My full project and interactive visualization can be found at this URL:

https://mybinder.org/v2/gh/jhoglund1/DTSA-5304/main?urlpath=voila%2Frender%2FDTSA5304-Final.ipynb

(It may take some time to load).

In case the URL doesn't load for some reason, I am pasting all my text below as well as screenshots of my visualization.

Data

After some research, I selected a dataset from ProRepublica's Data Store. This dataset is comprised of data relating to regular and alternative U.S. schools within the same school district. According to the Department of Education, an alternative school is established "to serve at-risk students whose needs are not being met in the traditional school setting". This data collected by ProRepublica is for the 2013-2014 school year is categorized by school districts and its state. Also included is data relating to pupil funding, the percentage of schools that have a counselor, the average student to teacher ratio, the percentage of teachers who are in their first or second year of teaching, the percentage of teachers who are absent more than 10 school days, and the graduation rates. These factors are included for both alternative schools and regular schools in the same school district.

Goals

After some preliminary investigation of the dataset, I was interested in examining the differences between the regular and alternative schools for each factor. Of all the available metrics, I thought that the graduation rate was the most general and the easiest to assess. I was curious to see if any of the other factors had any sort of effect on the graduation rates across the school districts.

Task

Goal: To create an exploratory visualization to gain knowledge about how alternative schools compare to regular schools across the U.S.

Means:
- Conducted through a scatter plot with graduation rates (y-axis) vs various factors users can select by hand
- Scatter plots for both regular and alternative schools

Characteristics:

- High level trend comparison between regular and alternative schools for each factor
- Low level point by point comparison of each school district

Target: Uses a relative reference frame, comparing against two points of data within the dataset

