Final Report

PART I: Introduction: Domain problem characterization

Data had been a huge driver in recent years on how principals structure their school and the decisions district leaders make now more than ever. This trickledown effect from the district leaders to the principals to the teachers has every employee in CMS handling data and using it regardless of how knowledgeable they are at statistics. Teachers are required to collect data on their students and prove to principals how they used this data to improve the students learning, and are even graded on this in their teacher evaluations. The principals then have to show the district leaders how their data is being used to improve their schools and it is just one big circle of data being used. Our goal is to show data that is already publicly available in an easy to use manner that anyone can understand. The metric we are tackling is diversity. Over the years CMS claims to be improving diversity according to their data but in reality is diversity really being improved? The thing about data is you can show it however you want and our goal is to provide an unbiased viewpoint on how diversity has changed over the past 12 years in CMS with applications that anyone can use.

PART II: Introduction: Domain problem characterization

When deciding to tackle diversity in CMS we started with CMS website. The CMS Office of Accountability keeps all data that is public, but only offers some data online. Requesting data that is not available on the website is available by paying a fee if you are not a CMS employee and you also have to agree to their terms on how you use the data provided as such. . .

All other data requests must meet both of the following criteria:

- Request is from a CMS employee/CMS Community Partner and/or for purposes that align with the CMS goals and priorities.
- Requests must be able to be completed in a reasonable time frame.

Please note the following:

- Data may only be used for the purpose stated on the data request. Use of data for any other purpose must also be approved by CMS Office of Accountability.
- Data may not be shared with a third party outside of CMS unless approved by the CMS Office of Accountability. Failure to comply will result in the loss of data
- Request privileges in the future.
- Data will not be provided for students whose student IDs, school IDs, or parent consent forms are not provided.
- Files will include only CMS data. Linking to external data is the responsibility of the requester.

The data needed for our project was available online but only for the years 2008-2020. Data earlier than 2008 would have to be requested.

Another interesting thing to note CMS only kept diversity data on the school years 2015-2020. As we could not find previous years we decided to contact CMS' Department of Accountability. We were then directed to the North Carolina Department of Instructions website where we could access 2008-2014.

Once all data was collected the hardest part began. Over 12 years many things happen with schools. They change locations, names, merge, even go from elementary to k-8 and vice versa. When collecting the data we created a column called School_Consistant as shown below with three different schools.

```
## # A tibble: 6 x 3
##    Year School                 School_Consistant
##   <dbl> <chr>                  <chr>
## 1  2010 ALBEMARLE ROAD ELEM    ALBEMARLE ROAD ELEMENTARY
## 2  2011 ALBEMARLE ROAD ELEMENTARY ALBEMARLE ROAD ELEMENTARY
## 3  2011 ASHLEY PARK ELEMENTARY    ASHLEY PARK PREK-8 SCHOOL
## 4  2012 ASHLEY PARK PREK-8 SCHOOL ASHLEY PARK PREK-8 SCHOOL
## 5  2011 BERRYHILL ELEMENTARY      BERRYHILL SCHOOL
## 6  2012 BERRYHILL SCHOOL          BERRYHILL SCHOOL
```

Using our current knowledge of CMS schools and some research we were able to figure out how each school was re-named, changed, or added in the 12 year span. When it came to races there were changes over the years. In 2011 the race "Two or More" was added to the demographics. This lowered the amount of Black students in the district as a majority of students classified as Two or More in 2011 were categorized as black in 2010. Once we generated the last step was gathering the latitude of each school using google maps.

PART III: Encoding & Interaction Design

A goal of this project was to equip public school leaders and policy makers with key insights into the state of diversity within the Charlotte-Mecklenburg School district over the past 10+ years in as concise manner as possible so not to overwhelm the target audience. We felt this could be accomplished satisfactorily with three plots: two reactive bar plots and an interactive reactive map. Each of the three plots is housed within the Shiny app in its own tab panel.

The first tab panel contains a district overview bar chart. It is important to give the app's users the ability to look at the makeup of the CMS student population as a whole. This information is conveyed in a bar plot in which each bar represents the number of students of each race/ethnicity in the CMS district in a given year (FIGURES 1 and 2). While the largest three population groups are Black, White, and Hispanic, these graphs suggest that a larger number of white students than other groups have been gradually leaving the CMS school system. Alternatively, instead of raw counts, the user has the option to view the percent of the total student population that was made up by a each race/ethnicity in a given year.

Bar charts were selected (as opposed to pie charts or other options for displaying counts and percentages of categories) because bar charts make it easy to compare different groups. This comparison is made even easier when the bars are ordered by count, which was adopted for these charts. A coordinate flip was used to make it easier to read the race/ethnicity axis labels. And each race/ethnicity was given its own color to help keep the groups visually distinct. The order from top to bottom ended up being Black, White, Hispanic, Asian, Indian, Pacific Islander, and Two or More. This same order and color scheme was also used in the individual school bar charts visible in FIGURES 6 and 7.

