Analysis of financial value and altruism of alumni at American colleges

36-315 Final Project

Colin Yip and Yutong Wu

June 24, 2022

Contents

Description of dataset

For high school seniors, choosing a college is a life-changing decision. Financially, some colleges are more expensive than others and some colleges lead to better job opportunities than others. Socially, the college a student chooses affects their happiness for the four years they attend, and leads to lifelong connections with peers. To determine which colleges provide the value and experience, we compare the **cost**, **expected career pay of alumni**, and **percent of alumni who think they are making the world a better place** of American colleges.

We will analyze a dataset which contains the tuition, diversity, and average pay of alumni for American colleges and universities. The dataset is compiled from various websites, but all originally comes from the US Department of Education. There are three tables in the dataset (tuition cost, salary potential, and diversity school). The tuition cost data comes from the National Center for Education Statistics and contains the in-state tuition, out-of-state tuition, and cost of room and board for each college. The salary potential data records the average early and mid career pay, and the percent of alumni who think they are making the world a better place for each college. The diversity school table contains the enrollment of each racial demographic at each institution.

Link to dataset: https://github.com/rfordatascience/tidytuesday/blob/master/data/2020/2020-03-10/readme.md#diversity_schoolcsv

Descriptions for the variables:

Tuition Cost

1. Name: School name
2. State: State name
3. State_code: State abbreviation
4. Type: Type of school (Private, public non-profit, or for-profit
5. Degree length: 4 year or 2 year degree
6. Room and Board: Room and Board in USD
7. In State Tuition: Tuition for in-state residents in USD
8. In State Total: Total cost for in-state residents (sum of room & board + 9. in-state tuition)
9. Out of State Tuition: Tuition for out-of-state residents in USD
10. Out of State Total: Total cost for out-of-state residents (sum of room & board + out-of-state tuition)

Salary Potential

1. Rank: Potential salary rank within state
2. Name: Name of school
3. State: State name
4. Early career pay: Estimated early career pay in USD
5. Mid career pay: Estimated mid career pay in USD
6. Make world better percent: Percent of alumni who think they are making the world a better place
7. Stem Percent: Percent of student body in STEM

Diversity School

1. Name: School name
2. Total enrollment: Total enrollment of students
3. State: State name
4. Category: Group/Racial/Gender category
5. Enrollment: Enrollment by category

Research Questions

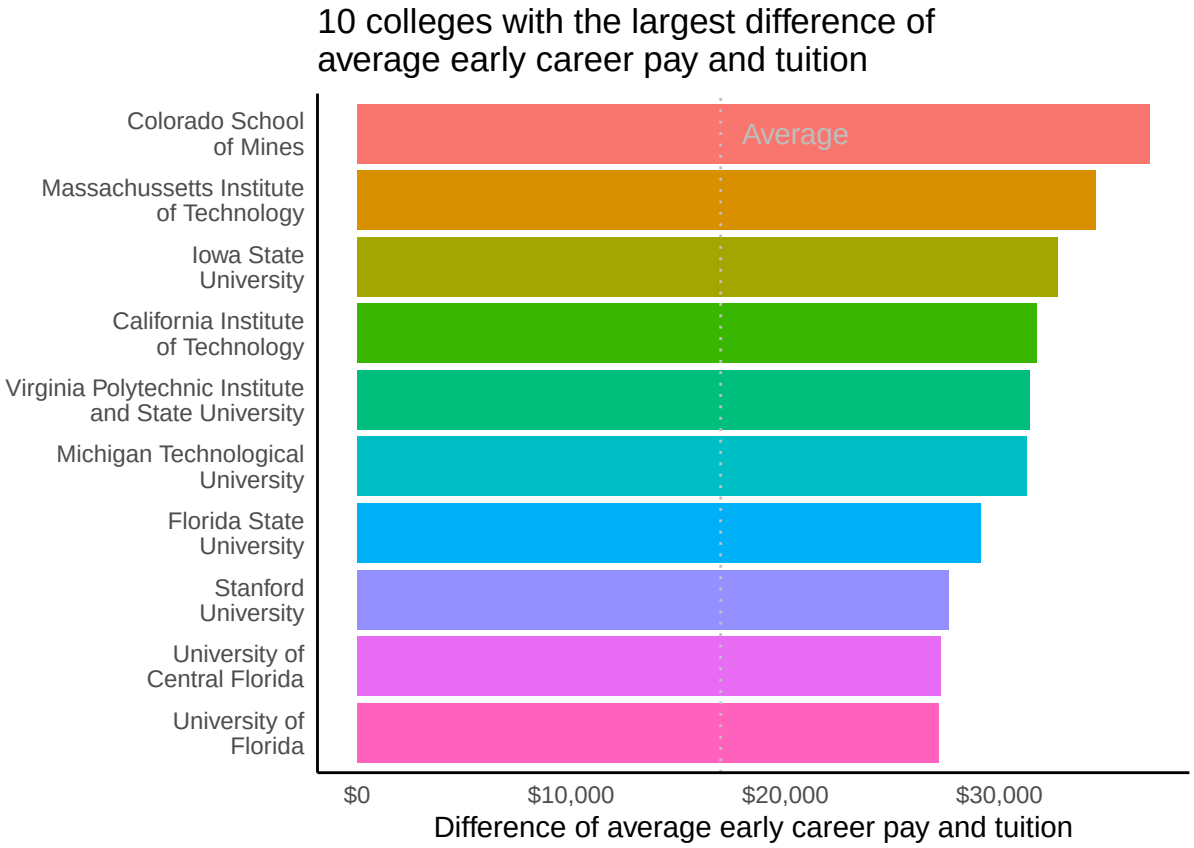
We plan to explore how the salary of alumni and percentage of alumni who think they are making the world a better place changes with other variables. Some variables include the diversity, percent of undergraduates majoring in STEM subjects, and whether the institution is a national university or liberal arts college. In turn, we want to provide some insights about how to choose a university.

Here are a few specific questions we plan to explore:

1. Which colleges have the largest positive difference between early career pay and tuition?
2. Do STEM-focused schools have higher early career and mid career pay?
3. Is there correlation between cost and salary?
4. How do average mid career pay and percent of alumni who think they are making the world a better place compare at liberal arts colleges and national universities?
5. Colin and Mark's story. Which colleges that Mark and Colin were accepted have the best financial value and most altruistic alumni?

