Life Expectancy Data Analysis Project

Analyzed By : Pulkit Mehrotra

E-mail Id: pulkitmehrotra246@gmail.com

Project Overview

The goal of this analysis is to **understand the key factors influencing life expectancy** across different countries and predict life expectancy values using relevant socio-economic, demographic, and health indicators.

Target variable:

Life expectancy

Key Questions:

- What features most strongly affect life expectancy?
- How do economic indicators (like GDP, schooling, income composition) impact it?
- Can we build a predictive model to estimate life expectancy?

Technology Stack Used

MS Excel

For Data Gathering

Python | Pandas | Seaborn | NumPy | Scikit-learn | Matplotlib

For Data Extraction, Cleaning, Manipulation, EDA, Feature Engineering, Model Selection, Training, Hyperparameter Tuning

Data Overview

Column	Description
Country	Name of the country
Year	Year of record
Status	Developed / Developing
Life expectancy	Average number of years a newborn is expected to live
Adult Mortality	Probability of dying between 15 and 60 years per 1000 population
Infant deaths	Number of infant deaths per 1000 population
Alcohol	Alcohol consumption per capita (litres)
BMI	Average body mass index of population
HIV/AIDS	Deaths due to HIV/AIDS per 1000 population
GDP	Gross Domestic Product per capita (in USD)
Schooling	Average number of years of schooling
Income composition of resources	Index (0–1) reflecting income equality and resources

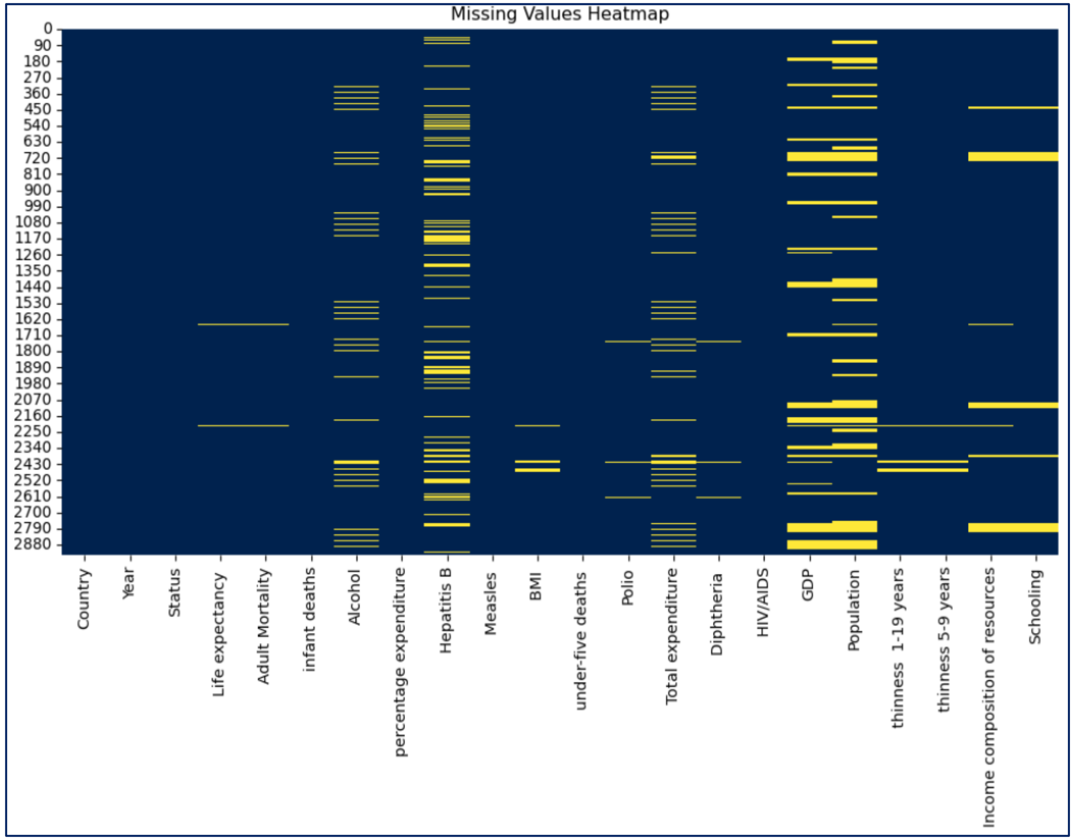
Dataset Overview

- Time Period: 2000-2015
- Countries Covered: 193
- Total Rows: 2,938
- Total Columns: 22

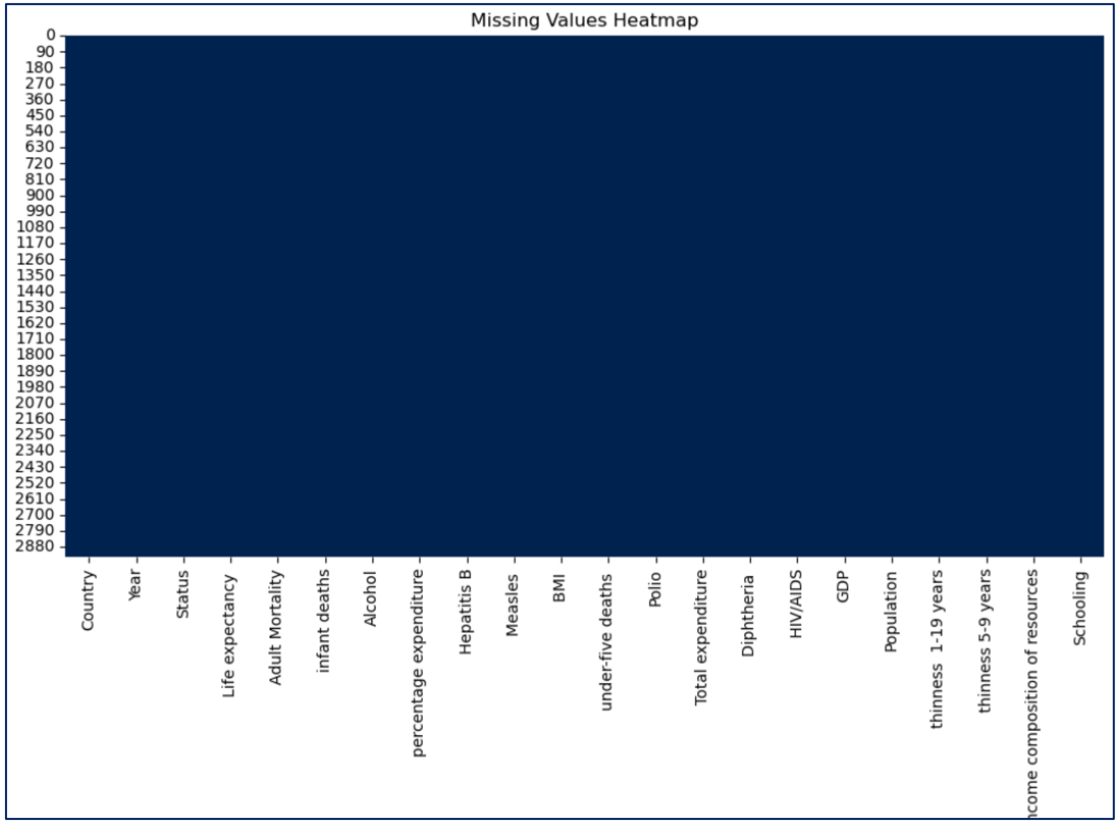
Category	Features
🏥 **Immunization Factors**	Hepatitis B, Polio, Diphtheria Coverage (%)
⚰️ **Mortality Factors**	Infant Mortality, Adult Mortality Rates
💰 **Economic Factors**	GDP, Healthcare Expenditure (%)
🏫 **Social Factors**	Education, Alcohol Consumption, Smoking, Exercise

Data Cleaning

Dataset with Missing Values

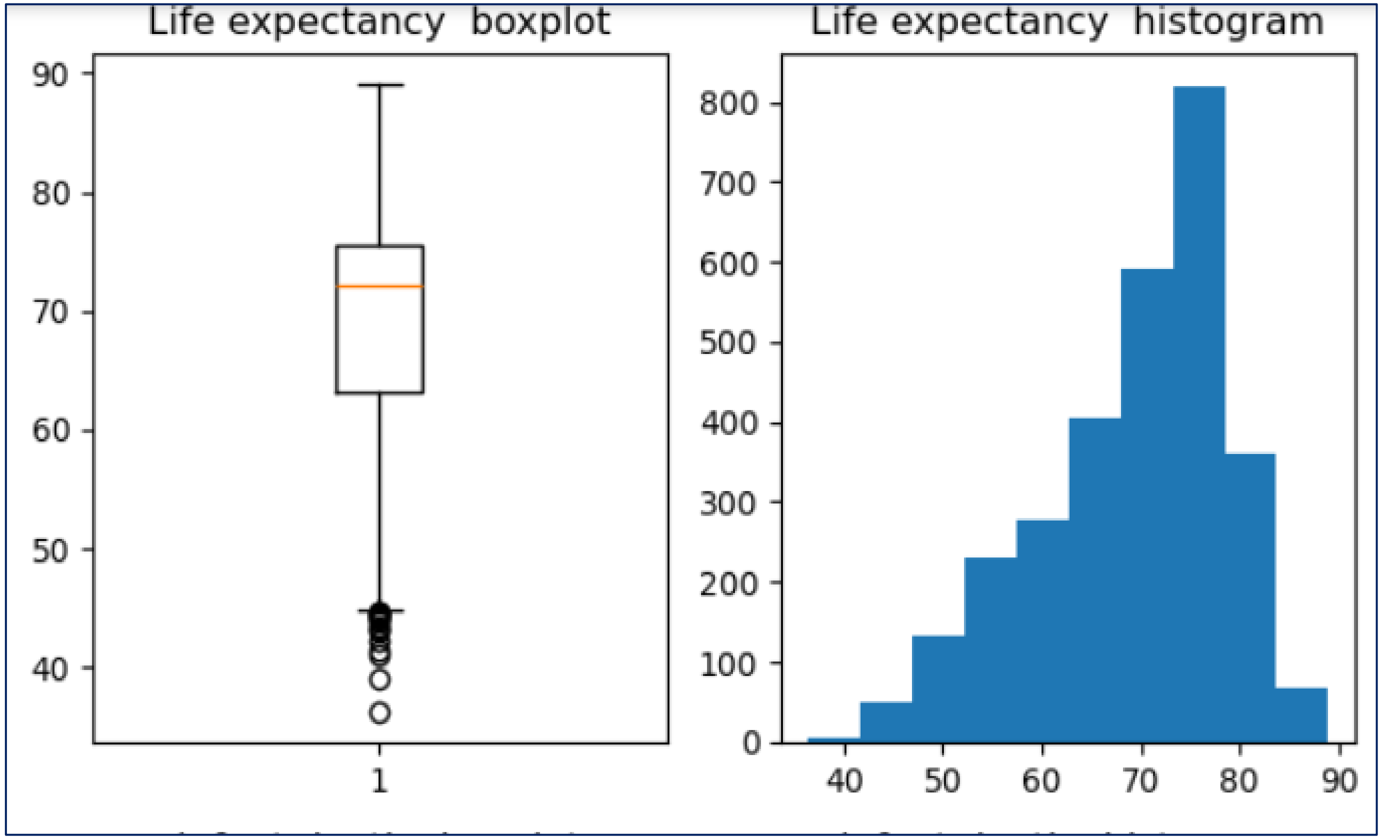


Cleaned Dataset



All numeric missing values were replaced with the median of that column (robust to outliers).

Outliers Detection

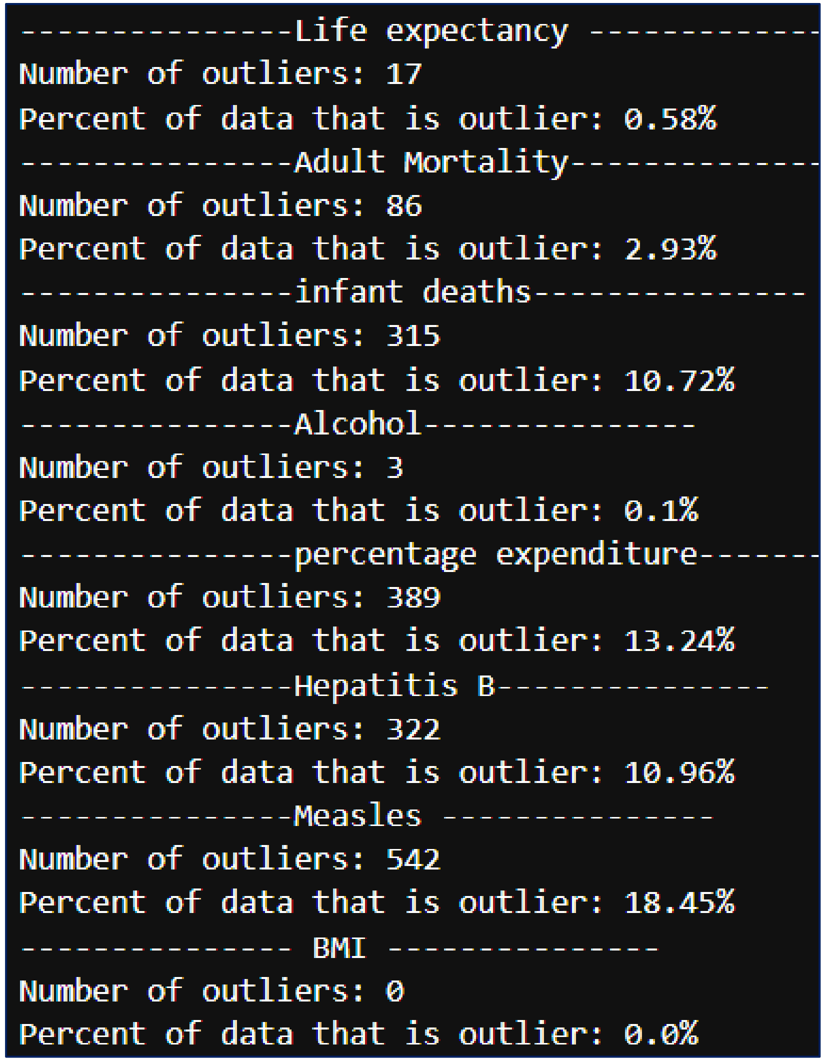


Visually, it can be seen that there are a number of outliers for all of the variables - including the target variable, life expectancy.

