Intersectional Data Analysis of Gun Violence in Boston: Teaching Data Activism to Mitigate Systemic Oppression

Zeynep Yalcin*
zy1@wellesley.edu
Wellesley College
Wellesley, MA, USA

Raechel Walker
raechelw@media.mit.edu
Massachusetts Institute of Technology Media Lab
Cambridge, MA, USA

Cynthia Breazeal
breazeal@mit.edu
Massachusetts Institute of Technology Media Lab
Cambridge, MA, USA

ABSTRACT

Biased data is increasingly becoming a part of algorithms that determine people's livelihood, such as predictive policing or recidivism predictors. One of the most effective ways of understanding how such algorithms work starts by examining the systems of oppression that lead to biased data. The lesson, "Intersectional Data Analysis: Examining Shootings in Boston", begins with examining the connection between racism, housing, and policing. Then, students use their data science skills to analyze how gun violence disproportionately harms African Americans. As a result, students examine the direct effects of historical bias embedded in data. The results show the student's ability to use data science and their knowledge of gun violence being a racial justice issue to create unbiased datasets, which may lead to fair algorithms.

CCS CONCEPTS

• **Social and professional topics → K-12 education**;

KEYWORDS

data science; historical bias; intersectional analysis; data activism; critical pedagogy; secondary education

ACM Reference Format:
Zeynep Yalcin, Raechel Walker, and Cynthia Breazeal. 2022. Intersectional Data Analysis of Gun Violence in Boston: Teaching Data Activism to Mitigate Systemic Oppression. In *Proceedings of ACM Student Research Competition (SIGCSE TS '23).* ACM, New York, NY, USA, 3 pages. https://doi.org/XXXXXXX.XXXXXXX

1 PROBLEM AND MOTIVATION

Predictive policing algorithms, such as the Los Angeles Strategic Extraction and Restoration (LASER) algorithm, are an example of algorithms that claim to be able to predict who will commit a crime and where the crime will happen [8]. However, the LASER algorithm trains on data that is biased, frequently resulting in Black and Latinx populations being unfairly targeted [3]. Algorithms make predictions based on patterns in the data that the algorithm is trained on. If the input data is prejudiced and racist, then the algorithm will make biased and racist decisions. The LASER program is a prime example of how the intersections of technology and race unfairly affect Black and Latinx communities [3]. The results show the student's ability to use data science and their knowledge of gun violence being a racial justice issue to create unbiased datasets, which may lead to fair algorithms. A way of teaching minoritized students to mitigate bias in AI, starts by examining the systems of oppression that lead to biased data through a liberatory computing framework. Liberatory computing is a framework that ensures African American students acquire a sound racial identity, critical consciousness, collective obligation, liberation centered academic/achievement identity, and activism skills [13]. Within the data activism curriculum, I created a lesson that taught students how to mitigate racial bias in a data set about Boston shootings. Data activism involves using data science to challenge power inequalities, such as racism [9].

Students completed the intersectional data analysis about gun violence in Boston lesson in three days. This was one of a series of lessons that was situated within a six week data activism curriculum. I analyzed how high school students used their understanding of the connection between racism, housing, and policing to create data visualizations that highlight racial disparities in gun violence.

2 BACKGROUND AND RELATED WORK

Some computing curricula build student's critical consciousness about how AI may amplify oppression of minoritized groups [5]. However, computing curricula do not ensure that students have the data science and critical thinking skills to analyze data sets that can cause harm [2, 4]. Although there are computing programs that give students projects that are culturally relevant [1], they do not connect historical facts to trends in data analysis. The historical and sociological context of a problem that affects their city, provided students with the skills to potentially audit and dismantle oppressive algorithmic systems [7]. This lesson plan not only ensures the education of diverse groups of students, but also that students apply their technical skills to understanding historically biased data in the context of algorithms.

