

Data Science

Practical No.3

DEPARTMENT OF COMPUTER SCIENCE

Name:	Saurabh Jha	Roll Number	TCS2425023
Paper Code:		Class	TYBSc(Computer Science)
Topic:	exploratory data analysis of mtcars.csv dataset in R	Batch	I
Date:	9/1/2025	Practical No	3

A) AIM: exploratory data analysis of mtcars.csv dataset in R

B) DESCRIPTION:.

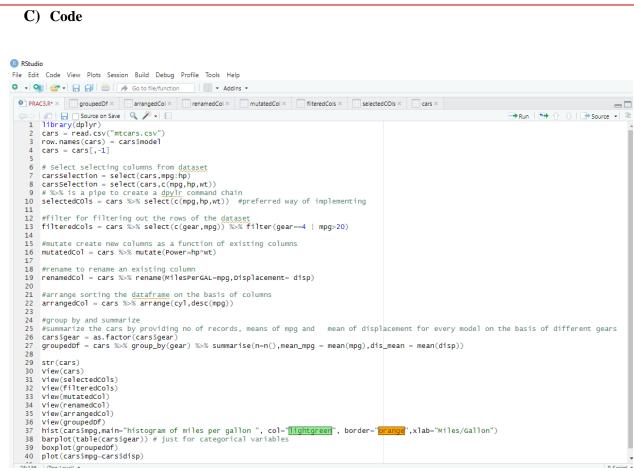
The mtcars.csv dataset consists of various specifications and performance measurements for 32 different car models from the 1970s. It includes both quantitative variables (such as miles per gallon, horsepower, weight, and acceleration) and categorical variables (like the number of cylinders and transmission type). The dataset provides a good mix of continuous and categorical features, making it suitable for performing a comprehensive exploratory data analysis (EDA).

During the EDA, key patterns such as the negative correlation between the car's weight and miles per gallon (mpg) were observed—heavier cars generally have lower fuel efficiency. Further analysis revealed distinct groupings based on the number of cylinders, with cars having more cylinders typically exhibiting higher horsepower but lower mpg. The dataset also allows for an exploration of the relationship between other variables, such as the impact of horsepower on fuel efficiency and how the car's transmission type (manual vs automatic) affects performance metrics.

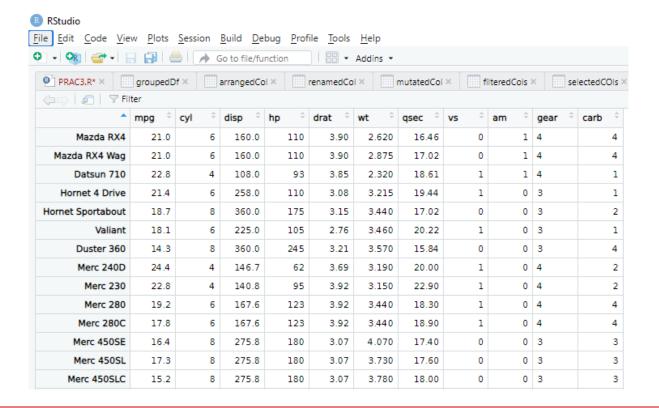
The Core Funtions of dplyr are:

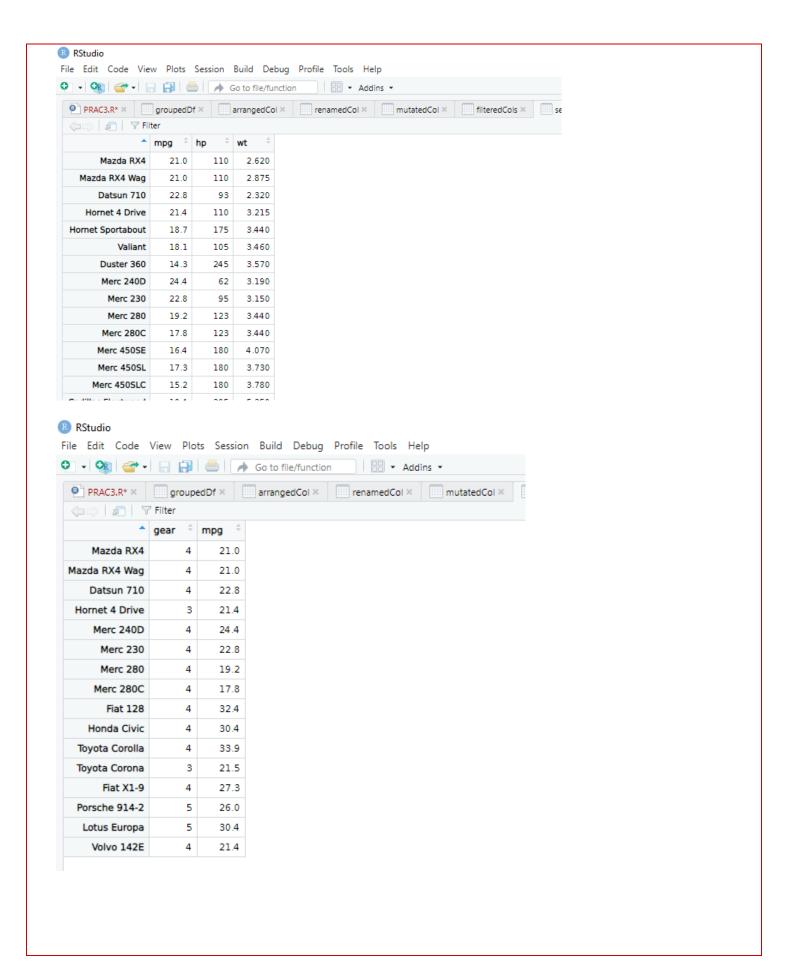
- filter() Selects rows based on conditions.
- select() Chooses specific columns from a dataset.
- mutate() Adds or modifies columns.
- arrange () Sorts rows by one or more columns.
- summarise() Calculates summary statistics for data.
- group by () Groups data by one or more columns.

