**Introduction:**

After researching many potential datasets to analyze, I settled on using the Department of Education’s CollegeScorecard data. The program was initiated on September 12, 2015 with the goal of informing consumers of the costs and value of higher education institutions in the United States. CollegeScorecard gives consumers data ranging from tuition rates to schools demographics to salaries after graduation. Using this data, consumers are able to become more informed to make decisions for their higher education.

**Procedure:**

Data was collected from <https://collegescorecard.ed.gov/data/> and dataset was from 2013 collection.

Population scatter will be created, but sample distribution will also be necessary to do analysis.

Regression test will be done on a sample of 100 public and private universities from the United States. Schools that do not report Admissions Rates or Completion Rates will have to be removed from the sample, and replaced with the next random value.

150 random numbers will be generated, from 1 to 1771 inclusive in order to randomize the selection of

Schools with null values for either admissions rates or completion rates will be removed from the selectable list.

Side-note on analysis: Many for-profit and small schools tend to avoid publishing their completion rates and acceptance rates, and few publish their tuition numbers relative to most public and private schools

From a population of 7804 schools, only 1771 schools reported data on their admission and completion rates. Schools that did not have data reported were removed from the sample. Due to this lack of reporting of completion rates and admissions rates by just under 80% of the listed schools in the U.S, the analysis may have a great deal of bias.

The data took hours to parse as CollegeScorecard listed 7804 schools that reported data to the study, and the datasets were tens of thousands of columns wide.

**Reflection**

A possible further study could analyze admissions rate and completion rate over 6 years.

A corollary study could analyze admissions rate and tuition.

The data shows…..