**AIML ASSIGNMENT – 2**

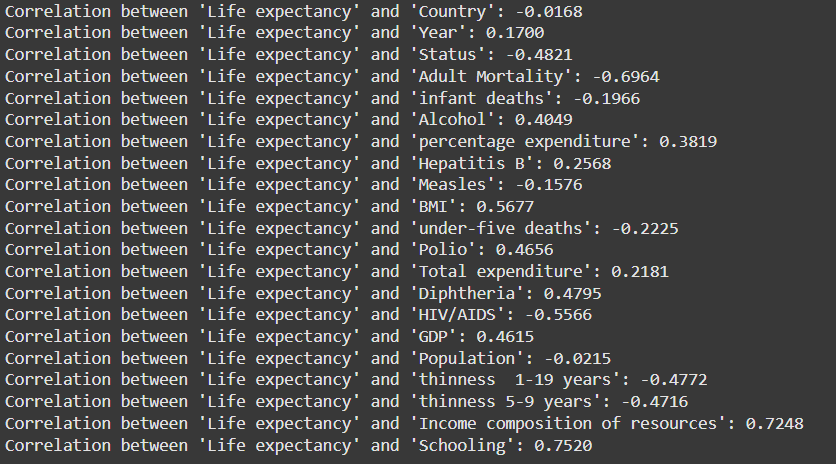
**Name and Roll No.:** Akilesh Kanna P M, 2313009

**Dataset:** Life Expectancy Data.csv

**Dependent Variable:** Life Expectancy

**Independent Variable (chosen for correlation):** Adult Mortality, BMI, HIV/AIDS, Income Composition and Schooling.

**Interpretation for Simple Linear Regression:**

****

From the above results of correlation, we can infer that only **Adult Mortality, BMI, HIV/AIDS, Income Composition and Schooling** are having a strong relationship with the dependent variable, saying they have impact over the dependent variable.

So, after conducting the Simple Linear Regression over the model, we have found the following results:

