**EDA COMPREHENSIVE PROJECT**

**Problem Statement:**

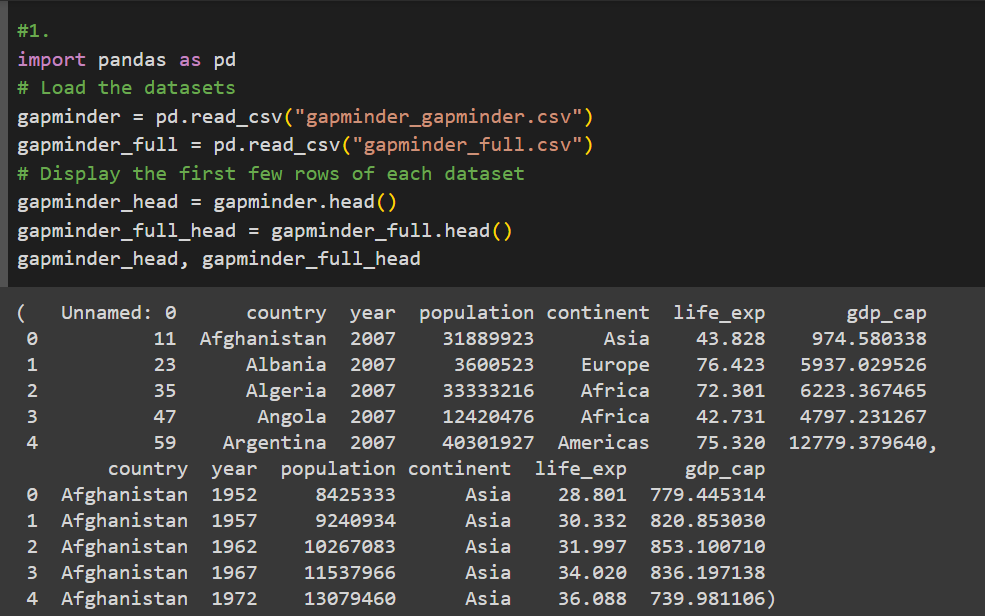
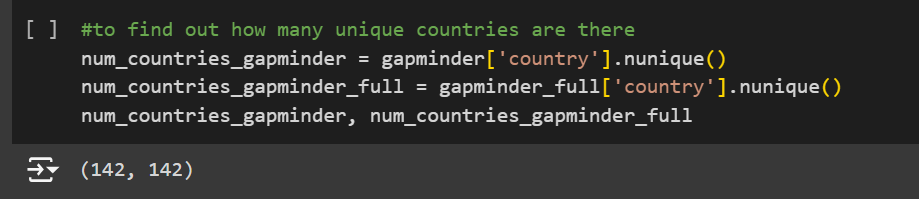
The modern world is shaped by complex dynamics in population, health, and economics, making understanding these trends vital for informed policy-making. Global Trends, a leading analytics firm, is dedicated to deciphering these patterns through a comprehensive analysis of the Gapminder dataset. Your role in this project is to conduct an in-depth Exploratory Data Analysis (EDA), uncovering the intricate relationships between demographic changes, economic development, and health advancements over recent decades.

**Data Overview:**

## The dataset used in this analysis is derived from the Gapminder data, which includes various socio-economic and health indicators for multiple countries over several years. The data primarily focuses on life expectancy, GDP per capita, population, and other relevant metrics.

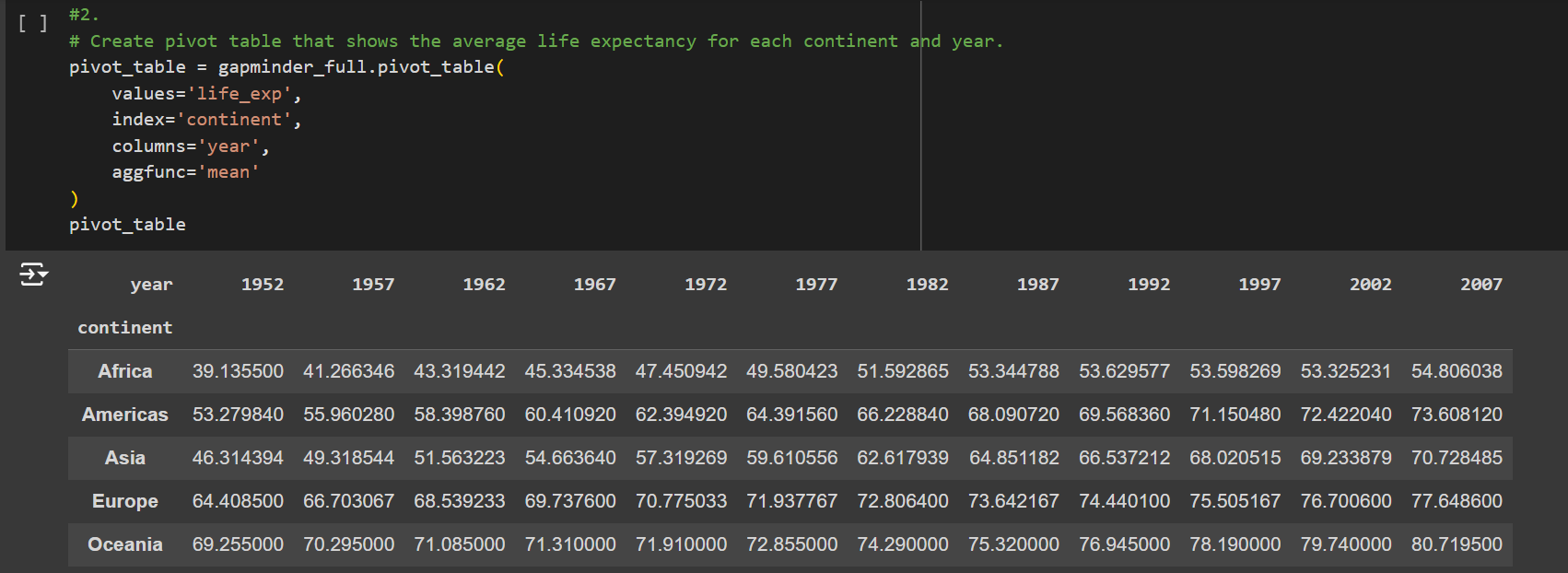
## Insights and Analysis:

### 1. Load the dataset and display the first few rows. How many countries does the dataset have?

** **

**Number of Countries:** 142

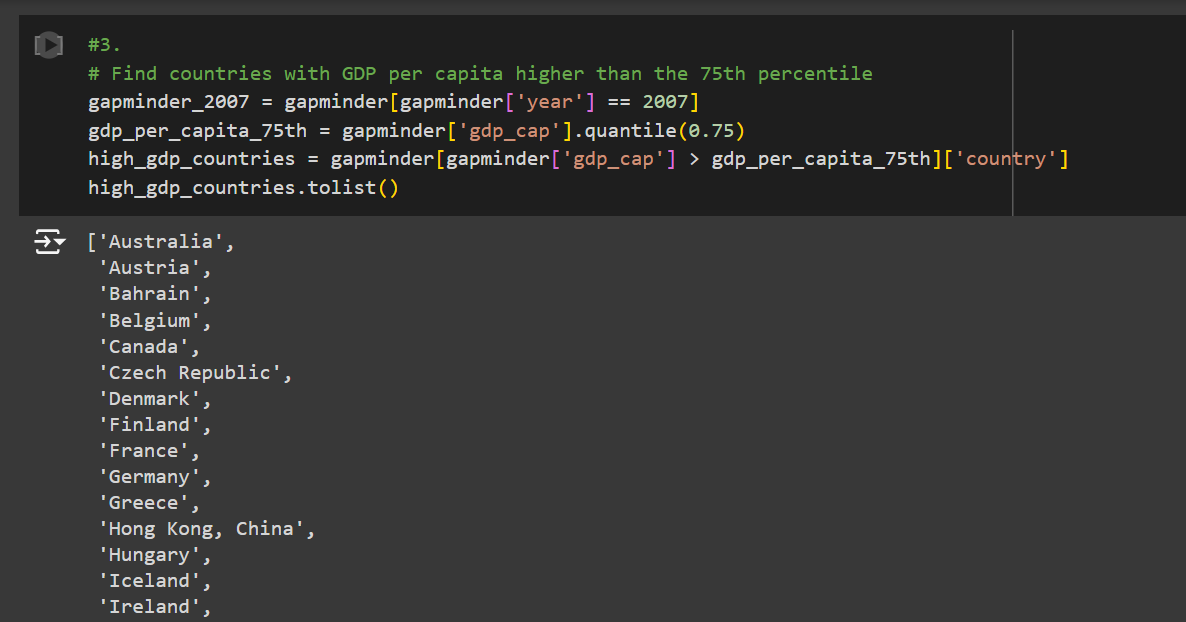
2. Create a pivot table showing the average life expectancy for each continent and year.



 Life expectancy has generally increased across all continents from 1952 to 2007.

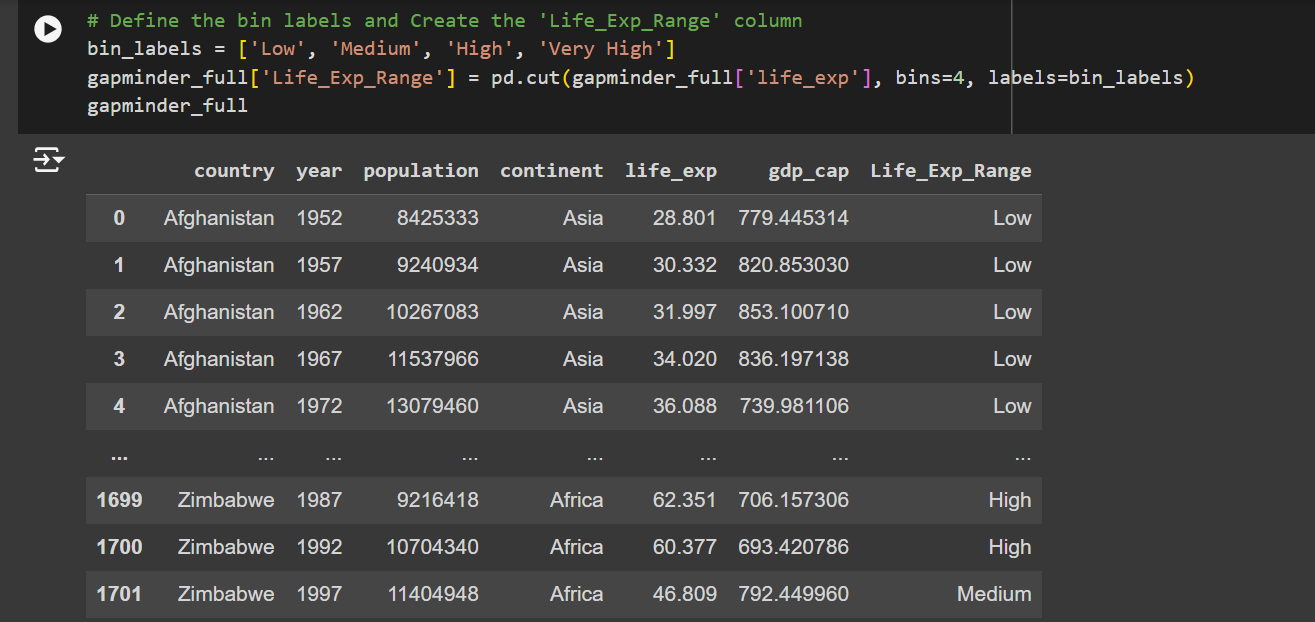
 Oceania consistently has the highest average life expectancy, while Africa has the lowest.

