Kate Lanman: lanman.k@northeastern.edu

Teera Tesharojanasup: tesharojanasup.t@northeastern.edu

The Rising Cost of Education

Introduction

Over the past few decades, the cost of higher-level education has risen at unprecedented rates. Since 1980, college costs have increased by more than 169% (cnbc.com). Generally, the rising cost of education is attributed to inflation rates, yet over the past 50 years, the cost of college has increased at a rate more than 4 times that of inflation (foxbusiness.com). Accompanying the tremendous increase in college costs is the question of how this will impact students' ability to pay for school. Specifically, there is the concern that as costs rise, students will be forced to take out greater loans and in turn, take on greater debt in order to pay for school. The rising cost of education then becomes an access issue looking at who is able to afford college and what repercussions taking on debt might have for certain students once they graduate. The current research aims to gain a better understanding of these increases through the question: how has the cost of higher education within the United States changed from 2014 to 2021, and what factors contribute to this change? This question is broken down more specifically:

1. How does the type of school (public, private for-profit, and private non-profit) affect the cost of education?
2. What is the correlation between the cost of education and the number of international students in the United States?
3. What is the correlation between the cost of education and the inflation rate in the United States?
4. How does the cost of education relate to student loan debts in the United States?

Gaining a deeper understanding of how college costs are changing as well as what is affecting these costs is a step towards a clearer understanding of the issues of affordability and accessibility of education.

The Data

In order to answer these questions, nine datasets were used. The first five of these were data from the Integrated Postsecondary Education System (IPEDS) containing information on education costs over the 2014 to 2020 academic years grouped by type of school (public, private non-profit, and private for-profit) as well as the type of cost (tuition, books and supplies, room and board, and other expenses). This data is reported by Title IV institutions within the United States. This data also filtered data by level of institution. However, for the purposes of this research, only four-year institutions were examined. Moreover, public school tuition was taken as the average of in-district, in-state, and out-of-state tuition to make this figure more comparable to the tuition at the two other types of schools. Also, where relevant, only on-campus data was used in an effort to avoid the influence of extraneous factors (i.e. school location) on these costs. These datasets were used for the bulk of our analysis with others included to support our understanding of the patterns this information presented.

To look at changes in United States inflation over time, data was pulled from the Federal Reserve Economic Database (FRED). This dataset contained data ranging from 1960 to 2020 which was taken from the World Bank. To be consistent with the IPEDS data, only data from academic years 2014 to 2020 was used.

A dataset from Statista was used to gather information on the number of international students within the United States by year. The data ranged from academic years 2003 to 2020 and was gathered from the Institute of International Education. As with the FRED data, only data from academic years 2014 to 2020 was used.

Additionally, a dataset from the Education Data Initiative was used to look at the average student loan debt in the United States by year. This data ranged from 1917 to 2021 and was taken from a report done by George Washington University. Once again, only data from 2014 to 2021 was used.

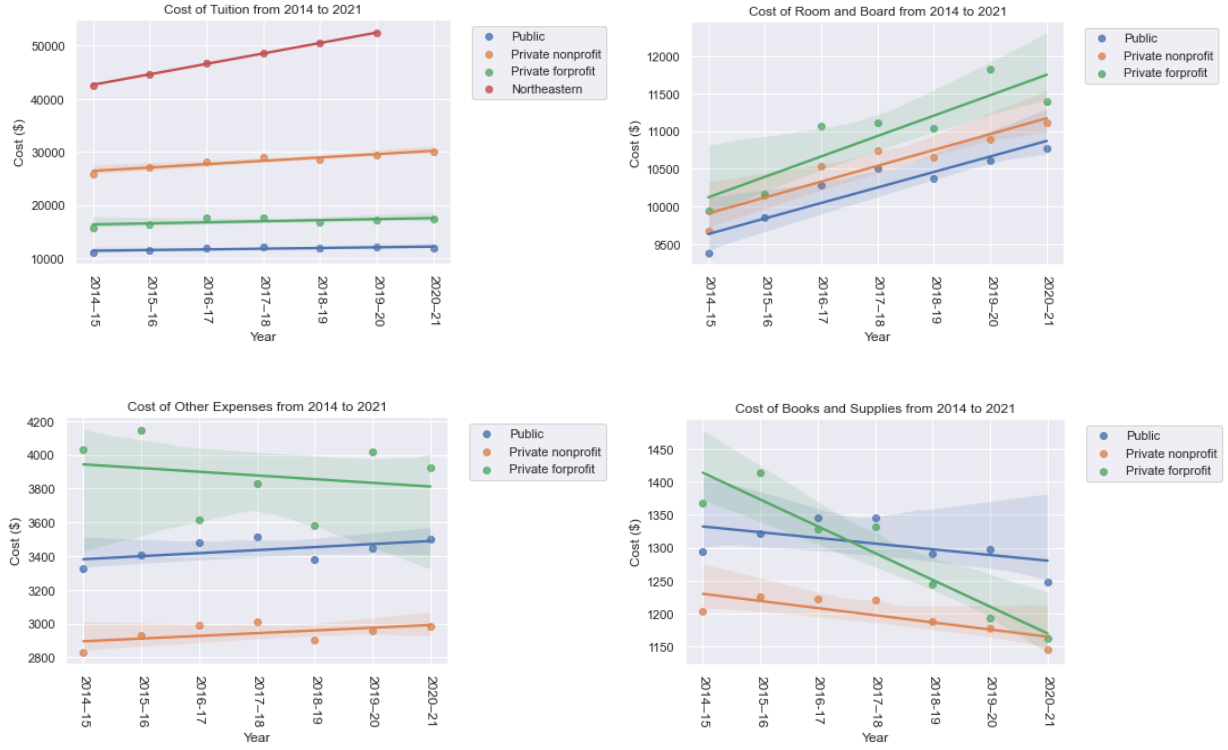
Finally, a dataset on tuition costs at various colleges over the academic years 2014 to 2019 was pulled from Data USA. This was used to specifically look at the changes in tuition at Northeastern University and how this compares to the changes in tuition at United States colleges as a whole.

