

Global Life Expectancy 2000-2020 WHO

By Paul-Moritz Jost for Ironhack Berlin

Goal:

Predict Life Expectancy and its Politico-Economic Antecedents

Global Life Expectancy 2000-2020 WHO

- 1) Introduction to Dataset**
- 2) Exploratory Data Analysis**
- 3) Target Variable Regression**
- 4) Power Law Distribution**
- 5) Prediction by Classification**
- 6) Practical recommendations**

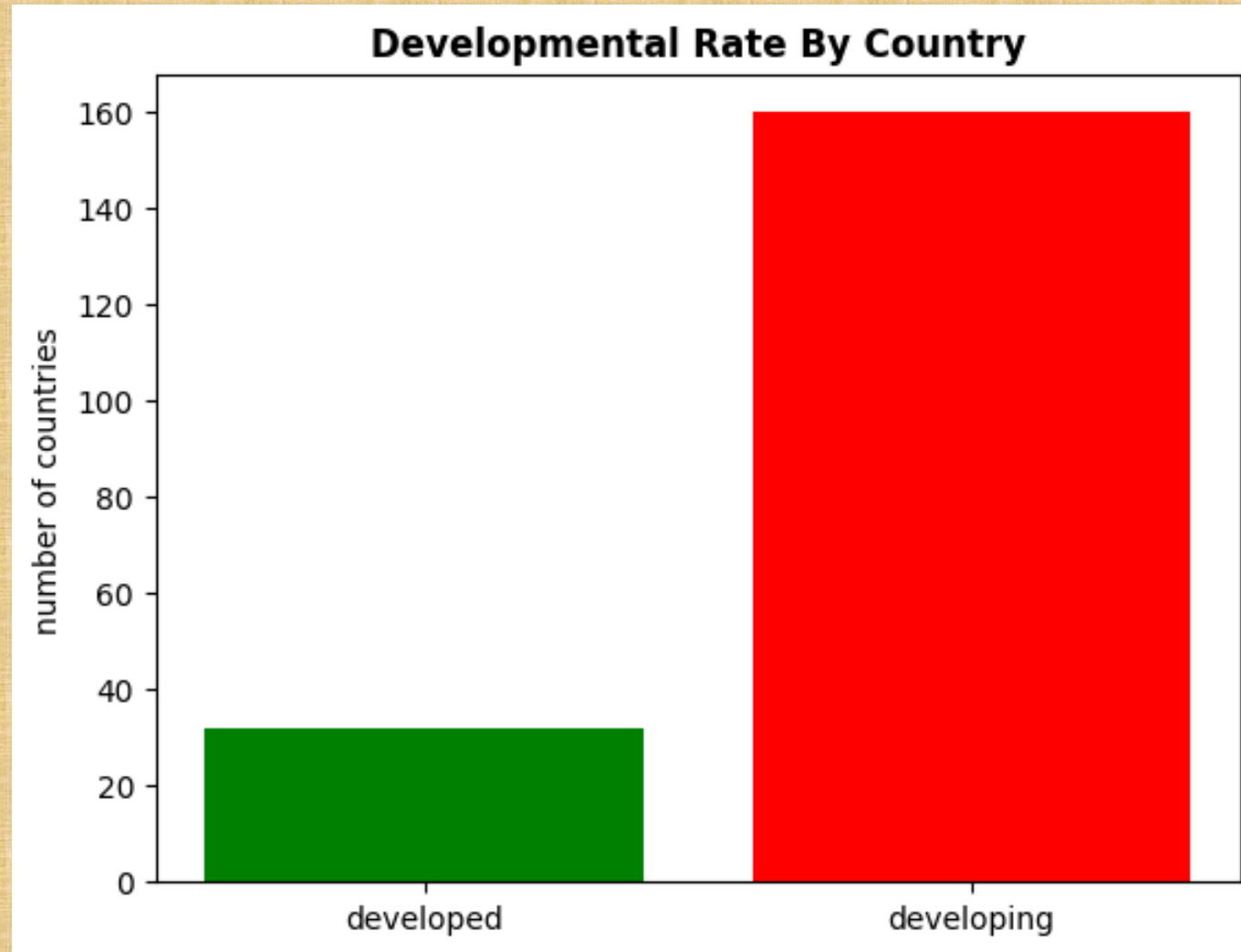
1.) Introduction to Dataset

- 1) 193 Countries
- 2) 20 Years
- 3) Life Expectancy
- 4) Disease
- 5) Population Size
- 6) Economic (In-)Equality
- 7) Schooling