FIGURE 1: Count of students of each race/ethnicity in the year 2008
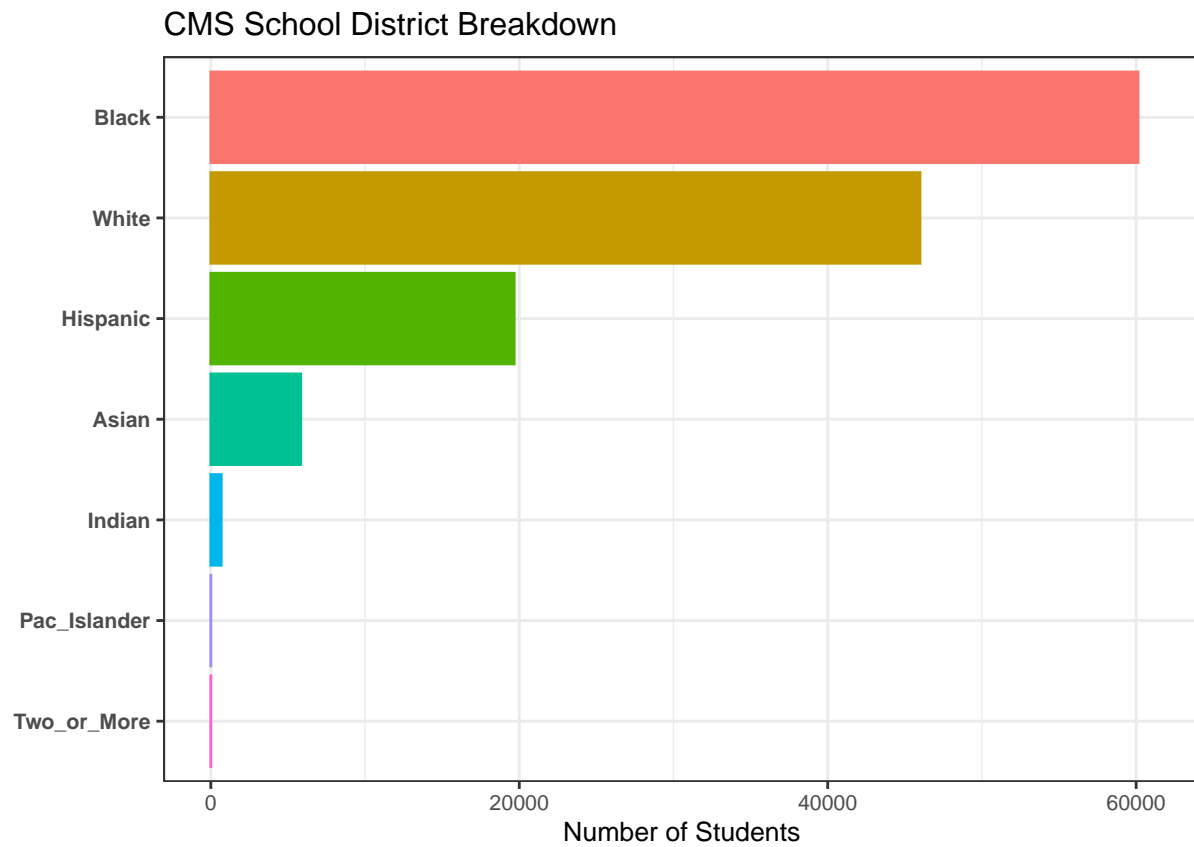
FIGURE 2: Count of students of each race/ethnicity in the year 2019
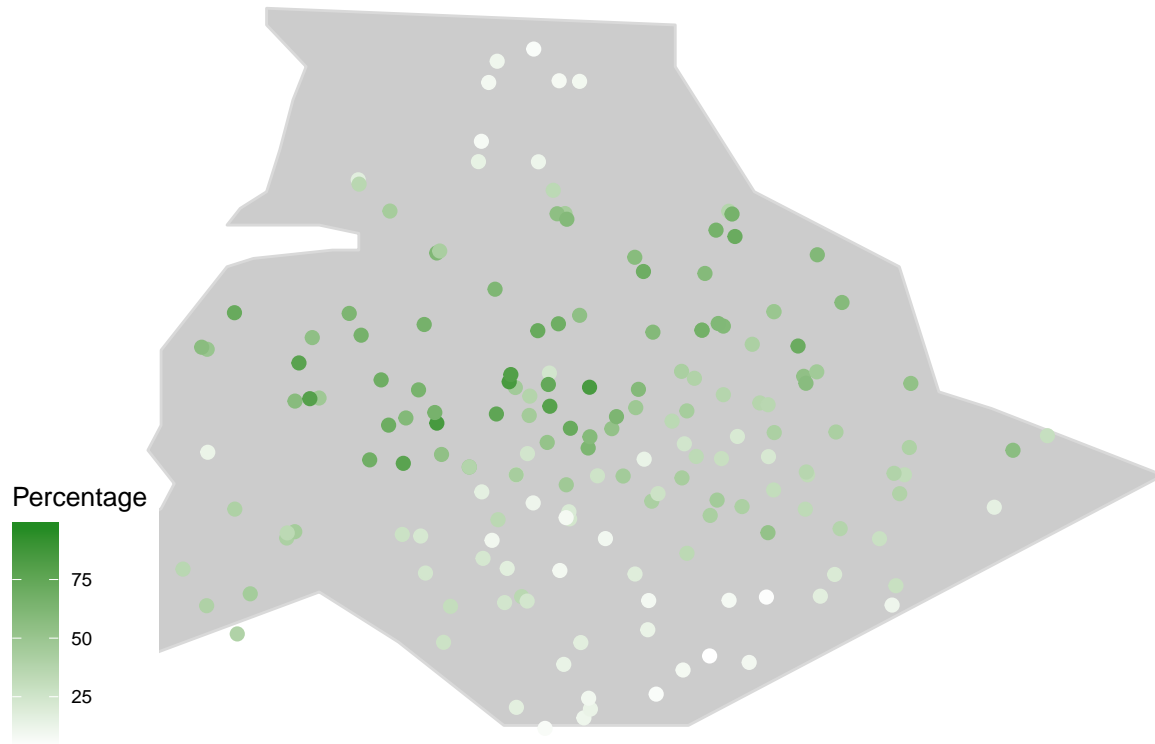
CMS School District Breakdown
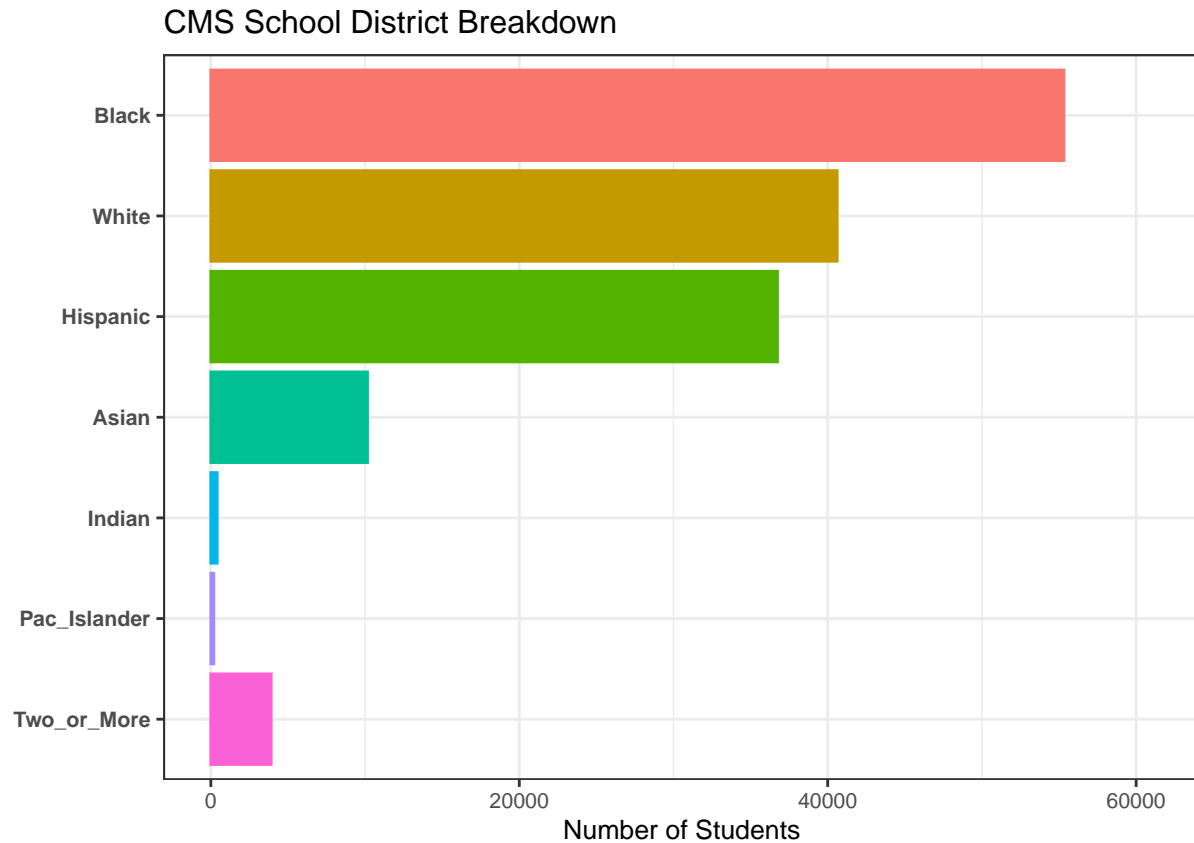
The second tab panel contains a geographical map. Additional insight can be given by providing the user with the ability to look at a map of the district and see its geographical makeup. The maps in FIGURES 3, 4, and 5 show each school as a dot, and the shading of each dot represents the percent of that school that is black, white, or hispanic respectively. This allows someone using the app to notice areas of the map in which a particular race/ethnicity makes up a larger percentage of the student population than other areas. From these maps it becomes clear that some geographical segregation exists. The user has the ability to select a year for any of the plots in the Shiny app, and for the maps the user has the additional ability to shade the points based on the percentages of one or of multiple races/ethnicities.

