

A comprehensive exploration of the dataset:

# Health and Demographics

through interactive storytelling...

Data Visualization – Master in Data Sciences and Business Analytics – April 2024  
Tableau Titans | Irene Sunny, Maria Hallak, Anna Parshikova, Rhianne Gonsalves



01.

# Introduction



## What does the project visualize?

- ✓ life expectancy disparities globally
- ✓ socioeconomic ties to longevity

## For whom it might be useful?

- ✓ policymakers
- ✓ health organizations
- ✓ researchers
- ✓ media
- ✓ anyone concerned with global health

## Why is it important?

- ✓ to **uncover insights** that can inform strategies to improve global health outcomes
- ✓ to **allocate resources** effectively, prioritize areas of need, and implement evidence-based strategies to enhance health outcomes
- ✓ to **raise awareness** about global health issues and promote understanding of the factors influencing health outcomes



**02.**

# **Project Description**

## Motivation

- ✓ addressing the need for data-driven insights
- ✓ exploring the socioeconomic factors that can improve long-term well-being, population health and life expectancy
- ✓ promoting collaboration between users through our tool for positive outcomes



## Tool

- ✓ utilizing a dataset compiled over 15 years from a wide range of countries worldwide
- ✓ offering an interactive dashboard with user-friendly interface to explore the data from different angles
- ✓ enabling users to make informed, evidence-based decisions
- ✓ facilitating the prioritization of efforts to address global health challenges

## Research Questions

01

How did life expectancy vary across countries and across years?

02

What is the relation among all the health-related factors and the life expectancy?

03

How does GDP, schooling and health expenditures relate to life expectancy?

**Data source:** Used data source contains of a global health and demographic data from 2000 to 2015, compiled from reputable international sources such as WHO, World Bank, and United Nations databases

Dataset [here](#)  
Github Repository [here](#)  
Dashboard [here](#)





**03.**

# **Tool and Framework**



# Leveraging Python and Streamlit for Interactive Data Visualization

- ✓ Python: A versatile programming language with strong support for data manipulation, statistical analysis, and visualization.
- ✓ Key Libraries Used:
  - Pandas: For data manipulation and analysis.
  - Plotly: Advanced interactive graphing library.
  - Seaborn/Matplotlib: For high-quality statistical graphics.
  - Numpy: For numerical operations.
- ✓ Streamlit: An open-source app framework specifically for Machine Learning and Data Science teams to create beautiful, performant apps quickly with pure Python.

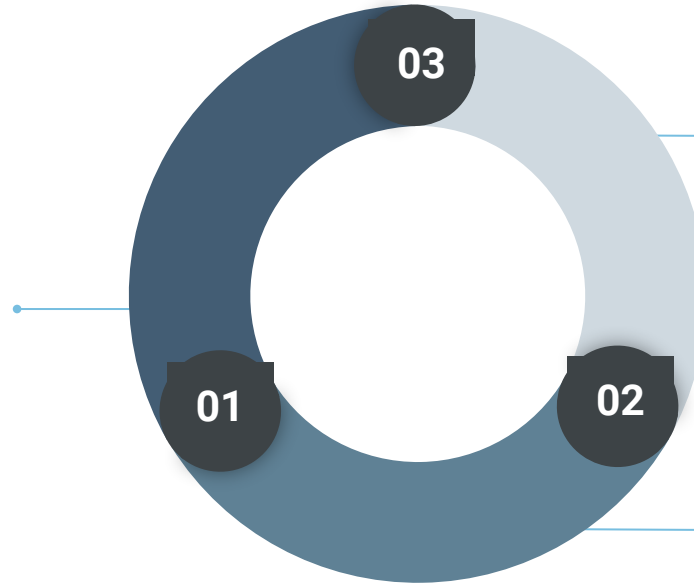


## Why Streamlit over other frameworks?

### Rapid

#### Development:

Allows for quick turnaround from concept to prototype to deployment of data applications.



#### Interactivity:

Easily integrates sliders, buttons, maps, and other widgets to make the dashboard interactive without the need for callbacks or backend engine

#### Simplicity:

Requires minimal setup to create interactive dashboards compared to other frameworks, making it accessible to analysts not deeply versed in web development.



04.

