Phase 3 Submission Report: Data Visualization Project

DATASET LINK:

- https://www.kaggle.com/datasets/mutindafestus/world-statistics-dataset-from-world-bank?resource=download
- https://www.kaggle.com/datasets/umairnasir14/population-growth-of-top-25-countries

DEMO VIDEO LINK:

- **https://drive.google.com/file/d/1amGiTIROzS2xJxjemn17c9aTPFZ3knPf/view?usp=sharing**

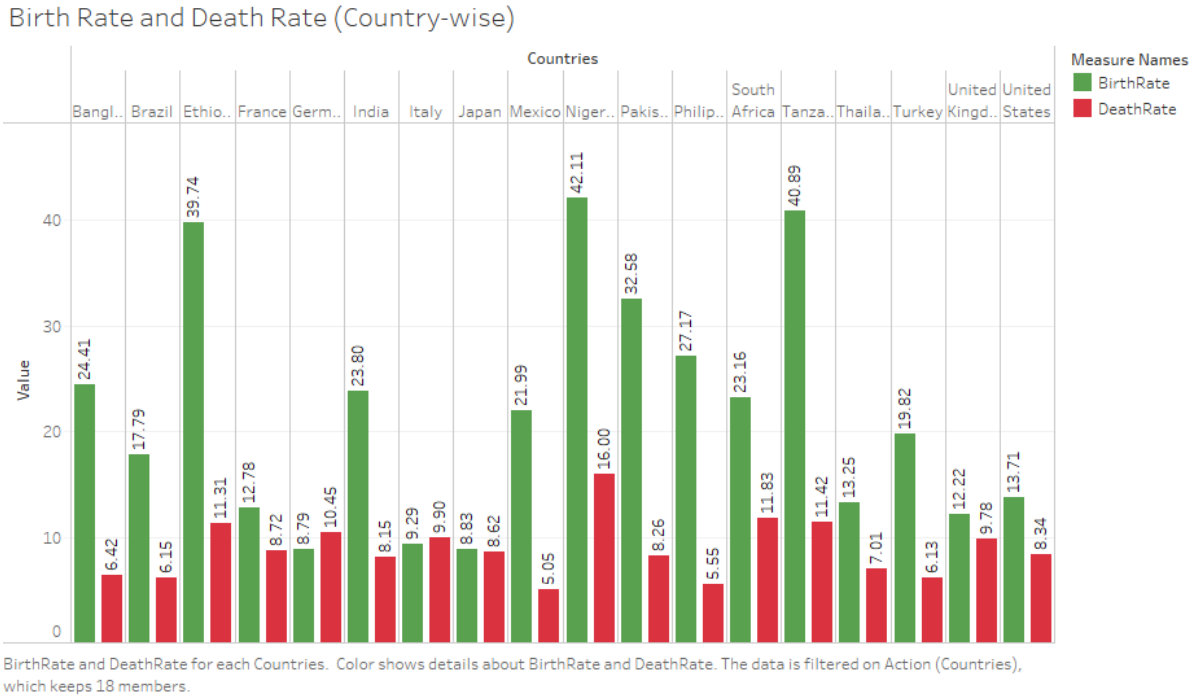
DATASET DESCRIPTION:

- The dataset used in this project was created by combining two sources: the World Bank's World Statistics data and a dataset ranking countries by population. The final dataset was created by merging the two data sources and thoroughly cleaning and pre-processing it using Tableau Prep Builder
- The goal of this dataset is to visualize various general statistics, including birth rate, death rate, population growth, agricultural land use, and more, over the period of 1995-2015. Post-preprocessing, the data is structured in CSV format to allow easy access and manipulation.

