

Assignment 3

STSM 2634

Date: 14 April 2025

Due by: 19 April 2025

Read the instructions carefully:

- **You must submit the word file generated by R-markdown (.rmd). Other types of files will not be accepted.**
- Your answer file should be named as 'Assignment3' as the title.
- In the author section, write down your student number.
- After creating each plot, write a brief description and critical insight about each plot based on your statistical thinking.
- **Use the package 'tidyverse' and the data "USArrests".**
- **Both the codes and outputs should be displayed after compiling the markdown file.**
- You are allowed to use the class notes, or any other help from the internet.
- All computations must be done with the help of suitable R functions. Manual or calculatorbased answers will not be accepted.
- 50% deduction in marks for not including the code or output in your answer script.
- **Create a report following this template with the same Section headers and required details of the analysis of the "USArrests" data.**

1. Description of the dataset

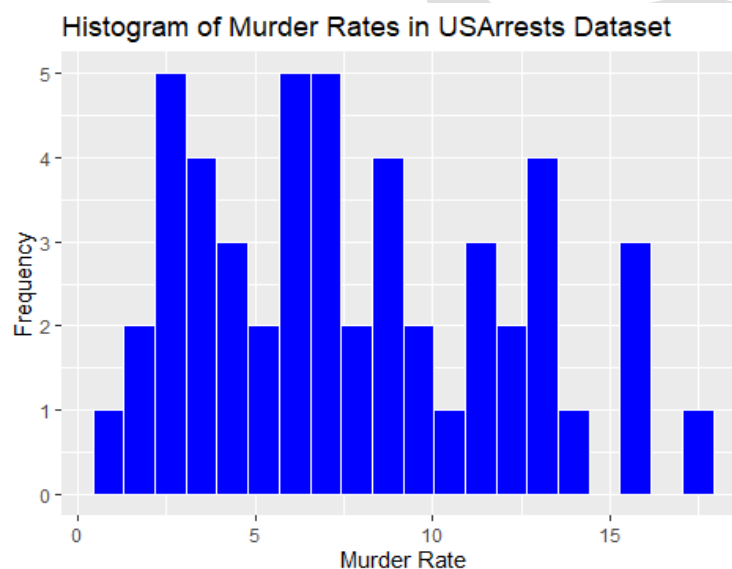
Q1. Begin with a description of the dataset USArrests.

2. Preview of the USArrests dataset

Q2. Then create a preview of the data with the first 6 observations and the structure of the dataset as given below. Add basic statistical summary of the dataset.

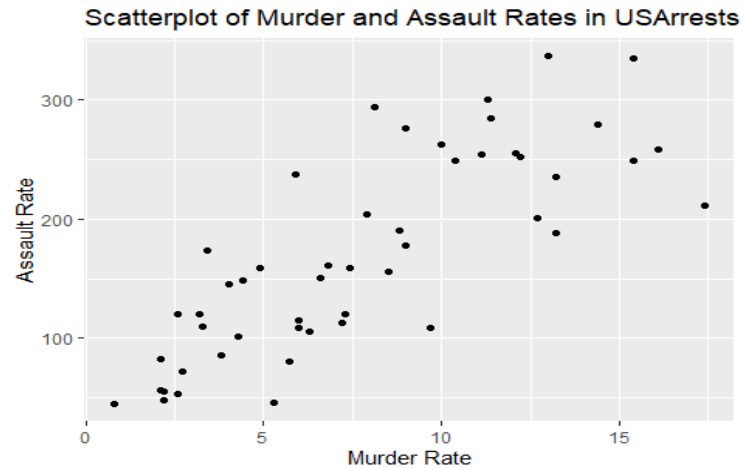
3. A Histogram of the Murder rate in USArrests dataset

Q3. Write a code to recreate the following histogram. Take the number of bins as 20. Provide insights about the histogram.



4. A scatterplot of Murder vs. Assault

Q4. Write a code to recreate the scatterplot. Provide a brief discussion about what you can understand from the plot.



5. A correlation between urban population and the rate of rape

Q5. Write a code to recreate the correlation plot and add some explanation.

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg    ggplot2
```

