**University of Sulaymaniyah**

**Collage of Science**

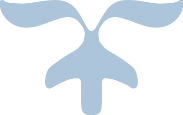
**Computer Department**

**Fourth Stage**



**Data Analysis Of Word Population Data from 2000 to 2023**

**(Data Science)**



**Prepared by:**

**Maria Osman Helin Yunis Omer Gashbin Mhamed Mahmud**

**2023-2024**

**Introduction**

*Overview* : This project to analyze global population trends from 2000 to 2023 and Understanding population dynamics is crucial for making informed decisions in various fields such as urban planning, resource allocation, and policy formulation.

We *chose* this project that provides insights into population growth, distribution and insight aids governments, organizations, and policymakers in making informed and strategic decisions that align with the evolving demographic landscape.

*Relevance to Society* : becauseunderstanding population dynamics is crucial for achieving sustainable development, our project contributes to the knowledge base needed to create and implement development plans that are environmentally, socially, and economically sustainable, ensuring the well-being of current and future generations and by distribution patterns, planners can design and implement infrastructure projects that meet the demands of growing urban populations, reducing congestion, and enhancing overall livability and Population data is essential for planning healthcare and education services the project's findings can be utilized to anticipate healthcare needs, plan for medical facilities, and tailor educational strategies to accommodate changing demographics and economic planning and employment strategies cause population trends have a direct impact on the economy ,and the benefits include improved decision-making, sustainable development, optimized resource allocation, enhanced urban planning, and better delivery of essential services, ultimately contributing to the overall well-being and progress of society.

**Problem statement**

The main problem addressed by this project is the lack of a comprehensive understanding of global population dynamics, the available knowledge lacks a detailed examination of how global populations have been changing over the specified period. there is a gap in understanding the patterns, variations, and accelerations of population growth across different regions and nations. So our project involves understanding population growth, identifying influential factors, and recognizing disparities among countries and continents .

**Methodology**

The project utilizes Python programming and data analysis libraries to process and visualize population data. The chosen methodology involves exploring key population indicators, such as growth rates and continent-wise distribution, to derive meaningful insights.

Design Process

The design process includes data cleaning, exploratory data analysis, and the application of statistical techniques to uncover patterns and correlations within the dataset.

This is a pseudocode :

1 . Import necessary libraries

2 . Read data from CSV file

3 . Display all columns

4 . Display first 5 rows of the DataFram

5 . Display last 5 rows of the DataFrame

6 . Display general information about the dataset

7 . Display summary statistics of the dataset

8 . Check for missing values and duplicates

9 . Select the years columns

10 . Finding mode and frequency table to find out which continents have the most countries

