

# Project Part 1

Risshi Naavaal, Caleb Adhikari, and Jonas Lee

## Introduction

Oftentimes, quality of life is measured by Real GDP per Capita; the higher this is for a country, the higher the perceived quality of life its citizens have. However, we believe it is important to look at health outcomes, specifically life expectancy and infant mortality, to measure quality of life. We will be looking at the relationship between health expenditure as a % of GDP and infant mortality and life expectancy at birth to see if additional spending is strongly correlated with better health outcomes.

- Do countries with health expenditure above the median have lower infant mortality rates than countries that spend below the median?
- Do countries with health expenditure above the median have higher life expectancy than those that spend below the median?
- Lastly, we will observe countries that have a below median GDP per capita and determine whether, from that set of countries, countries that spend more than 5% on health expenditure have a greater life expectancy than countries that spend less than 5% on health expenditure.

## Data Summary

The data used in our project represents a population. The data was collected from the World Bank Open Data Portal and contained data on territories, countries, and regional groupings across the world. We filtered the data to remove any data that did not represent an individual country, and collected data from 2022 only. Further, when merging the datasets

for Health Expenditure, Infant mortality Rate, GDP per Capita, and Life Expectancy, some countries were missing data so we excluded them from the dataset entirely. These countries were Cuba, Eritrea, North Korea, Kosovo, South Sudan, Ukraine, Venezuela, and Yemen. After we completed the data cleaning process, we had a total of 188 countries.

We removed regional groupings to ensure there was no double counting of data, and narrowed our research to countries because territories tend to be missing data.

The data comes from a population, but because we were missing some data for 8 countries, it is not a full population. There could be some missing data bias as the countries that are missing data are similar in that they have political instability (ie. Ukraine and North Korea). This means our data does not fully represent unstable countries. We chose the year 2022, and the data may reflect changes due to the Covid-19 Pandemic, so some of the data like health expenditure as a % of GDP may be inflated.

The data is perfect for our research purposes as we have standard units to measure across all countries, and official published data is available. We have a large dataset with 188 observations to reduce variability. Below is a data dictionary for clear understanding of the variables and units.

## Data Dictionary

Table 1: Data Dictionary

Variable		Unit of
Name	Description	Measure- ment
Country Name	Official name of the sovereign country	—
Country Code	Three-letter country code assigned to countries (e.g., USA for United States, FRA for France)	3-letter code
Health Expenditure	Total spending on health care as a percentage of gross domestic product.	Percentage of GDP
Infant Mortality Rate	The number of infant deaths (children under one year old) per 1,000 live births in 2022. Lower values indicate better child survival rates.	Deaths per 1,000 births

Variable Name	Description	Unit of Measurement
GDP per Capita	The total economic output of a country divided by its population, expressed in current U.S. dollars. Represents the average income per person.	U.S. dollars
Life Expectancy	The average number of years a newborn is expected to live at birth, measured in years.	Years