# Project Demo

## The Data

Period: 2004 to 2014

Countries: 133 (19 developed and 114 developing)

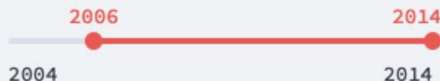
## Two Views

Select View

- ☒ Global Overview
- ☐ Deep Dive : Socio Economic Influencers

Both views feature a year range slider for analysis

Select Year Range



## Global Overview

- ✓ summarized information
- ✓ map visualization of the variation of life expectancy
- ✓ life expectancy distribution through a box plot
- ✓ the top- and bottom-ranking countries

Select Status

Developing x Developed x

Select Continent

Asia x Europe x Africa x

South America x Australia x

North America x Oceania x

## Socio-Economic Influencers

- ✓ a correlation matrix between different health metrics for developed and developing countries
- ✓ a linear regression between life expectancy and the factor

Select a factor to compare with Life Expectancy

- ☒ Income Composition Of Resources (how
- ☐ Average Years of Schooling
- ☐ Health Expenditure (% of GDP)

## Interactive Features

- ✓ radio buttons
- ✓ range slider
- ✓ tags selection
- ✓ interrelated visualizations based on selections and filters
- ✓ zoom and pan/move tools on the map
- ✓ interactive tooltips are shown when users hover over marks

Select View



Global Overview



Deep Dive : Socio Economic Influencers

## Explanation of what the tool shows: Global Overview

Summarized information based on users selection and filters:

- ✓ number of countries selected
- ✓ average life expectancy
- ✓ average GDP per capita
- ✓ health expenditures share of GDP

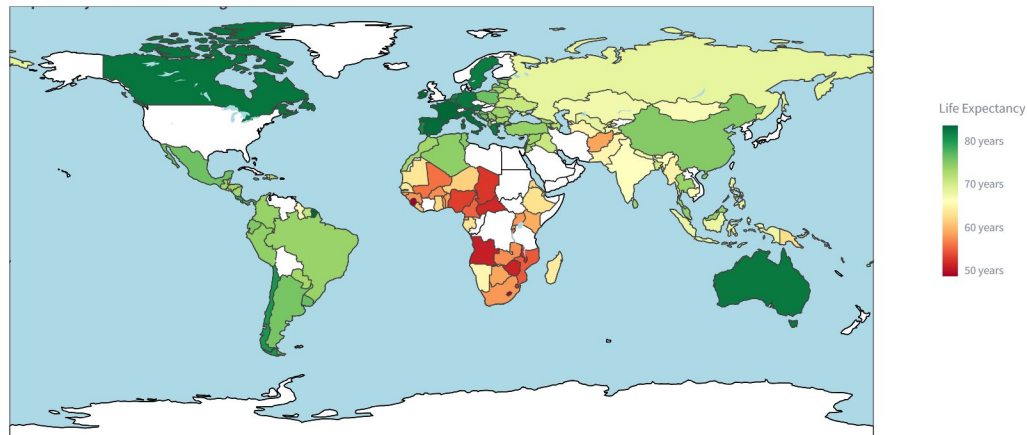
Map visualization:

- ✓ consistent with selection and filters
- ✓ colour legend as a visual aid
- ✓ country information display when users hover over marks

### Visual encodings used:

- ✓ spatial marks (position, shape and area of each country)
- ✓ continuous colour scale as a visual channel

