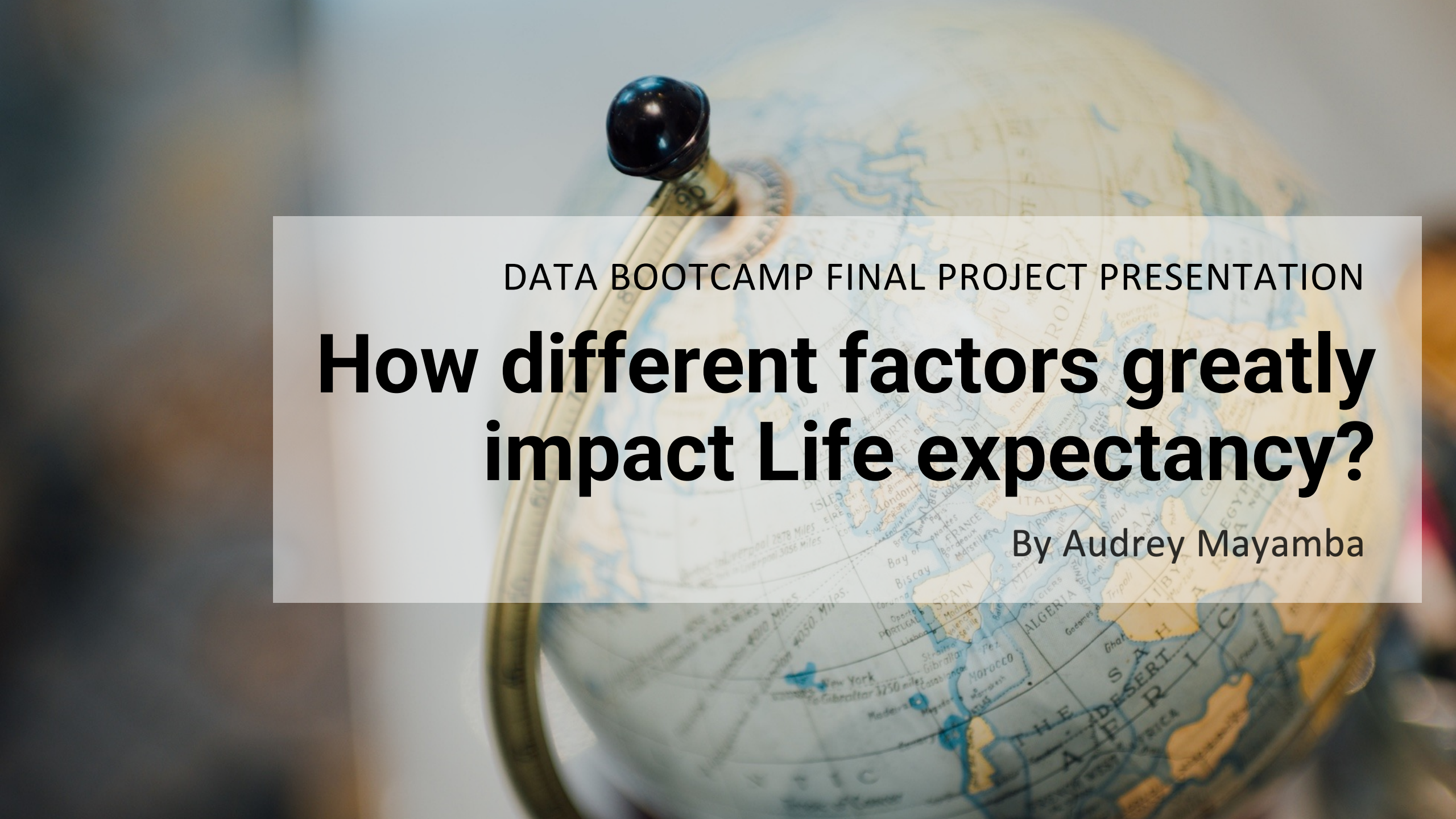


DATA BOOTCAMP FINAL PROJECT PRESENTATION

How different factors greatly impact Life expectancy?

By Audrey Mayamba

About me

Hello,
I am Audrey Mayamba.

- Biomedical scientist in Haematology and Blood transfusion
- I discovered Niyo bootcamp at Black girls in Tech fest and from then on, they have helped me in my journey.
- At this bootcamp I have been taught SQL, Power BI, Python and Project management.
- I have greatly enjoyed the whole experience and teachings of the boot camp. It has exercised and developed my technical skills and I cannot wait to explore what the future holds for me.

My Objectives

This project analysed **Life Expectancy WHO (World Health Organisation)** dataset from Kaggle.

The aim of the project was to explore the following questions;

1. Understanding the relationship between life expectancy and economic factors
2. The relationship between the status of a country and life expectancy
3. To analyse the correlation between expenditure on healthcare and education on life expectancy
4. To analyse the effects of HIV/AIDS, Polio and immunisations on life expectancy

Findings: The data highlighted the vast economic differences between countries and how GDP grossly impacts other factors, which in turn leads to lower average life expectancies.

Trends

What is Life Expectancy?

Life expectancy refers to measure of the average time a person is expected to live based on their age, gender, and other factors. It is an important indicator of the overall health of a population and can be used to track progress in improving healthcare and reducing mortality rates. Understanding life expectancy can help us identify areas where healthcare and public health interventions are needed most.

Global Trends in Life Expectancy:

Life expectancy has been steadily increasing around the world over the past century, with significant progress made in many countries. This increase is largely due to improvements in healthcare, sanitation, diet, as well as advances in medical technology.

Data cleaning using Excel power query

- The dataset contained a lot of missing data. Therefore, I used Excel power query to clean up the data and remove null values and columns that weren't needed such as alcohol consumption, thinness 1-19 and 5-9 year's old and BMI. The average values for BMI seemed too high for a lot of the countries to be accurate.

- I used the column quality option in power query to highlight potential errors and null values and then used the filter to highlight all the null values and those data were removed.

