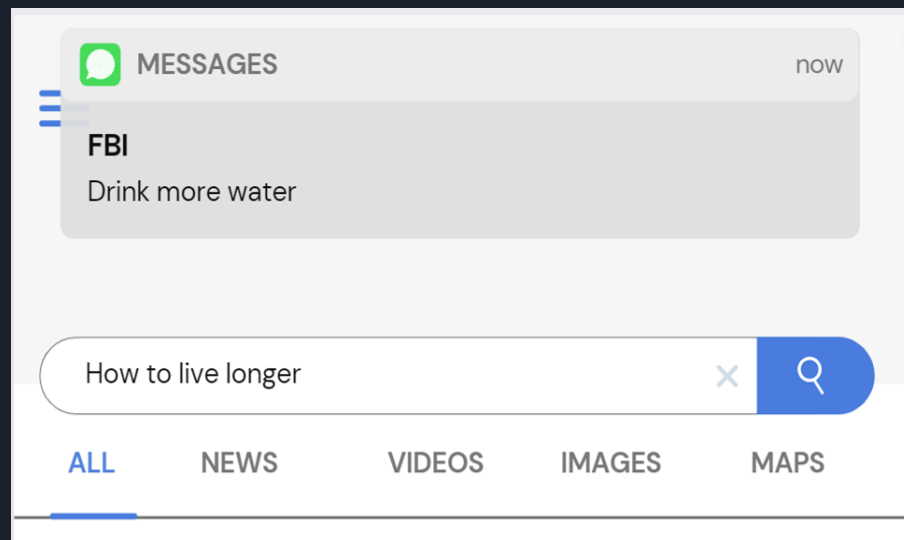




EDA on Life Expectancy

-Computing for Data Science

The usual question.. How longer
can humans live?



What we planned on researching?

The Global Health Observatory (GHO) data repository under World Health Organization (WHO) keeps track of the health status as well as many other related factors for all countries. The datasets are made available to public for the purpose of health data analysis. The dataset related to life expectancy, health factors for 193 countries has been collected from the same WHO data repository website and its corresponding economic data was collected from United Nation website.

It has been observed that in the past 15 years, there has been a huge development in health sector resulting in improvement of health conditions especially in the developing nations in comparison to the past 30 years.

Therefore, in this project we have considered data from year 2000-2015 for 193 countries for further analysis to understand the causes, and reasons of life expectancy, how it varies in different location, and how development has an impact on them.

Domain

Our dataset contains 22 features and 2938 rows. The features of the data set are:

1. Country
2. Year
3. Status : Developed or Developing status
4. Life expectancy : Life Expectancy in age
5. Adult Mortality : Adult Mortality Rates of both sexes (probability of dying between 15 and 60 years per 1000 population)
6. infant deaths : Number of Infant Deaths per 1000 population
7. Alcohol : Alcohol, recorded per capita (15+) consumption (in litres of pure alcohol)
8. percentage expenditure : Expenditure on health as a percentage of Gross Domestic Product per capita(%)
9. Hepatitis B : Hepatitis B (HepB) immunization coverage among 1-year-olds (%)
10. Measles : Measles - number of reported cases per 1000 population
11. BMI : Average Body Mass Index of entire population
12. under-five deaths : Number of under-five deaths per 1000 population
13. Polio : Polio (Pol3) immunization coverage among 1-year-olds (%)
14. Total expenditure : General government expenditure on health as a percentage of total government expenditure (%)
15. Diphtheria : Diphtheria tetanus toxoid and pertussis (DTP3) immunization coverage among 1-year-olds (%)
16. HIV/AIDS : Deaths per 1 000 live births HIV/AIDS (0-4 years)
17. GDP : Gross Domestic Product per capita (in USD)
18. Population : Population of the country
19. thinness 1-19 years : Prevalence of thinness among children and adolescents for Age 10 to 19 (%)
20. thinness 5-9 years : Prevalence of thinness among children for Age 5 to 9(%)
21. Income composition of resources : Human Development Index in terms of income composition of resources (index ranging from 0 to 1)
22. Schooling : Number of years of Schooling(years)

