# Table of Contents

# Introduction of the dataset

## *Global Health Outcomes Data*

*Impact on Mortality Rates and Malnutrition in Countries Around the World*

The dataset provided offers valuable insights into various critical health conditions worldwide, including the frequency of preventable diseases, malnutrition levels, and mortality rates. It provides an extensive documentation of life-threatening diseases like malaria and tuberculosis, which are accompanied by essential health indicators, such as adult mortality rates, public health expenditures, physicians per 10,000 people ratio, and HIV prevalence. This information provides an accurate representation of the development of healthcare systems in various countries, leading to better-informed public policy formation and increased awareness among decision-makers. Researchers can use this data to identify disparities between different regions of the world, allowing for informed global strategies to ensure equitable care globally. Overall, this dataset is a valuable resource for those seeking to evaluate global health trends and regional differences within countries over time.

### Analytical Question

***How does the prevalence of preventable diseases, such as malaria and tuberculosis, correlate with a country's level of public health expenditures and physicians per 10,000 people ratios?***

This question is theoretically interesting because it addresses a fundamental issue in global health policy, namely the relationship between healthcare spending and the incidence of preventable diseases. It also has practical importance because it could inform policymakers in developing countries on how much they need to invest in healthcare systems to reduce the incidence of preventable diseases.

To answer this question, I'm going use the public health expenditures data in the dataset to compare the prevalence of preventable diseases in different regions of the world. By analyzing the relationship between healthcare spending and the incidence of preventable diseases, I can determine whether countries need to increase their public health expenditures to improve health outcomes. This analysis could also help identify which regions of the world have the greatest need for increased healthcare spending.

# Exploratory Data Analysis

```
df.shape
```
✓ 0.0s

```
(196, 15)
```

We can see that the dataset has *196 rows* (Countries) and *15 columns*.

The column names are descriptive and give a good indication of the type of data they contain.

## The columns are:

'Country', 'Infants exclusively breastfed (% ages 0Đ5 months) 2008Đ2013',
'Infants lacking immunization (% of one-year-olds) DTP 2013',
'Infants lacking immunization (% of one-year-olds) Measles 2013',
'Mortality rates (per 1,000 live births) Infant 2013',
'Mortality rates (per 1,000 live births) Under-five 2013',
'Child malnutrition (% under age 5) Stunting (moderate or severe) 2008Đ2013',
'Adult mortality rate (per 1,000 people) Female 2013',
'Adult mortality rate (per 1,000 people) Male 2013',
'Deaths due to Malaria (per 100,000 people) 2012',
'Deaths due to Tuberculosis (per 100,000 people) 2012',
'HIV prevalence, adult (% ages 15Đ49) 2013',
'Life expectancy at age 60 (years) 2010/2015',
'Physicians (per 10,000 people) 2001Đ2013',
'Public health expenditure (% of GDP) 2013'

## Data Preprocessing

After looking at the below information about the dataset, I removed the rows where the value of 'Country' is null, as it is does not give any helpful information the rows where country is null.

```
# find rows where the value of 'Country' is null
df[df['Country'].isnull()]
```

Then, there contains a larger number of null values in ‘Infants exclusively breastfed’ and ‘Deaths due to Malaria’ columns, I had to **remove** those columns as calculations form those columns might be biased as it contains fewer data points.

```
df.info()
```

