# Global Population Trends Analysis (1960–2022) Using World Bank Data

## Objective

The goal of this project is to conduct a comprehensive analysis of the world’s population trends over a span of more than six decades (1960–2022). Using official World Bank datasets, this analysis seeks to identify:

* The countries with the highest and lowest populations at each decade mark.
* Long-term growth patterns and demographic shifts in different parts of the world.
* Emerging trends such as rapid population growth in certain regions.
* Potential implications of these trends for policy-making, resource distribution, and development planning.

The analysis also aims to present the results in a visual, intuitive format so that even non-technical audiences can grasp the insights.

## Dataset Details

* **Source**: World Bank Open Data Portal.
* **File Used**: *API\_SP.POP.TOTL\_DS2\_en\_csv\_v2\_45183.csv*
* **Data Scope**:
  + Population totals for individual countries and regions from 1960 to 2022.
  + Includes both annual values and aggregated metadata.
* **Key Columns**:
  + Country Name — full name of the country or region.
  + Country Code — three-letter ISO country code.
  + Year columns from 1960 to 2022 — total population per year.
  + Additional metadata rows for aggregates such as continents or global totals (removed during cleaning).
* **Data Format**: CSV, wide format (years as separate columns) initially.

## Methodology

#### **Step 1: Data Import**

#### A custom read\_data() function was implemented to allow flexible reading of .csv or .xlsx formats.

#### The dataset was imported into a Pandas DataFrame for further processing.

#### **Step 2: Data Cleaning**

#### Removed non-country rows (e.g., "World", "High income", "East Asia & Pacific") to focus on country-level analysis.

#### Dropped irrelevant metadata columns.

#### Converted population values to numeric types and handled missing entries.

#### **Step 3: Data Reshaping**

#### Transformed data from wide format (each year as a column) to long format (year as a single column) for easier grouping and filtering.

#### Converted Year column to integer and Population column to numeric.

#### **Step 4: Analytical Breakdown**

#### **Decade-wise Ranking**: Grouped data by decade (1960, 1970, …, 2020) and identified top and bottom populated countries for each decade.

#### **Growth Rate Calculation**: Compared population in 1960 with 2022 to determine the percentage growth for each country.

#### **Regional Aggregation**: Summed populations by region to visualize continent-level shifts over time.

#### **Step 5: Visualization**

#### Line Charts: Displayed population growth trends for selected countries.

#### Bar Charts: Compared decade-wise population rankings.

#### Heatmaps: Highlighted population density changes over time.

#### Styling: Used Matplotlib and Seaborn for consistent, visually clear outputs.

## Findings / Results

1. ***Top Populated Countries:*** 
   * **China and India** consistently dominate the list across all decades, with India showing a sharper upward curve in recent decades.
   * The gap between China and India has been closing, with India projected to surpass China (data trend).
2. ***Least Populated Countries:***
   * Nations like **Tuvalu, Nauru, and San Marino** consistently remain at the bottom of the population rankings.
   * These are mostly small island states or microstates.
3. ***Global Growth:***
   * World population grew from ~3 billion in 1960 to over 7.9 billion in 2022.
   * This represents a growth rate of over **160%** in just over six decades.
4. ***Regional Patterns****:*
   * **Asia** accounts for more than half of the global population, maintaining dominance across the time range.
   * **Africa** shows the **fastest growth rate**, especially in sub-Saharan countries, suggesting future demographic shifts.

## Conclusion

The world’s demographic landscape has changed significantly since 1960. While Asia remains the most populated region, Africa’s accelerated growth may redefine future global population distribution. This type of population analysis can inform governments, NGOs, and international bodies when creating strategies for infrastructure, healthcare, education, and sustainable development.