ASSIGNMENT -01

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 | 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 | 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 | ORDINAL |
| Celsius Temperature | INTERVAL |
| Weight | RATIO |
| Hair Color | NOMINAL |
| Socioeconomic Status | NOMINAL |
| Fahrenheit Temperature | INTERVAL |
| Height | RATIO |
| Type of living accommodation | NOMINAL |
| Level of Agreement | INTERVAL |
| IQ(Intelligence Scale) | INTERVAL |
| Sales Figures | RATIO |
| Blood Group | NOMINAL |
| Time Of Day | RATIO |
| Time on a Clock with Hands | INTEREVAL |
| Number of Children | RATIO |
| Religious Preference | ORDINAL |
| Barometer Pressure | RATIO |
| SAT Scores | INTERVAL |
| Years of Education | RATIO |

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



Ans:

Here we can see that the major Chick weights fall in the category of 50-100g(measures in x) as the maximum which is 200.The minimum weights have a frequency if less than or equal to 5.

The plot is Right skewed which show that there is lesser concentration of chick weights in the 300-400gram category.

The expected value should be above 46.45



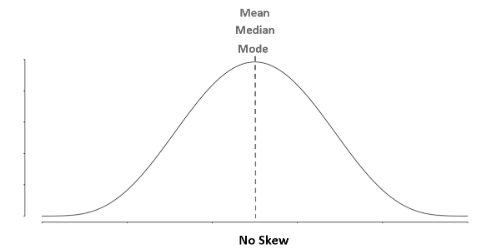
Ans:

Median is less than mean right skewed and we have outlier on the upperside of box plot and there is less data points between Q1 and bottom point.

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

ANS:

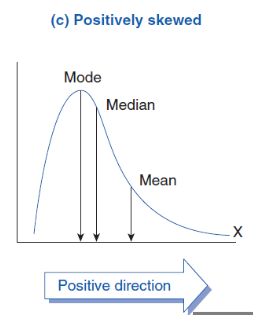
The nature of the skewness is when mean and median is equal is data skewness is zero, it mean our data is in symmetrical distribution, our data is in the bell shaped curve



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

ANS:

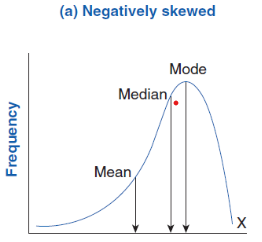
THE NATURE OF THE SKEWNESS WHEN MEAN >MEDIAN IS WE GET THE POSITIVELY SKEWED DISTRIBUTION



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

Ans:

We get the negatively skewed distribution



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

Ans:

A distribution with a positive kurtosis value indicates that the distribution has heavier tails than the normal distribution. For example, data that follow a t distribution have a positive kurtosis value.

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

Ans:

A distribution with a negative kurtosis value indicates that the distribution has lighter tails than the normal distribution.

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



1.What can we say about the distribution of the data?

Ans: unsymmetrically distributed

2.What is nature of skewness of the data?

Ans: negatively distributed

3.What will be the IQR of the data approximately?

Ans: IQR=Q3-Q1= 08

Q19) Comment on the below Boxplot visualizations?



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

Ans:

Here there is a representation of 2 box plots in which box plot 2) is highly distributed across the plane and 1) is slightly less distributed.(variance)Whiskers in these diagrams also show this.100% of the data is spread across values from 350 in 2 whereas its spread in range 250-290 app x in 1)

Here when we compare box plot 1 with box plot 2 we can say that the data in box plot 1 is widely spread. Here the main inference is that since the data range varies high in box plot 2 it is hard to make a prediction in box plot 2. The median in the 2 box plots are equal. And the data spread in both of them are symmetrical