**Adult Mortality:** 0.4849

**BMI:** 0.3223

**HIV/AIDS:** 0.3098

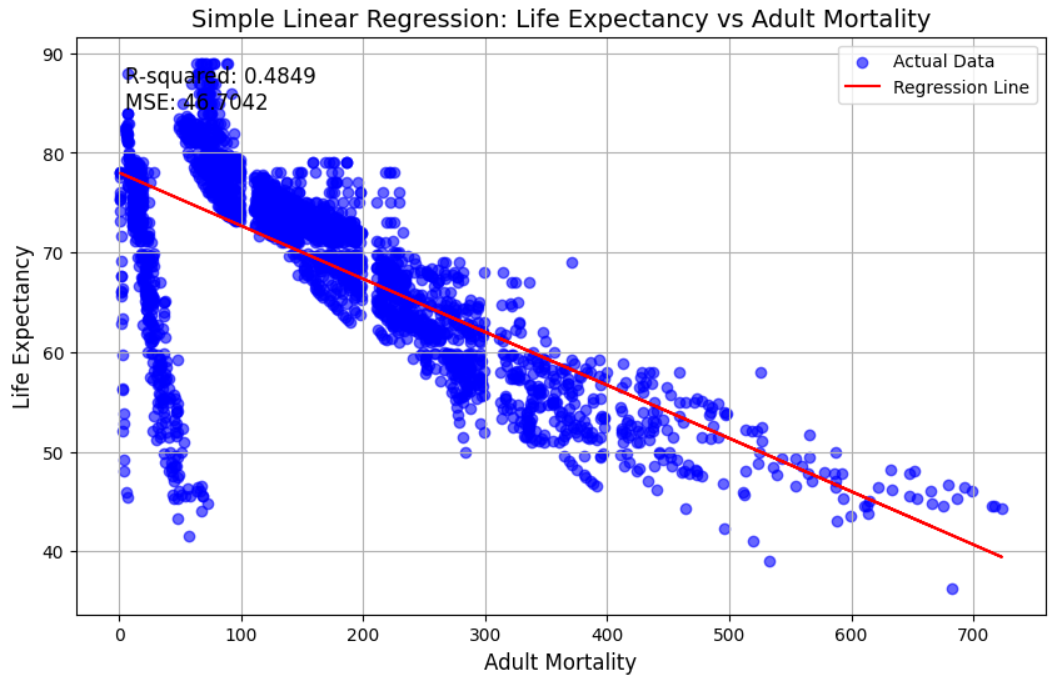
**Income Composition:** 0.5253

**Schooling:** 0.5655

From the above results, we can infer that the **Schooling and Income Composition** show the strongest positive relationship over the **Life Expectancy,** and **Adult Mortality, BMI and HIV/AIDS** are showing moderate positive correlations, but these relationships may be influenced by other external factors too.

The below given are the plots for the Simple Linear Regression of the above given variables, which gives the same inference as above, indicating Schooling and Income Composition have Strongest Positive relationship and others have medium positive relationships.

A graph with blue dots

Description automatically generated

A graph with a blue line

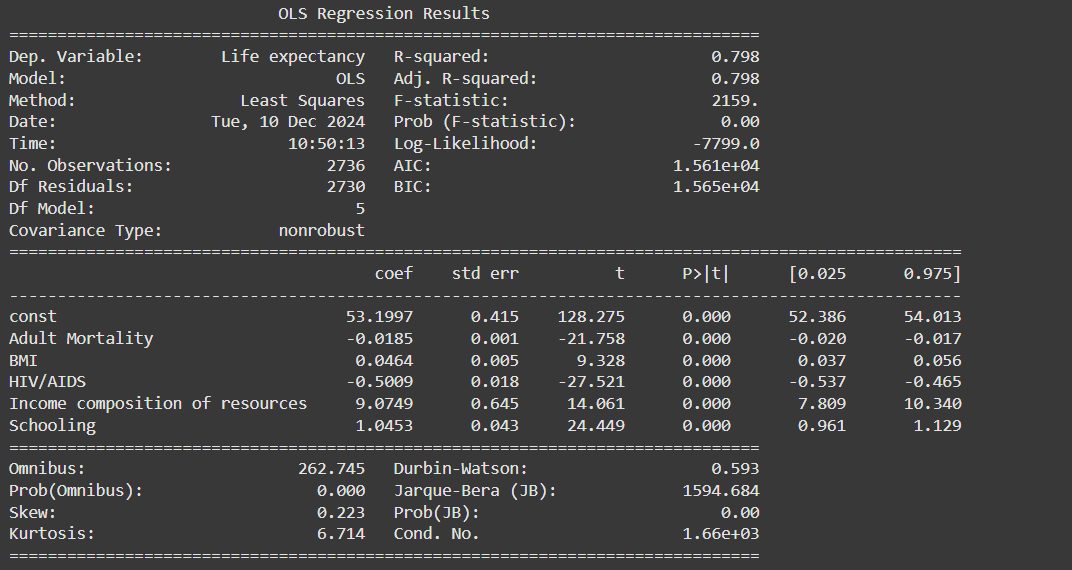
Description automatically generatedA graph with a red line and blue dots

Description automatically generated

A graph with a line drawn on it

Description automatically generated

**Inference for Multiple Linear Regression:**

****

The above given is the result of the Multiple Linear Regression, where the overall model is showing, it is fit for **Life Expectancy,** that it indicates that the Adult Mortality, BMI, HIV/AIDS, Income Composition and Schoolings.

**R^2 Value –** 0.798

**Adj. R^2 Value –** 0.798

The above R^2 value indicate that the 79.8% of the variance in **Life Expectancy** can be explained by the independent variables.

The Adj. R^2 value is same as R^2 value, which furthers confirms that the model does not suffer from overfitting and inclusion of multiple variables is justified.