Workflow:
- Users can:
  - Select different x-axes to look at
  - Click on points to identify other points in the same region
  - Click on points to do a regular vs alternative comparison of the specific school district
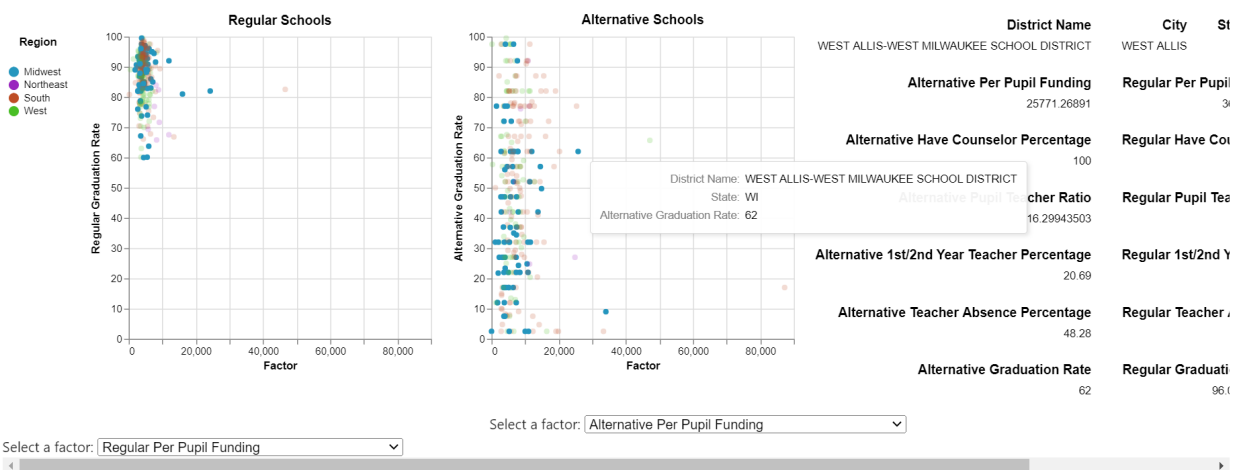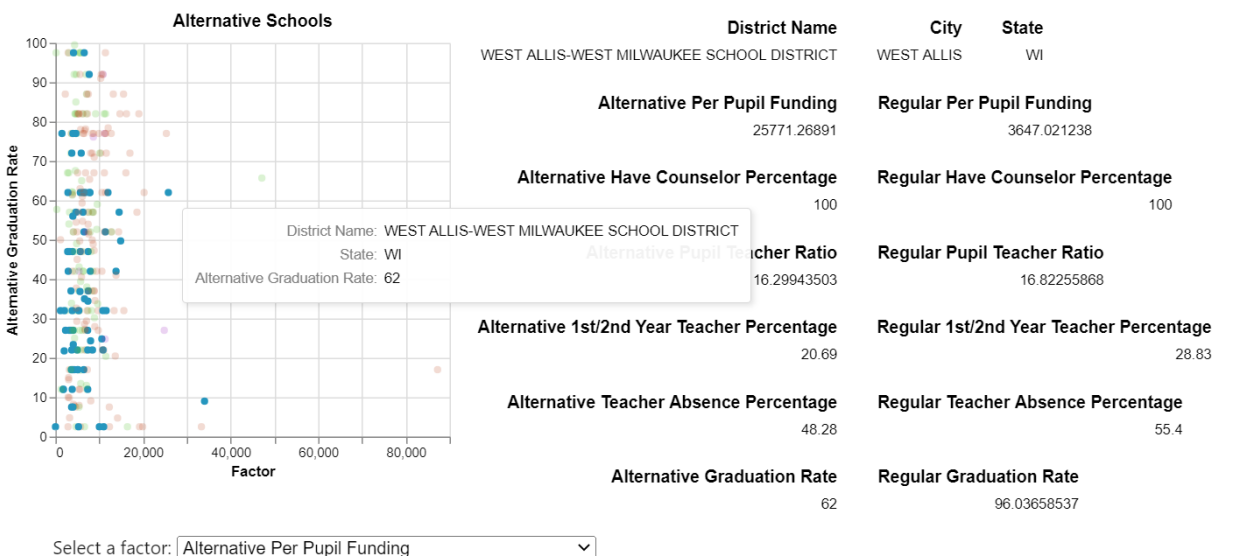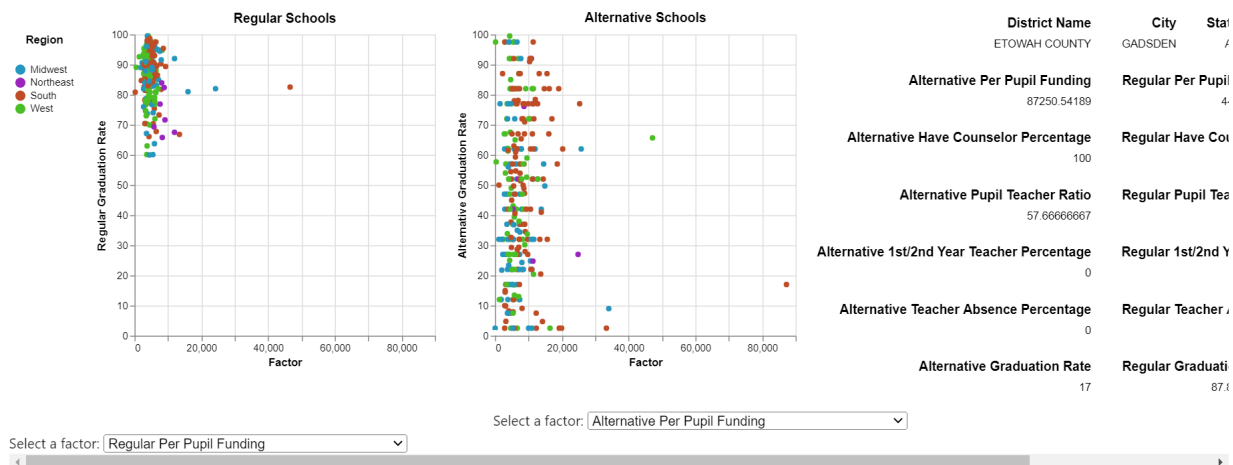
Roles: Anyone can execute this task

Visualization

How to Use:

- Users can hover over points in each scatterplot to receive preliminary information about the point.
- Users can click on individual points in each scatterplot.
  - This will highlight all the points in the scatterplot that are in the same region of the U.S. A legend is on the left-most side.
  - This will change the data table on the right-most side to present data for that individual school district.
- Users can click anywhere on a scatterplot that is not a point to remove the region highlighting.
- Users can select through the two dropdowns to change the x-axis for each scatterplot.
  - The left-most dropdown changes the left scatterplot (regular schools) and the right-most dropdown changes the right scatterplot (alternative schools).

