LEARN: Visualizing Unequal Access to Computer Science Education in Massachusetts

Alina Chiu*   Kreena Totala†   Julia Favis‡

Abstract

Jobs in technology are appearing faster than any other fields and paying much more. Yet, unequal access to computer science education in Massachusetts public high schools remains an issue. Despite access to national educational datasets and school-by-school course offerings, determining which schools teach computer science and the demographics of students taking these courses is arduous without proper tools. We propose a visualization tool, LEARN, for educators and members of Northeastern University's club, FirstByte, who hope to identify public schools in Massachusetts that do not teach computer science or do not provide equal access to computer science education. Our approach would allow users to easily view and understand the levels of accessibility to Computer Science courses in different public schools. Insights derived from our visualization are aimed at determining which schools need the most support to start teaching computer science. https://github.com/DS4200-S22/final-project-learn.

1 Introduction

Unequal access to STEM education is an issue that is often overlooked. Proper STEM education can improve problem solving and critical thinking skills, boost confidence, and open up opportunities for careers in higher-paying fields. Many of the fastest-growing jobs require a STEM background. Children who grow up with less access to STEM education may be put at an unfair disadvantage. One of the fastest growing fields of STEM is Computer Science. Not all public schools have the funding or resources to provide a Computer Science education. However, it is one of the most well-paid and fastest evolving fields in the 21st century. According to the Bureau of Labor Statistics [6], employment of computer and information research scientists is projected to grow 22 percent from 2020 to 2030, much faster than the average for all occupations. Children, and especially minorities exposed to computer science at a younger age are more likely to pursue a career in computer science. We want to do what we can to help our local schools provide this necessary computer science education to students. Northeastern's club FirstByte has the mission of making a Computer Science education more accessible to local public schools. Our goal is to create a visualization tool using data from the Computer Science Teachers Association to more easily identify the public schools in Massachusetts that lack a computer science education in order to further this goal.

2 Related Work

Given that unequal access to computer science education primarily affects minorities in the technological field, a firm understanding of what demographics should be displayed in our visualization tool is increasingly important. In Wang et al [7], it was found that Hispanic students and female students were less likely to have learned computer science or have confidence to learn computer science compared to their counterparts and that Black and Hispanic students reported less access to computer science education in school and computers at home. Additionally, it was also found that student's perceptions of those in the computer science field were narrow and stereotypical (White, male, smart). Since LEARN focuses on highlighting unequal access to computer science education in high school, these findings will help influence the design of our final project. More specifically, we will include data related to individuals and identities that lack access to computer science education in our final visualization.

To accurately display information in our visualization, other works that discuss making our visualization more effective such as Hullman et al [2] and Lisa Muth [4] aid understanding of different visualization strategies using structuring strategies. We will have many attributes to consider, for example: we may choose to show a choropleth of Massachusetts over time showing each region's access to computer science education. We may also want to show funding to these specific public schools. Understanding viewers' perceptions of the quality and intention of different sequences based on Hullman et al's work will guide our understanding of information presentation since Hullman et al's study was run with a user interface that had multiple choropleth graphs. Hullman et al specifically writes about four factors that influence effectiveness of the structure: time, geography, consistent measures, and hierarchy. Hullman's analysis of these factors directly relate to our data which relates to Massachusetts geography and educational measures. Muth's blog post also addresses the strength and weaknesses of a choropleth map. For example, she mentioned that these maps are good for studying one variable, and are best suited for relative data and showing the big picture in the form of regional trends. However, they often fail to convey subtle differences. Since we have data that relates to regions and specific public schools in the state of Massachusetts, a choropleth may be considered as a visualization tool. However, since we have access to specific data from different public high schools, a point map may be more effective.

We are not aware of any previous work that applies visualization techniques for the purpose of access to computer science education. However, existing work which focuses on general map visualizations related to other topics provides information on how do display information. For example, Çay et al [9] conducted a series of interviews to help understand user experiences and interactions with different COVID-19 data visualizations that included geographical location and number of cases. These visualizations identify the areas with more COVID-19 cases and those with different shades of the same color on a map or with interactive tabs that contain specific information for each state/region. Çay et al discovered that spatial exploration of map visualizations can be driven by a user's existing knowledge about global events, curiosity about the user's local situation, and personal connections with places. For both the WHO and Johns Hopkins visualizations, users were mostly interested in the overall picture and refrained from deeper analysis. Çay et al also identified color as an element that aids in the interpretation of data and in inducing positive or negative feelings towards the

*e-mail: chiu.a@northeastern.edu
†e-mail: totala.k@northeastern.edu
‡e-mail: favis.j@northeastern..edu

visualized topic. The information about user experience from Çay et al informs our understanding of user interaction with our visualization tool. For instance, since many users look for general trends and react to visualizations based on color, it may be helpful to create a visualization of Massachusetts that represents schools that have the least resources for computer science education in red and the schools that have the most in another more neutral color so that users can easily identify trends about which districts need the most support.

Another approach to map visualization is Tilt Map, proposed by Yang et al [8], which a new interaction technique that transitions between 2D and 3D map visualisations in immersive environments. Yang et al compared subjects' task performance interpreting population density data using 2D choropleth maps and 3D prism maps in virtual reality (Tilt Maps) and found benefits in time, accuracy and user preference for the 3D prism maps over 2D alternatives due to their interactive nature. However, since Tilt Maps primarily aid in displaying population density in general districts, and our visualization tool focuses on individual public high schools, this method of visualization may not be as effective as a point map. An additional approach to mapping and visualizing was written in Jo et al [3] which discusses the effectiveness of Multiclass Density Maps. Jo et al's method of generating Multiclass Density Maps explores how they can be an effective idiom of scalable visualization and displays a unique method to displaying data outside of point maps. Since our visualization tool makes use of multiple data types, the flexibility of a Multiclass Density Map may prove to be useful when structuring our data.

3 Use Case

To determine schools in the Massachusetts area that lack equal access to computer science education, we present a usage scenario where our visualization tool can be used to help Northeastern's computer science education club, FirstByte. FirstByte is a club that aims to provide educators with materials, curricula, and support to teach computer science and engineering in their classrooms. The club hopes to address larger issues such as gender inequality in the tech industry by introducing computer science to students at a younger age. Suppose a member of FirstByte would like to determine where to loan their two MayKey MayKey kits in Massachusetts. Since there are 437 public schools in Massachusetts and over 287,234 students [5], it is incredibly difficult for this member to determine which two schools to distribute kits and curricula to. They download the 2021 State of Computer Science Education: Accelerating Action Through Advocacy dataset, which contains information about the number of public schools nationwide that teach at least one foundational computer science course. They find that it is still difficult to determine which specific schools need the most educational support because over 18% of these schools do not teach computer science at all. They would like to prioritize supporting schools that lack diversity in their computer science courses.

They load their dataset into LEARN and are able to easily see a map of Massachusetts and the specific schools that do not have computer science courses as well as data about the school's demographics. The FirstByte member may notice that there are certain areas of Massachusetts that need more support than others based on color coding. For instance, a public school that may not teach computer science and may not have policies in place to begin teaching computer science courses would be highlighted in red since they have no plans to implement technical courses at all. This visualization tool would help the FirstByte member quickly determine which area to send their kit and curricula to next, so the students at that school could begin to gain foundational knowledge in computer science as soon as possible. The member may then reach out to the two schools that they think would be good targets and send the kits to them, along with a helpful guide on how to use them effectively.

4 Data

We will use data from the Computer Science Teachers Association, Expanding Computing Education Pathways, and Code.org's Advocacy Coalition called 2021 State of Computer Science Education: Accelerating Action Through Advocacy [1]. This dataset includes information about national and state-level computer science education policies, including policy trends, maps, state summaries, and data about which schools teach at least one foundational computer science course. There is also nationwide information and information specific to Massachusetts high schools that have responded to Code.org's survey. The data source is self-reported from different schools in the United States that fill out a Code.org survey with information about the number of students that take computer science courses or participate in Code.org curriculum. If a school is not listed in Code.org's database, then they need to fill out an additional survey to verify that their school exists. Because a school needs to be listed in Code.org's database and must know about Code.org's survey, this dataset may have an issue of non-response bias, where schools that do not teach computer science and have not responded to the survey are not included. Schools that have not responded to the survey cannot be displayed in our visualization tool because we will have no information on their current resources. Additionally, the survey asks about whether or not a school participates in Code.org's Hour of Code initiative which is not directly related to teaching foundational computer science courses and more directly related to the interests of Code.org's organization.

5 Task Analysis

Task ID #	Domain Task	Analyze Task (high level)	Search Task (mid level)	Analytic Task (low-level, "query")
1	I want to easily determine the schools in Massachusetts that do/do not offer computer science education.	Consume →Discover	Browse	Summarize
2	I want to be able to identify underrepresented communities in the Massachusetts schools.	Consume →Discover	Browse	Identify
3	I want to be able to compare a school's demographics to other schools in their county.	Consume →Discover	Lookup	Compare

The primary consumer of our visualization will be educational clubs that look to identify which schools need the most support in teaching computer science, more specifically, Northeastern's club FirstByte and its members. The type of consumption our visualization will be primarily developed for will be Discover because our visualization hopes to help FirstByte members discover and determine which schools in the Boston/Massachusetts area have the least access to computer science curricula. LEARN would allow its users to consume by discovering these schools with ease.

