MATHEMATICS TEACHING RESEARCH JOURNAL
GOLDEN FALL 2025
Vol 17 no 4

Designing Equity-Focused Pedagogy in Mathematics Teacher Preparation with Digital Data Tools

Nadia Stoyanova Kennedy[1], Boyan Kostadinov[1], Sandie Han[2]

[1]New York City College of Technology of the City University of New York, USA

[2]Medgar Evers College of the City University of New York, USA

nkennedy@citytech.cuny.edu*

Abstract: This study examines the implementation of a course module that integrated computing and data analysis. The module aimed to engage prospective mathematics teachers in the practice of working with datasets and analyzing data to investigate questions related to local Brooklyn schools, as well as to gain deeper insight into the students and communities where they conduct their field experiences. Throughout the module, teacher candidates explored issues such as student learning opportunities, achievement gaps, school segregation, educational (in)equity, and other social justice concerns through readings, data analysis, and direct engagement. They shared and reflected on their findings with peers, deepening their understanding of these topics through collective discussion. The teacher candidates expressed mixed perceptions about the potential of incorporating computing and data analysis into mathematics teaching and learning. While many acknowledged that such integration could enhance instruction by connecting mathematical concepts to real-world issues—such as poverty rates and rising global temperatures—and by helping students make sense of the world, relate mathematics to their own lives, and engage with social justice issues, there was also a shared concern. Most believed that implementing this approach would require substantial additional planning and would likely not be feasible during the early years of their teaching careers.

Keywords: mathematics teacher education, computational thinking, data analysis, critical dialogue, equitable teaching, culturally responsive pedagogy

INTRODUCTION

The authors are interested in enriching mathematics teacher education by including computational thinking and digital literacies (CDL) in teacher education and including more versatile and contemporary approaches to mathematical modeling, as well as incorporating critical perspec-

tives on computational thinking and mathematical modeling at its intersections with social issues.
We believe that computational thinking, digital literacies, and modeling can help prospective mathematics teachers engage their students with more authentic mathematics problem solving and modeling with more efficient and powerful technology tools. We are also persuaded of the importance of incorporating critical perspectives on computational and mathematical thinking and modeling as such perspective is essential in teaching that is equitable and culturally responsive and for participating in a democratic society.

In the summer of 2022, we joined the CUNY Computing Integrated Teacher Education (CITE)) -- a multi-year initiative to support CUNY faculty in integrating state standards and aligning computing content and pedagogy with required education courses, field work, and student teaching. The initiative focuses on supporting institutional change in teacher education programs; supporting faculty computing pedagogical content knowledge through the lens of culturally response-sustaining education; and faculty research in equitable computing education and STEM pedagogies. Its core value is equity, and "equitable pedagogies, which seek to empower learners and communities, promote meaningful and joyful learning, and transform institutions towards justice" (The CUNY CITE Equity Working Group, 2023). All CITE work is guided by the following design principles: "co-learning and co-construction of knowledge in communities; supporting learner agency in experimenting with, modifying and creating tools; encouraging creativity and expression; mobilizing computing for social action; vetting and critiquing tools and tech cultures; and adopting expansive notions of learning" (The CUNY CITE Equity Working Group, 2023).

Prior to the CITE initiative, our team had already embarked on computing integration by designing and piloting computational modeling projects with prospective mathematics teachers. Joining CITE helped us layer on computing and digital literacies (CDL) and get support from a diverse group of colleagues. Initially, we were focused exclusively on introducing teacher candidates to programming and coding. We consider that an expanded conception of computing integration would be much more valuable to prospective teachers, not only in terms of enhancing their mathematics teaching, but also in preparing them to use various technology tools critically. CITE offered us experimental space in which our team could learn, create, discuss, and try out our ideas and artifacts. The CITE initiative helped us reconceptualize computing integration as a larger domain, encompassing not only coding and computational modeling, but digital literacy, and the use of digital tools, accompanied by a critical stance on technology, mathematics, and their uses.

As part of our work related to CITE, we developed a course module to support prospective mathematics teachers in developing data analysis and critical thinking skills. The module was designed to engage teacher candidates in analyzing and visualizing data from education data sets, posing questions, analyzing and interpreting data to answer the questions, and visualizing interpretations to communicate findings. The module also engaged teacher candidates in exploring school segregation, educational opportunities, and education (in)equities in the context of these

data analyses and visualizations. In addition, we planned for discussion among teacher candidates of bias in databases and data representations, of how bias can be introduced or guarded against, and the impact of data technologies on society and individuals. The course module was informed by the *New York State K12 Computer Science and Digital Fluency Learning Standards* (2024), the *NYS Culturally Responsive-Sustaining Education Framework* (2024), and the *Equitable CITE Pedagogy and Design Principles Framework* (CUNY CITE Equity Working Group, 2023). The module was featured on the CUNY CITE website (Kennedy, 2023).

The purpose of our research project was two-fold. It aimed at: a) exploring the implementation of a module with integrated computing components designed for preservice mathematics teachers, and b) exploring their perceptions of the mathematics teacher candidates of the potential of computing integration in mathematics teaching and learning.

The research questions which guided the study were as follows:

RQ1: What are the characteristics of the implementation process of the course module with integrated computing component designed for mathematics preservice teachers? What kind of teacher learning did it facilitate, if any?

RQ2: How do preservice teachers perceive the potential of integrating computing, and data analysis in particular, in mathematics teaching and learning? What are their beliefs about the challenges and opportunities of such integration?

LITERATURE REVIEW

The emergence in recent years of multiple integrated fields such as computational biology, computational chemistry, computational geometry, and computational physics is a testimony to the power of computation and the potential of its integration with classical STEM fields in order to solve more complex problems. The notion of CDL integration into STEM classrooms is relatively new. Computational thinking (CT) is understood as the thought process involved in formulating problems whose solutions can be expressed as computational steps or algorithms to be carried out by a computer (Lee et al., 2019). CDL is a more expansive notion, which in addition to CT also includes computing and digital literacies such as tinkering, experimenting, digitally supported communication, critically and ethically navigating digital tools and more (Vogel, Patel, Yadav, 2025). Many researchers argue for the inclusion of CDL in mathematics and science classrooms (e.g., Weintrop et al., 2016; Lee et al., 2019). Some see such integration as creating the potential for students to build a strong and mutually supportive relationship between mathematical knowledge and computational knowledge (Weintrop et al., 2016). It has been suggested that students who learn to engage with and use computational methods and tools will advance their understanding of mathematics and will gain awareness of current applications of CT in mathematics and across STEM fields. Others insist that the integration of mathematics and CT allows students to engage with real-world problems, authentic tools, and authentic mathematics practices (Bain et al., 2021). There has in fact been progress in elucidating what the integration of CT and mathematics looks like in the classroom (e.g., Kennedy, Kostadinov, Masuda, 2024; Benakli et al., 2016; Bain et al., 2021) but work on the integration of CT in the mathematics

classroom is still needed, ranging from project designs to pedagogies. While some mathematical problems lend themselves to computational approaches, problems and course modules often need to be reframed so that existing computational tools can be utilized.

The integration of computational thinking and computer science in K-12 education has been an ongoing process for decades (Nouri et al., 2020; Wing, 2006; 2008). For CDL to become part of the K-12 curriculum, it is imperative that teachers are well trained to incorporate CDL in their teaching (Barr & Stephenson, 2011). Teacher education plays a critical role in equipping pre-service teachers with the knowledge and skills needed to teach CDL and integrate it into their practice (Yadav, Stephenson, & Hong, 2017). Before implementing CDL in teacher preparation programs, it is important to gain a better understanding of both the process of computing implementation, and the prospective teachers' perceptions of CDL integration. This study attempts to shed light on both of these issues.

METHODS

Design

We adopted an exploratory case study approach (Yin, 2017), centered around examining the implementation of the designed course module with a group of eight mathematics teacher candidates and the learning it engaged teacher candidates in. For this research study, we combine four data collection instruments (authors' descriptive and reflective notes, participants' work relevant to the study and transcripts from class discussions, and participants' focus-group interview).

Study Context and Participants

The participants in the study are eight prospective mathematics teachers taking a mathematics education course in Fall 2022, which was taught by the first author. The course participants are undergraduate prospective teachers in their second or third year of the undergraduate degree, who have completed an initial round of school observations of mathematics teaching and learning. At the completion of the program, they expect to obtain initial certificates and teach mathematics in grades 7 to 12 as full-time mathematics teachers in public NYC schools. The group of teacher candidates is invariably diverse in terms of race, ethnicity, education experiences, and age. A large percentage of these students have completed their middle and high school education outside of the United States.

The course module was designed for the aforementioned mathematics education course, which is part of the required undergraduate mathematics education program curriculum on campus. The course has a practice-oriented component, which includes 60 hours of intermediate field school-based experience, as well as lesson and unit planning, design of student assessment, and delivery of instruction. It involves micro-teaching – teaching small segments of lessons in real classrooms – which provides an appropriate context for prospective teachers to do some exploratory work

with school data sets and to perform data analysis in order to learn more about the schools and students they will work with.

As part of the course work the prospective mathematics teachers are engaged in a course module "*Behind the Data: Analysis and Visualization of School and Student Data*." This computing integration module focuses on engaging prospective mathematics teachers in using data sets to answer questions about local Brooklyn schools and learn more about the schools and students where they do field experience. The purpose of the module is to support prospective teachers in developing data analytical and critical thinking skills. For the purposes of this module, we used multiple databases and other data resources, including the NYC open database https://opendata.cityofnewyork.us/, the Office for Civil Rights' Data on Equal Access to Education, and ProPublica's Miseducation database (ProPublica.org).

