

U.S College Scorecards Report

This is a report for subset of the U.S. Department of Education's College Scorecard Data. The data set contains information for different colleges within the U.S.

Data Exploration and Dimension Analysis

Estimating Number of Distinct Colleges

There are 3,312 observations recorded in the dataset, out of which 2,431 represent a distinct colleges or college system.

```
## [1] 3312
```

```
## [1] 2431
```

Analysis of the Features

There are **51 features** describing each of the observation.

There are **9 factor variables** ("unit_id", "ope_id", "city, zip", and "branches"*) have been converted to factors, in addition, "online_only", and "main_campus" can be viewed as factors as well.)

There are **17 discrete** variables (excluding logical and factors,) and **21 continuous variables**.

Dataset contains **3 logical** variables [which can be viewed as factor variables with 3 levels: "True", "False", and "NA"] and **1 character variable** for name (which can also be converted into factor with unique levels if desired.)

*branches can be viewed as a discrete numeric or a factor variable.

```
## [1] 51
```

```
## [1] 9
```

Exploring the Missing Values

There are 23,197 missing values in the dataset, with avg_sat [the average SAT score] having the most missing values. Patterns: 'linked' variables like average and median family income tend to be missing for the same schools, as well as variables that come from the same source like pell grant and federal loan. Some schools are missing all demographics data.

```
## [1] 23197
```

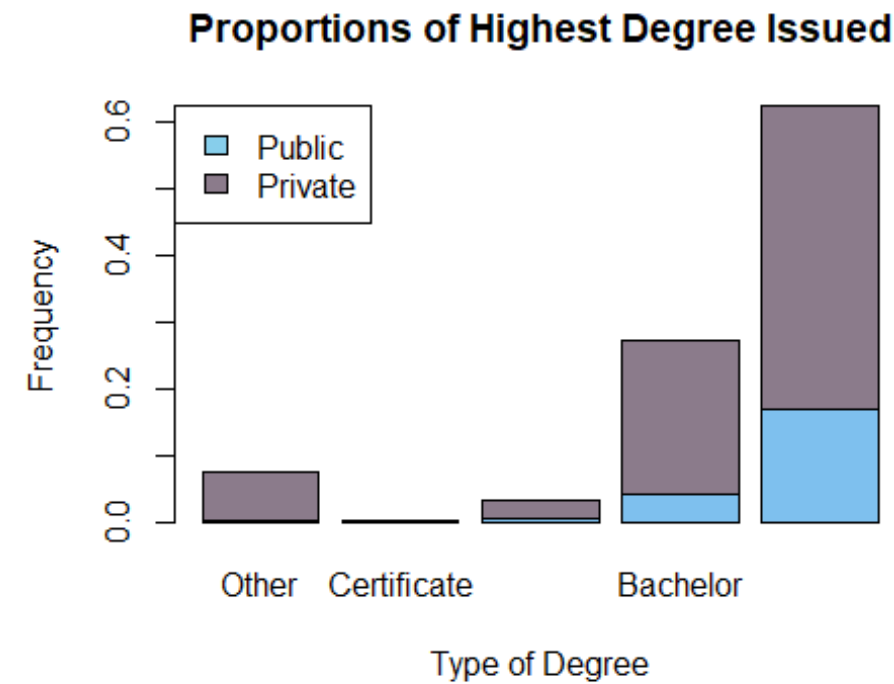
```
## [1] "avg_sat"
```

Estimating Number of Highest Degrees Issued by Private vs Public Schools

There are more **private (2596)** than **public (716)** colleges recorded.

We can see that private schools issue proportionally more of each degree type. In both categories the majority of schools offer a Graduate degree as their highest (approx. 60%), with Bachelors being the second highest (approx. %30)

