question: what should a data scientist learn? Data science programs could be housed in math or statistics departments, computer science departments, business departments, or information technology departments, and this central location often determines the emphasis placed on computational, statistical, and mathematical thinking.

Although it is becoming increasingly common that high school students interested in STEM fields are exposed to Python, HTML, and other programming languages, it is still rare to find a high school student who has worked with large data sets. High school students are more likely to have taken a computer science class, and possibly a statistics class, but using data sets small enough that they can be analyzed on the student's local computer. In computer science classes, students learn the basics of coding in languages such as Python or Java, or perhaps C# or Javascript. High school students will likely not (yet) have first-hand familiarity with data science. This is potentially the reason why so many data science students at the two-year college level are non-traditional adults, returning to college in the midst of their careers, as opposed to 18–20 year-olds pursuing a degree for the first time. Adults are more likely to have heard about data science at their places of employment. Data science program coordinators can directly promote their data science programs to students in high school computer science and statistics courses. They might even start an early college data science degree.

Another option for promoting data science courses and programs is to create summer internship experiences for high school students to learn some data science fundamentals. For instance, rising junior and senior high school students in Montgomery County, MD, may participate in a Summer R.I.S.E. (Reimagining an Innovative Summer Experience). As another example, The Data Mine at Purdue University offers summer courses for high school students that introduce students to the same material that they will learn as undergraduate students, including the opportunity to work in a high performance computing environment for the first time. In addition to traditional advertising, recruiting and retaining women into data science programs can be a challenge, because gender role conformity starts in early childhood [140]. It is imperative that two- and four-year higher education institutions begin outreach to girls in elementary and middle school, and involve women, including minority women, in the outreach efforts. When girls see women in the fields they aspire to, they gain a sense of confidence that they can also achieve success.

Faculty should be intentional about creating a welcoming environment for students who are curious, but hesitant to learn about data science, once they make the decision to take an introductory data science class. Many prospective students interested in data science have no programming experience and are, not surprisingly, eager to learn to code. Faculty can choose timely, newsworthy, entertaining, and relevant datasets to motivate them. Faculty can aim to make the material immersive and interactive, rather than relying on traditional methods of lecturing. The list of potential data science topics is endless. Students' comfort and familiarity with a topic will allow them to engage more successfully with the challenging new data science concepts. Exploring data science is possible on almost any category or topic. At the same time, teaching code should be scaffolded by domain-specific questions, developing skills methodically, modeling, teamwork, in-class presentations, and collaborative debugging and problem solving. As students develop skills to work on individual or group projects, they may present their projects in class. A key goal should be to develop communication skills and other soft skills required for a career in data science.

Data science contests and hackathons are a popular means of attracting students to an appetizer-sized taste of the data science revolution. An example of an annual competition is ASA's DataFest, held each spring. Students from around the country participate in small teams and work around the clock to find and share meaning in large, rich, and complex data sets. These types of events get students excited and have the potential to turn into longer-term interest in the discipline, but by their very nature, they are short and geared more toward recruiting rather than retention.

15.6 Diversity, Equity, and Inclusion

Due to their ability to engage students from a broad range of backgrounds, data science programs have the potential to attract students from diverse groups and promote equity and inclusion by: (1) increasing the recruitment and retention of women and minority students, (2) providing a safe forum for students to debate controversial ethical and social topics related to data science, and (3) providing a curriculum designed to promote student success, retention, completion, and access to career opportunities.

A commonality among data science programs at two-year colleges is that a majority of the students are non-traditional, working professionals. For these students, it is imperative that classes are held in the evenings. At the outset of the pandemic in early 2020, many schools transitioned to virtual classes out of necessity. A resulting outcome is that many data science departments decided to continue to teach virtually. Offering virtual options for traditional and non-traditional students has become a response to market demand, as more and more institutions are re-thinking ways to meet the needs of their students. Because many of these programs are relatively small, they may not be able to offer multiple sections of the courses. With virtual classes in the evenings, students do not have to miss work to attend a daytime class, commute to campus, or find evening child care while attending the classes. Four-year institutions are also creating more flexible pathways, albeit more slowly. These pathways include virtual options, including fully online programs through many notable universities as part of their primary degree programs.

Additionally, many lower income two-year college students may have to make difficult choices in order to afford course materials fees on top of tuition fees. To address this issue, many in the data science two-year have embraced free, open-source software and tools. Faculty can teach R and R Studio, Python, SQL, Tableau Public, and Github without requiring students to make any purchases. Further, there are many free, open-source textbooks such as Wickham and Grolemund's R for Data Science [205]. Some organizations offer their course content free to anyone with internet access, which includes project assignments and how-to videos, intended for students with no data science background. Eliminating the cost of course materials is an equity issue and should seriously be considered not just by data science faculty, but by faculty in all departments. Such considerations are crucial for improving recruiting and retention efforts.

Faculty at two-year colleges and four-year institutions must also reflect on who is being left out of STEM education and ensure that underrepresented students are provided with the skills and knowledge to participate in this technological and digital space. In *Race After Technology*, Benjamin opines, "Who gets muted in this brave new world? The view that 'technology is a neutral tool' ignores how race also functions like a tool, structuring whose literal voice gets embodied. . . " [17]. Women and girls of color encounter obstacles compounded by the intersectionalities they experience in their lives ([28]; [140]). These women's experiences must be examined beyond just their gender, skin-color, or religious and cultural beliefs [101]. Using the theory of intersectionality, they must be examined as a composite of the various attributes that make up their identity ([28]; [101]; [140]). Specifically, Black and Hispanic women are significantly underrepresented in science, technology, engineering, and math (STEM). Alfred, et.al., noted (from statistics provided by the NSF's 2017 Report on Women, Minorities, and Persons with Disabilities in Science and Engineering; see [130] that both Black and Hispanic women's participation rates in STEM fields are 2%, several percentage points lower than their "working age population" [3]. Low rates of participation in STEM are a result of the social and cultural gender norms these women and girls experience from early childhood ([3]; [141]). These low rates of participation in

STEM include data science.

Deaf, blind, and disabled learners each face barriers to accessibility, when it comes to receiving an equitable education in the data sciences. For example, data science terminology is developed in the English language. Although many scientific terms have been established in other areas of science, data science terminology is still emerging in ASL. This results in more difficult communication and interpreting due to the necessity of fingerspelling many words [85]. Although federally mandated, the reality for online learning content for data science is that it is often not transcribed or captioned sufficiently [27]. When interpreters are available for deaf learners in their courses, the interpreters are not typically versed in data science, so the interpreters may be less accurate, as compared to an interpreter trained in data science. An under-qualified interpreter can also create a liability for the hiring organization, further exacerbating fears of hiring a professional from the Deaf community. For blind learners, screen readers and devices to convert audio to braille are used, but these technologies may not be able to accurately convert the material due to the new, evolving nature of data science terminology. Assuming these students are adequately educated in data science and want to pursue a degree, the students are at a significant disadvantage because of inherent bias employers have regarding the Deaf community, and the implications that may come with hiring a deaf professional. Congress passed the Americans with Disabilities Act (ADA) in 1990. Although the intent was to improve the *under*-employment of deaf professionals, the reality is that the rate of *un*employed deaf professionals has increased [99]. This coincides with the researchers findings that the ADA has made it more difficult for deaf or disabled individuals to find employment because of employer bias and fears that come along with hiring someone who is deaf or disabled.

Data science education at all levels is responsible for providing students–especially from minority and underrepresented populations–access to paths into the digital workforce, including leadership and decision-making roles in this evolving field.

15.7 Conclusion

A growing number of national and international organizations dedicate significant resources towards helping two- and four-year colleges and universities build and develop data science programs, such as the Academic Data Science Alliance (ADSA), Association for Computing Machinery (ACM), American Statistical Association (ASA), American Mathematical Association of Two-Year Colleges (AMATYC), Big Data Hubs, Education Development Center's Oceans of Data Institute (ODI), Genomic Data Science Community Network (GDSCN), Mathematical Association of America (MAA), and the National Academy of Sciences (NAS). These organizations are creating spaces for people to share materials, present workshops and webinars, develop and promote best practices, and connect mentors and mentees.

Beyond these organizations, many other groups have created freely available, open source materials specifically to share with faculty teaching introductory statistics and data science classes. UC Berkeley's Data 8 created The Foundations of Data Science, Mine Çetinkaya-Rundel offers Data Science in a Box, OpenIntro offers a number of introductory statistics texts with other resources and labs, faculty at Johns Hopkins University created Open Case Studies, and The Data Mine at Purdue has created The Examples Book.

