**GDP Data Analysis**

**1. Project Overview:**

This project aims to conduct an in-depth analysis of global data concerning population, GDP, country-specific characteristics, and health indicators sourced from various datasets. We will be using Power BI to connect, clean, transform, analyze, and visualize the data to derive valuable insights. The final deliverable will be an interactive Power BI dashboard that effectively communicates key trends and findings from the data.

**2. Data Sources:**

**SQL Database (CountriesWorld):** This dataset contains critical economic, demographic, and health-related data across countries. The key fields provided are:

* Country Name and Country Code
* Region
* Population and Area (sq. miles)
* Population Density and Coastline Ratio
* Infant Mortality Rate (per 1000 births)
* GDP per capita
* Literacy Rate and Mobile Phones (per 1000 people)
* Agricultural, Industrial, and Service contributions to GDP
* Birthrate, Deathrate, and Climate data

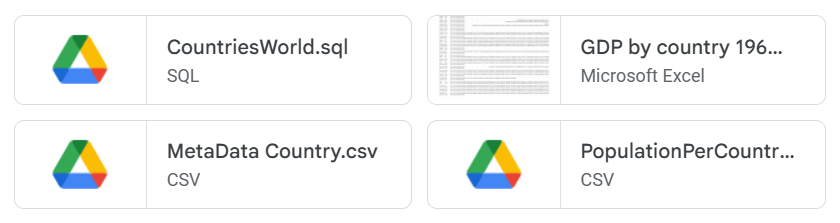
**Excel Files:**

1. **PopulationPerCountry**:
   * Contains data for population trends (1960-2017) by country for various indicators,such
   * as population size and growth.
2. **MetaData**:
   * Includes information like country code, region, income group, and special notes for each country.
3. **GDP by Country (1960-2016)**:
   * Provides GDP data across different countries for the years 1960 to 2016.

**3. Objective:**

The primary goal of this project is to develop a Power BI report and dashboard that extracts useful insights from the provided datasets. By connecting the SQL database (CountriesWorld) with Excel files (Population, MetaData, and GDP), the project will deliver a comprehensive analysis that highlights global economic, demographic, and health trends. Ultimately, the goal is to visualize these insights in a way that is both interactive and accessible.

