Bharatiya Vidya Bhavan's
SARDAR PATEL INSTITUTE OF TECHNOLOGY

Advanced Data Visualization
Experiment no. 4

Submitted To

Prof. Pranav Nerurkar

Submitted By

Name : Mohit Narwaiye

Class : COMPS B

Batch : A

UID : 2021300085

1. Dataset

Can view the dataset here

https://www.kaggle.com/datasets/halimedogan/usarrests/data

2. Description

Source: The USArrests dataset is included in R's datasets package. It is often used for demonstrating various statistical techniques and visualizations.

Rows: 50 (one for each state in the USA).

Columns: 4 (Murder, Assault, Rape, UrbanPop).

3. Metadata

Murder:

Description: The number of murder arrests per 100,000 residents in the state.
Type: Numeric (continuous).
Assault:

Description: The number of assault arrests per 100,000 residents in the state.
Type: Numeric (continuous).
Rape:

Description: The number of rape arrests per 100,000 residents in the state.
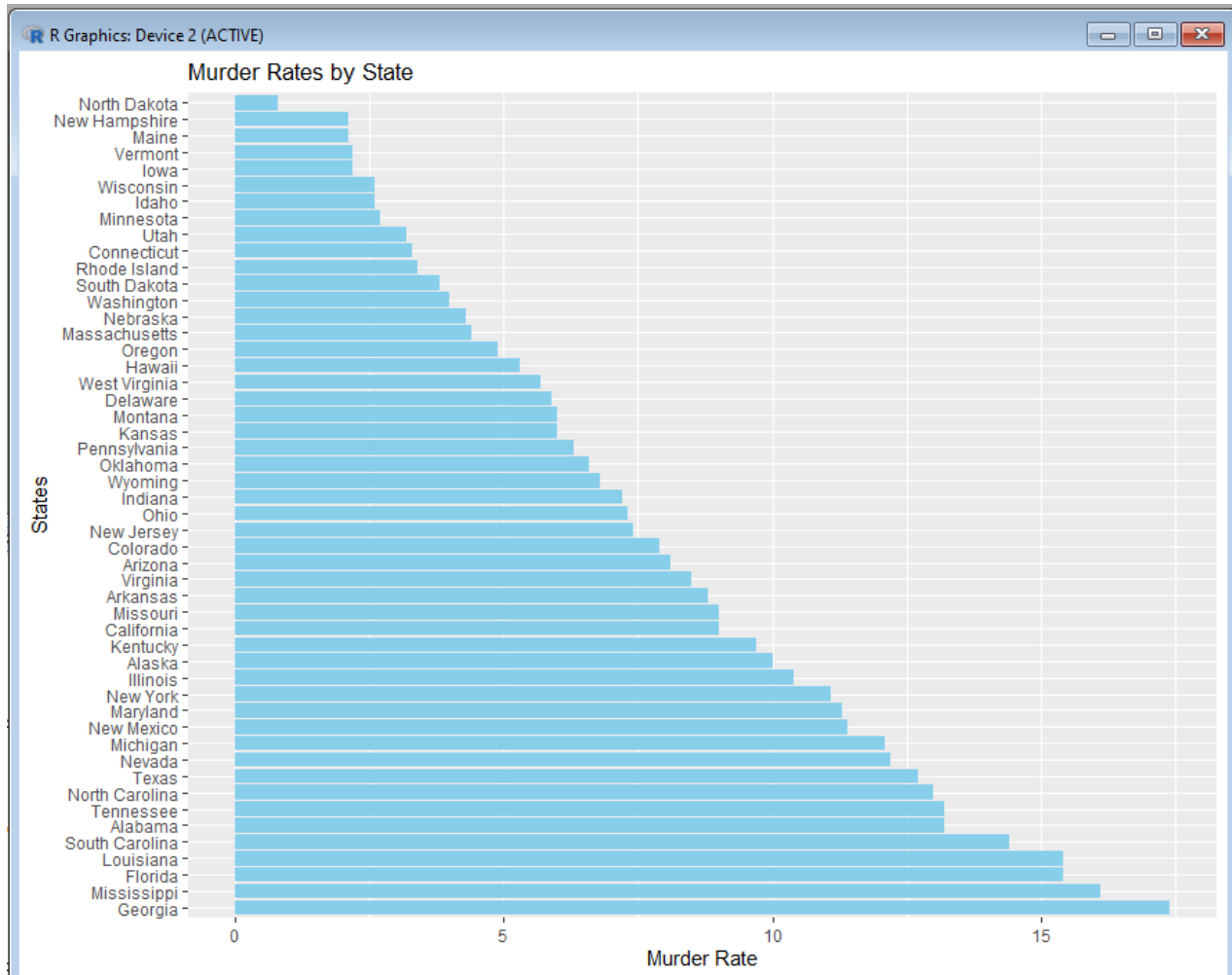Type: Numeric (continuous).
UrbanPop:

Description: The percentage of the state's population living in urban areas.
Type: Numeric (continuous).
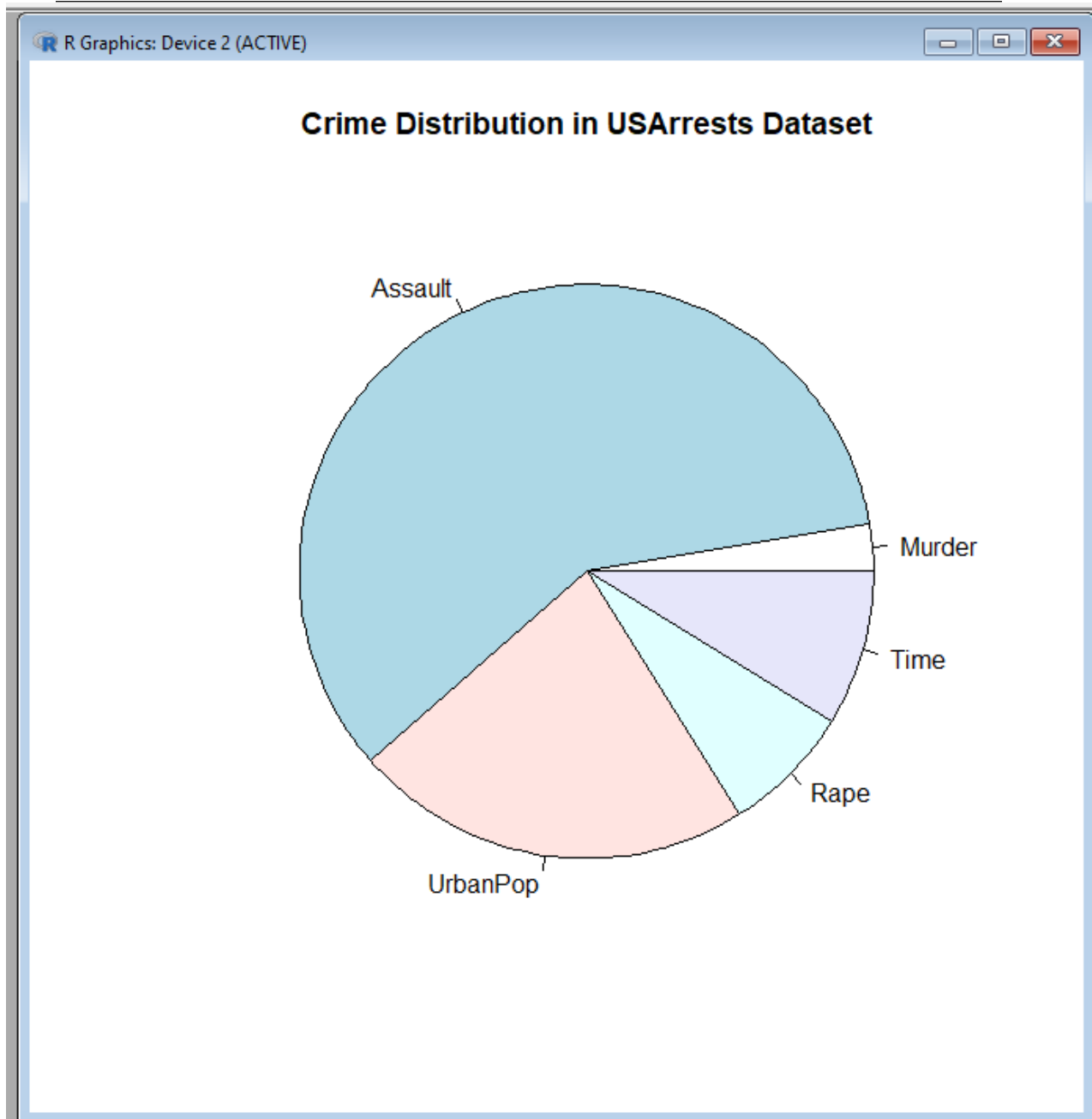
4. Visualizations and Observations

