|  |  |
| --- | --- |
| 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 |

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

Q3) Three Coins are tossed, find the probability that two heads and one tail are obtained?  
  
Ans- P=(HHH,TTT,HHT,THH,HTH,THT,HTT,TTH).  
Probability of getting two heads and one tail is 3 out of 8 combination.

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- P=(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) Equal to 1 = Probability 0  
b) Less than or equal to 4 = Probability 6  
c) Sum divisible by 2 = Probability 18  
 Sum divisible by 3 = Probability 12

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-(RR, RG, RB), (GG,GR,GB), (BG,BR,BB)  
Probability of getting none ball blue is 4.

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

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.

**Ans.**

A screenshot of a computer

Description automatically generated

Mode

A screenshot of a computer

Description automatically generated

Variances

A screenshot of a computer code

Description automatically generated

A screenshot of a computer

Description automatically generated

**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**

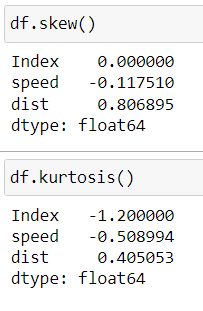
****

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

**Cars speed and distance**

**Use Q9\_a.csv**

**A screenshot of a computer

Description automatically generatedAns  
**

Median = Mean – Symmetric distribution

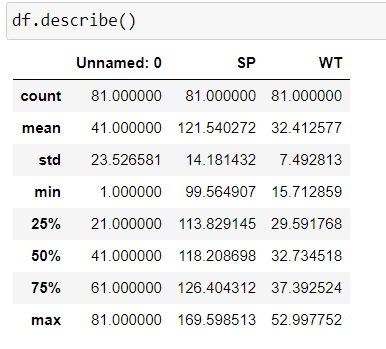
**SP and Weight(WT)**

**Use Q9\_b.csv**

**A screenshot of a computer code

Description automatically generatedA screenshot of a computer

Description automatically generated**

****

**Mean = Median – Symmetric distribution.**

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



ANS.

Chick Weight range between 50 to 100 is frequently occurring, and as weight increases occurrence is reduced.



Ans.   
According to box pot,

* There are more out liers/Single data points.
* Upper whisker is more.
* Lower Whisker is less
* 1st quartile has more values compared to 3rd quartile.
* Un symmetric data.

**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?

A screenshot of a formula

Description automatically generated  
ANS=

**Sample mean = 200**

**Z = 94% - 1.88, 96%-1.96, 98%-2.33**

**S = 30**

**n = 2000**

**CI – 94%**

**Lower Limit – 200 – 1.88 (30/**√2000**) = 198.74**

**Upper Limit – 200 + 1.88 (30/**√2000**) = 201.26**

**CI – 96%**

**Lower Limit – 200 – 1.96(30/**√2000**) = 198.69**

**Upper Limit – 200 + 1.96 (30/**√2000**) = 201.31**

**CI – 98%**

**Lower Limit – 200 – 2.33(30/**√2000**) = 198.44**

**Upper Limit – 200 + 2.33 (30/**√2000**) = 201.56**

**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=

Mean = sum of values/Number of vales = 41

Median value = 40.5



Variance = 25.52

Standard Deviation = 5.05

As standard deviation is height there will be more deviation in scores.

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

Ans – Mean=median, then data is symmetric data.

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

Ans – Mean > Median, data positively skewed and distribution has long tail on right side.

Q15) What is the nature of skewness when median > mean?  
Ans- Median > Mean, negatively skewed, distribution has long tail on left side.

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

Ans – The tail of the distribution is heaver or fatter than normal distribution.

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

Ans - The tail of the distribution is Lighter or thiner than normal distribution.

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)?

Ans - Negatively skewed, Median > Mean, IQR=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 –

* Both Box plot has median at 265,
* 1 box plot Iqr = 25, 2 box plot iqr = 75. Two different range of iqr,

2.IQR > 1.IQR

* No Outliers in both box plot.

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

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 🡪 pt(tscore,df)

df 🡪 degrees of freedom