##



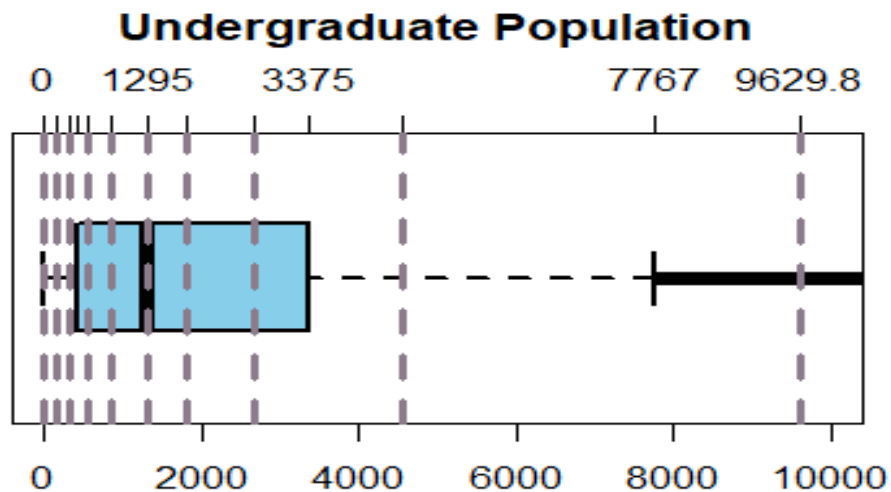
##	Public	Private
##	716	2596

Estimating Student Population

We see that over 40% of schools have student population below 1,000 students. Most of the schools [>90%] have a population of less than 10,000. The maximum student population [166,816] is found at "University of Phoenix-Online Campus."

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
##	0	428	1295	3600	3372	166816	490
##	0%	25%	50%	75%	100%		
##	0	428	1295	3372	166816		

```
## [1] "University of Phoenix-Online Campus"
```

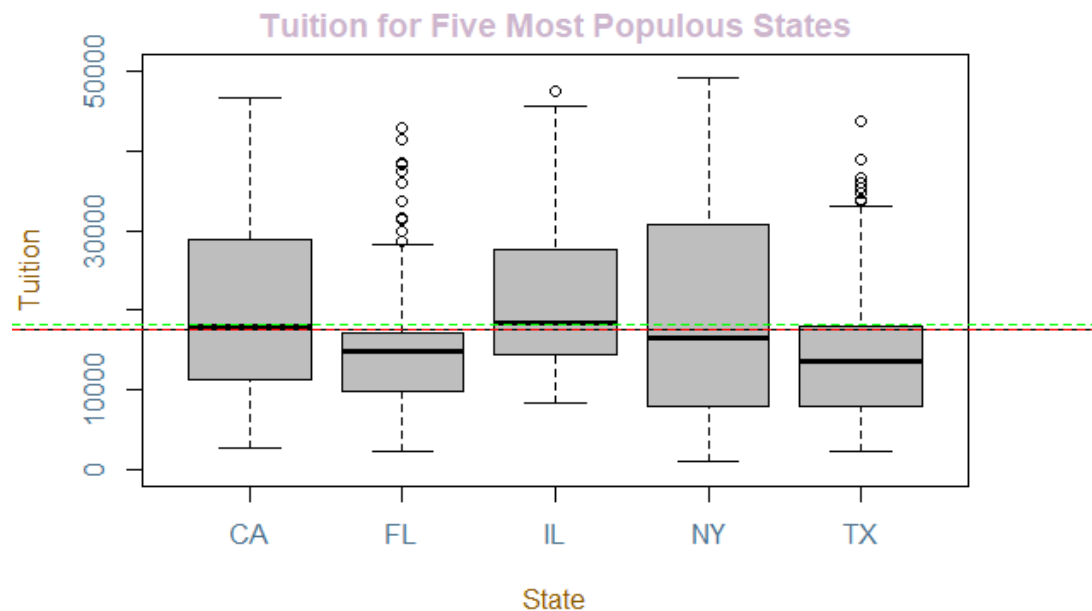


Estimating the Average Cost of Tuition for the Five Most Populous States

The five most populous states are: California, Texas, Florida, New York, and Illinois.

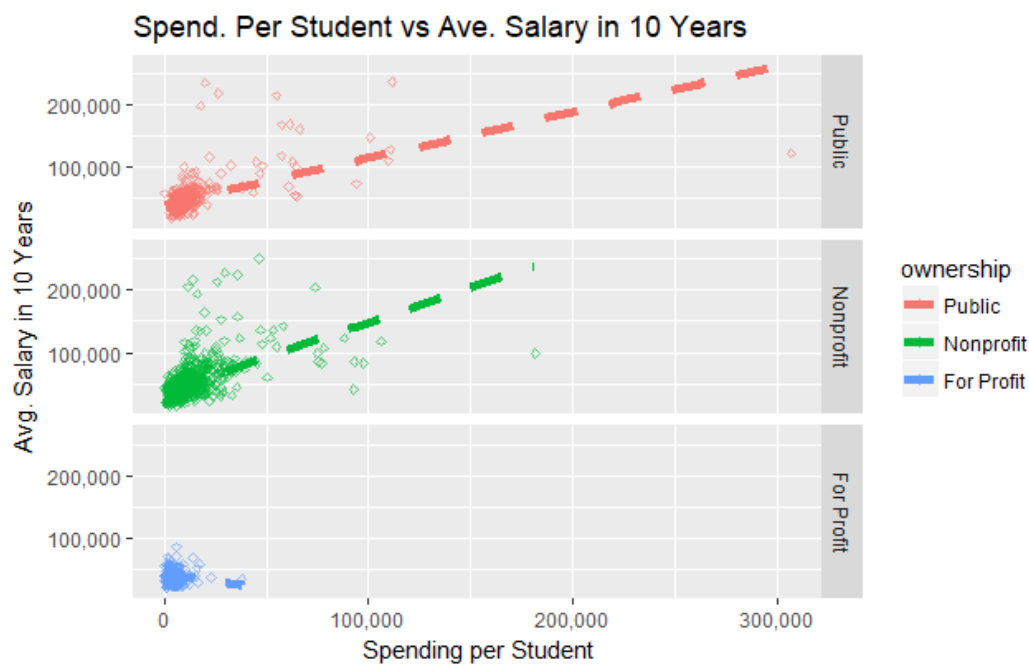
https://simple.wikipedia.org/wiki/List_of_U.S._states_by_population

