**Demographic Change Analysis and Insights (1995–2020)**

Population trends and demographic indicators across 235 countries  
Based on data from 1995 to 2020

# 1. Introduction

This report provides an investor‑friendly overview and explanation of a demographic analysis project. The project uses a dataset of country‑level population statistics covering 25 years (1995‑2020). It aims to understand how populations are changing across countries and to present the data in a clear, approachable manner. The accompanying Jupyter notebook (not altered here) performs data cleaning, exploration and visualization using Python. This report translates those technical steps into business‑friendly language, explains why each step was taken, and highlights the insights generated.

# 2. Dataset Source and Variables

The dataset comes from a public compilation of country populations and demographic indicators from 1995 to 2020. It contains 4,195 rows representing 25 years of data for 235 countries. Each row corresponds to one country in a given year. The following table summarises the key variables available:

|  |  |  |
| --- | --- | --- |
| Variable | Type | Description |
| Year | Integer | Calendar year |
| Country | Categorical | Name of the country |
| Population | Integer | Total population |
| Yearly % Change | Float | Annual population growth rate |
| Yearly Change | Integer | Change in population from prior year |
| Migrants (net) | Float | Net number of migrants |
| Median Age | Float | Median age of the population |
| Fertility Rate | Float | Average number of births per woman |
| Density (P/Km²) | Integer | People per square kilometre |
| Urban Pop % | Float | Share of people living in urban areas |
| Urban Population | Float | Urban population size |
| Country's Share of World Pop % | Float | Country’s share of world population |
| World Population | Integer | World population |
| Country Global Rank | Integer | Population rank among all countries |

Definitions:   
• Yearly % Change denotes the annual percentage change in a country’s total population; for most years it is computed as the homogeneous annual percentage change over the preceding five‑year period.  
• Yearly Change measures the absolute change in population from one year to the next (or the average annual numerical change over the preceding five years).  
• Net migration is the number of immigrants minus the number of emigrants during the period.  
• Median age divides the population into two equal halves by age.  
• Fertility Rate (Total Fertility Rate) is the average number of births per woman, assuming current age‑specific fertility rates remain constant.  
• Urban population refers to people living in areas classified as urban by national statistical offices.  
• Population density is mid‑year population divided by land area.  
• Country’s share of world population is the country’s population as a percentage of the world population.  
• Global rank indicates each country’s position in the world population ranking.

# 3. Data Preparation & Cleaning

The raw dataset contained a mix of numeric and textual values and some missing entries. In the notebook, Pandas is used to load the CSV file and inspect the first few rows ("head"), the structure of the data ("info"), and basic summary statistics ("describe"). These steps are standard in data analysis: they help you understand the shape of the data, spot anomalies and assess data types.  
  
Several cleaning operations were carried out:  
• Density values contained commas (e.g., ‘1,234’) which prevented them from being treated as numbers. The commas were removed and the column was converted to integers.  
• The Country column was ensured to be of string type to avoid unintended numeric conversion.  
• Missing values in Migrants (net), Median Age and Fertility Rate were imputed with the median of each column. The median is used because it is less sensitive to extreme values than the mean.  
• Missing values in Urban Pop % and Urban Population were filled with the mean of their respective columns. A mean is appropriate here because these columns represent percentages and continuous values.  
After cleaning, the dataset no longer contained missing values, which ensures that subsequent analyses and visualizations are based on complete data.

# 4. Exploratory Analysis & Visualizations

## 4.1 Population of the Top 30 Countries in 2020

To identify the largest populations in the most recent year, the notebook filters the data to 2020 and sorts countries by population. The top 30 countries are plotted using bar charts. This helps investors understand which countries have the greatest population scale in the present year. A vertical bar chart conveys the relative differences clearly, while a horizontal bar chart with labels makes precise values easy to read.

