## Background and Motivation

Finding a school for one's children in St. Louis can prove to be an overwhelming and time-consuming process as available information is often uncentralized and unstandardized, making a hard decision even more difficult. Access to a tool that interactively integrates various metrics in one location could potentially be useful to parents of students across different socio-economic and racial backgrounds.

Dr. Saras Chung is a Foundation Director of Educational Strategy and Innovation at SkipNV, an innovative St. Louis-based think tank that seeks to increase access to high-quality education for the nation's most vulnerable youth. One of her group's goals is to help parents find educational opportunities for their children in elementary and secondary school settings. The group is in need of a tool that helps parents visualize available schools and make informed decisions. Dr. Chung has found that school progress over several years is an informative approach to use and is particularly interested in enabling parents to use time series data in their decisions.

## Previous Work

The Harvard Visualization course had a past project similar to the one we are seeking to build: http://itisaasta.com/nycs/. However, the tool created there was built for New York City, is not terribly accessible for the average person, and does not emphasize school performance over time. Furthermore, we felt that the visualizations used in the above project, such as parallel coordinates, were not particularly intuitive to the average user; we wanted to maximize information gain and intuitiveness, rather than just the former.

Additionally, we felt that the Harvard Visualization lacked the ability to intuitively compare different values between schools over time. This posed a unique challenge as we were unable to find any existing visualizations that effectively compared multiple values across different entities over time (in this case, we wish to compare selected metrics across different schools over time). We also tried to avoid using 3D visualizations as a way of adding the time component to our data as we have found that 3D graphics are often difficult to read and do not simplify the ability to read a given set of data. Furthermore, 3D graphics can also be difficult for the average user to read as 3D graphics are not a common method of displaying data and these users lack experience in utilizing these types of visualizations.

In order to accomplish this goal, we've developed two types of visualizations that users can choose between: a heatmap that shows a gradient over time for multiple metrics and multiple

schools, and a series of line graphs where each metric has its own graph containing the appropriate values from every selected school.

## Project Objectives

We want to create a tool that can help parents decide which primary and secondary school they should send their child to, with particular emphasis on school progress over time. The main goals we would like our visualization to answer are as follows:

1. How can a user determine which school in the St. Louis metropolitan region is best suited for their child?
2. How can a user easily understand the ways in which a school has changed over time, even across multiple metrics (e.g. graduation rates, SAT scores, etc.)?
3. How can a parent easily compare statistics across multiple schools, especially given the multidimensional aspects of the data we are interested in presenting?
4. How can a parent incorporate their particular interests in selecting a school?

By answering these questions, we hope to accomplish the following:

1. Enable users to create a list of schools of interest based on their personal criteria.
2. Allow users to compare multiple metrics and multiple time series in an intuitive and interactive way.
3. Consolidate statistics that are hard to find or interpret for the average user.

## Data

Our data comes from the Missouri Department of Elementary and Secondary Education, which provides a series of public CSV files detailing statistics on schools across Missouri (https://mcds.dese.mo.gov/Pages/default.aspx). We are using data such as MAP Test "Proficiency" scores from each school, school district spending per student, number of disciplinary incidents per year, etc.

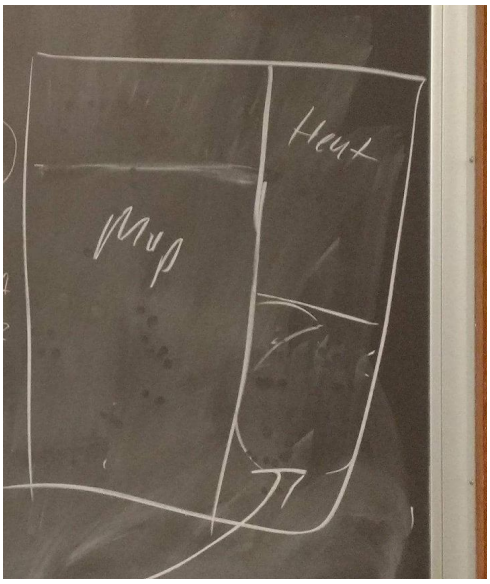
We also have geographic boundary information for each school district in St. Louis County and the St. Louis City that will be used to build a geographic layout of the St. Louis region (https://data-stlcogis.opendata.arcgis.com/datasets/school-districts and https://github.com/eleanortutt/stlcity-wards-2010/blob/master/stlcity-wards-2010.geojson).

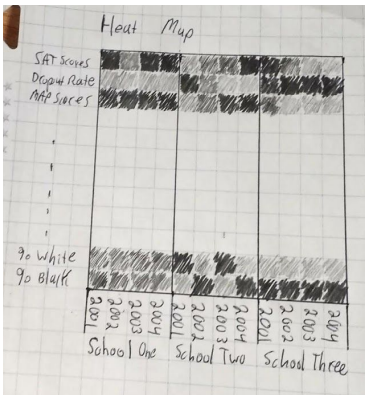
We created a python scripts to run the various raw data files and combine the information into JSON files for our visualizations.

## Exploratory Data Analysis

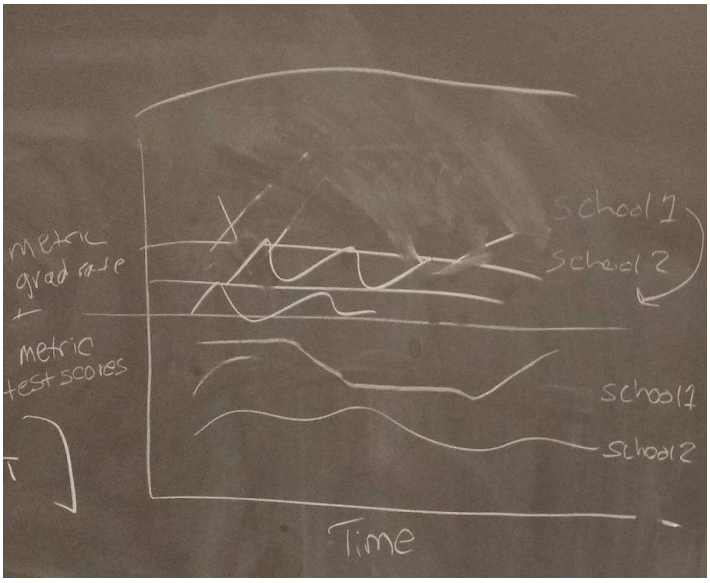
Most of our exploratory data analysis was performed in finding out what types of information we have access to, as we want to leave most of the exploration to parents for their use. We searched through more than a dozen spreadsheets containing different types of data our users could visualize and pared it down to that information which is most interesting.