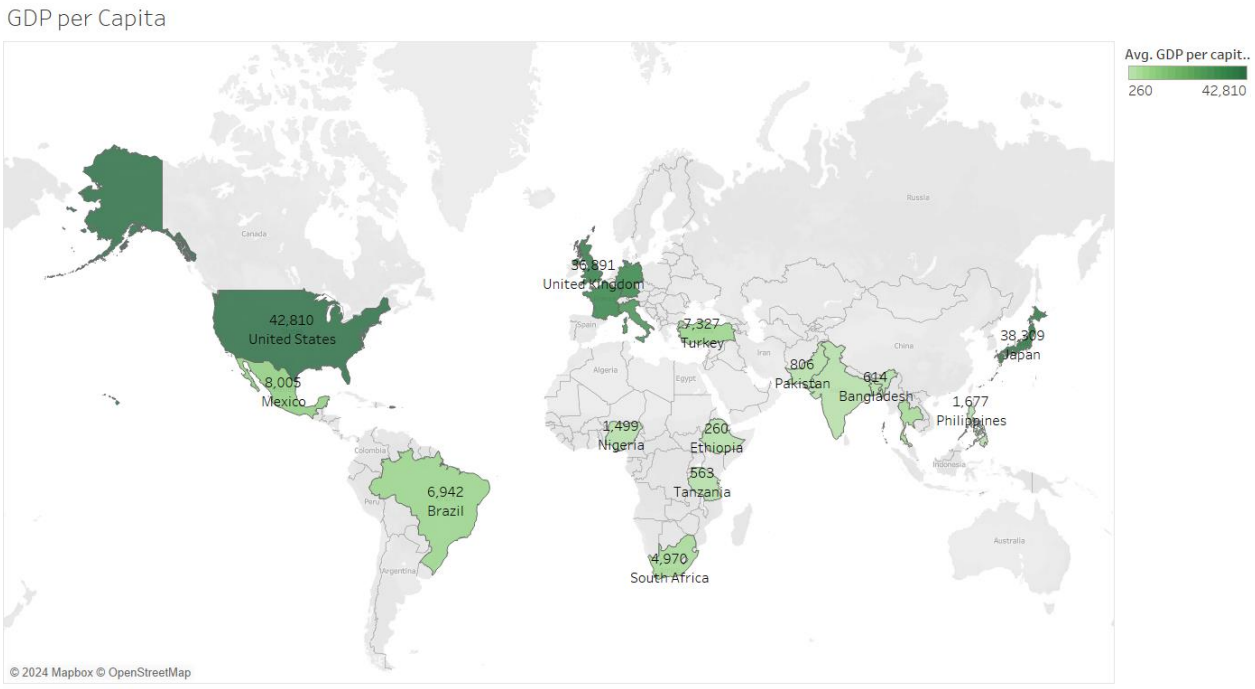
SYSTEM ARCHITECTURE AND DESCRIPTION:

Data from 18 countries has been analyzed, covering statistics over a 20-year period from 1995 to 2015

- **Birth Rate and Death Rate (Country-wise):** A clustered bar chart illustrates the relationship between the average birth rate and death rate, using green for birth rate and red for death rate to enhance pre-attentive processing. To further aid quick comprehension, the corresponding values are displayed above each bar for easy reading.

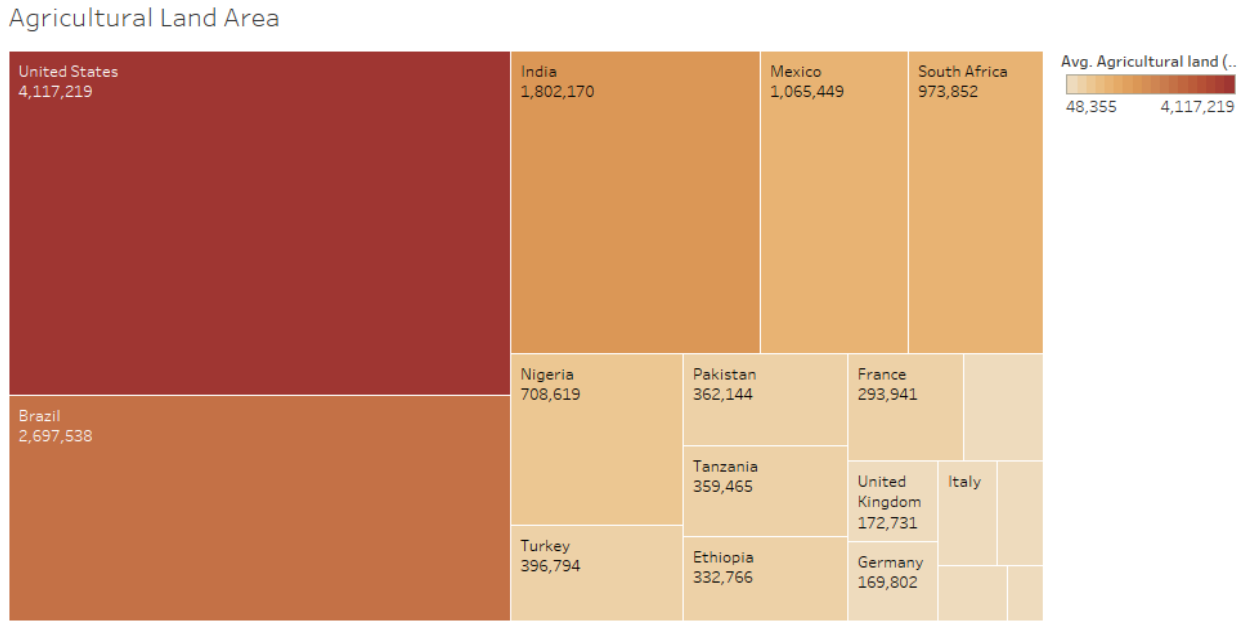


- **GDP per Capita:** A world map visualizes GDP per capita by country, using varying shades of the same color to enable instant comparison of GDP intensity across countries. For enhanced pre-attentive processing, the labels display both the values and corresponding country names.