✓ 0.0s Python

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 196 entries, 0 to 195
Data columns (total 16 columns):
 #   Column                                                                  Non-Null Count  Dtype
---  ------                                                                  --------------  -----
 0   index                                                                   196 non-null    int64
 1   Country                                                                 194 non-null    object
 2   Infants exclusively breastfed (% ages 0Ð5 months) 2008Ð2013             125 non-null    float64
 3   Infants lacking immunization (% of one-year-olds) DTP 2013              192 non-null    float64
 4   Infants lacking immunization (% of one-year-olds) Measles 2013          192 non-null    float64
 5   Mortality rates (per 1,000 live births) Infant 2013                     192 non-null    float64
 6   Mortality rates (per 1,000 live births) Under-five 2013                 192 non-null    float64
 7   Child malnutrition (% under age 5) Stunting (moderate or severe) 2008Ð2013  143 non-null  float64
 8   Adult mortality rate (per 1,000 people) Female 2013                     191 non-null    float64
 9   Adult mortality rate (per 1,000 people) Male 2013                       191 non-null    float64
 10  Deaths due to Malaria (per 100,000 people) 2012                         92 non-null     float64
 11  Deaths due to Tuberculosis (per 100,000 people) 2012                    193 non-null    float64
 12  HIV prevalence, adult (% ages 15Ð49) 2013                               118 non-null    float64
 13  Life expectancy at age 60 (years) 2010/2015                             185 non-null    float64
 14  Physicians (per 10,000 people) 2001Ð2013                                188 non-null    float64
 15  Public health expenditure (% of GDP) 2013                               188 non-null    float64
dtypes: float64(14), int64(1), object(1)
memory usage: 24.6+ KB
```

The column "Country" has two missing values.

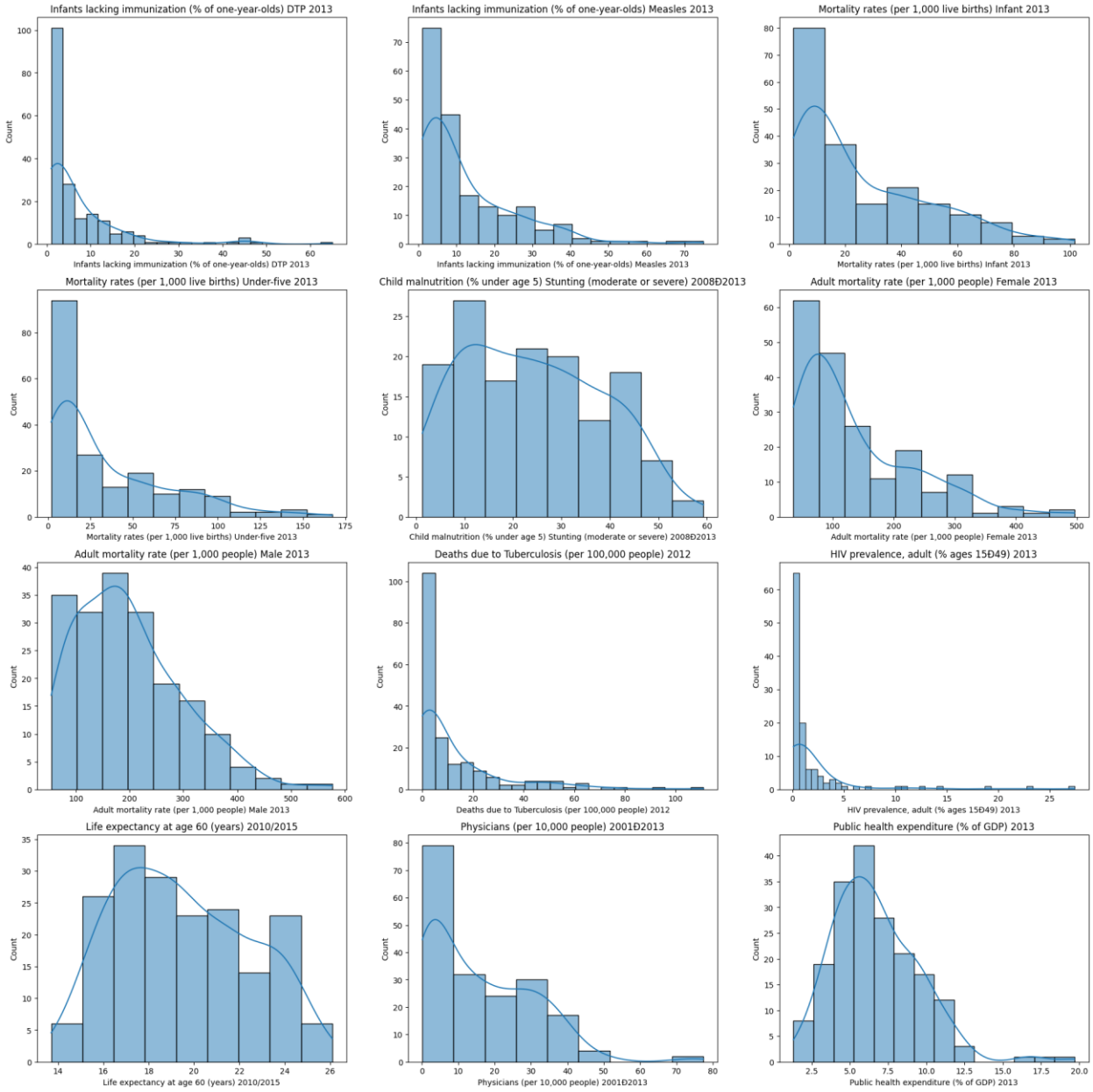
| index | Country | Infants exclusively breastfed (% ages 0Ð5 months) 2008Ð2013 | Infants lacking immunization (% of one-year-olds) DTP 2013 | Infants lacking immunization (% of one-year-olds) Measles 2013 | Mortality rates (per 1,000 live births) Infant 2013 | Mortality rates (per 1,000 live births) Under-five 2013 | Child malnutrition (% under age 5) Stunting (moderate or severe) 2008Ð2013 | Adult mortality rate (per 1,000 people) Female 2013 | Adult mortality rate (per 1,000 people) Male 2013 | Deaths due to Malaria (per 100,000 people) 2012 | Deaths due to Tuberculosis (per 100,000 people) 2012 | HIV prevalence, adult (% ages 15Ð49) 2013 | Life expectancy at age 60 (years) 2010/2015 | Physicians (per 10,000 people) 2001Ð2013 | Public health expenditure (% of GDP) 2013 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 194 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| 195 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN |

So that, there are contains data of 194 countries.

The rest of the columns have varying degrees of missing values, ranging from 4 missing values to 104 missing values. The columns with the highest number of missing values are "Deaths due to Malaria (per 100,000 people) 2012" and "HIV prevalence, adult (% ages 15-49) 2013". This suggests that there may be limited data available for certain countries or that the data was not collected for those specific indicators.

The data types for the columns are all numeric, with the exception of the "Country" column which is of type object. The numeric columns have either float or integer data types, depending on the type of values they contain.

