Economic Factors for Choosing a Major

Stephen O'Dea

4/30/2021

Introduction

Ellucian, a company that creates software and services to power higher education, published a survey about the career path of students. In the survey, it is shown that students are not confident in choosing their major, and that approximately two-thirds of them found the process too overwhelming ("Course Correction: Helping Students Find and Follow a Path to Success" 2019).

The inability to choose the right major may have serious ramifications. It may cause students to switch career paths without fully understanding the implications. It may also cause them to take irrelevant course work, leading to a unnecessarily prolonged and expensive college experience. Most importantly, one may be pidgen-holed into a career they are unhappy with.

It is to the benefit of all college prospects to be able to make informed decisions about their choice of college and what type of career to pursue. This paper aims to provide some objectivity towards the subject. The data analyzed is from a 2010 survey from Payscale Inc. The dataset shows the 10th, 25th, 50th, 75th, 90th, and entry level salary for 50 different majors based on school, school type, and location (split into 5 regions)

The data was extracted from Kaggle. For more information on the original source of the data, variable dictionary, and an overview of the variables, visit the following link: https://www.kaggle.com/wsj/college-salaries.

Data Exploration and Summary Statistics

There is a total of 5112 total observations between the three dataframes.

Table 1: Salaries Aggregated by Major

Starting Median Salary	Mid-Career Median Salary	Percent change from Starting to Mid-Career Salary	Mid-Career 10th Percentile Salary	Mid-Career 25th Percentile Salary	Mid-Career 75th Percentile Salary	Mid-Career 90th Percentile Salary
Min. :34000	Min. : 52000	Min. : 23.40	Min. :26700	Min. :36500	Min. : 70500	Min. : 96400
1st Qu.:37050	1st Qu.: 60825	1st Qu.: 59.12	1st Qu.:34825	1st Qu.:44975	1st Qu.: 83275	1st Qu.:124250
Median :40850	Median : 72000	Median : 67.80	Median :39400	Median :52450	Median : 99400	Median :145500
Mean :44310	Mean : 74786	Mean : 69.27	Mean :43408	Mean :55988	Mean :102138	Mean :142766
3rd Qu.:49875	3rd Qu.: 88750	3rd Qu.: 82.42	3rd Qu.:49850	3rd Qu.:63700	3rd Qu.:118750	3rd Qu.:161750
Max. :74300	Max. :107000	Max. :103.50	Max. :71900	Max. :87300	Max. :145000	Max. :210000

Table 2: Salaries Aggregated by Type

Starting Median Salary	Mid-Career Median Salary	Mid-Career 10th Percentile Salary	Mid-Career 25th Percentile Salary	Mid-Career 75th Percentile Salary	Mid-Career 90th Percentile Salary
Min. :34800	Min. : 43900	Min. :22600	Min. : 31800	Min. : 60900	Min. : 87600
1st Qu.:42000	1st Qu.: 74000	1st Qu.:39000	1st Qu.: 53200	1st Qu.:100000	1st Qu.:136000
Median :44700	Median : 81600	Median :43100	Median : 58400	Median :113000	Median :153000
Mean :46068	Mean : 83932	Mean :44251	Mean : 60373	Mean :116275	Mean :157706
3rd Qu.:48300	3rd Qu.: 92200	3rd Qu.:47400	3rd Qu.: 65100	3rd Qu.:126000	3rd Qu.:170500
Max. :75500	Max. :134000	Max. :80000	Max. :104000	Max. :234000	Max. :326000
NA	NA	NA's :38	NA	NA	NA's :38

Table 3: Salaries Aggregated by Region

Starting Median Salary	Mid-Career Median Salary	Mid-Career 10th Percentile Salary	Mid-Career 25th Percentile Salary	Mid-Career 75th Percentile Salary	Mid-Career 90th Percentile Salary
Min. :34500	Min. : 43900	Min. :25600	Min. : 31800	Min. : 60900	Min. : 85700
1st Qu.:42000	1st Qu.: 73725	1st Qu.:39500	1st Qu.: 53100	1st Qu.: 99825	1st Qu.:136000
Median :45100	Median : 82700	Median :43700	Median : 59400	Median :113000	Median :154000
Mean :46253	Mean : 83934	Mean :45253	Mean : 60614	Mean :116497	Mean :160442
3rd Qu.:48900	3rd Qu.: 93250	3rd Qu.:48900	3rd Qu.: 66025	3rd Qu.:129000	3rd Qu.:178000
Max. :75500	Max. :134000	Max. :80000	Max. :104000	Max. :234000	Max. :326000
NA	NA	NA's :47	NA	NA	NA's :47