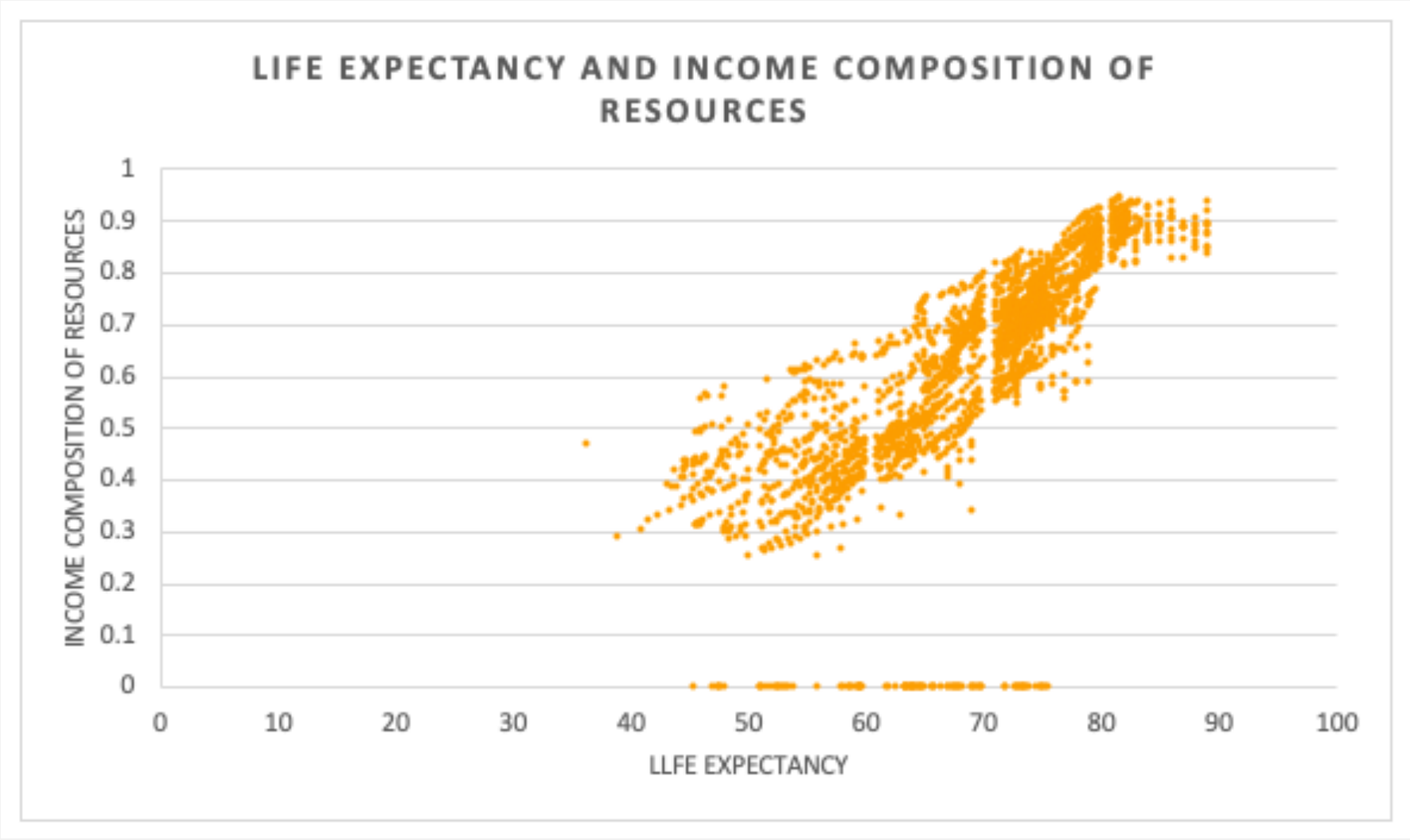
- The data ranged from the year 2000 to 2015, however most countries didn't have data in several of the categories for the year 2015. So the data is being explored from 2000 to 2014.