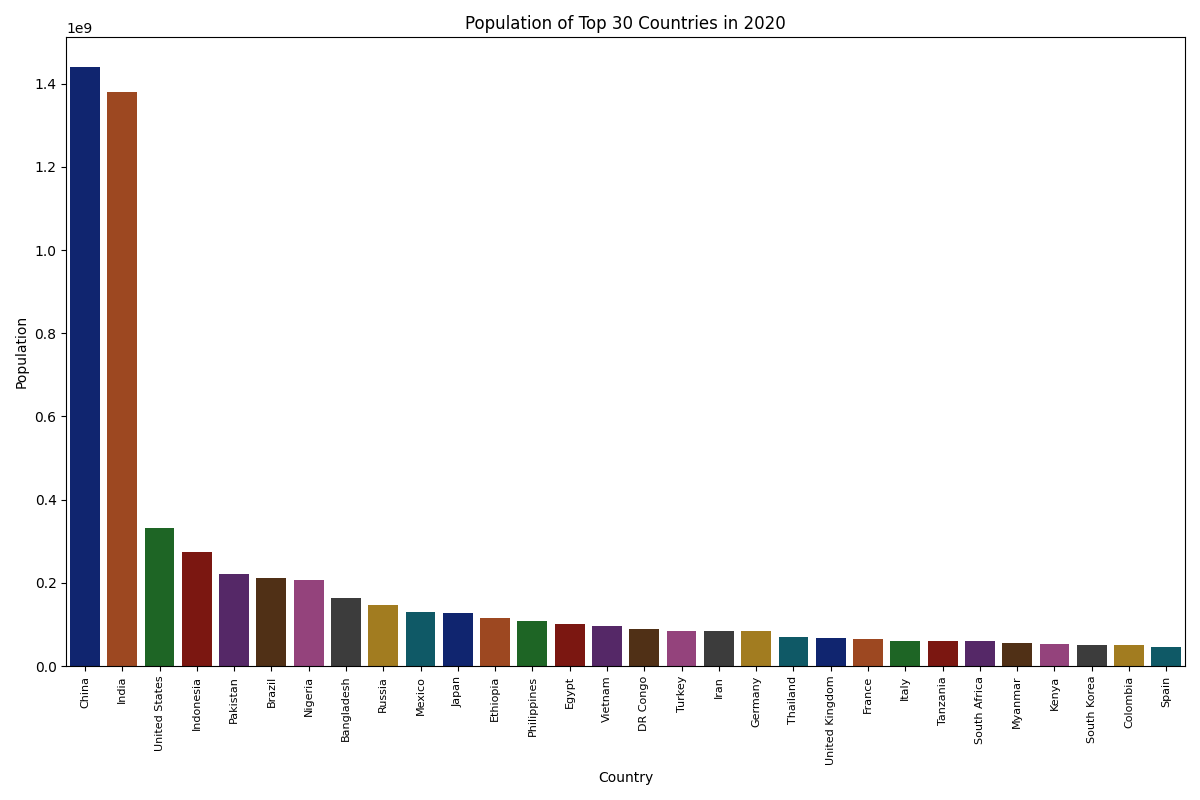


Figure 1: Vertical bar chart of the top 30 countries by population in 2020.

