Louie, J., Stiles, J., Fagan, E., Chance, B., & Roy, S. (2022). Building toward critical data literacy with investigations of income inequality. *Educational Technology & Society, 25* (4), 142-163.

Building toward Critical Data Literacy with Investigations of Income Inequality

Josephine Louie[1*], Jennifer Stiles[1], Emily Fagan[1], Beth Chance[2] and Soma Roy[2]
[1]Education Development Center, USA // [2]California Polytechnic State University, USA // jlouie@edc.org // jstiles@edc.org // efagan@edc.org // bchance@calpoly.edu // soroy@calpoly.edu
[*]Corresponding author

ABSTRACT: To promote understanding of and interest in working with data among diverse student populations, we developed and studied a high school mathematics curriculum module that examines income inequality in the United States. Designed as a multi-week set of applied data investigations, the module supports student analyses of income inequality using U.S. Census Bureau microdata and the online data analysis tool the Common Online Data Analysis Platform (CODAP). Pre- and post-module data show that use of this module was associated with statistically significant growth in students' understanding of fundamental data concepts and individual interests in statistics and data analysis, with small to moderate effect sizes. Student survey responses and interview data from students and teachers suggest that the topic of income inequality, features within CODAP, the use of person-level data, and opportunities to engage in multivariable thinking helped to support critical data literacy and its foundations among participating students. We describe our definitions of *data literacy* and *critical data literacy* and discuss curriculum strategies to develop them.

Keywords: K–12 education, Intercultural competence, 21st century skills, Inquiry learning

1. Introduction

In a world steeped in data, many voices have called on schools to strengthen students' data literacy. Researchers argue that a wide range of data literacy skills are needed to meet rapidly growing workforce demands (Henke et al., 2016; Manyika et al., 2011) and to participate in modern civic life (Bargagliotti et al., 2020; Wolff et al., 2016). As ever more aspects of our lives become captured in data, additional voices have called for greater critical data literacy. These calls emanate from concerns about the potential harms to individuals and society when powerful groups collect vast amounts of our personal data and shape its narratives (e.g., D'Ignazio & Klein, 2020; Raffaghelli, 2020). To equip all individuals with the skills required to thrive in today's world, K–12 education needs scalable research-based strategies to strengthen and support diverse student populations in developing critical data literacy. This article describes a curriculum intervention developed and implemented in U.S. high school non-Advanced Placement (AP) mathematics classes to study one such strategy.

2. Conceptual approach

2.1. Defining data literacy and critical data literacy

For decades, statistics educators have worked to define and advance skills to make sense of data (Donoho, 2017; Rubin, 2020). Like Gould (2017) and Weiland (2017), we consider data literacy to rest on a foundation of statistical literacy. This type of literacy involves the abilities to interpret, assess, and communicate understandings of data from our everyday lives (Gal, 2002). It also includes basic fluency with the process of data investigation. As articulated by leading statistics educators, this process involves four iterative steps: (1) formulating questions that can be answered with data, (2) assembling data to address one's questions, (3) using statistical and other tools to analyze the data, and (4) interpreting results to address one's original questions (Bargagliotti et al., 2020). We also associate data literacy with a disposition to interrogate each step within the data investigation process (Bargagliotti et al., 2020). Data literate individuals carry a habit of mind (Cuoco et al., 1996; Finzer, 2013) that routinely asks about the provenance or origins of the data; who defined, measured, and collected the data; what tools or methods were used to analyze the data; and the degree to which data interpretations or conclusions are valid. Furthermore, we consider data literacy to include an ability to display multivariable thinking. This type of thinking recognizes that the relationship between two variables may not be as it initially seems given the possible effects of other interacting or confounding variables. Because real-world data and phenomena are typically multivariable, all learners should develop a capacity for multivariable thinking (Bargagliotti et al., 2020; Engel, 2016; Ridgway, 2015).

Building on this foundation, we consider critical data literacy to incorporate the skills of data literacy along with ethical and sociopolitical perspectives toward data. Digital technologies today allow actors to collect massive amounts of personal data to predict our behaviors and to target us with messages that can shape our social perceptions and actions (D'Ignazio & Klein, 2020; Noble, 2018; O'Neil, 2016; Pangrazio & Selwyn, 2019). Critical data literacy requires ethical consideration of whether and how to collect and use data from others and to reduce potential harm to individual privacy and autonomy (Baumer et al., 2022; National Academies of Sciences, Engineering, and Medicine, 2018). It also involves an awareness of how data can be used to perpetuate or aggravate unequal power relations in society, such as among groups that differ by race, sex, or economic class (e.g., Bhargava et al., 2015; D'Ignazio & Klein, 2020; Noble, 2018; Philip et al., 2016). People with critical data literacy not only scrutinize data for biases and possible misdirection, they also use data to understand and challenge forms of social oppression and to work toward more equitable social outcomes (Bhargava et al., 2015; D'Ignazio & Klein, 2020; Louie, 2022; Pangrazio & Selwyn, 2019).

2.2. Building toward critical data literacy

Prior research points to the types of classroom environments that are most likely to help students build foundational data literacy skills. In such environments, students explore motivating questions with real data, engage actively in the four steps of the data investigation process, use digital tools to help visualize and analyze data, and participate in rich classroom discussions to make sense of data (Bargagliotti et al., 2020; Berry et al., 2020; Chance et al., 2007; Garfield & Ben-Zvi, 2008). To help students recognize how data sets are constructed and their limitations, scholars suggest that students engage in primary data collection or examine secondary data collected by others in familiar contexts (e.g., Enyedy & Mukhopadhyay, 2007; Lee & Wilkerson, 2018; Rubel et al., 2016; Van Wart et al., 2020).

Scholarship on teaching mathematics for social justice, or TMSJ (Berry et al., 2020), offers ideas for fostering critical data literacy. TMSJ seeks to build students' capacities to recognize and counter social injustices with mathematics and data. Gutstein (2003, 2006), drawing on the ideas of Freire (1970), describes teaching for social justice as having three primary goals: (1) developing students' social and political consciousness, (2) building agency to affect social change, and (3) fostering positive social and cultural identities. Bringing this approach to mathematics education, he describes TMSJ as helping students to "read" and "write" the world with mathematics—that is, to discern and devise solutions to pressing social problems with quantitative data and reasoning (Gutstein, 2003; Gutstein, 2006). TMSJ invites students to investigate complex, real-life issues using data that reveals the nature and scope of unequal socioeconomic outcomes among groups that differ by race, class, sex, or other social characteristics and to raise questions about the factors that contribute to these outcomes. Scholars of TMSJ have organized instruction around projects examining issues such as regressive utility rates that cause low-income families to pay higher electricity rates than wealthy energy users (Frankenstein, 2009; Frankenstein, 2013); racial disparities in police stops and access to housing markets (Gutstein, 2003; Gutstein, 2013); and disparate racial, class, and spatial patterns in public lottery participation (Rubel et al., 2016).

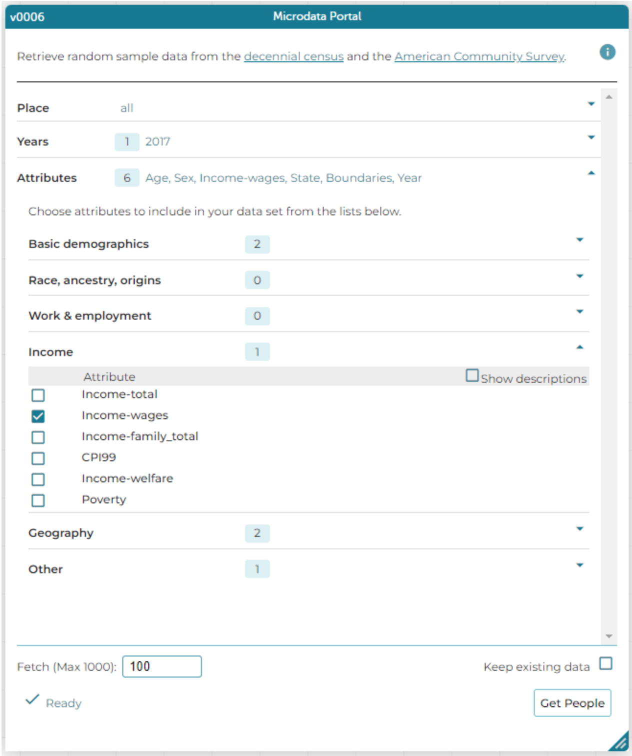
Another key feature of TMSJ is the use of culturally relevant pedagogy to advance equitable learning opportunities among diverse student populations (Berry et al., 2020). Culturally relevant pedagogy values and incorporates students' cultural backgrounds and lived experiences in classroom instruction, both to facilitate students' academic learning and to strengthen their identities (Gay, 2002; Gay, 2013; Ladson-Billings, 1995). It also emphasizes high expectations and standards-based instruction for students from historically marginalized groups to help them achieve academically and to access social and economic power (Berry et al., 2020; Delpit, 1988; Ladson-Billings, 1995). Research has found that culturally relevant pedagogy can lead to improved academic engagement and learning outcomes among diverse student populations, particularly those whom educators label as "at risk" (Aronson & Laughter, 2016; Dee & Penner, 2017). Because the goals of critical data literacy, TMSJ, and culturally relevant pedagogy are so highly aligned, employing strategies of TMSJ and culturally relevant pedagogy may be an effective way to promote critical data literacy among diverse populations of students.

3. Data literacy modules focused on social justice issues

Drawing on these ideas, we developed a set of high school curriculum modules that consist of data investigations examining social justice issues. A primary goal of our work has been to promote understanding of and interests in fundamental statistical thinking practices (i.e., data literacy) among diverse student populations with high proportions of learners from historically marginalized groups. A second goal has been to explore the extent to which module use may advance aspects of students' critical data literacy, focusing on two outcomes that Gutstein identified as central to social justice learning: acquiring social and political consciousness and gaining a sense of social agency. Designed as a series of lessons that teachers facilitate over 15 one-hour class periods, the modules guide students in exploring social and economic questions of direct relevance to themselves, their families, and their communities. Students address these questions by analyzing large-scale person-level microdata from the American Community Survey (ACS) and the U.S. decennial census, using the free and browser-based Common Online Data Analysis Platform (CODAP).

