|  |  |
| --- | --- |
| Activity | Data Type |
| Number of beatings from Wife | Discreate |
| Results of rolling a dice | Continuous |
| 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 | Discreate |
| Number of kids | Discreate |
| Number of tickets in Indian railways | Discreate |
| Number of times married | Discreate |
| Gender (Male or Female) | Discreate |

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) | Ordinal |
| Sales Figures | Ratio |
| Blood Group | Nominal |
| Time Of Day | Ordinal |
| Time on a Clock with Hands | Ordinal |
| Number of Children | Ordinal |
| Religious Preference | Ordinal |
| 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:  Answer given in python file

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

1. Equal to 1

Ans: Sum never be 1 in the rolling of two dice so the probability is zero

1. Less than or equal to 4

Outcomes are: (1,1),(1,2),(1,3),(2,1),(2,2),(3,1)

Probability = 1/36 + 1/36 + 1/36 + 1/36 + 1/36 + 1/36

= 6/36

= 1/6

1. Sum is divisible by 2 and 3

Ans: Outcomes are: (1,5),(2,4),(3,3),(4,2),(5,1),(6,6)

Probability = 1/36 + 1/36 + 1/36 + 1/36 + 1/36 + 1/36

= 6/36

= 1/6

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:** Probability of drawing 2 balls = 7C2 = = 21

Selected ball is either red or green = 5C2 = = 10

probability that none of the balls drawn is blue = 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:** Expected number of candies = Candies count × Probability

= (1×0.015) + (4×0.2) + (3×0.65) + (5×0.005) + (6×0.01) + (2×0.12)

= 3.09

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:** Solved in python file

**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:** Probability of selecting one patient = 1/9

Expected Value of weight = Weight probability × Probability

= (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

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

**Cars speed and distance**

**Use Q9\_a.csv**

**Ans:** Solved in Python file

**SP and Weight(WT)**

**Use Q9\_b.csv**

**Ans:** Solved in Python file

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



**Ans:** Data is right skewed and mean is greater than the median.



**Ans:** For the above plot all the outliers are lying above the maximum value or forth quartile.

**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?

**Ans:** Solved in Python file

**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.

**Ans:** Solved in python file

1. What can we say about the student marks?

**Ans:** Solved in python file

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

**Ans:** No skewness in the distribution at all, meaning the distribution is perfectly symmetrical.

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

**Ans:** The distribution is positively skewed.

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

**Ans:** The distribution is negatively skewed.

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

**Ans:** Positive kurtosis indicates a distribution where more of the numbers are located in the tails of the distribution instead of around the mean. It is known as Leptokurtic.

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.



What can we say about the distribution of the data?

**Ans:** Median is shifted toward the maximum side so the data is not normally distributed.

What is nature of skewness of the data?

**Ans:** Most of the is lying on the left side of the plot so the data is left skewed.

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

**Ans:** IQR = Q3 – Q1 = 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 the boxplots are normally distributed.

Median of both the plots is approximately equal to 263.

First quartile of the plot 1 and plot 2 is approximately 255 and 220 respectively.

Third quartile of the plot 1 and plot 2 is approximately 280 and 310 respectively.

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)
  3. P (20<MPG<50)

**Ans:** Solved in python file

Q 21) Check whether the data follows normal distribution

1. Check whether the MPG of Cars follows Normal Distribution

Dataset: Cars.csv

**Ans:** Solved in python file

1. Check Whether the Adipose Tissue (AT) and Waist Circumference(Waist) from wc-at data set follows Normal Distribution

Dataset: wc-at.csv

**Ans:** Solved in python file

Q 22) Calculate the Z scores of 90% confidence interval,94% confidence interval, 60% confidence interval

**Ans:** Solved in python file

Q 23) Calculate the t scores of 95% confidence interval, 96% confidence interval, 99% confidence interval for sample size of 25

**Ans:** Solved in python file

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

**Ans:** Solved in python file