

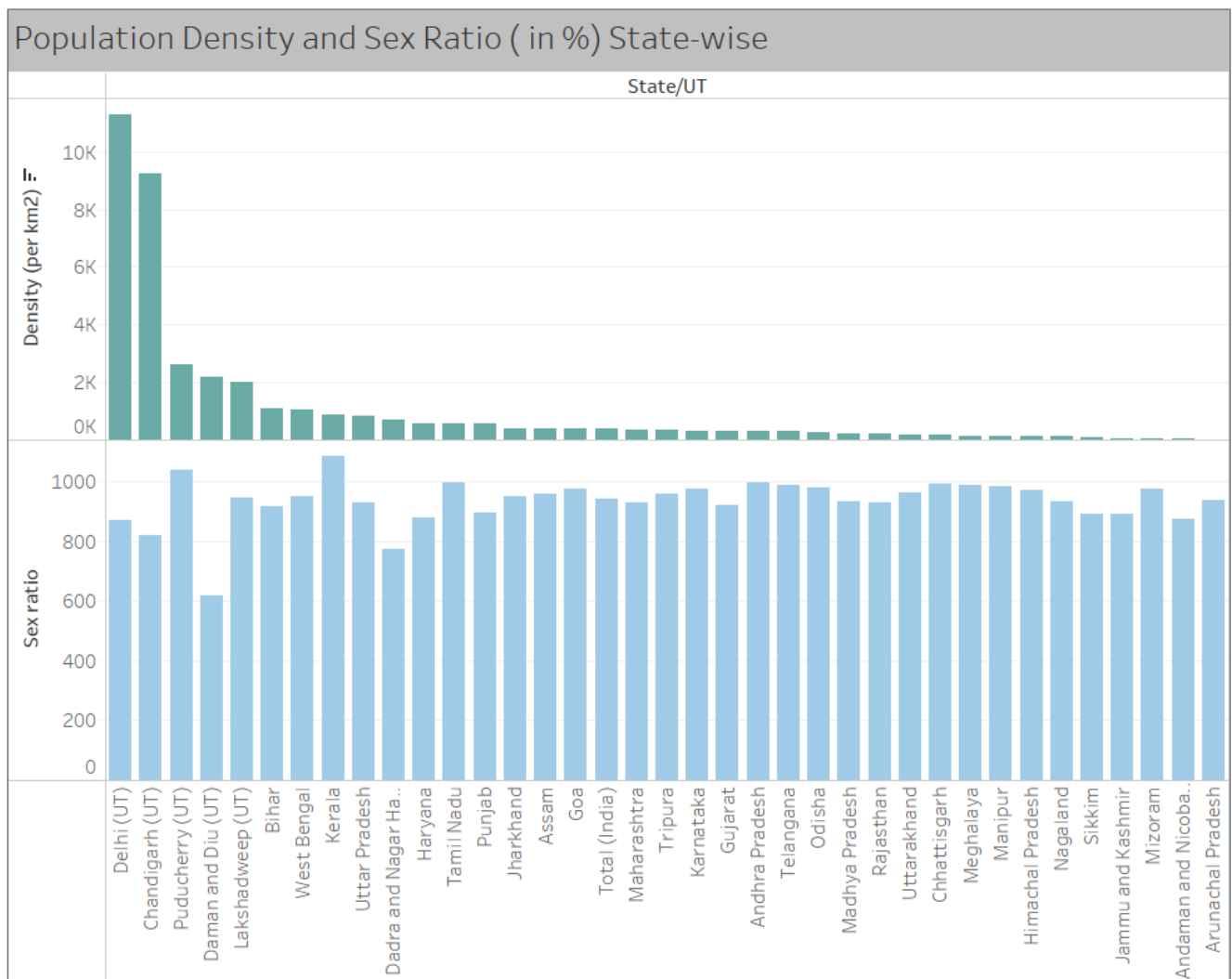
**Arnav Tumbde**  
**RCOEM Nagpur**

**Domain : Data Science**

### #Task 1

- Create a bar chart or histogram to visualize the distribution of a categorical or continuous variable, such as the distribution of ages or genders in a population

*Screenshots :*



## **a. About the Dataset**

For this task, I used a dataset that contains population statistics for various states and union territories. The dataset includes the following columns:

- State/UT: The name of the state or union territory.
- Population: The total population of each state or union territory.
- Percent (%): The percentage of the total population for each state.
- Male: The male population of each state or union territory.
- Female : The female population of each state or union territory.
- Difference between male and female : The numerical difference between the male and female populations.
- Sex ratio : The ratio of females to males per 1,000 males.
- Rural : The rural population of each state or union territory.
- Urban : The urban population of each state or union territory.
- Area (km<sup>2</sup>) : The area of each state or union territory in square kilometers.
- Density (per km<sup>2</sup>) : The population density calculated as population per square kilometer.

## **b. Explanation of the Concept**

The primary objective of this task was to create a bar chart or histogram to visualize the distribution of a continuous variable, such as the distribution of population sizes, genders, or densities across different states or union territories. This helps to understand how population figures vary across regions, which can reveal insights into demographic patterns, urbanization trends, and gender imbalances.

## **c. Outcome of the Analysis**

1. **Population Density vs. State/UT** : I created a bar chart showing the population density of each state and union territory. This visualization revealed significant differences in how densely populated certain regions are compared to others.

2. **Male vs. Female Population** : A side-by-side bar chart was created to compare the male and female populations across states. This chart highlighted regions with notable gender imbalances.

3. **Urban vs. Rural Population** : By visualizing the urban and rural populations in a dual-axis chart, I could compare the urbanization levels across states. Some states have a higher urban population, indicating more urbanized regions, while others have a predominant rural population.

4. **Sex Ratio vs. State/UT** : A bar chart showing the sex ratio for each state helped identify states with particularly high or low sex ratios, indicating gender distribution across regions.

#### **d. Conclusion**

- Population Density : There is a wide variation in population density, with some states being much more densely populated than others. This has implications for resource allocation, infrastructure development, and urban planning.
- Gender Distribution : Certain states show significant gender imbalances, which can impact social policies and gender-focused programs.
- Urban vs. Rural : The urbanization levels vary greatly, indicating different stages of economic development and urban planning needs.