5/29/2021

EDA REPORT

ON LIFE EXPECTANCY DATASET

Benedict Atta Biney, Alexander Kissiedu, Evans Ankomah

Group members

**TABLE OF CONTENT**

[INTRODUCTION 2](#_Toc73633298)

[HYPOTHESIS 2](#_Toc73633299)

[HYPOTHETICAL QUESTIONS 2](#_Toc73633300)

[METHODOLOGY 2](#_Toc73633301)

[Importing Liberties used 2](#_Toc73633302)

[Understanding and cleaning of Dataset 3](#_Toc73633303)

[Dealing with Missing Values 4](#_Toc73633304)

[Handling null values 5](#_Toc73633305)

[HYPOTHETICAL ANALYSIS 6](#_Toc73633306)

[**Question 1.** What is the distribution of countries by their status? 6](#_Toc73633307)

[**Question 2:** What is the life expectancy trend of over the years? 7](#_Toc73633308)

[**Question 4:** How do the various immunizations (Hepatitis B, Polio and Diphtheria) relate to life expectancy? 9](#_Toc73633309)

[**Question 5:** Which countries have high life expectancy despite having low schooling? 10](#_Toc73633310)

[CONCLUSION 11](#_Toc73633311)

[GROUP MEMBERS 11](#_Toc73633312)

# INTRODUCTION

According to Wikipedia, Life expectancy is a statistical measure of the average time an organism is expected to live, based on the year of its birth, its current age, and other demographic factors including biological sex.

This report is based on dataset from Global Health Observatory (GHO) data repository under World Health Organization (WHO) for the period of 2000-2015 for all countries. We will explore the life expectancy dataset which consists of different countries' data and gain insights from it. We used various tools and libraries such as such as numpy, pandas, matplotlib and seaborn for our exploration.

# HYPOTHESIS

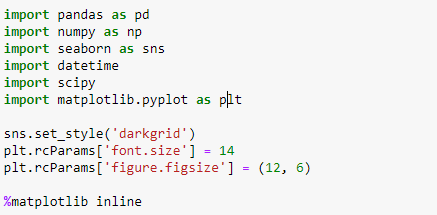
Life expectancy in developed counties seem to be better than that of developing countries.

# HYPOTHETICAL QUESTIONS

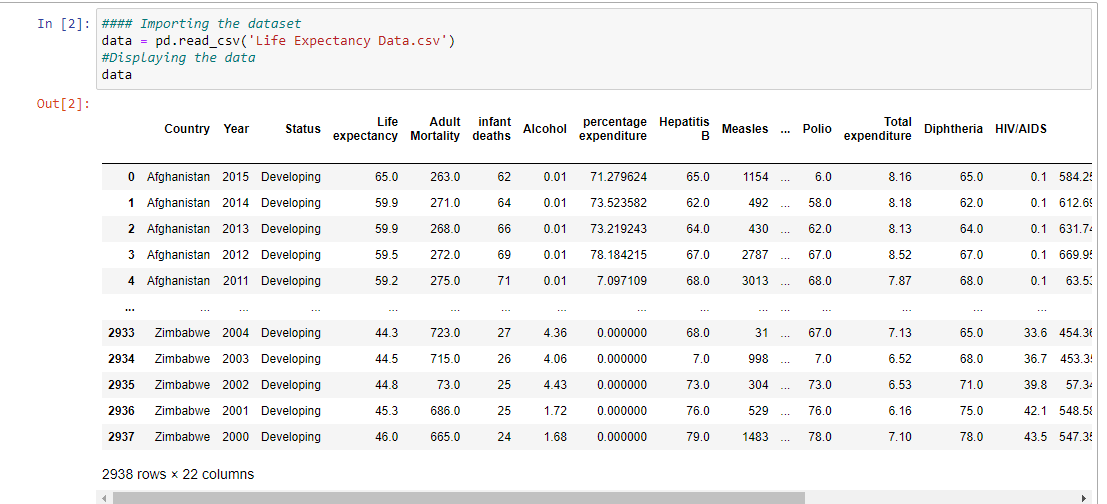
1. What is the distribution of countries by their status?
2. What is the life expectancy trend of over the years?
3. What is the average life expectancy in developing and developed countries each year?
4. How do the various immunizations (Hepatitis B, Polio and Diphtheria) relate to life expectancy?
5. Which countries have high life expectancy despite having low schooling?