## Design Evolution

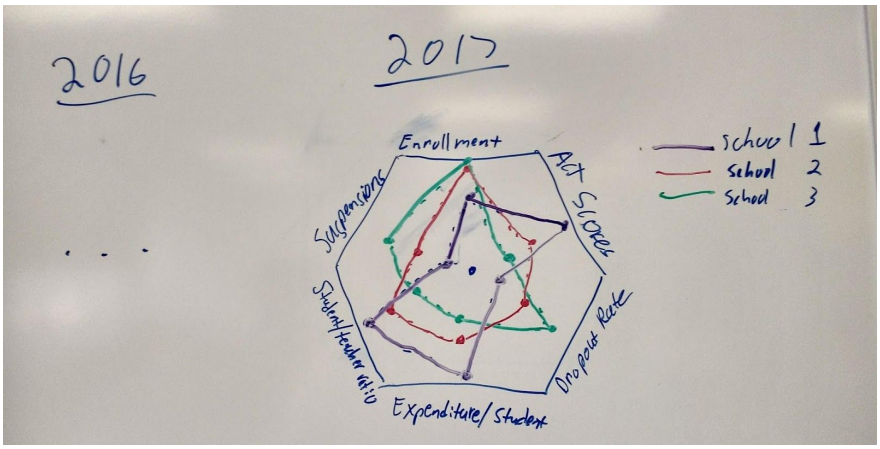
1. **Initial Ideas**
    a. General Layout
        i. 

    b. HeatMap
        i. 

    c. Line Charts

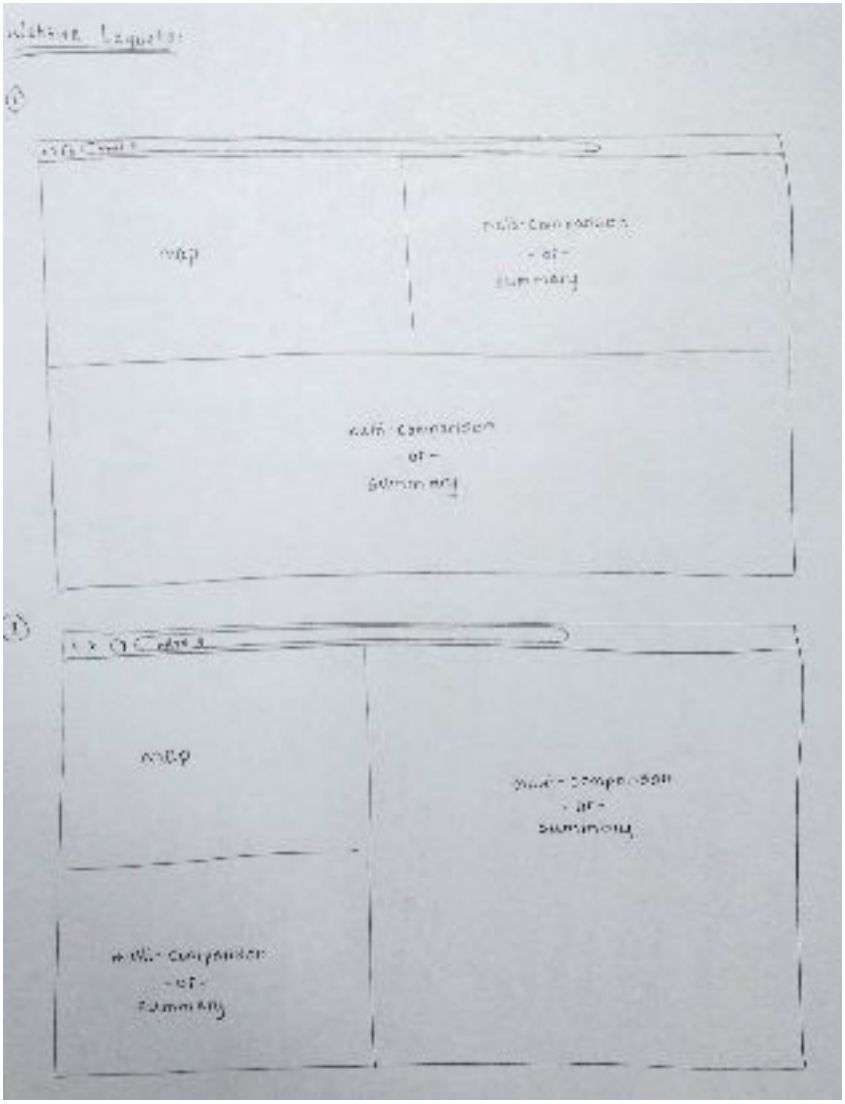
i. 


d. Star Charts

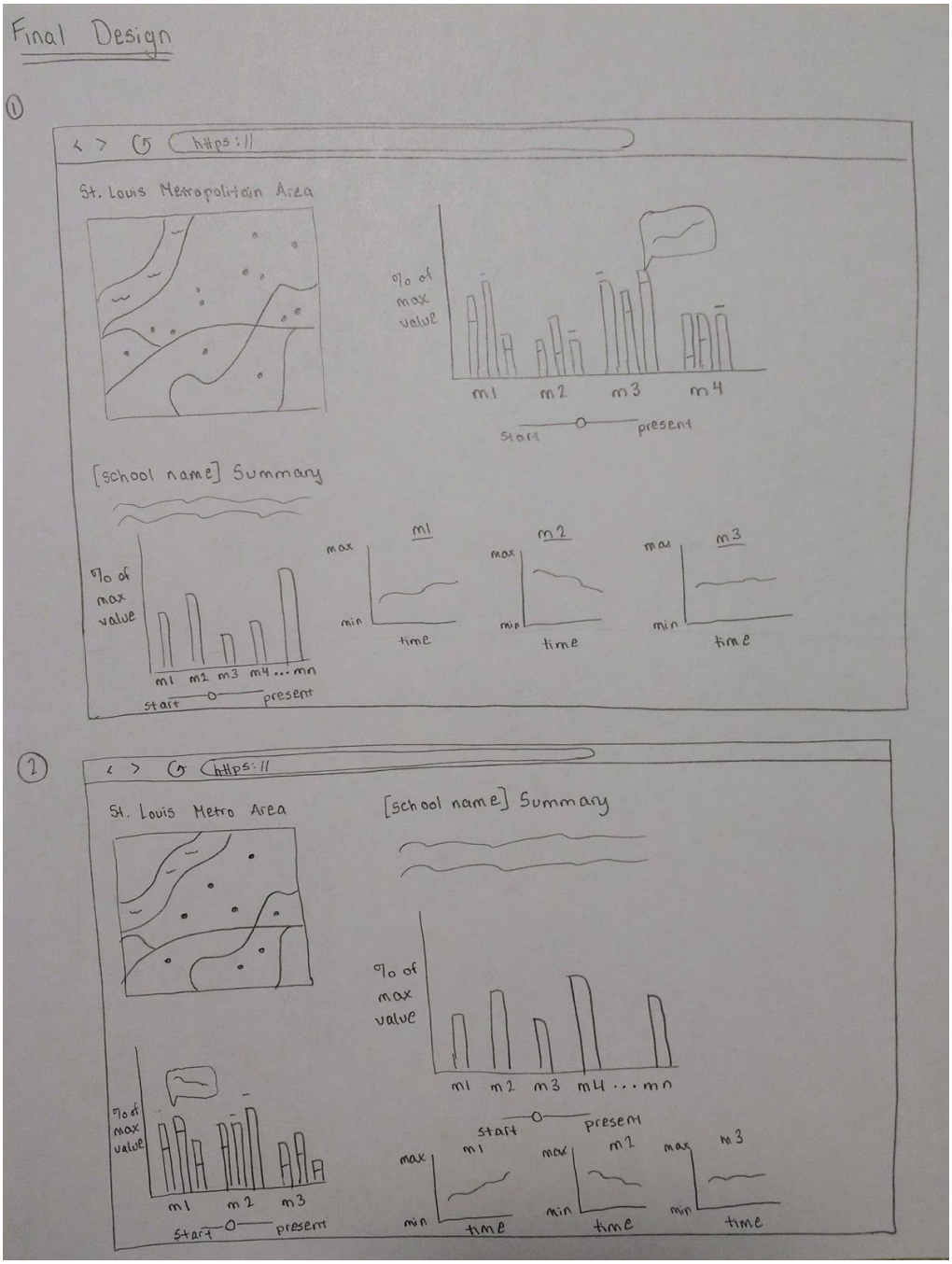
i. 