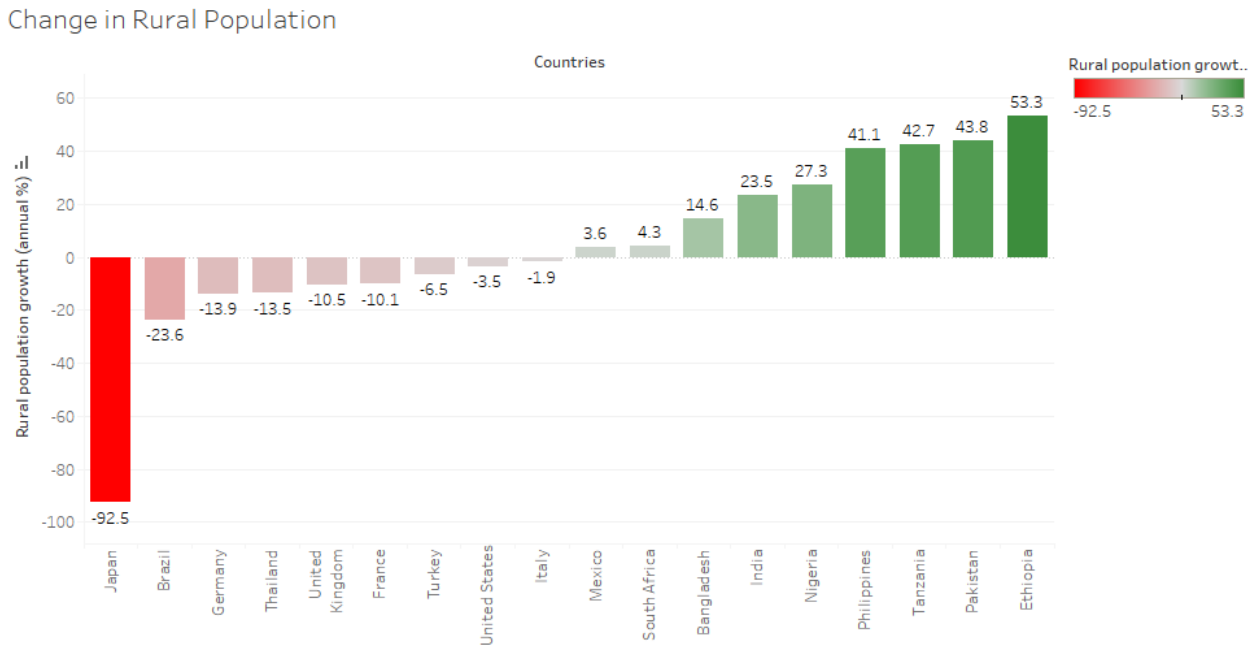
Map based on Longitude (generated) and Latitude (generated). Color shows average of GDP per capita (current US$). The marks are labeled by average of GDP per capita (current US$) and Countries. Details are shown for Countries.

- **Agricultural Land Use:** A tree-map with heat mapping is used to display the average agricultural land area for each country, with labels showing the country name and its agricultural land area. The tree-map enables instant pre-attentive processing, as the size of each block visually represents the area.



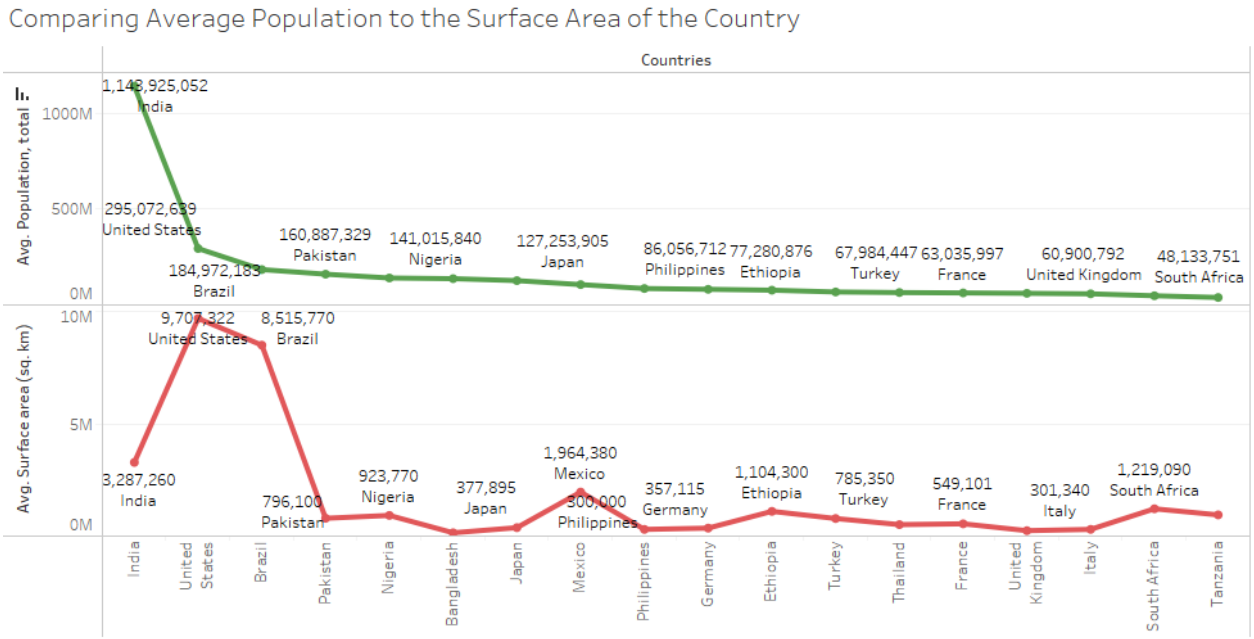
Countries and average of Agricultural land (sq. km). Color shows average of Agricultural land (sq. km). Size shows average of Agricultural land (sq. km). The marks are labeled by Countries and average of Agricultural land (sq. km). The data is filtered on Action (Countries), which keeps 18 members. The view is filtered on Countries, which keeps 18 of 18 members.