3. Which countries had a GDP per capita higher than the 75th percentile in 2007?



Most of the countries with high GDP in 2007 are from Europe, North America, and Oceania.

4. Categorize the 'life\_exp' into 4 equally ranged bins from 'Low' to 'Very High'.



The majority of life expectancies in the dataset fall into the "Low" and "Medium" categories, particularly in earlier years.

5. Identify the top 5 countries with the highest GDP per capita in 2007 using a horizontal bar chart.



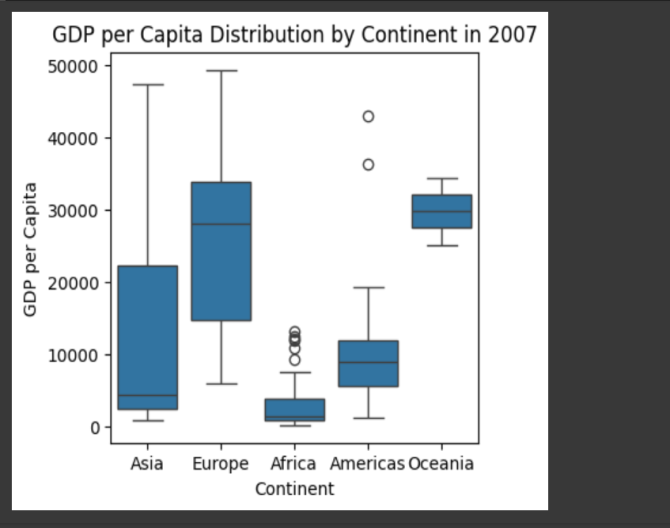
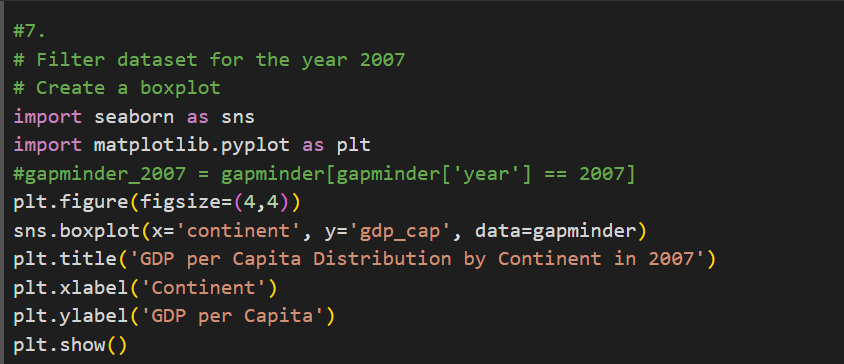
The top 5 countries by GDP per capita in 2007 are predominantly from Europe, with Kuwait being the exception from the Middle East.

6. Find all country names that start with "I" and end with "a" using regex.



Only two countries in the dataset match the regex pattern: India and Indonesia.