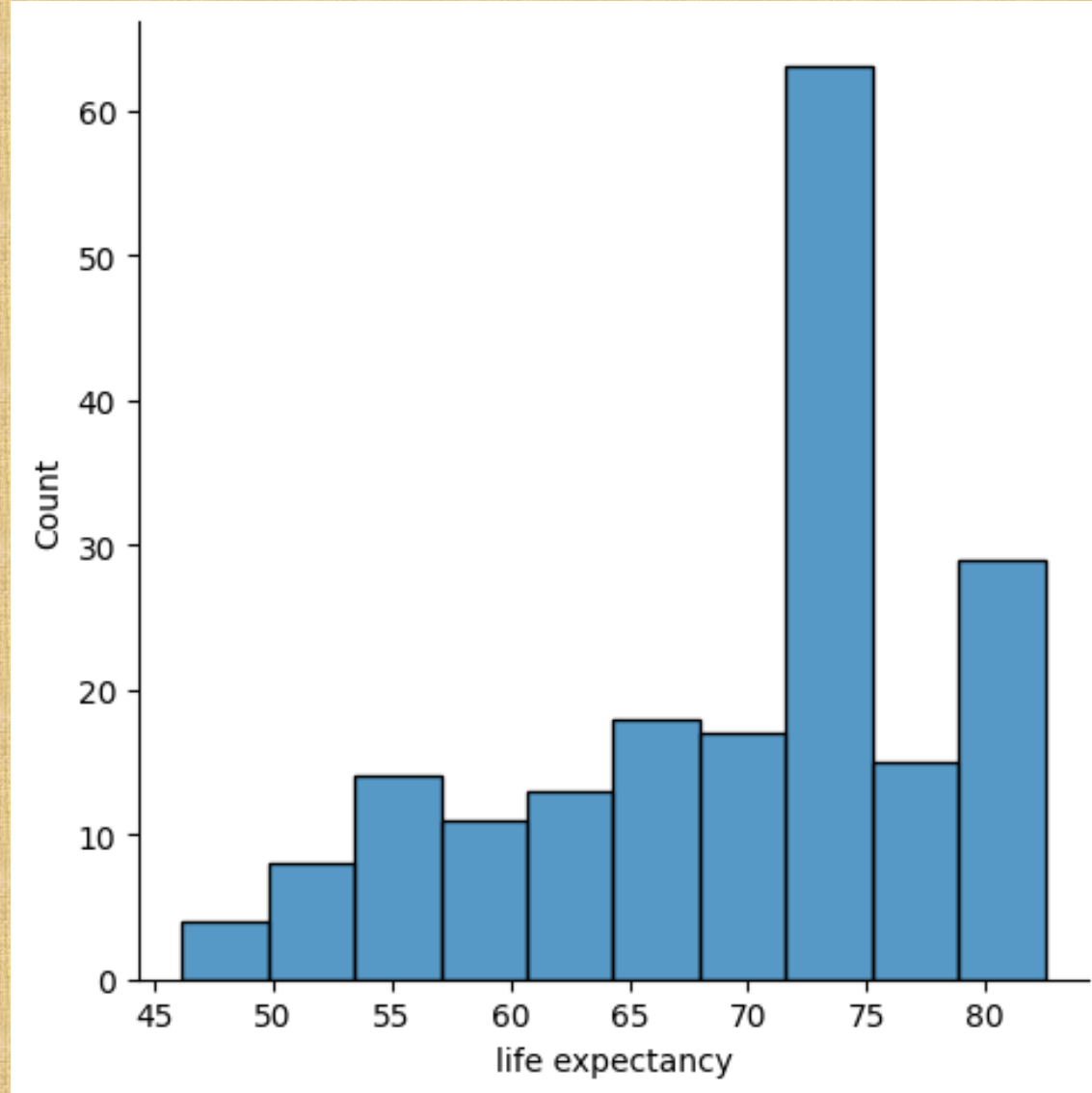
2.) Exploratory Data Analysis

- 1) 22 features
- 2) 2 categorical features, 20 numerical
- 3) 2938 entries
- 4) NaN's max 22.19% (population)
- 5) No duplicates

