Analysis of Various Factors Affecting the Life Expectancy

Submitted to
Prof. Devlina Chatterjee

Submitted by: Group - 12

- Mrudul Shivhare (241140606)
- Preeti Borkar (241140610)
- Jyoti Tewari (241140605)
- Kavita Omar (241290009)
- Praveen Sahu (241290014)
- Kshitij Raj (241140017)

Contents

- Objective
- Data Overview
- Scatter plots
- Correlation Matrix
- Regression Model (Results)
- Implications

Objective

- The study aims to investigate the factors influencing life expectancy across different countries, using data from the World Health Organization (WHO) and the World Bank for the year 2015.
- The analysis will build a regression model that considers both mixed effects and multi linear regression to evaluate the relationships between life expectancy and different factors. The study seeks to understand how the parameters like GDP, infant deaths, adult mortality, healthcare expenditure, BMI, population, and immunization (BCG Vaccine, PVC, MCV, DPT1) correlate with life expectancy outcomes.

- **Analysis on Relevant Factors Affecting Life Expectancy," *2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), DOI* 10.1109/IPEC54454.2022.9777372**
- **Factors affecting life expectancy: evidence from 1980-2009 data in Singapore, Malaysia, and Thailand. *Asia Pacific Journal of Public Health***
- **Changes in life expectancy 1900–1990 - https://doi.org/10.1093/ajcn/55.6.1196S**
- **Healthy life expectancy in 191 countries, 1999, Mathers, Colin D et al., The Lancet, Volume 357, Issue 9269, 1685 - 1691**

Data Overview

- **Brief Description of data**
  - Data consists of annual data of life expectancy of different countries for 2015
- **Parameters**
  - **Dependent Variable:**
  - **Life Expectancy** - Average number of years once could expect to live
  - **Independent Variable:**

Model 1

- **Adult Mortality** - Probability of dying (betw. age 15 to 60 yrs.) per 1000 people
- **Infant Deaths** - No. of infant deaths per 1000 live births
- **GDP per Capita** - Indicating the growth of the country
- **Population** - in Millions
- **BMI**
- **Expenditure on health -** as % of GDP

Model 2

- **Immunization - BCG vaccine, DPT1, MCV1, HEPB3 - % of children aged b/w 12-23 months**