11 . Finding median and mean of all numerical columns

12 . finding mean and median of population columns

13 . Select the numerical columns only for finding the correlation

14 . Find the correlation

15 . Set a larger figsize

16 . Find the total population for each year

17 . Increase population between the year of 2000 to 2023

18 . Find the total population by continent

19 . Display the total population by continent

20 . Find the top five countries with highest population for each year

21 . Find the five countries with the lowest population for each year

22 . Sort the dataset based on the (growth rate) column

23 . Select the top 10 rows

24 . predection about future population

25 . Visualization using matplotlib and seaborn

26 . Save visualizations as images

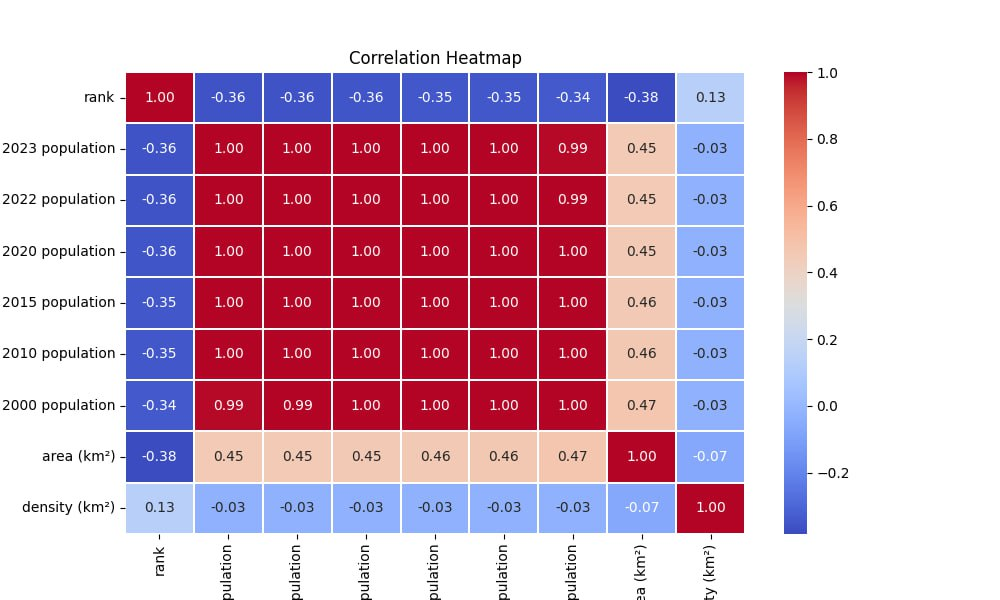
27 . Future population prediction

28 . Save the final projection as an image

**Implementation**

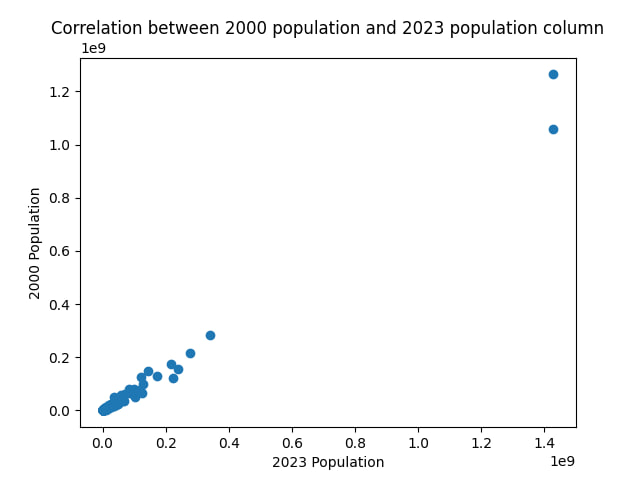
Correlatin Heatmap :

Correlation between numerical columns and correlation between population in difference years mostly have a good correlation



