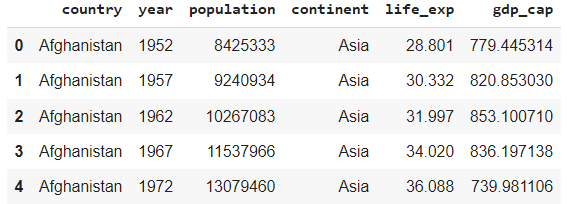
**Comprehensive Exploratory Data Analysis of Global Trends Using Gapminder**

**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.

**Objectives:**

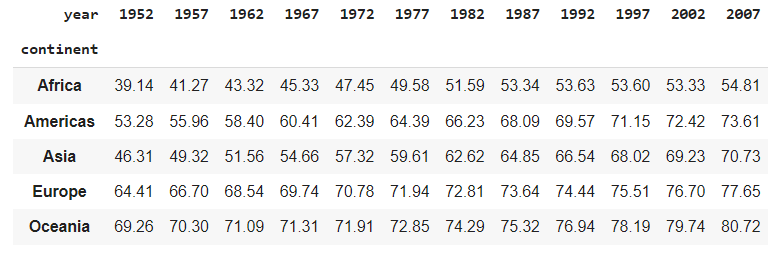
**1. Load the dataset and display the first few rows. How many countries does the dataset have?**



**Insight:** The dataset contains information on 142 unique countries.

**2.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.**

**Pivot Table**



**Insight:** Overall, life expectancy has improved across all continents from 1952 to 2007, reflecting global advancements in healthcare, economic development, and living conditions. However, disparities still exist, with continents like Africa lagging behind others, highlighting the need for continued efforts to address healthcare and socio-economic challenges.

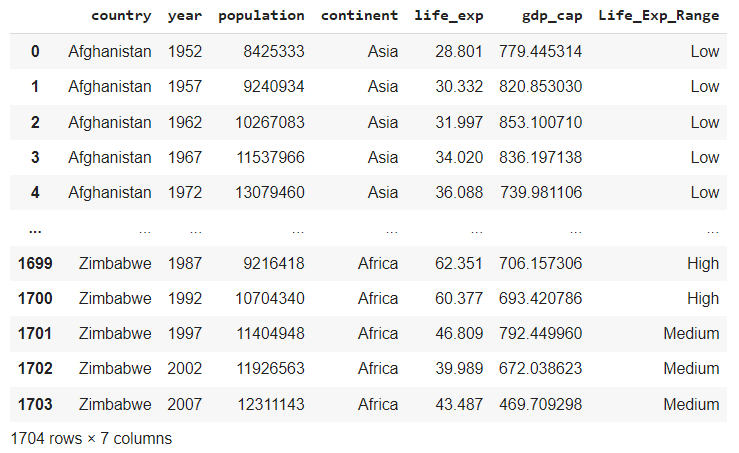
**3. Which countries had a GDP per capita higher than the 75th percentile in 2007?**

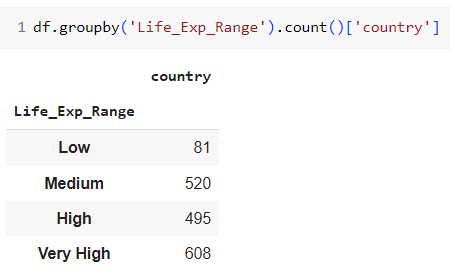
**75th percentile of gdp\_cap in 2007 is 18008.84**

**These countries represent the global economic performance in 2007, with a GDP per capita above 18,008.84.**



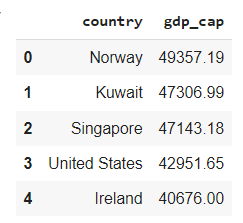
**4. Categorize the 'life\_exp' into 4 equally ranged bins from 'Low' to 'Very High'. Use cut to create these categorical life expectancy groups and add them as a new column 'Life\_Exp\_Range'.**