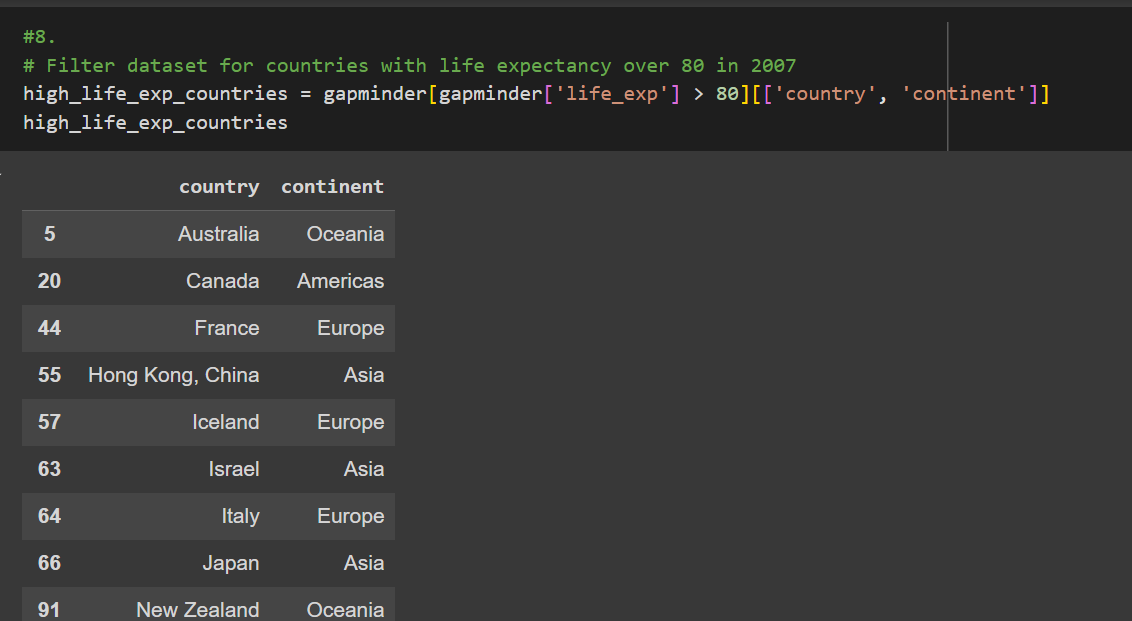
7. Create a boxplot using Seaborn to compare the distribution of GDP per capita for each continent in 2007.



 Oceania and Europe have the highest median GDP per capita.

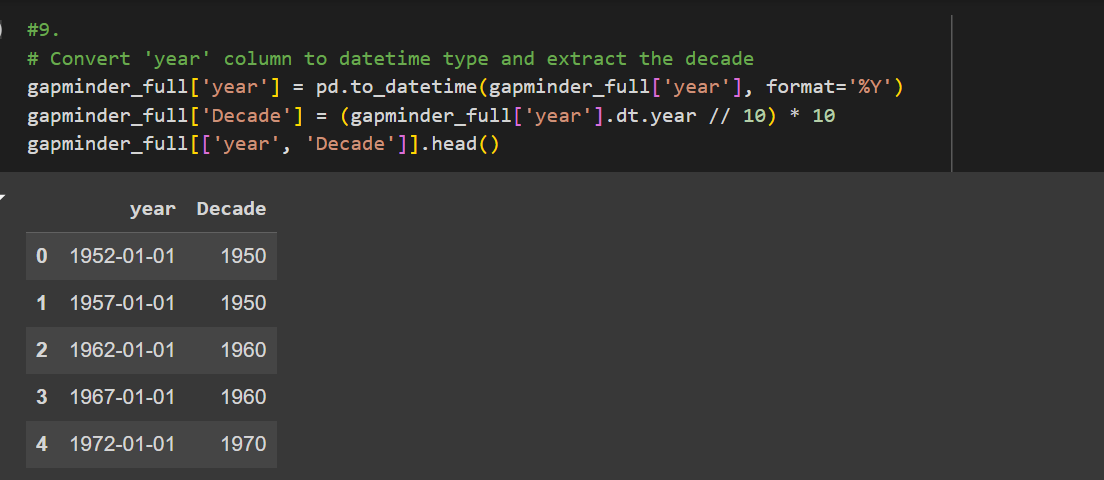
 Africa has the lowest median GDP per capita and the widest range, indicating significant economic disparity within the continent.

8. Find all countries with a life expectancy of over 80 years in 2007.



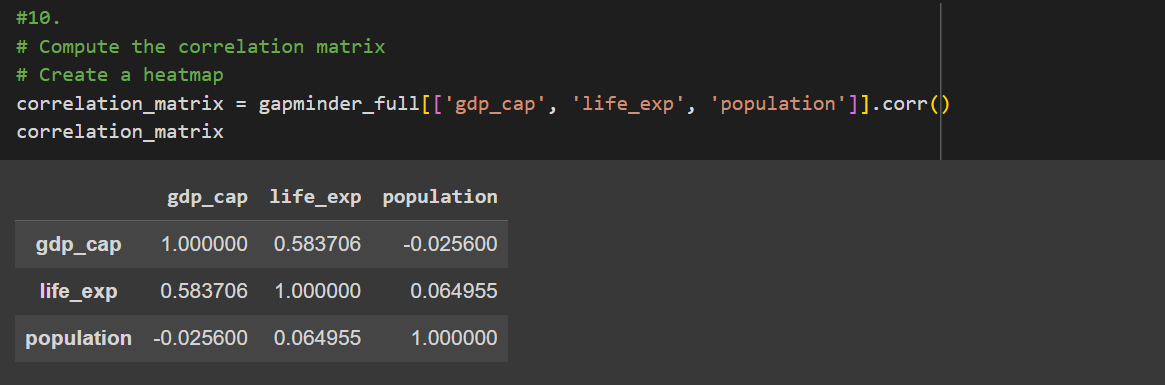
Countries with life expectancy over 80 years in 2007 are mostly from Europe, with a few from Oceania, the Americas, and Asia.