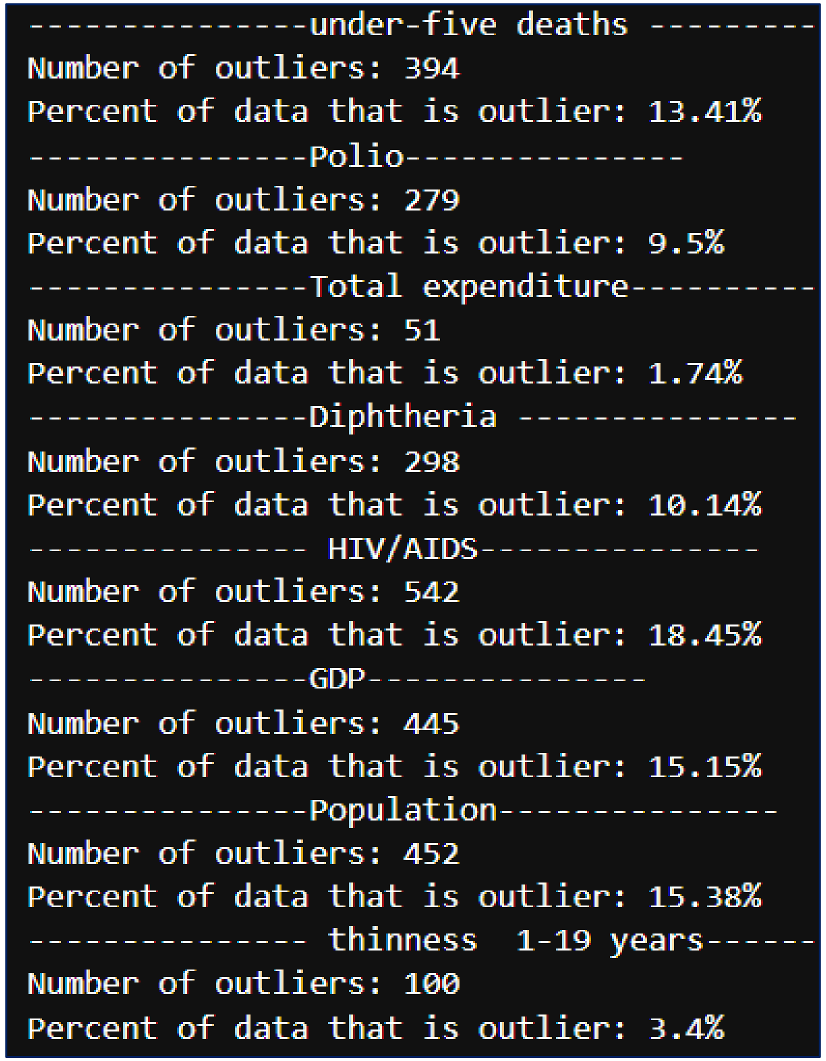
Insights:

- Median life expectancy ≈ 72 years
- Most countries have life expectancy between 65–80 years.
- Outliers exist below 50 years (low-health regions).
- Distribution is mostly normal but slightly right-skewed.

Outliers Detection



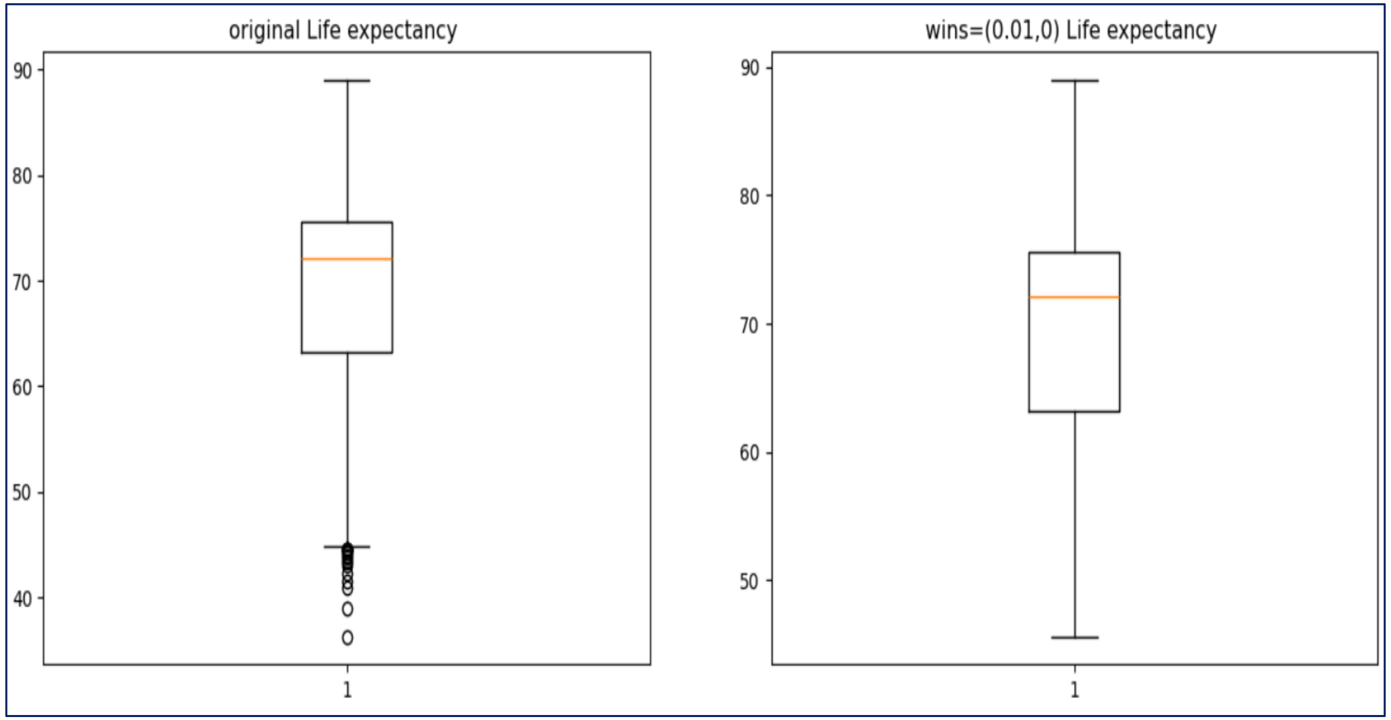
```
---------------Life expectancy -------------
Number of outliers: 17
Percent of data that is outlier: 0.58%
---------------Adult Mortality--------------
Number of outliers: 86
Percent of data that is outlier: 2.93%
---------------infant deaths---------------
Number of outliers: 315
Percent of data that is outlier: 10.72%
---------------Alcohol---------------
Number of outliers: 3
Percent of data that is outlier: 0.1%
---------------percentage expenditure-------
Number of outliers: 389
Percent of data that is outlier: 13.24%
---------------Hepatitis B---------------
Number of outliers: 322
Percent of data that is outlier: 10.96%
---------------Measles ---------------
Number of outliers: 542
Percent of data that is outlier: 18.45%
--------------- BMI ---------------
Number of outliers: 0
Percent of data that is outlier: 0.0%
```



```
---------------under-five deaths ---------
Number of outliers: 394
Percent of data that is outlier: 13.41%
---------------Polio---------------
Number of outliers: 279
Percent of data that is outlier: 9.5%
---------------Total expenditure----------
Number of outliers: 51
Percent of data that is outlier: 1.74%
---------------Diphtheria ---------------
Number of outliers: 298
Percent of data that is outlier: 10.14%
--------------- HIV/AIDS---------------
Number of outliers: 542
Percent of data that is outlier: 18.45%
---------------GDP---------------
Number of outliers: 445
Percent of data that is outlier: 15.15%
---------------Population---------------
Number of outliers: 452
Percent of data that is outlier: 15.38%
--------------- thinness  1-19 years------
Number of outliers: 100
Percent of data that is outlier: 3.4%
```

- Another method to detect outliers is done statistically using Tukey's method where - outliers being considered anything outside of 1.5 times the IQR.
- Through this, it appears that there are a decent amount of outliers in this dataset.

Dealing with Outliers



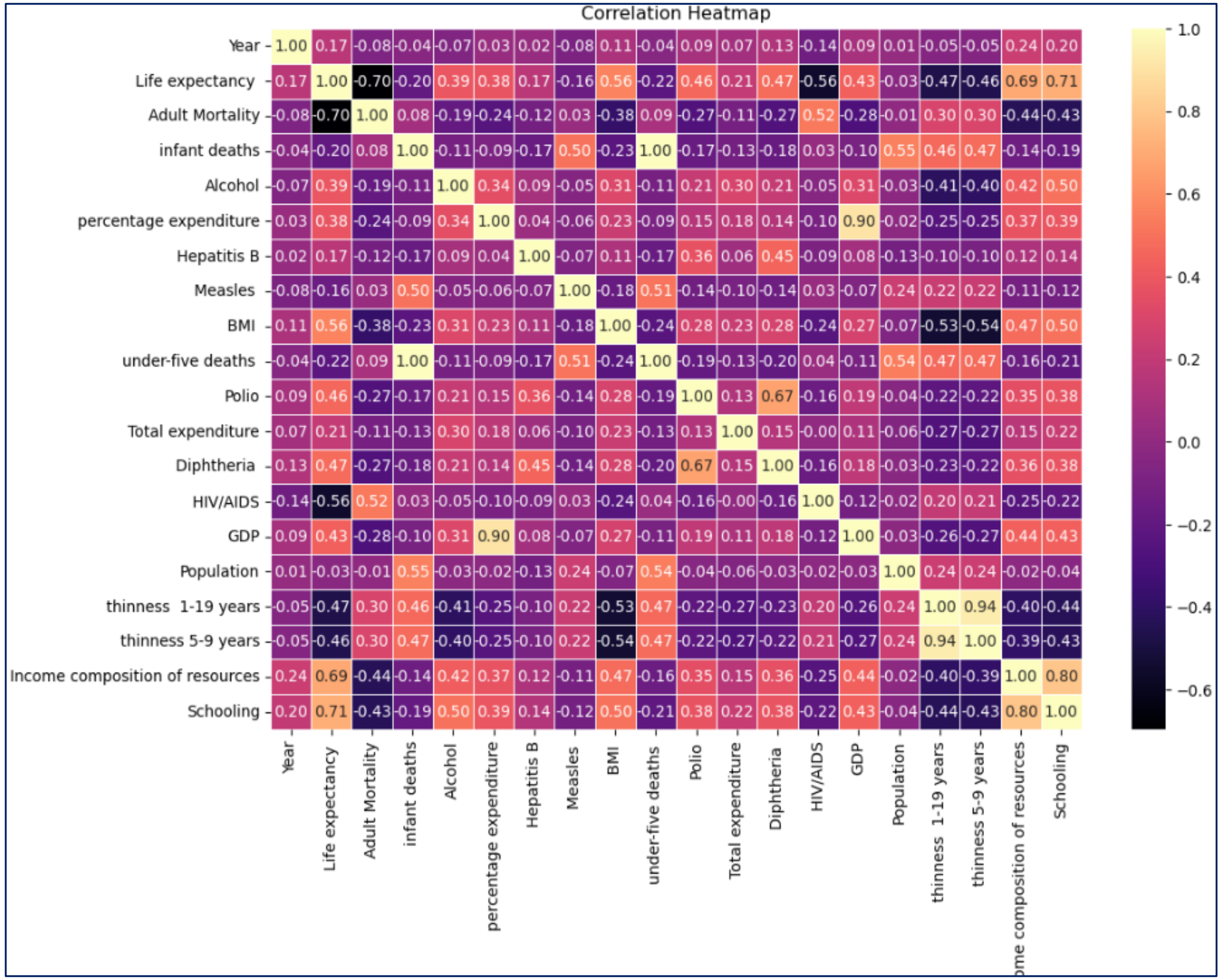
- Since each variable has a unique amount of outliers and also has outliers on different sides of the data, the best route to take is probably winsorizing (limiting) the values for each variable on its own until no outliers remain.

- The function allows to do exactly that by going variable by variable with the ability to use a lower limit and/or upper limit for winsorization.

- By default the function shows two boxplots side by side for the variable (one boxplot of the original data, and one with the winsorized change).

- Once a satisfactory limit is found (by visual analysis), the winsorized data is be saved in the wins_dict dictionary.

Exploratory Data Analysis (EDA)


Correlation Heatmap

	Year	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	BMI	under-five deaths	Polio	Total expenditure	Diphtheria	HIV/AIDS	GDP	Population	thinness 1-19 years	thinness 5-9 years	ome composition of resources	Schooling
Year	1.00	0.17	-0.08	-0.04	-0.07	0.03	0.02	-0.08	0.11	-0.04	0.09	0.07	0.13	-0.14	0.09	0.01	-0.05	-0.05	0.24	0.20
Life expectancy	0.17	1.00	-0.70	-0.20	0.39	0.38	0.17	-0.16	0.56	-0.22	0.46	0.21	0.47	-0.56	0.43	-0.03	-0.47	-0.46	0.69	0.71
Adult Mortality	-0.08	-0.70	1.00	0.08	-0.19	-0.24	-0.12	0.03	-0.38	0.09	-0.27	-0.11	-0.27	0.52	-0.28	-0.01	0.30	0.30	-0.44	-0.43
infant deaths	-0.04	-0.20	0.08	1.00	-0.11	-0.09	-0.17	0.50	-0.23	1.00	-0.17	-0.13	-0.18	0.03	-0.10	0.55	0.46	0.47	-0.14	-0.19
Alcohol	-0.07	0.39	-0.19	-0.11	1.00	0.34	0.09	-0.05	0.31	-0.11	0.21	0.30	0.21	-0.05	0.31	-0.03	-0.41	-0.40	0.42	0.50
percentage expenditure	0.03	0.38	-0.24	-0.09	0.34	1.00	0.04	-0.06	0.23	-0.09	0.15	0.18	0.14	-0.10	0.90	-0.02	-0.25	-0.25	0.37	0.39
Hepatitis B	0.02	0.17	-0.12	-0.17	0.09	0.04	1.00	-0.07	0.11	-0.17	0.36	0.06	0.45	-0.09	0.08	-0.13	-0.10	-0.10	0.12	0.14
Measles	-0.08	-0.16	0.03	0.50	-0.05	-0.06	-0.07	1.00	-0.18	0.51	-0.14	-0.10	-0.14	0.03	-0.07	0.24	0.22	0.22	-0.11	-0.12
BMI	0.11	0.56	-0.38	-0.23	0.31	0.23	0.11	-0.18	1.00	-0.24	0.28	0.23	0.28	-0.24	0.27	-0.07	-0.53	-0.54	0.47	0.50
under-five deaths	-0.04	-0.22	0.09	1.00	-0.11	-0.09	-0.17	0.51	-0.24	1.00	-0.19	-0.13	-0.20	0.04	-0.11	0.54	0.47	0.47	-0.16	-0.21
Polio	0.09	0.46	-0.27	-0.17	0.21	0.15	0.36	-0.14	0.28	-0.19	1.00	0.13	0.67	-0.16	0.19	-0.04	-0.22	-0.22	0.35	0.38
Total expenditure	0.07	0.21	-0.11	-0.13	0.30	0.18	0.06	-0.10	0.23	-0.13	0.13	1.00	0.15	-0.00	0.11	-0.06	-0.27	-0.27	0.15	0.22
Diphtheria	0.13	0.47	-0.27	-0.18	0.21	0.14	0.45	-0.14	0.28	-0.20	0.67	0.15	1.00	-0.16	0.18	-0.03	-0.23	-0.22	0.36	0.38
HIV/AIDS	-0.14	-0.56	0.52	0.03	-0.05	-0.10	-0.09	0.03	-0.24	0.04	-0.16	-0.00	-0.16	1.00	-0.12	-0.02	0.20	0.21	-0.25	-0.22
GDP	0.09	0.43	-0.28	-0.10	0.31	0.90	0.08	-0.07	0.27	-0.11	0.19	0.11	0.18	-0.12	1.00	-0.03	-0.26	-0.27	0.44	0.43
Population	0.01	-0.03	-0.01	0.55	-0.03	-0.02	-0.13	0.24	-0.07	0.54	-0.04	-0.06	-0.03	-0.02	-0.03	1.00	0.24	0.24	-0.02	-0.04
thinness 1-19 years	-0.05	-0.47	0.30	0.46	-0.41	-0.25	-0.10	0.22	-0.53	0.47	-0.22	-0.27	-0.23	0.20	-0.26	0.24	1.00	0.94	-0.40	-0.44
thinness 5-9 years	-0.05	-0.46	0.30	0.47	-0.40	-0.25	-0.10	0.22	-0.54	0.47	-0.22	-0.27	-0.22	0.21	-0.27	0.24	0.94	1.00	-0.39	-0.43
Income composition of resources	0.24	0.69	-0.44	-0.14	0.42	0.37	0.12	-0.11	0.47	-0.16	0.35	0.15	0.36	-0.25	0.44	-0.02	-0.40	-0.39	1.00	0.80
Schooling	0.20	0.71	-0.43	-0.19	0.50	0.39	0.14	-0.12	0.50	-0.21	0.38	0.22	0.38	-0.22	0.43	-0.04	-0.44	-0.43	0.80	1.00

