A Report on

STAT-S 670 Mini Project: Life Expectancy

By

Harshwardhan Patil (hrpatil@iu.edu)

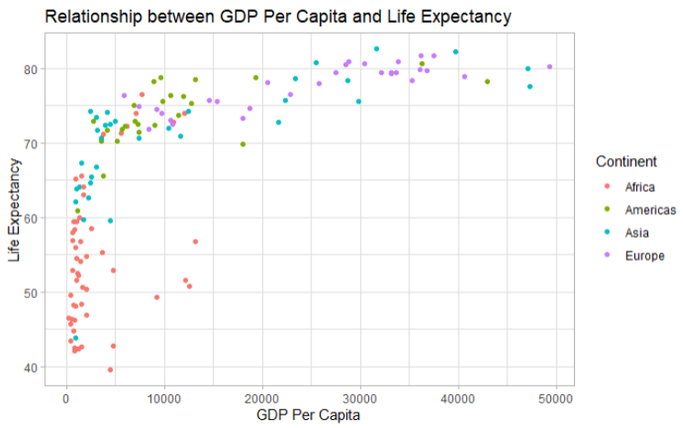
Manasa Gudise (mgudise@iu.edu)

Radhika Agarwal (radagar@iu.edu)

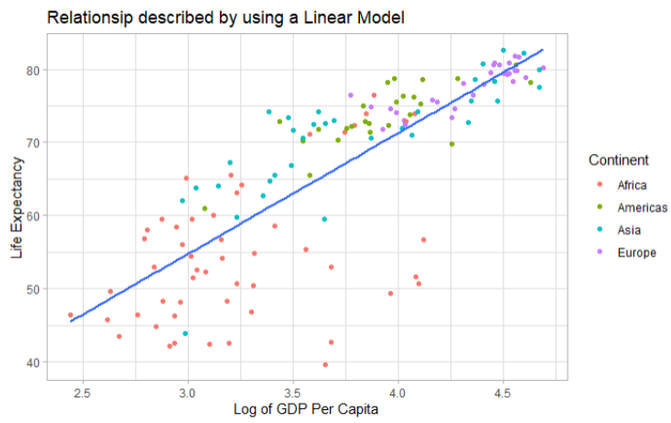
1. GDP and Life Expectancy in 2007:

Question 1.1: How does life expectancy vary with GDP per capita in 2007? Can the trends be well-described by a simple model such as a linear model, or is a more complicated model required?

To view the relationship between the GDP per Capita and life expectancy, we will plot a scatter plot. Scatter plot is used to describe relationship between two variables. Here, we will not consider continent Oceania, as it contains few countries compared to the other continents.

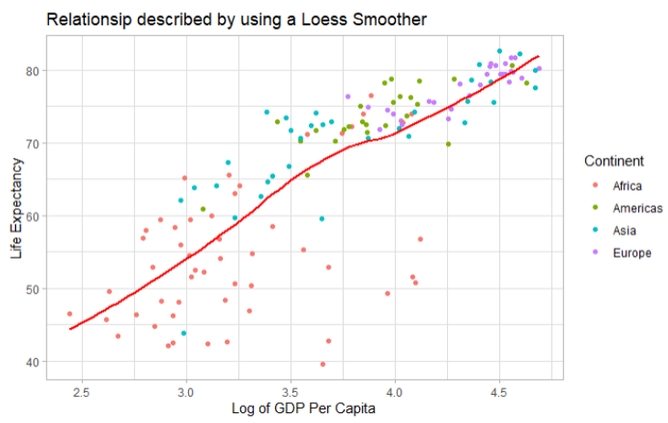


In general, for the year 2007, the life expectancy increases with the increase in the GDP per Capita. However, the relationship between GDP Per Capita and Life Expectancy is not linear. For low values of GDP, life expectancy increases rapidly, however for high values of GDP, it increases with mild progression. Therefore, the relationship will be linearly described by logarithmic transformation.



Here, the linear model explains the relationship well. However, from 3.25 to 4.0, it does a poor job in fitting the data. So, we have to look at different data fitting methods. One of the methods is to fit the data using LOESS smoother.

