**EDA COMPREHENSIVE PROJECT**

**Problem Statement:**

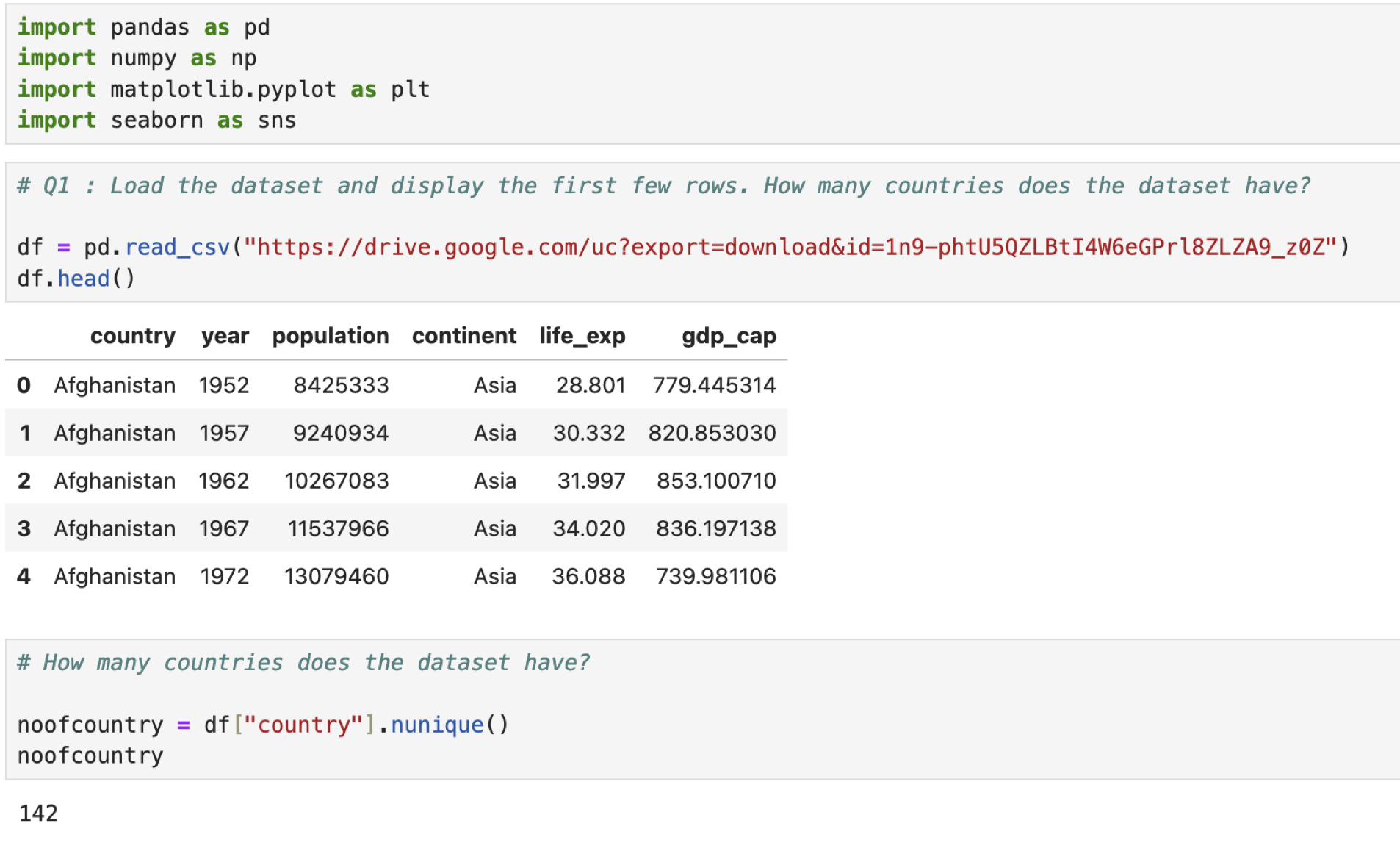
The modern world is shaped by complex dynamics in population, health, and economics, making understanding these trends vital for informed policy-making. GlobalTrends, a leading analytics firm, is dedicated to deciphering these patterns through a comprehensive analysis of the Gapminder dataset. Your role 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.

**NOTE:** For subjective questions, explain your answer in theory with clear examples and thoughtful analysis of the reasons behind the trends.

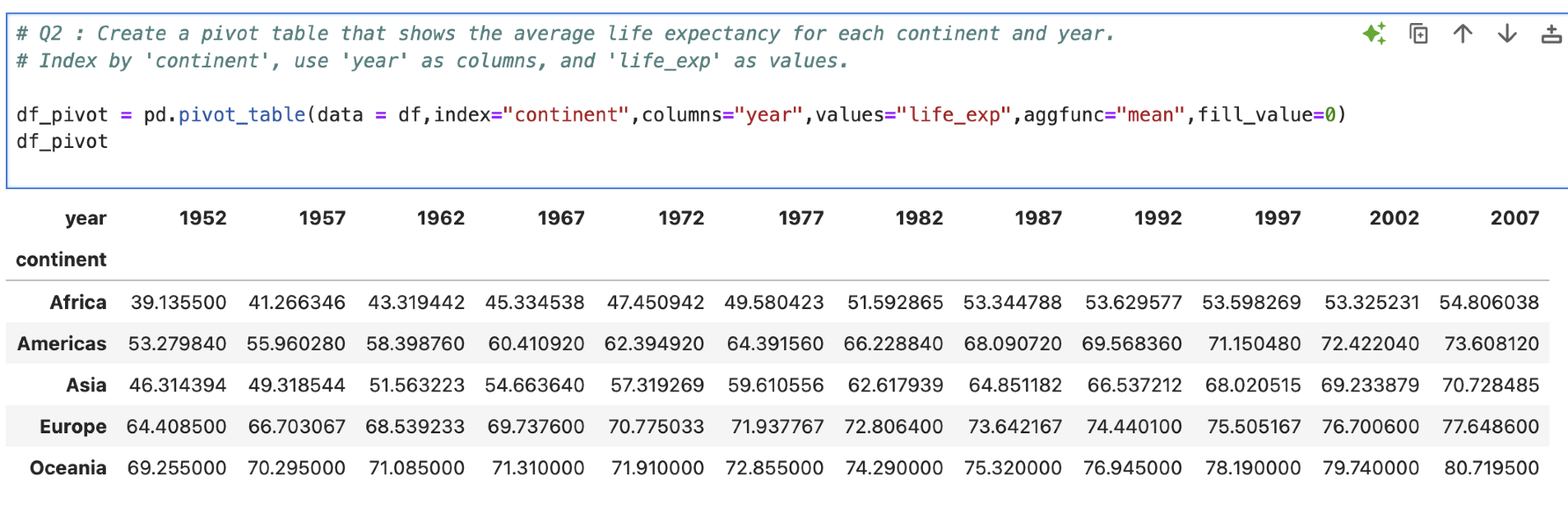
**Dataset:** [Gapminder World](https://drive.google.com/uc?id=1fDGZh86UPUkt2K6enlNQfB0mswU8pB_P)

**Objectives:**

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

Ans : 

