

# Predicting Economic Health with Financial Indicators

Luke Everson, Elwood Olson, Daniel Moultrie, and Douglas Byers

# Topic Overview

- The goal of this project is to use large-scale financial indicators to estimate and predict a country's economic health for individuals.
- For example, using indicators such as a country's GDP per capita and its total value of stocks traded to attempt to predict metrics like the country's unemployment rate.

# Exploratory Questions

- How well can quality of life for individuals be predicted by large scale economic indicators?
- How well do large scale economic indicators predict wealth inequality?
- Can short-term economic data be used to effectively predict quality of life indicators?
- What are the most important economic indicators for predicting quality of life and wealth inequality?
- Can we accurately fill in missing economic data such as the Gini Index from the World Bank using financial indicators?

# Research Process

- World Bank Databank
  - Researched various indicators to select input and response variables
  - Used the World Bank Free API to pull data
- Selected financial and socioeconomic factors that we deemed to be related as input and response variables



# Machine Learning Desired Outcomes

- **Gini Index**
  - Fill in missing values for the World Bank Site
- **Unemployment**
  - Create a model to predict unemployment rate for next year

# Machine Learning: Variables

## EXPLANATORY VARIABLES

- Stocks traded, total value (% of GDP)
- Inflation, consumer prices (annual %)
- Tax revenue (% of GDP)
- GDP per capita (current US\$)
- Broad money (% of GDP)
- Total reserves (includes gold, current US\$)
- Exports of goods and services (% of GDP)

## RESPONSE VARIABLES

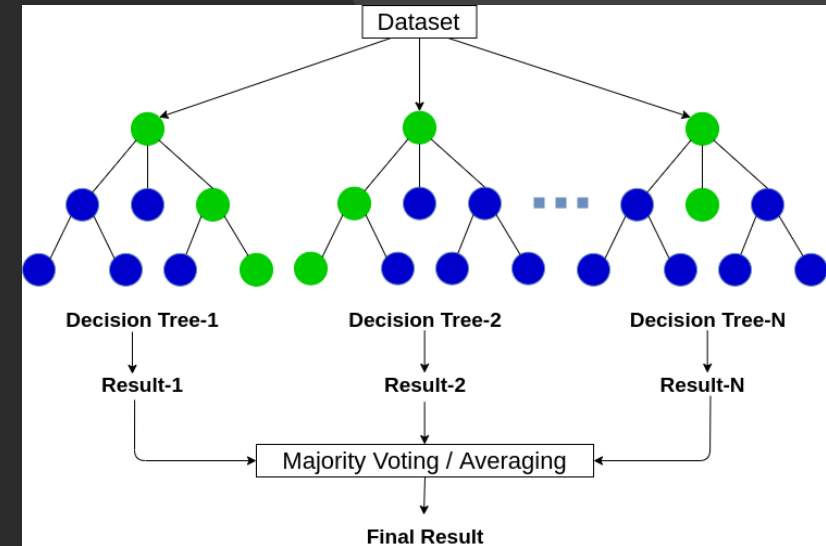
- Gini Index
- Unemployment, total (% of total labor force) (year  $n + 1$ )

# Key Term Definitions

- Gini Index - a statistical measure of a country's income inequality. A score of 100 indicates total income inequality, while a score of 0 indicates perfect income equality.
- Broad Money - the total amount of money in a national economy, including liquid money and other assets
- Total Reserves - the total of a country's holdings of monetary gold, special drawing rights, and other IMF reserves