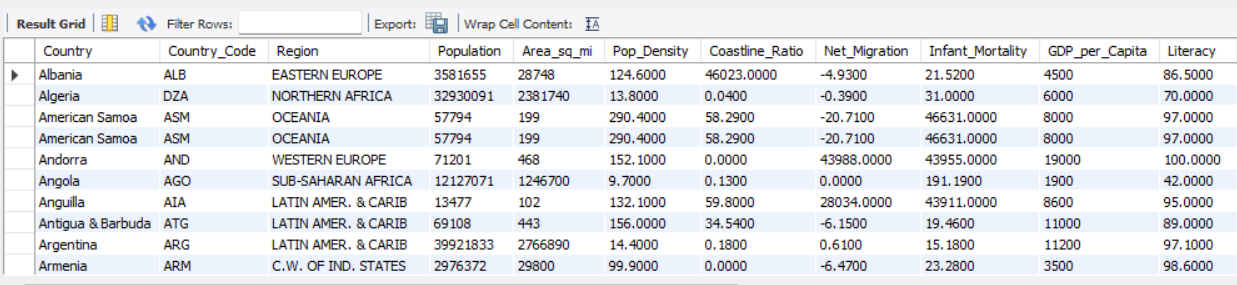
Given data

****

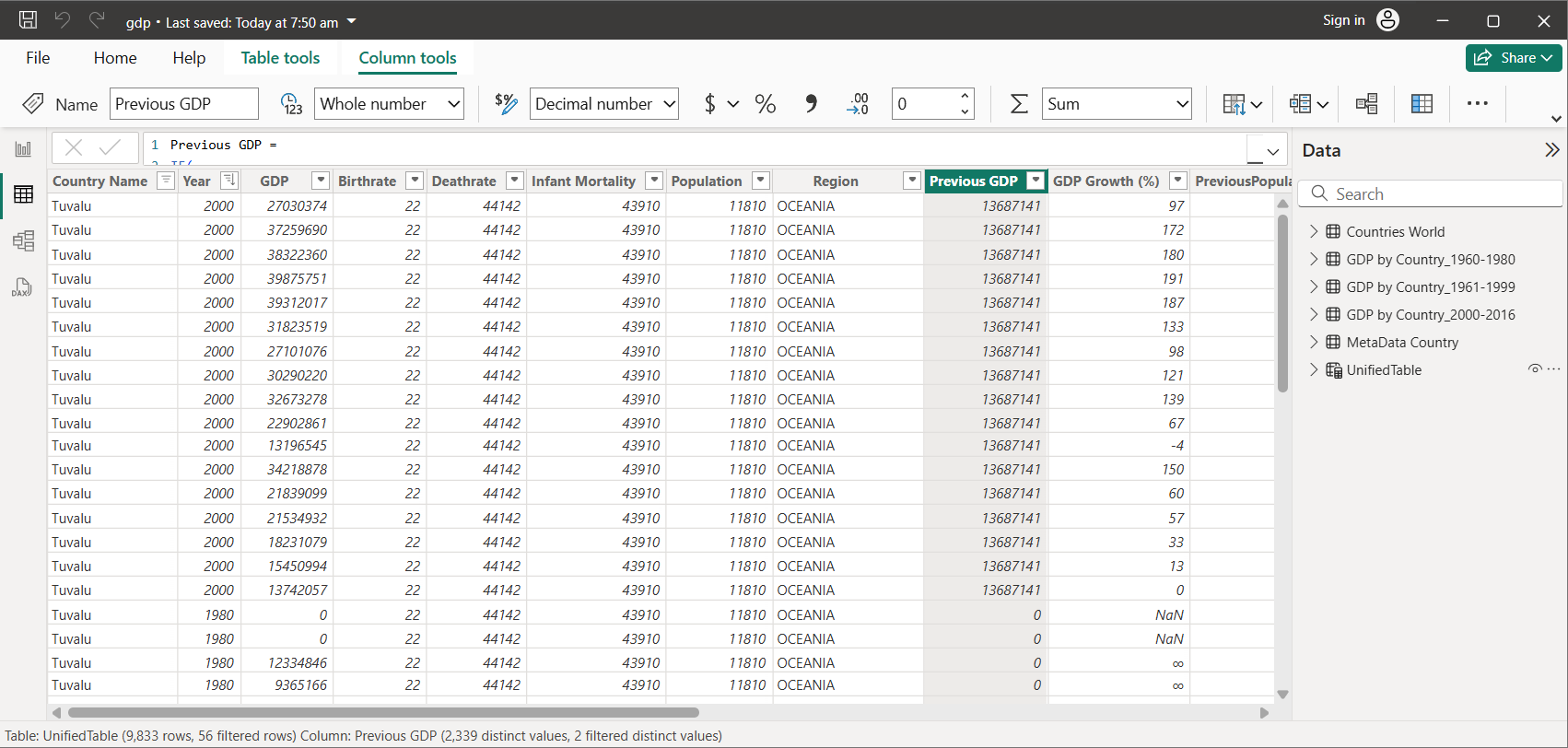
1. **Key Tasks and Analysis:**

**Data Import, Cleaning & Transformation:**

* Import all relevant datasets into Power BI.
* Cleanse and format the data to ensure consistency and usability by resolving missing values, data type transformations, deduplication.
* Link the datasets through common fields, such as Country Name or Country Code.