In one module, *Investigating Income Inequality in the U.S.* (the "Income Inequality module"; full module materials can be found at https://go.edc.org/ussdata), students work through seven lessons and a final team data investigation, discussing different forms of income inequality, its scope, and its causes. They begin in Lesson 1 by grappling with the following question: What is income inequality, and when might it be a concern for society? With teacher facilitation, they discuss how they think total income in the United States is distributed among lower- and higher-income groups, how this distribution has changed over time, the questions these topics raise for them, and whether they can answer their questions with data. In Lesson 2, students consider the provenance, strengths, and limitations of data from the U.S. decennial census and the ACS. They learn about the U.S. federal government's stated purposes and uses of these data sources, as well as the types of data that are collected (e.g., multiple measures of income but not wealth) by reviewing and attempting to answer ACS questionnaire items themselves. Students discuss the reasons why people may skip or have difficulty answering specific questions and the implications of missing or inaccurate responses.

Figure 1. The U.S. decennial census and ACS Microdata Portal in CODAP



Note. The Microdata Portal within CODAP provides access to person-level data from the U.S. decennial census for 11 decades (from 1860 to 2010) and from the ACS for 2017. Users can draw random samples of individuals from the United States as a whole as well as from each of the 50 states. They can draw samples of up to 1,000 people at a time, with the capacity to add cases to create larger samples. The data source is IPUMS-USA, University of Minnesota, www.ipums.org.

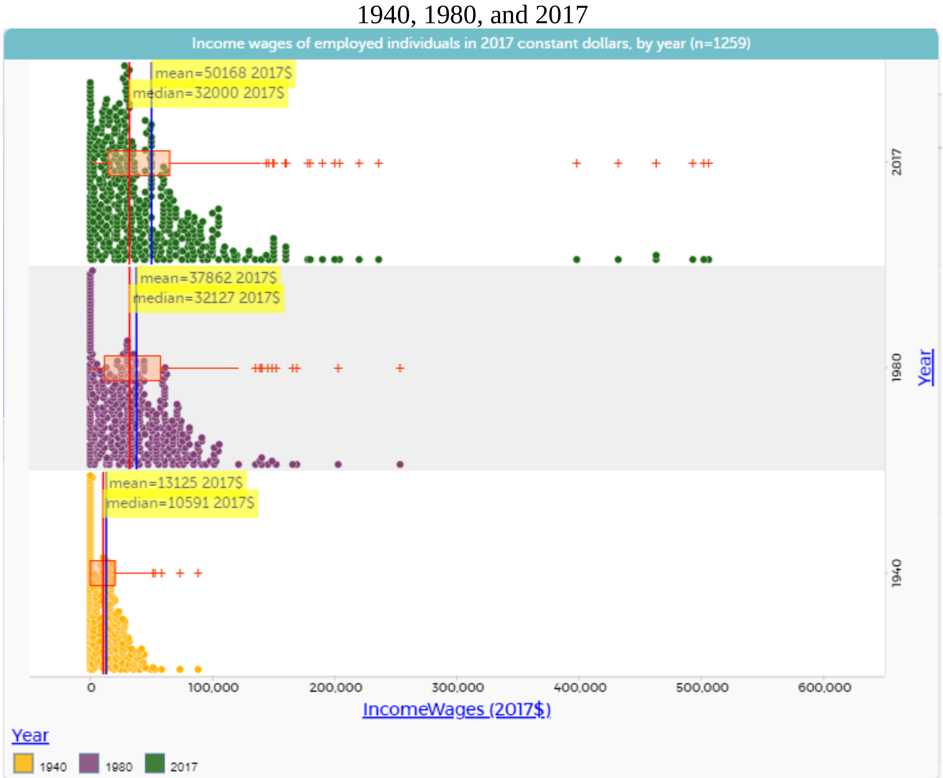
Subsequent lessons ask students to investigate questions about individual and aggregate incomes of U.S. wage earners, using data drawn from a U.S. census and ACS Microdata Portal that we developed as a permanent plug-in tool for CODAP (Figure 1). We provided students with access to person-level microdata with over three dozen individual attributes (or variables) for exploration, under the hypothesis that students would be better able to relate to the people in the data if they could see person-level records in their data sets. A person-level data set would also allow students the flexibility to examine a wide range of relationships among multiple variables of their own choosing, such as the relationship between income and education when controlling for race. Prior studies involving map-based investigations of social issues with large-scale social data (e.g., Kahn, 2020; Radinsky et al., 2014; Rubel et al., 2016) have allowed students to examine social and economic outcomes based on neighborhoods or higher units of geography, but not based on individuals as the unit of analysis.

In Lessons 3–5, students investigate how income equality in the United States has changed over time. They begin by creating data tables that contain random samples of individuals from the U.S. population in 2017, with information about each person's reported annual wages and other characteristics (such as age, sex, race/ethnicity, and occupation). They then create dot plots of the wage distribution by dragging and dropping the wage attribute from their data table to a graphing window and watching data cases populate the graph dynamically. Students can hover over or click on individual dots to see an individual's specific attributes in a pop-up window or in the linked data table (Figure 2).

Figure 2. A CODAP data table and dot plot of annual income from wages for a random sample of employed U.S. individuals in 2017

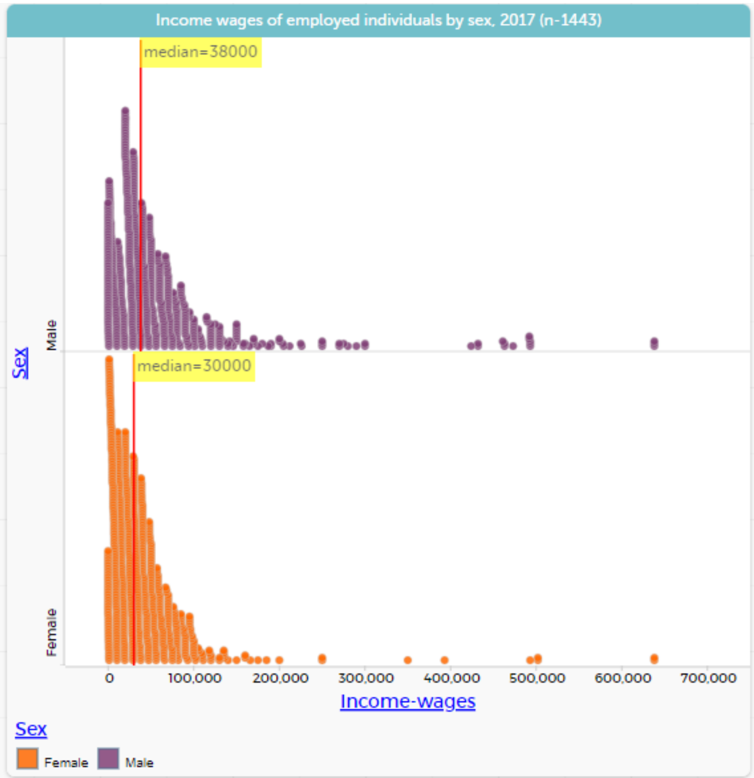
Note. CODAP allows users to hover over or click on individual cases in the graph to highlight case information within the graph itself and in the linked data table. A toolbar provides functions that allow users to display measures such as the mean and median of the distribution.

Figure 3. CODAP dot plots of annual income from wages for random samples of employed U.S. individuals in 1940, 1980, and 2017

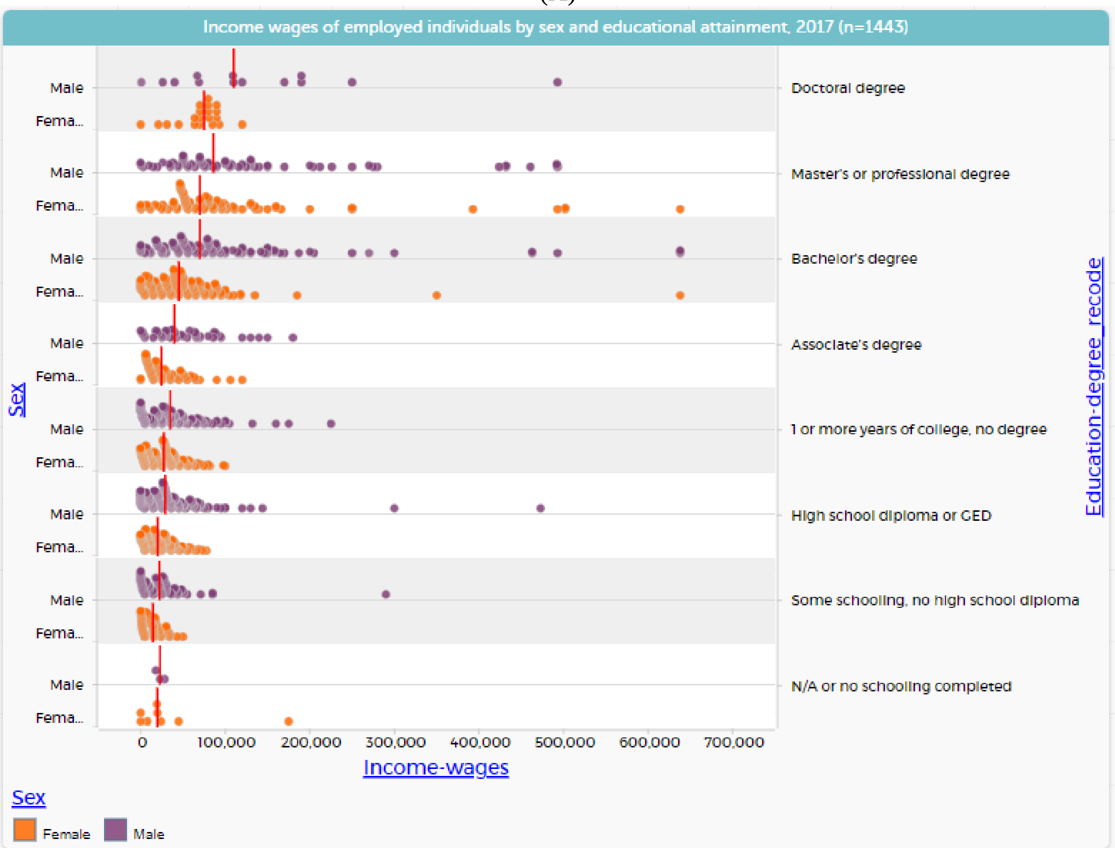


Note. Students analyze annual income from wages in each decade as measured in constant 2017 dollars.