6 Implementation Plan and Preliminary Work

In our visualization, we will have an interactive point map of Massachusetts that the user can can query to find out which public schools have or lack a certain curriculum, as well as demographic information about the schools. There is also an option for the user to toggle between the point map and table data views. The data set we will work with contains specifics about how and where computer science is taught in Massachusetts, and the user will be able to query the map based on those specifics.

The visual encodings included in our project include the point map of Massachusetts, a pie chart (for demographic info) and various box and whisker plots (for other statistics) to compile a summary for each school when it is selected.There are also filters and a table included in the table view to narrow a search based on specific criteria.

The interactive components include a mix of brushing and linking, zoom, highlighting, and filtering. Brushing and linking will be used to make the school and county maps interactive in the map window. There will also be an option to zoom into and out of the map for a better view of specific counties or regions. Once a school is selected, it will be highlighted and a popup with its demographic makeup and statistics will be displayed. The filtering feature will be utilized in the table view where the user can manually put in their desired filter criteria to find schools that match.

We will prioritize creating a map view and table view. We will also prioritize the filters in the table view and the colored points on the map view that indicate CS or no CS. We will then move on to creating as many of the statistic and demographic items as we can when a school is selected. Next, we will make the summary information of each of the counties in the map view. Lastly, we will implement the zoom feature in the map view. The summary information and the zoom features are nice to have if we have time to implement them.

The technologies we plan to use to create our visualization include: Visual Studio Code as our text editor, HTML/CSS to build our web page and JavaScript with D3 to implement the functionality of our visualization. We do not have any specific libraries or APIs that we will be using at the moment. Our final web page will be hosted on GitHub.

7 Visualization Design

The first view the user see when they open LEARN will be the side by side map views that include a point map of schools and a chloropleth map of counties in Massachusetts; they will hover over one of the maps with their cursor. The main purpose of the point map is to easily show which schools in Massachusetts offer/do not offer CS education where orange dots represent schools that do not offer CS education and blue dots represent schools that do offer CS education. The colors were chosen to be as distinct and color-blind friendly as possible. Displayed next to the School map will be a box containing information about the school (represented by a dot on the map) that their cursor is hovering over. In that box will be information about if the school has a CS program, the county it is located in, and the type of school it is, as well as 4 box and whisker plots each with the dot in them that indicates where in each statistic the selected school lies, (the 4 statistics include median income, number of students, number of students of color, and number of students in stem, as well as a pie chart that represents the racial distribution of the school. Each of these components allow users to easily compare the school they have selected to other schools in Massachusetts.

At the same time, brushing and linking is implemented when a school is hovered over so that its corresponding county and county information will be displayed simultaneously. The information that will be displayed in that box will be the number of high schools, the total number of high school students, and the percent of schools that offer CS in the county. Brushing and linking will allow users to easily understand how the school they have selected compares to the county it is in. Alternatively the user can also hover over a county on the county map where only the county information will be displayed. This option exists because a county cannot necessarily be linked to one specific school. Visualizing two side-by-side maps will allow for users to easily compare different schools to other schools in their county and to easily identify underrepresented communities in the Boston area and beyond by county.

The user can also navigate to the table view tab where they can filter which the schools they want to see by name, if they offer CS education, type, number of students, and school county. A list of schools and their information fitting the selected criteria will then appear in the table. The table view gives users an alternative format to query and interpret data by allowing them to filter using a variety of categories.

Acknowledgments

The authors wish to thank Prof. Ab Mosca, the students of DS4200, and contributors to open source datasets we use.

References

[1] E. C. E. P. Computer Science Teachers Association, Code.org. 2021 state of computer science education: Accelerating action through advocacy. `https://advocacy.code.org/stateofcs`, 2021.

[2] J. Hullman, R. Kosara, and H. Lam. Finding a clear path: Structuring strategies for visualization sequences. *Computer Graphics Forum*, pp. 365–375, July 2017. doi: 10.1111/cgf.13194

[3] J. Jo, F. Vernier, P. Dragicevic, and J.-D. Fekete. A declarative rendering model for multiclass density maps. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):470–480, 2019. doi: 10.1109/TVCG.2018.2865141

[4] L. Muth. What to consider when creating choropleth maps, March 2018.

[5] M. D. of Elementary and S. Education. School and district profiles., 2019.

[6] U. B. of Labor Statistics. Occupational outlook handbook, 2020.

[7] J. Wang, H. Hong, J. Ravitz, and S. H. Moghadam. Landscape of k-12 computer science education in the u.s.: Perceptions, access, and barriers. *SIGCSE '16: Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, 36:645–650, Feb. 2016. doi: 10.1145/2839509.2844628

[8] Y. Yang, T. Dwyer, K. Marriott, B. Jenny, and S. Goodwin. Tilt map: Interactive transitions between choropleth map, prism map and bar chart in immersive environments. *IEEE Transactions on Visualization and Computer Graphics*, 27(12):4507–4519, 2021. doi: 10.1109/TVCG.2020.3004137

[9] D. Çay, T. Nagel, and A. E. Yantaç. Understanding user experience of covid-19 maps through remote elicitation interviews. In *2020 IEEE Workshop on Evaluation and Beyond - Methodological Approaches to Visualization (BELIV)*, pp. 65–73, 2020. doi: 10.1109/BELIV51497.2020.00015

8 GROUP CHARTER

- Group Purpose:
  We formed this group to make computer science education more broadly available for students in Boston/Massachusetts and identify key areas where computer science education can be improved.

- Group Goals:
  Our group's project, process, and quality goals are to create a visualization tool that is high quality and easy to use. We are aiming for an A in the class as the course grade. We are planning to provide utmost performance and investing as much effort as we can!

- Group Member Roles/Responsibilities:

  Alina: group leader, point person for communications
  Julia: meeting facilitator, information manager
  Kreena: documentation coordinator

- Ground Rules:
  We are planning to meet via zoom so that we can easily have conversations and check in on our current progress. We will also meet based on the dynamic schedules. Our first meeting has been on Saturday (February 19, 2022) and we are meeting on Wednesday (February 23, 2022). We will likely continue to meet on Saturdays at 2pm to assign work and on Wednesdays at 4pm to make sure all work has been completed and reassign work if necessary. We are planning to divide up the work for assignments that involve writing different sections and then meeting again to proof read everything. For assignments where we are working on programming our visualization tool, we plan to split up the work into pairs and go over everything together after coming up with design implementations.
  We plan to handle different views by making sure that everyone has a voice and voting on areas where we may have disagreements after discussing the pros and cons of each problem. If someone feels strongly about an idea, we will try to come to a compromise that will satisfy all parties. We also plan to hold each other accountable by checking in over text, trying to complete the work by Monday/Tuesday (setting up deadlines). We expect that everyone will have an equal level of participation and commitment to this project. However, if someone cannot finish the work on time, the most important thing is that they let us know as soon as possible so that we can find a way to finish the assignment on time.

- Potential Barriers and Coping Strategies:

  A potential barrier would be if someone isn't able to finish their work on time by our deadline. In this case, we can redistribute the work to the other members. It would be best if someone can't finish on time to let us know as soon as possible. Another barrier could be if someone drops out of the class. In this case, we would talk to the teaching staff and see if we can add someone else to our group or redivide up the work if necessary. Finally, if we are unable to contact someone for several days, we will also try to contact the teaching staff to determine if they have dropped the class or are okay.

  – Have you all been abiding by your agreed-upon guidelines?

    Yes we have.

  – Do you all feel comfortable with the group roles? Are there any problems you need to troubleshoot, and if so how can you address them?
    No we are all good with our roles/

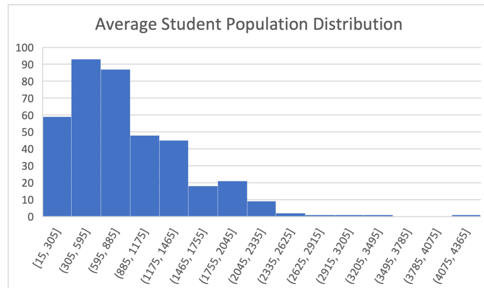
  – Each group member should write one positive thing they have seen other group members contribute to the project.
    Julia: Alina wrote the questions for the interview, Kreena made visualizations for the Data Review.
    Alina: Julia made sketches for the design sketches, Kreena wrote insights to the data review.
    Kreena: Alina made visualizations for the data explorations, Julia interviewed a person for the interview section.