Model 1 – Without Immunization

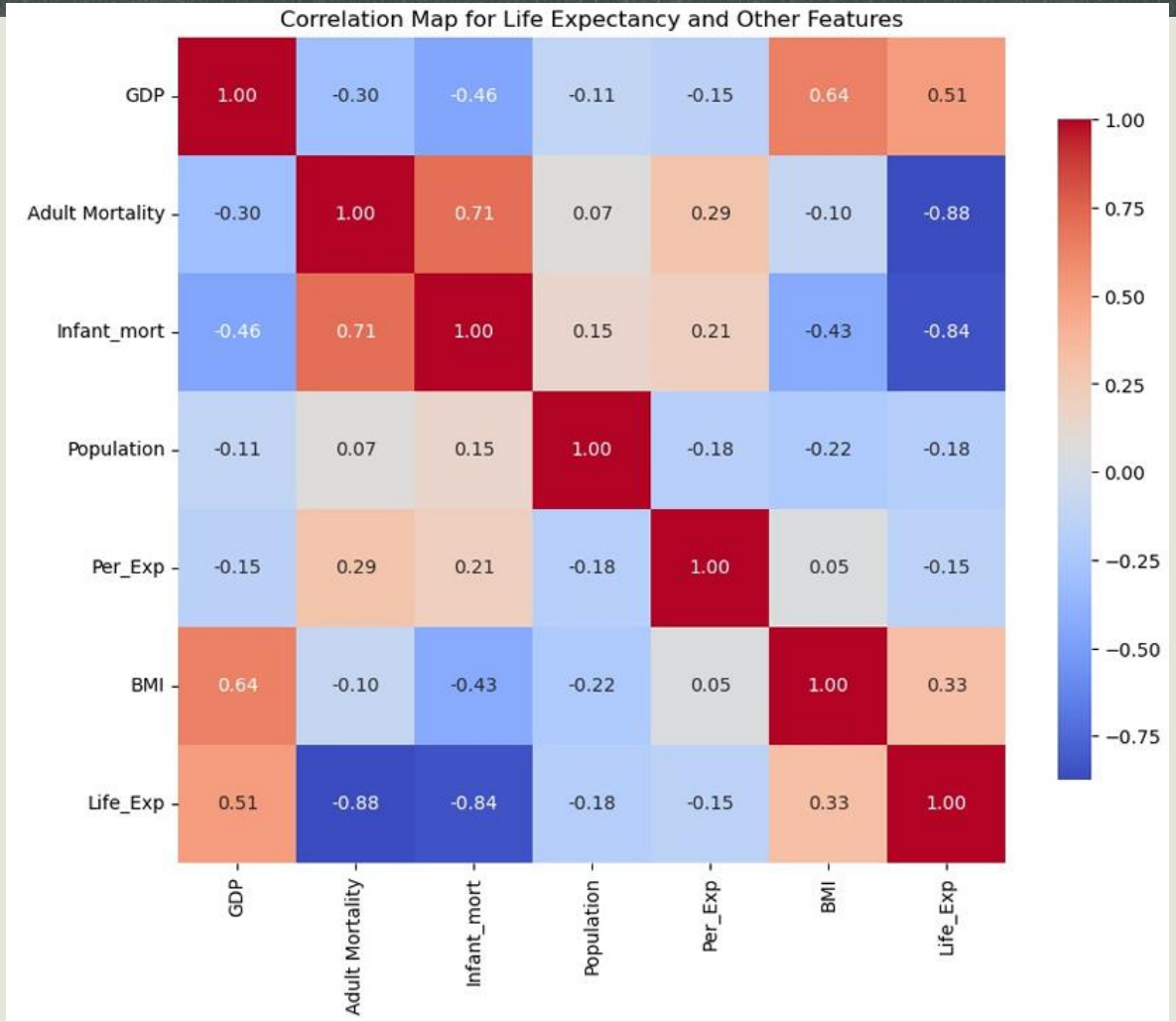
Correlation Matrix

Correlation of **Life Expectancy** with -

- Adult Mortality - High and Negative
- GDP - High and Positive
- Infant Deaths - High and Negative
- Population - Very Low and Negative
- Health Exp – Low and Negative

Other Imp Correlations

- Adult Mortality & GDP - Low and Negative
- Infant deaths and Population - Low and Positive
- Infant deaths and GDP - Negative and low



OLS Regression Results

$$LifeExp = 78.2715 + 0.2528 * \text{Per\_Exp} - 0.0413 * \text{AM} - 0.1075 * \text{infant}_{\text{deaths}} + 0.0004 * \text{GDP} + 8.045e^{-09} * \text{Popl.} - 0.0968 * \text{BMI}$$

OLS Regression Results

Dep. Variable:	Life_Exp	R-squared:	0.911
Model:	OLS	Adj. R-squared:	0.896
Method:	Least Squares	F-statistic:	58.26
Date:	Wed, 25 Sep 2024	Prob (F-statistic):	1.85e-16
Time:	16:29:34	Log-Likelihood:	-80.167
No. Observations:	41	AIC:	174.3
Df Residuals:	34	BIC:	186.3
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	78.2715	6.823	11.471	0.000	64.405	92.138
Adult Mortality	-0.0413	0.005	-7.576	0.000	-0.052	-0.030
Infant_mort	-0.1075	0.025	-4.285	0.000	-0.159	-0.057
GDP	0.0004	0.000	2.812	0.008	0.000	0.001
Population	-8.045e-09	9.58e-09	-0.840	0.407	-2.75e-08	1.14e-08
BMI	-0.0968	0.296	-0.327	0.746	-0.699	0.505
Per_Exp	0.2528	0.109	2.312	0.027	0.031	0.475

Omnibus:	2.437	Durbin-Watson:	1.212
Prob(Omnibus):	0.296	Jarque-Bera (JB):	1.559
Skew:	0.457	Prob(JB):	0.459
Kurtosis:	3.276	Cond. No.	8.88e+08