Figure 4. Graphs of wages by sex (A), then by educational attainment (B), for a random sample of employed U.S. individuals in 2017



(A)



(B)

The dot plots in CODAP allow students to see the shape of the entire wage distribution, with the bulk of individuals reporting wages at the low end of the income spectrum and a small number of individuals reporting wages at the upper end. The strong skew of the wage distribution and the CODAP tools help students build a conceptual understanding of the difference between the mean and the median—two concepts that students traditionally learn to compute without a deeper grasp of their meaning or when to use them (Konold & Higgins, 2003). In CODAP, students can turn on visual indicators of where the mean and median fall within the distribution; they can also hold and drag an outlying high-wage case further to the right to see the mean of the wage distribution increase while the median does not change. Students discuss whether the mean or median value is a better representation of the sample's "typical" wage, given the skew in the data. They also discuss how the distance between the two measures can indicate the presence of a small number of extreme cases "pulling" the mean away from the median, thus providing one way to measure the degree of income inequality in the sample.

Students use these ideas to examine how wages have diverged between lower- and higher-income earners since 1940 (Figure 3). They examine the changing shape, spread, whisker length, and number of statistical outliers associated with box plots superimposed on the data. They quantify growing gaps among higher- and lower-wage earners by calculating the ratio of annual wages of individuals at the top and bottom quartiles each decade. Furthermore, in a class activity, they build an intuition for the margin of error for their sample estimates by using the Microdata Portal to select multiple samples of varying sizes (starting from n = 100 and growing by increments until they reach 1,000 cases) and observing the level of variation across their estimates.

In Lesson 6, students turn to examine income inequality between male and female employees in the United States. Using data drawn from the Microdata Portal, students make graphs (e.g., Figure 4, graph A) to compare measures of center and variability for the wage distributions of male and female employees. In Lesson 7, after speculating about possible causes behind differences in wages by sex, they examine whether the gap in median wages for the two groups can be explained by education. They explore this question with a graph "splitting" function, which allows students to visualize how the relationship between two variables may change after adjusting for a third variable (Figure 4, graph B). Once they discover that the wage gap persists after controlling for education, students discuss if other individual attributes may explain the gap or if sex discrimination may be at play. In a final team data investigation, students collaborate in groups to choose and investigate another variable that may explain the gap in wages by sex. They enact, document, and present findings from the entire data investigation process. Through these activities, they deepen their abilities to describe and compare quantitative distributions using measures of center and variability—key practices in the Common Core State Standards for Mathematics (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) for secondary grade levels. They also build abilities to reason with multivariable data, a skill that should be but is not currently emphasized in K–12 education (e.g., Engel, 2016).

4. Module iterations and study research questions

With input from both high school mathematics and social studies teachers, we developed an "alpha" iteration of the Income Inequality module, which five high school mathematics teachers implemented with their students during the 2018-2019 academic year. We then revised the materials and tested a "beta" iteration of the module with seven high school mathematics teachers and their students during the 2019-2020 academic year. Testing of this module ended just at the onset of the COVID-19 pandemic.

The questions that drive this study derive from the project goals described earlier. We also aim to expand on existing research. Prior studies of mathematics or data-focused interventions with a social justice or critical literacy orientation have involved smaller-scale classroom or summer session case studies (e.g., Enyedy & Mukopadhya, 2007; Van Wart et al., 2020; Rubel et al., 2016) and research in which study leaders enacted the interventions themselves (e.g., Gutstein, 2003; Brantlinger, 2013). In contrast, we sought from the start to recruit multiple teachers to test our modules in authentic classroom settings with high proportions of students from historically marginalized groups, where we (as researchers, curriculum developers, and higher education faculty members) were not involved in direct module implementation. We also wanted to examine whether module use is associated with quantitative gains in students' learning and interest outcomes to provide evidence of promise for a curriculum approach that teachers could implement and test at a larger scale. We sought to address the following research questions:

- RQ1. To what extent do students in participating classrooms show increased understanding of and interests in data literacy after completing the Income Inequality module?
- RQ2. What aspects of the module appear to promote participating students' understanding of and interests in data literacy?
- RQ3. In what ways do participating students demonstrate critical data literacy or social justice learning outcomes after completing the Income Inequality module?

5. Methods

5.1. Study team, sample, and teacher preparation

The leaders of the project are an educational researcher (Asian American, female) with a background in social policy research and two university-level statistics educators (one White American, female; one South Asian, female). The other team members are a senior mathematics curriculum writer and a research associate (each White American, female). During the early years of the project, a senior scientist/technology developer (White American, male) was heavily involved in development of CODAP tools for the modules. All members of the project team participated in different aspects of module development and classroom research.

We recruited seven high school mathematics teachers from six cities in a northeast U.S. state to implement the Income Inequality module in non-Advanced Placement (AP) mathematics classes in fall 2019 and early 2020. We targeted public high schools with high proportions of students from Black, Latinx, and low-income families, sharing a letter seeking teachers "to help test and inform the development of a set of applied data investigations in high school non-AP mathematics classes" to support students in "investigating U.S. socioeconomic trends and questions related to social justice, using large-scale U.S. population and economic data sets from the U.S. Census Bureau and other government agencies." Unlike other studies (e.g., Kokka, 2020), teachers did not have to profess a social justice orientation in their instruction to participate. In screening interviews, teachers whom we accepted in our study expressed a strong interest in our curriculum approach, particularly its project-based orientation, use of real data, and investigation of real-life issues.

Table 1. Characteristics of study participants

A. Teacher, classroom, and student characteristics

Teachers[1]			Classrooms[2]		Students[3]				
Name	Highest degree	Years teaching math/ statistics	Course sections (#)	Students per section (#)	Age range	Black (%)	Latinx (%)	White (%)	Work to support self or family (%)
Eve	MA/CAGS	9/2	5	20	17–19	17	57	18	70
Rachel	MA/CAGS	22/10	1	30	16–19	10	50	25	75
Sasha	BA	3/2	2	10–15	17–19	30	30	20	60–70
Julie	BS	6/6	2	24	17–19	20	35	40	70
Anne	MA	16/14	1	25	16–18	20	30	40	80
Will	BA	9/5	2	25–30	16–19	20	40	25	50
Bella	MA	15/3	1	10–15	16–23	70	20	10	70

B. School characteristics[4]

School name	Teacher participant(s)	Enrollment (#)	Black (%)	Latinx (%)	White (%)	First Language not English (%)	Low income (%)
Elm	Eve	2,000	20	53	21	60	49
Rose	Rachel	2,020	4	58	31	67	45
Spruce	Sasha	1,230	13	47	32	58	45
Wood	Julie, Anne	1,650	10	45	39	51	38
Beech	Will	1,110	19	42	31	48	61
Ash	Bella	410	48	43	6	35	65

Note. [1]Data are from teacher self-report for the 2019-2020 school year. Teacher names are pseudonyms. All teachers self-identify as White and female except for Will (male) and Sasha (non-binary). All majored in mathematics as undergraduates except for Julie (engineering) and Bella (history). [2]Data are from teachers' reporting before or at the start of the 2019-2020 school year. [3]Data are from teachers' retrospective reporting during the 2021-2022 school year. [4]Data are from the state department of education for the 2019-2020 academic year. School names are pseudonyms. Enrollment figures are rounded to the nearest 10. All schools are public, include grades 9–12, and are in separate cities in the Northeast.

Table 1 provides background information on the seven participating teachers, as well as on their classrooms, students, and school contexts. All teachers taught grade 12 non-AP courses in statistics or data analysis and had the flexibility to incorporate multi-week curriculum modules into their course schedules. None of the teachers had prior experience working with large-scale data or with data analysis tools other than spreadsheet programs such as Microsoft Excel. Between 55% and 91% of students in the participating schools were Black or Latinx, and teachers estimated that 50%–80% of their students held jobs to help support themselves and their families. In each school, the percentages of students classified as low-income exceeded statewide percentages. The Ash school has a particularly high-needs population: It is an alternative public school that serves students who have had long-standing attendance issues or have not been successful in other schools.

We provided all participating teachers with an in-person seven-hour workshop in summer 2019 and a follow-up two-hour after-school session in the early fall to prepare them to implement the module with students. During these sessions, we provided a module overview, supported teachers in learning CODAP, and facilitated select lessons to allow teachers to experience the materials as learners. Three of the teachers (Rachel, Sasha, and Julie) had implemented the module's alpha iteration and provided data to inform revisions for the beta iteration. The remaining four teachers (Eve, Anne, Will, and Bella) were new to the project. During module implementation, project team members were available by phone and email to help teachers troubleshoot technology and other lesson questions.

5.2. Data sources and measures

To address RQ1, we invited all participating students to complete an online data literacy assessment and interest/affect survey immediately before starting and after completing the module. The assessment contained 19 multiple-choice items drawn primarily from item pools developed and validated by Jacobbe et al. (2014) and Garfield et al. (2006). We selected items to measure understanding in five domains that were aligned with the module's data literacy learning objectives: sampling and data collection, data representation, measures of center, measures of variability, and multivariable thinking (see Appendix 1). The interest/affect survey contained 23 seven-point Likert-type items adapted from validated scales by Linnenbrink-Garcia et al. (2010) and Sproesser et al. (2016). We assembled scales measuring students' temporary situational interest, longer-term individual interest, self-concept, and perceptions of the value in statistics and data analysis (see Appendix 2, Table 7 for sample items and Louie et al. (2021) for more on scale constructs and development). Cronbach's alpha for each scale ranged from 0.73 to 0.92, suggesting acceptable to very good reliability (DeVellis, 2012).