Average Life Expectancy Across Countries



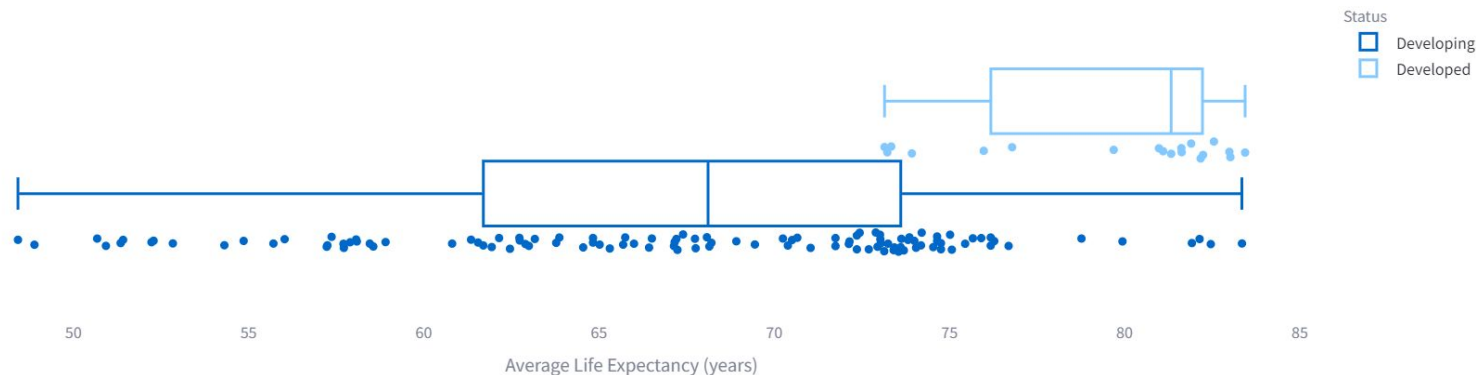
Select View

☒ Global Overview

☐ Deep Dive : Socio Economic Influencers

## Explanation of what the tool shows: Global Overview

### Distribution of Life Expectancy Across Countries



#### Box plot:

- ✓ consistent with selection and filters
- ✓ distribution of life expectancy
- ✓ country information display, median and quartiles when users hover over marks

#### Visual encodings used:

- ✓ points as marks
- ✓ position and size as visual channels

Select View



Global Overview



Deep Dive : Socio Economic Influencers

## Explanation of what the tool shows: Global Overview

### Best and Worst Countries by Average Life Expectancy

Contrasting presentation of five countries with the highest and lowest life expectancies, offering insights into their health landscapes.

Country	Continent	Status	Population (Million)	Life Expectancy (Years)	Adult Mortality (probability of death b/w ages 15 and 60 per 1,000 people)	Infant Deaths (per 1,000 live births)	Under Age Five Deaths (per 1,000 live births)	HIV/AIDS (% of population)
Ireland	Europe	Developed	3.77	83.44	55.40	0.00	0.00	0.10
France	Europe	Developing	27.08	83.35	86.91	3.00	3.27	0.10
Italy	Europe	Developed	39.05	83.02	46.82	2.00	2.09	0.10
Spain	Europe	Developed	26.33	82.99	56.91	1.55	1.91	0.10
Austria	Europe	Developed	7.66	82.55	63.82	0.00	0.00	0.10
Malawi	Africa	Developing	6.26	51.35	414.45	34.55	53.91	14.78
Zimbabwe	Africa	Developing	6.84	50.93	447.55	27.55	42.00	18.54
Angola	Africa	Developing	10.10	50.68	362.75	76.62	118.75	2.45
Lesotho	Africa	Developing	1.17	48.89	538.00	4.36	6.09	20.99
Sierra Leone	Africa	Developing	2.03	48.42	335.25	26.25	38.62	1.38

#### Table showing rankings:

- ✓ consistent with selection and filters
- ✓ additional country-specific information
- ✓ country information display when users hover over marks

#### Visual encodings used:

- ✓ positions as the visual channel
- ✓ colour code for top-ranking vs low-ranking as a visual channel



Select View

☒ Global Overview

☐ Deep Dive : Socio Economic Influencers

Line charts:

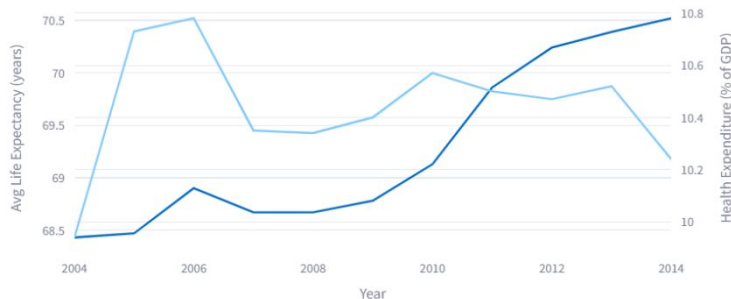
- ✓ consistent with selection and filters
- ✓ different trend views
- ✓ legends as a visual aid
- ✓ information display when users hover over marks

Visual encodings used:

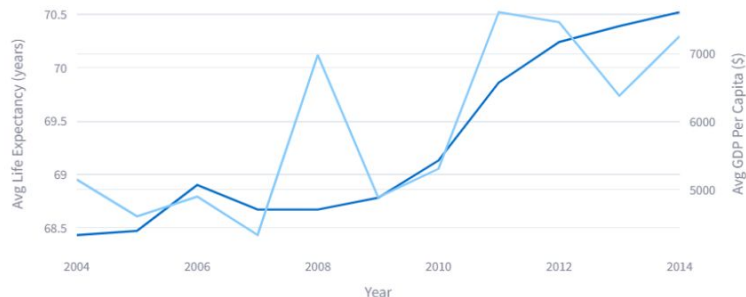
- ✓ points and lines as marks
- ✓ positions as a visual channel
- ✓ colour code as a visual channel

## Explanation of what the tool shows: Global Overview

Life Expectancy vs Health Expenditure



Life Expectancy vs GDP per Capita



Select View

☐ Global Overview

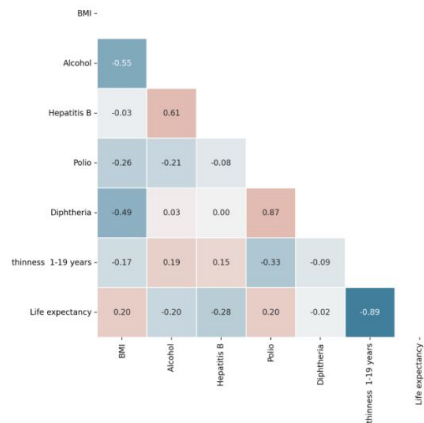
☒ Deep Dive : Socio Economic Influencers

## Explanation of what the tool shows: Socio-Economic Influencers

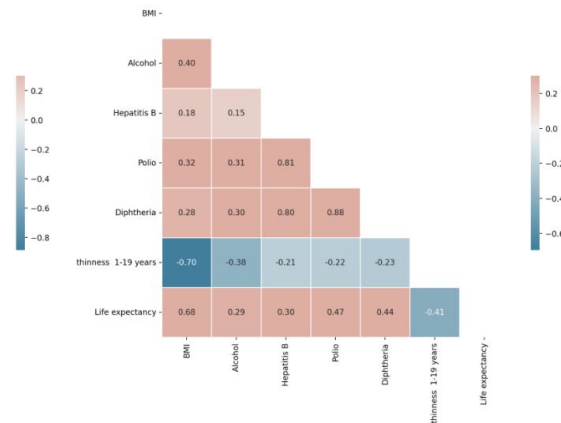
Correlation matrices:

- ✓ consistent with selection and filters
- ✓ different views for developed and developing countries
- ✓ colour legend as a visual aid

Developed Countries



Developing Countries



Visual encodings used:

- ✓ areas as marks
- ✓ positions as a visual channel
- ✓ colour as a visual channel

Select View

☐ Global Overview

☒ Deep Dive : Socio Economic Influencers

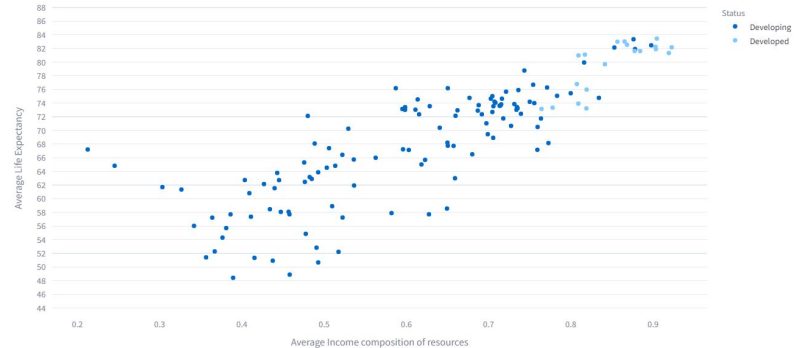
Correlation matrices:

- ✓ consistent with selection and filters
- ✓ different views for different socioeconomic factors
- ✓ colour legend as a visual aid
- ✓ information display when users hover over marks

## Explanation of what the tool shows: Socio-Economic Influencers

- ☒ Income Composition Of Resources (how well a country is utilising its resources, higher the better)
- ☐ Average Years of Schooling
- ☐ Health Expenditure (% of GDP)

Relationship Between Income composition of resources and Life Expectancy



Visual encodings used:

- ✓ points as marks
- ✓ positions as a visual channel
- ✓ colour as a visual channel



05.