Data Science Approaches

Our methods can be broken down generally into two parts: the reading and cleaning of the data and the data analysis. There were many challenges with gathering the required data to answer our questions, specifically with the formatting of the IPEDS dataset. Most notably, this data was split into five separate datasets each containing information on two academic years with some overlap between a few of the datasets. In order to merge these datasets into one, inconsistencies in the formatting of these as well as unnecessary rows and columns had to be accounted for. Solving this issue required broad use of the Pandas library. Overall, problems such as the removal of wrong type or unwanted data were solved similarly with the use of Pandas, lambdas, and various other built-in functions. Cleaning the datasets in these ways allowed for a more effective and efficient analysis to be conducted.

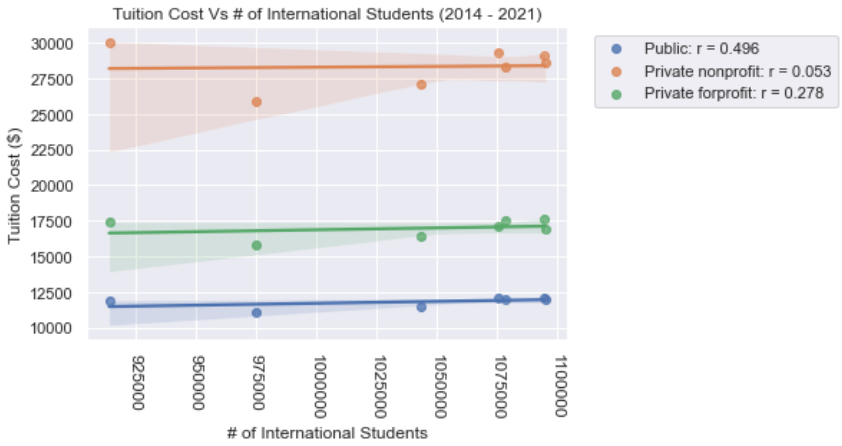
Our data analysis consists of seven graphs in total. Each graph is a scatter plot that contains three lines of best fit for public, private non-profit, and private for-profit schools. Here, we used the seaborn, matplotlib, and scipy modules. We used the seaborn and matplotlib modules for plotting the graph itself and creating the line of best fit and the scipy modules to find the correlation coefficient for each line of best fit.

Results and Discussion



Graphs 1-4

In graph 1, we plotted the cost of tuition for all three school types from 2014 to 2021 and Northeastern from 2014 to 2020 since we could not find data for the 2020-2021 year for Northeastern. From 2014 to 2021, we can see a clear and consistent upward trend for all types of universities where private non-profit schools seemed to have the greatest increase in costs as well as having the highest tuition costs. Northeastern (private non-profit) in comparison has a tuition cost far higher than the average tuition cost of a private non-profit institution. Northeastern also has the steepest increase in tuition costs over the years. The cost of room and board has been increasing from 2014 to 2021 too. However, the cost of books and supplies has been decreasing and other expenses have been seeing a downward trend for private for-profit universities.