To address RQ2 and RQ3, we drew upon several data sources. First, we collected student responses to an open-ended prompt in the post survey, which asked: *Tell us how much you liked the Income Inequality data lessons compared to your other statistics lessons. Please explain your reasons*. Second, we conducted five focus group interviews with n = 4–6 students per group at four schools (Elm, Rose, Wood, and Beech). The project's lead researcher conducted the semi-structured interviews asking questions such as the following: *In what ways were the Income Inequality data lessons different from your other work in this class? What aspects of the lessons were [most/least] interesting, and why? What do you think you learned from these lessons?* Each interview lasted 40–60 minutes and was audio recorded and transcribed for analysis. Third, we conducted 1- to 2-hour interviews with five teachers soon after they completed the module and again in early 2022 as an auxiliary source of data about students' responses to the module. All teacher interviews were conducted by videoconference and recorded and transcribed for analysis. Fourth, three project team members conducted classroom observations during module implementation, with five teachers observed twice and two teachers observed once. Observers recorded written field notes, documenting classroom environments and how teachers implemented and how students responded to module activities. Fifth, we collected online implementation logs from teachers for contextual information on how they facilitated each lesson.

5.3. Data analyses

We calculated student scores on the data literacy assessment as the total number of correct responses and on the interest/affect scales as the average rating among scale items. We compared students' pre- and post-module scores using paired *t*-tests. To analyze student responses to the open-ended prompt in the post-survey, two project researchers reviewed all student responses; developed a set of overarching codes under the general categories of "positive," "neutral," and "negative" responses to each module; and created a more detailed set of subcodes. For example, subcodes under the positive category included the focus on a relevant or real-life topic, the use of CODAP, and the ease of navigating the materials. Each researcher independently assigned codes to student responses and arrived through discussion at a consensus set of coded data.

To analyze the student focus group data, two project researchers developed a set of overarching codes identifying when students spoke about something they had learned or had found interesting in the module; subcodes identifying whether their learning or interest was related to data and statistics or the social justice topic; and subcodes about aspects of the module (e.g., CODAP, the pedagogical approach) that supported their learning or interest. A parallel set of codes described challenges to students' learning or interests. The two researchers independently coded common transcript excerpts, met to resolve disagreements and to refine code definitions, and iterated on this process until reaching 91% interrater agreement, after which one researcher coded the remaining data. The other researcher applied the same set of codes to the teacher interview data, focusing on teachers' descriptions of student learning and interest during the module. The project's lead researcher analyzed all coded statements to identify emergent themes by code, which were discussed with the project team and teachers for corroboration or refinement. Data from teacher implementation logs and classroom observations were also reviewed by the team to add context and to help support or qualify findings from student and teacher interview data.

6. Results

6.1. RQ1: Changes in student data literacy and interests

Based on total scores on the data literacy assessment, students' understandings of assessed statistical concepts grew between the start and end of the module, and the change was statistically significant at $p < .0001$ with a moderate effect size ($d = 0.43$; Table 2). Students showed the greatest growth in their understanding of measures of center ($p < .0001$, $d = 0.38$), multivariable thinking ($p < .0001$, $d = 0.35$), and data representation ($p < .0001$, $d = 0.19$). Growth was largest in Eve's classes ($d = 0.89$; Table 3), which also had the largest number of participants. Although we offered each student a \$20 Amazon gift card to complete both pre- and post-module assessments, attrition levels ranged from high to extremely high and varied widely by teacher (from 13% to 79%).

Table 2. Students' pre- and post-module data literacy scores, total and by domain

	Pre mean (*SD*)	Post mean (*SD*)	Post - Pre mean (*SD*)	*p*-value	Effect size (*d*)
Total (19 items)	9.66 (2.67)	10.94 (3.16)	1.28 (2.99)	< .0001	0.43
Sampling & Data Collection (3 items)	1.53 (0.80)	1.57 (0.79)	0.04 (0.92)	.560	0.04
Data Representation (4 items)	2.91 (0.91)	3.13 (0.97)	0.22 (1.15)	.011	0.19
Measures of Center (5 items)	2.15 (1.08)	2.68 (1.28)	0.53 (1.41)	< .0001	0.38
Measures of Variability (4 items)	1.57 (0.91)	1.71 (0.97)	0.13 (1.19)	.144	0.11
Multivariable thinking (3 items)	1.50 (0.86)	1.86 (0.89)	0.36 (1.03)	< .0001	0.35

Note. Items were multiple choice and scored as 1 = correct, 0 = incorrect or missing, with $n = 180$ students who provided both pre- and post-module responses.

Table 3. Students' pre- and post-module data literacy scores, total scores by teacher, with attrition rates

	Pre mean (*SD*)	Post mean (*SD*)	Post - Pre mean (*SD*)	*p*-value	Effect size (*d*)	Pre *n*	Post *n*	Attrition (%)
All teachers	9.66 (2.67)	10.94 (3.16)	1.28 (2.99)	< 0.0001	0.43	277	180	35
Eve	9.27 (2.68)	12.05 (3.15)	2.79 (3.13)	< 0.0001	0.89	100	75	25
Rachel	10.14 (1.21)	11.14 (2.41)	1.00 (2.83)	0.386	0.35	32	7	78
Sasha	6.60 (2.41)	6.40 (2.30)	- 0.20 (1.92)	0.828	- 0.10	24	5	79
Julie	9.51 (2.50)	10.43 (2.96)	0.91 (2.01)	0.011	0.46	47	35	26
Anne	9.58 (1.83)	8.92 (2.97)	- 0.67 (2.96)	0.452	- 0.23	20	12	40
Will	10.66 (2.73)	10.37 (2.75)	- 0.29 (2.39)	0.438	- 0.12	47	41	13
Bella	11.00 (3.87)	11.80 (2.59)	0.80 (1.79)	0.374	0.45	7	5	29

Note. Post *n* represents the number of students who provided both pre- and post-module responses.

Table 4. Students' pre- and post-module interest and affect scores related to data literacy, by scale

	n^1	Pre mean (*SD*)	Post mean (*SD*)	Post - Pre mean (SD)	*p*-value	Effect size (*d*)
Situational interest (8 items)	192	4.95 (0.97)	4.98 (1.05)	0.03 (1.01)	.689	0.03
Individual interest (4 items)	194	3.81 (1.20)	4.11 (1.23)	0.31 (1.16)	.000	0.26
Self-concept (5 items)	195	4.71 (1.04)	4.72 (1.05)	0.01 (0.81)	.873	0.01
Perceptions of value (6 items)	189	4.96 (0.94)	5.08 (0.93)	0.12 (0.83)	.052	0.14

Note. A student's scale score is the average rating for each scale item, which could range from 0 to 7. [1]Sample size *n* represents the number of students who provided pre- and post-module responses for each scale.

Results from the interest/affect survey suggest that changes in students' temporary situational interest, self-concept, and perceptions of the value in statistics and data analysis between the start and end of the module were not statistically significant at $p < .05$, whereas growth in students' deeper individual interest in these domains was significant with a small effect size (p = .001, d = 0.25; see Table 4). The overall results primarily reflect strong growth in ratings among Eve's students (p = .000, d = 0.59). Like before, attrition rates and effect sizes for the interest/affect scales varied widely by teacher (Table 5).

Table 5. Students' pre- and post-module scores on the individual interest scale, with attrition rates

	Pre mean (*SD*)		Post mean (*SD*)		Post - Pre mean (*SD*)		*p*-value	Effect size (*d*)	Pre *n*	Post *n*	Attrition (%)
All teachers	3.81	(1.20)	4.11	(1.23)	0.31	(1.16)	.000	0.26	277	194	30
Eve	3.86	(1.20)	4.55	(1.20)	0.69	(1.17)	.000	0.59	100	76	24
Rachel	3.75	(1.04)	3.66	(1.17)	- 0.09	(1.53)	.868	- 0.06	32	8	75
Sasha	3.84	(0.96)	4.22	(0.71)	0.38	(0.73)	.191	0.51	24	8	67
Julie	3.30	(1.07)	3.49	(1.11)	0.18	(0.79)	.137	0.23	47	42	11
Anne	4.46	(1.29)	4.21	(1.07)	- 0.25	(1.29)	.481	- 0.19	20	14	30
Will	3.96	(1.20)	3.98	(1.22)	0.02	(1.23)	.900	0.02	47	41	13
Bella	4.30	(1.63)	4.15	(1.81)	- 0.15	(1.42)	.825	- 0.11	7	5	29

Note. Post *n* represents the number of students who provided both pre- and post-module responses.

Although student outcomes varied by teacher, classroom observations and teacher logs did not uncover systematic differences in how teachers implemented the module. An exception was with Eve, who tended to keep her students working together at the same pace while other teachers let small groups proceed at different paces. Teachers reported that they made few modifications to the materials and skipped few activities. There were two exceptions: Sasha skipped the lesson that examines wage distributions over time, and neither Sasha nor Eve implemented the final team data investigation due to time constraints. All teachers described problems with student attendance, with absenteeism most acute in Sasha's and Bella's classes. Attendance was so irregular in Bella's class that she required six weeks to complete the module; other teachers completed the module in three weeks or less.

6.2. RQ2: Aspects of the module that supported students' understanding of and interests in data literacy

Students' responses to an open-ended prompt in the post-module survey provide insight into aspects of the module that played a salient role in supporting students' understandings of and interests in statistics and data analysis. When asked to state how much they liked the module compared to their other statistics lessons, 86% of respondents shared positive comments about the module (e.g., "I liked the income inequality lessons a lot"), while 5% of the responses were neutral (e.g., "I don't really have an opinion"), and 13% offered comments that were negative ("not that much"; Table 6). In some cases, students responded with both positive and negative comments (e.g., "it was definitely better but still kinda boring to me").