Model 2 – With Immunization

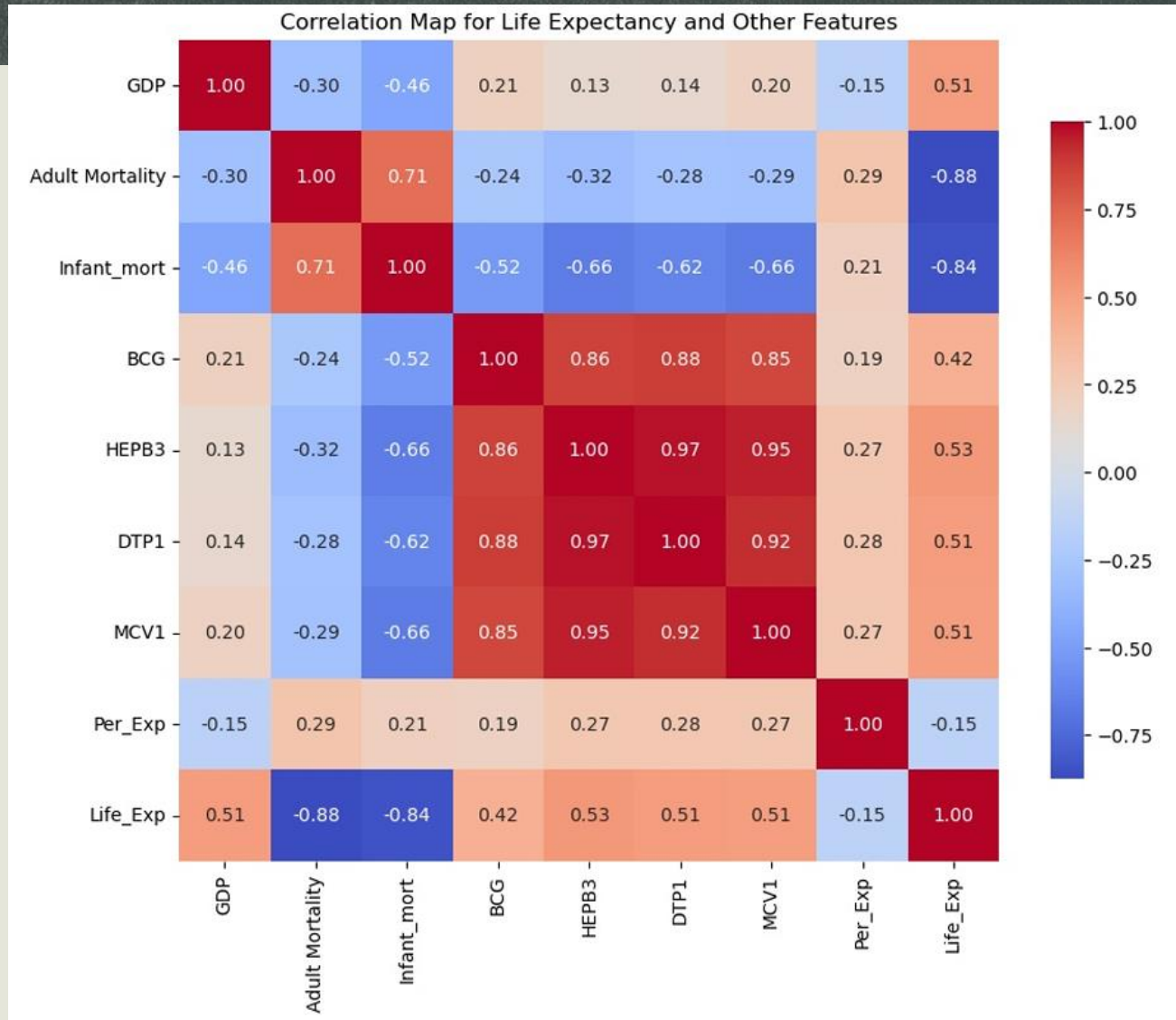
Correlation Matrix with added independent variables

Correlation of **Life Expectancy** with -

- Exp on Health Care - High and Positive
- BMI - High and Positive
- Immunization – High and Positive