Question 1: Which colleges have the largest positive difference between early career pay and tuition?

To investigate which colleges that provide the best value, we will first find the colleges with the largest difference between average early career pay and out of state tuition. Only national universities in the top 150 of the U.S. News and World Report Rankings were included.



Out of the top 150 ranked national universities, the ten colleges with the largest difference of average early career salary and tuition are the following:

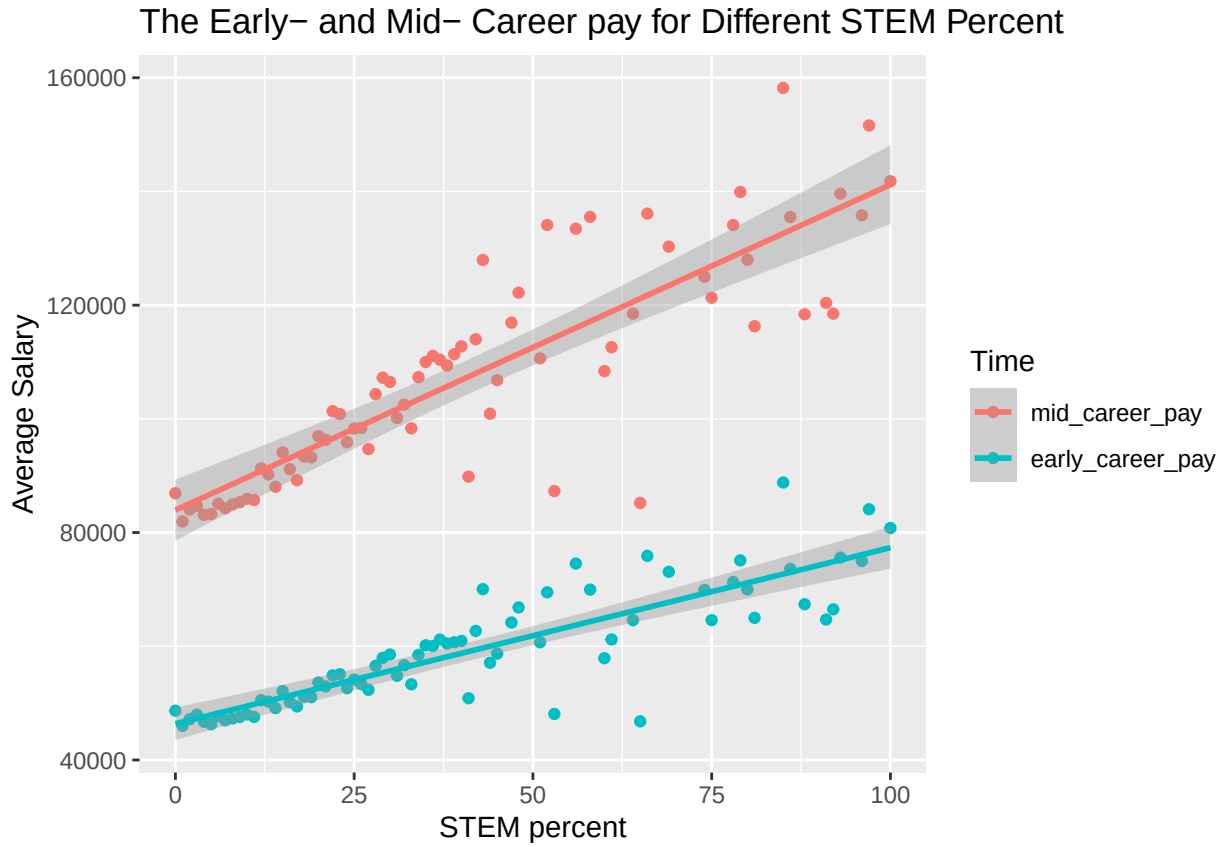
1. Colorado School Of Mines
2. Massachussetts Institute of Technology
3. Iowa State University
4. California Institute of Technology
5. Virginia Polytechnic Institute and State University
6. Michigan Technological University
7. Florida State University
8. Stanford university
9. University of Central Florida
10. University of Florida

These schools have a large difference for one of two reasons. They have either low tuition or high average early career pay compared to the rest of the top 150 national universities. The private universities (California Institute of Technology, Massachusetts Institute of Technology, and Stanford University) tend to have high early career pay, with an average pay of 82 thousand dollars among the three private universities to an average pay of 63 thousand dollars among the seven public universities. On the other hand, the public

universities tend to have lower tuition, with an average tution of 31 thousand dollars among the seven public universities to average tuition of 52 thousand dollars among the three private universities.

Question 2: Do STEM-focused schools have higher early career and mid career pay?

Many high paying careers such as software development, data science, and engineering require skills in STEM (Science, Technology, Engineering, and Math). As such, we are interested in investigating whether schools with a high percent of students majoring in STEM have high average salaries of alumni.



Schools are grouped based on the percent of the STEM (Science, Technology, Engineering, and Mathematics) degree. In both groups, each dot represents the average salary for all schools with a given percent of students majoring in STEM subjects. According to the scatter plot and the simple linear regression line, there is a positive correlation between the stem_percent and the pay, and the difference is more prominent in the mid-career than early-career. The regressed lines almost perfectly simulate for those under 40%, while has huge deviations for those higher than that threshold.

```
##
## Call:
## lm(formula = early_career_pay ~ stem_percent, data = salary_potential)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -21503  -3858   -729   2956  45415
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  45784.53     324.31  141.17   <2e-16 ***
## stem_percent   328.71      14.16   23.21   <2e-16 ***
```