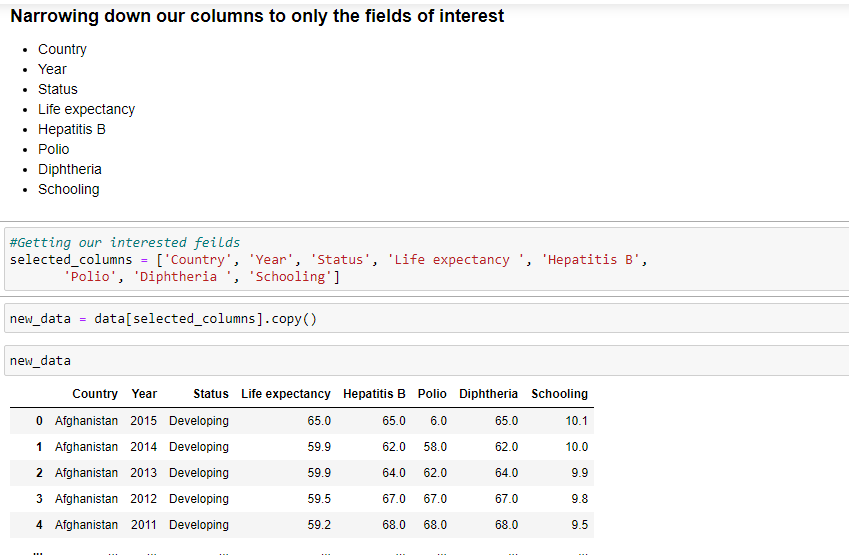
# METHODOLOGY

## Importing Liberties used

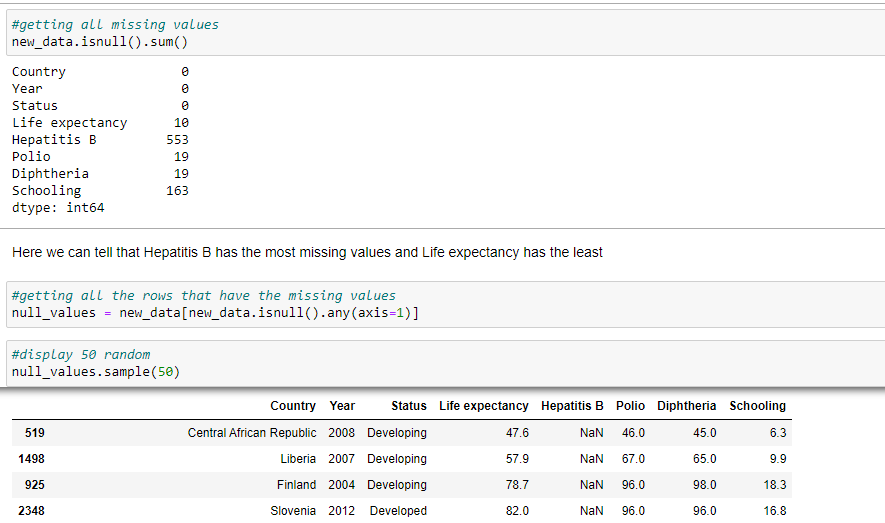


## Understanding and cleaning of Dataset



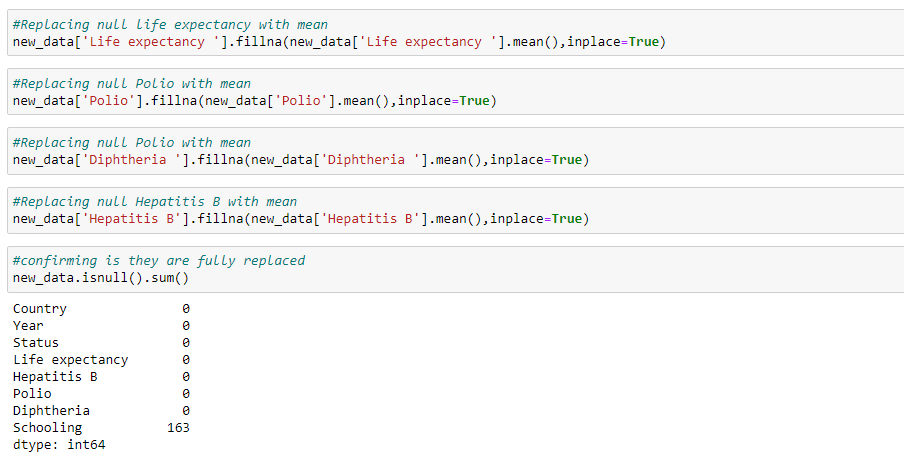
We limited our dataset to the following variables: Country, Year, Status, Life expectancy, Hepatitis B, Polio, Diphtheria, Schooling since those are the ones, we wanted to focus on.

