**Use the Automobiles Dataset.**

1. From given data set print first and last five rows

import pandas as pd

df = pd.read\_csv("C:\\Users\\91778\\Desktop\\python ExcelR\\Assignments\\Automobile\_data.csv")

df.head(5)

index company body-style ... horsepower average-mileage price

0 0 alfa-romero convertible ... 111 21 13495.0

1 1 alfa-romero convertible ... 111 21 16500.0

2 2 alfa-romero hatchback ... 154 19 16500.0

3 3 audi sedan ... 102 24 13950.0

4 4 audi sedan ... 115 18 17450.0

[5 rows x 10 columns]

import pandas as pd

df = pd.read\_csv("C:\\Users\\91778\\Desktop\\python ExcelR\\Assignments\\Automobile\_data.csv")

df.tail(5)

index company body-style ... horsepower average-mileage price

56 81 volkswagen sedan ... 85 27 7975.0

57 82 volkswagen sedan ... 52 37 7995.0

58 86 volkswagen sedan ... 100 26 9995.0

59 87 volvo sedan ... 114 23 12940.0

60 88 volvo wagon ... 114 23 13415.0

1. Print All Toyota Cars details (hint: use groupby function)

import pandas as pd

df = pd.read\_csv("C:\\Users\\91778\\Desktop\\python ExcelR\\Assignments\\Automobile\_data.csv")

a=df.groupby('company')

x=a.get\_group('toyota')

print(x)

index company body-style ... horsepower average-mileage price

48 66 toyota hatchback ... 62 35 5348.0

49 67 toyota hatchback ... 62 31 6338.0

50 68 toyota hatchback ... 62 31 6488.0

51 69 toyota wagon ... 62 31 6918.0

52 70 toyota wagon ... 62 27 7898.0

53 71 toyota wagon ... 62 27 8778.0

54 79 toyota wagon ... 156 19 15750.0

1. Count total cars per company (by using value\_counts function)

import pandas as pd

df = pd.read\_csv("C:\\Users\\91778\\Desktop\\python ExcelR\\Assignments\\Automobile\_data.csv")

a=df['company'].value\_counts()

print(a)

toyota 7

bmw 6

nissan 5

mazda 5

mitsubishi 4

audi 4

mercedes-benz 4

volkswagen 4

honda 3

isuzu 3

alfa-romero 3

jaguar 3

porsche 3

chevrolet 3

volvo 2

dodge 2

Name: company, dtype: int64

1. Find each company’s Higesht price car. (hint: Groupby function and then apply .max function)

import pandas as pd

df = pd.read\_csv("C:\\Users\\91778\\Desktop\\python ExcelR\\Assignments\\Automobile\_data.csv")

car\_Manufacturers = df.groupby('company')

priceDf = car\_Manufacturers['company','price'].max()

print(priceDf)

company

alfa-romero alfa-romero 16500.0

audi audi 18920.0

bmw bmw 41315.0

chevrolet chevrolet 6575.0

dodge dodge 6377.0

honda honda 12945.0

isuzu isuzu 6785.0

jaguar jaguar 36000.0

mazda mazda 18344.0

mercedes-benz mercedes-benz 45400.0

mitsubishi mitsubishi 8189.0

nissan nissan 13499.0

porsche porsche 37028.0

toyota toyota 15750.0

volkswagen volkswagen 9995.0

volvo volvo 13415.0

1. Find the average mileage of each car making company (hint: Groupby function and then apply .mean function)

import pandas as pd

df = pd.read\_csv("C:\\Users\\91778\\Desktop\\python ExcelR\\Assignments\\Automobile\_data.csv")

car\_Manufacturers = df.groupby('company')

priceDf = car\_Manufacturers['company','average-mileage'].mean()

print(priceDf)

average-mileage

company

alfa-romero 20.333333

audi 20.000000

bmw 19.000000

chevrolet 41.000000

dodge 31.000000

honda 26.333333

isuzu 33.333333

jaguar 14.333333

mazda 28.000000

mercedes-benz 18.000000

mitsubishi 29.500000

nissan 31.400000

porsche 17.000000

toyota 28.714286

volkswagen 31.750000

volvo 23.000000

1. Sort all cars by Price column (hint: using sort\_values function).

import pandas as pd

df = pd.read\_csv("C:\\Users\\91778\\Desktop\\python ExcelR\\Assignments\\Automobile\_data.csv")

priceDf = df.sort\_values(by=['price'], ascending=False)

print(priceDf)

index company body-style ... horsepower average-mileage price

35 47 mercedes-benz hardtop ... 184 14 45400.0

11 14 bmw sedan ... 182 16 41315.0

34 46 mercedes-benz sedan ... 184 14 40960.0

46 62 porsche convertible ... 207 17 37028.0

12 15 bmw sedan ... 182 15 36880.0

.. ... ... ... ... ... ... ...

27 36 mazda hatchback ... 68 30 5195.0

13 16 chevrolet hatchback ... 48 47 5151.0

22 31 isuzu sedan ... 70 38 NaN

23 32 isuzu sedan ... 70 38 NaN

47 63 porsche hatchback ... 288 17 NaN

1. Check for the null values for entire dataset.

import pandas as pd

df = pd.read\_csv("C:\\Users\\91778\\Desktop\\python ExcelR\\Assignments\\Automobile\_data.csv")

a=df.isnull()

print(a)

1. **Use Crime dataset.**

I) find the aggregations like all moments of business decisions for all columns,value counts.

II) do the plottings like plottings like histogram, boxplot, scatterplot, barplot, piechart,dot chart.

1. **use mtcars dataset.**

A) delete/ drop rows-10 to 15 of all columns

**import pandas as pd**

**df=pd.read\_csv("C:\\Users\\91778\\Desktop\\python ExcelR\\Assignments\\mtcars(1).csv")**

**update\_df=df.drop(df.index[10:16])**

**print(update\_df)**

B)drop the VOL column

There is no VOL column in dataset

1. **Use Bank Dataset.**
2. change all the categorical columns into numerical by creating Dummies and using label encoder.

B) rename all the column names in DF

C) Rename only one specific column in DF

D) After doing all the changes in bank dataset. save the file in your directory in Csv Format.(hint: by using .to\_csv)