Project Proposal

Andrew Kroening, Pooja Kabber, Dingkun Yang, Echo Chen

2022-10-03

Research Proposal

Selected Dataset

The team selected the World Health Organization's Life Expectancy Dataset. This dataset contains observations of 193 countries in 20 variables, across a roughly ten year period.

Research Question #1

Research Question #1 will be an inference question. Our group will explore Life Expectancy in the year 2014 as a function of these possible variables:

- *Percentage Expenditure*
- *Hepatitis B*
- *Measels*
- *Polio*
- *HIV/AIDS*
- *Diptheria*
- *GDP*
- *Schooling*
- *Income Composition of Resources*
- *Country Status (Developed/Developing)*
- *Population*

The specific question we will be answer is: "How did major disease, economic, and social factors impact life expectancy around the globe in 2014?"

Research Question #2

Research Question #2 will be predictive. Our group will seek to predict a coutry's category as developed or developing based on the best-fit model.

The specific question we will answer is: "Can observations of life expectancy-related data such as mortality rates and disease statistics, along with country-specific properties, be used to accurately predict a country's status as developed or developing?"