3 APPROACH AND UNIQUENESS

3.1 Participation

The lesson plan, "Intersectional Data Analysis: Examining Shootings in Boston", was a part of the data activism curriculum which was piloted to ten high school students. Nine of the students were African American and one student was Asian American. Three of the students identified as male, seven of the students identified as

SIGCSE TS '23, March 15–18, 2023, Toronto, Canada

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

female. In the six week duration of the program, four classes were held per-week for four hours. This lesson lasted three classes, which is twelve hours of lessons and coding. Additionally, the students received an hourly wage to participate in this research study.

3.2 Activities

The lesson plan consists of three separate parts: (1) examining the connection between racism, housing, and policing, (2) applying sociological theories to historical facts, and (3) conducting an intersectional data analysis about Boston shootings. This three tier approach gives the students an unique opportunity to understand the context and meaning of the data while examining it in intersectional ways. By deconstructing the history and sociological meaning of the data, the students are able to understand how this applies to all data sets and how an intersectional analysis can be applied to any algorithm.

3.2.1 Examining the Connection Between Racism, Housing, and Policing. First, students learn about the history of redlining and policing through educational videos and active discussion. Redlining is the systematic implementation of discriminatory lending practices that denied mortgages in neighborhoods of color [10]. Redlined areas are more likely to report gun-violence [12]. Students then learn about the history of policing in America, and the role of police as 'fine-collectors' from their low income constituents that are most likely to live in redlined areas [11]. After students learned about the the systemic and institutional reasons behind housing segregation and the location of police forces and they are able to understand why gun violence data in Boston is located in predominantly Black and Brown communities.

3.2.2 Applying Sociological Theories To Historical Facts. Next, students engage in conversations about sociological concepts that impact policing and housing by learning about Critical Race Theory (CRT) and the construction of American historiography [6] to emphasize the direct connection between America's history and the racist problems seen today. The lesson plan emphasizes how segregationist laws have led to under-funded communities of color that are over-policed and therefore over-represented in the data set of shootings in Boston.

3.2.3 Intersectional Analysis About Boston Shootings. Based on this knowledge, students are able to recognize the racial disparities in the dataset. This lesson puts students in a unique position to understand the historical bias embedded in the data. Students learned how to create table functions and comparison bar graphs with value counts, percentages, and stacked bar graphs to examine multiple variables at once. Table functions were used to produce tables with rows and columns of a specific aggregation, such as race, gender, or location. For example, students were able to compare two variables at once by creating bar graphs that examined the distribution of victims of gun violence by race and district.

4 RESULTS AND CONTRIBUTIONS

All students articulated how different forms of oppression led to racial disparities in gun violence in the data set they examined. Also, students completed the guided short answer responses and data visualization tasks. In order to evaluate the effectiveness of this

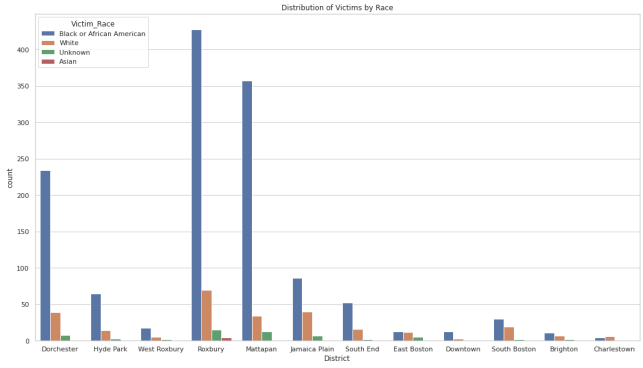

Figure 1: A bar chart that a student produced showcasing locations in Boston and race.

activity, I completed a thematic analysis of how students applied the skills they learned from the intersectional analysis on Boston shootings data set lesson to their final projects.

Students demonstrated their understanding of the historical bias embedded in the data through classroom discussions and written responses about predictive policing. For example, in response to videos about predictive policing, the history of policing, and redlining, nine students discussed how different forms of racism led to bias in the Boston shootings data set.