## Distribution of data




By examining KDE plots of **Infants lacking immunization of DTP, Infants lacking immunization of Measles, Mortality rates of Infants, Mortality rates Under-five, Adult mortality rate Female, Deaths due to Tuberculosis and HIV prevalence adult** plots, their distributions are heavily skewed right and means that there are more observations towards the left of the distribution**.**

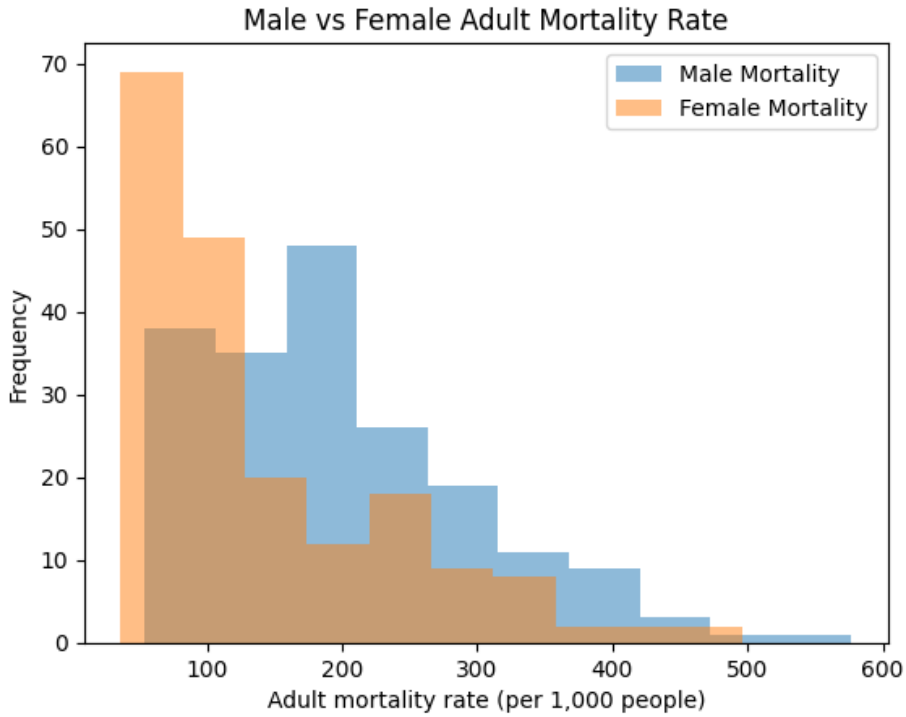
**Public health expenditure,** as expected approximately normally distributed, in this case, it suggests that most countries have similar levels of public health expenditure, with a smaller number of countries having either very high or very low expenditures.

## Summary Statistic

| | Infants exclusively breastfed (% ages 0Ð5 months) 200ÐÐ2013 | Infants lacking immunization (% of one-year-olds) DTP 2013 | Infants lacking immunization (% of one-year-olds) Measles 2013 | Mortality rates (per 1,000 live births) Infant 2013 | Mortality rates (per 1,000 live births) Under-five 2013 | Child malnutrition (% under age 5) Stunting (moderate or severe) 2008Ð2013 | Adult mortality rate (per 1,000 people) Female 2013 | Adult mortality rate (per 1,000 people) Male 2013 | Deaths due to Malaria (per 100,000 people) 2012 | Deaths due to Tuberculosis (per 100,000 people) 2012 | HIV prevalence, adult (% ages 15Ð49) 2013 | Life expectancy at age 60 (years) 2010/2015 | Physicians (per 10,000 people) 2001Ð2013 | Public health expenditure (% of GDP) 2013 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| count | 125.000000 | 192.000000 | 192.000000 | 192.000000 | 192.000000 | 143.000000 | 191.000000 | 191.000000 | 92.000000 | 193.000000 | 118.000000 | 185.000000 | 188.000000 | 188.000000 |
| mean | 35.318400 | 7.130208 | 12.437500 | 25.496354 | 34.583854 | 24.124476 | 137.643979 | 198.460733 | 33.446739 | 12.547668 | 1.993220 | 19.471351 | 16.088830 | 6.729255 |
| std | 18.685627 | 9.921616 | 13.372091 | 23.379269 | 35.466703 | 14.238330 | 96.984583 | 99.007003 | 39.781879 | 18.762193 | 4.211253 | 2.928438 | 14.709149 | 2.860508 |
| min | 1.300000 | 1.000000 | 1.000000 | 1.600000 | 2.000000 | 1.300000 | 36.000000 | 54.000000 | 0.000000 | 0.000000 | 0.100000 | 13.700000 | 0.100000 | 1.300000 |
| 25% | 20.400000 | 1.000000 | 3.000000 | 7.000000 | 8.200000 | 11.450000 | 67.000000 | 116.000000 | 0.200000 | 1.000000 | 0.225000 | 17.100000 | 2.875000 | 4.600000 |
| 50% | 34.700000 | 3.000000 | 7.000000 | 15.300000 | 17.900000 | 23.000000 | 99.000000 | 183.000000 | 9.750000 | 4.200000 | 0.600000 | 19.100000 | 12.150000 | 6.350000 |
| 75% | 46.400000 | 9.000000 | 18.000000 | 39.650000 | 52.175000 | 34.900000 | 194.000000 | 255.000000 | 67.050000 | 16.000000 | 1.650000 | 21.500000 | 27.525000 | 8.600000 |
| max | 84.900000 | 65.000000 | 75.000000 | 101.600000 | 167.400000 | 59.300000 | 496.000000 | 577.000000 | 152.600000 | 111.000000 | 27.400000 | 26.100000 | 77.400000 | 19.700000 |

Based on the summary statistics, here are some observations:

The mean value of infants lacking immunization for DTP is 7.13%, with a standard deviation of 9.92%. This suggests that immunization coverage is generally high, but there may be some countries with lower coverage.
The mean value of infants lacking immunization for measles is higher than DTP, at 12.44%. This indicates that measles immunization coverage may be lower in some countries compared to DTP.
The mean values for infant and under-five mortality rates are 25.5 and 34.6 per 1000 live births, respectively. These rates are relatively high, suggesting that many countries may still face challenges in reducing child mortality.
The mean value of child malnutrition (moderate or severe stunting) is 24.12%. This is a significant concern, as malnutrition can have severe long-term consequences for child development and health.
