STAT 670-Exploratory Data Analysis

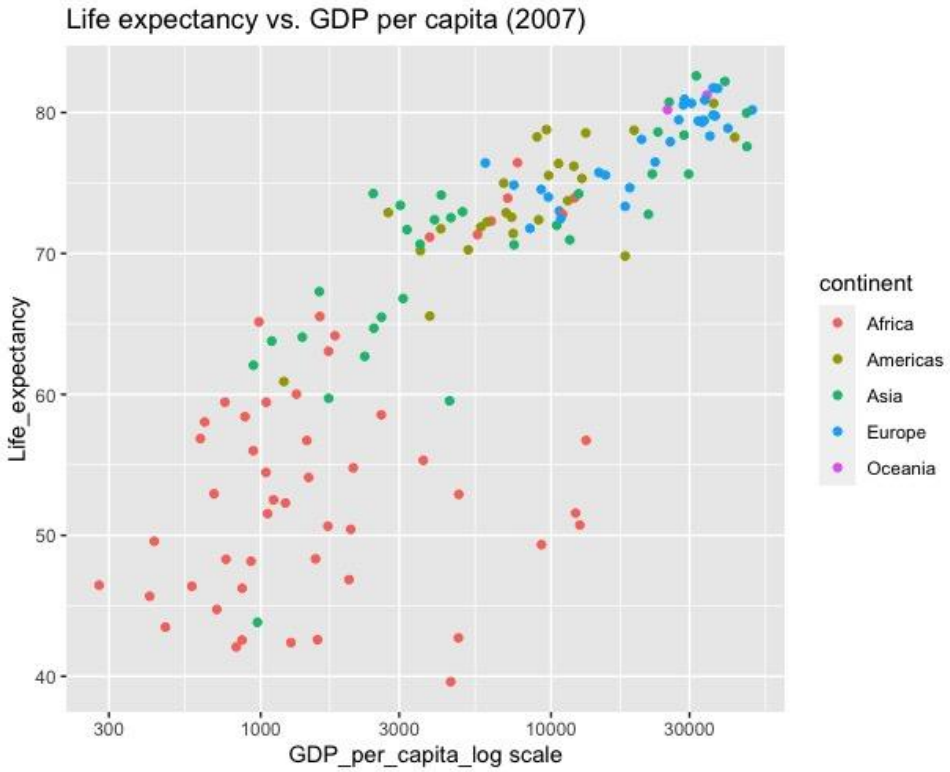
Mini Project: Life Expectancy

-Pavan Anvesh

-Netaji Sai Pavan

-Kaushik Parvathaneni

Question 1:



How does life expectancy vary with GDP per capita in 2007?

From the above plot we can see that there is a positive relationship between GDP per capita and life expectancy in 2007 and countries with higher GDP tend to have higher life expectancy.

Can the trends be well-described by a simple model such as a linear model, or is a more complicated model required?

To get to know more about the trends in particular , we tried fitting linear model and got the r squared value of 0.65 which explains that a linear model is not a good fit as it is yet to explain large portion of trend.So a complicated model is required.

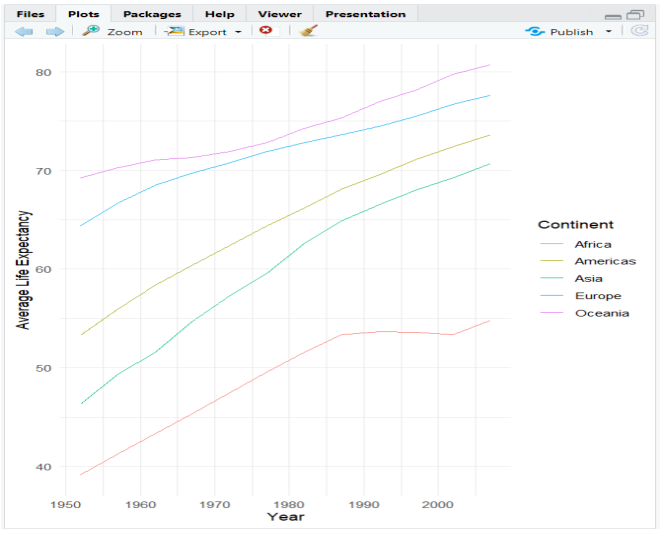
Is the pattern the same or different for every continent? If some continents are different, which ones?

