**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?
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.
3. Which countries had a GDP per capita higher than the 75th percentile in 2007?
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'.
5. Identify the top 5 countries with the highest GDP per capita in 2007. Use a horizontal bar chart to display this data.
6. Find all country names that start with "I" and end with "a" using regex.
7. Create a boxplot using Seaborn to compare the distribution of GDP per capita for each continent in 2007.
8. Find all countries with a life expectancy of over 80 years in 2007. List these countries and their respective continents.
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).
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.
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.

1. 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.
2. 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?

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.

* 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.