**2. For Project Proposal**

a. General Layout

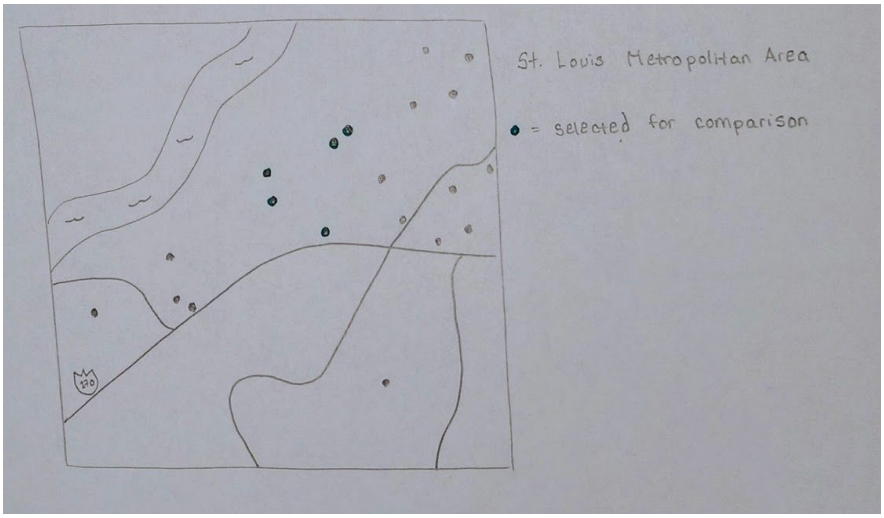
i. 

ii.


Final Design
1
https://
St. Louis Metropolitan Area
% of max value
m1 m2 m3 m4
start
present
[school name] Summary
% of max value
m1 m2 m3 m4 ... mn
start
present
m1
max
min
time
m2
max
min
time
m3
max
min
time
2
https://
St. Louis Metro Area
[school name] Summary
% of max value
m1 m2 m3
start
present
% of max value
m1 m2 m3 m4 ... mn
start
present
m1
max
min
time
m2
max
min
time
m3
max
min
time

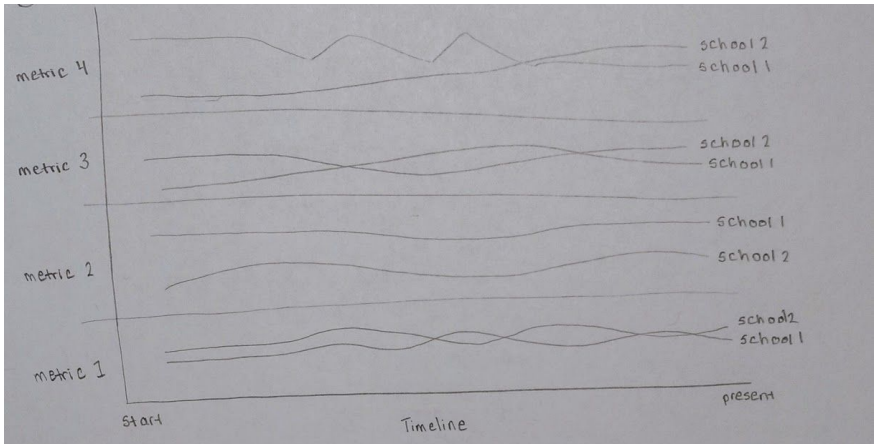
b. Map

i.



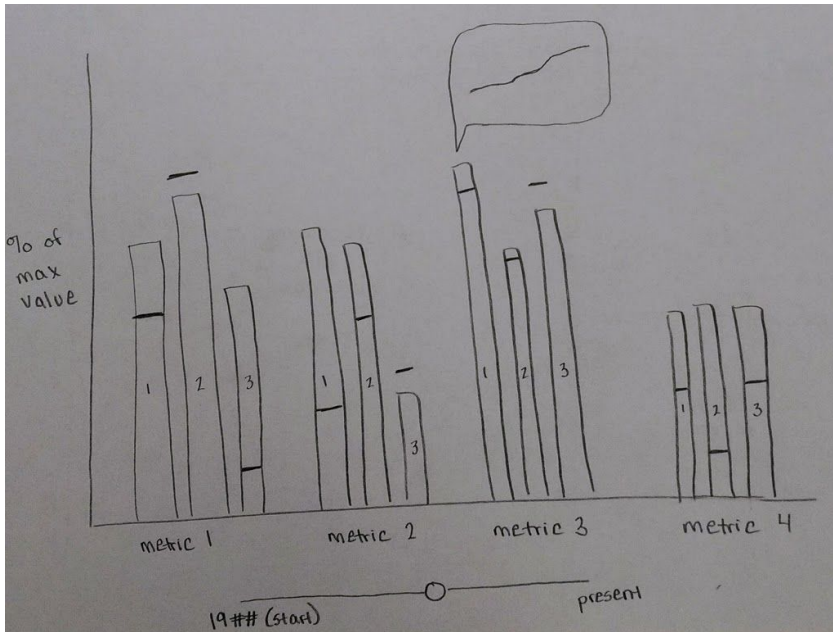

c. Multi-Comparison Line Chart

i.



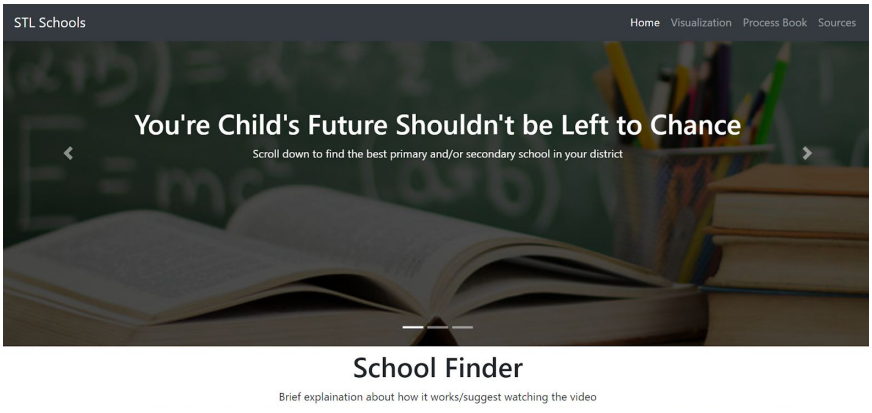

d. Bar and Spark Charts

i.