.	Freq
Engineering	0.0706320
Ivy League	0.0297398
Liberal Arts	0.1747212
Party	0.0743494
State	0.6505576

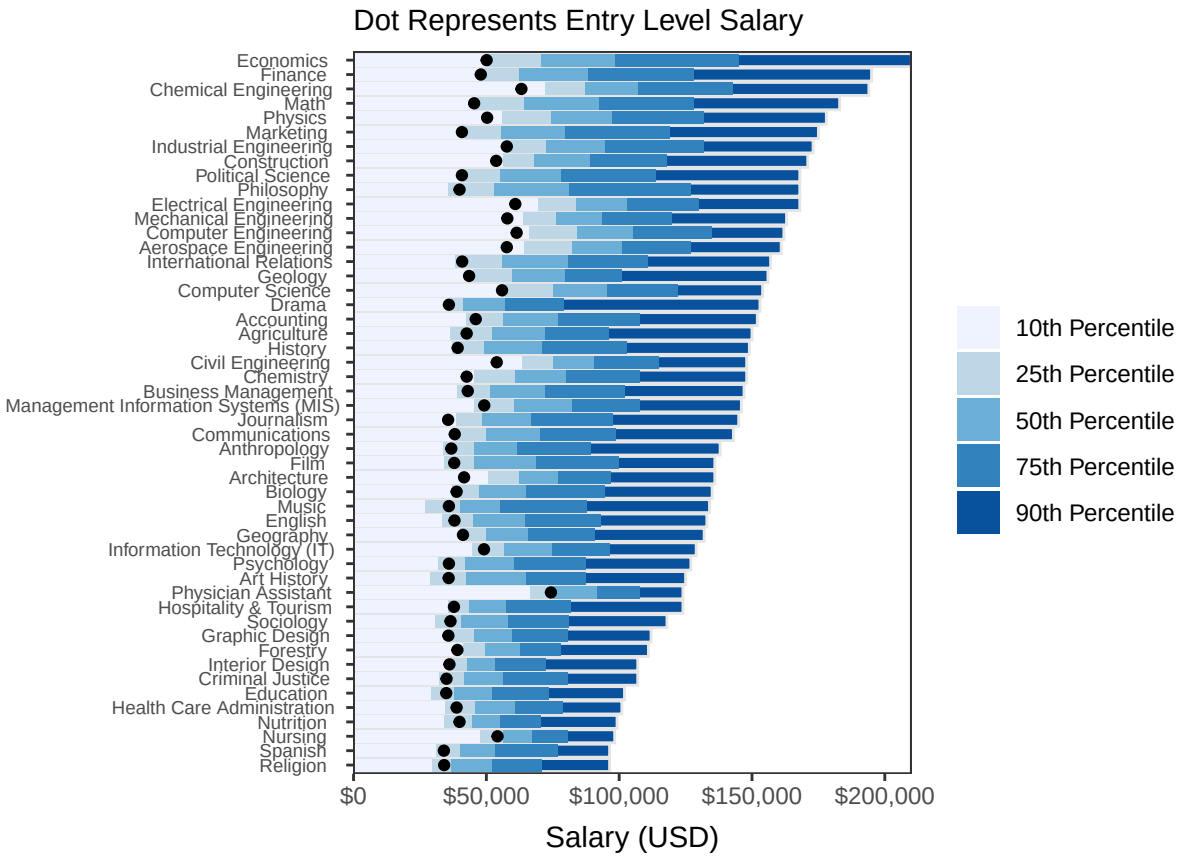
.	Freq
California	0.087500
Midwestern	0.221875
Northeastern	0.312500

.	Freq
Southern	0.246875
Western	0.131250

As expected, there exists a large range of salaries. Furthermore, note the skewness in the distribution of school types. State schools account for more than 65% of the observations. Lastly, it is interesting that California was given its own region. Though it is the least represented , bare in mind that it is the culmination of only one state.