1. Life Expectancy Relationships

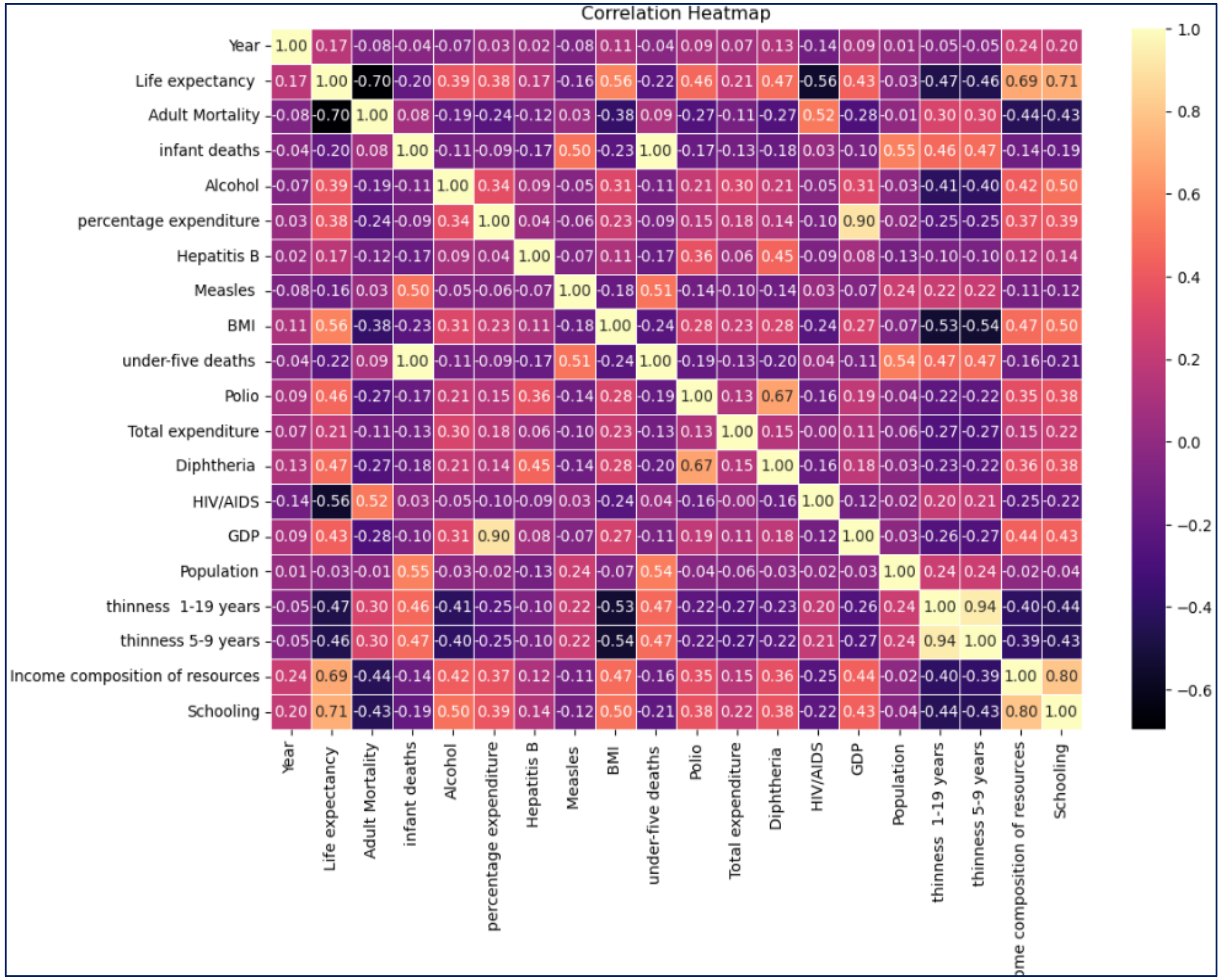
- **Strong Positive Correlation with:**

(i) **Schooling (0.71)** → More education is linked with longer life expectancy.

(ii) **Income composition of resources (0.69)** → Higher income and resource access increase life expectancy.

(iii) **Diphtheria (0.67) and GDP (0.43)** → Better vaccination coverage and stronger economies promote longer life spans.

Exploratory Data Analysis (EDA)



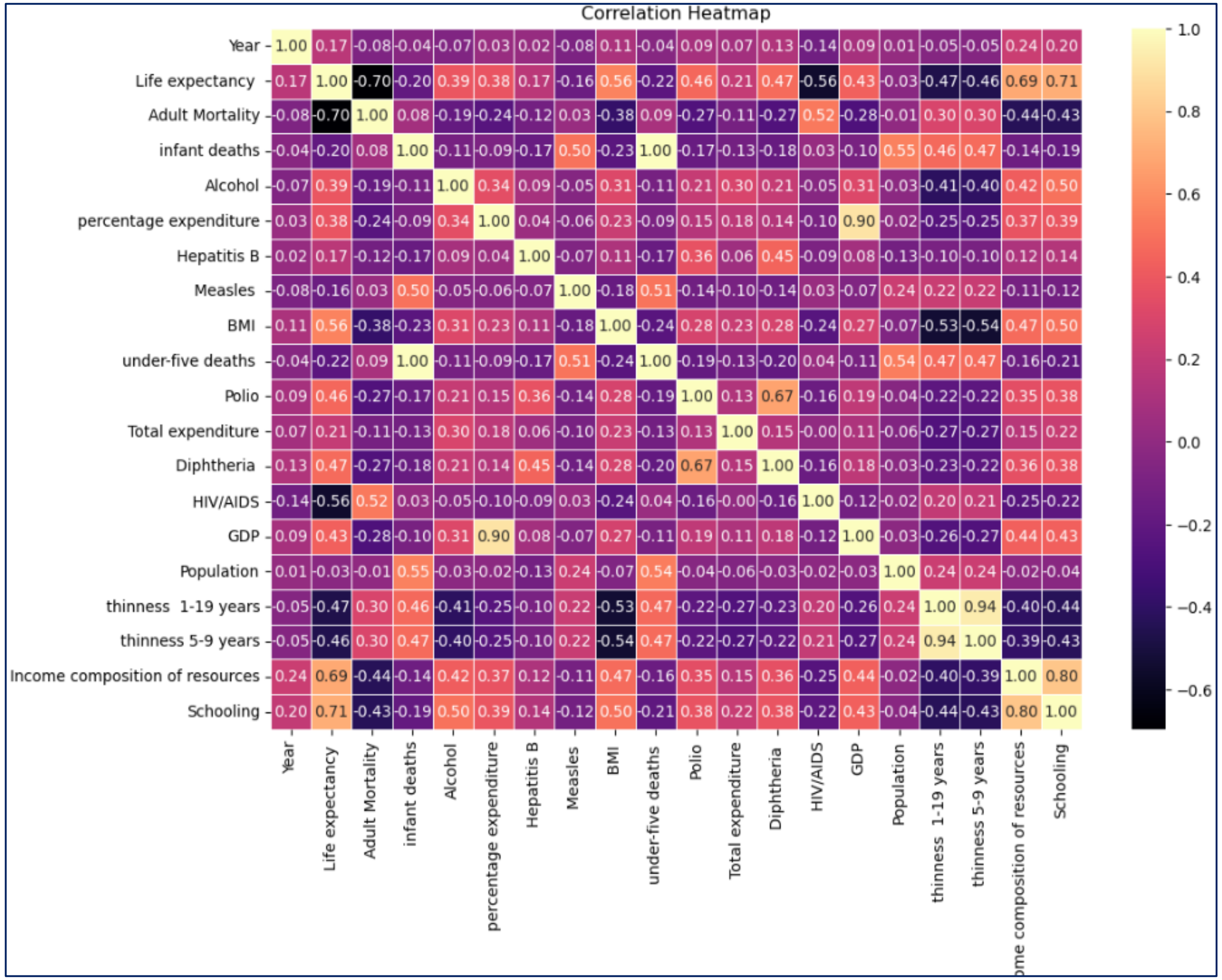
Correlation Heatmap

	Year	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	BMI	under-five deaths	Polio	Total expenditure	Diphtheria	HIV/AIDS	GDP	Population	thinness 1-19 years	thinness 5-9 years	ome composition of resources	Schooling
Year	1.00	0.17	-0.08	-0.04	-0.07	0.03	0.02	-0.08	0.11	-0.04	0.09	0.07	0.13	-0.14	0.09	0.01	-0.05	-0.05	0.24	0.20
Life expectancy	0.17	1.00	-0.70	-0.20	0.39	0.38	0.17	-0.16	0.56	-0.22	0.46	0.21	0.47	-0.56	0.43	-0.03	-0.47	-0.46	0.69	0.71
Adult Mortality	-0.08	-0.70	1.00	0.08	-0.19	-0.24	-0.12	0.03	-0.38	0.09	-0.27	-0.11	-0.27	0.52	-0.28	-0.01	0.30	0.30	-0.44	-0.43
infant deaths	-0.04	-0.20	0.08	1.00	-0.11	-0.09	-0.17	0.50	-0.23	1.00	-0.17	-0.13	-0.18	0.03	-0.10	0.55	0.46	0.47	-0.14	-0.19
Alcohol	-0.07	0.39	-0.19	-0.11	1.00	0.34	0.09	-0.05	0.31	-0.11	0.21	0.30	0.21	-0.05	0.31	-0.03	-0.41	-0.40	0.42	0.50
percentage expenditure	0.03	0.38	-0.24	-0.09	0.34	1.00	0.04	-0.06	0.23	-0.09	0.15	0.18	0.14	-0.10	0.90	-0.02	-0.25	-0.25	0.37	0.39
Hepatitis B	0.02	0.17	-0.12	-0.17	0.09	0.04	1.00	-0.07	0.11	-0.17	0.36	0.06	0.45	-0.09	0.08	-0.13	-0.10	-0.10	0.12	0.14
Measles	-0.08	-0.16	0.03	0.50	-0.05	-0.06	-0.07	1.00	-0.18	0.51	-0.14	-0.10	-0.14	0.03	-0.07	0.24	0.22	0.22	-0.11	-0.12
BMI	0.11	0.56	-0.38	-0.23	0.31	0.23	0.11	-0.18	1.00	-0.24	0.28	0.23	0.28	-0.24	0.27	-0.07	-0.53	-0.54	0.47	0.50
under-five deaths	-0.04	-0.22	0.09	1.00	-0.11	-0.09	-0.17	0.51	-0.24	1.00	-0.19	-0.13	-0.20	0.04	-0.11	0.54	0.47	0.47	-0.16	-0.21
Polio	0.09	0.46	-0.27	-0.17	0.21	0.15	0.36	-0.14	0.28	-0.19	1.00	0.13	0.67	-0.16	0.19	-0.04	-0.22	-0.22	0.35	0.38
Total expenditure	0.07	0.21	-0.11	-0.13	0.30	0.18	0.06	-0.10	0.23	-0.13	0.13	1.00	0.15	-0.00	0.11	-0.06	-0.27	-0.27	0.15	0.22
Diphtheria	0.13	0.47	-0.27	-0.18	0.21	0.14	0.45	-0.14	0.28	-0.20	0.67	0.15	1.00	-0.16	0.18	-0.03	-0.23	-0.22	0.36	0.38
HIV/AIDS	-0.14	-0.56	0.52	0.03	-0.05	-0.10	-0.09	0.03	-0.24	0.04	-0.16	-0.00	-0.16	1.00	-0.12	-0.02	0.20	0.21	-0.25	-0.22
GDP	0.09	0.43	-0.28	-0.10	0.31	0.90	0.08	-0.07	0.27	-0.11	0.19	0.11	0.18	-0.12	1.00	-0.03	-0.26	-0.27	0.44	0.43
Population	0.01	-0.03	-0.01	0.55	-0.03	-0.02	-0.13	0.24	-0.07	0.54	-0.04	-0.06	-0.03	-0.02	-0.03	1.00	0.24	0.24	-0.02	-0.04
thinness 1-19 years	-0.05	-0.47	0.30	0.46	-0.41	-0.25	-0.10	0.22	-0.53	0.47	-0.22	-0.27	-0.23	0.20	-0.26	0.24	1.00	0.94	-0.40	-0.44
thinness 5-9 years	-0.05	-0.46	0.30	0.47	-0.40	-0.25	-0.10	0.22	-0.54	0.47	-0.22	-0.27	-0.22	0.21	-0.27	0.24	0.94	1.00	-0.39	-0.43
Income composition of resources	0.24	0.69	-0.44	-0.14	0.42	0.37	0.12	-0.11	0.47	-0.16	0.35	0.15	0.36	-0.25	0.44	-0.02	-0.40	-0.39	1.00	0.80
Schooling	0.20	0.71	-0.43	-0.19	0.50	0.39	0.14	-0.12	0.50	-0.21	0.38	0.22	0.38	-0.22	0.43	-0.04	-0.44	-0.43	0.80	1.00

- **Strong Negative Correlation with:**
- **Adult Mortality (-0.70) and Infant deaths (-0.44)** → Higher mortality rates reduce life expectancy.
- **Under-five deaths (-0.44) → More child deaths** = lower life expectancy.

Interpretation:
Countries with higher education, income, and vaccination rates tend to have higher life expectancy, while those with high mortality and child death rates have lower life expectancy.

