Good Intentions, Uncertain Results:

An examination of Oakland Unified School District's suspension rates, racial disparities.

ARIJIT SEN

Introduction

There is a long-documented history of racial discrimination in school discipline. Black kids and in particular black boys are suspended at much higher rates than their White classmates. This disparity not only removes children from instructional time it also has numerous other pernicious effects including lower college attendance and academic performance. In fact, one study from University of Kentucky researchers concluded that a fifth of the Black-White achievement gap could be attributed to suspension disparities.[1]

Many would suspect that Oakland Unified School District would buck this trend. To address this issue, the OUSD has invested millions of dollars on several programs - they created an office catering just to Black male achievement[2], implemented restorative justice[3] and positive behavior intervention programs[4] and banned suspensions for "willful defiance"[5].

But despite historic efforts and millions of dollars invested, the suspension rate and racial disparity in suspensions has stayed basically flat in the past five school years.

Project Goals

I wanted to create a dynamic, interactive and attention-grabbing visualization to introduce this issue to an audience of news readers. The intent for this visualization is to serve as a supplement for an in-depth investigative story on the issue of suspensions in Oakland schools, which could be pitched to major regional and national journalistic publications like the *New York Times, Los Angeles Times, San Francisco Chronicle* or ProPublica.

Design Inspiration/Related Work

[1] The Punishment Gap: School Suspension and Racial Disparities in Achievement
[2] African American Male Achievement / Frequently Asked Questions
[3] Restorative Justice / Homepage
[4] Behavioral Health Services / Positive Behavioral Intervention and Support (PBIS)
[5] Oakland to halt school suspensions for willful defiance

In support of the project goals, I examined some of the visualizations from many of the top journalistic publications and noticed a trend among them: the use of "scrollytelly" interactive visualizations.

This relatively new type of visualization engages readers by progressively revealing information as the user scrolls through a webpage. I thought about how this design could be used for my purposes and came up with the idea of creating a map with color-coded dots representing each school in the district and its suspension rate. As the user scrolls through the design the dots change color to represent the suspension rate for the labeled school year.

In my research for this project, I discovered three pieces of journalistic work that served as design and content inspiration.

The first was "Failure Factories,"[6] a Pulitzer-Prize winning investigative series published by the *Tampa Bay Times*. In staggering detail, the reporters detail the abject failure of five elementary schools to educate students and keep them safe. Their findings include the fact that 95 percent of Black students are failing reading or math and more than half of teachers requesting a transfer from the schools in a single year. But, most relevantly to my project, in the fourth part of the series, the reporters document how Black students in the school district lost a combined 45,942 days of instructional time in five years because of suspensions.[7] Obviously Oakland Unified is not the same as Pinellas County schools. But, with their tireless reporting and endless hours of document and data analysis, the reporters at the *Tampa Bay Times* set the standard for what an investigation of school suspensions could be.

The second inspiration was ProPublica's collaboration with *The Times-Picayune* and *The Advocate* on an investigation into chemical plant pollution.[8] Although the subject matters are totally unrelated, when I saw the visualization created by ProPublica I knew it was a great model for what I wanted to achieve with my visualization. As users scroll through the page, points representing chemical plants appear on a map of southeastern Louisiana and change color and size to represent different variables. Explanatory text scrolls over the map on the left and right sides to explain the different sections to readers/viewers. Occasionally as the users scroll by a certain section of the map is zoomed in on to highlight readers/viewers' attention. Upon examination of the underlying code, it appears they used the same methods and tools (Mapbox GL, Scrollama) that I did to construct my final visualization.

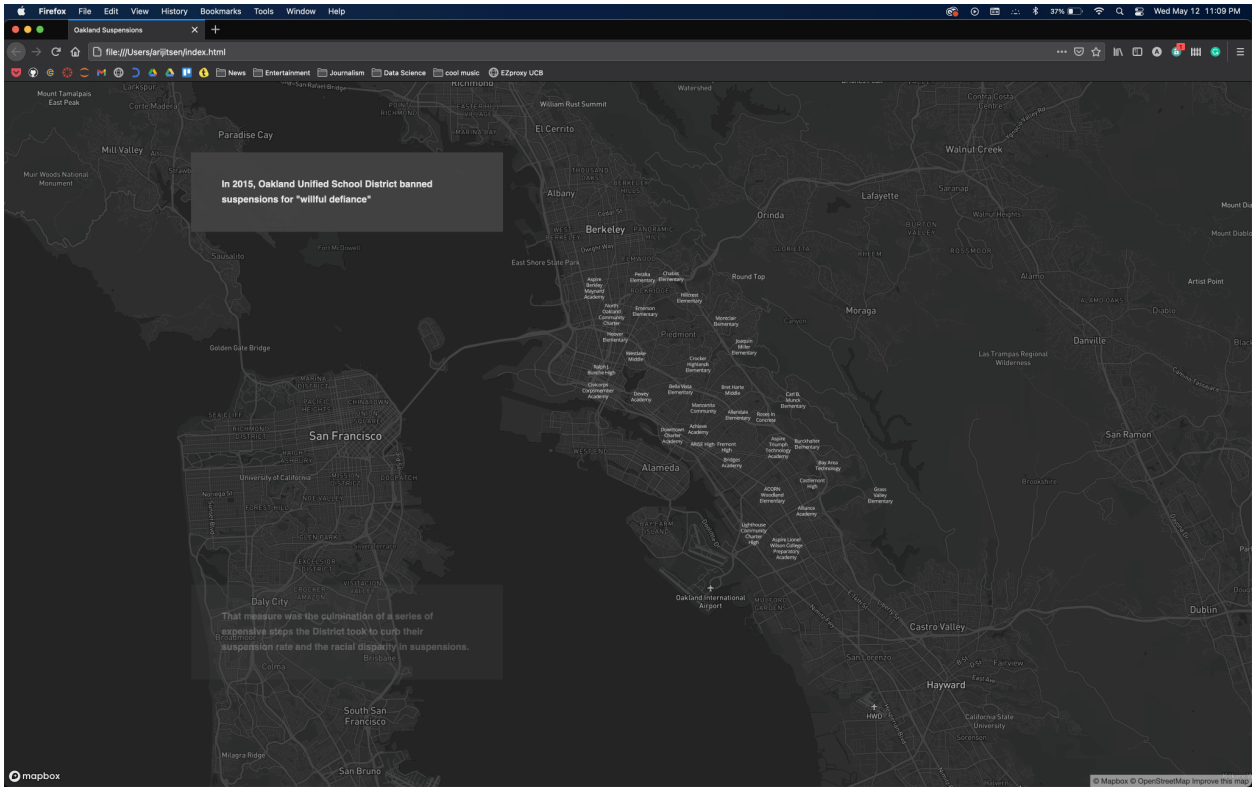
The final inspiration was the Reuters' multimedia story "Life in the Camps," which tells the story of the poor sanitation of Rohingya refugee camps in Bangladesh.[9] Roughly halfway through the page users are presented with a map of the camp, overlaid with orange dots representing latrines and blue

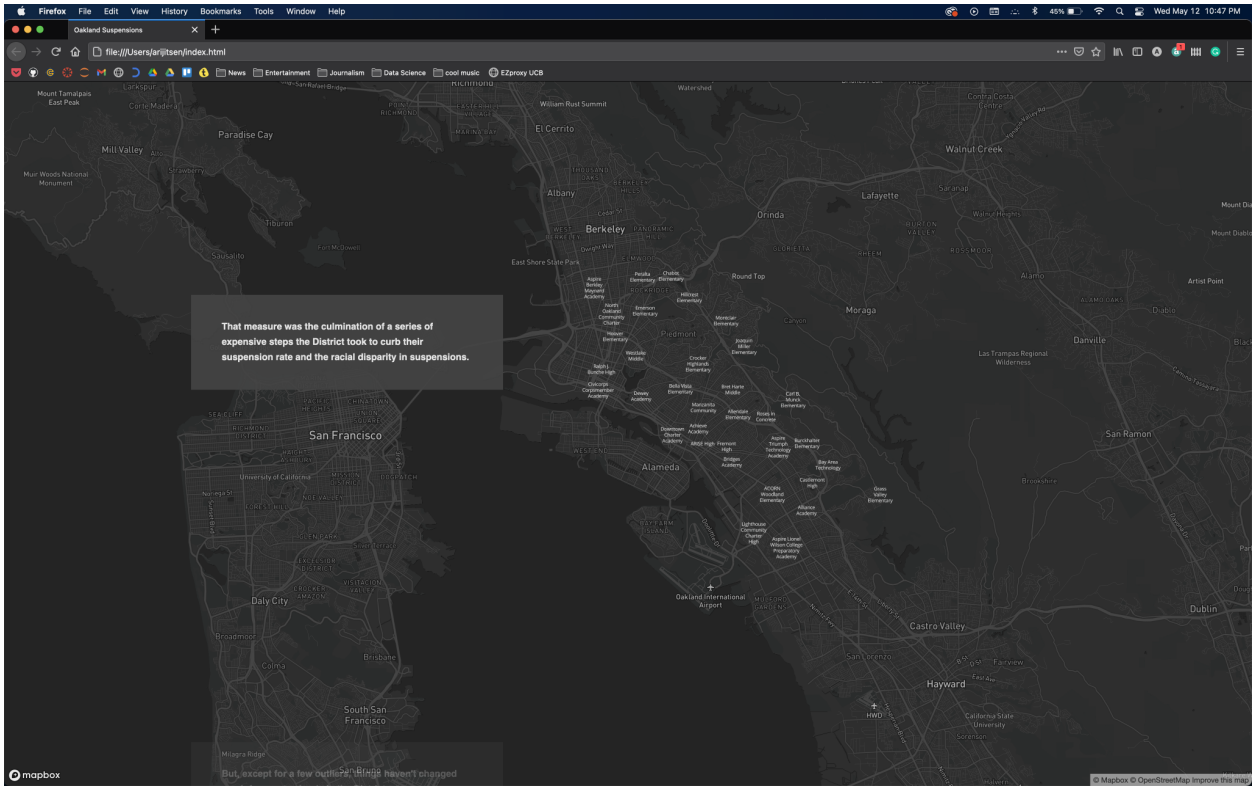
[6] Failure Factories | Tampa Bay Times
[7] Pinellas suspends black kids more than virtually every other big Florida county
[8] In a Notoriously Polluted Area of the Country, Massive New Chemical Plants Are Still Moving In
[9] Life in the camps

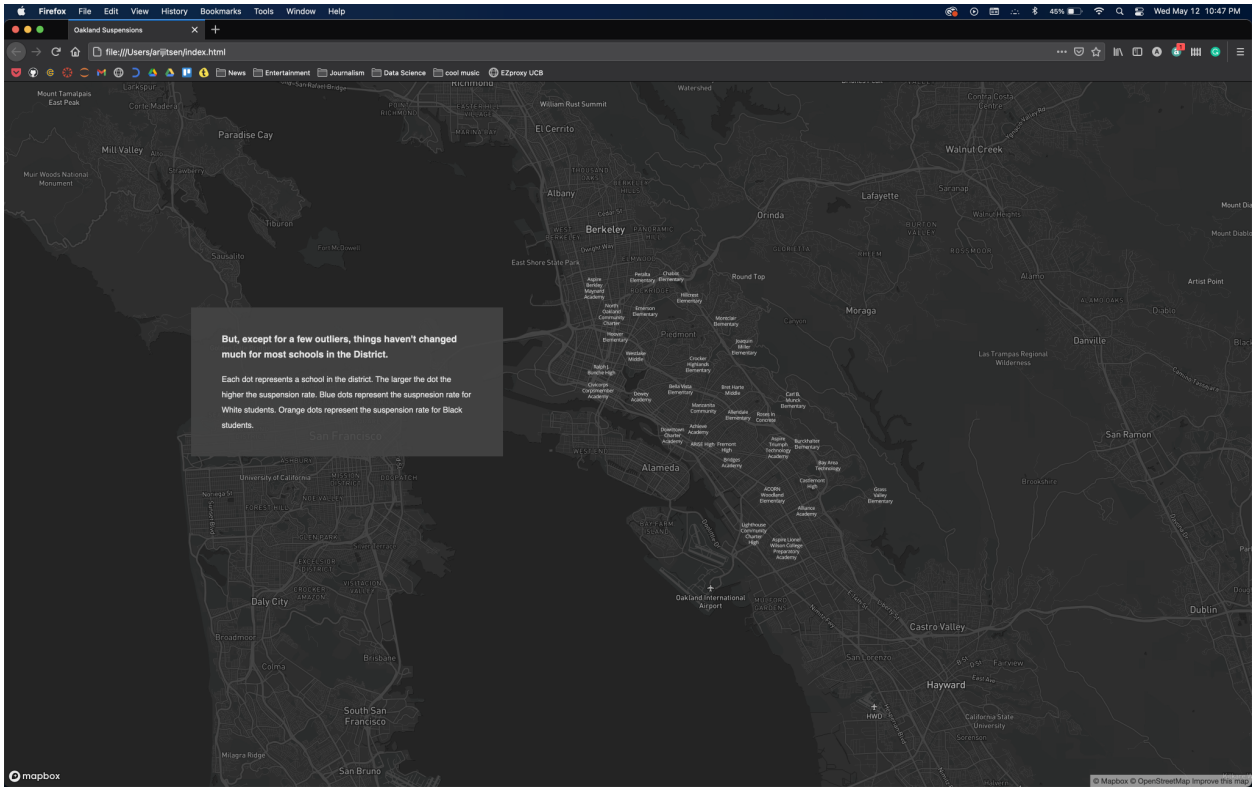
dots representing water pumps. The relative position of these dots are used to convey sanitary conditions— if a blue dot was too close to an orange dot there was a large possibility of cross-contamination. On inspection of the code, the Reuters visualization was also built with Javascript, but with slightly different libraries (Scrollmagic as opposed to Scrollama).