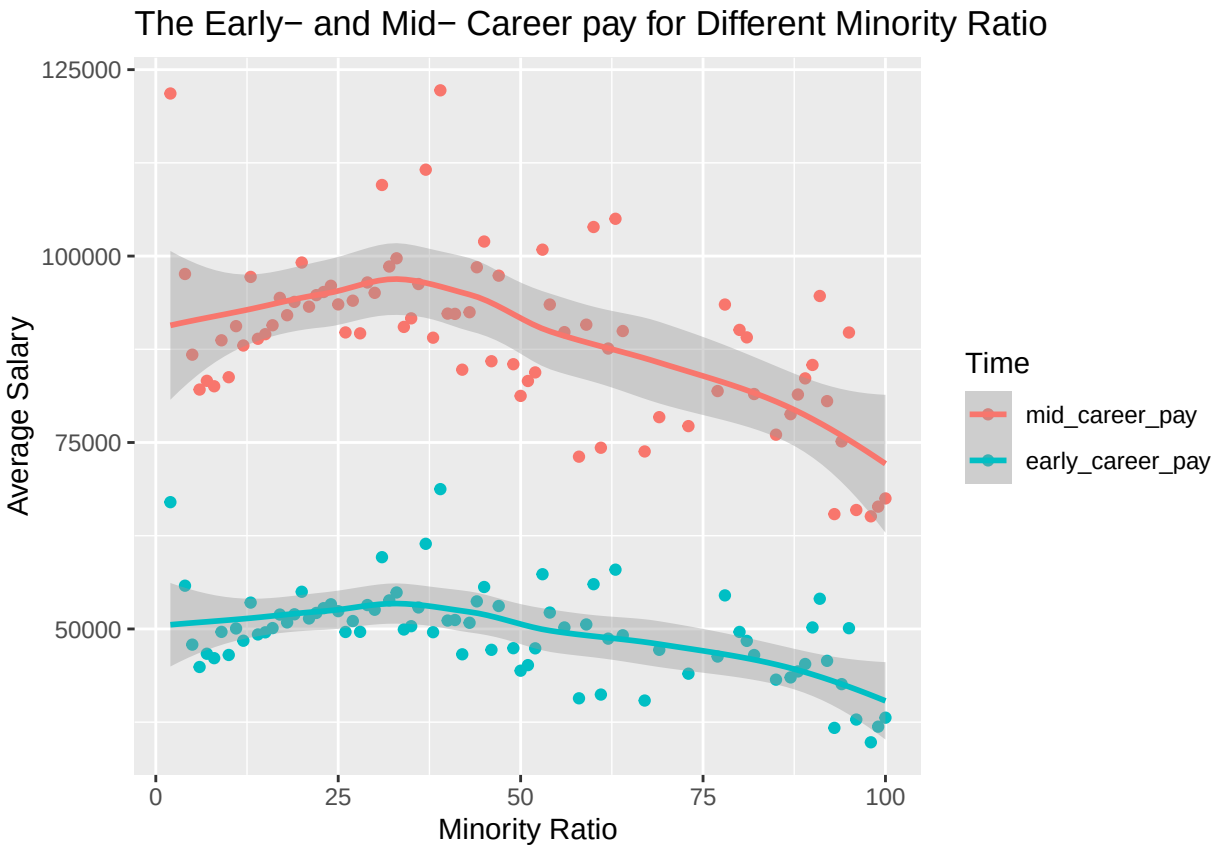
```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6664 on 933 degrees of freedom
## Multiple R-squared:  0.3661, Adjusted R-squared:  0.3654
## F-statistic: 538.8 on 1 and 933 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = mid_career_pay ~ stem_percent, data = salary_potential)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -40027  -7581  -1434   5997  72144
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  81956.19     605.05  135.46   <2e-16 ***
## stem_percent   639.70      26.42   24.21   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12430 on 933 degrees of freedom
## Multiple R-squared:  0.3859, Adjusted R-squared:  0.3852
## F-statistic: 586.3 on 1 and 933 DF,  p-value: < 2.2e-16
```

Given the infinitesimal p-value (<2e-16) for the slope, which is much smaller than 0.05, we would reject the null hypothesis, and determine that there are linear relationships between the salary and the percent of STEM degrees. In addition, there is a huge difference between the slope (whose ratio is around 1 to 2), that is to say the difference between the salaries for different STEM degree percentages is more prominent in the mid-career than early in the career.

Question 3: Is there correlation between diversity and salary?

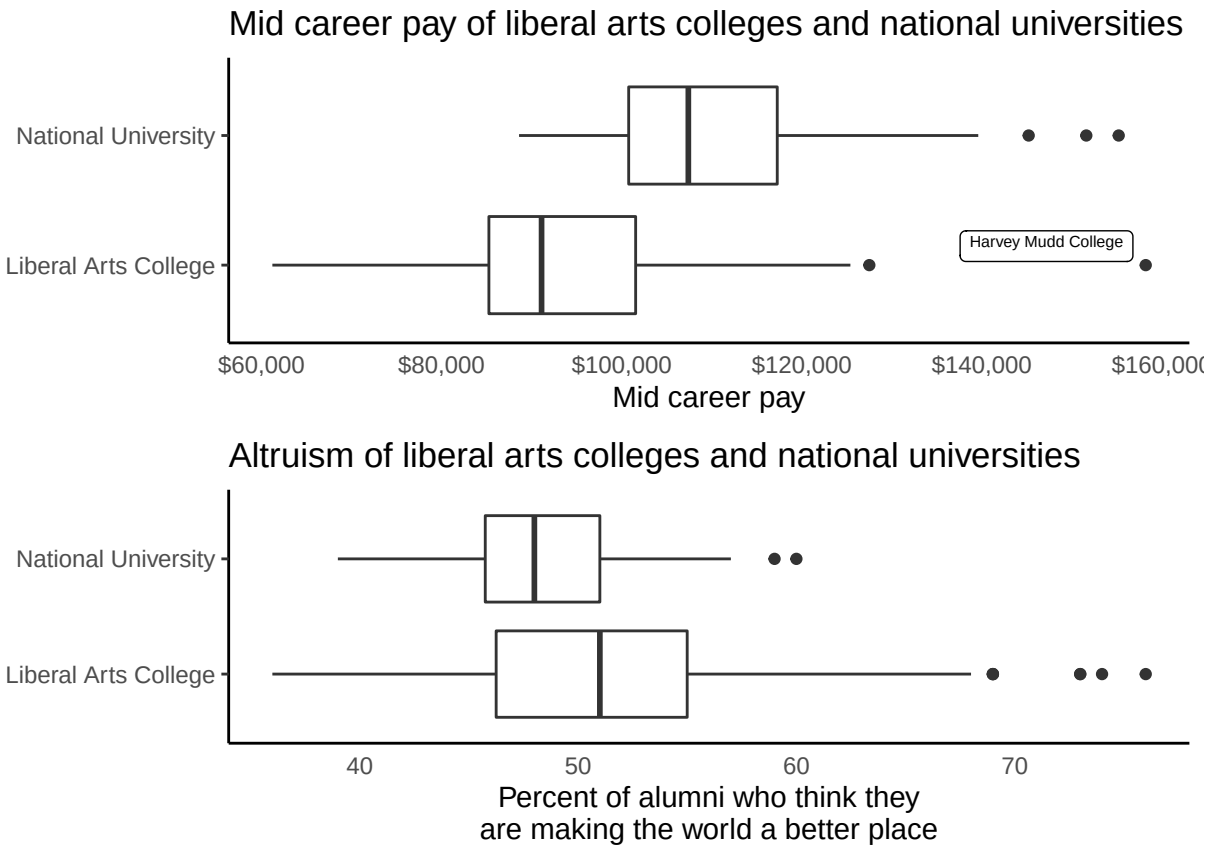
Ethnic diversity exposes students to new perspectives and helps students connect with cultures other than their own. We are interested in whether these benefits enhance a student's career prospects. We will examine whether diverse schools have higher average salaries for alumni than racially uniform schools.