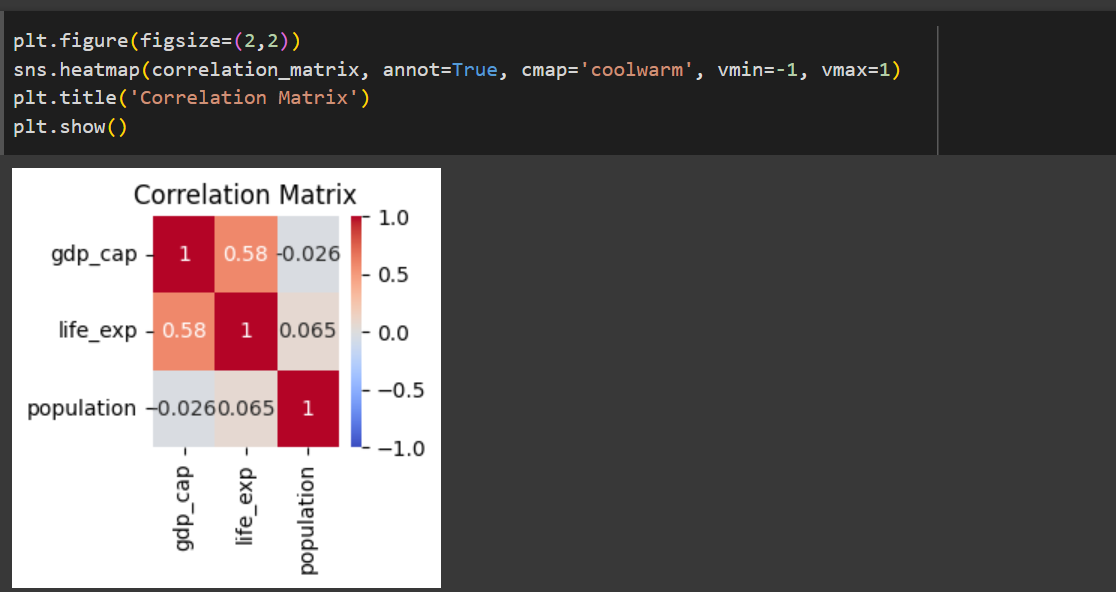
9. Convert the 'year' column to a datetime type and extract the decade.



The dataset now includes a 'Decade' column, grouping years into decades for easier temporal analysis.

10. Compute the correlation matrix between GDP per capita, life expectancy, and population.

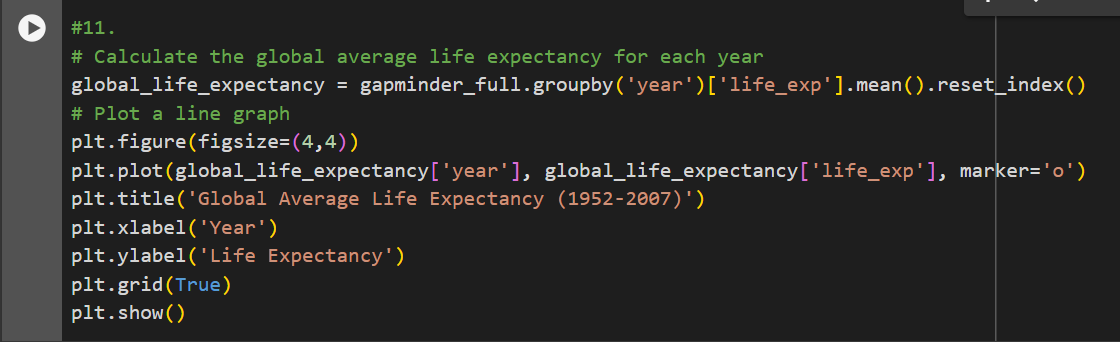


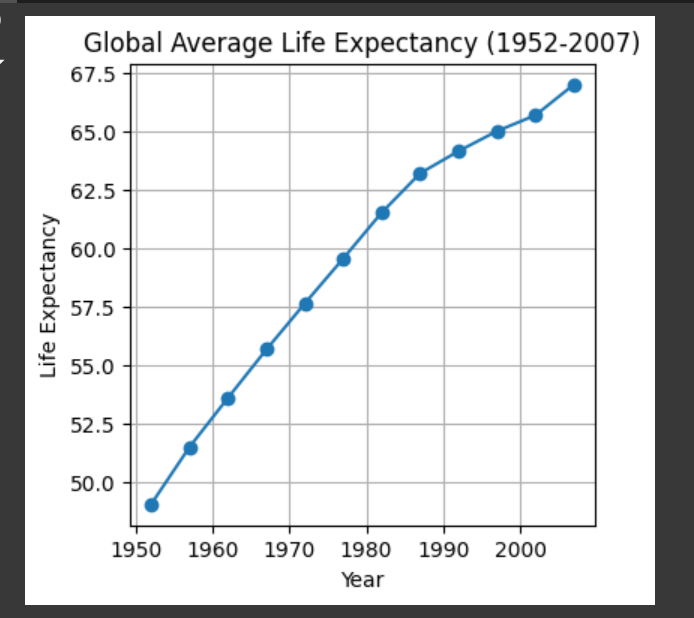


 There is a strong positive correlation between GDP per capita and life expectancy (0.60), indicating that higher economic prosperity is associated with longer life spans.

 Population has a weaker correlation with both GDP per capita and life expectancy.

11. How has the global average life expectancy changed from 1952 to 2007?





The global average life expectancy has increased from around 48 years in 1952 to about 70 years in 2007.

