DATA INVESTIGATION PROCESS FOR TEACHING STATISTICS FOR SOCIAL JUSTICE

Heather Barker
Elon University, Elon, North Carolina, United States
hbarker2@elon.edu

This paper delves into the integration of teaching mathematics for social justice principles and a data investigation process framework within a statistics investigation utilizing a dataset that documents police-involved deaths. It outlines the development of the statistics investigation, showcasing the application of statistical techniques to real-world data. Furthermore, it introduces the Social Justice Data Investigation Process framework, illustrating the interplay between teaching mathematics for social justice, and the data investigation process. This framework elucidates how these components interact to foster critical inquiry and promote a sense of social agency through statistical analysis, offering valuable insights into the integration of social justice principles into data-driven inquiry. Preliminary insights from high school and college students who participated in this data investigation are also shared.

INTRODUCTION

In recent years, there has been a growing emphasis on incorporating content about teaching social justice within mathematics education (Berry et al., 2020; Gutstein, 2006). Teaching mathematics for social justice involves contextualizing math concepts within students' cultural and social backgrounds, fostering a deeper comprehension of mathematics (Berry et al., 2020). This approach enables students to critique societal issues, understand their impact, and advocate for change. To implement this method effectively, educators must appreciate students' cultures, understand knowledge development within cultural contexts, and recognize the influence of cultural institutions on learning (Rubel, 2017). By positioning students as active participants, teaching mathematics for social justice becomes essential for meaningful education. Gutstein (2006) outlined mathematical and social justice pedagogical goals for *Teaching Mathematics for Social Justice (TMSJ)*.

This shift towards a more inclusive and socially conscious pedagogy has extended to the field of statistics as well (Buell et al., 2019; Dogucu et al., 2023). Lesser (2007) suggested "[o]ne way to move towards understanding 'teaching statistics for social justice' (TSSJ) is mentally to substitute 'statistics' for 'mathematics'" (p. 3) in Gutstein's description of TMSJ. To integrate social justice principles into statistics education effectively, it is crucial to interpret this framework through a statistical lens using "conceptual and computational proficiency goals recommended by most leading statistics educators" (Lesser, 2007, p. 3).

Though the goals for TMSJ are useful in developing mathematical investigations, there has not been research to show how these goals can be used for statistics investigations specifically. Weiland and Williams (2023) coined the term *culturally relevant data* (CRD) "as data that are relevant and meaningful to students or relate to issues of social justice" (p. 2). CRD is built upon the seminal work of *culturally relevant pedagogy* in mathematics education defined by Ladson-Billings (1995). This paper will explore how creating a statistics investigation with CRD was used to develop a process for creating a Social Justice Data Investigation using the TMSJ pedagogical goals (Gutstein, 2006) as well as a framework built on prior statistics education research, the Data Investigations Process framework (Lee et al., 2022). Hopefully describing this process will prove useful for educators who want to use social justice teaching pedagogy in their statistics classes.

THEORETICAL FRAMEWORKS

Using Lesser's (2007) suggestion of combining Gutstein's TMSJ framework with recommendations from leading statistics educators, I chose to use the Data Investigation Process (DIP) framework developed by Lee et al. (2022) as a foundation for developing a statistics investigation using CRD. The DIP framework was informed from literature across decades of work from statistics educators and researchers who have proposed practices for statistics to describe data (Lee et al., 2022).

Data Investigation Process (DIP) Framework

Numerous frameworks have emerged to guide statistical investigations in recent decades. Lee et al. (2022) used decades of frameworks from statistics education and interviews with active data scientists to describe practices and processes when working with data. These frameworks range from the PCAI (pose question, collect data, analyze data, and interpret results) Model of Statistic Investigation (Graham, 1987) to the revised K-12 Guidelines for Assessment and Instruction in Statistics Education (GAISE II) framework (Bargagliotti et al., 2020). This informed the Data Investigation Process (DIP) Framework. The DIP framework encompasses six phases of a data investigation: problem framing, data consideration and collection, data processing, exploration and visualization, model consideration, and communication and action proposals. Lee et al., (2022) emphasized that these practices are not absent from existing frameworks, but rather the "purpose is to elevate specific practices and processes since they may be implicit in other frameworks" (p. 11).

