According to Amy Fontinelle in her blogpost on massmutual.com, taking on student loans judiciously for a graduate degree can be a worthy endeavor for a student (2020). One of the factors she advises to consider is the financial compensation in that field (Fontinelle, 2020). This advice seemed intriguing as a current Master’s program student living in Arizona.

The question I tried to answer with Exploratory Data Analysis was “do Arizonans with a master’s degree make more money than those with a bachelor’s?”. The dataset I used came from the United States Census Bureau and is from the year 2018, which is the most recent year Arizona data was available. I used this data and selected the variables I felt were the most useful. I included respondent income, age, marital status, educational attainment, and occupation.

What I found through all of the required analysis was that Arizonans with a master’s degree earn $11,313 more than Arizonans with a bachelor’s degree on average. What came along with that was a good deal of skepticism. Only about 17% of the variability in earnings could be attributed to controlling for marital status, occupation, age, and education. This tells me that there are must be other factors in life that influence earnings. It was also interesting to see that certain occupations in Arizona have positive or negative relationships with earnings. The occupation code for elementary and middle schoolteachers had a negative relationship with earnings with a coefficient of -2.794e04. Finding that teachers make less money is not a new concept in our modern society, but it was still fascinating to see it in the data.

There must be other variables that can help account for more variability in earnings, but looking at the original dataset from the US Census is exhausting. There were originally more than 270 attributes to sift through. If this was a project I would continue or revisit in the future, then I would like to contact someone from the US Census in order to get a deeper understanding or at least solidify my own understandings. There were numerous challenges for me in this project. One is the sheer size of the dataset. Due to that, I have attached a “cleaned” version of the dataset. Another challenge was taking the time to clean the data properly. There were times when a certain variable needed to be a different type and that was an arduous process. This project also aimed a bright spotlight on my lack of understanding of modeling distributions. The pareto distribution I have in my code does not feel up to snuff for what I should have taken away from this course.

References:

Fontinelle, A. (2020, July 17). Is student debt worth it? A look at the numbers | MassMutual. MassMutual. https://blog.massmutual.com/post/college-student-debt-worth

US Census Bureau. (2020, September 21). PUMS Documentation. The United States Census Bureau. https://www.census.gov/programs-surveys/acs/microdata/documentation.html