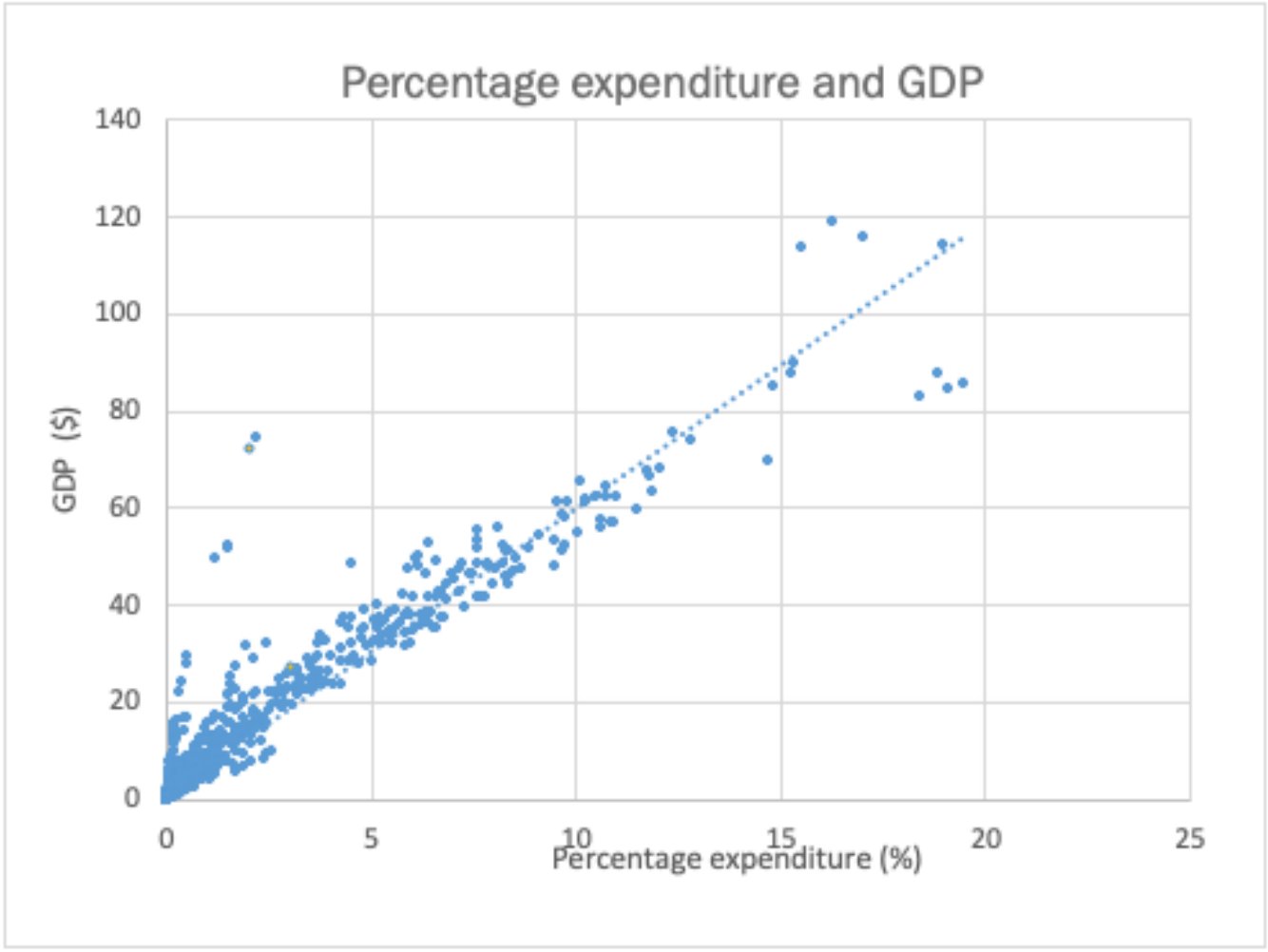
Excel

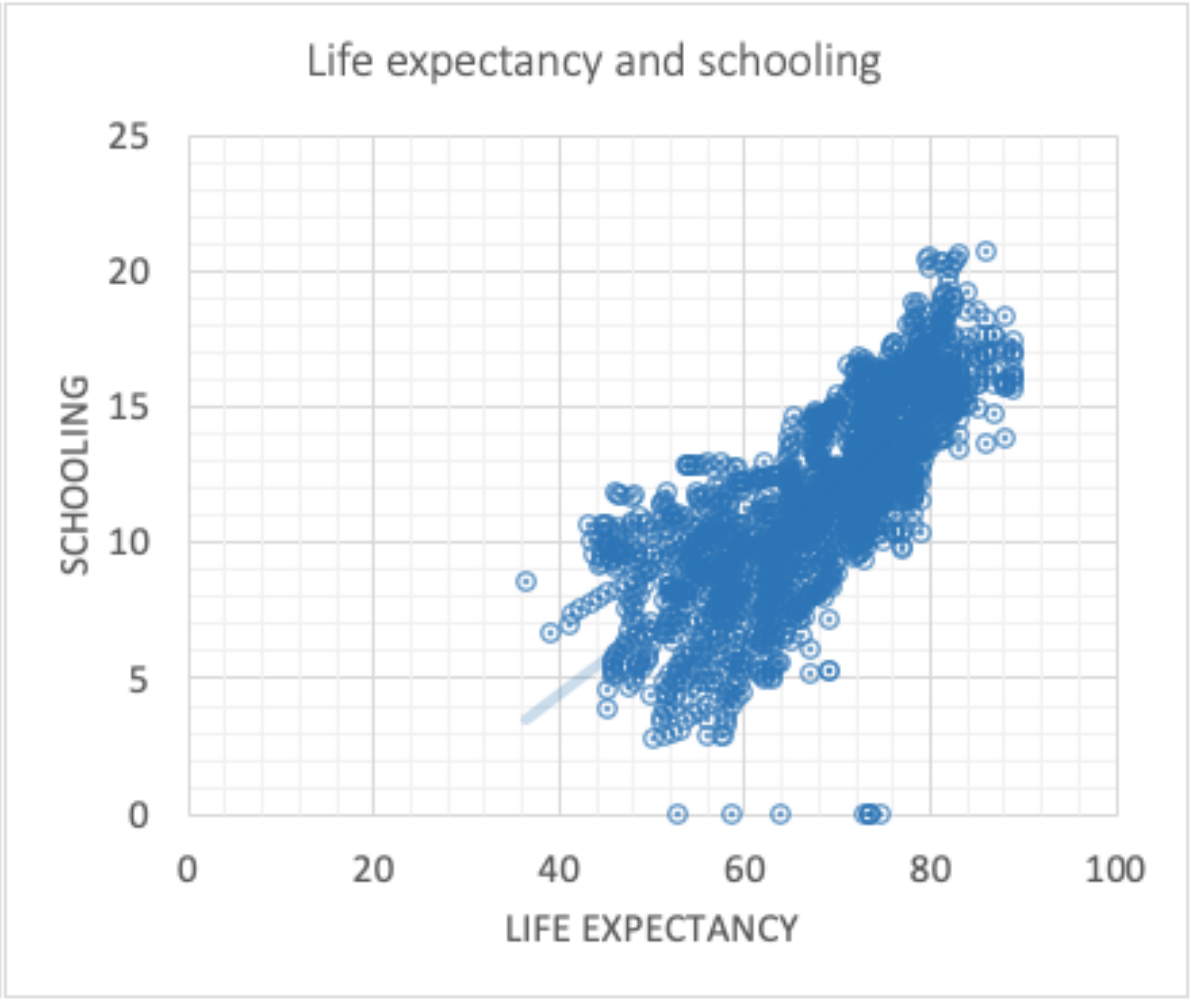
With the cleaned data I created pivot tables exploring different aspects of the data.

It highlighted a key point;

There was a strong positive correlation with a lot of the factors and life expectancy.