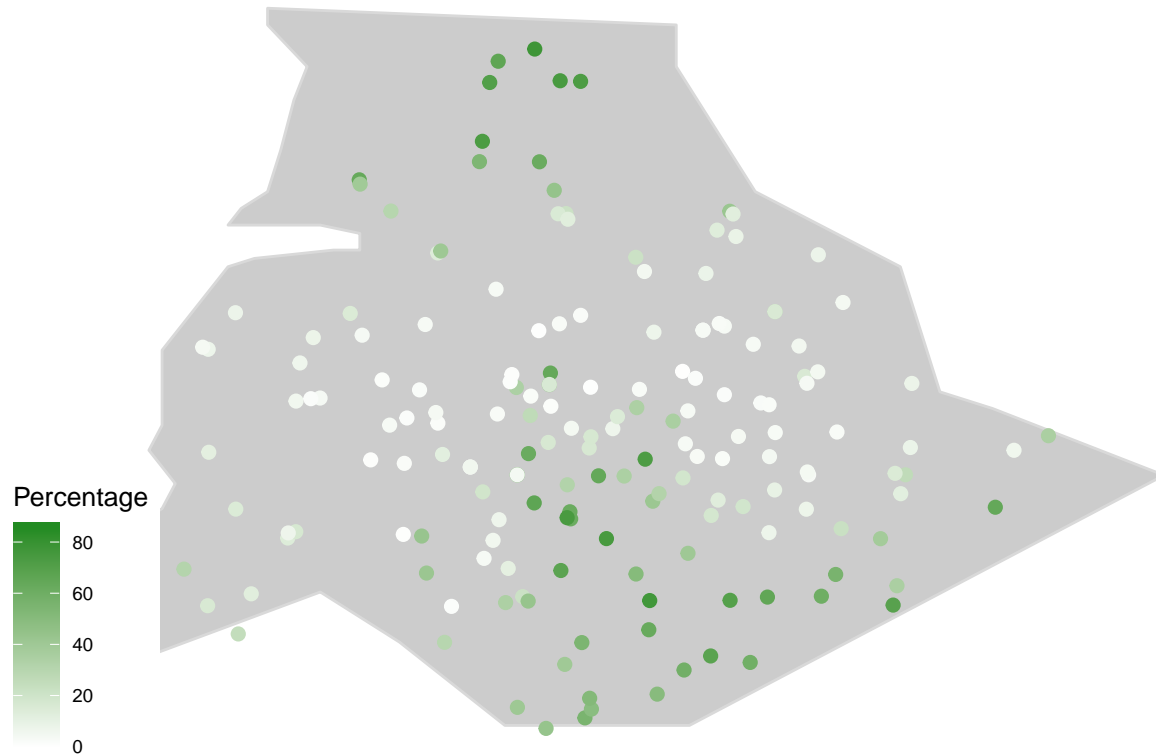**FIGURE 3: Geographical Map of CMS district in 2019 portraying percent Black**

