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

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

**Ans: probability that two heads and one tail are = 0.375**

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

**Ans:**

1. **Equal to 1 = 0**
2. **Less than or equal to 4 = 0.166**
3. **Sum is divisible by 2 and 3 = 0.166**

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(E) = 0.476**

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 for a randomly selected child  = 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:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Points** | **Score** | **weigh** |
| **Mean** | **3.59653** | **3.21725** | **17.84875** |  |
| **Median** | **3.695** | **3.325** | **17.71** |
| **Mode** | **3.92** | **3.44** | **17.02** |
| **Variance** | **0.285881** | **0.957379** | **3.193166** |
| **Standard Deviation** | **0.53468** | **0.978457** | **1.798122** |
| **Range** | **2.17** | **3.911** | **8.4** |

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 the Weight of that patient = 145.33**

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

Cars speed and distance

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **HP** | **MPG** | **VOL** | **SP** | **WT** |
| **SKEW** | **1.716216** | **-0.17795** | **-0.5902** | **1.61145** | **-0.61475** |
| **KURT** | **2.960025** | **-0.61168** | **0.920229** | **2.977329** | **0.950291** |

**Ans:**

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



**Ans : Most of the data points are concerated in the range 50-100 with frequency 200.Least range of weight is 400 somewere with frequency 0-10. So the value above distribution is 75. It is heavily right skewed.**



**Ans: Median is less than mean hence right skewed and we have outlier on the upper side of boxplot and there is less data points.**

**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:**  **confidence interval for 94 % is (201.03 , 198.97)**

**confidence interval for 98 % is (201.37 , 198.63)**

**confidence interval for 96 % is (201.17 , 198.83)**

**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=41**

**Median=40.5**

**Variance=25.52**

**Standard deviation=4.91**

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

**ANS:** **The nature of skewness is zero when mean, median of data are equal .**

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

**ANS: The nature of skewness is positive or right skewed when mean > median**

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

**ANS: The nature of skewness is negative or left skewed when mean > median**

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

**ANS: Positive kurtosis indicates that data has taller tails and sharper peak.**

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

**ANS: Negative values of kurtosis indicate that a distribution is flat and has thin tails.**

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



What can we say about the distribution of the data?

**Ans: Lets assume above box plot is about age of the student in a school.50% of the people are above 10 yrs old and remaining are less and student who’s age is above 15 are approx 40%.**

What is nature of skewness of the data?

**Ans: Left skewed, median is greater than mean.**

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

**Ans:-8**

Q19) Comment on the below Boxplot visualizations?



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

**Ans: 1.By observing both boxplot whisker level is high in boxplot 2.**

**2.mean and median are equal hence distribution is equal.**

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:**

* 1. **P(MPG>38) = 0.65=65%**
  2. **P(MPG<40) = 0.72 = 72%**
  3. **P (20<MPG<50) = 0.89 =89%**

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

**Ans:**

**The Z scores of 90% confidence interval is – 1.28**

**The Z scores of 90% confidence interval is – 1.55**

**The Z scores of 60% confidence interval is – 0.25**

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

**Ans:**

**The t scores of 95% confidence interval – 1.71**

**The t scores of 96% confidence interval – 1.83**

**The t scores of 99% confidence interval –2.49**

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

**Ans: t=0.4714**

**Stats.t.cdf(-0.4714,17) =32.18%**

**The probability of the bulbs lasting less than 260 days on average of 0.3218 .**