The original module is comprised of four activities, all of which are based on explorations of data sets. As part of these three activities, prospective teachers compare and contrast schools and educational services in different parts of NYC. They generate a profile of the school where they do field experience by analyzing and visualizing data from an education data set. Teacher candidates articulate a question about their school's student population and find, analyze, and interpret data to answer the question, and visualize interpretations to communicate findings. They explore data about schools and students, as well as relationships between variables (i.e., test scores, demographics, grades, school funding, inequity, and achievements). They are also tasked to investigate the relationship between segregation, educational opportunities, and (in)equities. Each activity ends with a discussion focused on questions related to segregation and (in)equity in education. The fourth activity was structured as a workshop on using R for data analysis with NYC Open database. It ended with a discussion about bias in databases and data representations, how bias can be introduced or guarded against, and the impact of data technologies on society and individuals. In the second iteration of the module design, we replaced the data analysis using of R —which required more extensive learning—with an activity built in CODAP, a more accessible data analysis platform.

Data Collection

The researchers took descriptive and reflective notes during the implementation of the course module. Teacher candidates' artifacts developed in relation to the course module were collected as well. All participants were invited to take part in a focus group interview, which took 60 minutes. The focus group interview started with an invitation to talk freely about their experience with the computational module. The interview was semi-structured, which allowed for follow up questions. The interview was recorded and transcribed for analysis. Consent forms were secured from all participants. All names of the participants were changed to ensure confidentiality.

RESULTS

Our course module integrated computing into an existing mathematics education course. Teacher candidates have read about the "achievement gap" and pervasive inequalities in learning oppor-

tunities in readings for previous mathematics education courses. This module, we thought, would offer them the occasion not only to learn more about their placement schools and their students, but also the chance to encounter head-on the drastic difference in the learning opportunities that NYC schools do (or don't) provide and reflect on them.

Our Intentions

Ultimately, we wanted to encourage prospective teachers to learn more about and to raise questions about their field-work schools and students and to draw conclusions about the educational opportunities available to students. The course module was also intended to challenge prospective teachers' beliefs and conceptions about education and schools. We grounded our work in themes that included (in)equity, school segregation, (in)equitable school funding, and availability of educational opportunities. The team wanted the prospective teachers to understand the deeper social issues that affect schools, education, and learning outcomes. We thought that it would be helpful for teacher candidates to leverage both qualitative methods – conducting observations and reading articles about NYC schools – and quantitative methods – analyzing numerical data about Brooklyn schools' funding, demographics, and other academic characteristics. On the quantitative side, we wanted to support prospective mathematics teachers in developing data analysis and critical thinking skills. For the purposes of this artifact, the team used various open data sets and data resources (see Appendix). The team also wanted to expose prospective mathematics teachers to "R" -- an integrated suite of software for data manipulation, calculation, and graphical display-- so that prospective mathematics teachers would experience "the power of R," and be motivated to learn more about using R in the future. We hoped that prospective teachers would experience how data analysis and data literacy could put their concrete experience into perspective and that it would help them understand it in the context of large social and educational phenomena. Towards those ends, we conducted discussions in line with a pedagogical model called "community of inquiry," which is a model for collective, democratic, and collaborative deliberation. We hoped that this approach would foster participant agency.

The Course Module

The course module is designed to engage prospective mathematics teachers in critical inquiry using real-world educational data sets. Centered around the Brooklyn schools where teacher candidates are placed for their micro-teaching field experience, it includes opportunities for prospective mathematics teachers to read articles, explore data sets, wonder, pose and answer questions, and engage in critical collective discussions. The module is implemented over four consecutive synchronous sessions held over Zoom. Assignments involve readings and a few homework tasks. During the first three 1.5-hour sessions, prospective mathematics teachers explore databases and engage in discussions guided by key questions. Prospective teachers investigate school demographics and educational opportunities (e.g., whether high-level math and statistics classes were available, the percentage of certified teachers, the available technology, etc.). Collective discussions are guided around data visualization, interpretation, and analysis, supporting students in making sense of the readings and of the contexts prospective teachers encountered in the

schools and in drawing conclusions from the data. The final session, is designed as a 2.5-hour immersive workshop focused on the statistical package, R, introducing teacher candidates to basic data manipulation, filtering, and visualization techniques, allowing teacher candidates to extend their inquiry through coding.

Planned Activities at a Glance:

- ***Creation of school profiles through data analysis and visualizations:*** Teacher candidates explore education data sets, and analyze and visualize data related to their assigned field-work schools. They are guided to formulate questions about their schools' student populations and educational opportunities, and use data to analyze and interpret their findings. Each candidate is expected to developed a short presentation – including graphs and visualizations – about their school and its students and share with peers.
- ***Engagement and inquiry with data to understand educational inequity:*** Faculty and teacher candidates discussed the data that had already been collected about schools and students. Teacher candidates interpreted visualizations from various databases to compare and contrast schools and educational services in different parts of the country and in New York City in particular, where they explored relationships between variables (test scores, demographics, grades, etc.) related to social justice, (in)equity, and segregation.
- ***Critical discussions and reflections with data analysis:*** Each session ended with a discussion about critical educational issues such as quality education for all NYC students; available learning opportunities, school segregation, school funding, the impact of funding on student achievement, the broader consequences of segregated schooling for individuals and society, and potential strategies to address these inequities.
- ***Engagement with coding and data analysis with R:*** The final session introduced mathematics teacher candidates to R, guiding them in sorting and filtering data to answer specific questions, and visualize their findings.
- ***Discussions "for" and "against" particular uses of big data:*** Faculty facilitated conversations about the potential benefits of using educational data sets and data analysis in mathematics teaching and learning. Teacher candidates also examined the limitations of big data, interrogating assumptions that may have been made in data collection and visualization. Discussions highlighted what data sets often fail to capture—specifically, the nuanced details of communities and the individual characteristics of students.

In the course design iteration that followed the module's implementation in Fall 2022, the fourth activity, which originally involved data analysis using R, was replaced with an activity utilizing CODAP. This change was made to enhance accessibility and ease of use for students, particularly those with limited prior experience in programming. By integrating CODAP—a more intuitive, web-based data exploration tool—the course aimed to lower technical barriers while still achieving the intended learning objectives related to data analysis and interpretation. The new CODAP-based activity has been designed to engage teacher candidates in data exploration by mapping mathematics achievement scores of New York City schools, allowing comparisons

across neighborhoods and against each neighborhood's Economic Index. The resulting graphs and maps reveal equity gaps that reflect broader economic disparities among ethnic demographics.

Course Module Implementation

The following section discusses the implementation of the first iteration of the designed course module, conducted in Fall 2022.

Creation of School Profile Through Data Analysis and Creating Visualizations

Prior to class sessions, prospective teachers were given an assignment to use the interactive map of the United States in and explore a few education data bases, including the Civil Rights Data Collection and Miseducation (ProPublica). They were instructed to review the "Opportunity," "Discipline," and "Achievement Gap" tabs, as well as the "Black" and "Hispanic," demographic tabs. Additionally, they were asked to search for their field experience school and district in the databases, make notes of what they noticed in the results and write down questions that arose from the data. Prospective teachers were asked to prepare a one-page report summarizing their discoveries about their schools and school districts. In their reflection, they were instructed to include detailed results, graphs and tables from the databases, and be ready to share their findings in class.

At the beginning of the session, the prospective teachers were invited to share their reports and findings, particularly noting anything they found unexpected and surprising. The guiding question was: What did you find about your school/school district? They were also invited to comment on how the database explorations helped make previously unseen patterns or disparities visible to them. The excerpts of candidates' findings, given below, show some very different school profiles. The names of the specific schools have been withheld or changed.

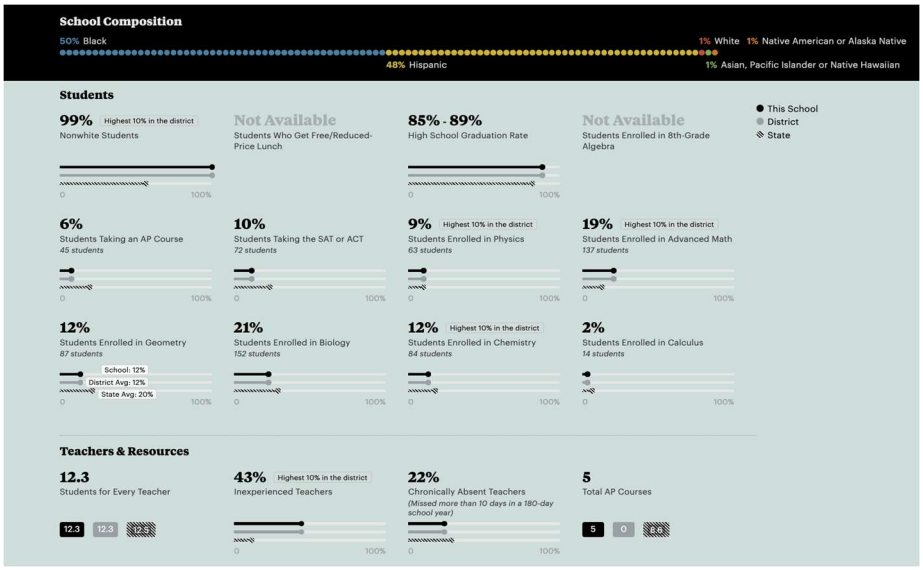


Figure 1: Excerpt from teacher candidate A's profile of their NYC field placement school