Description

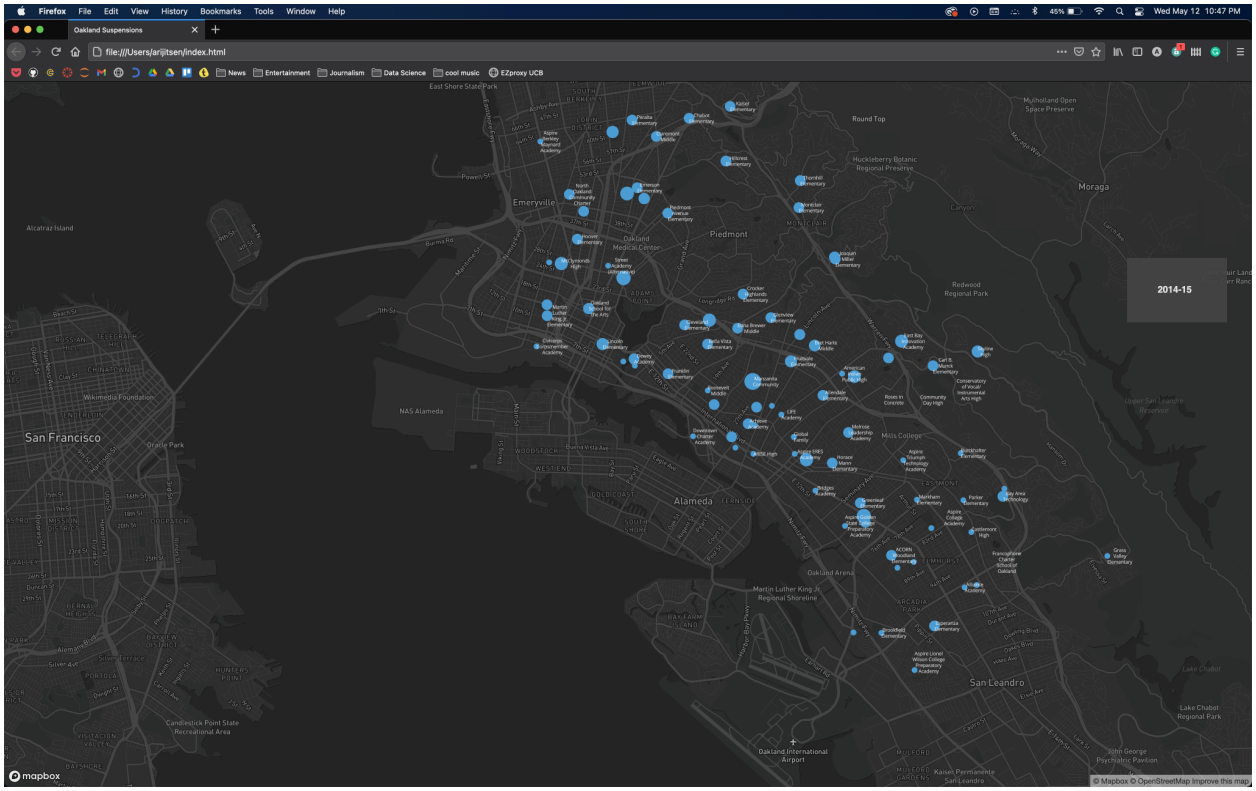
The final visualization is a web-hosted "scrollytelly" geo-visualization, consisting of 14 panels or chapters. The first three chapters consist of explanatory text which appears on the left side of the screen.

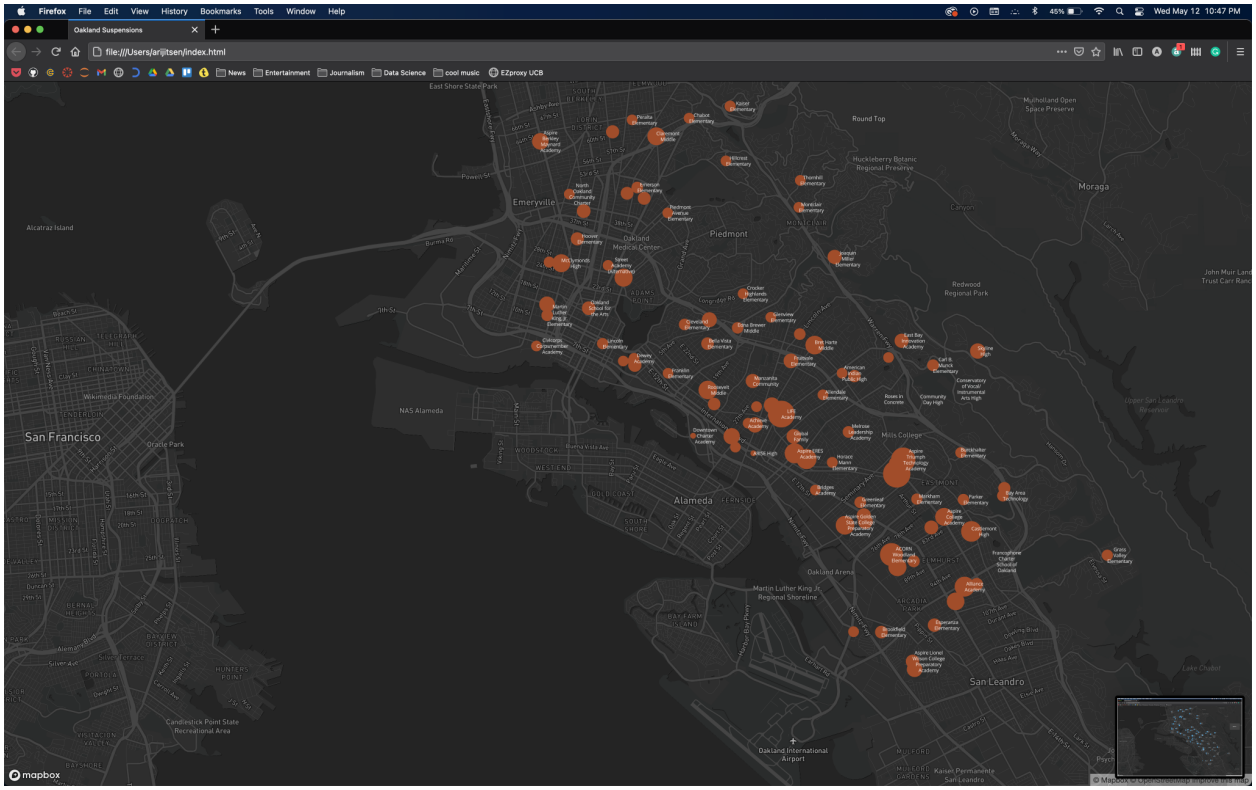


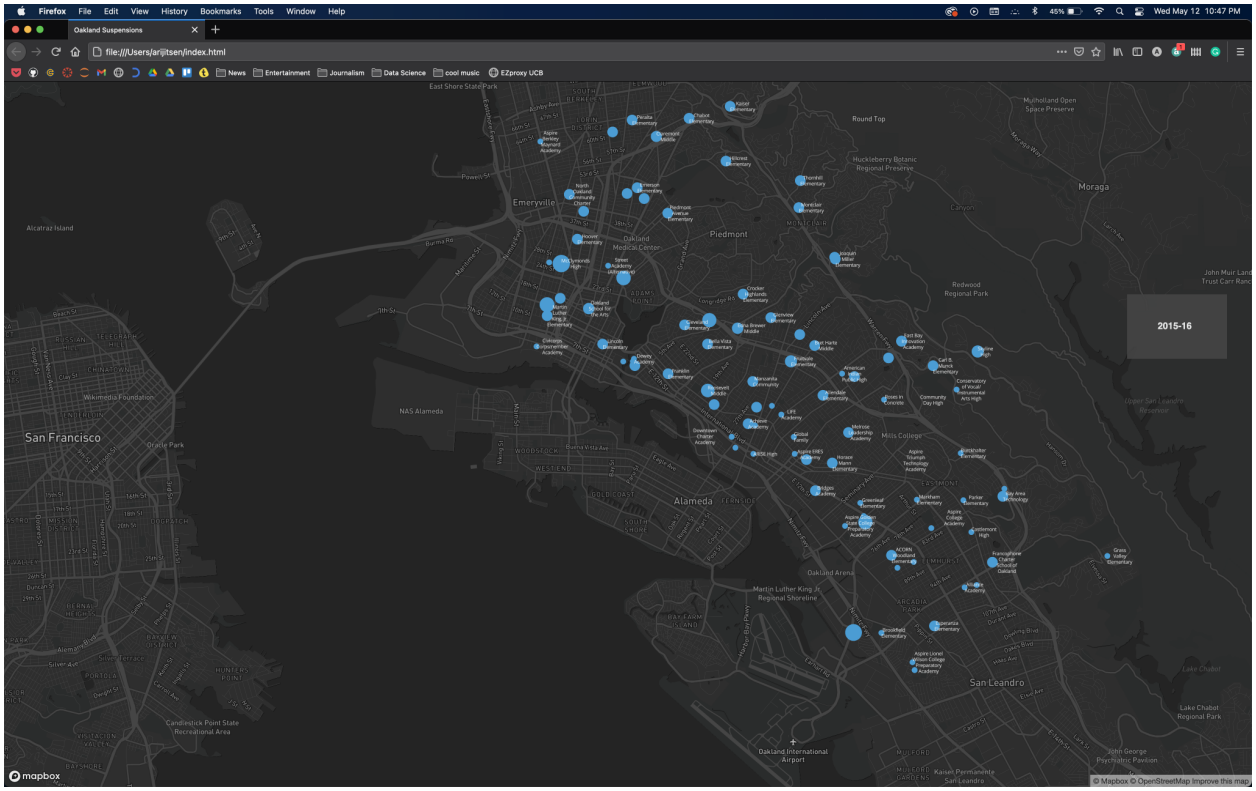

Oakland Suspensions
file:///Users/arjitsen/index.html
That measure was the culmination of a series of expensive steps the District took to curb their suspension rate and the racial disparity in suspensions.
But, except for a few outliers, things haven't changed
mapbox