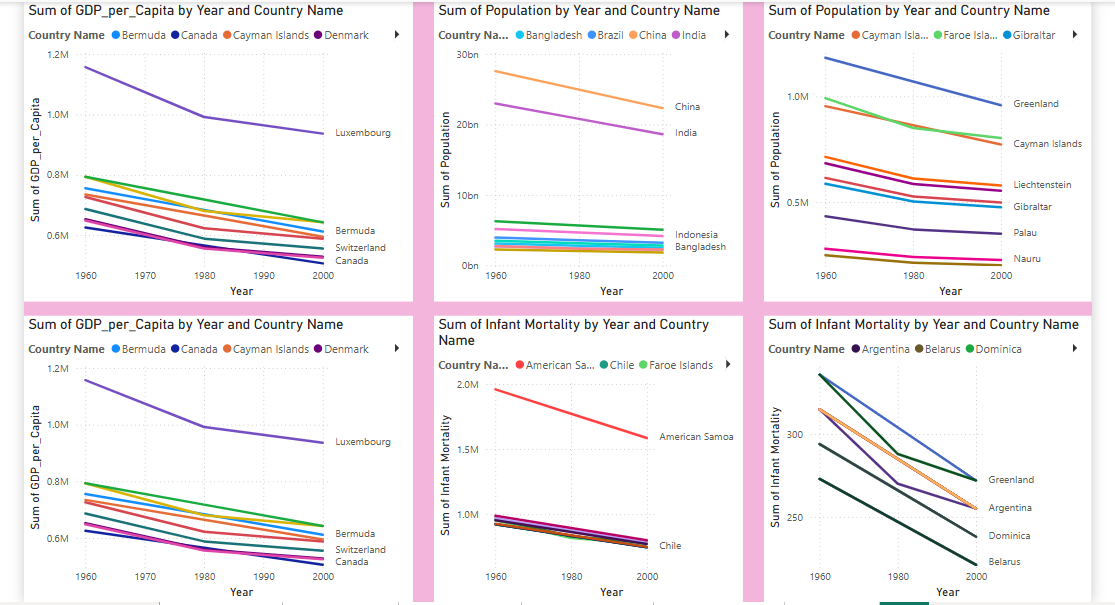


Imported SQL file to Power BI. Similarly csv and excel files were imported into Power BI



**Descriptive Statistical Analysis:**



* **Population Trends:**
  + Computed key statistics (mean, median, standard deviation) for population data across countries.
  + **CHINA** POPULATION 1313973713
  + AREA 9596960
  + POPULATION-DENSITY 136.9
  + INFANT\_MORTALITY 24.18
  + BIRTHRATE 13.25
  + DEATHRATE 35582
  + **ST PIERRE & MIQUELON 7026**
  + AREA 242
  + POPULATION DENSITY 29
  + INFANT MORTALITY 19906
  + BIRTHRATE 13.52
  + DEATHRATE 30468
  + **AMERICAN SAMOA** 57794
  + AREA 199
  + POPULATION DENSITY 290.4
* **GDP Analysis:**
  + Calculated average GDP per capita for countries and regions.
  + Identified countries with the highest and lowest GDP.
* **Infant Mortality:**
  + Assessed the distribution of infant mortality rates.
  + Calculated key statistics like range, mean, and standard deviation.
* **Literacy Rates:**
  + Analyzed average literacy rates for countries and regions.
* **Growth Trends (1960-2017):**
  + Investigated growth trends for variables like population size, GDP, and infant mortality over time.
  + A table was created to do growth calculations
  + 

**DAX Calculations:**

Calculated growth rates for population and GDP.

UnifiedTable =

UNION(

    SELECTCOLUMNS(

        'GDP by Country\_1960-1980',

        "Country Name", 'GDP by Country\_1960-1980'[Country Name],

        "Year", 1960,"GDP\_per\_Capita",'GDP by Country\_1960-1980'[GDP\_per\_Capita],

        "GDP", 'GDP by Country\_1960-1980'[GDP I],"Birthrate", 'GDP by Country\_1960-1980'[Birthrate],

        "Deathrate", 'GDP by Country\_1960-1980'[Deathrate],

        "Infant Mortality", 'GDP by Country\_1960-1980'[Infant\_Mortality],

        "Population", 'GDP by Country\_1960-1980'[Population],

        "Region", 'GDP by Country\_1960-1980'[Region]