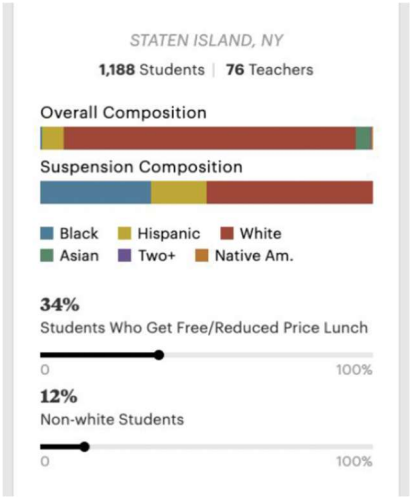
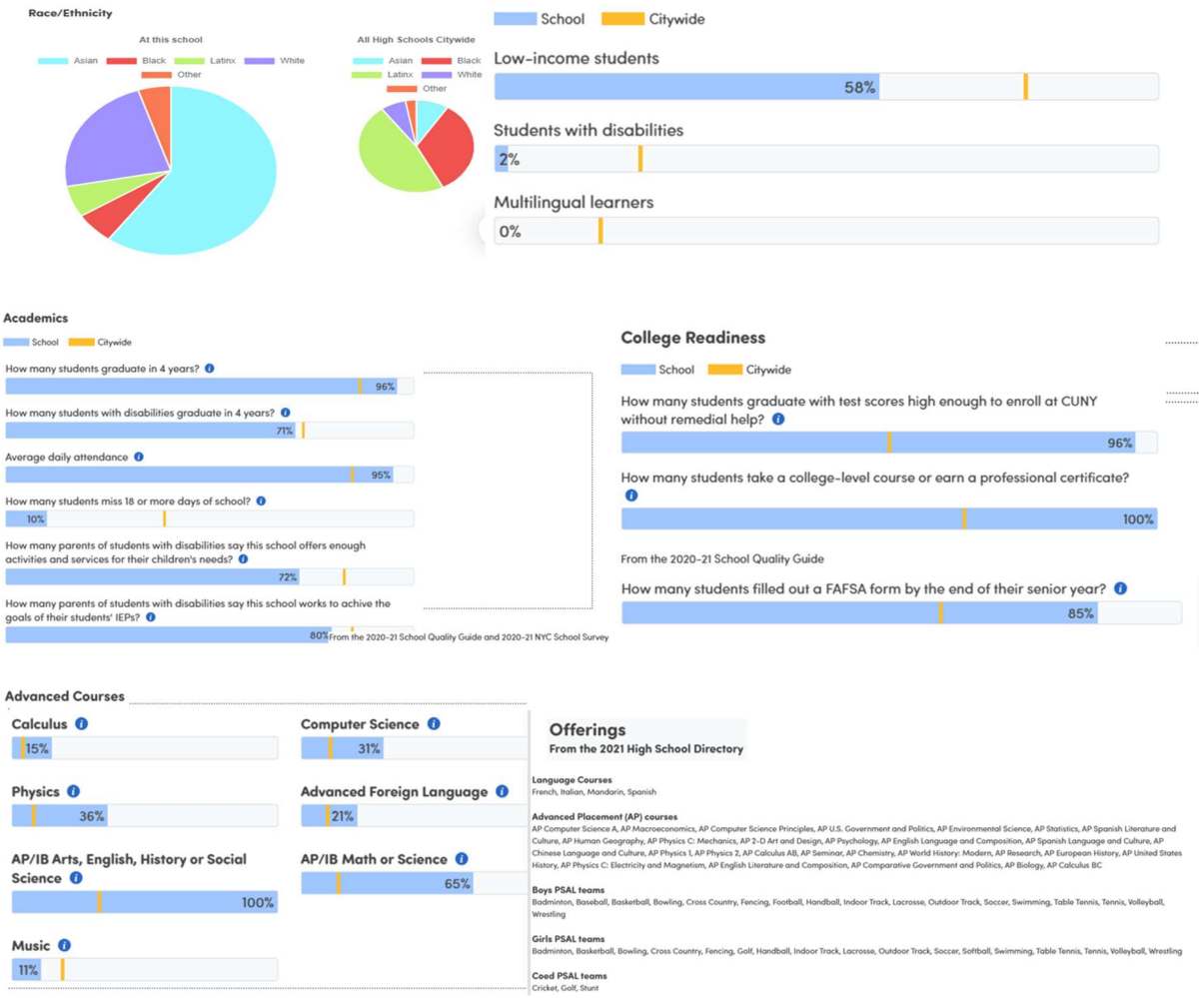


Figure 2: Excerpt from teacher candidate B's profile of their NYC field placement school



"Based on the research on *Brooklyn Star high school*, students have more opportunities than NYC citywide students…. The students take more AP courses and have more options to choose from four foreign languages and 20 AP courses (vs. citywide average 4.9). Their achievement is in the highest 10% in the NYC DOE schools even though their student teacher ratio is twice higher than citywide. The student population of *Brooklyn Star high school* is mainly Asian (Chinese 60%) and White (23 %) much higher compared to the percentage of Asian and White student population in NYC, which are 9% and 7% respectively. There are no ELL students admitted in this school and the school has very few students with SPED. Most students (96%) graduate on time. Every student takes at least one college course."

Note: The name of the school has been changed.

Figure 3: Excerpt from teacher candidate C's profile their NYC field placement school

Engagement and Inquiry with Data to Understand Educational Inequity

In the next phase of the module, the teacher candidates engaged in reading journal and newspaper articles, listening to an episode of The Daily and in using the data sets to investigate relationships between inequity, segregation, school funding, and achievement. During the second session, the candidates' inquiries focused on identifying key variables to explore, supported by a modeled demonstration of how to navigate data sets and data analysis. Through these analyses, the candidates found that about 51% of the nation's school students are in "racially concentrated districts," where over 75 percent of students are either white or nonwhite. They also found that both NYC schools, as well as many of the schools where they were placed for field experience mirrored these patterns of racial concentration and segregation. Data overwhelmingly supported the conclusions drawn by the group: Black and Hispanic students are several times more likely to get suspended than White students. White students have significantly higher chances of being in advanced placement classes than Black or Hispanic students. Schools with higher percentages of Black and Hispanic students are likely to have lower graduation rates. The candidates wondered whether the "achievement gap" varied significantly between urban and rural contexts, and how these differences might have been exacerbated during and after the COVID-19 pandemic. One candidate shared insights gained from watching a PBS Frontline documentary (Growing Up Poor in America, FRONTLINE, noting that while New York State was able to provide students with laptops and tablets during remote learning, students in Alabama had to travel to their schools to use computers on-site or access wifi. She concluded that lack of access to basic resources directly impacts educational quality, and students outcomes, perpetuating cycles of disadvantage. Candidates also posed their own questions and used data analysis to investigate and answer them. Below is one example:

Question: Is there a correlation between the % of white students in a district and students' proficiency ELA and Math scores?

Data will be used for School districts 1-5.

Table For School Districts 1-5

	D1	D2	D3	D4	D5
White	15.7%	12.8%	26.3%	4.2%	7.3%
ELA Level 4	26.9%	22.3%	35.5%	24.5%	20.7%
Math Level 4	15.7%	11.7%	15.4%	15.1	8.5%

Source: NYC Open data

There is a strong correlation between ELA proficiency level 4 and % of White students in a school district, and a weak correlation between Math proficiency level 4 and % of White students.

Figure 4: Excerpt from teacher candidate D's work on posing and answering their own question

Next, the teacher candidates explored the learning opportunities offered across various schools and districts, and in order to determine whether NYC students received the same quality of education. Based on their findings, the candidates determined that schools with predominantly Black and Latinx student populations often offered fewer Advanced Placement (AP) classes, had lower enrollment rate in AP classes, and had higher percentages of inexperienced or frequently absent teachers, as well and higher numbers of total days missed due to out-of-school suspensions. Two candidates found that their school offered no AP Mathematics classes, and that students' college readiness rate (measured by percentage of students graduating with test scores high enough to enroll at CUNY without needing remedial coursework) was well below the city average.
As part of their inquiry into educational equity, the candidates were also tasked to find out how much less total funding is allocated to school districts serving predominantly students of color compared to school districts serving predominantly white students at national level, and how these national findings compare to those in NYC. To do so, candidates researched their local school districts' budgets, using public records or local media such as newspapers or television reporting, and with determining the budget per student.

Candidates were also guided to summarize their findings by comparing data on local school budget, available learning opportunities, and teacher qualifications, with their prior research on segregation and student achievements. The goal of this comparative analysis was to determine whether data suggested correlation between some variables, such as the relationship between funding levels and educational outcomes. During class discussions, the candidates shared their reflections and key findings. Candidates commented that they were shocked to read that the differences in annual funding between nonwhite and white school districts nationwide amounts to 23 billion (EdBuild, 2019), which translates to a difference of $2300 per student.

One candidate, who explored the EdBuild database (2019), found that nonwhite school districts in New York State receive, on average, 10% less funding than predominantly white school districts. Additionally, she found that high-poverty white school districts receive about $4000 more per student compared to high-poverty nonwhite school districts. The visualization below, created

by this candidate, illustrates these disparities, and initiated a deeper conversation about structural inequities in school funding.

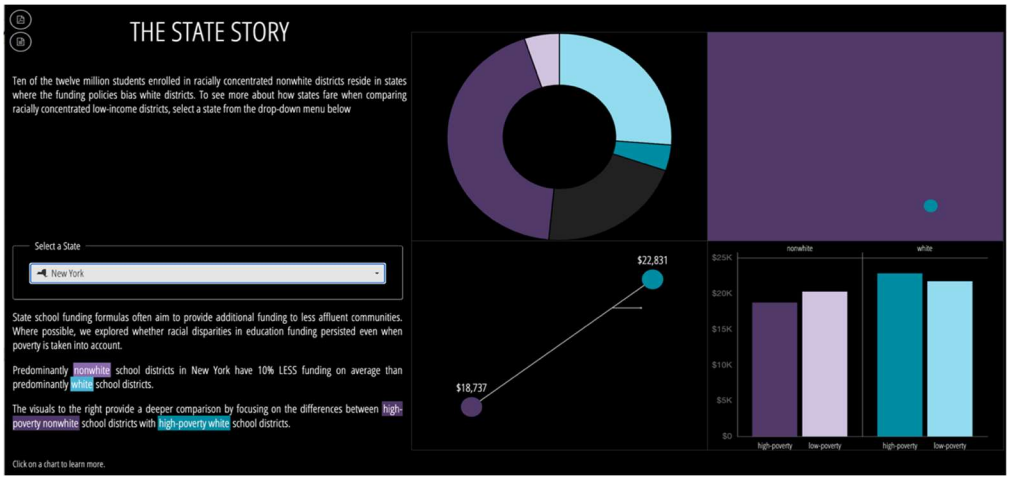
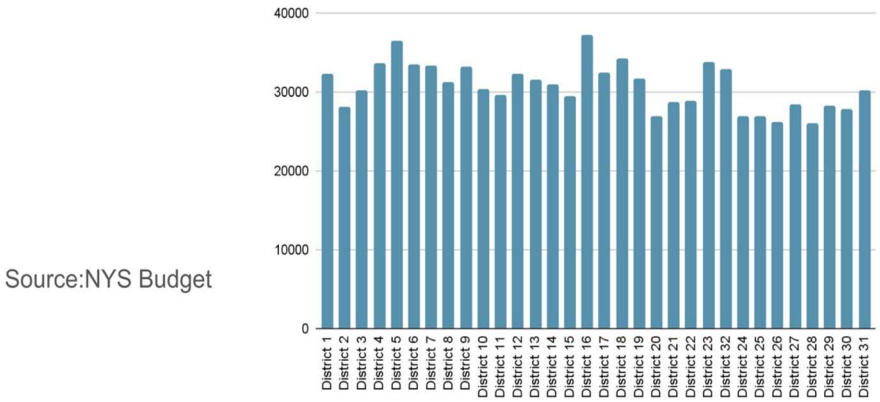


Figure 5: Teacher candidate E's diagrams of high-poverty white and non-white school districts funding

Another teacher candidate shared his findings on the budget allocations within New York City school districts.