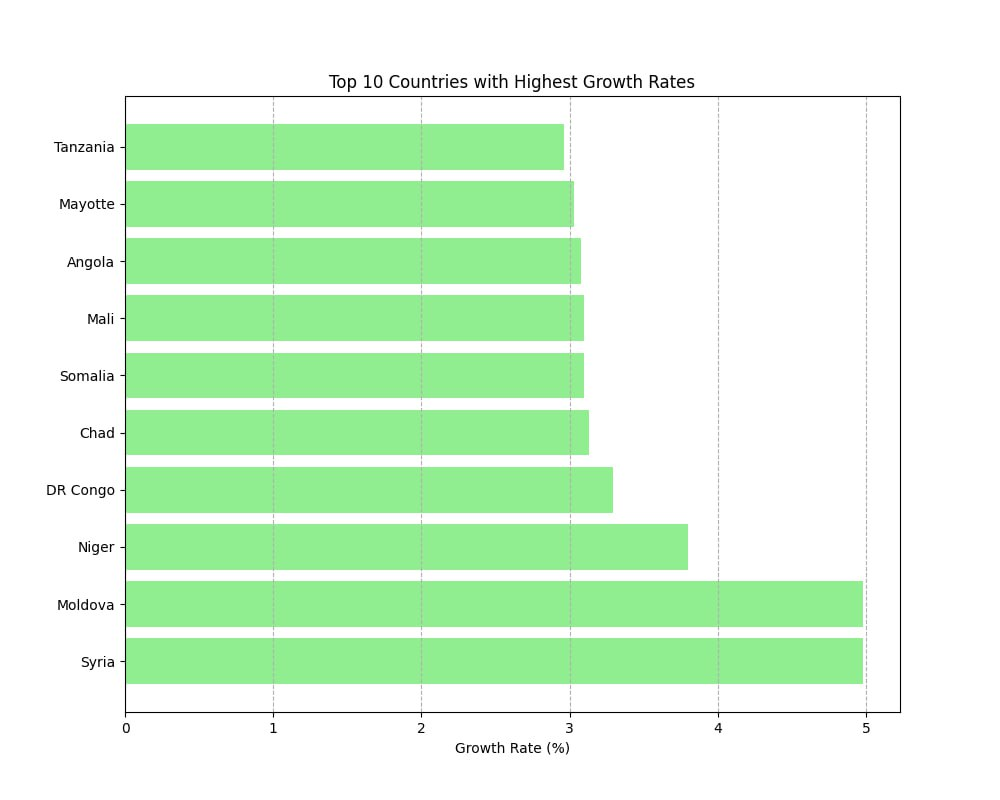
Correlation between 2000 & 2023 population column :

This is a sccater plot o show correlation between 2000 & 2023 population column near the linear line , it means that a good correlation



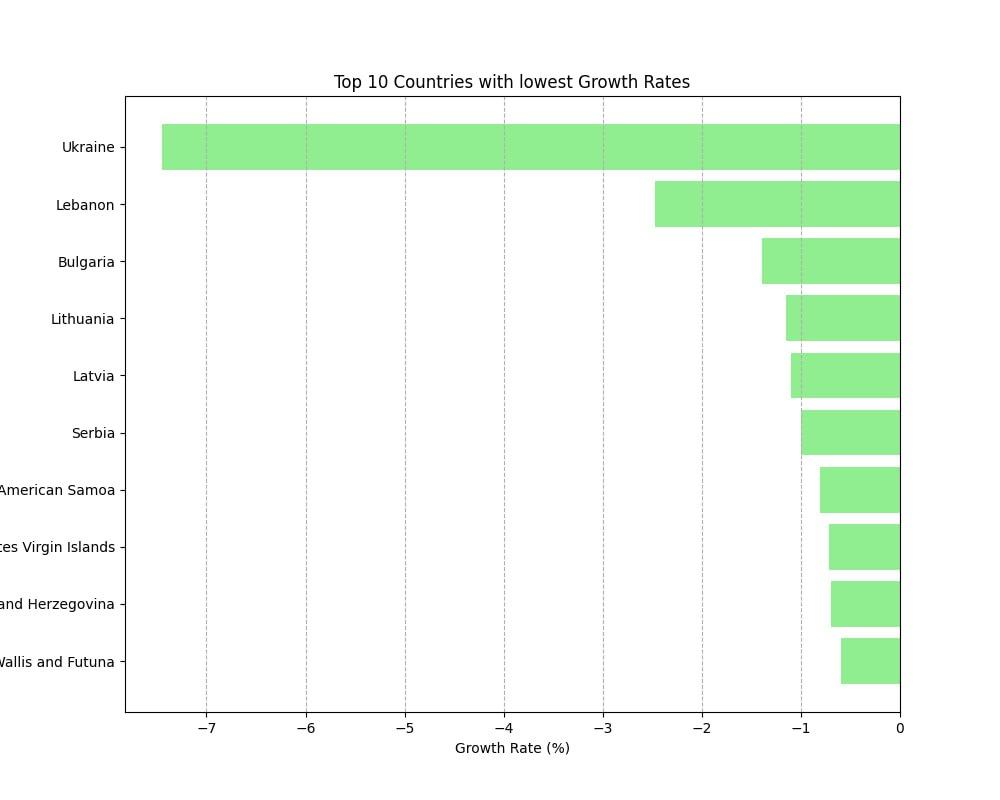
Top 10 columns with highest growth rates :

