What Factors Impact Student Success in Washington State?

Gautam Moogimane, Alec Olschner, Tucker Standlee

VISUALIZATION

Tableau Online: https://goo.gl/bpgStK (Tableau online account required)
Tableau Public: https://goo.gl/R27Rkb (Viewable by anyone)

EXECUTIVE SUMMARY

The purpose of our project was to create a visualization that can be used by decision makers in education for the state of Washington to explore what factors within state or district control might influence student success. We used data primarily pertaining to high school students to measure student success at the district level and connected these outcomes with potential factors that may lead to the level of student success.

The intended users of this visualization would be decision-makers in education in the state of Washington. This could include district and building level administrators, student advocate organizations, parent-teacher associations (PTAs), teachers' unions, as well as state officials in education and government.

BACKGROUND

To research the specific data that potential users would wish to see displayed in our visualization, we conducted user interviews and submitted online surveys to building and school district administrators as well as teachers' union representatives to identify the following:

- What measures of student success they feel are most important.
- What types of decisions they make that could benefit from access to more data.
- What hypotheses they have about specific factors that might influence student achievement, especially factors that could be influenced or controlled.
- What elements of school life and resources offered to students beyond the classroom they feel may contribute to student success.

Based on the findings from our research, we selected specific data from the State of Washington's Office of the Superintendent for Public Instruction (OSPI) via its website: http://www.k12.wa.us/DataAdmin/default.aspx

The specific data points that we selected were not only in line with the ones expressed as crucial in our research, but also listed as specific indicators and success rates that OSPI specifically cited as their own goals for student success. The data points that we collected were the following ratio data:

- Cohort size
- Graduation rate
- College acceptance rate
- Dual credit enrollment
- Ratio of students per teacher
- 9th grade failure rate
- Chronic absenteeism
- Discipline rate

This data was collected for all 322 school districts within the state over the last five years (2012-2016).

DESIGN PROCESS

To determine the best visual representation of our data, our team individually brainstormed potential methods and formats that would allow educational leaders to discover trends and relationships between the success factors and indicators we selected from our data. From here we quickly shared our ideas and grouped similar ideas together. We then categorized the types of visualizations we had come up with and selected a few to explore in more detail (see figures 1-3 in the appendix). We explored geographic mapping with a multivariate visualization, parallel coordinates with filters to select the districts and variables being shown, and a small-multiples dashboard that would allow the user to focus on one primary district and compare it to other similar districts. After discussing each concept and weighing the pros and cons, we decided to focus on the parallel coordinates visualization as we felt it did the best job of visualizing a large amount of quantitative data and filtering, along with other interactions, would allow the user to focus on a specific district and selected variables. We also knew that next to the scatter plot, parallel coordinates are one of the best visualization methods for allowing users to compare multiple types of data at once. This was in line with our goal: to allow these comparisons to aid users in the discovery of potential correlations between success factors and indicators.

After deciding upon a visualization method, we created a paper prototype of our design to assess the viability of the design concept and conducted pilot usability testing with three users (see figure 4 in the appendix). The pilot testing showed that parallel coordinates were an appropriate way to visualize the data and also revealed user's expectations for how they would interact with the visualization if it were digital. From here, we developed an interactive prototype of our parallel coordinate visualization using the Tableau platform in order to conduct usability studies with potential users. The goal of the sessions was to evaluate the usability and utility of the prototype based on the feedback of domain-specific target users.

Research Questions

- Is the design and functionality of the visualization understandable?
- Does the visualization present student success data efficiently?
- Does the visualization lead to insights into the factors impacting student success?

Participant Characteristics

Participant 1:
- High school assistant principal
- 6th year in current position, mid-career
- Technology competency: 3 (intermediate)

Participant 2:
- High school teacher; former teacher's union president for the district
- Late career
- Technology competency: 3 (intermediate)

Participant 3:
- High school assistant principal
- 6th year in current position, nearing retirement
- Technology competency: 2 (novice)

Testing Methods

Moderated Usability Testing with Post-Task Interviews:
- Three scenario-based tasks (printed for the participant to read)
- "Think-Aloud" protocol
- Screen recordings for asynchronous data collection
- Post-task follow-up questions to gauge the difficulty of the task, the usability and utility of the visualization, and any feedback the participant has about improvements or changes.