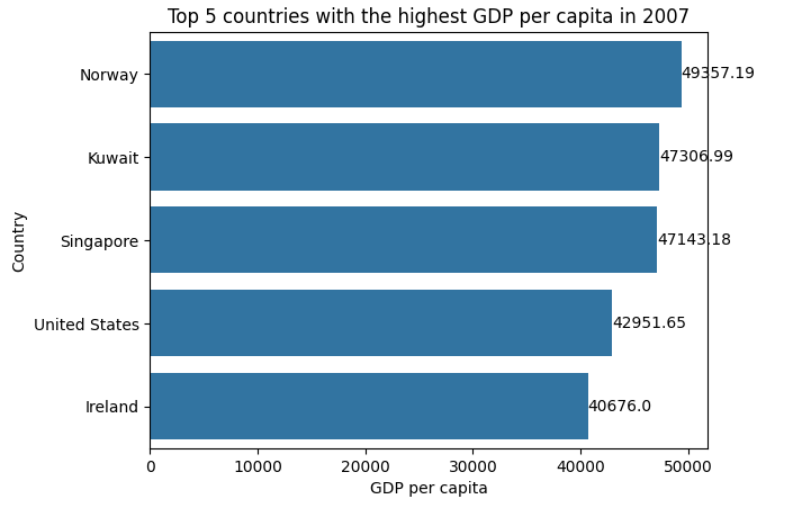




**Insight:** By categorizing life expectancy into bins, we can quickly assess which countries fall into different life expectancy ranges, making it easier to identify trends and outliers. For example, a significant number of countries might fall into the 'Medium' category, indicating average life expectancy globally.

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





**Insight:**

Top Economies in 2007: The countries listed—Norway, Kuwait, Singapore, the United States, and Ireland

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

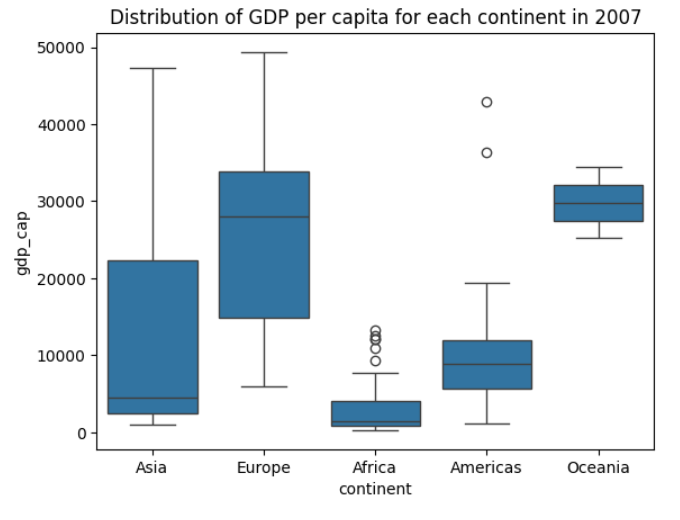
**Code:**



**Output:**

****

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



**Insights:**

* Africa shows several high outliers, possibly countries with unique economic advantages.
* The Americas have a few high outliers, likely representing wealthier nations.
* Europe and Oceania show no significant outliers, suggesting more uniformity in GDP per capita within these regions.
* The boxplot illustrates the significant economic disparities across continents. Europe and Oceania are at the higher end of the spectrum, reflecting stronger and more consistent economic performance. In contrast, Africa shows the lowest GDP per capita with a narrow distribution, highlighting the continent's economic challenges. Asia and the Americas show more variability, indicating a mix of wealthier and less wealthy nations within these regions.