Methodology



1

Pandas

2

Numpy

3

Matplotlib

4

PyLab

5

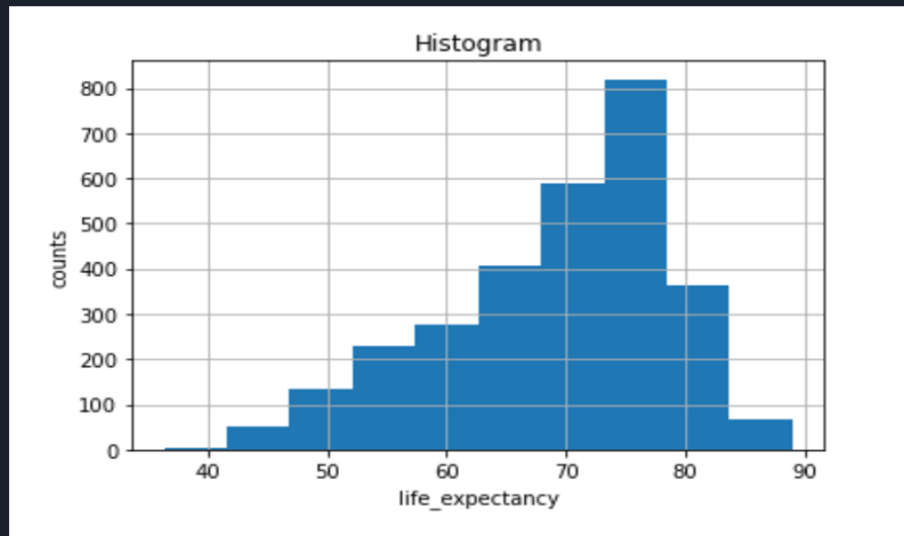
Seaborn

6

Plotly.express

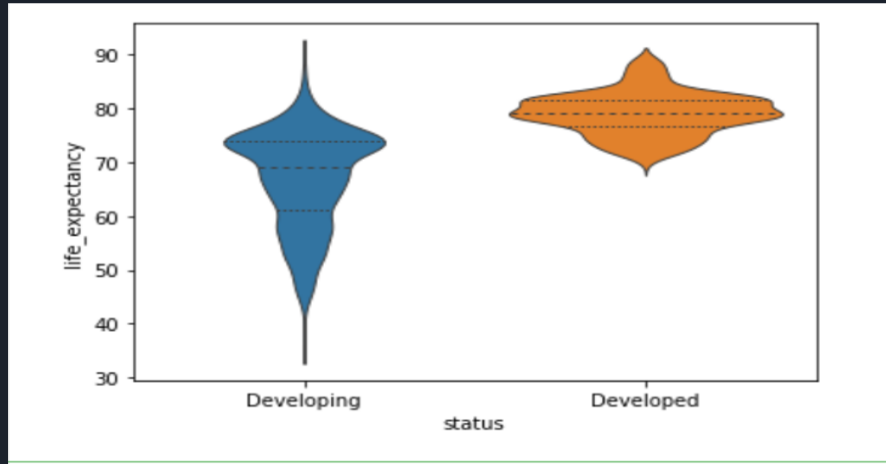


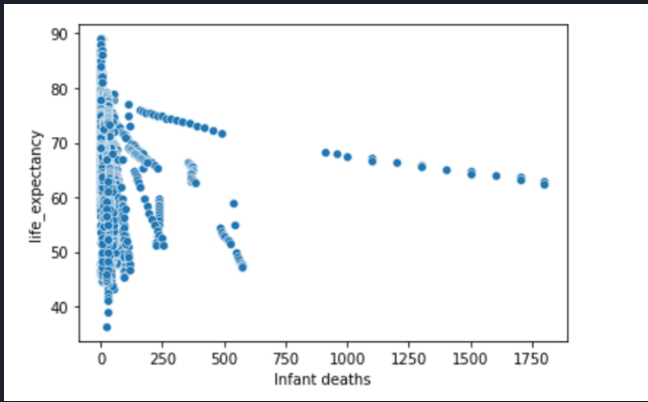
Graphs - Distribution of Life Expectancy according to the age