Visualizations

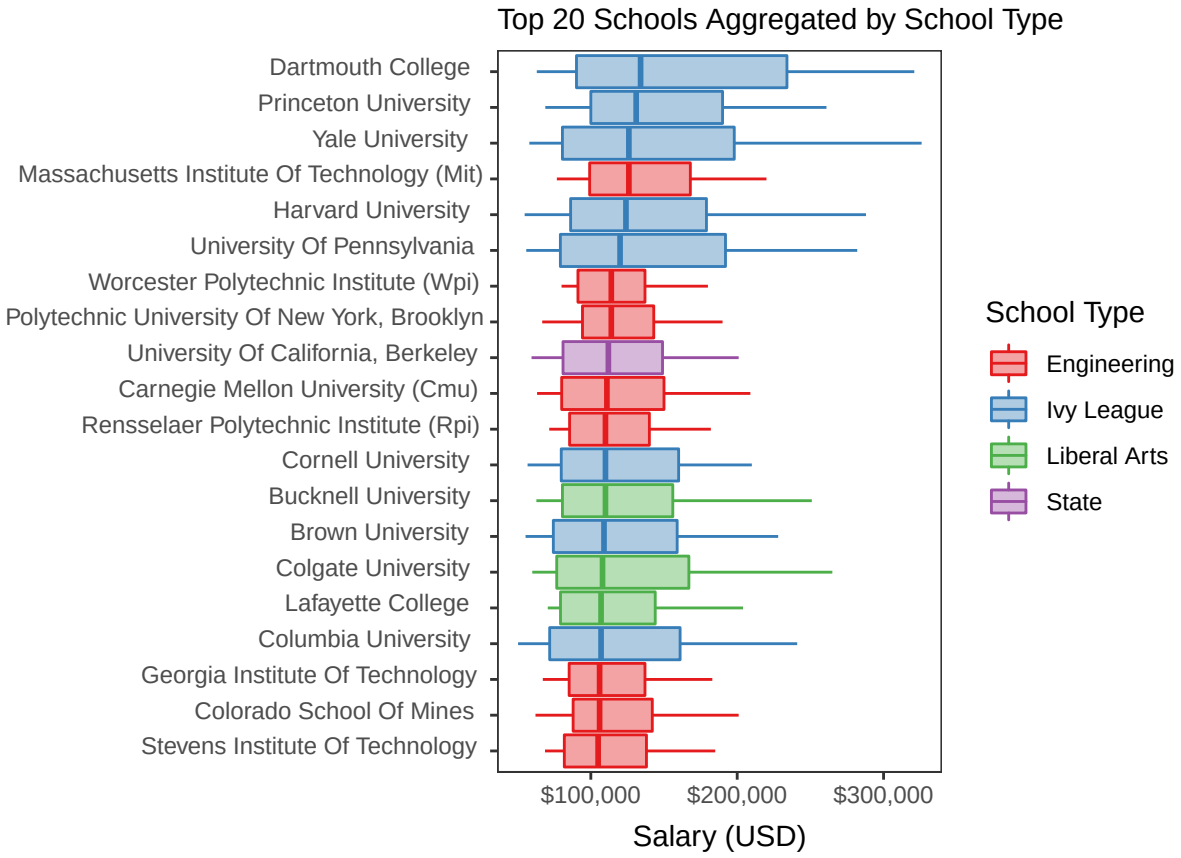
Degrees That Pay Back



There are some interesting things to point out about this data. First, the best degrees in terms of mid-career salaries goes to most of the STEM majors (predominately the types of engineering) as well as economics and finance. Furthermore, these careers exhibited some of the highest pay increases.

In terms of fresh out of school, the degree with the highest entry level salary is Physician Assistant. This job starts at $73,000, but only has an earning power of roughly 23%. A similar circumstance can be seen for Nursing, which ranks low when aggregated by 90th percentile salaries.