The graph shows the relationship between the minority ratio and average salary, and it shows that there is a approximately negative quadratic relationship. According to the graph and regression line, colleges with 25% to 37.5% minority ratio have the highest payment, although there is no clear difference for those under 50%.

Question 4: How do average mid career pay and percent of alumni who think they are making the world a better place compare at liberal arts colleges and national universities?

U.S. news categorizes institutions into two broad categories, national universities and liberal arts colleges. National universities typically focus on research, offer graduate and Ph.D. programs, and have large enrollment. Liberal arts colleges usually focus on undergraduate education, do not offer graduate programs, and have smaller enrollment. We will investigate how their alumni compare with regard to average mid career pay and percent of alumni who think they are making the world a better place.

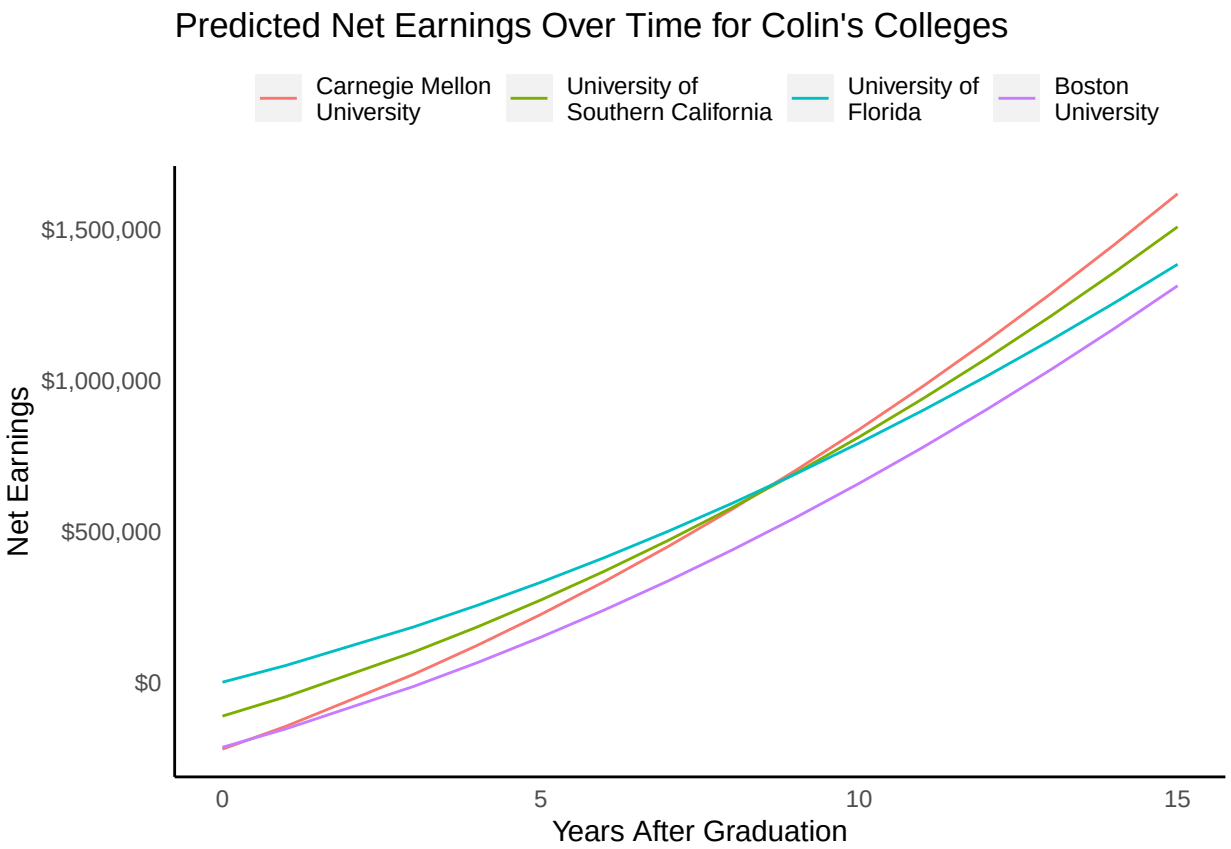


In general, average mid career pay of alumni is lower for liberal arts colleges compared to national universities. However, a higher percentage of alumni think they are making the world a better place at liberal arts colleges than at national universities.

One liberal arts college, Harvey Mudd College, has an unusually high average mid career pay of $158,200, over $30,000 more than the second highest liberal arts college. Upon further analysis, Harvey Mudd has a stem percent of 85%. In the top ten schools by mid career pay (both national universities and liberal arts colleges included), eight have a stem percent over 50%, far higher than the national average of 23.4%. This suggests that stem percent is more predictive of a college's expected mid mid career pay than whether the college is a liberal arts college or national university.

Question 5: Colin and Mark's story. Which colleges that Mark and Colin were accepted have the best financial value and most altruistic alumni?

