

Q1) Identify the Data type for the Following:

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

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

Nominal, Ordinal, Interval, Ratio.

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

| | |
|----------------------|----------|
| Religious Preference | Ratio |
| Barometer Pressure | Interval |
| SAT Scores | Ratio |
| Years of Education | Nominal |

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

Answer :- three coins are toss = { HHH,THH,HTH,HHT,TTT,HTT,THT,TTH }

$$\begin{aligned}\text{Probability} &= \text{favourable condition} / \text{total toss outcome} \\ &= 3/8\end{aligned}$$

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

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

Answer :- total outcome = 36 [(1,1) (1,2) (1,3) (1,4) (1,5) (1,6)

(2,1) (2,2) (2,3) (2,4) (2,5) (2,6)

(3,1) (3,2) (3,3) (3,4) (3,5) (3,6)

(4,1) (4,2) (4,3) (4,4) (4,5) (4,6)

(5,1) (5,2) (5,3) (5,4) (5,5) (5,6)

(6,1) (6,2) (6,3) (6,4) (6,5) (6,6)

a)

$$\begin{aligned}\text{b) Sum(less than or equal to 4)} &= [(1,1) (1,2) (1,3) (2,1) (2,2) (3,1)] \\ &= 6/36\end{aligned}$$

$$= 1/6$$

$$\text{c) sum(divisible by 2 and 3)} = [(1,5) (2,4) (3,3) (4,2) (5,1) (6,6)]$$

$$=6/36$$

$$=1/36$$

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?

Answer:- we have total 7 ball , if two ball drawn at random

$$\text{Then Case 1} = 2 \text{ red} + 3 \text{ green} / 7$$

$$= 5/7$$

Suppose 1 of red or green ball are drawn

$$\text{Case 2} = 1 \text{ red} + 3 \text{ green}$$

$$= 4/7$$

$$\text{Then, } = 5/7 \times 4/6$$

$$= 10/21$$

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

Answer :- Child A = 0.015

Child B = 0.20

$$\begin{aligned} &= 1 \times 0.015 + 4 \times 0.20 + 3 \times 0.65 + 5 \times 0.005 + 6 \times 0.01 + 2 \times 0.120 \\ &= 0.015 + 0.80 + 1.95 + 0.025 + 0.06 + 0.24 \\ &= 3.09 \end{aligned}$$

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 :- answer in the statistic assignment.py file

Use Q7.csv file

Q8) Calculate Expected Value for the problem below

a) 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 :-

Expected value = $\sum [\text{probability} \times \text{value}]$

9 patient and 1 is selected at random

Probability = $1/9$

$$\begin{aligned} &= 1/9 \times (108 + 110 + 123 + 134 + 135 + 145 + 167 + 187 + 199) \\ &= 1/9 \times 1308 \\ &= 145.333 \end{aligned}$$