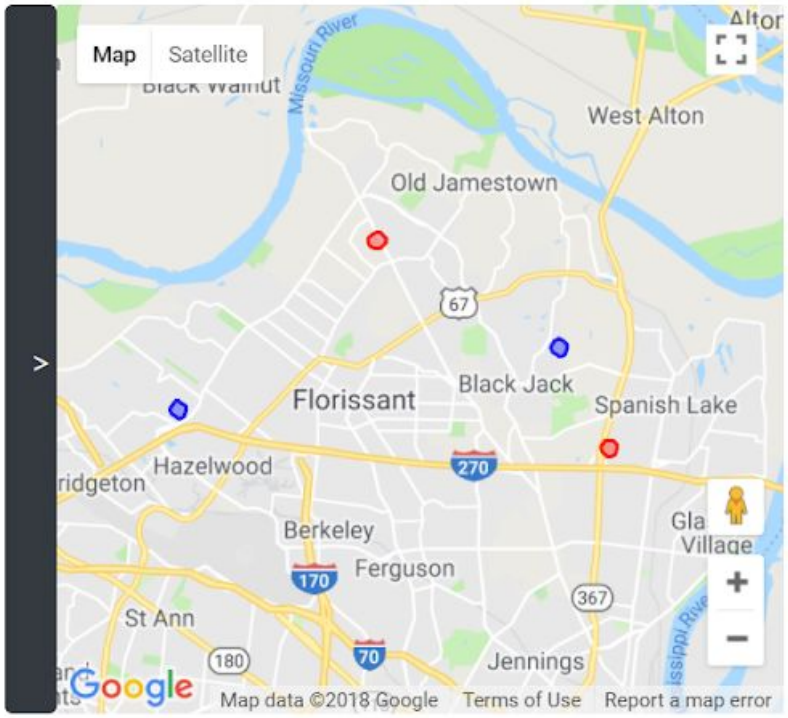
## 3. Prototypes

a. Website Template and Basic Layout




i.

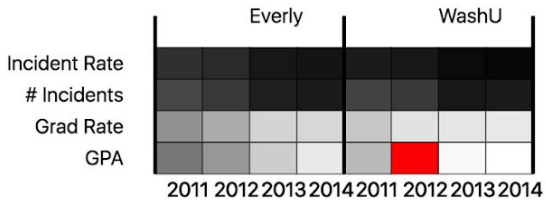
b. Map




i.

1. Blue dot if selected, red otherwise

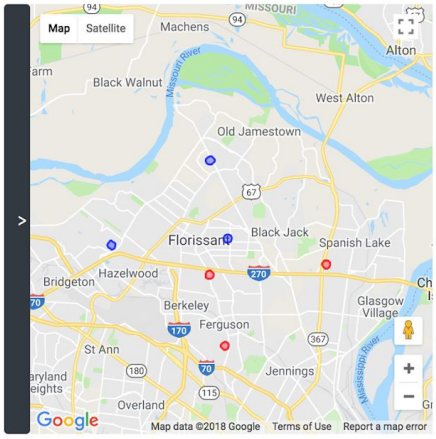
c. Heatmap

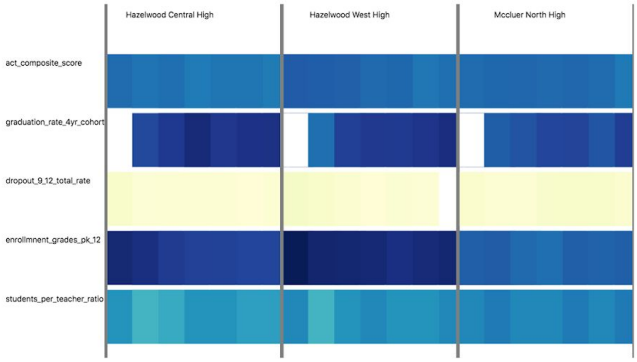



i.

1. White if high, black if low, red if data is missing or N/A

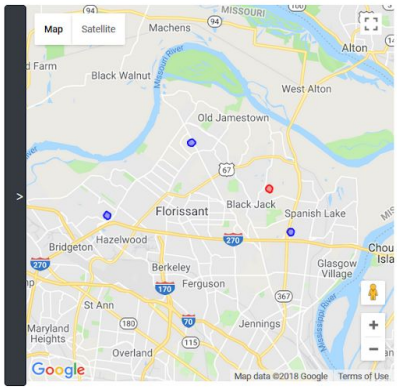
ii.

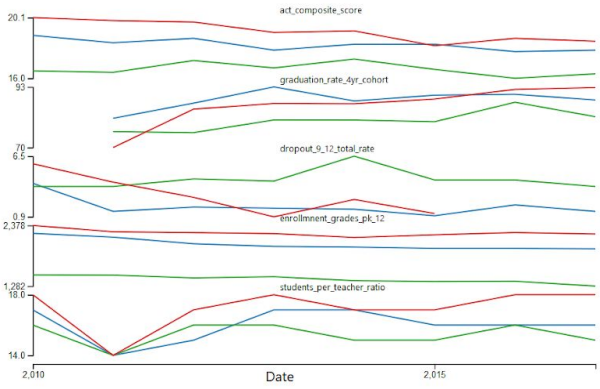



1. Blue if high, yellow if low, white is data is missing or N/A

d. Multi-Comparison Line Charts

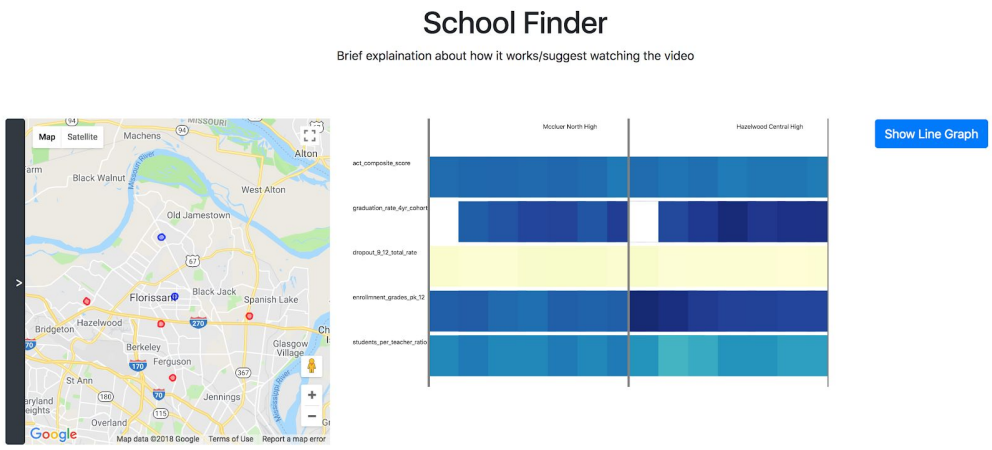


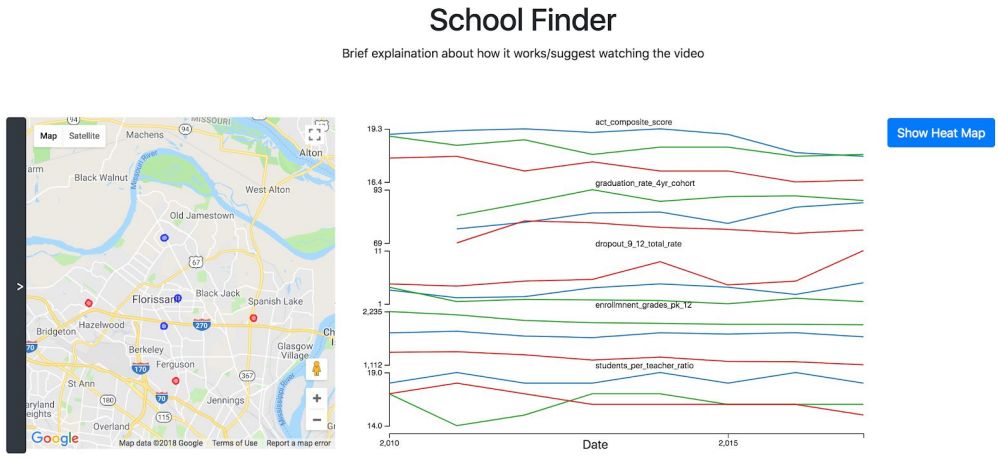

i.

e. Multi-Comparison with Toggle Button

i.