We can see that the highest tuition is found in Illinois and it's slightly above the overall average (the red dotted line,) with California being a close second. Texas has the lowest tuition among the five states. The average tuition among the five most populous states (the green dotted line) is slightly higher than the average tuition overall, both are around \$17,500, but not significantly.



Exploring Relationship between Spending per Student and the Average Salary in 10 Years for Public, Private, and Non-Profit schools

There seem to be a positive correlation between spending per student and average salary in 10 years for Public and Nonprofit schools. However, in For Profit schools that relationship seems to be reversed with a small negative correlation. That means the higher spending per student by Private Schools do not necessarily improve student earnings 10 years after starting school. Overall, Nonprofit schools indicate higher 10-year average earnings per dollar spend for each student.



Top 10 Schools with the Highest Earnings per Cost of Tuition

The main indicator of student earnings in our data set is the average salary in 10 years after school (we can also use median, which is less affected by outliers, but in our case these two values are very close. Therefore, we can use a ratio of the average 10-year salary per estimated total cost without financial aid as our metric. We could also use a ratio of the estimated total cost w/o fin. aid per lifelong [or careerlong] earnings (if such data was available) to see which schools provide the highest payoff over a students lifetime. Top ten are shown in the table below. On average, the earnings 10 years after school are 1.5-1.7 times higher than the annual cost of tuition. The Top 10 Schools far exceed that level.

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
##	0.460	1.180	1.520	1.704	2.160	13.990	886

Rank	School Name	State	Earnings to Cost Ratio
1	United States Merchant Marine Academy	NY	13.99
2	Augusta University	GA	8.20
3	South Texas College	TX	5.33
4	University of Connecticut-Avery Point	CT	4.79
5	University of Connecticut-Stamford	CT	4.76
6	University of Connecticut-Tri-Campus	CT	4.73
7	Indian River State College	FL	4.71
8	Palm Beach State College	FL	4.61
9	Virginia College-Columbus	GA	4.50
10	Brigham Young University-Idaho	ID	4.47

Top 10 Most Racially Diverse Schools

To determine the most racially diverse school I decided to use the variance of the race-representing variables for each school (i.e. to select schools for which the variance among the percentages of each race is the lowest.) The underlying assumption is that the more evenly the racial groups are represented in a particular school (w/o being heavily dominated by a single race or two races) the more racially diverse it is. Limitations: 1) we don't know exactly what information is stored in the race_other variable, 2) there are 490 schools for which no racial demographics are available. Top ten racially diverse schools in the country are shown in the table below. Five of the top ten, including the most racially diverse school, are in California (4 of them in Bay Area.)

Rank	School Name	State
1	Holy Names University	CA
2	Pacific Rim Christian University	HI
3	California State University-East Bay	CA
4	Golden Gate University-San Francisco	CA
5	The University of Texas MD Anderson Cancer Center	TX
6	ITT Technical Institute-Concord	CA
7	Andrews University	MI
8	Pacific Union College	CA
9	University of Hawaii Maui College	HI
10	Rutgers University-Newark	NJ

Comparing UC Davis with Other Schools

The three metrics I would like to look at when comparing UC Davis and other schools are: average 10-year salary, cost of attendance, and spending per student. As we can see from the plot below, UC Davis has a higher than the mean across all schools cost of attendance equal to 31,803 US dollars, which lies in the 3rd quantile. However, it has higher than the mean average 10-year salary of 65,200 US dollars, which lies in the 4th quantile. Also its spending per student of 21,365 US dollars is significantly higher than of the most schools in the country and lies within highest 10%.

UC Davis vs Other Schools