## Data Exploration

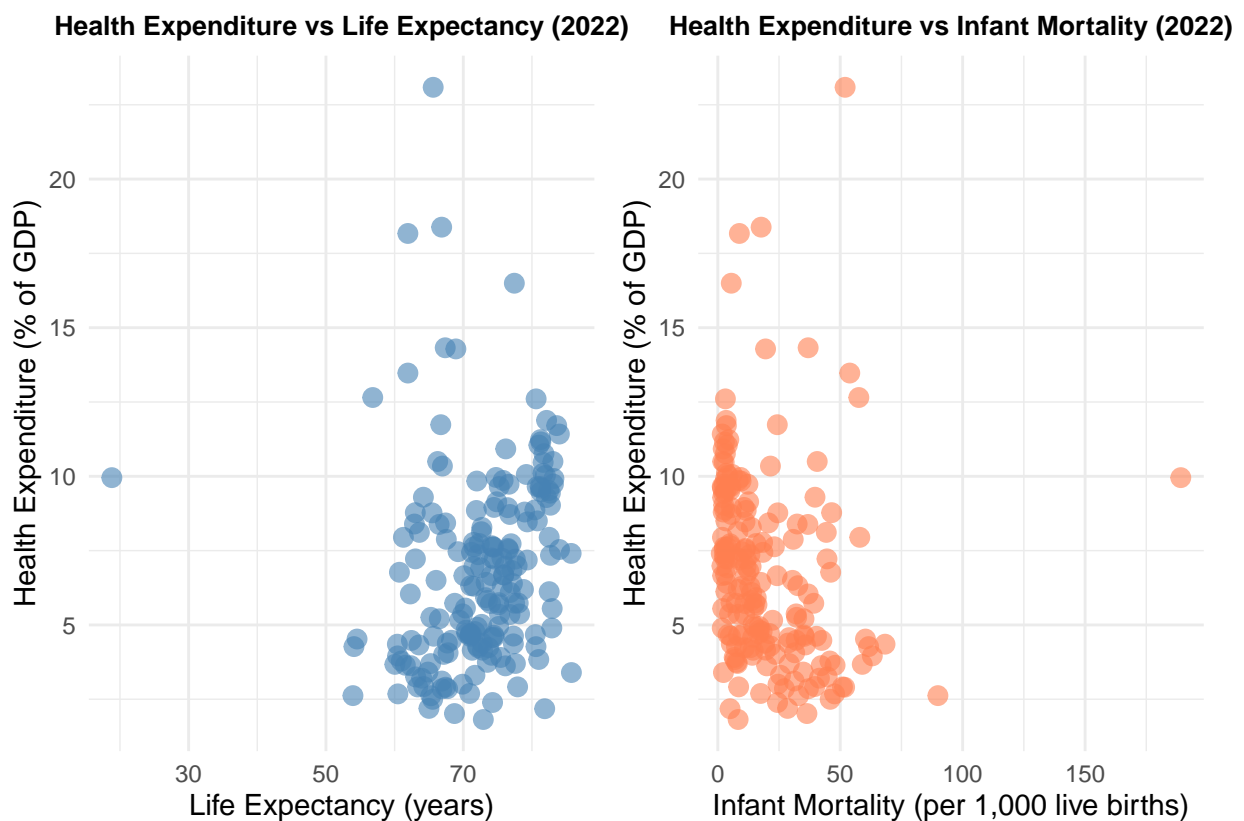


Table 2: Summary Statistics for Health and Economic Indicators (2022)

Variable	Mean	Median	Min	Max
Health Expenditure (% of GDP)	6.85	6.43	1.82	23.09
Infant Mortality (per 1,000 births)	20.03	13.60	1.40	189.10
GDP per Capita (US\$)	17496.38	6523.41	250.63	226052.00
Life Expectancy (years)	72.40	73.25	18.82	85.75