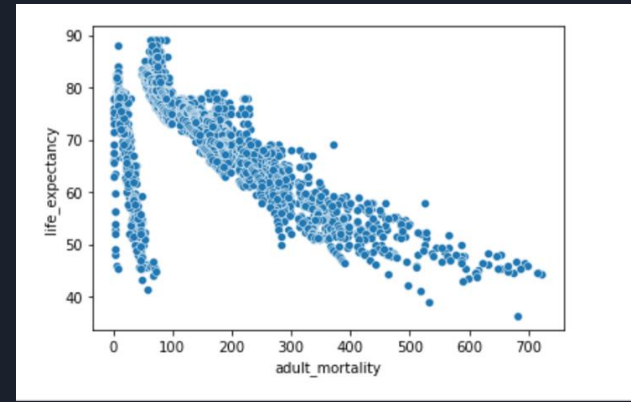


Comparing the life expectancy of Developing and Developed Countries using violin plot

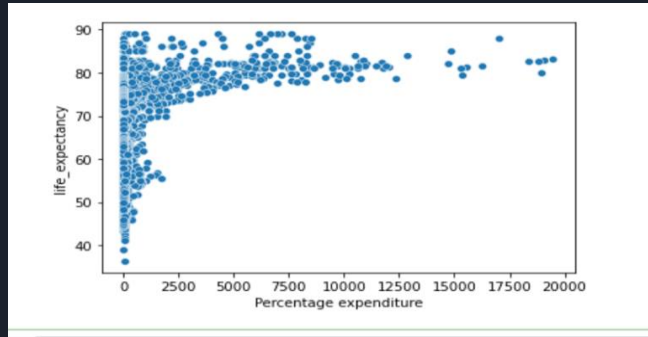




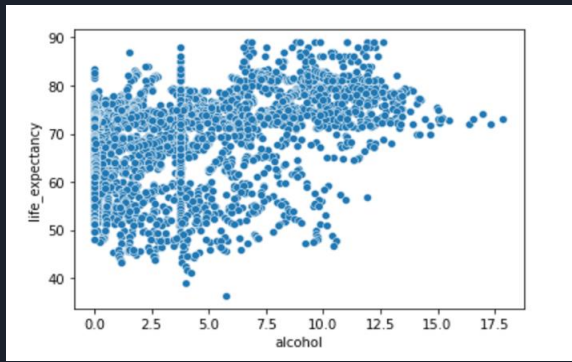
Life expectancy v/s the infant death in countries every year



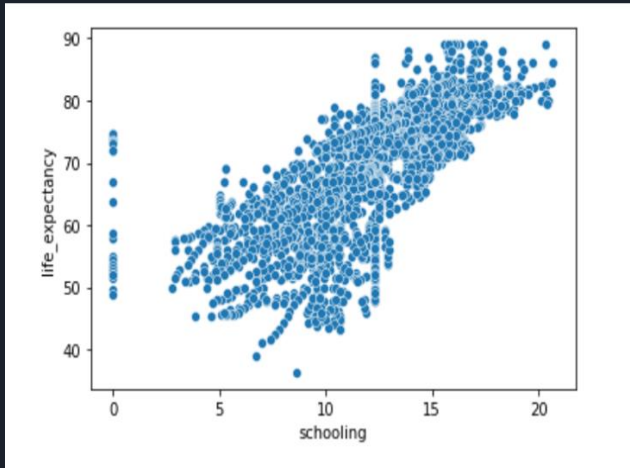
Life Expectancy v/s the adult Mortality in different countries every year



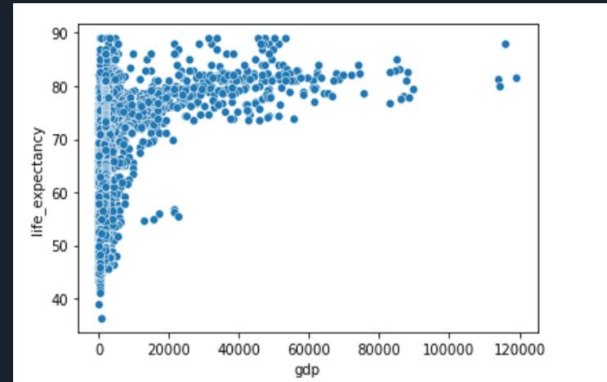
Life Expectancy Vs Percentage Expenditure