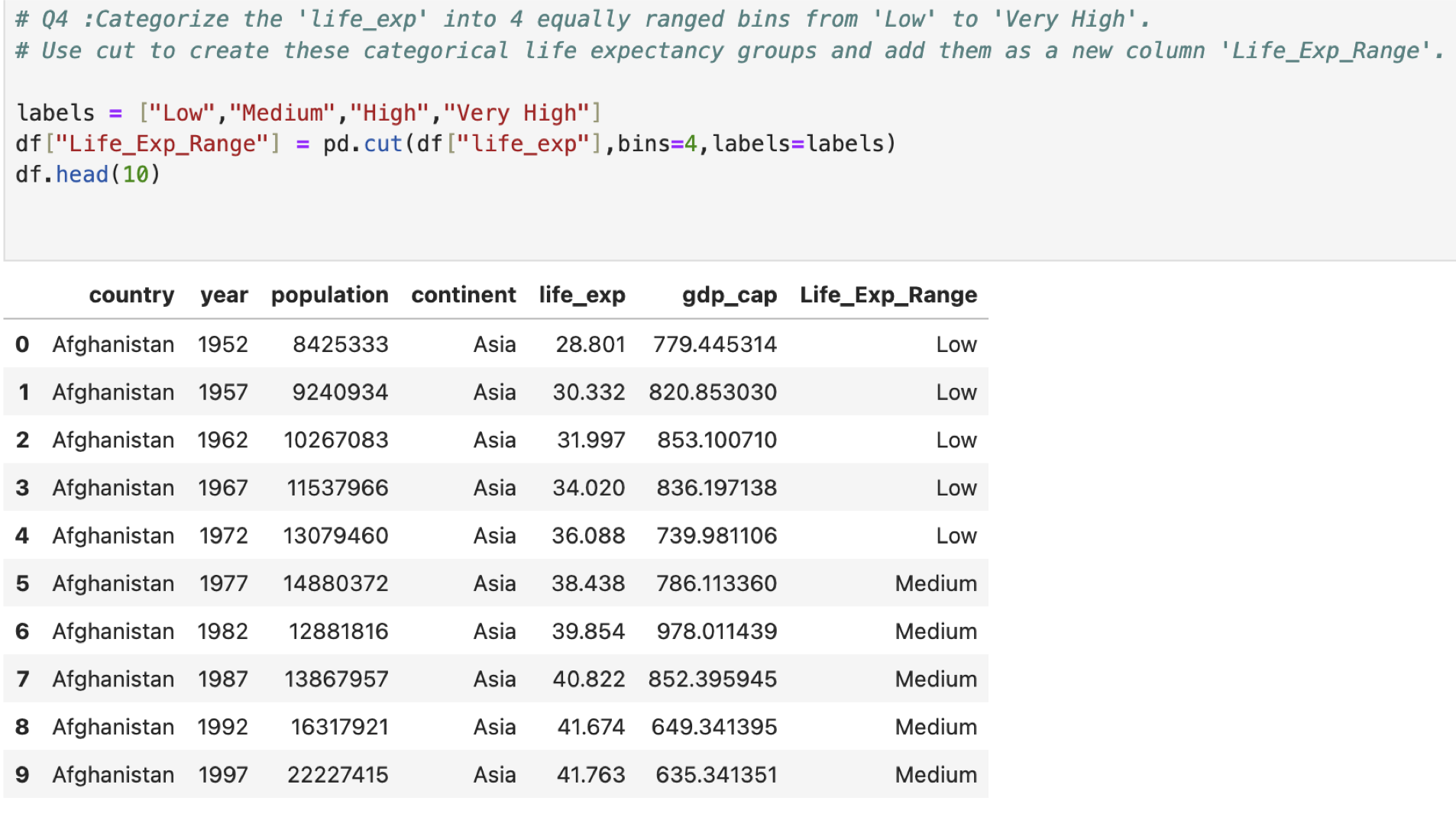
1. Create a pivot table that shows the average life expectancy for each continent and year. Index by 'continent', use 'year' as columns, and 'life\_exp' as values.



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

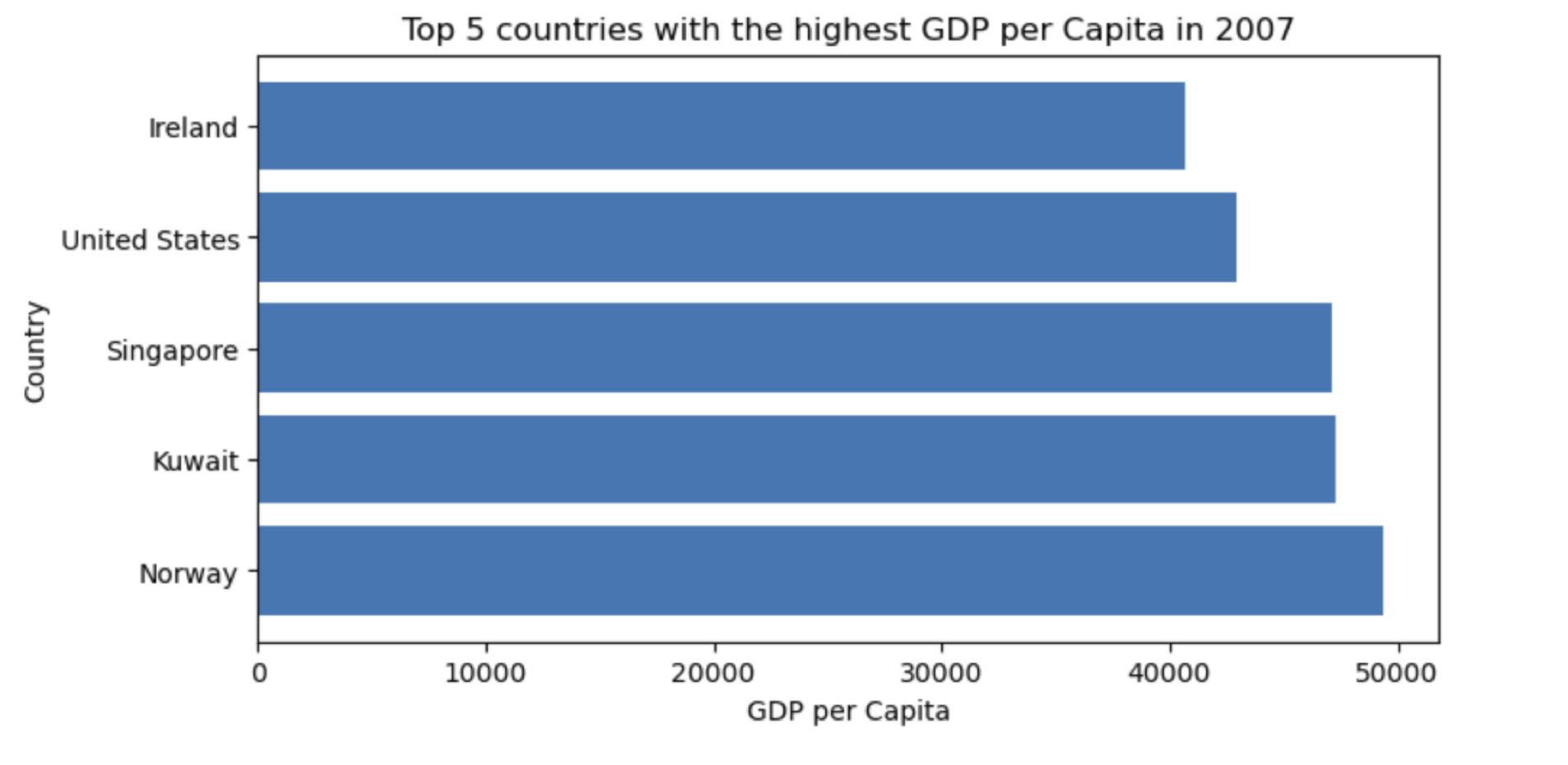


1. Categorize the 'life\_exp' into 4 equally ranged bins from 'Low' to 'Very High'. Use cut to create these categorical life expectancy groups and add them as a new column 'Life\_Exp\_Range'.

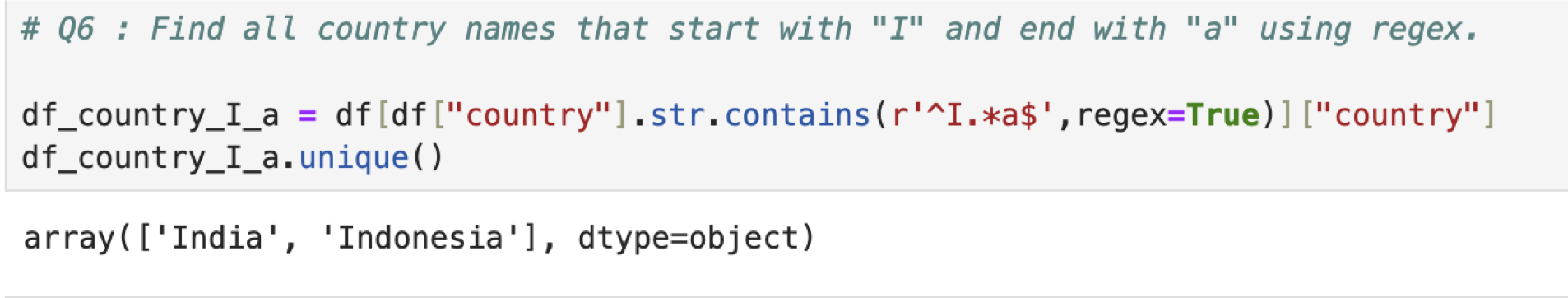


1. Identify the top 5 countries with the highest GDP per capita in 2007. Use a horizontal bar chart to display this data.

