Stat 470/670 Mini Project 1: Life Expectancy

Group Members: Rajesh Sharma: rajeshar
Ashish Patidar: apatida

GDP and Life Expectancy in 2007

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

Figure 1.1 shows the graph between life expectancy and GDP per capita for the year 2007. As we can see, Life expectancy increases with an increase in GDP per capita, but the points are scattered all over; therefore, we can conclude that this data does not follow any pattern, and hence fitting a linear model is not the best choice here. The blue line in the graph shows the linear model, and we can clearly see it is not a good fit. To reinforce this statement can also see the Homoscedasticity and normality using QQ plots of residuals, referring to appendix figure A.1.1 and figure A.1.2. respectively.

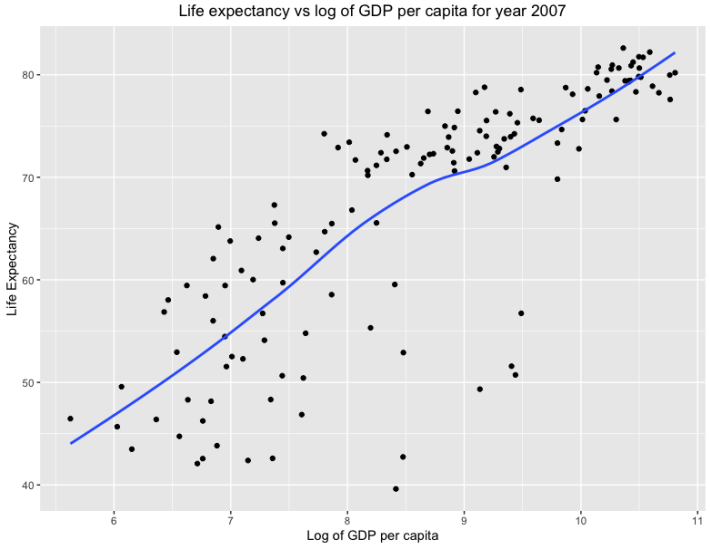

Fig 1.1.- Life Expectancy VS GDP per capita in 2007

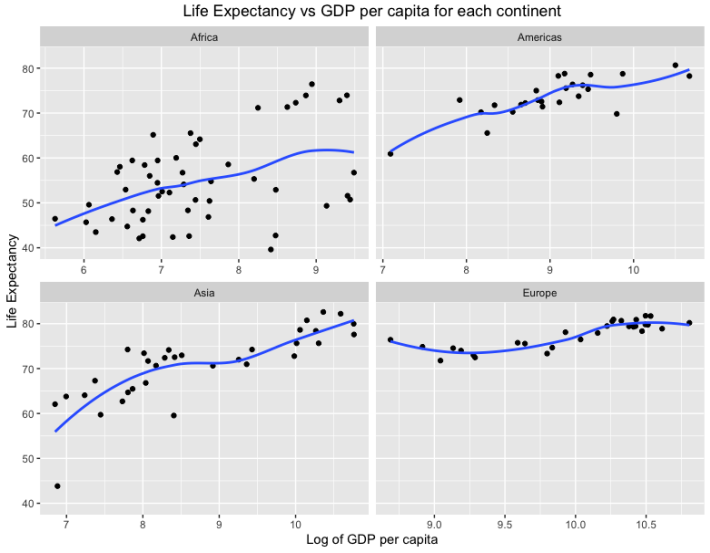

Fig 1.2.- Life Expectancy VS GDP per capita for each continent in 2007

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

From figure 1.2, we can say that the general pattern is the same for all continents: Life expectancy is increasing with an increase in GDP per capita. However, the rate of increase for each continent is different from one another, and the x-axis of each continent is different, so we can conclude the pattern is not the same for each continent.

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

From figure 1.2, it could be seen that each continent follows a different pattern, and describing them by an additive or multiplicative shift with others is not possible. So, we can conclude that describing the pattern for one continent with respect to another is much more complicated.

Life Expectancy over time by Continent

How has average life expectancy changed over time in each continent?

For each continent, overall the average life expectancy has increased over time but from figure 2.1 we can say that continents like the Americas, Europe, and Oceania have an approximately linear rate of increase but for Africa, the rate was linear till 1987 then a sudden decrease in that rate was seen. If we talk about Asia, we can see kinks for the years 1957 and 1962 where the rate suddenly dropped in the year 1957 and then increased suddenly in the year 1962.