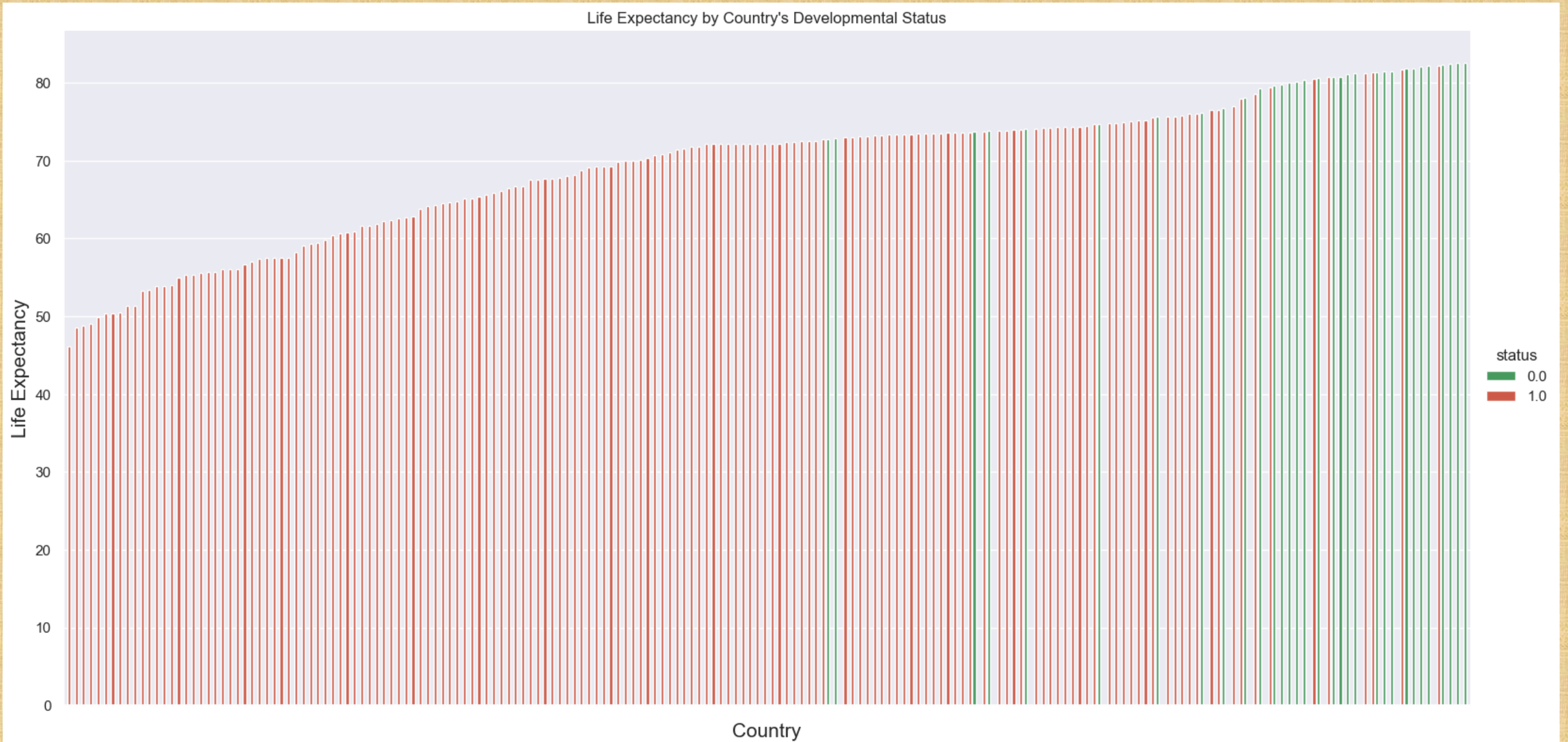
2.) Exploratory Data Analysis



2.) Exploratory Data Analysis



