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.

Ans:

Step 1: Collect and preprocess the dataset.

Step 2: Calculated the percentile(Q1,Q2,Q3,Q4)

Step 3: Calculate IQR = Q3-Q1

Step 4: Calculate lower bound (Q1-1.5*IQR) and upper bound(Q3+1.5*IQR).

Step 5: To find outlier, Use condition if the value of the data is less than lower bound or greater than upper bound, the outlier is present in the dataset.

Step 6: Use histogram to show the overall distribution and box plot to show outliers.

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.

Ans:

Step 1: Collect and preprocess the dataset.

Step 2: Calculate Mean, Median and Mode for the dataset to understand the central tendency.

Step 3: To understand skewness, apply condition such as Mean>Median, its positive skewness, Mean<Median - Negative Skewness, and Mean=Median - Symmetric distribution.

Step 4: Use Median for better outlier representation as its include outlier.

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.

Ans:

Step 1: Collect and preprocess the dataset.

Step 2: Calculate frequency distribution

Step 3: Use Histogram and Bar plot for data visualization of 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?

Ans:

Step 1: Collect and preprocess the dataset.

Step 2: For descriptive statistics, Calculate the mean, median and mode and standard deviation to understand the central tendency and dispersion.

Step 3: For data visualization, use scatter plot for relationship between the subjects, use box plot for outliers, and use histogram for normal distribution of data.

Step 4: Use Correlation method, to check the relationship between the subjects. And use heatmap for correlation visualization.

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.

Ans:

Step 1: Collect and preprocess the dataset.

Step 2: Perform T-Test to check the medication works effectively or not.

- **H_0 (Null Hypothesis):** No difference between medication and placebo.
- **H_1 (Alternative Hypothesis):** Medication lowers blood sugar more than the placebo.

Step 3: Use condition,

if p-value <0.05,rejects H0, there is a effectiveness of the medicine.

If p-value >0.05, fails to reject H0. there is no significant effectiveness of the medicine