Approximate life expectancy graph

LIFE EXPECTANCY (approx)	1952	2007
Africa	38	54
Americas	60	75
Asia	43	68
Europe	65	77
Oceania	69	81

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?

Referring to figure 2.1. Asia has caught up with the Americas and others or not exactly caught up but tried to catch them. This was possible because of countries like India, China, and Japan. As we are taking a weighted average here based on the population of countries, increase in life expectancy and population of these countries over the years was one primary reason for the catch-up. Can also visualise individual trends for each continent as visualised in appendix figure A.2.1.

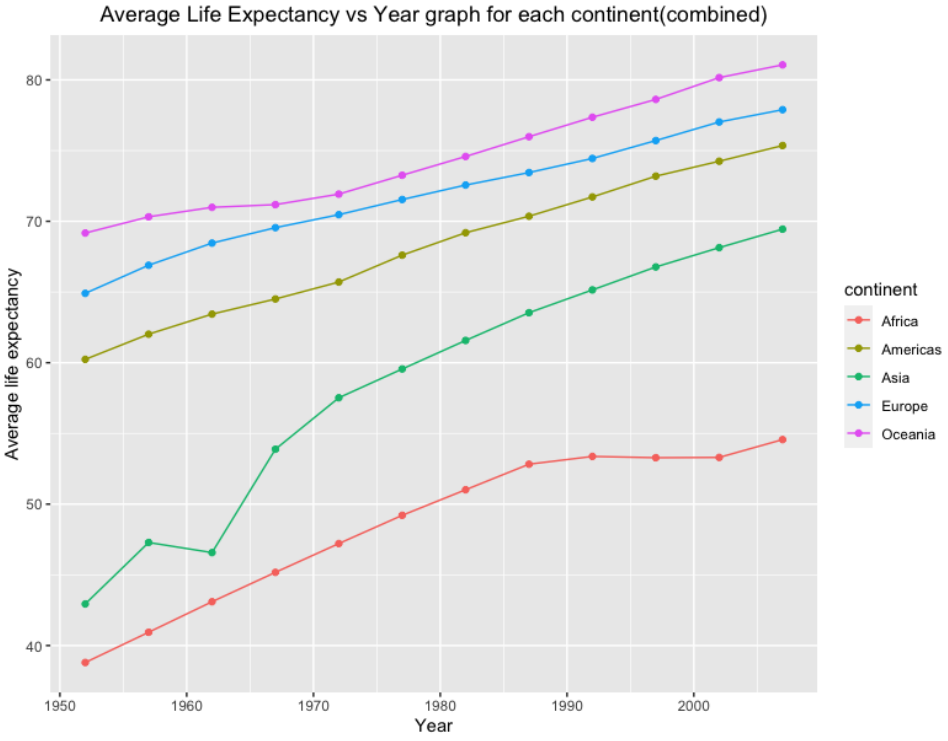

Fig 2.1.-Average Life Expectancy over time by continent

Have the changes been linear, or has it been faster/slower in some periods for some continents? What might explain periods of faster/slower change?

Other than Africa and Asia, all the continents have approximately linear increases. However, Africa saw a decline in life expectancy because some countries like Zambia, Uganda, Zimbabwe, Rwanda, Namibia, Lesotho, Liberia, and Botswana faced a decrease in life expectancy between the years 1975 - 2000. These are the observations from figure 2.2.

Asia saw ups and down between the time frame of 1952-1962. The primary reason can be that China's life expectancy changed in the same manner as overall Asia. This is mainly because China is the most populated country globally and was for the years 1952-1962; figure 2.3 is used for this conclusion.