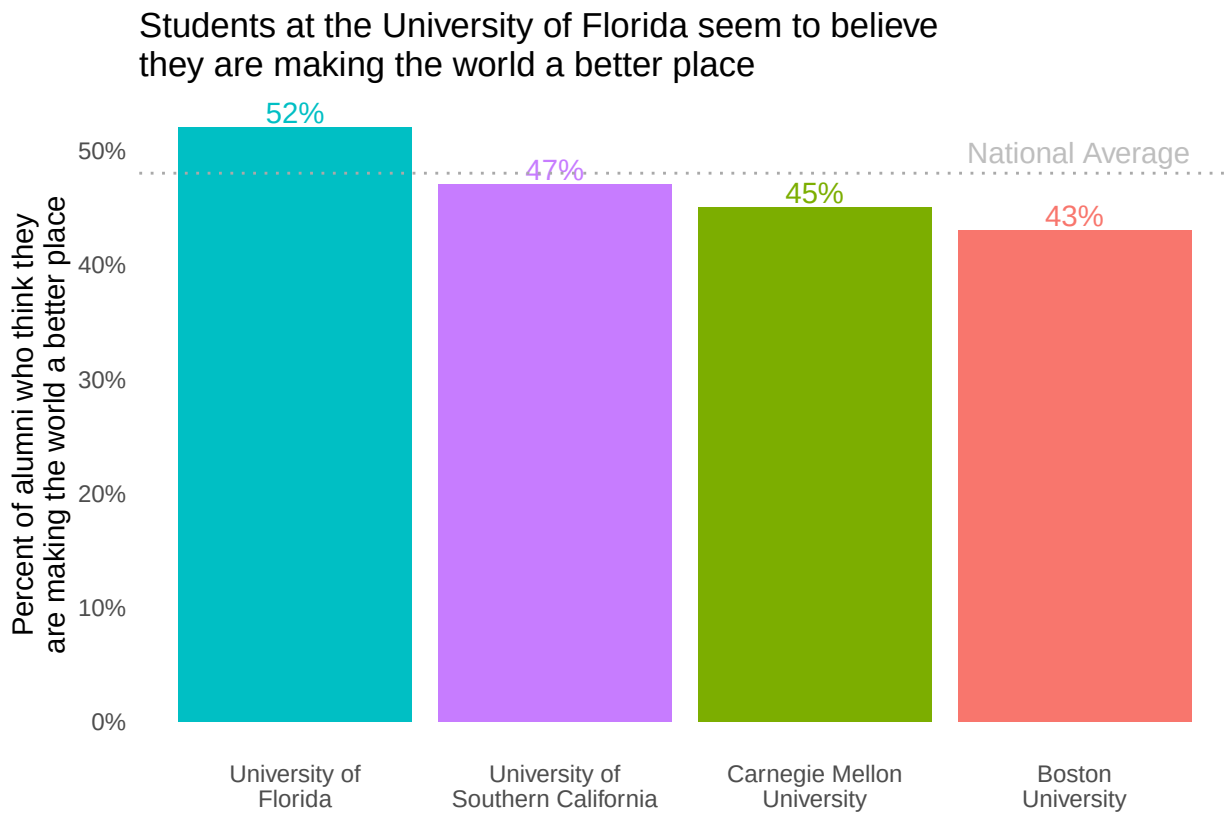
First, we will examine the the financial value and altruism of alumni (quantified by percent of alumni who believe they are making the world a better place) at each of the schools we were accepted at. In particular, we seek to evaluate how the average pay and altruism of alumni of CMU compare to other colleges we had the choice to attend.



Predicted net earnings was calculated by subtracting the cost of four years of tuition from the total income. For total income, we used the average early career salary in year one, with salary increasing in even increments each year until it reaches the school's mid career salary in year ten. Salary then continued to grow linearly unitl year fifteen.

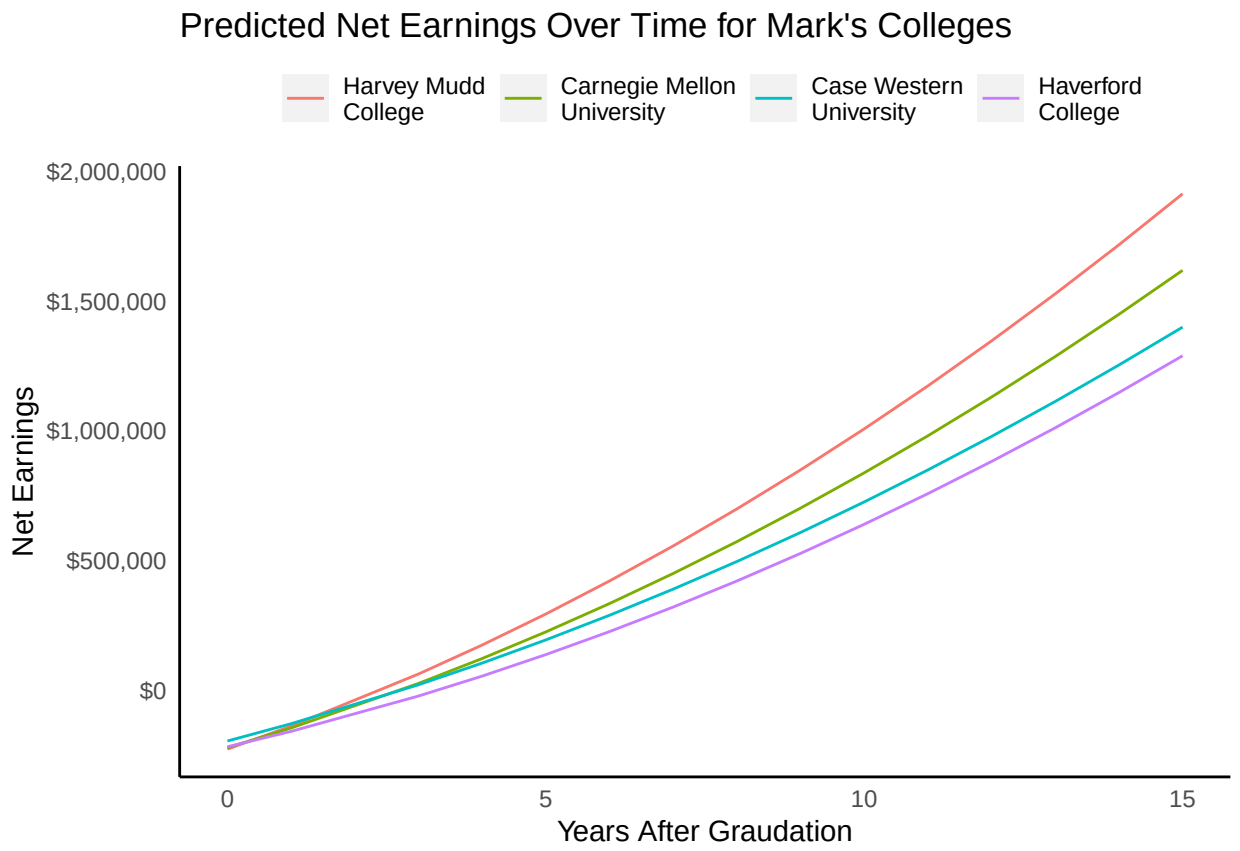
Due to scholarships, the University of Florida and the University of Southern California led to the highest net earnings for the first five years. However, the higher estimated early and mid career pay of Carnegie Mellon University led it to be the highest after fifteen years. Thus, using our metrics, Carnegie Mellon University is worth the initial investment for Colin in the financial long term.