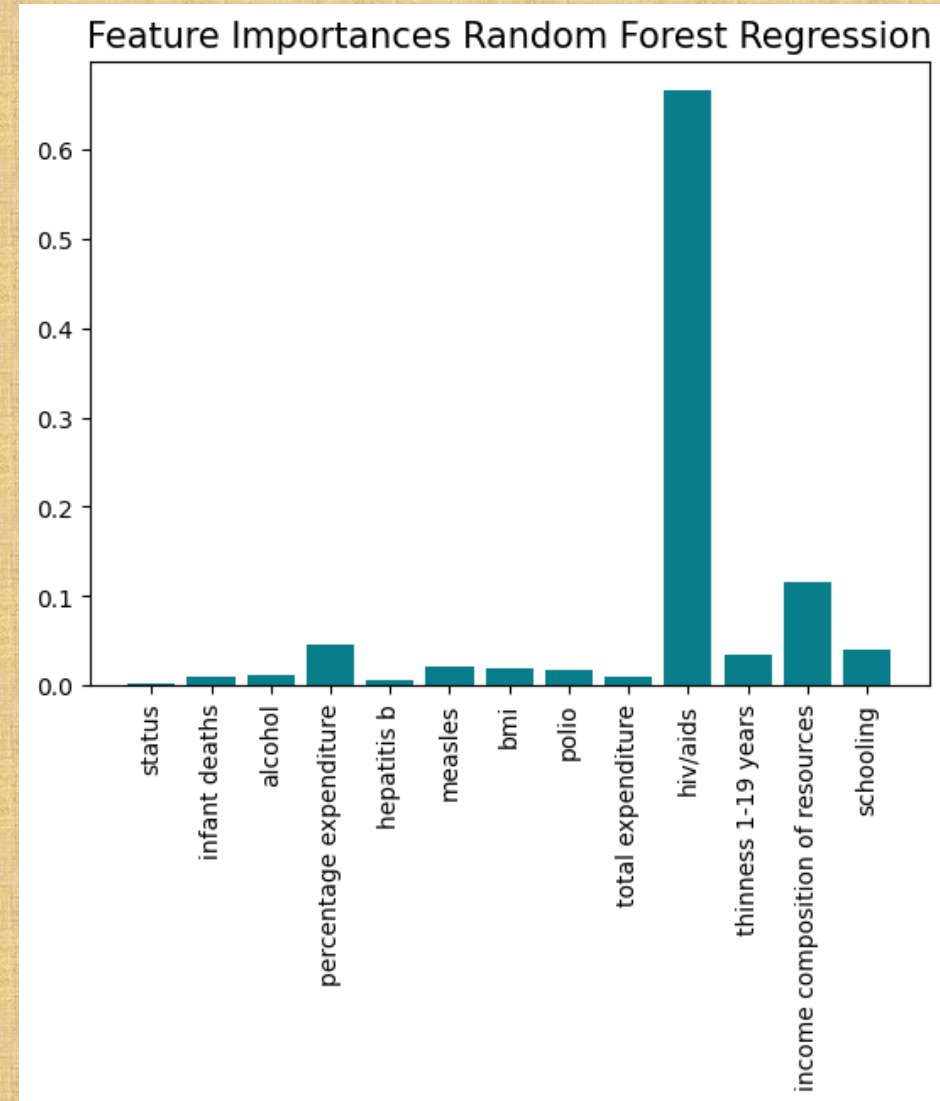
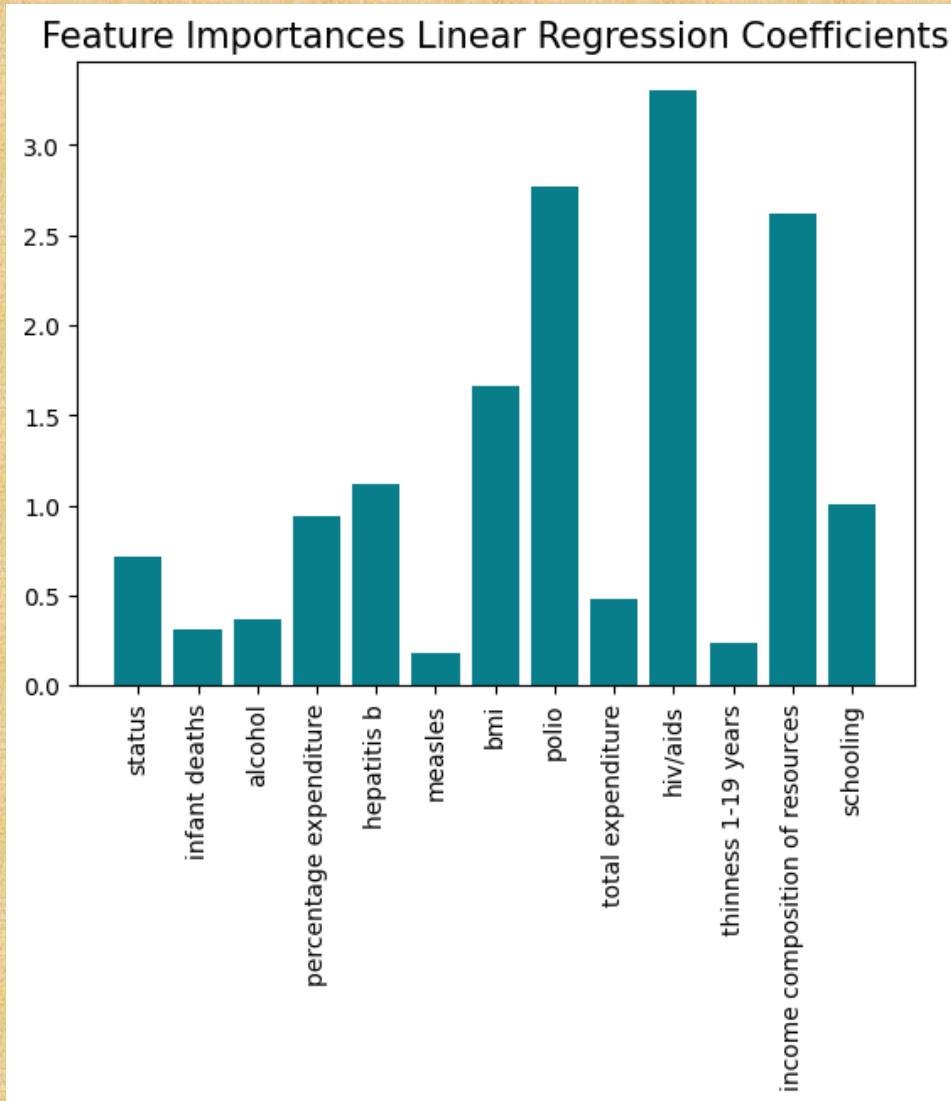
2.) Exploratory Data Analysis



