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

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

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

Answer :1/3

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

Answer :

a) not possible

b) (1,1)(1,2)(1,3)(2,1)(2,2)(3,1)

6/36 =1/6

c)(1,5)(2,4)(3,3)(4,2)(5,1)(6,6)

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?

Answer :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

Expected number of candies=0.515

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.

POINTS:-

Mean:3.596563 SCORE :

Median:3.695 Mean: 3.21725

Mode:3.92 Median:3.325

Variance:0.2858814 Mode: 3.44

Standard deviation:0.5346787 Variance:0.957379

Range: 2.76 4.93 Standard deviation:0.9784574

Range: 1.513 5.424

WEIGH:-

Mean:17.84875

Median:17.71

Mode:17.02

Variance:3.193166

Standard deviation:1.786943

Range: 14.5 22.9



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 of the weight of that patient=145

(9 patients are there..so we add 9 patients weight and divided by 9)

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

**Cars speed and distance**



Speed distance

Skewness :-0.8448909 1.217917

Kurtosis :2.991396 4.816933

**SP and Weight(WT)**

SP weight

Skewness :-0.4076971 5.398582

Kurtosis : 2.086739 30.82725



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



Answer : Positive histogram



Answer :left skewed distribution with Outliers

**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 ?

For 94% → 2003.98748 1996.01252

For 98% → 2004.933446 1995.066554

For 96% → 2004.354413 1995.645587

**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

1. What can we say about the student marks?

The most students marks are below the mean value…so improve the student marks

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

When the mean,median of data are equal,then its tends to be a symmetric distribution with two modes.

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

The nature of skewness when mean>median is Right-skewed distribution.

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

The nature of skewness when median>mean is left-skewed distribution.

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

A positive kurtosis implies a more peaked distribution than the normal distribution.

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

A negative kurtosis implies a ﬂatter distribution than the normal distribution.

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



What can we say about the distribution of the data?

The distribution of data’s,a longer left side of the box implies a left-skewed data set.

What is nature of skewness of the data?

Left-skewed data

What will be the IQR of the data (approximately)?   
IQR is equal to 8(approx)

Q19) Comment on the below Boxplot visualizations?



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

Both boxplot 1,2 have same median point.the boxplot 2 is the symmetical distribution because both sides are symmetrically distributed.but boxplot1 is left-skewd distribution.

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)

P=33/81

* 1. P(MPG<40)

P=61/81

c. P (20<MPG<50)

p=68/81

Q 21) Check whether the data follows normal distribution

1. Check whether the MPG of Cars follows Normal Distribution

Dataset: Cars.csv

The MPG of cars not following a normal distribution.

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

Dataset: wc-at.csv

Both Adipose Tissue(AT) and Waist Circumference(Waist) not folllowing a normal distribution.

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

90% confidence interval → 1.644

94% confidence interval → 1.880

60% confidence interval → 0.8416

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

95% confidence interval for sample size 25 → 2.059

96% confidence interval for sample size 25 → 2.166

99% confidence interval for sample size 25-→ 2.787

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

p(x≤260)=p(t≤260-µ/sd)=(260-270/90)=-0.1111

degrees of freedom=18-1=17

T-score for degrees of freedom=-0.1111

this shows cumulative probability=0.8888

so 88% of average bulb life for 18 randomly selected bulbs would be less than or equal to 260 days