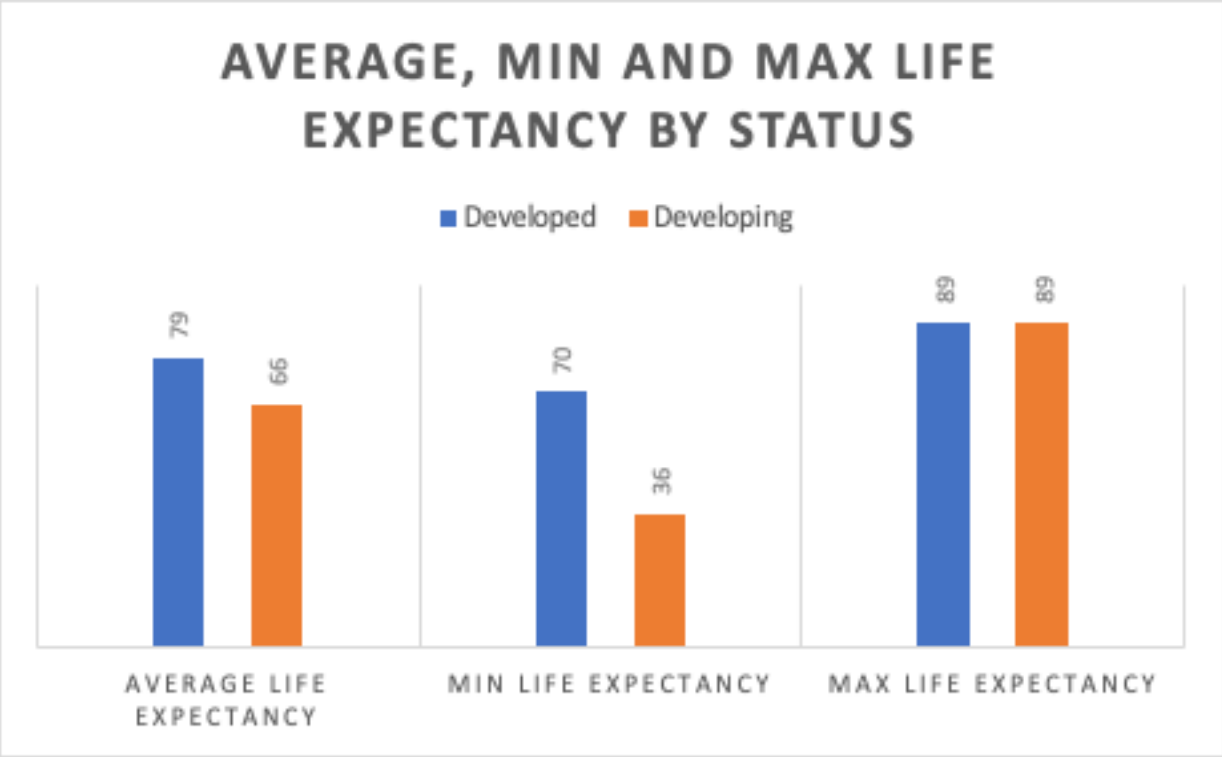


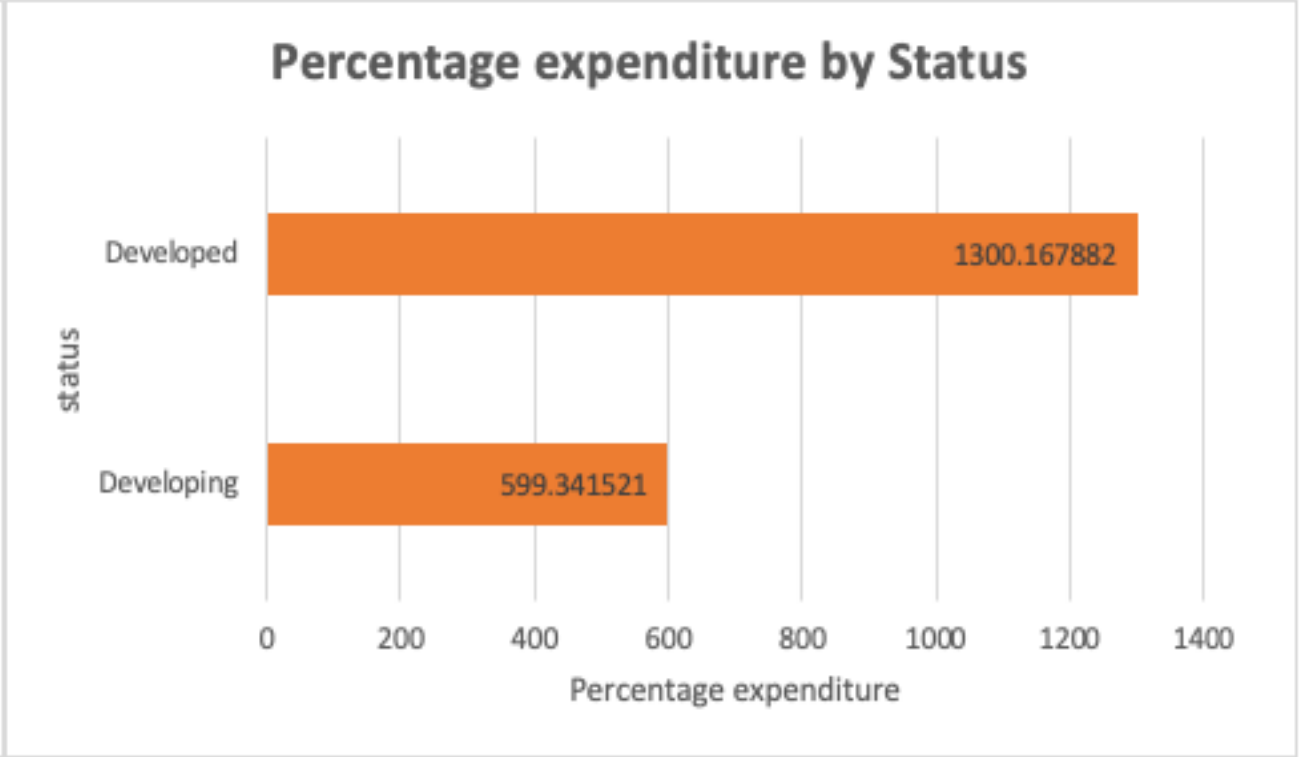




In both graphs there were strong positive correlations in between the two variables.

SQL Analysis






```
SELECT status, AVG(life_expectancy) AS Average_life_expectancy,MIN(life_expectancy) AS
Min_life_expectancy, MAX(life_expectancy) AS Max_life_expectancy
FROM `level-facility-394518.Life_expectancy.Life_Expectancy`
GROUP BY  status;
```

