STATISTICS ASSIGNMENT
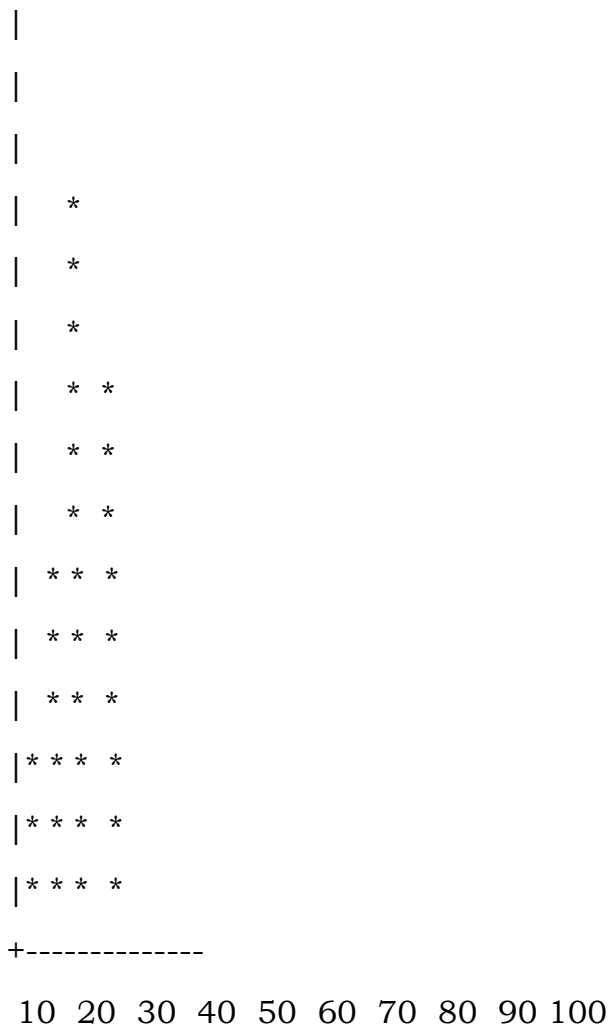
Que 1) Plot a histogram,

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

```
Frequency
  |
  |
  |
  |    *
  |    *
  |    *
  |    *  *
  |    *  *
  |     *  *
  |   *  *  *
  |   *  *  *
  |   *  *  *
  |*  *  *  *
  |*  *  *  *
  |*  *  *  *
  +--------------
   10 20 30 40 50 60 70 80 90 100
```

Que 2) 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.

To construct an 80% confidence interval (CI) about the mean of a sample, you can use the following formula:

$$ \text{CI} = \text{Sample Mean} \pm \left(\text{Critical Value} \times \frac{\text{Population Standard Deviation}}{\sqrt{\text{Sample Size}}} \right) $$

Here:

- Sample Mean = 520 (given)

- Population Standard Deviation = 100 (given)

- Sample Size (n) = 25 (given)

The critical value for an 80% confidence interval depends on the distribution. For a standard normal distribution (Z-distribution), you can use the Z-score associated with an 80% confidence level. The Z-score for an 80% confidence level can be found using a Z-table or calculator.

For an 80% confidence level, the Z-score is approximately 1.28.

Substitute the values into the formula:

$$ \text{CI} = 520 \pm \left(1.28 \times \frac{100}{\sqrt{25}} \right) $$

Simplify:

$$ \text{CI} = 520 \pm 1.28 \times 20 $$

Calculate the bounds of the confidence interval:

Lower bound: $520 - 1.28 \times 20 = 494.4$

Upper bound: $520 + 1.28 \times 20 = 545.6$

So, the 80% confidence interval about the mean is approximately 494.4 to 545.6. This means you can be 80% confident that the true population mean falls within this interval based on the given sample.

Que 3) 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.

a) State the null & alternate hypothesis.
b) At a 10% significance level, is there enough evidence to support the idea that vehicle owner in ABC city is 60% or less.

In hypothesis testing, you start by defining a null hypothesis (H_0) and an alternative hypothesis (H_1 or H_a) based on the question you're trying to answer. Then, you analyze the sample data to determine whether there's enough evidence to reject the null hypothesis in favor of the alternative hypothesis.

Given the situation, let's define the null and alternative hypotheses:

Null Hypothesis (H_0): The percentage of citizens in city ABC who own a vehicle is 60% or less.

Alternative Hypothesis (H_1): The percentage of citizens in city ABC who own a vehicle is greater than 60%.