## Dealing with Missing Values

We first checked for null values 

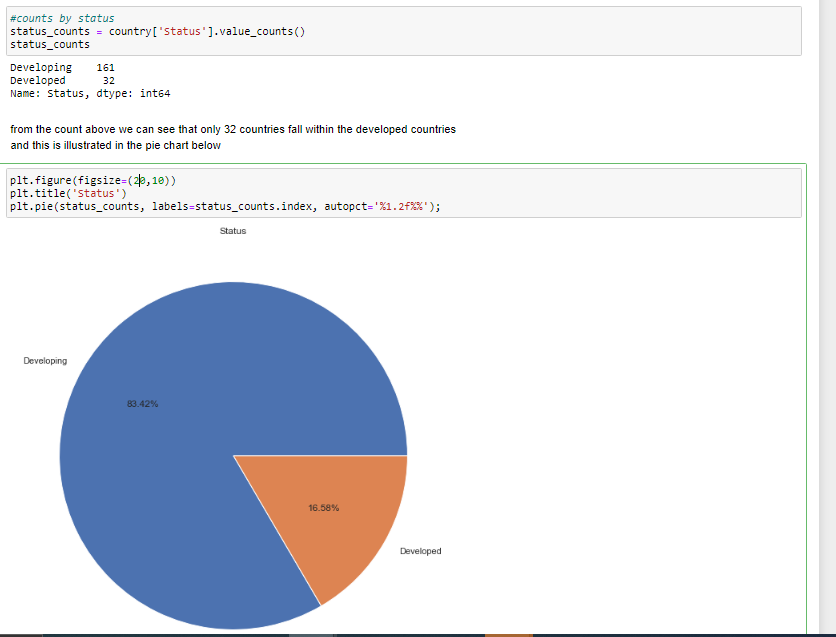
## Handling null values

After observing our dataset, we decided to do replacement for life expectancy, Polio and Hepatitis B using their mean values because they are lower than the 75th percentile and hence will not introduce outliers

