Project Milestone

CS – 5630 / CS – 6630
University of Utah
Team:
Dylan Zwick
dylanzwick@gmail.com
UID: u0075213

Clark Barnett
twentyonetroy@gmail.com
UID: u1119424

As the process book is not a journal, but a writeup, this will be a writeup of the current status of our project at the time of the milestone.

Data

The data we use is from the college scorecard dataset, which is available online:

https://collegescorecard.ed.gov/data/

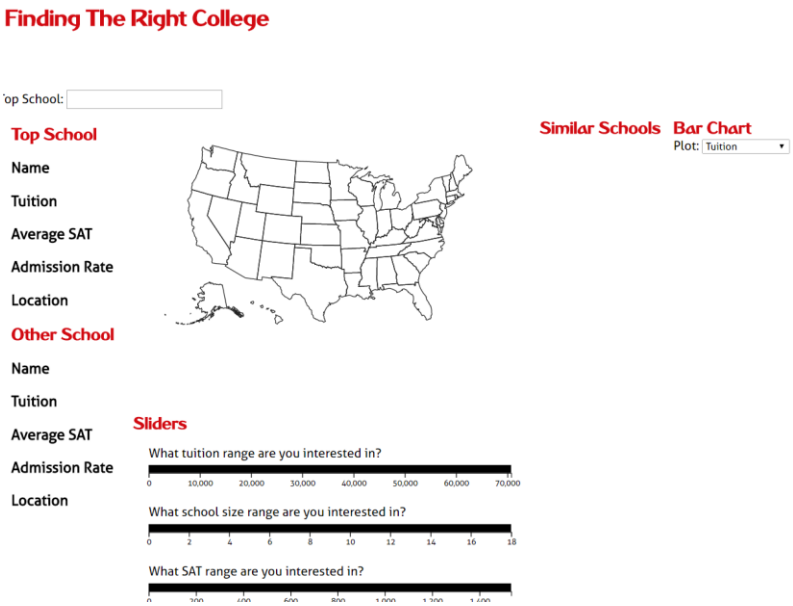
The entire dataset is quite large (246 MB zipped). The data for the visualization is only a much smaller, processed subset of this data. The data used by the visualization are:

- School_Data.csv : A file containing relevant data about each school under consideration.
- similarity_rankings.csv : A file containing, for each school under consideration, the 15 schools most similar to it according to the similarity metric.
- us-states.json – A .json data file providing the information necessary for the US map layout.