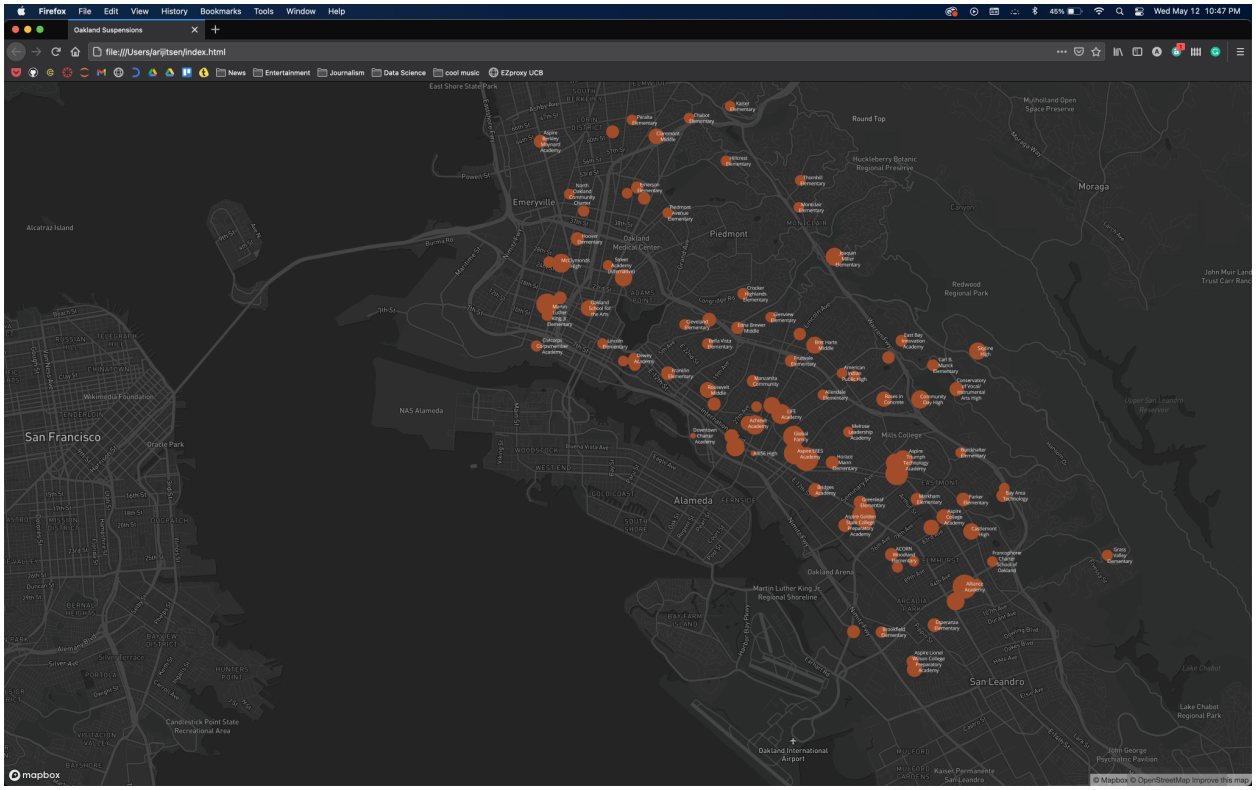

Oakland Suspensions
file:///Users/arjitsen/index.html
But, except for a few outliers, things haven't changed much for most schools in the District.
Each dot represents a school in the district. The larger the dot the higher the suspension rate. Blue dots represent the suspension rate for White students. Orange dots represent the suspension rate for Black students.
mapbox

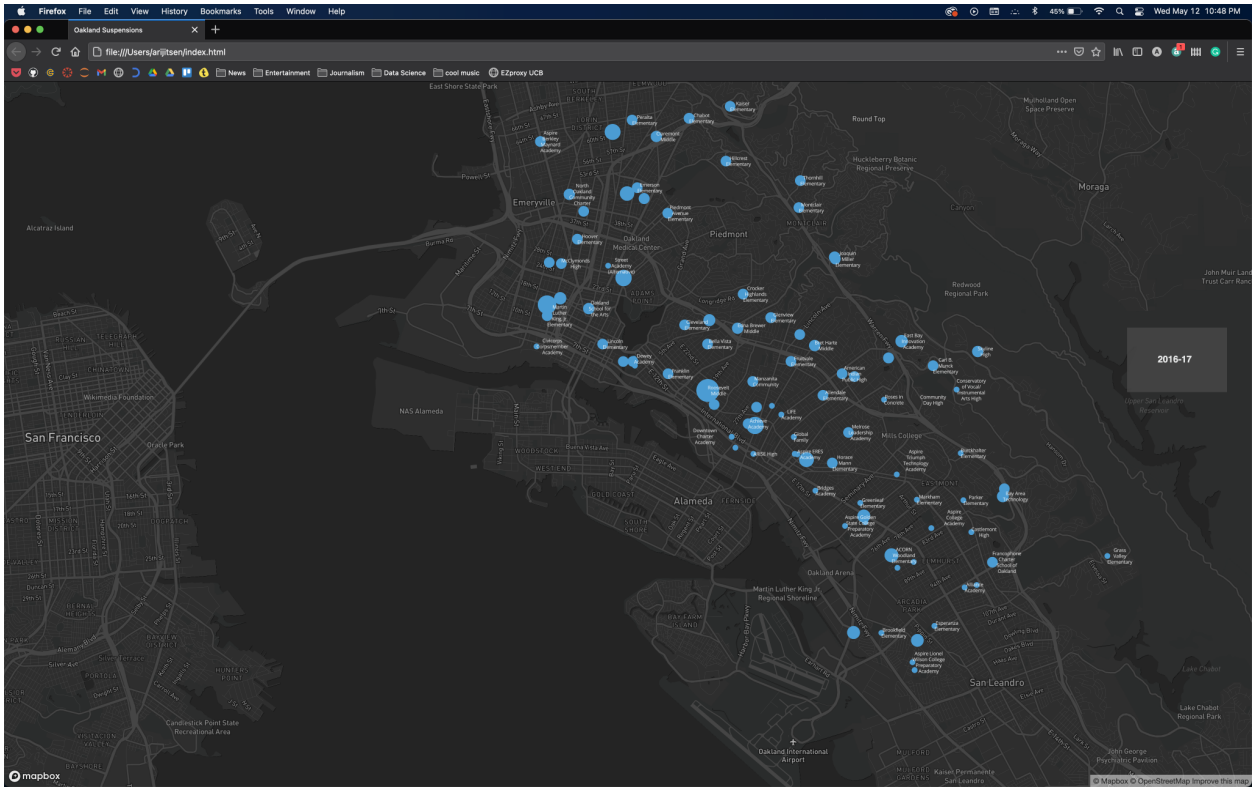
Then the map zooms in and the dots appear. For each school year, there are two sets of visualization— one with blue dots representing the White suspension rate and one with orange dots representing the Black suspension rate.

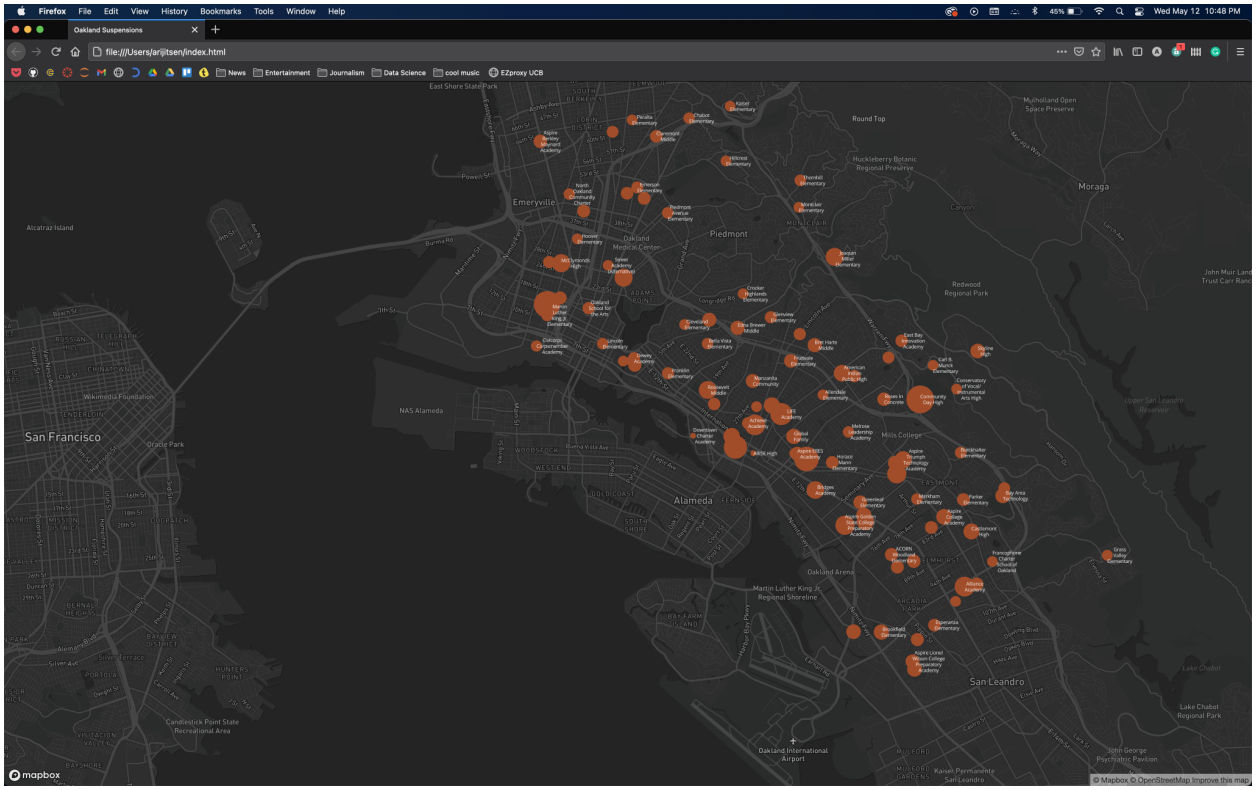


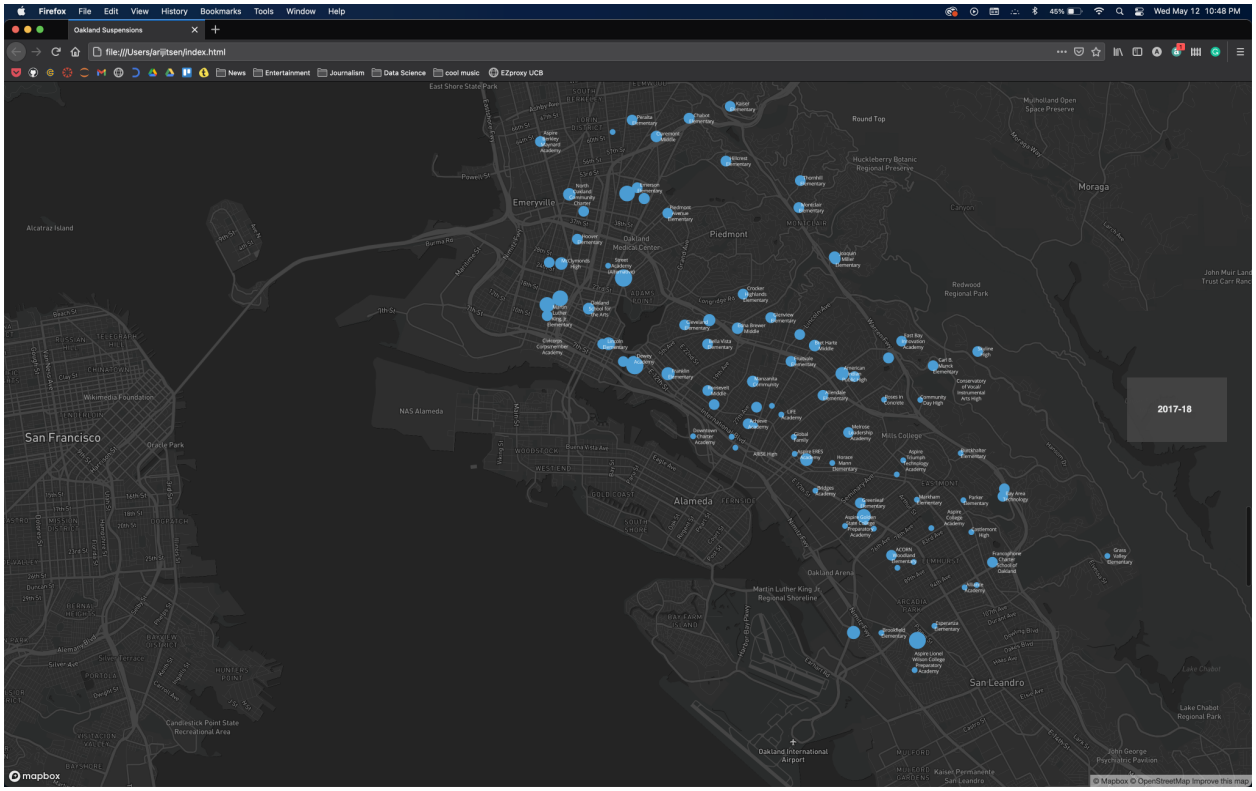

Oakland Suspensions
file:///Users/arjitsen/index.html
San Francisco
Alameda
San Leandro
Emeryville
Piedmont
Oakland Arena
Oakland International Airport
mapbox