(https://openbudget.ny.gov/schoolFundingTransparency.html)

There is a large difference in funding between NYC school districts 5 and 16, for example, and districts 26 and 28.

Figure 6: Excerpt from teacher candidate F's visualization of NYC school district budgets

Additionally, the candidates were also asked to formulate their own inquiry questions, explore

the relevant data sets to investigate those questions, and then share their findings with the group. Close exploration of the data sets and analysis of relationships between variables such as race, teacher quality, suspensions, and educational opportunities such as AP courses, Computer Science Courses, and Gifted and Talented Programs, led the candidates to conclude that there is a clear correlation between race and educational opportunities.

Engagement with Coding and Data Analysis With R

The final session of the module was structured as an immersive workshop led by the second author. The purpose of this session was to introduce teacher candidates to R, a programming language for statistical computing and data analysis, allowing them to engage with data analysis through a different lens. This experience provided teacher candidates with new tools to engage with data analysis enabling them to revisit and further investigate some of the questions that the candidates had previously explored. The second author introduced candidates to R Studio. He guided candidates through the process of downloading R and RStudio, creating a free account, and setting up their workspaces. Candidates were then introduced to the NYC Department of Education (DOE) data available through the NYC Open Data portal. The session continued with instruction on using the Tidyverse—a collection of R packages designed for data manipulation and visualization. Through hands-on practice, candidates learned to sort, filter, and analyze large datasets and to generate visual representations of trends and relationships relevant to their research questions.

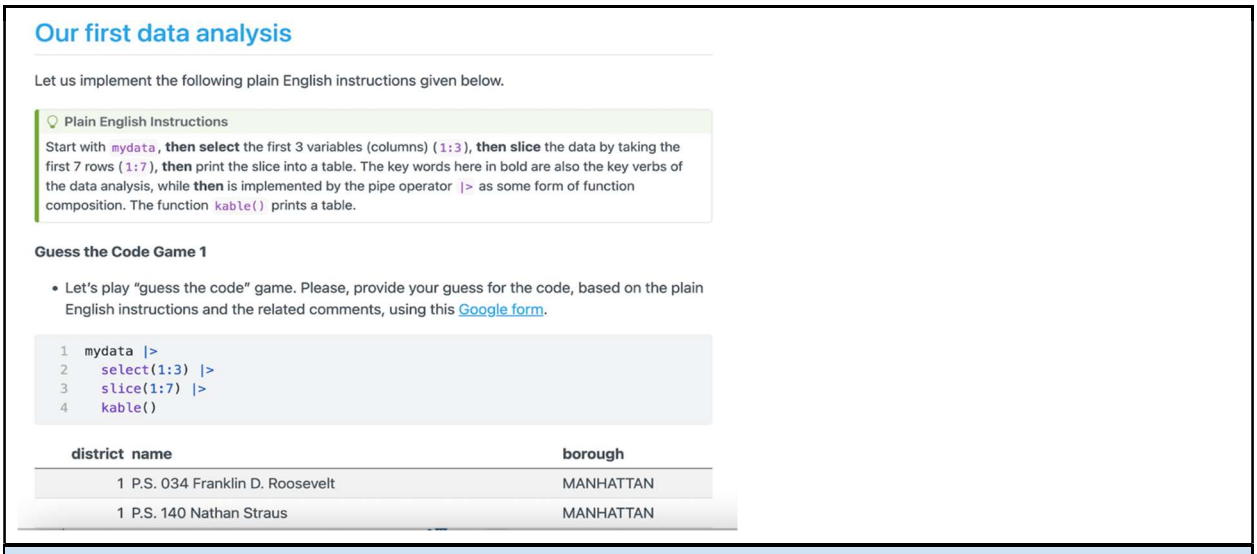


Figure 7: Excerpt from a collaborative "guess the code" exercise with R.

Next, the second author guided the candidates about filtering data.

Figure 8: Excerpt from a collaborative filtering data exercise with R.

He then guided the candidates in exploratory data analysis to basic model fitting. The question that was explored was "Is there a relationship between student ELA proficiency and math proficiency?"

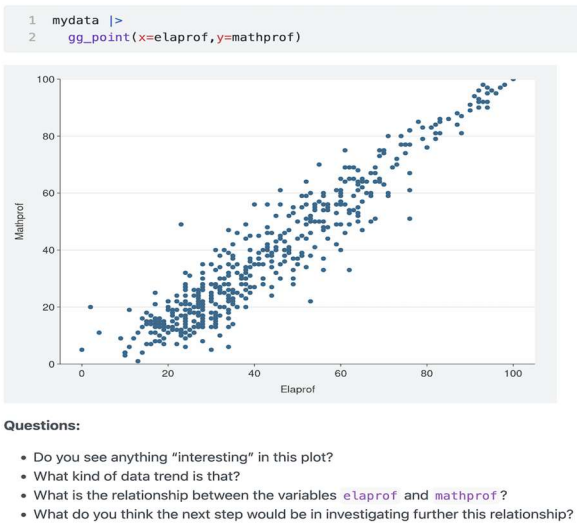


Figure 9: Excerpt from collaborative exploratory data analysis to model fitting

Finally, the candidates were guided in writing a code, which functioned to select a specific number of schools from the dataset and to create a chart documenting the percentage of students with math proficiency. The chart provided a visual representation of school-level variation in student achievement, demonstrating the importance of data visualization in identifying patterns and disparities within educational outcomes.

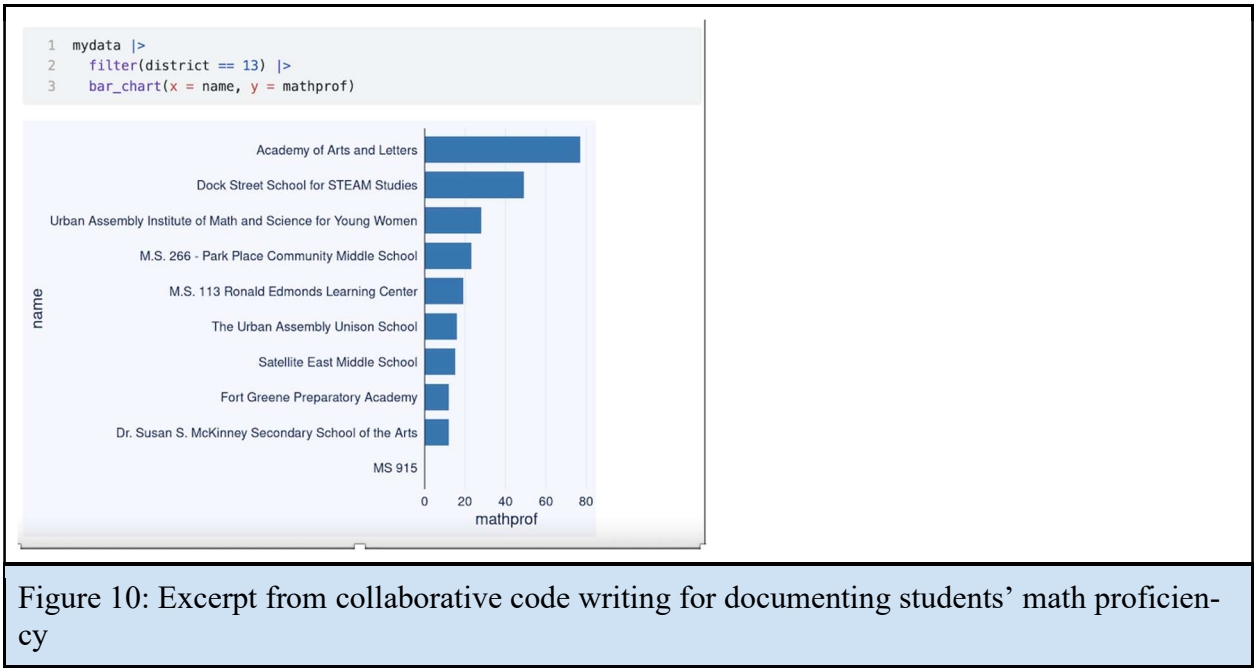


Figure 10: Excerpt from collaborative code writing for documenting students' math proficiency

Discussing "for" and "against" Uses of Big Data in Education

In the final phase of the module, the first author prompted the teacher candidates to reflect on the uses and limitations of technology and big data in education decision making. The discussion encouraged critical reflection on the limitations of data, and specifically on what data can and cannot capture, particularly in relation to students, schools, and communities. The candidates observed that important aspects of schools and student experiences are often absent from data sets, as the data reflects only information that the designers choose to collect. Because datasets are shaped by the interests, priorities, and assumptions of their creators, they may omit elements that others would consider essential. For example, one candidate expressed frustration that he was unable to find any data on extracurricular activities in the middle schools he was reviewing and comparing. Despite their significance in shaping school culture and student engagement, such qualitative features were not included in the available datasets.