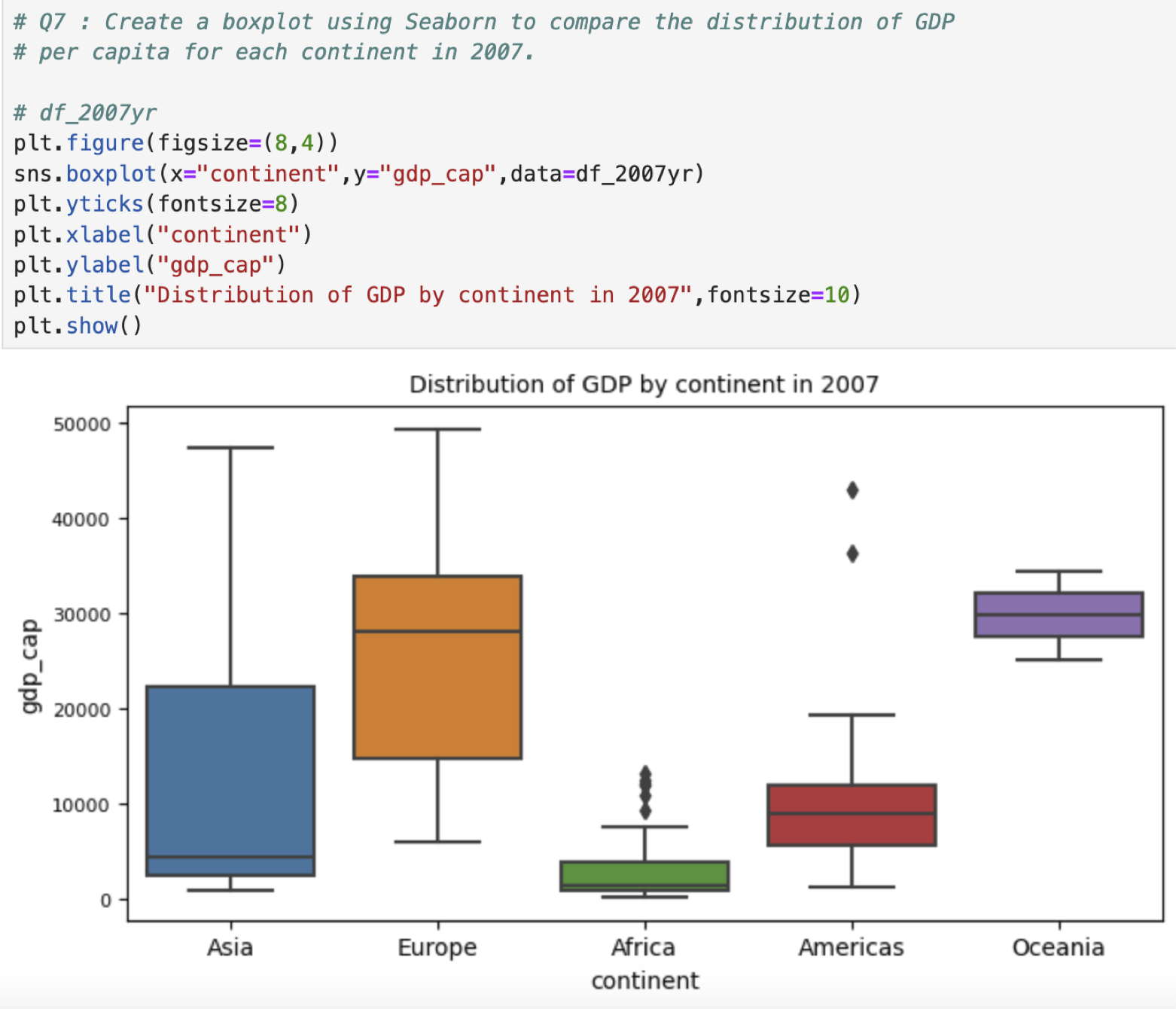




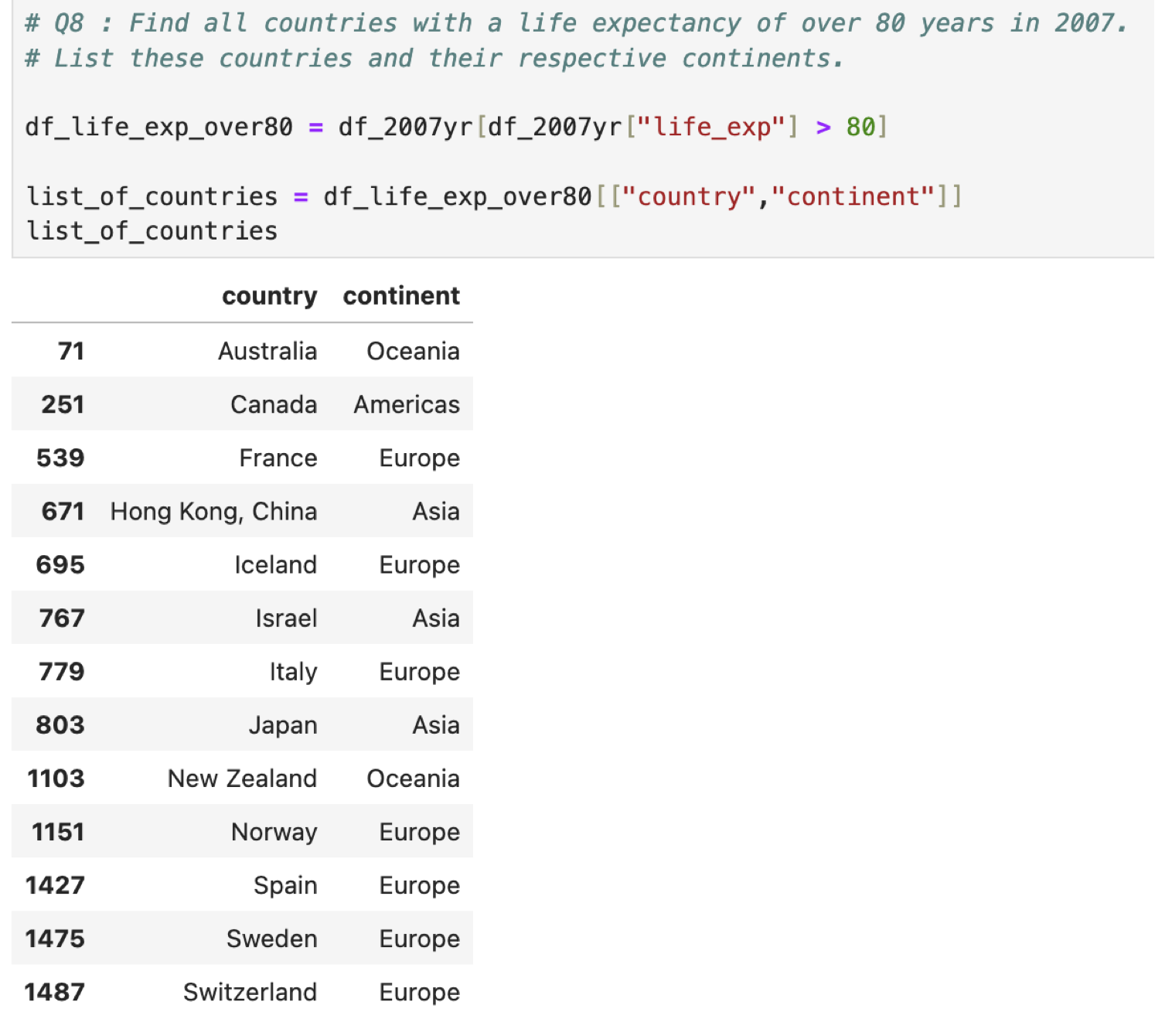
1. Find all country names that start with "I" and end with "a" using regex.