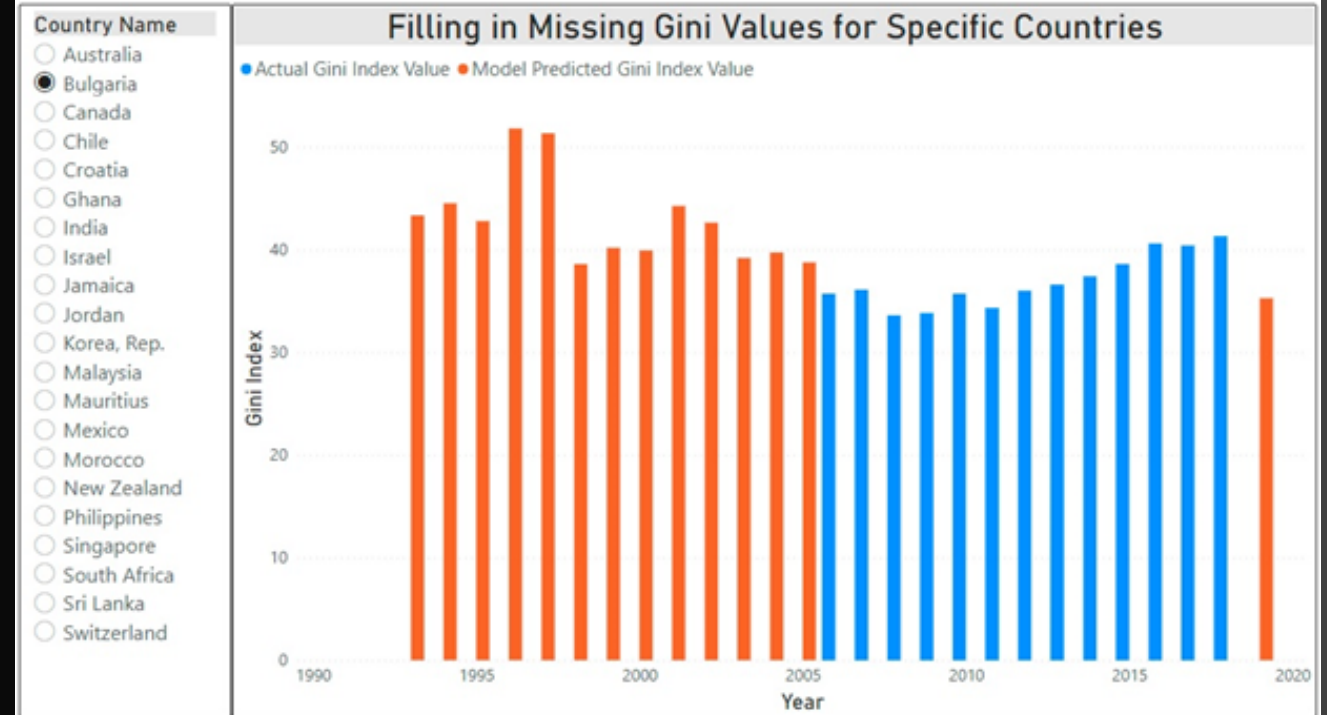
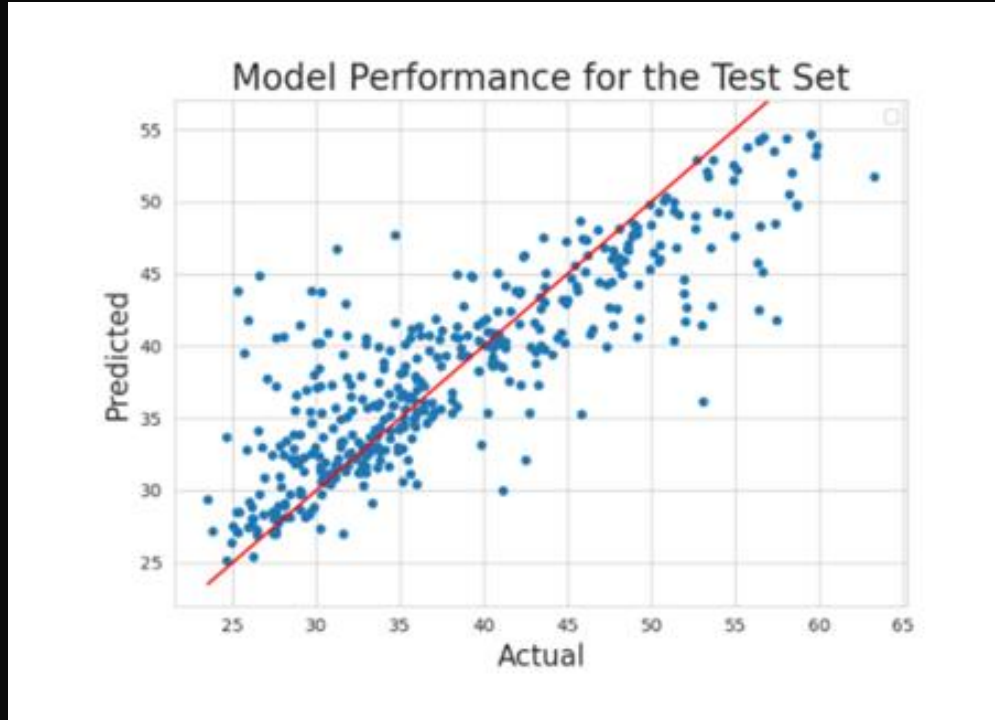
# Machine Learning: Random Forest Model