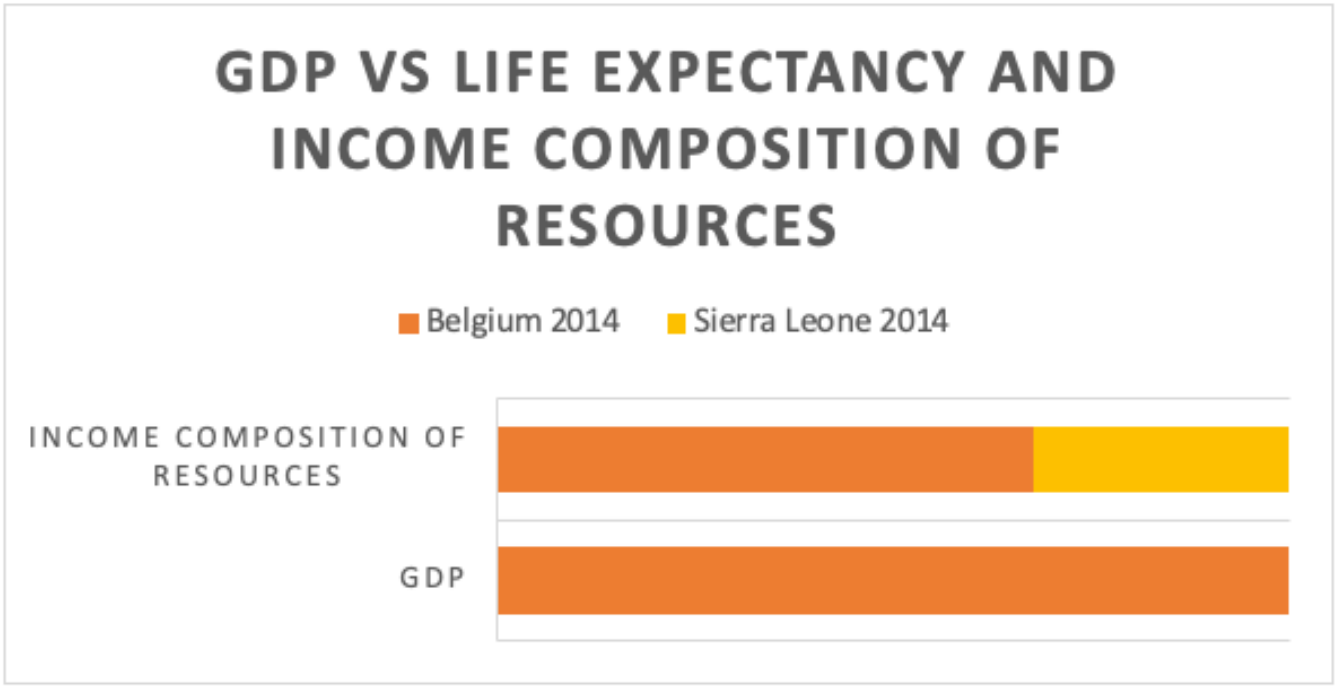
SQL

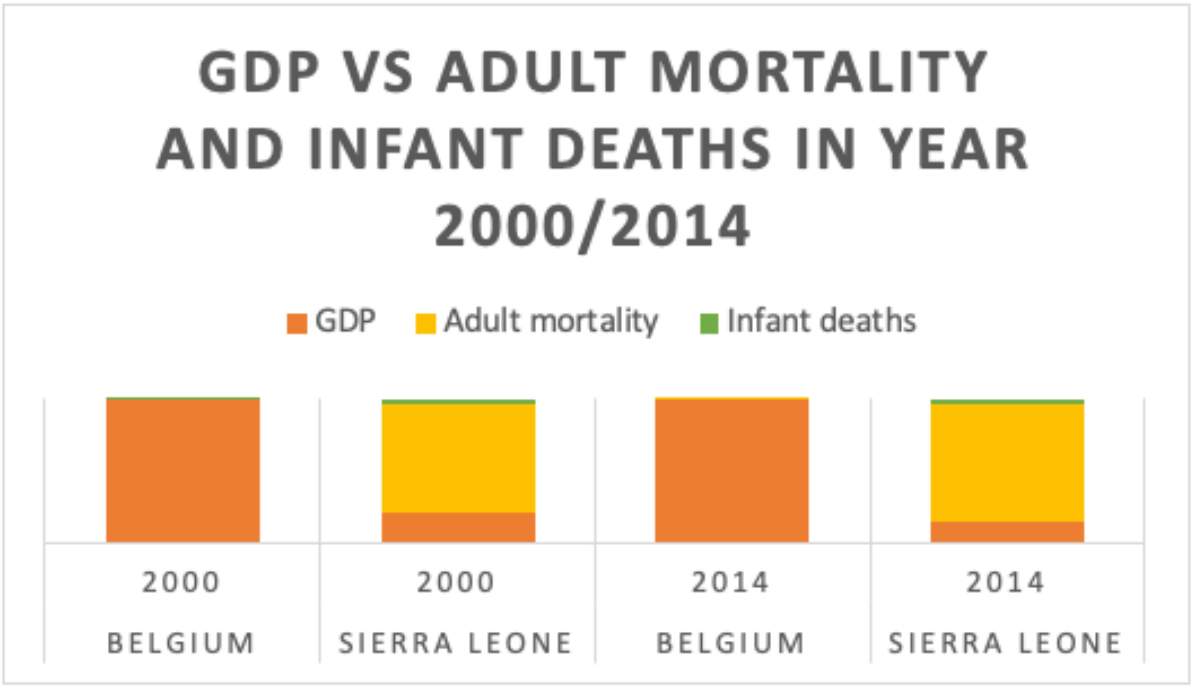
- I then used SQL to examine the data further focusing primarily on the years 2000 and 2014.
- I found that in 2000 that Japan had the highest life expectancy, whereas Sierra Leone had the lowest.
- I also found that Belgium had the highest life expectancy in 2014, when compared with Sierra Leone in the same year.
- I explored the differences between the two countries further to examine the impact different factors had on life expectancy

```
SELECT country
FROM `level-facility-394518.Life_expectancy.Life_Expectancy`
WHERE year = 2000
ORDER BY life_expectancy DESC
LIMIT 1;
```

```
SELECT country
FROM `level-facility-394518.Life_expectancy.Life_Expectancy`
WHERE year = 2014
ORDER BY life_expectancy ASC
LIMIT 1;
```