We will next analyze the level of altruism of alumni at schools Colin was accepted to.

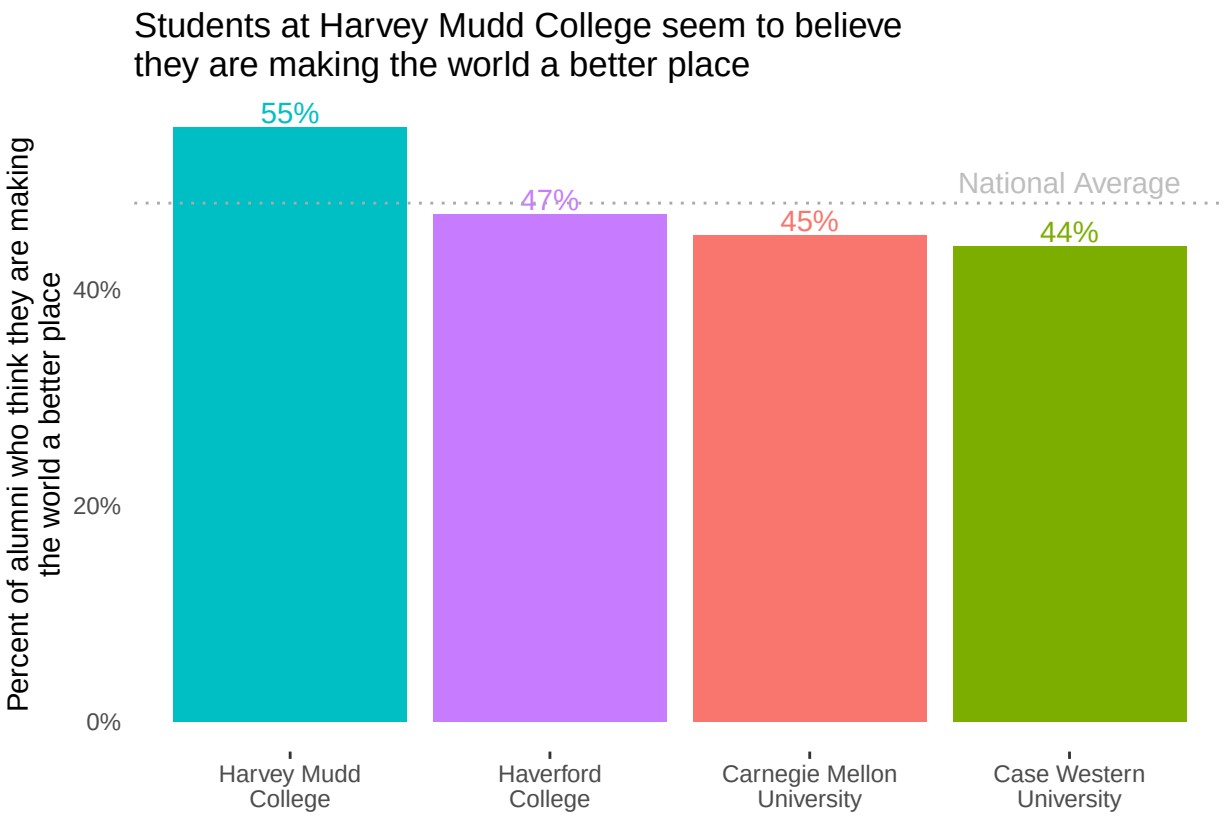


Out of the four colleges Colin was accepted to, the University of Florida has the highest percentage of alumni who think they are making the world a better place (52%), followed by the University of Southern California (47%), Carnegie Mellon University (45%), and Boston University (43%). Only the university of Florida is above the national average of 48%.

Next, we will analyze the financial value and fullfillment of alumni at schools Mark was accepted to.



Based upon the simulation, it illustrates that the predicted net earning is highest for Harvey Mudd College, and the second is Carnegie Mellon University. Despite this fact, Carnegie Mellon University is famous for its interdisciplinary education. Students could savor this advantage to have deep insights into both humanities and sciences.



According to the data of these four school, it seems that the two liberal arts colleges (Harvey Mudd College and Haverford College) have a high rate of alum who think they are making the world a better place, possibly due to the size of each class and the strong cohesion. Carnegie Mellon University ranked third, with a lower percentage of alumni thinking that they make the world a better place than at Haverford College and Harvey Mudd College.

Conclusion

This exploration compares specific schools with highest return, the relationship between the salary and STEM degree percent and minority ratio (two common indicators for university rankings). We found that the national universities with the highest difference of early career pay and tuition were either highly selective private universities or public universities with lower tuition. Next, the positive correlation between salary and STEM degree percent conforms with the statement that STEM students are easier to have higher payment than non-STEM students. Third, the approximately negative quadratic relationship with the minority ratio suggests that communicating with peer students with different background could be beneficial for their future career; it is also true that high ratio of minority is another type of homo-cultural. Fourth, alumni are generally less highly paid, but more altruistic, at liberal arts colleges than at national universities. Finally, out of the colleges that Colin and Mark were accepted to, Carnegie Mellon University and Harvey Mudd College provide the best financial value, meanwhile alumni are most altruistic at the University of Florida and Harvey Mudd College.