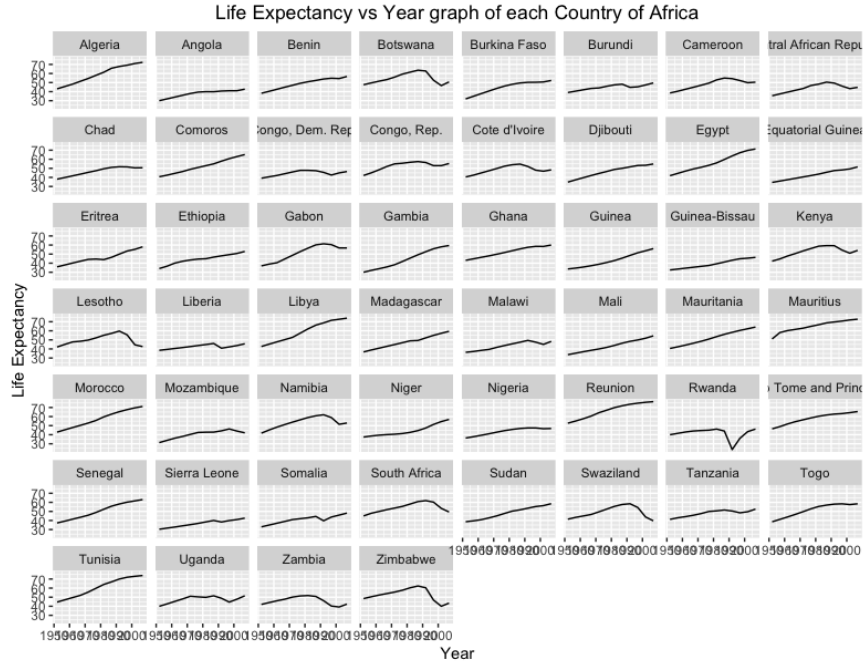

Fig 2.2.-Life Expectancy VS Year of African countries

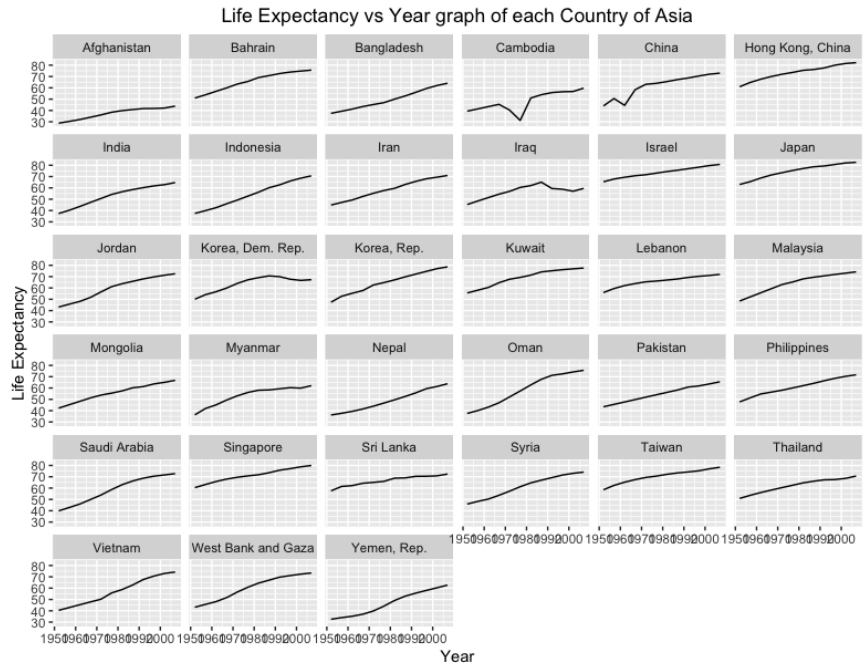

Fig 2.3. -Life Expectancy VS Year of Asian countries

Changes in the relationship between GDP and Life Expectancy over time

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

From figure 3.1, we can say that GDP and Life expectancy both increased over the years for each continent, and it is a general observation. However, for Africa, as we know, the growth rate of Life expectancy with respect to GDP per capita was not linear(figure 2.1), because of which it is complicated to find this trend in the scatter plot of Africa; therefore, we can see points being scattered all over the places, same can be the reason for distortion in trend for Asia. Europe follows this trend, and we can see this relationship could be defined for Europe using a linear model but for other continents, we have to fit a more complex model.