Teaching Mathematics for Social Justice

Gutstein (2006), in his book, *Reading and Writing the World with Mathematics,* outlined social justice pedagogical goals and practices along with mathematical goals and practices to describe a framework for TMSJ pedagogy. The pedagogical goals are:

- Reading the world with mathematics – Students can use mathematics to analyze power dynamics, resource disparities, and discrimination (based on factors like race, class, gender, and language), and foster a critical examination of societal phenomena and their interconnections.
- Writing the world with mathematics – Students begin to feel they can become an advocate for social justice and begin to develop a sense of social agency.
- Developing positive social and cultural identities – Students recognize and value their own cultural and social identities, as well as begin to understand how they may become agents of social change.

Reading the world is considered more of a reflective process, whereas writing the world would require students to take some sort of action. Developing positive social and cultural identities should be attended to by educators by helping students honor their own backgrounds.

FATAL ENCOUNTERS DATA INVESTIGATION

The following shows how the pieces of the DIP framework were used to create a statistics investigation using CRD appropriate for a high school or college level statistics classroom. This section will also describe the TMSJ considerations used when creating the investigation. These considerations are then combined to describe a data investigation process for teaching statistics for social justice.

For the investigation, students explored data from a site called Fatal Encounters (https://fatalencounters.org/). The dataset included people who were killed in any interactions with police in the United States, not just directly during an arrest. The dataset contained people who may have been shot and killed by a police officer during a violent interaction, but also other indirect deaths such as people who may have been involved in a fatal car accident with an off-duty officer. The data spanned from January 1, 2000 through December 31, 2021 with approximately 35,000 records. Students were warned before engaging in the investigation that the data may be triggering and have an option to do another assignment.

Frame the Problem

To give the students purpose for exploring the data, they considered a then current social justice issue. In May 2020, George Floyd (a Black American citizen) died as a direct result of police brutality during a routine traffic stop in Minnesota. Though this was not the first time a Black citizen had been unlawfully killed at the hands of the police, the death shocked the nation when footage of Floyd's death became viral (Hill et al., 2020). Because of the civil unrest as a result of this injustice, considering fatal encounters with police is an example of CRD (Weiland & Williams, 2024) for students in the United States. Students were then introduced to the idea of de-escalation training. De-escalation training is a process designed to prevent intensification of a conflict (such as interactions between police and civilians) that, ideally, resolves conflicts peacefully. Anecdotally, people hope that de-escalation training works, though there are few studies showing that it has helped in reducing fatal

encounters with police. To illustrate how this may relate to their own identity, students were shown a map of the United States where de-escalation training was required for law enforcement officers (Stockton, 2021) to see whether their own state required the training.

Additionally, students were shown a newsclip about de-escalation training. Students then engaged in class discussion to consider questions such as: What are the benefits or drawbacks of de-escalation training? How could we determine if de-escalation training is making a difference in encounters among civilians and officers? For this investigation students are given the following initial investigative question: *Do states that require de-escalation training show a decrease in the number of fatal encounters with police compared to states that do not require de-escalation training?*

