

## STAT 225 - Data Read In - Group D

Our project is about analyzing the differences in 10 years post graduation compensation among colleges.

### Data Sources

Our only data source is College Scorecard.

### Workflow to Generate your Data Set

We downloaded the data from the following link: <https://collegescorecard.ed.gov/data/>

(Midway down the page on the left hand side)

```
data<-read.csv("Scorecard.csv")  
data2<-data[,c(-6, -18, -19, -(27:34), -(42:89))] #simplifying and only selecting certain columns
```