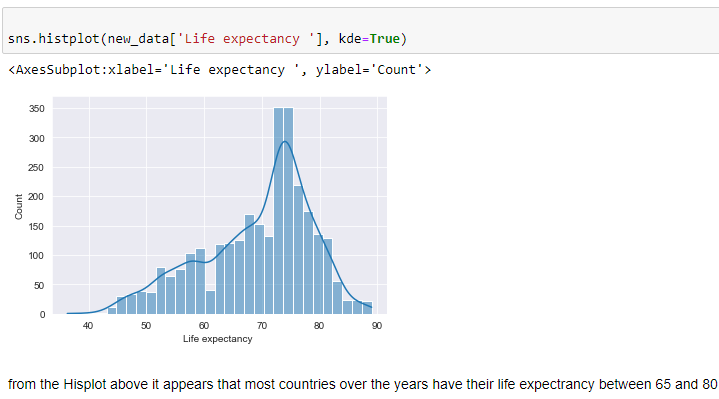


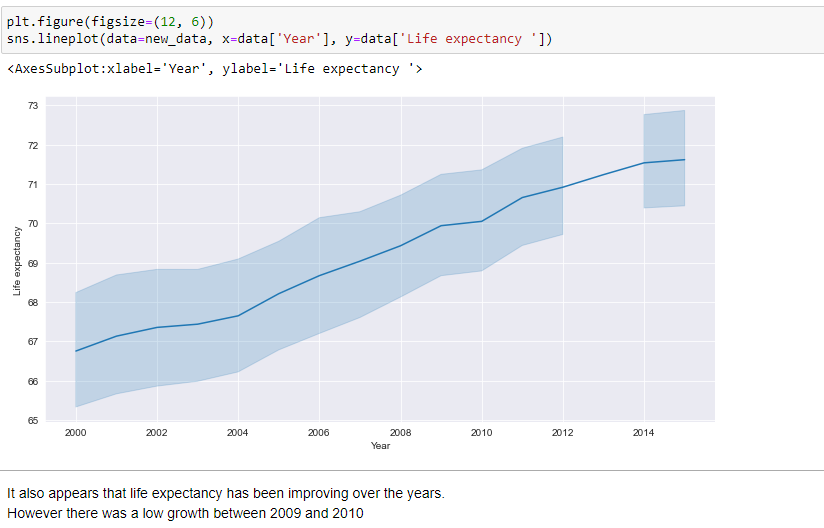
# HYPOTHETICAL ANALYSIS

## **Question 1.** What is the distribution of countries by their status?

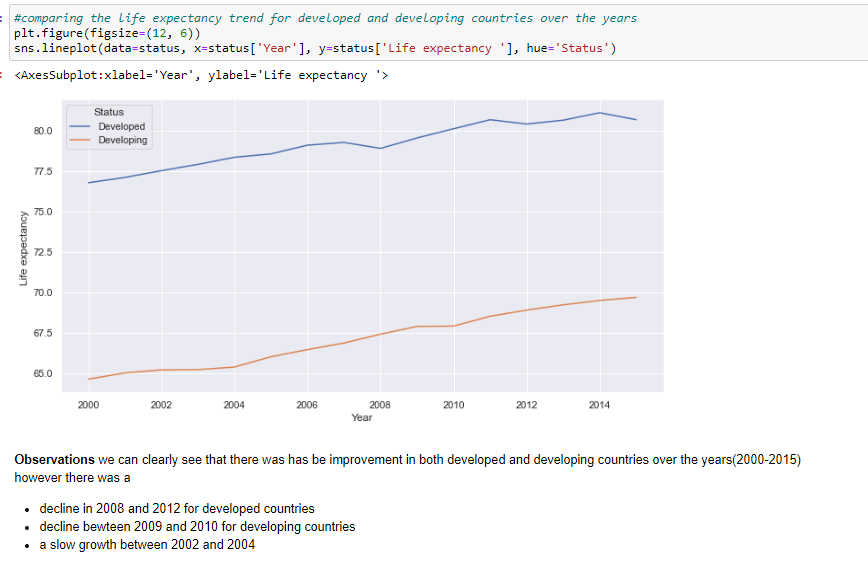
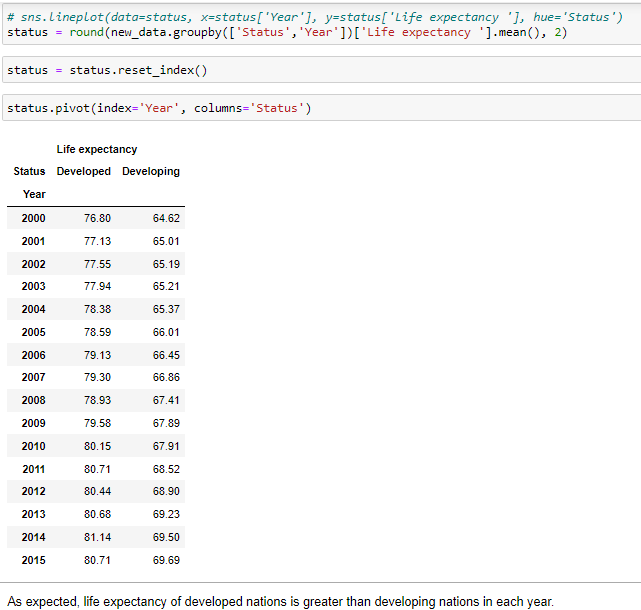