Task List

Task 1:
You are a member of the Washington State Board of Education and want to get a brief overview of which districts stand out from the rest in certain elements. Using the visualization, determine which district has:
- The highest college acceptance rate
- The lowest absentee rate
- The smallest student to teacher ratio

Task 2:
You are an assistant principal in the Kent school district. Your goal for this year is to improve your school's percentage of students accepted into college/secondary education. In order to do so, you wish to identify what the rate was for your whole district in the most recent record, the year 2016. Using the visualization, determine what the college acceptance rate was for the students in your district during this year.

Task 3:
You are a member of the local district's School Board. Some parents have been complaining about class sizes being too large and that students would be fairing better in school if there were smaller class sizes and more teachers per student. Hoping to investigate this matter, use the visualization to identify if there is indeed any sort of correlation between student success rates and the amount of students per teacher.

Evaluation Results

The evaluation provided a large range of findings that support some initial design decisions as well as pointing to changes and improvements that could be made. For the findings that indicate the need for changes and improvements, we have grouped them by theme and prioritized creating solutions for them.
(Priority Rating: ★★★ = high, ★★ = medium, ★ = low)

Positive Findings

- Performing tasks related to finding the maximum and minimum value of the success indicators were easily accomplished using the visualization.
- Tasks related to filtering data to a particular year, or highlighting values for an individual district, were easy to perform.
- Users found the visualization more helpful and user friendly as compared to other versions that are publicly available.

Problematic Findings

Functionality & Accuracy Issues

★★★ Users found some data points that were misaligned to the vertical scales.

★★★ Users were confused by the values displayed on the data labels. Vertical coordinate values displayed on hover were normalized data points vs. actual values.

★★★ Users were uncertain of the meaning of some measurements.

★★ Users had difficulty in completing the third task efficiently and found it difficult to determine correlations as the task requested.

★ Users had difficulty using some of the filters, and none of them used the functionality to type in values, and instead choose to manually scroll down the list.

★★ Users referenced student income levels as an important factor in student success, but that data was not shown in the visualization.

Visual Characteristics

★★★ Users didn't always use the filters to select a district or change the year based on the prompt given for the task (low discoverability of filters).

★★ Users had difficulty understanding how to scroll through the districts in the filter/ scroll bars in filters.

★★ Users scrolled rather than typing to search for a district in the filter.

★★ One user (P1) complained about the lack of visual aesthetics and found the visualization a little bland.

★★ Users had never heard of some districts that they found in the course of completing Task 1 (probably due to their small size). Also, the cohort size wasn't referenced by users, but provides a helpful way of comparing districts.

★★ Users had to pause when deciding what constituted a measure of student success.

Adjustments

Based on the feedback that we received from the users, we incorporated the following improvements to our parallel coordinate visualization:

- Amended the filter section so that when the user clicks in the specific filter section, a permanent search bar function is listed right below and is always visible while a drop down menu of all of the potential selections appears and clearly indicates the ability to scroll when necessary.
- Included data on the number of students that qualified for Free or Reduced Price Lunches within a school district as a manner of representing the percentage of low income students in each district.
- Created headers over the specific data points that were either success factors or indicators in order to assist users with distinguishing the two different types of data.
- Created information icons below each axis label that offered a more detailed explanation of the data and what it was measuring when clicked or hovered on.
- Included circular marks at the intersection of each district line and y-axis to clearly display the exact location and measurement value for each district and the respective data being measured.
- Included the ability for the user to color code the range of any of the success factors or indicators based on a blue hue for the lower values and and an orange hue for the higher values. This not only allowed users to identify districts that were similar in value within the measurements of the specific factor or indicator, it also improved the visual aesthetic and allowed users to quickly identify general trends in the data.
- Created a second dashboard in the visualization that displayed a scatter plot to assist users in identifying potential correlations more clearly by selecting a desired factor and indicator for any number of districts in any given year to see how they relate and if any direct correlation exists between the factor and indicator.

FINAL EVALUATION

Overall, our team is quite satisfied with the results of our visualization and are confident that it achieves the main functions that we envisioned for it. Our main goals were to take all of the data that we collected about the student success indicators in Washington state and present it all together in a single cohesive display that would allow decision makers within Washington state education to identify potential correlations between success indicators and factors that might prove supportive in their efforts to improve education in the state.

