|  |  |
| --- | --- |
| Activity | Data Type |
| Number of beatings from Wife | Discrete |
| Results of rolling a dice | Discrete |
| Weight of a person | continuous |
| Weight of Gold | Continuous |
| Distance between two places | Continuous |
| Length of a leaf | Continuous |
| Dog's weight | Continuous |
| Blue Color | Discrete |
| Number of kids | continuous |
| Number of tickets in Indian railways | Discrete |
| Number of times married | Discrete |
| Gender (Male or Female) | Continuous |

Q1) Identify the Data type for the Following:

Q2) Identify the Data types, which were among the following

Nominal, Ordinal, Interval, Ratio.

|  |  |
| --- | --- |
| Data | Data type |
| Gender | Nominal |
| High School Class Ranking | Ordinal |
| Celsius Temperature | Interval |
| Weight | Ratio |
| Hair Color | Nominal |
| Socioeconomic Status | Ordinal |
| Fahrenheit Temperature | Interval |
| Height | Ratio |
| Type of living accommodation | Ordinal |
| Level of Agreement | Ordinal |
| IQ(Intelligence Scale) | Ordinal |
| Sales Figures | Ratio |
| Blood Group | nominal |
| Time Of Day | Interval |
| Time on a Clock with Hands | Interval |
| Number of Children | Ordinal |
| Religious Preference | nominal |
| Barometer Pressure | Interval |
| SAT Scores | Ratio |
| Years of Education | Interval |

Q3) Three Coins are tossed, find the probability that two heads and one tail are obtained?

**Ans** Three coins are tossed the total number of possible combination are (23) = 8.the possible outcomes are HHH,HHT,HTH,THH,TTH,THT,HHT,TTT.

We can see that the possibility of two head and one tail is 3 times in the possible outcomes.

So we can say that The probability of getting two head and one tail is 3\8.

Q4) Two Dice are rolled, find the probability that sum is

1. Equal to 1
2. Less than or equal to 4
3. Sum is divisible by 2 and 3

**Ans**

a) 0\36

b)6\36

c)24\36

Q5) A bag contains 2 red, 3 green and 2 blue balls. Two balls are drawn at random. What is the probability that none of the balls drawn is blue?

**Ans**

None of the balls drawn is blue, this can only happen when the two balls drawn at random are either red and green or both.

Total number of balls = 2 + 3 + 2 = 7

Required Probability = 10/21

Q6) Calculate the Expected number of candies for a randomly selected child

Below are the probabilities of count of candies for children (ignoring the nature of the child-Generalized view)

|  |  |  |
| --- | --- | --- |
| CHILD | Candies count | Probability |
| A | 1 | 0.015 |
| B | 4 | 0.20 |
| C | 3 | 0.65 |
| D | 5 | 0.005 |
| E | 6 | 0.01 |
| F | 2 | 0.120 |

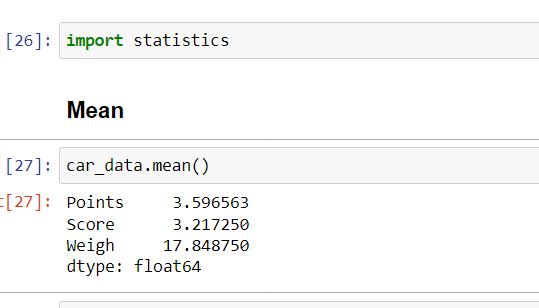
Child A – probability of having 1 candy = 0.015.

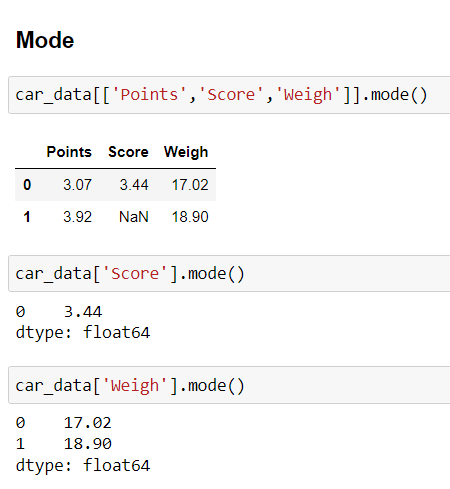
Child B – probability of having 4 candies = 0.20