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

Cars speed and distance

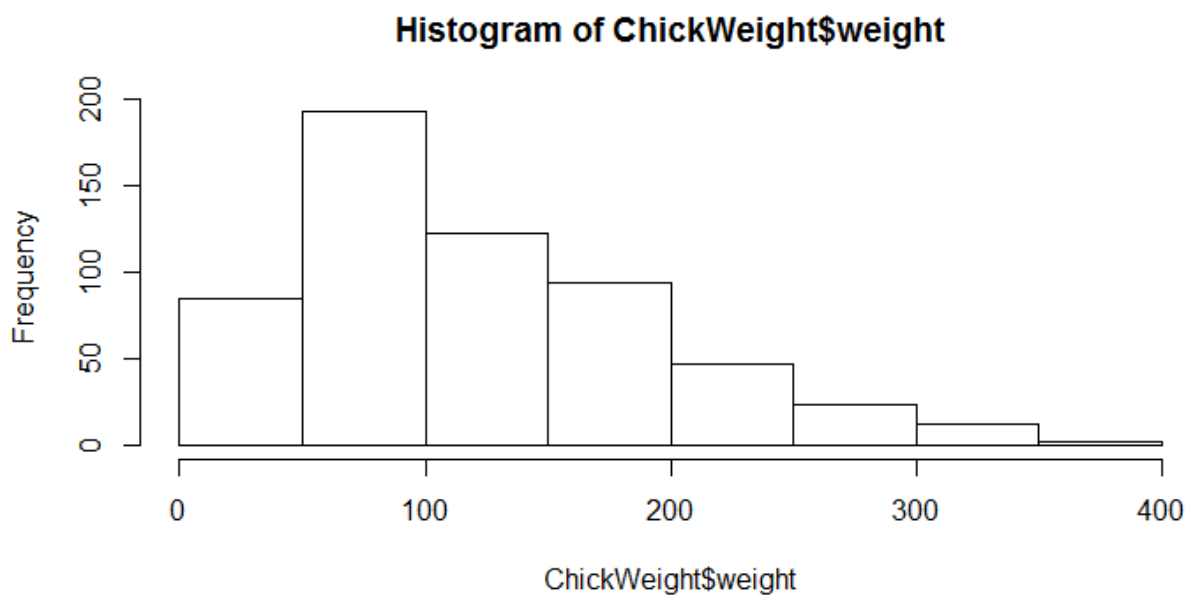
Use Q9_a.csv

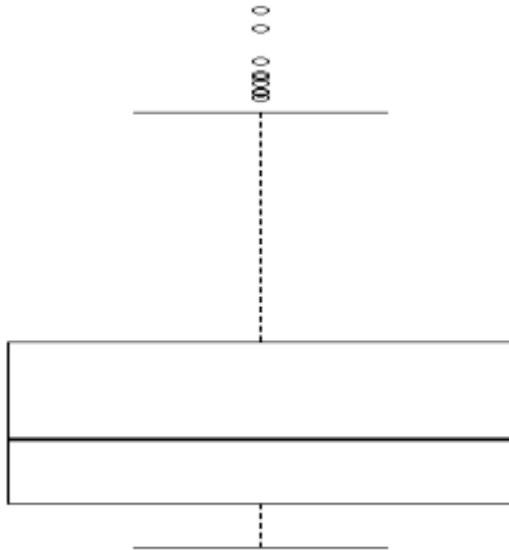
SP and Weight(WT)

Use Q9_b.csv

Answer:- statistic assignment-Q9[a,b]. ipyb file

Q10) Draw inferences about the following boxplot & histogram





Answer:-

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?

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 :- statistic assingment-1(Q12) [NOTE :- 2 -> WHAT IS THAT MEAN,second ans not written]

Q13) What is the nature of skewness when mean, median of data are equal?

Answer:- skewness can be positive negative or zero,when mean median mode are equal then there is no skewness.

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

Answer= If the mean is greater than the median, the distribution is positively skewed. If the mean is less than the median, the distribution is negatively skewed.

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

Answer:- if the mean is less than the median, the distribution is negatively skewed.

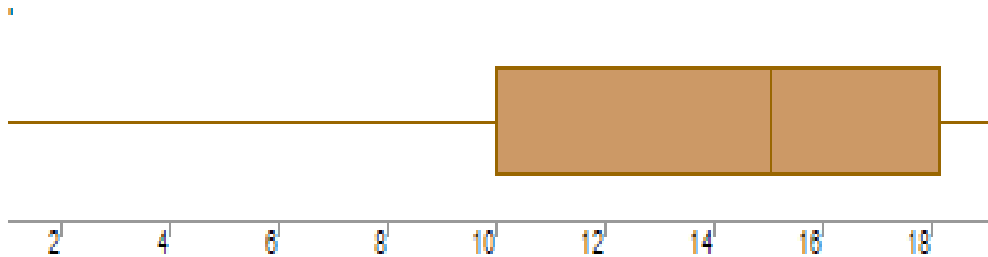
Q16) What does positive kurtosis value indicates for a data ?

Answer:- Positive values of kurtosis indicate that distribution is peaked and possesses thick tails.

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

Answer:- negative kurtosis indicate that a **distribution is flat and thin tails**.

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

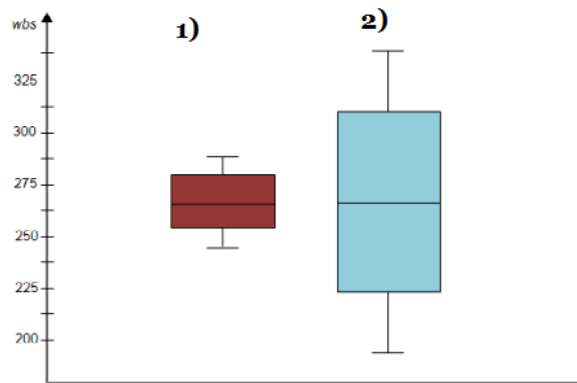


What can we say about the distribution of the data?

What is nature of skewness of the data?

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

Q19) Comment on the below Boxplot visualizations?



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

Answer :- By observing both the plots whisker's level is high in boxplot 2, mean and median are equal hence distribution is symmetrical

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`

- a. $P(\text{MPG} > 38)$
- b. $P(\text{MPG} < 40)$
- c. $P(20 < \text{MPG} < 50)$

Q 21) Check whether the data follows normal distribution

a) Check whether the MPG of Cars follows Normal Distribution

Dataset: Cars.csv

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

Dataset: wc-at.csv

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

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

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 \rightarrow pt(tscore,df)

df \rightarrow degrees of freedom