**Introduction**

The College Scorecard project is designed to increase the transparency of the college data, working as a navigator for students to make informed decisions about their choice for higher education. The information gathered in the dataset also gives students and families more power to differentiate postsecondary institutions and get an idea about how well colleges are preparing their students to be successful. The dataset is created to address the lack of information about college costs, earnings, federal student aid, admission rate etc. The data provides insights into the performance of institutions that receive federal aid dollars, and the outcomes of the students of those institutions. It also gives extra support for students and families to weight the tradeoffs on choosing the right college according to their own needs and educational goals.

These data are collected by United States Department of Education with the goal of helping students and families make more educated decisions about college and compare vital information easily. The data elements are associated with those currently operating institutions whose predominant awards are:

* Associate degrees
* Bachelor’s degrees
* Certificate degrees
* Graduate degrees

Data elements with an ‘IPEDS’ value in the ‘Source’ field are derived directly from the Integrated Postsecondary Education Data System (IPEDS). In our team research, we mainly focus on the MERGED2012\_13\_PP.csv data file which is related to the data elements collected during the IPEDS 2012-13 data collection year.

**The following variables are used for analysis:**

* NPT4\_COMBINE
* DEP\_STAT\_PCT\_IND
* PAR\_ED\_PCT\_1STGEN
* MD\_FAMINC
* PCTPELL
* PCTFLOAN
* GRAD\_DEBT\_MDN
* WDRAW\_DEBT\_MDN
* APPL\_SCH\_PCT\_GE2
* MD\_EARN\_WNE\_P6
* MD\_EARN\_WNE\_P10
* RPY\_3YR\_RT\_SUPP
* CDR3

**Data cleaning Method:**

* Variables having more than 20% missing value are excluded from the dataset

**Reference(MLA):**

"Data Documentation for College Scorecard (Version: January 2017)." *United States Department of Education* (2017): 1+. *CollegeScorecard\_Raw\_Data*. Web. 13 Feb. 2017. <https://collegescorecard.ed.gov/assets/FullDataDocumentation.pdf>.