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

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

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

P(c) = [HHH,HHT,HTH,THH,TTT,TTH,THT,HTT]

=1/8

ANS= 3/8=0.375

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 -total possible outcome = 6^2=36

Ans-a=0\36=0

Ans-b=6\36=0.17

Ans-c=6\36=0.17

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-total no of ball=7,Drawing 2 ball out of 7=7c2=21,drawing two balls out of 5=5c2=10,

ans=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

Ans-1\*0.015+4\*0.20+3\*0.65+5\*0.005+6\*0.01+2\*0.120

=3.09

**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?

Ans-expected value of a descrete random is actually the mean

Mean is 1308\9

=145.33

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

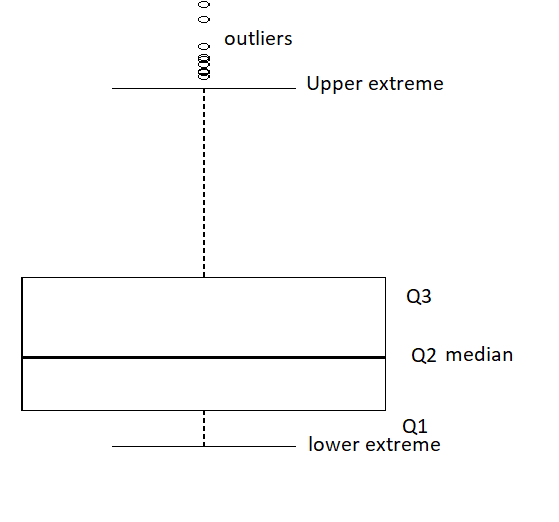
**Positively skewed**



Infrences of histogram

1.Right Skewed,value is positive

2.Mass is concentrated on left hand side



**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?

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

Ans- symmetric

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

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

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

Ans- sharper peak

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

Ans-flatter peak

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



What can we say about the distribution of the data?

Ans-data is not distributed equally,outliers influencing the data ,median=14.7

What is nature of skewness of the data?

Ans-left skew

What will be the IQR of the data (approximately)?  
iqr = q3-q1=18-10=8

Q19) Comment on the below Boxplot visualizations?

boxplot1-is less distributed



boxplot2-is highly distributed

Median are equal for both

Data spread in both of them are symmetrical

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)

Q 21) Check whether the data follows normal distribution

1. Check whether the MPG of Cars follows Normal Distribution

Dataset: Cars.csv

1. 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