

Capstone Project 1 Proposal

I would like to propose a project that will model which areas of the U.S. are becoming more or less impoverished and educated over time.

My intended client would be government policymakers at all three levels (local, state, and federal) that would use my analysis and results in developing public policy. My project would serve as a reference and basis for implementing income/tax credits, subsidies, financial aid, public education programs, and other incentives supporting increases in education levels and eliminating poverty.

The source data I plan to clean and analyze are income statistics publicly provided by the Internal Revenue Service and educational attainment data publicly provided by the U.S. Census Bureau. Calculating these two together as a ratio would measure how the overall standard of living changes over time.

After cleaning the data, I would predict areas of growing/contracting economic inequality by plotting the education and income based descriptive statistics over time for each zip code. Additionally, using K-Means Clustering I would attempt to segment the data and draw insights from trends over a six year period.

To explain my process and findings, I will submit a Jupyter notebook outlining the steps/lines of code leading up to generating the final visualization, a paper explaining all the steps and results taken to generate the project, and a slide deck that is a brief, summarized version of my paper.