- **Change in Rural Population:** A bar graph is used to illustrate the rate of change in the rural population, with values represented as both positive and negative. To enhance immediate pre-attentive processing, the values are also labeled on the graph.



Sum of Rural population growth (annual %) for each Countries. Color shows sum of Rural population growth (annual %). The marks are labeled by sum of Rural population growth (annual %). The data is filtered on Action (Countries), which keeps 18 members.

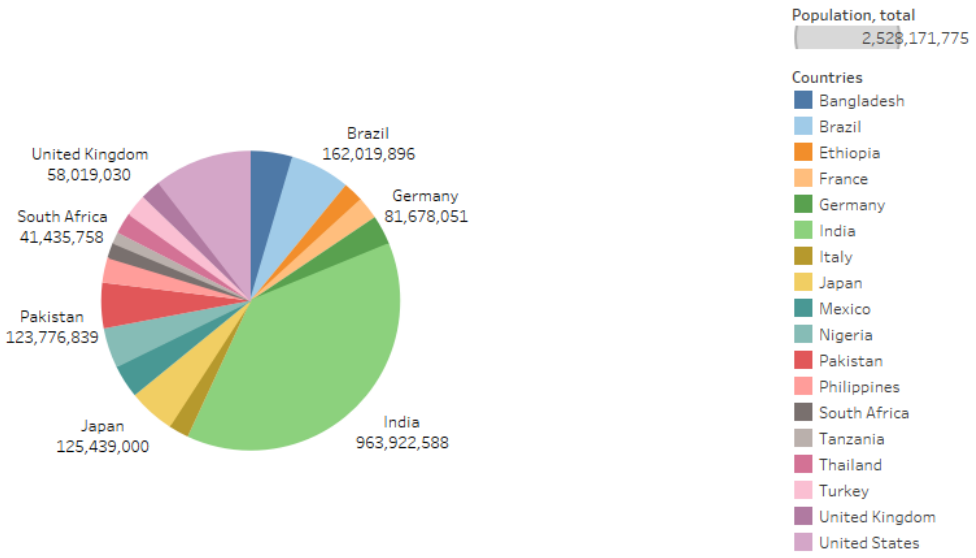
- **Comparing Average Population to the Surface Area:** Two line-graphs are used to compare the total average population with the surface area of each country. Together, they highlight which countries have a higher population density by examining both graphs separately. For enhanced pre-attentive processing, the countries and their respective values are labeled.



The trends of average of Population, total and average of Surface area (sq. km) for Countries. For pane Average of Surface area (sq. km): The marks are labeled by average of Surface area (sq. km) and Countries. For pane Average of Population, total: The marks are labeled by average of Population, total and Countries. The data is filtered on Action (Countries), which keeps 18 members. The view is filtered on Countries, which keeps 18 of 18 members.