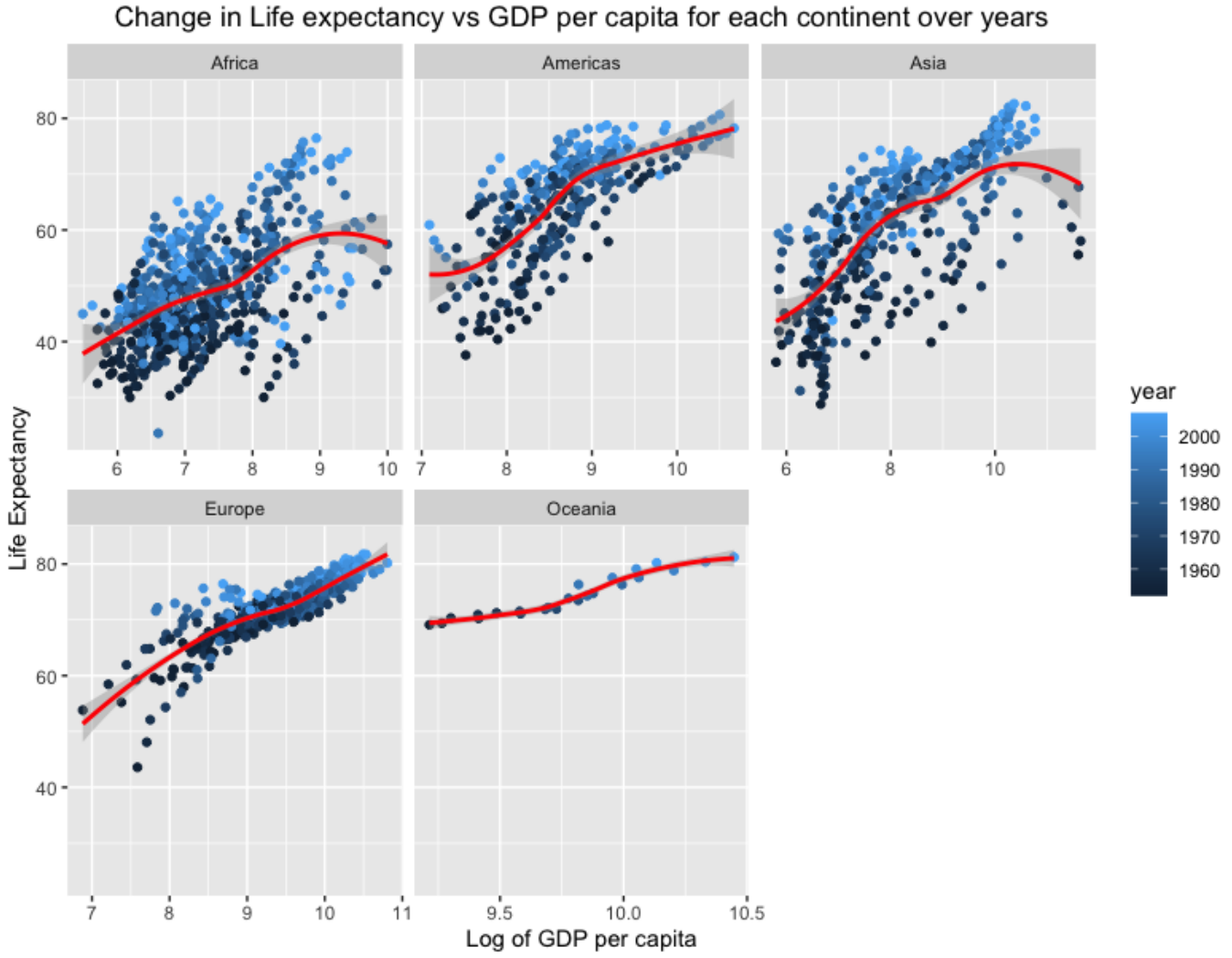

Fig 3.1. - Change in Life Expectancy

Can changes in life expectancy be entirely explained by changes in GDP per capita?

In figure 3.2 we can see a linear model is fitted over life expectancy vs GDP per capita graph, but to see if this linear model is the best fit for data, we have plotted a QQ plot for residuals as shown in figure 3.3 and a point plot for residuals to check homoscedasticity in appendix figure A.3.1. From this, we can see the model is pretty well fitted, but the r-squared value is 0.7054 suggests that life expectancy cannot be entirely explained by GDP per capita, and there are some other factors responsible for the same.

