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 | Ordinal |
| Fahrenheit Temperature | Interval |
| Height | Ratio |
| Type of living accommodation | Ordinal |
| Level of Agreement | Ordinal |
| IQ(Intelligence Scale) | Ratio |
| Sales Figures | Interval |
| Blood Group | Nominal |
| Time Of Day | Ratio |
| Time on a Clock with Hands | Ratio |
| Number of Children | Ordinal |
| Religious Preference | Nominal |
| Barometer Pressure | Ratio |
| SAT Scores | Ratio |
| Years of Education | Nominal |

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

Ans: Total number of possible combinations = 2n= 23=8

The combinations are HHH, HHT, HTH, THH, TTH, THT, HTT, TTT.

Number of combinations that have two heads and one tail = 3,

i.e., HHT, HTH, TTH

P=3/8.

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

1. Equal to 1

Ans: If two dices were rolled, then total possible cases =36

The combinations are=

[(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)].

As minimum sum are 2 for outcome (1, 1).

Hence, probability is 0.

1. Less than or equal to 4

Ans: combinations are= [(1, 3), (2, 2), (3, 1)].

Hence, probability is=3/36=1/12.

1. Sum is divisible by 2 and 3

Ans: combinations are= [(1 , 5) , (3 , 3) , (4 , 2) , (5 , 1) , (6 , 6)].

Hence, probability is=5/36.

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: P=5c2/7c2=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 |

Ans: Expected number of candies for a randomly selected child

= 1 \* 0.015 + 4\*0.20 + 3 \*0.65 + 5\*0.005 + 6 \*0.01 + 2 \* 0.12

= 0.015 + 0.8 + 1.95 + 0.025 + 0.06 + 0.24

= 3.09

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?

Ans: 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

Q10) Draw inferences about the following boxplot & histogram





Ans: The boxplot has outliers on the maximum side.

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?

Ans: from above plot we can say that mean of marks of student is 41 which is slightly greater than median.

Most of the students got marks in between 41-42, there are two outlier 49, 56.

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

Ans: Nature of skewness, skewness can be positive, negative or Zero. When the value of mean, median and mode are equal there is no skewness. And skewness is a measure of the asymmetric of probability distribution of real valued mean.

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

Ans: If the mean is greater than the median, the distribution is positively skewed.

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

Ans: If the mean is less than the median, the distribution is negatively skewed.

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

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

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.



a).What can we say about the distribution of the data?

Ans: The above Boxplot is not normally distributed the median is towards the higher value

b).What is nature of skewness of the data?

Ans: The data is a skewed towards left. The whisker range of minimum value is greater than maximum.

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

Ans: The Inter Quantile Range = Q3 Upper quartile – Q1 Lower Quartile = 18 – 10 =8

Q19) Comment on the below Boxplot visualizations?



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

Ans: First there are no outliers. Second both the box plot shares the same median that is approximately in a range between 275 to 250and they are normally distributed with zero to no skewness neither at the minimum or maximum whisker range