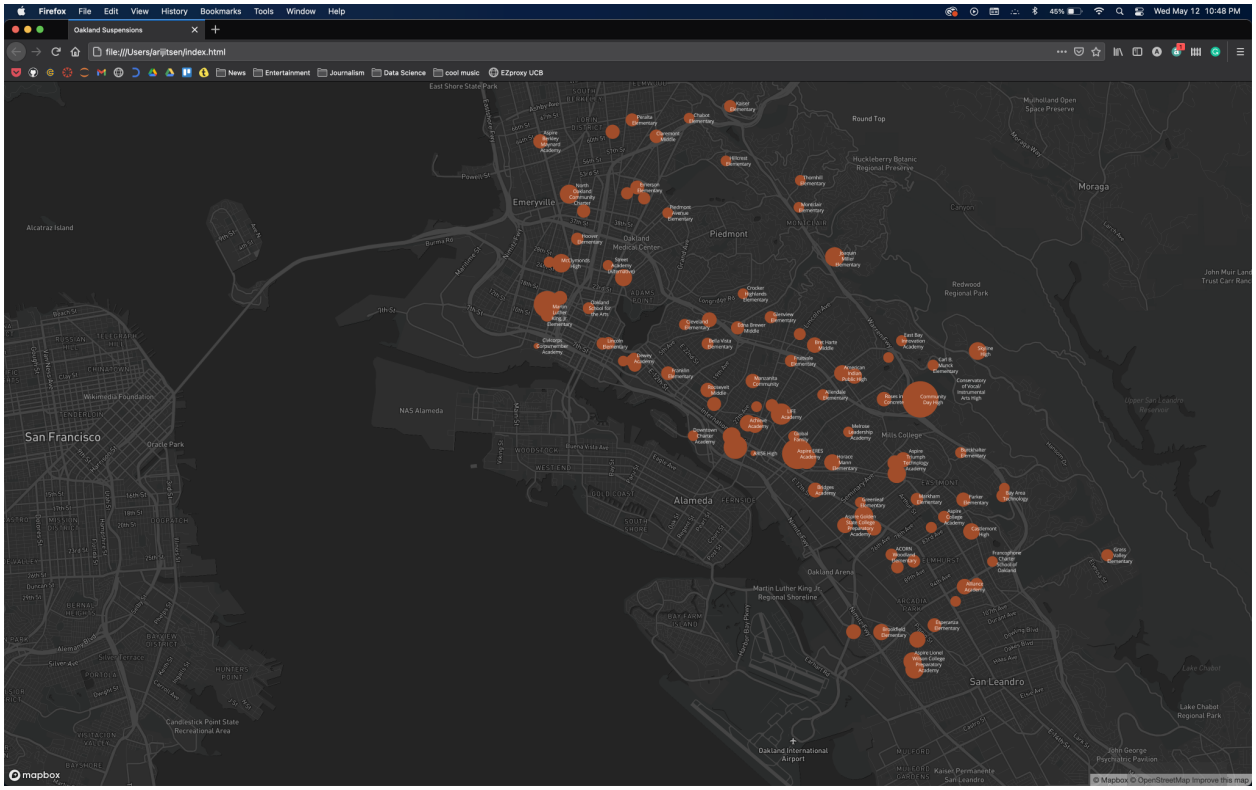

Oakland Suspensions
file:///Users/arjitsen/index.html
2015-16
San Francisco
Alameda
San Leandro
Emeryville
Piedmont
Oakland Arena
Oakland International Airport
mapbox

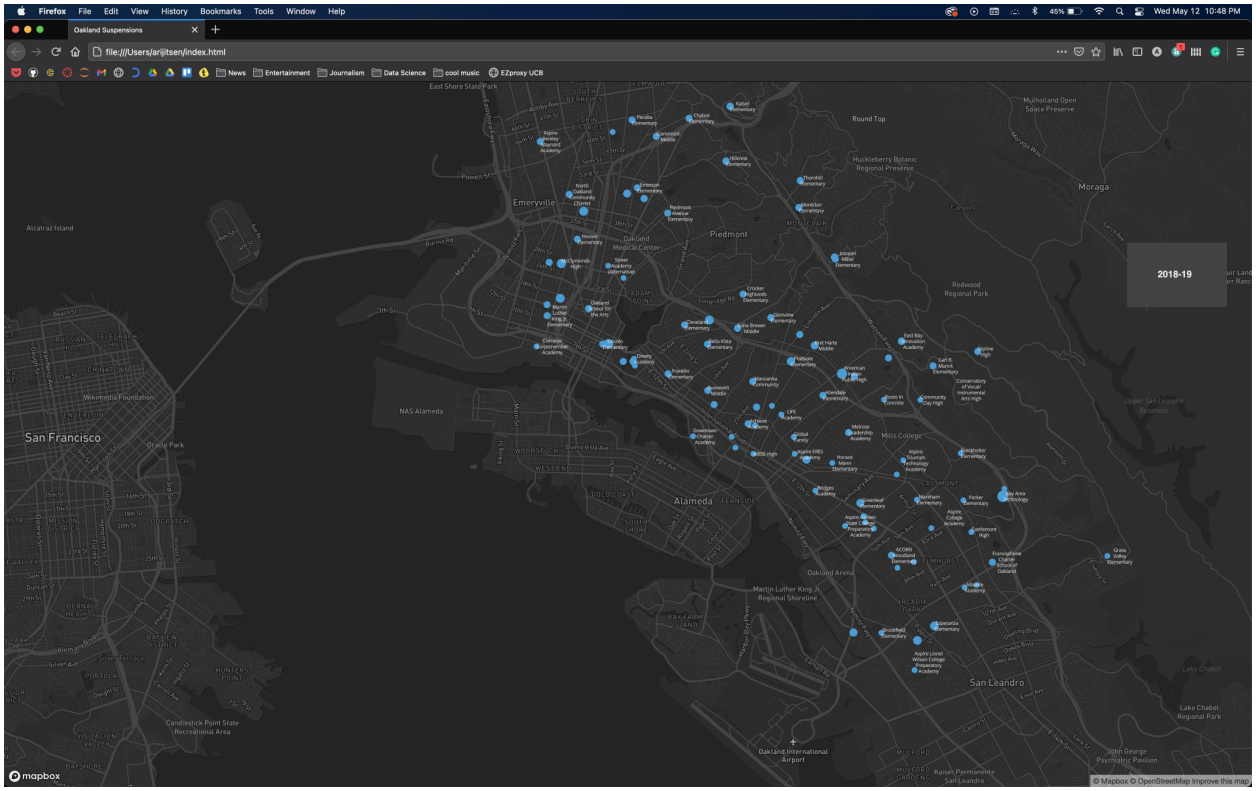

Oakland Suspensions
file:///Users/arjitsen/index.html
San Francisco
Alameda
San Leandro
Emeryville
Piedmont
Oakland Arena
Oakland International Airport
mapbox

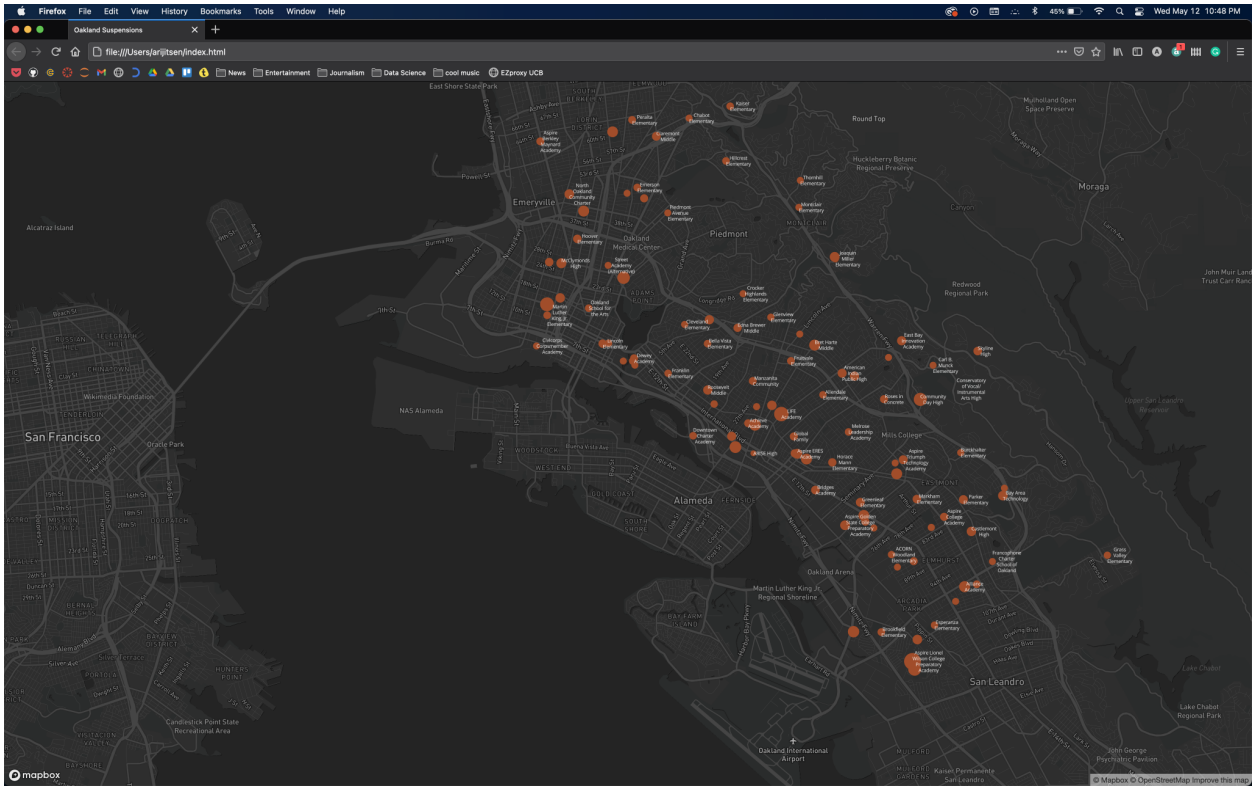

Oakland Suspensions
file:///Users/arjitsen/index.html
2016-17
San Francisco
Alameda
San Leandro
Emeryville
Piedmont
Oakland Arena
Oakland International Airport
mapbox


Oakland Suspensions
file:///Users/arjitsen/index.html
San Francisco
Alameda
San Leandro
Emeryville
Piedmont
Oakland Arena
Oakland International Airport
mapbox

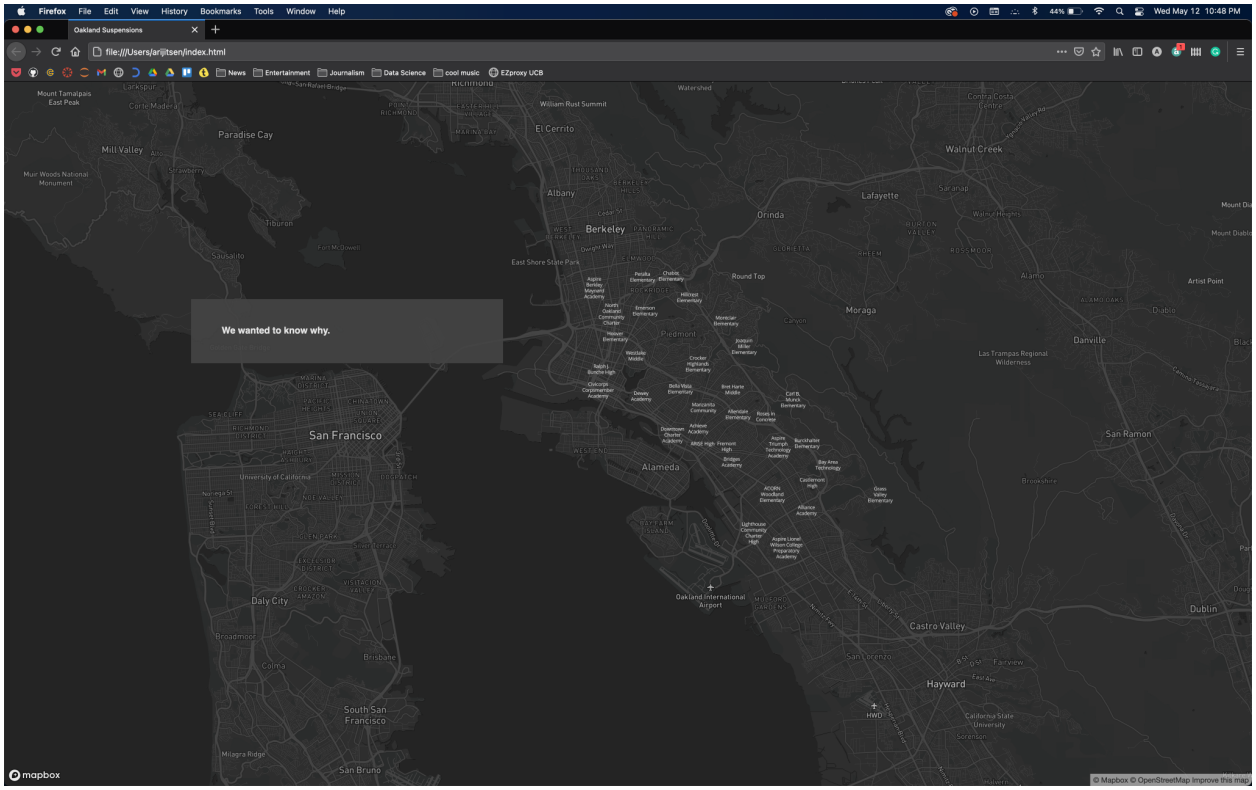

Oakland Suspensions
file:///Users/arjitsen/index.html
2017-18
San Francisco
Alameda
San Leandro
Emeryville
Piedmont
Oakland Arena
Oakland International Airport
mapbox


Oakland Suspensions
file:///Users/arjitsen/index.html
San Francisco
Alameda
Emeryville
Piedmont
Oakland Medical Center
Oakland International Airport
San Leandro
Moraga
Round Top
Redwood Regional Park
Mills College
Oakland Arena
Martin Luther King Jr. Regional Shoreline
Lake Chabot Regional Park
mapbox


Oakland Suspensions
file:///Users/arjitsen/index.html
2018-19
San Francisco
Alameda
Emeryville
Piedmont
Oakland Medical Center
Oakland International Airport
San Leandro
Moraga
Round Top
Redwood Regional Park
Mills College
Oakland Arena
Martin Luther King Jr. Regional Shoreline
Lake Chabot Regional Park
mapbox



This is followed by an attention-grabbing final piece of text on the left which would lead the reader into the written.