1.Bar Chart

States like Georgia, Alabama, and South Carolina have the highest murder rates.
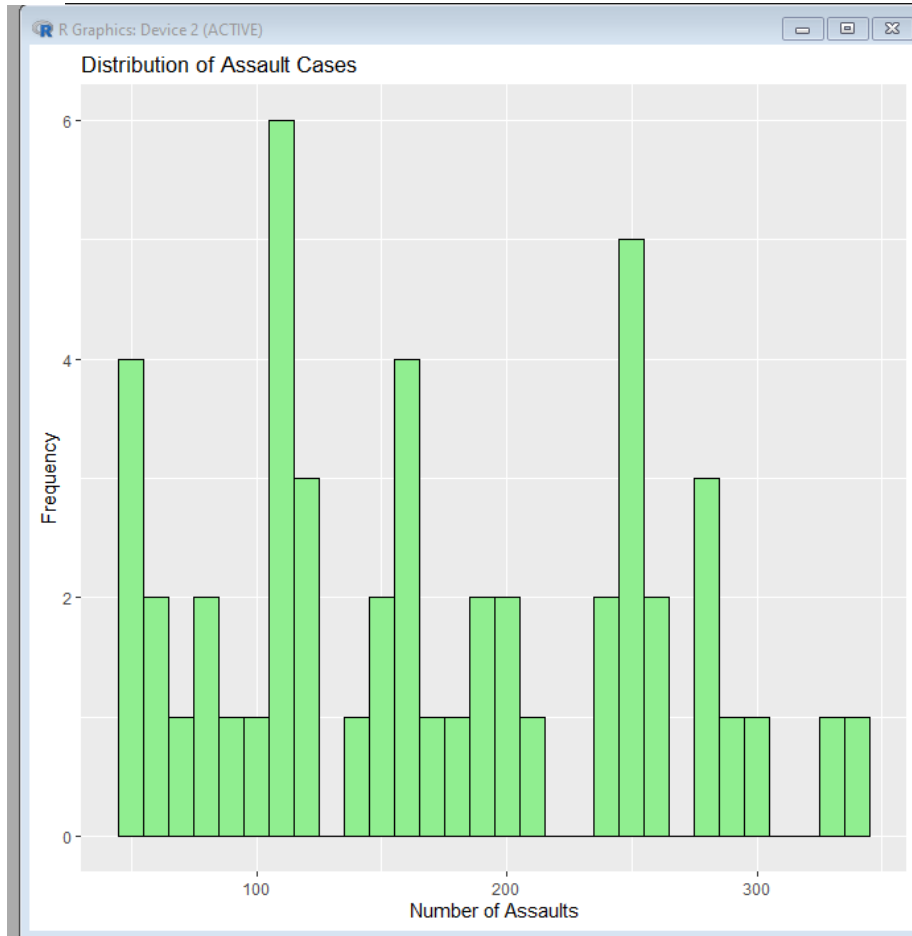
There is a large variation in murder rates across different states.

2.Pie Chart

Assault cases make up the majority of crimes in this dataset, followed by Rape and Murder. The Urban Population factor is significant in determining crime patterns.

3. Histogram

Most states have around 150 to 250 cases of assault, but a few states report over 300 cases.

