Paul Glenn
Shir Nehama
Info 247

D3 Assignment

Final Project Web Link:
http://people.ischool.berkeley.edu/~paul.glenn/i247s16/

Overview

We attempted to incorporate both the parallel coordinates and radar visualizations into a single interface that gave the user the ability to view broad trends across all universities, as well as make a specific comparison between two universities. We specifically implemented this via a side-by-side view of both charts. A pair of drop-down menus allows the user to select specific universities to focus on. Color is used to draw a thread between the two visualizations by highlighting the selected universities in the parallel coordinates visualization in a different color than is used for the rest of universities in that view.

Parallel Coordinates Analysis - Coded by Shir Nehama

Visualization Objective

In this visualization we wanted to present the relationship between the Total University Score with the scores that contribute to this overall score. The idea here is to show that Universities with a high Total University Score will also have high scores in other categories.

For each category, the higher the score on the axis, the "better" that score is. We imagined that the path lines will have a generally horizontal path across the 7 categories since high scoring universities will have other high scoring categories.

From the final outcome of the visualization we saw that there is a positive correlation of the University Score with the Teaching Score, Research Score, and Citation Score. However, this trend was less clear with the International Score, Income Score, and Staff to Student Ratio.

Data Source

Using the Time Data, the parallel coordinates visualization includes the following attributes for each University from 2011.

Total University Score	Teaching Score	Research Score	Citation Score	International Score	Income Score	Staff to Student Ratio

There were a few data manipulations implemented. Firstly, null values in the Income, Staff to Student Ratio, and International Score were set to the average of values of that dimension. We spent a lot of time discussing how best to deal with null values, and other option we considered were setting the

values to zero. Ultimately, we settled on using the average primarily for visual reasons and because we thought that a very low value like zero was potentially more misleading than a calculated value like the average.

Secondly, we inverted the Student to Staff attribute in the raw data to a Staff to Student Ratio. We did this since we need a higher score to indicate a "better" score and a higher staff to student ratio is more desirable than a higher student to staff ratio. This allows the viewer to recognize quickly that higher is better across all axes in the parallel coordinates chart.

Finally, we removed all special characters from the names of the universities. The csv file provided did not correctly represent special characters like accented letters, and commas are not legal in html id's and were thus removed. A consequence of this is that some of the university names are incomplete, and a better solution would have been to process these manually to make a closer approximation of the university name.

Code Modification

For this code, we included a hover feature that not only highlights the path that is being moused over, but all displays the University Name corresponding to the path. This adds valuable qualitative information to help the audience understand which university the data corresponds to. We also added code to highlight those universities selected for comparison in the radar chart. Additionally, when the user selects one of the comparison universities, its line in the parallel coordinates chart is "brought to front" in order to make it easier to target it for mouseover.

Code Contributions

- The original parallel coordinates code was borrowed from Jason Davies open source code: http://bl.ocks.org/jasondavies/1341281

- Some of the coding format for the hover feature was borrowed from the parallel coordinates visualization as seen in the Jeopardy Final Project http://people.ischool.berkeley.edu/~japple/jeopardy/ by Anand Rajagopal, Anubhav Gupta, Joshua Appleman, and Juan Shishido.

Strengths and Weaknesses

This visualization is helpful in seeing the parameters that may contribute to a University's high ranking. It is easy for the reader to be able to see horizontal lines which would correspond to a positively linear relationship between the overall score with the other scores. The parallel line coordinates fits really well with this trend that we seek due to the horizontal and parallel alignment of the axis across the window.

The biggest weakness of this visualization is the fact that there are a lot of paths on the graph. The parallel coordinates is good at showing the overall line trends with the data as a group. However, it is less easy to trace the path of a certain university. Adding the highlight and name hover feature helps alleviate this problem to a certain degree.

Radar Chart Analysis - Coded by Paul Glenn

Visualization Objective

We added the highlight feature to the parallel coordinates chart to make it easier to see the full path of a single university, but it still lacked a way to see and compare the paths of two universities at once. The primary objective of the radar chart was to create this opportunity, and we believe it generally succeeds, although we will acknowledge that one of the last changes we made to the parallel coordinates chart in linking the radar chart and parallel coordinates chart by color does, perhaps redundantly, add that capability to the parallel coordinates chart.

Data Source

The data source is the same as for the parallel coordinates section and has the same caveats and modifications. A close investigation of the csv file will show that much of the data is duplicated and the various charts draw from different columns, however, all of that data was duplicated with differing headers for convenience to merge our code.

Code Modification

The primary modification of this chart over the stock radar chart was the addition of the drop-down menus that modify the chart using d3.dispatch as soon as they are changed. In retrospect, this was far more complicated than we expected it to be, and a button would have been a lot easier. We additionally added tooltips that show the university name when hovering inside the bubble and the specific value of a datapoint when hovering over that point. Finally, modifications were made to the original colors, opacities, etc. Originally, we struggled with whether to order the universities alphabetically or by ranking in the drop down, but ultimately decided to go with ranking, without including that specific number in the dropdown, so that typing the name of the university would function as a search in the dropdown. However, later investigation revealed the jQuery plugin Chosen, which adds a search to a drop-down box, and allowed us to include the rankings in the drop-down as well as the tooltip.

Code Contributions

- The main RadarChart functionality comes from https://github.com/alangrafu/radar-chart-d3 and its demonstration at http://bl.ocks.org/tpreusse/2bc99d74a461b8c0acb1.

- The functionality to move the parallel coordinates lines to the front of the canvas when they are selected on the radar dropdowns was inspired from http://bl.ocks.org/eesur/4e0a69d57d3bfc8a82c2.

- The overall framework of using d3.dispatch to automatically modify the visualizations when the drop-downs are modified is based on https://bl.ocks.org/mbostock/5872848 as is the code that creates and populates the selection boxes.

- The regex that converts spaces in the university names to underscores for use in HTML ids so they can be identified by the dispatch function to modify them when the dropdown changes comes from http://stackoverflow.com/questions/441018/replacing-spaces-with-underscores-in-javascript

- The jQuery plugin that provides the search capabilities of the drop-down menus is Chosen, http://harvesthq.github.io/chosen/, and some guidance on how to get it to trigger events in the d3 dispatch was found at http://stackoverflow.com/questions/8980131/changing-selection-in-a-select-with-the-chosen-plugin as well as http://stackoverflow.com/questions/11279898/binding-event-to-chosen-select

Strengths and Weaknesses

This chart provides a very easy way to compare two universities on specific metrics. A key weakness is that it removes all of the context for that comparison, which can be found in the parallel coordinates chart. It does however, provide a good understanding of the totality of the score. For example, in a comparison of 1. Harvard University and 147. Alexandria University, although Alexandria edges Harvard in two categories, the overwhelming area of Harvard's chart makes it clear why there is such a wide difference in ranking -- something that is less clear in the parallel coordinates chart.

Partner contributions:

As indicated above, the work was split equally, with Shir focusing primarily on the data cleaning and parallel coordinates chart, and Paul focusing on the radar and dropdowns, and wiring the charts together and laying out the final page.