**Q12)**Below are the scores obtained by a student in tests

**34,36,36,38,38,39,39,40,40,41,41,41,41,42,42,45,49,56**

1. Find mean,median,variance,standard deviation.

Ans :-

import pandas as pd

import numpy as np

data = pd.read\_csv("D:\\ExcelR Data\\Data Sets\\data.csv")

data.mean()

data["Score"]

np.mean(data.Score)

data['Score'].mean()

data['Score'].median()

data['Score'].mode()

data['Score'].var()

data['Score'].std()

data = pd.read\_csv("D:\\ExcelR Data\\Data Sets\\data.csv")

Output:-

data["Score"]

Out[13]:

0 34

1 36

2 36

3 38

4 38

5 39

6 39

7 40

8 40

9 41

10 41

11 41

12 41

13 42

14 42

15 45

16 49

17 56

Name: Score, dtype: int64

np.mean(data.Score)

Out[14]: 41.0

data['Score'].mean()

Out[15]: 41.0

data['Score'].median()

Out[16]: 40.5

data['Score'].mode()

Out[17]:

0 41

dtype: int64

data['Score'].var()

Out[18]: 25.529411764705884

data['Score'].std()

Out[19]: 5.05266382858645

Q 20) Calculate probability from the given dataset for the below cases

Data \_set: Cars.csv

Calculate the probability of MPG of Cars for the below cases.

MPG<- Cars$MPG

* 1. P(MPG>38)
  2. P(MPG<40)
  3. P (20<MPG<50)

Ans:-

>mean(Cars$MPG)

[1] 34.42208

P(MPG>38):

>sd(Cars$MPG)

[1] 9.131445

>pnorm(38,34.42,9.13)

[1] 0.652513

P(MPG>38)=1-P(MPG<38)(PS: Z-table gives you only less than probabilities)

>1 - 0.65

[1] 0.35

P(MPG<40):

>pnorm(40,34.42,9.13)

[1] 0.7294571

P(20<MPG<50):

>pnorm(50,34.42,9.13)-pnorm(20,34.42,9.13)

[1] 0.8989178

Q 21) Check whether the data follows normal distribution

1. Check whether the MPG of Cars follows Normal Distribution

Dataset: Cars.csv

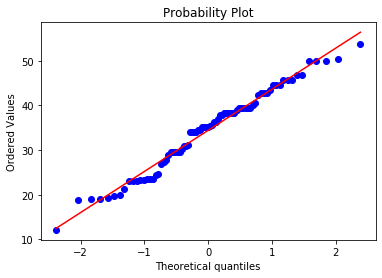
Ans:-

import pandas as pd

cars = pd.read\_csv("D:\\ExcelR Data\\Assignments\\1\\cars.csv")

# Checking Whether data is normally distributed

st.probplot(cars['MPG'], dist="norm",plot=pylab)



1. Check Whether the Adipose Tissue (AT) and Waist Circumference(Waist) from wc-at data set follows Normal Distribution

Dataset: wc-at.csv

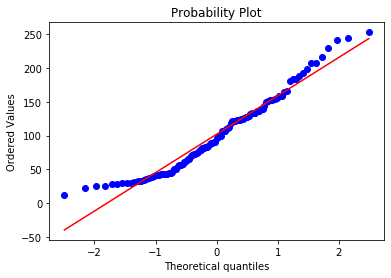
Ans:-for AT=

import pandas as pd

wc\_at = pd.read\_csv("D:\\ExcelR Data\\Assignments\\1\\wc\_at.csv")

# Checking Whether data is normally distributed

st.probplot(wc\_at['AT'], dist="norm",plot=pylab)



Fot Waist=

import pandas as pd

wc\_at = pd.read\_csv("D:\\ExcelR Data\\Assignments\\1\\wc\_at.csv")

# Checking Whether data is normally distributed

st.probplot(wc\_at['Waist'], dist="norm",plot=pylab)