For institutions considering the establishment of data science certificates, degrees, majors, or minors:

- Be sure to garner support from administrators, because they are absolutely necessary to shepherding the certificate or degree through all levels of approval – curriculum committees, Board of Trustees, state higher education committees, etc.
- Be comfortable being uncomfortable. Data science is constantly evolving, so faculty developing data science initiatives need to be comfortable working in an ever-changing discipline.
- Establish an advisory board with members from diverse settings – their expertise can be of great value.
- Join local and national organizations and become involved – American Statistical Association (ASA), American Mathematical Association of Two-Year Colleges (AMATYC), Academic Data Science Alliance (ADSA), Mathematical Association of American (MAA), to name a few.
- Attend webinars, conferences, symposiums, meetings – anything to learn more about data science education and get to know people who are also working in this area.

On a local, national, and global level, there is great demand and interest in the field of data science. One of the most valued talent pools will include those with data science skills, and their abilities will become more vital for organizations of all types. Meeting this demand to provide a new generation of students with data literacy and data acumen is imperative. The teaching of mathematics and statistics using traditional pedagogical approaches is no longer sufficient. Ideally, nationwide K–16 education must change to reflect the new challenges posed by the fact that all citizens are both data consumers and data producers on a massive scale. StatPREP provides faculty the professional development to transform their statistics pedagogy to use data-centered methods and act as a driving force of innovation in teaching modern statistics and data science.

Conclusion: Looking Ahead

Donna LaLonde, *American Statistical Association*

Were we successful? Throughout this volume we have shared our accomplishments and acknowledged the goals we did not meet. That reflection is important. In this Conclusion, we want to ask a different question. We do this to look to the future. As the StatPREP project wraps up, we are asking: How has the project contributed to advancing the visions of AMATYC, ASA, and MAA? As a leadership team we believe this is the measure that is future-focused.

The vision of AMATYC is to be the leading voice and resource for excellence and inclusion in the first two years of mathematics in colleges and universities. The ASA vision imagines a world that relies on data and statistical thinking to drive discovery and inform decisions. The MAA envisions a society that values the power and beauty of mathematics and fully realizes its potential to promote human flourishing. Taken collectively we believe this is the future for which StatPREP has made an important and sustainable contribution.

One theme that emerged as this volume came to fruition was the power of small changes. This is explicitly addressed in Chapter 6, but the theme is present throughout the volume. The other chapters in Part I provide additional context for understanding this theme by describing the philosophy, community, and pedagogy that supported its realization in both the workshops and in their classes—past, present, and future. In each of the chapters in Part II, the authors describe opportunities for change that span the gamut from utilizing one activity in one class to revamping an entire course. The chapters in Part III provide the future focus for the theme. Moving forward, the online library will serve as a digital repository for the activities and supporting documentation described in this volume. The online community will continue to provide resources, and we hope that the opportunities for small changes will continue to thrive.

Hundreds of StatPREP workshop participants were challenged to shift from traditional teaching methods to dynamic, hands-on approaches that prioritize active learning and critical thinking. The pedagogical approaches introduced in the summer workshops and in subsequent webinars, enables instructors to cultivate a deeper understanding of statistical concepts and foster invaluable skills that transcend the classroom. The Little Apps and supporting activities make it possible to immerse students in authentic data sets and practical scenarios. This volume, together with the online StatPREP Hub Community and online library, will serve as a resource not only to document the work of the project but to support the visions of supporting organizations.

For many of the StatPREP participants and the colleagues with whom they collaborated, the integration of technology changed the way statistics was learned. Offering interactive tools and resources enhanced accessibility and comprehension for students of diverse backgrounds and learning styles. A data-centric approach not only enhances statistical literacy but also cultivates a generation of informed citizens equipped to navigate the complexities of our data-driven world. It is our hope that this volume and the complementary online community will serve as a guiding light for educators seeking to transform their teaching methodologies and inspire a newfound passion for statistics among their students.

As we write this conclusion, the impact of ChatGPT and other Large Language Models is being discussed in formal and informal environments. What will the impact be on the nature of work? On education? Some of us remember similar questions when the graphing calculator was introduced and when Computer Algebra Systems became widely available. One does not need to be prescient about the future to be confident that technology will continue to change the nature of work and education. We are confident our collective visions and the strength of our collaboration will ensure that we are able to meet the challenges and take advantage of the opportunities.

Contributing Authors and Editors

Contributing Authors

Anna Bargagliotti. Anna Bargagliotti is a Professor and MAT Director in the Department of Mathematics, Statistics, and Data Science at Loyola Marymount University (LMU). She has extensive experience working on projects related to undergraduate STEM recruitment and retention, and she has expertise in nonparametric statistics, generalized linear models, data visualization, and statistics education. She teaches all statistics courses ranging from introductory statistics to graduate level statistics. She has over 50 scholarly publications, has published one book, given over 100 presentations, and won numerous awards related to her research as well as her teaching of statistics.

Megan Briet-Goodwin. Megan Breit-Goodwin teaches statistics and mathematics at Anoka-Ramsey Community College in Minnesota. She was a participant in the Minneapolis and Saint Paul StatPREP Hub. Her greatest joy in teaching is learning alongside her students. In addition to teaching, Megan is a coordinator of the Minnesota State REFLECT Program which engages 2 year college and 4 year university faculty in the Scholarship of Teaching and Learning across the Minnesota State system of colleges and universities.

Helen Burn. Helen Burn, Ph.D., is faculty member in the department of mathematics and director of the Curriculum Research Group at Highline College. She served as StatPREP Pacific Northwest Hub Leader and workshop facilitator. Her educational background includes a B.S. from The Evergreen State College, an M.S. in mathematics from Western Washington University and a Ph.D. in higher education from the University of Michigan Center for the Study of Higher and Postsecondary Education. She is current chair of the Mathematical Association of America's Special Interest Group (SIGMAA) on Statistics and Data Science.

Carol Howald. Dr. Carol Howald is currently a mathematics professor at Howard Community College in Columbia, Maryland. She regularly teaches introductory statistics and strives to develop and incorporate ways to engage her students actively in every class. Her role as StatPREP project hub leader in the mid-Atlantic region was a great experience and contributed to those goals!

Daniel Kaplan. Danny Kaplan has taught applied math and statistics for about 30 years, mostly at Macalester College (where he is an emeritus professor) and most recently for a three-year appointment at the US Air Force Academy as a Distinguished Visiting Professor. He's the author of several textbooks, the most recent of which are *MOSAIC Calculus* and *Lessons in Statistical Thinking*. He co-wrote the NSF proposal that started StatPREP and initiated Project MOSAIC, which worked to unify the introductory college-level curriculum in modeling, statistics, computing, and calculus. He is a Fellow of the American Statistical Association and was awarded the USCOTS Lifetime Achievement Award.

Kelly McConville. Kelly McConville is the inaugural Director for the Dominguez Center for Data Science at Bucknell University. She has taught courses across the undergraduate statistics and data science curriculum and is passionate about constructing introductory courses that lower the barrier to entry and broaden participation in statistics and data science. She is a Fellow of the American Statistical Association and served as a presenter for multiple StatPREP workshops.

Flora P. McMartin. Dr. Flora P. McMartin, Broad-based Knowledge (BbK), LLC, led the external evaluation team for the duration of the StatPREP project. McMartin founded BbK to assist STEM educators in evaluating the deployment of technology-assisted teaching and learning. For StatPREP, McMartin drew on her extensive experience with MAA projects such as WebWork, MathDL, Calc Visualizations and PIC. Recently retired, McMartin teaches yoga and is an avid cyclist.

Amelia McNamara. Amelia McNamara (she/her) is an Associate Professor of Data Science at the University of St Thomas, in Minnesota. She received her bachelor's degree in English and mathematics from Macalester College, and her PhD in statistics from UCLA. Her research interests include statistical computing, statistics education, data visualization, and spatial statistics. She has helped lead StatPrep workshops since 2017.

Susan A. Peters. Susan A. Peters is a Professor of Mathematics Education in the Department of Elementary, Middle, and Secondary Teacher Education at the University of Louisville, where she works with middle and high school mathematics teachers. She is chief editor of the Statistics Education Research Journal and a member of the American Statistical Association/National Council of Teachers of Mathematics Joint Committee on K-12 Education in Statistics and Probability. Her research focuses on mathematics teacher education in statistics.

Joseph Roith. Joe Roith is an Associate Professor of Practice in Mathematics, Statistics, and Computer Science at St. Olaf College in Northfield, MN. He joined the StatPREP team as Midwest Regional Hub Leader and workshop facilitator at the beginning in 2017. Joe has taught simulation based introductory level statistics courses using real data for the past ten years.

Rachel Saidi. Rachel Saidi is a professor of Math, Statistics, and Data Science at Montgomery College and the Data Science Program Director. She is the director of AMATYC's annual DataFest for Two-Year Colleges, an editorial board member for MAA's Scatterplot Journal, and a member of ASA/AMATYC Joint committee, MAA CRAFTY committee, and ADSA. She was a panelist for a number of data science workshops, including the StatPREP workshop. She was awarded Outstanding Faculty for Excellence in Teaching (2022) and the NISOD Award for Excellence (2023).

