



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 Functions 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.

C) Code

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
PRAC3.R* groupedDf arrangedCol renamedCol mutatedCol filteredCols selectedCOIs cars
1 library(dplyr)
2 cars = read.csv("mtcars.csv")
3 row.names(cars) = cars$model
4 cars = cars[,-1]
5
6 # Select selecting columns from dataset
7 carsSelection = select(cars,mpg:hp)
8 carsSelection = select(cars,c(mpg,hp,wt))
9 # %>% is a pipe to create a dplyr command chain
10 selectedCols = cars %>% select(c(mpg,hp,wt)) #preferred way of implementing
11
12 #filter for filtering out the rows of the dataset
13 filteredCols = cars %>% select(c(gear,mpg)) %>% filter(gear==4 | mpg>20)
14
15 #mutate create new columns as a function of existing columns
16 mutatedCol = cars %>% mutate(Power=hp*wt)
17
18 #rename to rename an existing column
19 renamedCol = cars %>% rename(MilesPerGAL=mpg,Displacement= disp)
20
21 #arrange sorting the dataframe on the basis of columns
22 arrangedCol = cars %>% arrange(cyl,desc(mpg))
23
24 #group by and summarize
25 #summarize the cars by providing no of records, means of mpg and mean of displacement for every model on the basis of different gears
26 cars$gear = as.factor(cars$gear)
27 groupedDf = cars %>% group_by(gear) %>% summarise(n=n(),mean_mpg = mean(mpg),dis_mean = mean(displacement))
28
29 str(cars)
30 View(cars)
31 View(selectedCols)
32 View(filteredCols)
33 View(mutatedCol)
34 View(renamedCol)
35 View(arrangedCol)
36 View(groupedDf)
37 hist(cars$mpg,main="histogram of miles per gallon ", col="lightgreen", border="orange",xlab="Miles/Gallon")
38 barplot(table(cars$gear)) # just for categorical variables
39 boxplot(groupedDf)
40 plot(cars$mpg~cars$displacement)
```

D) Output

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

PRAC3.R* groupedDf arrangedCol renamedCol mutatedCol filteredCols selectedCOIs

Filter

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4
Merc 280C	17.8	6	167.6	123	3.92	3.440	18.90	1	0	4	4
Merc 450SE	16.4	8	275.8	180	3.07	4.070	17.40	0	0	3	3
Merc 450SL	17.3	8	275.8	180	3.07	3.730	17.60	0	0	3	3
Merc 450SLC	15.2	8	275.8	180	3.07	3.780	18.00	0	0	3	3

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

PRAC3.R* groupedDf arrangedCol renamedCol mutatedCol filteredCols se

Filter

	mpg	hp	wt
Mazda RX4	21.0	110	2.620
Mazda RX4 Wag	21.0	110	2.875
Datsun 710	22.8	93	2.320
Hornet 4 Drive	21.4	110	3.215
Hornet Sportabout	18.7	175	3.440
Valiant	18.1	105	3.460
Duster 360	14.3	245	3.570
Merc 240D	24.4	62	3.190
Merc 230	22.8	95	3.150
Merc 280	19.2	123	3.440
Merc 280C	17.8	123	3.440
Merc 450SE	16.4	180	4.070
Merc 450SL	17.3	180	3.730
Merc 450SLC	15.2	180	3.780

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

PRAC3.R* groupedDf arrangedCol renamedCol mutatedCol

Filter

	gear	mpg
Mazda RX4	4	21.0
Mazda RX4 Wag	4	21.0
Datsun 710	4	22.8
Hornet 4 Drive	3	21.4
Merc 240D	4	24.4
Merc 230	4	22.8
Merc 280	4	19.2
Merc 280C	4	17.8
Fiat 128	4	32.4
Honda Civic	4	30.4
Toyota Corolla	4	33.9
Toyota Corona	3	21.5
Fiat X1-9	4	27.3
Porsche 914-2	5	26.0
Lotus Europa	5	30.4
Volvo 142E	4	21.4

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

PRAC3.R* groupedDf arrangedCol renamedCol mutatedCol filteredCols selectedCOLs cars

Filter

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb	Power
Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4	288.200
Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4	316.250
Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1	215.760
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1	353.650
Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2	602.000
Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1	363.300
Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4	874.650
Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2	197.780
Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2	299.250
Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4	423.120
Merc 280C	17.8	6	167.6	123	3.92	3.440	18.90	1	0	4	4	423.120
Merc 450SE	16.4	8	275.8	180	3.07	4.070	17.40	0	0	3	3	732.600
Merc 450SL	17.3	8	275.8	180	3.07	3.730	17.60	0	0	3	3	671.400
Merc 450SLC	15.2	8	275.8	180	3.07	3.780	18.00	0	0	3	3	680.400
Cadillac Fleetwood	10.4	8	472.0	205	2.93	5.250	17.98	0	0	3	4	1076.250

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

PRAC3.R* groupedDf arrangedCol renamedCol mutatedCol filteredCols selectedCOLs cars

Filter

	MilesPerGAL	cyl	Displacement	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

PRAC3.R* groupedDf arrangedCol renamedCol mutatedCol filteredCols selectedCOLs cars

Filter

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Cadillac Fleetwood	10.4	8	472.0	205	2.93	5.250	17.98	0	0	3	4
Lincoln Continental	10.4	8	460.0	215	3.00	5.424	17.82	0	0	3	4
Camaro Z28	13.3	8	350.0	245	3.73	3.840	15.41	0	0	3	4
Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
Chrysler Imperial	14.7	8	440.0	230	3.23	5.345	17.42	0	0	3	4
Maserati Bora	15.0	8	301.0	335	3.54	3.570	14.60	0	1	5	8
Merc 450SLC	15.2	8	275.8	180	3.07	3.780	18.00	0	0	3	3
AMC Javelin	15.2	8	304.0	150	3.15	3.435	17.30	0	0	3	2
Dodge Challenger	15.5	8	318.0	150	2.76	3.520	16.87	0	0	3	2

	gear	n	mean_mpg	dis_mean
1	3	15	16.10667	326.3000
2	4	12	24.53333	123.0167
3	5	5	21.38000	202.4800

Name of Instructor: Prof. maya Nai