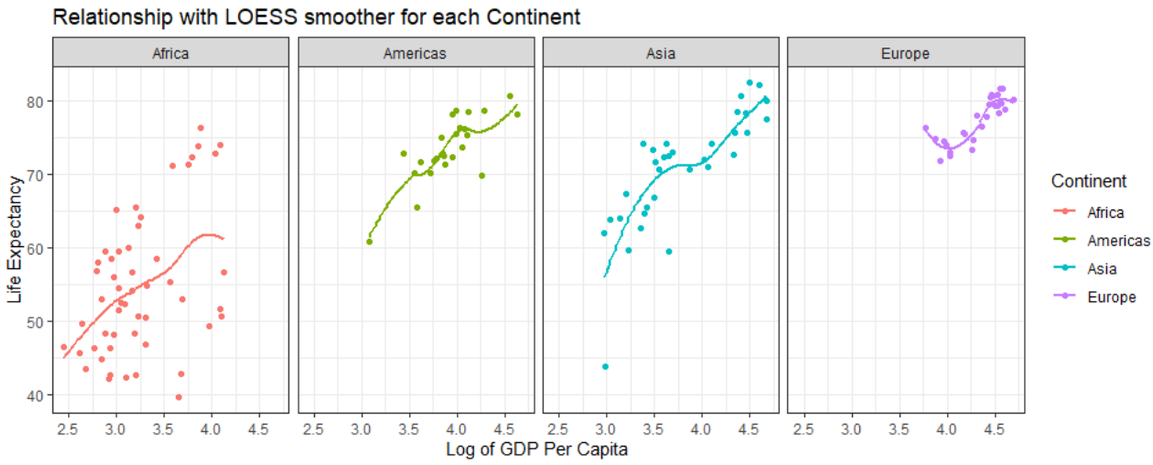
LOESS smoother is a nonparametric method, which means it will take no assumptions about the distribution of the data. The basic idea behind LOESS is to fit a weighted regression line to the data, where the weights are higher for points that are closer to the point being estimated and lower for points that are further away. This results in a smooth curve that captures the underlying trend in the data while reducing the impact of random fluctuations or noise.



From the plot, we can conclude that, the LOESS does a reasonably good job than linear model for fitting the data.

Question 1.2: Is the pattern the same or different for every continent? If some continents are different, which ones?

As mentioned in 1.1, we can fit the LOESS smoother to view the relationship. Additionally, we can visualize relationship for each continent with the help of facet grid.

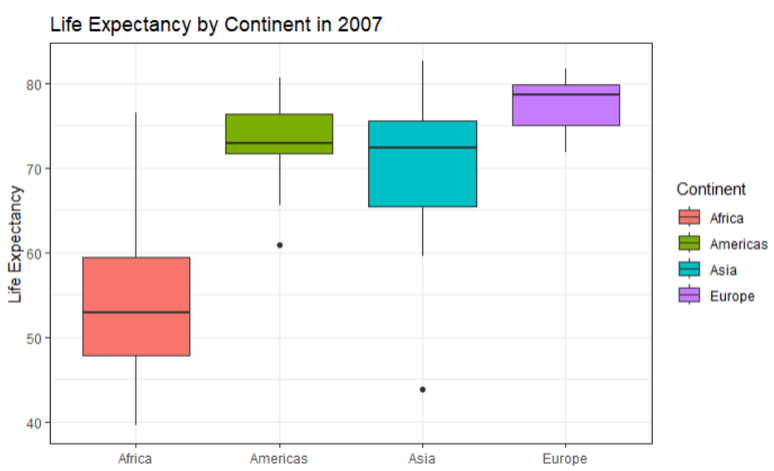


From the above plots, we can infer that the general pattern is same for each continent, meaning that with increase in the GDP per Capita, the life expectancy also increases. However, the rate of increase in life expectancy is different for each continent. Additionally, there are few points to consider.

- The relationship trend for Asia and America is almost similar, however it is different for Africa and Europe.
- Countries in Africa generally have a lower GDP per capita and life expectancy than rest of the continents and Europe has the higher GDP per capita.
- The pattern explained by the Africa, America, and Asia for lower values log GDP is fairly linear, but that's not the case with Europe.

Question 1.3: Can differences between continents be simply described by an additive or multiplicative shift, or is it more complicated than that?