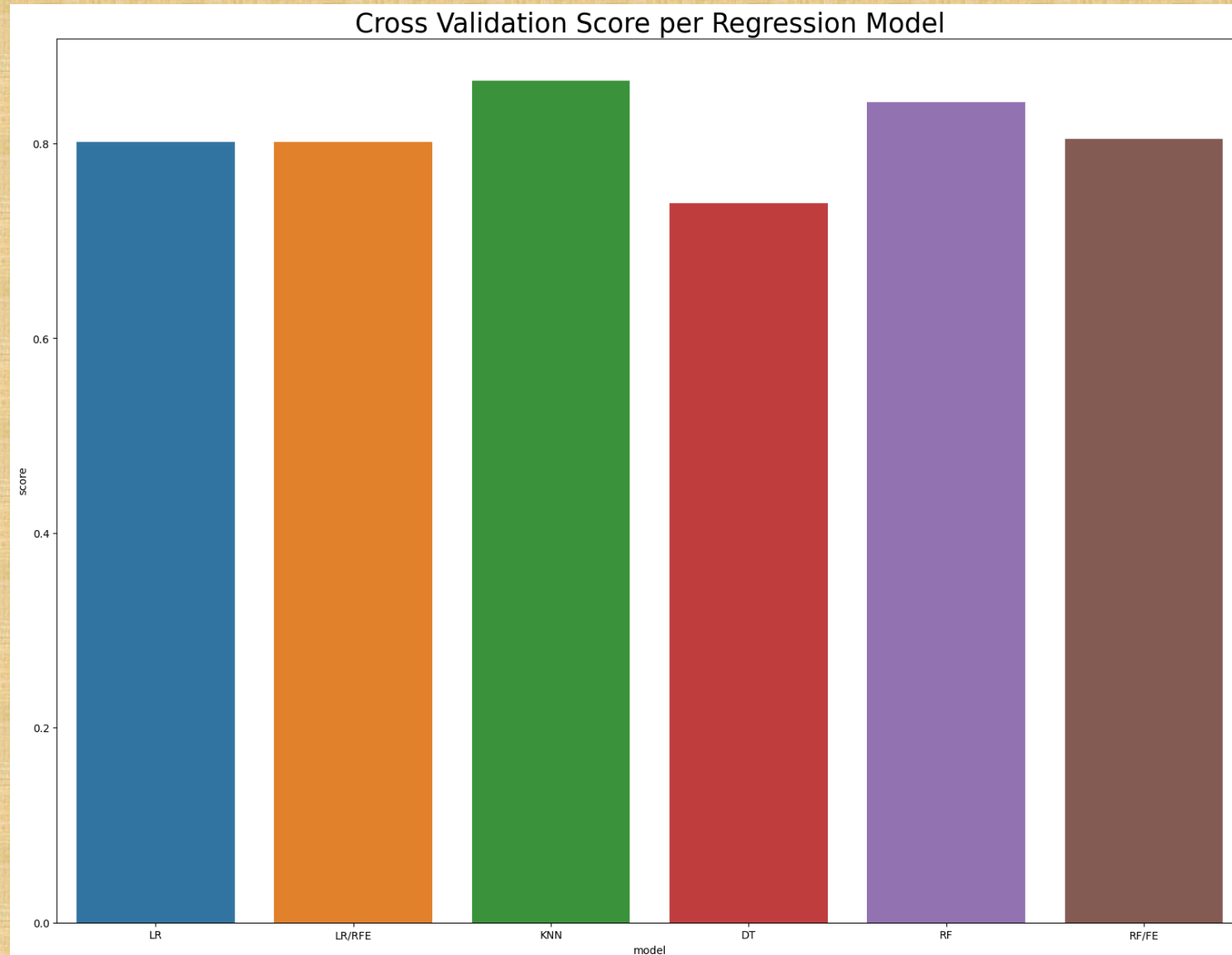
3.) Target Variable Regression

- 1) Linear Regression
- 2) KNN
- 3) Decision Tree
- 4) Random Forest

3.) Target Variable Regression

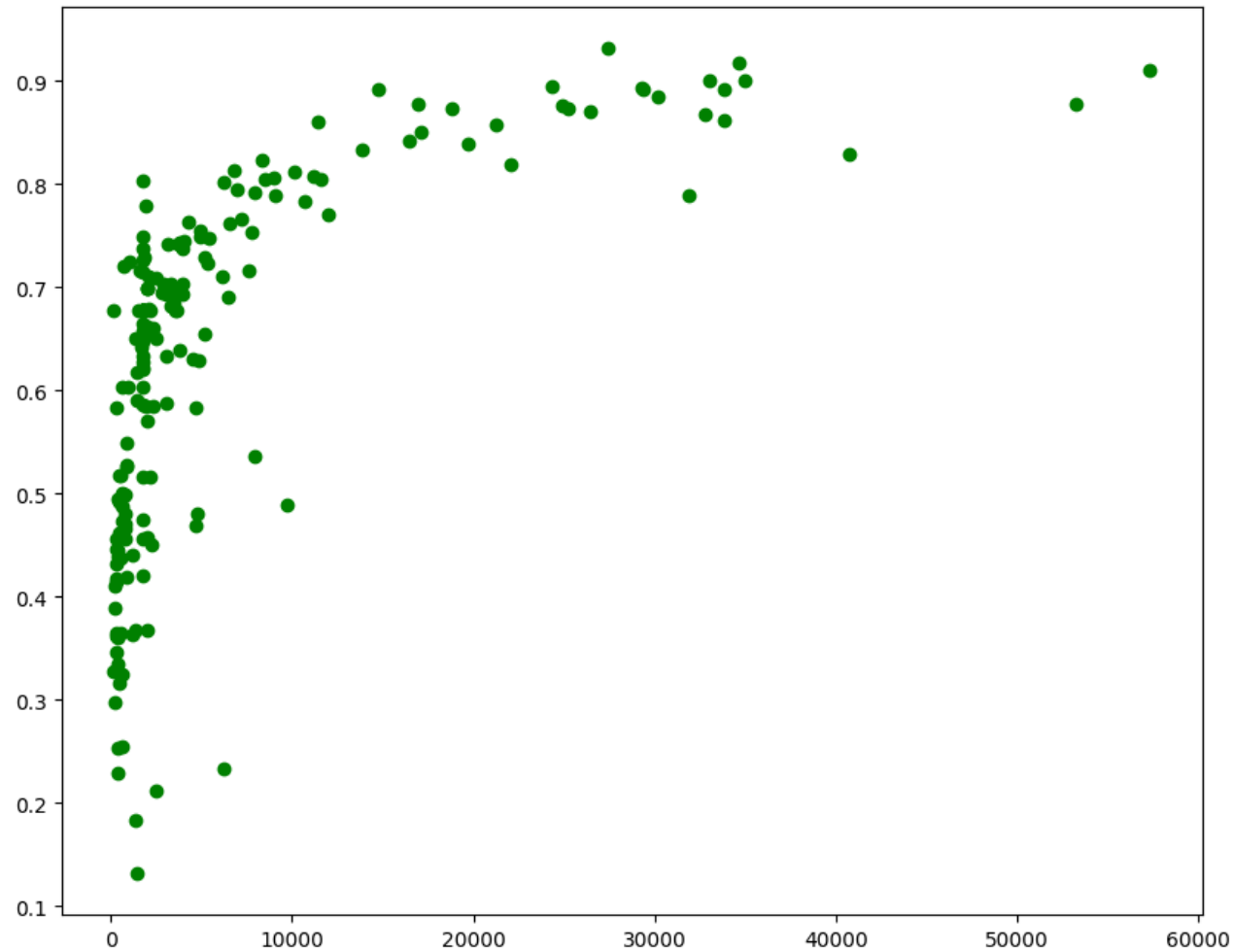


3.) Target Variable Regression