Life Expectancy versus Alcohol



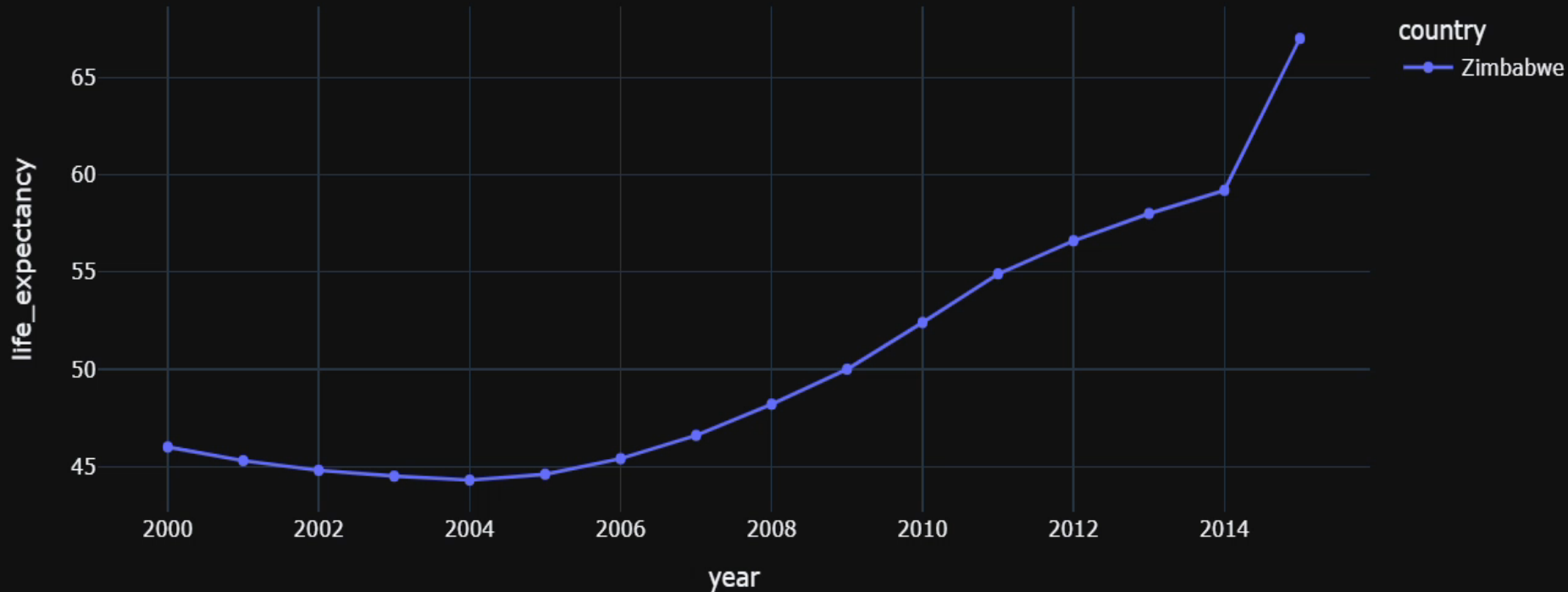
Life Expectancy Versus
Schooling



Life Expectancy v/s GDP of
Countries all over the World



Country Wise Life Expectancy over the years



country=Zimbabwe



Zimbabwe

Croatia

San Marino

Life Expectancy over the World Map



country

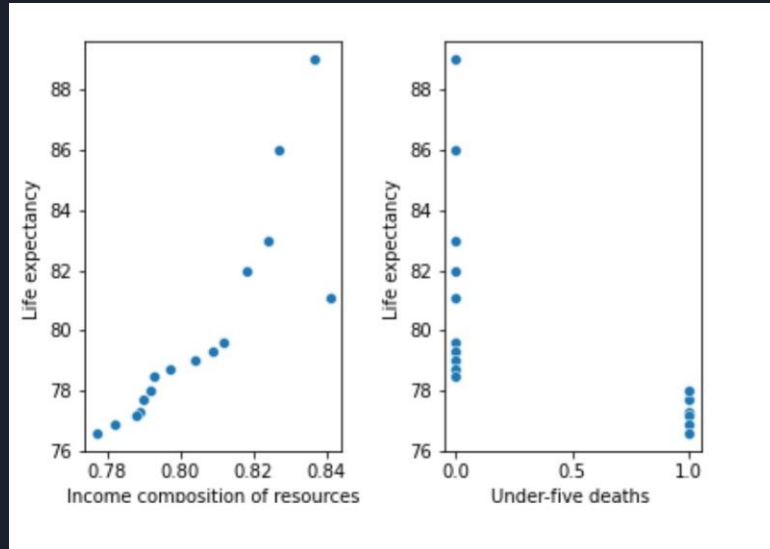
- Afghanistan
- Albania
- Algeria
- Angola
- Argentina
- Australia
- Austria
- Bahrain
- Bangladesh
- Belgium
- Benin
- Bolivia
- Bosnia and Herzegovina
- Botswana
- Brazil
- Bulgaria
- Burkina Faso

High Life Expectancy in PORTUGAL



Income composition of resources highly positively correlated.

Under-five deaths & Thinness5-9years is highly negatively correlated.



Scatter Plot for Belgium

Repetition of more Life Expectancy in France



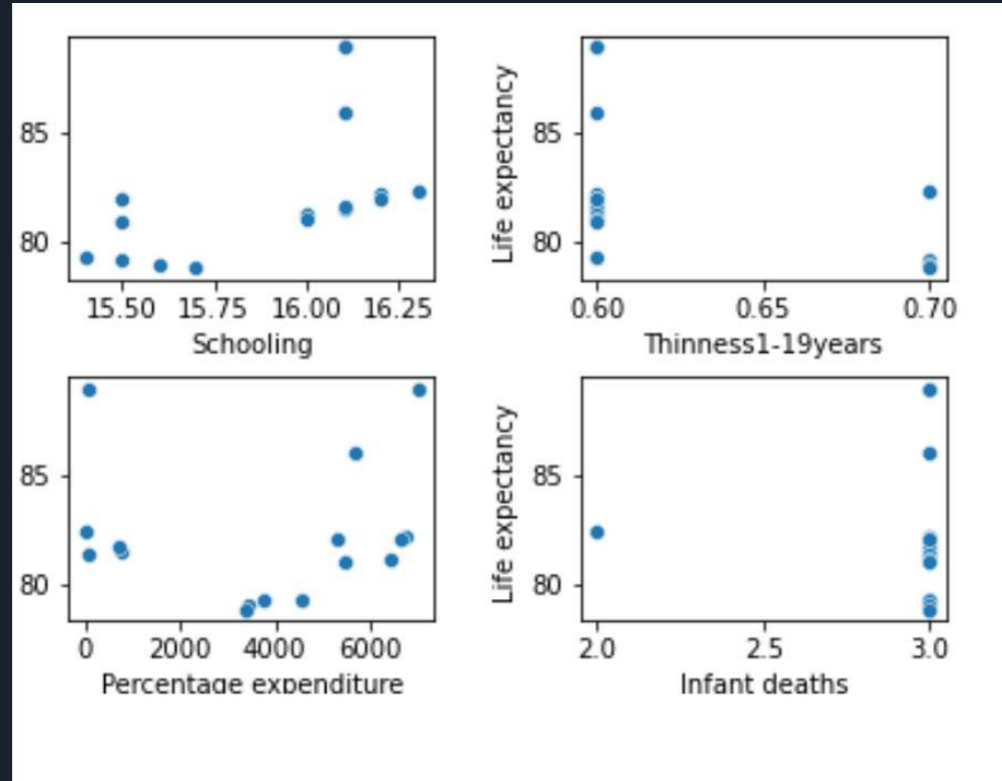
- In the mid eighteenth century, half of all children died before age ten, and life expectancy was just 25 years.
- It reached 30 years at the end of the century, then jumped to 37 years in 1810, partly thanks to smallpox vaccination.
- It continued to increase slowly during the nineteenth century, reaching 45 years in 1900.
- During the Napoleonic wars and the war of 1870, life expectancy plummeted temporarily to below 30 years once again.
- During the twentieth century, life expectancy rose more quickly, except during the two World Wars.
- Infant and child deaths became much rarer: 15% of all babies died before age one in 1900, 5% in 1950 and just 0.4% (3,5 for thousand exactly) born in 2015.
- Life expectancy is still increasing, thanks to progress in the fight against cardiovascular diseases and cancer.
- In 2020, life expectancy at birth in France was 79,2 years old for men and 85,3 years old for women).



