Data for social good in the classroom

July 2021

Data and statistics have been used to help describe and understand a variety of social phenomena in recent years, from helping society understand the COVID-19 pandemic and vaccines, to election forecasting, to understanding trends in policing and the criminal justice system in the United States. As statistics has been more widely used to understand social phenomena, it is important that students are equipped with skills to engage critically with statistics used in the world around them. These skills are outlined in two goals for introductory statistics classes from the 2016 Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report

- *Goal 1: Students should become critical consumers of statistically-based results reported in popular media, recognizing whether reported results reasonably follow from the study and analysis conducted.*
- *Goal 9: Students should demonstrate an awareness of ethical issues associated with sound statistical practice.*

With a mind towards achieving these goals, one of the GAISE (2016) recommendations for introductory statistics courses is to "integrate real data with context and purpose." In addition to equipping students with the skills outlined in the GAISE (2016) guidelines, using real data in the classroom helps keep students engaged as they are able to better connect with the examples they're presented and can more easily understand the relevance of statistics in their daily lives.

More specifically, when we have a data for social good approach, we help students understand the power of statistics and how it can impact their daily lives .

In this article we focus on two courses that takes a social good approach to teaching statistics. Then we share resources to help instructors start incorporating real data and social issues in their classrooms.

Consulting and Service Learning

Service Learning has been a cornerstone of all degree programs at California State University, Monterey Bay since its inception in 1994. Service learning is experience-based learning, where the student takes an active role in an experience or project that benefits the community while also deepening their own understanding of a particular curriculum. Our course, Service Learning for Mathematics and Statistics Consultants, began in 2013 at the request of mathematics majors who wished to engage in service that prepared them for careers other than teaching. We developed the course so that students engage with the question "How can

mathematicians and statisticians contribute to furthering social justice and equity through their profession?" Since it began, our 5-unit course has averaged 14 students each spring semester. Students are expected to have completed, at minimum, an introductory level statistics course. Most students are in their third year as Statistics and Mathematics majors at the point they enroll in the course. Our course is designed to meet the major learning outcome: Students demonstrate ethical reasoning, global awareness, and civic and intercultural engagement through statistical practice in ways that promote sustainability, and social justice and equity across diverse communities. The course-specific learning outcomes expect that at the end of this course, students will be able to

1. Explain the relationship between their social and cultural identities and their identity as a mathematics or statistics professional.

2. Identify and address the social inequities created by applications of mathematical and statistical methods.

3. Describe the professional responsibility of mathematicians and statisticians to promote social justice and equity.

4. Follow the statistical investigative process to collect, manipulate, and explore data, and analyze and conduct inference to address the need of a community partner.

5. Demonstrate ethics in statistical practice and communication through community-based projects.

6. Communicate with clients effectively and respectfully while considering the dynamics of power, privilege, knowledge of the community and its history.

7. Apply mathematical and statistical methods to expose social injustices, propose solutions, and take action to initiate change.

The course is structured around two main components: 1) the service learning project where students consult with local nonprofits and schools on data collection and analysis projects for a minimum of 30 hours, and 2) the classroom discussions. The service learning component is discussed in more detail in Unfried and Canner (2019), as are some of the details related to classroom discussions, so here I will focus on some more recent additions to the classroom components of the course.

Our classroom discussions focus on issues related to social justice and equity as it pertains to the ways statistics and algorithms impact communities and individuals. We build our classroom discussions, in part, around the topics addressed in Cathy O'Neil's book, Weapons of Math Destruction (2016). Students read a chapter each week, write a reflective journal, and then engage in class discussions on topics such as targeted-ads, insurance rates, and predictive policing. The book is beginning to show its age, though the general themes are still relevant. The book also focuses on the destructive aspects of statistics and data science. Therefore, we supplement the book with additional readings, podcasts, and videos to update

students on the topics or to address data for good topics. Most of the students have never before engaged in discussions about social justice or equity in the context of mathematics and statistics, therefore we must create a framework for our discussions early and reinforce it often to ensure productive conversations. The LARA (Listen, Affirm, Respond, Add/Ask) method provides a framework for discussions to create space for differing perspectives while reinforcing similar viewpoints (Zomorodi, 2014). The method is practical to prepare students for both classroom conversations and client conversations in the service learning projects. It is important to reinforce LARA and model it as an instructor throughout the course to ensure students feel comfortable sharing their views. It is also important, especially as an instructor from a privileged background, to approach each discussion and topic with cultural humility - giving voice to the students rather than imparting personal views.