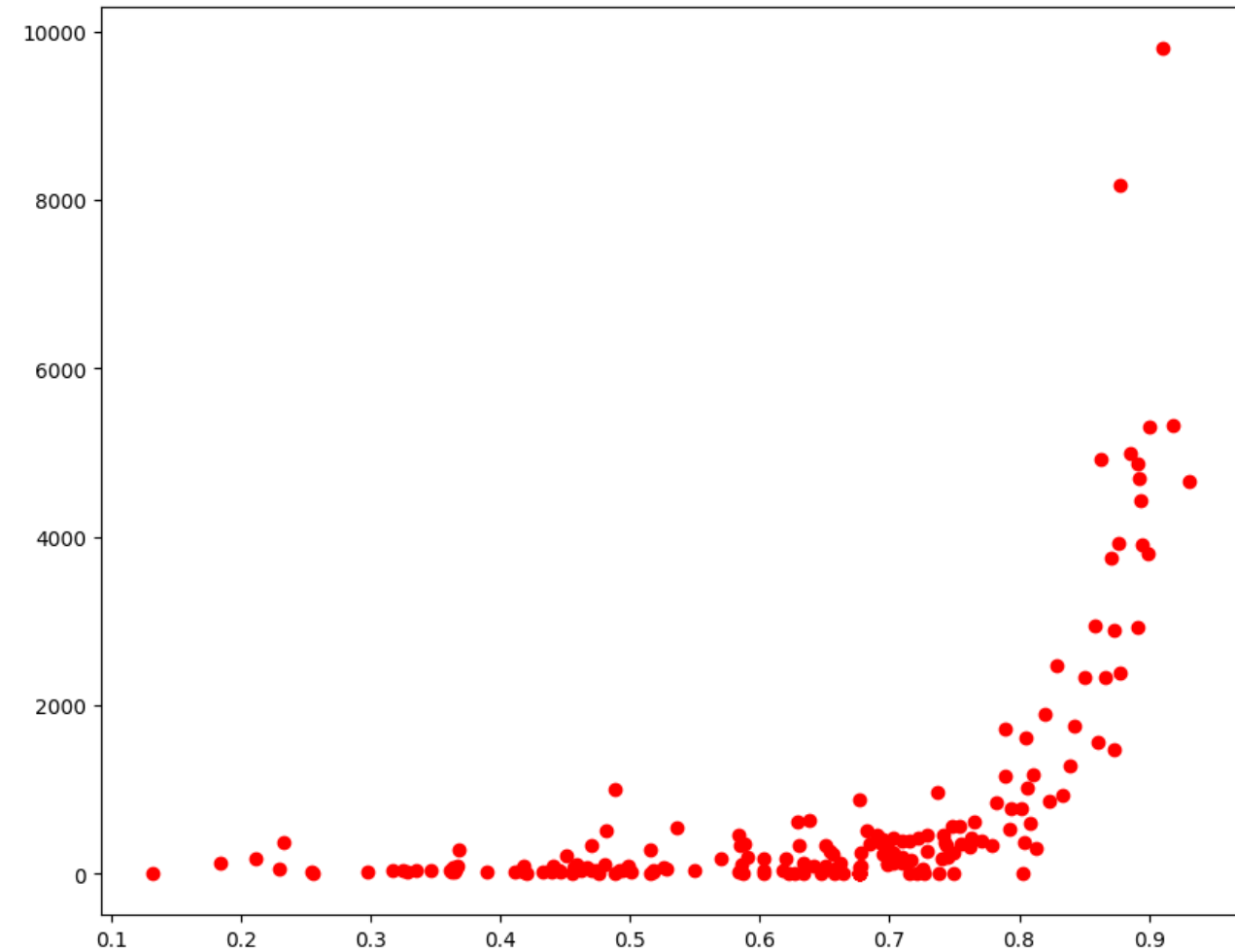


4.) Power Law Distributions

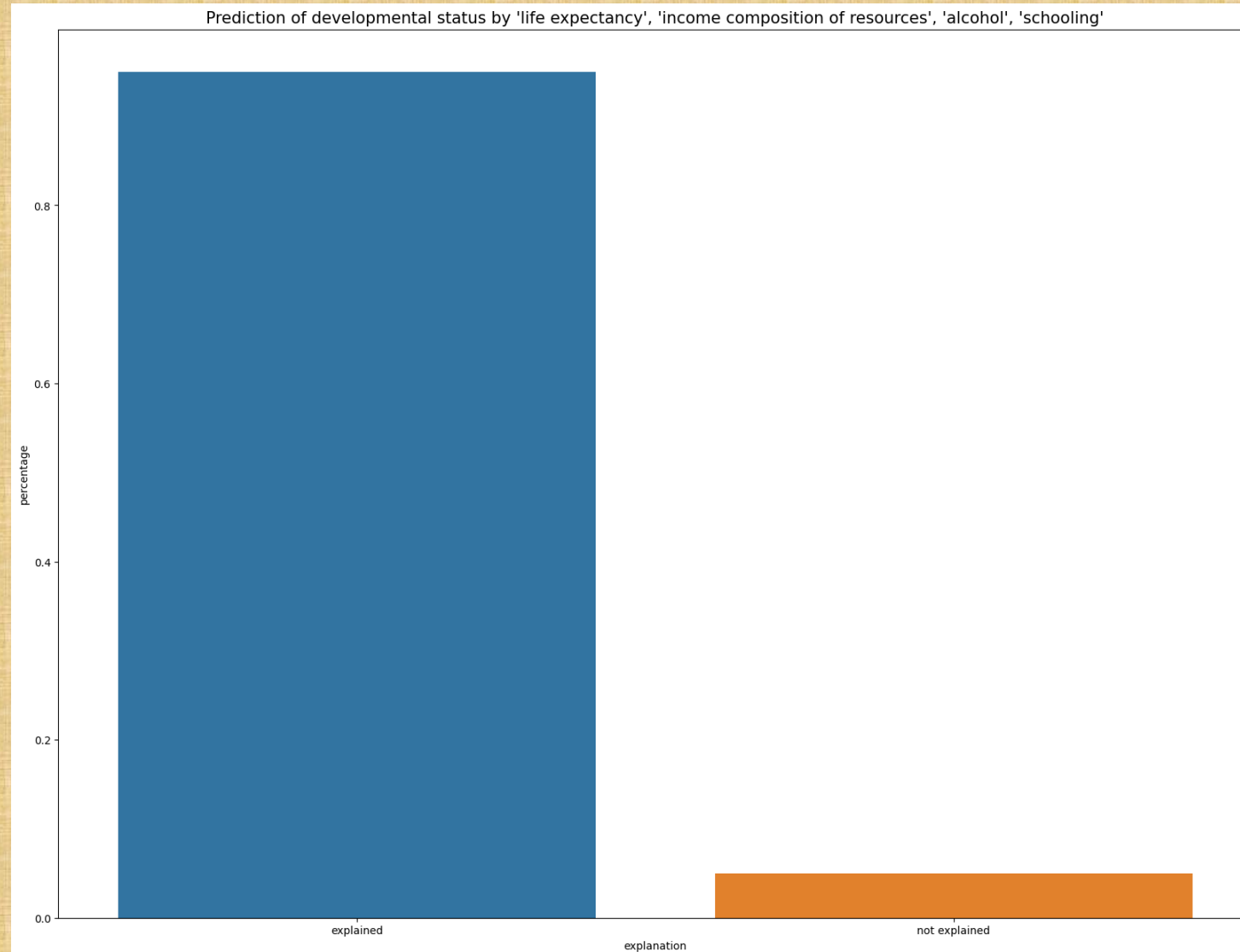
income composition of resources by gdp



percentage expenditure by income composition of resources



5.) Prediction by Classification



6.) Practical Recommendations

- 1) **Gathering data that's less skewed towards developing countries**
- 2) **Explore reasons for high life expectancy in certain developing countries**
- 3) **Compare algorithmic architecture of ML models to politico-economic goals**
- 4) **Relate data of life expectancy to data of wealth distribution within countries**
- 5) **Gather an exponential regressor compatible with scikit-learn**

Thank you for listening!