

Final Project Report

CS 2704 - Data Analytics using Python

Title: Analyzing the Impact of Healthcare Expenditure on Life Expectancy Across Countries

Author: Kuvalesh Parsad

1. Introduction and Background

Life expectancy serves as a critical indicator of a nation's overall well-being, public health infrastructure and socioeconomic status. It reflects a population's access to medical care, sanitation, education and preventive health measures. One of the most intuitive drivers of life expectancy is healthcare expenditure, which means how much a country spends on health services per person.

This project explores whether there is a statistically meaningful relationship between a country's healthcare expenditure per capita and its average life expectancy at birth. By leveraging publicly available data from the World Bank, I aim to use data-driven analysis to test this hypothesis and explore global patterns that link investment in health with long-term outcomes.

2. Hypothesis

Countries with higher healthcare expenditure per capita have higher life expectancy rates.

The underlying assumption is that countries that invest more in their healthcare systems will provide better medical services, improved disease prevention and overall better health infrastructure. This should result in longer average lifespans. However, life expectancy is influenced by more than just spending on healthcare, it can also depend on factors such as education, environment and lifestyle.

This project seeks to test whether there is a significant relationship between spending on healthcare and increased life expectancy using descriptive and predictive analytics.

3. Analysis and Implication

Two datasets were obtained from the World Bank:

- **Life Expectancy at Birth (Years):** Measures the average number of years a newborn is expected to live.
- **Healthcare Expenditure per Capita (US\$):** Measures how much a country spends on health per person annually.

Data was extracted for multiple years and cleaned using a Python function named `load_and_clean_data()`, which merged the datasets, standardized column names, and removed missing values.

Descriptive Analysis

- A **scatter plot** which was created with `sns.scatterplot()` showed a clear upward trend which means higher spending tends to align with higher life expectancy.
- A **Heatmap** created using `sns.heatmap()` showed a positive correlation between the two variables.

These visualizations confirm that, generally, countries that spend more on healthcare have longer-lived populations.

Predictive Analysis

A linear regression model was used to predict life expectancy based on healthcare spending:

- **Response variable:** Life Expectancy
- **Predictor variable:** Healthcare Expenditure per Capita
- **R² Score:** 0.316

This means that healthcare expenditure explains about 31.6% of the variation in life expectancy. While the relationship is statically significant, it suggests that other factors also play a large role in determining life expectancy.

4. Conclusion

The analysis supports our hypothesis: there is a positive relationship between healthcare expenditure and life expectancy. Countries that spend more on healthcare tend to have populations that live longer.

However, the predictive power of the model ($R^2 = 0.316$) reveals that healthcare spending alone does not fully explain differences in life expectancy. Other variables, including socioeconomic conditions, education, diet, environmental factors, and public health policies, must also be considered.

5. References

Healthcare Expenditure per Capita

<https://data.worldbank.org/indicator/SH.XPD.CHEX.PC.CD>

Life Expectancy at Birth

<https://data.worldbank.org/indicator/SP.DYN.LE00.IN>

Github

<https://github.com/Nileshp1007/Final-Presentation>