SQL



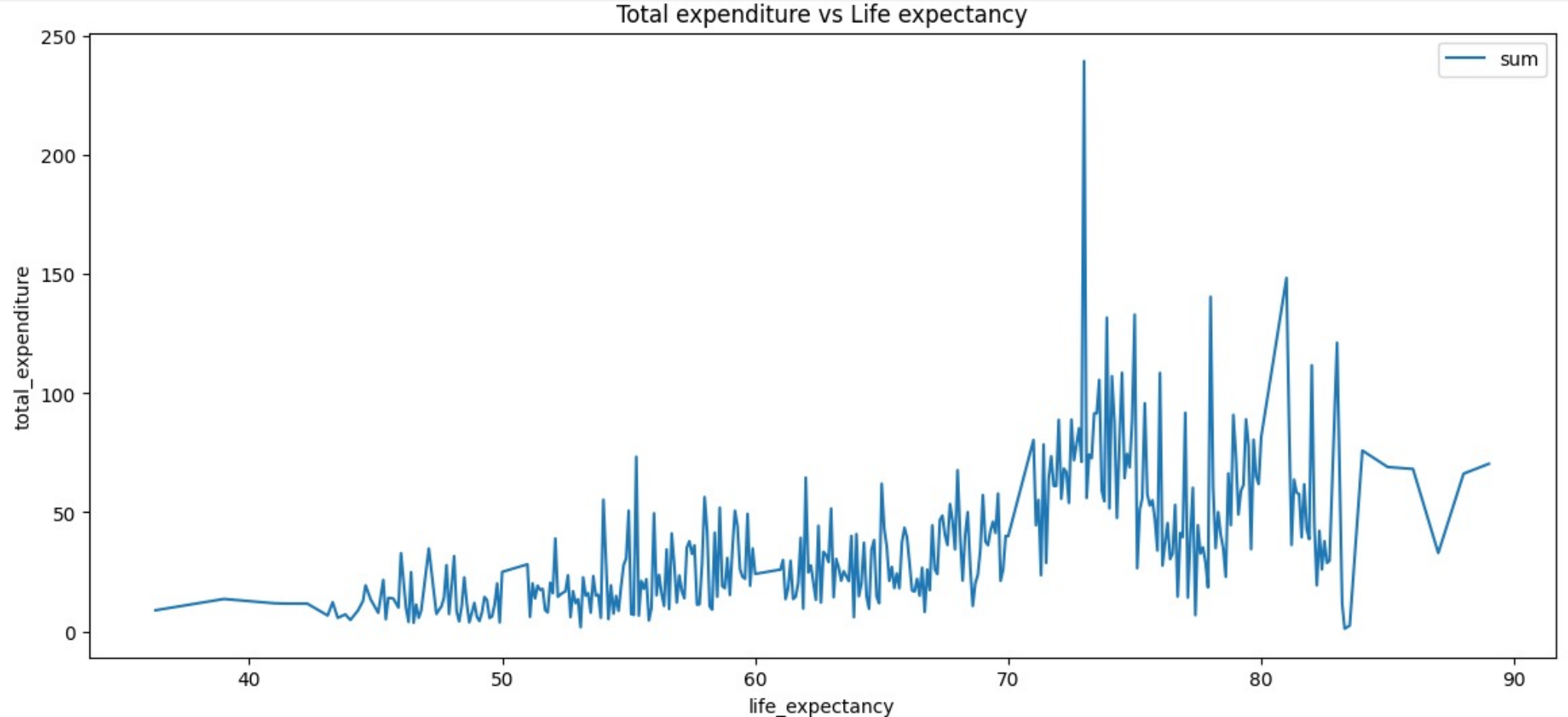


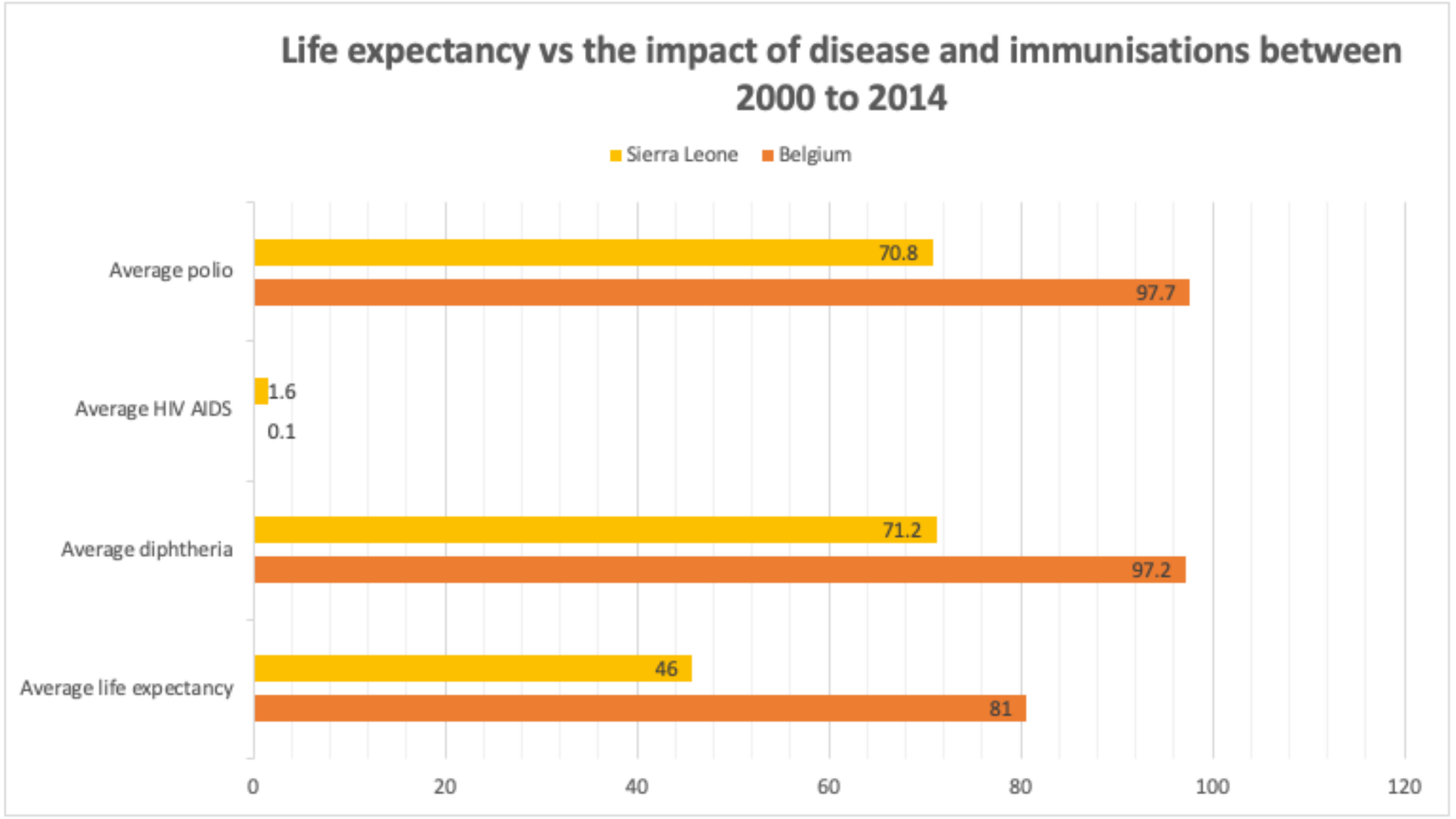
There is a strong correlation between GDP and mortality rates