Rebecca Sharples. Rebecca Sharples is a professional specializing in information systems and organizational change. She earned her master's degree in information systems from Bryant University and her doctorate in learning and organizational change from Baylor University. She has extensive managerial and consulting experience in information systems, business processes, and risk management. She is particularly interested in advancing equity for women in STEM professions.

Dustin Silva. Dr. Dustin Silva, a mathematics professor with a focus on statistics education, teaches at the College of the Canyons in Santa Clarita, CA. He actively participated in the inaugural StatPREP workshop. As a fervent advocate for statistics reform in two-year colleges, Dustin has dedicated over a decade to leading and participating in professional development workshops, demonstrating his commitment to enhancing statistical education.

Maria Tackett. Maria Tackett is an Assistant Professor of the Practice in the Department of Statistical Science at Duke University. Her work focuses on curriculum development for undergraduate statistics courses and research examining factors that impact students' sense of belonging in introductory math and statistics courses. She presented sessions on teaching introductory data science courses and computing using R at the 2022 StatPrep Summer Workshop.

Jennifer Ward. Jennifer Ward, holding an MS in Statistics from Portland State University, serves as an adjunct statistics instructor at Clark College in Vancouver, WA (a two-year college). As a member of the StatPREP 2018-2019 Seattle/Tacoma Hub, Jennifer specializes in teaching online, infusing her courses with active learning methodologies. The integration of applets facilitates seamless data summarization and exploration for her students. Contact Jennifer at jswardclark.edu.

Mark Daniel Ward. Mark Daniel Ward is Professor of Statistics and (by courtesy) of Agricultural & Biological Engineering, Computer Science, Mathematics, and Public Health at Purdue University, in West Lafayette, Indiana, USA. His research is in analysis of algorithms. Ward is also interested in data science, science of information, game theory, and large-scale computation. He also serves as Executive Director of The Data Mine.

Editorial Team

Michael Brilleslyper. Dr. Michael Brilleslyper is professor and chair of the applied mathematics department at Florida Polytechnic University. Prior to his current post, he was a professor of mathematics at the U.S. Air Force Academy. He has been involved with StatPREP since its inception and has been an active MAA member for 28 years. He strongly supports the use of student-centered learning and providing professional development opportunities to the entire mathematics community.

Jenna Carpenter. Dr. Jenna P. Carpenter is Founding Dean and Professor of Engineering at Campbell University and President-Elect of the Mathematical Association of America. She is a past president and Fellow of the American Society for Engineering Education (ASEE). Dr. Carpenter was one of four recipients awarded the 2022 Bernard M. Gordon Prize for Innovation in Engineering and Technology Education from the National Academy of Engineering. She received the 2023 Claire Felbinger Award for Women in Engineering from ABET and was inducted into the ASEE Hall of Fame in 2023. Carpenter, an expert on innovative STEM curricula and on issues related to DEI in STEM fields, served as a co-PI on the NSF-funded StatPREP Project.

Sarah Holsted. Sarah Holsted, MSLS, was part of the external evaluation team, led by Broad-based Knowledge, LLC, for the duration of the StatPREP project. Previously, Holsted conducted project evaluations, literature reviews, and guided the development of research repositories and digital libraries in higher education. Holsted is the Hospital Library Services Program Manager for Southeastern NY Library Resources Council.

Kathryn Kozak. For the last 30 years, Kathryn Kozak has been teaching mathematics and statistics at Coconino Community College, in Flagstaff, AZ. Currently she is the lead faculty for the Mathematics and Accounting Department. She was one of the authors of AMATYC's IMPACT document. She was the co-PI on the StatPREP grant, and is a co-PI on the CURM grant and the co-PI on several other grants. She is a past president of AMATYC.

Donna LaLonde. Donna LaLonde is associate executive director of the American Statistical Association, where she works with colleagues to advance the association's vision and mission and supports activities associated with presidential initiatives, professional development, education, and accreditation. Before joining the ASA, LaLonde was a faculty member at Washburn University and served in various administrative positions, including interim chair of the education department and associate vice president for academic affairs.

Ambika Silva. Dr. Ambika Silva is a mathematics professor and the statistics faculty coordinator at College of the Canyons in Santa Clarita, CA. She was involved with StatPREP as a hub leader at the beginning of the grant, and later was added to the leadership team. A strong advocate for statistics reform at two year colleges, Ambika has been involved with and led professional development workshops for over a decade. She also was the Data Science Sub-Committee leader for AMATYC.

Deirdre Longacher Smeltzer. Deirdre Longacher Smeltzer gained experience teaching a full spectrum of undergraduate mathematics courses while on the faculties at the University of St. Thomas in St. Paul, MN, and at Eastern Mennonite University (EMU) in Harrisonburg, VA. Subsequently, Smeltzer served for six years as Vice President and Undergraduate Dean at EMU before moving into the Senior Director for Programs position at the MAA. The MAA Senior Director position afforded Smeltzer the opportunity to join the leadership teams of several NSF-funded projects, including StatPREP.

References

15.8 Bibliography

[1] Abell, Martha L., Linda Braddy, Douglass Ensley, Lewis Ludwig, and Hortensia Soto, 2018. *MAA Instructional Practices Guide*. Washington, DC: Mathematical Association of America.

[2] Adams, Bryan, Daniel Baller, Bryan Jonas, Anny-Claude Joseph and Kevin Cummiskey. 2021. Computational skills for multivariable thinking in introductory statistics. *Journal of Statistics and Data Science Education* 29(sup1): S123-S131.

[3] Alfred, Mary V., Sarah M. Ray, Michele A. Johnson. 2019. Advancing Women of Color in STEM: An Imperative for U.S. Global Competitiveness. *Advances in Developing Human Resources* 21(1): 114–132. doi.org/10.1177/1523422318814551.

[4] Aliaga, Martha, George Cobb, Carolyn Cuff, Joan Garfield, Rob Gould, Robin Lock, Tom Moore, et al. 2005. *Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report*. Alexandria, VA: American Statistical Association.

[5] Allaire, Joseph J., Christophe Dervieux, Carlos Scheidegger, Charles Teague, and Yihui Xie. 2022. *Quarto: Open Source Tools for Scientific and Technical Publishing*. quarto.org/.

[6] American Mathematical Association of Two-Year Colleges. 2021. *Useful Websites for Teachers*. www.amatyc.org/?page=StatsResources.

[7] American Statistical Association. 2022. *American Statistical Association Ethical Guidelines for Statistical Practice*. Zenodo. doi.org/10.5281/zenodo.7092386.

[8] American Statistical Association. 2021. *Useful Websites for Teachers*. www.amstat.org/education/useful-websites-for-teachers.

[9] American Statistical Association. 2014. Curriculum Guidelines for Undergraduate Programs in Statistical Science. www.amstat.org/education/curriculum-guidelines-for-undergraduate-programs-in-statistical-science-.

[10] Aschwanden, Christie. 2016. Statisticians found one thing they can agree on: It's time to stop misusing p-values. FiveThirtyEight. fivethirtyeight.com/features/statisticians-found-one-thing-they-can-agree-on-its-time-to-stop-misusing-p-values/. March, 7, 2016.

[11] Ball, Richard, Norm Medeiros, Nicholas W. Bussberg, and Aneta Piekut. 2022. An Invitation to Teaching Reproducible Research: Lessons from a Symposium. *Journal of Statistics and Data Science Education*, 30(3): 209-218. doi.org/10.1080/26939169.2022.2099489.

[12] Bargagliotti, Anna, Wendy Binder, Lance Blakesley, Zaki Eusufzai, Ben Fitzpatrick, Maire Ford, ... and Thomas Zachariah. 2020. Undergraduate learning outcomes for achieving data acumen. *Journal of Statistics Education*, 28(2): 197-211.

[13] Bargagliotti, Anna, Christine Franklin, Pip Arnold, Rob Gould, Sheri Johnson, Leticia Perez, Denise A. Spangler. 2020. *Pre-K12 Guidelines for Assessment and Instruction in Statistics Education (GAISE) report II*. Alexandria, VA: American Statistical Association and National Council of Teachers of Mathematics.

[14] Baumer, Ben, Mine Çetinkaya-Rundel, Andrew Bray, Linda Loi, Nicholas J. Horton. 2014. R Markdown: Integrating A Reproducible Analysis Tool into Introductory Statistics. *Technology Innovations in Statistics Education*. 8 (1). doi.org/10.5070/T581020118.