Data & Tools

A variety of methods were used to obtain, clean, analyze and visualize the data necessary for this project.

I obtained school suspension data from the California Department of Education website using a Selenium scraper written in Python. The script scraped five school years' worth of suspension data (2014-15 to 2018-19) from every school in Oakland Unified School District. These five years were chosen because they were the most recent data available at the outset of my research and they begin the year that willful defiance suspensions were banned. The data come from each school's self-reported count of suspensions for each year. The data is displayed on the California DoE website as HTML tables but was transformed into CSVs by the web scraping script. Unfortunately, some of the enrollment and suspension data is redacted because of an overly broad and legally dubious interpretation of the Federal Education Rights and Privacy Act.

Once I acquired the data for each school for each school year with the scraper, I unioned all the tables for the respective school year together with DB Browser.

To get addresses for each school, I used a Python script to feed the name of every school in the 2014-15 sheet concatenated with the string "oakland" into the Google Maps API. These addresses

were then saved into a list object in Python and appended as a column next to the corresponding school in a Pandas dataframe made from that school year's sheet. This dataframe was then exported as a CSV and used as a template for the addresses in the proceeding school years. For the small number of schools that did not exist in the 2014-15 school year or where the Google Maps API did not work, I searched for the address manually using Google Maps or the OUSD website and inputted it into the spreadsheet. Each of the five sheets was then given a column containing the respective school year.

I then inputted additional variables about each school (charter vs public, grades served etc.) manually using information from the OUSD website and greatschools.org. All five of these sheets were then unioned together again using DB Browser. Finally, I inputted the master sheet into Geocodio to get precise latitude and longitude coordinates for each school.

To produce the scrolling visualization, I uploaded the master sheet to Mapbox Studio and created 15 layers (one for White students, one for Black students and one with the text labels for every school year in the dataset) by filtering the data on the "Ethnicity" and "school_year" fields of the database and then styled them. Using templates provided by the Mapbox team, I then modified some HTML, CSS and JavaScript code to take advantage of the styled base map and layers. The Javascript code uses the Mapbox GL JS, D3 and Scrollama libraries to construct the final "scrollytelly" visualization. The final visualization was hosted on GitHub Pages. Earlier iterations of this visualization were created in D3, Tableau and Carto.

Results

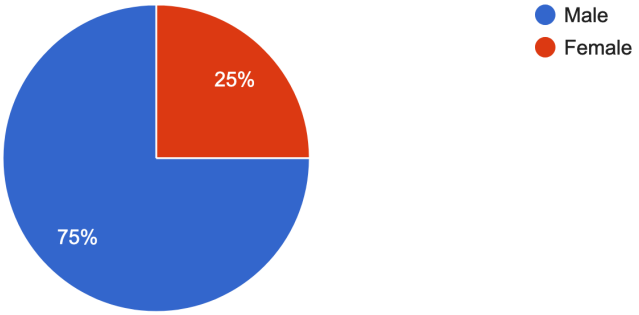
Usability Testing

For the usability evaluation, I deciding to collect both qualitative and quantitative information using a survey design, which was given to users after they evaluated the visualization. In order to conduct the study, I recruited eight volunteers (six in the initial survey, two a day later) to evaluate the "scrollytelly" visualization. Because of the short time frame needed to complete the study, a convenience sampling design was used.

Although the sample size was small, the participants were quite diverse across a number of variables. On the racial dimension, testers identified as Asian, Hispanic, Black and White, with some checking multiple categories indicating that they were multiracial. Participants also ranged in age from early twenties to late sixties. The study did have a bias towards men (6 to 2) over women, however.

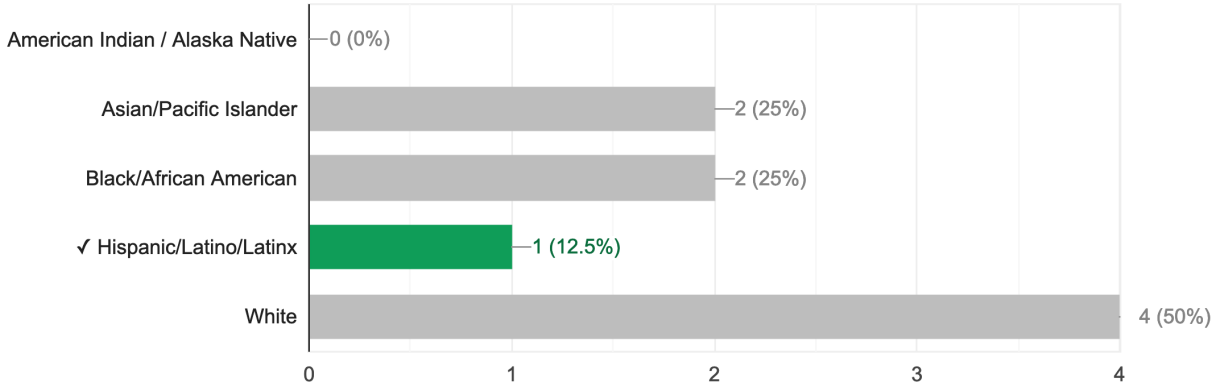
What is your gender?

8 responses

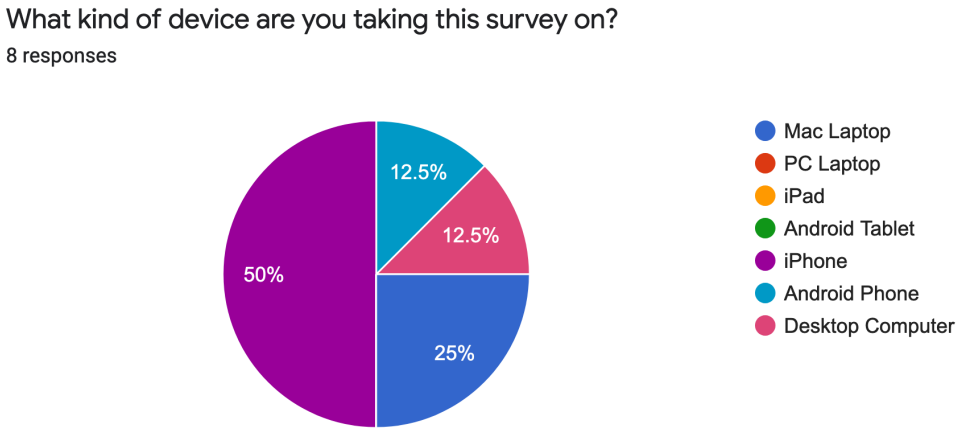


What is your race?