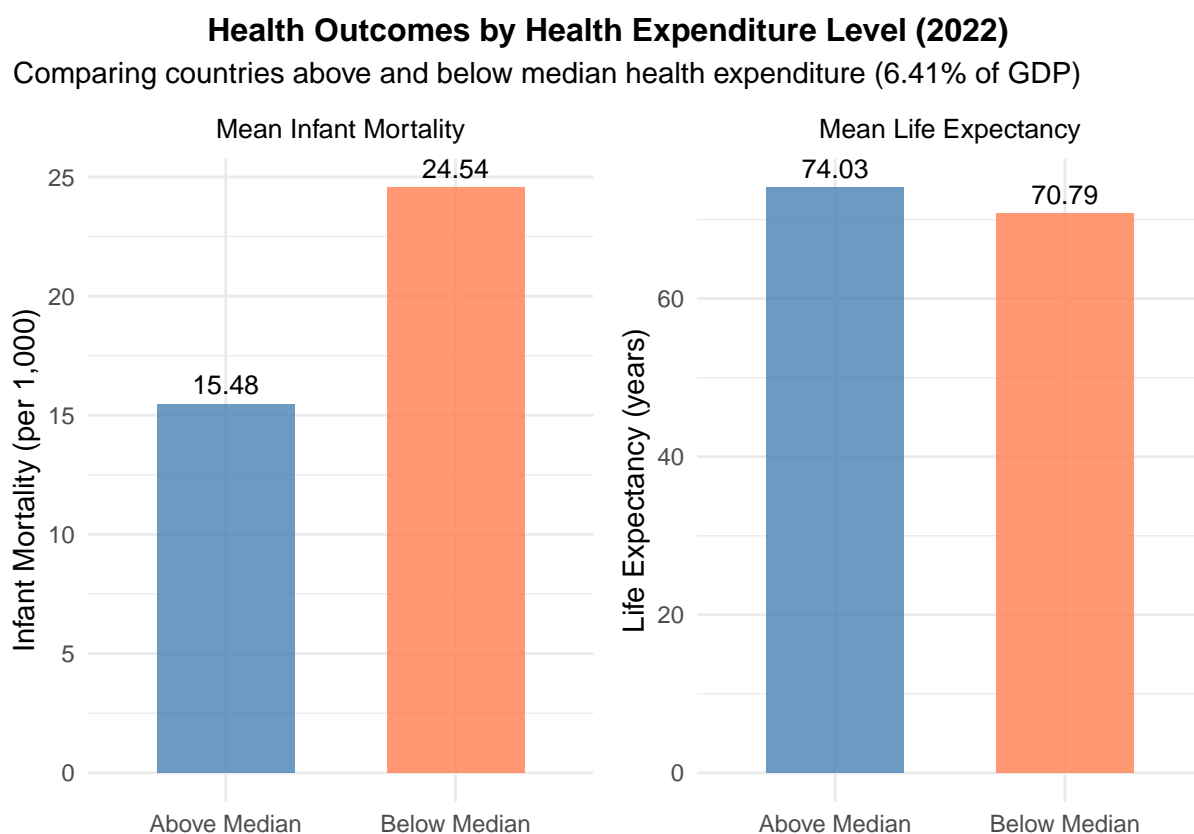
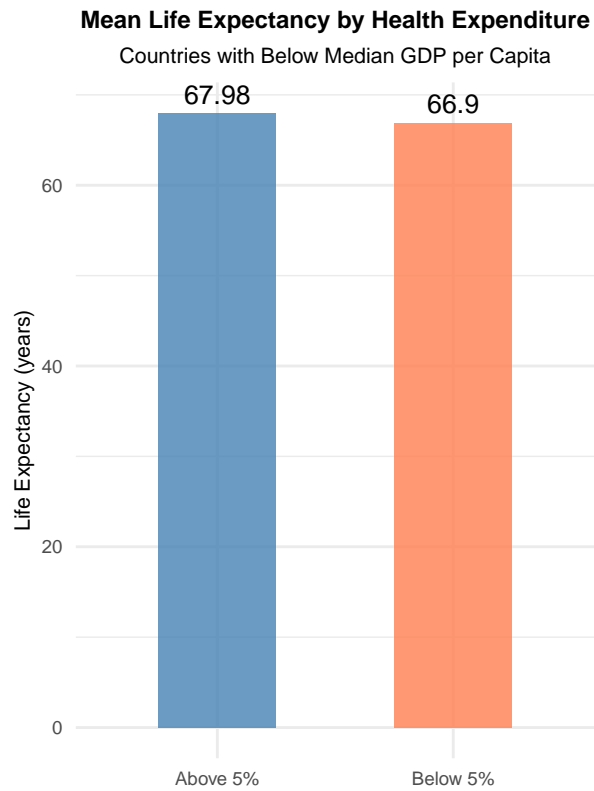
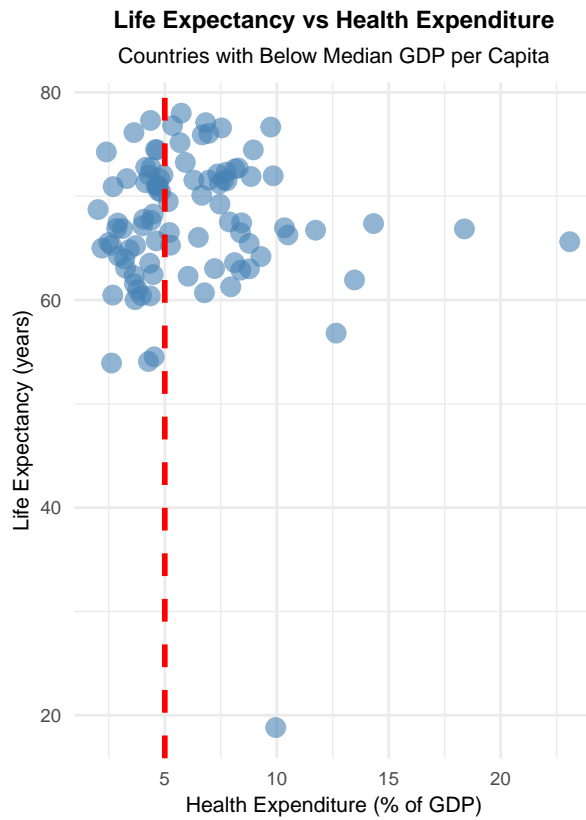


Table 3: Correlation Matrix of Health and Economic Indicators (2022)

	Health Exp.	Infant Mort.	GDP per Cap.	Life Exp.
Health Exp.	1.000	-0.162	0.149	0.153
Infant Mort.	-0.162	1.000	-0.412	-0.916
GDP per Cap.	0.149	-0.412	1.000	0.567

	Health Exp.	Infant Mort.	GDP per Cap.	Life Exp.
Life Exp.	0.153	-0.916	0.567	1.000



## Conclusions

### Research Question #1: Health Expenditure and Infant Mortality

Countries that spend more than the median on health (6.43% of GDP) have relatively lower infant mortality rates—15.48 deaths per 1,000 births versus 24.54 for below-median spenders. This 37% difference suggests that higher health spending indeed buys better child survival. However, the correlation is weak at -0.162, and the scatter plot shows a great deal of variation around the trend, including several low-spending countries with good outcomes and some high-spending countries with poor results. This indicates that it is not the level of spending that determines infant mortality but rather how the resources are used.