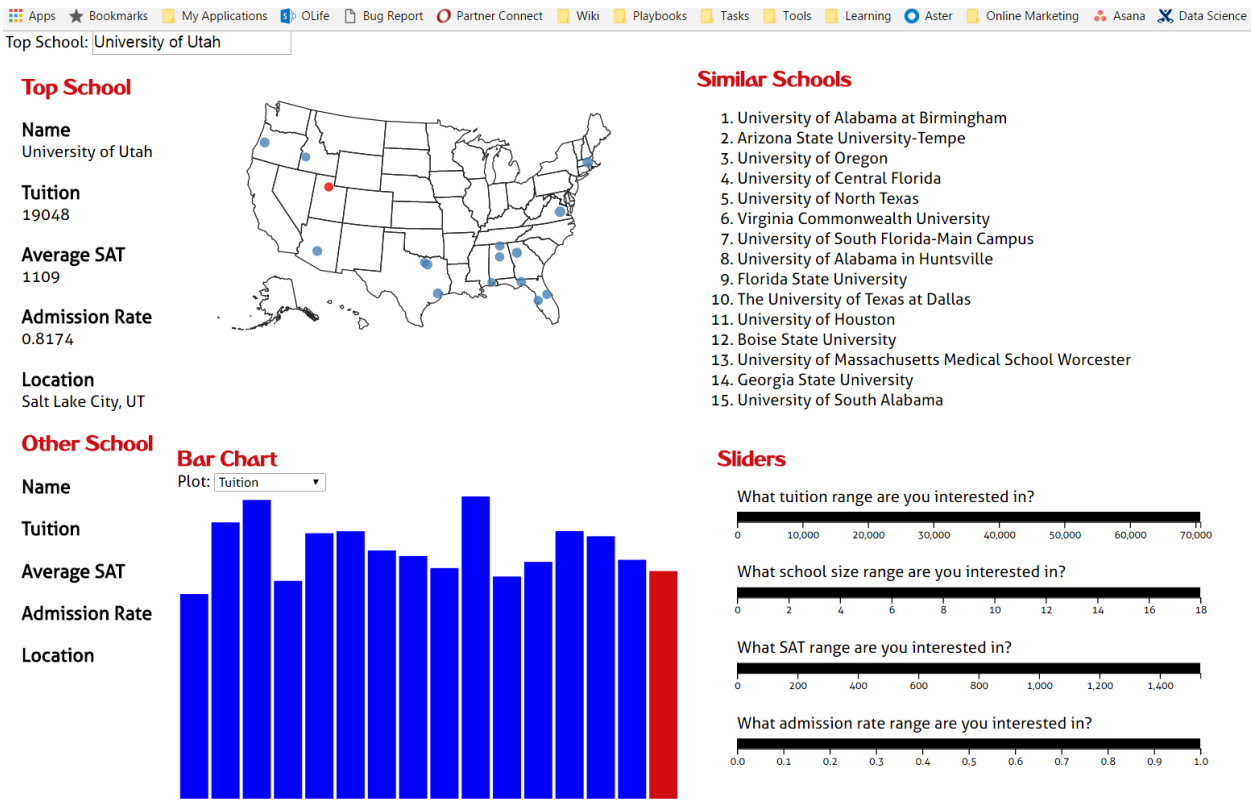
The similarity metric, and how it was derived, will be explained on the final project website, but the details behind it are not important for the visualization. The important thing is just that there is a metric.

Visualization

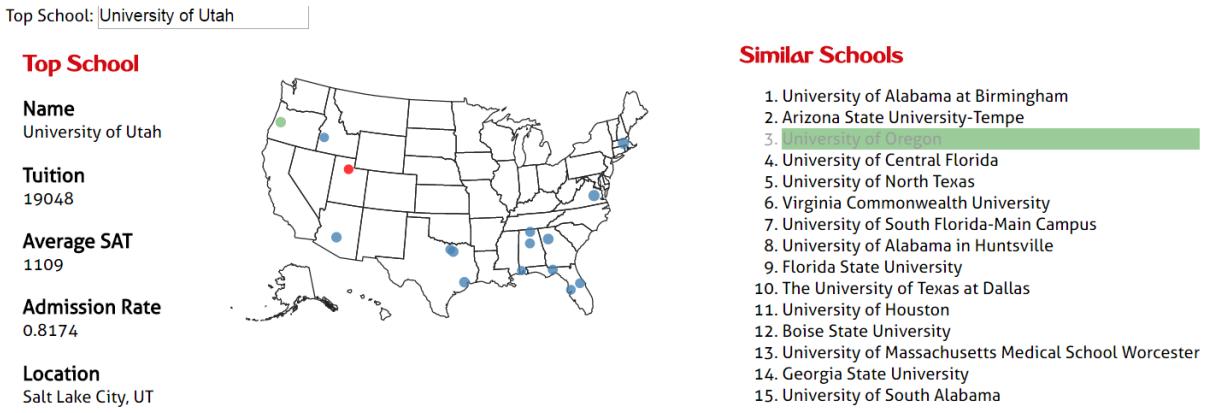
The visualization begins with a screen waiting for data, and a search bar:



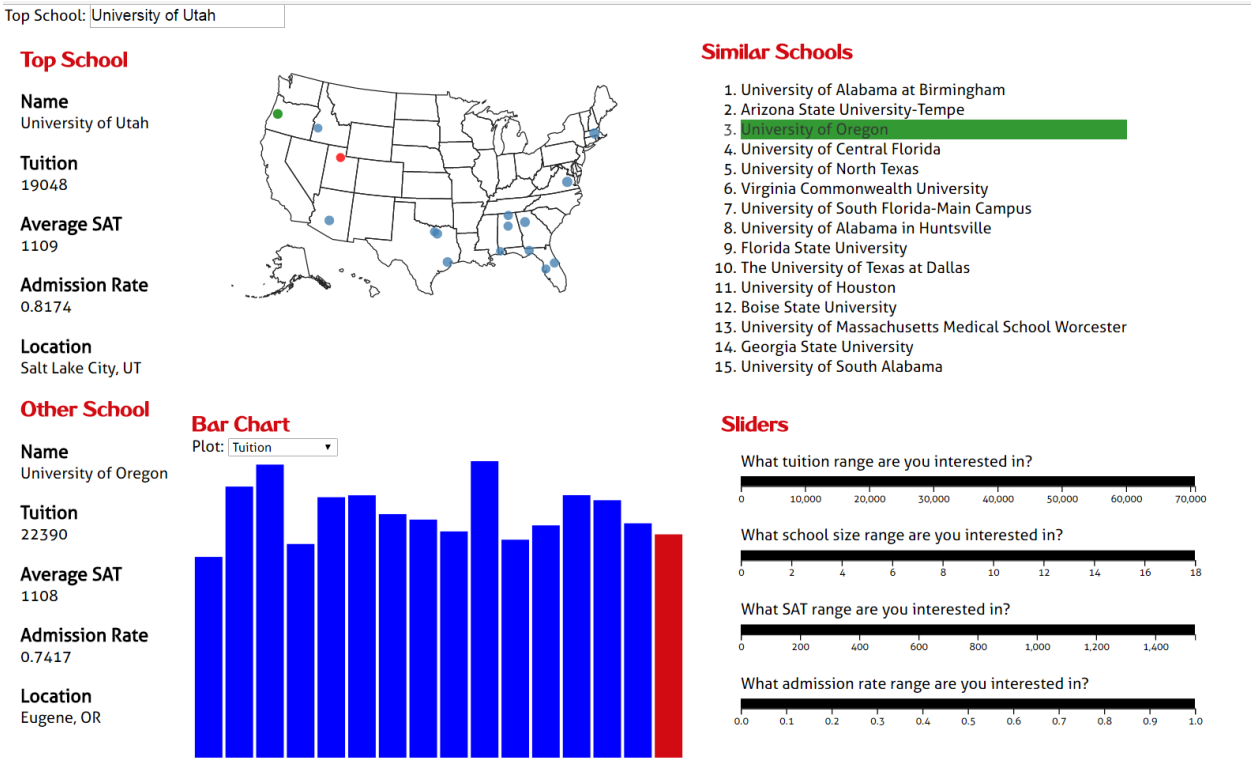
If the user types a top school into the (autocomplete) search bar, the l location of the top school is populated in the map layout with a red circle, while the top 15 similar schools are shown with blue dots on the map, and are also listed in order of similarity:



If the user hovers over a similar school, either on the map or on the list, that school highlights in light green on both:



If the user then selects the school with a mouse click, the school highlights in darker green, and additional information about the similar school is provided:



A bar chart visualization provides data about the similar schools, where the data is selected from a pulldown menu. Sliders can also be used to filter the similar schools under consideration.

Where This Is Going

- **Additional Data on Map Layout** – The user will be able to visualize data about the similar schools on the map layout as well, where the size of the school's circle will represent the give metric, be it tuition, average SAT, etc…
- **General Data Exploration –** The user will be able to explore the data without the context of a favorite school.
- **Visualization Cleanup** – The visualization, in its current state, is obviously a work in progress. It will look much, much nicer as a final product.