- Filtered down dataset to include only rows with a value for our desired response variable
- Used a KNN Imputing process to fill in gaps in predictive variables
- Used random forest regression, an ensemble machine learning model based on iterating decision tree machine learning
- Hyper-tuned the models using random-search and grid-search consecutively





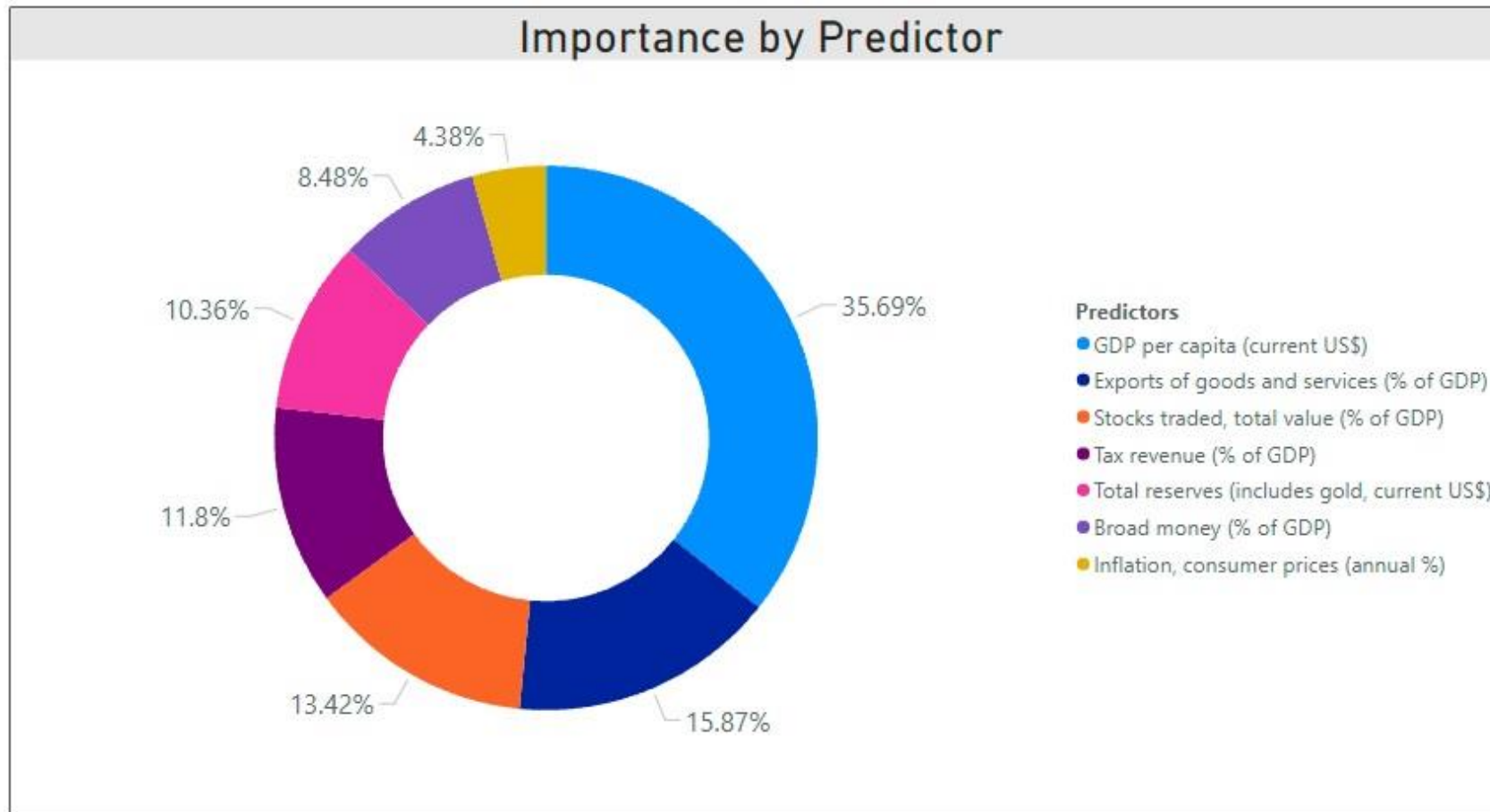
# Machine Learning: Gini Index



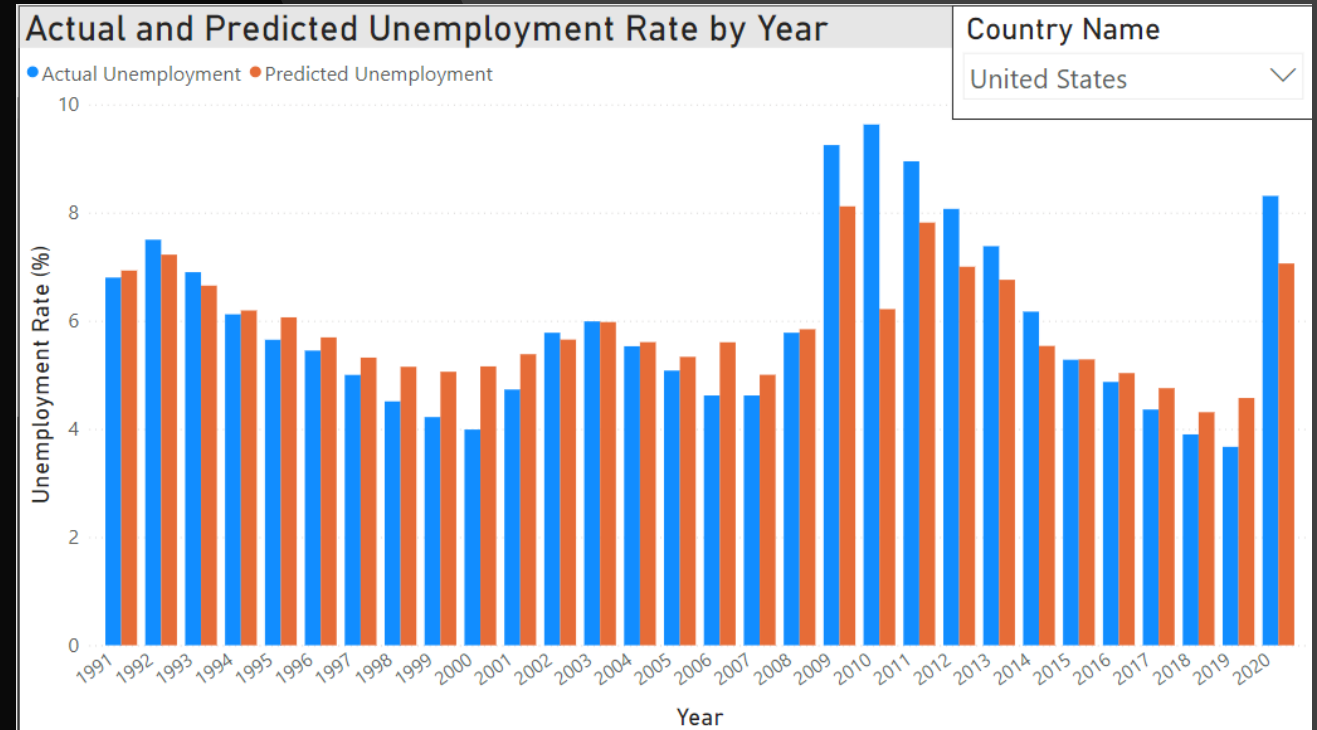
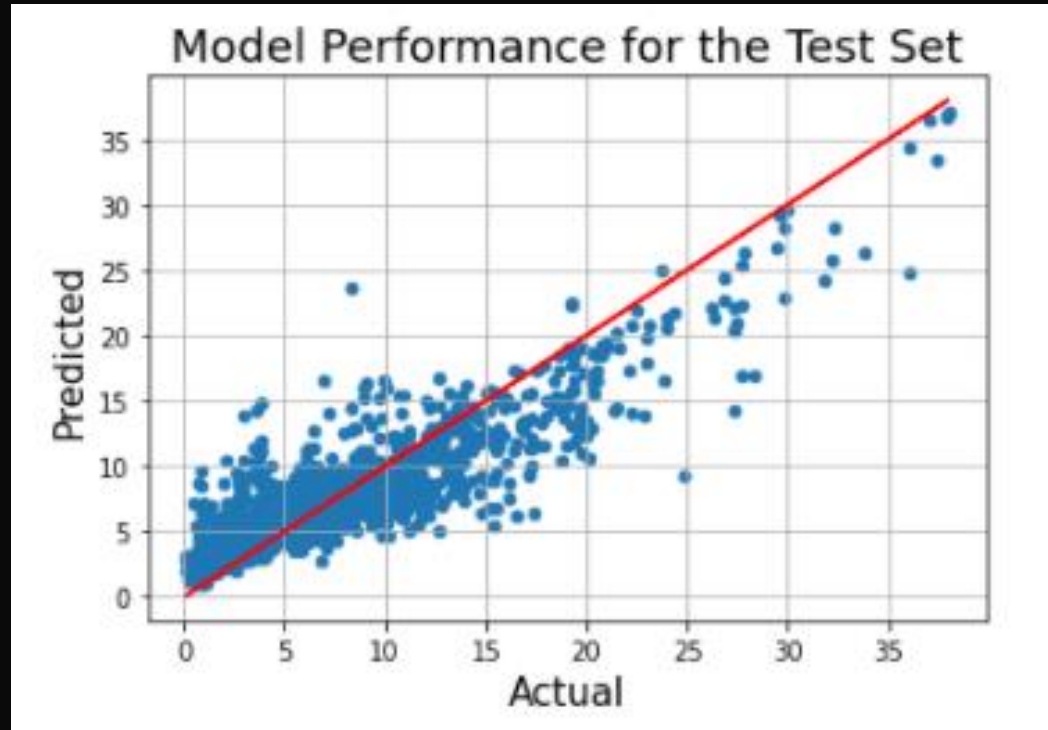
Model accuracy: 0.71