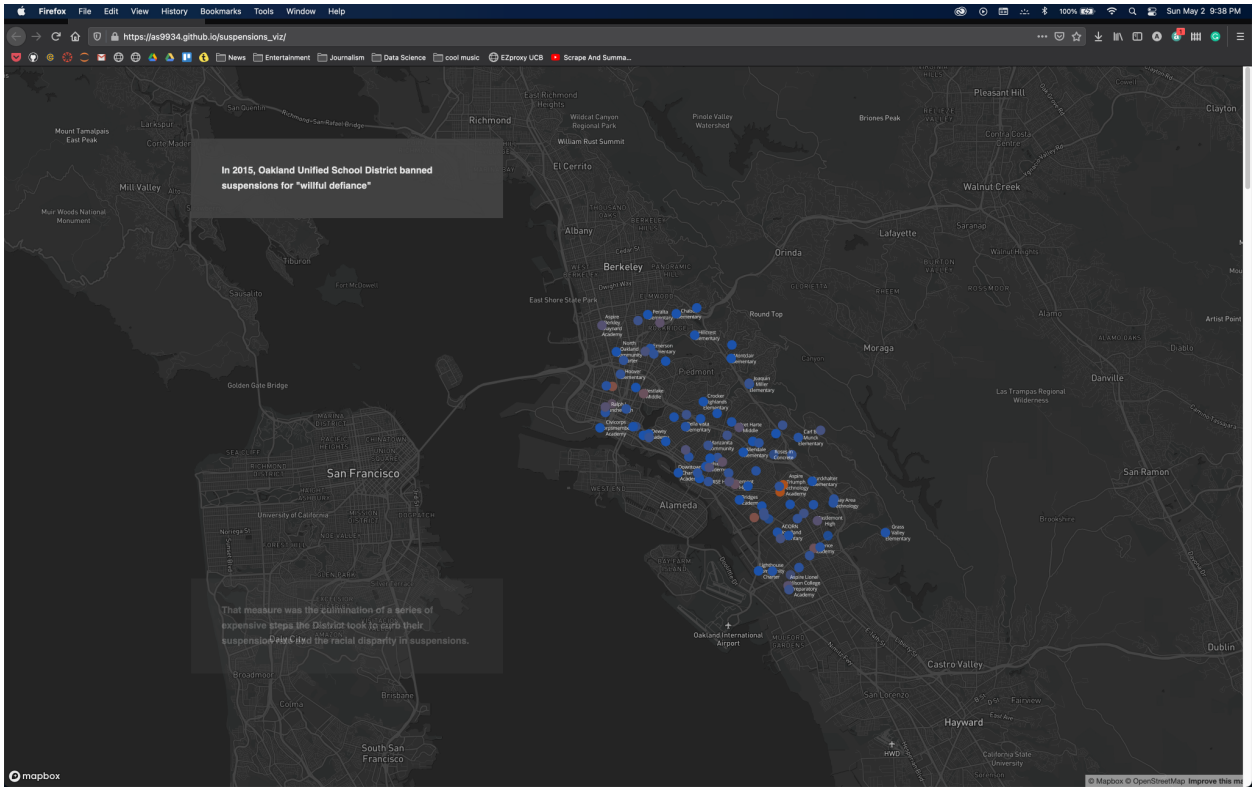
0 / 8 correct responses

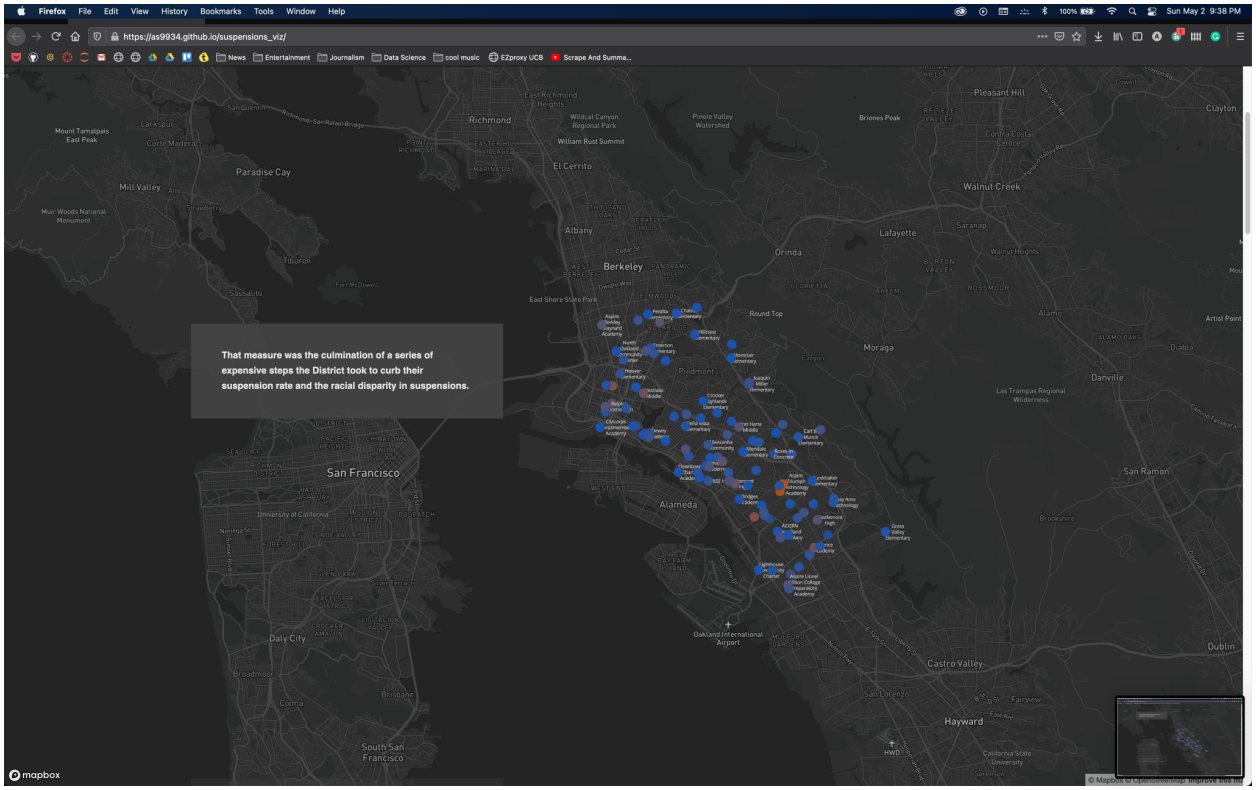


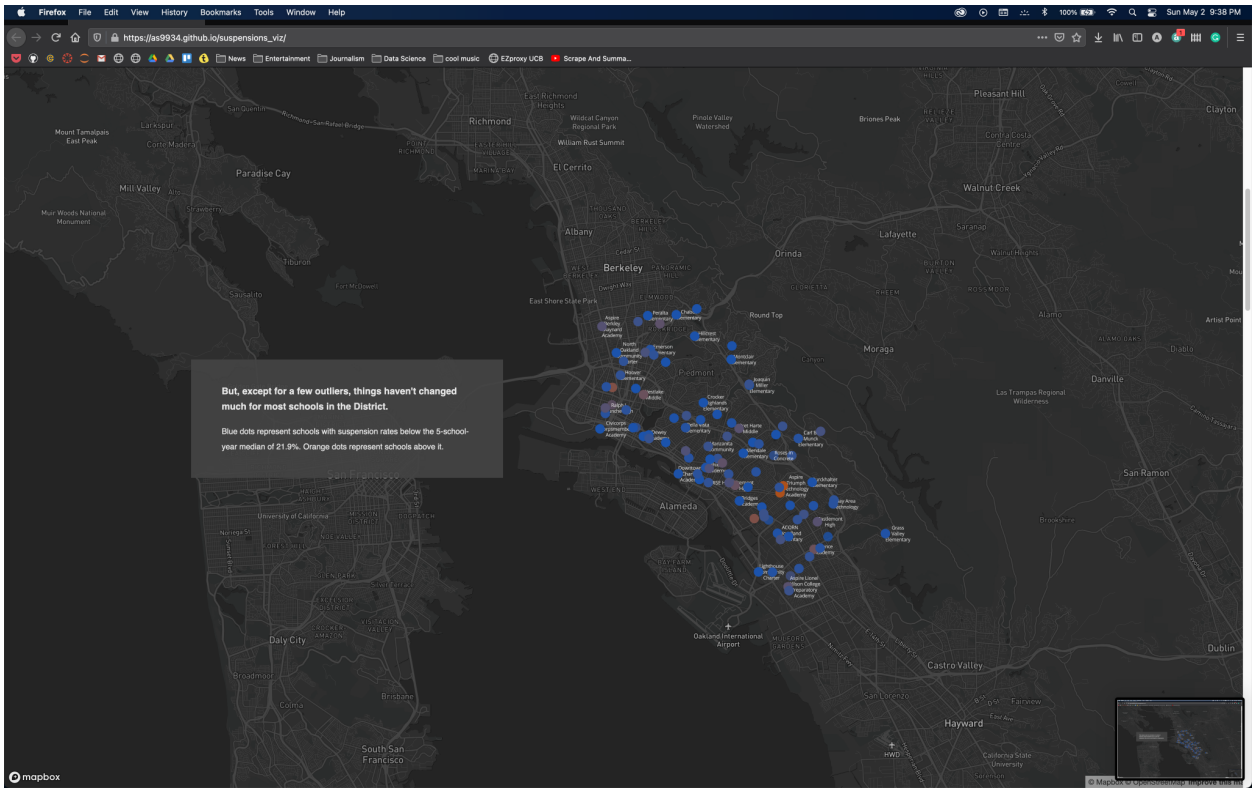
More relevantly to the task at hand, most of the participants (5/8) used mobile phones to evaluate the visualization. Four out of five using phones evaluated the visualization on iPhones, while one tested on an Android phone. The remaining users used Apple laptops or desktop computers. No testers used a Windows PC or tablets for their evaluation, and I do not have access to these devices myself, which is a potential limitation of the study.

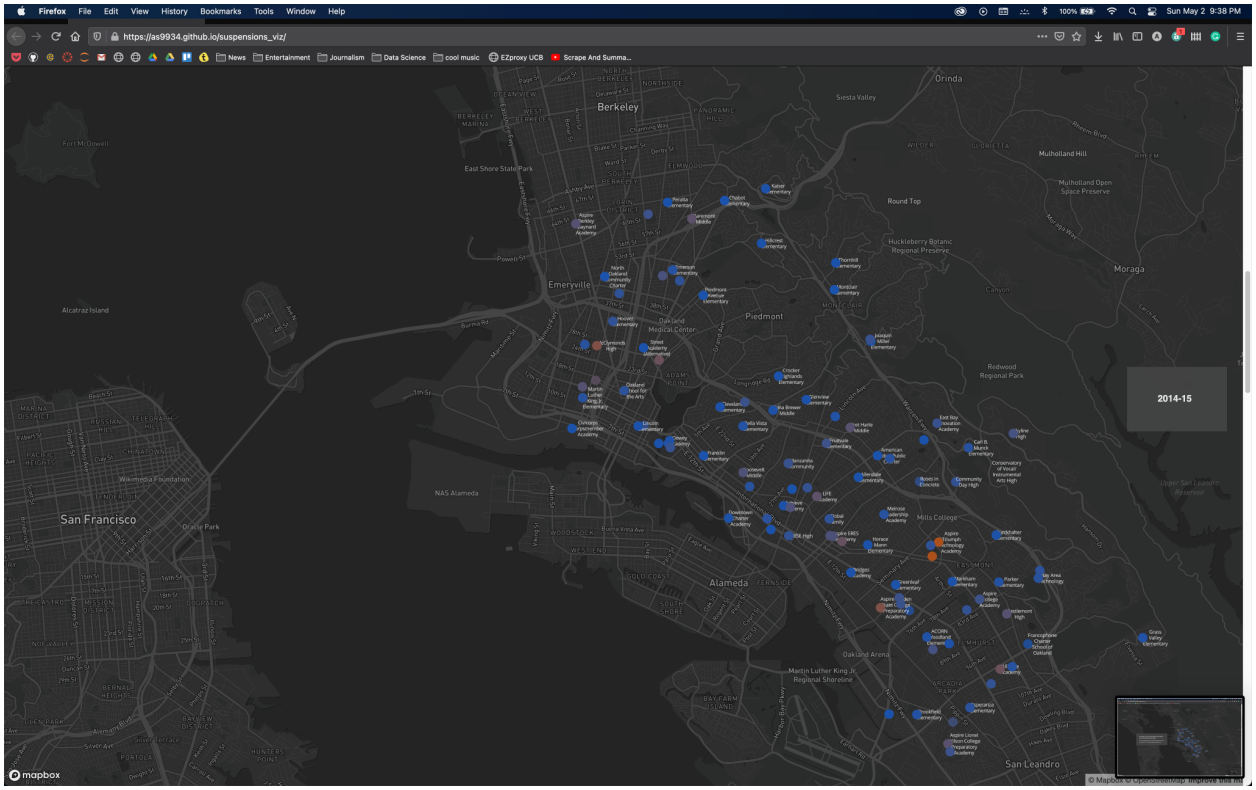


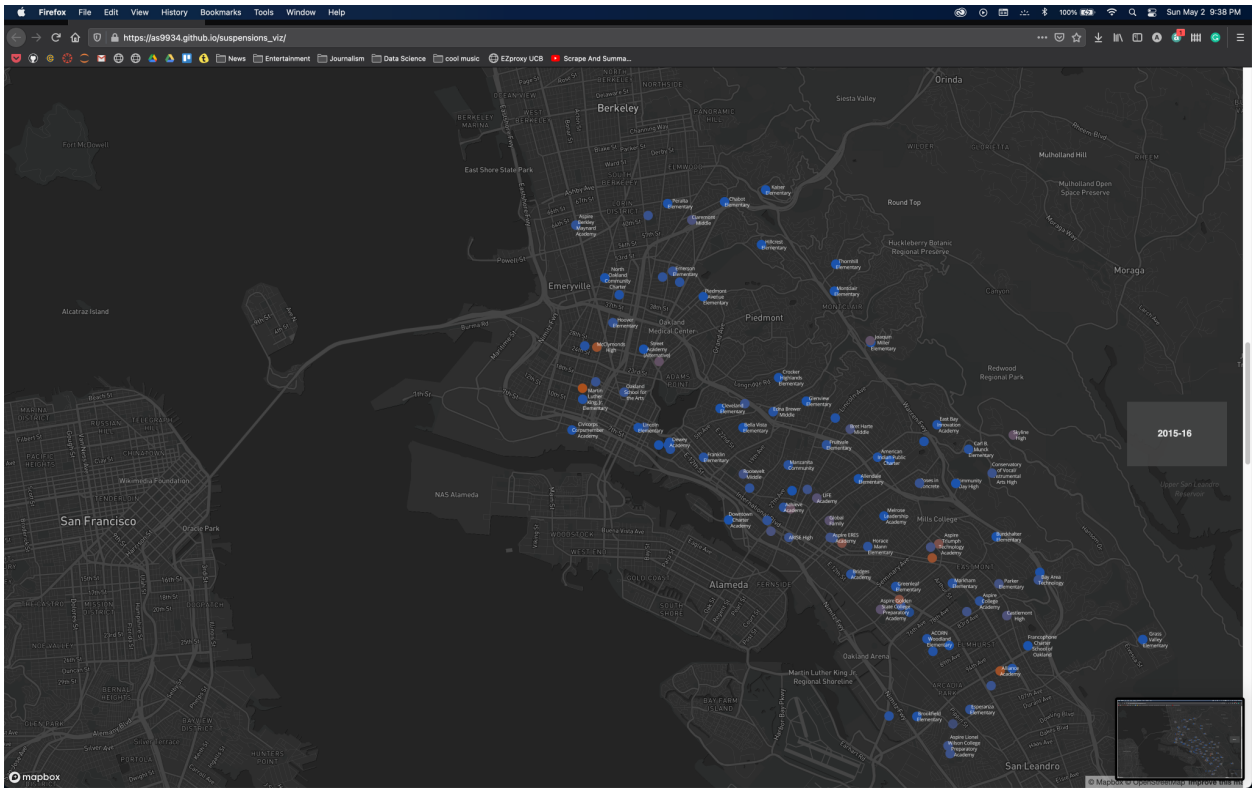
I instructed each of the participants to scroll through the panes/chapters of the visualization using their device and evaluate the text, dots and scrolling experience. In particular, I was looking for users to report noticing the dots change color, a smooth scrolling experience and legible and understandable text.