Although many students did not provide reasons for their responses, a plurality of students indicated that they liked the module due to its focus on income inequality. About 32% of students who spoke positively of the module wrote that the topic of income inequality was "real," "relevant," or "relatable," and another 32% said the topic was "interesting." Some students said they liked examining an authentic problem that exists in their own lives, unlike the artificial problems or foreign contexts they encounter in textbooks. In a focus group interview, some of Will's students gave examples of these textbook problems, such as "comparing GPA to height," looking at "Olympic 100-meter sprint times over the past 25 years," and examining "the longevity of a cowboy." Students described these examples as "so boring," and one student asked: "What does this have to do with what we're gonna need?" In line with this last comment, other students suggested that they liked the module because the topic had practical value for their lives. A student from Eve's class wrote: "I actually like [the module] because we are learning about business money, wages… [and] I will like to take a business career some day." Similarly, one of Rachel's students wrote that the module "was the most interesting because it had real data that was useful for the future to see how I could make money." Other students suggested that they found the social justice aspect of the topic compelling. One of Eve's students wrote: "I really liked the income inequality data lessons a lot more than regular statistics lessons because income inequality is something that I find very interesting and I feel passionate about."

Another set of reasons that students gave for liking the Income Inequality module involved using CODAP. Working with this tool was a novel experience for students in each participating classroom. Only two teachers (Sasha and Bella) reported any type of student computer use in their statistics classes, where students occasionally used an online spreadsheet program (Google sheets) to analyze small data sets compiled from class surveys or online sources. The remaining teachers reported that their students used graphing calculators to enter data (typically no more than 40 cases) and to compute statistics by hand. Julie's typical class routine, which other teachers said they shared, was to discuss statistics concepts from a textbook and then to have students engage in "bookwork," that is, solve practice problems on paper with the aid of their calculators.

Table 6. Frequency and percent of coded student responses to the Income Inequality module[1]

Response codes	*n*	% of total[3]	% of subtotal[4]
Total student responses[2]	209	100	--
Positive response to the module	180	86	100
Module was "fun," "engaging," "interesting," or students "liked" or "enjoyed" it.	145		81
The topic of income inequality was "real," "relevant," or "relatable."	58		32
The topic of income or income inequality was interesting.	58		32
Enjoyed using CODAP, the interactive graphs, or the hands-on learning.	30		17
Learned or gained understanding, especially about society.	25		14
Found the lessons clear, well-structured, or easy to follow.	22		12
Not sure or neutral reactions to the module	10	5	--
Negative response to the module	28	13	100
Lessons were difficult or confusing.	6		21
Did not like working with CODAP.	3		11
Lessons or questions were too repetitive.	3		11

Note. [1]In the post-module survey, students responded to the following prompt: "Tell us how much you liked the Income Inequality data lessons compared to your other statistics lessons. Please explain your reasons." [2]Includes all students who provided a response to the open-ended prompt in the post-module survey. [3]Responses do not sum to 100% because some responses included both positive and negative comments. [4]Responses do not sum to 100% of each subtotal because coding categories are not mutually exclusive, and this table displays only those response codes that were applied to at least 10% of all positive or negative comments.

In this context, one of Eve's students wrote that the module "felt like an advanced lesson because we had to use our laptops," and another noted that "it was nice to work on a computer rather than stare at a white board." A classmate commented that using CODAP helped the module feel like "more of an interactive lesson where you are more engaged into the data and [you're] really doing and analyzing the data." Yet another of Eve's students explained that the ability to manipulate and visualize data in CODAP was engaging and supported her learning. She wrote:

> I feel that the income inequality data lessons provided more hands-on learning with the teacher and classmates. Making it very easy for everyone to learn and understand. As well as giving a very good visual representation of the data. I really like learning this way, which is why I've enjoyed this statistics class more than my other statistics classes I've taken before. *(Written survey response from one of Eve's students)*

Teachers indicated that they too thought that CODAP features helped to deepen students' understandings of key data concepts. Will stated in an interview that "CODAP and the hands-on activity…really started showing them what a skewed distribution does. And how mean and median change, why it's sometimes more appropriate to use [the] median." Bella shared that previously, she would simply tell her students that "you really only need 1,000 people to get good data about the whole U.S." With the Microdata Portal, however, students could "experientially see it in terms of the data that each group is generating and how little variation there is of a sample size of 1,000." For her, CODAP tools and activities "really helped build students' conceptual understanding" of central statistical ideas.

In addition, the use of person-level microdata, as well as the ability to see people represented as individual dots in a graph and as discrete records in a linked data table, helped students understand the structure of the data and connect to its human referents. A student in Will's class wrote in the post-module survey that "using the [CODAP] software made it easier to understand where data comes from." Teachers corroborated the value of person-level data as well as accompanying module activities. Both Julie and Rachel had implemented the alpha version of the module, when students had jumped immediately into making graphs with a random sample of 100 people from the ACS. In the beta version, students began by examining the ACS questionnaire and conducting a scavenger hunt to locate individuals with specific demographic characteristics in a CODAP table and a dot plot of 12 cases from the ACS. During a teacher focus group interview, Julie observed:

> [The] most powerful thing about this lesson was connecting it back to each dot on the graph representing a person and [students] being able to identify that—'Wait, that's me… That's my person.' And last year, we didn't have this activity and I realized later on in the modules that they could not connect the fact that each

> dot was a person. And so doing this up front made it a lot easier throughout the entire module. *(Comment from Julie during a teacher focus group interview)*

Rachel concurred, saying, "My students seemed to be pretty engaged…I just kept telling them this is a real person… [like] there is someone who lives in New Jersey... with this salary who is this age… They did like looking to find their person or people who were similar or dissimilar to them."

Another CODAP feature that students liked was the ability to explore relationships among multiple variables at the same time. Students expressed curiosity and motivation to uncover whether another variable such as education could explain the gap in median wages among male and female employees. CODAP allowed them to drag and drop a third variable into a graph of wages by sex and watch the graph split to illustrate new wage distributions by sex and the third variable. This CODAP feature helped students deepen their understanding of how graphs can be organized and structured. Some of the third variables that students chose to explore were a person's occupation, part- or full-time status, race/ethnicity, marital/family status, and age. A student from one of Julie's classes wrote in the post-module survey:

> I liked [the module] because we were able to investigate different attributes to try and figure out what causes the income gap between males and females... Also making different graphs to those different attributes was pretty cool to try and figure out so we have a better understanding on how to work and create graphs. *(Written survey response from one of Julie's students)*

During a focus group interview, a student from one of Julie's classes echoed these comments, noting that in their mathematics classes they normally examine only bivariate relationships and do not use tools that allow them to see data relationships change dynamically in multivariable contexts. The student explained:

> When we had the income and an attribute, and then we had the gender, we're not used to having so many different variables on the actual graph. Usually, it's just two on the graph and two on a separate graph. This way we could have one graph with all the things... It was cool watching how the data would in front of us change while we added an attribute. (*Focus group comment from one of Julie's students*)

6.3. RQ3: Students' critical data literacy or social justice learning outcomes

Although we did not embed explicit critical data literacy learning objectives in the module lessons, we were interested in exploring whether and in what ways students demonstrated outcomes that Gutstein (2003, 2006) describes as goals of social justice pedagogy. Based on both student and teacher data, there are signs that the module contributed to at least one goal—building students' social and political awareness. For example, in response to the post-module survey, a student from Eve's class wrote that the module "tied into the real-life world… and brought awareness to certain things," while another of her students wrote that the module "helped me understand both statistics and society in more in-depth ways." Other students discussed how they had heard about income inequality from social media or had observed social inequalities in their daily lives. Seeing statistics and larger-scale data on the topic, however, helped to make the phenomenon more real. One of Eve's students wrote in the post-module survey: "I thought that there wasn't as much of a difference in the wage gaps between males and females let alone between the different races," but then "it was enjoyable to calculate out the numbers to find out the honest truth behind income inequality." In a focus group interview, several students of different racial and ethnic backgrounds from Will's class shared similar reflections (with exchanges about media sources omitted):

133 Interviewer:	What did you guys know about income inequality before you jumped into these lessons?
134 Student 1:	To be honest, I thought it was a myth.
135 Interviewer:	You thought it was a myth?
136 Student 1:	Yeah.
137 Interviewer:	How about for others?

[….]

142 Student 2: I agree that I saw it in the media a lot, but I don't think we were ever presented with proper information behind it. Like you always see on the news, oh the wage gap between men and women, but there would never be any statistics behind it. So I think this is an interesting input to actually see the statistics behind it.

[...]

149 Interviewer: How about others… where have you heard about income inequality before?

[...]

153 Student 3: From like real life. Like I would obviously see that colored people are not making, or not just colored, but like Hispanics don't make as much as others.

154 Interviewer: How can you tell?

155 Student 3: Just because of the jobs that some are forced to do. Like you just know that they're not making as much as like a lawyer would make, and it's obvious. But if I …*[inaudible]* see a percentage or see something saying about income inequality, I'm just like okay, so it does exist. And then I see more and more, and it's obviously real.

Not only did analyzing data help to cement the reality of income inequality in some students' minds, but the scale of the inequalities helped to awaken awareness of the depth of the issue. One of Will's students wrote in the post-module survey: "The overall idea of having attributes to really compare [income] data from 2017 is what astonished me, I didn't realize so many people were making $10,000 to nothing." In her debriefing interview, Bella observed the same reaction among her students: "They really latched onto the fact that 25 percent of the country makes less than $15,000 a year and what that means, and really how shocking that is." The ability to explore multiple explanatory variables also helped to heighten students' sense of the persistence and deep-rooted nature of the wage gap between male and female employees. Rachel described her students' reactions to the last lesson of the module: "They got more and more outraged that no matter what the education level was, the male median [income] was always larger than the female median and sometimes quite a bit larger. And they just kept going, 'But that's not fair.'"