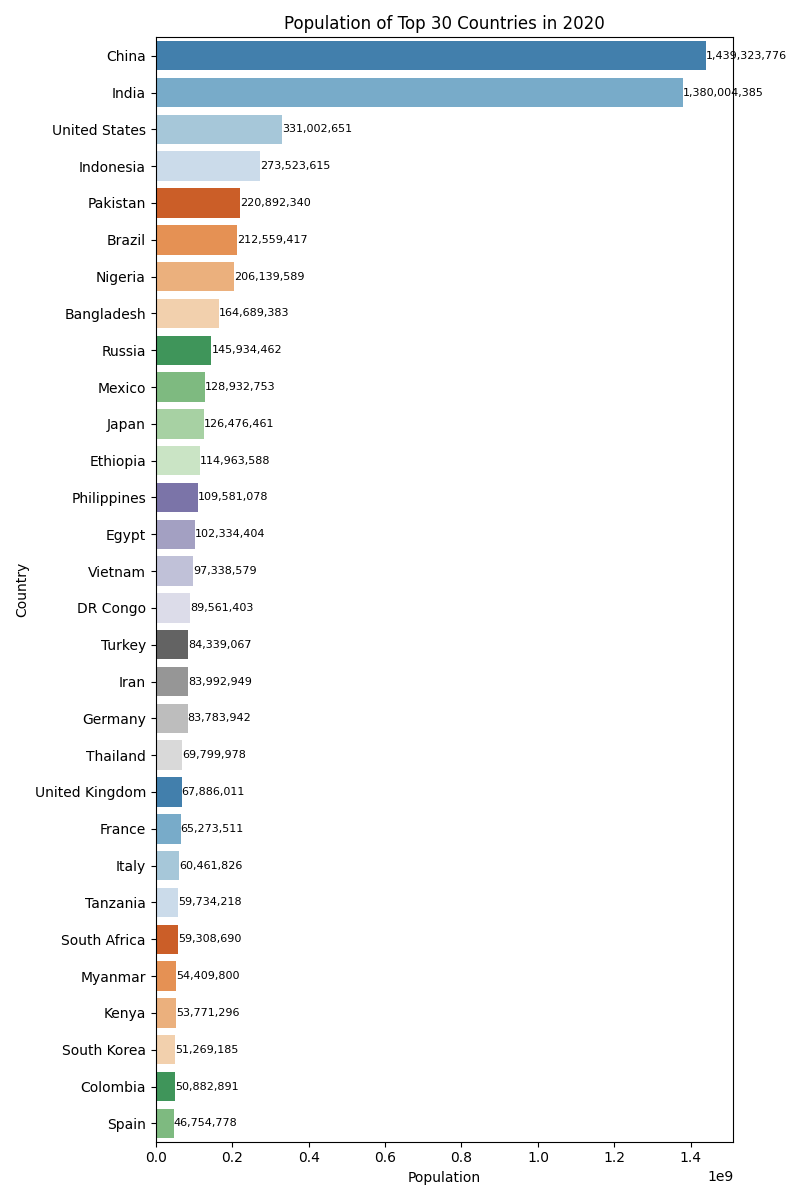


Figure 2: Horizontal bar chart of the top 30 countries by population in 2020 with value labels.

## 4.2 Share of World Population for Top 10 Countries (2020)

The pie chart aggregates the populations of all countries in 2020 and highlights the top ten. The remaining countries are grouped into an "Other" category. The chart makes it clear how dominant a few countries are in terms of global population. China and India together account for a significant share, with the next eight countries holding much smaller slices.

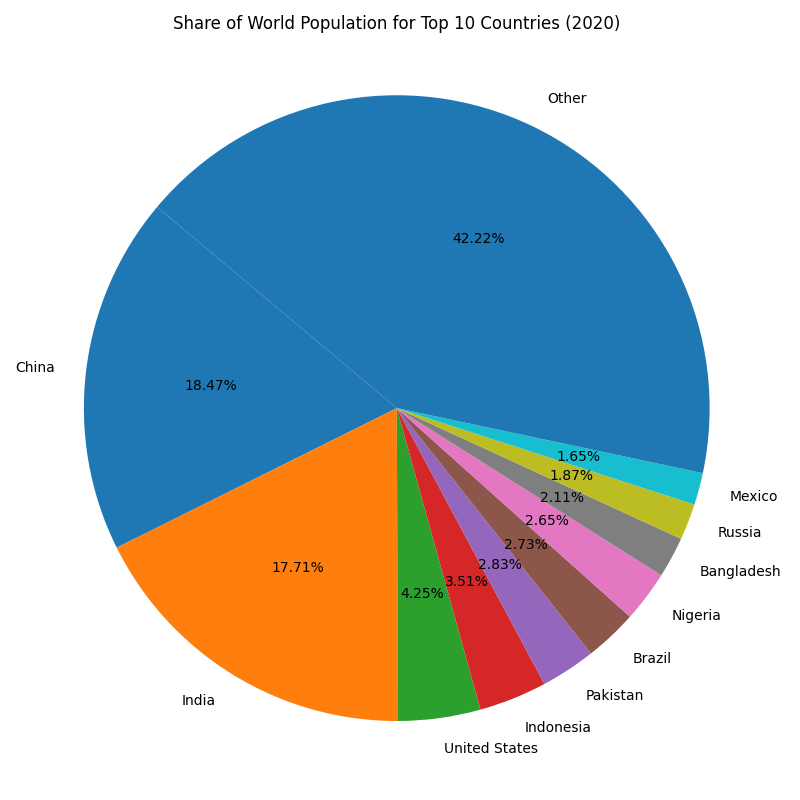


Figure 3: Pie chart showing the percentage share of the world population held by the top 10 countries in 2020.