# Project Walkthrough

Dashboard [here](#)



**06.**

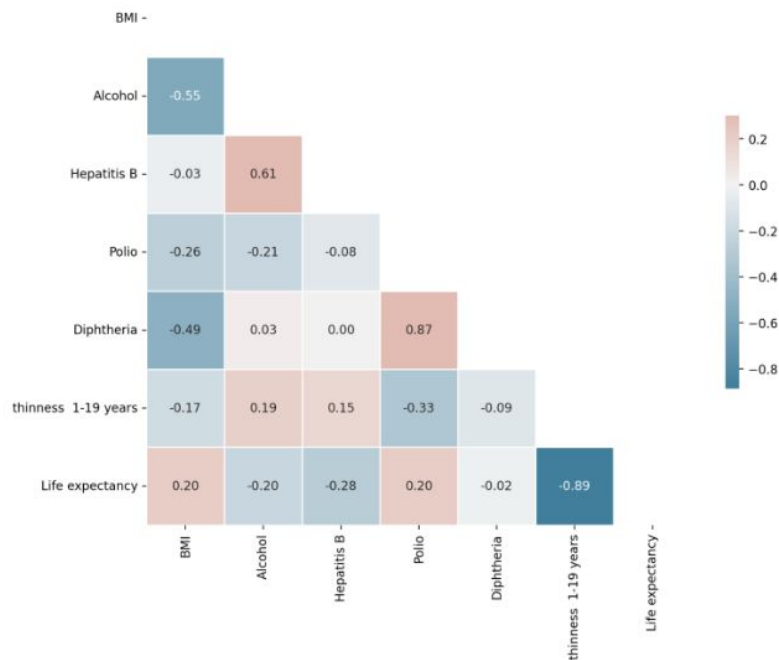
# **Key Insights**

## 01. How did life expectancy vary across countries and across years?

- We can utilize the heat map graph displaying life expectancy for developed and developing countries over the years.
- The map colors indicate life expectancy with greens showing higher and reds showing lower life expectancy. Regions like North America, Europe, Australia, and some parts of South America and Asia have greener shades, indicating higher life expectancy. This can be associated with better healthcare systems, higher standards of living, and better health awareness.
- In contrast, many parts of Africa and some regions in Asia are colored orange to red, indicating lower life expectancy.
- In addition, box-plot to quantify our conclusions by providing statistical distribution bas, thus allowing us to corroborate our observations from the heat map and gain a deeper understanding of the disparities in life expectancy across different socio-economic contexts through the analysis of the median, quartiles, and extremes depicted in the box plot.

## 02. What is the relation between health-related factors and the life expectancy?

### Developed Countries

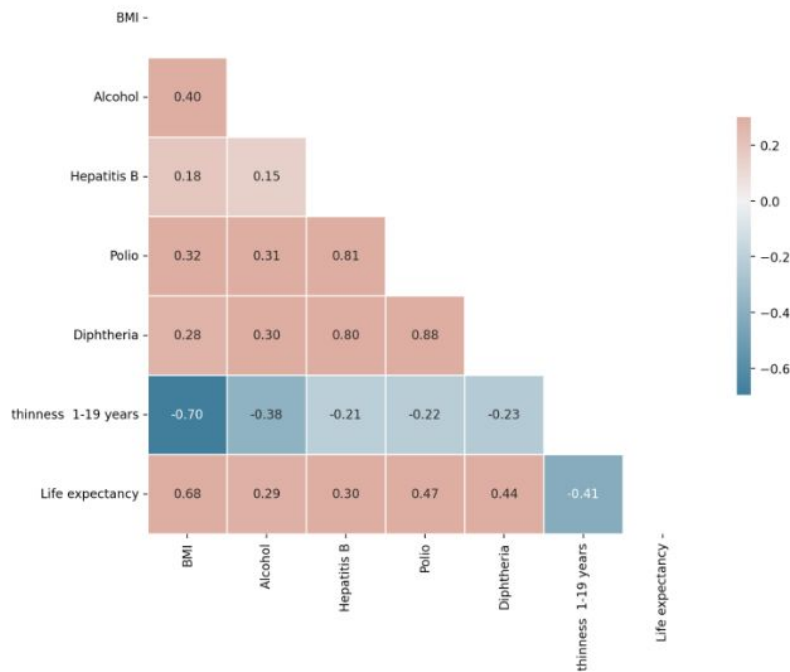


### High Negative Correlation between Thinness 1-19 years and Life Expectancy:

A strong negative correlation exists between thinness in 1-19-year-olds and life expectancy (-0.89), indicating that malnutrition or underweight issues among the youth significantly impact life expectancy in developing countries.