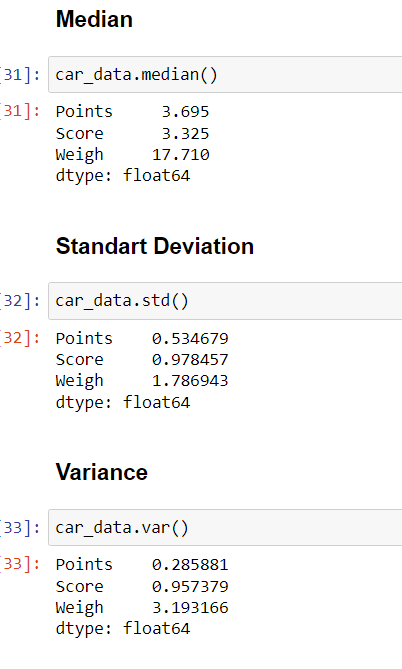
Q7) Calculate Mean, Median, Mode, Variance, Standard Deviation, Range & comment about the values / draw inferences, for the given dataset

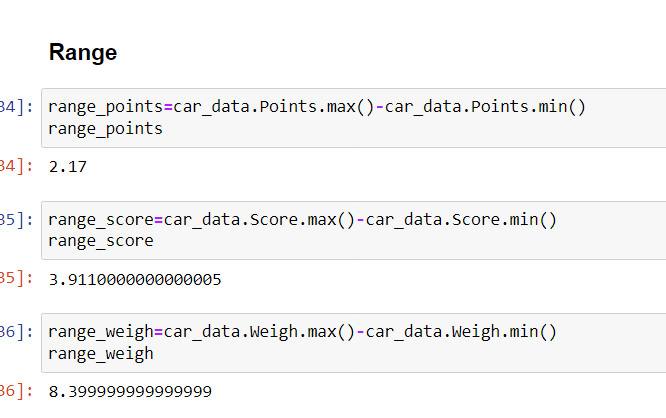
* For Points,Score,Weigh>

Find Mean, Median, Mode, Variance, Standard Deviation, and Range and also Comment about the values/ Draw some inferences.

Answer:- 







**Use Q7.csv file**

Q8) Calculate Expected Value for the problem below

1. The weights (X) of patients at a clinic (in pounds), are

108, 110, 123, 134, 135, 145, 167, 187, 199

Assume one of the patients is chosen at random. What is the Expected Value of the Weight of that patient?

# Answer:- - The chance of getting one patient is 1\9 in each value

weight of patients = 108, 110, 123, 134, 135, 145, 167, 187, 199

probability = 1\9,1/9,1\9,1/9,1/9,1/9,1/9,1/9,1/9

expected value = (1/9)(108) + (1/9)110 + (1/9)123 + (1/9)134 + (1/9)135 + (1/9)145 + (1/9(167) + (1/9)187 + (1/9)199

(1/9) ( 108 + 110 + 123 + 134 + 135 + 145 + 167 + 187 + 199)

(1/9)(1308)

145.33

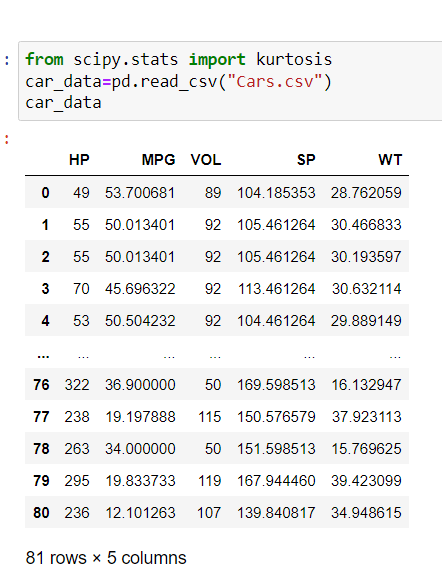
**Q9) Calculate Skewness, Kurtosis & draw inferences on the following data**