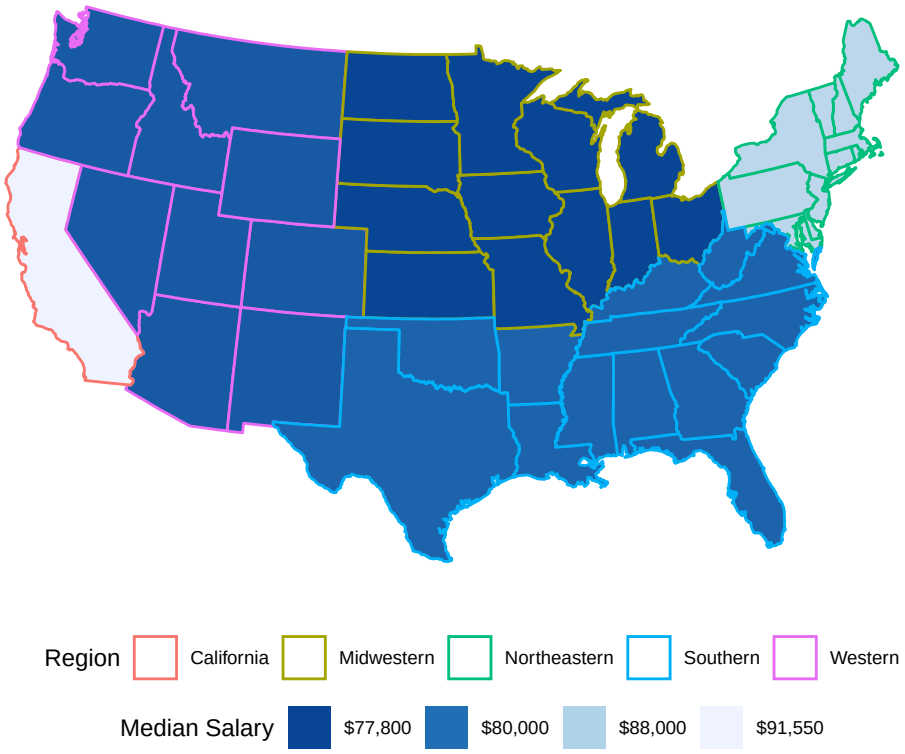
Where It Pays to Attend School



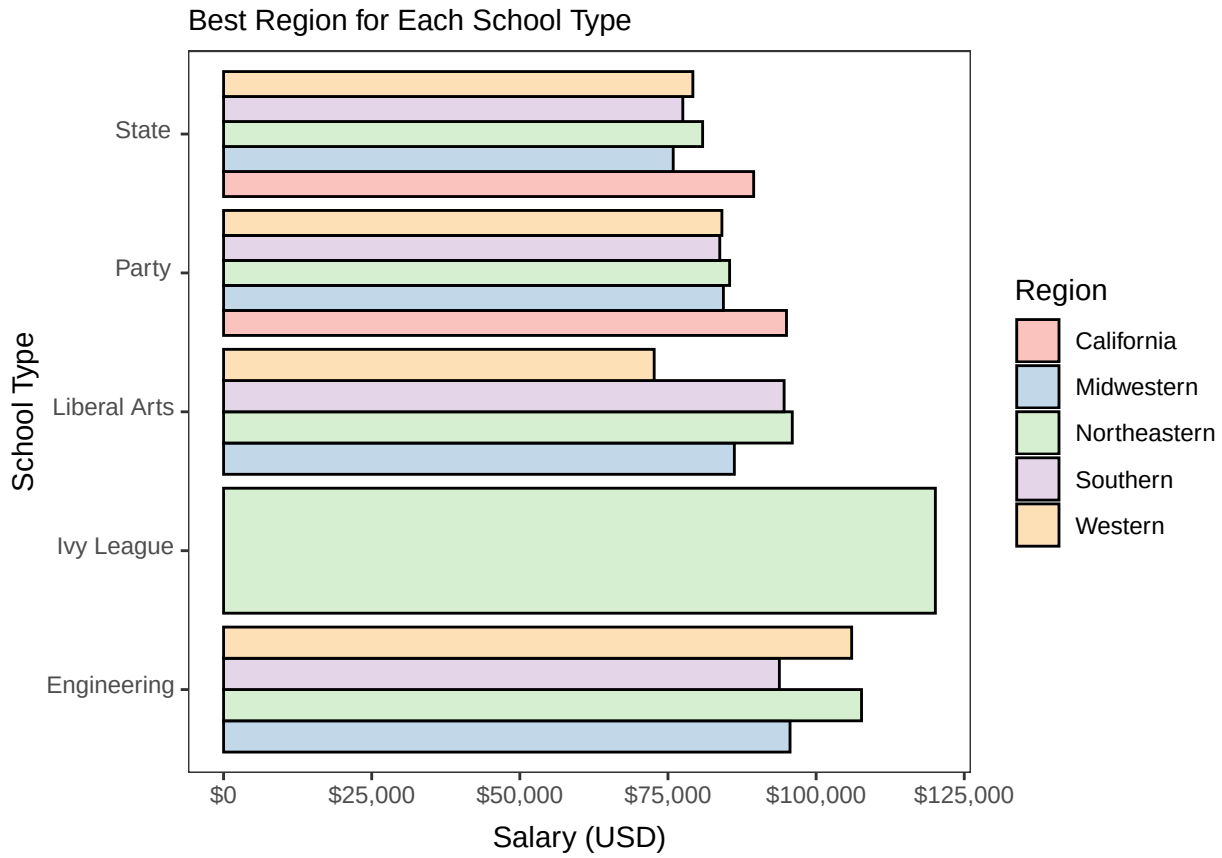
As shown, ivy league schools exhibit the highest salaries with many engineering schools in the top 20 as well. However, consider the distribution of these two school types. The interquartile range for the engineering colleges is much smaller than the interquartile range for the ivy league schools. This is likely caused from the types of degrees offered at these schools.