Scatter plot for France

schooling highly
positively correlated
Thinness 1-19years
death is highly
negatively correlated.

Percentage
expenditure is lowly
positively correlated.
Infants deaths is the
lowly
negatively correlated.





Less Life Expectancy in Haiti

- ❖ LE and HALE increased substantially in Haiti. People may hope to live longer in 2017, but in poor health.
- ❖ The Caribbean countries had significantly lower YLLs rates than Haiti for ischemic heart disease, stroke, lower respiratory infections, and diarrheal diseases.
- ❖ Road injuries were the leading cause of DALYs for people aged 5-14 years. Road injuries and HIV/AIDS were the leading causes of DALYs for men and women aged 15-49 years, respectively.
- ❖ Ischemic heart disease was the main cause of DALYs for people older than 50 years. Maternal and child malnutrition were the leading risk factors for DALYs in both sexes.



Scatter plot for Haiti

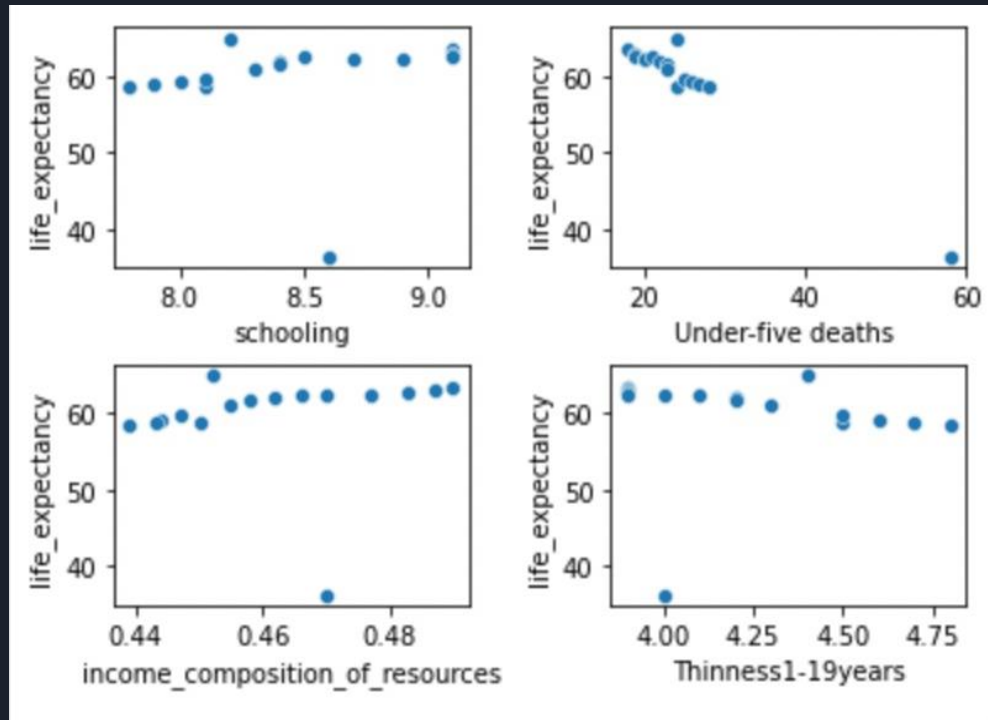
schooling highly positively correlate

under-five death is highly negatively correlated.

Income composition of resources is lowly positively correlated.

