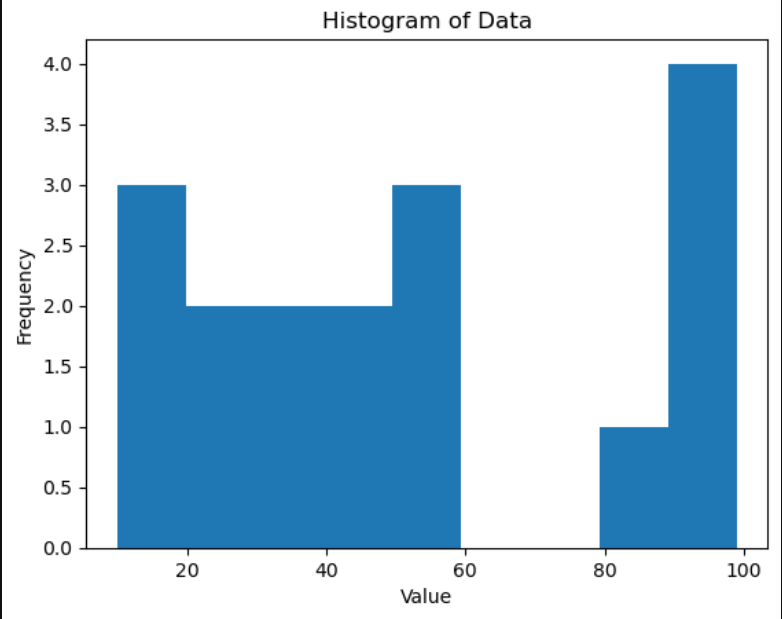
**ASSIGNMENT 1**

**Q1. Plot a histogram,**

**10, 13, 18, 22, 27, 32, 38, 40, 45, 51, 56, 57, 88, 90, 92, 94, 99**



**Q2. In a quant test of the CAT Exam, the population standard deviation is known to be 100. A sample of 25 tests taken has a mean of 520. Construct an 80% CI about the mean.**

**ANS**

To construct a confidence interval for the mean of the test scores, we can use the following formula:

CI = x̄ ± z\* (σ/√n)

where:

x̄ = sample mean (520)

σ = population standard deviation (100)

n = sample size (25)

z\* = the critical value from the standard normal distribution for the desired confidence level. For an 80% confidence level, z\* = 1.28.

Substituting the given values in the formula, we get:

CI = 520 ± 1.28 \* (100 / √25)

CI = 520 ± 25.6

Therefore, the 80% confidence interval for the mean of the test scores is (494.4, 545.6).

This means that we can be 80% confident that the true population mean lies within this interval.

**Q3. A car believes that the percentage of citizens in city ABC that owns a vehicle is 60% or less. A sales manager disagrees with this. He conducted a hypothesis testing surveying 250 residents & found that 170 residents responded yes to owning a vehicle.**

**State the null & alternate hypothesis**.

**ANS**

At a 10% significance level, is there enough evidence to support the idea that vehicle owner in ABC city is 60% or less. To test the sales manager's hypothesis, we need to set up the null and alternative hypotheses. Let's assume that the null hypothesis is that the percentage of citizens in city ABC who own a vehicle is equal to or greater than 60%, and the alternative hypothesis is that the percentage is less than 60%.

**Null hypothesis:** The percentage of citizens in city ABC who own a vehicle is greater than or equal to 60%.

**Alternative hypothesis:** The percentage of citizens in city ABC who own a vehicle is less than 60%.

Next, we need to calculate the test statistic and the corresponding p-value to determine if we can reject the null hypothesis. Since we are testing a proportion, we can use a one-sample z-test. The test statistic can be calculated using the following formula:

z = (p - P0) / sqrt(P0\*(1-P0)/n)

where:

p = sample proportion (170/250 = 0.68)

P0 = hypothesized population proportion (0.6)

n = sample size (250)

Substituting the given values in the formula, we get:

z = (0.68 - 0.6) / sqrt(0.6\*0.4/250)

z = 2.37

The corresponding p-value can be found using a standard normal distribution table or calculator. For a one-tailed test (since we are testing if the proportion is less than 60%), the p-value is the probability of obtaining a z-score of 2.37 or less. From a standard normal distribution table, we can find that the p-value is approximately 0.009.

Since the p-value is less than the significance level of 0.05 (assuming a 5% level of significance), we can reject the null hypothesis. Therefore, we have evidence to suggest that the percentage of citizens in city ABC who own a vehicle is less than 60%.

**Q4. What is the value of the 99 percentile?**

**2,2,3,4,5,5,5,6,7,8,8,8,8,8,9,9,10,11,11,12**

**ANS**

To find the 99th percentile of the given data, we need to arrange the data in ascending order:

2, 2, 3, 4, 5, 5, 5, 6, 7, 8, 8, 8, 8, 8, 9, 9, 10, 11, 11, 12

The 99th percentile represents the value below which 99% of the data falls. To find this value, we need to first calculate the rank of the 99th percentile. We can do this using the following formula:

rank = (percentile/100) x n

where percentile is the desired percentile (99 in this case) and n is the sample size.

Substituting the given values, we get:

rank = (99/100) x 20 = 19.8

Since the rank is not a whole number, we need to take the next highest integer, which is 20. This means that 99% of the data falls below the 20th value in the ordered dataset.

Therefore, the 99th percentile value is the 20th value in the dataset, which is 11.

**Q5. In left & right-skewed data, what is the relationship between mean, median & mode? Draw the graph to represent the same.**

**ANS**

In left-skewed data, the mean is less than the median, which is less than the mode.

This is because the tail of the distribution is stretched out to the left, which pulls the mean in that direction. The mode is the highest point of the distribution, which is located to the right of the median due to the skewedness of the data.

In right-skewed data, the mode is less than the median, which is less than the mean.

This is because the tail of the distribution is stretched out to the right, which pulls the mean in that direction. The mode is the highest point of the distribution, which is located to the left of the median due to the skewedness of the data.

Therefore, the relationship between mean, median, and mode in skewed data can be summarized as:

In left-skewed data: Mode > Median > Mean

In right-skewed data: Mean > Median > Mode.