Screenshots of Visualization:

Regular Schools

Alternative Schools

	District Name	City	Stat
	ETOWAH COUNTY	GADSDEN	A

Alternative Per Pupil Funding	Regular Per Pupil
87250.54189	4

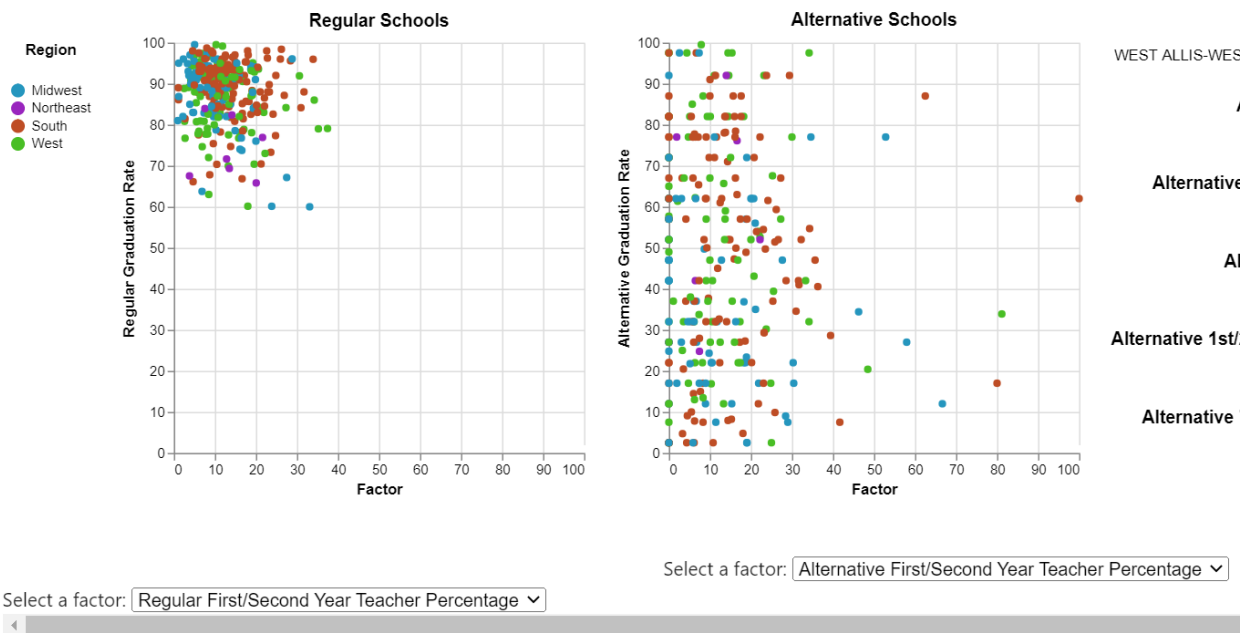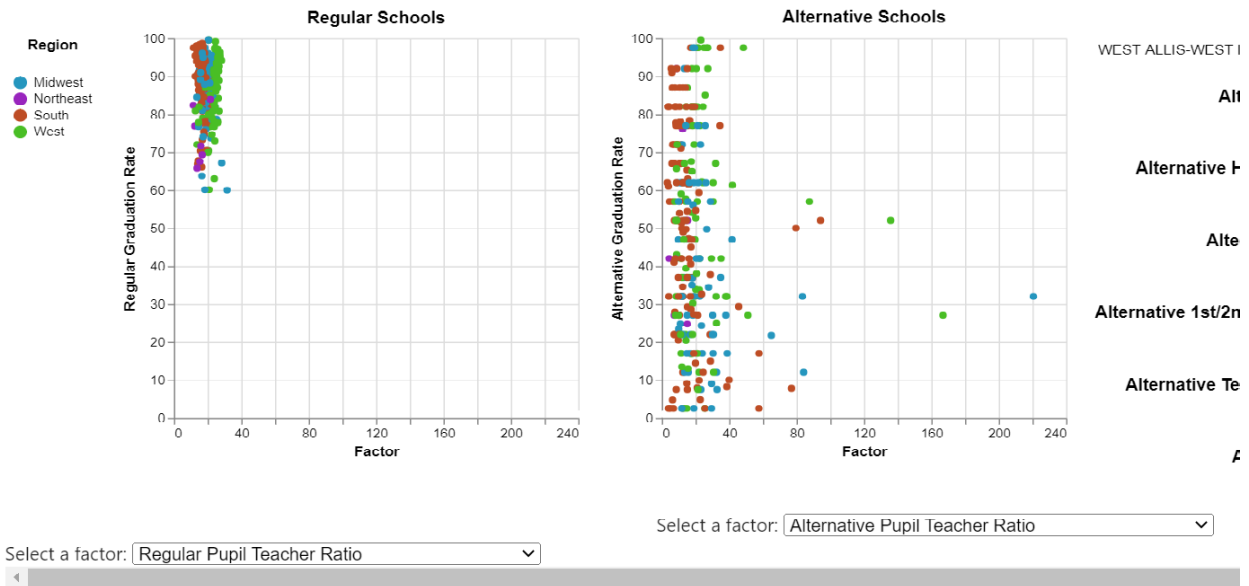| Alternative Have Counselor Percentage	Regular Have Cou
100	