### Research Question #2: Health Expenditure and Life Expectancy

The above-median health spenders have a 3.2-year life expectancy advantage: 74.03 versus 70.79 years. Again, this correlation is weak at 0.153. The scatter plot shows country life expectancies from 50 to 80 years when health spending as a percent of GDP is the same. More interestingly, GDP per capita has a much stronger correlation with life expectancy (0.567) than health expenditure as a percent of GDP does, indicating overall economic success improves longevity more than healthcare spending as a percent of GDP.

### Research Question #3: Health Expenditure Among Lower-Income Countries

In countries below the median GDP per capita, those that spend above 5% on health only have a life expectancy advantage of 1.1 years (67.98 versus 66.9 years). The scatter plot shows most lower-income countries clustering between 60-75 years regardless of spending level, with no clear separation at thresholds around 5%. This minimal difference suggests that other factors beyond health expenditure percentage have a more influential and larger role.

### Story in the Data

This analysis suggests that higher health spending is associated with better outcomes, but the relationships are surprisingly weak. Spending more on health generally helps but is no guarantee of better results—many countries achieve strong health outcomes at modest spending levels while others spend heavily with disappointing returns. Economic development would appear more influential than health expenditure percentage, with stronger

associations of GDP per capita with both infant mortality and life expectancy. The very strong link between infant mortality and life expectancy suggests that systems effective at keeping children alive also support health throughout life. The critical insight is that how money is spent—through efficient systems, equitable distribution, and strategic allocation—matters more than how much is spent. For policymakers, adequate funding is a necessary but not sufficient condition; countries must also focus their attention on healthcare system efficiency, primary care, preventive services, and broader social determinants of health.

## Data Appendix

##	Country.Name	Country.Code	Health_Expenditure_2022
## 1	Afghanistan	AFG	23.088169
## 2	Albania	ALB	6.193681
## 3	Algeria	DZA	3.623043
## 4	Andorra	AND	7.536788
## 5	Angola	AGO	2.927376
## 6	Antigua and Barbuda	ATG	5.729994
## 7	Argentina	ARG	9.863053
## 8	Armenia	ARM	9.962129
## 9	Australia	AUS	9.926879
## 10	Austria	AUT	11.157293
## 11	Azerbaijan	AZE	3.978406
## 12	Bahamas, The	BHS	7.216920
## 13	Bahrain	BHR	3.837453
## 14	Bangladesh	BGD	2.390386
## 15	Barbados	BRB	6.293892
##	Infant_Mortality_2022	GDP_per_Capita_2022	Life_Expectancy_2022
## 1	52.0	357.2612	65.61700
## 2	8.3	6846.4267	78.76900
## 3	20.0	4961.5526	76.12900
## 4	2.6	42414.0590	84.01600
## 5	39.6	2929.6945	64.24600
## 6	7.5	20105.1989	77.48300
## 7	8.3	13935.6811	75.80600
## 8	9.3	6571.9745	74.76585
## 9	3.2	64997.0137	83.20000
## 10	2.7	52176.6649	81.29512
## 11	14.1	7770.5942	74.12500
## 12	11.6	34957.1613	74.49100
## 13	7.0	30470.5219	80.99200
## 14	24.4	2716.4859	74.26500
## 15	9.6	22164.0260	75.67600
##	Health_Exp_Group	GDP_Group	
## 1	Above Median	Below Median	



## 2	Below Median Above Median
## 3	Below Median Below Median
## 4	Above Median Above Median
## 5	Below Median Below Median
## 6	Below Median Above Median
## 7	Above Median Above Median
## 8	Above Median Above Median
## 9	Above Median Above Median
## 10	Above Median Above Median
## 11	Below Median Above Median
## 12	Above Median Above Median
## 13	Below Median Above Median
## 14	Below Median Below Median
## 15	Below Median Above Median

## References

- Anthropic. (2024). Claude (Version 3.5 and Sonnet 4.5) [Large language model]. <https://www.anthropic.com/claude/newline>
- World Bank. (2024). GDP per capita (current US\$) [Data set]. World Bank Open Data. <https://data.worldbank.org/indicator/NY.GDP.PCAP.CD/newline>
- World Bank. (2024). Life expectancy at birth, total (years) [Data set]. World Bank Open Data. <https://data.worldbank.org/indicator/SP.DYN.LE00.IN/newline>
- World Bank. (2024). Mortality rate, infant (per 1,000 live births) [Data set]. World Bank Open Data. <https://data.worldbank.org/indicator/SP.DYN.IMRT.IN/newline>
- World Bank. (2024). Current health expenditure (% of GDP) [Data set]. World Bank Open Data. <https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS>
- Link to GitHub Repository: <https://github.com/RisshiN24/3080-project>