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



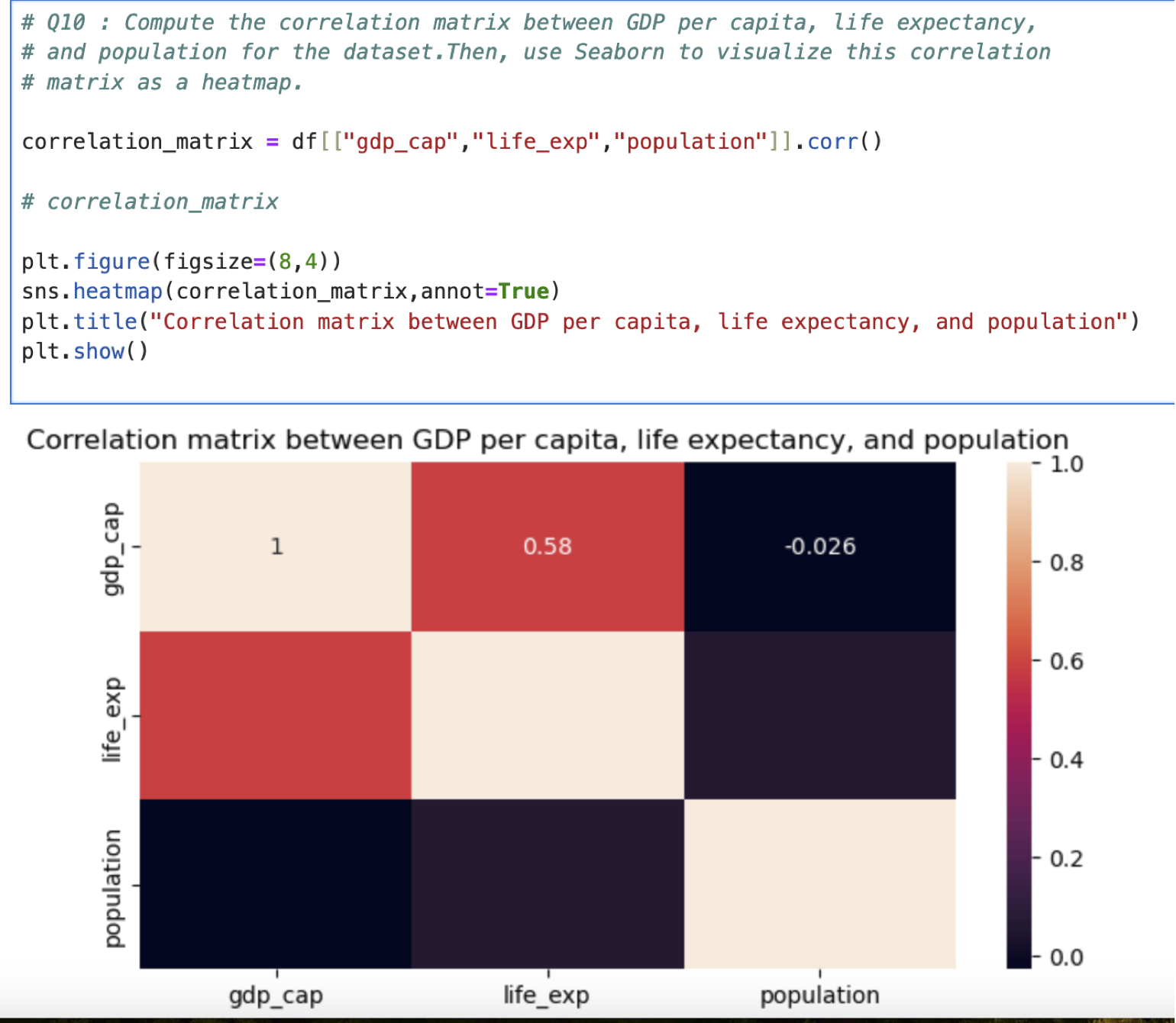
1. Find all countries with a life expectancy of over 80 years in 2007. List these countries and their respective continents.



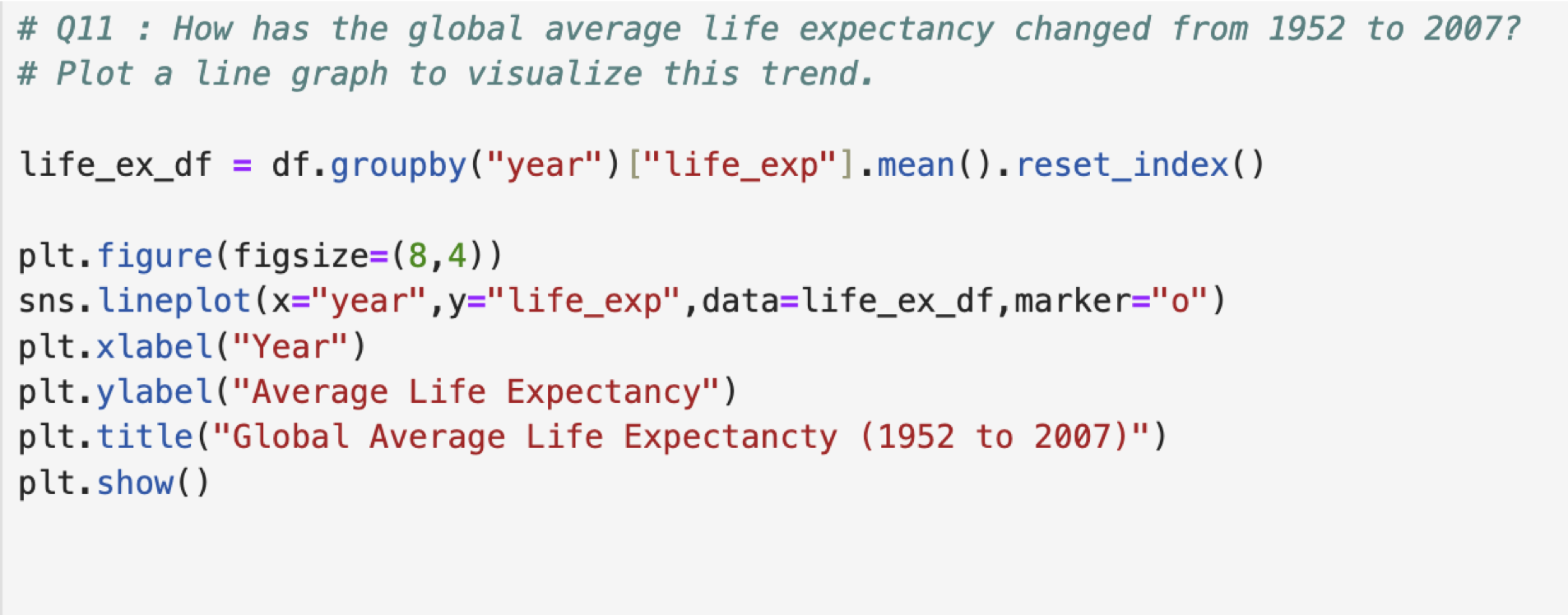
1. Convert the 'year' column to a datetime type and extract the decade. Create a new column 'Decade' that groups the years into decades (e.g., the 1950s, 1960s).