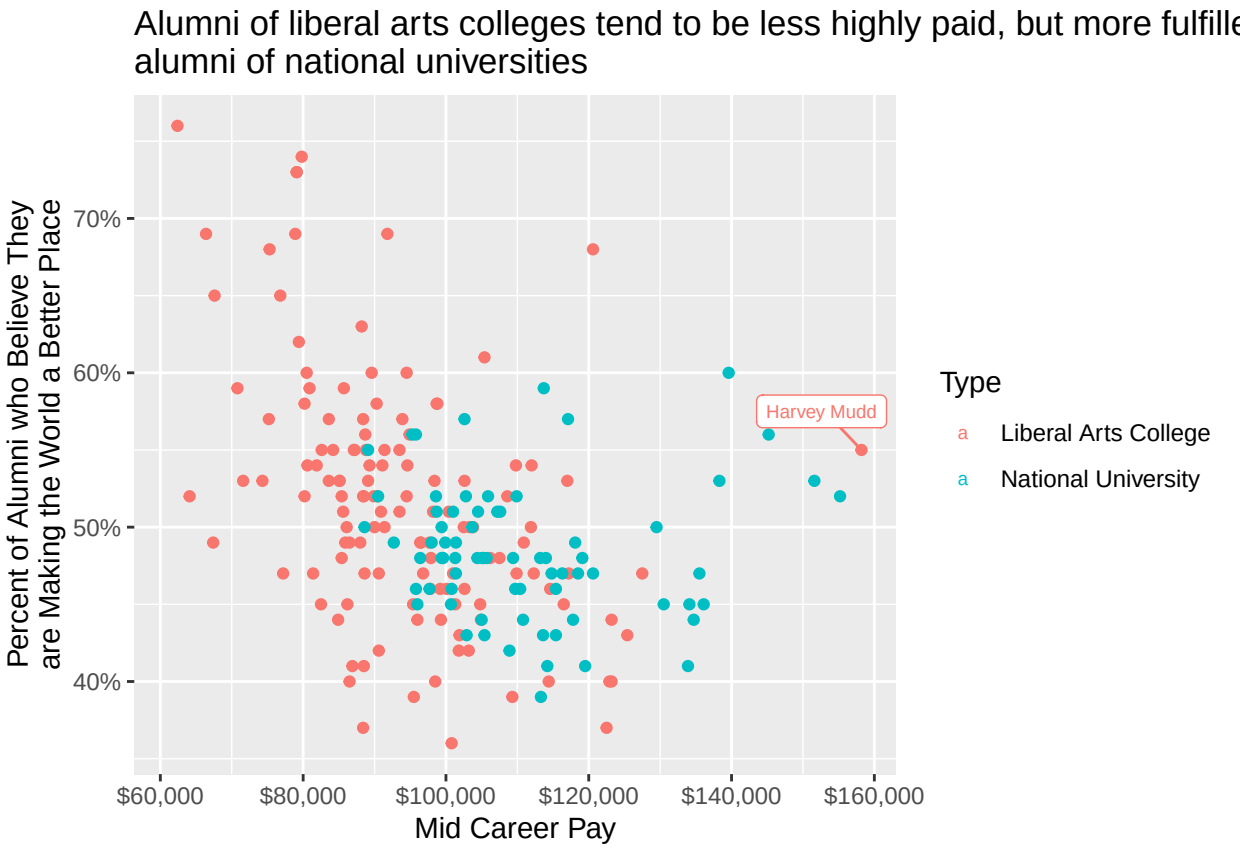
One limitation of our analysis is that it does not take financial aid and scholarships into account. Since many highly ranked colleges offer generous financial aid, our analysis may underestimate the financial value these schools provide. Another limitation is that we used out-of-state tuition as the tuition for all public colleges. This may make state colleges and universities appear to have worse value than they actually do,

since in-state tuition is much cheaper than out-of-state.

As a follow-up experiment, we could find the colleges with the largest difference of early career pay and average net cost (after financial aid). This is identical to question 1, except that it uses average net cost instead of tuition. Since this accounts for financial aid, it may more accurately determine which colleges provide the best financial value. As another possible study, we could compare the value of attending an in-state public college vs. out-of-state or private college. Third, since financial aid depends on family income, we could analyze which colleges provide the best value for families of each income bracket.

Appendix

First attempt at question 4: Do alumni make more money and feel more fulfilled for liberal arts colleges than for national universities?