The candidates also interrogated the assumptions embedded in data collection and visualization practices. Candidates observed that comparing schools based solely on quantitative data assumes that the schools are somewhat equivalent in relation to all "invisible dimensions," such as school climate, leadership, community engagement, or teacher expectations. They also noted that in the "standardization" of individuals in a data set, all featured students and schools lose their unique

characteristics, reducing them to variables that can be measured, ranked and compared. Qualitative features such as school culture or teacher expectations are difficult if not impossible to quantify and cannot be captured in databases. Overall, candidates thought that data analysis was helpful and potentially advantageous but that it had to be used with care, and that the potential biases and limitations of datasets and data analysis needed to be kept in mind. They concluded that while data can be a powerful tool for equity-oriented decision-making, the importance of remaining mindful of biases, oversimplifications, and omissions cannot be overstated. Data must be interpreted critically and complemented by qualitative insights to account for the complexities and nuances of educational systems.

From Data to Critical Inquiry: Reflecting on Segregation, Inequity, and the Future of Public Education

Overall, the teacher candidates reported that combining data analysis with readings and collaborative discussions deepened their understanding of the schools where they were placed for field experience, as well as of the students in those schools. They appreciated the opportunity to compare their placement schools to others in New York City and across the nation. Many noted that they had developed practical skills for examining school data and were able to uncover important, and sometimes surprising, information. One candidate shared: *"I found [the data analysis] useful, because you can see all this bunch of data, but until you organize it you won't be able to understand it really. I found more information about the school I was in."*

Another reflected on his current school placement: *"What I found [about the school I'm at now] was that the majority of the teachers were not certified. I was really surprised—as they've been working there for years and still didn't have certification. And then I looked at how students were doing and compared that school with my previous placement, which was the second in District xxx [with the second-highest math proficiency level]. The current school—everyone says it's good—had very low math and ELL proficiency levels. But interestingly, the data showed its funding was higher than most schools in the neighborhood."*

In some cases, candidates reported that these new analytical skills influenced not only their development as educators but also their roles as parents or community members. One candidate shared: *"I have three kids in District x and I wanted to find more about the schools there… My daughter was going to middle school, and I wanted to know what kind of school she was going to. I learned a lot by looking at the data. I saw how wealthy the parents [in one school] were compared to the other two schools. I was shocked by the different economic level of the parents… and the school is whiter."*

The teacher candidates largely agreed that the data analysis activities revealed patterns of school inequity and segregation, as well as the negative impact these have on students from high-poverty or minoritized communities. As one candidate remarked: *"…for me, it helped me understand and see correlations as to why some schools have students who are struggling with math*

and others don't. It seems that it has a lot to do with literacy rates—the higher the literacy, the higher the math proficiency. I also found that schools with more white or Asian students tend to perform better than schools with more Latinx and Black students."

Through data exploration, candidates discovered wide disparities in access to advanced coursework—particularly in Precalculus, Calculus, and AP courses—as well as extracurricular opportunities. These disparities correlated strongly with differences in school funding, teacher certification, student demographics, and geographic location. One teacher candidate concluded: *"I never believed there were such great differences between schools in NYS and the country. But with the data, I really had the chance to go inside it, and the data showed that equal opportunity in education doesn't exist."*

Other candidates noted how the process helped challenge long-held assumptions: *"I assumed those elite schools had more funding per student because students were performing better—but I was wrong. That isn't true. Maybe it was true back then, but when I looked at the data, they were getting $11K per student, whereas some of my neighborhood schools are getting around $30K per student."*

The close data analysis conducted by the teacher candidates, along with the various empirical questions they pursued, laid the groundwork for a broader and more critical inquiry around more general and philosophical questions such as: What repercussions do school segregation and educational inequity have for students in society? What are potential remedies for addressing school segregation and inequity? What is fair school funding and why does it matter?

Teacher candidates concluded that school segregation nowadays is driven by systemic inequality, and existing residential segregation in NYC (and the US) results in schools divided by race and socioeconomic status. The candidates' analysis led to the conclusions that segregation and inequity have a profound effect on minority students' learning outcomes and thus act to limit their future educational and job opportunities. The teacher candidates were asked to make an argument for what should— or should not —be done to make New York City's public schools more diverse and inclusive. Below are excerpts from several of their reflections.

Teacher Candidate C: *"If I were a mayor and had the power to change the NYC public education system of NYC, I would get rid of gifted & talented programs and elite high school programs. Those special schools would accept students without testing and would accept students that represent the community. If the community is diverse, the school should accept students whose composition mirrors the community."*

Teacher Candidate D: "*One thing that I would like to see change is the way schools are funded. If a school is located in a wealthy neighborhood or where the taxes are high, that school would have a better chance to receive more funds than a school in a poor neighborhood. Therefore, the first school will be more likely to recruit qualified and experienced teachers and have more*

school supplies. Moreover, the same school will have more resources such as a bigger library and better equipped computer labs. The current process of funding schools is outdated, not equitable and perpetuates school segregation. The government must reformulate the way schools are being funded, for example by using entirely federal funding. It is imperative for schools from low-income neighborhoods to receive sufficient funds in order to provide the necessary resources and supplies needed."

Teacher Candidate H: "*New York City should make changes to ensure that every school accepts at least 15 percent of the students from every major ethnic group in the school neighborhood. This law should be applied to every elementary, middle, and high school to make sure students from different races and backgrounds are not excluded and get access to sufficient resources. Usually, students who aren't native English speakers struggle due to language barriers, therefore every school should also ensure additional support for those students. Furthermore, I also think that as we promote schools' diversity in NYC, we need more teacher diversity. We need to recruit teachers from different cultural backgrounds, whose presence may be important for students to build relationships and to be more successful*."

Mathematics Teacher Candidates Perceptions of the Potential of Computing and Data Analysis in Mathematics Teaching and Learning

Based on a focus group interview conducted after the completion of the course, we found that the teacher candidates' perceptions of the potential of computing and data analysis in mathematics teaching and learning, although overall positive, were nuanced and mixed. Several candidates thought that using real-world data sets in mathematics instruction could promote new and useful skills among students, equipping them to independently engage with data analysis in the future. A teacher candidate expressed the view that data analysis can be used by teachers to learn more about the student population one teaches. He noted: "…you can use data to learn more about the students [in your school] and understand better where they are coming from." Several candidates viewed data analysis as a promising pedagogical tool, capable of making math activities more relevant to students, and fostering engaging and meaningful discussions. They emphasized that integrating data-driven tasks could enhance mathematics lesson by connecting them to real-life issues-- such as poverty rates, student loans, and rising climatic temperature – and by helping students make sense of the world around them. It promises to provide relevance to students' lives, and may act to increase their motivation. One candidate commented: "One can bring social justice issues to the classroom, for example collaborate with the social study teacher and maybe they can do a study of neighborhoods or comparing schools." Another noted the potential for data to highlight inequities in STEM fields, stating: "[Data sets and analyses] could be used showing… because woman predominantly make up the smallest number of engineers, and Latinx women make the smallest number of women in STEM."

Several teacher candidates believed that engaging students in data analysis can introduce social justice issues, and can be empowering and lead to activism in their communities. As one re-

marked: "it [data analysis] helps students to know more about the world around them and can lead them to get involved in advocacy work." The integration of computing and data analysis in the mathematics classroom was viewed not juts as a technical skill, but as a means to of developing a critical stance and provide a context for engaging with important social issues. However, some concerns were also expressed. One candidate voiced hesitation about addressing social justice issues through data analysis, remarking: "I don't think I'll use this specifically; I heard that there is push back in the schools if the lessons are too political." Another reflected on the practical challenges of implementing data-driven lessons, including the time and effort to curate data sets, additional preparation to adapt lessons, as well as extra time needed for students working with data sets and analysis. The perceptions of most candidates were that using data analysis can enhance mathematics lessons and be worthwhile, but it may not be realistic in the first one or two years of teaching. In sum, candidates recognized the pedagogical and societal potential of computing and data analysis in mathematics education, particularly in promoting relevance and critical engagement. However, their reflections also highlighted the need for adequate support, time, and teaching experience in order to fully realize this potential in classroom practice.

DISCUSSION

Overall, out team found that the course module effectively engaged teacher candidates in data exploration and analyses related to NYC schools – particularly the schools where they were placed for field experience. The module succeeded in fostering critical discussions and equity-oriented inquiry through data. However, our reflections also pointed to several areas for improvement in future iterations. First, we observed that some candidates struggled with posing their own research questions and identifying the type of questions that could be meaningfully addressed though data analysis. More scaffolding is necessary in that regard in the next iteration of the module implementation. In general, we recommend that collectively exploration of data sets precede individual data analysis. In the next iteration of the module implementation, instructors should model question posing and data analysis before assigning individual assignments. This shared experience can build candidates' confidence and understanding before moving into individual or small-group work.

Second, we found that introducing R and RStudio required more time and support than the course schedule allowed. Thus, we think it would be more appropriate that a next iteration of the teaching module use CODAP for data analysis rather than R. Some module activities have been modified to utilize CODAP. Our experimentation shows that R studio would need more time to introduce, and that more than one course might be necessary for teacher candidates to begin comfortably using it for data analysis and computational modeling. While R is a powerful tool for data analysis, we found that it may not be the most accessible platform for all students—particularly within the limited timeframe of a multi-week module. Based on our experience with the first iteration of the module, we recommend using CODAP (Common Online Data Analysis Platform) for future implementations. CODAP provides a more user-friendly and intuitive environment for beginners, while still supporting meaningful engagement with data analysis.