Next, every student conducted an intersectional examination of the data in terms of race, gender, and location. Students typed Python functions that produced tables and bar graphs to highlight the bias in the Boston shootings data set. Specifically, seven of the students created two bar graphs that showed the number of people from each race that suffered a fatal and non fatal shooting. One student was able to connect their historical knowledge with the data visualizations they created by stating, "One of the main reasons why my data visualizations depict African Americans being the most negatively affected by gun violence is because of the over-policing of communities of color."

The data activism final project gave students the freedom to choose a project of interest. Since students chose to further examine housing as well as gun violence in their final projects for the data activism curriculum, the lesson engaged student's unique interests. For the final project, one student decided to further study gun violence, while two other students completed projects that focused on housing accessibility. Moreover, two students utilized the table functions and the comparison bar graphs, which was a part of this lesson plan. This demonstrates that students were able to apply their data science skills to their own projects.

The discussion posts and data visualizations students created showcase the effectiveness of this lesson plan, as students were able to make a clear connection between housing, racism, and policing within the data set they examined. Using this knowledge, students are able to mitigate bias in data sets and potentially apply this skill set to other areas such as algorithm analysis. Additionally, teachers can use this lesson plan to teach minoritized students data science skills for social justice.

REFERENCES

[1] L. Alvarez, I. Gransbury, V. Cateté, T. Barnes, Ákos Ledéczi, and S Grover. 2022. A Socially Relevant Focused AI Curriculum Designed for Female High School Students. *Proceedings of the AAAI Conference on Artificial Intelligence* 36, 11 (2022), 12698–12705. https://doi.org/10.1609/aaai.v36i11.21546

[2] M. A. Beasley and M. J. Fischer. 2012. Why they leave: The impact of stereotype threat on the attrition of women and minorities from science, math and engineering majors. *Social Psychology of Education* 15, 4 (2012), 427–448.

[3] Arrigo B Browning, M. 2021. Stop and Risk: Policing, Data, and the Digital Age of Discrimination. *Am J Crim Just* (2021), 298–316.

[4] Veronica Catete, Amy Isvik, and Marnie Hill. 2022. A Framework for Socially-Relevant Service-Learning Internship Experiences for High School Students. *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education* 1 (2022), 815–821. https://doi.org/10.1145/3478431.3499355

[5] Kapor Center. 2021. Culturally responsive-sustaining cs education: A Framework. (2021).

[6] Hannah-Jones Nikole et al. 2021. The 1619 Project : A New Origin Story. (2021).

[7] Jayne Everson, F. Megumi Kivuva, and Amy J. Ko. 2022. "A Key to Reducing Inequities in Like, AI, is by Reducing Inequities Everywhere First": Emerging Critical Consciousness in a Co-Constructed Secondary CS Classroom. *In Proceedings of the 53rd ACM Technical Symposium on Computer Science Education* 1 (2022), 209–215. https://doi.org/10.1145/3478431.3499395

[8] Andrew Guthriel Ferguson. 2016. Policing predictive policing. *Wash. UL Rev.* (2016).

[9] M. Gutierrez. 2018. Data activism and social change. *London: Palgrave Macmillan* (2018).

[10] Emily E et al Lynch. 2021. The legacy of structural racism: Associations between historic redlining, current mortgage lending, and health. *SSM - population health* 14, 100793 (2021). https://doi.org/doi:10.1016/j.ssmph.2021.100793

[11] Shaw Theodore M and United States. The Ferguson Report. 2015. The Ferguson Report : Department of Justice Investigation of the Ferguson Police Department. (2015).

[12] Tracey Dechert Lisa Allee Kelly M. Kenzik Michael Poulson, Miriam Y Neufeld. 2021. Historic redlining, structural racism, and firearm violence: A structural equation modeling approach. *The Lancet Regional Health - Americas* 3 (2021), 12698–12705. https://doi.org/10.1016/j.lana.2021.100052

[13] R. Walker, E. Sherif, and C. Breazeal. 2022. Liberatory computing education for African American students. *2022 Conference on Research in Equitable and Sustained Participation in Engineering, Computing, and Technology (RESPECT)* (2022), 85–89.