By choosing to create a parallel coordinate graph, we achieved our function of displaying all of the data at once in a coherent manner. With nine quantitative variables to display, parallel coordinates allows each to be mapped using position, which is the strongest encoding for this type of data. Furthermore, each one of the success indicators and factors is measured on different scales with different unit measures and exist independently of one another. The parallel coordinates design allows for an individual Y-axis for each indicator and factor to be scaled and displayed independently. Additionally, a line drawn between each data point across all of the Y-axes for each district visually connects these measures, allowing them to be analyzed at once. For more precise analysis, the user can color the district lines based on a chosen indicator or factor which may reveal patterns in the relationship between the potential contributing factors and the student success indicators. For example, when coloring the visualization based on 9th grade failure rate, it is clear to see that for high 9th grade failure rates, colored in orange, there also tends to be lower dual credit enrollment, higher free and reduced lunch percentage, as well as lower graduation and college attendance rates (see figure 5 in the appendix). This color scheme was chosen specifically because the contrast between colors is strong and can still be seen by users

with the most common types of color blindness. All together, the parallel coordinates graph allows the user to get a sense of the big picture and have an idea of how each success factor and indicator stands throughout the state and can see how all of the districts compare with one another.

While the parallel coordinates can be used to some extent to achieve our essential goal of identifying potential correlations, it is the inclusion of the scatter plot that really makes this goal not only possible, but very simple to achieve. By reducing the axes to only one success indicator and only one contributing factor, the user is able to focus solely on the selected data points and find the exact correlation between the two measurements without being distracted or confused by the other data points. This keeps with Shneiderman's philosophy of details on demand in that the user hone in on a single relationship between two variables in the data. The trend line that autopopulates in the display also creates a clear indication for the user what the specific trend is between the selected factor and indicator, even if a very strong correlation is not clearly present. By hovering over the trend line the user can identify important statistical values such as r-squared values to assess the degree of correlation and p-value to assess the confidence interval of the trend. Showing these details on hover keeps the scatter plot uncluttered while giving the expert user the level of detail they would expect. Overall, the scatter plot gives the user an idea of just how much a certain contributing factor affects a specific success indicator, and therefore can assist the user with determining whether or not it is worthwhile to focus attention and efforts on improving a specific factor in order to improve student success in general.

Finally, it was our team's goal for this visualization to be understandable and useful to actual individuals who worked in education in Washington state, and we were pleased with how the results of our usability testing with such individuals proved that our display was just that. The highest compliment that we were able to receive was from our first usability participant, an assistant principal in a school district near Tacoma, who stated that she felt that our visualization was much more effective than OSPI's own representation of the same data on their website.

> "This tool, as...compared to the OSPI website, way more user friendly, this is way more user friendly because everything is pretty much there. When you go to OSPI you have to type in a district and you can't just cross compare." – Participant 1

Our other participants cited the visualization as "pretty slick" and "super easy" to use, and the fact that these educational leaders were not only able to find the display easy to use but also found it genuinely useful was indeed one of the highest marks of success in our eyes.

In the end our team set out to create a tool for Washington state decision makers in education to improve student success, and we strongly believe that our visualization is a tool that can help achieve this goal.