There are also signs that for some students, the module may have helped to instill a greater sense of agency in their learning. One of Will's focus group students said: "I thought [the module] treated us more like students and less like children… because it asked for our input as opposed to telling us, hey, this is [it], learn about it." Another of his students wrote in the post-module survey that "I liked the income inequality data more than the statistics lessons we usually do because of the fact that it felt like we were our own statisticians analyzing data." Although the module did not include activities specifically designed to promote students' collective agency, at least one of Julie's students suggested that seeing evidence of income inequality in the data helped to spark such agency. In a focus group interview, she said, "You always hear about it, like there's income inequality between different genders, and when we actually did the data it was like, oh wow, this is real. People aren't just making things up. This is a real problem, and maybe *hopefully we can figure something out"* (emphasis added).

At the same time, some challenges may have hindered critical data literacy or social justice learning outcomes. One challenge involved feelings of disempowerment among some students when studying income inequality. In her debriefing interview, Bella shared: "I was a little worried during the project that there was this sense of this is inevitable." Although she wanted to address this sentiment, she said that "I didn't really have any time to talk about what's being done… [like] pay equity, laws, and activism." Sasha shared similar reflections in her interview. She had noticed a "fatalism" among some students, which data on income inequality had reinforced. She voiced what she believed were students' thoughts: "Okay, now I can see that this is how it is for all Black men. So why do I bother? All the… crashing, crushing sense of inevitability, and I'm never going to be able to get out of the system." When asked how she handled these sentiments, Sasha said that she raised examples of social activism and protest movements that have made a difference in U.S. history, but she did not have time to go in depth with these topics.

7. Discussion

Prior research efforts to promote students' critical data literacy or to teach mathematics for social justice have typically involved smaller-scale classroom case studies and interventions in which developers have played an active role in facilitating learning activities (e.g., Brantlinger, 2013; Gutstein, 2003; Gutstein, 2006; Kahn, 2020; Rubel, 2016; Van Wart et al., 2020). This study adds to the literature by examining a critical data literacy intervention that was developed for implementation by high school mathematics teachers, without classroom facilitation support by project developers. Employing both quantitative measures and qualitative data, we explored changes in learning and affective outcomes among approximately 200 students of seven teachers of grade 12 non-AP mathematics classes in schools with high proportions of students from historically marginalized groups. We found that teachers were able to implement the module after receiving less than two days of professional development, and module use was associated with statistically significant growth in students' understanding of fundamental data literacy concepts and interests in statistics and data analysis. Our results suggest that the curriculum approach has potential to strengthen important *data literacy* outcomes—which we consider to be the foundation for critical data literacy—for diverse student populations. Study designs with randomized control group comparisons and strategies to reduce student attrition are needed, however, to establish the true efficacy of the module. Additional qualitative research would also help us understand why student outcomes may vary so widely across teachers and under what conditions the module may hold the greatest promise. Educators need such interventions to help all students of all backgrounds gain knowledge and practices that are essential for participation in today's data-driven society.

Several module components appeared to play large roles in supporting students' data literacy interests and understandings. One component involved the module's topical focus: income inequality. Consistent with theories of culturally relevant pedagogy (e.g., Aronson & Laughter, 2016; Ladson-Billings, 1995), students shared that the topic was compelling because it was a real-life issue relevant to their lives. What makes a topic or learning activity relevant, however, may have different interpretations (e.g., Ladson-Billings, 2014; Paris, 2012). Enyedy and Mukhopadhyay (2007) found that learning activities may be socially or culturally relevant in at least three ways: by drawing on contexts from students' daily lives, by attending to topics that students find valuable in their lives outside of school, and by adopting learning processes and norms that are familiar to students. The students in our study suggested additional nuances and details to the concept of relevance. A topic like income inequality may be relevant because it is an authentic problem that students feel they may face in real life—not an artificial question such as the relationship between GPA and height. It may be relevant by providing students with practical information to guide their own future life decisions, such as where to live or what occupations to consider for maximizing future earnings. Relevant topics may also involve social injustices that arouse one's moral indignation. For some students in our study, the topic of income inequality was relevant because it served individualistic or instrumental goals; for others, relevance connected to larger social, emotional, or moral concerns. Lee et al. (2021) recommend that learning designers seek data sources and topics with which students can connect at personal, cultural, or social and political levels. Students in our study appeared to find the topic of income inequality meaningful at all three of these levels.

Dynamic and interactive data features within CODAP also appeared to help engage students in data practices and build their understandings of important data concepts. In fall 2019 and early 2020, prior to the onset of the COVID-19 pandemic and during the period of this study, students and teachers said that statistics learning typically involved teacher lectures, independent problem-solving with textbook-based problems, and computations performed with graphing calculators. The prevalence of more didactic pedagogy and older forms of technology is in keeping with research that has documented the slow pace of reform in statistics education (Zieffler et al., 2018). The teachers who chose to participate in our study were attracted to the inquiry-based approach of our project model and may have needed the curriculum and professional development supports that we offered to move toward more student-centered forms of instruction. In these contexts, many students found it novel and exciting to investigate official and large-scale data from the U.S. Census Bureau using CODAP's interactive computer-based data visualization tools. Both students and teachers suggested that the ability to display, explore, and even manipulate the shapes and statistical measures of large data distributions—particularly the provocative and highly skewed distribution of individual incomes in the United States—helped students gain a deeper understanding of basic statistical concepts, such as when and why the mean and median may differ. Traditional methods of instruction typically fail to support such understandings (Konold & Higgins, 2003). Students showed the greatest growth in their understanding of measures of center based on their pre- and post-module assessment scores, suggesting that CODAP tools and module activities may have been especially helpful for learning in this domain. In addition, the Microdata Portal enabled students to draw and investigate repeated samples of varying sizes from available decennial census and ACS data. At least one teacher felt that this activity helped students develop an intuition for the level of accuracy associated with statistical estimates from samples of different sizes. This is an intuition that is challenging and complex to build (Chance et al., 2004; Garfield et al., 2015).

The use of person-level microdata also appeared to enhance students' data interests and learning. Public data that non-experts can access and analyze from sources like the U.S. Census Bureau typically provide information about people and their characteristics at aggregate levels, involving units of analysis such as states, metropolitan areas, counties, or neighborhoods. Although neighborhood-level or other geographic-level data can provide rich opportunities to examine spatial patterns in aggregate outcomes of interest (e.g., Kahn, 2020; Radinsky, 2014; Rubel et al., 2016; Van Wart et al., 2020), units of geography obscure the individuals within them and may make it more difficult for students to relate to the data. Students and teachers in our study indicated that by working with person-level data in CODAP, students could see and identify with the individual people in their data samples. By engaging in activities that focused attention on the original ACS questionnaire and on CODAP's linked displays of individuals in table rows and as dots in dot plots, at least some students seemed to develop a stronger understanding of the provenance of the data and the multiple ways in which data can be represented. Prior research points to the barriers that learning designers may face when trying to help students connect with secondary data that they have not collected themselves (e.g., Lee & Wilkerson, 2018). Learning activities involving person-level microdata, when available in student-friendly formats, may help remove some of these barriers.

The tools and activities that we employed to advance multivariable thinking—an important aspect of data literacy—may have had an impact within our study sample, given positive findings from our pre- and post-module assessment data. The opportunity to examine relationships among multiple variables also appeared to support aspects of *critical data literacy* in at least some of our participants, as suggested by our qualitative data. Some students indicated that they developed greater awareness of income inequality between higher- and lower-income earners by seeing the large numbers of individuals earning low wages, the low numbers earning high wages, and how this pattern has persisted (and worsened) over multiple decades. Uncovering large-scale and persistent data patterns helped to validate some students' experiences, showing them that observations of income disparities from their own lives were not flukes but were systemic in society. Students who had been previously skeptical of claims about income inequality also found the data patterns credible and convincing, perhaps because they had conducted in-depth activities considering the origins and limitations of their data as well as the levels of accuracy afforded by their data samples.

Students also became much more convinced that sex discrimination may exist by examining differences in the typical wages of male and female employees and finding that the wage gap does not disappear (and sometimes grows) after they had the opportunity to control for multiple other variables. Enyedy and Mukhopadhyay (2007) raised warnings that in certain contexts, students may not examine data critically and may simply draw conclusions that accord with their own preconceptions. Our work suggests that one way to guard against this outcome is to design activities in which students investigate claims about social disparities—just as scholars of TMSJ advocate—with tools and supports that encourage students to explore multiple competing explanations with multivariable data. Such activities have strong potential to advance aspects of critical data literacy that align with two of Gutstein's (2003, 2006) goals for social justice pedagogy: building students' awareness of larger social and political forces in their lives and fostering their agency to advocate for social change with data.

Designing and implementing such learning activities in classrooms with diverse populations will face challenges. Philip et al. (2016) note how discussions of social inequality in the United States almost inevitably raise issues of race and how conversations about race without skilled facilitators can be harmful to students. Although we did not observe negative exchanges involving race in our participating classrooms, and although teachers and students did not report any such exchanges, we cannot be sure that our classrooms avoided them entirely. Educators who wish to promote critical data literacy among students need to know not only how to facilitate students' learning of technical data literacy skills, but also how to guide students through conversations that explore why and how different groups may experience unequal social, economic, or political outcomes. In addition, there is a risk that extended investigations of topics that highlight systemic social injustices can feel demoralizing for students, especially for those students from historically marginalized groups. As suggested by some of our participating teachers, a strategy for countering such responses may be to include examples of individual and collective actions that have led to positive social change. Within the microcosm of classrooms, teachers can also work to establish norms and facilitate learning activities that value and build upon the contributions of students from all backgrounds to help students experience and envision more equitable social interactions and arrangements (Alim & Paris, 2017). Educators will need time, materials, and support to pursue these strategies, however, if efforts to promote critical data literacy are to reach their full potential.

Acknowledgement

This work was supported by the National Science Foundation (NSF) under Grant No. DRL 1813956. Opinions, findings, conclusions, or recommendations are those of the authors and do not necessarily reflect the views of the NSF.

References

Alim, H. S., & Paris, D. (2017). What is culturally sustaining pedagogy and why does it matter? In D. Paris & H. S. Alim (Eds.), *Culturally sustaining pedagogies: Teaching and learning for justice in a changing world* (pp. 1–24). Teachers College Press.