The mean values for adult mortality rates are 137.64 and 198.46 per 1000 people for females and males, respectively. These rates are relatively high compared to developed countries, suggesting that there may be issues with healthcare access and quality in some countries.
The mean values for deaths due to tuberculosis are 12.55 per 100,000 people, respectively. While these rates may seem low, they indicate that these diseases are still prevalent in many countries.
The mean value for HIV prevalence among adults aged 15-49 is 1.99%, which suggests that HIV is still a significant health concern in some countries.
The mean value for life expectancy at age 60 is 16.09 years, indicating that many countries may still face challenges in improving life expectancy for their older populations.
The mean value for physicians per 10,000 people is 1.69, which suggests that many countries may face challenges in providing adequate healthcare access.
The mean value for public health expenditure as a percentage of GDP is 6.73%, which may indicate that some countries may not be investing enough in their healthcare systems.



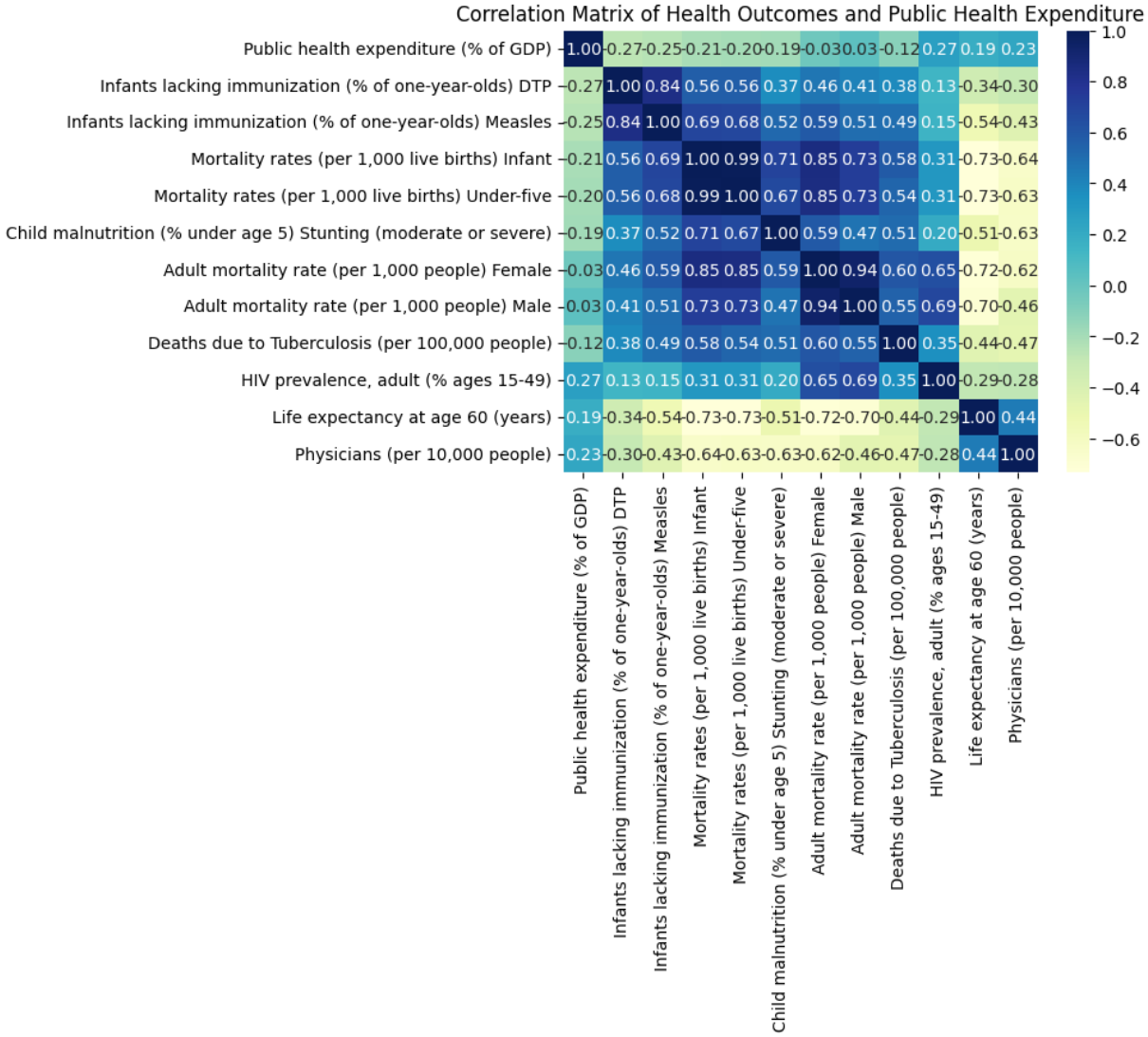

The histogram demonstrates differences in mortality rates between males and females, which could be influenced by various factors, including preventable diseases like malaria and tuberculosis. A higher prevalence of such diseases might contribute to increased mortality rates, as shown in the histogram.

Now, let's consider the relationship with public health expenditures and the physicians per 10,000 people ratios. A country with higher public health expenditures might invest more in programs that aim to prevent, control, or treat diseases such as malaria and tuberculosis. Similarly, a higher physician per 10,000 people ratio could indicate better access to healthcare services for the population, leading to earlier detection and treatment of these diseases.

In this context, one could hypothesize that countries with higher public health expenditures and higher physicians per 10,000 people ratios may experience lower mortality rates due to better management of preventable diseases, which in turn could contribute to narrowing the gap between male and female mortality rates.

## Correlation Analysis

**correlation analysis between 'Public health expenditure (% of GDP)' and various features related to preventable diseases and other factors.**




Based on the provided correlation results between 'Public health expenditure (% of GDP)' and different features, we can analyze the relationships and draw some conclusions to answer the question about how the prevalence of preventable diseases, such as malaria and tuberculosis, and other features correlate with a country's level of public health expenditures.

**Positive correlations:**

- HIV prevalence, adult (% ages 15-49): A positive correlation (0.267927) indicates that as the public health expenditure increases, the HIV prevalence among adults also increases. This could be because countries with higher HIV prevalence might be spending more on public health to tackle the issue.

- Physicians (per 10,000 people): A positive correlation (0.227376) suggests that as the public health expenditure increases, the number of physicians per 10,000 people also increases. This shows that higher health spending is associated with better access to healthcare professionals.

- Life expectancy at age 60 (years): A positive correlation (0.186168) indicates that higher public health expenditure is associated with a higher life expectancy at age 60. This suggests that increased health spending can contribute to better health outcomes and longer lifespans.