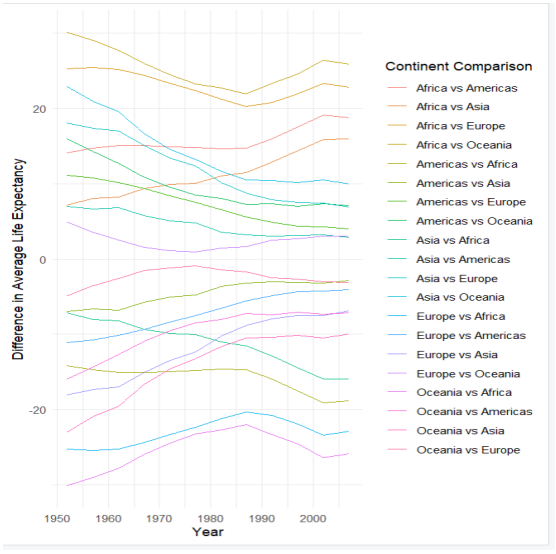
The relationship between GDP per capita and life expectancy is similar across all continents except Africa.Higher GDP is associated with higher life expectancy across all continents except Africa.

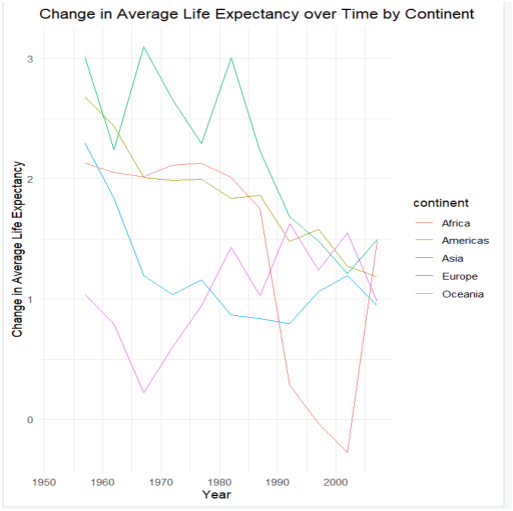
Can differences between continents be simply described by an additive or multiplicative shift, or is it more complicated than that?

To explain the differences between continents by additive or multiplicative shift,by fitting an linear regression and adding an interaction term,we can see that GDP and life expectancy across continents varies in a non additive way. GDP effect in life expectancy is strongest for Europe and weakest for Africa.

Question 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, some continents have partially caught up to others in terms of life expectancy over time.

This catch-up is not just because of some countries in the continent, but is more general. We can confirm this by looking at the data at the country level. For example, we can calculate the mean life expectancy for each country and year and then plot the average life expectancy for the top 10 countries with the highest life expectancy in each continent.

From this graph, we can see that while the top-performing countries in each continent have generally maintained their high life expectancy over time, the overall improvement in life expectancy in each continent is driven by improvements in the majority of the countries, rather than just a few high-performing ones.

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

The changes in life expectancy over time have not been linear, and the rate of change has varied across continents and time periods.

What might explain periods of faster/slower change?