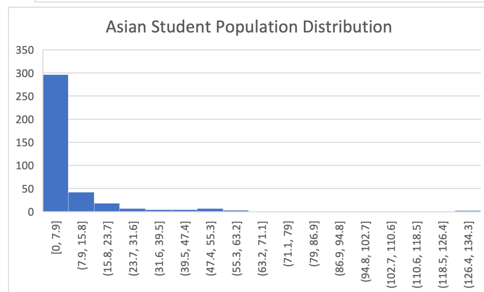
9 Appendix B: Data Exploration

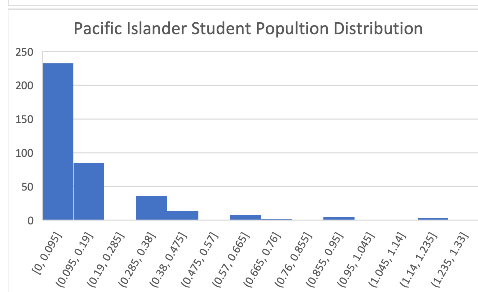
9.1 Data Review

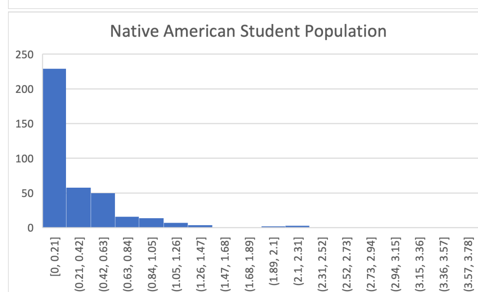
The data types present in our dataset include attributes and positions. The position data in our dataset are the Latitude and Longitude of each school. The remaining data in the dataset are categorical or ordered attributes. The categorical attributes and levels in this dataset include County Name (levels: counties in each state in the US), School District Name (levels: school districts in each state in the US), School Name (levels: school names in each state in the US), School Type (group) (levels: public charter, private), StateLowerChamber (levels: lower chambers in each state in the US), StateUpperChamber (levels: upper chambers in each state in the US), School Type (levels: charter, public, private), State (levels: abbreviations of all 50 states in the US), State Name (levels: full names of each state in the US), Audit Data (levels: the metadata/JSON object from each audit for each school), Year Spelled out (levels: 2019-20), Urban v. Rural (levels: urban, rural), Community Type (group) (levels: city, suburban, rural, town), Congressional Districts where the schools are located (levels: numbers of congressional districts from each state in the US), Community Type Grouping (levels: city, suburban, rural, town), Grade Band when graduation/start dates are (levels: 2022, mixed), Stage Hi, if this school teaches high school (levels: 0, 1), Stage Mi, if this school teaches middle school (levels: 0, 1), Stage Lo, if this school teaches lower school (levels: 0, 1), Offers CS?, if this school offers CS (levels: yes, no), and Teaches CS, if this school teaches CS (levels: yes, no), the School ID associated with each school (levels: valid school ID numbers for schools throughout the US), and Gradespan, the grade range at a school from lowest to highest grade (levels: KG-09, KG-12, PK-11, PK-12). Quantitative attributes include Avg Students/School, Students - Asian, number of Asian students at a school; Students - Black, number of Black students at a school; Students - Hispanic/Latinx, number of Hispanic/Latinx students at a school; Students - Native American, number of Native American students at a school; Students - Pacific Islander, number of Pacific Islander students at a school; Students - Two or More Races, number of students with two or more races at a school; and Students - White, number of white students at a school. Ordinal attributes include Year Spelled out (levels: year when the access report was given to Code.org), Grades Hi, the highest grade a school teaches (levels: 1 - 13), Grades Lo, the lowest grade a school teaches (levels: PK-11), School Year, the current school year at time of submission (levels: current school year at time of submission), and Year, the year the data was collected (levels: year data was collected).

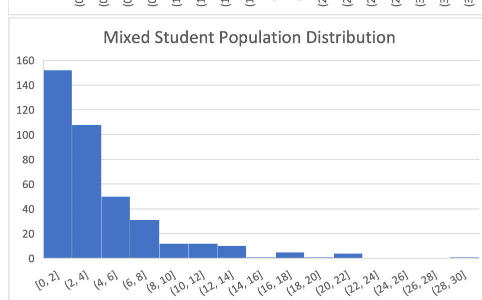
The ranges and levels of the categorical and ordinal data are listed above. Additionally, ranges of quantitative and ordinal data, mean and median values of quantitative and ordinal data include Avg Students/School (range: 4062, min: 15, max: 4077, mean: 848.1036269, median: 703), Students - Asian (range: 133.2857143, min: 0, max: 133.2857143, mean: 7.968992248, median: 2.285714286), Students - Black (range: 368.8571429, min: 0, max: 368.8571429, mean: 12.95348837, median: 4), Students - Hispanic/Latinx (range: 444.2857143, min: 0, max: 444.2857143, mean: 24.42488003, median: 8.857142857), Students - Native American (range: 3.714285714, min: 0, max: 3.714285714, mean: 0.294573644, median: 0.142857143), Students - Pacific Islander (range: 1.285714286, min: 0, max: 1.285714286, mean: 0.112218531, median: 0), Students - Two or More Races (range: 29.14285714, min: 0, max: 29.14285714, mean: 3.951642673, median: 2.714285714), Students - White (range: 254.5714286, min: 0, max: 254.5714286, mean: 69.86046512, median: 66.71428571), School Year (range: 0, min: 2019, max: 2019, mean: 2019, median: 2019, distribution: all 2019), Year (range: 0, min: 2019, max: 2019, mean: 2019, median: 2019, distribution: all 2019), Grades Hi (range: 3, min: 9, max: 12, mean: 11.9742, median: 12, distribution: heavily skewed right, mostly grade 12 but a few are grade 8), Grades Lo (range: 10, min: 1, max: 11, mean: 8.42663, median: 9, distribution: heavily skewed right, mostly grade 9 but some are KG/PK). Finally, the distributions of quantitative and ordinal data that have not been mentioned yet are represented below with charts:

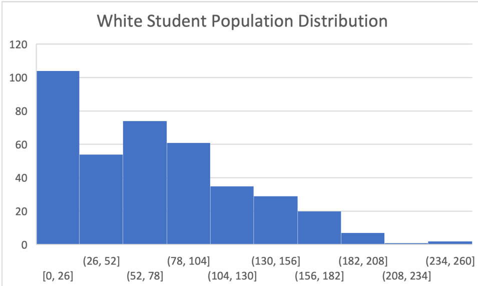


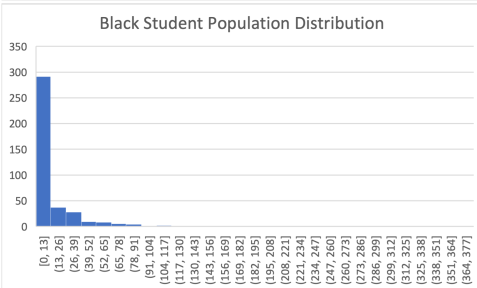












There were a few quality issues that we noticed while performing our initial review of the dataset including:

- There is a column called Audit Data which is filled with JSON data about each school and is confusing because, to extract the data we will need to parse through each individual JSON file to learn more about what audit data was given. However, all of the data provided in each JSON file overlaps with existing columns in the dataset such as County Name and Community Type. Because the dataset already includes the information in each JSON file in a less confusing format, we will clean the data by removing the Audit Data column.

- Most columns are entirely filled in, but some columns including County Name, statelowerchamber, and stateupperchamber all have blanks for missing data. The amount of missing data in these columns is relatively low, so, when we do use these columns, our solution strategy would be to represent the data that does not exist as "Unknown." This is because some of the columns in the dataset already represent unknown values with "Unknown," so representing missing data would remain relatively consistent.

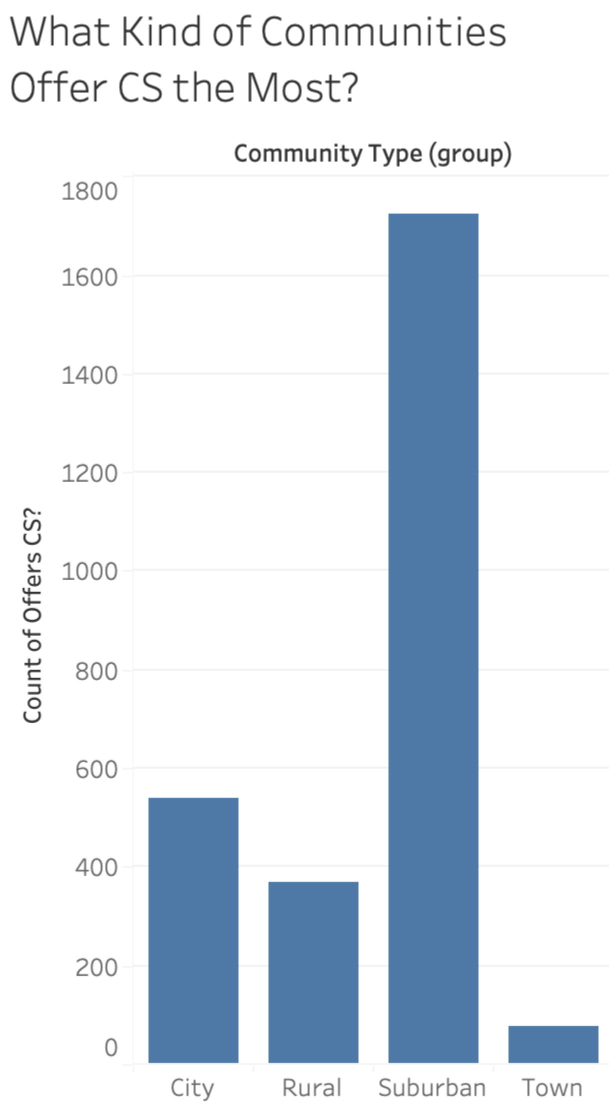
- There is also data included from other states/cities outside of Massachusetts. Our solution strategy for this is to filter the dataset so that only the data about Massachusetts schools remains since our visualization is focused on access to CS education in Massachusetts.