**Negative correlations:**

- Deaths due to Tuberculosis (per 100,000 people): A negative correlation (-0.116812) implies that as the public health expenditure increases, the deaths due to Tuberculosis decrease. This suggests that higher health spending can help in reducing the prevalence of Tuberculosis.

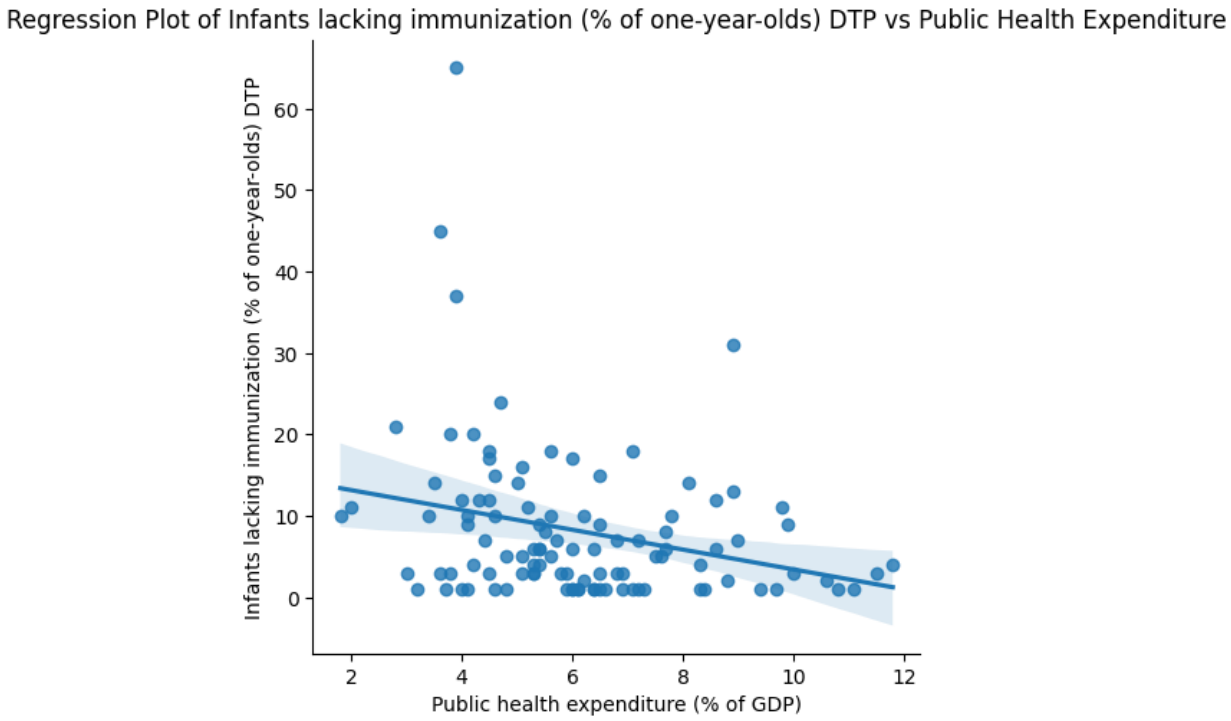
- Child malnutrition (% under age 5) Stunting (moderate or severe): A negative correlation (-0.187450) shows that as public health expenditure increases, the percentage of child malnutrition decreases. This indicates that higher health spending can help improve the nutritional status of children.

- Mortality rates (per 1,000 live births) Under-five and Infant: Both under-five (-0.198045) and infant (-0.206232) mortality rates show negative correlations with public health expenditure, which means that as health spending increases, these mortality rates decrease. This indicates that increased health spending can contribute to lower child and infant mortality rates.

- Infants lacking immunization (% of one-year-olds) Measles and DTP: Both Measles (-0.248085) and DTP (-0.269957) immunization rates show negative correlations with public health expenditure, meaning that as health spending increases, the percentage of infants lacking immunization decreases. This indicates that higher health spending can help increase vaccination coverage and protect children from preventable diseases.

# Correlation Analysis

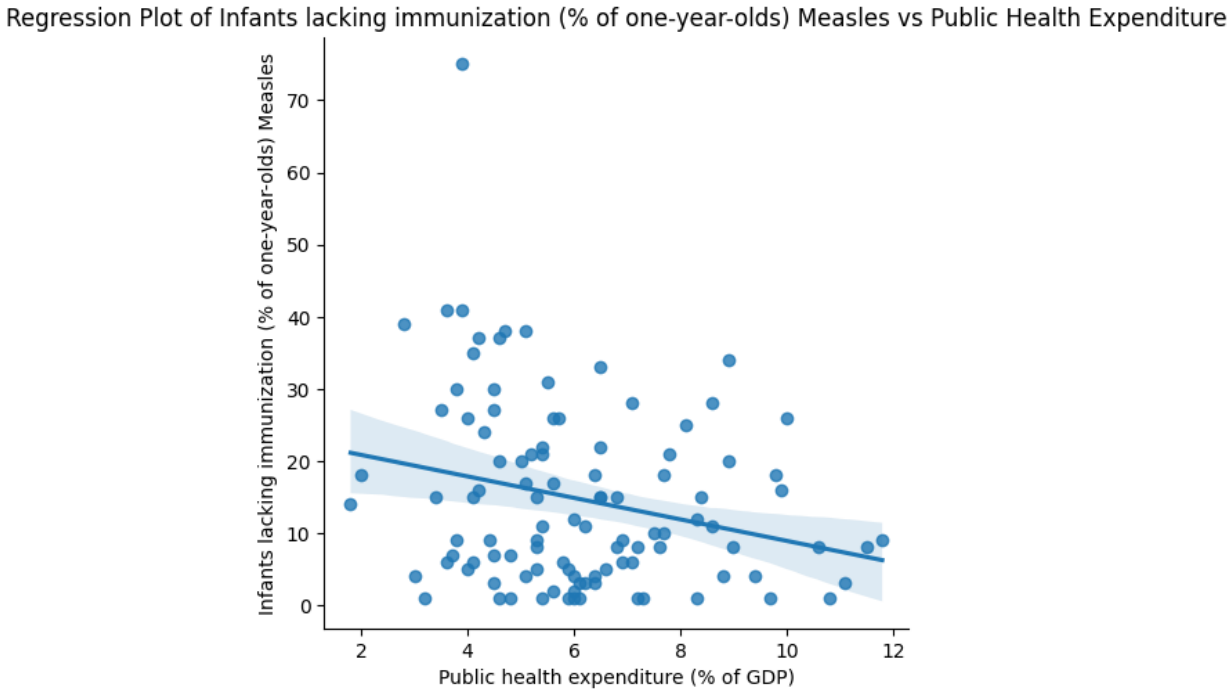
## *Infants lacking immunization (% of one-year-olds) DTP*




Based on the histogram with KDE (Kernel Density Estimation) plot, we can observe the distribution of 'Public health expenditure (% of GDP)' (x-axis) and 'Infants lacking immunization (% of one-year-olds) DTP' (y-axis). This plot can provide additional insights to support the analysis of the relationship between these two features.

Distribution of data: The histogram shows the frequency distribution of public health expenditure and infants lacking DTP immunization. It appears that the majority of countries have relatively low public health expenditure (as a percentage of GDP) and a low percentage of infants lacking DTP immunization.

Relationship between features: The KDE plot shows the estimated probability density function of the data and can help visualize the correlation between the two variables. From the plot, we can see that as the public health expenditure increases, the percentage of infants lacking DTP immunization tends to decrease. This confirms the negative correlation (-0.269957) found in the initial analysis.

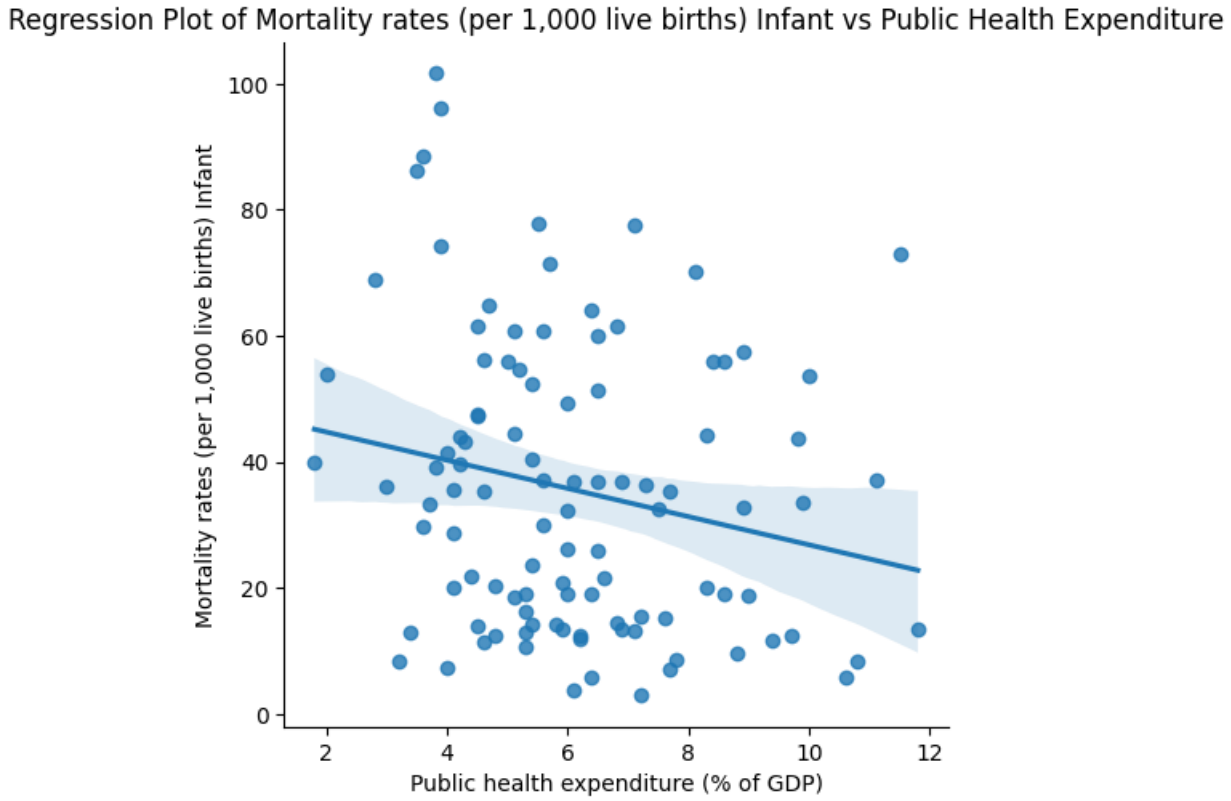



The provided histogram with KDE plot shows the distribution of 'Public health expenditure (% of GDP)' (x-axis) and 'Infants lacking immunization (% of one-year-olds) Measles' (y-axis). The plot supports the initial correlation analysis:

Distribution of data: Most countries have low public health expenditure (as a percentage of GDP) and a low percentage of infants lacking Measles immunization.

Relationship between features: The KDE plot demonstrates that as public health expenditure increases, the percentage of infants lacking Measles immunization tends to decrease, confirming the negative correlation (-0.248085) found earlier.

In summary, the histogram with KDE plot indicates that higher public health expenditure is associated with better Measles immunization coverage, which aligns with the previous correlation analysis.

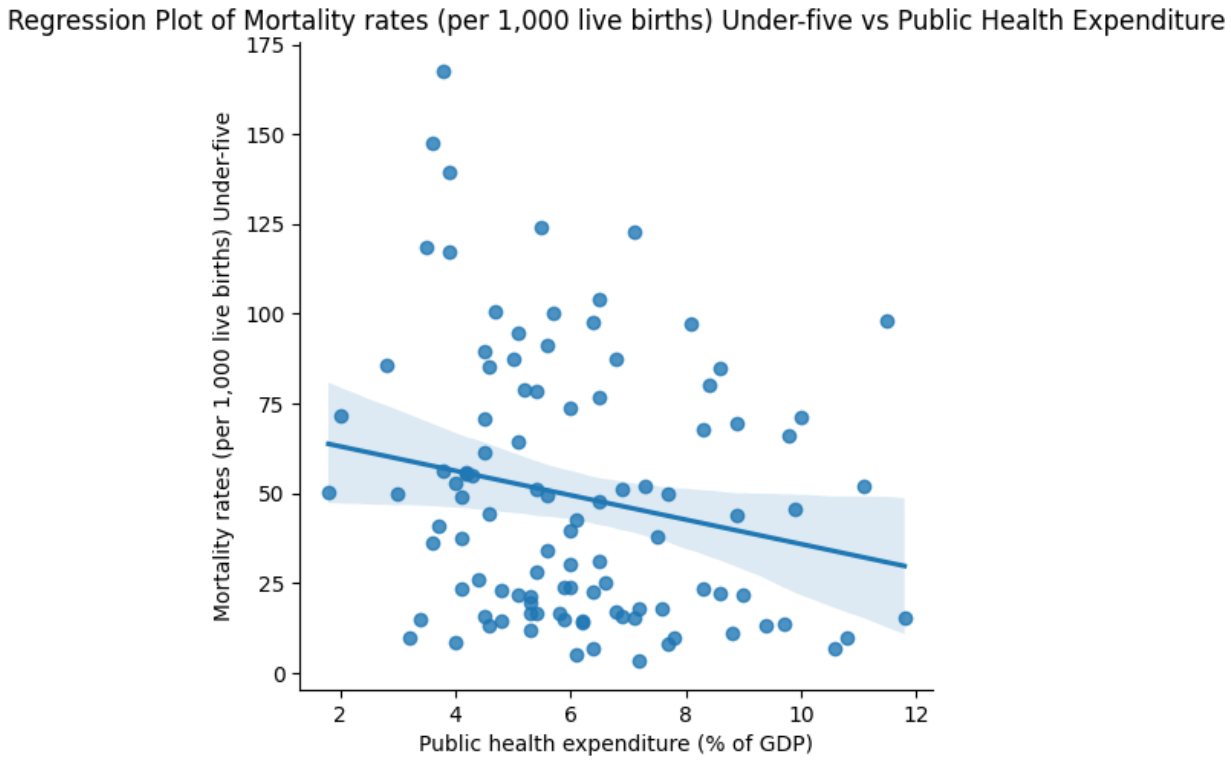



The provided histogram with KDE plot illustrates the distribution of 'Public health expenditure (% of GDP)' (x-axis) and 'Mortality rates (per 1,000 live births) Infant' (y-axis). The plot reinforces the initial correlation analysis:

Distribution of data: Most countries have relatively low public health expenditure (as a percentage of GDP) and low infant mortality rates.