Activity 4 of the course module has already been revised to incorporate CODAP (see Appendix for details), although the adapted version is yet to be implemented in the next module cycle. This CODAP-based activity uses NYC School Open Data to help teacher candidates visually and analytically explore educational issues. Integrated into a math education course, it promises to help teacher candidates builds skills in data analysis and visualization, raise awareness of equity gaps, and prepares candidates to apply data-driven approaches in their teaching. We believe this adaptation will enable teacher candidates to engage more deeply in independent inquiry and data interpretation by reducing the technical barriers associated with learning R.

CONCLUSIONS

This study examined the implementation of a course module which integrated computing and data analyses into a mathematics education course. Designed for prospective mathematics teachers, the module engaged candidates in exploring datasets to better understand their field placement contexts, and support them in developing data analysis and critical thinking skills. Candidates engaged in topics related to student learning opportunities, achievement gaps, school segregation, and broader issues of education (in)equity and social justice. The module emphasized a commitment to equity, and encouraged deep explorations of these topics through texts, data analysis, collaborative discussions, and direct experience. By sharing and reflecting on their findings, candidates developed a deeper understanding of the principles and uses of data analysis in addressing issues of social justice, and of its potential role in the mathematics classroom.

ACKNOWLEDGMENTS

This work was supported by The City University of New York (CUNY) Computing Integrated Teacher Education (CITE) grant # 7W 194-12 01 A/2023.

REFERENCES

[1] Bain, C. P., Anton, G., Horn, M., & Wilensky, U. (2021). From "Authentic Tools" to Authenticity: Using CT to Enable Discovery in Statistics Classrooms. In E. de Vries, Y. Hod, & J. Ahn (Eds.), *Proceedings of the International Society of the Learning Sciences (ICLS 2021)* (pp. 1115–1116). International Society of the Learning Sciences. https://repository.isls.org/bitstream/1/7425/1/1115-1116.pdf

[2] Barr, V., & Stephenson, C. (2011). Bringing computational thinking to K-12: What is involved and what is the role of the computer science community? *ACM Inroads*, *2*(1), 48–54.

[3] Benakli, N. et al. (2017). Introducing Computational Thinking through Hands-on Projects Using R with Applications to Calculus, Probability and Data Analysis, *International Journal of Mathematical Education in Science and Technology, 48*(3), 393–427, doi: 10.1080/0020739X.2016.1254296.

[4] Kennedy, N. S., Kostadinov, B., Masuda, A. (2024). Estimating the Area of a Lung Nodule from an X-ray Image. In E. Donovan, L. Hoots, & L. Wiglesworth (Eds.), *Cross-Curricular Applications for Pure Mathematics Courses* (pp. 219-232). Mathematical Association of America (MAA).

[5] Kennedy, N. S. (2023). *Behind the data: Analysis and visualization of school and student data.* CITE Learning https://citelearning.commons.gc.cuny.edu/drs-kennedy-kostadinov-masuda/

[6] Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., ... & Werner, L. (2020). Computational thinking from a disciplinary perspective: Integrating Computational Thinking in K-12 Science, Technology, Engineering, and Mathematics Education. *Journal of Science Education and Technology*, *29*(1), 1–8, doi: 10.1007/s10956-019-09803-w.

[7] Nouri, J., Zhang, L., Mannila, L. & Noren, E. (2020). Development of computational thinking, digital competence and 21st century skills when learning programming in K-9, *Education Inquiry*, *11*(1), 1-17.

[8] *NYS Culturally-Responsive Sustaining Education Framework*. (2024). Retrieved from https://www.nysed.gov/sites/default/files/programs/crs/culturally-responsive-sustaining-education-framework.pdf

[9] The CUNY CITE Equity Working Group. (2023). *Equitable CITE Pedagogy: Putting it into Praxis.* https://cuny.manifoldapp.org/read/equitable-cite-pedagogy-putting-it-into-praxis/section/45bc24dc-6adb-49a9-96b8-25dfd185e37a

[10] *The New York State K12 Computer Science and Digital Fluency Learning Standards.* (2024). Retrieved from https://www.nysed.gov/curriculum-instruction/computer-science-and-digital-fluency-learning-standards

[11] Vogel, S., Patel, A., & Yadav, A. (2025, May). *CITE Framework*. CITE Learning. CUNY.

[12] Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms, *Journal of Science Education and Technology*, *25*(1), 127-147, https://doi.org/10.1007/s10956-015-9581-5

[13] Wing, J.M. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33-35.

[14] Wing, J.M. (2008), Computational thinking and thinking about computing. *Philosophical Transactions of the Royal Society*, 366, 3717–3725, doi:10.1098/rsta.2008.0118

[15] Yadav, A., Stephenson, C., & Hong, H. (2017). *Computational thinking for teacher education.* Communications of the ACM, *60*(4), 55-62. https://doi.org/10.1145/2994591

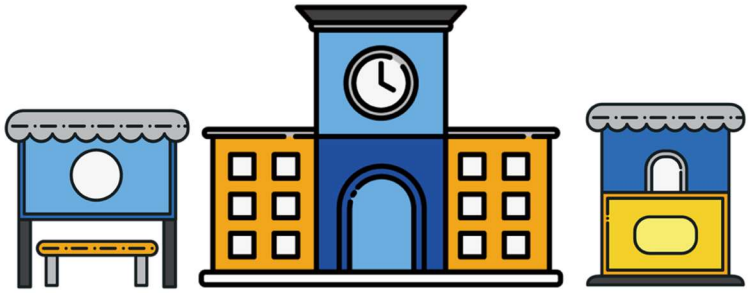
[16] Yin, R. K. (2017). *Case study research and applications: Design and methods*. Sage.

APPENDIX

Course Module: Behind the Data: Analysis and Visualization of School and Student Data

Activity 1: Warm-Up: Visualize Data

Visual Portrait of Your School: Learn about Your Field Experience School and School District



Source: Pixabay

Search through the NYC, NYS and national education data bases below and search for information about your school and about your school district.

Additionally:
Scroll down to the interactive map of the United States in the Miseducation (ProPublica) database and then answer the following questions:
1. Click the tabs "Opportunity," "Discipline," "Segregation" and "Achievement Gap" and answer these two questions: What do you notice? What do you wonder about?
2. Next, click the tabs "Black" and "Hispanic." What do you notice? What do you wonder?
3. Search for your school or district in the database. What do you notice in the results? What questions do you have?

Explore the Civil Rights Data Collection and check:
- School and district search
- Detailed data tables
- Data analytic tools
- Special reports for schools and districts (English Language Learner report, discipline report, education equity report, state and national estimations)

Assignment:
Prepare a report (minimum 1 page) on what you have found about your school and about your school district. In your reflection, include detailed results, e.g., "School district ….'s composition is 24 % Black, 41%, Hispanic, 16% White, 17% Asian, 1% Native American, 1% Two or more races." Or "White students are 1.7 times more likely to be enrolled in at least one AP class as Black students." Include graphs and tables from the database. Be ready to share your findings in class.

Possible Data Sources to Use

Key National Education Data Sources

Data Source	What It Covers	Why It's Useful	Link
Common Core of Data (CCD) – NCES	Public elementary & secondary schools & districts: enrollment, staff, school types, fiscal data. NCES	Good for longitudinal and cross-state comparisons of K-12 schooling.	CCD, NCES NCES
EDFacts	PreK-12 data submitted by state education agencies: demographics, performance & assessments, program participation, funding/grants data. U.S. Department of Education+2U.S. Department of Education+2	Centralized source for many metrics at school and district level. Useful for policy analysis, funding, accountability.	EDFacts Initiative U.S. Department of Education
NCES / Administrative Data Collections	NCES collects non-fiscal and fiscal data about K-12 via CCD; other state/local K-12 and postsecondary data as well. NCES+1	For basic data: student counts, school/district descriptors, finances, staffing.	NCES Administrative Data Collections NCES
National Assessment of Educational Progress (NAEP)	Nationally representative assessments in reading, math, science etc.; long-term trends in student achievement.	Strong source for understanding achievement over time, comparing states.	Available via the U.S. Department of Education / NCES "Nation's Report Card" portal. (via ED Data / NCES) U.S. Department of Education
Open Data Platform / U.S. Department of Education	Multiple datasets: civil rights data, school finance, performance, student outcomes, etc. U.S. Department of Education+1	Good starting place for locating various national-level datasets in one place.	ED's Open Data Platform U.S. Department of Education
Office for Civil Rights' Data on Equal Access to Education	The Civil Rights Data Collection (CRDC) is a mandatory survey conducted by the U.S. Department of Education's Office for Civil Rights. It covers public schools (pre-K through grade 12) in all 50 states, Washington DC, and Puerto Rico.	It is biennial (collected every two years), gathering information on a wide array of indicators relevant to civil rights in education — whether students of different races, sexes, abilities, etc., get equal educational opportunities.	Civil Rights Data Collection (CRDC)
Miseducation (ProPublica)	Covers nearly all public and charter schools in the U.S., it allows comparisons across and within states, districts, and schools.	Focuses on equity and racial disparities	Miseducation (ProPublica)

New York State (NYSED) Data Sources

Data Source	Link	What It Offers
NYSED Data Site	data.nysed.gov New York State Education Department Data	Public reporting of educational data for NY: student & school data, etc. New York State Education Department Data
School & District Account-	NYSED Accountability & ESSA data	State accountability data under ESSA: perfor-

Data Source	Link	What It Offers
ability Resources and Data (ESSA)	New York State Education Department	mance, reporting, etc. New York State Education Department
Basic Education Data System (BEDS)	NYSED / IRS / BEDS New York State Education Department	Institutional school/district master data, enrollment, staffing, etc.

NYC / Local NYC Public Schools Data Sources

Data Source	Link	What It Offers
School Quality Snapshots	NYC School Quality Snapshots via NYC Schools website	Summary reports of school environment & student achievement, etc.
NYC School Survey	NYC School Survey web	Feedback from families, students, and staff on school environment, etc.
NYC InfoHub – School Quality Reports & Resources	NYC School Quality Reports & Resources InfoHub	Multi-year school performance data, visualizations; guides; downloadable data. InfoHub
NYC Insights & Reporting / Data at a Glance	NYC DOE Insights & Reporting web	Snapshot summaries, reports on student population, etc.
NYC Open Data	NYC Open Data website	Free public data published by NYC agencies and partners

In-Class

Gallery Walk

Teacher candidates share their findings.