[15] Baumer, Benjamin S., Randi L. Garcia, Albert Y. Kim, Katherine M. Kinnaird, and Miles Q. Ott. 2022. Integrating data science ethics into an undergraduate major: A case study. *Journal of Statistics and Data Science Education*, 30(1): 15-28.

[16] Beach, Andrea L., Charles Henderson, and Noah Finkelstein. 2012. Facilitating Change in Undergraduate STEM Education. *Change: The Magazine of Higher Learning*, 44(6): 52-59. doi.org/10.1080/00091383.2012.728955

[17] Benjamin, Ruha. 2019. *Race after Technology: Abolitionist Tools for the New Jim Code*. Cambridge, England: Polity Press.

[18] Birman, Beatrice F., Laura Desimone, Andrew C. Porter, and Michael S. Garet. 2000. Designing Professional Development That Works. *Educational Leadership*, 57(8): 28-33. files.ascd.org/staticfiles/ascd/pdf/journals/ed_lead/el200005_birman.pdf

[19] Blanton, Maria L., and Despina A. Stylianou. 2009. Interpreting a Community of Practice Perspective in Discipline-Specific Professional Development in Higher Education. *Innovative Higher Education*, 34(2): 79-92. doi.org/10.1007/s10755-008-9094-8

[20] Börner, Katy, Andreas Bueckle, and Michael Ginda. 2019. Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences* 116, no. 6 (2019): 1857-1864.

[21] Bouwma-Gearhart, Jana. 2012. Science Faculty Improving Teaching Practice: Identifying Needs and Finding Meaningful Professional Development. *International Journal of Teaching and Learning in Higher Education*, 24(2): 180-88. www.isetl.org/ijtlhe/pdf/IJTLHE1201.pdf

[22] Bradstrett, T.E. 1996. Teaching introductory statistics courses so that nonstatisticians experience statistical reasoning. *The American Statisticia*, 50(1): 69-78.

[23] Burckhardt, Philipp, Rebecca Nugent, and Christopher R. Genovese. 2021. Teaching Statistical Concepts and Modern Data Analysis With a Computing-Integrated Learning Environment. *Journal of Statistics and Data Science Education*, 29(sup1): S61–S73. doi.org/10.1080/10691898.2020.1854637

[24] Burn, Helen. 2017. *Math pathways to completion: Introductory Statistics (Math& 146) recommendations*. Highline College Curriculum Research Group. documents.highline.edu/crg/mpc_math146recommendations_dec2017.pdf

[25] Burn, Helen, and Daniel Kaplan. 2019. *Instructor guide to linear regression*. Washington, DC: Mathematical Association of America. connect.maa.org/viewdocument/instructor-guide-to-linear-regressi?CommunityKey=6be8da59-c237-4569-bc85-b4ab59cbaf91&tab=librarydocuments.

[26] Burrill, Gail, and Rolf Biehler. "Fundamental statistical ideas in the school curriculum and in training teachers." *Teaching statistics in school mathematics-challenges for teaching and teacher education: A Joint ICMI/IASE study: The 18th ICMI study* (2011): 57-69.

[27] Campoverde-Molina, Milton, Sergio Lujan-Mora, Llorenc Valverde. 2021. Accessibility of University Websites Worldwide: A Systematic Literature Review. *Universal Access in the Information Society*. doi.org/10.1007/s10209-021-00825-z.

[28] Carastathis, Anna. 2014. The Concept of Intersectionality in Feminist Theory. *Philosophy Compass* 9(5): 304–314. doi.org/10.1111/phc3.12129.

[29] Carlson, Kieth A., and Jennifer R. Winquist. 2011. Evaluating an active learning approach to teaching introductory statistics: A classroom workbook approach. *Journal of Statistics Education* 19(1). doi.org/10.1080/10691898.2011.11889596

[30] Centers for Disease Control and Prevention (CDC) National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. CDC NHANES Data.

[31] Çetinkaya-Rundel, Mine. 2021. Data Science in a Box. Retrieved online April 27, 2024. datasciencebox.org/.

[32] Çetinkaya-Rundel, Mine, David M. Diez, and Christopher Barr. 2023. OpenIntro Statistics. Retrieved online April 27, 2024. www.openintro.org/book/os/..

[33] Çetinkaya-Rundel, Mine, and Victoria Ellison. 2020. A Fresh Look at Introductory Data Science. *Journal of Statistics and Data Science Education*, 29(sup1): S16–S26. doi.org/10.1080/10691898.2020.1804497

[34] Çetinkaya-Rundel, Mine and Johanna Hardin. 2023. *Introduction to Modern Statistics*. 1st ed. Retrieved online April 27, 2024. openintro-ims.netlify.app/.

[35] Çetinkaya-Rundel, Mine, Johanna Hardin, Benjamin S. Baumer, Amelia McNamara, Nicholas J. Horton, and Colin W. Rundel. 2022. "An Educator's Perspective of the Tidyverse." *Technology Innovations in Statistics Education* 14 (1). /doi.org/http://dx.doi.org/10.5070/T514154352.

[36] Çetinkaya-Rundel, Mine, and Colin Rundel. 2018. Infrastructure and Tools for Teaching Computing Throughout the Statistical Curriculum. *The American Statistician*, 72(1): 58–65. doi.org/10.1080/00031305.2017.1397549

[37] Chance, Beth, Dani Ben-Zvi, Joan Garfield, and Elsa Medina. 2007 The role of technology in improving student learning of statistics. *Technology Innovations in Statistics Education* 1(1).

[38] Chance, Beth L., and Joan B. Garfield. 2002. New approaches to gathering data on student learning for research in statistics education. *Statistics Education Research Journal* 1(2): 38-44.

[39] Chance, Beth and Allan Rossman. 2015. *Introduction to Statistical Concepts, Applications, and Theory*, 3rd edition. Self-published.

[40] Chance, Beth, Jimmy Wong, and Nathan Tintle. 2016. Student performance in curricula centered on simulation-based inference: A preliminary report. *Journal of Statistics Education* 24(3): 114-126.

[41] Chin, Monica. 2021. Students Who Grew up with Search Engines Might Change STEM Education Forever. *The Verge*. https://www.theverge.com/a/verge-2021.

[42] Clark, M. Carolyn, Jeffrey Froyd, Prudence Merton, and Jim Richardson. 2004. The Evolution of Curricular Change Models within the Foundation Coalition. *Journal of Engineering Education* 93(1): 37-47. doi.org/10.1002/j.2168-9830.2004.tb00786.x.

[43] Clayton, Aubrey. 2020. How eugenics shaped statistics. *Nautilus*, 29. nautil.us/issue/92/frontiers/how-eugenics-shaped-statistics.

[44] Cobb, George. 2015. Mere Renovation is Too Little Too Late: We Need to Rethink our Undergraduate Curriculum from the Ground Up. *The American Statistician*. 69(4):266-282.

[45] Cobb, George. 1992. Teaching Statistics. *Heeding the call for change: Suggestions for curricular action*, 22: 3-43.

[46] Cohen, David K., and Heather C. Hill. 2008 *Learning policy: When state education reform works*. Yale University Press.

[47] Cohen, Jacob. 1969. *Statistical Power Analysis for the Behavioral Sciences.*. 1st. New York: Academic Press.

[48] Common Core State Standards Initiative. 2022. *Mathematics: High School-Statistic and Probability*. www.corestandards.org/.

[49] Darling-Hammond, Linda, and Deborah Loewenberg Ball. 1998. "Teaching for high standards: What policymakers need to know and be able to do. CPRE Joint Report Series." *CPRE Publications*, November 30, 1997. eric.ed.gov/?id=ED426491.

[50] Davenport, Thomas H., and DJ Patil. 2012. ”Data Scientist: The Sexiest Job of the 21st Century.” *Harvard Business Review*. hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century.

[51] Davies, Harry. 2015. “Ted Cruz Campaign Using Firm that Harvested Data on Millions of Unwitting Facebook Users,” The Guardian.

[52] Jacqueline A., Greg Benfield, and Richard Francis. 2012. An Academic Development Model for Fostering Innovation and Sharing in Curriculum Design. *Innovations in Education and Teaching International* 49(2): 135-47. doi.org/10.1080/14703297.2012.677595.

[53] De Veaux, Richard. “How Much Is a Fireplace Worth?” *Stats 101: A Resource for Teaching Introductory Statistics*. Accessed April 28, 2024. community.amstat.org/stats101/home.

[54] Diamond, Neil T., and Ewa M. Sztendur. 2002. Simplifying consulting problems for use in introductory statistics lectures. In *Proceedings of the Sixth International Conference on Teaching of Statistics, Cape Town. Voorburg, The Netherlands: International Statistical Institute.*

[55] Doi, Jimmy, Gail Potter, Jimmy Wong, Irvin Alcaraz, and Peter Chi. 2016. Web application teaching tools for statistics using R and shiny. *Technology Innovations in Statistics Education* 9(1).