## **Question 2:** What is the life expectancy trend of over the years?

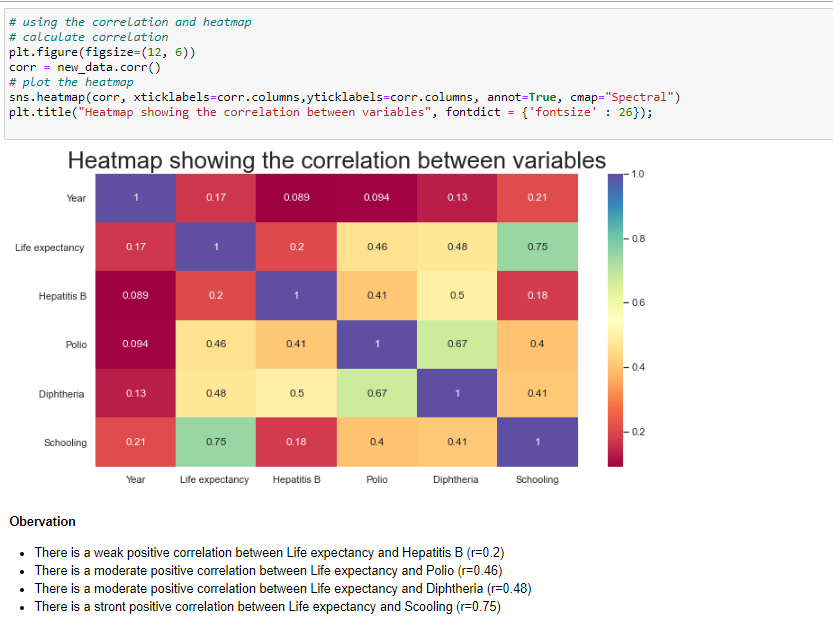




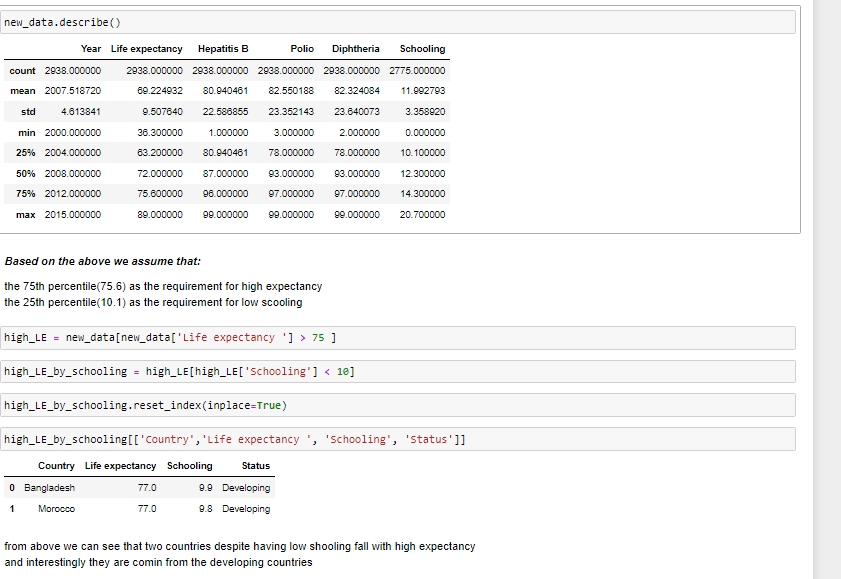
**Question 3:** What is the average life expectancy in developing and developed countries each year?



## **Question 4:** How do the various immunizations (Hepatitis B, Polio and Diphtheria) relate to life expectancy?



## **Question 5:** Which countries have high life expectancy despite having low schooling?



# CONCLUSION

From the above exploration analysis, we can confirm that life expectancy rate is better among developed countries (16.58% of our dataset) than developing countries. However, we also made the following interesting observations

1. Above 75% of the countries have life expectancy above 60. And around 25% of the countries have life expectancy ranging from 70 to 75.
2. More than half of the countries reached their highest life expectancy in 2015.
3. Two countries have high life expectancy despite having low schooling rate and all of them are from developing countries.
4. Each year from 2000 - 2015, the average life expectancy for developing countries was below 68 whiles it was above 79 for developed ones.

So, to conclude we can strongly say that schooling and immunization interventions (Polio, Hepatitis B, and Diphtheria) are somehow linked to life expectancy rate, that is an increase in Schooling and immunization intervention (Polio, Hepatitis B, and Diphtheria) will improve life expectancy. Therefore, life expectancy is better in developed countries than developing ones.