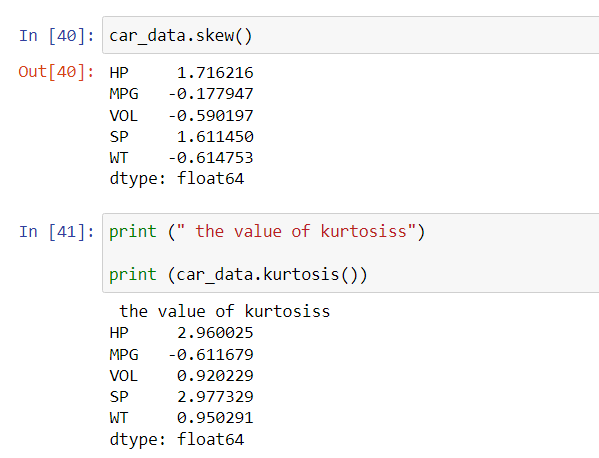
**Cars speed and distance**

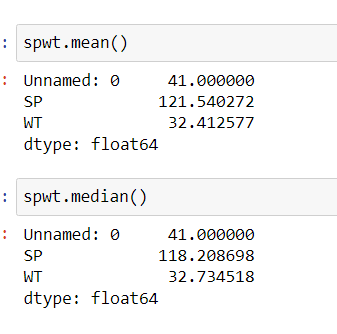
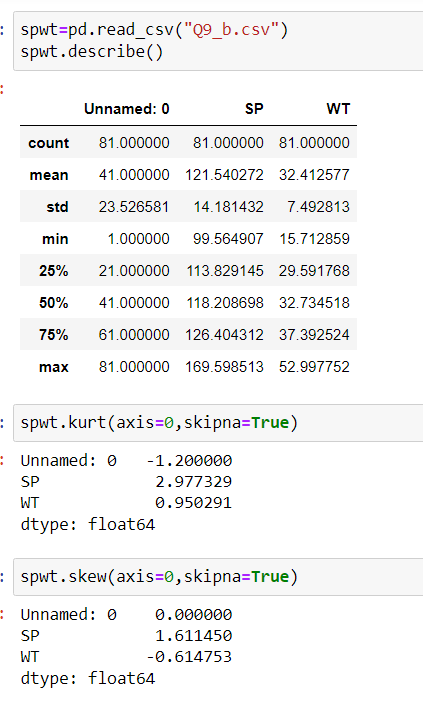
**Use Q9\_a.csv**

**SP and Weight(WT)**

**Use Q9\_b.csv**

**Answer:- **

****

****

**Q10) Draw inferences about the following boxplot & histogram**



Answer:- This graph histogram showing that the distribution of the data is asymmetric.this is right skewed and we also say that the this is positively skewed.

The boxplot showing positively skewness of the distribution.



**Q11)** Suppose we want to estimate the average weight of an adult male in Mexico. We draw a random sample of 2,000 men from a population of 3,000,000 men and weigh them. We find that the average person in our sample weighs 200 pounds, and the standard deviation of the sample is 30 pounds. Calculate 94%,98%,96% confidence interval?

Answer:- population size=3,000,000 sample size(n)=2,000

Sample mean=200 sample standard deviation(s)=30

For calculating the t value:

From scipy import stats

From math import sqrt

stats.t.ppf(CI,df)

t value for 94%= 1.8818614764780113

t value for 96%=2.055089962825778

t value for 98%= 2.328214776106972

For calculating the confidence interval Stats.norm.interval(x,mean,scale=(s/√n))

Confidence interval of 94% is [198.738325292158, 201.261674707842]

Confidence interval of 98% is [198.62230334813333, 201.37769665186667]

Confidence interval of 96% is [198.43943840429978, 201.56056159570022]

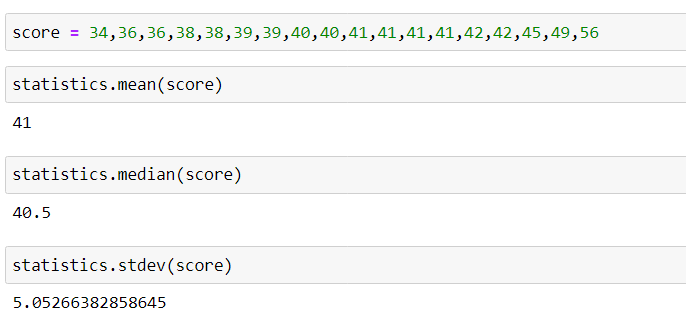
.