[56] Dunn, Dana S., Randolph A. Smith, and Bernard C. Beins (Eds). 2007. *Best practices for teaching statistics and research methods in the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

[57] Edwards, Ann R., and Rachel L. Beattie. 2016. Promoting Student Learning and Productive Persistence in Developmental Mathematics: Research Frameworks Informing the Carnegie Pathways. *NADE Digest*, 9(1): 30-39.

[58] Edwards, B. 2015. *External Evaluation of PREP Program (2004–2015)*. Washington, DC: Mathematical Association of America.

[59] Finzer, William. 2013. The data science education dilemma. *Technology Innovations in Statistics Education*, 7(2).

[60] Franconeri, Steven L., Lace M. Padilla, Priti Shah, Jeffrey M. Zacks, and Jessica Hullman. 2021. The science of visual data communication: What works. *Psychological Science in the public interest* 22(3): 110-161.

[61] Franklin, Christine, Gary Kader, Denise Mewborn, Jerry MOreno, Roxy Pack, Mike Perry, Richard Scheaffer. 2007. *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K–12 Curriculum Framework*. American Statistical Association. Alexandria, VA.

[62] Freeman, Scott, Sarah L. Eddy, Miles McDonough, Michelle K. Smith, Nnadozie Okoroafor, Hannah Jordt, and Mary Pat Wenderoth. 2014. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, 111(23): 8410–8415. https://doi.org/10.1073/pnas.1319030111

[63] Friedman, Jerome. 2001, The Role of Statistics in the Data Revolution? *International Statistics Review*, 69: 5–10.

[64] Froyd, Jeffrey, Andrea Beach, Charles Henderson, and Noah Finkelstein. 2008. Improving Educational Change Agents’ Efficacy in Science, Engineering, and Mathematics Education. In *Integrating the Sciences and Society: Challenges, Practices, and Potentials*, edited by H. Hartman, 16: 227-55. Research in Social Problems and Public Policy. Leeds: Emerald Group Publishing Limited. doi.org/10.1016/S0196-1152(08)16009-4

[65] Froyd, Jeffrey, Charles Henderson, Jean Layne, Andrea Beach, Noah Finkelstein, and R. Sam Larson. 2008. More Than Good Curricula: A Guide For Curricular Change Agents. In *2008 ASEE Annual Conference & Exposition Proceedings*, 13.906.1-13.906.14.Pittsburgh, Pennsylvania: American Society for Engineering Education. doi.org/10.18260/1-2--3539

[66] Funk, Cary. 2021. “Key Findings: How Americans’ Attitudes about Climate Change Differ by Generation, Party and Other Factors.” Pew Research Center, May 26, 2021. www.pewresearch.org/short-reads/2021/05/26/key-findings-how-americans-attitudes-about-climate-change-differ-by-generation-party-and-other-factors/.

[67] GAISE College Report ASA Committee. 2016. *Guidelines for Assessment and Instruction in Statistics Education College Report 2016*. Alexandria, VA: American Statistical Association. www.amstat.org/education/gaise.

[68] Gardner, Martin J, and Douglas G Altman. 1986. Confidence Intervals Rather than P Values: Estimation Rather than Hypothesis Testing. *BMJ* 292, no. 6522: 746–50. doi.org/10.1136/bmj.292.6522.746.

[69] Garet, Michael S., Andrew C. Porter, Laura Desimone, Beatrice F. Birman, and Kwang Suk Yoon. 2001. What makes professional development effective? Results from a national sample of teachers. *American educational research journal* 38(4): 915-945. doi.org/10.3102/00028312038004915.

[70] Garfield, Joan. 1993. Teaching Statistics Using Small-Group Cooperative Learning. *Journal of Statistics Education*, 1(1). https://doi.org/10.1080/10691898.1993.11910455

[71] Garfield, Joan, and Dani Ben-Zvi. 2008. *Preparing school teachers to develop students’ statistical reasoning. Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study: The 18th ICMI Study*, 299-310.

[72] Garfield, Joan, and Dani Ben-Zvi. 2009. Helping students develop statistical reasoning: Implementing a statistical reasoning learning environment. *Teaching Statistics*, 31(3): 72-77.

[73] Giambalvo, Ornella, and Linda Gattuso. 2008 ”Teachers training in a realistic context.” *International Association for Statistical Education*.

[74] Giersch, Sarah, Flora McMartin, Elizabeth Nilsen, Sheri Sheppard, and Phil Weilerstein. 2014. “Supporting Change in Entrepreneurship Education: Creating a Faculty Development Program Grounded in Results from a Literature Review.” In *2014 ASEE Annual Conference & Exposition Proceedings*, 24.1144.1-24.1144.12. Indianapolis, Indiana: ASEE Conferences. doi.org/10.18260/1-2--23077

[75] Gonchar, Michael. “Teach about Inequality with These 28 New York Times Graphs.” *The New York Times*, May 12, 2021. www.nytimes.com/2021/05/11/learning/lesson-plans/teach-about-inequality-with-these-28-new-york-times-graphs.html.

[76] Goos, Merrilyn, Shelley Dole, and Katie Makar. 2007. Designing professional development to support teachers’ learning in complex environments. *Mathematics Teacher Education and Development* 8: 23-47.

[77] Gould, Robert. 2010. Statistics and the Modern Student. *International Statistical Review / Revue Internationale de Statistique* 78(2): 297–315.

[78] Gould, Robert, Roxy Peck, Julie Hanson, Nicholas Horton, Brian Kotz, Kathy Kubo, Joyce Malyn-Smith, Mary Rudis, Brad Thompson, Mark Daniel Ward, and Rebecca Wong. 2018. “The Two-Year College Data Science Summit.” American Statistical Association. www.amstat.org/asa/files/pdfs/2018TYCDS-Final-Report.pd.

[79] Greenfield, Tony. 1993. Communicating statistics. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 156(2): 287-297.

[80] Gundlach, Ellen, and Mark Daniel Ward. 2021. The Data Mine: Enabling Data Science Across the Curriculum. *Journal of Statistics and Data Science Education*, 29(sup1): S74–S82. doi.org/10.1080/10691898.2020.1848484

[81] Hancock, Chris, James J. Kaput, and Lynn T. Goldsmith. 1992. Authentic inquiry with data: Critical barriers to classroom implementation. *Educational Psychologist* 27(3): 337-364.

[82] Hardin, Johanna, Nicholas J. Horton, Deborah Nolan, and Duncan Temple Lang. 2021. Computing in the statistics curricula: A 10-year retrospective. *Journal of Statistics and Data Science Education* 29, no. sup1: S4-S6.

[83] Harris, Alma, and Michelle Jones. 2010. Professional learning communities and system improvement. *Improving Schools*, 13(2): 172-181. doi.org/10.1177/1365480210376487

[84] Hart, Lynn C. 2002. A four year follow-up study of teachers' beliefs after participating in a teacher enhancement project. In: Leder, G.C., Pehkonen, E., Törner, G. (eds) *Beliefs: A hidden variable in mathematics education?*: 161-176. Mathematics Education Library, vol 31. Springer, Dordrecht. doi.org/10.1007/0-306-47958-3_10

[85] Hauser, Peter C., Karen L. Finch, and Angela B. Hauser. 2008. *Deaf Professionals and Designated Interpreters: A New Paradigm*. Washington, DC: Gallaudet University Press.

[86] Heaton, Ruth M., and William T. Mickelson. 2002. The learning and teaching of statistical investigation in teaching and teacher education. *Journal of Mathematics Teacher Education* 5: 35–59.

[87] Heck, Daniel J., Courtney L. Plumley, Despina A. Stylianou, Adrienne A. Smith, and Gwendolyn Moffett. 2019. Scaling up innovative learning in mathematics: exploring the effect of different professional development approaches on teacher knowledge, beliefs, and instructional practice. *Educational Studies in Mathematics* 102: 319–342.

[88] Henderson, Charles, Andrea Beach, and Noah Finkelstein. 2011. Facilitating Change in Undergraduate STEM Instructional Practices: An Analytic Review of the Literature. *Journal of Research in Science Teaching*. 48(8): 952–84. doi.org/10.1002/tea.20439.

[89] Hesterberg, Tim C. 1998. *American Statistical Association Proceedings of the Section on Statistical Education*, 44-52. American Statistical Association, Alexandria, VA.

[90] Hilliard, Ann Toler, and Edward Newsome, Jr. 2013. Effective Communication and Creating Professional Learning Communities Is a Valuable Practice for Superintendents. *Contemporary Issues in Education Research*. 6(4): 353–64.

[91] Hodgson, Ted, and Maurice Burke. 2000. On simulation and the teaching of statistics. *Teaching Statistics*, 22(3): 91–96.