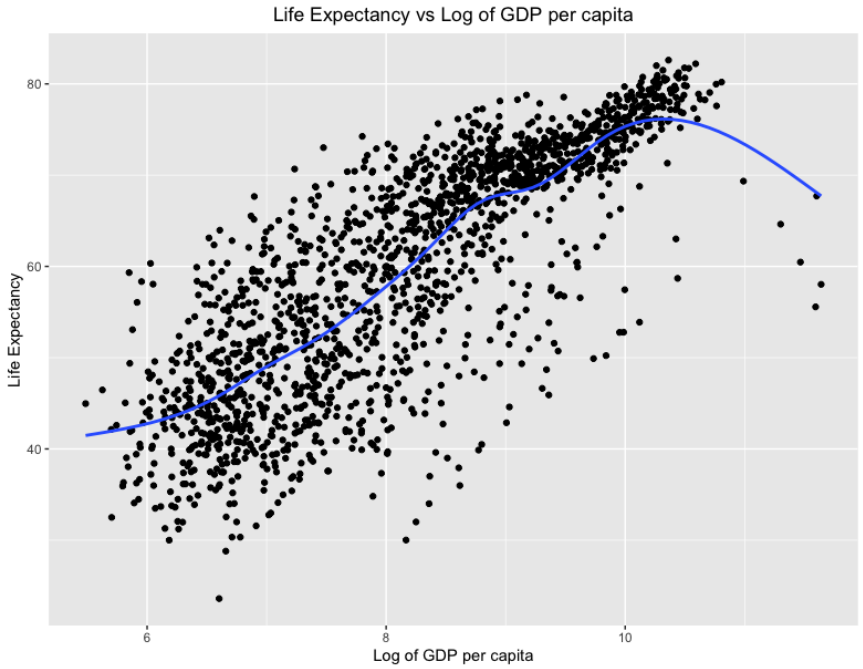

Fig 3.2.- Life Expectancy VS GDP per capita

Does it look like there's a time effect on life expectancy in addition to a GDP effect?

We fitted a linear model lifeExp ~ (gdp_log + year) to see if we get a better fit than lifeExp ~ gdp_log, and the difference between fitting both models is very small, which we can compare by QQ plot of residuals for both models in figure 3.3. Furthermore, in figure 3.4., the point plots of residuals are slightly different, but it does not suggest that the latter model is not a good fit and r-squared values for both the models are almost the same.

Based on the above observations, we can conclude that the time effect is less or almost negligible compared to GDP per capita.

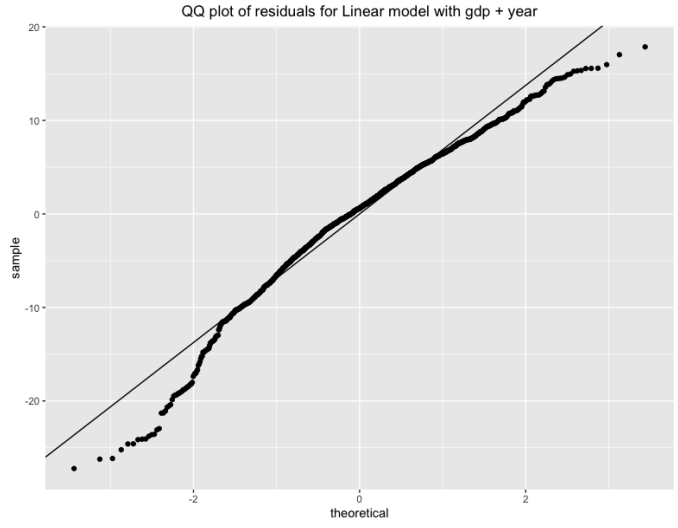

Fig 3.3.- QQ plot of residuals gdp+year

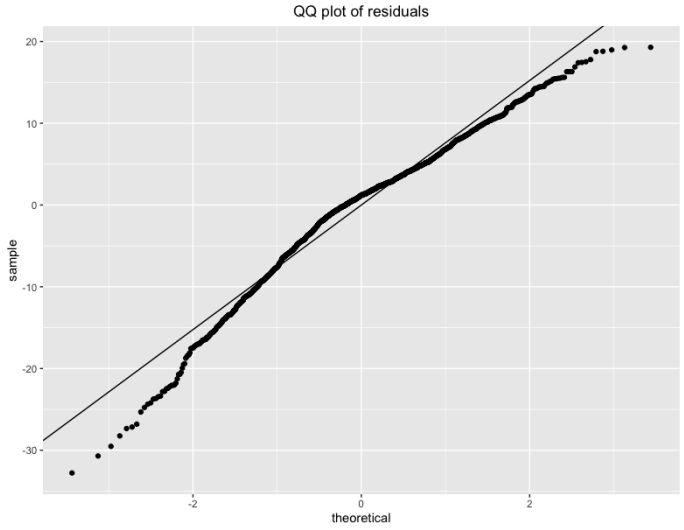

Fig 3.4.- QQ plot of residuals linear model

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?

There is convergence in GDP as there is less spread in life expectancy for the log of GDP with years. As we can see in figure 3.5, the spread is a constraint to a limited area in the region with the majority of light blue color points, and referring to the legend of the figure, we know light color signifies the latest years. So with time, GDP is converging.