As we know, we can determine whether the differences between continents described by an additive or multiplicative shift using a Box plot to compare the distribution. If the box plots for each continent would have similar shapes but varying center values, then we can implement additive shift and if the box plots for each continent would have comparable spreads but varying positions on horizontal axis then we can implement a multiplicative shift.

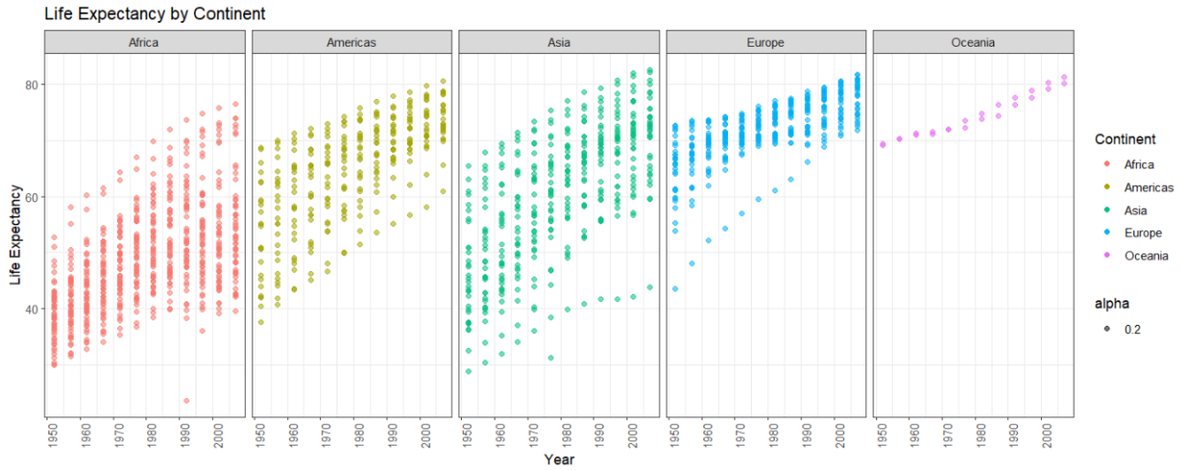


From the plot, we can see that neither the spread nor shape of box plot looks similar. So, it is evident that, the differences between continents are more complicated than a simple additive or multiplicative shift, we may need to use more complex models to describe the relationship between GDP per capita and life expectancy.

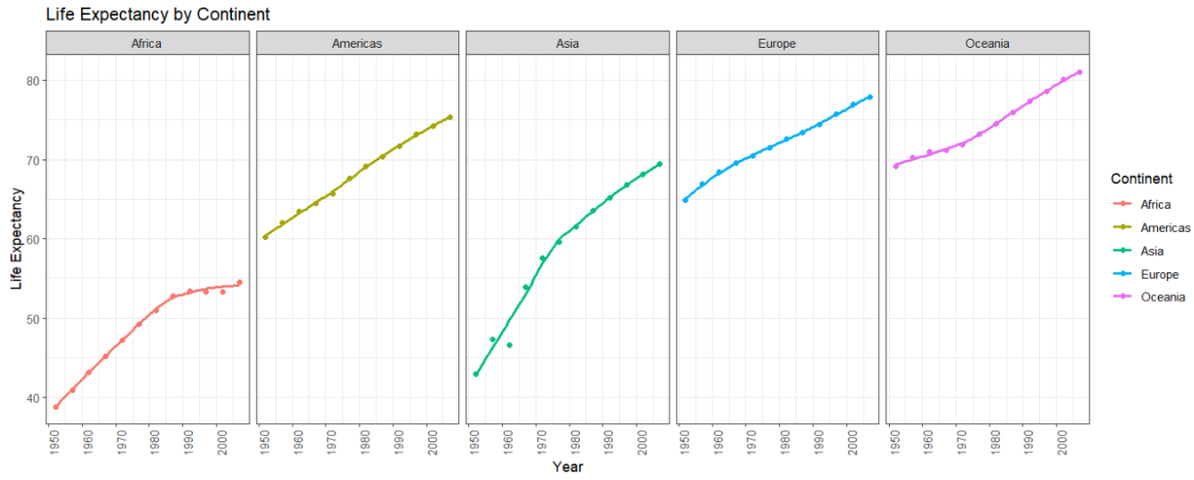
2. Life expectancy over time by continent:

Question 2.1: Has average life expectancy changed over time in each continent?