Exploratory Data Analysis (EDA)

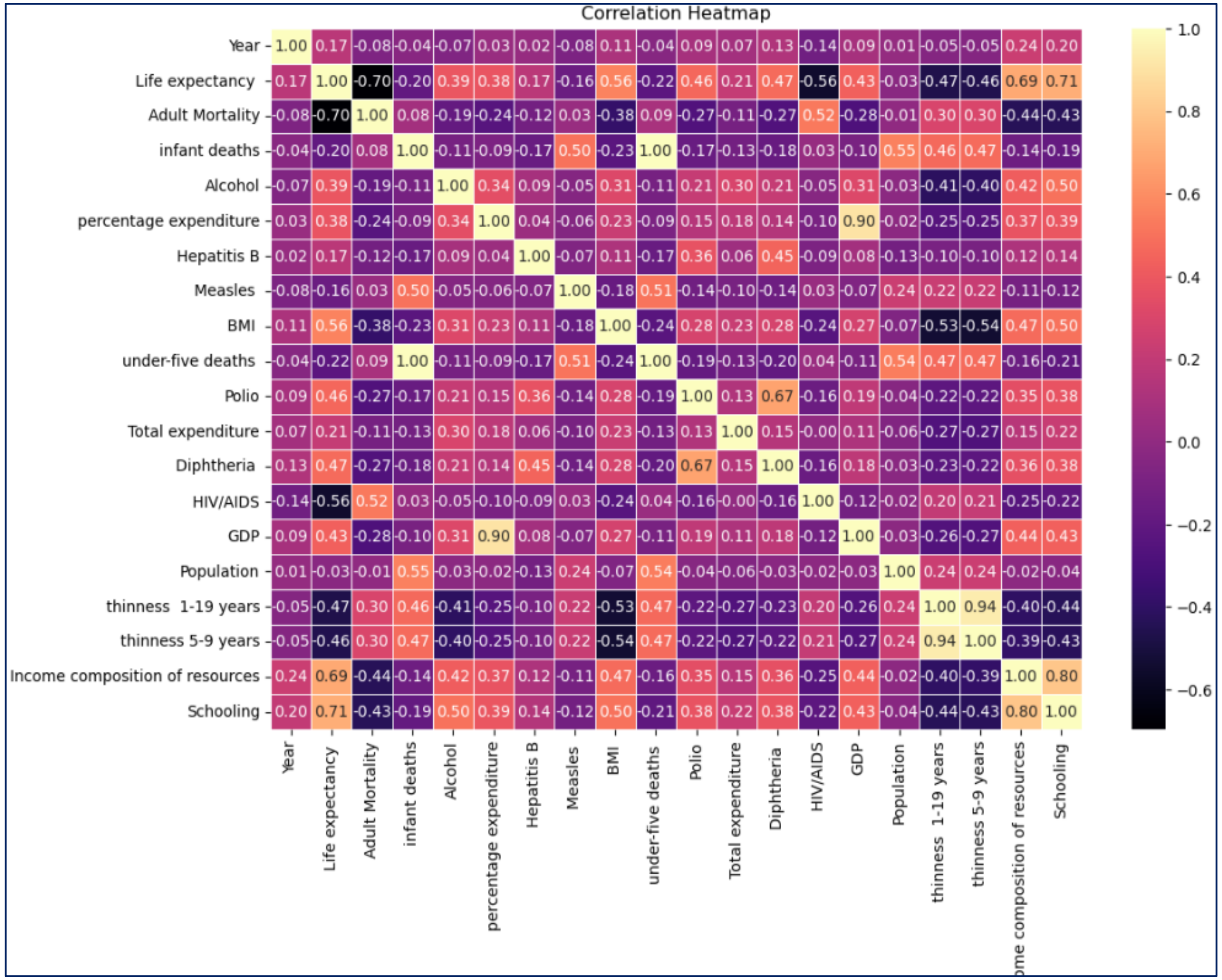


	Year	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	BMI	under-five deaths	Polio	Total expenditure	Diphtheria	HIV/AIDS	GDP	Population	thinness 1-19 years	thinness 5-9 years	ome composition of resources	Schooling
Year	1.00	0.17	-0.08	-0.04	-0.07	0.03	0.02	-0.08	0.11	-0.04	0.09	0.07	0.13	-0.14	0.09	0.01	-0.05	-0.05	0.24	0.20
Life expectancy	0.17	1.00	-0.70	-0.20	0.39	0.38	0.17	-0.16	0.56	-0.22	0.46	0.21	0.47	-0.56	0.43	-0.03	-0.47	-0.46	0.69	0.71
Adult Mortality	-0.08	-0.70	1.00	0.08	-0.19	-0.24	-0.12	0.03	-0.38	0.09	-0.27	-0.11	-0.27	0.52	-0.28	-0.01	0.30	0.30	-0.44	-0.43
infant deaths	-0.04	-0.20	0.08	1.00	-0.11	-0.09	-0.17	0.50	-0.23	1.00	-0.17	-0.13	-0.18	0.03	-0.10	0.55	0.46	0.47	-0.14	-0.19
Alcohol	-0.07	0.39	-0.19	-0.11	1.00	0.34	0.09	-0.05	0.31	-0.11	0.21	0.30	0.21	-0.05	0.31	-0.03	-0.41	-0.40	0.42	0.50
percentage expenditure	0.03	0.38	-0.24	-0.09	0.34	1.00	0.04	-0.06	0.23	-0.09	0.15	0.18	0.14	-0.10	0.90	-0.02	-0.25	-0.25	0.37	0.39
Hepatitis B	0.02	0.17	-0.12	-0.17	0.09	0.04	1.00	-0.07	0.11	-0.17	0.36	0.06	0.45	-0.09	0.08	-0.13	-0.10	-0.10	0.12	0.14
Measles	-0.08	-0.16	0.03	0.50	-0.05	-0.06	-0.07	1.00	-0.18	0.51	-0.14	-0.10	-0.14	0.03	-0.07	0.24	0.22	0.22	-0.11	-0.12
BMI	0.11	0.56	-0.38	-0.23	0.31	0.23	0.11	-0.18	1.00	-0.24	0.28	0.23	0.28	-0.24	0.27	-0.07	-0.53	-0.54	0.47	0.50
under-five deaths	-0.04	-0.22	0.09	1.00	-0.11	-0.09	-0.17	0.51	-0.24	1.00	-0.19	-0.13	-0.20	0.04	-0.11	0.54	0.47	0.47	-0.16	-0.21
Polio	0.09	0.46	-0.27	-0.17	0.21	0.15	0.36	-0.14	0.28	-0.19	1.00	0.13	0.67	-0.16	0.19	-0.04	-0.22	-0.22	0.35	0.38
Total expenditure	0.07	0.21	-0.11	-0.13	0.30	0.18	0.06	-0.10	0.23	-0.13	0.13	1.00	0.15	-0.00	0.11	-0.06	-0.27	-0.27	0.15	0.22
Diphtheria	0.13	0.47	-0.27	-0.18	0.21	0.14	0.45	-0.14	0.28	-0.20	0.67	0.15	1.00	-0.16	0.18	-0.03	-0.23	-0.22	0.36	0.38
HIV/AIDS	-0.14	-0.56	0.52	0.03	-0.05	-0.10	-0.09	0.03	-0.24	0.04	-0.16	-0.00	-0.16	1.00	-0.12	-0.02	0.20	0.21	-0.25	-0.22
GDP	0.09	0.43	-0.28	-0.10	0.31	0.90	0.08	-0.07	0.27	-0.11	0.19	0.11	0.18	-0.12	1.00	-0.03	-0.26	-0.27	0.44	0.43
Population	0.01	-0.03	-0.01	0.55	-0.03	-0.02	-0.13	0.24	-0.07	0.54	-0.04	-0.06	-0.03	-0.02	-0.03	1.00	0.24	0.24	-0.02	-0.04
thinness 1-19 years	-0.05	-0.47	0.30	0.46	-0.41	-0.25	-0.10	0.22	-0.53	0.47	-0.22	-0.27	-0.23	0.20	-0.26	0.24	1.00	0.94	-0.40	-0.44
thinness 5-9 years	-0.05	-0.46	0.30	0.47	-0.40	-0.25	-0.10	0.22	-0.54	0.47	-0.22	-0.27	-0.22	0.21	-0.27	0.24	0.94	1.00	-0.39	-0.43
Income composition of resources	0.24	0.69	-0.44	-0.14	0.42	0.37	0.12	-0.11	0.47	-0.16	0.35	0.15	0.36	-0.25	0.44	-0.02	-0.40	-0.39	1.00	0.80
Schooling	0.20	0.71	-0.43	-0.19	0.50	0.39	0.14	-0.12	0.50	-0.21	0.38	0.22	0.38	-0.22	0.43	-0.04	-0.44	-0.43	0.80	1.00

2. Mortality and Child Deaths

- **Adult Mortality, Infant deaths, and Under-five deaths are highly positively correlated (~0.9) →** Suggests that these indicators move together — if one is high, others tend to be high too.
- **Indicates overlapping information** → these could be redundant features in a model (high multicollinearity).

Exploratory Data Analysis (EDA)



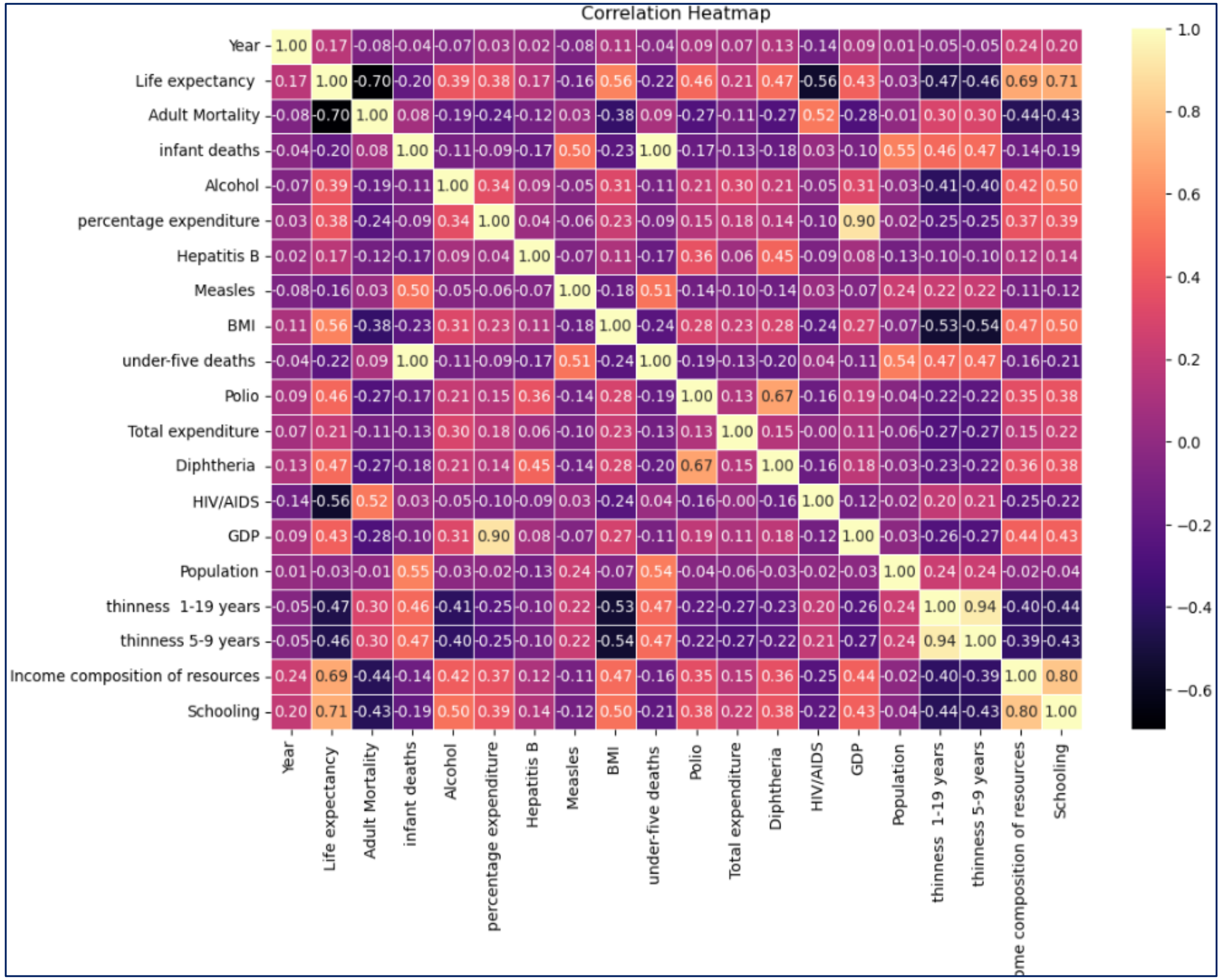
Correlation Heatmap

	Year	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	BMI	under-five deaths	Polio	Total expenditure	Diphtheria	HIV/AIDS	GDP	Population	thinness 1-19 years	thinness 5-9 years	ome composition of resources	Schooling
Year	1.00	0.17	-0.08	-0.04	-0.07	0.03	0.02	-0.08	0.11	-0.04	0.09	0.07	0.13	-0.14	0.09	0.01	-0.05	-0.05	0.24	0.20
Life expectancy	0.17	1.00	-0.70	-0.20	0.39	0.38	0.17	-0.16	0.56	-0.22	0.46	0.21	0.47	-0.56	0.43	-0.03	-0.47	-0.46	0.69	0.71
Adult Mortality	-0.08	-0.70	1.00	0.08	-0.19	-0.24	-0.12	0.03	-0.38	0.09	-0.27	-0.11	-0.27	0.52	-0.28	-0.01	0.30	0.30	-0.44	-0.43
infant deaths	-0.04	-0.20	0.08	1.00	-0.11	-0.09	-0.17	0.50	-0.23	1.00	-0.17	-0.13	-0.18	0.03	-0.10	0.55	0.46	0.47	-0.14	-0.19
Alcohol	-0.07	0.39	-0.19	-0.11	1.00	0.34	0.09	-0.05	0.31	-0.11	0.21	0.30	0.21	-0.05	0.31	-0.03	-0.41	-0.40	0.42	0.50
percentage expenditure	0.03	0.38	-0.24	-0.09	0.34	1.00	0.04	-0.06	0.23	-0.09	0.15	0.18	0.14	-0.10	0.90	-0.02	-0.25	-0.25	0.37	0.39
Hepatitis B	0.02	0.17	-0.12	-0.17	0.09	0.04	1.00	-0.07	0.11	-0.17	0.36	0.06	0.45	-0.09	0.08	-0.13	-0.10	-0.10	0.12	0.14
Measles	-0.08	-0.16	0.03	0.50	-0.05	-0.06	-0.07	1.00	-0.18	0.51	-0.14	-0.10	-0.14	0.03	-0.07	0.24	0.22	0.22	-0.11	-0.12
BMI	0.11	0.56	-0.38	-0.23	0.31	0.23	0.11	-0.18	1.00	-0.24	0.28	0.23	0.28	-0.24	0.27	-0.07	-0.53	-0.54	0.47	0.50
under-five deaths	-0.04	-0.22	0.09	1.00	-0.11	-0.09	-0.17	0.51	-0.24	1.00	-0.19	-0.13	-0.20	0.04	-0.11	0.54	0.47	0.47	-0.16	-0.21
Polio	0.09	0.46	-0.27	-0.17	0.21	0.15	0.36	-0.14	0.28	-0.19	1.00	0.13	0.67	-0.16	0.19	-0.04	-0.22	-0.22	0.35	0.38
Total expenditure	0.07	0.21	-0.11	-0.13	0.30	0.18	0.06	-0.10	0.23	-0.13	0.13	1.00	0.15	-0.00	0.11	-0.06	-0.27	-0.27	0.15	0.22
Diphtheria	0.13	0.47	-0.27	-0.18	0.21	0.14	0.45	-0.14	0.28	-0.20	0.67	0.15	1.00	-0.16	0.18	-0.03	-0.23	-0.22	0.36	0.38
HIV/AIDS	-0.14	-0.56	0.52	0.03	-0.05	-0.10	-0.09	0.03	-0.24	0.04	-0.16	-0.00	-0.16	1.00	-0.12	-0.02	0.20	0.21	-0.25	-0.22
GDP	0.09	0.43	-0.28	-0.10	0.31	0.90	0.08	-0.07	0.27	-0.11	0.19	0.11	0.18	-0.12	1.00	-0.03	-0.26	-0.27	0.44	0.43
Population	0.01	-0.03	-0.01	0.55	-0.03	-0.02	-0.13	0.24	-0.07	0.54	-0.04	-0.06	-0.03	-0.02	-0.03	1.00	0.24	0.24	-0.02	-0.04
thinness 1-19 years	-0.05	-0.47	0.30	0.46	-0.41	-0.25	-0.10	0.22	-0.53	0.47	-0.22	-0.27	-0.23	0.20	-0.26	0.24	1.00	0.94	-0.40	-0.44
thinness 5-9 years	-0.05	-0.46	0.30	0.47	-0.40	-0.25	-0.10	0.22	-0.54	0.47	-0.22	-0.27	-0.22	0.21	-0.27	0.24	0.94	1.00	-0.39	-0.43
Income composition of resources	0.24	0.69	-0.44	-0.14	0.42	0.37	0.12	-0.11	0.47	-0.16	0.35	0.15	0.36	-0.25	0.44	-0.02	-0.40	-0.39	1.00	0.80
Schooling	0.20	0.71	-0.43	-0.19	0.50	0.39	0.14	-0.12	0.50	-0.21	0.38	0.22	0.38	-0.22	0.43	-0.04	-0.44	-0.43	0.80	1.00

3. Education and Income

- **Schooling ↔ Income composition of resources (0.64)**
  → Countries with better education levels tend to have higher income equality and resource access.
- **Schooling ↔ GDP (0.43)**
  → Education correlates moderately with national income, reflecting economic development.

