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

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) | Ratio |
| Sales Figures | Ratio |
| Blood Group | Nominal |
| Time Of Day | Ordinal |
| Time on a Clock with Hands | Interval |
| Number of Children | Nominal |
| Religious Preference | Nominal |
| Barometer Pressure | Interval |
| SAT Scores | Interval |
| Years of Education | Ordinal |

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

=3/8

= 0.375 =37.5%

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
4. 0/36 =0
5. 6/36=0.166=16.66%
6. 6/36=0.166=16.66%

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?

=10/21 =0.4761

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

= (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

= 3.090

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?

Expected value=sum(probability \* value)

Probability= 1/9

= (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**

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



Histogram:-

1.Chick weight data histogram has right skewed or positively skewed.

2.More than 50% Chick Weight is between 50 to 150.

3.Most of the chick weight is between 50 to 100.



The histograms peak has right skew and tail is on right. Mean > Median. We have outliers on the higher side.

The boxplot has outliers on the maximum side.

Boxplot=

1.The data is right skewed.

2.There are outliers at 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.

Mean=41 Median=40.5 variance =25.52941 Standard deviation=5.052664

2)What can we say about the student marks?

There are no outliers because mean is greater than median

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

= Zero skewed

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

=Positively skewed/Right

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

=Negatively skewed/Left

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

= Distribution is peaked/Leptokurtic

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

= Distribution is flat/Platykurtic

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



What can we say about the distribution of the data?

= The above Boxplot is not normally distributed.

What is nature of skewness of the data?

= The data is left skewed the median is greater than mean

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

- In above visualization mean are approximate same Boxplot 2 has higher range campare to Boxplot 1

-The maximum and minimum values has much more difference in Boxplot 1 and Boxplot 2

- Both Boxplot are Normally Distributed

- No outlier’s in both boxplot