

# DATASCI Assignment 5

Anny Huang

2026-02-25

## Load the dataset

```
library(tidyverse)

-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr     1.1.4      v readr     2.1.5
vforcats   1.0.0      v stringr   1.5.1
v ggplot2   4.0.0      v tibble    3.2.1
v lubridate 1.9.3      v tidyr    1.3.1
v purrr    1.0.2

-- Conflicts ----- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become non-conflicting
```

```
library(ggplot2)
wdi<-read.csv("/Users/annyhuang/Desktop/AH/QTM350/wdi.csv")
#view(wdi)
colnames(wdi)
```

```
[1] "country"                      "inflation_rate"
[3] "exports_gdp_share"            "gdp_growth_rate"
[5] "gdp_per_capita"               "adult_literacy_rate"
[7] "primary_school_enrolment_rate" "education_expenditure_gdp_share"
[9] "measles_immunisation_rate"    "health_expenditure_gdp_share"
[11] "income_inequality"             "unemployment_rate"
[13] "life_expectancy"               "total_population"
```

The dataset is derived from the World Development Indicators provided by the World Bank (World Bank 2022).

## Exploratory Data Analysis (Q3)

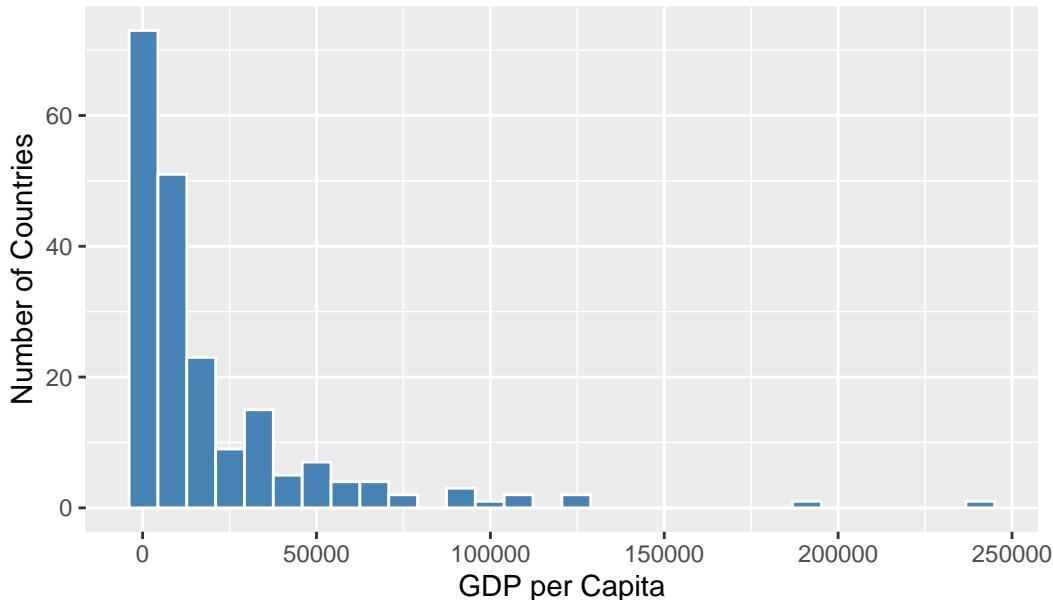
In this section, I plan to look at three key development indicators for 2022: GDP per capita, life expectancy, and unemployment rate.