In our data set syria has a greatest grouth rate over the countries



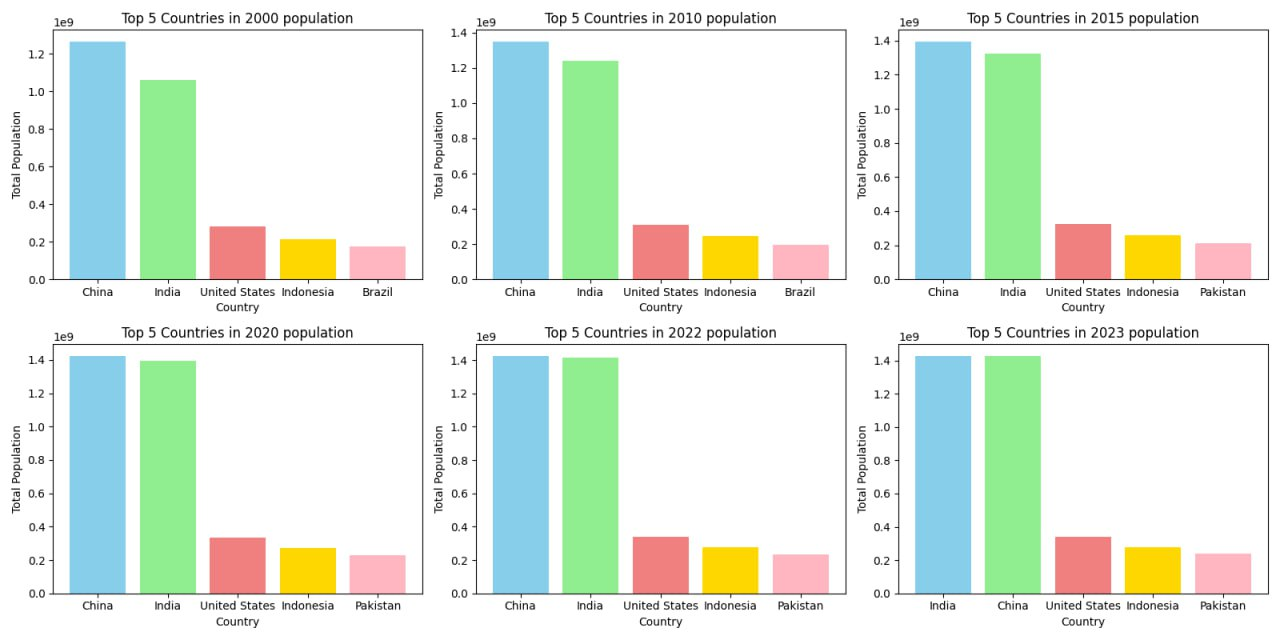
Top 10 columns with lowest growth rates :