Aronson, B., & Laughter, J. (2016). The Theory and practice of culturally relevant education: A Synthesis of research across content areas. *Review of Educational Research, 86*(1), 163–206. https://doi.org/10.3102/0034654315582066

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. A. (2020). *Pre-K–12 guidelines for assessment and instruction in statistics education II (GAISE II)* (2nd ed.). American Statistical Association.

Baumer, B. S., Garcia, R. L., Kim, A. Y., Kinnaird, K. M., & Ott, M. Q. (2022). Integrating data science ethics into an undergraduate major: A Case study. *Journal of Statistics and Data Science Education, 30*(1),15–28. https://doi.org/10.1080/26939169.2022.2038041

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.

Bhargava, R., Deahl, E., Letouze, E., Noonan, A., Sangokoya, D., & Shoup, N. (2015). *Beyond data literacy: Reinventing community engagement and empowerment in the age of data.* Data-Pop Alliance (Harvard Humanitarian Initiative, MIT Media Lab and Overseas Development Institute) and Internews.

Brantlinger, A. (2013). Between politics and equations: Teaching critical mathematics in a remedial secondary classroom. *American Educational Research Journal, 50*(5), 1050–1080. https://doi.org/10.3102/0002831213487195

Chance, B., Ben-Zvi, D., Garfield, J., & Medina, E. (2007). The Role of technology in improving student learning of statistics. *Technology Innovations in Statistics Education, 1*(1). https://doi.org/10.5070/T511000026

Chance, B., delMas, R., & Garfield, J. (2004). Reasoning about sampling distributions. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of developing statistical literacy, reasoning, and thinking* (pp. 295–323). Springer.

Cuoco, A., Goldenberg, E. P., & Mark, J. (1996). Habits of mind: An Organizing principle for mathematics curriculum. *Journal of Mathematical Behavior*, *15*(4), 375–402. https://doi.org/10.1016/S0732-3123(96)90023-1

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press.

Dee, T., & Penner, E. (2016). *The Causal effects of cultural relevance: Evidence from an ethnic studies curriculum* (CEPA Working Paper No.16-01). Stanford Center for Education Policy Analysis. http://cepa.stanford.edu/wp16-01

delMas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal, 6*(2), 28–58. https://doi.org/10.52041/serj.v6i2.483

Delpit, L. D. (1988). The Silenced dialogue: Power and pedagogy in educating other people's children. *Harvard Educational Review, 53*(3), 280–298. https://doi.org/10.17763/haer.58.3.c43481778r528qw4

DeVellis, R. F. (2012). *Scale development: Theory and applications* (3rd ed.). Sage Publications.

Donoho, D. (2017). 50 years of data science. *Journal of Computational and Graphical Statistics, 26*(4), 745–766. https://doi.org/10.1080/10618600.2017.1384734

Engel, J. (2016). Open data at the interface of mathematics and civics education: Challenges of the data revolution for the statistics curriculum. *Journal of Mathematics and Statistical Science*, *2*(5), 264–273.

Enyedy, N., & Mukhopadhyay, S. (2007). They don't show nothing I didn't know: Emergent tensions between culturally relevant pedagogy and mathematics pedagogy. *Journal of the Learning Sciences, 16*(2), 139–174. https://doi.org/10.1080/10508400701193671

Finzer, W. (2013). The Data science education dilemma. *Technology Innovations in Statistics Education, 7*(2), 1–9. https://doi.org/10.5070/T572013891

Frankenstein, M. (2009). Developing a criticalmathematical numeracy through *real* real-life word problems. In *Words and Worlds: Modelling Verbal Descriptions of Situations* (pp. 111–130). Sense Publishers.

Frankenstein, M. (2013). Reading the world with math. In E. Gutstein & B. Peterson (Eds.), *Rethinking mathematics: Teaching social justice by the numbers* (pp. 30–39). Rethinking Schools.

Freire, P. (1970). *Pedagogy of the oppressed* (30th ed.). Continuum International.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review, 7*(1), 1–51. https://doi.org/10.1111/j.1751-5823.2002.tb00336.x

Garfield, J. B., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Springer.

Garfield, J., delMas, B., Chance, B., & Ooms, A. (2006). *Summary statistics for a national sample of undergraduates: Comprehensive assessment of outcomes for a first course in statistics (CAOS).* The Web ARTIST Project.

Garfield, J., Le, L., Zieffler, A., & Ben-Zvi, D. (2015). Developing students' reasoning about samples and sampling variability as a path to expert statistical thinking. *Educational Studies in Mathematics, 88,* 327–342. https://doi.org/10.1007/s10649-014-9541-7

Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education, 53,* (2), 106–116. https://doi.org/10.1177/0022487102053002003

Gay, G. (2013). Teaching to and through cultural diversity. *Curriculum Inquiry, 43*(1), 48–70. https://doi.org/10.1111/curi.12002

Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal, 16*(1), 22–25. https://doi.org/10.52041/serj.v16i1.209

Gutstein, E. (2003). Teaching and learning mathematics for social justice in an urban, Latino school. *Journal for Research in Mathematics Education, 34*(1), 37–73. https://doi.org/10.2307/30034699

Gutstein, E. (2006). *Reading and writing the world with mathematics.* Routledge.

Gutstein, E. (2013). Home buying while Brown or Black. In E. Gutstein & B. Peterson (Eds.), *Rethinking mathematics: Teaching social justice by the numbers* (pp. 61–66). Rethinking Schools.

Henke, N., Bughin, J., Chui, M., Manyika, J., Saleh, T., Wiseman, B., & Sethupathy, G. (2016). *The Age of analytics: Competing in a data-driven world.* McKinsey Global Institute.

Jacobbe, T., Case, C., Whitaker, D., & Foti, S. (2014). Establishing the validity of the LOCUS assessments through an evidenced-centered design approach. In *Sustainability in statistics education. Proceedings of the ninth international conference on teaching statistics (ICOTS9).* International Statistical Institute. https://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_7C2_JACOBBE.pdf?1405041692

Kahn, J. (2020). Learning at the intersection of self and society: The Family geobiography as a context for data science education. *Journal of the Learning Sciences, 29*(1), 57–80. https://doi.org/10.1080/10508406.2019.1693377

Kokka, K. (2020). Social justice pedagogy for whom? Developing privileged students' critical mathematics consciousness. *The Urban Review, 52*(2), 778–803. https://doi.org/10.1007/s11256-020-00578-8

Konold, C., & Higgins, T. (2003). Reasoning about data. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 193–215). National Council of Teachers of Mathematics.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465–491. https://doi.org/10.3102/00028312032003465

Ladson-Billings, G. (2014). Culturally relevant pedagogy 2.0: A.k.a. the remix. *Harvard Educational Review, 84*(1), 74–84. https://doi.org/10.17763/haer.84.1.p2rj131485484751

Lee, V. R., & Wilkerson, M. (2018). *Data use by middle and secondary students in the digital age: A status report and future prospects.* National Academies of Sciences, Engineering, and Medicine. https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_189500.pdf

Lee, V.R., Wilkerson, M., & Lanouette, K. (2021). A Call for a humanistic stance toward K-12 data science education. *Educational Researcher, 50*(9), 664-672. https://doi.org/10.3102/0013189X211048810

Levels of conceptual Understanding in Statistics [LOCUS] (2002). *Question browser.* https://locus.statisticseducation.org/professional-development/questions/by-grade/high-school-%28grades-9-12%29?page=&type=prodev_const_response_question

Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., & Harackiewicz, J. M. (2010). Measuring situational interest in academic domains. *Educational and Psychological Measurement, 70*, 647–671. https://doi.org/10.1177/0013164409355699

Louie, J. (2022). *Critical data literacy: Creating a more just world with data.* National Academies of Sciences, Engineering, and Medicine. https://www.nationalacademies.org/documents/embed/link/LF2255DA3DD1C41C0A42D3BEF0989ACAECE3053A6A9B/file/D16254F310D01BBDA873920E4EFB8151F2D8334181AA

Louie, J., Roy, S., Chance, B., Stiles, J., & Fagan, E. (2021). Promoting interest and skills in statistical and multivariable thinking with social justice data investigations. In R. Helenius & E. Falck (Eds.), *Statistics education in the era of data science: Proceedings of the International Association for Statistical Education (IASE) 2021 Satellite Conference*. https://iase-web.org/documents/papers/sat2021/IASE2021%20Satellite%20133_LOUIE.pdf?1649974212

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big data: The Next frontier for innovation, competition, and productivity*. McKinsey Global Institute. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation

National Academies of Sciences, Engineering, and Medicine. (2018). *Data science for undergraduates: Opportunities and options*. National Academies Press. https://doi.org/10.17226/25104

National Governors Association Center for Best Practices & Council of Chief State School Officers [NGA Center & CCSSO]. (2010). *Common core state standards for mathematics*.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Pangrazio, L., & Selwyn, N. (2019). 'Personal data literacies': A Critical literacies approach to enhancing understandings of personal digital data. *New Media and Society, 21*(2), 419437. https://doi.org/10.1177/1461444818799523

Paris, D. (2012). Culturally sustaining pedagogy: A Needed change in stance, terminology, and practice. *Educational Researcher, 41*(3), 93–97. https://doi.org/10.3102/0013189X12441244

Philip, T. M., Olivares-Pasillas, M. C., & Rocha, J. (2016). Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-contestations. *Cognition and Instruction, 34*(4), 361–388. https://doi.org/10.1080/07370008.2016.1210418

Radinsky, J., Hospelhorn, E., Melendez, J. W., Riel, J., & Washington, S. (2014). Teaching American migrations with GOS census webmaps: A Modified "backwards design" approach in middle-school and college classrooms. *Journal of Social Studies Research, 38, 143-158.* https://doi.org/10.1016/j.jssr.2014.02.002

Raffaghelli, E. J. (2020). Is data literacy a catalyst of social justice? A Response from nine data literacy initiatives in higher education. *Education Sciences, 10*(9), 233. https://doi.org/10.3390/educsci10090233