### 1. GDP per Capita

GDP per capita measures the average economic output per person and is often used as an indicator of a country's standard of living.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
259	2571	7588	20346	25983	240862	14

Distribution of GDP per Capita (2022)

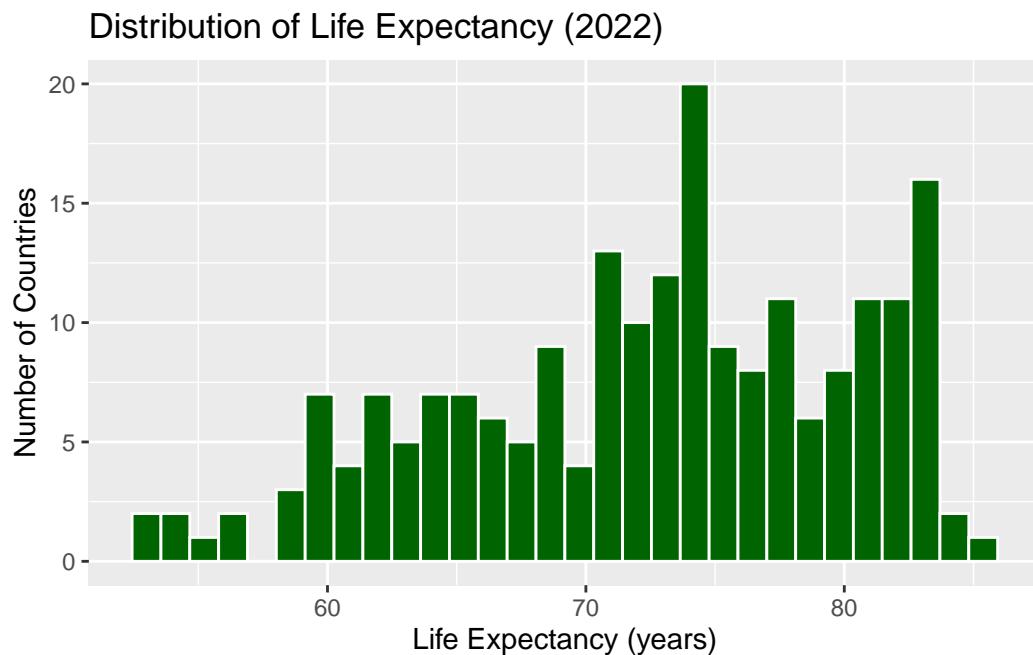


We can see that the distribution is strongly right-skewed, indicating that most countries have relatively low to middle income levels, while a small number of countries have very high GDP per capita.

### 2. Life Expectancy

Life expectancy represents the average number of years a newborn is expected to live under current mortality conditions.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
53.00	66.78	73.51	72.42	78.47	85.38	8



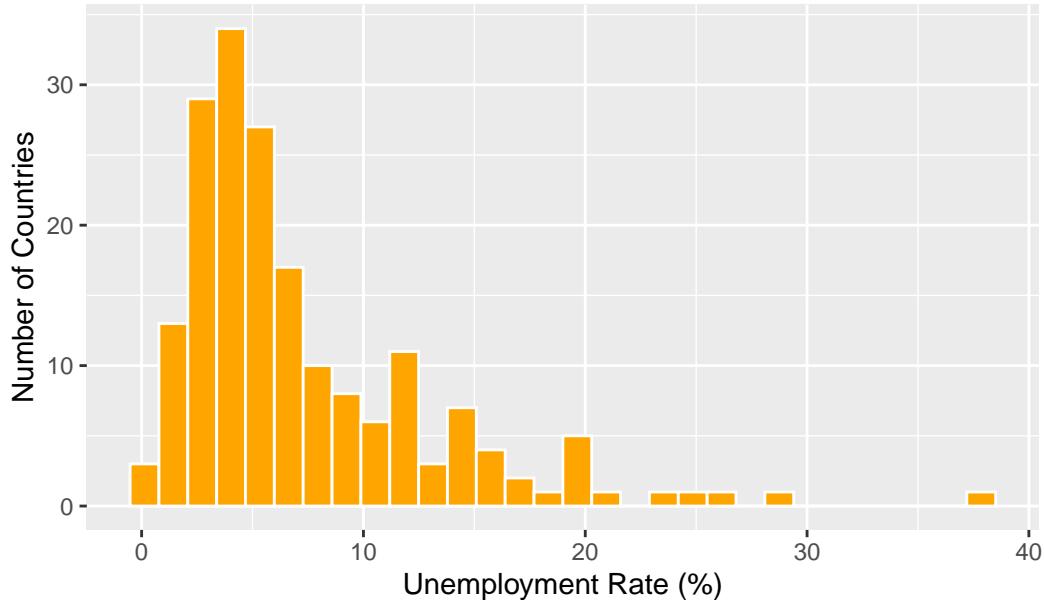
Life expectancy appears more normally distributed compared to GDP per capita. Most countries cluster between 65 and 80 years, with fewer countries experiencing extremely low life expectancy.