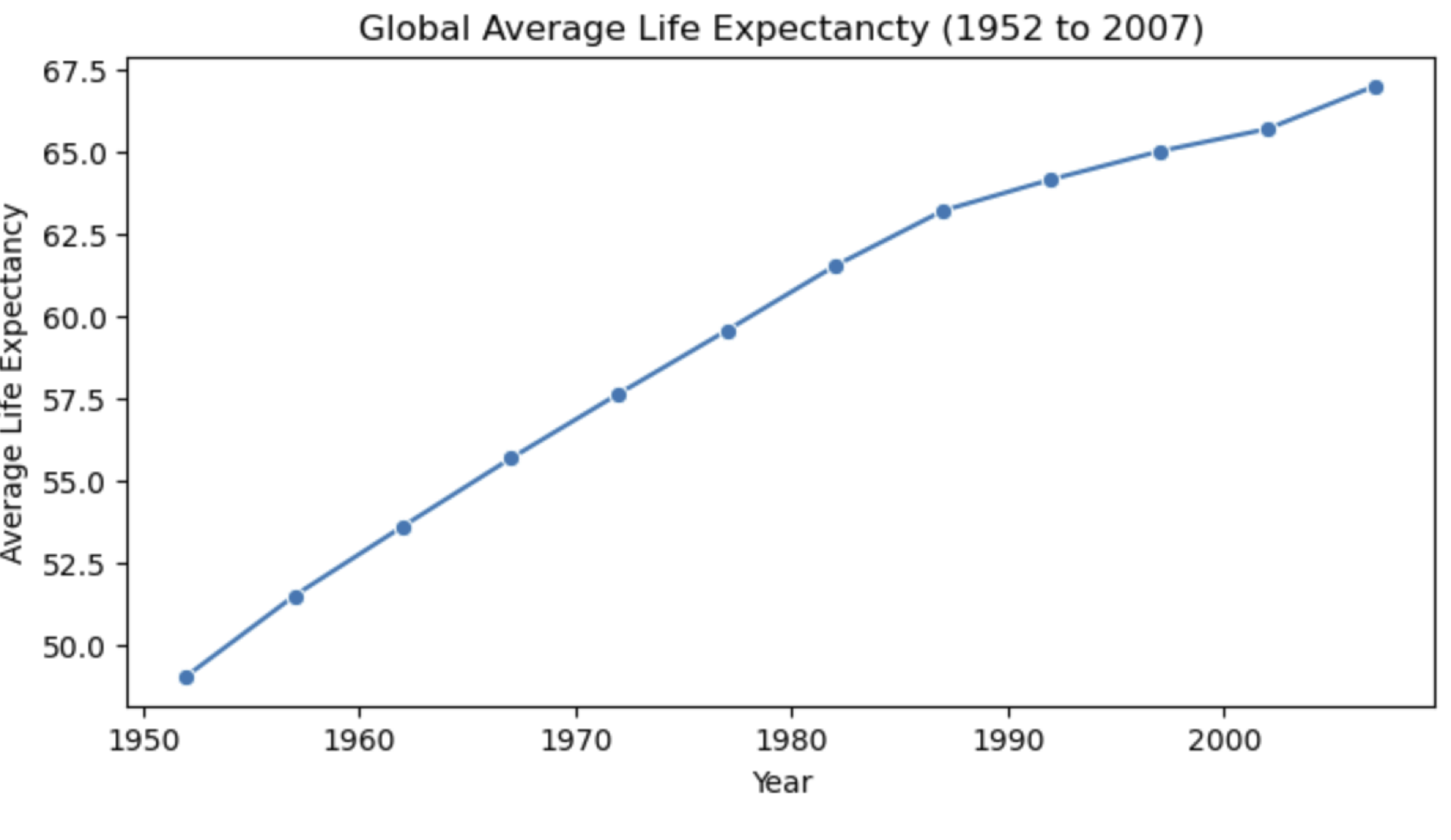


1. Compute the correlation matrix between GDP per capita, life expectancy, and population for the dataset. Then, use Seaborn to visualize this correlation matrix as a heatmap.



1. How has the global average life expectancy changed from 1952 to 2007? Plot a line graph to visualize this trend.

Ans : 



Subjective Question: Discuss the various reasons that could have contributed to the change.

1. The graph avove reprsents an upward trend indicating improvements in life expectancy over the years.

2. This is a result of a combination of various factors

3. This includes medical advancements, economic development, improved public health initiatives and better nutrition

4. Adavance in Healthcare The development in vaccines, antibiotics and other medical treatments and advanced surgical techniques have signifacantly reduced the infectious diseasses, contributing to increased life expectancy.

5. Health initiatives. Increase in awareness and education about health practices, such as vaccination, proper nutrition, and exercise, Improved santation,clean drinking water, and better waste management practices have helped improve overall health outcomes.

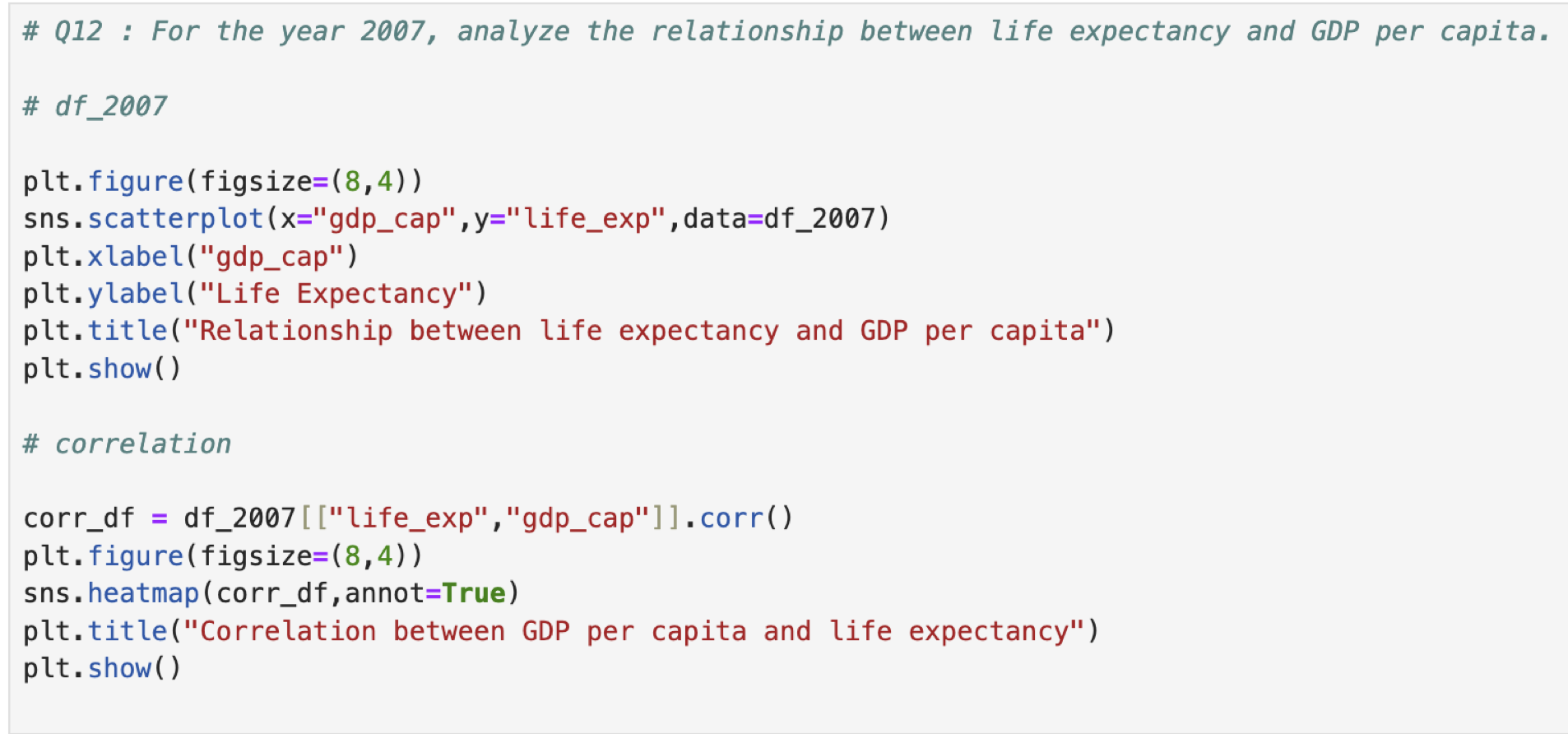
6. Nutrition and Food: The Green Revolution and advances in agricultural practices have increased food production, reducing malnutrition, which are critical factor affecting life expectancy.