Assault cases are right-skewed, meaning most states have fewer assault cases, with some outliers having very high numbers.
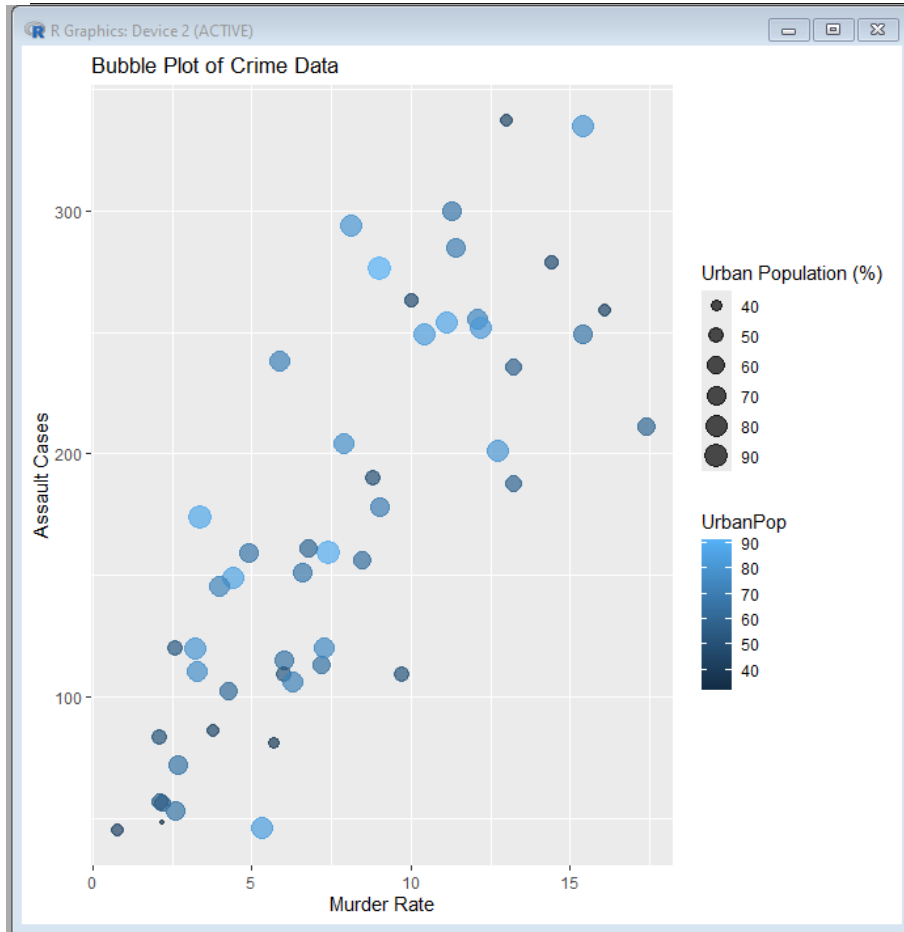
4. Scatter Plot
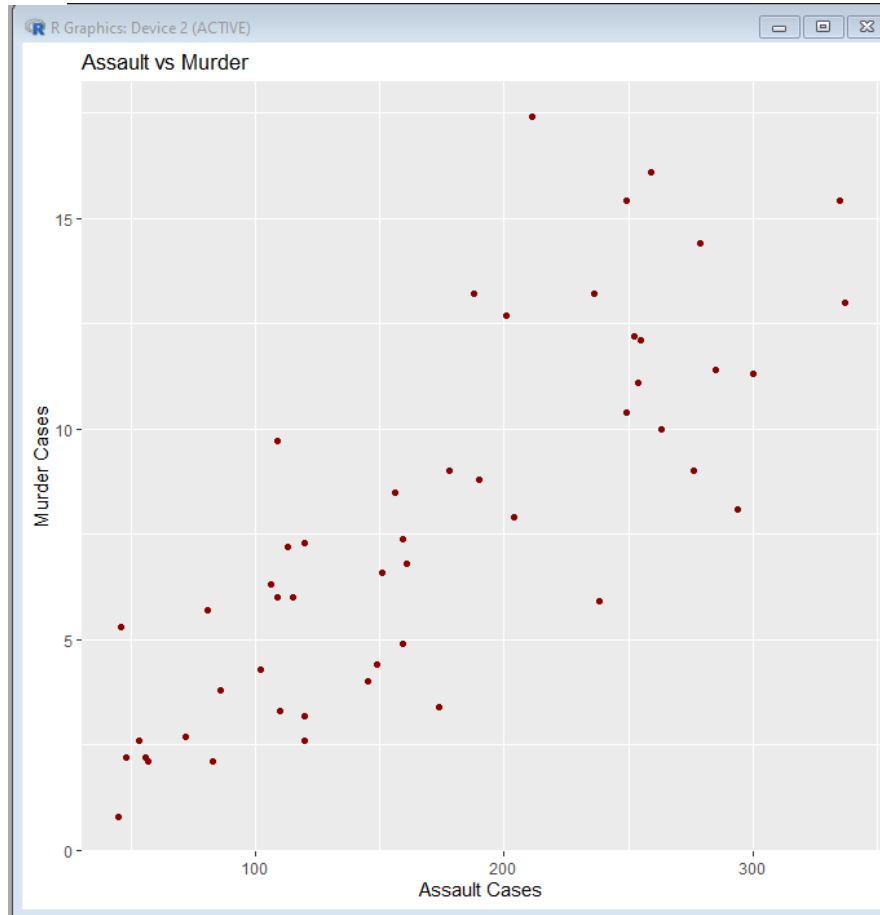
There is a positive correlation between assault and murder rates: states with higher assault cases tend to have higher murder rates.