Ridgway, J. (2015). Implications of the data revolution for statistics education: The Data revolution and statistics education. *International Statistical Review, 84*(3), 528–549. https://doi.org/10.1111/insr.12110

Rubel, L. H., Lim, V. Y., Hall-Wieckert, M., & Sullivan, M. (2016). Teaching mathematics for spatial justice: An Investigation of the lottery. *Cognition and Instruction, 34*(1), 1–26. https://doi.org/10.1080/07370008.2015.1118691

Rubin, A. (2020). Learning to reason with data: How did we get here and what do we know? *Journal of the Learning Sciences, 29*(1), 154–164. https://doi.org/10.1080/10508406.2019.1705665

Sproesser, U., Engel, J., & Kuntze, S. (2016). Fostering self-concept and interest for statistics through specific learning environments. *Statistics Education Research Journal, 15*(1), 28–54. https://doi.org/10.52041/serj.v15i1.256

Van Wart, S., Lanouette, K., & Parikh, T. (2020). Scripts and counterscripts in community-based data science: Participatory digital mapping and the pursuit of a third space. *Journal of the Learning Sciences, 29*(1), 127–153. https://doi.org/10.1080/10508406.2019.1693378

Weiland, T. (2017). Problematizing statistical literacy: An Intersection of critical and statistical literacies. *Educational Studies in Mathematics, 96*, 33–47. https://doi.org/10.1007/s10649-017-9764-5

Wolff, A., Gooch, D., Cavero Montaner, J. J., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics*, *12*(3). https://doi.org/10.15353/joci.v12i3.3275

Zieffler, A., Garfield, J., & Fry, E. (2018). What is statistics education? In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education*. *Springer International Handbooks of Education* (pp. 37–70). Springer. https://doi.org/10.1007/978-3-319-66195-7_2

Appendix 1

Sample Items from the Pre- and Post-Module Data Literacy Assessment

I. Sampling and Data Collection

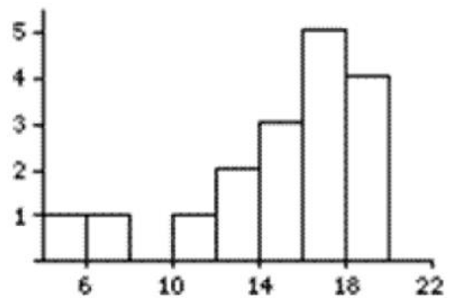
Nathan read that students tend to carry backpacks that are very heavy. He decides to collect data to determine if this is true for his school. Which of the following would be the most appropriate data for Nathan to collect?

a. The number of books in students' backpacks
b. Whether or not students carry a backpack to school
c. The weight of the backpacks that students carry to school
d. The number of days per week students carry a backpack to school

Source: Adapted from Levels of Conceptual Understanding in Statistics (LOCUS), 2022.

II. Measures of Center

In the graph of Quiz Scores below, which estimates of the mean and median are most likely to be correct?



Quiz Scores

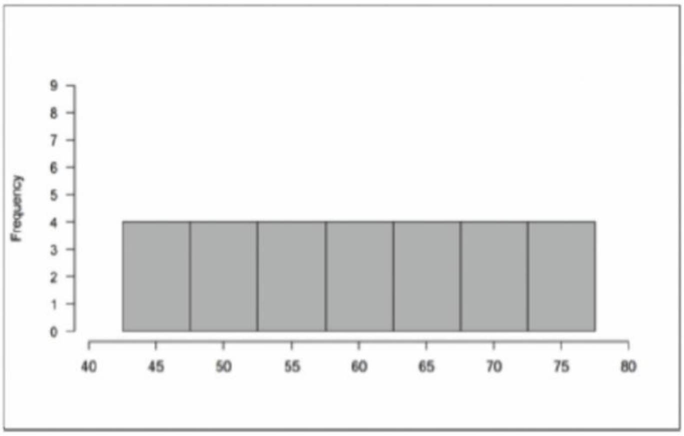
a. median = 13.0 and mean = 12.0
b. median = 14.0 and mean = 15.0
c. median = 16.0 and mean = 14.3
d. median = 16.5 and mean = 16.2

Source: Adapted from del Mas et al. (2007).

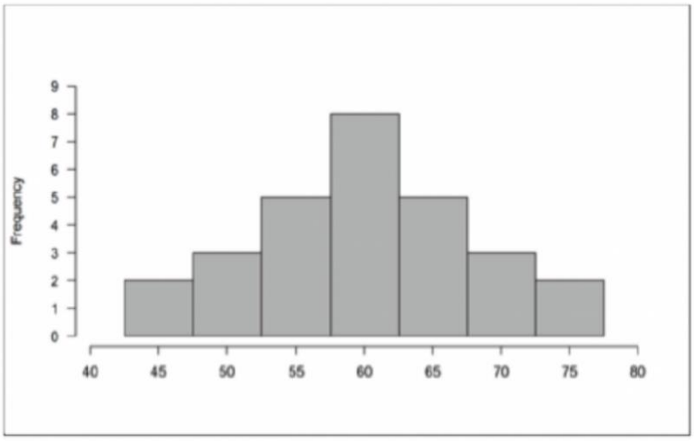
III. Measures of Variability

Three data sets are summarized in the histograms below.

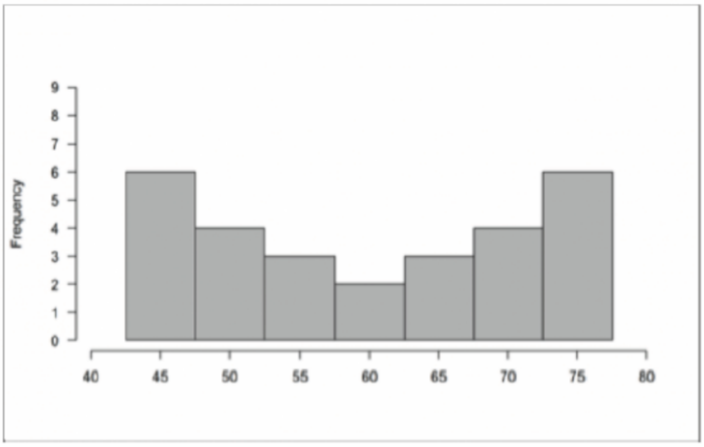
Data Set A



Data Set B



Data Set C



Which set of data varies the most from its mean?

a. Data set A

b. Data set B

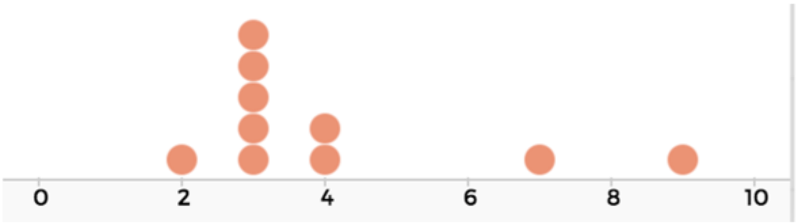
c. Data set C

d. The variability from the mean is the same for all three data sets.

Source: LOCUS (2022).

IV. Data Representation

The dotplot below shows the quiz scores for each student who took a quiz. The quiz had 10 questions.

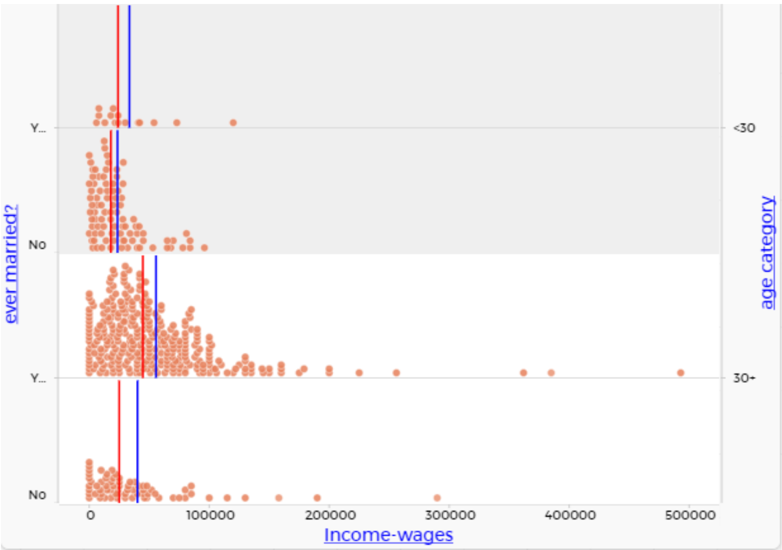


How many students received scores higher than 4?
a. 1
b. 2
c. 3
d. 4

Source: Authors.

V. Multivariable Thinking/Reasoning

Below is a set of graphs that add data on a third variable: age. The same 476 employed individuals in the graph above are now split into two age categories: whether they are under 30 years old (<30), or 30 years old or above (30+).



Source: Authors.

Based on the additional information in the graph above, which of the following is an appropriate statement?

a. There is no difference in income-wages for people who have been married compared to people who have never been married once you control for a person's age category.

b. There is no difference in income-wages for people under age 30 compared to people age 30+ once you control for whether they have ever been married.

c. People who have been married earn higher income-wages, on average, than people who have never been married, even after controlling for their age category.

d. People age 30+ earn lower income-wages, on average, than people under age 30, even after controlling for whether or not they have ever been married.

Appendix 2

Table 7. Example interest survey items and associated subscales

Item	Scale
Our statistics lessons over the past couple of weeks sparked my interest in investigating data.	Situational interest
What we learned in *statistics class over the past couple of weeks* is important to me.	
Investigating data is one of my favorite activities.	Individual interest
I like to think about statistics and data even when I'm outside of my statistics class.	
Understanding tasks with diagrams and statistical data is easy for me.	Self-concept
I am good at solving statistical problems.	
It is important to me to be a person who can analyze data statistically.	Perceptions of value
Statistics and working with data are useful for me to know.	

Note. Phrases in italics were replaced with "the Income Inequality data lessons" in the post-module survey. Each item had a seven-point rating scale, ranging from 1 (Strongly disagree) to 7 (Strongly agree).