We provide an opportunity for students to address topics of social justice and equity of interest through mini TED Talks. Students choose a topic of interest and present a data supported 5 minute presentation to the class, followed by a class discussion on the topic. The details of the assignment implementation are included in Unfried and Canner (2019), but a new addition to the project is a plan for civic action. We want to empower students to take action in different ways. Our first attempt at a plan for civic action was a simple assignment - write a letter, reinforced with the research they completed for the talk, to a person or organization in a position of power to enact change related to their TED Talk topic. Students wrote letters to politicians, community leaders, and industry leaders to ask them to consider how their decisions, policies, or businesses impacted various communities or contributed to systemic racism or exclusion. The impact of such action is difficult to measure, but the experience allows students to take personal action even if it does not yield immediate, or even positive, results. We look forward to how we might evolve the role of personal civic action in our course in the future.

For many of our students, our course is the first time they have confronted the fact that mathematics and statistics are not neutral. In the course, they begin to evaluate their social and professional responsibilities, regardless of future profession, and we culminate our discussions with the creation of a Hippocratic Oath" as statisticians, mathematicians, and data scientists. Though we have kept the structure of the course consistent the last few years, the conversations continue to evolve as students bring their experiences and backgrounds to the discussion.

Note: we are happy to share our course materials upon request jcanner@csumb.edu

U of Chicago

In 2014, in the midst of a budget crisis, Flint, Michigan switched its municipal water supply from treated water imported from Detroit to locally sourced water from the Flint River. Unfortunately, officials chose not to include standard corrosion control chemicals in the local water treatment. This led to degradation of lead service lines and an increase in water lead content to levels far above federally mandated limits. State and local officials initially denied that there was a problem, but a coalition of Flint citizens gathered evidence of the increased lead levels in both the water and in the blood of Flint children. Eventually, the activists

succeeded in pressuring the government to switch the water back and supply clean water in the meantime.

In 2019, I developed and first taught a 10-week introductory statistics class at the University of Chicago which uses this water crisis as a case study. The course serves undergraduate non-STEM majors, primarily first and second-year students who want to fulfill a core mathematics requirement. It covers typical intro statistics topics, including descriptive statistics, probability, correlation and regression, and inference, but does not aim to prepare students for higher-level statistics courses.

Students in the Flint course use statistical concepts to answer key questions about the water crisis and the conditions leading up to it. In the first unit, students use descriptive statistics to examine income inequality between the city and the surrounding suburbs. Using census data, they explore how this inequality has changed over time and between different racial groups. We pair these exercises with readings about historical drivers of such inequality, such as redlining. In the second unit, they learn about percentiles and sample size as they see how the government was able to manipulate early lead sampling data to avoid passing federal lead thresholds. In the third unit, students use probability and inference to evaluate study results showing an increase in the percentage of Flint children with elevated blood lead levels (eBLLs). In the fourth unit, students learn about regression and use this to read studies about the health effects of childhood lead exposure. Finally, in the fifth unit, students make predictions about the effects of the Flint crisis on children's future, and compare the size of the lead effect to the effects of other causes such as poverty.

During the course, students read a memoir by the Flint pediatrician who recognized the increase in eBLLs, Dr Mona Hanna-Attisha. They also have an introductory statistics textbook as a reference. But they primarily spend their time interacting with primary sources, such as census data or lead measurements, or reading scientific articles. Because these sources are difficult, students spend a great deal of class time working in small groups to interpret their readings. For homework and final projects, students write essays in which they must use statistics in making their arguments.

My biggest challenge when teaching course has come because the Flint course touches on important but difficult issues such as poverty and structural racism. As a statistics professor, I did not come in with training in addressing such sensitive topics in the classroom. I have found my university's Center for Teaching to be very helpful resource in giving feedback and ideas for making my classroom a safe space for learning, even as we tackle difficult concepts.

My goal in developing and teaching this course is to help students develop a sense of agency, with the citizen-scientists of Flint as a model. I hope that in their future lives, my students will have the confidence to believe they can use statistical tools and language when analyzing and addressing problems in their own communities.

Resources for teaching

References

- Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., ... & Wood, B. (2016). Guidelines for assessment and instruction in statistics education (GAISE) college report 2016. Available at amstat.org/asa/files/pdfs/GAISE/GaiseCollege_Full.pdf.
- Clark, A. (2018). *The poisoned city: Flint's water and the American urban tragedy*. Metropolitan Books.
- Hanna-Attisha, M. (2019). *What the eyes don't see: A story of crisis, resistance, and hope in an American city*. One World.
- Langkjær-Bain, R. (2017). The murky tale of Flint's deceptive water data. *Significance*, 14(2), 16-21.
- O'neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- Unfried, A., & Canner, J. (2019). Doing social justice: turning talk into action in a mathematics service learning course. *PRIMUS*, 29(3-4), 210-227.
- Wicklin, R. (2017, May 17). Quantiles and the Flint water crisis. SAS. https://blogs.sas.com/content/iml/2017/05/17/quantiles-flint-water-crisis.html
- Zomorodi, M. (2014, June 8). Your Facebook friend said something racist. Now What?. In *Note to Self*. WNYC Studios. https://www.wnycstudios.org/podcasts/notetoself/episodes/your-facebook-friend-said-something-racist-now-what