    ),

    SELECTCOLUMNS(

        'GDP by Country\_1961-1999',

        "Country Name", 'GDP by Country\_1961-1999'[Country Name],

        "Year", 1980,"GDP\_per\_Capita",'GDP by Country\_1961-1999'[GDP\_per\_Capita],

        "GDP", 'GDP by Country\_1961-1999'[GDP II],

        "Birthrate", 'GDP by Country\_1961-1999'[Birthrate],

        "Deathrate", 'GDP by Country\_1961-1999'[Deathrate],

        "Infant Mortality", 'GDP by Country\_1961-1999'[Infant\_Mortality],

        "Population", 'GDP by Country\_1961-1999'[Population],

        "Region", 'GDP by Country\_1961-1999'[Region]

    ),

    SELECTCOLUMNS(

        'GDP by Country\_2000-2016',

        "Country Name", 'GDP by Country\_2000-2016'[Country Name],

        "Year", 2000,"GDP\_per\_Capita",'GDP by Country\_2000-2016'[GDP\_per\_Capita],

        "GDP", 'GDP by Country\_2000-2016'[GDPIII],"Birthrate", 'GDP by Country\_2000-2016'[Birthrate],

        "Deathrate", 'GDP by Country\_2000-2016'[Deathrate],

        "Infant Mortality", 'GDP by Country\_2000-2016'[Infant\_Mortality],

        "Population", 'GDP by Country\_2000-2016'[Population],

        "Region", 'GDP by Country\_2000-2016'[Region]

    )

)

As the population was static new column adjusted population was created

AdjustedPopulation =

IF('UnifiedTable'[Year] < 2000,

   'UnifiedTable'[Population] / (1 + (2000 - 'UnifiedTable'[Year]) \* 0.1),

   'UnifiedTable'[Population]

)

Similarly Infant Mortality column was created

Then previous and growth rate was calculated for infant mortality and population

PreviousPopulation =

CALCULATE(

    MAX(UnifiedTable[AdjustedPopulation]),

    FILTER(

        UnifiedTable,

        UnifiedTable[Country Name] = EARLIER(UnifiedTable[Country Name]) &&

        UnifiedTable[Year] = EARLIER(UnifiedTable[Year])

    )

)