Exploratory Data Analysis (EDA)



Correlation Heatmap

	Year	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	BMI	under-five deaths	Polio	Total expenditure	Diphtheria	HIV/AIDS	GDP	Population	thinness 1-19 years	thinness 5-9 years	ome composition of resources	Schooling
Year	1.00	0.17	-0.08	-0.04	-0.07	0.03	0.02	-0.08	0.11	-0.04	0.09	0.07	0.13	-0.14	0.09	0.01	-0.05	-0.05	0.24	0.20
Life expectancy	0.17	1.00	-0.70	-0.20	0.39	0.38	0.17	-0.16	0.56	-0.22	0.46	0.21	0.47	-0.56	0.43	-0.03	-0.47	-0.46	0.69	0.71
Adult Mortality	-0.08	-0.70	1.00	0.08	-0.19	-0.24	-0.12	0.03	-0.38	0.09	-0.27	-0.11	-0.27	0.52	-0.28	-0.01	0.30	0.30	-0.44	-0.43
infant deaths	-0.04	-0.20	0.08	1.00	-0.11	-0.09	-0.17	0.50	-0.23	1.00	-0.17	-0.13	-0.18	0.03	-0.10	0.55	0.46	0.47	-0.14	-0.19
Alcohol	-0.07	0.39	-0.19	-0.11	1.00	0.34	0.09	-0.05	0.31	-0.11	0.21	0.30	0.21	-0.05	0.31	-0.03	-0.41	-0.40	0.42	0.50
percentage expenditure	0.03	0.38	-0.24	-0.09	0.34	1.00	0.04	-0.06	0.23	-0.09	0.15	0.18	0.14	-0.10	0.90	-0.02	-0.25	-0.25	0.37	0.39
Hepatitis B	0.02	0.17	-0.12	-0.17	0.09	0.04	1.00	-0.07	0.11	-0.17	0.36	0.06	0.45	-0.09	0.08	-0.13	-0.10	-0.10	0.12	0.14
Measles	-0.08	-0.16	0.03	0.50	-0.05	-0.06	-0.07	1.00	-0.18	0.51	-0.14	-0.10	-0.14	0.03	-0.07	0.24	0.22	0.22	-0.11	-0.12
BMI	0.11	0.56	-0.38	-0.23	0.31	0.23	0.11	-0.18	1.00	-0.24	0.28	0.23	0.28	-0.24	0.27	-0.07	-0.53	-0.54	0.47	0.50
under-five deaths	-0.04	-0.22	0.09	1.00	-0.11	-0.09	-0.17	0.51	-0.24	1.00	-0.19	-0.13	-0.20	0.04	-0.11	0.54	0.47	0.47	-0.16	-0.21
Polio	0.09	0.46	-0.27	-0.17	0.21	0.15	0.36	-0.14	0.28	-0.19	1.00	0.13	0.67	-0.16	0.19	-0.04	-0.22	-0.22	0.35	0.38
Total expenditure	0.07	0.21	-0.11	-0.13	0.30	0.18	0.06	-0.10	0.23	-0.13	0.13	1.00	0.15	-0.00	0.11	-0.06	-0.27	-0.27	0.15	0.22
Diphtheria	0.13	0.47	-0.27	-0.18	0.21	0.14	0.45	-0.14	0.28	-0.20	0.67	0.15	1.00	-0.16	0.18	-0.03	-0.23	-0.22	0.36	0.38
HIV/AIDS	-0.14	-0.56	0.52	0.03	-0.05	-0.10	-0.09	0.03	-0.24	0.04	-0.16	-0.00	-0.16	1.00	-0.12	-0.02	0.20	0.21	-0.25	-0.22
GDP	0.09	0.43	-0.28	-0.10	0.31	0.90	0.08	-0.07	0.27	-0.11	0.19	0.11	0.18	-0.12	1.00	-0.03	-0.26	-0.27	0.44	0.43
Population	0.01	-0.03	-0.01	0.55	-0.03	-0.02	-0.13	0.24	-0.07	0.54	-0.04	-0.06	-0.03	-0.02	-0.03	1.00	0.24	0.24	-0.02	-0.04
thinness 1-19 years	-0.05	-0.47	0.30	0.46	-0.41	-0.25	-0.10	0.22	-0.53	0.47	-0.22	-0.27	-0.23	0.20	-0.26	0.24	1.00	0.94	-0.40	-0.44
thinness 5-9 years	-0.05	-0.46	0.30	0.47	-0.40	-0.25	-0.10	0.22	-0.54	0.47	-0.22	-0.27	-0.22	0.21	-0.27	0.24	0.94	1.00	-0.39	-0.43
Income composition of resources	0.24	0.69	-0.44	-0.14	0.42	0.37	0.12	-0.11	0.47	-0.16	0.35	0.15	0.36	-0.25	0.44	-0.02	-0.40	-0.39	1.00	0.80
Schooling	0.20	0.71	-0.43	-0.19	0.50	0.39	0.14	-0.12	0.50	-0.21	0.38	0.22	0.38	-0.22	0.43	-0.04	-0.44	-0.43	0.80	1.00

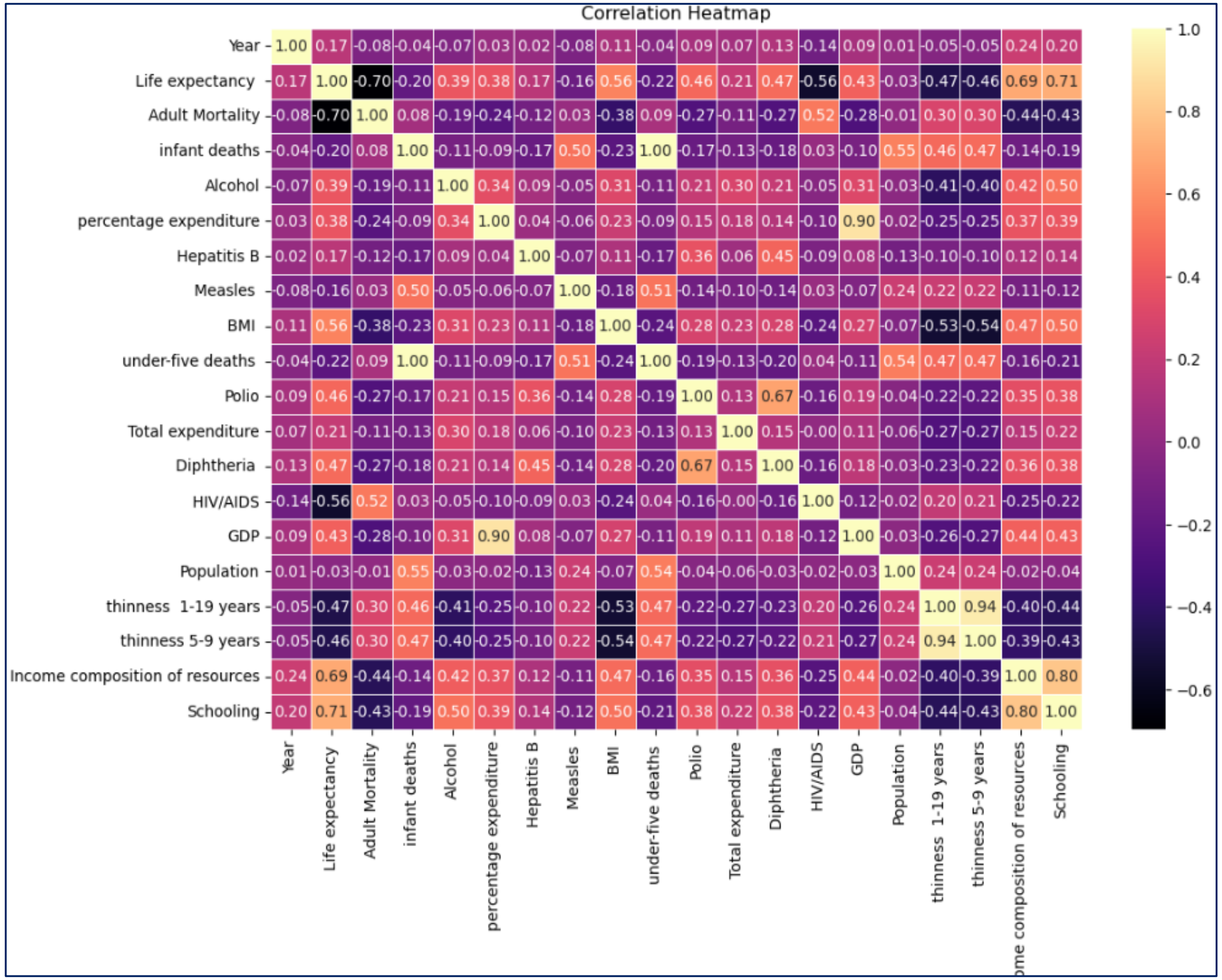
4. Vaccination (Diphtheria, Polio, Hepatitis B)

- **Diphtheria ↔ Polio (0.87)** → Very high correlation.
- **Diphtheria ↔ Hepatitis B (0.67)** → Indicates that countries with strong immunization programs perform well across multiple vaccines.

Interpretation:

These vaccine coverage indicators might represent a single latent factor: “immunization coverage.”

Exploratory Data Analysis (EDA)