Alternative Pupil Teacher Ratio	Regular Pupil Tea
57.66666667	

Alternative 1st/2nd Year Teacher Percentage	Regular 1st/2nd Y
0	

Alternative Teacher Absence Percentage	Regular Teacher A
0	

Alternative Graduation Rate	Regular Graduati
17	87.8

Select a factor: Alternative Per Pupil Funding ▼

Select a factor: Regular Per Pupil Funding ▼

Alternative Schools

District Name	City	State
WEST ALLIS-WEST MILWAUKEE SCHOOL DISTRICT	WEST ALLIS	WI

Alternative Per Pupil Funding	Regular Per Pupil Funding
25771.26891	3647.021238

Alternative Have Counselor Percentage	Regular Have Counselor Percentage
100	100

Alternative Pupil Teacher Ratio	Regular Pupil Teacher Ratio
16.29943503	16.82255868

Alternative 1st/2nd Year Teacher Percentage	Regular 1st/2nd Year Teacher Percentage
20.69	28.83

Alternative Teacher Absence Percentage	Regular Teacher Absence Percentage
48.28	55.4

Alternative Graduation Rate	Regular Graduation Rate
62	96.03658537

Tooltip: District Name: WEST ALLIS-WEST MILWAUKEE SCHOOL DISTRICT | State: WI | Alternative Graduation Rate: 62

Select a factor: Alternative Per Pupil Funding ▼

Regular Schools

Alternative Schools

	District Name	City	St
	WEST ALLIS-WEST MILWAUKEE SCHOOL DISTRICT	WEST ALLIS	

Alternative Per Pupil Funding	Regular Per Pupil
25771.26891	3

Alternative Have Counselor Percentage	Regular Have Cou
100	

Alternative Pupil Teacher Ratio	Regular Pupil Tea
16.29943503	

Alternative 1st/2nd Year Teacher Percentage	Regular 1st/2nd Y
20.69	

Alternative Teacher Absence Percentage	Regular Teacher A
48.28	

Alternative Graduation Rate	Regular Graduati
62	96.0

Tooltip: District Name: WEST ALLIS-WEST MILWAUKEE SCHOOL DISTRICT | State: WI | Alternative Graduation Rate: 62