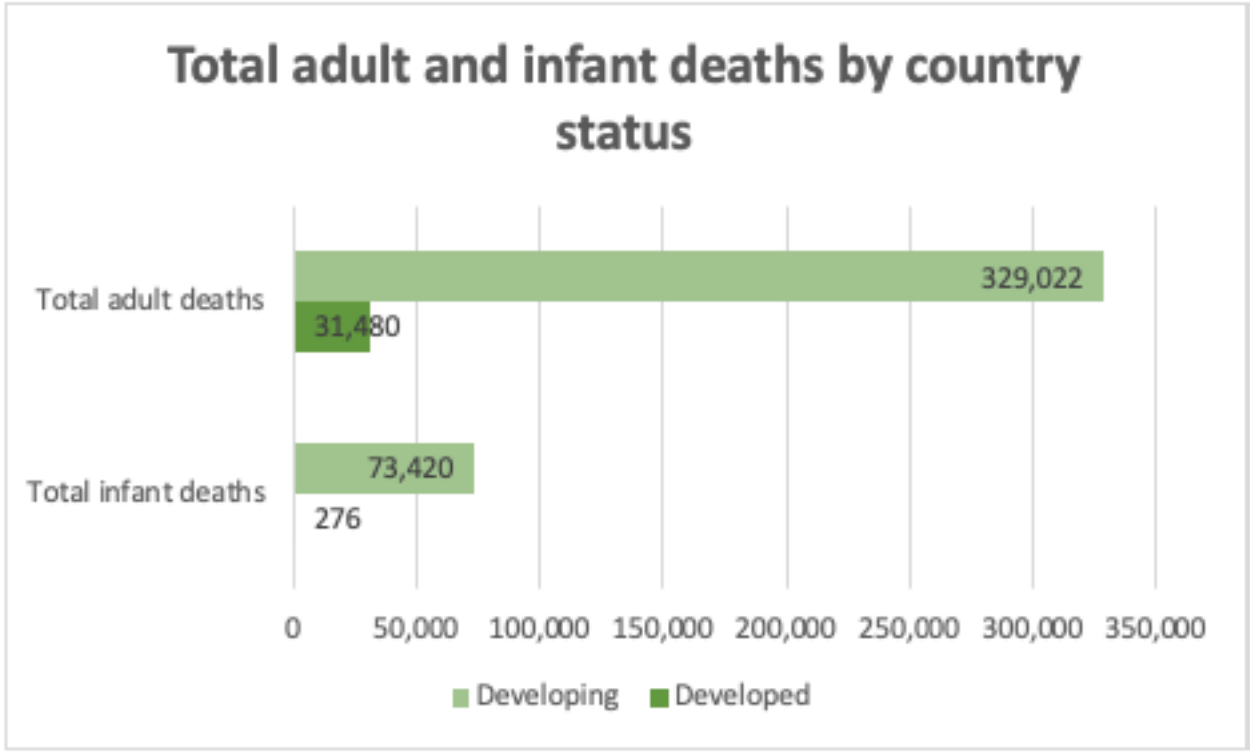
PYTHON

I used python seaborn to create a line graph for total expenditure and life expectancy the for all countries.

```
plt.figure(figsize=(14,6)) plt.title("Total expenditure vs Life expectancy")
sns.lineplot(x='life_expectancy', y='total_expenditure', data=LE, estimator=np.sum,
label='sum', error bar=None)
```




Life expectancy vs the impact of disease and immunisations between 2000 to 2014
Sierra Leone
Belgium
Average polio
70.8
97.7
Average HIV AIDS
1.6
0.1
Average diphtheria
71.2
97.2
Average life expectancy
46
81
0
20
40
60
80
100
120


Total adult and infant deaths by country status
Total adult deaths
329,022
31,480
Total infant deaths
73,420
276
0
50,000
100,000
150,000
200,000
250,000
300,000
350,000
Developing
Developed