APPENDIX

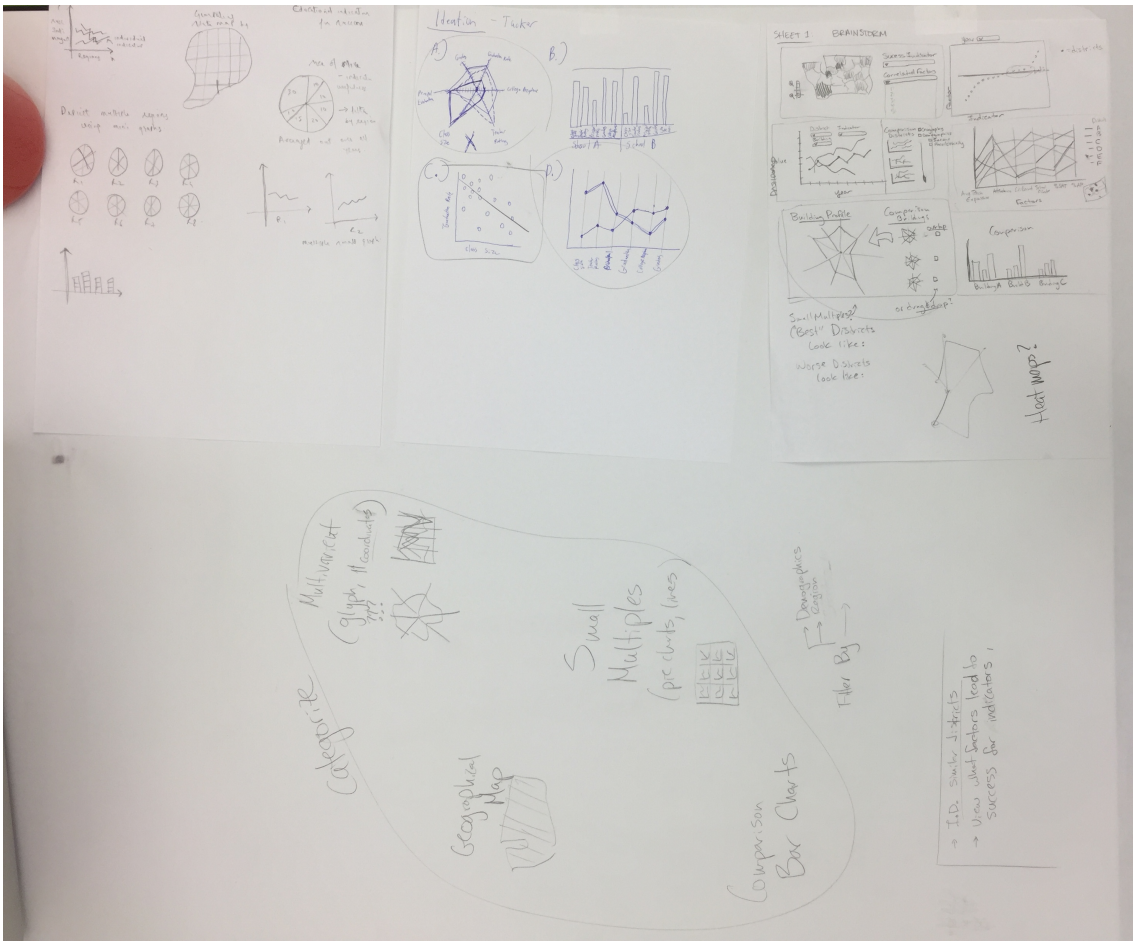

Figure 1:
Design Sheet 1 of the ideation exercise, this includes the individual brainstorms as well as some categorization and sorting of the ideas.