(Shading represents the percent of each school's population that is Black.)

4

CMS School District

FIGURE 4: Geographical Map of CMS district in 2019 portraying percent White

(Shading represents the percent of each school's population that is White.)

CMS School District

FIGURE 5: Geographical Map of CMS district in 2019 portraying percent Hispanic

(Shading represents the percent of each school's population that is Hispanic)

CMS School District
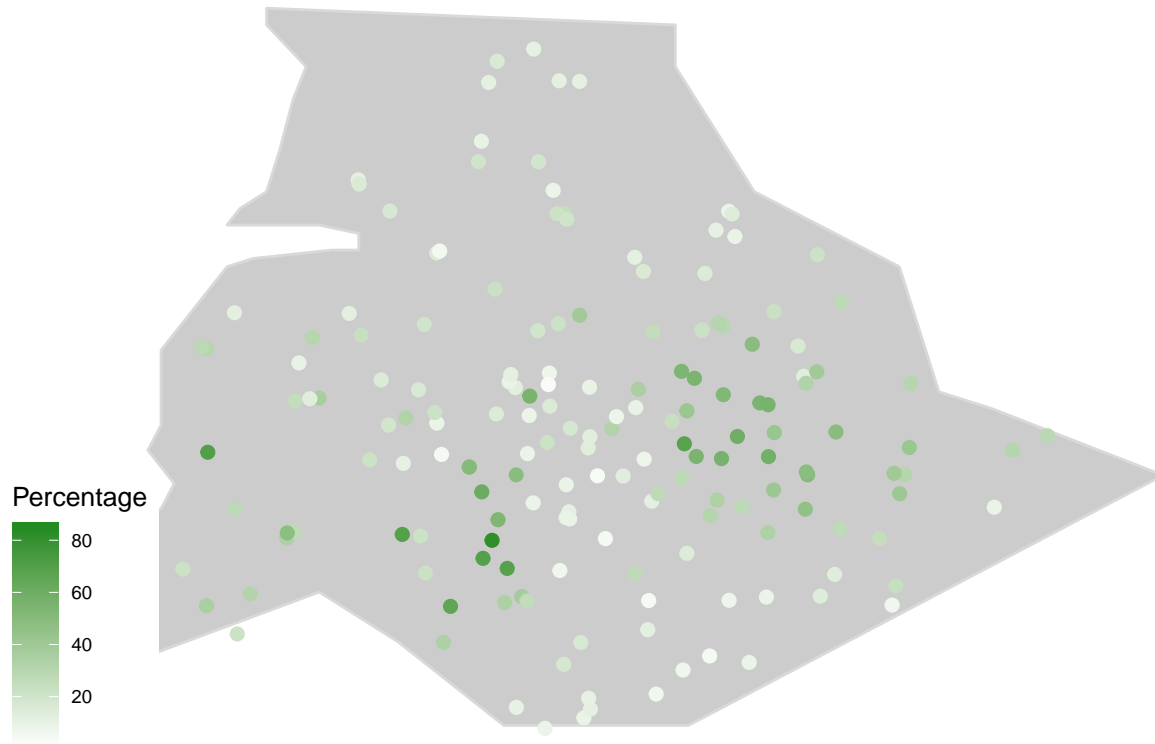
The third tab panel of the Shiny app contains another bar chart which offers a more granular view. In this panel the user has the ability select any school within the CMS system and look at a bar chart of that school's integretion across time. Using the play button on the Year slider input bar is particularly illuminating for this plot, as it allows the user to watch a school's makeup change throughout the years. For example, North Mecklenburg High had a significant change in the makeup of its student body across time. FIGURES 6 and 7 show the bar charts for the years 2008 and 2020 respectively. As previously mentioned, the order and coloring of the race/ethnicity categories is consistent with that of the district panel bar charts. This makes comparisons between the two panels easier, even though it means that for many of the charts the bars are not ordered by descending size.

FIGURE 6: Percent of student population made up by each race/ethnicity

North Mecklenburg High in 2008
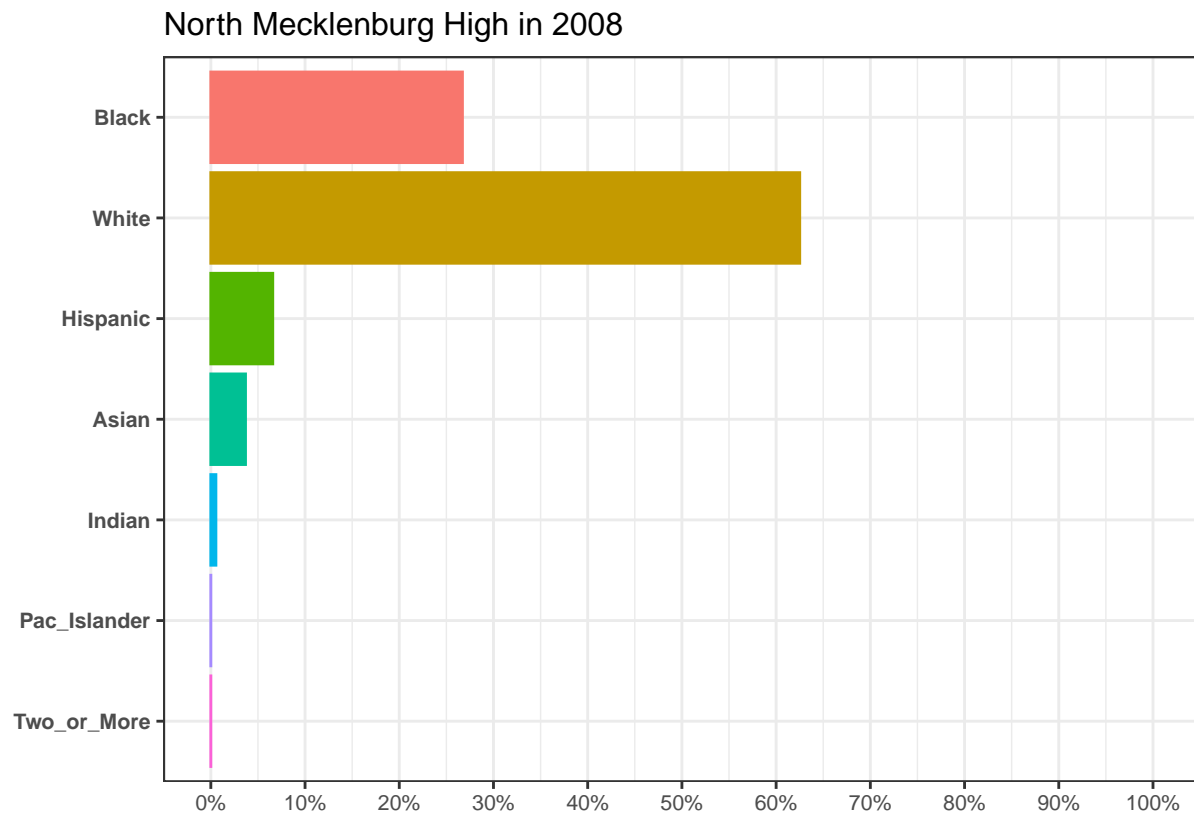
FIGURE 7: Percent of student population made up by each race/ethnicity