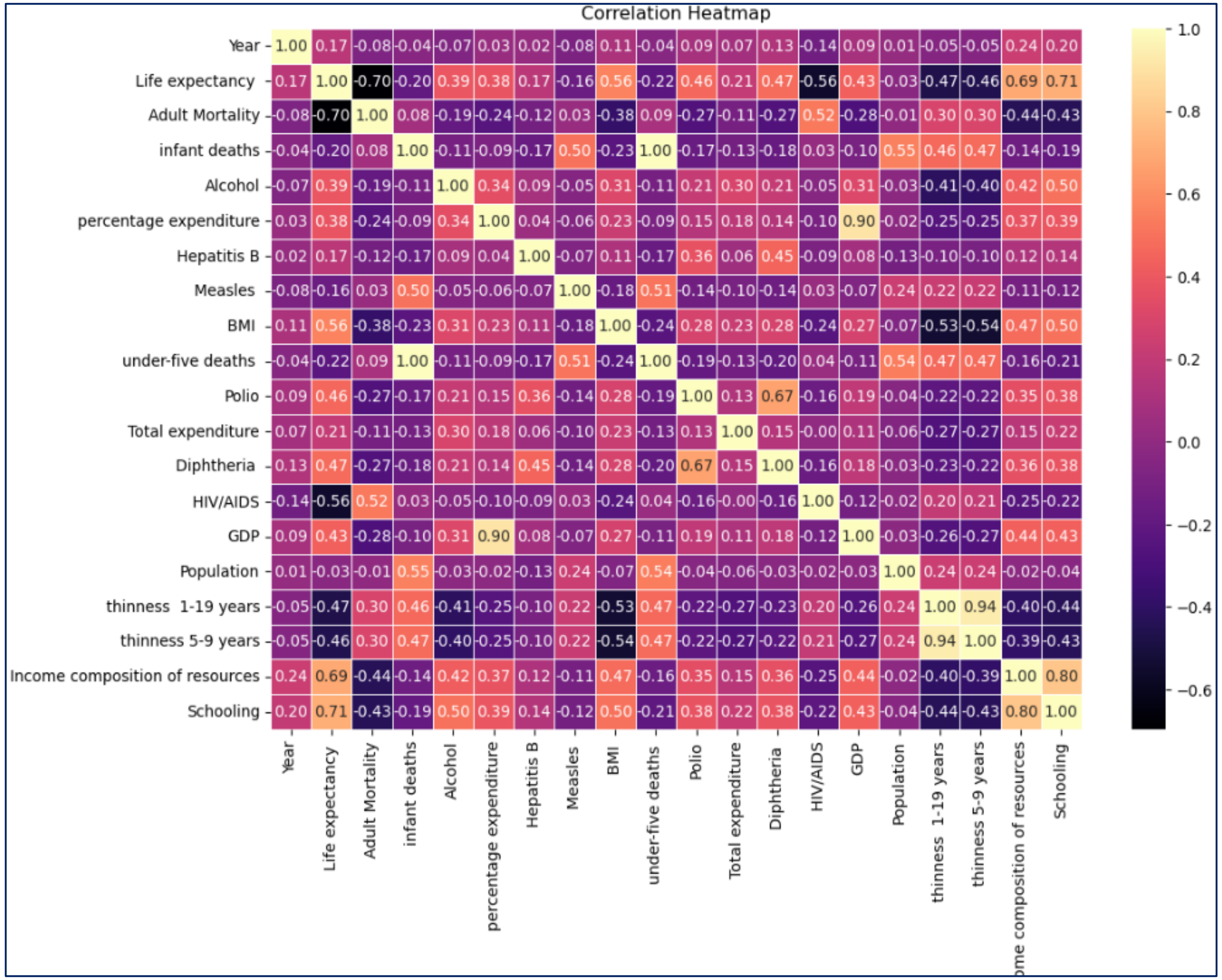


	Year	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	BMI	under-five deaths	Polio	Total expenditure	Diphtheria	HIV/AIDS	GDP	Population	thinness 1-19 years	thinness 5-9 years	ome composition of resources	Schooling
Year	1.00	0.17	-0.08	-0.04	-0.07	0.03	0.02	-0.08	0.11	-0.04	0.09	0.07	0.13	-0.14	0.09	0.01	-0.05	-0.05	0.24	0.20
Life expectancy	0.17	1.00	-0.70	-0.20	0.39	0.38	0.17	-0.16	0.56	-0.22	0.46	0.21	0.47	-0.56	0.43	-0.03	-0.47	-0.46	0.69	0.71
Adult Mortality	-0.08	-0.70	1.00	0.08	-0.19	-0.24	-0.12	0.03	-0.38	0.09	-0.27	-0.11	-0.27	0.52	-0.28	-0.01	0.30	0.30	-0.44	-0.43
infant deaths	-0.04	-0.20	0.08	1.00	-0.11	-0.09	-0.17	0.50	-0.23	1.00	-0.17	-0.13	-0.18	0.03	-0.10	0.55	0.46	0.47	-0.14	-0.19
Alcohol	-0.07	0.39	-0.19	-0.11	1.00	0.34	0.09	-0.05	0.31	-0.11	0.21	0.30	0.21	-0.05	0.31	-0.03	-0.41	-0.40	0.42	0.50
percentage expenditure	0.03	0.38	-0.24	-0.09	0.34	1.00	0.04	-0.06	0.23	-0.09	0.15	0.18	0.14	-0.10	0.90	-0.02	-0.25	-0.25	0.37	0.39
Hepatitis B	0.02	0.17	-0.12	-0.17	0.09	0.04	1.00	-0.07	0.11	-0.17	0.36	0.06	0.45	-0.09	0.08	-0.13	-0.10	-0.10	0.12	0.14
Measles	-0.08	-0.16	0.03	0.50	-0.05	-0.06	-0.07	1.00	-0.18	0.51	-0.14	-0.10	-0.14	0.03	-0.07	0.24	0.22	0.22	-0.11	-0.12
BMI	0.11	0.56	-0.38	-0.23	0.31	0.23	0.11	-0.18	1.00	-0.24	0.28	0.23	0.28	-0.24	0.27	-0.07	-0.53	-0.54	0.47	0.50
under-five deaths	-0.04	-0.22	0.09	1.00	-0.11	-0.09	-0.17	0.51	-0.24	1.00	-0.19	-0.13	-0.20	0.04	-0.11	0.54	0.47	0.47	-0.16	-0.21
Polio	0.09	0.46	-0.27	-0.17	0.21	0.15	0.36	-0.14	0.28	-0.19	1.00	0.13	0.67	-0.16	0.19	-0.04	-0.22	-0.22	0.35	0.38
Total expenditure	0.07	0.21	-0.11	-0.13	0.30	0.18	0.06	-0.10	0.23	-0.13	0.13	1.00	0.15	-0.00	0.11	-0.06	-0.27	-0.27	0.15	0.22
Diphtheria	0.13	0.47	-0.27	-0.18	0.21	0.14	0.45	-0.14	0.28	-0.20	0.67	0.15	1.00	-0.16	0.18	-0.03	-0.23	-0.22	0.36	0.38
HIV/AIDS	-0.14	-0.56	0.52	0.03	-0.05	-0.10	-0.09	0.03	-0.24	0.04	-0.16	-0.00	-0.16	1.00	-0.12	-0.02	0.20	0.21	-0.25	-0.22
GDP	0.09	0.43	-0.28	-0.10	0.31	0.90	0.08	-0.07	0.27	-0.11	0.19	0.11	0.18	-0.12	1.00	-0.03	-0.26	-0.27	0.44	0.43
Population	0.01	-0.03	-0.01	0.55	-0.03	-0.02	-0.13	0.24	-0.07	0.54	-0.04	-0.06	-0.03	-0.02	-0.03	1.00	0.24	0.24	-0.02	-0.04
thinness 1-19 years	-0.05	-0.47	0.30	0.46	-0.41	-0.25	-0.10	0.22	-0.53	0.47	-0.22	-0.27	-0.23	0.20	-0.26	0.24	1.00	0.94	-0.40	-0.44
thinness 5-9 years	-0.05	-0.46	0.30	0.47	-0.40	-0.25	-0.10	0.22	-0.54	0.47	-0.22	-0.27	-0.22	0.21	-0.27	0.24	0.94	1.00	-0.39	-0.43
Income composition of resources	0.24	0.69	-0.44	-0.14	0.42	0.37	0.12	-0.11	0.47	-0.16	0.35	0.15	0.36	-0.25	0.44	-0.02	-0.40	-0.39	1.00	0.80
Schooling	0.20	0.71	-0.43	-0.19	0.50	0.39	0.14	-0.12	0.50	-0.21	0.38	0.22	0.38	-0.22	0.43	-0.04	-0.44	-0.43	0.80	1.00

5. Negative Health Indicators

- **HIV/AIDS** has a **negative correlation** with **Life expectancy (-0.56)**.
  → Regions with higher HIV rates have lower life expectancy.
- **HIV/AIDS** also correlates negatively with **GDP (-0.43)** and **Schooling (-0.46)**.
  → Reflects the socioeconomic burden of the disease.

Exploratory Data Analysis (EDA)



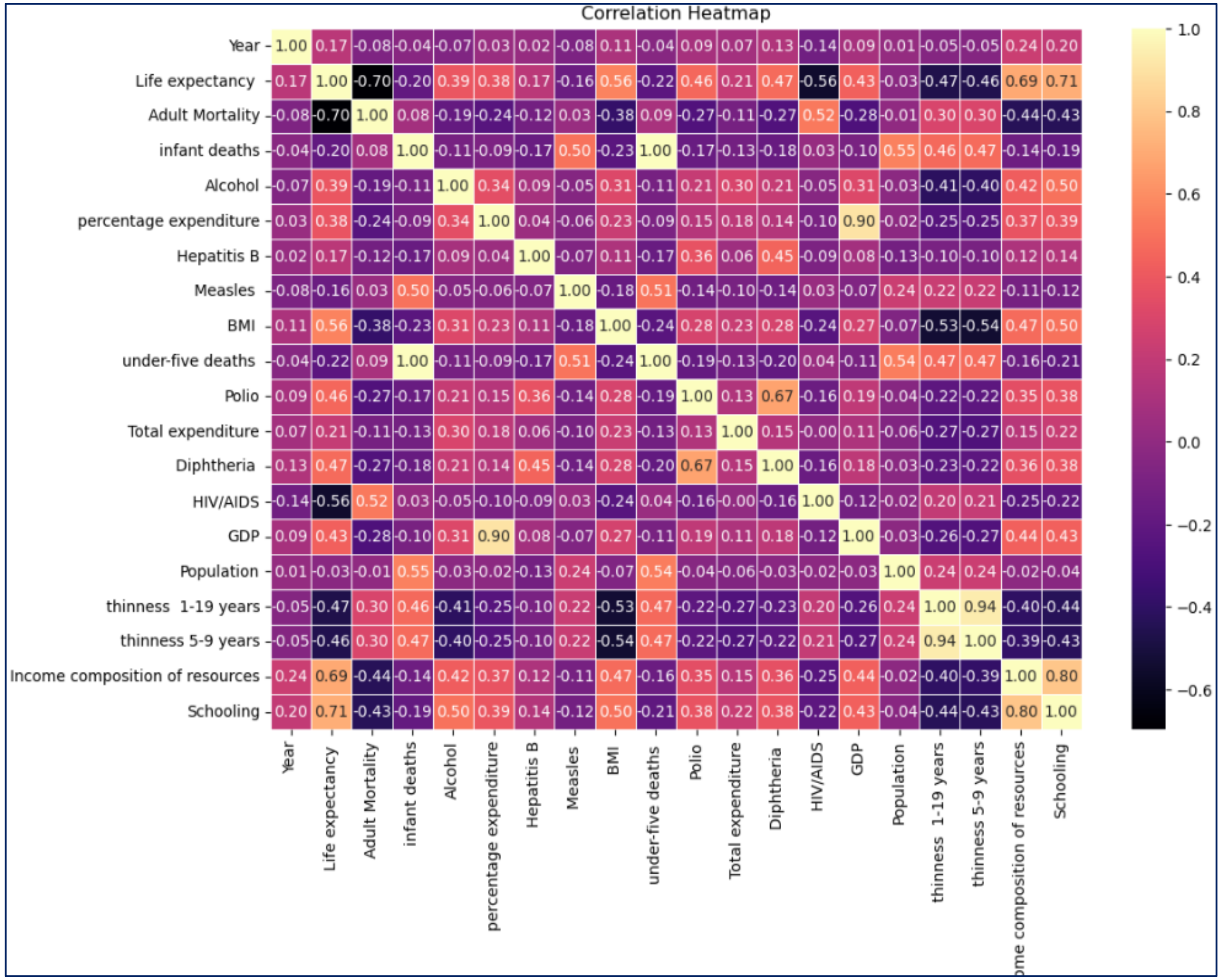
Correlation Heatmap

	Year	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	BMI	under-five deaths	Polio	Total expenditure	Diphtheria	HIV/AIDS	GDP	Population	thinness 1-19 years	thinness 5-9 years	ome composition of resources	Schooling
Year	1.00	0.17	-0.08	-0.04	-0.07	0.03	0.02	-0.08	0.11	-0.04	0.09	0.07	0.13	-0.14	0.09	0.01	-0.05	-0.05	0.24	0.20
Life expectancy	0.17	1.00	-0.70	-0.20	0.39	0.38	0.17	-0.16	0.56	-0.22	0.46	0.21	0.47	-0.56	0.43	-0.03	-0.47	-0.46	0.69	0.71
Adult Mortality	-0.08	-0.70	1.00	0.08	-0.19	-0.24	-0.12	0.03	-0.38	0.09	-0.27	-0.11	-0.27	0.52	-0.28	-0.01	0.30	0.30	-0.44	-0.43
infant deaths	-0.04	-0.20	0.08	1.00	-0.11	-0.09	-0.17	0.50	-0.23	1.00	-0.17	-0.13	-0.18	0.03	-0.10	0.55	0.46	0.47	-0.14	-0.19
Alcohol	-0.07	0.39	-0.19	-0.11	1.00	0.34	0.09	-0.05	0.31	-0.11	0.21	0.30	0.21	-0.05	0.31	-0.03	-0.41	-0.40	0.42	0.50
percentage expenditure	0.03	0.38	-0.24	-0.09	0.34	1.00	0.04	-0.06	0.23	-0.09	0.15	0.18	0.14	-0.10	0.90	-0.02	-0.25	-0.25	0.37	0.39
Hepatitis B	0.02	0.17	-0.12	-0.17	0.09	0.04	1.00	-0.07	0.11	-0.17	0.36	0.06	0.45	-0.09	0.08	-0.13	-0.10	-0.10	0.12	0.14
Measles	-0.08	-0.16	0.03	0.50	-0.05	-0.06	-0.07	1.00	-0.18	0.51	-0.14	-0.10	-0.14	0.03	-0.07	0.24	0.22	0.22	-0.11	-0.12
BMI	0.11	0.56	-0.38	-0.23	0.31	0.23	0.11	-0.18	1.00	-0.24	0.28	0.23	0.28	-0.24	0.27	-0.07	-0.53	-0.54	0.47	0.50
under-five deaths	-0.04	-0.22	0.09	1.00	-0.11	-0.09	-0.17	0.51	-0.24	1.00	-0.19	-0.13	-0.20	0.04	-0.11	0.54	0.47	0.47	-0.16	-0.21
Polio	0.09	0.46	-0.27	-0.17	0.21	0.15	0.36	-0.14	0.28	-0.19	1.00	0.13	0.67	-0.16	0.19	-0.04	-0.22	-0.22	0.35	0.38
Total expenditure	0.07	0.21	-0.11	-0.13	0.30	0.18	0.06	-0.10	0.23	-0.13	0.13	1.00	0.15	-0.00	0.11	-0.06	-0.27	-0.27	0.15	0.22
Diphtheria	0.13	0.47	-0.27	-0.18	0.21	0.14	0.45	-0.14	0.28	-0.20	0.67	0.15	1.00	-0.16	0.18	-0.03	-0.23	-0.22	0.36	0.38
HIV/AIDS	-0.14	-0.56	0.52	0.03	-0.05	-0.10	-0.09	0.03	-0.24	0.04	-0.16	-0.00	-0.16	1.00	-0.12	-0.02	0.20	0.21	-0.25	-0.22
GDP	0.09	0.43	-0.28	-0.10	0.31	0.90	0.08	-0.07	0.27	-0.11	0.19	0.11	0.18	-0.12	1.00	-0.03	-0.26	-0.27	0.44	0.43
Population	0.01	-0.03	-0.01	0.55	-0.03	-0.02	-0.13	0.24	-0.07	0.54	-0.04	-0.06	-0.03	-0.02	-0.03	1.00	0.24	0.24	-0.02	-0.04
thinness 1-19 years	-0.05	-0.47	0.30	0.46	-0.41	-0.25	-0.10	0.22	-0.53	0.47	-0.22	-0.27	-0.23	0.20	-0.26	0.24	1.00	0.94	-0.40	-0.44
thinness 5-9 years	-0.05	-0.46	0.30	0.47	-0.40	-0.25	-0.10	0.22	-0.54	0.47	-0.22	-0.27	-0.22	0.21	-0.27	0.24	0.94	1.00	-0.39	-0.43
Income composition of resources	0.24	0.69	-0.44	-0.14	0.42	0.37	0.12	-0.11	0.47	-0.16	0.35	0.15	0.36	-0.25	0.44	-0.02	-0.40	-0.39	1.00	0.80
Schooling	0.20	0.71	-0.43	-0.19	0.50	0.39	0.14	-0.12	0.50	-0.21	0.38	0.22	0.38	-0.22	0.43	-0.04	-0.44	-0.43	0.80	1.00

6. BMI and Alcohol

- **BMI ↔ Life expectancy (0.56)** → Higher BMI (to a limit) may indicate better nutrition, hence higher life expectancy.
- **Alcohol ↔ Life expectancy (0.39)** → Weak to moderate positive link; likely confounded by richer countries having both higher alcohol consumption and higher life expectancy.

Exploratory Data Analysis (EDA)