Consider and Gather Data

Students accessed the Fatal Encounters data through a data portal that was created for CODAP (Common Online Data Analysis Platform; https://codap.concord.org/), a free, online dynamic statistical software tool. Students were first given a sample of 200 random individuals in the dataset (Figure 1). This gave students an opportunity to familiarize themselves with the variables in the dataset, as well as the capabilities of CODAP if they had not previously used the tool. Students were presented with several questions, such as which parts of the country show clusters of fatal encounters? Are those areas that require de-escalation training?
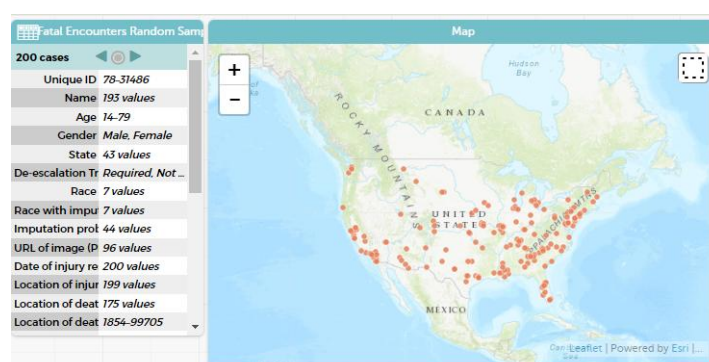

Figure 1. Initial sample of 200 fatal encounters with United States map.

After considering the sample of 200 people and the variables present, students were then directed to the Fatal Encounters data portal (https://concord-consortium.github.io/codap-data-interactives/FatalEncounters/) . This portal allowed students to collect a subset of the entire dataset by either year or by state. Students were then instructed to choose data from two states, one that required de-escalation training and one that did not. Students were asked to create a graph of the variable *gender* (gender of the victim) to see how the counts for each gender compared, then separate data based on the states they chose. This then led them to explore other variables of interest in the dataset to see how the states compare. Students shared their exploration with a partner. This inherently introduced the idea of variability across the chosen datasets since different peers in class were likely to choose different states to compare as well as different variables to investigate.

Process Data

Referring to the driving statistical investigation question, students tried to determine if states that do or do not require de-escalation training had seen a difference in the number of fatal encounters over the years. To do this, students had to count the number of fatal encounters by year. In CODAP, they first grouped the data into a hierarchical table by year, then created a variable to count the number of cases for each year (Figure 2).

Explore and Visualize Data

After creating the count variable, students were tasked with generating a scatterplot illustrating the correlation between year and count. They were presented with the option to segment this scatterplot by state to discern potential variations (Figure 3). Following this analysis, students were encouraged to explore how additional variables (such as *Cause of Death, Race, Intended Use of*

Force) may have influenced the frequency of fatal encounters across different states. Students were encouraged to create more graphs and try to begin to think about an argument for or against de-escalation training based on the graphs they created. They could then go back to consider and gather data phase if desired to explore different states.
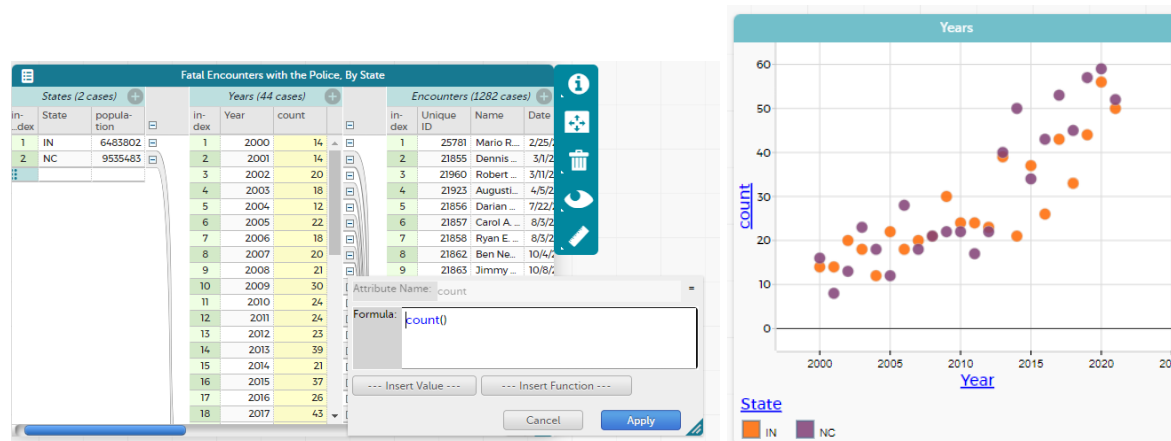
Figures 2 and 3. Using CODAP to create a hierarchical table by year and create a count variable and a scatterplot comparing Year, Count, and State.