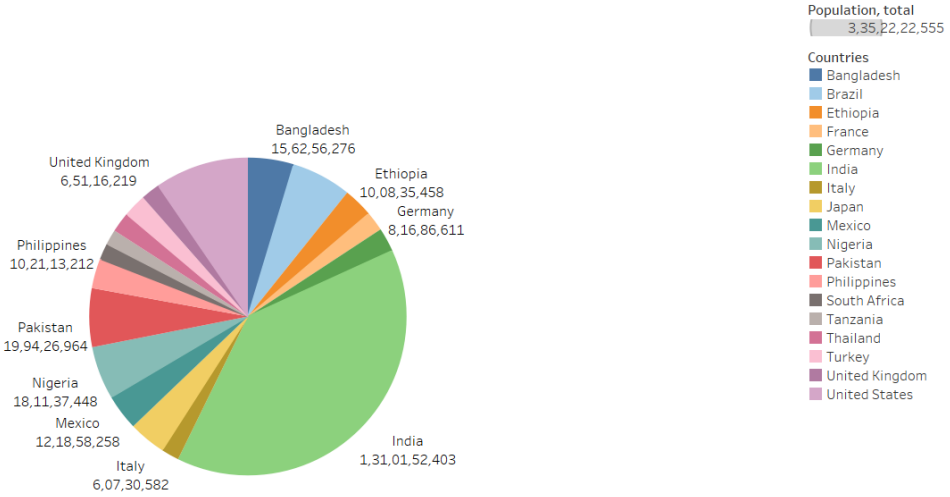
- **Total Population 1995 and 2015:** Two pie charts are used to display a single variable—the total population of each country in 1995 and 2015. The size of each pie segment clearly highlights which countries have the largest populations. For instant comprehension, the countries and their respective population values are labeled.

Total Population 1995



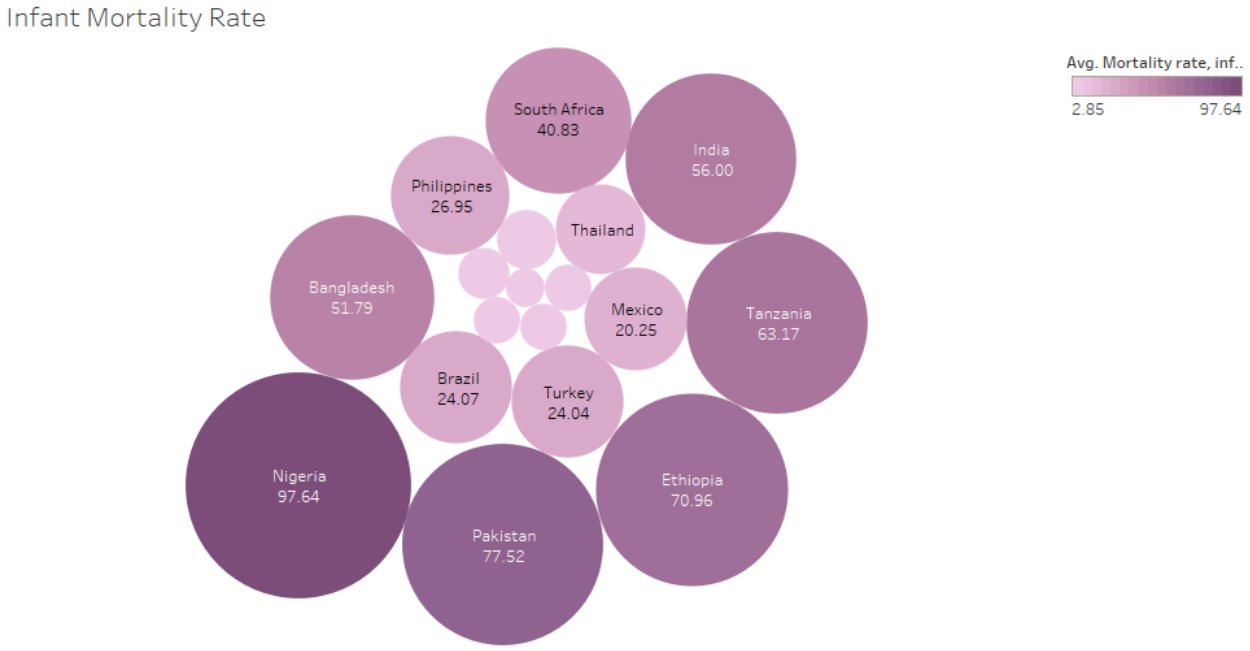
Countries and Population, total. Color shows details about Countries. Size shows Population, total. The marks are labeled by Countries and Population, total. The data is filtered on Year, which ranges from 1995 to 1995.

Total Population 2015



Countries and Population, total. Color shows details about Countries. Size shows Population, total. The marks are labeled by Countries and Population, total. The data is filtered on Year, which ranges from 2015 to 2015.

- **Infant Mortality Rate:** A bubble chart with varying color intensity is used to visualize the infant mortality rate. Both the size of the bubbles and the color shading enhance pre-attentive processing for quick interpretation. For easy comprehension, the mortality rates and country names are also labeled.

