Group 1 Capstone Executive Summary: Predicting Individual Economic Health and Quality of Life with Financial Indicators

Overview

Broadly speaking, this project aims to use a variety of and financial factors to forecast different measures of economic and societal well-being. Our data, drawn primarily from the World Bank databank, draws upon different measures including government debt, tax revenue, stock value within an economy, inflation, and GDP, among others, to predict values such as poverty, unemployment, wealth disparity, life expectancy, and infant mortality rate.

Using an automated API call from the World Bank, we were able to collect data from every registered country from the years 1990-2020. The goal of our analysis is to be able to produce actionable results by building models and visualizations. We can then use the data from one country in one particular year to predict whichever dependent variable we choose for the following year in that country. For example, we could use GDP, central government debt, tax revenue, and total value of stocks traded in the United States in 2000 to predict the unemployment rate in the United States in 2001. With this goal in mind, by the end of our process, we should be able to create predictions on a global scale for how these socioeconomic dependent variables will end up at the end of 2021 using the 2020 data. Additionally, these predictions may also be useful for filling in missing data in the World Bank's databank itself.

Guiding Questions

1. How well can quality of life for individuals be predicted by large scale economic indicators?
2. How well do large scale economic indicators predict wealth inequality and poverty?
3. Can short-term economic data be used to effectively predict quality of life indicators?
4. What are the most important economic indicators for predicting quality of life and wealth inequality?
5. Can countries be grouped into clusters based on different economic attributes?
6. How does wealth inequality vary between countries (which countries have the greatest and lowest average levels of inequality)?
7. Can we apply some of these same questions we are posing on a global scale to a national and state scale for the United States using census data?