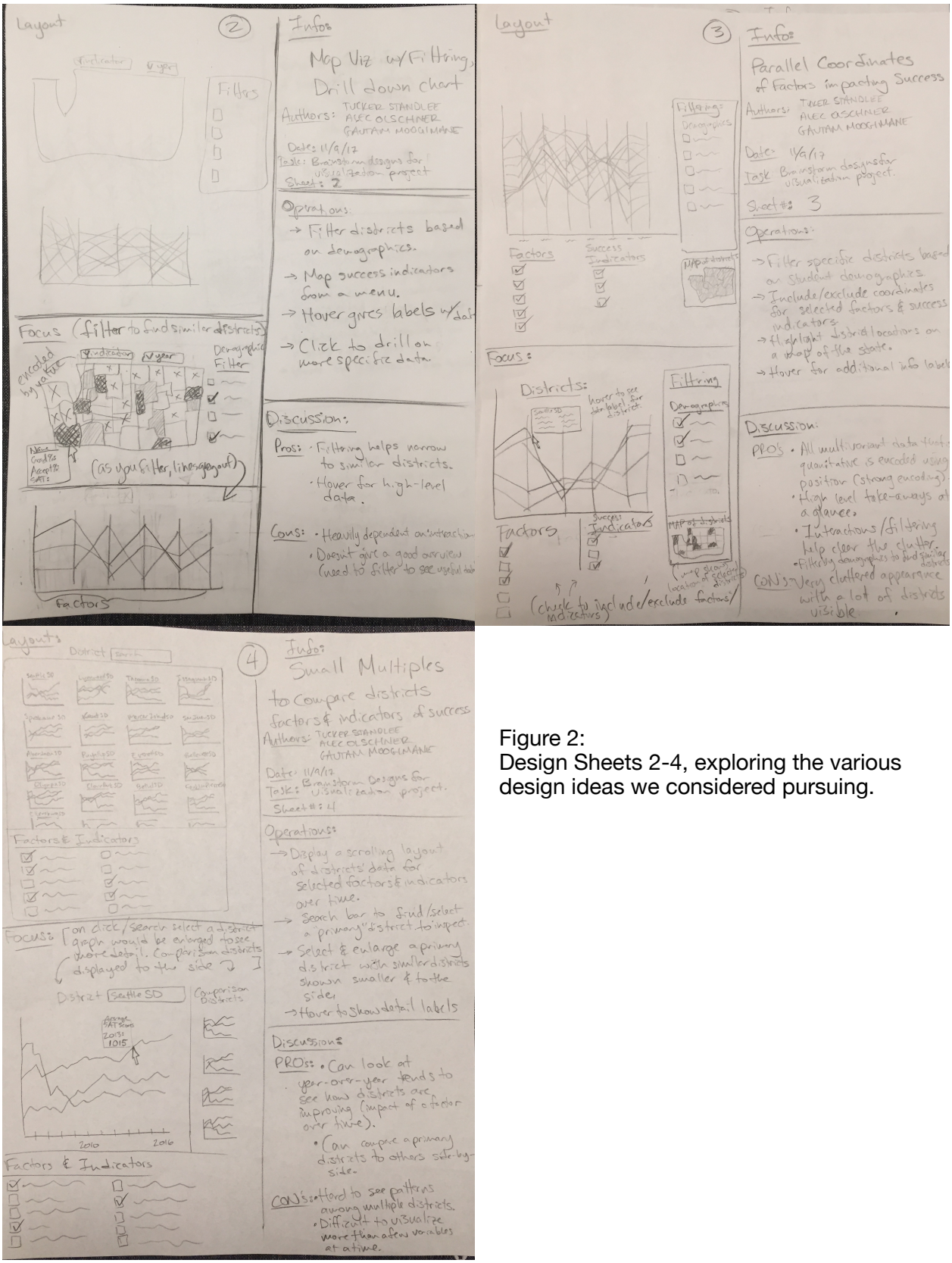

Figure 2:
Design Sheets 2-4, exploring the various design ideas we considered pursuing.