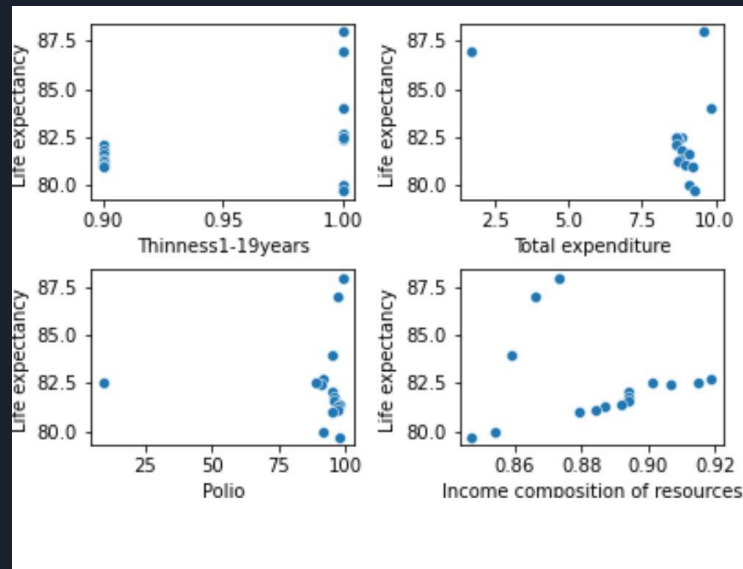
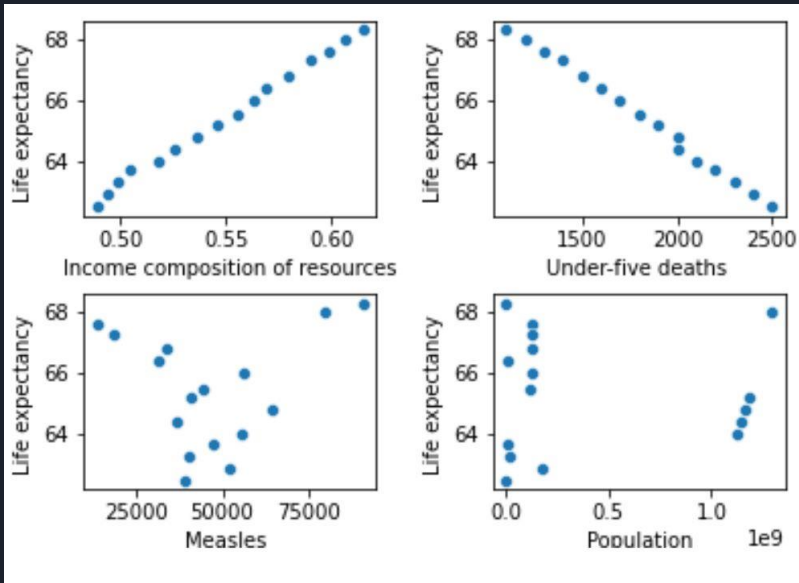
thinness 5-9 years is the lowly negatively correlated.

Here the correlation is 36.3



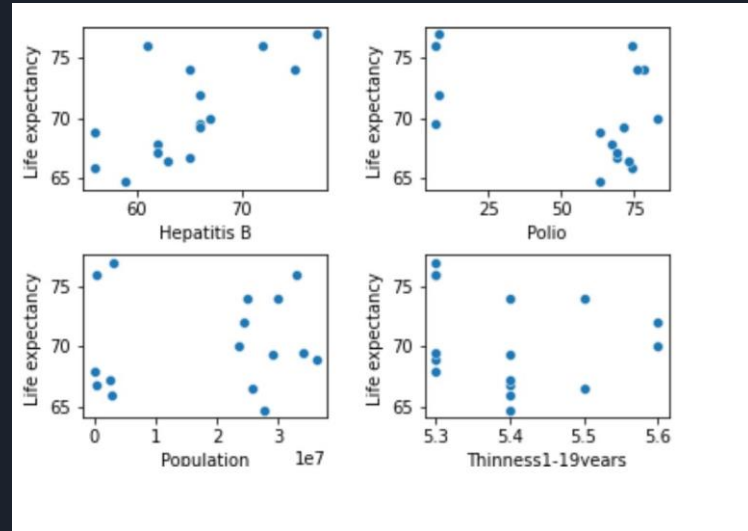
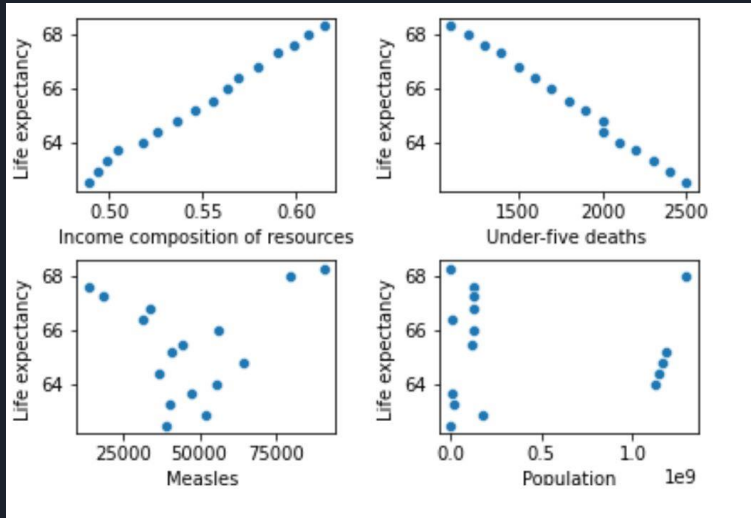


