Capstone Project Proposal

Varun Nadgir

August 7, 2017

The College Scorecard

The Problem

Using the data available from The College Scorecard, I would like to explore how US universities compare in terms of admission, affordability, predominant majors, completion, and any other interesting data points. The target of these data are college-seeking students and the family members supporting them financially. By finding a good pick, students can set themselves up for future success by going to schools with a history of being strong in a certain subject, or they can be better prepared for dealing with their student loans after graduating. If the suggestions are helpful, it could be possible to develop a generation of students who major in fields they are comfortable in, are known to have good debt repayment rates, and are able to provide a survivable income.

The Approach

This dataset is broken up into 19 .csv files - one for each academic year from 1996-'97 to 2014-'15, with approximately ~7,500 schools (rows) in each and 1,744 recorded datapoints (columns). To get a better grasp of this dataset, I will need to refer to the data dictionary, so that I can break up the larger, full dataset into smaller, workable chunks. I will begin by adding a column to the 19 original .csv files that marks the data collection year (from the filename). Then I will merge them all into one large .csv (about 2 GB), from which I will build the smaller datasets. For example, I will find columns related to **Location** (state, zip, # of branches), **Admission** (rates, demographics, avg. SAT), **Finances** (cost, aid, debt, repayment), and **Education** (predominant majors/minors, completion rates). I will be able to compare these data across schools and academic years and find potential trends for certain schools or regions

The Result

Ultimately, I would like to provide charts and graphs that can assist in the decision, whether by suggesting schools that fit a provided demographic or by informing students what qualities are important when thinking about their dream school. If possible, it would be great to include a web app that allows the user to enter demographic information (minus personal info like name, address, phone) and have the app give suggestions of schools that are good fits.