Most important feature: GDP per Capita – 35%

# Machine Learning: Gini Index



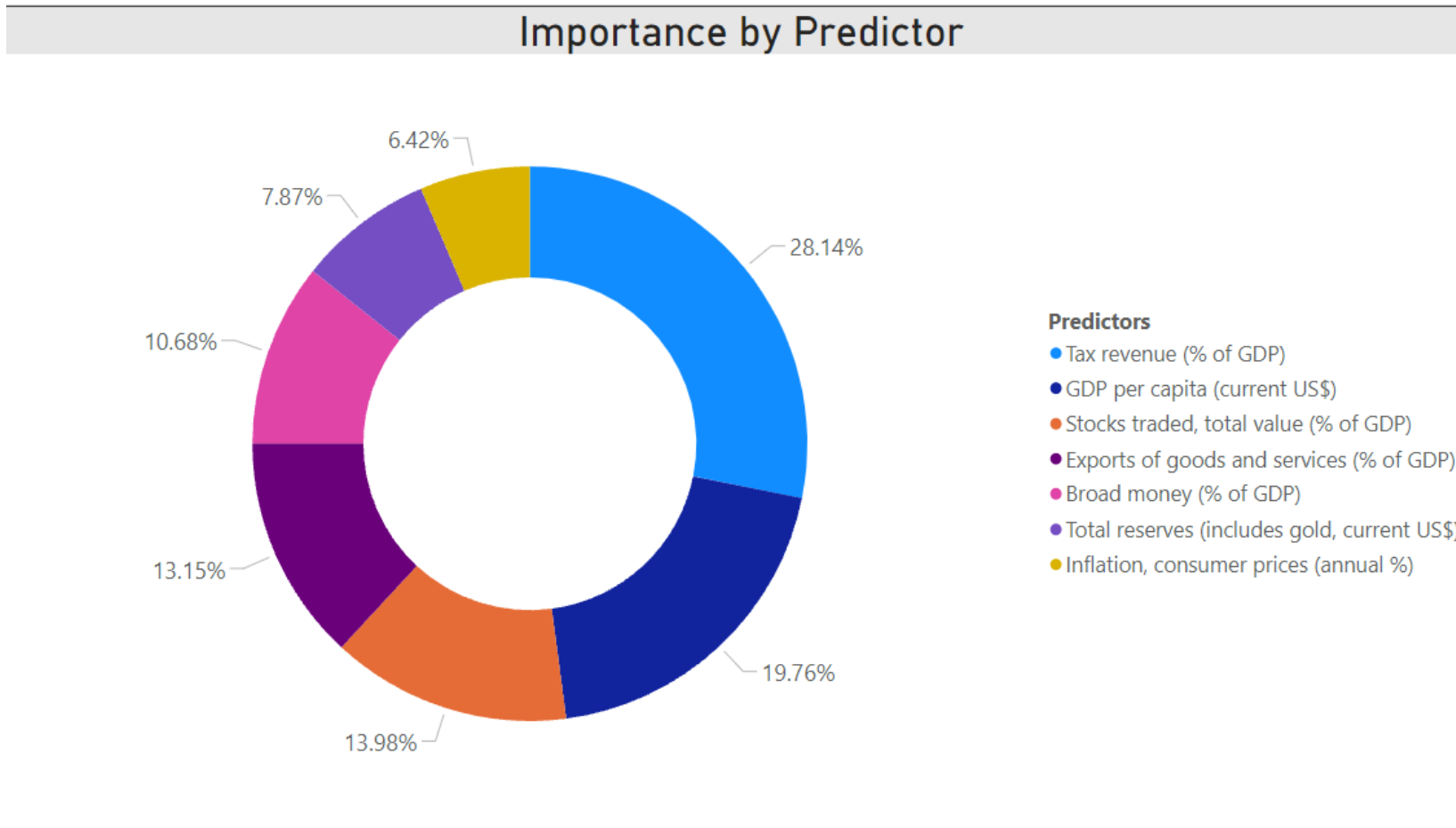
# Machine Learning: Unemployment Rate



Model Score: 0.76

Most Important Feature: Tax revenue (% of GDP) - 28%

# Machine Learning: Unemployment Rate



## Takeaways and Recommendations

- Although our models are imperfect, short-term financial data is remarkably effective at predicting large-scale economic and quality of life indicators
  - Many of the errors seem to come from major social and political discontinuities, events that all analysts struggle to predict.
- We recommend that the World Bank develop a method of filling in blank data like the one we created. Such projections could provide businesses, social scientists, and policy makers with a wealth of otherwise inaccessible information.

# Sources

- [World Bank Indicators](#)
- [Guide to Random Forests](#)
- [Random Forest Documentation](#)
- [Grid Search/Random Search Tutorial](#)
- [Shifting the Data to Create Leading Variables](#)