Project Proposal

The idea I have in mind is to predict demographic trends based on human geography factors such as birth rate, death rate, replacement rate, as well as other statistics like fertility rate, diseases, and sanitation accessibility. By this, I want to predict population growth and other factors by using the [CIA World Factbook Dataset](https://www.kaggle.com/datasets/lucafrance/the-world-factbook-by-cia). This dataset contains population growth trends as well as data for several factors that affect population growth in various countries. There are many different factors to consider, so the exploratory data analysis here will try to draw conclusions and comparisons between each factor and how it explains the demographic trends of a country.

In addition to this, several machine learning algorithms can be used to do this prediction – linear regression, support vector regression, k-nearest neighbors (to predict data points), trees and random forests, etc. This will help me have a variety of algorithms so that I can compare them and conclude which algorithm performs the best and gives a good prediction for the growth trends in a given country.  
 I would like to evaluate each algorithm and its performance using criteria like rmse, f1-score, confusion matrices, etc. I would love to get feedback here on what other criteria I can use to compare the performances of algorithms as well. As for tools, I plan to use Python and some libraries to help me load the dataset and do data analysis on them. There are also a variety of online resources that I would like to utilize to help me understand how to use each algorithm as well as learn how to use their results to conduct a meaningful data analysis.