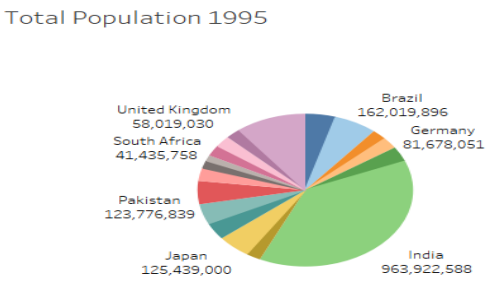


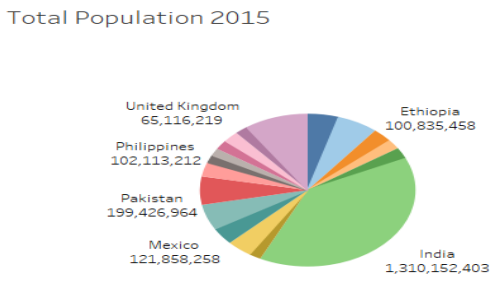
Countries and average of Mortality rate, infant (per 1,000 live births). Color shows average of Mortality rate, infant (per 1,000 live births). Size shows average of Mortality rate, infant (per 1,000 live births). The marks are labeled by Countries and average of Mortality rate, infant (per 1,000 live births). The data is filtered on Action (Countries), which keeps 18 members.

DASHBOARDS

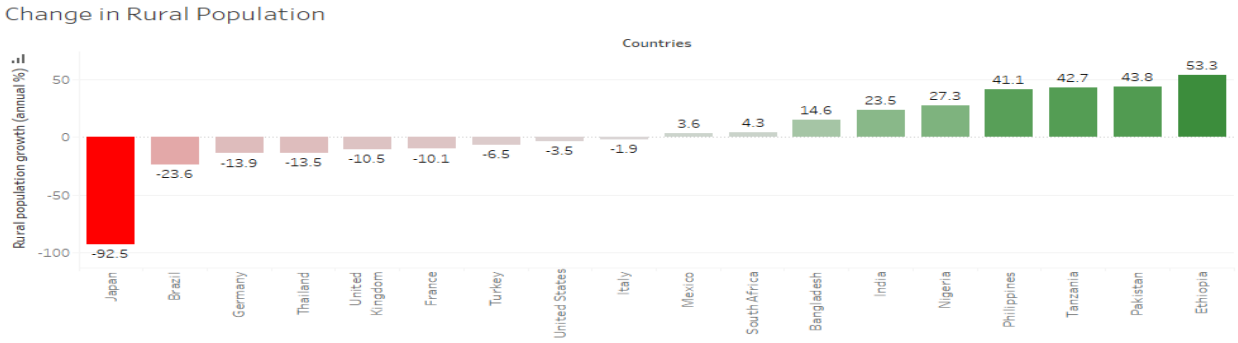
- **Population Dashboard**

This dashboard links the two variables—total population and the change in rural population—for each country.





This interactive Tableau dashboard presents



two Key visualizations related to population statistics:

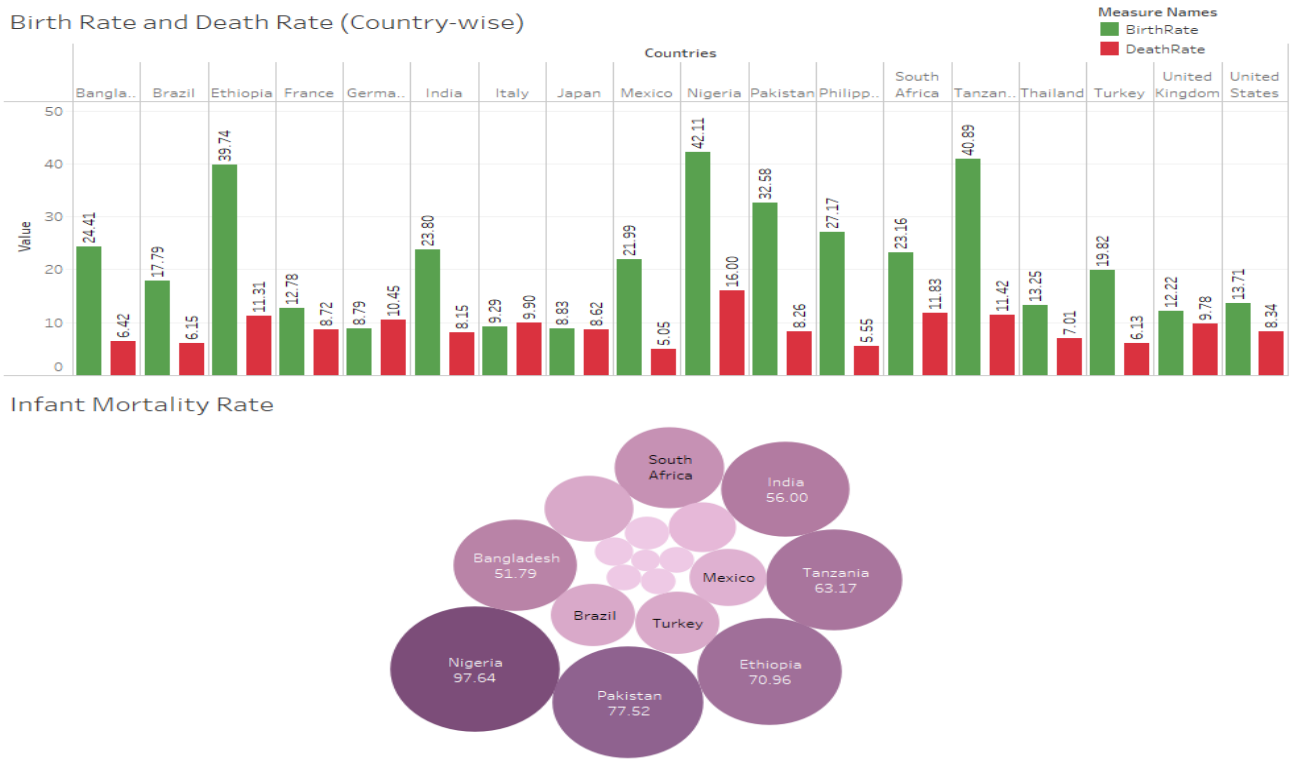
1. **Total Population (1995 vs 2015):**
    - **Pie charts** show the total population of various countries for the years 1995 and 2015.
    - The charts highlight countries such as India, Brazil, Germany, and Pakistan.
    - The size of each slice corresponds to the population of each country, with labels showing the exact population figures.
    - By comparing the two pie charts, we can visually assess changes in the population distribution over the 20-year period.
2. **Change in Rural Population:**
    - This **horizontal bar graph** displays the change in rural population growth (in annual percentage) for each country from 1995 to 2015.
    - The bars are color-coded, with negative values shown in red, indicating a decrease in rural population, and positive values shown in green, indicating an increase.
    - The countries are arranged from the largest decrease to the largest increase in rural population growth.
    - Japan shows the largest decrease (-92.5%), while Ethiopia shows the largest increase (53.3%).
    - Labels on each bar provide the exact percentage values for quick comprehension.