North Mecklenburg High in 2020
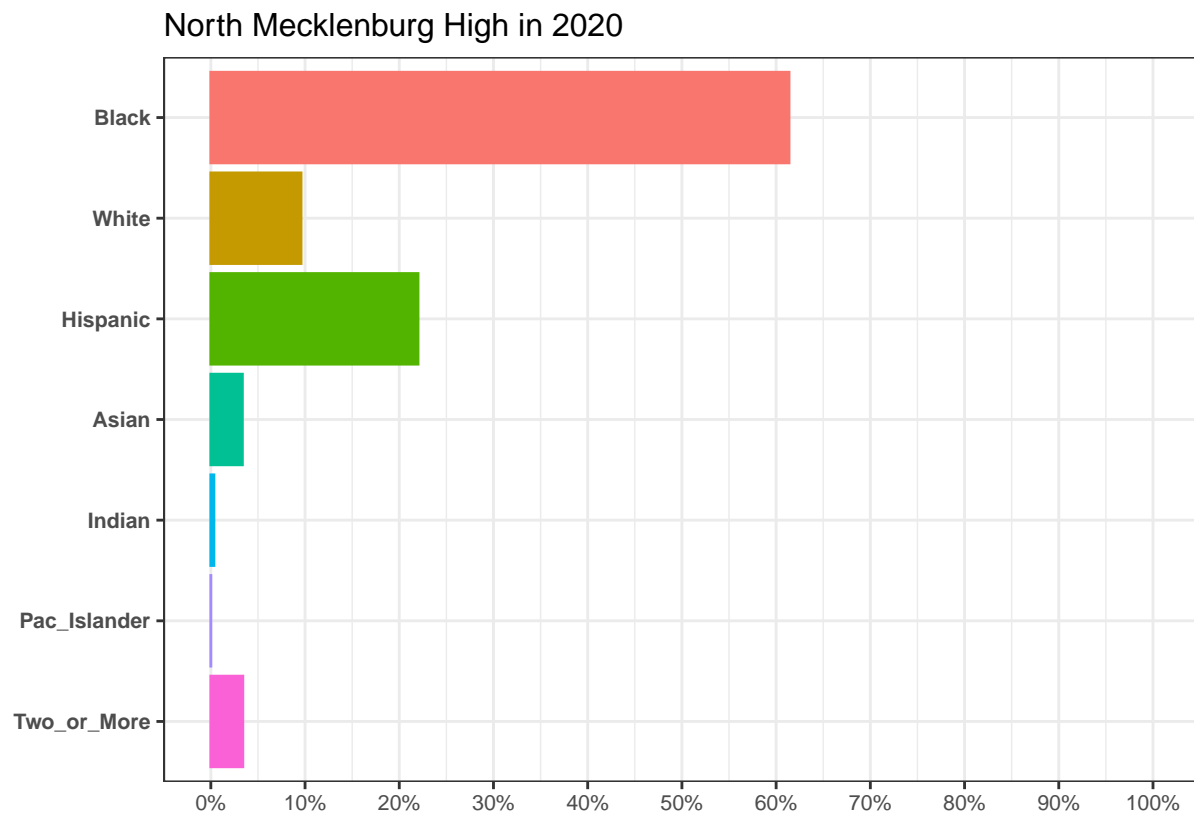
PART IV: Algorithmic design

Step 1: Start

Step 2: Prepare the tidydata. 2.1 Read SchoolData.csv and filter schools which are exist all 13 years and flag as 0 and 1. 2.2 Create a dataframe containing the percentages of each group across all students in the district each year 2.3 Calculate tidydistrict_counts and tidydistrict_prcnts for district wise count and percent values 2.3 Gather the groups into a single column with their percent in a separate column

Step 3: Initialize variables

Step 4: Design the UI with - Slider Input - Sidebar panel and Main Panel with Tabsetpanel and Four tabpanel– Overview, Map and Individual Schools and Schools Resembling District

Step 5: Read year from slider input if user clicks it.

Step 6: Read the respective tabpanel information

Step 7: Repeat the steps until year <=2020 or user stops slider input –

```
7.1 If Tab selected is Overview
  7.1.1 If Option selected is Counts
```

- Plot tidydistrict_counts variable on main panel by year on the slider input

7.1.2 If Option selected is Percentage - Plot bar chart with tidydistrict_prcnts variable on main panel by year on the slider input value.

