Daodao Zhang
daodaoz@umich.edu

Assignment 2

Problem Description

When applicants research for universities they are interested in applying, they are likely to have several schools or some general criteria in mind. In this assignment, I want to create a simple and clear interactive data visualization to enable applicants compare and evaluate university statistics from different aspects.
The visualization of US university rankings and statistics aims to answer the following questions:
As an applicant who is considering applying for universities, what are the general rankings and other statistics of schools I'm interested?

1. How is one university compares to another in different aspects such as ranking, tuition, acceptance rate and general enrollment?
2. If a user has some criteria in school rankings, tuition or acceptance rate, what are the universities fit such criteria?
3. What are the positions of the universities a user wants to apply among all universities' statistics? Is it above or below average?

Rationale

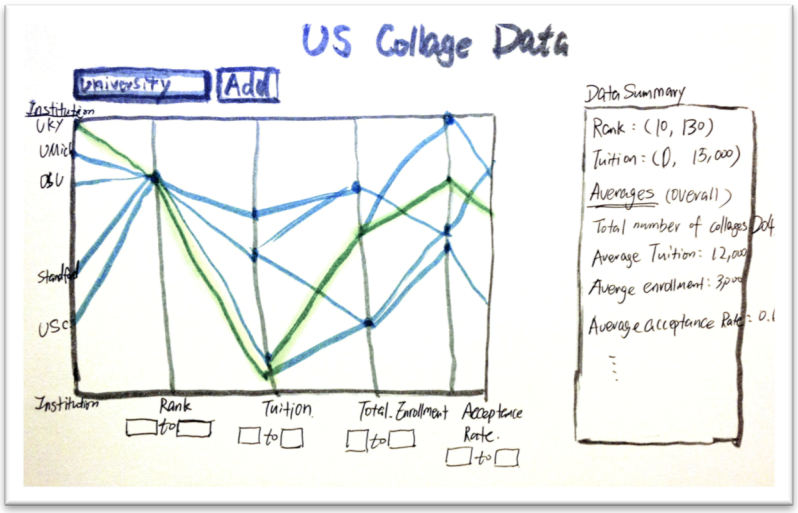
The variables for the university dataset are institution names (nominal), university general rankings (ordinal), tuitions (interval), general enrollment (interval) and acceptance rate (ratio). Even though the variables consist of several different data types, the main purposes of the visualization are universities' absolute positions on each variable, and comparisons of universities of the same variable.
After careful consideration, I've decided to use parallel coordinates as the basic form to present university data, and add interactions to enable users with comparisons and data look up in different dimensions. The reason why I chose parallel coordinates chart is because I believe that users would want to look at variables of universities as a whole instead of comparing data of variables individually. Parallel coordinates chart provides the possibility to display data with different scales and types, and overall compares multiple data entries in an intuitive manner.
A user can look up general statistics and rankings of several selected universities from a parallel coordinates chart, and decides whether a school is generally outstanding, or observes a school has an unusual low score on one aspect. None of bar chart, scatter plot, pie chart or line chart has a clearer vision and advantages to answer such questions like parallel coordinates chart does.
Besides using a chart to visualize university statistics, I would also provide a table

Daodao Zhang
daodaoz@umich.edu

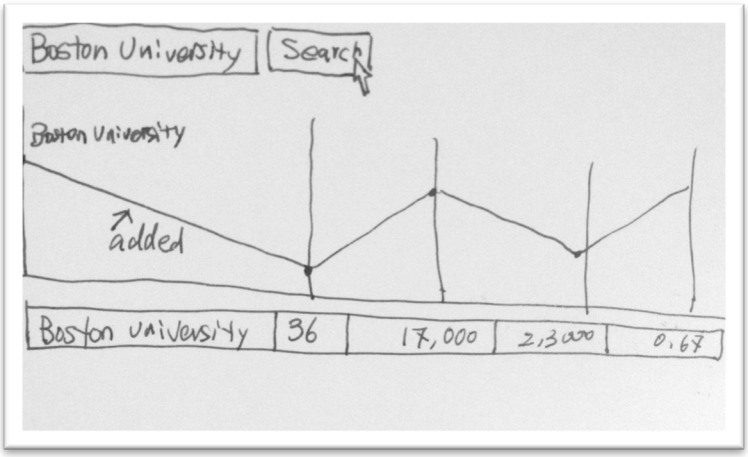
of general summaries and individual data of schools selected. As mentioned earlier, another major purpose of the visualization besides comparisons is to display the absolute position of selected universities. For example, when a user compares two universities, he or she might also like to know what positions these two universities stand among all universities. This can be done either by monitoring all university data in a chart and highlight one specific, or simply have interested university data displayed on the side.

Scenario

A user would first see a visualization of all universities data on five variables, and get a sense of value ranges, distributions and number of data from the parallel coordinate chart. Users are allowed to hover over the chart and view particular

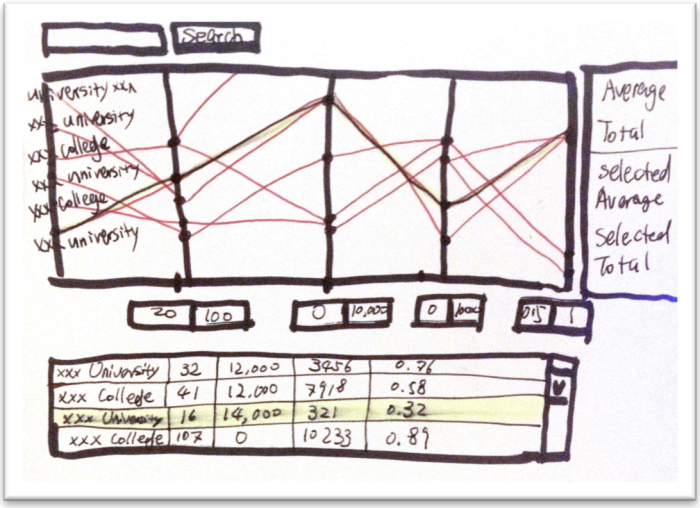


school with that line of school highlighted. They can also add universities to an empty chart by search university names and append lines to the chart, or filter the



Daodao Zhang
daodaoz@umich.edu

lines by variable ranges. (e.g. I want to view university data for all institutions between rank 10 to 100, tuitions < 20,000 USD, acceptance rate > 0.5, etc.).



After universities are selected, users can view statistics, numbers of data they selected, and learn their positions among all universities.

Tools & Data Source

In this assignment I will use D3.js and some additional packages in javascript to implement visualizations and interactions. My primary data resource will be university-ranking data from USnews.