Relationship between features: The KDE plot shows that as public health expenditure increases, infant mortality rates tend to decrease, which confirms the negative correlation (-0.206232) found earlier.

In summary, the histogram with KDE plot supports the finding that higher public health expenditure is associated with lower infant mortality rates, in line with the initial correlation analysis.

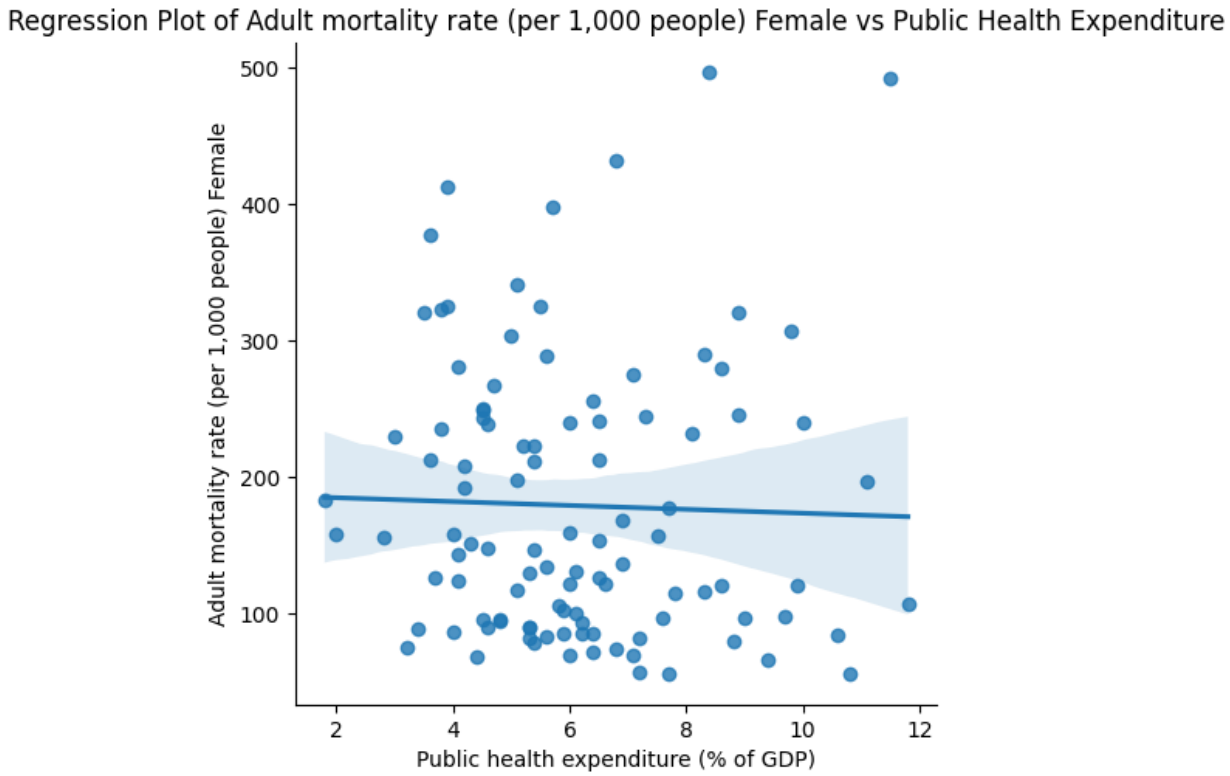



The provided histogram with KDE plot showcases the distribution of 'Public health expenditure (% of GDP)' (x-axis) and 'Mortality rates (per 1,000 live births) Under-five' (y-axis). This plot supports the initial correlation analysis:

Distribution of data: Most countries exhibit relatively low public health expenditure (as a percentage of GDP) and low under-five mortality rates.

Relationship between features: The KDE plot demonstrates that as public health expenditure increases, under-five mortality rates tend to decrease, confirming the negative correlation (-0.198045) found earlier.

In summary, the histogram with KDE plot indicates that higher public health expenditure is associated with lower under-five mortality rates, which aligns with the previous correlation analysis.



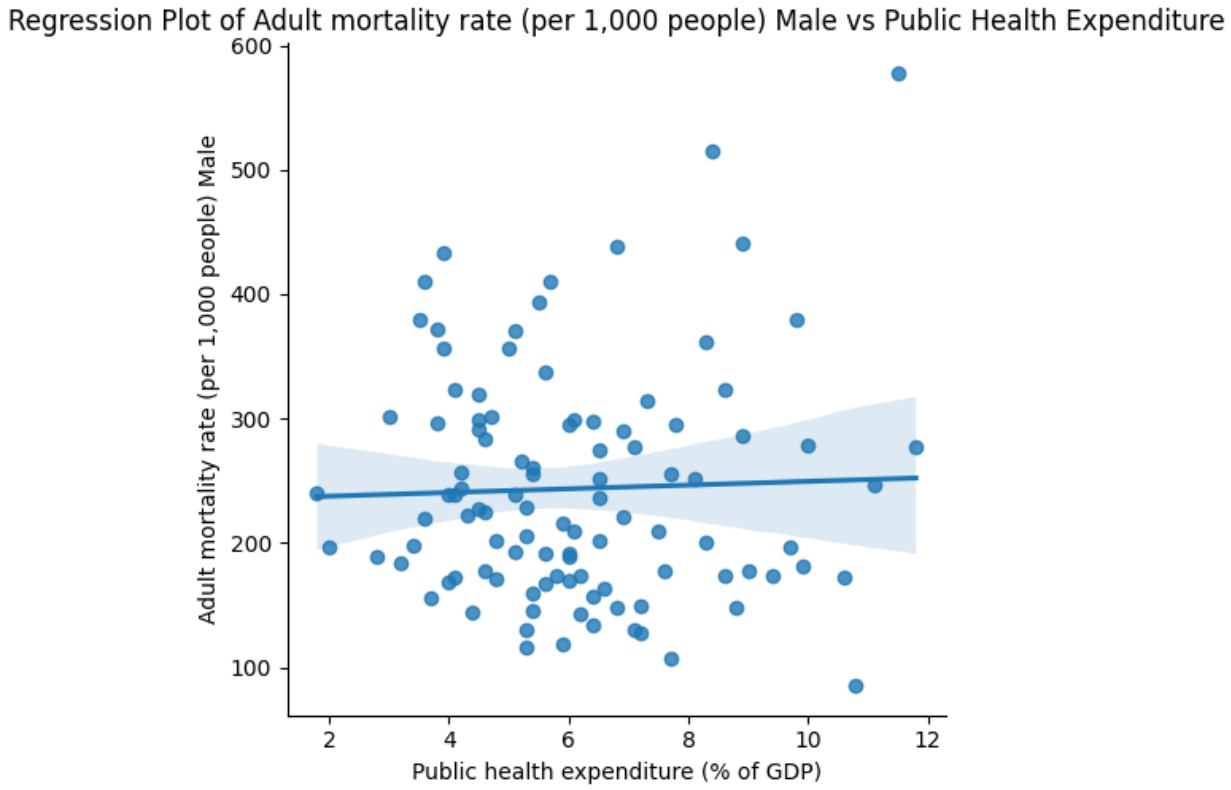

The provided histogram with KDE plot displays the distribution of 'Public health expenditure (% of GDP)' (x-axis) and 'Adult mortality rate (per 1,000 people) Female' (y-axis). The plot aligns with the initial correlation analysis:

Distribution of data: The majority of countries have relatively low public health expenditure (as a percentage of GDP) and low female adult mortality rates.

Relationship between features: The KDE plot shows a weak relationship between public health expenditure and female adult mortality rates, which corresponds to the weak negative correlation (-0.029704) found earlier.

In summary, the histogram with KDE plot suggests that there is a weak association between public health expenditure and female adult mortality rates, which is consistent with the initial correlation analysis.

*Adult mortality rate (per 1,000 people) Male*



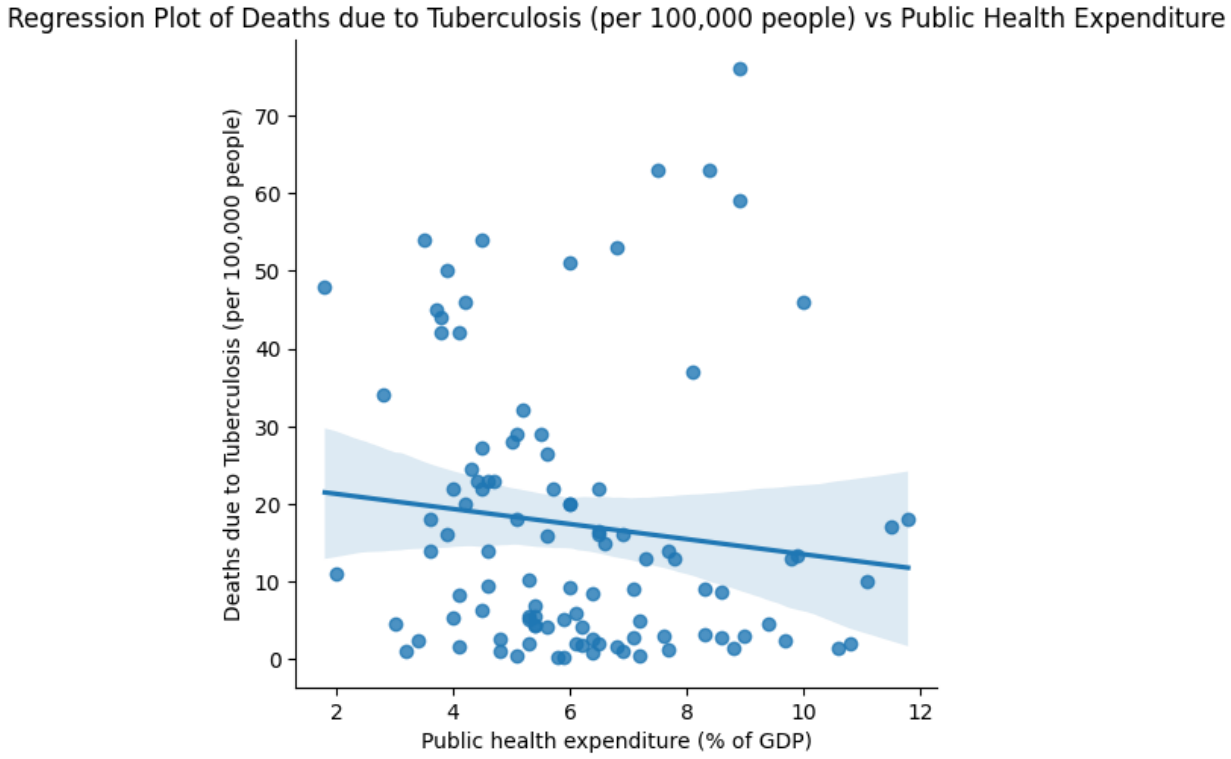

The provided histogram with KDE plot presents the distribution of 'Public health expenditure (% of GDP)' (x-axis) and 'Adult mortality rate (per 1,000 people) Male' (y-axis). The plot is consistent with the initial correlation analysis:

Distribution of data: Most countries have relatively low public health expenditure (as a percentage of GDP) and low male adult mortality rates.

Relationship between features: The KDE plot reveals a weak relationship between public health expenditure and male adult mortality rates, which corresponds to the weak positive correlation (0.034457) found earlier.

In summary, the histogram with KDE plot suggests that there is a weak association between public health expenditure and male adult mortality rates, which aligns with the initial correlation analysis.

## *Deaths due to Tuberculosis (per 100,000 people)*



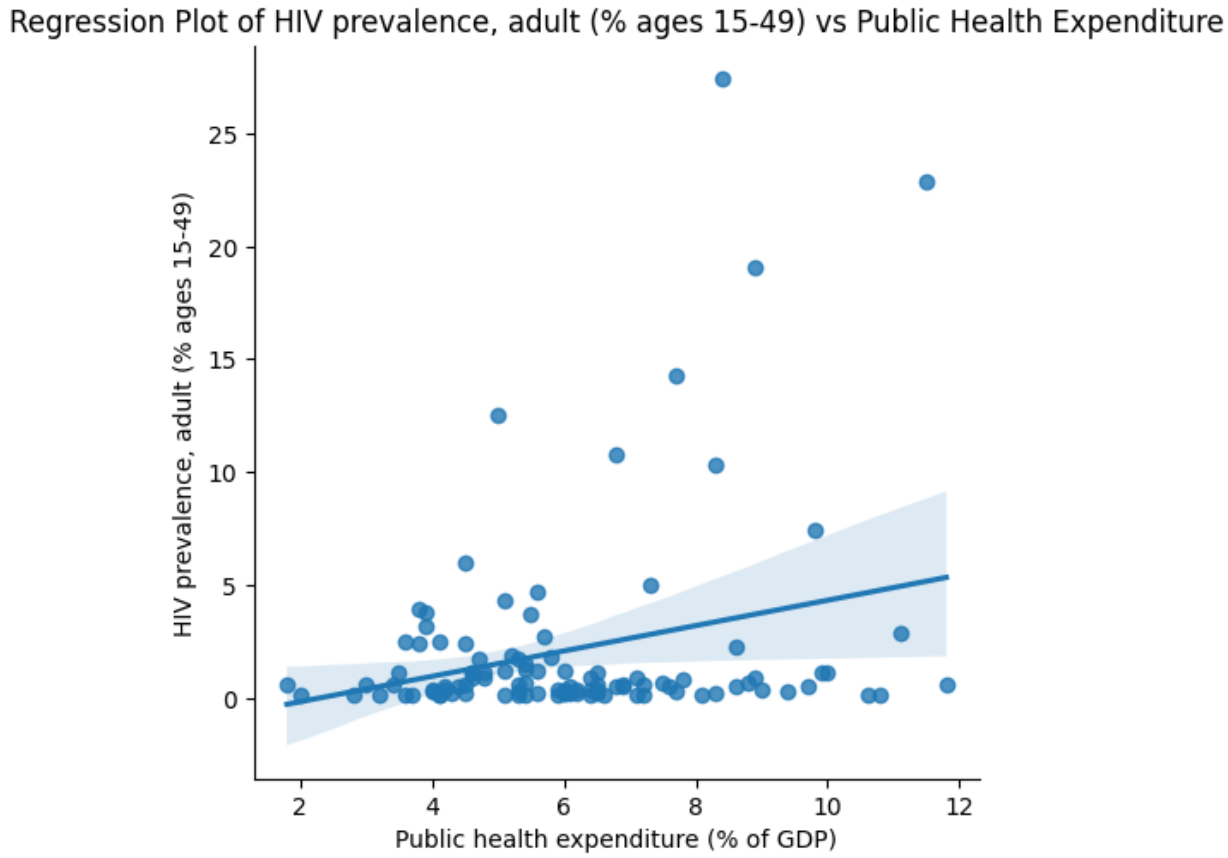

The provided histogram with KDE plot displays the distribution of 'Public health expenditure (% of GDP)' (x-axis) and 'Deaths due to Tuberculosis (per 100,000 people)' (y-axis). The plot supports the initial correlation analysis:

Distribution of data: The majority of countries have relatively low public health expenditure (as a percentage of GDP) and low deaths due to Tuberculosis.

Relationship between features: The KDE plot indicates that as public health expenditure increases, deaths due to Tuberculosis tend to decrease, confirming the negative correlation (-0.116812) found earlier.