In mathematical notation:

- $H_0: p \leq 0.60$

- $H_1: p > 0.60$

Where p is the population proportion of citizens in city ABC who own a vehicle.

Next, you need to perform a hypothesis test using the sample data. The sales manager surveyed 250 residents and found that 170 residents own a vehicle. To test the hypothesis, you can use the z-test for proportions since you have a large sample size (250) and can assume the data is approximately normally distributed.

Calculate the sample proportion ($\hat{p}$):

$$\hat{p} = \frac{\text{Number of "Yes" responses}}{\text{Sample size}} = \frac{170}{250} = 0.68$$

Now, calculate the test statistic (z-score) using the sample proportion, the hypothesized population proportion ($p_0 = 0.60$), and the standard error of the sample proportion ($SE(\hat{p})$):

$$SE(\hat{p}) = \sqrt{\frac{p_0 \cdot (1 - p_0)}{n}} = \sqrt{\frac{0.60 \cdot 0.40}{250}} \approx 0.0490$$

$$z = \frac{\hat{p} - p_0}{SE(\hat{p})} = \frac{0.68 - 0.60}{0.0490} \approx 1.6327$$

Now, at a 10% significance level, find the critical z-value. For a one-tailed test with a significance level of 0.10, the critical z-value is approximately 1.282.

Since the calculated z-score (1.6327) is greater than the critical z-value (1.282), we can reject the null hypothesis.

Conclusion: There is enough evidence to support the idea that the percentage of citizens who own a vehicle in city ABC is greater than 60% at the 10% significance level.

Que 4) 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

To find the value of the 99th percentile from the given data set, you need to determine the value below which 99% of the data falls. Here's the step-by-step process:

 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

Que 5) In left & right-skewed data, what is the relationship between mean, median & mode?

Draw the graph to represent the same.

In left-skewed (negatively skewed) data distributions, the mean tends to be less than the median, and the mode is often greater than the median.

In right-skewed (positively skewed) data distributions, the mean tends to be greater than the median, and the mode is often less than the median.
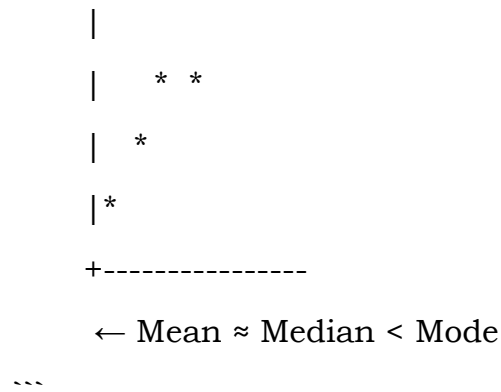
Let's visualize these relationships with graphs for both left-skewed and right-skewed distributions.

Left-Skewed Distribution:

In a left-skewed distribution, the tail of the distribution extends towards the left side.

```
    |
    |
    |
```

```
    |
    |      *  *
    |    *
    |*
    +----------------
       ← Mean ≈ Median < Mode
```
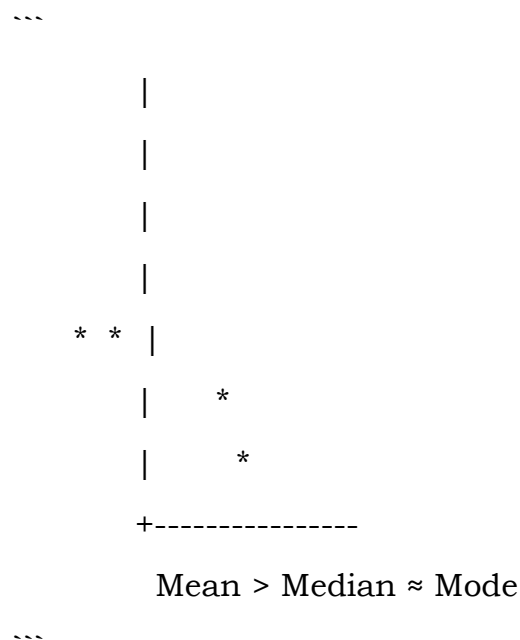
In this graph:

- The mean is approximately at the left of the median.

- The mode is greater than the median.

Right-Skewed Distribution:

In a right-skewed distribution, the tail of the distribution extends towards the right side.

```
            |
            |
            |
            |
     *  *   |
            |      *
            |        *
        +----------------
            Mean > Median ≈ Mode
```

In this graph:

- The mean is greater than the median.

- The mode is approximately at the right of the median.

These relationships hold true on average for skewed distributions, but there can be exceptions or variations depending on the specific shape of the distribution.