## 4.3 Yearly Population Change in China

To understand temporal trends within a single country, the notebook defines a simple function to filter data by country and then extracts China’s data. A line chart of the annual absolute population change illustrates whether the population is growing rapidly or slowing down over time. Between 1995 and 2020 China’s annual population change declined steadily: growth peaked in the late 1990s and has been slowing since, reflecting the effects of family‑planning policies and an ageing population.

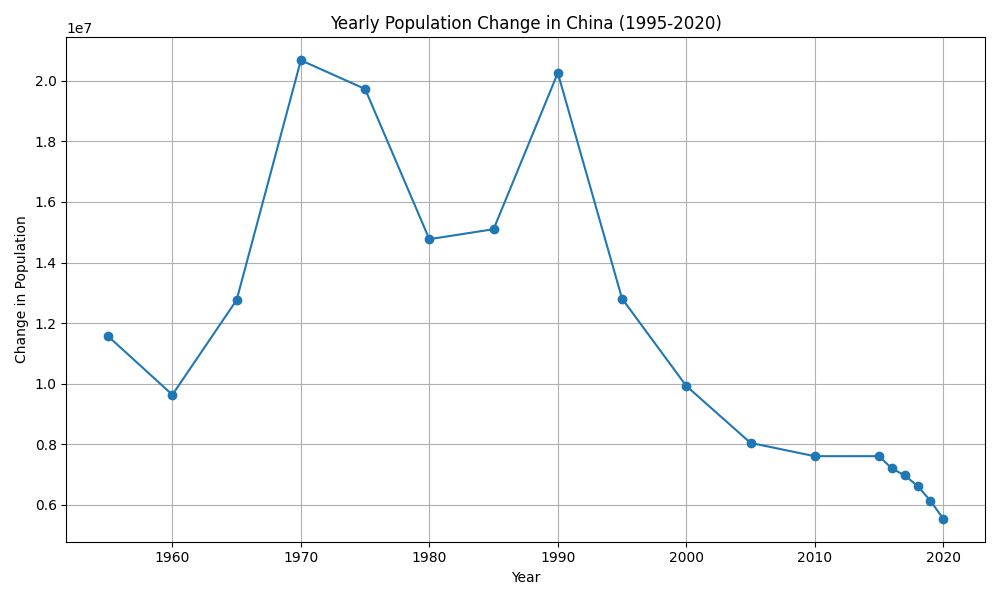


Figure 4: Annual change in China’s population (1995–2020).

## 4.4 Population Trends for the Top Five Countries

Beyond a single snapshot year, investors may want to know how population trajectories compare across major markets. The notebook identifies the five most populated countries in 2020 and then tracks their populations from 1995 to 2020. The line chart below shows that China and India dominate in size but India’s population curve is steeper, indicating faster growth. The United States shows steady growth, while Indonesia and Pakistan grow more modestly but consistently.