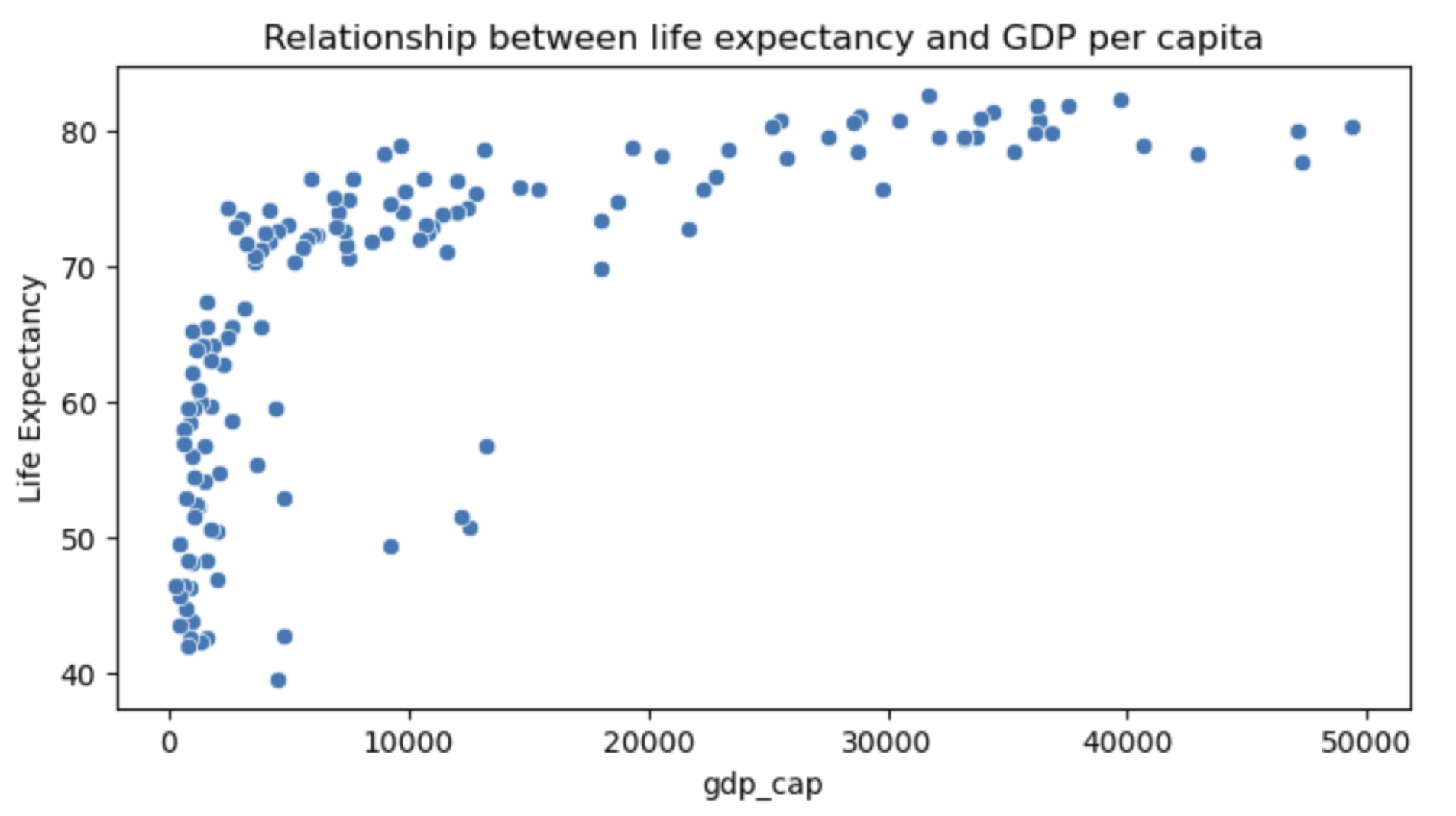
7. Economic Development: Economic growth in many parts of the world has led to higher income levels, which in turn Rave Improved access to healthcare better nutrition and healthier living conditions.

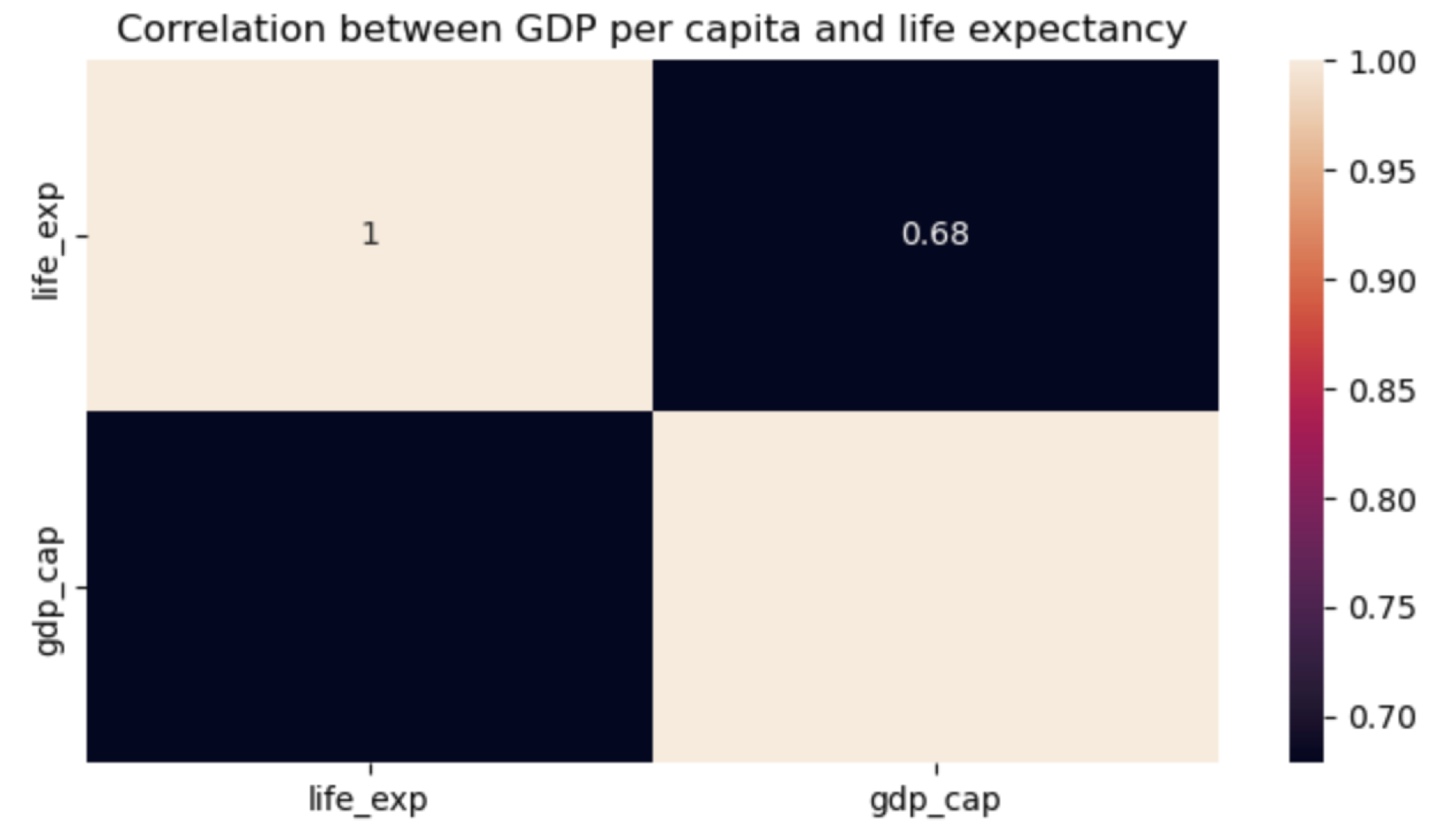
8. Better infrastructure, including transportation and communication systems has facilaged the delivery of healthcare services especially in remote areas.

These factors have collectively contributed to longer and healthier lives across much of the world

1. For the year 2007, analyze the relationship between life expectancy and GDP per capita.

Ans : 