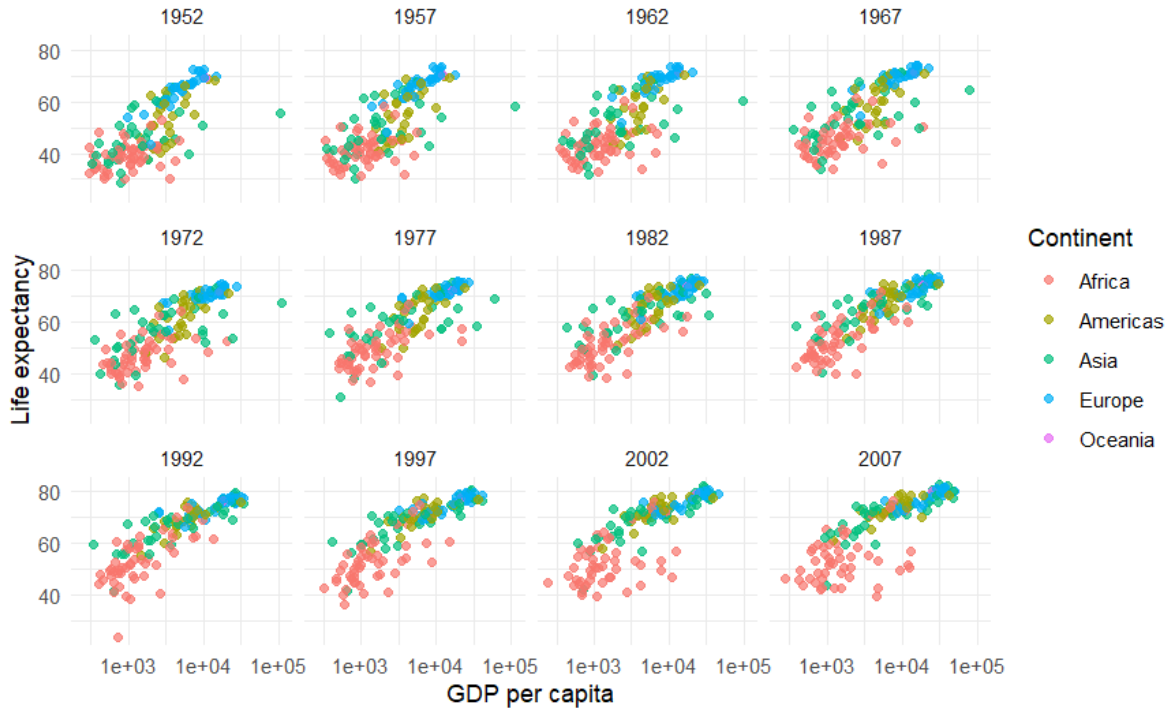
There are many factors that can explain periods of faster or slower change in life expectancy over time, including changes in healthcare, education, economic growth, and public health policies.

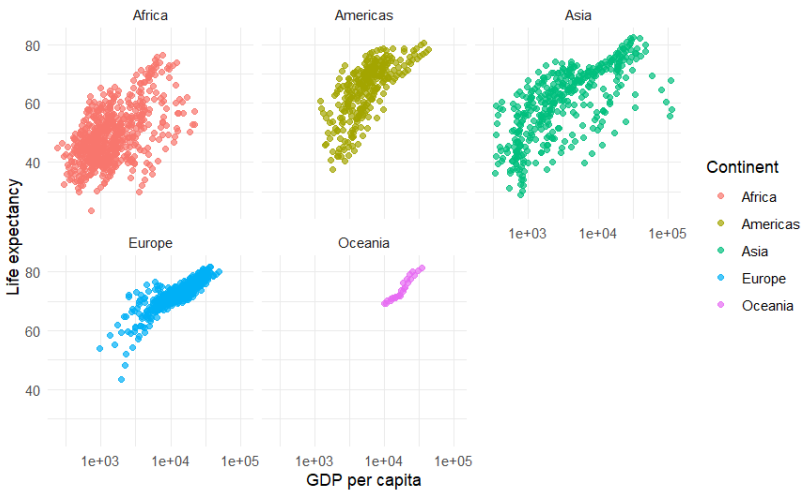
For example, improvements in healthcare can lead to better disease prevention and treatment, which can improve life expectancy. Similarly, increased access to education can lead to better health knowledge and behaviors, which can also improve life expectancy. Economic growth can also lead to improvements in health, as countries with higher GDP per capita may have more resources to invest in healthcare and public health programs.