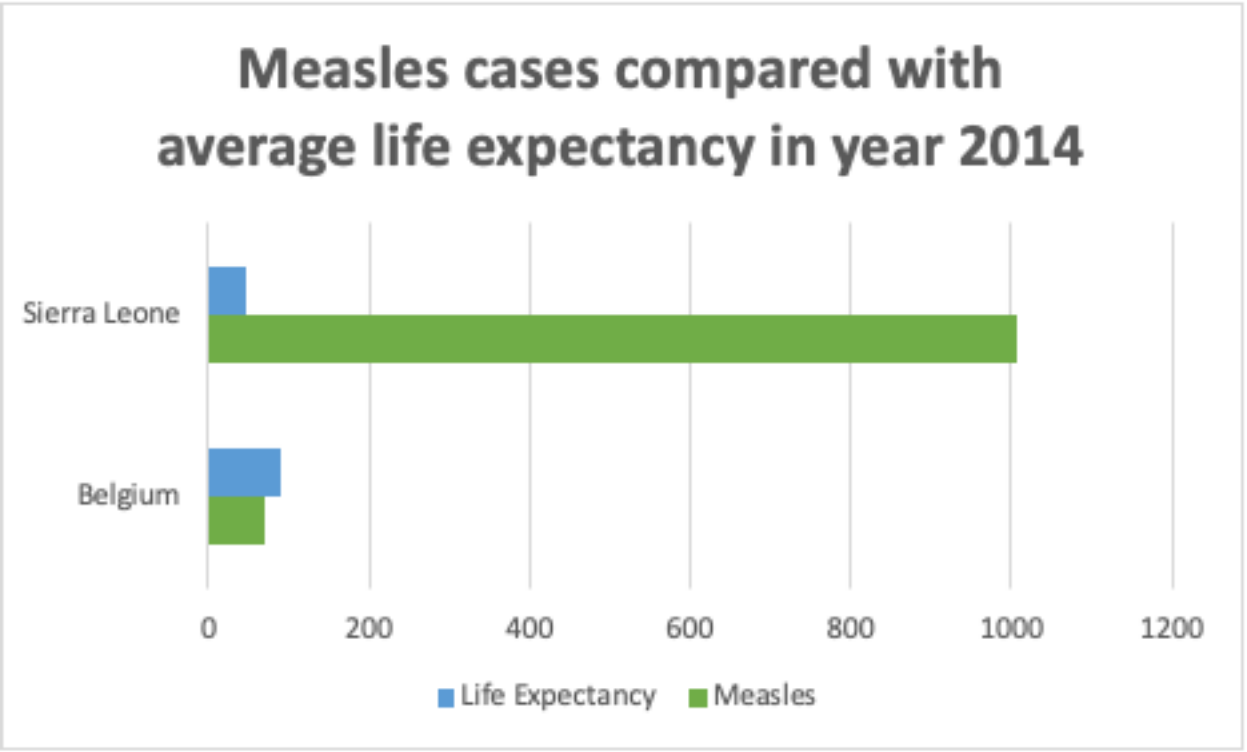
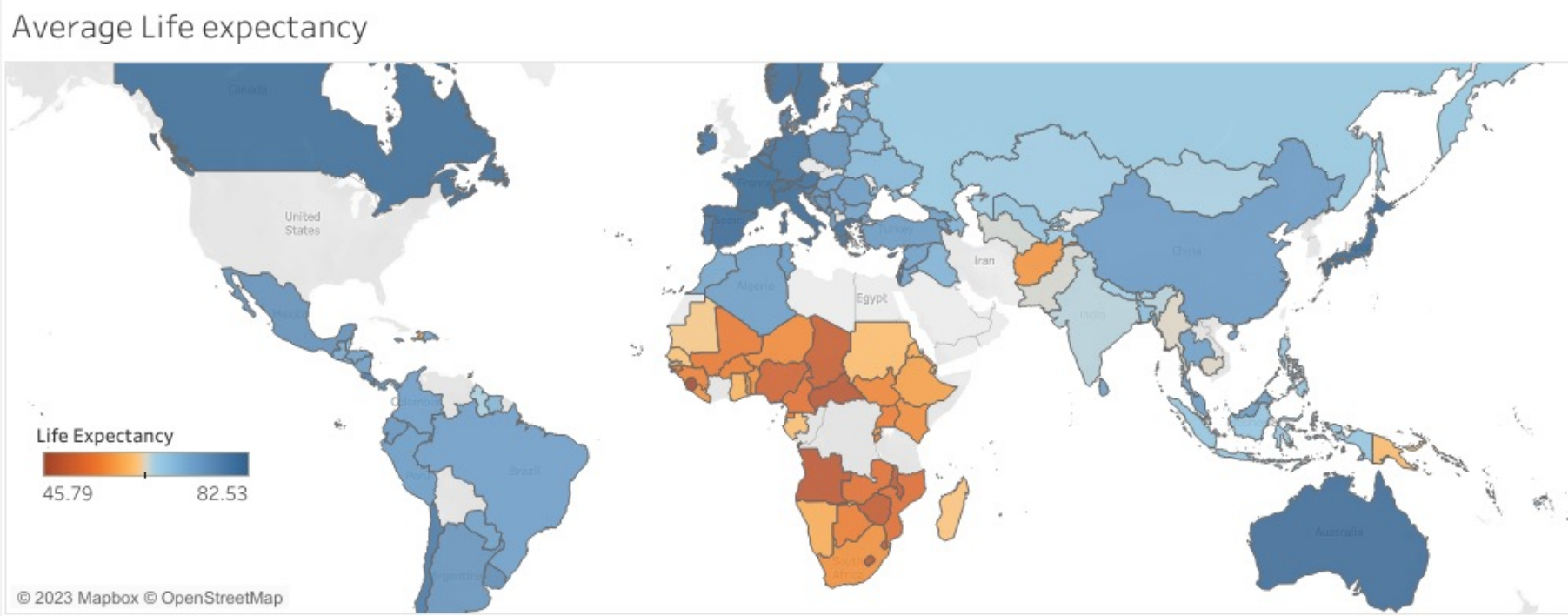
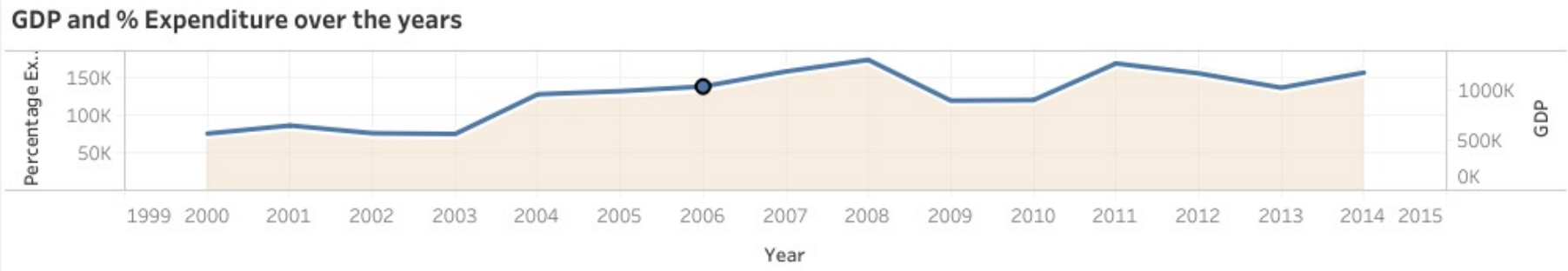
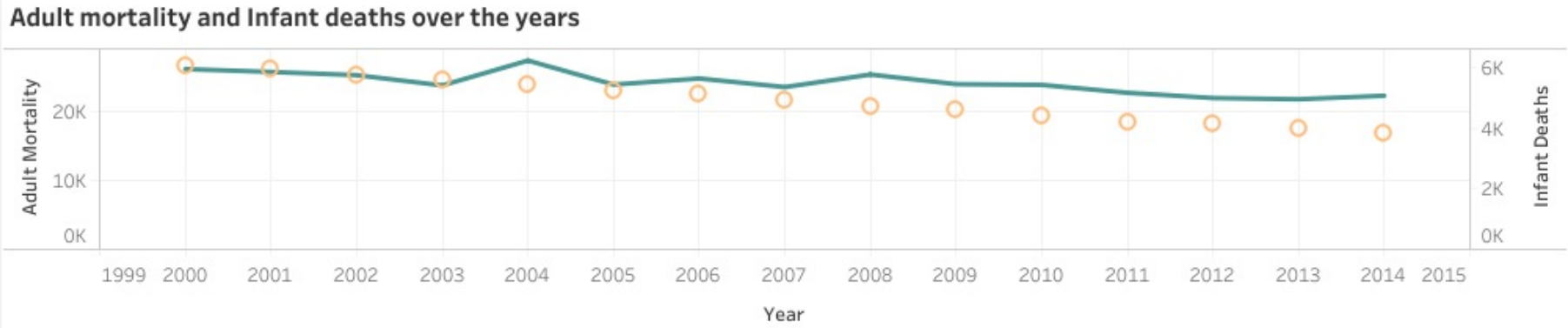

Measles cases compared with average life expectancy in year 2014
Sierra Leone
Belgium
0
200
400
600
800
1000
1200
Life Expectancy
Measles

Tableau Dashboard







Findings: The Importance of Life Expectancy

In Conclusion:

- **L**ife expectancy is a crucial metric for measuring the health and wellbeing of populations around the world. By examining global trends, exploring the factors that can impact life expectancy, and discussing the challenges and opportunities for improving it, we have gained a deeper understanding of this complex issue and how many of the factors are intertwined.

- There are many obstacles to achieving high levels of life expectancy, including poverty, inequality, and lack of access to healthcare.

- In summary, the economic status of a country is the biggest factor that effects life expectancy, as it is the precursor for many other factors.

Challenges

- A lot of the developing countries had significant amounts of missing data. In particularly when trying to explore life expectancy regarding other factors such percentage expenditure, GDP, total expenditure, population and more. Therefore, in the bulk of my analysis I couldn’t analyse D.R.Congo data as I had previously intended. So DRC data wasn’t included in the main analysis.

- Countries that do contain key data for key factors such total expenditure and GDP, should not be included.

- Explore schooling in regard to illiteracy further, in order to gage the true impact on life expectancy and education.

Linkedin : audreymayamba

Github : AudreyMx