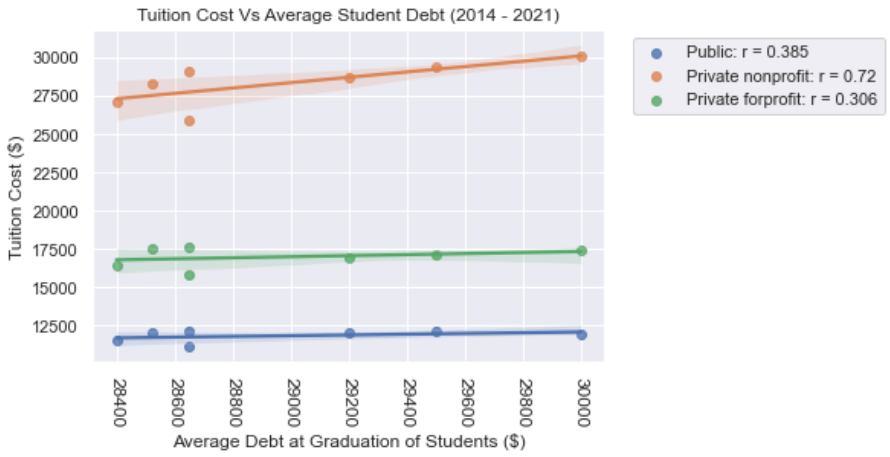


Graph 5

In graph 5, we plotted the tuition cost of all three types of universities compared to the number of international students within the United States. Here, we can see a weak/moderate positive correlation between the two variables for the three types of universities. Public universities seem to have a stronger correlation than private non-profit and private for-profit universities. We suggest this could be due to the fact that public universities have a much greater range in pre scholarship tuition costs such that a greater number of international students at a public university would have a greater effect on the tuition. That is, there is a difference of tens of thousands of dollars between the students paying the least (in-state) and the students paying the most (out-of-state) where international students count as out-of-state students.

Graph 6

In graph 6, we can see a weak positive correlation between inflation and tuition cost. Past research has shown that as inflation increases, products and services also increase in price, and tuition cost is not exempt from this. Due to the few data points, we are unable to show a stronger correlation between inflation rate and tuition cost. However, inflation rate can be a useful predictor of future tuition costs, especially for public universities. We speculate that this is because public universities' set their tuition cost based on the amount of funding the state and federal government gives them, which is all in US dollars. However, private for-profit and non-profit universities receive funding from some investors who are abroad and therefore not subject to the US inflation rate which could lead to the US inflation rate becoming a less reliable predictor of the tuition cost for non-profit and for-profit private universities.



Graph 7

Graph 7 shows us that as tuition cost rises, average student debt also rises. The line of best fit for private non-profit universities is the steepest and also has the strongest correlation whereas public and private for-profit universities have a weak correlation. This result was logical since, with higher tuition costs, students would be forced to borrow more money which would lead to higher average student debt.

Limitations

Due to our dataset containing few data points, we are unable to come to a strong conclusion from our seven graphs. Due to the small number of data points, it can cause the linear regression plot to have more variability, therefore, generating weaker conclusions. The dataset has such a small number of data points because the US only recently started tracking the average tuition cost of universities. Another limitation is that our school groups are very broad. For example, the number of public universities across the US is more than a thousand and they are all in different states so they receive different amounts of funding. This greatly affects their tuition cost so it would be better to look at the tuition costs of public universities in different states separately but our dataset looks at all public universities across the US.

Conclusion

After we finished our code to display the seven graphs, we learned that the tuition cost of all three types of universities has been increasing over the years. When we added Northeastern (private non-profit) as a comparison, we thought it would be quite high, but we were surprised that it almost doubled the average private non-profit universities tuition cost and has seen the steepest increase. When we looked at the correlation between the number of international students and tuition costs, we learned that there was not a strong correlation between the two which we thought was startling. We initially believed that there would be a stronger correlation between those two variables.

For the correlation between inflation rates and tuition cost, we thought that the correlation between the two should have been much stronger due to past research showing that the two variables correlate. However, our graph only shows a weak to moderate correlation which most likely stems from the fact that there are too few data points on the scatter plot. We also expected a strong correlation between average student debt and tuition costs which our graph did show but only for private non-profits which made sense since students would have to borrow the most money to attend

one on average. To summarize, our conclusions from our graph should be taken with a grain of salt since there were few data points plotted and should not be used for accurate predictions. Our research, unfortunately, revealed unsatisfactory results which was the result of our dataset that contained fewer data points than we expected.

Future Work

Our project only gave us a glimpse at the US's education costs. Further work should try to include more years if possible to be able to plot more data points which would yield better results. To explore this topic even further, we could also look at universities from across the globe and compare their education costs with US universities. We could also get more specific by looking at return on investment (ROI), by comparing university tuition compared to how much a graduate makes on average. We can analyze this and see which types of universities (private for-profit, private non-profit, or public) have the greatest ROI and even predict which type of college would lead to the most successful graduates in the near future. We can even go one step further by creating a monte carlo simulation of how much money a graduate is expected to make by graduating from a certain type of university.

Sources

https://www.cnbc.com/2021/11/02/the-gap-in-college-costs-and-earnings-for-young-workers-since-1980.html

https://www.foxbusiness.com/personal-finance/rising-college-costs-over-time-inflation-report

https://nces.ed.gov/ipeds/search?query=&query2=&resultType=all&page=1&sortBy=date_desc&overlayTableId=25002

https://fred.stlouisfed.org/series/FPCPITOTLZGUSA

https://www.statista.com/statistics/237681/international-students-in-the-us/

https://datausa.io/profile/university/northeastern-university#costs

https://educationdata.org/average-student-loan-debt-by-year