In our data set vallis and futana has a lowest grouth rate over the countries



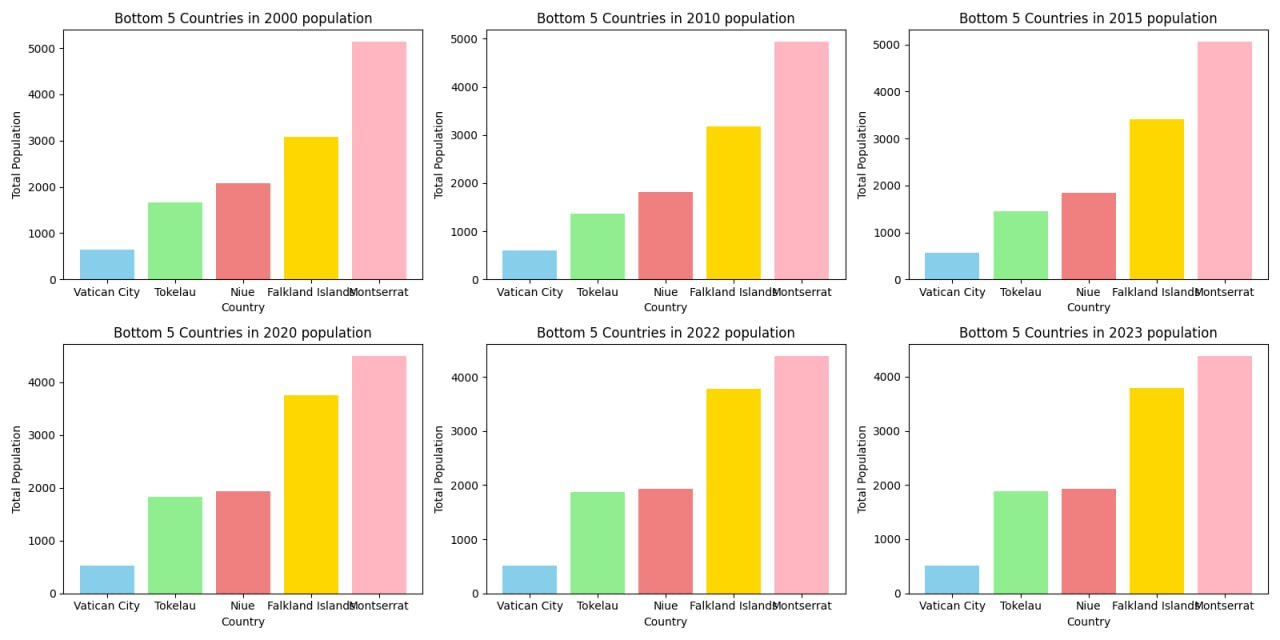
Top 5 countries for total population :

In 2000 China is top in number of population but in 2023 india is in top

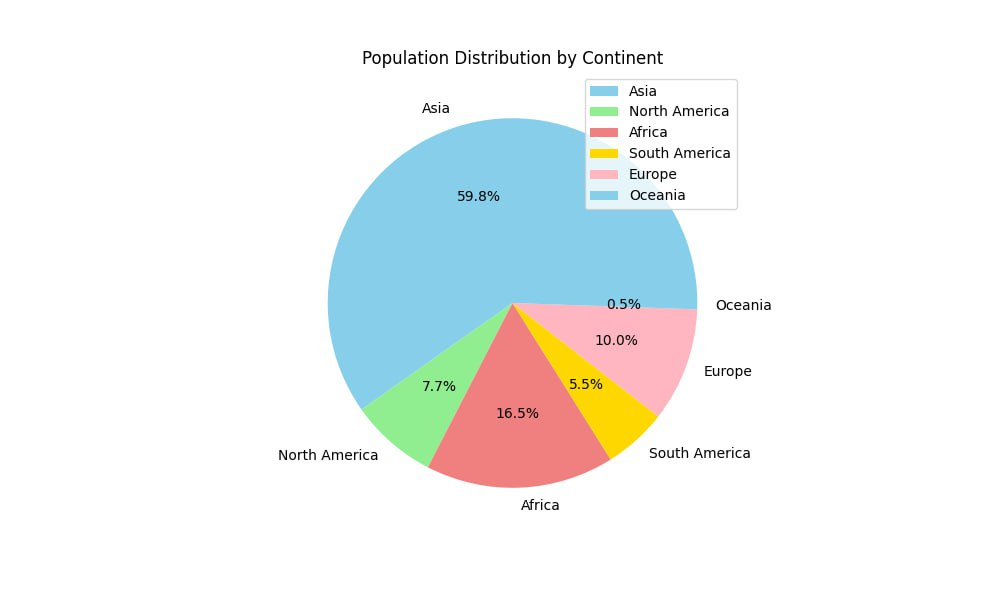


Buttom 5 countries for total population :