GrowthRate Population =

DIVIDE(

    UnifiedTable[Population] - UnifiedTable[PreviousPopulation],

    UnifiedTable[PreviousPopulation],0

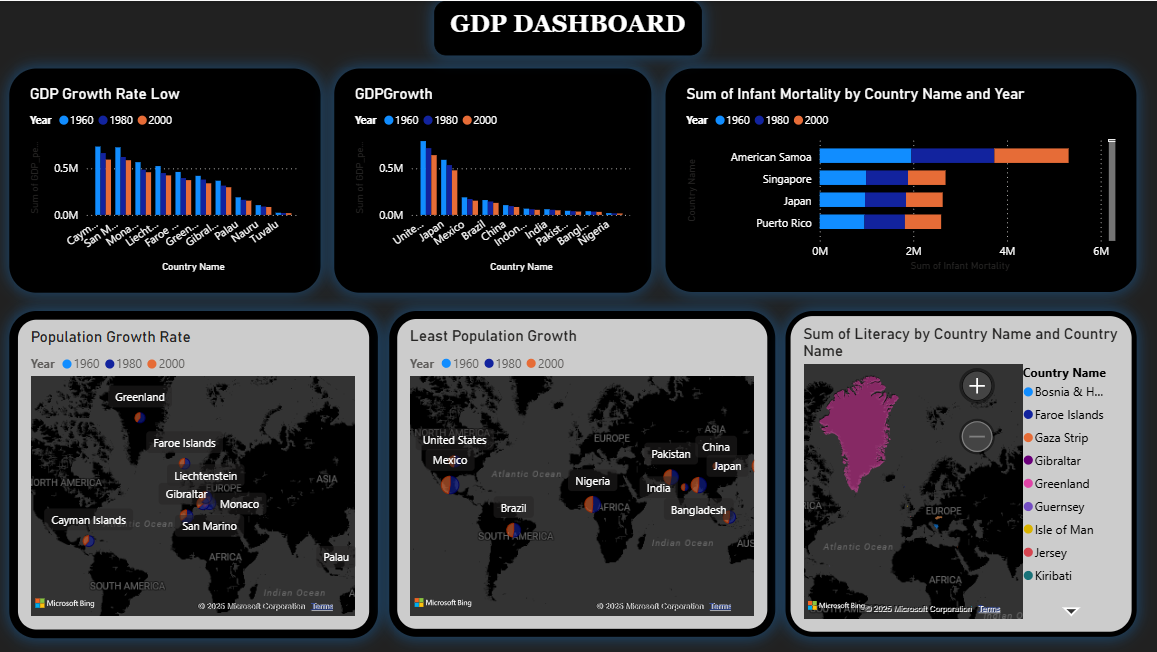
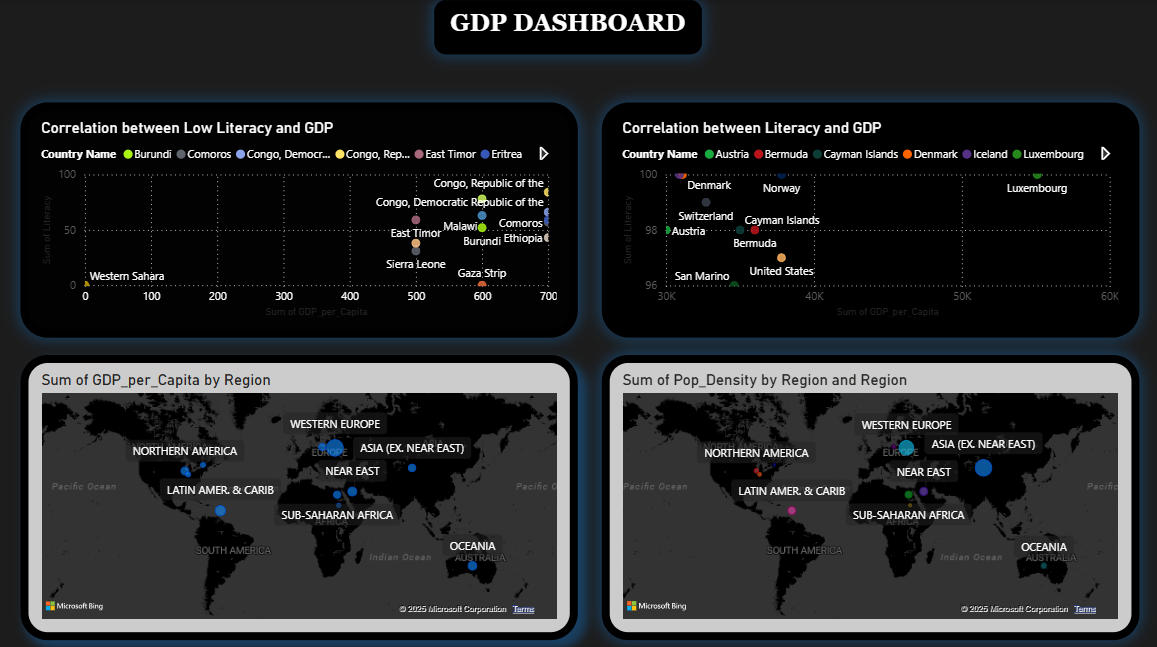
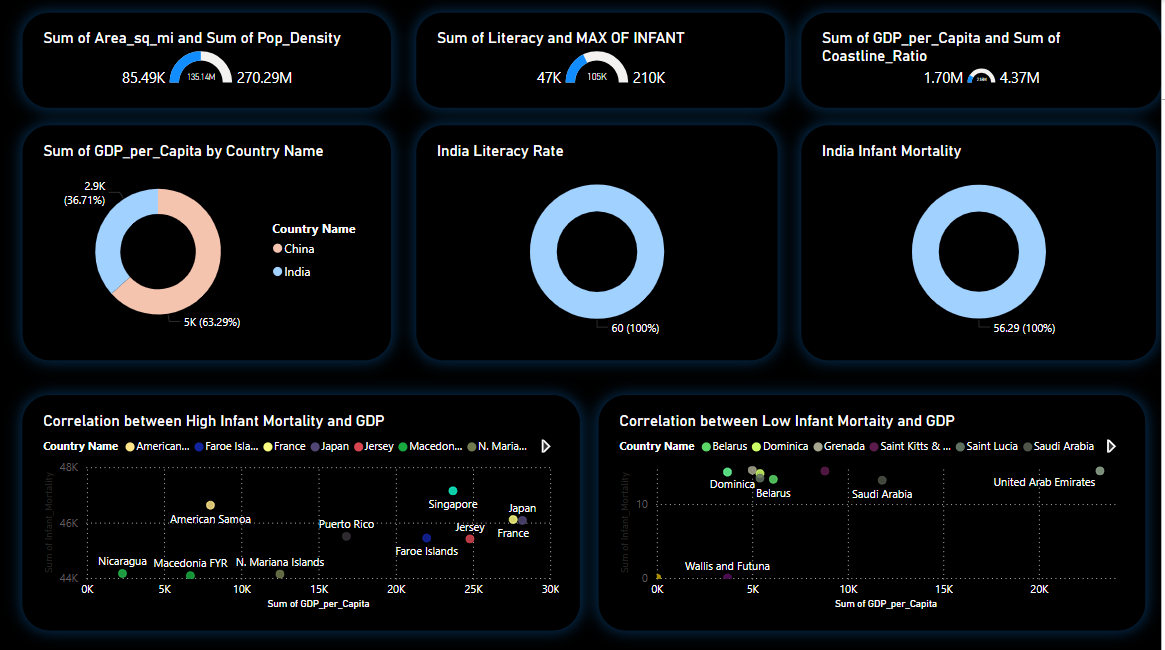
) \* 100

