DS- Scenario Set Question 2

Scenario 1:

Flight Delay Analysis

An airline tracks flight delays (in minutes) for 20 flights. Analyze the flight delays to calculate percentiles, detect outliers, and evaluate the overall distribution.

Answer:

Method:

1. Arrange all flight delay times in order.
2. Find **Q1 (25th percentile)**, **Median (50th percentile)**, and **Q3 (75th percentile)**.
3. Calculate **IQR = Q3 − Q1**.
4. Find **Lower Bound = Q1 − 1.5 × IQR** and **Upper Bound = Q3 + 1.5 × IQR**.
5. Delays outside this range are **outliers**.
6. Use a **boxplot or histogram** to check the delay pattern and distribution.

Scenario 2:

Employee Salary Analysis

A company wants to analyze the salary distribution of its employees to understand the central tendency and determine whether the data is skewed.

Answer:

Method:

1. **Calculate Central Tendency:**
   - Find **Mean**, **Median**, and **Mode** of salaries.
2. **Compare Mean and Median:**
   - If **Mean > Median** → data is **right-skewed** (few very high salaries).
   - If **Mean < Median** → data is **left-skewed** (few very low salaries).
3. **Check Spread:**
   - Calculate **Standard Deviation** to see how much salaries vary.
4. **Visualize:**
   - Use a **histogram** or **boxplot** to see the distribution shape.
5. **Conclusion:**
   - Identify if salaries are mostly balanced or affected by extreme values.

Scenario 3:

Product Sales Analysis

A retail store records product sales over 15 days. Create a frequency distribution table and visualize the sales data using appropriate charts.

Answer:

Method:

1. Arrange all sales data in order.
2. Group sales into ranges (like 0–10, 11–20, etc.).
3. Count how many days fall in each range to make a **frequency table**.
4. Show the data using a **bar chart** or **histogram**.
5. Observe which range has the highest sales.

Scenario 4:

Student Exam Performance Analysis

A school wants to analyze the exam performance of students across three subjects: Mathematics, Science, and English. How can Data Science concepts be applied to understand their performance?

Answer:

Method:

1. **Collect Data:** Gather marks for Mathematics, Science, and English.
2. **Calculate Summary Statistics:** Find **mean**, **median**, and **standard deviation** for each subject.
3. **Compare Subjects:** Identify which subject students perform best or worst in.
4. **Check Correlation:** See if performance in one subject relates to another.
5. **Visualize Data:** Use **bar charts**, **boxplots**, or a **heatmap** to show comparisons clearly.
6. **Draw Insights:** Find patterns, strengths, and areas where students need improvement

Scenario 5:

Clinical Trial for Diabetes Medication

A pharmaceutical company conducted a clinical trial with two groups: one receiving medication and the other a placebo. Perform a hypothesis test to determine the effectiveness of the medication.

Answer:

Method:

1. **Collect Data:** Record blood sugar levels for both medication and placebo groups.
2. **Set Hypotheses:**
    - **Null Hypothesis (H_0):** Medication has no effect (mean change same as placebo).
    - **Alternative Hypothesis (H_1):** Medication lowers blood sugar (mean change is different).
3. **Choose Statistical Test:** Use an **independent t-test** to compare the two groups.
4. **Calculate p-value:**
    - If **p-value < 0.05**, reject H_0 → medication is effective.
    - If **p-value ≥ 0.05**, fail to reject H_0 → no significant effect.
5. **Visualize Results (Optional):** Use **boxplots** or **bar charts** to show group differences.