- Some schools have submitted their information multiple times to Code.org, meaning that each school has a different number of submissions for each data type. This means that there is repeated data throughout the dataset that could skew results since some will have more weight than others. Our solution strategy is to remove repeated data, taking the averages of each category so that each school will only have one row of data.

9.2 Insights

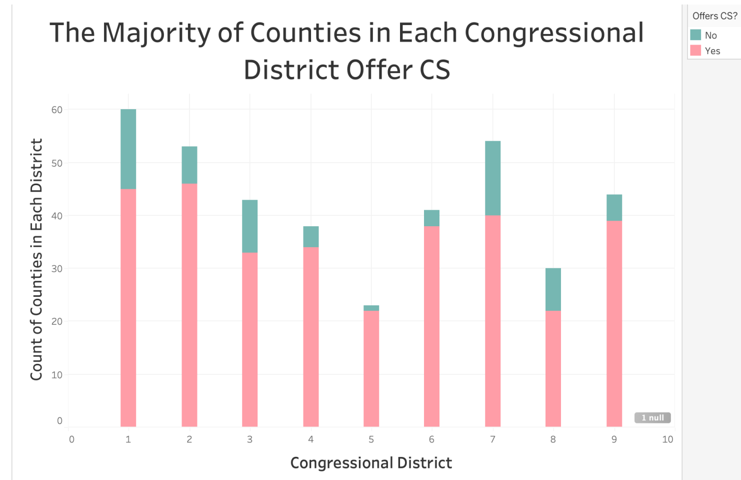
While exploring the data, our team found that the access to computer science education is not distributed equally in the state of Massachusetts. Our purpose when creating visualizations of the data was to point point exactly which categories of schools/students have the most and least access to computer science education. Some trends and patterns that we see are that suburban schools have the most access to computer science education by far compared to their city/rural/town counterparts. The state of Massachusetts is also comprised most significantly of city/suburban community types. Although it does it make sense that the larger population groups will have more access to CS education, it is also important to note that the rural and town groups are disproportionately lacking in CS programs when comparing amount of programs to number of students ratios. Massachusetts has more white students than any other race, and the data shows that white students have the most access to CS education in proportion to population, with only 7 percent of white students in MA not having access, compared to about 15 percent of black students not having access to CS education. About 8 percent of Asian students and 13 percent of Hispanic students do not have access to CS education. The data set includes the categories of public and charter schools, with significantly more public schools included. We were surprised that these were the only types of schools included in the data set, as we had previously thought that the discrepancy in CS education available in public vs. private schools would be a significant data point. This limits our ability to compare education access by school type, therefore community type is our most significant metric in determining what type of schools have access or not.

9.3 Screenshots

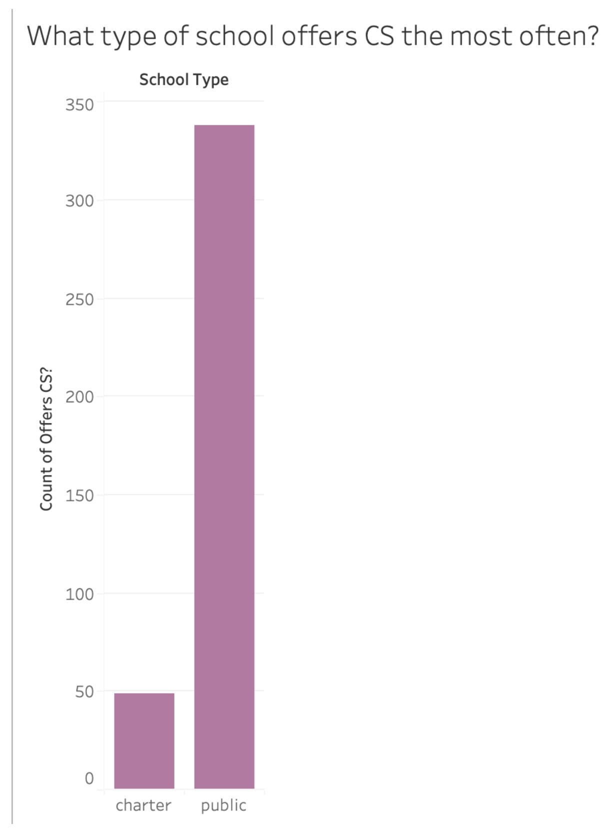


Count of Offers CS? for each Community Type (group).

In this visualization, we are exploring community type data (count of each community type included in the data set, and the count of how many of each community type offer CS education. The visual encodings used in this graph include size and position, as larger bars indicate more CS programs offered and smaller bars indicate less CS programs offered, and the bars are all positioned next to each other to show the contrast in size between them. We used these encodings to make it clear to the audience that different communities have different amounts of CS programs offered. The trend we see that suburban communities have the most access to computer science education by far, followed by cities, then rural, then township communities.



This visualization shows the number of counties in each congressional district. It also shows how many counties in each district offer CS and how many do not. The visual encodings are size, color, and position. The size of the bars represent how many counties are in a congressional district. The bars are positioned so that each represents a different district, and they are positioned next to each other so they can be easily compared. The color of the bars differentiate the counties that offer CS from the counties that do not. These encodings show the viewers the different numbers of counties in each district, as well as how many counties offer CS and how many do not offer CS. The trend is that all the districts have more counties that offer CS than counties that do not.



This visualization explores the difference in CS offerings based on school type. It specifically displays the count of how many schools of each type (public or charter) offer CS in Massachusetts. The visual encodings in this graph are size and position. The size of the bars indicates the quantity of schools that offer CS. The larger the bar, the more schools offer CS. The position indicates whether the bar represents public schools or charter schools. Additionally, the bars are positioned next to each other such that we can clearly see the contrast in size.These encodings are used to show a difference between the amount of public schools that offer CS and the amount of charter schools that offer CS. There is a clear trend that more public schools offer CS than charter schools.

9.4 Data Snippet

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	congressiona	County Name	Grade Band	Offers CS?	School Distric	School Id	School Name	School Type (	state_lower_	state_upper_	Year Spelled	Years Access	Grades Hi	Grades Lo	Gradespan	School Type	School Year	Stage El	Stage Hi	Stage Mi	State	State name	Teaches Cs	Year
2	7	Suffolk Coun	Mixed	Yes	BOSTON	2.50279E+11	O'BRYANT SC	public & char	7th Suffolk	Second Suffc	2019-20	1	12	7	12-Jul	public	2019	0	1	1	MA	Massachuse	Yes	2019
3	7	Suffolk Coun	Mixed	Yes	BOSTON	2.50279E+11	O'BRYANT SC	public & char	7th Suffolk	Second Suffc	2019-20	1	12	7	12-Jul	public	2019	0	1	1	MA	Massachuse	Yes	2019
4	7	Suffolk Coun	Mixed	Yes	BOSTON	2.50279E+11	O'BRYANT SC	public & char	7th Suffolk	Second Suffc	2019-20	1	12	7	12-Jul	public	2019	0	1	1	MA	Massachuse	Yes	2019
5	7	Suffolk Coun	Mixed	Yes	BOSTON	2.50279E+11	O'BRYANT SC	public & char	7th Suffolk	Second Suffc	2019-20	1	12	7	12-Jul	public	2019	0	1	1	MA	Massachuse	Yes	2019

Urban v. Rur	Avg Students	Latitude	Longitude	Community	Community	Community	Students - Asian	Students - Black	Students - Hispanic/Latinx	Students - Native American	Students - Pacific Islander	Students - Two or More Races	Students - White
Urban	1589	42.33136	-71.08892	city_large	City	City	46.28571429	74.14285714	75	0.428571429	0.428571429	4.428571429	26.28571429
Urban	1589	42.33136	-71.08892	city_large	City	City	46.28571429	74.14285714	75	0.428571429	0.428571429	4.428571429	26.28571429
Urban	1589	42.33136	-71.08892	city_large	City	City	46.28571429	74.14285714	75	0.428571429	0.428571429	4.428571429	26.28571429
Urban	1589	42.33136	-71.08892	city_large	City	City	46.28571429	74.14285714	75	0.428571429	0.428571429	4.428571429	26.28571429

	A	B	C	D	E	F
1	congressional_district	County Name	Grade Band	Offers CS?	School District Name	School Id
2	2	Worcester County	Mixed	Yes	ABBY KELLEY FOSTER CHARTER PUBLIC (DISTRICT)	2.50E+11
3	8	Plymouth County	12-Sep	Yes	ABINGTON	2.50E+11
4	7	Suffolk County	Mixed	Yes	ACADEMY OF THE PACIFIC RIM CHARTER PUBLIC (DISTRICT)	2.50E+11
5	3	Middlesex County	12-Sep	Yes	ACTON-BOXBOROUGH	2.50E+11

	G	H	I	J
1	School Name	School Type (group)	state_lower_chamber	state_upper_chamber
2	ABBY KELLEY FOSTER CHARTER PUBLIC SCHOOL	public & charter	13th Worcester	First Worcester
3	ABINGTON HIGH	public & charter	7th Plymouth	Norfolk & Plymouth
4	ACADEMY OF THE PACIFIC RIM CHARTER PUBLIC SCHOOL	public & charter	14th Suffolk	Norfolk & Suffolk
5	ACTON-BOXBOROUGH REGIONAL HIGH	public & charter	37th Middlesex	Middlesex & Worcester

