India Census Analysis

A detailed data analysis and visualization project using India Census data to understand population distribution, demographics, age groups, gender ratio, working population, and more.

Project Overview

This project explores the India Census dataset to uncover meaningful insights about population patterns across states and districts.
It includes data preprocessing, demographic calculations, and multiple visualizations to represent key findings clearly.

Technologies Used

- **Python**
- **NumPy**
- **Pandas**
- **Matplotlib**
- **Seaborn**

Data Preprocessing

Before analysis, the dataset was cleaned and prepared by:

- Inspecting data structure and info
- Checking and handling **null values**
- Identifying **duplicate records**
- Ensuring column consistency
- Understanding data types

Analysis Performed

1. Population Insights

- Total **state-wise population**
- Total **district-wise population** with district codes

- **Top 10 most populated states**
- District code–wise population distribution

2. Religion-Based Analysis

- Population counts across major **religions**

3. Age Group Analysis

- Age distribution **state-wise** and **district-wise**
- Minimum and maximum values in age groups

4. Gender Distribution

- Total **male population** in each state
- Total **female population** in each state
- **Working males** and **working females** by state
- **Non-working males** and **non-working females**

5. Education Status

- Male vs female comparison for different **education categories**

Visualizations

The following charts were created using Matplotlib & Seaborn:

- Line plot for population trends
- Distribution plot (**displot**) for demographic distribution
- Bar plot comparing male & female education status
- Pie chart for state-wise population share
- Top 10 populated states bar chart
- District code–wise population visualization
- Correlation **heatmap**

Conclusion

This project provides a complete exploration of India's census data—covering data cleaning, demographic analysis, gender comparisons, and meaningful visualizations. It highlights how population, age groups, religion, and working status vary across different states and districts.