Subjective Question:

1. Discuss the various reasons that could have contributed to the change.

The improvement in global life expectancy can be attributed to:

 Advancements in medical science and technology.

 Improved public health infrastructure.

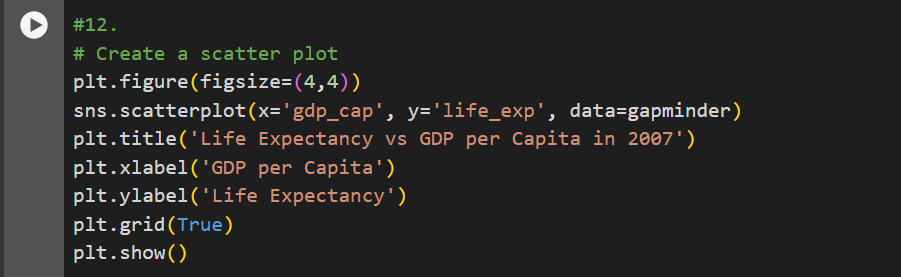
 Better nutrition and living conditions.

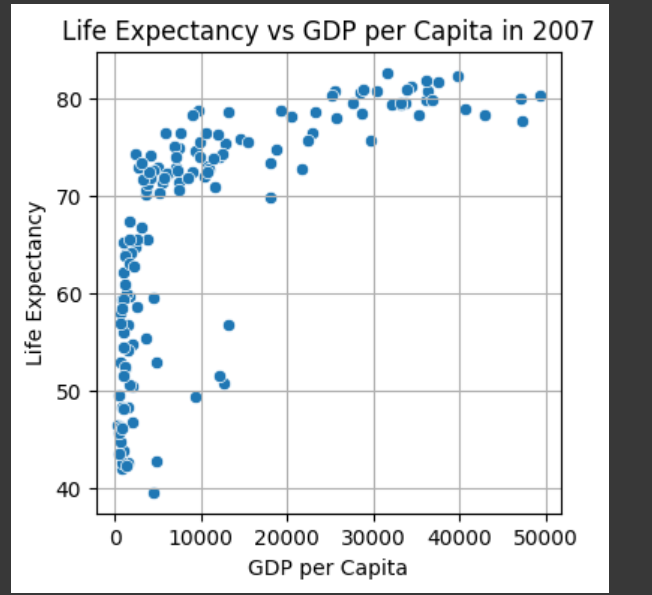
 Increased education and awareness.

 Economic development and poverty reduction.

 Political stability and global health initiatives.

12. Analyze the relationship between life expectancy and GDP per capita in 2007.





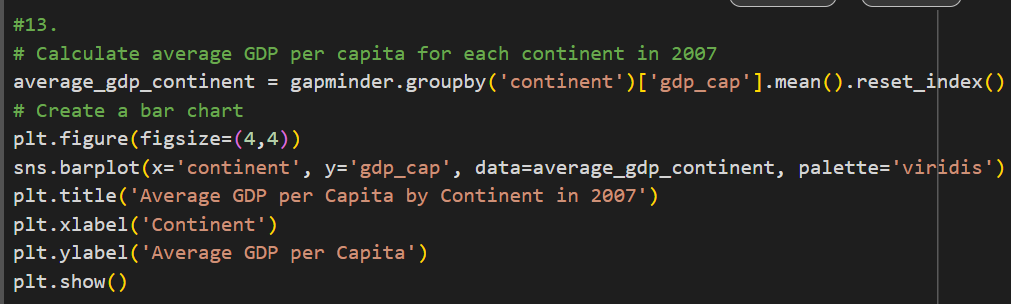
There is a noticeable positive correlation between GDP per capita and life expectancy. Higher GDP per capita is generally associated with higher life expectancy.

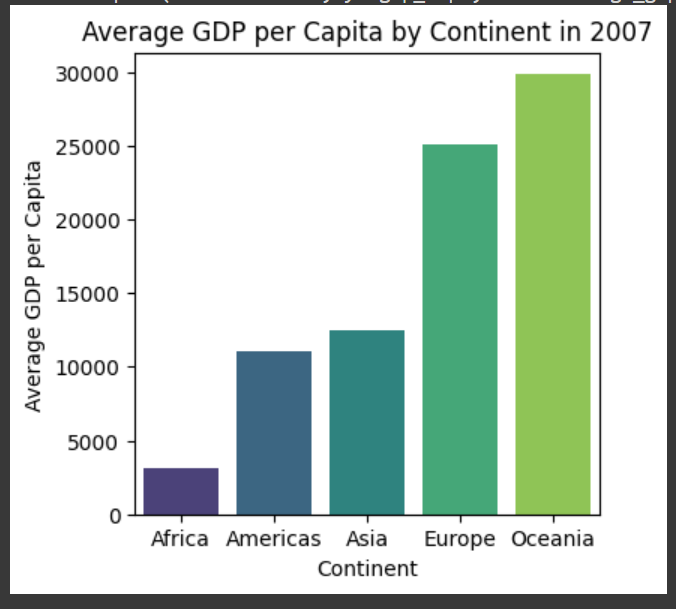
Subjective Question:

2. Is there a noticeable trend or correlation? Represent this using a scatter plot.

The trend indicates that wealthier countries tend to have better healthcare, education, and living conditions, which contribute to higher life expectancy. Economic prosperity enables countries to invest in public health and social services, improving the overall quality of life for their citizens.