### 3. Unemployment Rate

The unemployment rate measures the percentage of the labor force that is without work but actively seeking employment.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
0.130	3.501	5.537	7.269	9.455	37.852	31

## Distribution of Unemployment Rate (2022)



Unemployment rates vary considerably across countries. While many countries report moderate unemployment levels, a subset experiences very high unemployment, suggesting structural labor market challenges.

## Relationship Between Economic and Health Outcomes (Q4)

Figure 1 shows that countries with higher GDP per capita tend to have higher life expectancy, with diminishing returns at the high-income end.

Figure 2 presents the top 10 countries ranked by GDP per capita in 2022. The distribution highlights substantial disparities in income levels across countries. The highest-income countries significantly outperform others, indicating strong economic productivity and high standards of living. The large gaps between ranks suggest that global income distribution remains highly unequal. This positive association between income and life expectancy has been widely documented in the literature (Preston 1975).

## Highlighting key statistics & Conclusions (Q5)

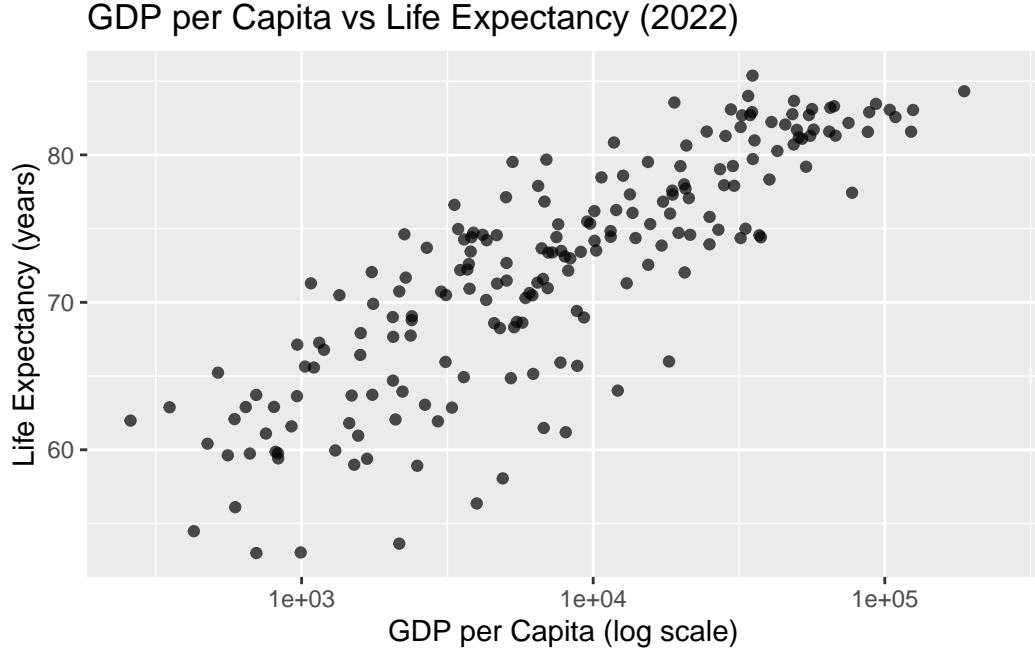


Figure 1: Relationship between GDP per capita and life expectancy in 2022. Data source: World Bank WDI (sample in wdi.csv).