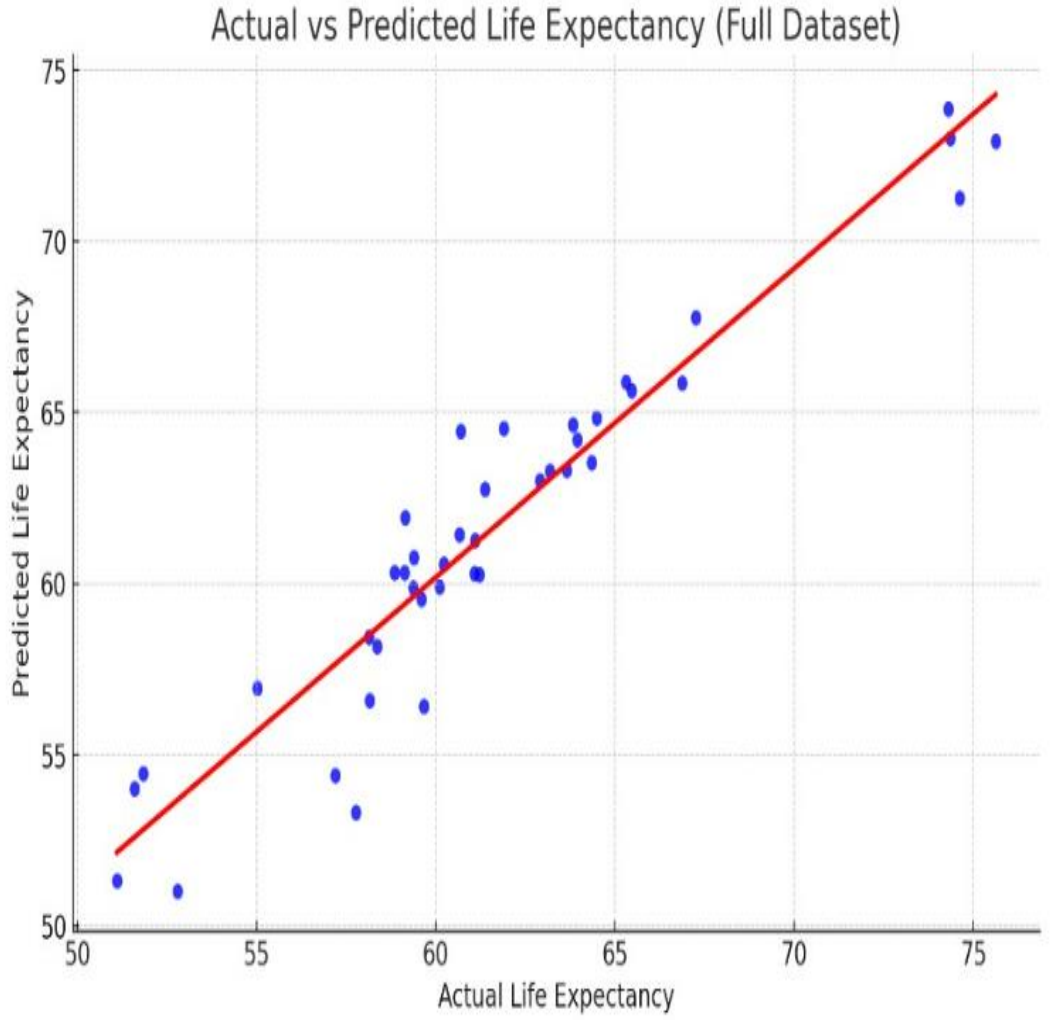
Other Imp Correlations

- GDP & Exp on Health Care - High and Positive
- Adult Mortality & GDP - Moderate and Negative
- Infant deaths and GDP - Negative and low



Regression Results

$$LifeExp = 69.89 - 0.0439 * AM - 0.0106 * infant_{deaths} + 0.0002 * GDP + 9.101e^{-09} * Popl. + 0.1448 * Per_{Exp} - 0.0676 * BCG + 0.0005 * HEPB + 0.1469 * DTP1 - 0.0246 * MCV1$$