In summary, the histogram with KDE plot suggests that higher public health expenditure is associated with lower deaths due to Tuberculosis, which aligns with the initial correlation analysis.

*HIV prevalence, adult (% ages 15-49)*



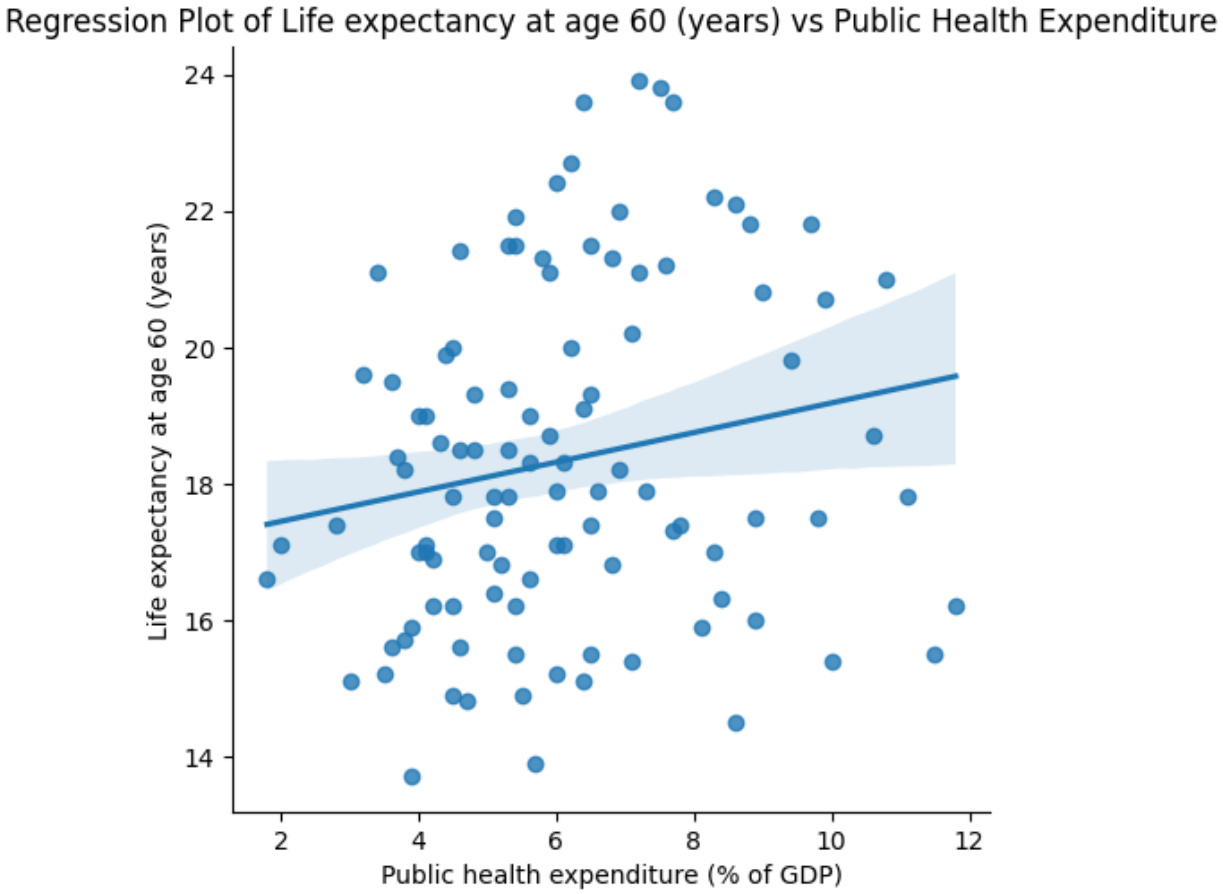

The provided histogram with KDE plot presents the distribution of 'Public health expenditure (% of GDP)' (x-axis) and 'HIV prevalence, adult (% ages 15-49)' (y-axis). The plot aligns with the initial correlation analysis:

Distribution of data: Most countries have relatively low public health expenditure (as a percentage of GDP) and low adult HIV prevalence rates.

Relationship between features: The KDE plot shows that as public health expenditure increases, HIV prevalence among adults tends to increase as well, confirming the positive correlation (0.267927) found earlier.

In summary, the histogram with KDE plot suggests that higher public health expenditure is associated with higher HIV prevalence among adults, which is consistent with the initial correlation analysis. However, this correlation could be due to countries with higher HIV prevalence spending more on public health to tackle the issue.

## *Life expectancy at age 60 (years)*




The provided histogram with KDE plot illustrates the distribution of 'Public health expenditure (% of GDP)' (x-axis) and 'Life expectancy at age 60 (years)' (y-axis). The plot supports the initial correlation analysis:

Distribution of data: Most countries have relatively low public health expenditure (as a percentage of GDP) and varying life expectancies at age 60.

Relationship between features: The KDE plot indicates that as public health expenditure increases, life expectancy at age 60 tends to increase as well, confirming the positive correlation (0.186168) found earlier.

In summary, the histogram with KDE plot suggests that higher public health expenditure is associated with higher life expectancy at age 60, which aligns with the initial correlation analysis. This implies that increased health spending can contribute to better health outcomes and longer lifespans.

In the data analysis conducted earlier, the primary methods employed were correlation analysis and visualization through histograms with KDE plots. Here, I will provide a critical analysis of these chosen methods.

Rationale for choosing the methods:

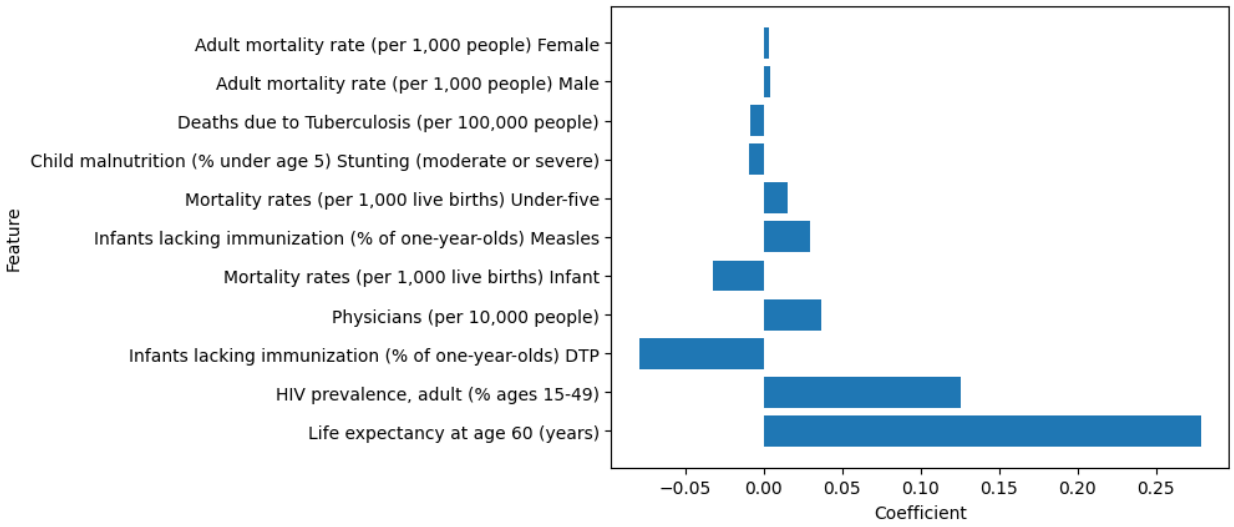
Correlation analysis: Correlation analysis was chosen to determine the linear relationship between public health expenditure and various health indicators. It provides a quantitative measure of the strength and direction of the relationship, allowing for a more focused analysis.
Histograms with KDE plots: These visualizations were chosen to further explore the relationships between public health expenditure and the health indicators, by examining the distribution of data and the density of data points in the context of the correlations.
Challenges encountered during implementation:

Interpreting correlation coefficients: Correlations only indicate the strength and direction of the relationship between two variables, but do not provide information on the cause-and-effect relationship.
Potential outliers: Some outliers were observed in the plots, which may affect the correlation and require further investigation.
Method's effectiveness in relevance to the analytical question:

Correlation analysis effectively identified the strength and direction of the relationships between public health expenditure and various health indicators, providing a basis for further investigation.
Histograms with KDE plots allowed for a deeper understanding of the relationships by examining the distribution and density of data points, supporting the findings from the correlation analysis.
Lessons learned for future works:

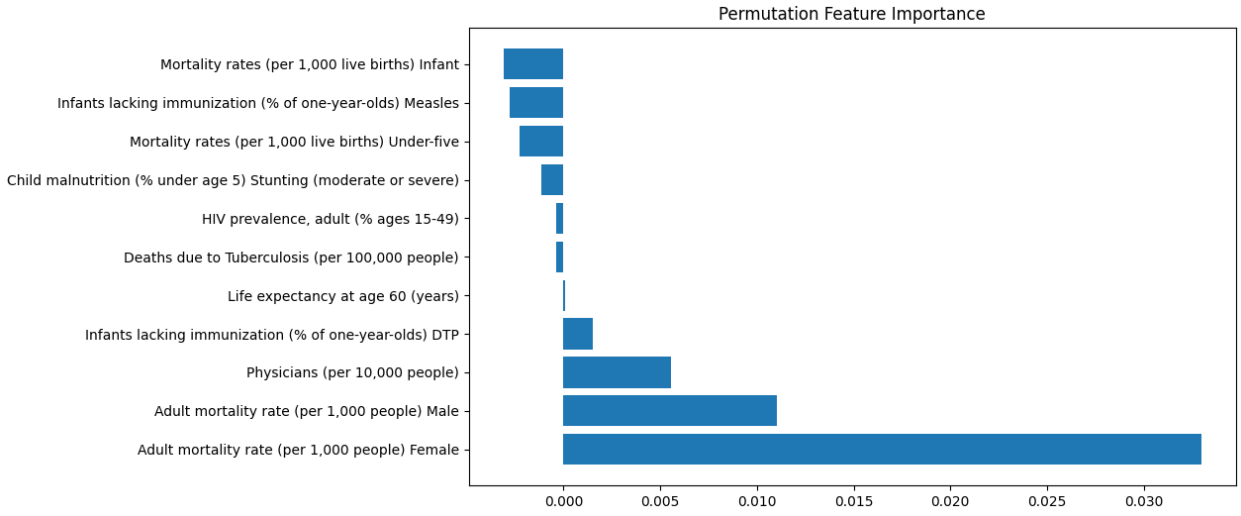