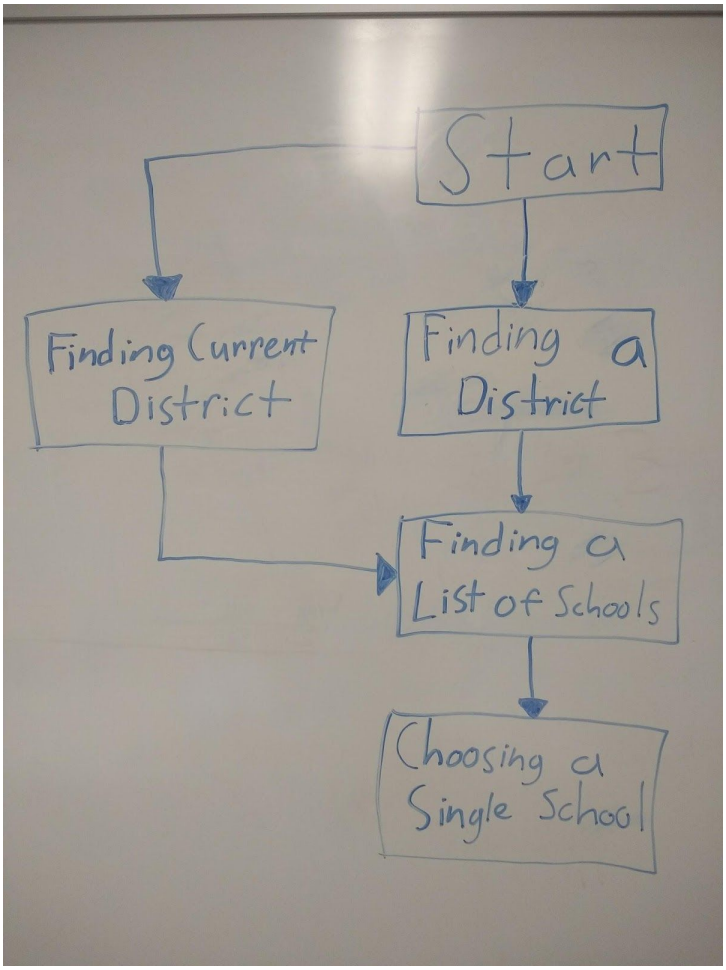



ii.

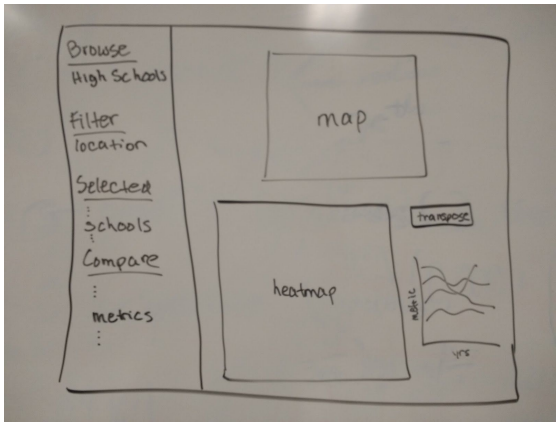


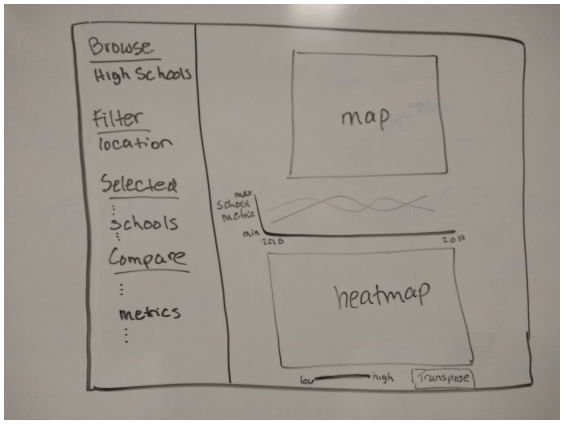

## 4. Final Visualization and Website

a. Anticipated Workflow



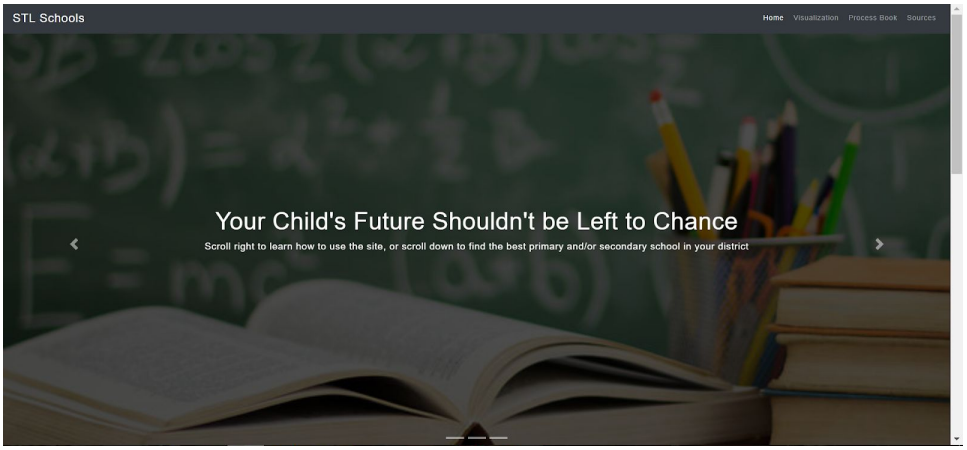

b. Website Redesign given Anticipated Workflow



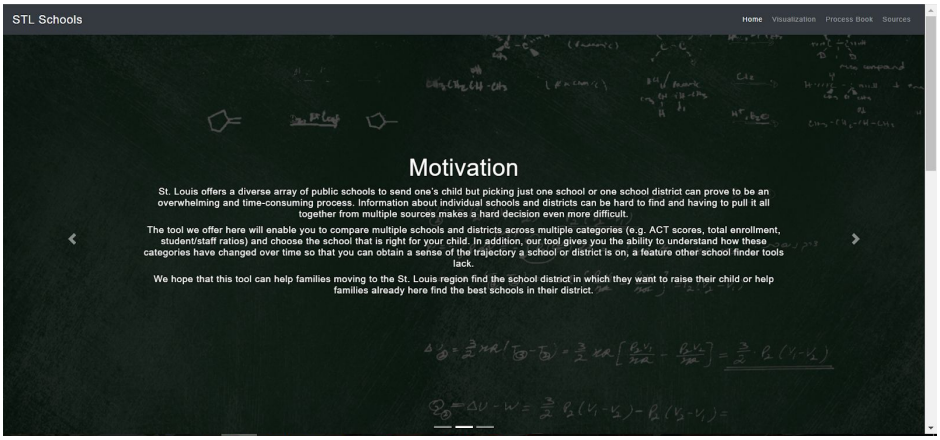




c. Introduction to Website

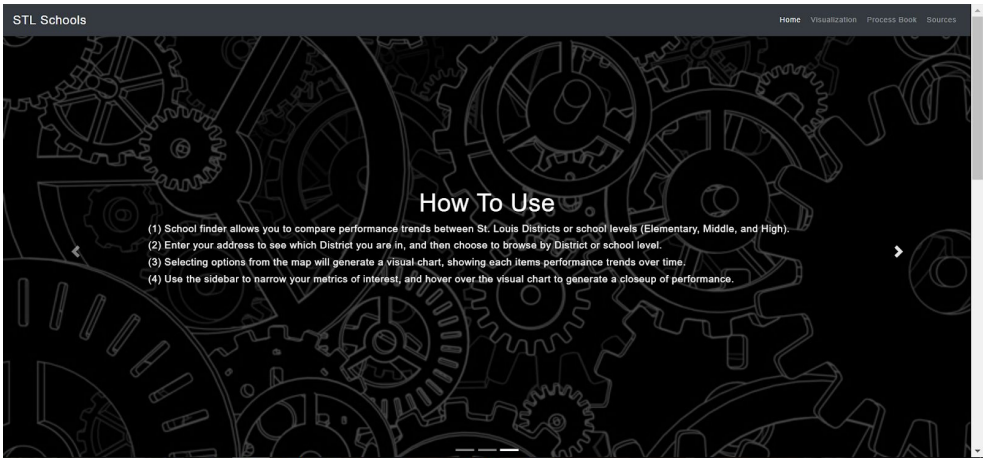
i.