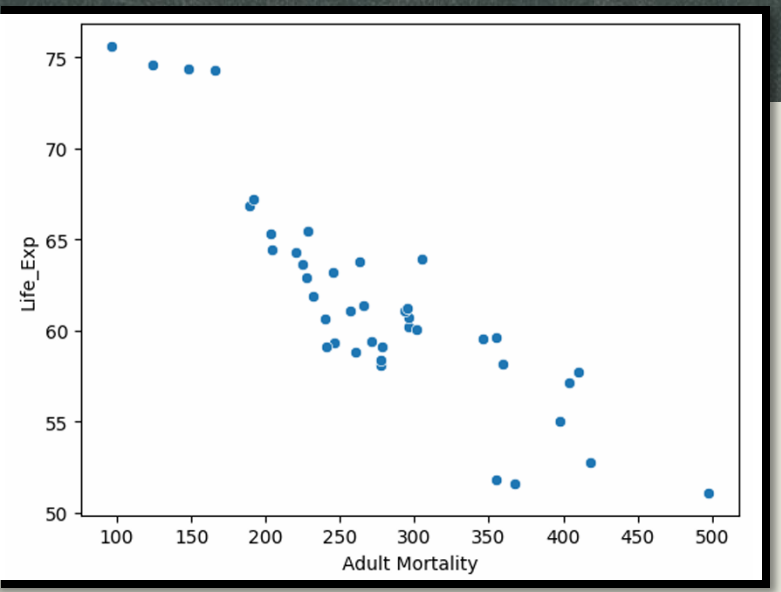
OLS Regression Results

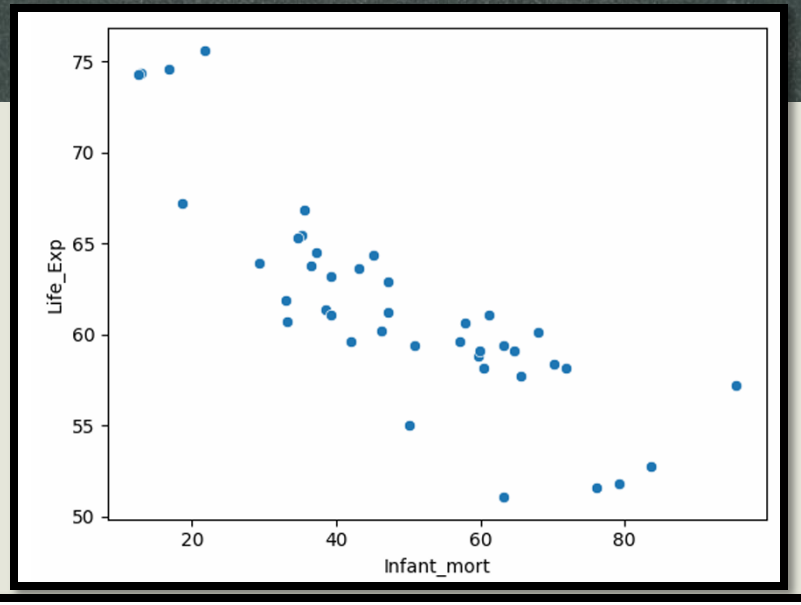
Dep. Variable:	Life_Exp	R-squared:	0.922
Model:	OLS	Adj. R-squared:	0.903
Method:	Least Squares	F-statistic:	47.45
Date:	Wed, 25 Sep 2024	Prob (F-statistic):	1.37e-15
Time:	16:29:36	Log-Likelihood:	-77.476
No. Observations:	41	AIC:	173.0
Df Residuals:	32	BIC:	188.4
Df Model:	8		
Covariance Type:	nonrobust		