	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Year Spelled out	Years Access Report	Grades Hi	Grades Lo	Gradespan	School Type	School Year	Stage El	Stage Hi	Stage Mi	State	State name	Teaches Cs
2	2019-20	1	12	KG	KG - 12	charter	2019	1	1	1	MA	Massachusetts	Yes
3	2019-20	1	12	9	12-Sep	public	2019	0	1	0	MA	Massachusetts	Yes
4	2019-20	1	12	5	12-May	charter	2019	1	1	1	MA	Massachusetts	Yes
5	2019-20	1	12	9	12-Sep	public	2019	0	1	0	MA	Massachusetts	Yes

	X	Y	Z	AA	AB	AC	AD	AE	AF
1	Year	Audit Data	Urban v. Rural	Avg Students/School	Latitude	Longitude	Community Type	Community Type (group)	Community type grouping
2	2019	{"version":1.0,"stat:	Urban	1425	42.305635	-71.801599	city_midsize	City	City
3	2019	{"version":1.0,"stat:	Urban	593	42.119058	-70.954519	suburban_large	Suburban	Suburban
4	2019	{"version":1.0,"stat:	Urban	527	42.244384	-71.132363	city_large	City	City
5	2019	{"version":1.0,"stat:	Urban	1814	42.4801	-71.45815	suburban_large	Suburban	Suburban

	AG	AH	AI	AJ	AK	AL	AM
1	Students - Asian	Students - Black	Students - Hispanic/Latinx	Students - Native American	Students - Pacific Islander	Students - Two or More Races	Students - White
2	7.142857143	106.7142857	42.42857143	0.714285714	0.142857143	6.571428571	39.85714286
3	2.428571429	4	7.714285714	0	0	1.142857143	69.42857143
4	0.857142857	45.42857143	21.57142857	0	0	1.142857143	6.285714286
5	86.42857143	6.714285714	11.42857143	0.571428571	0	11.57142857	142.2857143

10 Appendix C: Interview

10.1 Interview Script

Questions We Plan to Ask:

- What does your day to day consist of as FirstByte's president?
- Could you tell me how you currently determine which schools in Massachusetts need the most support in terms of computer science in five steps?
- What tools do you need to be more successful in identifying schools that do not currently teach computer science?
- What kinds of comparisons or statistics do you need to help more easily determine which schools need support to teach computer science?
- Please walk me through your thought process for this scenario. If you were to choose between two schools that both do not teach computer science, how would you determine which one to partner with for FirstByte's Loaner Program?
- Please describe two positive aspects of the way you currently locate educators/schools in the Massachusetts area.
- Could you also describe two pain points in your current process?

Potential Follow Up Questions:

- You mentioned that you typically do ¡X¿ as FirstByte's president. Could you explain to me how ¡X¿ shapes the club's goals?
- Can you explain more about ¡X¿ part of how you currently locate Massachusetts schools?
- In addition to ¡X/Y/Z¿ tool/statistic that you mentioned previously, can you name two more tools or statistics that may be helpful to have for the visualization tool? (Some examples may include schools in urban or rural areas, statistics about race, statistics about how far away a school is from Boston)
- Could you elaborate more on why you would choose ¡X¿ school over ¡Y¿ school in our given scenario?
- When you explained what works well for you now, you mentioned ¡X¿. Could you tell me about ¡X¿ in more detail?
- You mentioned ¡X¿ as a current pain point in communicating with educators. In an ideal world, how would ¡X¿ be alleviated?

10.2 Interview Notes

- What does your day to day consist of as FirstByte's president?
  - Connection side: look for potential organizations go collaborate with
  - Look for high/middle schools to share curriculum with
  - Meet directly with teachers and explain curriculum Clubs at NEU /schools
  - Work with other cubs to develop partnerships to give opportunities to teach (Neptun etc.)
  - Goal: figure out which clubs fit the needs/ increase opportunities for students to teach
- Could you tell me how you currently determine which schools in Massachusetts need the most support in terms of computer science?
  - Searching for public schools that are underfunded, that have/ don't have cs programs.
  - Quantity over quality, currently focusing on finding schools to reach out to in the first place.
- What tools do you need to be more successful in identifying schools that do not currently teach computer science?
  - Main issue: teachers should be willing to participate in the program System that keeps track of communication/keeps track of what schools lack what in terms of program.
- What kinds of comparisons or statistics do you need to help more easily determine which schools need support to teach computer science?
  - FirstByte leans towards schools with more low-income / underrepresented minority students.
- Please walk me through your thought process for this scenario. If you were to choose between two schools that both do not teach computer science, how would you determine which one to partner with for FirstByte's Loaner Program?
  - Has CS - firstByte evaluates the current curriculum Gives customized or default kit to school based on needs
  - No CS - firstByte volunteers to teach program
  - Accept or reject: based on excitement of the teacher
  - Identify teachers who are passionate about STEM
- Could you also describe two pain points in your current process?
  - Hard to find teachers that are willing to continue in the process. Retention rates for teching is bad.
- Additional notes:
  - FirstByte - only will look at Boston/Camby/Dorchester on a map - not all of MA
  - What is the target audience criteria:
  - K-6, middle high cs/ no cs - and how developed it is
  - Population data is helpful (race/income metrics) FirstByte only look at public schools Prioritze low income minorities
- SUGGESTION: look at bigger picture
- Target FirstByte first, other orgs next. We could impact a lot of clubs who work with education

10.3 Interview Results

- What does your day to day consist of as FirstByte's president?
  - The day to day activities of the president of FirstByte include finding schools to send curriculum to, talking with teachers and onboarding them into the program, and finding other clubs at NEU to collaborate with.
- Could you tell me how you currently determine which schools in Massachusetts need the most support in terms of computer science?

  - The club is currently focusing on increasing the quantity of schools they are partnering with, but in general they look for underfunded high minority public high schools that can or cannot have computer science programs that have a strong interest in promoting STEM curriculum.

- What tools do you need to be more successful in identifying schools that do not currently teach computer science?

  - It is less about the statistics of the school and more about if the teacher is actually willing to follow through with the program and introduce the curriculum, the main issue is that teachers should be willing to participate in the program. Th club also expressed interest in a tool that helps them keep track of communication between them and teachers. and what schools have/lack which programs.

- What kinds of comparisons or statistics do you need to help more easily determine which schools need support to teach computer science?

  - FirstByte leans towards schools with more low-income / underrepresented minority students.

- Please walk me through your thought process for this scenario. If you were to choose between two schools that both do not teach computer science, how would you determine which one to partner with for FirstByte's Loaner Program?

  - FirstByte reaches our to schools that both have/don't have an a CS education program that are excited about making STEM education more widely available, they are only likely to reject a school if they have no interest in the program or curriculum.

- Could you also describe two pain points in your current process?

  - Its is hard to find teachers that are willing to continue in the process. Retention rates for teaching are bad.

- CONCLUSION: From this interview we have gathered that FirstByte (and other NEU clubs focused on middle/high school education) would be able to utilize a visualization tool that shows a wide array of public schools across Boston and their population demographics (race, income) to find schools that they can potentially contact to give curriculum kits to.

11 APPENDIX D: DESIGN SKETCHES

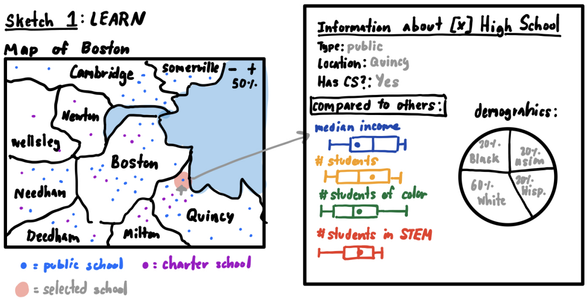


Figure 1. **Favorite** Drawn by: Julia.

Marks and Channels:
Marks: area, points, lines
Channels: position, color, area, angle

Encodings: map: the map can be zoomed in and out of, which changes its layout density, and the value of whether a school is public or charter is expressed by a colored dot.
box and whisker plots: values are expressed in each box and whisker plot by a dot that appears at the place in the plot that represents the data value, color is also used to differentiate what data the box plots represent from each other.
pie chart: values are expressed by the proportion of the pie chart that the percentage is, ideally each portion of the pie chart will be a different color. the angle of each line on the pie chart indicates how the regions will be aligned based on the data.

Marks, channels, and encodings are were chosen to accurately display the data in a clean manner, map with points to easily display the location of the school, box plots to show the place of each metric compared to other schools and a pie chart to give a ethnicity distribution. Color distinguishes the box plots from each other.

Tasks:
Easily determine schools in Boston area that do/do not offer CS. (shown at top)
Identify Underrepresented communities in Boston area.(shown through metrics in each box plot)
Compare schools to others in their county. (shown at top)

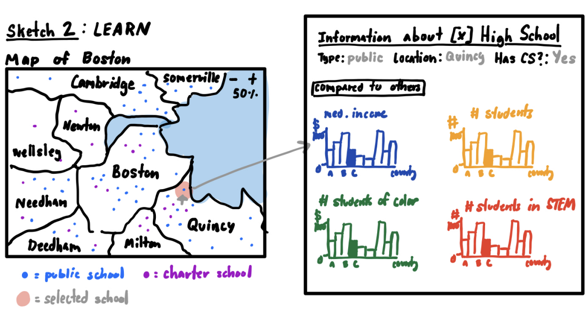


Figure 2. Drawn by: Julia.