**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.
2. What can we say about the student marks?

Answer:-

1. Average student marks is 41

Q13) What is the nature of skewness when mean, median of data are equal? Answer:- Skewness is zero

Q14) What is the nature of skewness when mean > median ?

Answer: positively or right skewed

Q15) What is the nature of skewness when median > mean?

Answer:- Negatively skewed or left skewed

Q16) What does positive kurtosis value indicates for a data ?

Answer:- This shows that the distribution has heavier tails than the normal distribution.

Q17) What does negative kurtosis value indicates for a data?

Answer:- This shows that the distribution has lighter tails than the normal distribution.

Q18) Answer the below questions using the below boxplot visualization.



What can we say about the distribution of the data?

The distribution of data on the high end of the scale

What is nature of skewness of the data?

The distrubution is skewed left

What will be the IQR of the data (approximately)?

The iqr is 18 minus 10

Q19) Comment on the below Boxplot visualizations?



Answer:- Both the boxplots show some symmetry means that they are normally distributed. There is no outlier, median is around260.

Draw an Inference from the distribution of data for Boxplot 1 with respect Boxplot 2.

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

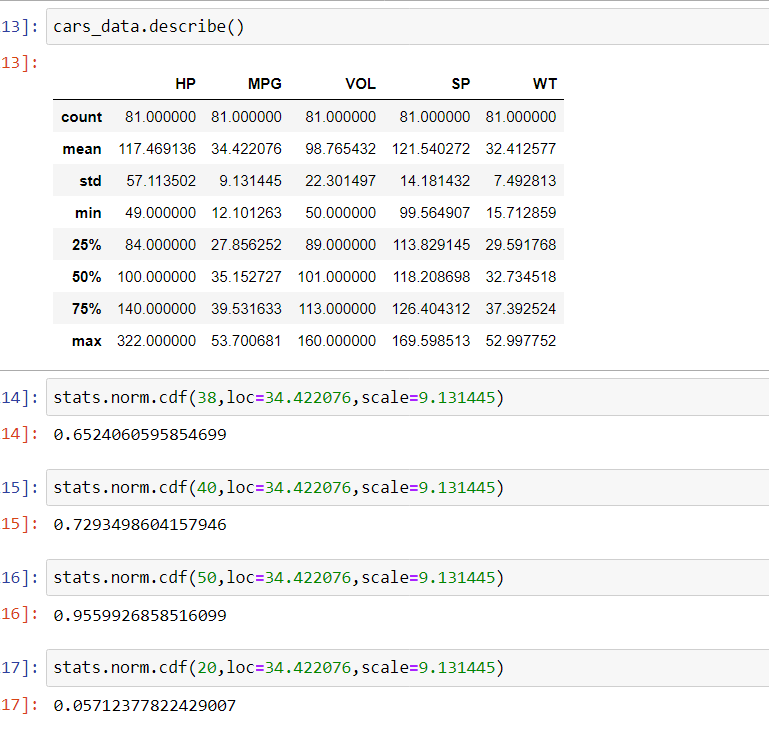
Data \_set: Cars.csv

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

MPG <- Cars$MPG

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

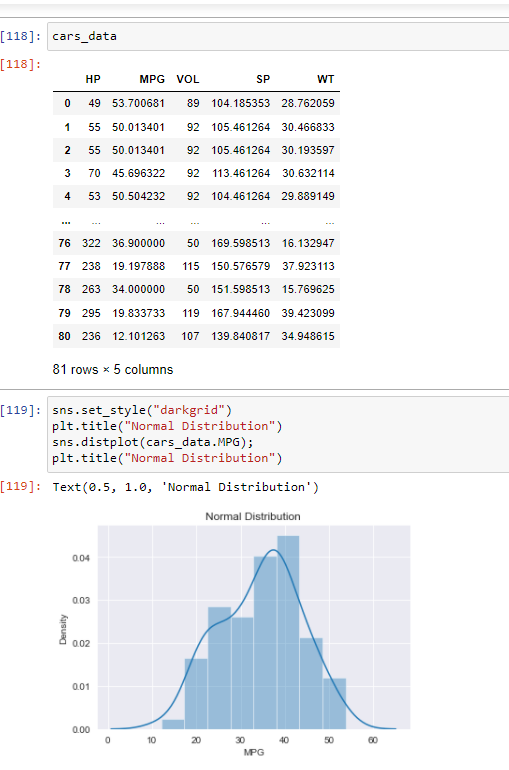
c. P (20<MPG<50)