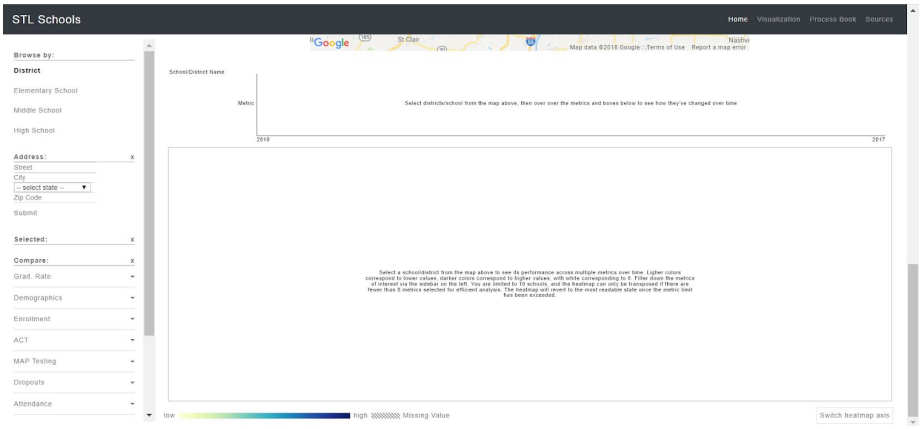

STL Schools
Home
Visualization
Process Book
Sources
Your Child's Future Shouldn't be Left to Chance
Scroll right to learn how to use the site, or scroll down to find the best primary and/or secondary school in your district

ii.

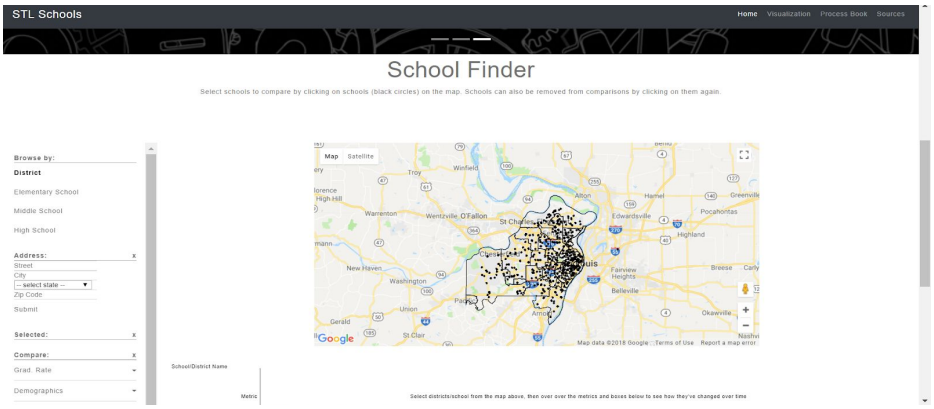

STL Schools
Home
Visualization
Process Book
Sources
Motivation
St. Louis offers a diverse array of public schools to send one's child but picking just one school or one school district can prove to be an overwhelming and time-consuming process. Information about individual schools and districts can be hard to find and having to pull it all together from multiple sources makes a hard decision even more difficult.
The tool we offer here will enable you to compare multiple schools and districts across multiple categories (e.g. ACT scores, total enrollment, student/staff ratios) and choose the school that is right for your child. In addition, our tool gives you the ability to understand how these categories have changed over time so that you can obtain a sense of the trajectory a school or district is on, a feature other school finder tools lack.
We hope that this tool can help families moving to the St. Louis region find the school district in which they want to raise their child or help families already here find the best schools in their district.

iii.


STL Schools
Home
Visualization
Process Book
Sources
How To Use
(1) School finder allows you to compare performance trends between St. Louis Districts or school levels (Elementary, Middle, and High).
(2) Enter your address to see which District you are in, and then choose to browse by District or school level.
(3) Selecting options from the map will generate a visual chart, showing each items performance trends over time.
(4) Use the sidebar to narrow your metrics of interest, and hover over the visual chart to generate a closeup of performance.

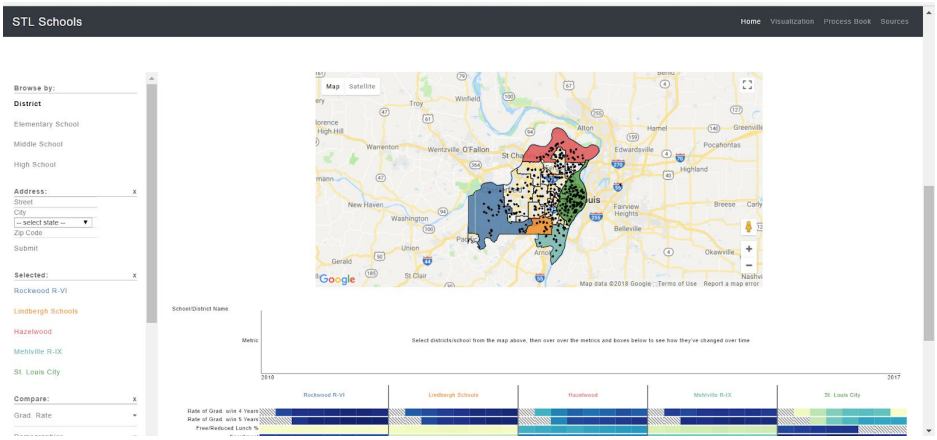
iv. 


1. View before data is selected

v. 


1. View before data is selected

d. Districts selected

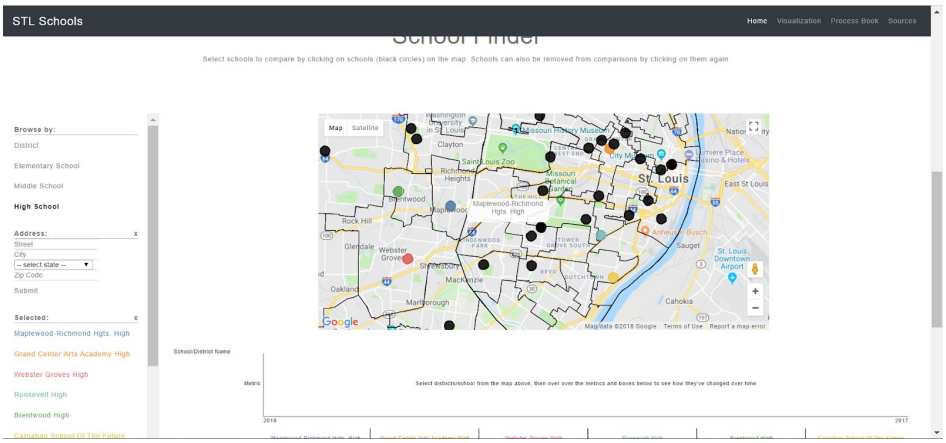
i. 