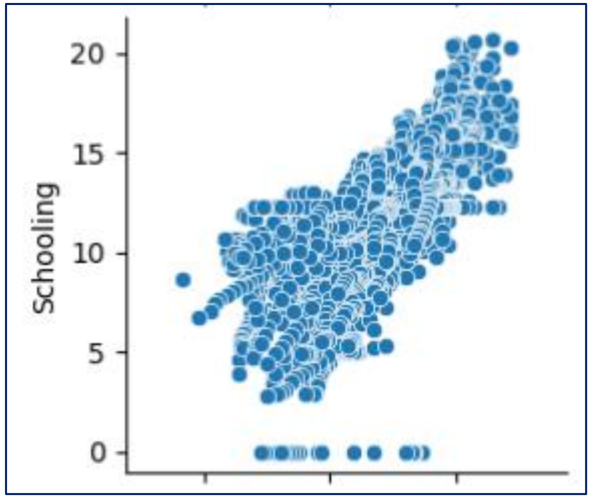


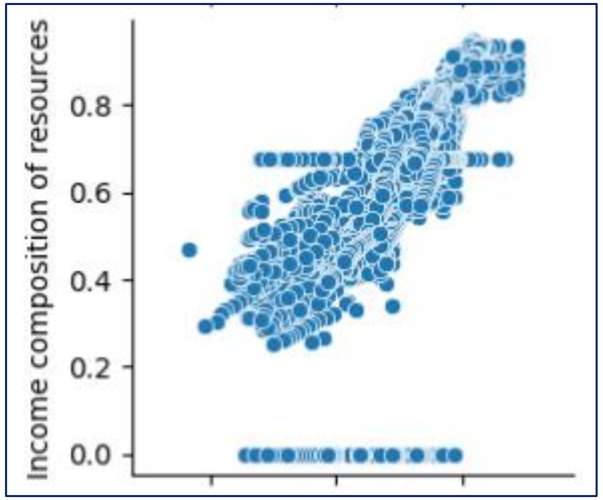
	Year	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	BMI	under-five deaths	Polio	Total expenditure	Diphtheria	HIV/AIDS	GDP	Population	thinness 1-19 years	thinness 5-9 years	ome composition of resources	Schooling
Year	1.00	0.17	-0.08	-0.04	-0.07	0.03	0.02	-0.08	0.11	-0.04	0.09	0.07	0.13	-0.14	0.09	0.01	-0.05	-0.05	0.24	0.20
Life expectancy	0.17	1.00	-0.70	-0.20	0.39	0.38	0.17	-0.16	0.56	-0.22	0.46	0.21	0.47	-0.56	0.43	-0.03	-0.47	-0.46	0.69	0.71
Adult Mortality	-0.08	-0.70	1.00	0.08	-0.19	-0.24	-0.12	0.03	-0.38	0.09	-0.27	-0.11	-0.27	0.52	-0.28	-0.01	0.30	0.30	-0.44	-0.43
infant deaths	-0.04	-0.20	0.08	1.00	-0.11	-0.09	-0.17	0.50	-0.23	1.00	-0.17	-0.13	-0.18	0.03	-0.10	0.55	0.46	0.47	-0.14	-0.19
Alcohol	-0.07	0.39	-0.19	-0.11	1.00	0.34	0.09	-0.05	0.31	-0.11	0.21	0.30	0.21	-0.05	0.31	-0.03	-0.41	-0.40	0.42	0.50
percentage expenditure	0.03	0.38	-0.24	-0.09	0.34	1.00	0.04	-0.06	0.23	-0.09	0.15	0.18	0.14	-0.10	0.90	-0.02	-0.25	-0.25	0.37	0.39
Hepatitis B	0.02	0.17	-0.12	-0.17	0.09	0.04	1.00	-0.07	0.11	-0.17	0.36	0.06	0.45	-0.09	0.08	-0.13	-0.10	-0.10	0.12	0.14
Measles	-0.08	-0.16	0.03	0.50	-0.05	-0.06	-0.07	1.00	-0.18	0.51	-0.14	-0.10	-0.14	0.03	-0.07	0.24	0.22	0.22	-0.11	-0.12
BMI	0.11	0.56	-0.38	-0.23	0.31	0.23	0.11	-0.18	1.00	-0.24	0.28	0.23	0.28	-0.24	0.27	-0.07	-0.53	-0.54	0.47	0.50
under-five deaths	-0.04	-0.22	0.09	1.00	-0.11	-0.09	-0.17	0.51	-0.24	1.00	-0.19	-0.13	-0.20	0.04	-0.11	0.54	0.47	0.47	-0.16	-0.21
Polio	0.09	0.46	-0.27	-0.17	0.21	0.15	0.36	-0.14	0.28	-0.19	1.00	0.13	0.67	-0.16	0.19	-0.04	-0.22	-0.22	0.35	0.38
Total expenditure	0.07	0.21	-0.11	-0.13	0.30	0.18	0.06	-0.10	0.23	-0.13	0.13	1.00	0.15	-0.00	0.11	-0.06	-0.27	-0.27	0.15	0.22
Diphtheria	0.13	0.47	-0.27	-0.18	0.21	0.14	0.45	-0.14	0.28	-0.20	0.67	0.15	1.00	-0.16	0.18	-0.03	-0.23	-0.22	0.36	0.38
HIV/AIDS	-0.14	-0.56	0.52	0.03	-0.05	-0.10	-0.09	0.03	-0.24	0.04	-0.16	-0.00	-0.16	1.00	-0.12	-0.02	0.20	0.21	-0.25	-0.22
GDP	0.09	0.43	-0.28	-0.10	0.31	0.90	0.08	-0.07	0.27	-0.11	0.19	0.11	0.18	-0.12	1.00	-0.03	-0.26	-0.27	0.44	0.43
Population	0.01	-0.03	-0.01	0.55	-0.03	-0.02	-0.13	0.24	-0.07	0.54	-0.04	-0.06	-0.03	-0.02	-0.03	1.00	0.24	0.24	-0.02	-0.04
thinness 1-19 years	-0.05	-0.47	0.30	0.46	-0.41	-0.25	-0.10	0.22	-0.53	0.47	-0.22	-0.27	-0.23	0.20	-0.26	0.24	1.00	0.94	-0.40	-0.44
thinness 5-9 years	-0.05	-0.46	0.30	0.47	-0.40	-0.25	-0.10	0.22	-0.54	0.47	-0.22	-0.27	-0.22	0.21	-0.27	0.24	0.94	1.00	-0.39	-0.43
Income composition of resources	0.24	0.69	-0.44	-0.14	0.42	0.37	0.12	-0.11	0.47	-0.16	0.35	0.15	0.36	-0.25	0.44	-0.02	-0.40	-0.39	1.00	0.80
Schooling	0.20	0.71	-0.43	-0.19	0.50	0.39	0.14	-0.12	0.50	-0.21	0.38	0.22	0.38	-0.22	0.43	-0.04	-0.44	-0.43	0.80	1.00

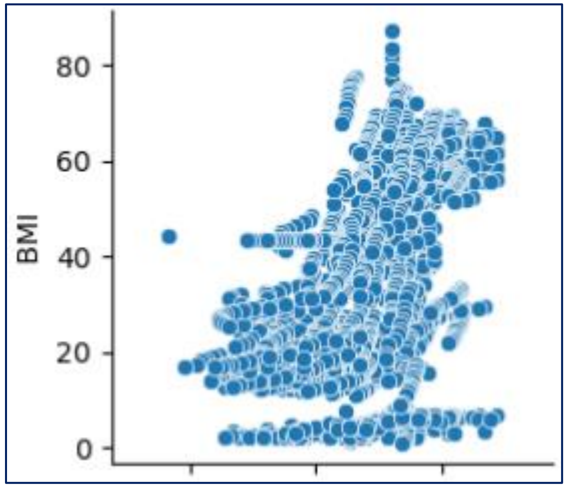
7. Year Variable

- **Year ↔ Life expectancy (0.17)** → Slight positive correlation, suggesting general improvement over time.

Exploratory Data Analysis (EDA)



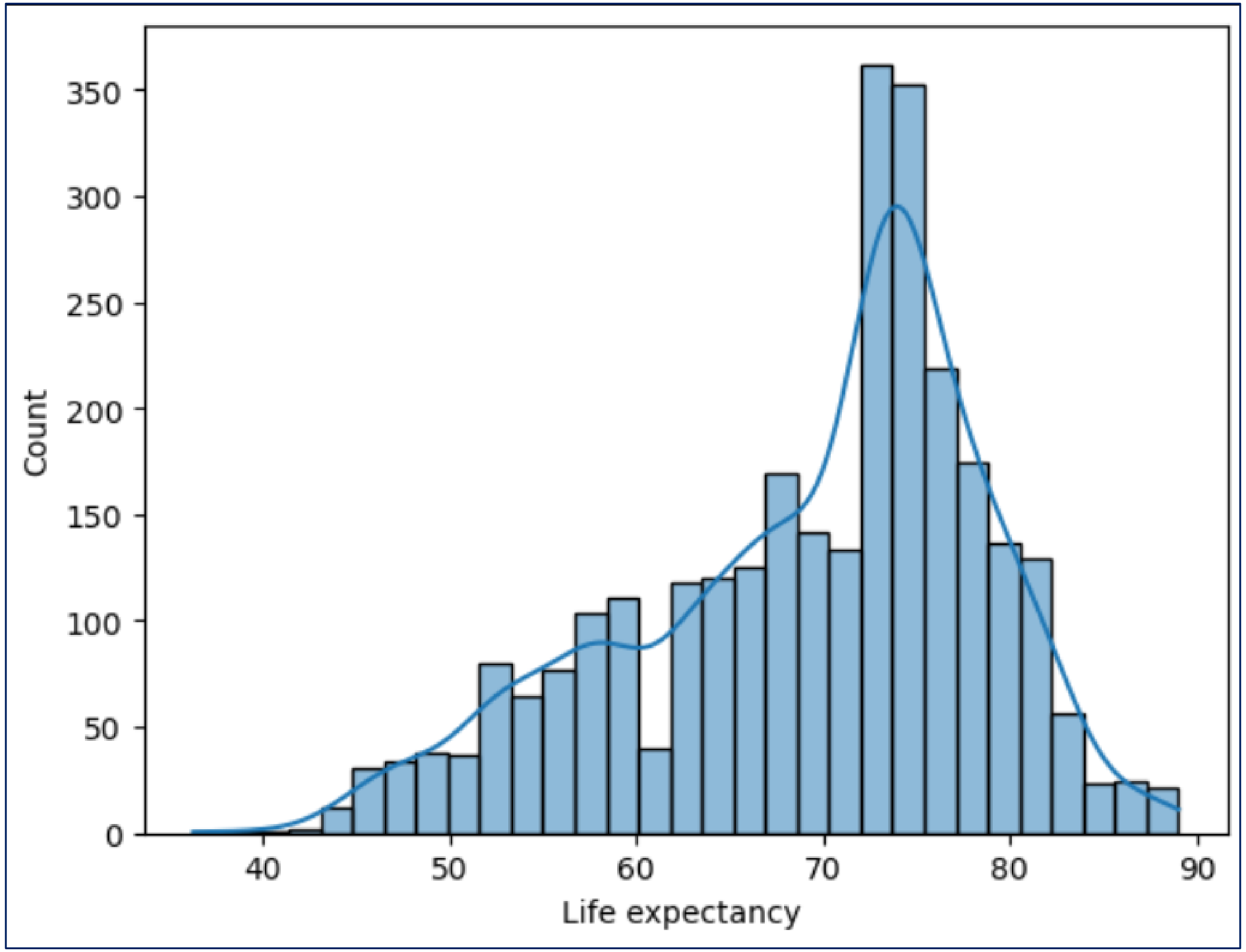




1. Life Expectancy Relationships

- **Strong Positive Relationship** with:
  - **(i) Schooling** → As average years of schooling increase, life expectancy rises almost linearly.
  - **(ii) Income composition of resources** → Countries with higher income equality and access to resources have higher life expectancy.
  - **(iii) BMI (moderate positive trend)** → Up to a healthy limit, better nutrition (reflected by BMI) is linked with higher life expectancy.

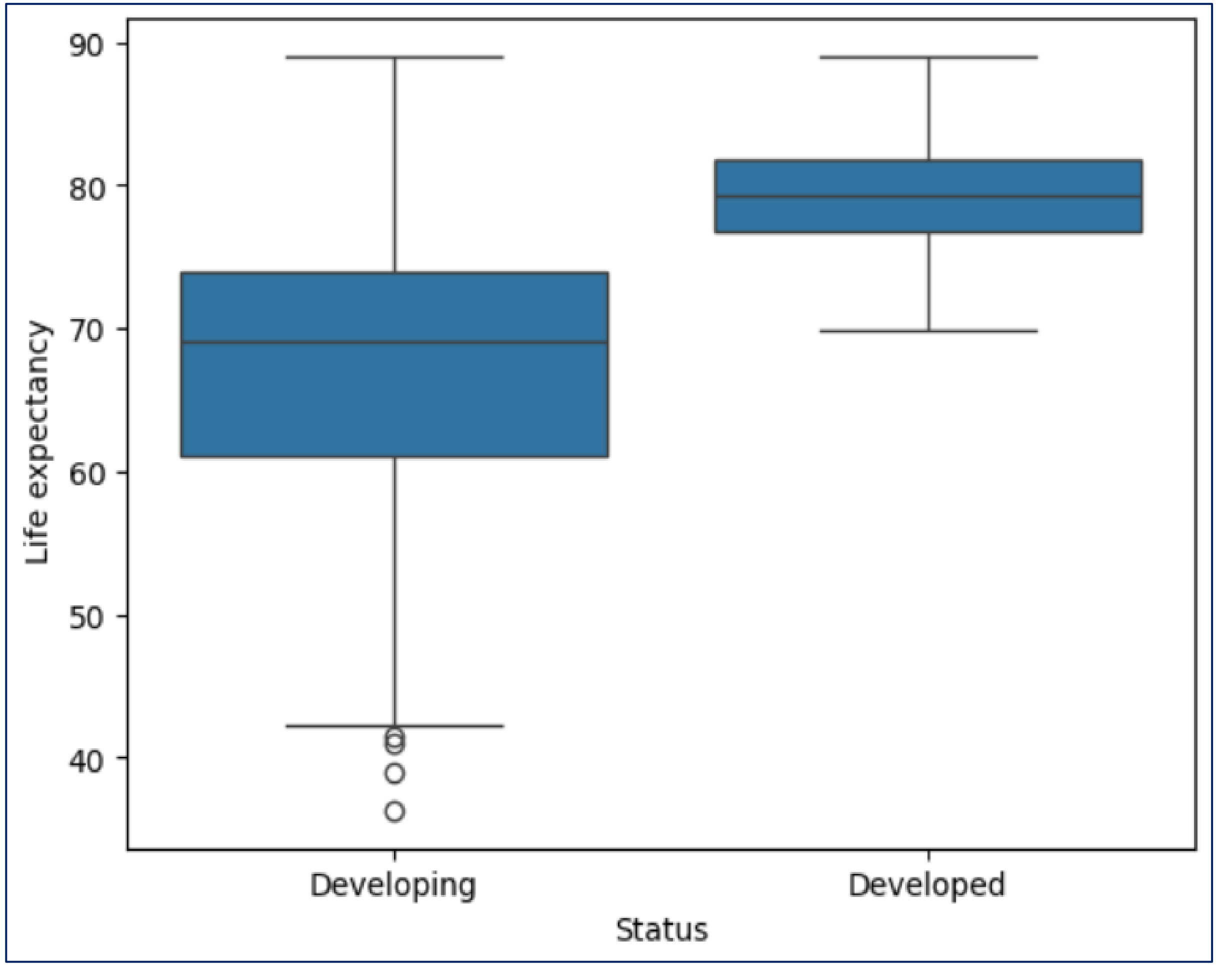
Exploratory Data Analysis (EDA)



Distribution Plots

- Life expectancy is slightly left-skewed, centered around 65–75 years.
- Developed nations cluster around 75–85 years.
- Developing nations show wider spread (40–70 years).

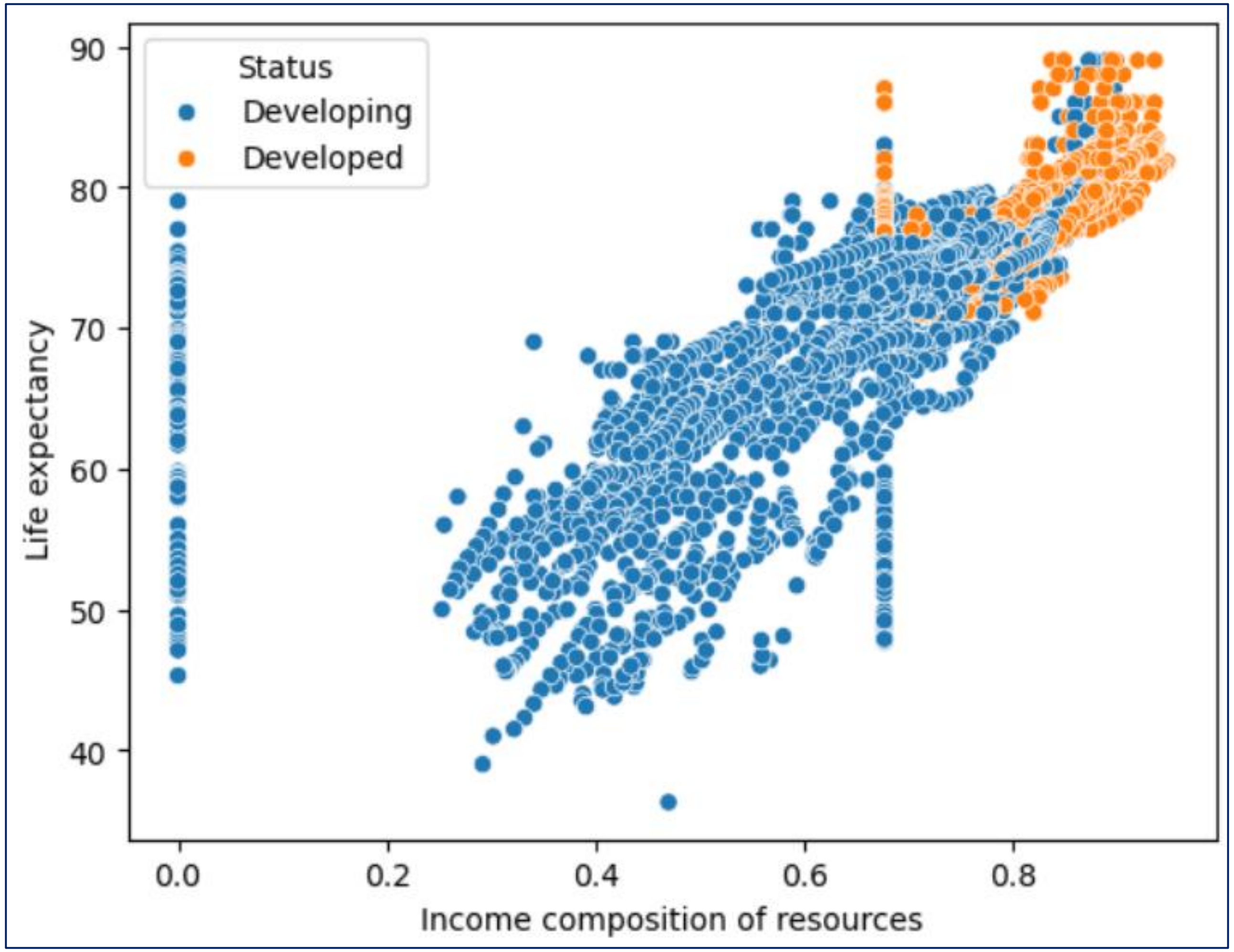
Exploratory Data Analysis (EDA)



Categorical Analysis

- Developed countries have significantly higher life expectancy.
- The gap is about 10–15 years between categories.