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



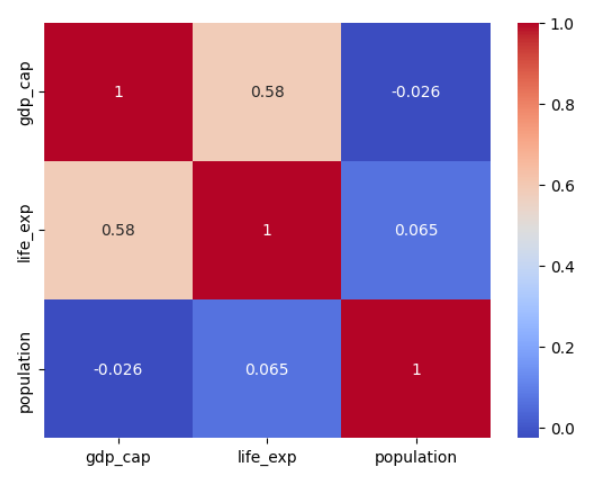
**Insight:** Countries with life expectancy over 80 years are typically developed nations with advanced healthcare systems, low infant mortality rates, and high standards of living. These factors contribute to longer life spans.

**9. 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).**



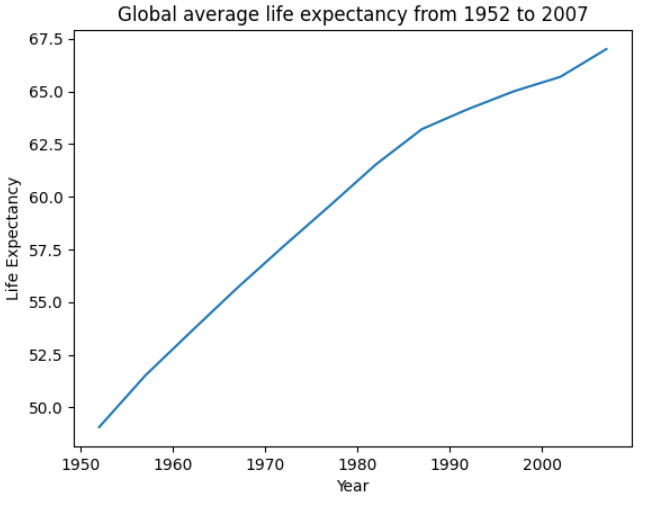
**Insight:** Grouping years into decades allows us to analyze trends over longer periods, smoothing out short-term fluctuations and focusing on significant long-term patterns.

**10. 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.**



**Insight:** The heatmap reveals the relationships between GDP per capita, life expectancy, and population. A strong positive correlation between GDP per capita and life expectancy might indicate that wealthier countries tend to have better healthcare and living conditions, leading to longer life expectancy.

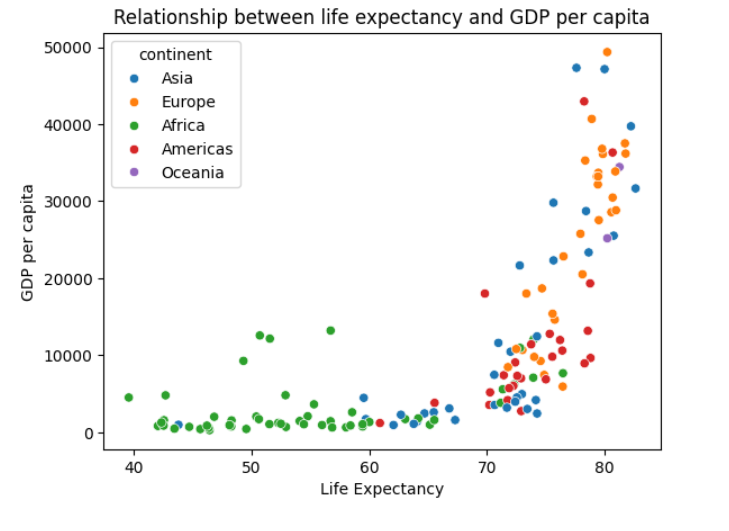
**11. How has the global average life expectancy changed from 1952 to 2007? Plot a line graph to visualize this trend.**



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

**Discussion:** The global average life expectancy has shown a steady increase from 1952 to 2007. Several factors likely contributed to this trend, including advancements in medical technology, improved access to healthcare, better sanitation and nutrition, and overall economic development. Efforts to combat infectious diseases, reduce child mortality, and promote healthier lifestyles have also played a significant role. However, the rate of improvement may vary across regions, reflecting disparities in development and healthcare infrastructure.