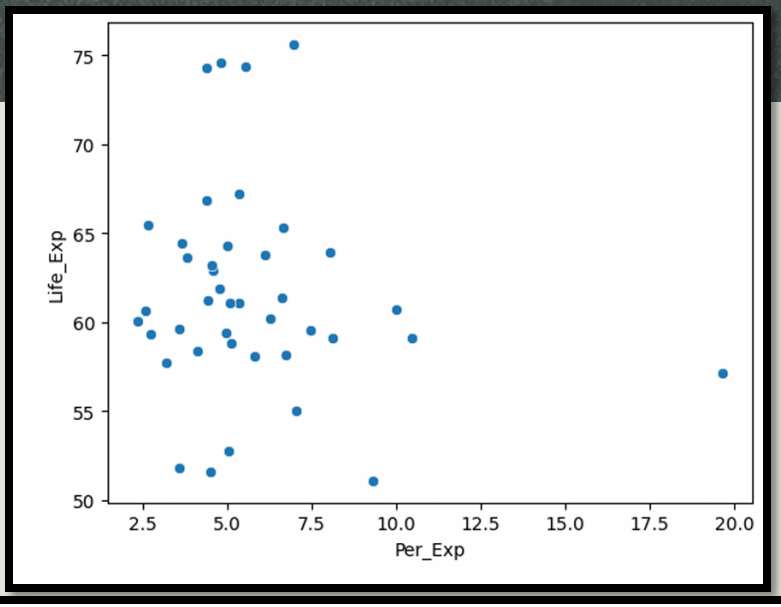
	coef	std err	t	P>|t|	[0.025	0.975]
const	69.8900	3.801	18.386	0.000	62.147	77.633
Adult Mortality	-0.0439	0.005	-8.300	0.000	-0.055	-0.033
Infant_mort	-0.0645	0.036	-1.774	0.086	-0.138	0.010
GDP	0.0004	0.000	3.768	0.001	0.000	0.001
BCG	-0.0676	0.054	-1.254	0.219	-0.178	0.042
HEPB3	0.0005	0.098	0.005	0.996	-0.200	0.201
DTP1	0.1469	0.100	1.467	0.152	-0.057	0.351
MCV1	-0.0246	0.062	-0.394	0.696	-0.152	0.103
Per_Exp	0.1448	0.128	1.133	0.266	-0.115	0.405

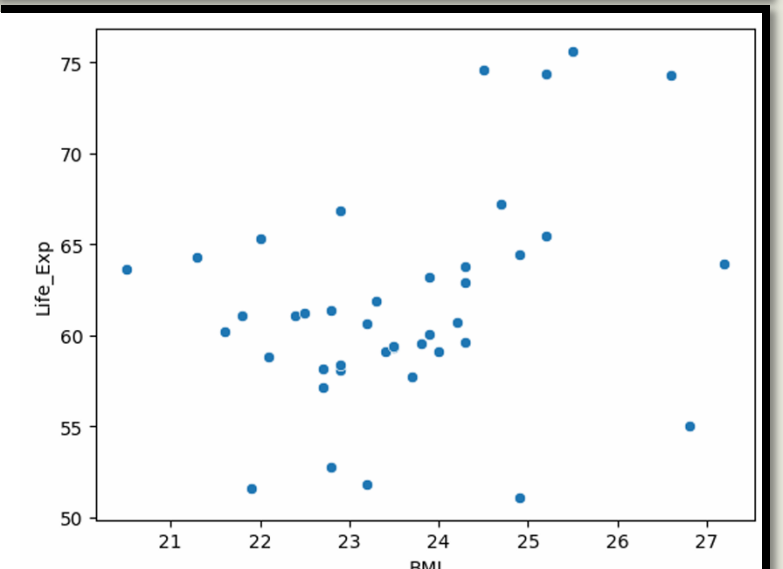
Omnibus:	0.291	Durbin-Watson:	1.371
Prob(Omnibus):	0.864	Jarque-Bera (JB):	0.202
Skew:	0.160	Prob(JB):	0.904
Kurtosis:	2.877	Cond. No.	5.41e+04