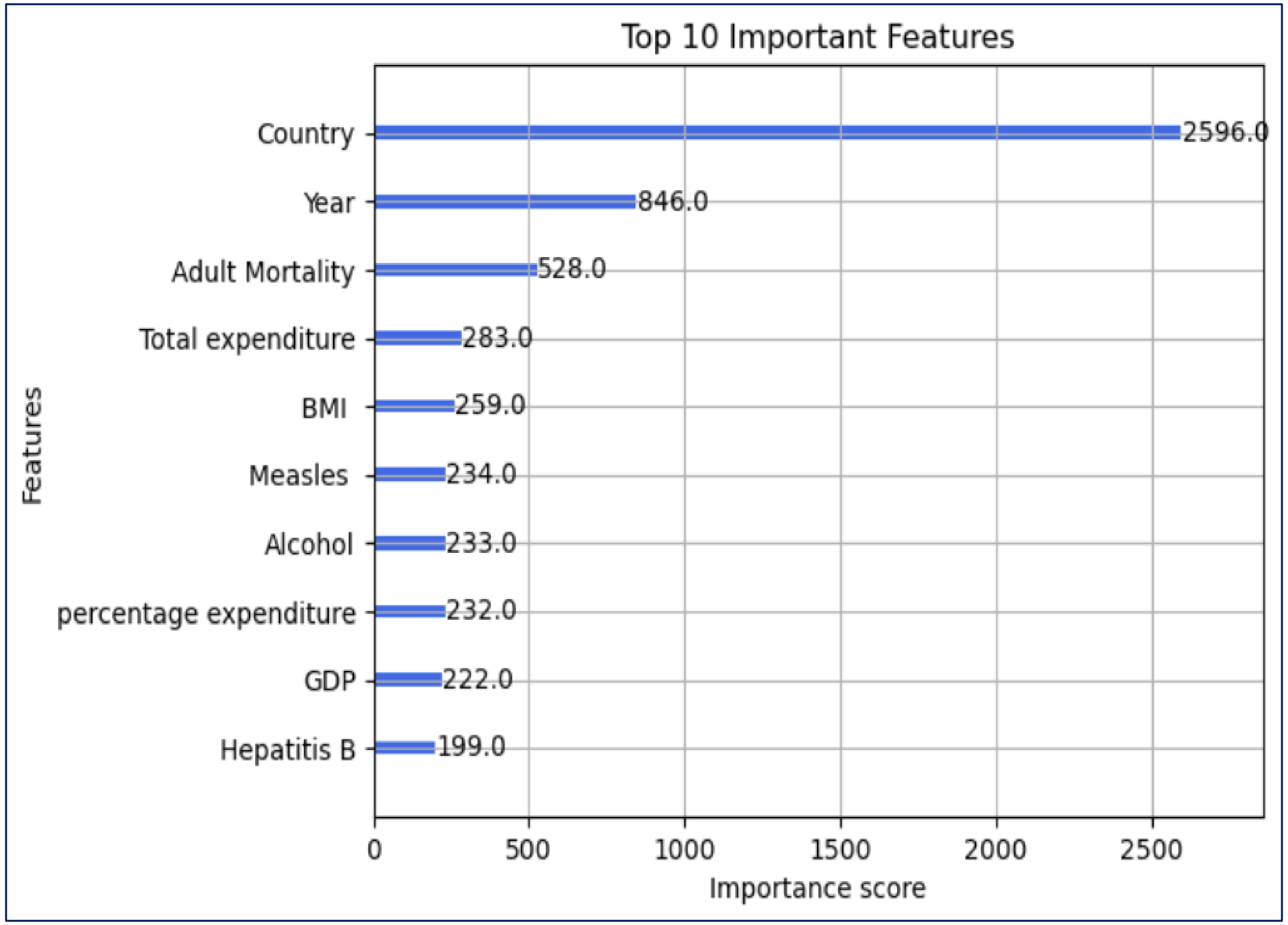
Exploratory Data Analysis (EDA)



Income & Education Effects

- Strong upward trend — higher income equality leads to longer lives.
- Developed countries cluster in top-right corner (high income, high life expectancy).

Exploratory Data Analysis (EDA)

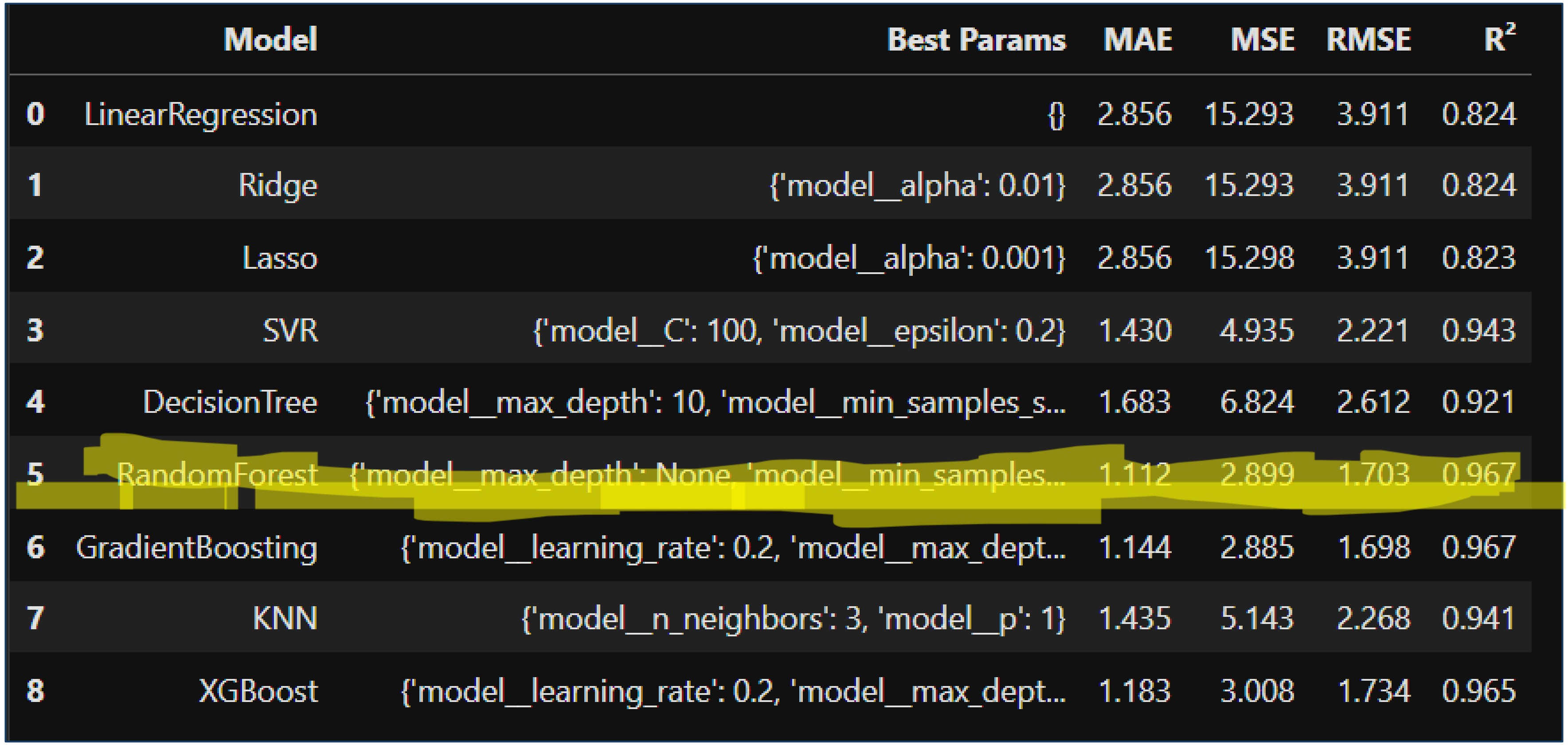


Top Contributing Features

- Country is the Highest contributing Feature to Life Expectancy in terms of importance score.
- Disease – Hepatitis B is the Lowest contributing Feature to Life Expectancy in terms of importance score.

Model Selection & Training

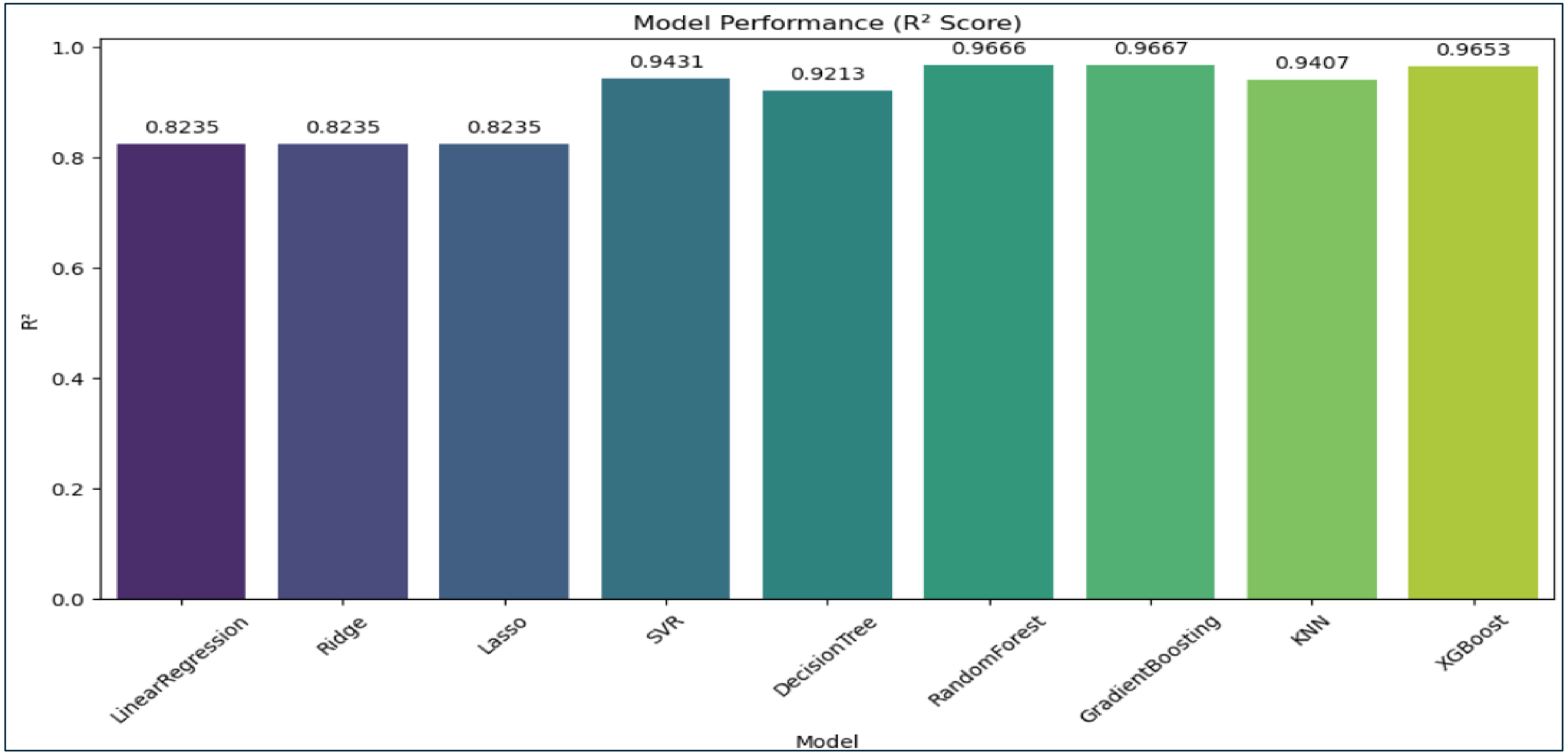
I selected the **RandomForest Regressor model** due to its efficiency in handling regression problems. The model is trained using the training dataset, and its performance is evaluated on the test set.



	Model	Best Params	MAE	MSE	RMSE	R^2
0	LinearRegression	{}	2.856	15.293	3.911	0.824
1	Ridge	{'model__alpha': 0.01}	2.856	15.293	3.911	0.824
2	Lasso	{'model__alpha': 0.001}	2.856	15.298	3.911	0.823
3	SVR	{'model__C': 100, 'model__epsilon': 0.2}	1.430	4.935	2.221	0.943
4	DecisionTree	{'model__max_depth': 10, 'model__min_samples_s...	1.683	6.824	2.612	0.921
5	RandomForest	{'model__max_depth': None, 'model__min_samples...	1.112	2.899	1.703	0.967
6	GradientBoosting	{'model__learning_rate': 0.2, 'model__max_dept...	1.144	2.885	1.698	0.967
7	KNN	{'model__n_neighbors': 3, 'model__p': 1}	1.435	5.143	2.268	0.941
8	XGBoost	{'model__learning_rate': 0.2, 'model__max_dept...	1.183	3.008	1.734	0.965

Model Selection & Training

RandomForest Regressor Model is the best model and top the leaderboard with $R^2 \approx 0.967$ and lowest RMSE (~1.7). This models is clearly capturing complex non-linear relationships.



Final Recommendations

- **Invest in education:** Implement policies to increase school enrollment and literacy rates and promote health awareness programs in schools to encourage healthy behaviors from an early age.
- **Enhance economic resources:** Introduce social welfare programs, subsidies, and skill development initiatives to improve income levels and access to essential resources for low-income populations.
- **Promote sustainable economic growth:** Encourage government and private sector investment in economic development projects that generate employment and fund public healthcare and social programs.
- **Reduce adult mortality:** Expand access to preventive healthcare, regular health screenings, and vaccination programs to lower the risk of early deaths.
- **Maintain healthy BMI ranges:** Launch national nutrition and fitness campaigns and provide community-level programs to promote balanced diets and regular physical activity.
- **Control HIV/AIDS prevalence:** Strengthen awareness campaigns, provide free or affordable HIV testing, and ensure availability of antiretroviral treatment to reduce prevalence.
- **Invest in healthcare expenditure:** Allocate higher budgets for healthcare infrastructure, train medical professionals, and ensure equitable access to medicines and healthcare services across regions.

Thankyou !