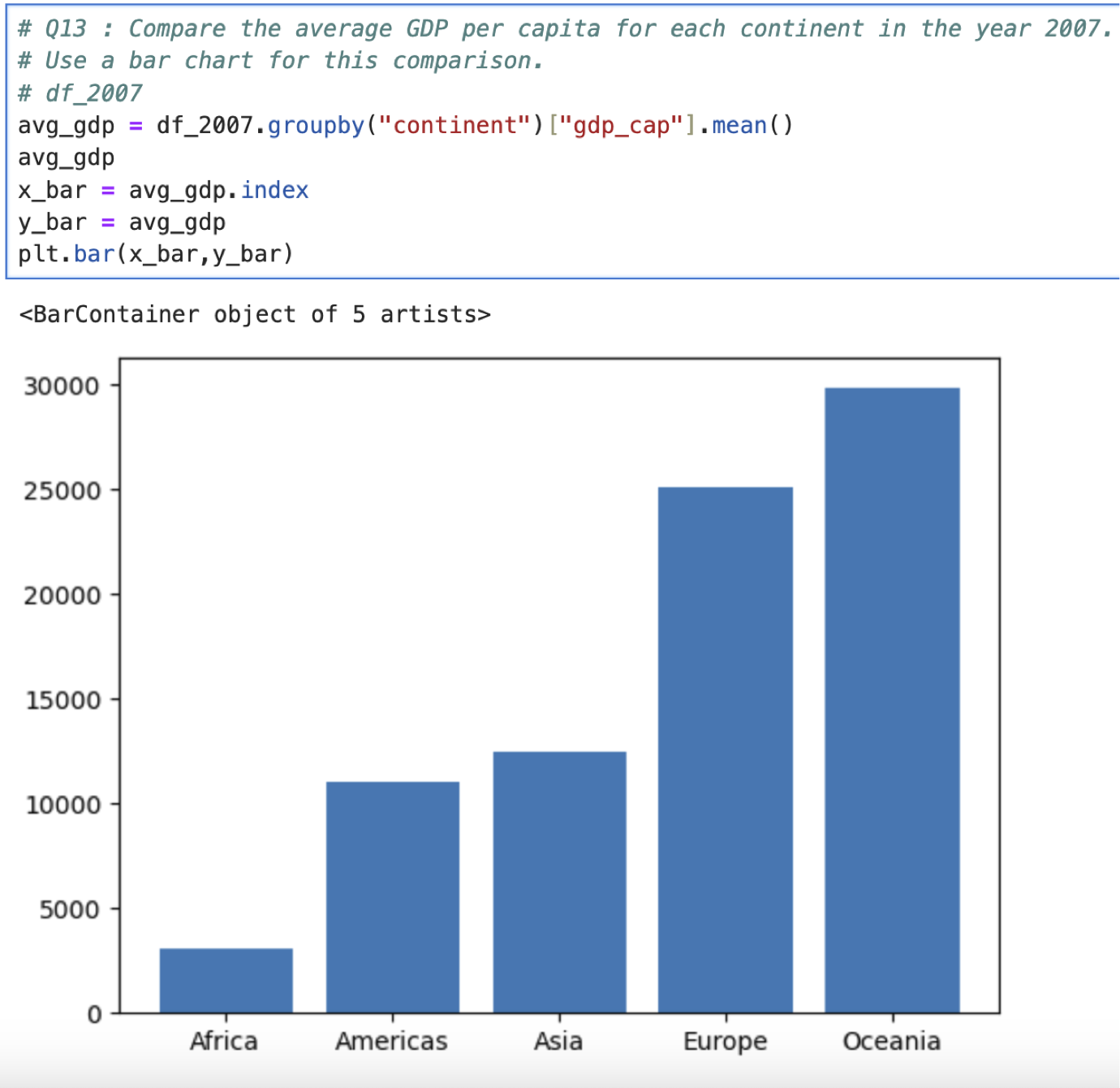
Subjective Question: Is there a noticeable trend or correlation? Represent this using a scatter plot.

1. Both scatter plot and correlation analysis reveal a positive relationship between GDP per capita and life expectancy, with a correlation of 0.68 suggesting a moderate positive relationship.

2. While higher GDP per capita generally leads to better living conditions and healthcare, and thus higher life expectancy, it's important to consider other factors that can influence this relationship.

3. Wealth alone does not guarantee a longer life, but it certainly contributes to creating the conditions for a healthier and longer life.

1. Compare the average GDP per capita for each continent in the year 2007. Use a bar chart for this comparison.

Ans : 

Subjective Question: Why is the average GDP per capita for Oceania higher than the Americas even though the Americas have more countries?

1. Oceania represents Australia and New Zealand, both of which are developed, high-income countries with strong economies. Their GDP per capita is significantly higher compared to many other countries worldwide.

2. Since these two countries dominate the continent in terms of population and economic output, their high GDP per capita strongly influences Oceania's average, pushing it higher.

3. The America includes a diverse range of countries with varying levels of economic development. North America (primarily the United States and Canada) has a high GDP per capita, but Latin American countries, which make up a significant portion of the Americas, generally have lower GDP per capita.

4. The presence of many lower and middle-income countries in Latin America pulls down the overall average GDP per capita for the Americas, despite the high figures from North America.

5. These two countries, with their advanced economies, drive up the average for the entire continent. In contrast, the Americas have a more diverse range of economies, with many countries in Latin America having lower GDP per capita, which brings down the overall average for the continent

6. The number of countries and their respective economic weights.

1. Compare the life expectancy and GDP per capita of Afghanistan (a country known for its historical conflicts) and Switzerland (representing a peaceful and economically prosperous country) using the dataset provided.

Subjective Question: 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.

* Firstly, for the year 2007, use a bar chart to directly compare the life expectancy and GDP per capita between these two countries.



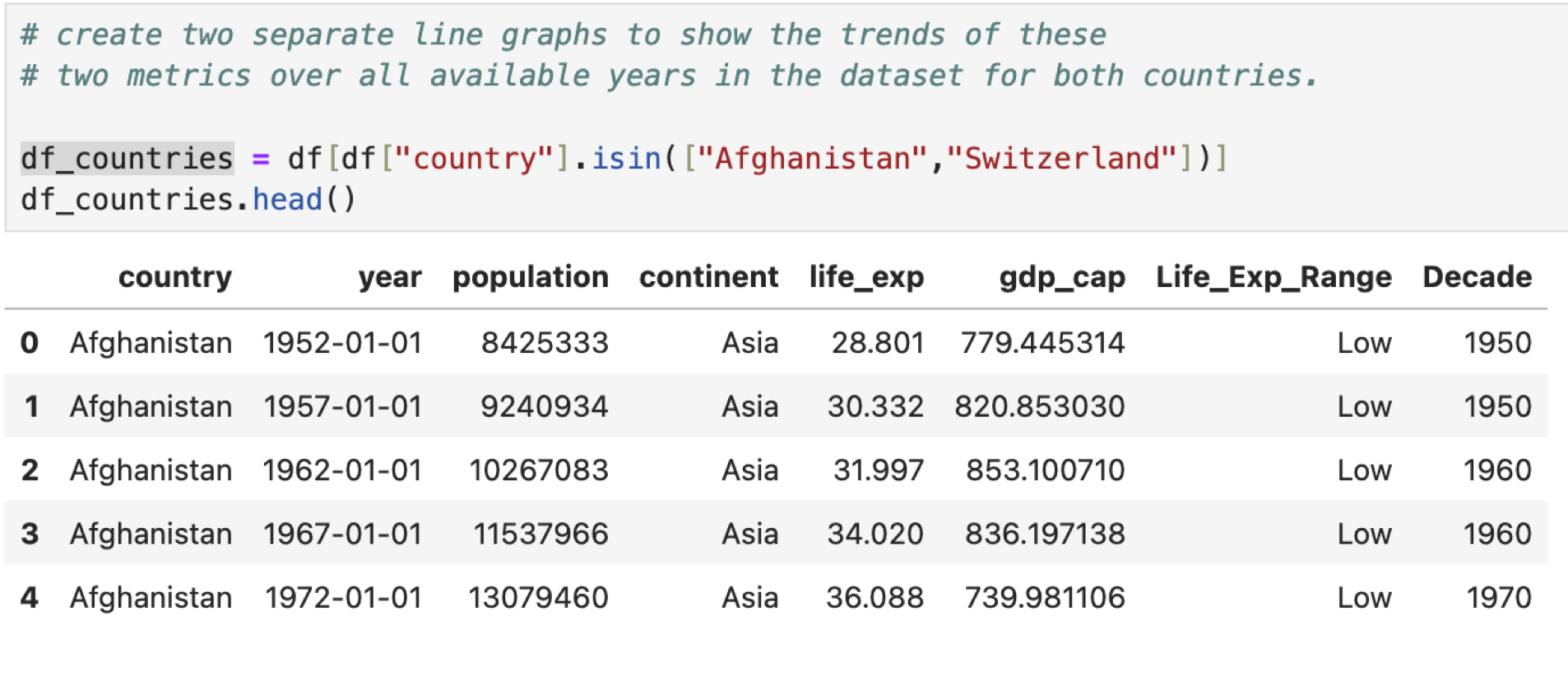


1. The bar chart comparison reveals that Afghanistan, a country affected by prolonged conflicts, has a life expectancy nearly half that of Switzerland highlighting the severe impact of instability on public health