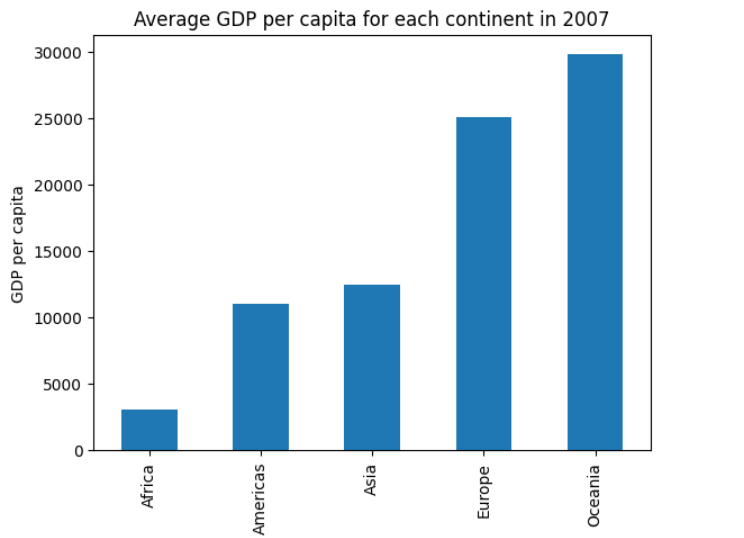
**12.For the year 2007, analyze the relationship between life expectancy and GDP per capita.**



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

**Discussion:** A scatter plot of life expectancy versus GDP per capita for the year 2007 typically shows a positive correlation, indicating that countries with higher GDP per capita tend to have longer life expectancies. This trend can be attributed to the fact that wealthier countries can afford better healthcare systems, higher standards of living, and more robust social safety nets. However, there may be exceptions where countries with moderate GDP per capita achieve high life expectancy due to efficient healthcare systems or social policies.