Table 1: Key summary statistics for selected indicators (2022).

Table 1: Key summary statistics for selected indicators (2022). Data source: World Bank WDI (sample in wdi.csv).

GDP per capita (mean)	GDP per capita (median)	Life expectancy (mean)	Life expectancy (median)	Unemployment rate (mean)	Unemployment rate (median)
20345.71	7587.59	72.42	73.51	7.27	5.54

Table 1 summarises key descriptive statistics for selected development indicators in 2022. The mean and median GDP per capita differ substantially, indicating a right-skewed distribution driven by a small number of high-income countries. In contrast, life expectancy shows less dispersion between the mean and median, suggesting a more concentrated global distribution. Unemployment rates display moderate variation, reflecting differences in labour market conditions across countries.

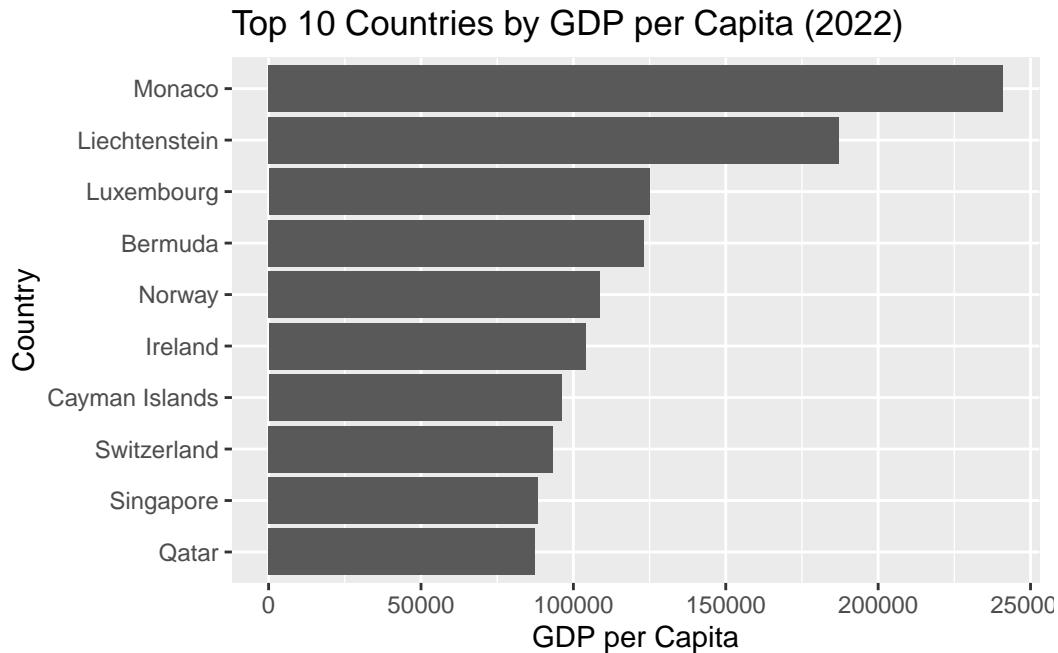


Figure 2: Top 10 countries by GDP per capita in 2022. Data source: World Bank WDI (sample in wdi.csv).

## Conclusion

This analysis of the 2022 World Development Indicators highlights substantial variation in economic and social outcomes across countries. GDP per capita exhibits strong inequality, with a small number of high-income countries significantly outperforming others. Life expectancy appears more evenly distributed, though it remains positively associated with economic prosperity. Unemployment rates vary across countries, reflecting structural and economic differences in labour markets.

Overall, the findings suggest that economic development remains closely linked to health outcomes, while disparities in income and employment conditions continue to shape global inequality patterns.

## References

- Preston, Samuel H. 1975. “The Changing Relation Between Mortality and Level of Economic Development.” *Population Studies* 29 (2): 231–48.  
 World Bank. 2022. “World Development Indicators.” <https://databank.worldbank.org/source/world-development-indicators>.