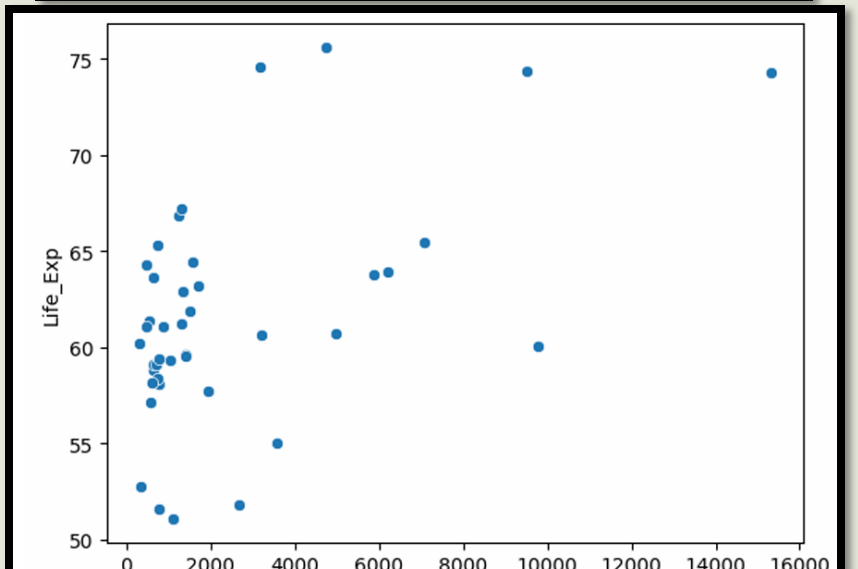
Scatter Plots

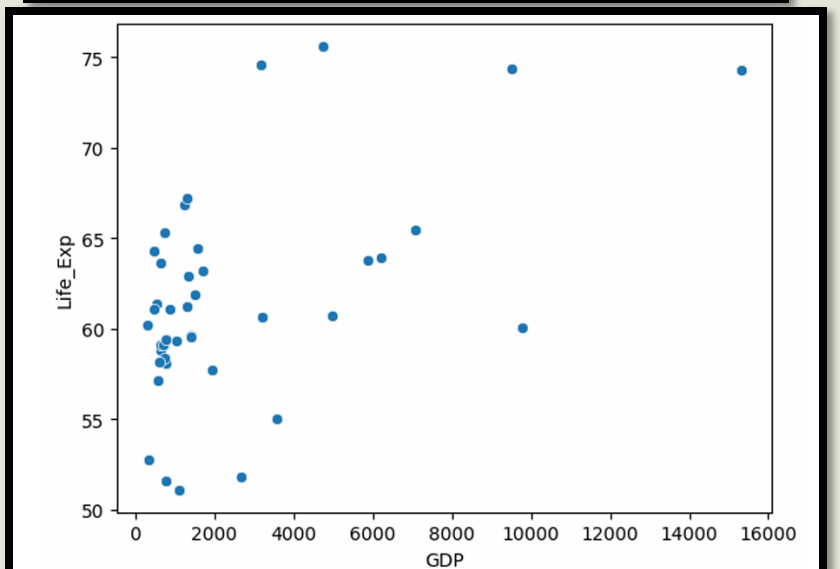












Observation and Analysis

- As the relevant independent variables increase, the ability of the model to predict the dependent variable improves
- Lower p-values indicate the parameters taken are statistically significant.
- Lower Standard error - data points are closer to regression line
- Increase in R-squared and Adjusted R-squared – due to the added independent variables.
- Omitted Variable Bias – Factors that may affect the life exp
  - Death due to unnatural events (e.g. Suicide/ Traffic/ Other accidents)
  - Social parameters like schooling etc.
  - Lifestyle of People
  - Other Diseases
  - Pollution Levels, etc.
- Model Restriction due to limited data availability for countries for all the parameters considered, the corelation between GDP and per expenditure on health is showing negative relationship.

References

- Chan, M. F., & Kamala Devi, M. (2015). Factors affecting life expectancy: evidence from 1980-2009 data in Singapore, Malaysia, and Thailand. *Asia Pacific Journal of Public Health*, *27*(2), 136-146.
- Pathirathne, L., & Sooriyarachchi, M. R. (2019). Factors Affecting Life Expectancy: A Global Perspective.
- X. He, J. Hu, C. Liu and Y. Zhang, "Analysis on Relevant Factors Affecting Life Expectancy," *2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*, Dalian, China, 2022, pp. 569-572, doi: 10.1109/IPEC54454.2022.9777372.
- https://data.who.int/countries/
- https://data.worldbank.org/
- https://catalog.data.gov/dataset/?tags=life-expectancy