Public health policies, such as vaccination campaigns, disease surveillance, and clean water and sanitation programs, can also have a significant impact on life expectancy.

Overall, the rate of change in life expectancy over time is influenced by a complex interplay of social, economic, and environmental factors, and understanding these factors can help to guide public health policies and interventions aimed at improving population health.

Question 3:

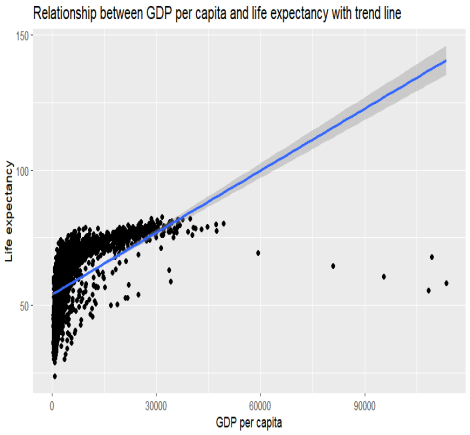


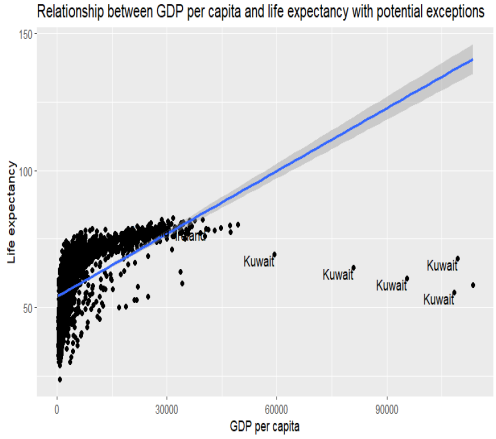



```
gdpPercap   7.649e-04  2.579e-05   29.66   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 10.49 on 1702 degrees of freedom
Multiple R-squared:  0.3407,    Adjusted R-squared:  0.3403
F-statistic: 879.6 on 1 and 1702 DF,  p-value: < 2.2e-16

[1] 0.5837062
gapminder$continent: Africa
[1] 0.4256076
------------------------------------------------------------------------------------------
----
gapminder$continent: Americas
[1] 0.5583655
------------------------------------------------------------------------------------------
----
gapminder$continent: Asia
[1] 0.3820476
------------------------------------------------------------------------------------------
----
gapminder$continent: Europe
[1] 0.7807831
------------------------------------------------------------------------------------------
----
gapminder$continent: Oceania
[1] 0.9564738
```





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

We can see that for some continents, such as Asia and Europe, there has been a clear positive relationship between the two variables throughout the years, while for others, such as Africa and Oceania, the relationship has been more variable. We can also see that for some continents, such as Europe, the relationship has become stronger over time, while for others, such as Africa and Oceania, it has become weaker. This suggests that there may be other factors beyond GDP per capita that are driving changes in life expectancy over time.

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

To assess whether changes in life expectancy can be entirely explained by changes in GDP per capita, we can fit a linear regression model to the data and examine the coefficient of determination (R-squared) of the model. The R-squared value indicates the proportion of the variation in life expectancy that can be explained by variation in GDP per capita. If the R-squared value is high, it suggests that changes in GDP per capita can explain a large portion of the changes in life expectancy, while a low R-squared value suggests that other factors may be more important.

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

There may be other factors beyond GDP per capita and time that are also important like Medical facilities, Increase in income, accessible to good food and better living conditions

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?

Yes there is Convergence and however, there are still some exceptions to the general patterns. For example, some African countries still have relatively low life expectancies even in the more recent years, despite having relatively high GDP per capita compared to other African countries. Additionally, there are some European countries that have relatively low life expectancies despite having relatively high GDP per capita. Therefore, while there is evidence of convergence over time, there are still exceptions to the general patterns, and other factors beyond GDP per capita are likely to be important in explaining differences in life expectancy between countries.