These visualizations are interactive, allowing users to explore the data for specific countries and analyze how their populations and rural growth rates have evolved over time.

Use cases:

- **Demographic and Population Analysis**: Compare population growth trends across countries to understand global changes and identify areas with rapid growth for resource allocation.
- **Urbanization and Rural Migration Studies**: Evaluate shifts from rural to urban areas and analyze rural population declines to support urban planning and agriculture-dependent economies.
- **Economic and Policy Planning**: Use population data to assess market potential, forecast economic trends, and develop policies for rapidly growing or declining populations.
- **Sustainability and Infrastructure Development**: Manage environmental resources and plan infrastructure in response to urbanization and rural population growth.

❑ Birth, Death Mortality Rate Dashboard



This dashboard connects the variables of birth rate, death rate, and infant mortality rate for each country.

This interactive Tableau dashboard consists of two key visualizations focusing on demographic indicators for various countries:

1. Birth Rate and Death Rate (Country-wise):

- This **bar chart** shows the **birth rate** (green bars) and **death rate** (red bars) for various countries.
- The y-axis represents the value of the birth rate and death rate, while the x-axis shows the countries.
- Some countries like **Ethiopia** and **Nigeria** have high birth rates, while countries like **Germany** and **United Kingdom** have relatively low birth rates.
- The **red bars** (death rate) are often lower compared to the **green bars** (birth rate), showing a clear difference in birth and death rates.
- The chart allows users to compare the **birth rate** and **death rate** for multiple countries at once.

2. Infant Mortality Rate:

- This **bubble chart** illustrates the **infant mortality rate** for several countries.
- Each bubble represents a country, with the **size** of the bubble corresponding to the mortality rate, and the **color intensity** indicates the level of the mortality rate (darker shades for higher rates).
- Countries like **Nigeria**, **Pakistan**, and **Ethiopia** have the highest infant mortality rates, as seen from the large dark purple bubbles.
- Smaller bubbles like those for **South Africa** and **India** reflect relatively lower infant mortality rates, but they still remain significant compared to others.