Discussion Questions:

- What did you find about your school/school district?
- Do all students in NYC receive the same quality of education?
- Do all students in America receive the same quality of education?
- Do you think that there is a correlation between students' race and the quality of education they receive?
- What is the purpose of public education?
- Is receiving a quality public education a right (for everyone) or a privilege (for some)

Activity 2: (In)equitable School Funding

Investigate relationships between school funding, inequity, and achievement

Source: Pixabay

Readings for the session:

EdBuild (2019). *23 Billion*. https://edbuild.org/content/23-billion

Lombardo. C. (2019). *Why white school districts have so much more money*. https://www.npr.org/2019/02/26/696794821/why-white-school-districts-have-so-much-more-money.

Mervosh, S. (2019). How much wealthier are white school districts than nonwhite ones? $23 Billion, reports. *New York Times*, https://www.nytimes.com/2019/02/27/education/school-districts-funding-white-minorities.html

Assignment:

Choose one of the following ideas (or generate another) to investigate the interrelationship among school segregation, funding and inequality.

1. Research your local school district budget, using public records or local media, such as newspapers or television reporting. What is the budget per student? How does that budget compare with the state average? The national average?
2. Compare your findings about your local school budget to your research about segregation and student achievements, using the Miseducation database. Do the results of your research suggest any correlations?
3. Prepare a report and share with peers.

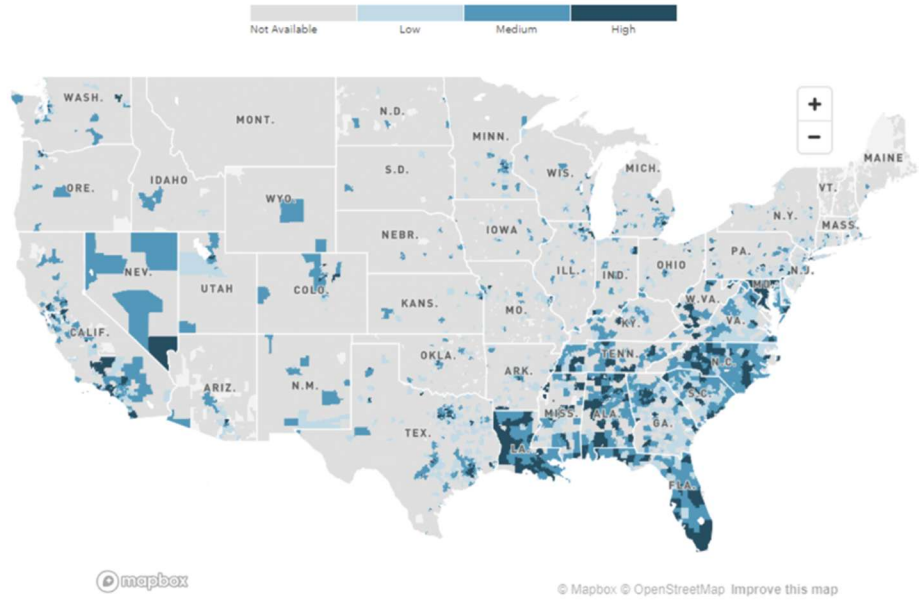
Discussion Questions:

1. How much less total funding do school districts that serve predominantly students of color receive compared to school districts that serve predominantly white students?
2. Why are school district borders problematic?
3. How many of the nation's schoolchildren are in "racially concentrated districts, where over 75

percent of students are either white or nonwhite"?
4. How much less money, on average, do nonwhite districts receive than white districts?
5. How are school districts funded?
6. How does lack of school funding affect classrooms?
7. What is fair school funding and why does it matter?

Activity 3: Segregation and Educational (In)equities

Investigate the relationship between segregation, educational opportunities, and (in)equities



Using the original interactive map on ProPublica, you can search for individual school districts and change data categories by race and measure.

Image from ProPublica

Readings for the session

Readings: NYC Schools (In)Equity
Read at least two of the articles below:

Lefty, L. (2021). *The long fight for educational equity in NYC*. Retrieved from https://www.mcny.org/story/long-fight-educational-equity-nyc on August 1, 2022.

Futterman, K. (2020). *Education and equality in NYC: The problem at hand*. Retrieved from https://polygonnews.org/2299/showcase/education-and-equality-in-nyc-the-problem-at-hand/ on August 7, 2022.

Robinson, G. (2013). *Class in the classroom: The income gap and NYC's schools*. https://citylimits.org/2013/09/25/class-in-the-classroom-the-income-gap-and-nycs-schools/ on August 1, 2022.

Readings: School Segregation (In NYC, America and Abroad)

Read at least one of the articles below:

Gould, J. (2021). *New York's schools are still the most segregated in the nation: Report*. Retrieved from https://gothamist.com/news/new-yorks-schools-are-still-the-most-segregated-in-the-nation-report on July 25, 2022.

Hannah-Jones, H. (2016). Choosing a school for my daughter in a segregated city, *New York Times Magazine*, https://www.nytimes.com/2016/06/12/magazine/choosing-a-school-for-my-daughter-in-a-segregated-city.html

Shapiro, E. (2019). Facing segregated schools, parents took integration into their own hands. It's working, *New York Times, https://www.nytimes.com/2019/04/16/nyregion/new-york-city-school-segregation.html*

Shapiro, E. (2021). Lawsuit challenging N.Y.C. school segregation targets gifted programs, *New York Times*, https://www.nytimes.com/2021/03/09/nyregion/nyc-schools-segregation-lawsuit.html

School Segregation in Other Countries

Read at least one of the articles below:

Surk, B. (2018). In a divided Bosnia, segregated schools persist, *The New York Times*, https://www.nytimes.com/2018/12/01/world/europe/bosnia-schools-segregated-ethnic.html

Maragkidou, M. (2015). Segregation, bullying and fear: The stunted education of Romani children in Europe. https://www.amnesty.org/en/latest/news/2015/04/the-stunted-education-of-romani-children-in-europe/

Assignment:

1. Only a tiny number of black students are admitted to the highly selective public high schools in New York City (e.g., 2019, 2020, 2021) raising the pressure on officials to confront the decades-old challenge of integrating New York's elite public schools. To learn more about this story, listen to this episode of The Daily. For more information, read this essay offering different perspectives on the problem and possible solutions.

Make a case for what should be done — or not done — to make New York's elite public schools more diverse.

2. Pose a question in relation to NYC school segregation and inequities that you would like to answer. Use NYC Open database https://opendata.cityofnewyork.us/ to find an answer to your question. Be ready to share in class your question, results of your research and visual representations.

Gallery Walk
Share your findings.

Discussion Questions:

- How and why are schools still segregated in 2022 (in NYC, America and the world)?
- What repercussions do segregated schools have for students and society?
- What are potential remedies to address school segregation?

Activity 4: Using R for Data Analysis and Discussions of Data Bias

Free image from Pixabay

Readings for the session:

Haagort, M. (2021). *Every database is biased*. https://becominghuman.ai/every-database-is-biased-6a402224b8a9.

R- Workshop Structure

The workshop is framed as a hands-on introduction to exploratory data analysis (EDA) using R and RStudio (now Posit). It is organized into the following stages:

1. Introduction – Setting goals (working with NYC DOE middle school data) and setting up R/RStudio/Posit.
2. Loading the Data – Importing via API with RSocrata or via CSV download with readr.
3. Preparing the Data – Selecting relevant variables from a large dataset (327 → 18 variables).
4. Exploratory Data Analysis – Using Tidyverse pipelines (dplyr, ggblanket, ggcharts) for filtering, slicing, selecting, and visualizing.
5. Progressive Practice ("Guess the Code" Games) – Translating plain English instructions into R code.
6. Modeling – Moving from scatterplots to regression (line of best fit).
7. Advanced Analysis – Sorting, mutating variable types, filtering, and producing ranked tables and bar charts.

The Data

The data comes from the 2021 DOE Middle School Directory (NYC Open Data Portal).

Size**:** 474 rows (schools) × 327 columns (variables).

Selected Variables (18 used): district, name, borough, latitude, longitude, diversityinadmissions, coursepassrate, elaprof (ELA proficiency), mathprof (Math proficiency), surveysafety, totalstudents, gradespan, acceleratedclasses, electiveclasses, languageclasses, tophs1, tophs2, tophs3.

Questions Explored

Participants investigate progressively deeper questions such as:

Basic exploration

- What are the first few schools in the dataset?
- How do district and borough variables look in small slices of data?

Filtering

- What happens when you plots Math vs. ELA proficiency with and without filtering by district?
- What trends or relationships exist between ELA and Math proficiency?

Exploring Relationships and Modeling

- Is there a linear relationship between Math and ELA proficiency?
- How do we calculate the center of mass and fit a regression line?
- How can the regression line be used to make predictions?

Advanced exploration

- Which schools have the highest math proficiency scores?
- Which schools have the lowest survey safety ratings?

- How do results compare across boroughs and districts?
- What does filtering on surveysafety or mathprof reveal about specific districts (e.g., District 13)?
- Which middle schools send the most of their students to the specialized math and science high schools in?

Coding Approach

Packages used: RSocrata, tidyverse, ggcharts, ggblanket, knitr

Workflow:

Import → Select → Filter → Summarize → Arrange → Slice → Visualize.

Heavy use of *pipes* **(|>)** for chaining steps in plain English style.

Type handling with mutate(as.numeric()) when variables import as characters.

Pedagogical technique: "Guess the Code" games encourage learners to translate natural-language tasks into R pipelines.

Visualization Techniques

Tabular outputs: knitr::kable() for clean tables.

Scatterplots: gg_point() to explore relationships (ELA vs. Math).

Regression fits: gg_smooth(method="lm") with overlays (line of best fit, center of mass)

Bar charts: bar_chart() for safety and proficiency comparisons across schools.