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

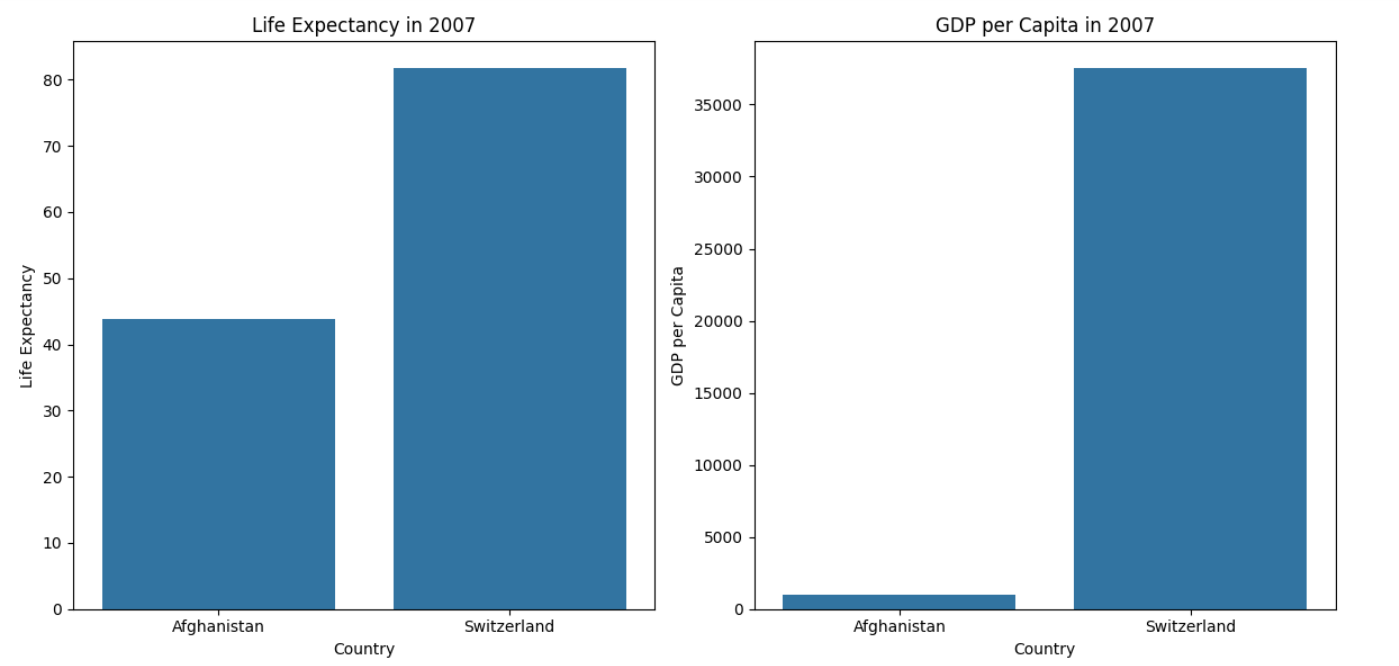


**Subjective Question:  Why is the average GDP per capita for Oceania higher than the Americas even though the Americas have more countries?**

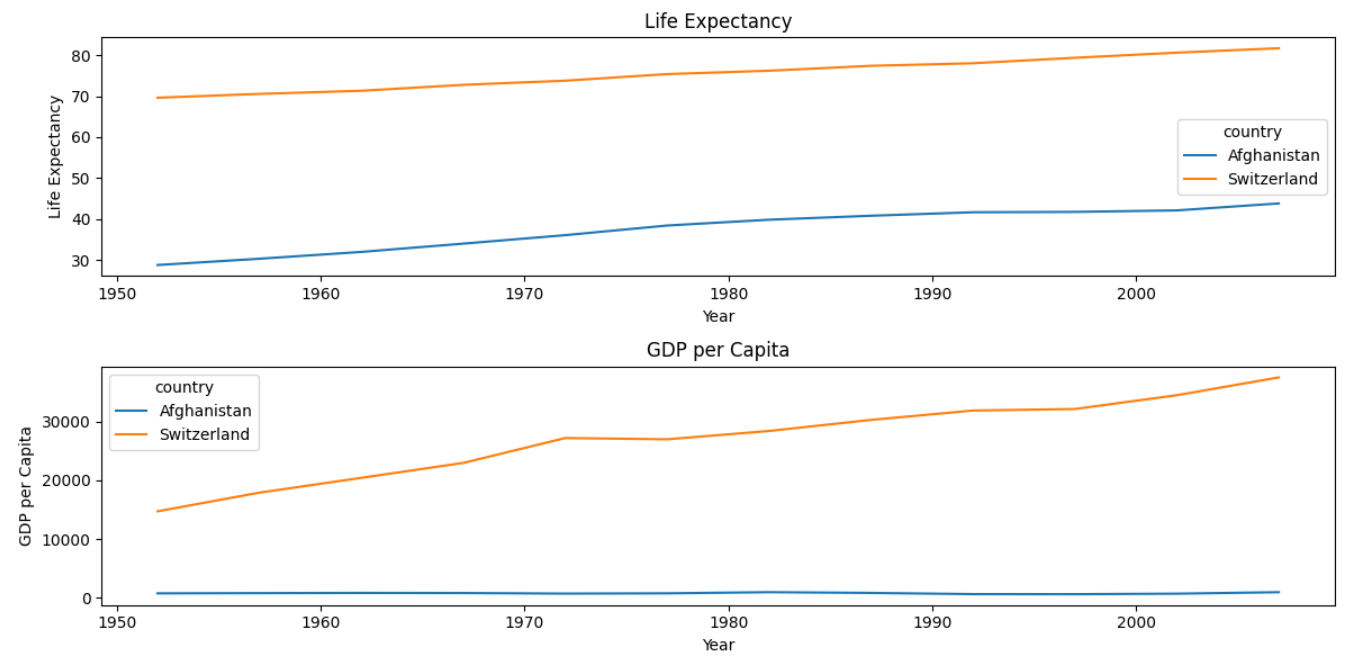
**Discussion:** Oceania's higher average GDP per capita compared to the Americas, despite having fewer countries, can be attributed to the economic performance of Australia and New Zealand, which dominate the region. These countries have strong, stable economies with significant natural resources, advanced industries, and well-developed infrastructure. In contrast, the Americas include a mix of high-income countries like the United States and Canada and lower-income countries in Central and South America, which lowers the overall average GDP per capita for the continent.