[92] Holdren, John P. and Eric Lander. 2012. Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. *PCAST Report to the President. Executive Office of the President: President's Council of Advisors on Science and Technology*.

[93] Holmes, Peter. 2002. Some lessons to be learned from curriculum developments in statistics. In *Sixth International Conference on Teaching Statistics, CapeTown, South Africa*.

[94] Horst, Allison M, Alison Presmanes Hill, and Kristen B Gorman. 2020. "Palmerpenguins: Palmer Achipelago (Antarctica) Penguin Data. R Package Version 0.1.0." Zenodo. doi.org/10.5281/ZENODO.3960218.

[95] Horton, Nicholas J. 2015. Challenges and opportunities for statistics and statistical education: looking back, looking forward. *The American Statistician* 69(2): 138–145.

[96] Horton, Nicholas J., and Johanna S. Hardin. 2021. Integrating Computing in the Statistics and Data Science Curriculum: Creative Structures, Novel Skills and Habits, and Ways to Teach Computational Thinking. *Journal of Statistics and Data Science Education*, 29(sup1): S1–S3. doi.org/10.1080/10691898.2020.1870416

[97] Horton, Nicholas J. and Suzanne Switzer. 2005. Statistical Methods in the Journal. *The New England Journal of Medicine*. 355(18): 1977–1979.

[98] Houghton, Lisa A., Andrew R. Gray, Michelle J. Harper, Pattanee Winichagoon, Tippawan Pongcharoen, Sueppong Gowachirapant, and Rosalind S. Gibson. 2014. Vitamin D Status among Thai School Children and the Association with 1,25-Dihydroxyvitamin D and Parathyroid Hormone Levels. *PloS one* 9(8): e104825.

[99] Houston, Kyle, H. Bruce Lammers, and Shirley Svorny. 2010. Perceptions of the Effect of Public Policy on Employment Opportunities for Individuals Who are Deaf or Hard of Hearing. *Journal of Disability Policy Studies* 21(1): 9-21. doi.org/10.1177/1044207309357428.

[100] Hulsizer, Michael R and Linda M. Woolf. 2009. A guide to teaching statistics: Innovations and best practices. John Wiley & Sons.

[101] Ireland, Danyelle T., Kimberley Edelin Freeman, Cynthia E. Winston-Proctor, Kendra D. DeLaine, Stacey McDonald Lowe, and Kamilah M. Woodson. 2018. (Un) Hidden Figures: A Synthesis of Research Examining the Intersectional Experiences of Black Women and Girls in STEM Education. *Review of Research in Education* 42(1): 226–254. doi.org/10.3102/0091732X18759072.

[102] Ismay, Chester and Albert Y. Kim. *Statistical Inference via Data Science: A Modern Dive Into R and the Tidyverse*. Retrieved online April 27, 2024. moderndive.com/.

[103] Jumping Rivers. 2023. *R User Groups — A List of R Conferences and Meetings*. jumpingrivers.github.io/meetingsR/r-user-groups.html.

[104] Keeler, Carolyn M., and R. Kirk Steinhorst. 1995. Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education* 3(2).

[105] Kezar, Adrianna. 2011. What Is the Best Way to Achieve Broader Reach of Improved Practices in Higher Education? *Innovative Higher Education*, 36(4): 235-47. doi.org/10.1007/s10755-011-9174-z.

[106] Kezar, Adrianna, and Susan Elrod. 2012. Facilitating Interdisciplinary Learning: Lessons from Project Kaleidoscope. *Change: The Magazine of Higher Learning*, 44(1): 16-25. doi.org/10.1080/00091383.2012.635999.

[107] Kezar, Adrianna, and Jaime Lester. 2009.Promoting Grassroots Change in Higher Education: The Promise of Virtual Networks. *Change: The Magazine of Higher Learning*, 41(2): 44-51. doi.org/10.3200/CHNG.41.2.44-51.

[108] King, Kathleen P. 2002. Educational Technology Professional Development as Transformative Learning Opportunities. *Computers & Education*, 39(3): 283–97. doi.org/10.1016/S0360-1315(02)00073-8.

[109] Knuth, Donald E. 1984. "Literate Programming." *The Computer Journal* 27 (2): 97–111.

[110] Koehler, Jonathan J. 2011. Misconceptions about statistics and statistical evidence. In *Handbook of trial consulting* (pp. 121–133). Boston, MA: Springer US.

[111] Kotter, John P., and Leonard A. Schlesinger. 2008. Choosing Strategies for Change. Harvard Business Review.

[112] Kozak, Kathryn. 2020. Statistics Using Technology. 3rd ed. Retrieved online April 27, 2024. Statistics Using Technology.

[113] Kvam, Paul H. 2000. The effect of active learning methods on student retention in engineering statistics. *The American Statistician* 54(2): 136-140.

[114] Lave, Jean, and Etienne Wenger. 1991. *Situated Learning: Legitimate Peripheral Participation*. 1st ed. Hoboken, NJ: Cambridge University Press. doi.org/10.1017/CBO9780511815355.

[115] Lee, Hollylynne S., and Karen F. Hollebrands. 2008. Preparing to teach data analysis and probability with technology. *Joint ICMI/IASE study: Teaching statistics in school mathematics. Challenges for teaching and teacher education. Proceedings of the ICMI Study* 18.

[116] Lee, Hollylynne, and Dung Tran. 2015. Statistical habits of mind. *Teaching statistics through data investigations MOOC-Ed*: 1–4.

[117] Lock, Robin H., Patti Frazer Lock, Kari Lock Morgan, Eric F. Lock, and Dennis F. Lock, (2020). *Statistics: Unlocking the Power of Data*. 3rd ed. Hoboken, NJ: Wiley.

[118] Makar, Katie, and Jere Confrey. 2007. Moving the context of modelling to the forefront: Preservice teachers' investigations of equity in testing. In *Modelling and applications in mathematics education: The 14th ICMI Study* (pp. 485-490). Boston, MA: Springer US.

[119] Margolis, Jane, and Allan Fisher. 2002. *Unlocking the Clubhouse*. Cambridge, MA: MIT Press.

[120] McDonald, Sarah, and Nicholas J. Horton. 2019. Data scraping, ingestation, and modeling: bringing data from cars.com into the intro stats class. *CHANCE*, 32(4): 57–64.

[121] McKenna, Ann F., and Bugrahan Yalvac. 2007. Characterizing Engineering Faculty's Teaching Approaches. *Teaching in Higher Education*, 12(3): 405–18. doi.org/10.1080/13562510701278765

[122] McNamara, Amelia. 2018. R Syntax Comparison Cheatsheet.

[123] —. 2022a. Teaching Modeling in Introductory Statistics: A Comparison of Formula and Tidyverse Syntaxes. *arXiv:2201.12960 [Stat]*, January. arxiv.org/abs/2201.12960

[124] —. 2022b. STAT 220 Labs (Formula and Tidyverse Versions). www.amelia.mn/STAT220labs.

[125] Mei, Katelyn, and Alex Lyford. 2022. Where is data ethics in postsecondary data science courses? *Proceedings of the Eleventh International Conference on Teaching Statistics (ICOTS11, Rosario, Argentina*. ISI/IASE.

[126] Meletiou-Mavrotheris, Maria. 2003. Technological Tools in the Introductory Statistics Classroom: Effects on Student Understanding of Inferential Statistics. *International Journal of Computers for Mathematical Learning* 8: 265–297. doi.org/10.1023/B:IJCO.0000021794.08422.65

[127] Mills, Jamie D. 2004. Learning Abstract Statistics Concepts Using Simulation. *Educational Research Quarterly*, 28(4): 18–33.

[128] Movshovitz-Hadar, Nitsa, and Rina Hadass. 1990. Preservice education of math teachers using paradoxes. *Educational Studies in Mathematics* 21(3): 265–287.

[129] National Council of Teachers of Mathematics. 2000. *Principles and Standard for School Mathematics*. Reston, VA: NCTM.

[130] National Center for Science and Engineering Statistics (NCSES). 2017. "Women, Minorities, and Persons with Disabilities in Science and Engineering." National Science Foundation. www.nsf.gov/statistics/2017/nsf17310/

[131] National Science Foundation. February 29, 2016. Award Abstract # 0920350 - Phase II: Building a Community around Modeling, Statistics, Computation, and Calculus. www.nsf.gov/awardsearch/showAward?AWD_ID=0920350&HistoricalAwards=false.

[132] —. September 5, 2019. Award Abstract # 1432251 - Training A New Generation Of Statistics Educators (TANGO Stat Ed). www.nsf.gov/awardsearch/showAward?AWD_ID=1432251&HistoricalAwards=false.

[133] —. January 18, 2024. Award Abstract # 1626337 - Professional Development Emphasizing Data-Centered Resources and Pedagogies for Instructors of Undergraduate Introductory Statistics (StatPREP). www.nsf.gov/awardsearch/showAward?AWD_ID=1626337.