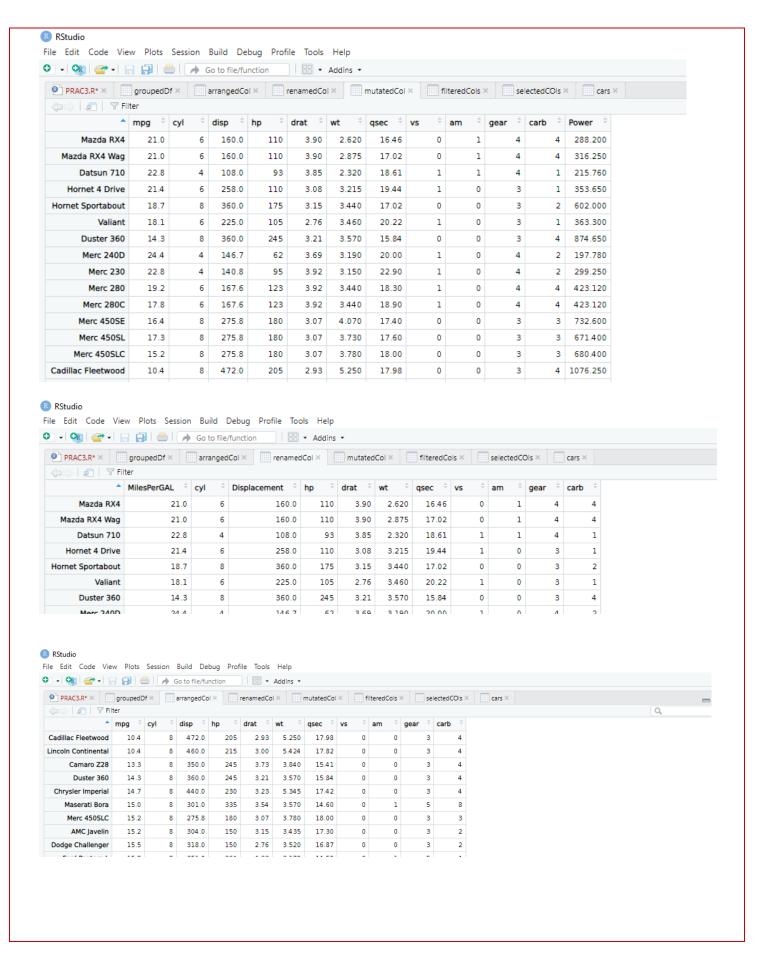
Name of Instructor: Prof. maya Nai

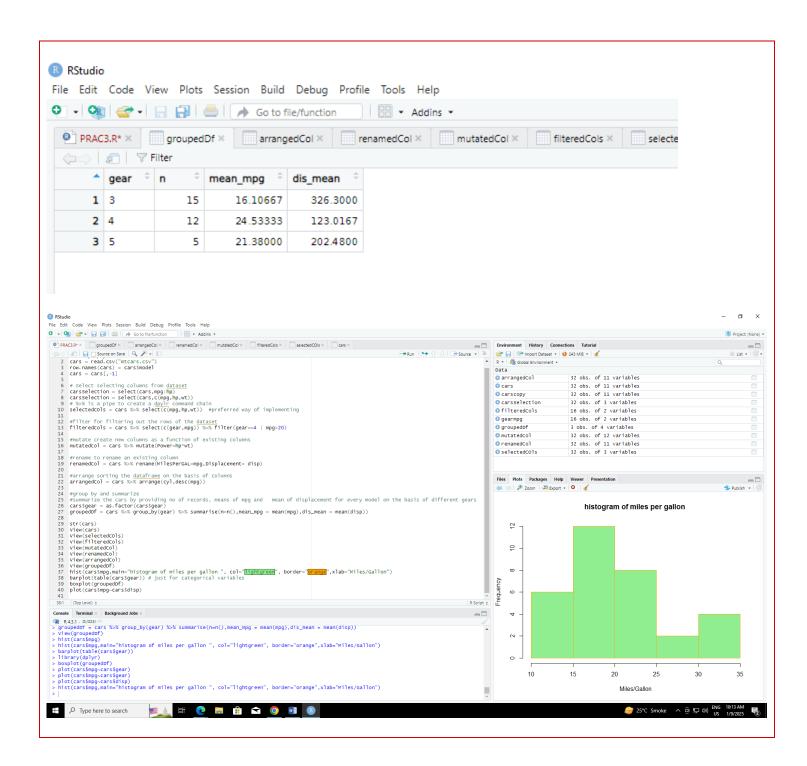


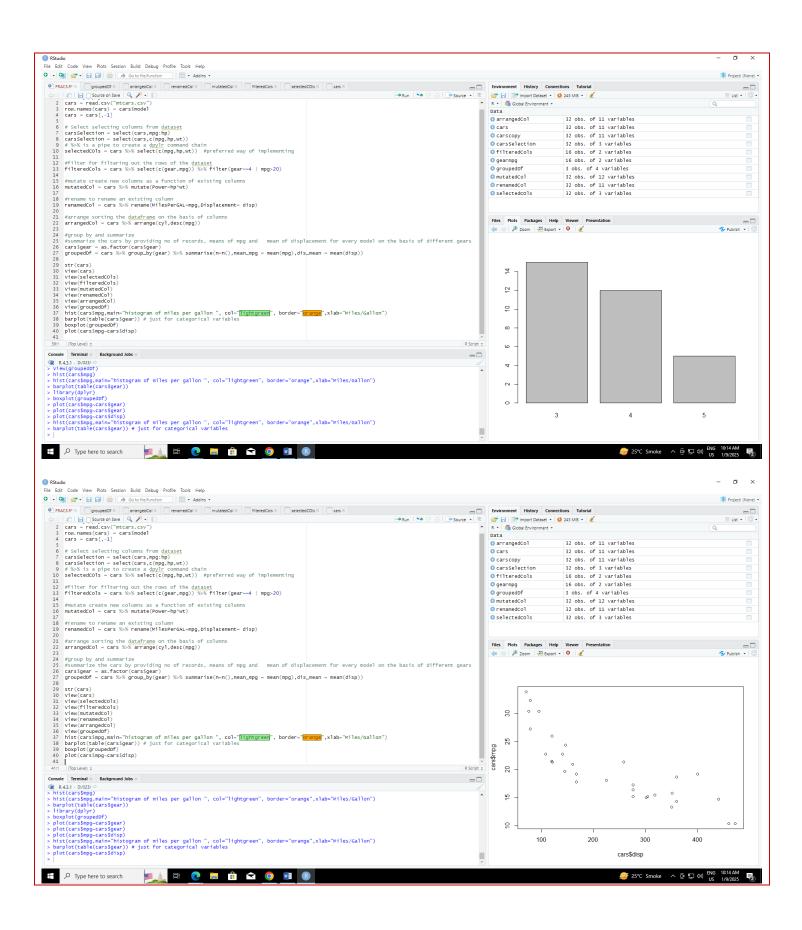
D) Output











Name of Instructor: Prof. Soni Yadav

Name of Instructor: Prof. Soni Yadav