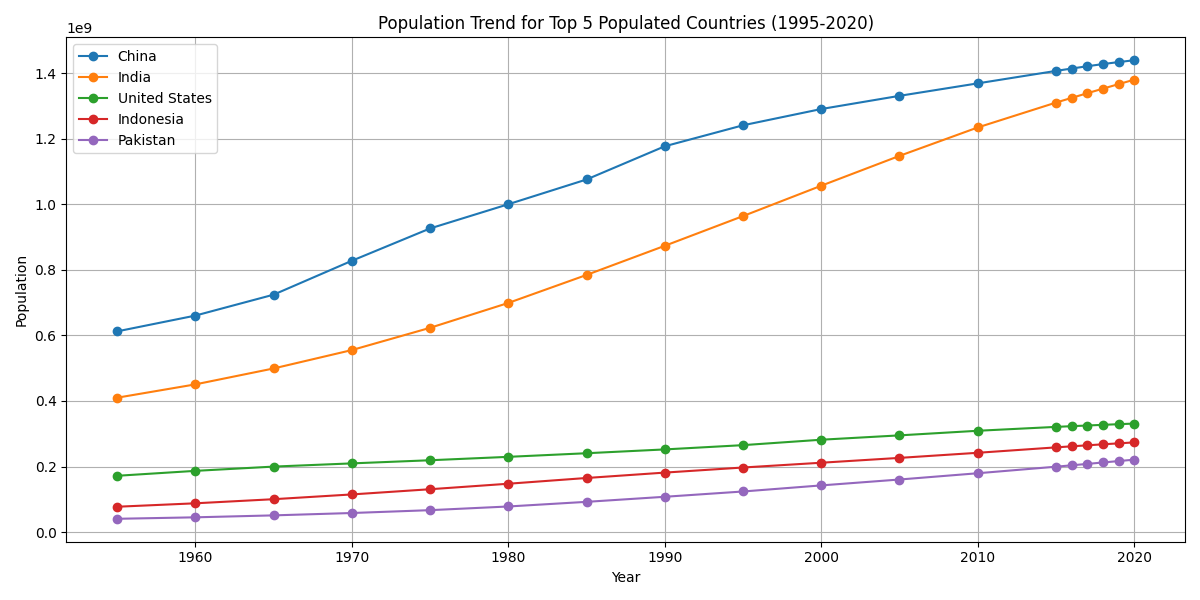


Figure 5: Population trends of the five most populous countries (1995–2020).

# 5. Key Insights and Business Implications

• Concentration of population: A handful of countries account for a large share of the world population. In 2020, China and India together represented nearly 36% of the global total. Understanding these concentrations helps businesses prioritise market entry and resource allocation.  
• Growth dynamics: Populations are not growing uniformly. China’s annual population change has slowed considerably, whereas India’s population continues to grow strongly and may overtake China in the coming years. Countries such as Nigeria and Pakistan, although smaller today, have high growth rates and could represent emerging markets with expanding labour forces.  
• Urbanisation trends: The dataset includes urban population metrics, reflecting the shift from rural to urban living. High urbanisation (e.g., above 80% in the U.S.) suggests mature infrastructure markets, while countries with lower urban shares may see future investment in housing, transport and utilities.  
• Age structure and fertility: Median age and fertility rate provide insight into demographic maturity. Countries with low median age and high fertility (many in Africa and South Asia) will have youthful populations entering the workforce. Countries with higher median age and low fertility (e.g., Europe) face ageing populations and potential labour shortages; these markets may focus on automation, healthcare and retirement services.  
• Migration effects: Net migration can affect population growth independently of births and deaths. Positive net migration (more immigrants than emigrants) supports workforce expansion, while negative net migration may signal challenges or instability.  
• Business strategy: For investors and companies, demographic indicators inform market potential, workforce availability and consumer demand. Rapidly growing, youthful populations can present opportunities for consumer goods, education and employment services. Ageing societies may prioritise healthcare, pensions and robotics.

# 6. Conclusion

This project demonstrates how open demographic data can be transformed into actionable insights through systematic cleaning, exploration and visualisation. By understanding the definitions behind each variable and carefully preparing the data, we can build accurate charts and detect meaningful patterns. The analysis indicates that demographic forces – population size, growth, urbanisation, age structure and migration – vary widely across countries. These forces shape economic opportunities and risks. For stakeholders seeking to invest, expand into new markets or anticipate future demand, demographic intelligence is an essential tool.  
  
The original notebook contains the full Python code used to produce these figures. This report translates its technical details into plain language and highlights why each step was necessary. All definitions used in this report are sourced from authoritative organisations such as the United Nations, World Bank and Our World in Data.