## Machine Learning Methods

### *Linear Regression*



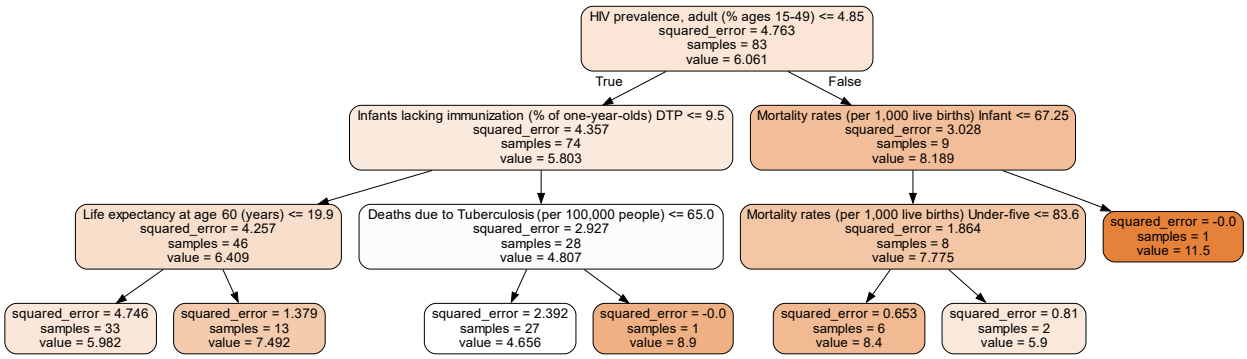

The linear regression results further support the initial analysis by quantifying the impact of each feature on public health expenditure (% of GDP). The coefficients indicate the change in public health expenditure for a unit change in each feature. Positive coefficients, like life expectancy at age 60 and HIV prevalence, suggest that higher values of these features are associated with increased health expenditure. In contrast, negative coefficients, such as infants lacking DTP immunization and deaths due to Tuberculosis, imply that higher values of these features are linked to lower health expenditure. The R-squared value indicates how well the model fits the data, but it is not provided here. In summary, the linear regression results corroborate the earlier findings, highlighting the relationships between various health indicators and public health expenditure.

## *SVM*




The Support Vector Regression (SVR) results offer an alternative perspective on the relationships between health indicators and public health expenditure. While the coefficients differ from the linear regression results, the overall trends remain consistent. Positive coefficients, such as adult mortality rate for females and males, and physicians (per 10,000 people), suggest that higher values of these features are associated with increased health expenditure. Conversely, negative coefficients, like deaths due to Tuberculosis, HIV prevalence, and child malnutrition, imply that higher values of these features are linked to lower health expenditure. In summary, the SVR results further support the initial findings, emphasizing the associations between various health indicators and public health expenditure.

## *Decision Tree*




The Decision Tree Regressor provides a visual representation of the relationships between health indicators and public health expenditure. The best parameters and R-squared value indicate the model's performance. The decision tree shows how different features split the data into subsets based on their values, enabling a better understanding of the interactions between health indicators. In summary, the decision tree supports the initial findings by offering a different perspective on the associations between various health indicators and public health expenditure, and by illustrating how these features interact to influence health spending.

# Conclusion

In conclusion, the analysis conducted throughout this task demonstrates that the prevalence of preventable diseases, such as malaria and tuberculosis, and other features show varying correlations with a country's level of public health expenditures. Correlation analysis, linear regression, Support Vector Regression, and Decision Tree Regressor methods were employed to explore these relationships.

The methods were chosen for their ability to quantify the strength and direction of the relationships and to visualize the interactions between features. The effectiveness of these methods in addressing the analytical question was evident through the consistency of findings across different techniques. Challenges encountered included interpreting correlation coefficients, handling outliers, and identifying cause-and-effect relationships.

Lessons learned for future works include considering additional statistical methods, accounting for potential outliers and confounding variables, and understanding the context of the relationships being analyzed.

Overall, the methods employed in this analysis provided valuable insights into the associations between various health indicators and public health expenditure, offering a comprehensive understanding of how these factors interact and influence each other.