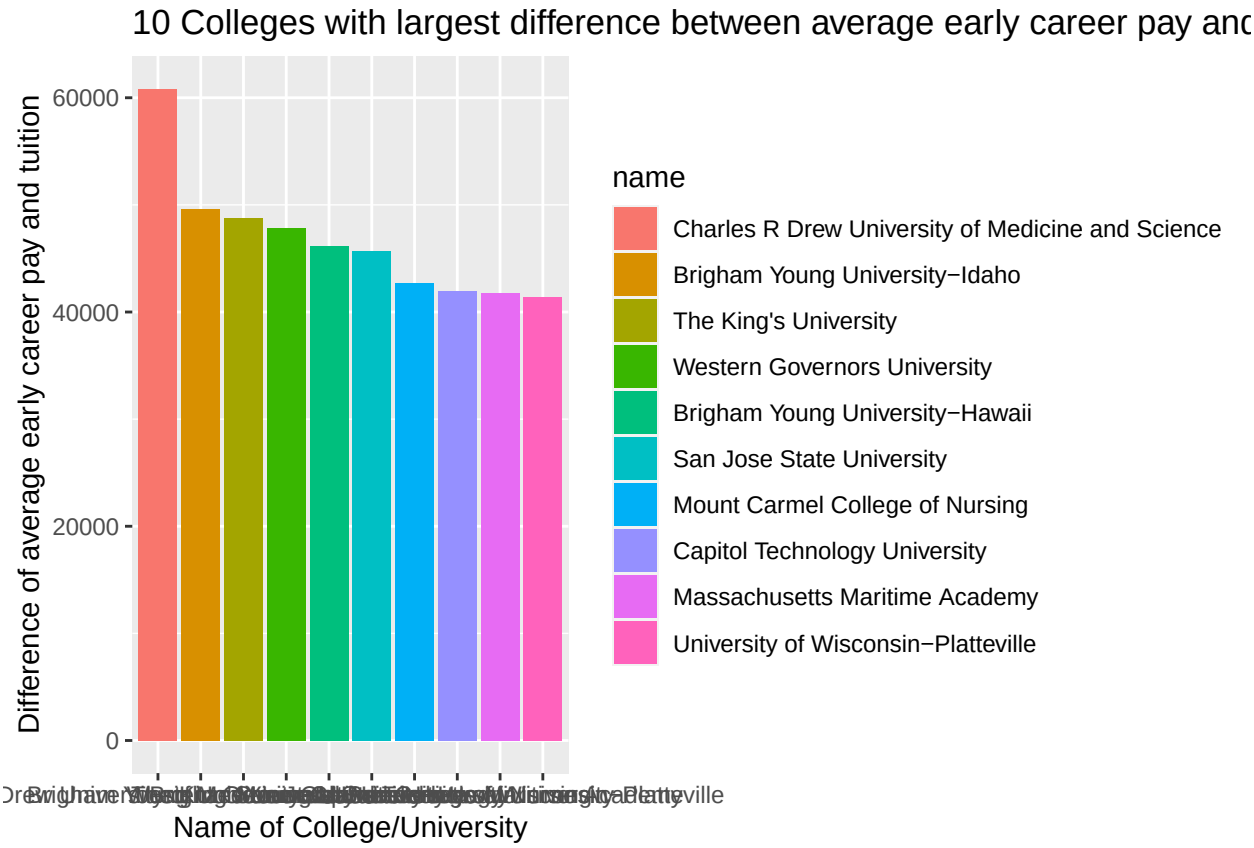


Mid career pay and percentage of alumni who believe they are making the world a better place, for national universities and liberal arts colleges. Among liberal arts colleges, it appears that as mid career pay increases, fulfillment decreases. Among national universities, there appears to be little association between mid career pay and fulfillment. In general, alumni of liberal arts colleges tend to have lower mid career pay than alumni of national universities. One liberal arts college, Harvey Mudd College, has an average mid career pay of $158,200, over $30,000 more than the second highest liberal arts college, Williams College.

I did not include this plot and caption in the final report because it does not show the median mid career pay and median percentage for liberal arts colleges and national universities. The two boxplots, which I included in the final report, do.

Attempt at question 1: Which colleges have the largest positive difference between early career pay and tuition?

To investigate which colleges that provide the best value, we will first find the colleges with the largest difference between average early career pay and out of state tuition.



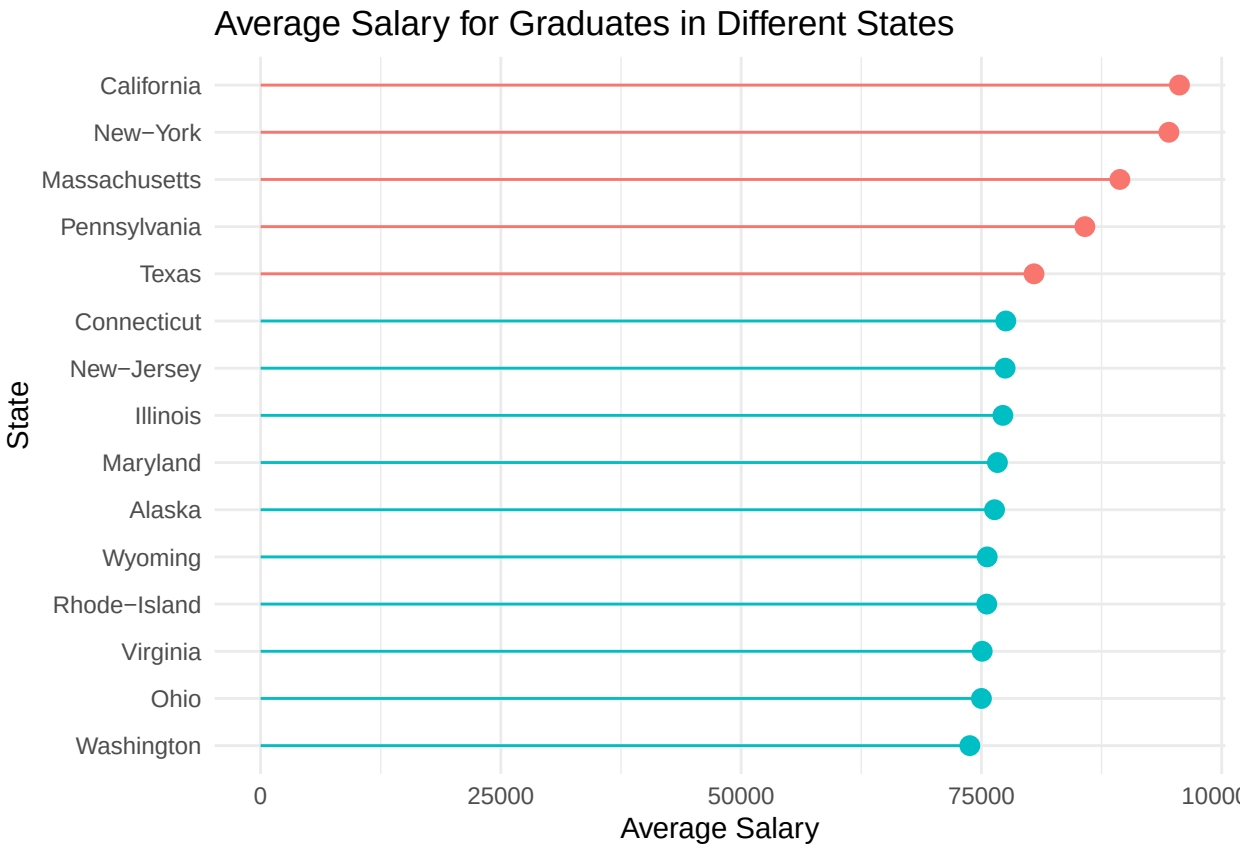
The ten colleges with the largest difference of average early career salary and tuition are the following.

1. Charles R Drew University of Medicine and Science
2. Brigham Young University-Idaho
3. The King's University
4. Western Governors University
5. Brigham Young University-Hawaii
6. San Jose State University
7. Mount Carmel College of Nursing
8. Capitol Technology University
9. Massuchussetts Maritime Academy
10. University of Wisconsin-Platteville

However, this list may be misleading when determining which colleges provide the best value for a student applying to college. First, some of the colleges in this list are not comparable to traditional four year institutions. For example, Western Governor's University (4th in diff) is an online only college and Charles R Drew University of Medicine and Science is a graduate college. Second, most of the colleges in the top 10 in diff have low graduation rates (most under 50%). This makes the average early career pay misleading, as that statistic only includes those who graduated from the institution.

For the bar chart in question one of the final report, we brought in a dataset containing the college ranking of each school, according to U.S. News and World Report. We then filtered the data to only the top 150 national universities.

Additional question: Is there any correlation between the salary and where the university is located?



Students who graduated from California, New-York, Massachusetts, Pennsylvania, and Texas earn the most and are much higher than those who graduated in the rest states.

We excluded this visualization from the final project because the question was less interesting than the other questions.