## Developing Countries



**Negative Correlation between Thinness 1-19 years and Life Expectancy:** A strong negative correlation exists between thinness in 1-19-year-olds and life expectancy (-0.48), indicating that malnutrition or underweight issues among the youth significantly impact life expectancy in developing countries.

The positive impact of vaccines suggests a continued investment in healthcare infrastructure is crucial

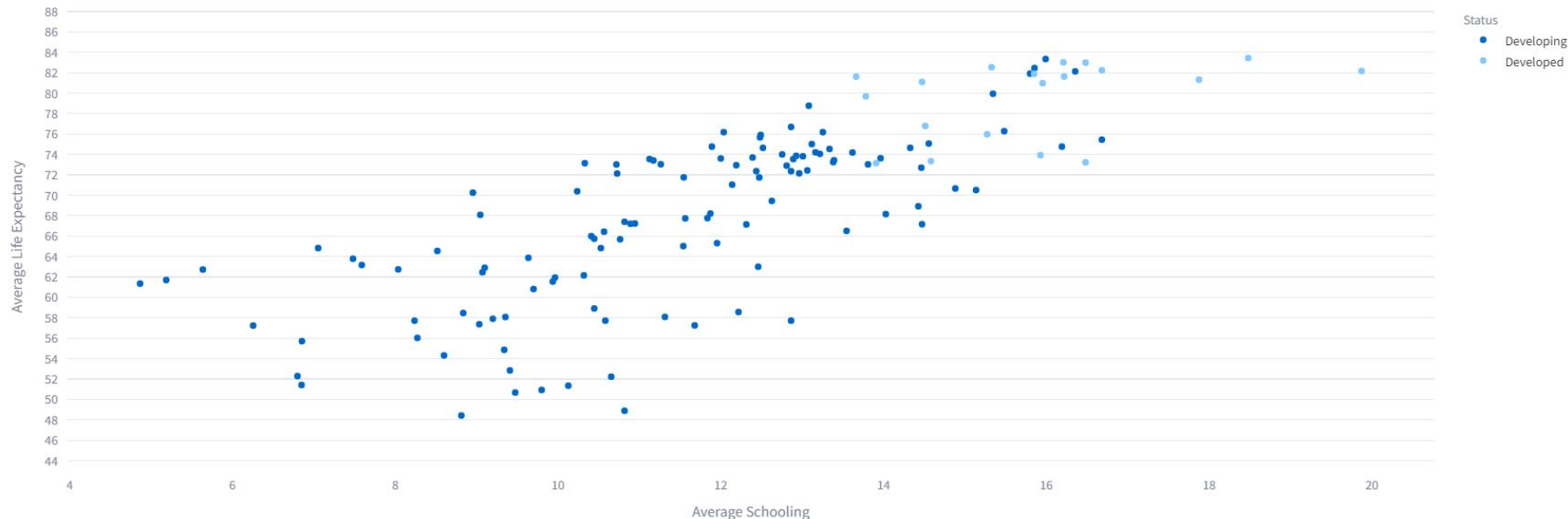
### **03. How are key socio economic factors related to life expectancy?**

To answer this question we can refer to the scatter plot and two line graphs depicting the necessary variables. From these visualizations, we can infer several insights:

1. Higher GDP per capita tends to correlate with greater life expectancy, likely due to improved access to healthcare, better living conditions, and overall higher quality of life.
2. Education level, as represented by average years of schooling, shows a strong association with life expectancy.
3. Investments in healthcare systems, as indicated by health expenditure as a percentage of GDP, also correlate positively with life expectancy.