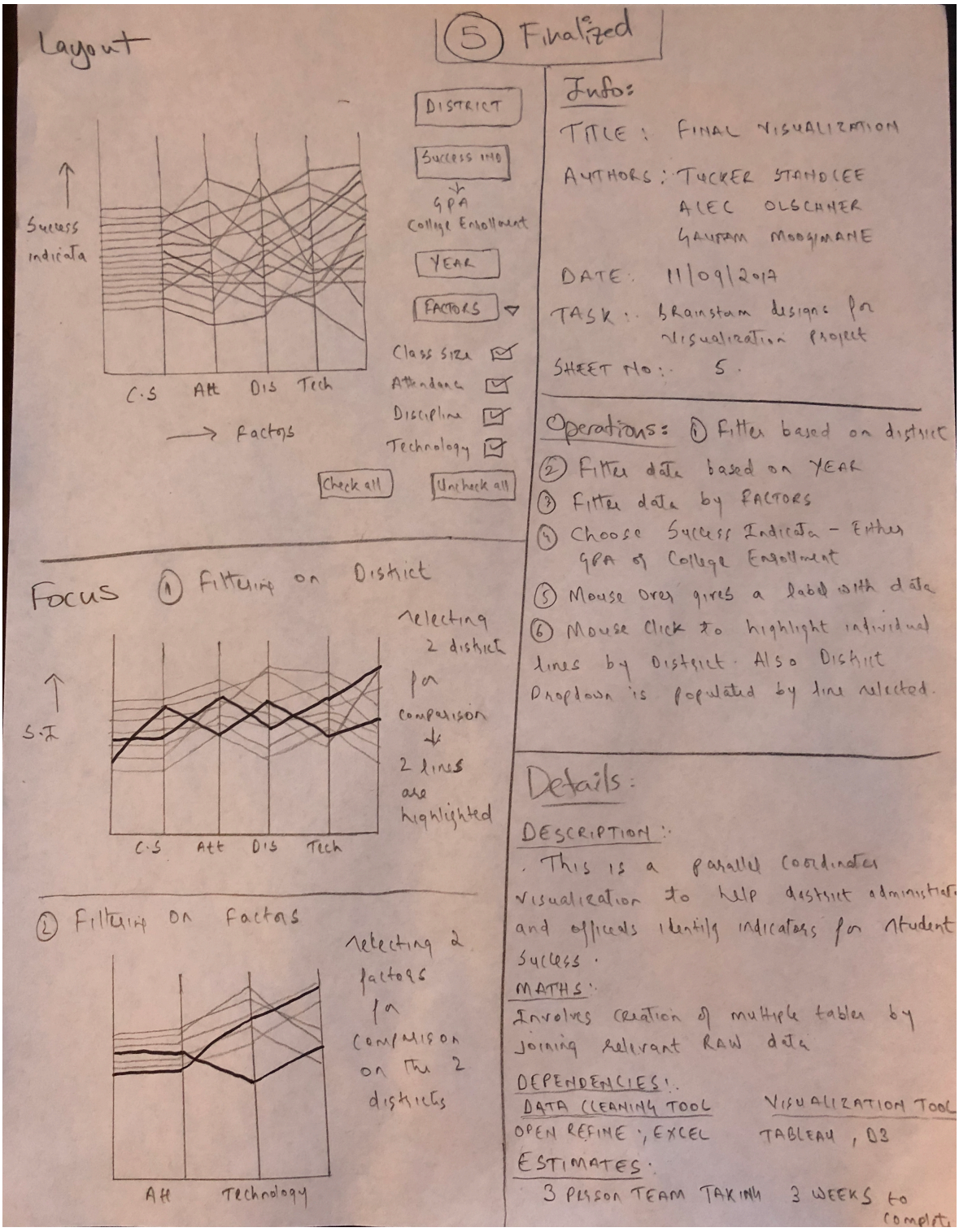

Figure 3:
Design Sheet 5, a more detailed look at the design we decided on.