13. Compare the average GDP per capita for each continent in 2007.





Oceania has the highest average GDP per capita, followed by Europe and the Americas.

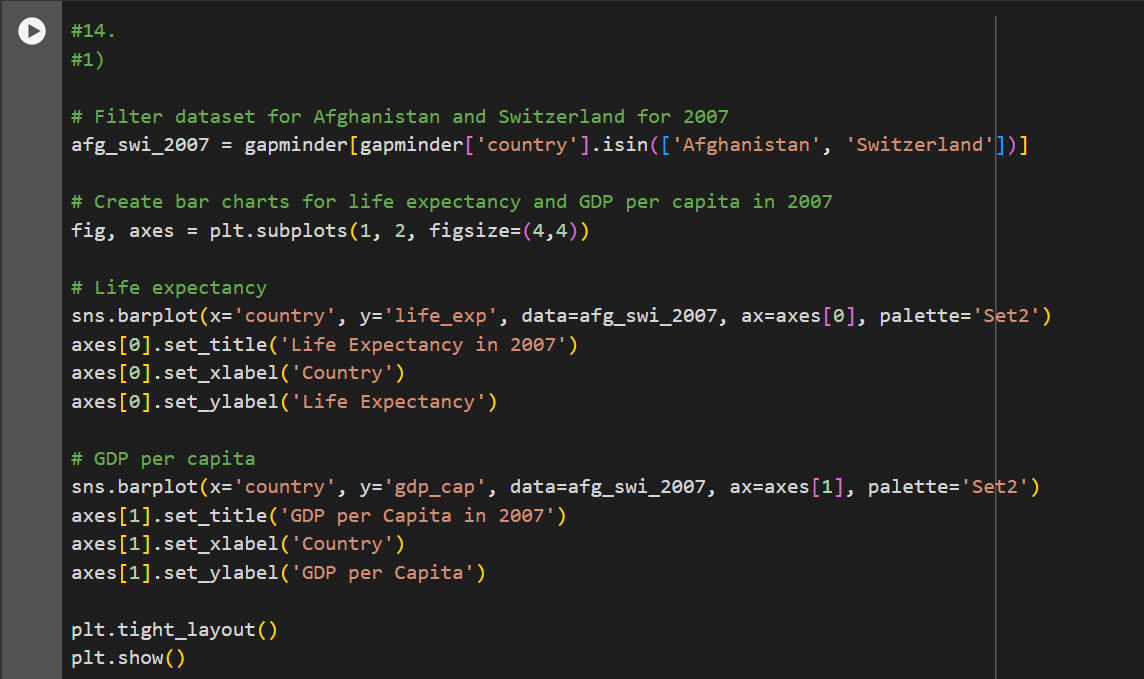
Subjective Question:

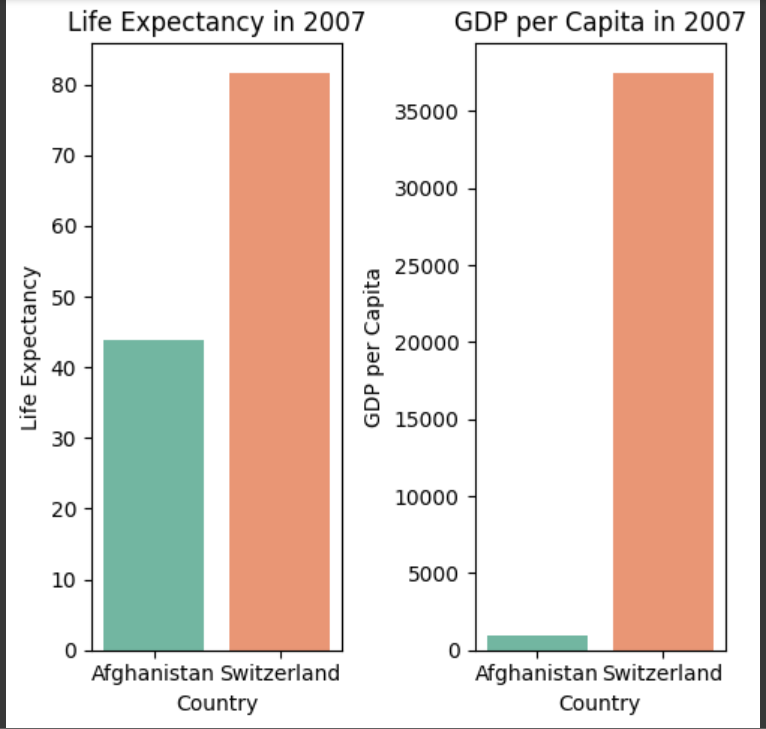
3. Why is the average GDP per capita for Oceania higher than the Americas even though the Americas have more countries?

Oceania's high average GDP per capita is driven by highly developed countries like Australia and New Zealand. The Americas, despite having more countries, have a mix of developed and developing nations, which lowers the overall average GDP per capita.