Furthermore, since 83% of the observations in the data are comprised of state and liberal arts colleges, it is quite interesting to see only 4 observations from these groups make the top 20. Lastly, and not as surprising, no party schools are represented in the graph.

Best Places to Live



The regions that have the highest median salaries are California and the northeast. However, the results of the choropleth can be improved upon. Not all college prospects are willing to relocate or are able to attend any type of school. By joining the salary_by_region and salary_by_type datasets, the same analysis can be performed while aggregating by school type. This will allow students to visualize the best type of school for their respective region.



The best region for each school type is listed below.

Table 6: Summary of Results

School Type	Region	Average Salary
Engineering	Northeastern	107662.5
Ivy League	Northeastern	120125.0
Liberal Arts	Northeastern	95975.0
Party	California	95000.0
State	California	89460.0

A Note on Interpretation

It is important to realize that the information from the data is not indicative of current salaries since this survey published in 2010. Though it is likely that the rank of the majors and the schools would not drastically, it would be interesting to investigate the impact the pandemic has had on some of these figures.

An important variable when choosing a major is the availability of work. Such parameters were not included in the data.

Furthermore, there exists some skepticism when examining salaries by region. Cost of living (COL) is calculated by the amount of money needed to reside in a particular city. Places such as New York and California have much higher costs of living than, for example, states in the midwest. Since COL is a factor when negotiating salaries, there is a chance that results from the previous analysis are inflated.

Lastly, a more cohesive rubric for picking majors and colleges could be constructed if a similar study were to

be published in such a way that the salary information for majors was integrated into each of the individual schools. This would allow for a more exhaustive interpretation of the graphs above.

Future Work / Conclusion

In conclusion, all three of the variables considered (school type, region, and major) proved to be significant features for determining salary. These conclusions were based off of exploratory analysis on the distributions of the variables. To summarize each finding:

- If a student were looking to make the most money out of college, they should pursue a career as a physician's assistant, nursing, or any of the engineering majors (they also exhibit some of the best mid-career salaries).
- When considering a long-term approach, the best areas of study go to math and philosophy, which show a percentage increase of over 100% from their entry-level to mid-career salaries. For highest potential salary, consider a career in finance or economics.

- If possible, college prospects should attend ivy league or engineering colleges. If not possible, state and liberal arts schools are also a wise choice given the major is as well. The worst performing colleges are party schools. If a student were to attend one of those, the best region to do so would be in California.
- Schools in the northeast and California outperform schools in the other regions.

For future work, it would be beneficial to construct an ANOVA test to determine if median salaries differ from majors, college types, and regions. In terms of machine learning, unsupervised learning techniques could be implemented to understand and discover patterns using association rules.

Source code for the document may be found on github at https://github.com/stephenodea54/Degrees-That-Pay-You-Back.

Bibliography

"Course Correction: Helping Students Find and Follow a Path to Success." 2019. Ellucian.