Marks and Channels:
Marks: area, lines
Channels: position, color, area

Encodings:
map: the map can be zoomed in and out of, which changes its layout density, and the value of whether a school is public or charter is expressed by a colored dot.
bar charts: values are expressed in each bar chart by a highlighed bar that represents the bracket that the school belongs to in relation to the data, color is also used to differentiate what data the bar charts represent from each other.

Marks, channels, and encodings are were chosen to accurately display the data in a clean manner, map with points to easily display the location of the school, and bar charts to show the place of each metric compared to other schools. Color distinguishes the bar charts from each other.

Tasks:
Easily determine schools in Boston area that do/do not offer CS. (shown at top)
Identify Underrepresented communities in Boston area.(shown through metrics in each bar chart)
Compare schools to others in their county. (shown at top)

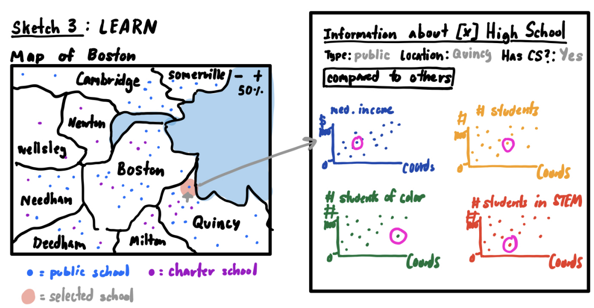


Figure 3. Drawn by: Julia

Marks and Channels:
Marks: area, points
Channels: position, color, area

Encodings:
map: the map can be zoomed in and out of, which changes its layout density, and the value of whether a school is public or charter is expressed by a colored dot.
scatter plots: values are expressed in each scatter plot by a highlighed point that represents the position of the school in relation to the other schools, color is also used to differentiate what data the scatter plots represent from each other.
Marks, channels, and encodings are were chosen to accurately display the data in a clean manner, map with points to easily display the location of the school, and scatter plots to show the place of each metric compared to other schools. Color distinguishes the scatter plots from each other.

Tasks:
Easily determine schools in Boston area that do/do not offer CS. (shown at top)
Identify Underrepresented communities in Boston area.(shown through metrics in each scatter plot)
Compare schools to others in their county. (shown at top)

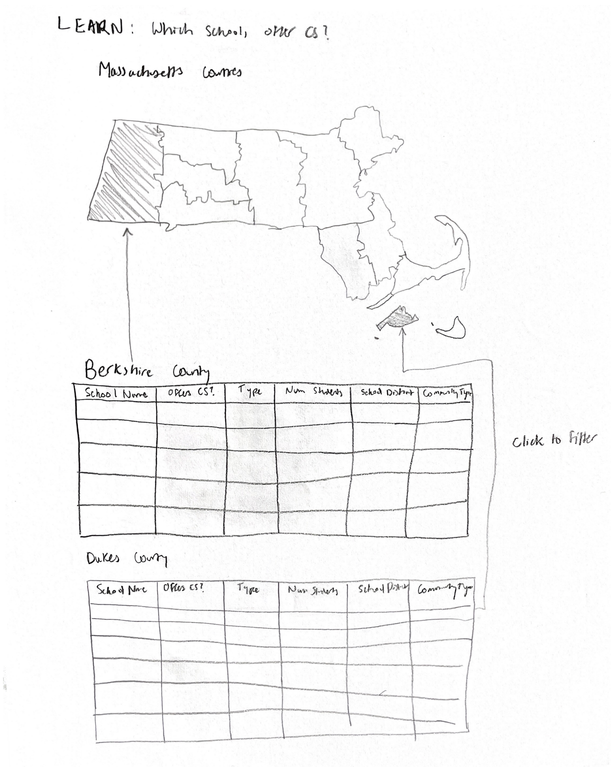


Figure 4. Drawn by: Alina

The marks and channels chosen were marks: points, lines, area and channels: position, size, shape, and color. The encodings are related to the map and table(s). For the map, the marks chosen were areas with gray color channels if selected/hovered over, white color channels if not selected/hovered over. When hovered over, county data appears. When school points are selected, the county data is also linked and displayed. There is also a tab where table data is displayed which is represented in a standard fashion, with uniform rows and columns and a position in the center of the screen. Each county area has position, shape, and size channels that are based on its geographical location. Each table uses keys and values with standardized space filling so that users can easily digest this information. Additionally, the lines pointing from each table to each county help the user easily understand which table and county are associated (space filled completely at each state). The size of each county and the map are determined based on geographic data, and the tables would be responsively resized based on how many counties are selected. The county data is also mapped based on the associated geographical position.

This visualization addresses tasks 1, 2, and 3 because it allows users to click on different counties to see whether or not they offer CS education (task 1) and the tables allow users to easily compare different schools in terms of grade level by filtering the tables by the type column (task 3). Task 2 is addressed by the table as well since people can filter by community type to determine underrepresented communities in the Boston area.

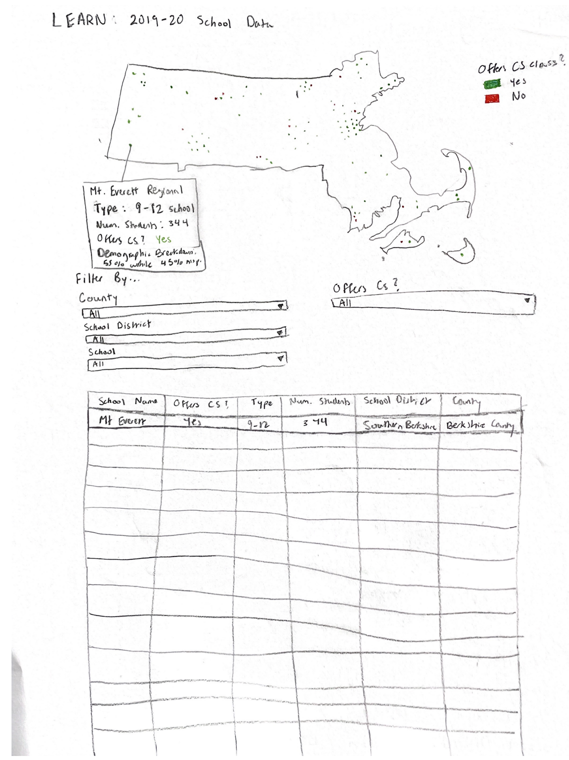


Figure 5. Drawn by: Alina

The marks chosen for the map were points with different color channels (red and green) where green represents schools that offer CS and red represents schools that do not offer CS. The position of each mark is based on geographical data (where each school is located in Massachusetts). All marks are the same size and shape (small, round dots). The actual map shape is based on the way that Massachusetts looks on a standard map. The encodings for map layout space and axis orientation are based on existing geographical data. When points are clicked on/hovered over, a tooltip with information will appear. The table is represented in a standard fashion, with uniform rows and columns and a position in the center of the screen. The encodings are related the the key/value representations used for categorical and quantitative data to help easily display the information. When a school is selected, it becomes a highlighted row on the table.

This visualization addresses tasks 1, 2, and 3 by allowing users to determine which areas do/do not offer CS education with green marks representing schools that do offer CS and red mars representing schools that do not offer CS on the Massachusetts map at the top. Task 2 is covered with tool tips for each school. When a user clicks/hovers over a school mark, they will see a tool tip that has a demographic breakdown of the school, showing the percentage of underrepresented communities. Finally, task 3 is covered by the filterable table that allows users to compare different schools to others in their county by clicking on each column to filter. A user can also click on multiple schools to see the differences between each school.

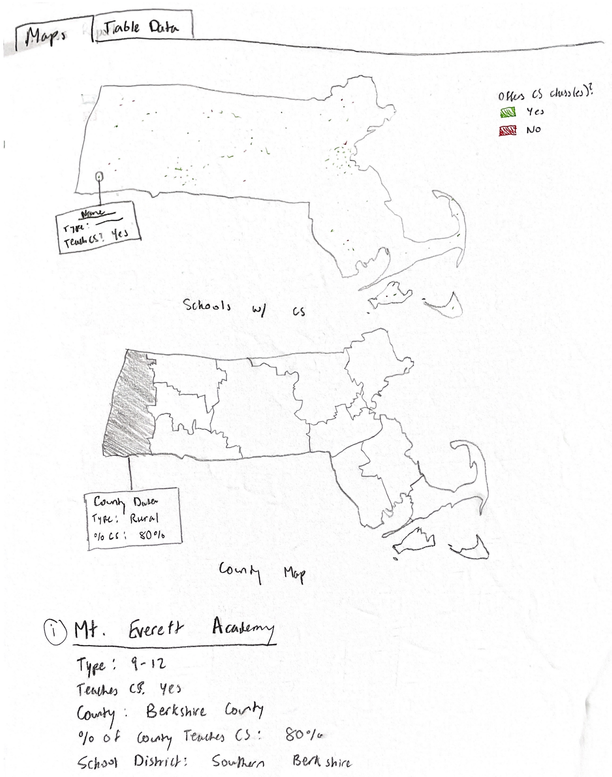


Figure 6. **Favorite** Drawn by: Alina

The marks chosen for the first map were points with different color channels (red and green) where green represents schools that offer CS and red represents schools that do not offer CS. The position of each mark is based on geographical data (where each school is located in Massachusetts). All marks are the same size and shape (small, round dots). The actual map shape is based on the way that Massachusetts looks on a standard map. The encodings for map layout space and axis orientation are based on existing geographical data. When points are clicked on/hovered over, a tooltip with information will appear. The marks chosen for the second map were areas with gray color channels if selected/hovered over, white color channels if not selected/hovered over. When hovered over, county data appears. When school points are selected, the county data is also linked and displayed. There is also a tab where table data is displayed which is represented in a standard fashion, with uniform rows and columns and a position in the center of the screen. The encodings are related the the key/value representations used for categorical and quantitative data to help easily display the information. This combination of marks, channels, and encodings should help the users easily understand the information that is presented to them.