Consider Models

Though it was not explicitly required for this investigation, students could get to the phase of considering models in their exploration. They might have been interested in creating statistical measures, such as measures of centers. For these data the only quantitative variable in the original data was age, though they might have created other variables that could have then been used to create statistical measures. Students might also have chosen to create a linear model to describe the relationship between year and count to predict the number of fatal encounters for subsequent years.

Communicating and Proposing Action

As students were exploring the data, they were tasked to produce a letter to a senator to communicate their findings. When this investigation was created, the *Law Enforcement De-Escalation Training Bill* was an active bill in the United States Congress. This bill would provide funding for state and local law enforcement to train officers, and mental health professionals who support law enforcement, in de-escalation tactics. This bill would authorize $70 million in grant funding for training. Students were tasked to write a letter to their senator voicing their support for or against the bill. The letter had to include statistics and at least one graph from their statistics investigation to support their opinion. (Note: this bill was passed into law in 2022.)

Teaching Mathematics for Social Justice Considerations

The TMSJ pedagogical goals were utilized in the development of this investigation. Students were asked to *read the world with statistics* by using the fatal encounters data to analyze and make sense of the phenomenon of police involved deaths in the United States. They initially made sense of this phenomenon in a smaller dataset, then expanded on their understanding by looking at their own dataset. In the larger investigation they continued making sense of the world by creating visualizations and other statistical measures. Students were asked to *write the world with statistics* by using the graphs and measures they found to make an argument for or against the passage of a bill. They were then tasked with writing a letter to a senator as a step to show them how they can take action themselves to bring about change. Finally, an attempt to *develop positive social and cultural identities* was accomplished by encouraging students to find data from their own state as well as using their own findings to voice their opinion for or against the bill, giving them a chance to form their own voice instead of echoing the views of the teacher.

Teaching Statistics for Social Justice

To make the TMSJ pedagogical goals more relevant for teaching statistics, I propose the following pedagogical goals for Teaching Statistics for Social Justice (TSSJ) built upon TMSJ goals.

- *Discovering Social Realities with Statistics*: This goal uses statistical analysis to uncover and illuminate underlying patterns and disparities within social contexts.
- *Empowering Learners to use Statistics for Social Change*: This goal equips learners with the statistical tools and knowledge necessary to advocate for and enact positive social change.
- *Reflecting on Personal Identity within Social Issues using Data*: This goal encourages learners to reflect on their own identities and perspectives within the context of social justice issues, fostering personal growth and empathy.
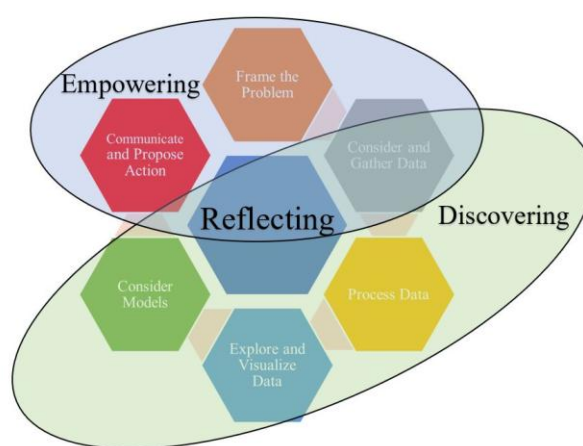

Figure 4. Social Justice Data Investigation Process

When comparing the data investigation process framework to the TSSJ pedagogical goals, there seems to be defined overlap. The *discovering* goals encompass consider and gather data, process data, explore and visualize data, and consider models pieces of the process. The *empowering* goal contains the frame the problem, communicate and propose action, and consider/gather data pieces. *Reflecting* is entrenched in each piece of the process as students make decisions about how they engage in the investigation based on how they perceive their world. Perhaps the intersection of TSSJ and the DIP framework can be identified as the Social Justice Data Investigation Process (Figure 4).