Created calculated columns for GDP per capita and population growth.

Aggregated data based on region for comparative insights.

Used DAX to calculate correlations between literacy rates and GDP

* **Data Visualizations and Dashboard Creation**

1. **Power BI Dashboard:**

Dashboard was created based on

Which regions have the highest population growth rates?

Africa has the highest population growth rate globally, with an annual growth rate of approximately 2.3%. This is driven by high fertility rates and a young population.

What is the relationship between GDP growth and literacy

rates?

Western Europe Countries and American with higher literacy rates tend to have higher GDP growth. Literacy improves workforce productivity, innovation, and economic participation. Conversely, low literacy rates can hinder economic development by limiting access to education and skilled jobs.

Which countries have the highest and lowest GDP per capita?

Highest GDP Per Capita: Luxembourg leads with a GDP per capita of $141,080, followed by Switzerland and Ireland.

Lowest GDP Per Capita: Countries like Burundi and Somalia have some of the lowest GDP per capita, often below $500, due to economic instability and limited resources.

Are there any notable trends in infant mortality rates over time

in different countries?

Infant mortality rates have been steadily declining worldwide due to improved healthcare and sanitation.

India’s infant mortality rate dropped from 39 per 1,000 live births in 2014

countries like Afghanistan and Somalia still face high rates due to healthcare challenges

Identify regions with the most and least population density.

Most Dense: China has the highest population density, with 22,000 people per square kilometer

Least Dense: Greenland is the least densely populated region, with only 0.03 people per square kilometer

Conclusion:

* Population Trend- Countries had significant rise in population but mortality rates were making a impact
* GDP Trend- United States and other countries had growth due to fiscal support and economic resilence. China growth remained steady (2000-2016- 3.8%)
* Infant Mortality –Countries like afghanisthan and Somalia and other African countries had more than 1000 deaths
* Literacy Trends-Sub-Saharan Africa and South Asia face the most significant challenges, with some countries reporting adult literacy rates below 30%