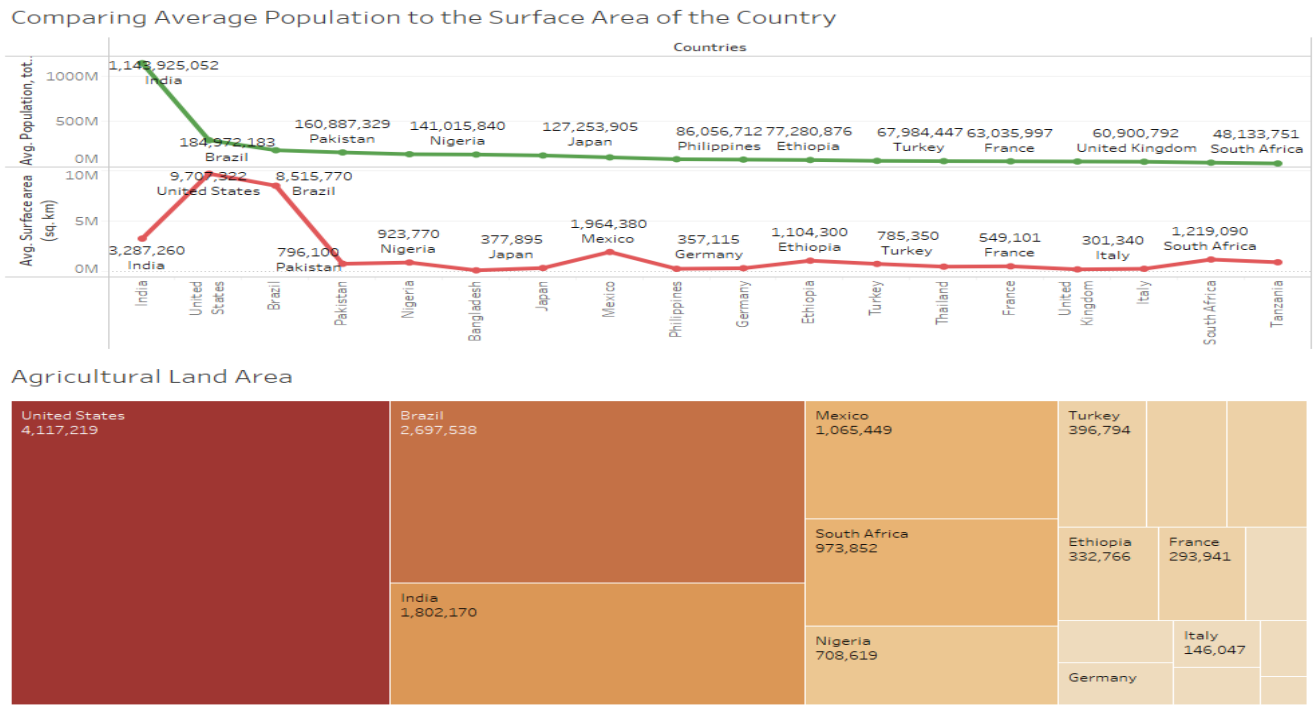
- This visualization allows for easy comparison of infant mortality rates by country based on both size and color.

These visualizations are interactive, enabling users to filter and explore specific countries, making it easy to understand global trends related to birth rates, death rates, and infant mortality.

Use cases:

- **Policy Makers**: Prioritize countries with high infant mortality for international aid or healthcare improvements.
- **Healthcare Analysis**: Identify regions with discrepancies in birth and death rates to allocate resources efficiently.
- **Population Studies**: Analyze trends in fertility and mortality to understand demographic transitions.

❑ Surface Area Dashboard



This dashboard connects the variables of average population, surface area, and agricultural land use.

1. Line Chart: Comparing Population and Surface Area

- **Purpose**: The line chart compares each country's **average population** (green line) against its **average surface area** (red line) to highlight disparities between population density and land availability.
- **Insights**:
    - **India** has the highest population (~1.15 billion) but a relatively smaller surface area (3.29 million sq. km), indicating high population density.
    - **The United States** has a large surface area (~9.7 million sq. km) but a smaller population compared to India.

- Countries like **Japan** and **Germany** show smaller surface areas and moderate populations, indicating higher urban density.
- **Brazil** and **Nigeria** balance relatively high populations with large surface areas.

2. Tree Map: Agricultural Land Area

- **Purpose**: This visual compares the **total agricultural land area** of each country, with the size of the rectangles representing the land area values.
- **Insights**:
  - **The United States** has the largest agricultural land area (~4.1 million sq. km), highlighting its agricultural capacity.
  - **Brazil** (~2.7 million sq. km) and **India** (~1.8 million sq. km) also have significant agricultural land resources.
  - Smaller agricultural land areas are seen in **Germany**, **France**, and **Italy**, represented by smaller rectangles.

Use cases:

- **Policy Analysis**: Understand population density relative to land and agricultural resources.
- **Economic Insights**: Highlight countries with high agricultural potential.
- **Resource Allocation**: Inform decisions regarding food security or urban planning

CONCLUSION:

- ❑ The visualizations present various statistics across different countries, comparing key factors such as GDP, birth rate, death rate, and more using diverse graph types. These visual representations simplify the understanding of multiple variables and country-specific data. The use of color and varying graph sizes enhances pre-attentive processing, making the information easy to grasp at a glance.
- ❑ The interactive dashboard allows users to view the values of all variables simultaneously for any selected country.
- ❑ Our primary goal with this visualization is to simplify the interpretation of these variables and provide insight into how countries are progressing across various fields.

GITHUB REPOSITORY LINK:

- ❑ https://github.com/GJaideepReddy/Data-Visualization-Project-World-Statistics-Data