1. Notice that the selected districts get added to a “Selected” list. Users can click on items in this list to remove them from the list, and to remove them from the map and heatmap visualizations.

ii. 


1. Hovering over the category "% Proficient in Math" for the Lindbergh Schools district. Notice that the line chart above the heatmap highlights the line corresponding to the district, and a point with the information corresponding to the box being hovered over is graphed.

e. Schools selected

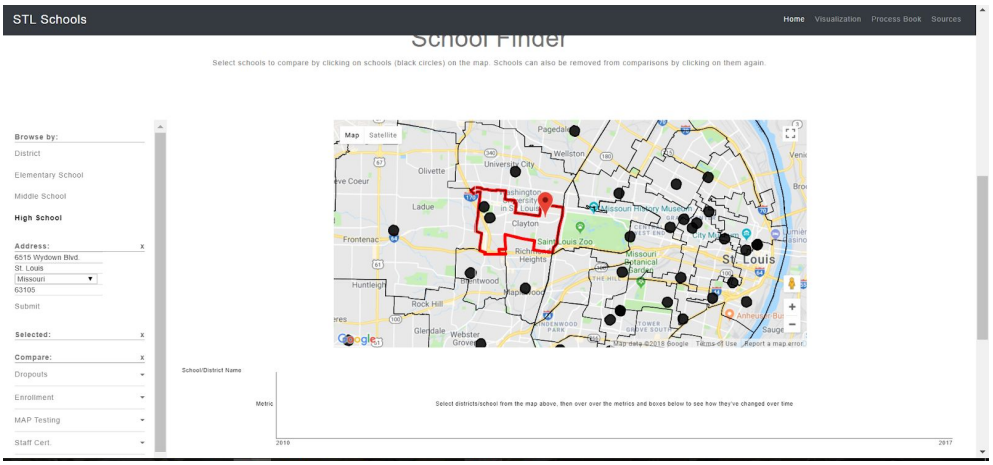
i. 


1. The map with high schools selected--each school gets added to a list of "selected schools" and gets a unique color.

ii. 

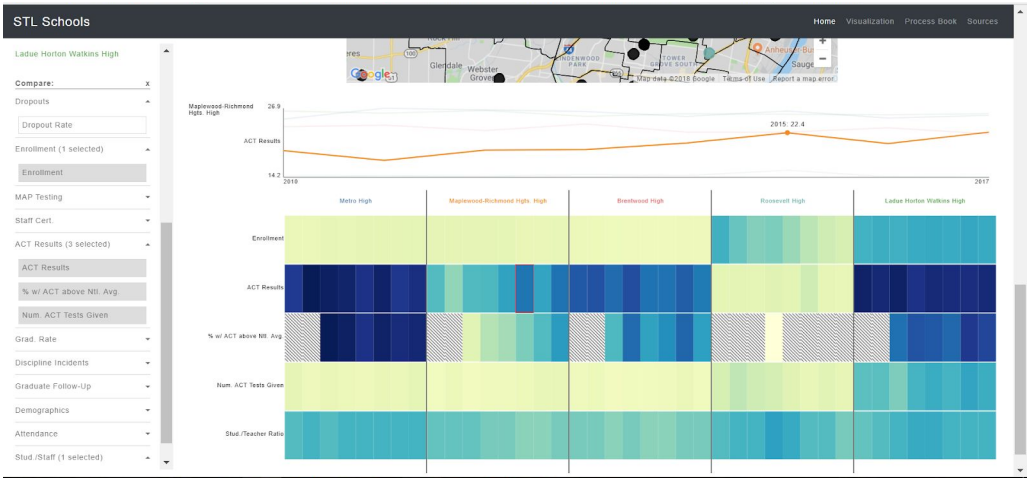
1. Heatmap with high schools selected. Some schools are newer or do not report a lot of information, so years for which there is no value are given a hatching.

f. Finding a District Based on an Address

i. 

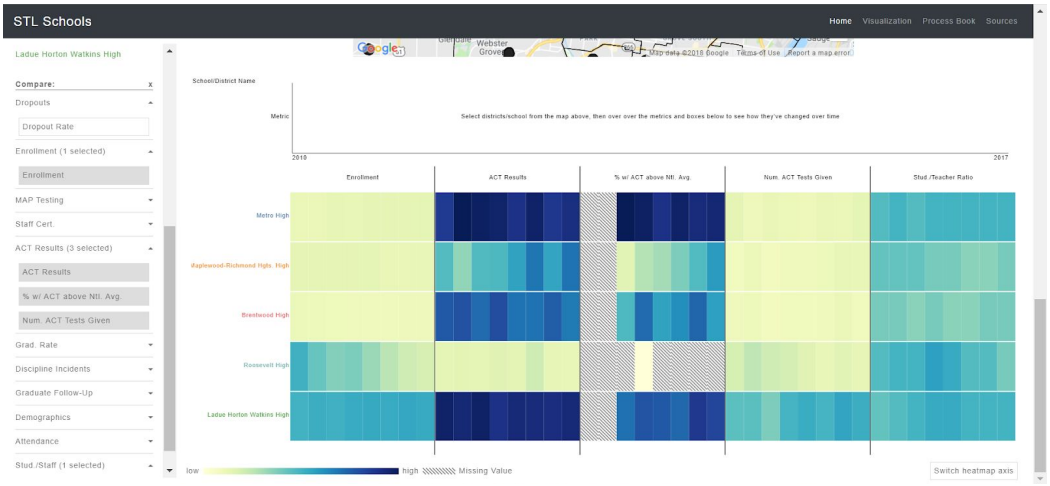

1. After entering an address in the sidebar, a pin will be placed on the map and the surrounding school district will be highlighted, indicating what school district that address would belong to

g. Comparing Schools Based on Certain Metrics

i. 


1. Here, “Enrollment”, “ACT Results”, “% w/ ACT above Ntl. Avg.”, “Num. ACT Tests Given”, and “Stud./Teacher Ratio” are selected

ii.




1. Transposed version of the previous image. Only heatmaps with 8 or fewer metrics can be transposed, however transposing it so that metrics are in columns can make it easier to compare schools across a single metric across time.

# Implementation

## Ideas

1. Implement a St. Louis-shaped map as opposed to a Google Maps-based map.
2. Ability to select school zone/district based on user's address.
3. Enabling ability to filter schools by distance from user's address and by metrics over a certain threshold.
4. Enabling users to add schools to a 'Favorites' list that get displayed in a heatmap or multiple line-graphs.

## Actually Implemented

Our visualizations are meant to guide a parent through the process of discovering and pinning down a particular school of interest for their child. We seek to mimic the natural workflow a parent might use in choosing schools: finding what is out there, looking at broad summary statistics, and finally investigating more detailed statistics about an individual school. We hope to accomplish this through our three visualizations: a map of the St. Louis region with dots representing individual schools that users can select and add to a list of cross-school comparisons, a multi-school comparison heatmap graphs individual metrics of interest across time and across schools, and a line chart that compares all schools for the currently selected metric (based on what metric is hovered over in the heatmap).

