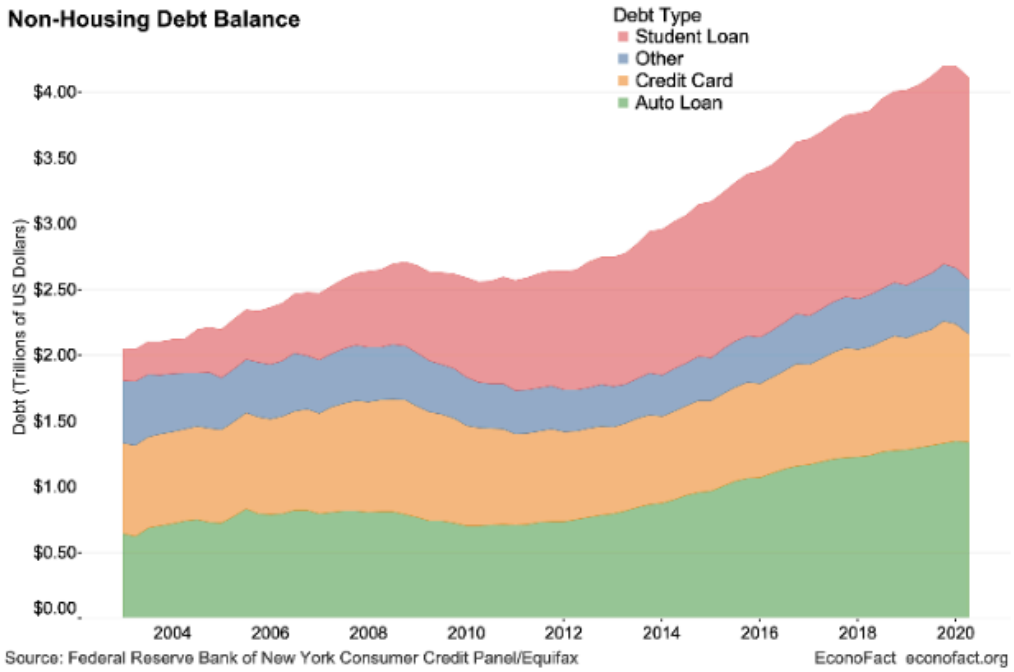
Project Proposal

Client:

The clients are High School seniors continuing their education paying for the cost with an educational loan. According to the National Center for Education Statistics, the United States had 3.4 million high school students graduated the school year 2020-2021. Of the 3.4 million, it is estimated 2.2 million of the high school graduates will attend some form of higher education in the form of a 2-year college or 4-year university.

Student loan debt in the United States is $1.75 trillion and grows about 5 times the nation's economy which is a problem for our young generation. By performing an exploratory analysis combined with qualitative and quantitative analysis, students will get an understanding and avoid the harmful effects of incurring large student debt at a young age by selecting an institution with low cost and competitive student earnings. Student loan tops the list of different types of loans in the country.



Question/Need:

- ### What is the question behind your analysis or model and what practical impact will your work have?

The student often makes the decision with the parents and the main point of discussion is often how to deal with the financial aspect of paying for years of education and boarding. Is the high cost of a 4-year university worth the amount of debt a student gets into?

This EDA will evaluate if incurring a minimal student debt can be associated with high earnings. As well as the relationship between low cost institution and high earnings. The practical impact will be for the student to save money and still have a career with competitive earnings.

- ### What is the purpose of the EDA you will do and the data science model you will propose?

Identify - What factors play a significant role in graduating from a university that has high earnings but low cost. In other words, which universities are worth the student debt based on what the student earns after graduation.

Relationships - Identify what university characteristics have the highest student earnings. What factors are important for a student in selecting a university?

Grouping or Segmentation - Which group has the most student debt? Group student debt by gender, race and income. Understanding this will help to address the specific community.

Impact

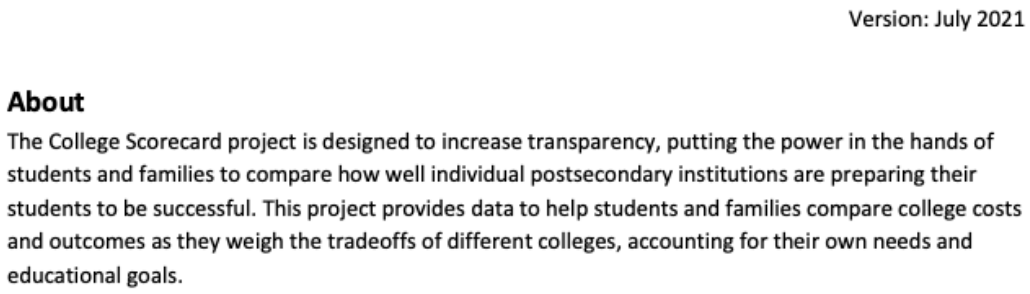
In 2022, there are many alternative methods for education or making a living. The traditional ways of having to go to the top university and being in debt the amount of a mortgage is not required to live as an average person in the United States. Therefore, by performing this EDA and bringing an awareness to the student debt crisis, we can have less students stuck in debt at a young age for a degree they might not be using.

Forecast future earnings so students can make better informed decisions about taking on student loan debt.

Data Description:

Website - https://collegescorecard.ed.gov/data/

- College Scorecard - is a dataset provided by the U.S. Department of Education. It contains institution-level data for all accredited institutions in the United States offering undergraduate degrees. It contains institutional admissions and academics data, student loan and student earnings data.
- Use of Python and API to retrieve the data



Version: July 2021

About

The College Scorecard project is designed to increase transparency, putting the power in the hands of students and families to compare how well individual postsecondary institutions are preparing their students to be successful. This project provides data to help students and families compare college costs and outcomes as they weigh the tradeoffs of different colleges, accounting for their own needs and educational goals.

Website - https://www.bls.gov/bls/blswage.htm

- U.S. Bureau Of Labor Statistics - provides occupational wage data in many formats such as by state or job characteristics with degree requirements.

Solution Path

Perform exploratory analysis to see the relationships of student loan data with student earnings data based on the universities characteristics analyzed by state and region for better comparison.

Use the institutional data with admissions criteria and academics records including cost to better understand the institutions.Then, identify institutions that have high earnings but low cost, so students can attend without having to incur significant student loan to earn a degree.

Alternatively, we have many other schooling options at a cheaper cost. Students can go the route of community college and add an industry required certificate for their resume or taking advantage of the open online courses platform to gain the knowledge and basic skills to put their foot in the door.

Criteria for Success

Success is having the amount of debt students take on decreases but they still have competitive earnings. For earnings to be considered competitive, a student will need to earn above their state's average monthly budget.

Assumptions and Risks:

Most students that go to college have not fully calculated what their earnings will be vs the debt they will incur.

A good portion of the population does not attend college to purse their passion, but rather to just find a career to make a living. For such people, the massive educational loan is not worth the earnings.

For those that are passionate about a specific specialized field, this analysis will not be useful to them as their focus is on the passion and not cost. Identifying the groups of students that are attending a university for basic survival need is key when promoting the results of this EDA.

Tools:

I will be using Python and API to retrieve the data and use Excel and Tableau for EDA.

MVP Goal:

> Visualize and understand cost of university with debt relationship as well as other university features including earnings to further build a regression model.