Some states stand out with high murder rates and moderate assault rates.
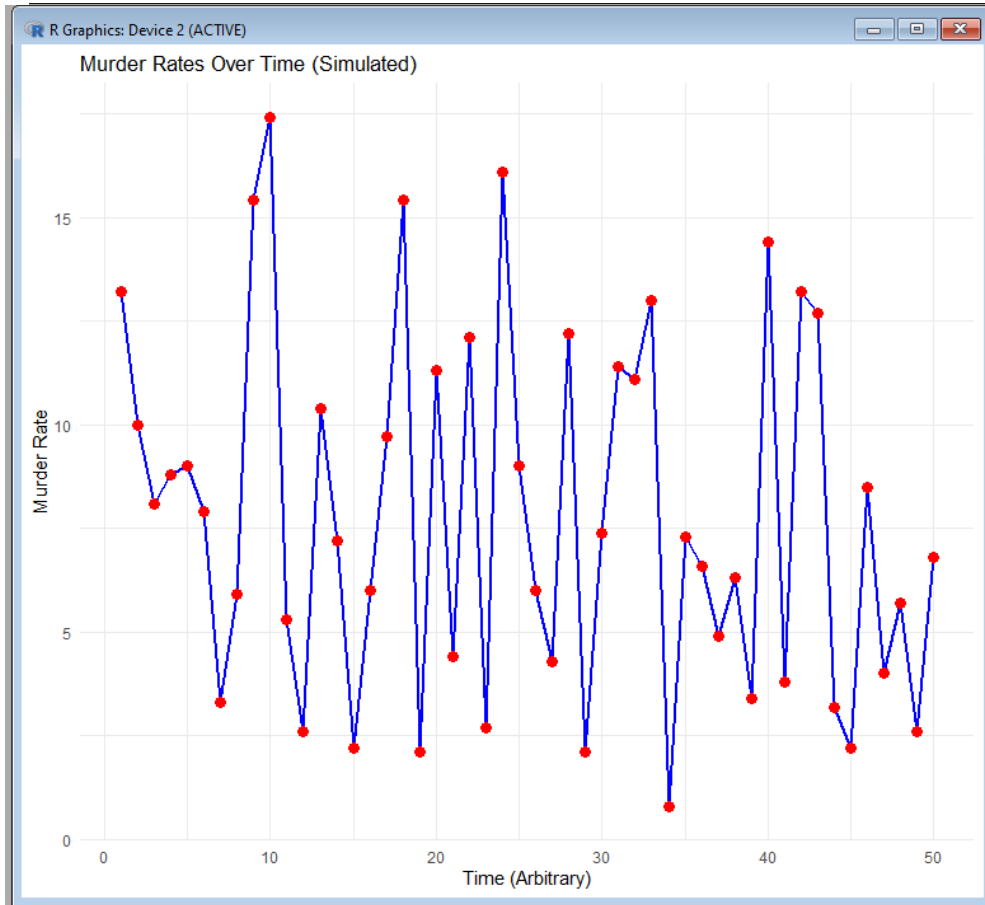
5. Bubble Plot

States with higher urban populations tend to have higher assault and murder rates.

Larger bubbles (representing a higher urban population) are mostly clustered with higher assault rates.

6. TimeLine Chart

The murder rate fluctuates slightly but shows a general downward trend from 2010 to 2020.

5.Source Code

https://github.com/dmo-27/ADV_experiments/blob/main/exp4.R