Select a factor: Alternative Per Pupil Funding ▼

Select a factor: Regular Per Pupil Funding ▼

Regular Schools

Regular Graduation Rate

Region
- Midwest
- Northeast
- South
- West

Factor

Alternative Schools

Alternative Graduation Rate

Factor

WEST ALLIS-WEST

A

Alternative

Alt

Alternative 1st/2

Alternative T

Select a factor: | Regular Have Counselor Percentage ⌄ |

Select a factor: | Alternative Have Counselor Percentage ⌄ |

Alternative Per Pupil Funding
Alternative Have Counselor Percentage
Alternative Pupil Teacher Ratio
Alternative First/Second Year Teacher Percentage
Alternative Teacher Absence Percentage

Design Summary

Design Summary

Key Elements

1. High level Regular vs Alternative School comparison
   - The largest part of the visualization is the two scatterplots depicting data relating to regular schools and alternative schools. The y-axis is the graduation rate and the x-axis is a dropdown of other pieces of data that can be assessed against the y-axis. These charts allow the user to see a high level overview of general trends

for regular school and alternative schools as well as make comparisons between the two.
2. Low level Point by Point Regular vs Alternative School comparison
    ○ Within each scatterplot is the ability to click on an individual point. This action gives a hover overview of the point with information on the schools' location and its graduation rate. This action also updates the data table on the right side with specific information for that selected school district. This allows the user to identify specific information about individual points and make point to point comparisons between the regular and alternative schools.
3. Variable X-axis
    ○ For both scatterplots, the x-axis is determined by a dropdown menu that contains other information from the dataset. This allows the user to compare different factors against the graduation rate and determine if there are any correlations between the factors and the graduation rate.
4. Regional selection
    ○ When a point is selected from within a scatterplot, all the other points from that same region of the U.S. are also highlighted. This gives the user another avenue of exploration to see how and if the region has any correlation to the graduation rate and how it compares between regular and alternative schools.

Evaluation

Target Question: Since this is an exploratory visualization for this dataset, I am interested in determining if users are able to navigate and understand the visualization well and if they are able to gain meaningful insights from visualization. I determine "meaningful" to be anything that they learn that is new information or is unique or interesting to them.

Testers: I recruited my family and work colleagues to evaluate my design.

Evaluation Measures / Approach / Instantiation: I used insight depth as my primary evaluation measure. With this method, I was able to understand how easy my design was to follow and what kind of insights my visualization produces. I employed journaling as my main evaluation technique. I gave my testers access to my visualization with a few instructions on how to use the tool. I asked them to explore and note any insights they discover and also to note any deficiencies and/or bugs in the design.

Criteria / Results:
My evaluation criteria include:
- testers are able to quickly and fully use and understand all parts of the visualization
- testers are able to gain information and understanding about the data that they did not know previously

Based on these criteria, my visualization accomplished its goal. Users were able to explore and understand the visualization with little difficulty as well as gain new and interesting knowledge about the data. The users also provided good feedback about ways to improve the design, which I will discuss in my Synthesis section.

Synthesis

Findings: From this visualization, myself and my users were able to learn some interesting things about this data. Some common threads include gaining an understanding that there is very little correlation between each of the given x-axis factors and graduation rates for both alternative schools and regular schools. There is, however, a much larger variance in the results for alternative schools pretty much across the board for all the factors when compared to regular schools. From my journaling study, one user noted that region is the most influential factor related to graduation rates. They hypothesize that culture and environment may contribute to alternative school graduation rates.

What worked well: I think the best element of my visualization is having both a high level and low level approach the same data. Users are able to look at broad trends as well as specific numbers for an individual school. This gives the users the ability to learn more about specific points, but it doesn't force them to. I think another element that worked well was the variable x-axis. This allows the user to have the freedom to investigate several different avenues of trends and relationships in an interactive and concise way.

Ways to improve: There are several ways that this visualization could be better. One way includes making it so clicking on a point in one chart will highlight the same point in the other chart. This would help with visual point to point comparisons on the scatterplots themselves and help the user not be as reliant on the data table. Another improvement could include linking the data together enough behind the scenes so that only one x-axis dropdown is needed. This would save the user some time having to update both dropdowns to do comparisons. There are also several visual improvements that I was unable to get to work, like aligning the text properly in the data table and lining up the dropdown menus better.