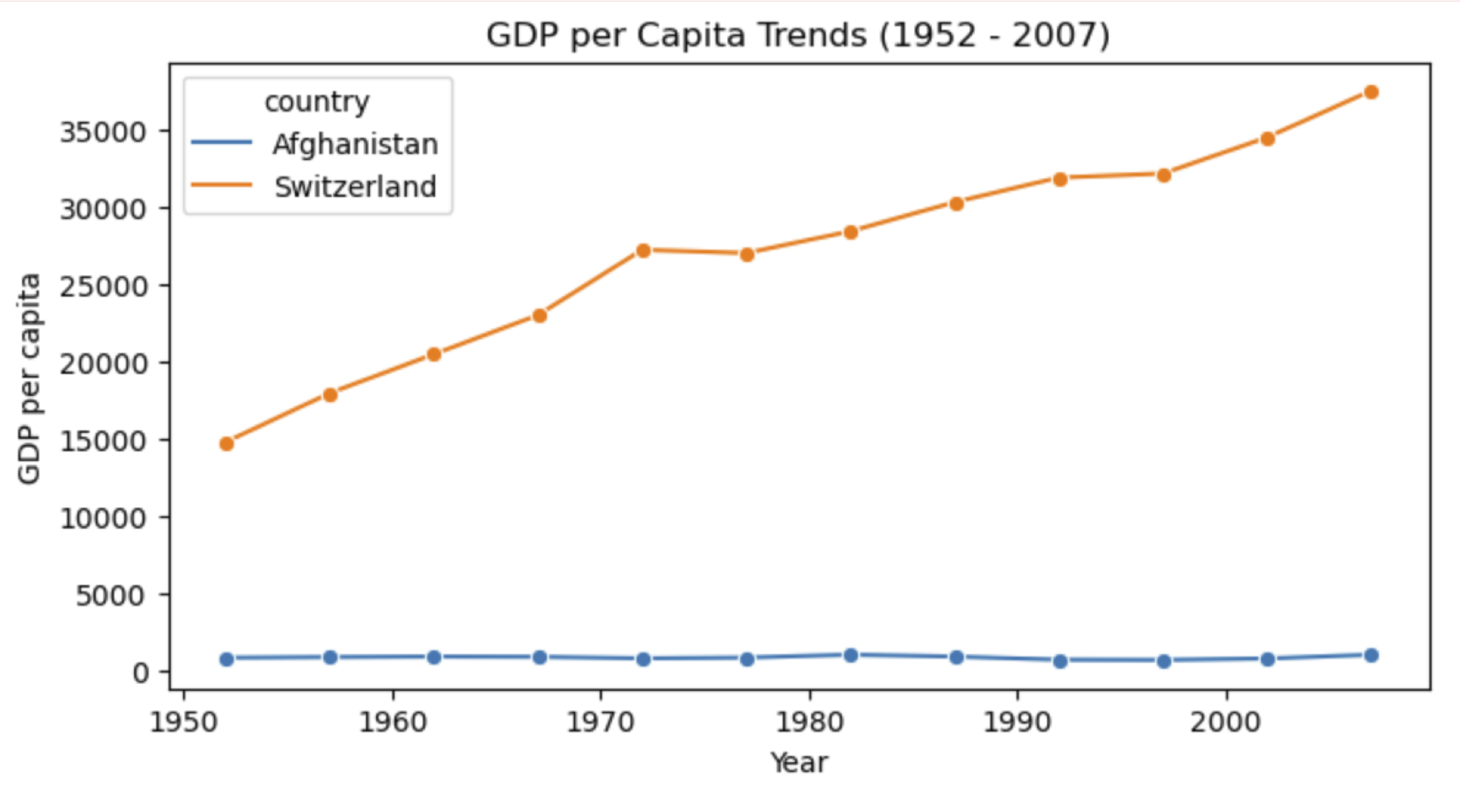
2. Additionally, Afghanistan's GDP per capita is drastically lower, almost negligible when compared to Switzerland, underscoring the economic disparity between a conflict-ridden nation and a peaceful, economically prosperous one.

3. These differences illustrate how peace and stability are critical for both the economic well-being and overall health of a population

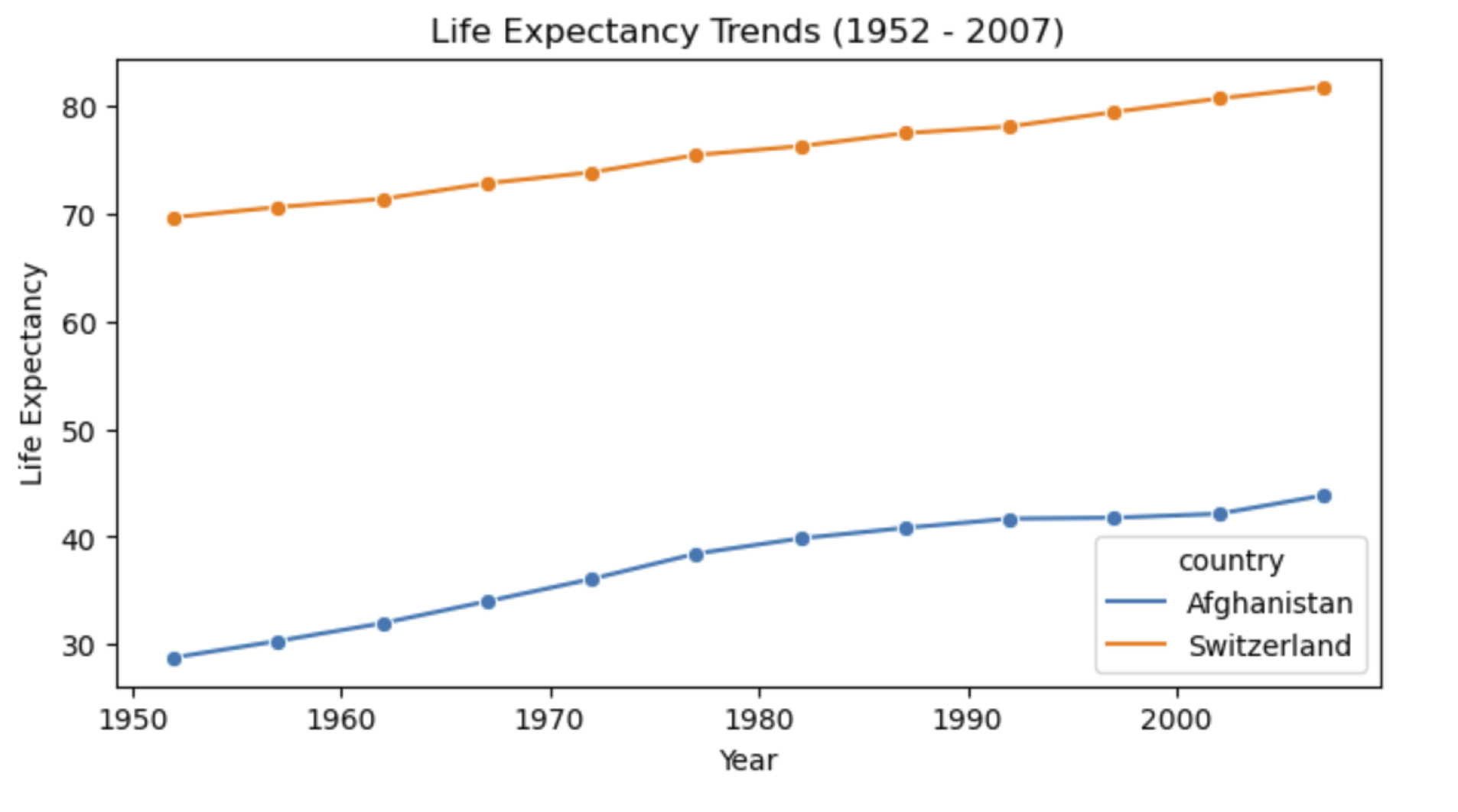
* Then, create two separate line graphs to show the trends of these two metrics over all available years in the dataset for both countries.











1. The increasing life expectancy in both Afghanistan and Switzerland can be attributed to global advancements in healthcare, public health initiatives, and improved access to medical services.

2. However Switzerland's significant rise in GDP per capita reflects its stable govemance, diversified economy, and strong global trade connections

3. In contrast, Afghanistan's stagnant GDP per capita is a result of prolonged coicts, political instability, and limited infrastructure, which have hindered economic growth and development

4. These factors underscore the crucial role of peace, stability, and effective governance in driving both economic and social progress.