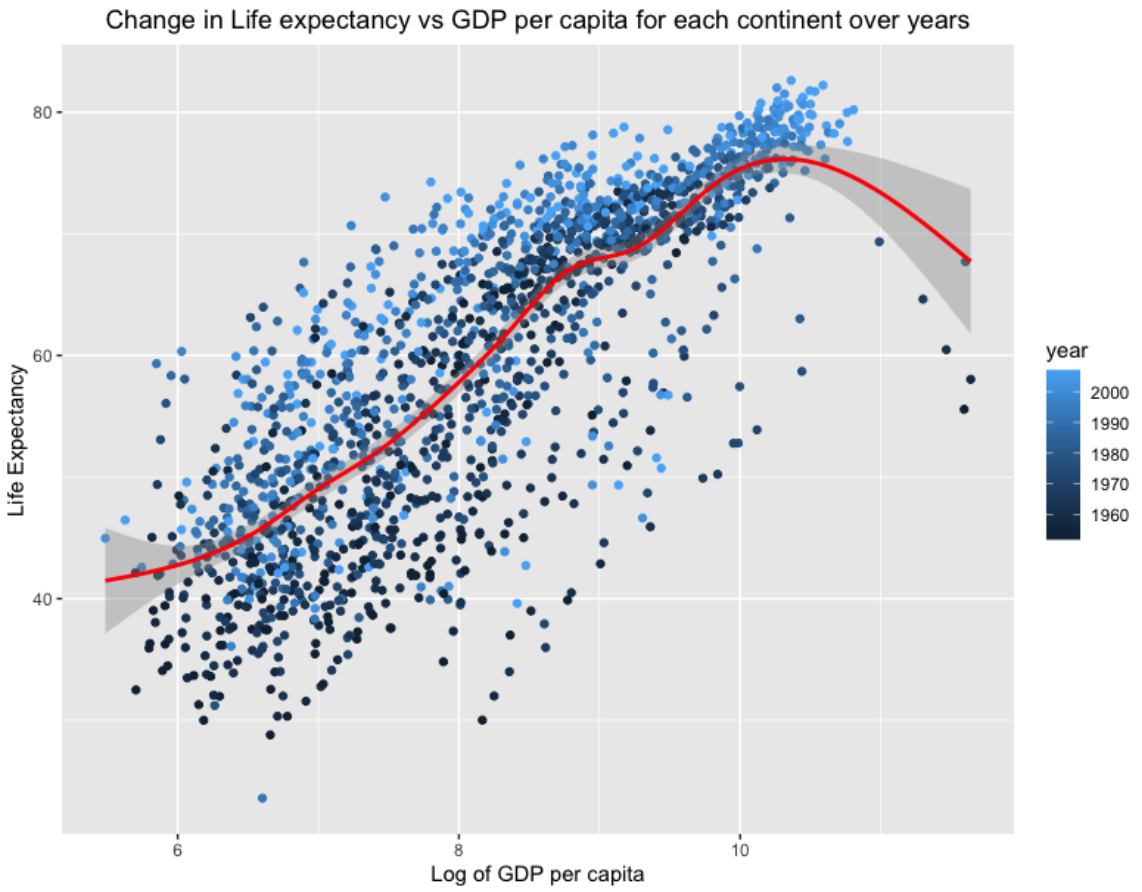

Fig 3.5. - Life Expectancy VS Log of GDP per capita (color according to the year)

All the continents other than Africa follow the same pattern of convergence in GDP per capita. We can refer to figure 3.1 and see Asia, Americas, and Europe follow the general pattern, it is difficult to conclude about Oceania since we have very little data regarding it.

For Africa, from figure 2.1 and figure 3.1, we can say that convergence of GDP over the years is not that much and GDP still matters much for the analysis of life expectancy there. For all other continents we can see the spread of life expectancy for GDP over years is converging or we can say for recent years it is very less but this is not the case for Africa and hence, it is an exception from the general pattern.