**14. 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.**

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



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



**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.**

**Discussion:** Afghanistan and Switzerland present stark contrasts in both life expectancy and GDP per capita. Afghanistan, with its history of conflict, instability, and economic challenges, shows significantly lower values in both metrics. In contrast, Switzerland, known for its peace, stability, and economic prosperity, exhibits high life expectancy and GDP per capita. The data highlights how a nation's stability directly impacts its economic development and the health and well-being of its population. Continuous conflict disrupts economic growth, healthcare systems, and social structures, leading to poorer outcomes in these key metrics over time.

**Conclusion and Integrated Insights**

The comprehensive exploratory data analysis of the Gapminder dataset provides a deep understanding of global demographic, economic, and health trends from 1952 to 2007. This analysis reveals several key insights:

**Global Life Expectancy Trends:** Over the years, life expectancy has generally improved across all continents, driven by advancements in healthcare, better nutrition, and increased access to clean water and sanitation. However, the data also highlights significant disparities between continents, with Africa consistently lagging behind due to persistent challenges such as poverty, disease, and political instability.

**Economic Disparities Across Continents:** The analysis of GDP per capita shows stark differences in economic development between regions. Continents like Europe and Oceania exhibit higher and more uniform GDP per capita, while Africa and Asia display greater variability, reflecting the diverse economic landscapes within these regions. This disparity is particularly evident when comparing countries like Switzerland and Afghanistan, where the former's economic stability and growth have led to high GDP per capita and life expectancy, while the latter's ongoing conflicts have severely hindered its development.

**Correlation Between Wealth and Health:** The correlation matrix analysis reveals a strong positive relationship between GDP per capita and life expectancy, suggesting that wealthier countries tend to have better healthcare systems, higher standards of living, and, consequently, longer life expectancies. This correlation is crucial for policymakers, indicating that economic growth and investments in healthcare can lead to significant improvements in population health.

**Regional Differences in Development:** The regional analysis of GDP per capita and life expectancy underscores the impact of historical, social, and political factors on development. For instance, countries in Europe and North America, which have enjoyed long periods of stability, show higher development indicators compared to regions that have experienced prolonged conflicts or economic instability.

**Impact of Conflict on Development:** The stark contrast between Afghanistan and Switzerland illustrates the profound impact of conflict on a nation's development. Afghanistan's struggles with low life expectancy and GDP per capita are direct consequences of its turbulent history, while Switzerland's peaceful and prosperous environment has enabled continuous improvement in these metrics. This comparison emphasizes the importance of stability and peace in fostering economic growth and improving quality of life.

**Long-Term Trends and Implications:** The decade-based analysis of life expectancy and GDP per capita trends highlights the importance of long-term strategies in addressing global challenges. While short-term fluctuations may occur due to various factors, sustained efforts in improving healthcare, education, and economic policies are essential for achieving lasting progress.

**Final Thoughts:**

This EDA not only provides valuable insights into historical trends but also serves as a foundation for future research and policy development. Understanding the relationships between economic development, health, and demographic changes is critical for creating effective strategies that promote global well-being and reduce disparities. The data highlights the importance of investing in healthcare, ensuring political stability, and fostering economic growth to achieve better outcomes for populations worldwide.