When framing a problem for a social justice data investigation, there may be questions that can be posed or answered using a social justice dataset that may not be one that can influence social change. When this is the case, framing the problem may lie in *discovering social realities with statistics*. Additionally, the consider/gather data piece of the process may be part of the *empowering* process as students may want to gather more data to help a hypothesis about social change. As much as possible, to make the problems meaningful for students, the investigation should be centered around investigative questions that could lead to social change.

PRELIMINARY FINDINGS FROM STUDENT INVESTIGATIONS

Data for a research study has been collected from 8 high school students and 6 college students who engaged in the fatal encounters investigation. These students were given pre- and post-surveys about their experience with the investigation and screencasts of them working as partners through the investigation were recorded (6 total screencasts, ~1 hour each). These data were used to analyze the parts of a CRD were of most interest to students and what parts of the DIP framework students engaged with. Below are some preliminary findings about their experiences.

In the post-survey all 14 participants agreed that exploring data about a relevant topic such as the fatal encounters data would be useful to them as statistics students. Students also voiced that this sort of data can be more engaging than other data. One of the college students stated that when they

are "working with boring datasets I usually lose focus or stop paying attention, but with this dataset, I was really intrigued by what happened".

All 6 pairs tended to spend most of their time in the Explore and Visualize Data phase, which is not surprising given the dynamic nature of CODAP. Figure 5 shows how a pair of college students engaged in this phase by creating multiple graphs to visualize different variables.
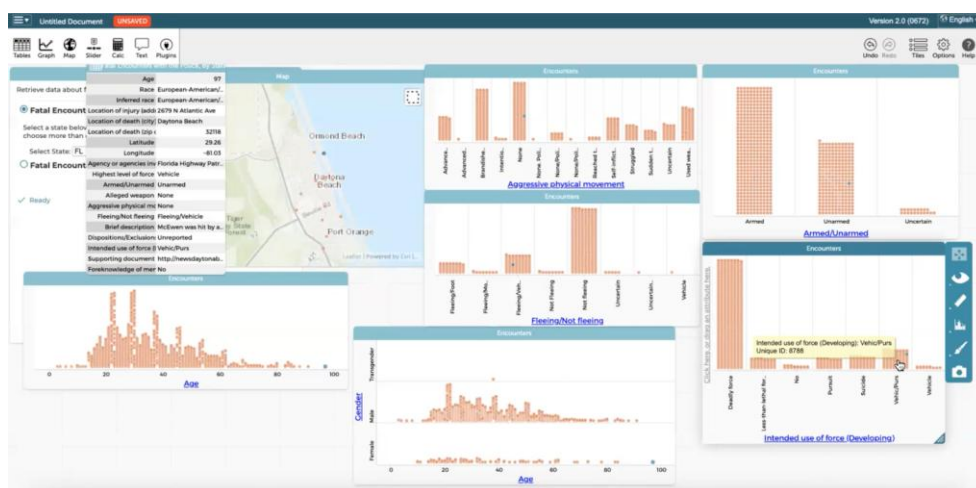

Figure 5.College students using CODAP to compare several variables of the data set.

Additionally, students were able to see the connection of how data can be used as a tool to Propose Action. When students were asked about their general thoughts about this data investigation, most shared positivity for the process. One high school student stated, "I would love to learn more on this topic considering that social injustice is something that I want to help slowly correct over time". Another college student stated that when looking at the data they wondered how it can be used in the real world "to hopefully reduce the amounts of the incidents".

