Capstone 1: Project Proposal

Investigating student Loan Data and a Borrowers Ability to Resolve the Debt.

- What is the problem you want to solve?
    - Help students obtain a better understanding of the burden student loans impose.
    - Many people blindly take on student loans without an accurate or realistic understanding of how the debt will be repaid. Many, if not all, students do not know their approximate after college income before starting their education. Therefore it is difficult to predict the effect of student loans once a student enters the repayment period.
- Who is your client and why do they care about the problem? In other words, what will your client do or decide based on your analysis that they wouldn't have done otherwise?
    - The client is any person applying for a student loan or any person currently enrolled in a student loan program. The client expected to take on student loans will have a strong sense of what field of study and line of work they will go into. They will know which city or region they want to live and work in after graduation. The client will know how much of the tuition costs they will be able to pay and how much they are expecting to borrow in the form of student loans.
    - This model will help the client by providing meaningful insight and guidance about choosing a field of study, how much debt they could potentially handle, and if the costs associated with their school of choice are a financially stable investment. The student will be informed about what their expected payments will be once they enter repayment and the progress of other borrowers who are already in repayment. Knowing the likelihood of early repayment or becoming delinquent and defaulting on loans could strongly influence a student's decision to attend a particular school and take on student debt.
- what data are you using?
    - Student Loan Landscape Data from data.world
    - National Student Loan Data System from data.gov
    - Census Income Data Set from UCI
    - average salary by profession: still need to find
- How will you acquire the data?
    - Much of the data is readily available through federal databases. There is a large variety of data available about student loans and the income of various demographics that can be used to describe the current

Capstone 1: Project Proposal

Investigating student Loan Data and a Borrowers Ability to Resolve the Debt.

status of student loan borrowers. Much of the data is available for download in .csv or .xlsx format. The challenge will be identifying key aspects of the data and making meaningful connections between the different resources.

- Briefly outline how you will solve this problem? your approach may change later, but this is a good first step to get you thinking about a method and solution.
    - This project will help students better understand the risk and responsibilities associated with student loan debt.
    - The problem will be solved by analyzing a few key factors that best define a borrower. The model will consider the expected amount borrowed, interest rates, field of study, average salary within the field, years spent in school, cost of living for their region, and if payments will be made before the repayment period.
    - The model will use these parameters to predict the student's expected monthly payment once they enter the repayment period, how long they can expect to be making payments, and what the borrower can afford to pay based on their expected field of study, salary, and cost of living.
    - Additionally, the model will show milestones for a similar demographic who are further along in the repayment period. For example, the borrower can see an approximation of others who are 5 and 10 years ahead in the repayment period. The model will highlight within the borrower's field of study the ratio of borrowers who have successfully resolved their debt vs. borrowers that are delinquent and have defaulted on their loans.
- What are your deliverables? Typically this includes code, a paper, or a slide deck.
    - The deliverables will be a collection of jupyter notebooks, a description of the model, instructions for using the model, expected input specific to the user, descriptions of each graph and plot, a guide for interpreting output from the model, and a presentation of a particular case study that utilizes the model.