[134] Nederbragt, Alexander, Rayna Michelle Harris, Alison Presmanes Hill, and Greg Wilson. 2020. Ten Quick Tips for Teaching with Participatory Live Coding. Edited by Francis Ouel lette. *PLOS Computational Biology* 16 (9): e1008090. doi.org/10.1371/journal.pcbi.1008090.

[135] Neumann, David L., Michelle Hood, and Michelle M. Neumann. 2013. Using real-life data when teaching statistics: Student perceptions of this strategy in an introductory statistics course. *Statistics Education Research Journal*, 12(2): 59–70.

[136] Neumann, David L., Michelle M. Neumann, and Michelle Hood. 2010. The development and evaluation of a survey that makes use of student data to teach statistics. *Journal of Statistics Education*, 18(1).

[137] Nolan, Deborah, and Duncan Temple Lang. 2010. Computing in the statistics curricula. *The American Statistician*, 64(2): 97–107.

[138] Nolan, Deborah, and Duncan Temple Lang. 2015. Explorations in statistics research: An approach to expose undergraduates to authentic data analysis. *The American Statistician*, 69(4): 292–299.

[139] Nolan, Deborah, and Jamis Perrett. 2016. Teaching and Learning Data Visualization: Ideas and Assignments. *The American Statistician*, 70(3): 260–269. doi.org/10.1080/00031305.2015.1123651

[140] Ong, Maria, Janet M. Smith, and Lily T. Ko. 2018. Counterspaces for Women of Color in STEM Higher Education: Marginal and Central Spaces for Persistence and Success. *Journal of Research in Science Teaching* 55(2): 206–245. doi.org/10.1002/tea.21417.

[141] Panitsides, Eugenia, and Eirini Andromachi Kiouka. 2018. Muslim Minority Women in Western Thrace: Any Room for Transformative Learning? *Education + Training* 60(6): 596–607. doi.org/10.1108/ET-02-2018-0052.

[142] Parker, Micaela S., Arlyn E. Burgess, and Philip E. Bourne. 2021. Ten Simple Rules for Starting (and Sustaining) an Academic Data Science Initiative. *PLOS Computational Biology* 17(2): e1008628. doi.org/10.1371/journal.pcbi.1008628.

[143] Peters, Tim. 2004. PEP 20 – The Zen of Python. *Python.org*. peps.python.org/pep-0020/.

[144] Peterson, Anna D. and Laura Ziegler. 2021. Building a Multiple Linear Regression Model With LEGO Brick Data. *Journal of Statistics and Data Science Education* 29(3): 297–303. doi.org/10.1080/26939169.2021.1946450

[145] Pfaff, Thomas J., and Aaron Weinberg. 2017. Do Hands-on Activities Increase Student Understanding?: A Case Study. *Journal of Statistics Education* 17(3). doi.org/10.1080/10691898.2009.11889536

[146] Plumb, Carolyn, and Richard M. Reis. 2007. Creating Change in Engineering Education: A Model for Collaboration Among Institutions. *Change: The Magazine of Higher Learning* 39(3): 22–31. doi.org/10.3200/CHNG.39.3.22-31

[147] Posit PBC. 2023. RStudio Community. *RStudio Community*. community.rstudio.com/.

[148] Prosser, Michael, and Keith Trigwell. 1999. *Understanding Learning and Teaching: The Experience in Higher Education*. Open University Press, Philadelphia, PA.

[149] Pruim, Randall. NHANES: Data from the US National Health and Nutrition Examination Study. R package version 2.1.0. Retrieved online April 27, 2024. cran.r-project.org/package=NHANES.

[150] Pruim, Randall, Daniel Kaplan, and Nicholas J. Horton. 2017. The mosaic Package: Helping Students 'Think with Data' Using R. *The R Journal* 9 (1).

[151] —. 2023. *mosaicData: Project MOSAIC Data Sets*. R package version 0.20.4. github.com/ProjectMOSAIC/mosaicData.

[152] R Core Team. 2009. An Introduction to R. rstudio.github.io/r-manuals/r-intro/.

[153] —. 2015. *Comprehensive R Archive Network*. cran.r-project.org/.

[154] —. 2022. *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna, Austria. www.R-project.org/.

[155] R4DS. 2023. R for Data Science Online Learning Community.

[156] Radke-Sharpe, Norean. 1991. Writing as a component of statistics education. *The American Statistician 45*, 4: 292–293.

[157] Rafalski, Timothy, P. Merlin Uesbeck, Cristina Panks-Meloney, Patrick Daleiden, William Allee, Amelia McNamara, and Andreas Stefik. 2019. A Randomized Controlled Trial on the Wild Wild West of Scientific Computing with Student Learners. *Proceedings of the 2019 ACM Conference on International Computing Education Research*, 239–47.

[158] Rawlings-Goss, Renata, Lillian Cassel, Melissa Cragin, Catherine Cramer, Angela Dingle, Shawnta Friday-Stroud, Al Herron et al. 2018. "Keeping Data Science Broad: Negotiating the Digital and Data Divide Among Higher-Education Institutions." National Science Foundation. scholar.valpo.edu/math_stat_fac_pubs/64/.

[159] Ridgeway, Jim. 2016. Implications of the Data Revolution for Statistics Education. *International Statistical Review* 84(4): 528-549.

[160] Rieley, Michael. 2018. Big Data Adds up to Opportunities in Math Careers. *Beyond the Numbers: Employment & Unemployment* 7(8). U.S. Bureau of Labor Statistics. www.bls.gov/opub/btn/volume-7/big-data-adds-up.htm.

[161] Roble, Dennis B. 2017. Communicating and valuing students' productive struggle and creativity in calculus. *Turkish Online Journal of Design Art and Communication*, 7(2): 255–263.

[162] Rogers, Meredith Park, Sandra Abell, John Lannin, Chia-Yu Wang, Kusalin Musikul, David Barker, and Shannon Dingman. 2007. Effective professional development in science and mathematics education: Teachers' and facilitators' views. *International journal of science and mathematics education* 5: 507-532.

[163] RStudio Team. 2014. *RStudio: Integrated Development for R*. www.rstudio.com/products/rstudio/.

[164] —. 2018. *R Markdown*. rmarkdown.rstudio.com/.

[165] Rumsey, Deborah J. 2002. Statistical literacy as a goal for introductory statistics courses. *Journal of statistics education*, 10(3).

[166] Rutz, Carol, William Condon, Ellen R. Iverson, Cathryn A. Manduca, and Gudrun Willett. 2012. Faculty Professional Development and Student Learning: What Is the Relationship. *Change: The Magazine of Higher Learning* 44(3): 40–47. doi.org/10.1080/00091383.2012.672915

[167] Saavedra, Elizabeth. 1996. Teachers study groups: Contexts for transformative learning and action. *Theory into Practice* 35(4): 271–277.

[168] Saibene, Yanina Bellini, Claudia Vitolo, Erin LeDell, Hannah Frick, and Laura Acion. 2020. R-Ladies Global, a Worldwide Organisation to Promote Gender Diversity in the R Community. EGU General Assembly. May 2020. doi.org/10.5194/egusphere-egu2020-20530.

[169] Saxe, Karen, and Linda Braddy. 2015. *A Common* Vision for Undergraduate Mathematical Sciences Programs in 2025. Washington, DC: Mathematical Association of America.

[170] Scheaffer, R. L. 2001. Statistics education: Perusing the past, embracing the present, and charting the future. *Newsletter for the section on statistical education*, 7(1).

[171] Schlager, Mark S., and Judith Fusco. 2003. Teacher professional development, technology, and communities of practice: Are we putting the cart before the horse?. *The information society* 19(3): 203–220.

[172] Schneiter, Kady, and Jurgen Symanzik. 2013. An Applet for the Investigation of Simpson's Paradox, *Journal of Statistics and Data Science Education* 21(1). doi.org/10.1080/10691898.2013.11889661.

[173] Schwab-McCoy, Aimee, Catherine M. Baker,and Rebecca E. Gasper. 2021. Data Science in 2020: Computing, Curricula, and Challenges for the Next 10 Years. *Journal of Statistics and Data Science Education* 29(sup1): S40–S50. doi.org/10.1080/10691898.2020.1851159

[174] Silva, Dustin. "Affective Domain and Its Impact on Community College Success in Math." PhD dissertation, University of La Verne, 2020.

[175] Schultheis, Elizabeth H., and Melissa K. Kjelvik. 2020. Using messy, authentic data to promote data literacy & reveal the nature of science. *The American Biology Teacher*, 82(7): 439–446.

[176] Singer, Judith D., and John B. Willett. 1990. Improving the teaching of applied statistics: Putting the data back into data analysis. *The American Statistician*, 44(3): 223–230.