To visualize the change in life expectancy over the time in each continent, firstly, we will plot a scatterplot for life expectancy and the year for each continent using facet grid.



As we can see, this plot is distorted as it is showing values of life expectancy for each country present in the respective continent for every year. Also, the average population in China is greater than other continents, we have used weighted average of life expectancy by population to fit a loess smoother. The final plot will be given as follows:

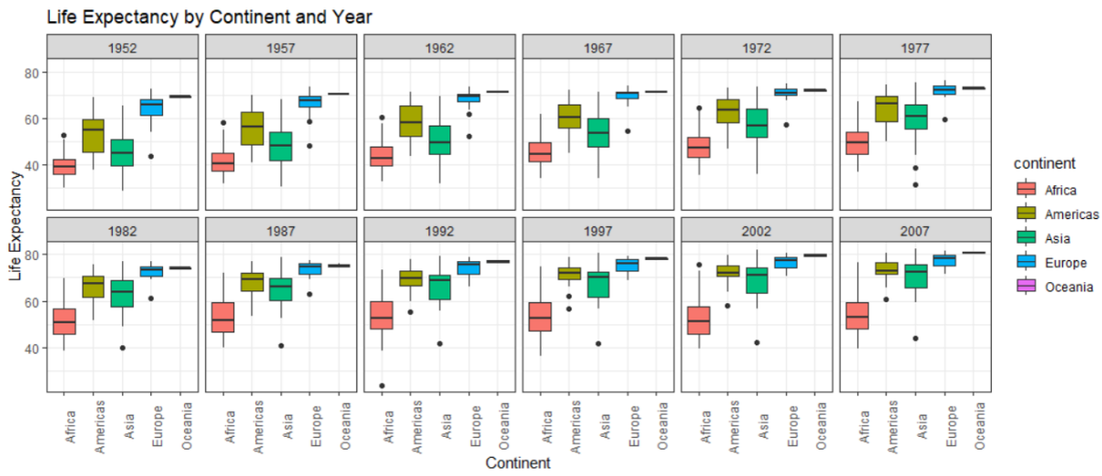


Finally, we can draw a conclusion that, the average life expectancy has almost linearly changed over time in each continent, except for the fact that the rate of increase is different in every continent and different time periods.

Question 2.2: Have some continents caught up (at least partially) to others? If so, is this just because of some countries in the continent, or is it more general?

- Yes, there has been some instances where some countries partially caught up. Asia caught up in life expectancy with America. If we have a look at the average life expectancy, in the early part (year 1950) people in America had higher average life expectancy than people in Asia. But this has been changed over time and we can say that Asia managed to increase its life expectancy over time, and thus caught up with America.
- This is not the case with people in Europe and Africa. We can see that from plot in 2.1, few countries in Africa have caught up with countries in the America and Asia despite having much lower life expectancy in the beginning.
- Also, for few countries in Asia, life expectancy was below than the countries in America and Europe, but at the end, these countries had higher life expectancy than rest of the continent.

To get in-depth understanding of the variation of life expectancy over the years, we can plot a box plot and look at the distribution.

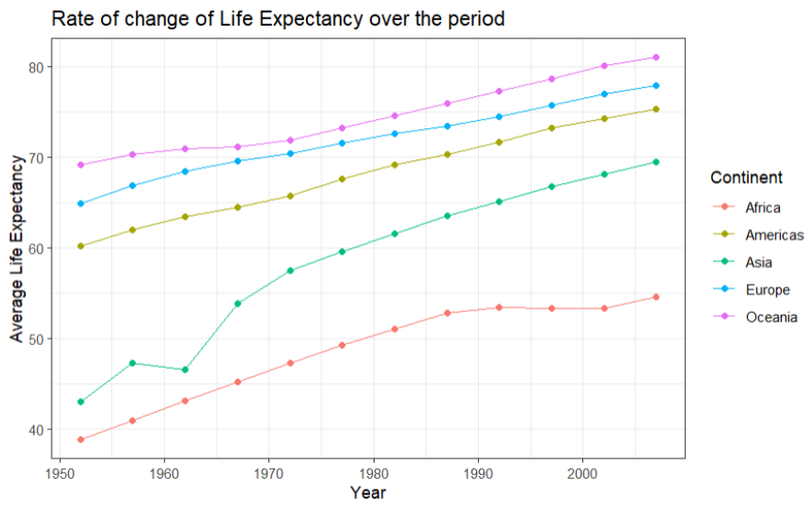


Looking at the above box plots, we can surely say that in terms of life expectancy the continent of Asia is caught up with America. Also, Europe managed to catch up with Oceania till the year 1987.

