

World Life Expectancy Data Cleaning and Economic Correlations Analysis

Objective

The objective of this analysis was to clean and preprocess the World Life Expectancy dataset and explore the relationships between life expectancy and economic indicators such as GDP, BMI, and adult mortality. The analysis aims to provide insights into global life expectancy trends and how economic factors influence health outcomes across countries.

Data Cleaning Steps

1. Duplicate Records Removal

- **Process:** Duplicates were identified by checking the combination of the "Country" and "Year" columns.
- **Action:** A query was executed to group data by "Country" and "Year" and count occurrences. Duplicates were found for countries such as Ireland, Senegal, and Zimbabwe. These duplicates were removed using a subquery to identify the Row_IDs and delete the redundant records.

2. Handling Missing Values in the 'Status' Column

- **Process:** The 'Status' column, which categorizes countries as either "Developing" or "Developed," had missing values.
- **Action:** A query identified rows with empty 'Status' values. For countries categorized as "Developing," the missing values were populated with "Developing," and for the remaining missing rows, the 'Status' was updated to "Developed."

3. Handling Missing Life Expectancy Data

- **Process:** Missing values in the "Life expectancy" column were identified.
- **Action:** For each missing life expectancy value, the average of the previous and next year's values was calculated to estimate the missing data, based on the assumption that life expectancy typically changes gradually over time.

Exploratory Data Analysis (EDA) on Life Expectancy and Economic Correlations

Key Findings

1. Life Expectancy Trends by Country

- **Min/Max Life Expectancy:** Significant increases in life expectancy were observed in countries like Zimbabwe, Haiti, and Eritrea.
- **Global Trends:** Life expectancy globally increased from 66.75 years in 2007 to 71.62 years in 2022, showing a steady upward trend over the period.

2. Correlation Between GDP and Life Expectancy

- **Positive Correlation:** Higher GDP is associated with higher life expectancy. Countries with GDPs greater than \$1500 had an average life expectancy of 74.20 years, while those with lower GDPs had an average life expectancy of 64.69 years.

3. GDP Categorization and Life Expectancy

- **High vs. Low GDP:** Of the countries with a GDP higher than \$1500 (1,326 countries), the average life expectancy was 74.20 years. In contrast, countries with a GDP lower than \$1500 (1,612 countries) had an average life expectancy of 64.69 years, highlighting the significant impact of economic development on health outcomes.

4. Developed vs. Developing Countries

- **Life Expectancy Gap:** Developing countries had an average life expectancy of 66.8 years, while developed countries had an average of 79.2 years. This gap reflects the disparities in healthcare quality, economic conditions, and overall development.

5. Life Expectancy and BMI

- **Positive Correlation:** A positive correlation between BMI and life expectancy was found. Countries with higher average BMI tended to have higher life expectancy, suggesting that better nutrition and overall health are linked to longer lifespans.

6. Adult Mortality and Life Expectancy

- **Cumulative Impact:** The rolling total of adult mortality rates showed significant fluctuations over time. These fluctuations were found to be closely tied to life expectancy trends, with countries like the United States experiencing notable shifts in mortality rates impacting life expectancy.

Conclusion

The analysis reveals that economic factors such as GDP and BMI play a critical role in shaping life expectancy outcomes. Countries with higher GDP tend to have longer life expectancies, reflecting better access to healthcare, nutrition, and overall living conditions. Conversely, lower GDP countries face challenges in improving life expectancy, often due to limited healthcare resources and poorer living conditions.

The relationship between adult mortality and life expectancy further emphasizes the importance of addressing health issues such as adult mortality and nutrition to improve life expectancy outcomes. The disparity in life expectancy between developed and developing countries calls for ongoing efforts to improve global healthcare systems, particularly in lower-income regions.