**Que 1) Plot a histogram,**

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

import matplotlib.pyplot as plt

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

num\_bins = 5 # You can adjust the number of bins as needed

# Create histogram

plt.hist(data, bins=num\_bins, edgecolor='k')

# Add labels and title

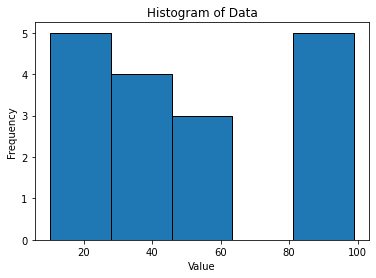
plt.xlabel('Value')

plt.ylabel('Frequency')

plt.title('Histogram of Data')

# Show the plot

plt.show()



**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 when the population standard deviation is known, you can use the formula for a CI:

\[ \text{Confidence Interval (CI)} = \bar{X} \pm Z \left( \frac{\sigma}{\sqrt{n}} \right) \]

Where:

- \(\bar{X}\) is the sample mean,

- \(\sigma\) is the population standard deviation,

- \(n\) is the sample size,

- \(Z\) is the Z-score associated with the desired confidence level.

In this case, you want to construct an 80% CI, so you need to find the Z-score corresponding to an 80% confidence level. You can use a Z-table or a calculator to find this value. The Z-score for an 80% confidence level is approximately 1.282.

Given:

- \(\sigma\) (population standard deviation) = 100

- \(\bar{X}\) (sample mean) = 520

- \(n\) (sample size) = 25

- \(Z\) (Z-score for 80% CI) ≈ 1.282

Now, plug these values into the formula to calculate the 80% CI:

\[ \text{CI} = 520 \pm 1.282 \left( \frac{100}{\sqrt{25}} \right) \]

\[ \text{CI} = 520 \pm 1.282 \left( \frac{100}{5} \right) \]

\[ \text{CI} = 520 \pm 1.282 \times 20 \]

Lower Bound: \(520 - 1.282 \times 20 \)

Upper Bound: \(520 + 1.282 \times 20 \)

Lower Bound ≈ 520 - 25.64 ≈ 494.36

Upper Bound ≈ 520 + 25.64 ≈ 545.64

So, the 80% confidence interval about the mean is approximately (494.36, 545.64). This means you can be 80% confident that the true population mean falls within this interval based on your sample data.

**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.**

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

a) Hypotheses:

Null Hypothesis (\(H\_0\)): The percentage of citizens in city ABC who own a vehicle is 60% or less.

\(H\_0: p \leq 0.60\)

Alternate Hypothesis (\(H\_1\)): The percentage of citizens in city ABC who own a vehicle is greater than 60%.

\(H\_1: p > 0.60\)

Where:

- \(p\) represents the proportion of citizens in city ABC who own a vehicle.

b) To determine whether there is enough evidence to support the idea that the vehicle ownership rate in ABC city is greater than 60%, you can perform a hypothesis test using a significance level of 10% (\(\alpha = 0.10\)).

Here are the steps for the hypothesis test:

1. Calculate the sample proportion (\(\hat{p}\)) of residents who responded "yes" to owning a vehicle:

\(\hat{p} = \frac{\text{Number of "yes" responses}}{\text{Total number of responses}} = \frac{170}{250} = 0.68\)

2. Calculate the standard error (\(SE\)) of the sample proportion:

\(SE = \sqrt{\frac{p(1-p)}{n}}\)

\(SE = \sqrt{\frac{0.60(1-0.60)}{250}} \approx 0.0490\)

3. Calculate the test statistic (\(Z\)) using the following formula:

\(Z = \frac{\hat{p} - p}{SE}\)

\(Z = \frac{0.68 - 0.60}{0.0490} \approx 1.6327\)

4. Find the critical value for a one-tailed test at the 10% significance level. You can use a Z-table or a calculator. For a 10% significance level, the critical value is approximately 1.282 (corresponding to the upper 10% of the standard normal distribution).

5. Compare the test statistic (\(Z\)) to the critical value:

- If \(Z\) is greater than the critical value, reject the null hypothesis.

- If \(Z\) is less than or equal to the critical value, fail to reject the null hypothesis.

In this case, \(1.6327 > 1.282\), so you reject the null hypothesis.

Conclusion:

At a 10% significance level, there is enough evidence to support the idea that the vehicle ownership rate in ABC city is greater than 60%.

**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 99th percentile of a dataset, you need to determine the value below which 99% of the data falls. In other words, you want to find the value that separates the top 1% of the data.

First, you should sort 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

Now, you have 20 data points, and you want to find the value below which 99% of the data falls. To do this, you can use the formula:

Percentile Value=Value at rank=10099​×(Total number of data points+1)

In this case:

Percentile Value=99100×(20+1)Percentile Value=10099​×(20+1)

Percentile Value=99100×21Percentile Value=10099​×21

Percentile Value=0.99×21Percentile Value=0.99×21

Percentile Value=20.79Percentile Value=20.79

Since percentiles usually represent whole values, you can round up to the nearest whole number. So, the 99th percentile of the given dataset is approximately 21. This means that 99% of the data values in the dataset are less than or equal to 21.

**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:

- The mean is typically less than the median, and both are less than the mode.

In right-skewed (positively skewed) data:

- The mean is typically greater than the median, and both are greater than the mode.

#left skewed data

\_\_\_\_\_\_

\_/

/

\_\_\_\_\_/\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

^ ^ ^

Mode Median Mean

```

- Mode: The mode is the highest point in the distribution, located to the right of the median.

- Median: The median is closer to the right of the distribution than the mode.

- Mean: The mean is typically the furthest to the left, influenced by the presence of extreme values.

### Right-Skewed Data:

```

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\

\\_

\

\_\_\_\_\_\_

Mode Median Mean

```

- Mode: The mode is the highest point in the distribution, located to the left of the median.

- Median: The median is closer to the left of the distribution than the mode.

- Mean: The mean is typically the furthest to the right, influenced by the presence of extreme values.

These relationships hold true for most cases, but there can be exceptions depending on the specific characteristics of the data distribution.