Further discussion revealed a different workflow than initially anticipated: we thought parents would mainly consider schools within their district--due to the difficulty of switching districts--and therefore would care more about the multi-school, multi-metric comparison and

school summary than districts. In reality, however, many parents may not know what district they are in, or consider moving districts for schools. Therefore, we modified our layout to account for the workflow. Now, parents can browse by district or school type (elementary, middle, or high school) and enter in their current address to see their district. As schools or districts are added, the heatmap is updated, initially showing all metrics as an overview. As parents identify and select metrics of interest, they can use the sidebar to narrow down which metrics are displayed on the heatmap. We removed the summary altogether, as the heatmap and corresponding line graph show the same information, especially if only one school is selected.

Users can decide between whether to compare schools directly or compare entire districts to one another. This allows users to take a depth or breadth approach when searching for schools, and lets users customize their search to their needs.

Users can also enter their address in the sidebar, automatically highlighting the school district associated with that address and allowing them to quickly narrow down potential schools to compare if they are only interested in staying within their current district. Users may also select specific metrics to compare between schools or between districts by using the "Compare" options on the sidebar which gives more flexibility if the user is only interested in a specific set of metrics. When comparing between these metrics, if a user has 8 or fewer selected, they can choose to transpose the heatmap to make it easier to compare schools across a single metric. This transpose function places metrics in the columns and schools in the rows. Schools and districts can quickly be removed from the list of selected items by clicking on the name of the item in the sidebar.

User testing provided two key insights that were then integrated into the final presentation: line chart use and transpose. Novice heatmap users showed a strong preference for the line graph versus heatmap due to its interpretability and clarity. These benefits, however, were constrained by the number of schools selected for comparison: once more than two or three were selected, there were too many lines and the visualization became confusing. With a greater understanding of the heatmap and multiple schools, individuals began to prefer the heatmap. To balance the pros and cons of both visualizations, we decided to use line graphs as a supplementary visualization to our heatmap: as users hover over different schools/metrics/year boxes, a line graph appears, highlighting the hovered schools trend over the corresponding metric and providing absolute numbers for the given year. Thus, individuals can use the line graph to confirm their interpretations of data trends identified in the heatmap. To see all schools across a given metric, users can also hover over the metric name.

Next, users found comparing multiple schools across a single metric difficult in the heatmap, as the boxes were not close together and therefore minute color differences were easily lost. To simplify these comparisons, we added a transpose button, which swaps the x and y axes. Users are only able to transform the data if less than 8 metrics are selected to maximize interpretability and comprehension.

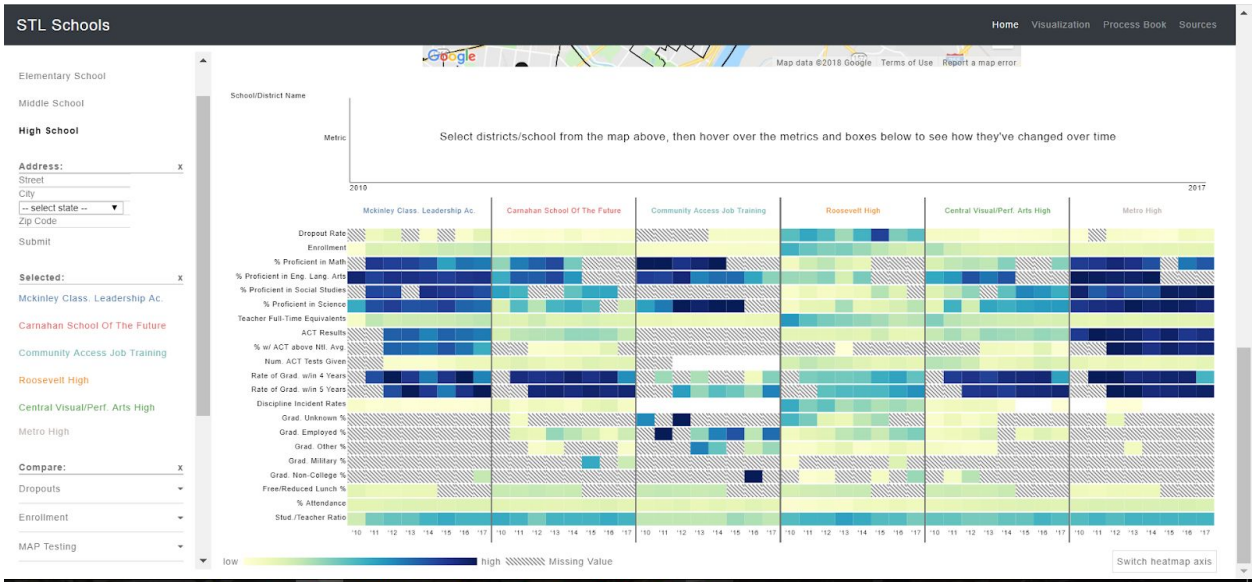
## Evaluation

Playing around with the heatmap with our initial color scheme (grayscale), we did not see much difference between schools in our toy data set and were somewhat disappointed with the information that could be learned from it. However, when we changed to a blue-purple and finally to a yellow-green color scheme, differences between schools immediately became apparent and we were able to make quick judgments about which schools were the best. Color plays an important part in our heatmap implementation. We also removed the grid lines in the heatmap to help highlight the change in colors between years.

Using the heatmap, we can see interesting, and usually expected, correlations in the data. For example, schools with higher average ACT scores over the years also tend to have higher graduation rates. One school (McCluer High) with a slightly declining average ACT score also has a somewhat declining graduation rate, a declining enrollment, and, surprisingly, a declining students-per-teacher-ratio.

Using the district-level information, we can see that the Clayton school district is a better district than the other districts selected (figure below). Generally, it has higher test scores, higher graduation rates, lower numbers of disciplinary incidents, and lower student:teacher ratios than the other districts. For parents looking to move into a new school district, the Clayton area might be a good place to go.


A comparison of high schools within the St. Louis City district shows that several visual areas of the heatmap make comparisons the easiest (figure below). In general, more blue in rows 3-12 is indicative of better schools, as this area contains Missouri testing and information about the ACT. However, trends can also be seen in the top row, which contains dropout rate information, and the row called “Discipline Incident Rates,” both of which are indicative of schools that do not have as much success training their students. These rows are also visually correlated with colors indicating lower test scores.


STL Schools
Home
Visualization
Process Book
Sources
Elementary School
Middle School
High School
Address:
Street
City
-- select state --
Zip Code
Submit
Selected:
Mckinley Class. Leadership Ac.
Carnahan School Of The Future
Community Access Job Training
Roosevelt High
Central Visual/Perf. Arts High
Metro High
Compare:
Dropouts
Enrollment
MAP Testing
Map data ©2018 Google
Terms of Use
Report a map error
School/District Name
Metric
Select districts/school from the map above, then hover over the metrics and boxes below to see how they've changed over time
2010
2017
Mckinley Class. Leadership Ac.
Carnahan School Of The Future
Community Access Job Training
Roosevelt High
Central Visual/Perf. Arts High
Metro High
Dropout Rate
Enrollment
% Proficient in Math
% Proficient in Eng. Lang. Arts
% Proficient in Social Studies
% Proficient in Science
Teacher Full-Time Equivalents
ACT Results
% w/ ACT above Ntl. Avg.
Num. ACT Tests Given
Rate of Grad. w/in 4 Years
Rate of Grad. w/in 5 Years
Discipline Incident Rates
Grad. Unknown %
Grad. Employed %
Grad. Other %
Grad. Military %
Grad. Non-College %
Free/Reduced Lunch %
% Attendance
Stud./Teacher Ratio
'10 '11 '12 '13 '14 '15 '16 '17
low
high
Missing Value
Switch heatmap axis