[177] Sirkin, Harold L., Perry Keenan, and Alan Jackson. 2005. The Hard Side of Change Management. *Harvard Business Review*, 83(10): 108–18, 158.

[178] Sly, Liz. "U.S. Soldiers Are Revealing Sensitive and Dangerous Information by Jogging." The Washington Post, January 29, 2018.

[179] Smith, Thomas M., Laura M. Desimone, and Koji Ueno. 2005. "Highly qualified" to do what? The relationship between NCLB teacher quality mandates and the use of reform-oriented instruction in middle school mathematics. *Educational evaluation and policy analysis* 27(1): 75–109.

[180] Sowey, Eric R. 1995. Teaching statistics: Making it memorable. *Journal of Statistics and Data Science Education* 3(2).

[181] StatPREP. 2018. *Little Apps for Teaching Stats*. www.statprep.org/LittleAppSite/. Support for StatPREP is provided by the National Science Foundation, Award DUE-1626337.

[182] Steinert, Yvonne, Richard L. Cruess, Sylvia R. Cruess, J. Donald Boudreau, and Abraham Fuks. 2007. Faculty Development as an Instrument of Change: A Case Study on Teaching Professionalism. *Academic Medicine: Journal of the Association of American Medical Colleges*, 82(11): 11057–64. doi.org/10.1097/01.ACM.0000285346.87708.67

[183] Steinert, Yvonne, Karen Mann, Angel Centeno, Diana Dolmans, John Spencer, Mark Gelula, and David Prideaux. 2006. A Systematic Review of Faculty Development Initiatives Designed to Improve Teaching Effectiveness in Medical Education: BEME Guide No. 8. *Medical Teacher*, 28(6): 497–526. doi.org/10.1080/01421590600902976

[184] Stoll, Louise, Ray Bolam, Agnes McMahon, Mike Wallace, and Sally Thomas. 2006. Professional learning communities: A review of the literature. *Journal of educational change* 7(4): 221-258.

[185] Strayer, Jeremy F., Natasha E. Gerstenschlager, Lisa B. Green, Nancy Mccormick, Scott McDaniel, and Ginger Holmes Rowell. 2019. Toward a full (er) implementation of active learning. *Statistics Education Research Journal* 18(1): 63-82.

[186] Tableau online. "Guide to Data Cleaning: Definition, Benefits, Components, and How to Clean Your Data." Retrieved online April 28, 2024. www.tableau.com/learn/articles/what-is-data-cleaning.

[187] Tackett, Maria, and Mine Çetinkaya-Rundel. 2020. COVID-19 Data in the Classroom. *CHANCE*. 33:3, 63–66.

[188] Taylor, Stanley A., and Amy E. Mickel. 2014. Simpson's Paradox: A Data Set and Discrimination Case Study Exercise. *Journal of Statistics and Data Science Education* 22(1).

[189] Tintle, Nathan L., Beth L. Chance, George W. Cobb, Allan J. Rossman, Soma Roy, Todd Swanson, and Jill VanderStoep. 2020. *Introduction to statistical investigations*. 2nd ed. Hoboken, NJ: Wiley.

[190] Tintle, Nathan, Jake Clark, Karen Fischer, Beth Chance, George Cobb, Soma Roy, Todd Swanson, and Jill VanderStoep. 2018. Assessing the association between precourse metrics of student preparation and student performance in introductory statistics: Results from early data on simulation-based inference vs. nonsimulation-based inference. *Journal of Statistics Education* 26(2): 103–109.

[191] Tractenberg, Rochelle E. 2020. Ten Simple Rules for Integrating Ethics into Statistics and Data Science Instruction, 2E. *SocArXiv*. August 15. doi.org/10.31235/osf.io/z9uej.

[192] Trigwell, Keith, Michael Prosser, and Fiona Waterhouse. 1999. Relations between Teachers' Approaches to Teaching and Students' Approaches to Learning. *Higher Education* 37(1): 57–70. doi.org/10.1023/A:1003548313194.

[193] Trufelman, Avery. "226: On Average." Produced by Roman Mars. 99% Invisible. August 23, 2016. 99percentinvisible.org/episode/on-average/.

[194] "Fastest Growing Occupations." U.S. Bureau of Labor Statistics, April 17, 2024. www.bls.gov/emp/tables/fastest-growing-occupations.htm

[195] Utts, Jessica. 2021. Enhancing data science ethics through statistical education and practice. *International Statistical Review*, 89(1): 1–17.

[196] Utts, Jessica. 2015. The Many Facets of Statistics Education: 175 Years of Common Themes. *The American Statistician* 69(2): 100–107.

[197] Utts, Jessica. 2003. What Educated Citizens Should Know About Statistics and Probability. *The American Statistician* 57(2): 74–79.

[198] Vescio, Vicki, Dorene Ross, and Alyson Adams. 2008. A Review of Research on the Impact of Professional Learning Communities on Teaching Practices and Student Learning. *Teaching and Teacher Education*, 24: 80–91. doi.org/10.1016/j.tate.2007.01.004

[199] von Kugelgen, Julius, Luigi Gresele, and Bernhard Scholköpf. 2021. Simpson's Paradox in COVID-19 Case Fatality Rates: A Mediation Analysis of Age-Related Causal Effects. *IEEE Trans Artif Intell*. 2(1): 18–27.

[200] Wang, Xiaofei, Cynthia Rush, and Nicholas Jon Horton. 2017. Data Visualization on Day One: Bringing Big Ideas into Intro Stats Early and Often. *Technology Innovations in Statistics Education* 10 (1). doi.org/10.5070/T5101031737.

[201] Wasserstein, Ronald L., Allen L. Schirm, and Nicole A. Lazar. 2019. Statistical Inference in the 21st Century: A World Beyond p¡ 0.05. *The American Statistician*, 73.

[202] Watkins, Ann, Anna Bargagliotti, and Christine Franklin. 2014. Simulation of the sampling distribution of the mean can mislead. *Journal of Statistics Education*, 22(3): 1–21.

[203] Wenger, Etienne. 1998. *Communities of Practice: Learning, Meaning, and Identity*. 1st ed. Cambridge University Press. doi.org/10.1017/CBO9780511803932

[204] Wenger, Etienne, Beverly Trayner, and Maarten de Laat. 2011. *Promoting and Assessing Value Creation in Communities and Networks: A Conceptual Framework.* Ruud de Moor Centrum Rapporten 18. Netherlands: Ruud de Moor Centrum. www.open.ou.nl/rslmlt/Wenger_Trayner_DeLaat_Value_creation.pdf

[205] Wickham, Hadley, Mine Çetinkaya-Rundel, and Garrett Grolemund. 2016. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. "O'Reilly Media, Inc." Retrieved online April 28, 2024. r4ds.hadley.nz/.

[206] Wickham, Hadley, and Danielle Navarro. *ggplot2:Elegant Graphics for Data Analysis*. 3rd ed. Accessed April 27, 2024. ggplot2-book.org/.

[207] Wild, Chris J., and Maxine Pfannkuch. 1999. Statistical thinking in empirical enquiry. *International statistical review* 67(3): 223–248.

[208] Wild, C.J., Maxine Pfannkuch, Meredith Regan, and Nicholas J. Horton. 2011. Towards more accessible conceptions of statistical inference. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 174: 247–295. doi.org/10.1111/j.1467-985X.2010.00678.x

[209] Wilson, Greg. 2019. *Teaching Tech Together: How to create and deliver lessons that work and build a teaching community around them*. Taylor & Francis.

[210] Witmer, Jeff. 2015. How Much Do Minority Lives Matter? *Journal of Statistics and Data Science Education*, 23(2).

[211] Witt, Gary. 2013. Using Data from Climate Science to Teach Introductory Statistics. *Journal of Statistics Education* 21:1. doi.org/10.1080/10691898.2013.11889667

[212] Wolf, Fredric. Meta-Analysis. 1986, 2455 Teller Road, Newbury Park California 91320 United States of America: SAGE Publications, Inc. doi.org/10.4135/9781412984980.

[213] Woolston, Chris. (2015). Psychology journal bans P values. *Nature*, 519(7541): 9-9.

[214] Zieffler, Andrew, Jiyoon Park, Joan Garfield, Robert delMas, and Audbjorg Bjornsdottir. 2012. The Statistics Teaching Inventory: A Survey on Statistics Teachers' Classroom Practices and Beliefs. *Journal of Statistics Education* 20(1):3. doi.org/10.1080/10691898.2012.11889632

[215] Zorn, Paul, John Bailer, Linda Braddy, Jenna Carpenter, William Jaco, and Peter Turner. 2014. *The INGenIOuS Project: Mathematics, Statistics, and Preparing the 21st Century Workforce.* Washington, DC: Mathematical Association of America.