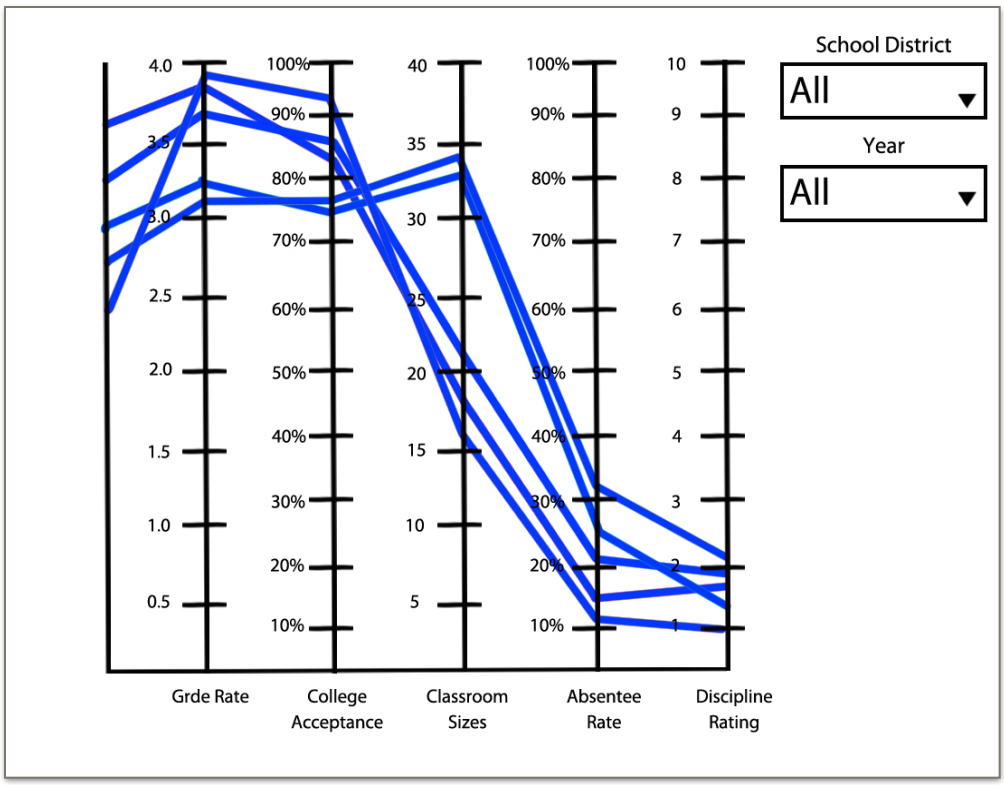

Figure 4: An example of the paper prototype used for pilot usability testing.

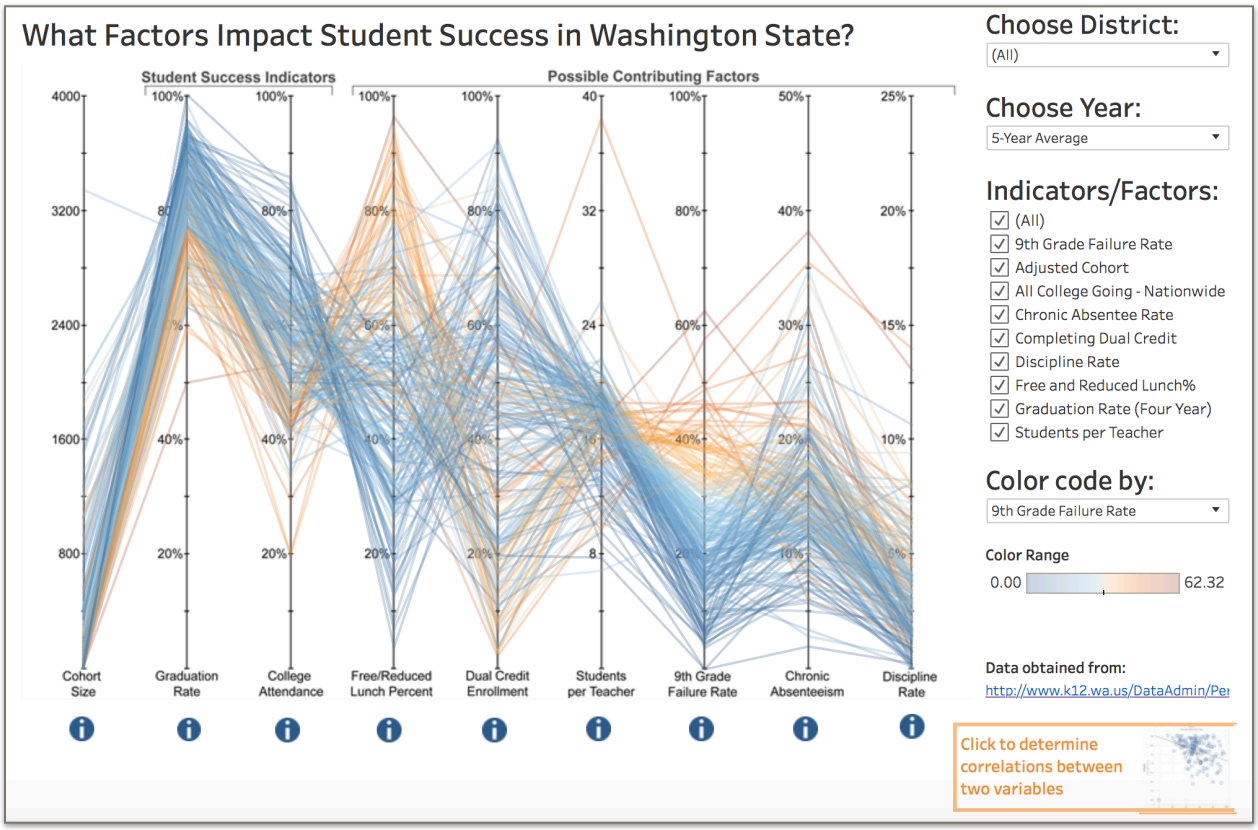

Figure 5: An example of the final visualization, colored based on 9th Grade Failure Rate.