INDIA VS ICELAND





INDIA VS IRAQ





Correlation values

```
In [45]: cor_mat = df.corr()["Life expectancy"].sort_values(ascending=False)
```

```
In [46]: cor_mat
```

```
Out[46]: Life expectancy      1.000000
          Schooling          0.713054
          Income composition of resources 0.688662
          BMI                0.556901
          Diphtheria         0.472211
          Polio              0.458399
          GDP                0.430461
          Alcohol            0.388918
          Percentage expenditure 0.381418
          Total expenditure   0.208844
          Year               0.170819
          Hepatitis B        0.170219
          Population         -0.029014
          Measles            -0.157767
          Infant deaths      -0.196769
          Under-five deaths  -0.222738
          Thinness5-9years   -0.462473
          Thinness1-19years -0.468002
          HIV/AIDS           -0.556703
          Adult Mortality    -0.696390
          Name: Life expectancy , dtype: float64
```

We see that over time, Life expectancy has a positive trend in both developed and developing countries
Developed countries have more years of life expectancy than developing countries.



we can say that the variables with moderate to high correlation coefficients, may be good predictors of life expectancy, this variables are : Adult Mortality, HIV/AIDS, thinness 5-9 years, thinness 1-19 years, schooling, Income composition, GDP, Total expenditure.

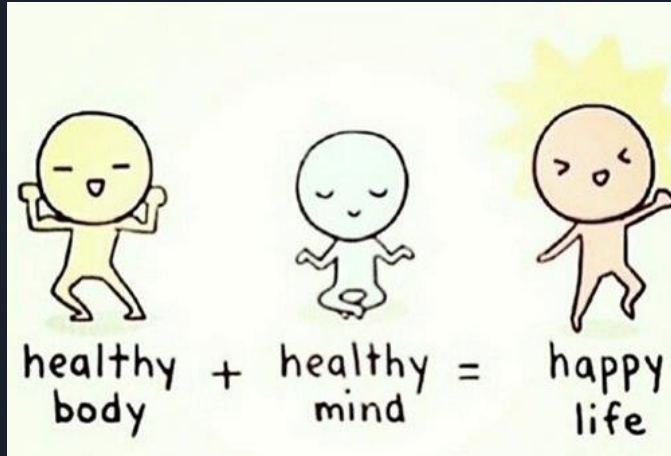
Alcohol has a positive unexpected relationship, we suspect that the data can reveal this misleading result due to outliers, we investigate the distribution properties of Alcohol feature.

Infant death shows a negative but weak correlation, we suspect that this can be too due to outliers



Apparently according to the scatter plot and correlation coefficient of (-0.1965) there's a weak correlation between infant deaths and life expectancy. But this result may be misleading, we should investigate the existence of outliers.

There is a negative correlation between Adult Mortality and Life expectancy, in both developed and developing countries



Presented by:

Jaini Maria John-222BDA06

Mareena Polin-222BDA07

Romel-222BDA16

Anjali. A. Sunil-222BDA34

Karanam Sreenidhi 222BDA20

Rachel Roshani J-222BDA42