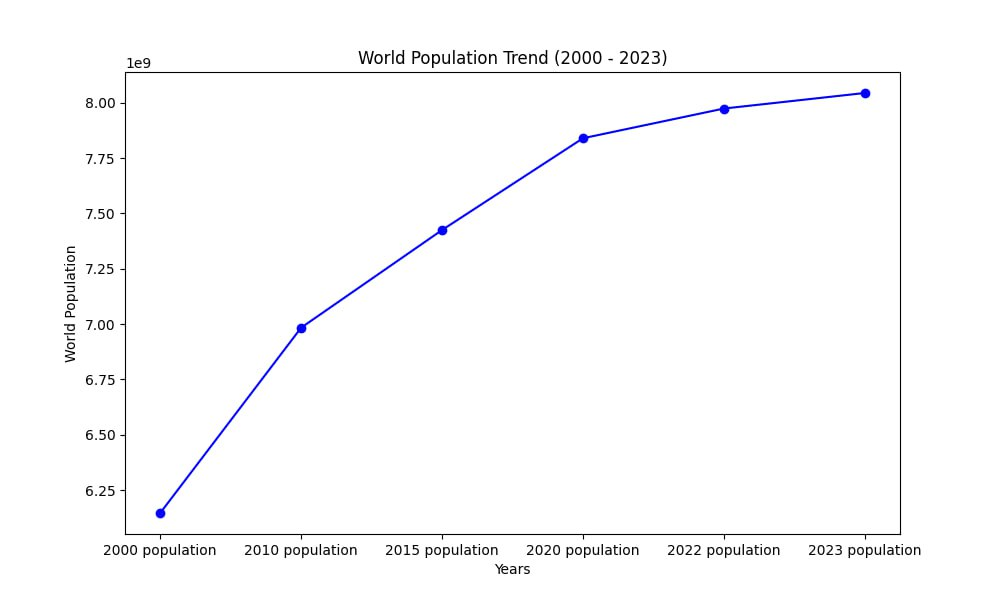
In 2000 to 2023 vatican city is bottum in number of population



Population distibution by continent :

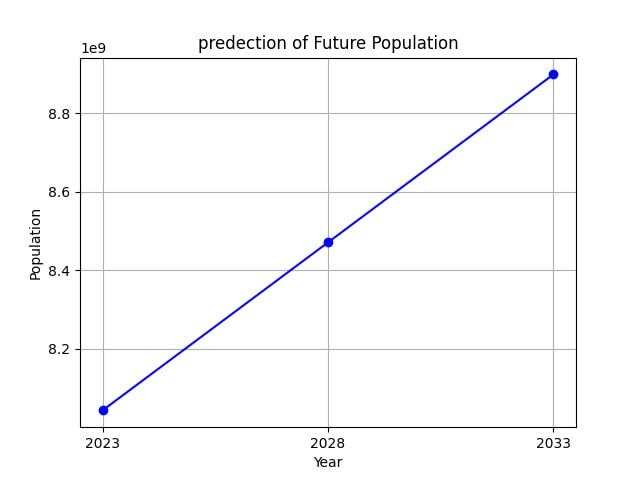


Word population trend (2000 \_ 2023) :



Future predection for population :

By using mean of increase population from 2010 to 2015 and from 2015 to 2020 we are predect this result for 5 years that population increase 427732802 in the world



**Results Discussion**

Overall, our project provides a comprehensive analysis of world population trends, distribution, and growth rates over the specified years, along with visualizations to aid in understanding the patterns in the data. And we know greatest range of population are in witch continent in difference years and we prediction to know how many numbers of people after 5 years in future .

**Project Conclusion**

In conclusion, this project lays the foundation for a comprehensive understanding of global population trends. Future developments can further refine and expand the analysis, making it a valuable resource for policymakers, researchers, and anyone interested in the dynamic field of population .