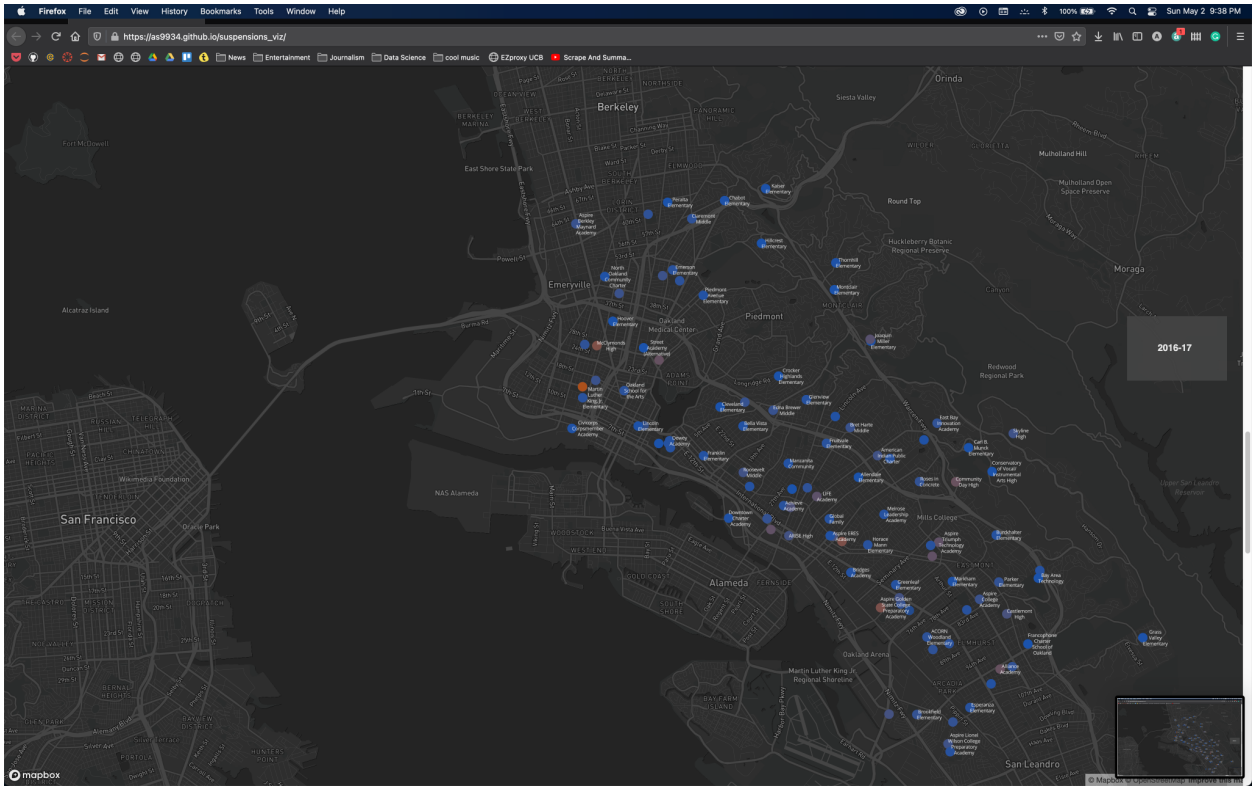


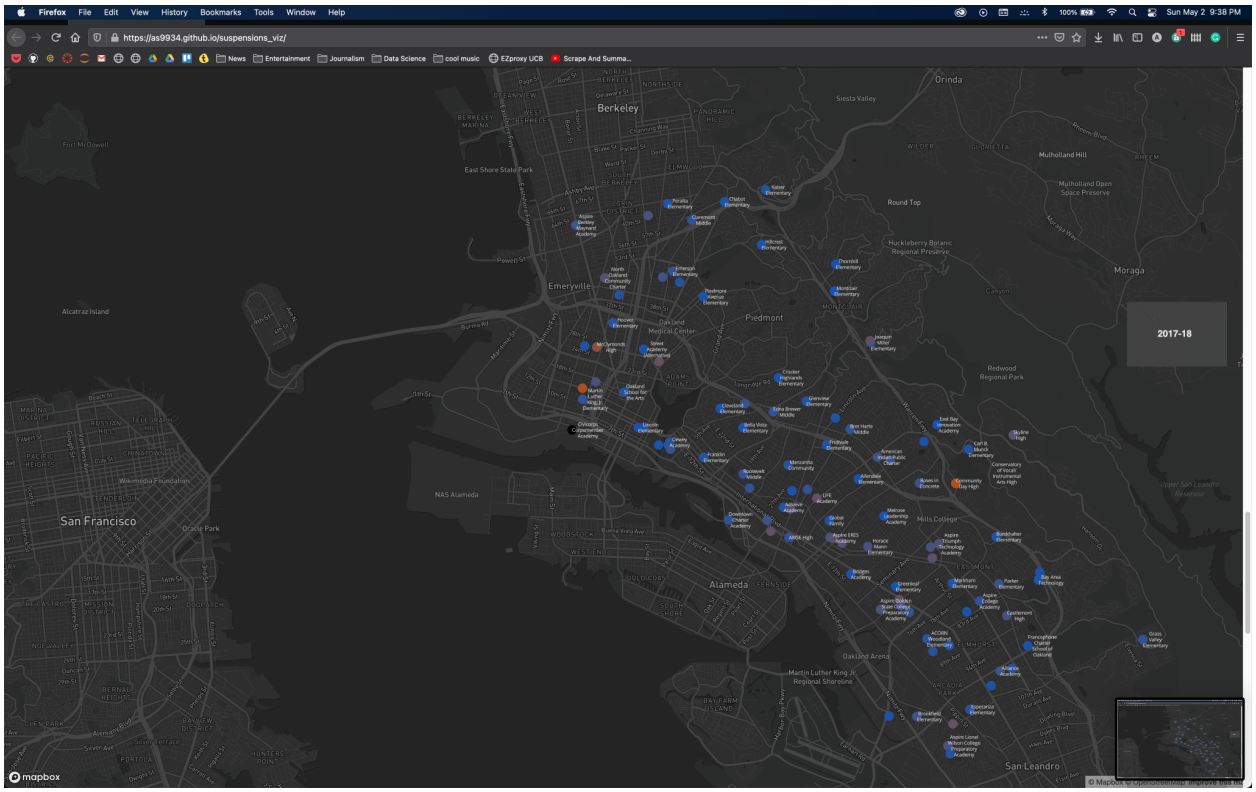

That measure was the culmination of a series of expensive steps the District took to curb their suspension rate and the racial disparity in suspensions.
mapbox

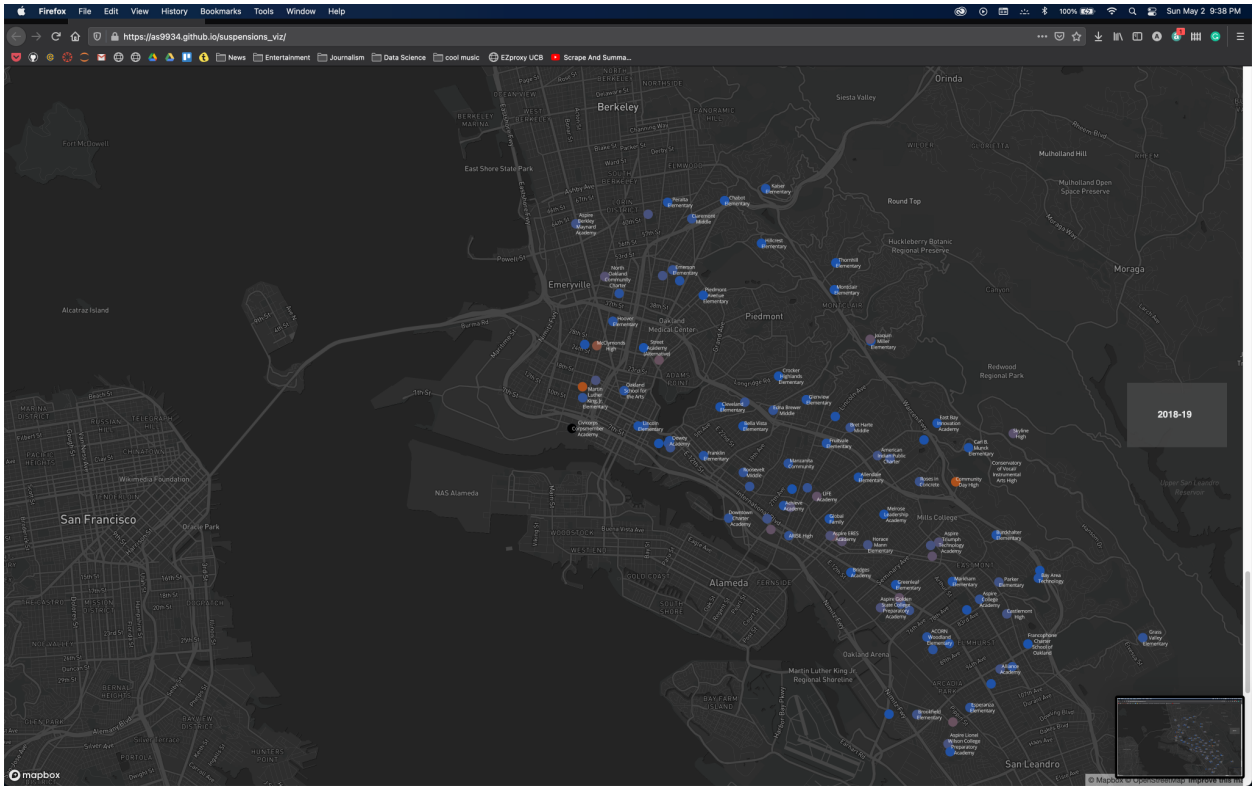

But, except for a few outliers, things haven't changed much for most schools in the District.
Blue dots represent schools with suspension rates below the 5-school-year median of 21.9%. Orange dots represent schools above it.
mapbox

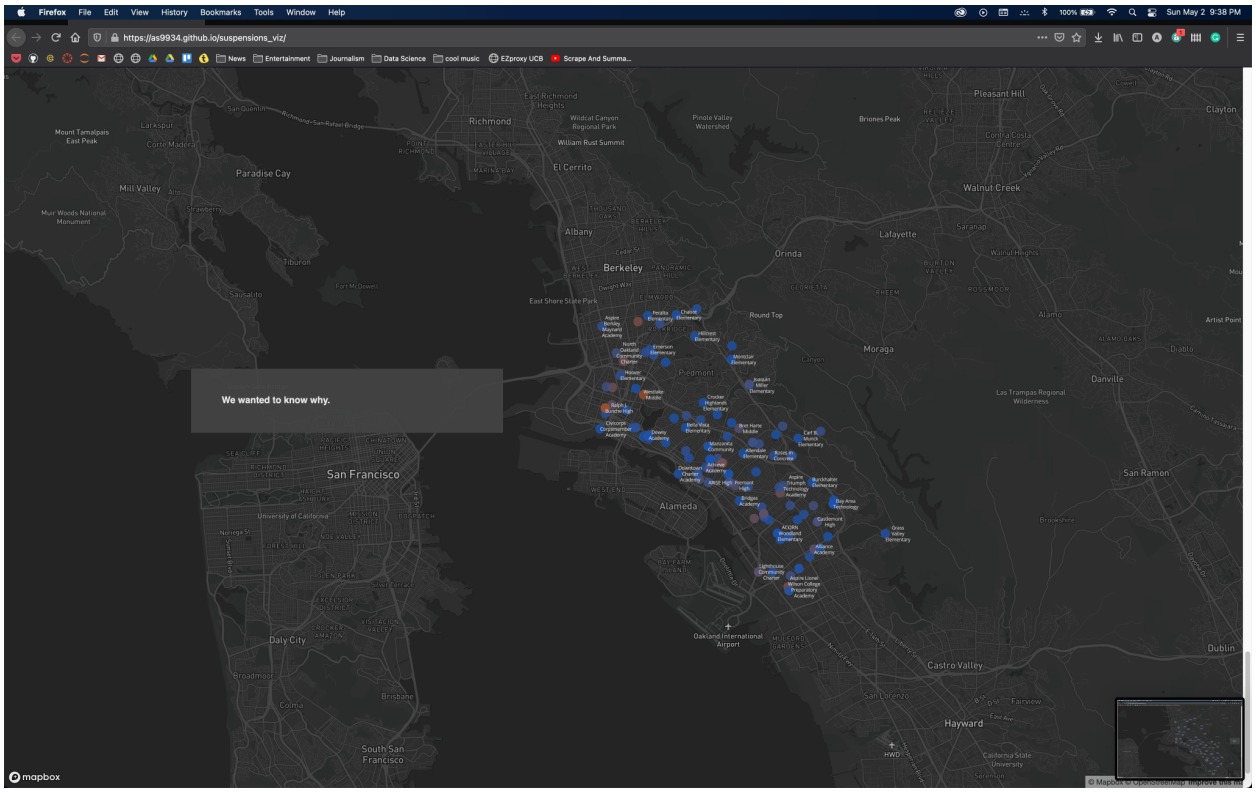

Firefox File Edit View History Bookmarks Tools Window Help
Sun May 2 9:38 PM
https://as9934.github.io/suspensions_viz/
News Entertainment Journalism Data Science cool music EZproxy UCB Scrape And Summa...
Berkeley
Orinda
Emeryville
Piedmont
Alameda
San Francisco
Oakland Arena
San Leandro
Moraga
Mulholland Hill
Round Top
Redwood Regional Park
Huckleberry Botanic Regional Preserve
Siesta Valley
Alcatraz Island
East Shore State Park
NAS Alameda
Martin Luther King Jr. Regional Shoreline
Mills College
2014-15
mapbox