CONCLUSION

As statistics instructors, we have a unique opportunity to embed CRD, particularly data tied to social justice, into our courses. But this is not an easy task for instructors who may worry that discussing these issues will cause discomfort. Lesser (2007) warns that this sort of detachment, or assuming these ideas are brought up in other departments, may lead to students missing out on growth of their own social agency. Instructors should be open to learning as part of this process as well. Engaging in social justice can also include data that uplifts marginalized groups such as educational data showing a decreasing achievement gap in students of color and white students.

Statistics instructors interested in teaching using social justice data should refer to others sharing work in this area (Berry et al., 2020; Skew the Script, www.skewthescript.org), but also realize that starting out one should not be expected to engage in all aspects of social justice investigation process for all investigations. The example of this paper is to only be used as a model. There will often be times when students only read the world with statistics and still help towards furthering their sense of social agency. To teach for social justice effectively, however, there should be times when students have opportunities to take further action to help empower their sense of social agency using statistics as a tool.

FUTURE RESEARCH

The screencast and pre- and post-survey data described previously will be used to create a case study of student experiences. Additionally, a collection of K-12 lessons has been collected from various sources to determine which parts of the social justice data investigation process are present. We hope this review of lessons will help further inform Teaching Statistics for Social Justice.

REFERENCES

Bargagliotti, A., Johnson, S., Franklin, C., Perez, L., Arnold, P.,, Spangler, D.A. & Gould, R. (2020). *Pre-K-12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II)*. American Statistical Association. https://www.amstat.org/asa/files/pdfs/GAISE/GAISEIIPreK-12_Full.pdf

Berry, R. Q., Conway, B.M., Lawler, B.R., & Staley, J. W. (2020). *High school mathematics lessons to explore, understand, and respond to social injustice*. Corwin Press.

Buell, C. A., & Shulman, B. (2019). An introduction to mathematics for social justice. *PRIMUS*, *29*(3-4), 205-209. https://doi.org/10.1080/10511970.2018.1530707

Dogucu, M., Johnson, A. A., & Ott, M. (2023). Framework for accessible and inclusive teaching materials for statistics and data science courses. *Journal of Statistics and Data Science Education*, *31*(2), 144-150. https://doi.org/10.1080/26939169.2023.2165988

Gutstein, E. (2006). *Reading and writing the world with mathematics: Toward a pedagogy for social justice.* Routledge.

Hill, E., Tiefenthäler, A., Triebert, C., Jordan, D., Willis, H., & Stein, R. (2020, May 31). How George Floyd was Killed in Police Custody. *The New York Times.* https://www.nytimes.com/2020/05/31/us/george-floyd-investigation.html

Ladson-Billings, G. (1995). But that's just good teaching! The case for culturally relevant pedagogy. *Theory Into Practice, 34*(3), 159-165. https://doi.org/10.1080/00405849509543675

Lee, H.S., Mojica, G.M., Thrasher, E., & Baumgartner, P. (2022). Investigating data like a data scientist: Key practices and processes. *Statistics Education Research Journal,* 21(2), 1 – 23. https://doi.org/10.52041/serj.v21i2.41

Lesser, L. (2007). Critical values and transforming data: Teaching statistics with social justice. *Journal of Statistics Education*, *15*(1), 1-22. https://doi.org/10.1080/10691898.2007.11889454

Rubel, L. (2017). Equity-directed instructional practices. *Journal of Urban Mathematics Education*, *10*(2), 66-105. https://doi.org/10.21423/jume-v10i2a324

Stockton, G. (2021). 21 states that still don't require de-escalation training for police. *APM Reports*. https://www.apmreports.org/story/2021/06/24/21-states-still-dont-require-deescalation-training-for-police

Weiland, T. & Williams, I. (2023). Culturally relevant data in teaching statistics and data science courses. *Journal of Statistics and Data Science Education*, *32*(3), 1-16. https://doi.org/10.1080/26939169.2023.2249969