14. Compare the life expectancy and GDP per capita of Afghanistan and Switzerland.

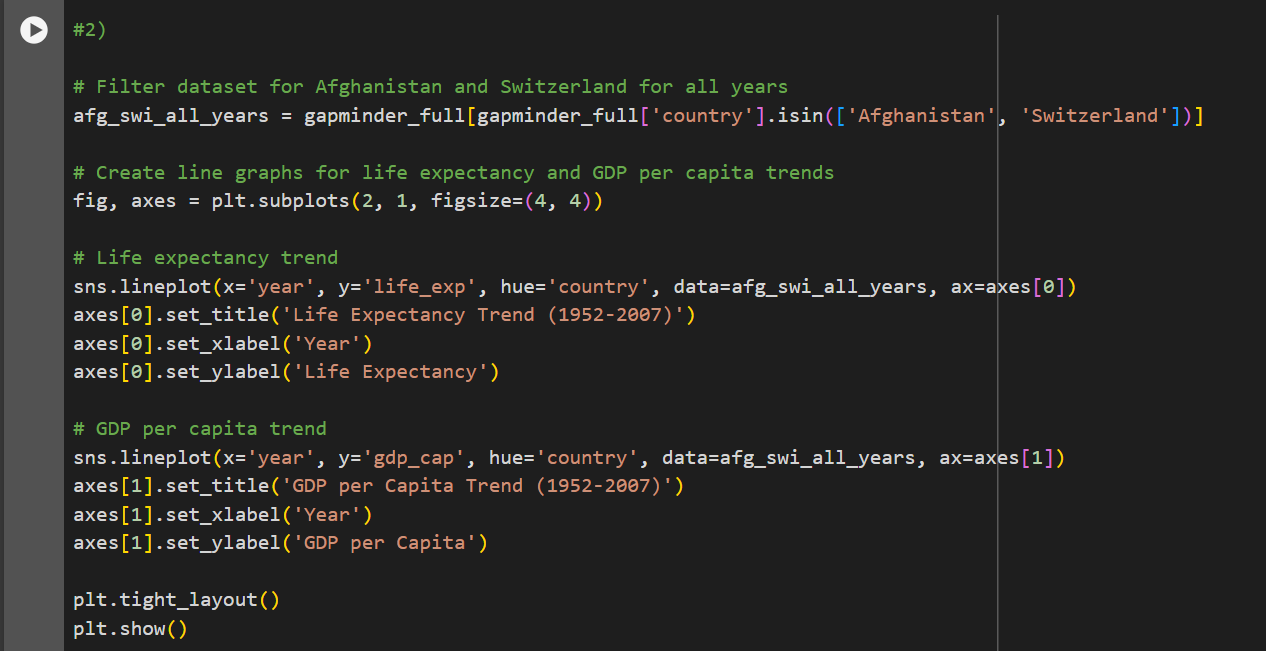
Bar chart for life expectancy and GDP per capita in 2007:

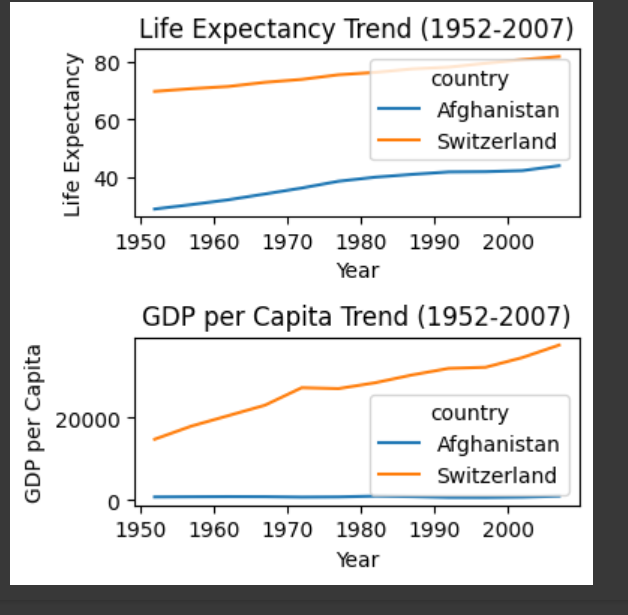




Switzerland has significantly higher life expectancy and GDP per capita compared to Afghanistan.

Line Graphs for Trends:





Switzerland's metrics have steadily increased over time, while Afghanistan shows much lower and more fluctuating values.

Subjective Question:

4. What differences do you observe in terms of life expectancy and economic development? How might the stability or instability of a country influence these key metrics over time? Analyze the data through these visualizations and discuss your inferences.

The differences between Afghanistan and Switzerland highlight the impact of stability on socio- economic indicators. Switzerland's consistent growth in life expectancy and GDP per capita is due to its political stability, robust economy, and advanced healthcare system. In contrast, Afghanistan's historical conflicts and instability have hindered its development, resulting in lower life expectancy and GDP per capita. Stability allows for sustained investment in healthcare, education, and economic development, leading to better overall outcomes.

**Summary and Important Insights:**

* The dataset reveals significant disparities in socio-economic indicators across different regions and countries.
* Life expectancy has generally improved globally, with notable advancements in healthcare and living standards.
* Wealthier countries exhibit higher life expectancy, underscoring the link between economic prosperity and health outcomes.
* Political stability and economic development play crucial roles in determining a country's health and economic indicators.
* The analysis highlights the need for targeted interventions in less developed regions to address disparities and improve quality of life.