This visualization addresses task 1 by letting users look at a map which displays green markings for when CS classes are offered and red markings for when CS classes are not offered at a school. They can also look at the table data to filter and determine this information as well. Task 2 is addressed because, when a user clicks on a school, it will link to the associated county and show information about the communities in that area. This can also be shown with the table data. Finally, Task 3 is addressed with the filterable table tab which would look similar to the table from the other designs.

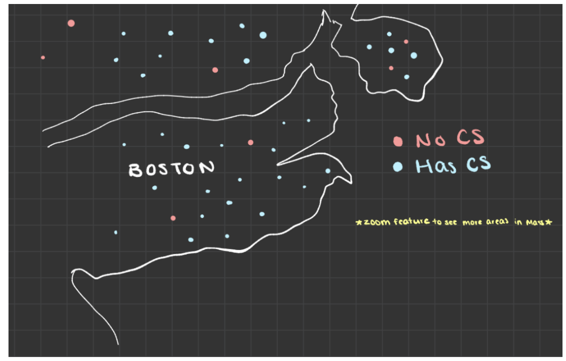


Figure 7. **Favorite** Drawn by: Kreena

The marks chosen for this sketch were points and the channels chosen were position and color (blue and pink). Blue represents the schools that offer CS and red represents the schools that do not offer CS. All the points are the same size and shape. The encodings include the map layout and its ability to be zoomed in and out of. Zooming will change the layout density of the points. These marks, channels, and encoding were chosen to display the data in the most user-friendly and digestible format. It was mainly designed to highlight the difference between schools that offer and do not offer CS.

The task that is mainly addressed in this visualization is Task 1. We can easily visualize which schools in the area offer CS and which do not based on the color of the points. This fulfils Task 1, and can be used as a component in the final visualization.

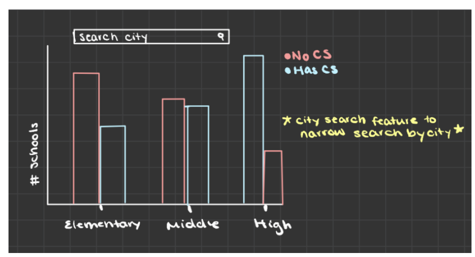


Figure 8. Drawn by: Kreena

The marks chosen for this sketch were areas (bars) and the channels chosen were position (elementary, middle, high), color (blue and pink), and size (length of bars). Blue represents the schools that offer CS and red represents the schools that do not offer CS. Bars that are positioned close to each other are schools of the same age level (elementary, middle, or high). The encodings include the filter for the bar chart and the colors to differentiate the bars. The filter allows the user to search for a specific city in Massachusetts and see its statistics. These marks, channels, and encoding were chosen to display the data in the most user-friendly and digestible format. It was specifically designed to highlight the different statistics by city, and emphasize the difference between the schools that offer CS and the schools that do not.

As such, the tasks addressed are Tasks 1 and 3. This visualization fulfills Task 1 of being able to see the distinction between schools that offer CS and do not offer CS. It does this by color: Blue bars are schools that offer CS and pink bars are schools that do not offer CS. It also fulfills Task 3 of being able to see which age group schools teach. This is done by separating the bars by the type of school (elementary, middle, and high school). This idea of a filtered bar chart can be applied as a component in the full visualization.

Figure 9. Drawn by: Kreena

The marks chosen for this sketch were points (representing a community) and the channels chosen were position, color (blue and pink), and size (of schools). Blue represents the schools that offer CS and pink represents the schools that do not offer CS. The encodings include the zoom feature on the map that can alter the layout density of the points. Additionally, there is a filter for community type (suburban, city, and rural). These marks, channels, and encoding were chosen to display the data in the most user-friendly and digestible format. It is designed so that the areas with the most schools easily catch the attention of the user. It is also very easy to tell the difference between schools that offer and do not offer CS.

The task that is targeted the most in this visual is Task 1. We can easily tell the difference between schools that offer and do not offer CS based on the color of the points.

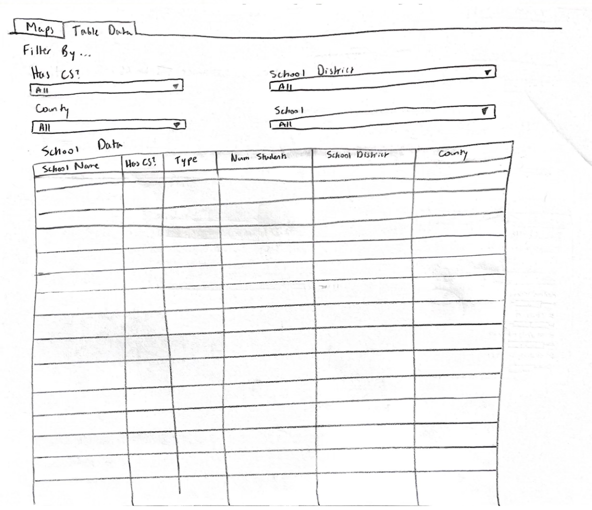
Favorite Sketches

Our first favorite sketch (Figure 1) was chosen because we like how it displays both the box and whisker plots of each relevant statistics and a pie chart detailing the ethnicity breakdown of each school. This allows the user to see the most amount of information possible at once very efficiently. The visual encodings are very effective; The use of color to differentiate the box plots from each other allows for clear distinction between the statistics, and the way the dots for each school are different colors based on if they are public or charter make it easy for the viewer to tell what kind of school it is.

Our second favorite sketch (Figure 6) was chosen because we like the use of tabs as a way for the user to explore the dataset in different forms. We also felt like the use of points and color channels for showing if a map offers CS classes was helpful and could be easily understood by users. Additionally, having school specific and county specific data allows for the user to interpret and compare data easily and helps accomplish our domain tasks

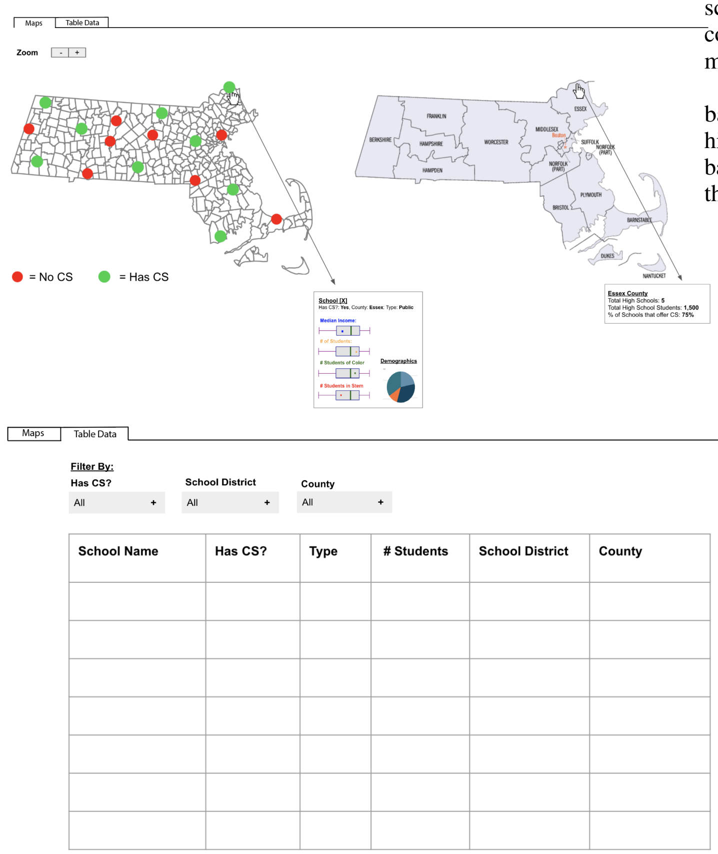
Our third favorite sketch (Figure 7) was chosen because the usage of color channels to represent whether or not a school has CS education is incredibly clear and would help users understand the primary goal of our visualization which is to determine which schools offer CS education. Additionally, a zoom feature would also allow users to easily explore specific counties in Massachusetts based on their interests.