APPENDIX

Figure names are according to the format:

Fig A.<Question which it is related to><Figure number>

Question 1

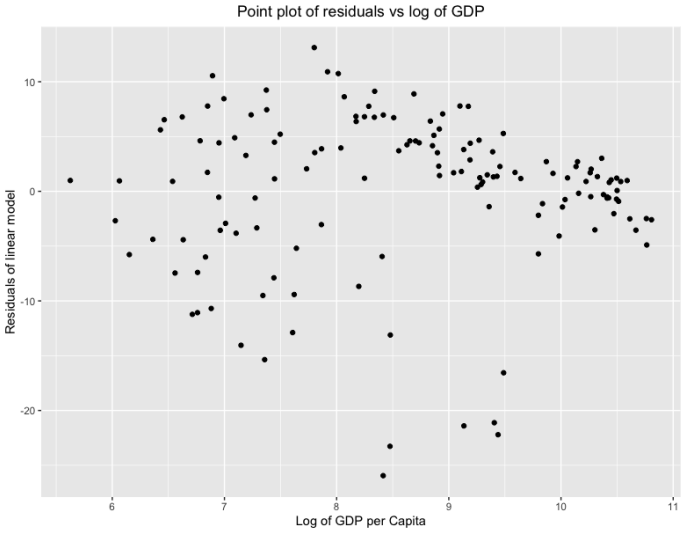

Fig A.1.1.- Homoscedasticity of residuals in 2007

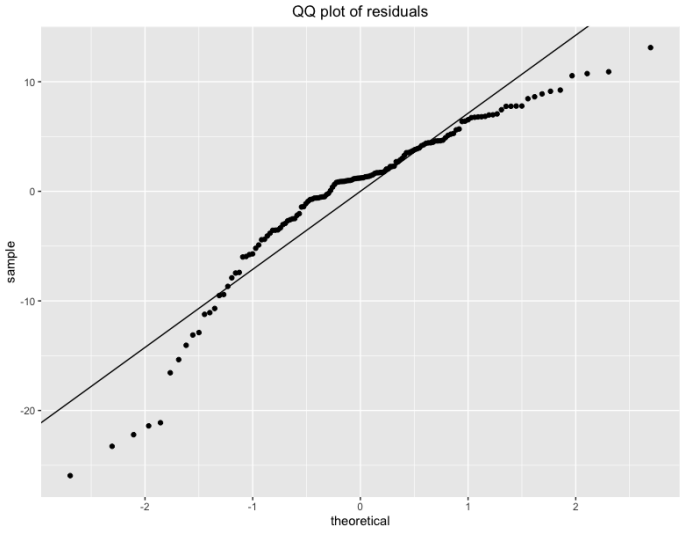

Fig A.1.2.- QQ plot of Residual in 2007

Question 2

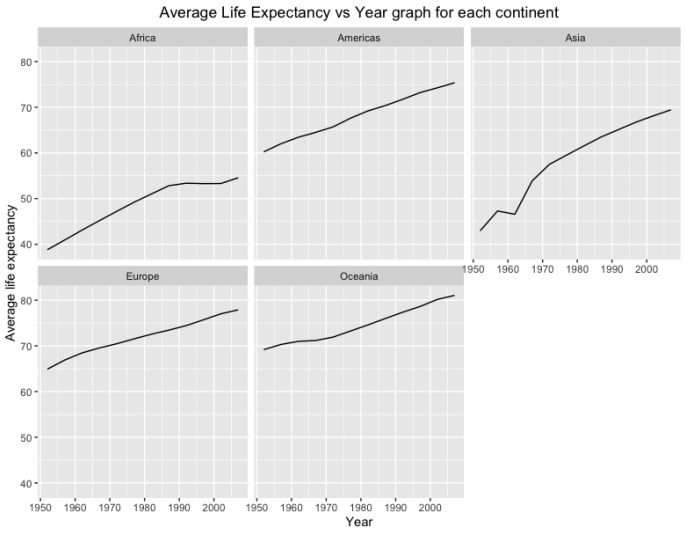

Fig A.2.1.- Average Life Expectancy VS Year for each continent