This trend is more general and not just due to a few countries within each continent. However, there still exists variations within continents, particularly in Asia, Europe, and America where some countries have much higher life expectancy than others.

Question 2.3: Have the changes been linear, or has it been faster/slower in some periods for some continents?

To visualize the change in the life expectancy for each continent as the time progresses, we can use a line plot.



For most of the years, the changes in the average life expectancy have been linear. However, the plot shows that the rate of change in average life expectancy has not been constant over time and across continents.

- In Africa, a rapid increase in the average life expectancy is observed till year 1987 and after that it become steady for a while and then saw a gradual increase after year 2000.
- For Oceania, the increment in life expectancy was much slower in the initial years, but after 1972, the life expectancy started to climb rapidly.
- Europe and America shown a consistent but slower improvement in life expectancy over time.
- A steady increase in life expectancy is observed in Asia, before dropping in 1957. This decrease in life expectancy continued till 1962. From 1962 to 1967, a steep increase in the life expectancy. After 1967, the life expectancy increased gradually.

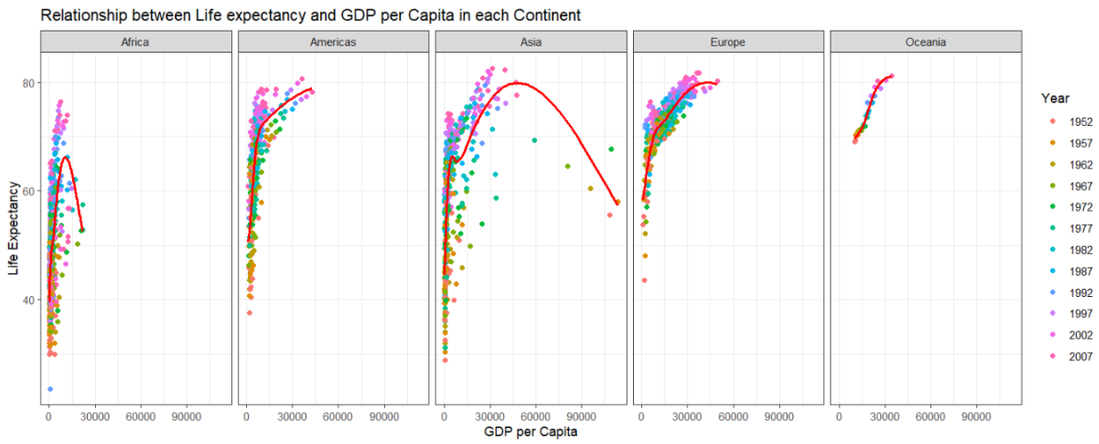
Question 2.4: What might explain periods of faster/slower change?

There could be many factors that explain periods of faster or slower change in average life expectancy. Some factors like changes in population and its distribution could be evident. Also, higher GDP per capita often correlates with better access to healthcare and other resources that impact life expectancy. Additionally, technological advancements, political and social factors such as wars, epidemics, and changes in government policies might have also impacted the rate of increase in life expectancy.

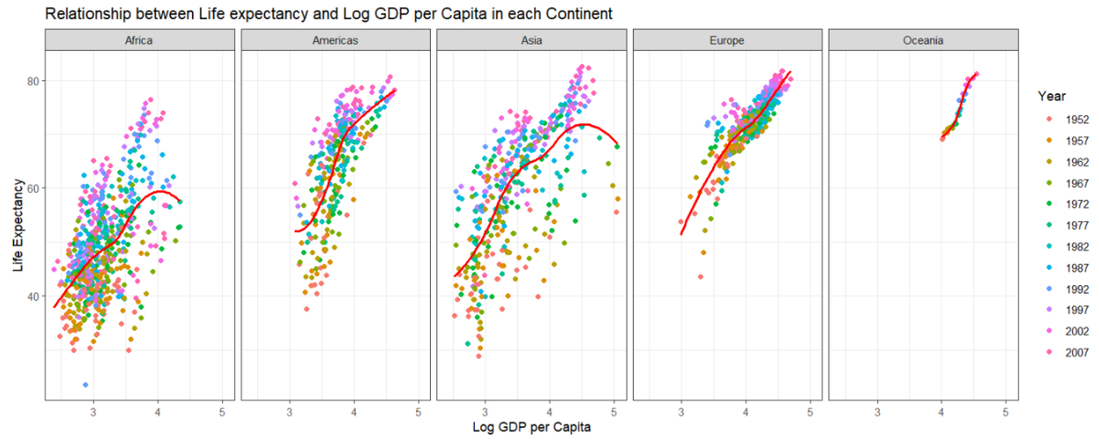
3. Changes in the relationship between GDP and life expectancy over time:

Question 3.1: How has the relationship between GDP and life expectancy changed in each continent?

To analyze the relationship between GDP and life expectancy over time, we can create scatter plots of life expectancy versus GDP per capita for each continent and color the points by year.



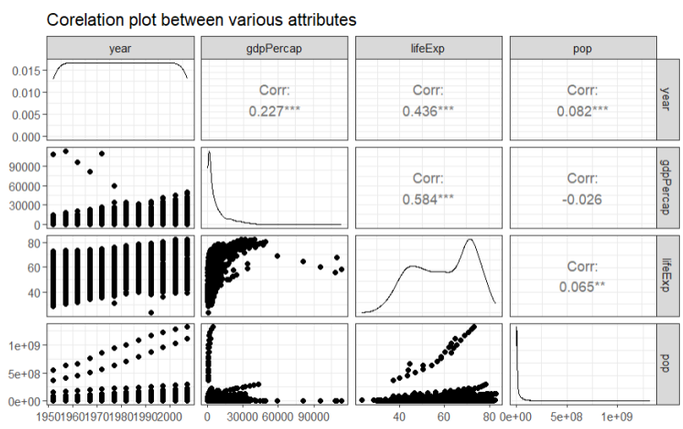
Here looking at the plots, it looks like the GDP per Capita has no significant impact on the changes in the life expectancy in Africa, America and Asia. Looking at the scatter plot, the relationship between GDP and life expectancy in each continent will be explained more concisely by using a logarithmic transformation. So, we will use log transformation of GDP per Capita to make the relationship linear.



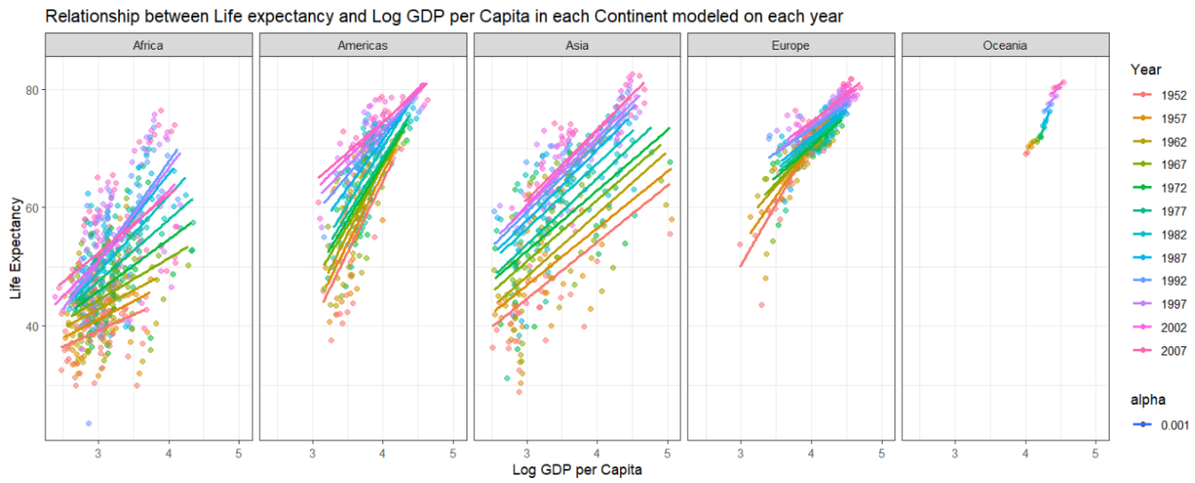
Looking at each facet, we can see that there is general trend of increase in Life expectancy with respect to GDP per capita over the years for each continent. However, for Africa, the growth rate of Life expectancy with respect to GDP per capita is not linear, thus it is difficult to find a conclusive trend in Africa. Asia has also lot of distortion in the points, and the variation in the life expectancy with respect to GDP per Capita cannot be explained by a simple linear model. However, in America, Europe and Oceania the relationship is fairly linear.