Firefox File Edit View History Bookmarks Tools Window Help
Sun May 2 9:38 PM
https://as9934.github.io/suspensions_viz/
News Entertainment Journalism Data Science cool music EZproxy UCB Scrape And Summa...
Berkeley
Orinda
Emeryville
Piedmont
Alameda
San Francisco
Oakland Arena
San Leandro
Moraga
Mulholland Hill
Round Top
Redwood Regional Park
Huckleberry Botanic Regional Preserve
Siesta Valley
Alcatraz Island
East Shore State Park
NAS Alameda
Martin Luther King Jr. Regional Shoreline
Mills College
2015-16
mapbox


Firefox File Edit View History Bookmarks Tools Window Help
Sun May 2 9:38 PM
https://as9934.github.io/suspensions_viz/
Berkeley
Orinda
Emeryville
Piedmont
Moraga
San Francisco
Alameda
San Leandro
Oakland Arena
Martin Luther King Jr. Regional Shoreline
2016-17
mapbox


Firefox File Edit View History Bookmarks Tools Window Help
Sun May 2 9:38 PM
https://as9934.github.io/suspensions_viz/
Berkeley
Orinda
Emeryville
Piedmont
Moraga
San Francisco
Alameda
San Leandro
Oakland Arena
Martin Luther King Jr. Regional Shoreline
2017-18
mapbox

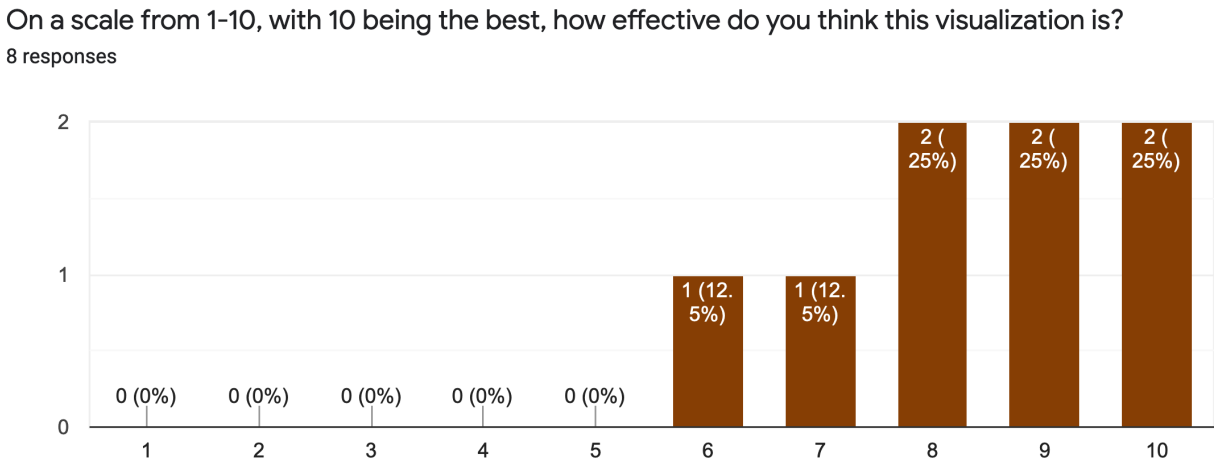

2018-19
Berkeley
Emeryville
Piedmont
Oakland Medical Center
Alameda
San Francisco
Oakland Arena
Martin Luther King Jr. Regional Shoreline
San Leandro
Orinda
Moraga
Mulholland Hill
Round Top
Siesta Valley
Redwood Regional Park
Mills College
NAS Alameda
East Shore State Park
Alcatraz Island
Fort McDowell
mapbox


We wanted to know why.
Berkeley
Richmond
San Francisco
Oakland
Alameda
Albany
El Cerrito
Orinda
Moraga
Lafayette
Walnut Creek
Pleasant Hill
Clayton
Danville
San Ramon
Castro Valley
Hayward
Daly City
Mill Valley
Paradise Cay
Tiburon
Sausalito
Oakland International Airport
mapbox

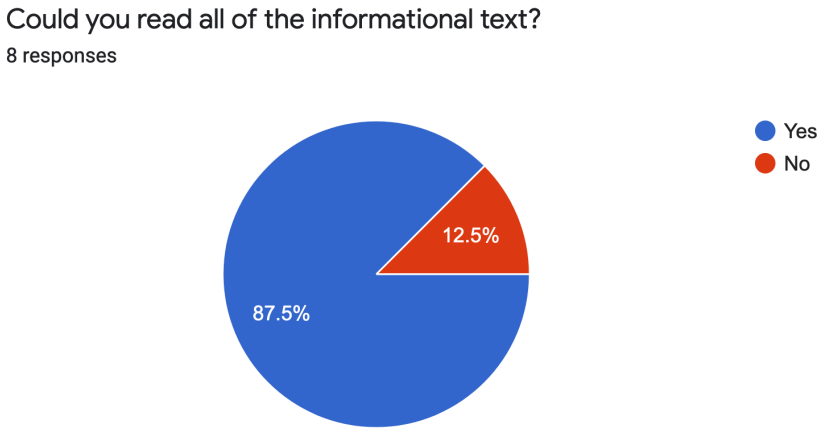
In the survey, I asked each user to rank their experience on a scale from 1 to 10, 10 being the best, with the text, scrolling experience and dots. I also asked users to provide qualitative feedback about each of these dimensions and for overall feedback on the visualization.

Specifically for the text elements, I asked users whether they could read the text and how understandable the information it conveyed was. For the dots, I asked users whether they could see the dots and whether they saw them change color. On the scrolling dimension, I asked users only whether their experience was smooth and to rank how smooth it was.

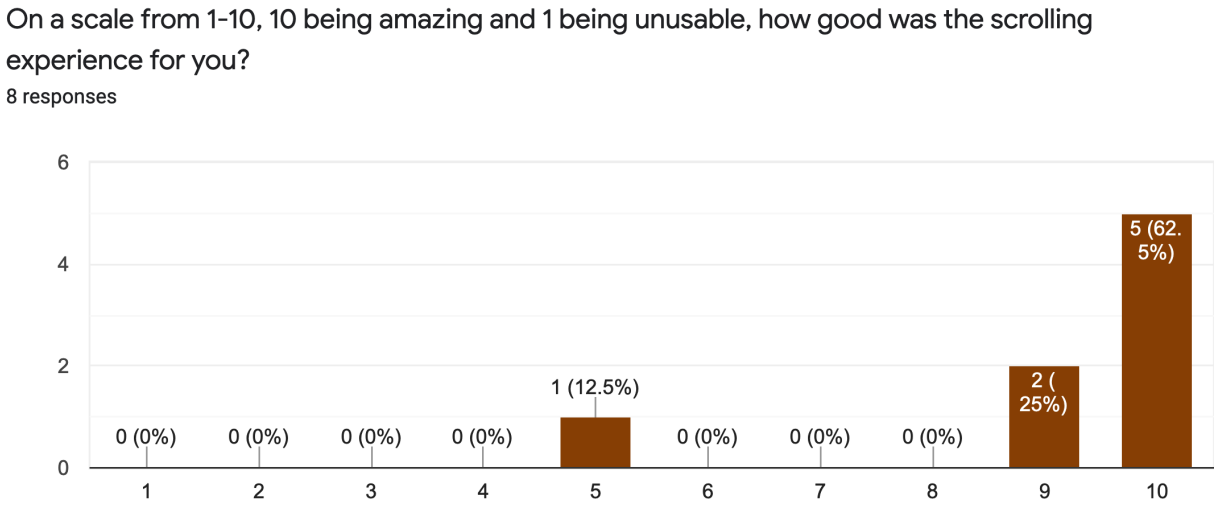
Overall study testers rated the visualization highly, with all users giving the overall design and ranking between 6 and 10 on a 10-point scale.



On the text dimensions, these scores held, with three of the six users giving the descriptions a rank of 9 or 10, denoting that the text was "incredibly clear." One user said they were unable to read all of the informational text, however.



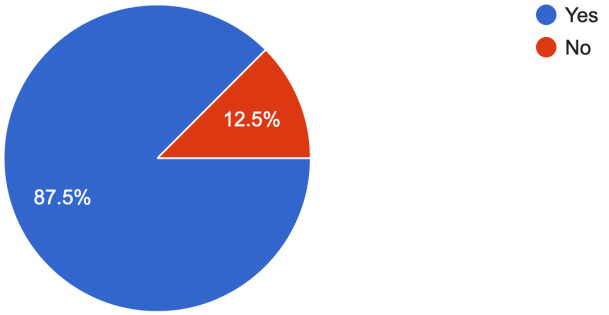
Participants near-universally enjoyed the scrolling experience with all seven of eight saying that the scrolling experience was smooth. When asked to rank their experience five of the eight participants gave the visualization a 10/10 on the scrolling dimension.



The experience of the dots was more of a mixed bag.Although seven of the eight participants said they could see the dots clearly, only half noticed them changing color as they scrolled through.

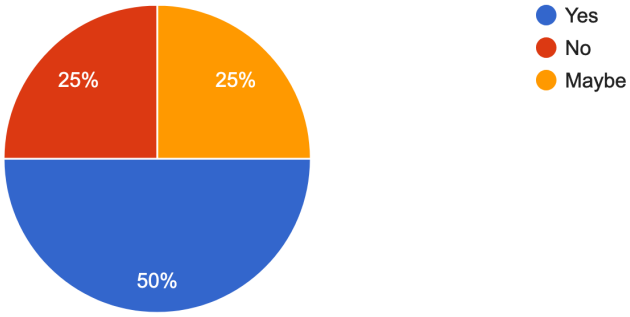
Were you able to see the colorful dots clearly?

8 responses



Did you see the dots change color as you scrolled?

8 responses



When asked to provide qualitative feedback on the various elements of the visualization, participants praised its mobile-friendly design and innovative user experience. But there were a few areas of constructive criticism that were raised.

The one user on an Android phone complained that the semi-transparency of some of the elements made text difficult to read. There were also several comments on the dots, with some users reporting that some of the blue dots appeared purple on their screen, which caused them to wonder whether this represented a third bucket in the suspension rates. Other users suggested that the color transition could be highlighted much more for the few dots that do change color and that the existing dots could be made brighter. I also received feedback suggesting that using this type of design to show a lack of significant change was perhaps not the most effective use of the scrollytelling form.

In response to these critiques, I first made the decision to have the dots represent the suspension rates for a single racial group, instead of the overall rate for the school that school year. I then color-coded the dots to represent that racial group (blue dots for White students and orange dots for Black students), so that the user could see the (quite dramatic) differences between the suspension rates of White and Black students for each school year. Instead of representing the size of the suspension rate with a certain color value, I instead chose to bind it to the radius of the point, with small dots representing lower suspension rates and large dots representing higher suspension rates. I figured this would be a good way to address the critiques about the lack of change and the issue where users perceived the dots as different colors depending on the device they were using.

Appendices

Visualization Link: https://as9934.github.io/suspensions_viz/

Visualization GitHub Repository: https://github.com/as9934/suspensions_viz

Scaper Python Code (GitHub):

https://github.com/as9934/oaklandsuspensions/blob/main/Oakland%20School%20Suspensions.ipynb

Content Inspiration: Pinellas suspends black kids more than virtually every other big Florida county, *The Tampa Bay Times*

Design Inspiration #1: In a Notoriously Polluted Area of the Country, Massive New Chemical Plants Are Still Moving In, ProPublica/The *Times-Picayune*/The Advocate

Design Inspiration #2: Life in the Camps, Reuters

Mapbox Template Explanation: Interactive Storytelling

Mapbox Templat Repository: "Storytelling"