7.2 If Tab selected is Map 7.2.1 Read the Race info from DropDown - Plot Map on main panel to show the percentage of student by race and year

7.3 If Tab selected is Individual Schools 7.3.1 Read the School info from DropDown - Plot bar chart on main panel to show the percentage of student by race and year per school

7.4 If Tab selected is Schools Resembling District 7.4.1 Read the Level of Resemblance Percentage from Side bar - List the Schools in Table format which are percentage resemblance with the District level percentage.

Step 8: Stop

Performance improvements: 1. Prepare data in advance: We have a separate script which is called at the beginning of the app to prepare the tidy data required for plotting school data. 2. Tried ShinyDashboard, Leaflet and Google Map options but basic Shiny App looks better for user
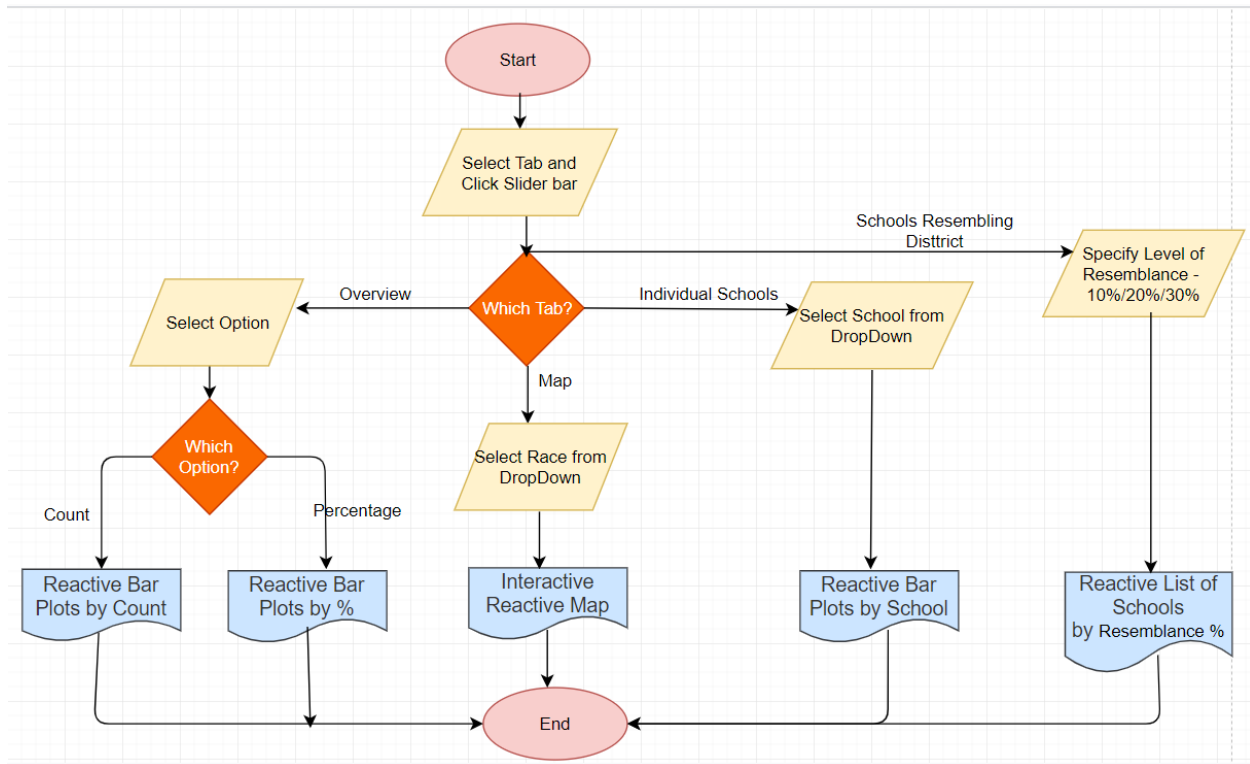
Perfromance Workflow

To assess the performance of our app we would potentially use shinyloadtest to get an idea of its speed, and we could use profvis to determine what elements if any might be slowing it down. From there we could decide where to focus our optimization efforts. We suspect that caching will be effective in our case, as the app contains many plots, but several get used repeatedly. These are all efforts that could be during future phases of the project.

PART V: User Evaluation

We sent the Proof of Concept to many different types of users. These Users included..

- Data Management and Communications Specialist from the Levine Museum of the New South
- Assistant Principal in CMS
- Parent of a CMS student
- Director of Research, Evaluation, and Analytics for CMS

Response from Cathy Doheny (Data Management and Communications Specialist)

"We needed this very tool just last month when planning marketing for a program. We will definitely keep it handy to use in the future. I would think that other organizations, including CMS, would also find it very helpful. I had a difficult time finding current stats online last month. You have created such an easy-to-use tool that consolidates all the data into one place"

Feedback from Lindsay Messinger, Ph.D. Director of Research, Evaluation, and Analytics

Your visuals are very interesting. The only suggestion that I have is to create an option so that the user can select "All minority" or "All Students of color" or "Black/Hispanic", rather than only being able to select one race at a time.