Question 3

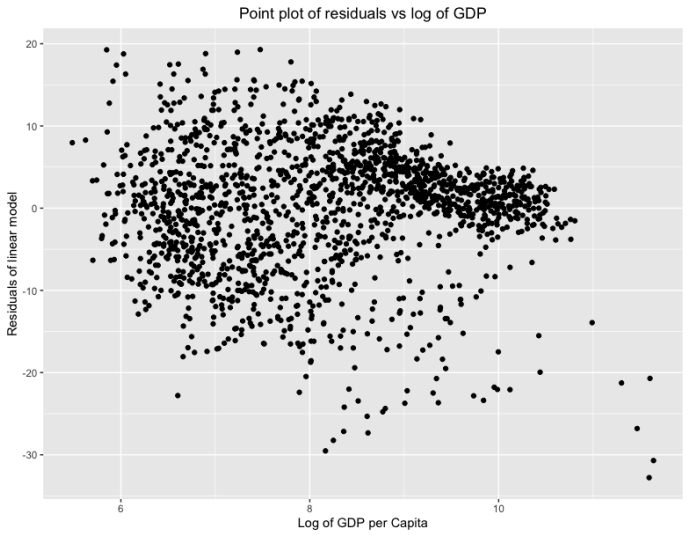

Fig A.3.1.- Residual plots linear model

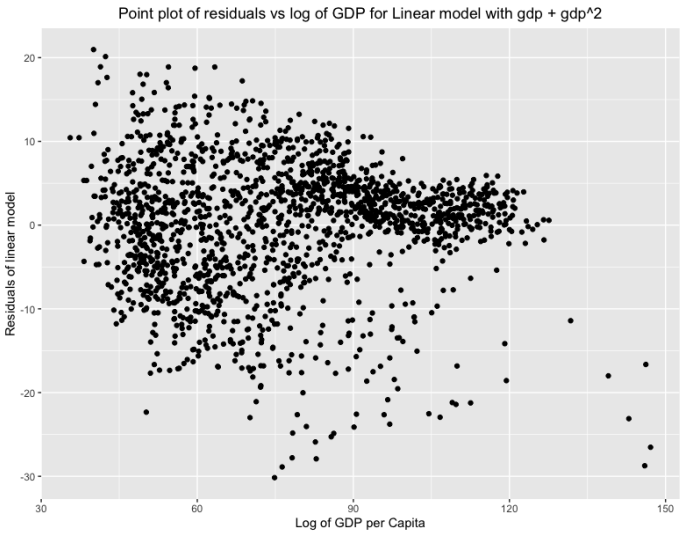

Fig A.3.2.- Residual plots quadratic model

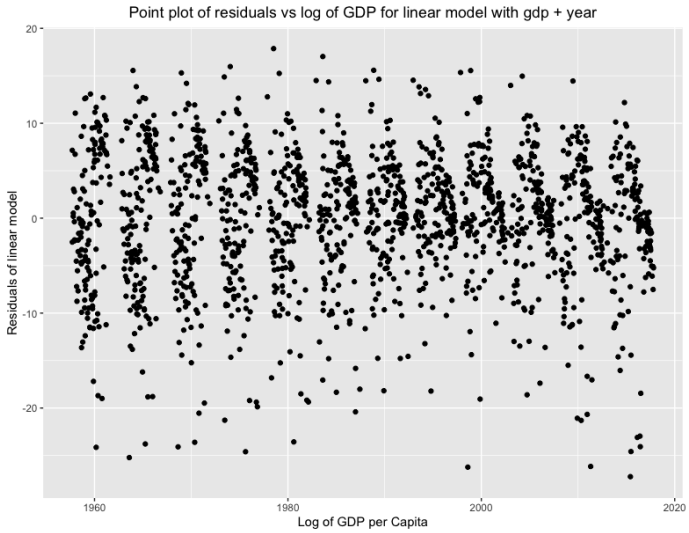

Fig A.3.3.- Residual plots VS log GDP per capita

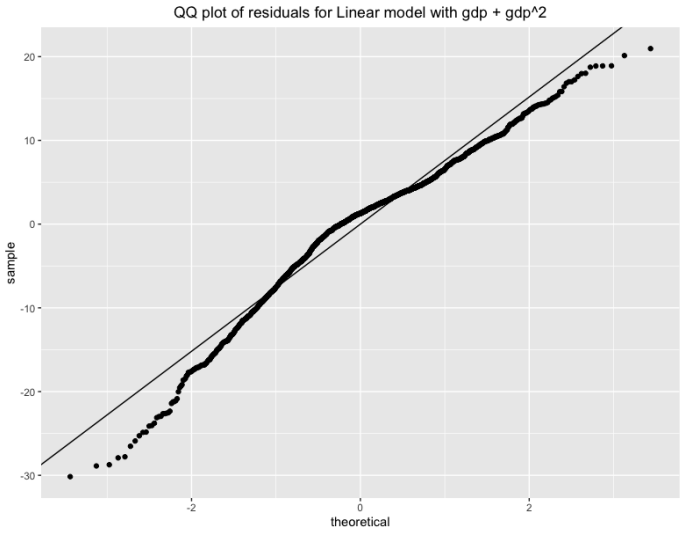

Fig A.3.4.- QQ plot of residuals quadratic model