Data slicing: visual confirmation of subsets (e.g., top-performing schools, low-safety schools).

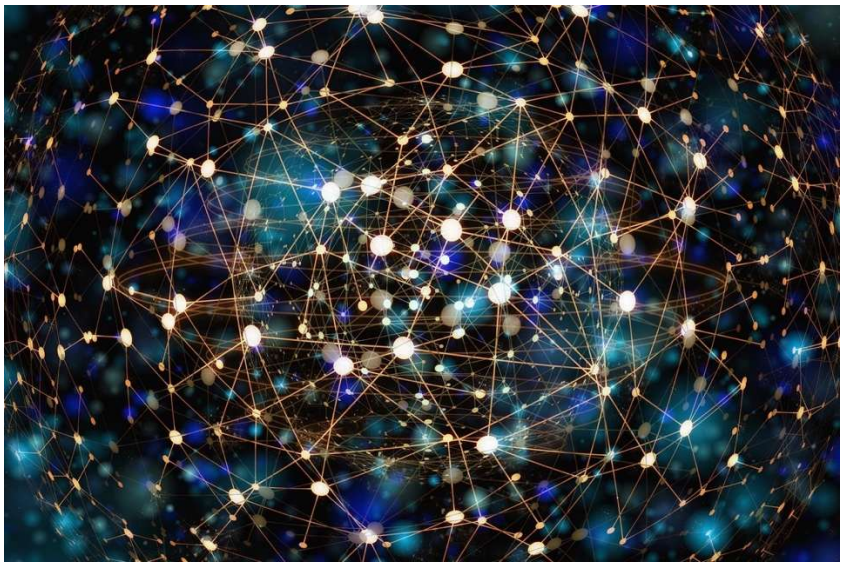
Overall, the workshop combines data literacy (interpreting variables), coding practice in R (tidyverse pipelines), and visual reasoning (plots + regression), and is structured as an engaging exploratory activity on NYC school data for mathematics teacher candidates.

Discussion Questions:

1. What is bias?
2. Could a database be biased? Explain.
3. If so, what might be some sources of bias?
4. What can introduce bias in data representation and interpretation?
5. What can guard against bias?
6. In what ways data technologies/ data analytical tools can be helpful?
7. Are there ways in which data technologies/ data analytical tools be harmful?
8. What is the impact of data technologies on society and individuals?

Activity 5: Using CODAP for Data Visualization and Analysis with NYC Open database

(May be used as an addition or alternative to Activity 4)



Free image from Pixabay

CODAP Activity

Explore this interactive CODAP activity to analyze mathematics achievement data from New York City schools. Investigate patterns, compare neighborhoods, and uncover equity gaps using real-world data and dynamic visualizations.

https://codap.concord.org/app/static/dg/en/cert/index.html#shared=https%3A%2F%2Fcfm-shared.concord.org%2FqL13dtOqVcnqN6hnredH%2Ffile.json.

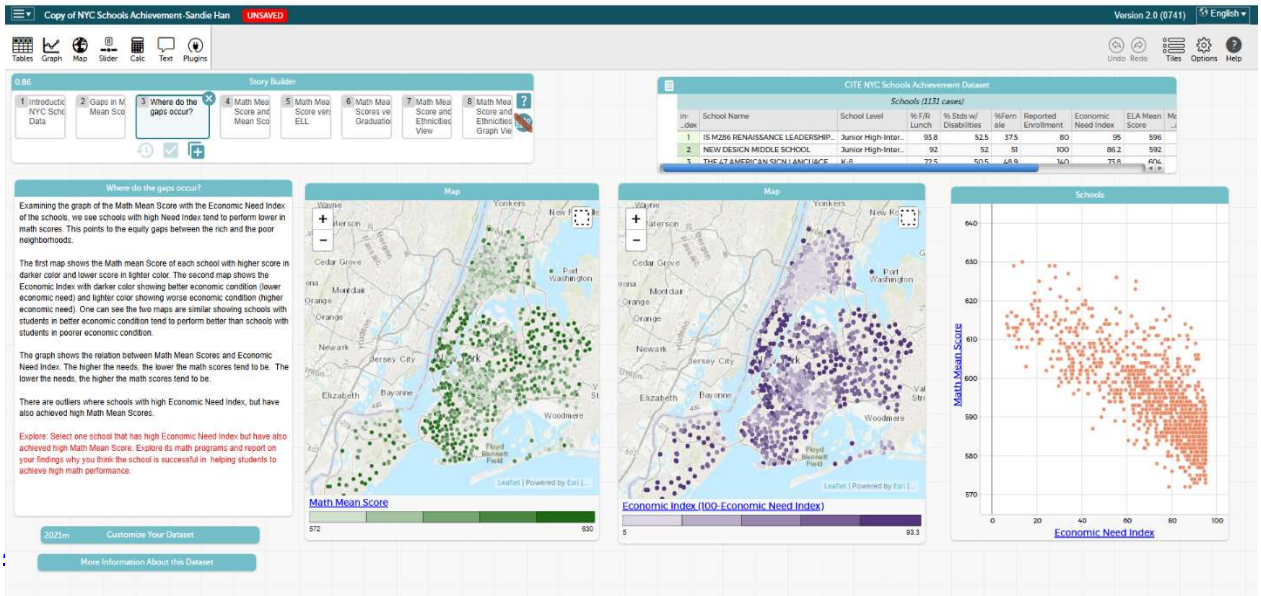


Figure 11: A CODAP activity exploring mathematics achievement gaps across New York City schools.

Following the activity, teacher candidates are encouraged to reflect on the following discussion questions:

Discussion questions:

1. What do the schools near City Tech look like in terms of selected characteristics (accessibility, course pass rate; ELA and math proficiency, total # students, graduation rates, ELL programs, other features, accelerated classes (APA classes), elective classes, survey safety) in comparison to schools in other neighborhoods (Manhattan, Brooklyn, Park Slope; Bronx)?
2. Can any inferences be made about what may account for learning gaps and differences in learning achievements?
3. What are some potential barriers or challenges to student success in mathematics and overall academic achievements in New York City schools, particularly in high-needs schools?
4. What evidence-based approaches/strategies may help improve the mathematics skills/engagement among students in high-need schools?
5. How would you design culturally responsive pedagogy to engage students in mathematics learning?
6. What activities would you design to increase students' interest and engagement in computational thinking and data analysis?

Course Module Readings

EdBuild (2019). *23 Billion*. https://edbuild.org/content/23-billion

Futterman, K. (2020). *Education and equality in NYC: The problem at hand.* https://polygonnews.org/2299/showcase/education-and-equality-in-nyc-the-problem-at-hand/ on August 7, 2022.

Hannah-Jones, H. (2016). Choosing a school for my daughter in a segregated city, *New York Times Magazine*, https://www.nytimes.com/2016/06/12/magazine/choosing-a-school-for-my-daughter-in-a-segregated-city.html

Lefty, L. (2021). The long fight for educational equity in NYC. https://www.mcny.org/story/long-fight-educational-equity-nyc on August 1, 2022.

Lombardo. C. (2019). *Why white school districts have so much more money*. https://www.npr.org/2019/02/26/696794821/why-white-school-districts-have-so-much-more-money.

Maragkidou, M. (2015). Segregation, bullying and fear: The stunted education of Romani children in Europe. https://www.amnesty.org/en/latest/news/2015/04/the-stunted-education-of-romani-children-in-europe/

Mervosh, S. (2019). How much wealthier are white school districts than nonwhite ones? $23 Billion, Report Says. *New York Times*, https://www.nytimes.com/2019/02/27/education/school-districts-funding-white-minorities.html

Gould, J. (2021). *New York's schools are still the most segregated in the nation: Report.* https://gothamist.com/news/new-yorks-schools-are-still-the-most-segregated-in-the-nation-report on July 25, 2022.

Robinson, G. (2013). *Class in the classroom: The income gap and NYC's schools*. https://citylimits.org/2013/09/25/class-in-the-classroom-the-income-gap-and-nycs-schools/ on August 1, 2022.

Shapiro, E. (2019). Facing segregated schools, parents took integration into their own hands. It's working, *New York Times, https://www.nytimes.com/2019/04/16/nyregion/new-york-city-school-segregation.html*

Shapiro, E. (2021). Lawsuit challenging N.Y.C. school segregation targets gifted programs, *New York Times, https://www.nytimes.com/2021/03/09/nyregion/nyc-schools-segregation-lawsuit.html*

Surk, B. (2018). In a divided Bosnia, segregated schools persist, *The New York Times*, https://www.nytimes.com/2018/12/01/world/europe/bosnia-schools-segregated-ethnic.html

ABOUT THE AUTHORS

Nadia Stoyanova Kennedy is Professor of Mathematics Education at CUNY's New York City College of Technology. Her research focuses on philosophy of mathematics education, dialogic teaching, mathematics identity, and critical approaches to mathematics and teacher education. She has contributed to CUNY's CITE initiative for integrating computing in teacher education programs. She has also served as PI and Co-PI on

NYSED and NSF grants supporting teacher development and the retention and STEM student success.

Boyan Kostadinov is Professor of Mathematics at New York City College of Technology, CUNY. He received his PhD in Mathematics from UCLA in 2005. His dedication to teaching was rewarded by receiving the prestigious Robert Sorgenfrey Distinguished Teaching Award from UCLA and the 2020 Distinguished Teaching Award of the Mathematical Association of America Metro New York Section. His main scholarly interests are computational thinking and data science at any level.

Sandie Han is Professor of Mathematics and Dean of the School of Science & Allied Health at Medgar Evers College, with over 30 years of experience in higher education. She is deeply committed to advancing STEM education and academic excellence, having played a pivotal role in curriculum reform and faculty professional development. Her impactful career includes serving as Chair of the Mathematics Department and spearheading CUNY-wide initiatives such as the New Lecturer Initiative for 250 new faculty, training for new department chairs, and a summer institute on STEM teaching and learning.