Other suggestions included

- Clarifying races such as indian. Some users thought indian was the race representing the country India when in fact Indian represents the american indian population.
- Adding landmarks or streets to the original map to make it clearer where a school was.
- Adding directions on how to use the tools

PART VI: Future Work

There are two major additions that could be made to the app that would provided expanded insights: school funding and a larger historical perspective. Regarding the former, we would like to acquire school funding data and use it to assess any correlations between an individual school's racial/ethnic makeup and its level of funding. It would also be useful to incorporate a way to determine what comes first: the change in a school's makeup or the change in its level of funding.

The second major addition would be to explore data that goes much farther back than 2008. Data from years earlier than 2008 was not obtainable online without submitting an official request. However, it would be beneficial to explore the changes that have occured in the CMS diversity landscape since the beginning of desegregation. Ideally this data would go back to the 1954 Supreme Court Brown v. Board of Education decision, but realistically we would hope to get data starting in the late 1960s when the Charlotte busing program began to be implemented.

A possible agreggated statistic that we could use to investigate historic diversity would be the yearly percent of schools in the district for which the White percentage was less than 10% or the Black percentage was less than 10% or the Hispanic percentage was less than 10%. Another useful statistic might be the percent of schools for which the makeup of Black, White, and Hispanic is within 10%, 20%, or 30% of the district makeup. These numbers could help the app's user assess how the landscape changed when busing was enforced beginning in the late 1960s to when busing ended at the turn of the century.

PART VII: Appendix

We also created a Shiny Dashboard version of the app (FIGURE 8). We liked the overall appearance of this version better than the simple Shiny version, however in the end we decided that the simple Shiny version was easier for our target audience to follow.

FIGURE 8: Shiny Dashboard prototype
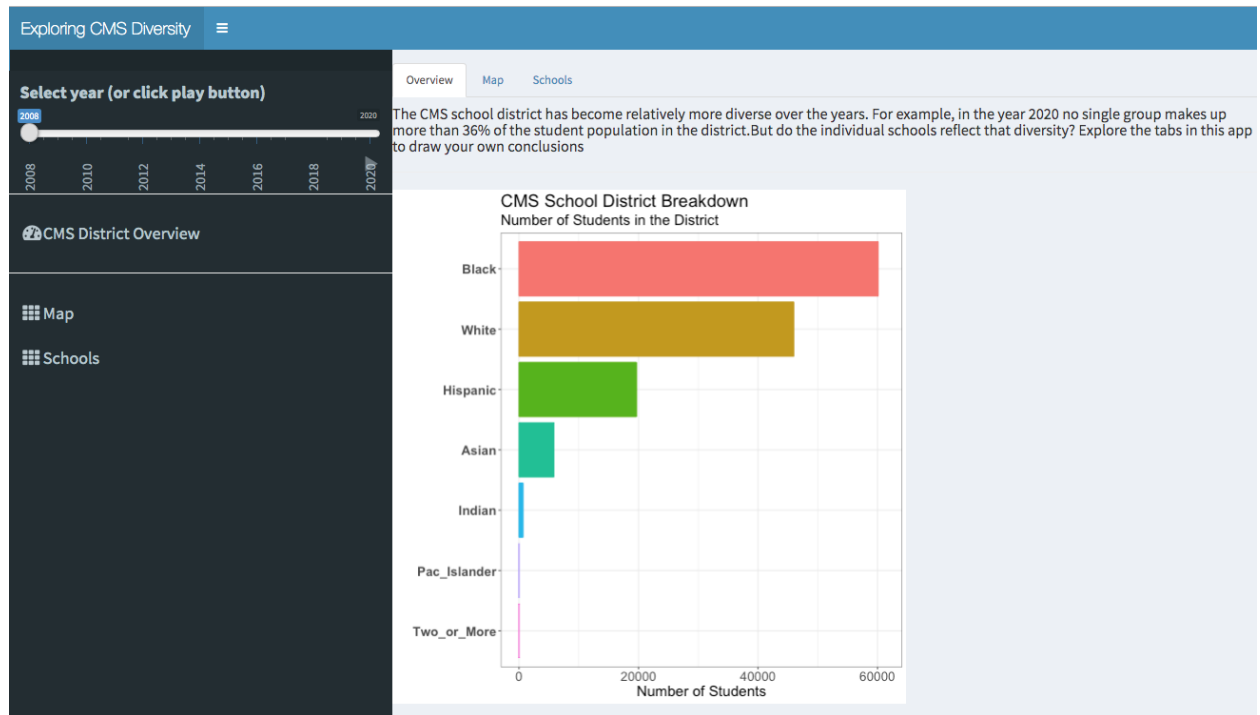
In the future we might like to refine the Shiny Dashboard version. Ideally we would like to do A/B testing to see which version is actually preferred by our users.