Scenario Based - Set Qn 1

Scenario 1: Delivery Time Analysis for an E-commerce Company

An e-commerce company tracks delivery times (in minutes) for 15 orders:
[25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95]

The company wants to analyze the delivery performance using percentiles and detect if there are any unusual delivery times.

Question 1:
Calculate Q1 and Q3.

First_half = [25, 30, 35, 40, 45, 50, 55]

Q1 = median of this = 40

Last_half = [65, 70, 75, 80, 85, 90, 95]

Q3 = median of this = 80

Question 2:
Find the Interquartile Range (IQR).

IQR = Q3 - Q1 = 80-40 = 40

Question 3:
Detect Outliers using the IQR method.

Upper = Q3+1.5*IQR = 80 +60 = 140

Lower= Q3+1.5*IQR = 40−60=−20

Any value **< -20** or **> 140** is considered an outlier.

Scenario 2: Student Score Analysis

A teacher is analyzing the mathematics scores of students in her class. The scores are:
[45, 50, 55, 60, 60, 62, 63, 65, 90, 95]

Question 1:
Calculate the mean, median, and mode of the scores.

Mean : 45+50+55+60+60+62+63+65+90+95/10 = 64.5

Median: Middle value = (60+62)/2=61(60 + 62) / 2 = 61(60+62)/2=61

Mode: 60 (Occurs twice)

Question 2:
Explain why the median might be a better representation than the mean in this case.

The **median** is a better measure here because the high outliers (90 and 95) skew the mean. The median is not affected by extreme values, providing a better reflection of student performance.

Scenario 3:

A grocery store manager tracks how many customers visit the store daily for a month:
[5, 10, 8, 15, 20, 5, 12, 14, 10, 18]

Question 1:

- Create a frequency distribution table for this data.

Number of Customers	Frequency
5	2
10	2
8	1
12	1
14	1
15	1
18	1
20	1

Scenario 4:

A real estate model has three variables:

- House Size
- Number of Rooms
- Number of Bathrooms

Question 1:

- How can you detect multicollinearity?
    - Calculate the Variance Inflation Factor (VIF).
    - VIF > 10 indicates multicollinearity.

High VIF means the variables are correlated, impacting model accuracy.

Scenario 5:

A company made a new medicine to lower blood pressure. They gave it to one group and gave a fake pill (placebo) to another group.

Question 1:

How can the company check if the new medicine works?

- **Hypothesis:**

    - **H0:** The medicine doesn't lower blood pressure.
    - **H1:** The medicine lowers blood pressure.
- **T-Test:**
    - Find the **p-value** (a number that shows how likely the result happened by chance).
    - If **p-value < 0.05**, it means the medicine likely works.
    - If the **p-value** is small, the medicine is effective.

Scenario 6: Identifying Outliers in Sales Data

A company wants to find any unusual spikes in sales.

Question 1:
How can the company detect outliers in their sales data?

Step 1: Calculate the Interquartile Range (IQR).

Step 2: Identify outliers using the formula:

$$\text{Outliers} = (\text{Data} < Q1 - 1.5 \times IQR) \text{ or } (\text{Data} > Q3 + 1.5 \times IQR)$$

Scenario 7: Understanding Customer Satisfaction

A restaurant conducted a survey to rate customer satisfaction on a scale of 1 to 5:
[5, 4, 4, 5, 3, 4, 5, 2, 4, 3]

Question 1: How can the restaurant summarize the overall satisfaction?

- Calculate the Mean and Median for further insights.
- Find the Mode to see the most common rating.
- If most ratings are 4 or 5, satisfaction is generally high.