Hand Drawn Design for Entire Coordinated View Visualization Tool



The marks chosen for the left map were points with different color channels (red and green) where green represents schools that offer CS and red represents schools that do not offer CS. The position of each mark is based on geographical data (where each school is located in Massachusetts). All marks are the same size and shape (small, round dots). The actual map shape is based on the way that Massachusetts looks on a standard map. The encodings for map layout space and axis orientation are based on existing geographical data. When points are clicked on/hovered over, a tooltip with information will appear. The information is represented with box plots with different colors so that each box plot is easily distinguishable and a pie chart with areas that are colored differently to represent each race. The marks chosen for the right map were areas with gray color channels if selected/hovered over, white color channels if not selected/hovered over. When hovered over, county data appears related to the number of schools that offer CS, the number of students in the county, and the number of schools in the county. When school points are selected, the county data is also linked and displayed. When a school is selected, the county associated with it has brushing and linking implemented so that the county and school data are shown at the same time. When zoom is pressed, both maps zoom in and out and can be dragged/explored. Maps were chosen so that the data was easily recognizable and could be explored based on positional data. The maps were placed side-by-side and are the same size so that they could be easily compared. They also have no axial tilt because tilted maps are a bit harder to view clearly. There is also a tab where table data is displayed which is represented in a standard fashion, with uniform rows and columns and a position in the center of the screen. The encodings are related the the key/value representations used for categorical and quantitative data because then schools with similar backgrounds can be compared in another format. This visualization addresses task 1 by letting users look at a map which displays green markings for when CS classes are offered and red markings for when CS classes are not offered at a school. They can also look at the table data to filter and determine this information as well. Task 2 is addressed because, when a user clicks on a school, it will link to the associated county and show information about the communities in that area. There will also be information related to racial demographics, median income, and more. This can also be seen in the table data tab. Finally, Task 3 is addressed with the school tool tip that compares the school to other schools and the filterable table tab which lets users look at different schools all at once and compare them.

12 Appendix E: Digital Sketch



The user, a FirstByte club member, would open LEARN in search for a new school to send a kit to. Suppose they are looking for a school that is specifically low income and majority underrepresented minority students that has a CS program. The first view they see when they open LEARN will be the side by side map views of the School Map and the County Map; they will hover over one of the maps with their cursor. Displayed next to the School map will be a box containing information about the school (represented by a dot on the map) that their cursor is hovering over. In that box will be information about if the school has a CS program, the county it is located in, and the type of school it is, as well as 4 box and whisker plots each with the dot in them that indicates where in each statistic the selected school lies, (the 4 statistics include median income, number of students, number of students of color, and number of students in stem, as well as a pie chart that represents the racial distribution of the school.

At the same time, beside the county map will a box displaying information about the county that the school selected in the previous map is in. The information that will be displayed in that box will be the number of high schools, the total number of high school students, and the percent of schools that offer CS in the county. Alternatively the user could just over over a county in the county map and only the county information will be displayed.

The user can also navigate to the table view tab where they can filter which the schools they want to see the name, if they have CS, type, number of students, school district, and county based on if they have CS, their school district, and their county.

After the user has found a school that is specifically low income and majority underrepresented minority students that has a CS program by using LEARN to compare different schools, they can then contact that school and the tool has successfully served its purpose.

They have then accomplished the task of determining which schools have and do not have CS education by seeing the color coded dots on the map, and identified schools with underrepresented minorities by hovering over schools and seeing their statistics.

The one task we changed was the ability to compare the schools based on others in their county since the club was looking only at high schools. We have since changed this one to comparing schools based if the type of school is public or charter, which is displayed in the information box while hovering over a school.

13 APPENDIX F: USABILITY TESTING

13.1 Preparation

Introduction/Overview of LEARN

Jobs in technology are appearing faster than any other fields and paying much more, but unequal access to computer science education in Massachusetts public high schools is still an issue. LEARN is a visualization for educators and members of Northeastern University's club, FirstByte, who hope to identify public schools in Massachusetts that do not teach computer science or do not provide equal access to computer science education. LEARN's intended use is to allow individuals to easily determine which schools in Massachusetts do/do not offer computer science education, to identify underrepresented communities in the Massachusetts schools, and to be able to compare a school's demographics to other schools in their county. The data it showcases is from Code.org's dataset that includes national and state-level computer science education policy, including policy trends, maps, state summaries, and implementation data.

Three Specific Tasks

- **Task 1: Navigate to and Explore Table Tab, Search Table by School Name, Filter Table by Offering CS, Type of School, County, and Number of Students, Highlight Rows**
We are testing this task to make sure that the subject understands how to switch from one tab to another to get to the table view, that they think having multiple tabs for the data is practical, that the tabs work correctly and the table displays all data correctly, and that the search by school name function works correctly and is efficient. Some specific outcomes we would like to look for include: the user will be able to easily navigate to the table view using the tabs and will be able to scroll through and compare at least two schools to one another, the amount of data displayed is not too overwhelming for the user, the visual encoding is effective at conveying and comparing data, the user should search for at least one school by name, and the user should find the process of searching for a school to be simple and for the data lookup to be efficient. We are testing filtering and highlighting as well to make sure that filtering tables by various categories works because we want to make sure that the subject can easily navigate and understand our user interface and to make sure that the data is queried efficiently since it is a relatively large dataset. We also want to test that rows can be highlighted when hovered over before and after filtration because we want to make sure it works and that the feature is useful for interpreting data. Some specific outcomes we would like to look for include: the user will hover over at least one row that will be highlighted on its hover state, the marks and channels are pleasing to the user and make sense, the user understands how to navigate our the filtration tabs easily, a user will use at least one filter button to filter the table data, the user thinks that this encoding is effective, and, when a user filters that the data with multiple queries, it is displayed efficiently.

- **Task 2: Explore Point Map and Determine which Schools Offer Computer Science Education**
We are testing this task to make sure that our chosen encoding is effective for showing which schools offer/do not offer computer science. We would also like to make sure that the user can easily explore the map and understands the color channels that we chose for our points. Finally, we want to make sure that the point map is interactive and links to the respective county associated with the school hovered over. Some specific outcomes we would like to look for include: user can easily identify which map is the point map, user can easily identify which schools offer/do not offer CS, user understands why this particular encoding was used, user can hover over points to learn more about each school, user can use brushing and linking to further explore point and county maps.

- **Task 3: Explore County Map to See Tool Tips With County Data**
We are testing this task because we want to see that a user will intuitively hover over various counties, that the tooltips will display county data pertaining to CS education and demographics, and that it is easy to compare and interpret the side-by-side data. Specific outcomes that we are expecting include: the user will hover over and read about at least three Massachusetts counties, the user will be able to compare the county data to the data shown on the point map to its left, and the user will know which county corresponds with the data shown based on which county is highlighted.

13.2 Results

The results of our in-class usability testing indicated the that the participants who reviewed our visualization were overall satisfied with its capabilities. Note that questions were asked about the functionality of the visualization once it is fully implemented while our current implementation is unfinished.

For task one, which was to Navigate to and Explore Table Tab, Search Table by School Name, Filter Table by Offering CS, Type of School, County, and Number of Students, Highlight Rows, the participants indicated that it is a useful functionality for identifying schools that fit a certain criteria. The metrics we were looking for told us that if the user was able to easily navigate the tab and search functionality to filter the table to find results, they would have a successful experience with the product. The participants indicated that it would be useful to be able to filter the table by multiple criteria at once, therefore, we will change our design to accommodate for this.

For task two, which was to Explore the Point Map and Determine which Schools Offer Computer Science Education. The participants indicated that is feature was useful because it lets them easily identify which schools have and do not have CS education at a glance, and then see basic information about them that they can use to find information about specific schools in the table. The metrics were looking for told us that the user if the user was to easily find which schools have and do not have CS through our marks and visual encodings then it would be successful. The participant indicated that it would be useful to have a feature where the you can chose to only display dots that have CS and only display dots that so not have CS, therefore, we will change our design to accommodate for this. The participants also noted that the tooltips had a lot of information.

For task three, which was to Explore the County Map to See Tool Tips With County Data, The participant indicated that this was useful especially because of the bushing and linking with the map next to it. The metrics were looking for told us that if the user was to hover over and read about Massachusetts counties, be able to compare the county data to the data shown on the point map to its left, and know which county corresponds with the data shown based on which county is highlighted, then the visualization would be successful. The participants did not indicate that any changed should be made to this portion of the visualization since it performs its task proficiently.

Based on the usability testing, we noted that there were no major issues with our overall visualization design because the overall layout made sense to the users. However, there were some minor issues based on specific user feedback that we would like to address. For instance, we plan to modify our visualization based on specific feedback. For the table portion of our visualization, we plan to add functionality to filter the table by multiple tab criteria at once to enable the user to select a more specific and detailed subsection of schools. For the first map visualization, we plan to add the functionality for the user to be able to chose to display all the schools, or only the schools with our without a CS program. We also plan to separate the brushing and linking into a separate map from the point map with tooltips since having both brushing and linking and tooltips on one map was overwhelming for the users and to reduce the amount of content on the tooltips to just text summaries. One thing that the users enjoyed was the point map with the tooltips of information, so we will maintain the overall structure of this map. In terms of dislikes, the participants did not like the fact that the brushing and linking and tooltips had a bit of lag on the point map before loading, so we will work on making the load time more efficient.