Answer:- 

Q 21) Check whether the data follows normal distribution

1. Check whether the MPG of Cars follows Normal Distribution

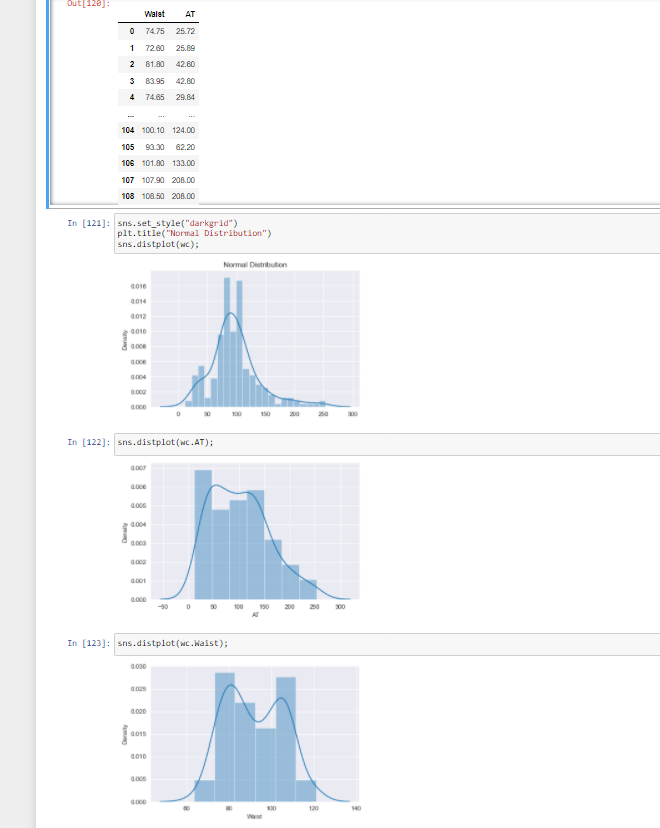
Dataset: Cars.csv

Answer:- 

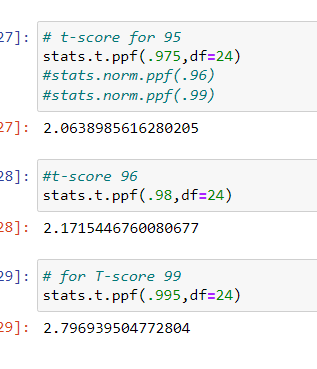
1. Check Whether the Adipose Tissue (AT) and Waist Circumference(Waist) from wc-at data set follows Normal Distribution

Dataset: wc-at.csv

Answer:-

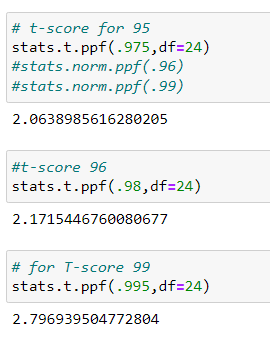


Q 22) Calculate the Z scores of 90% confidence interval,94% confidence interval, 60% confidence interval

Answer:- 

Q 23) Calculate the t scores of 95% confidence interval, 96% confidence interval, 99% confidence interval for sample size of 25

Answer:-

- 

Q 24**)** A Government company claims that an average light bulb lasts 270 days. A researcher randomly selects 18 bulbs for testing. The sampled bulbs last an average of 260 days, with a standard deviation of 90 days. If the CEO's claim were true, what is the probability that 18 randomly selected bulbs would have an average life of no more than 260 days

Hint:

rcode 🡪 pt(tscore,df)

df 🡪 degrees of freedom

Answer:-