Question 3.2: Can changes in life expectancy be entirely explained by changes in GDP per capita? Does it look like there's a time effect on life expectancy in addition to a GDP effect?

To explain the changes, we can plot a correlation matrix and find out, whether there is strong correlation exists between life expectancy and other variables, such as year and population.



So, the changes in life expectancy cannot be entirely explained by GDP per Capita. It appears that there is strong correlation exists between life expectancy and year as well.

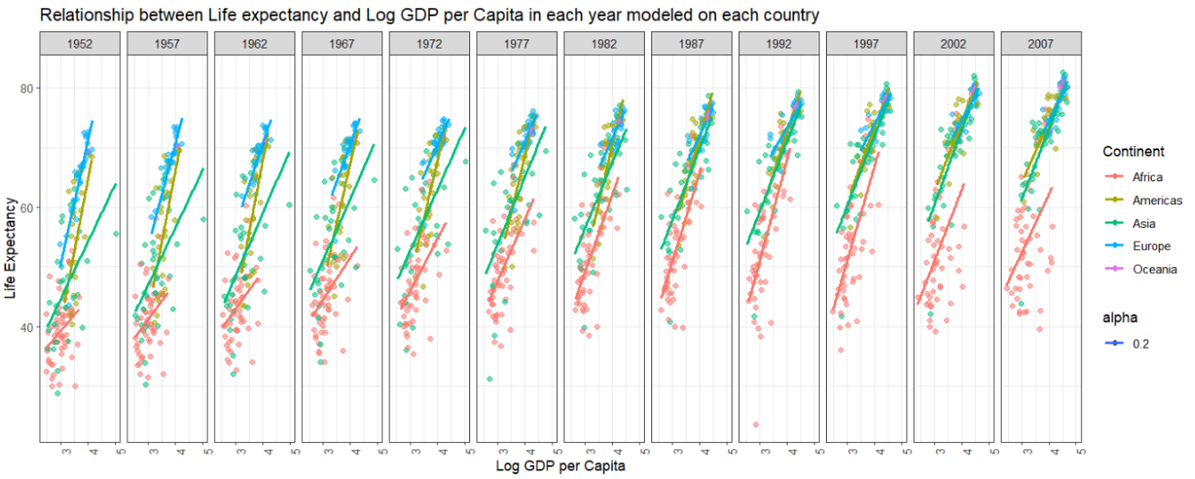


For each continent, the slope of the smoother is different which shows that the changes in life expectancy are not similar and there exists a variation which can be explained by other factors such as time and population. As the time increases the life expectancy tends to increase as well, even though there is no increment in the log GDP per Capita.

Question 3.3: Has there been "convergence" in the sense that perhaps GDP and/or continent don't matter as much as it used to? Are there exceptions to the general patterns?

We can state that there is some evidence of convergence in the relationship between GDP per capita and life expectancy over time. The scatter plots for different continents do show some similarities in terms of the overall trend, although there are still some differences in the strength of the relationship between GDP per capita and life expectancy.

However, there are also some exceptions to the general patterns. For example, some countries in Africa continue to have very low life expectancies despite increases in GDP per capita. In addition, there are some countries in Asia and Europe that have higher life expectancies than we would expect based on their GDP per capita.



Overall, if we look at different years, the slope of loess smoothers is not always similar for each continent over the period. Till 1987 there has been no convergence observed. But After 1987 there has been convergence as the slope of the smoothers for continents America, Asia, Europe, and Oceania is almost similar, Africa being an exception.

Conclusion:

The increase in life expectancy since World War 2 can be explained by increases in GDP per capita. However, the trend is not always increasing and linear. The rate of increase varies in different time periods and in different continents as well. Additionally, external factors such as technological advancements, political and social factors such as wars, epidemics, and changes in government policies might have also impacted the rate of increase in life expectancy.