# Impact of Education on Life Expectancy

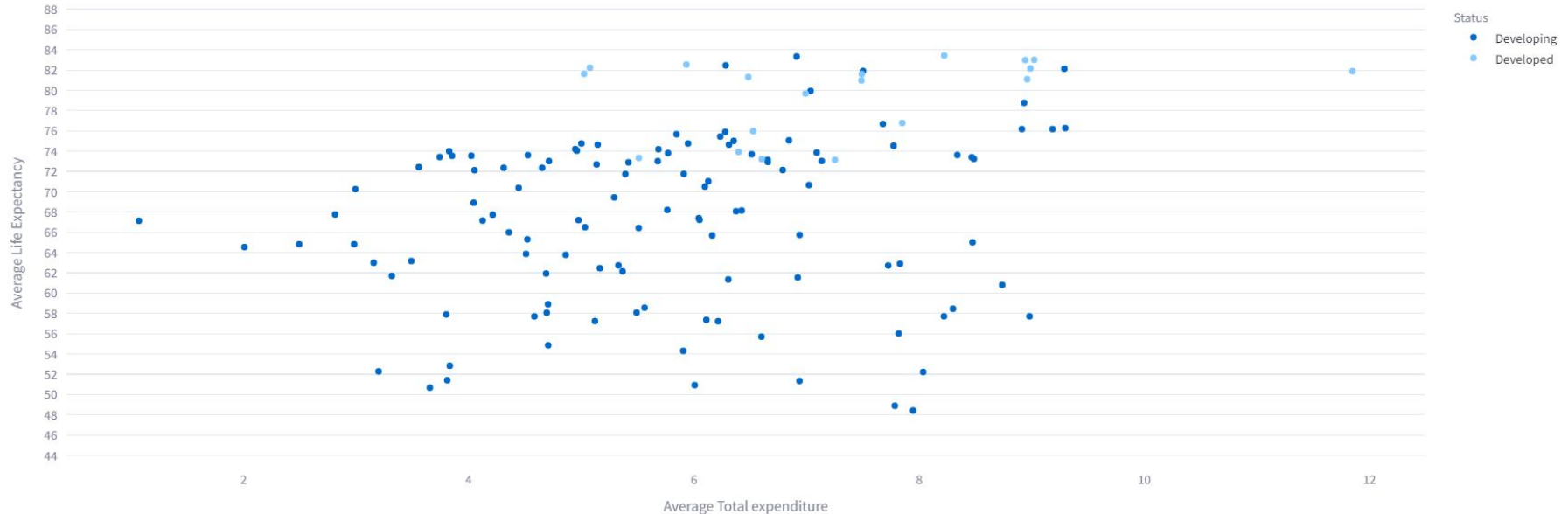
Relationship Between Schooling and Life Expectancy



An important takeaway is that developed countries and the developing countries with better life expectancy than the rest predominantly have above-average schooling years, suggesting a strong emphasis on education within these nations

# Impact of Health Expenditure on Life Expectancy

Relationship Between Total expenditure and Life Expectancy



Against intuition, not all developed countries and the developing countries with better life expectancy have an above average health expenditure. This observation may suggest that developed countries have optimized their health spending or achieved efficiencies in their healthcare systems that allow them to spend less while still maintaining high life expectancies

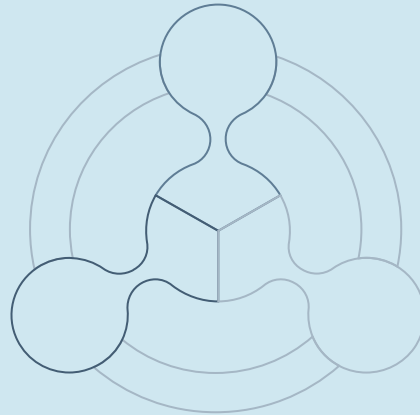


**07.**

# **Limitations**

## Limitations of the Dashboard

**Incomplete Data:** The data does not have data for every country for every year in the dataset. Thus preventing us from analysing all the countries



### **Feasibility to handle large data**

While streamlit was the best solution for the dataset we used. It takes a long time to load when there is a larger dataset and there are complex data transformations.

**Lack of Causality:** As correlation does not imply causation, while the visualizations may show associations between factors and life expectancy, they do not establish a causal relationship.

# Team Work

Dashboard

Powerpoint Presentation

Submission

Framework : Irene | Maria

EDA & Data Processing : Irene  
| Rhianne

Development : Irene

Deployment : Irene

Quality Check : Rhianne

Feedback Loop : All

Narrative : Anna | Maria

PPT Creation : Anna | Maria

Quality Check : All

Video Presentation : All

GitHub and Files Readiness :  
Anna | Rhianne

**Every team member contributed to the best of their abilities and played by their strengths to ensure timely submission.**





**Thank you!**

# Deliverables

- 1) Deploying the project on the web - youtube link [here](#)
- 2) Publishing your source code to GitHub - link [here](#)
- 3) Interactive Dashboard [here](#)
- 4) Dataset [here](#)