Data Science scenario Ans S2:2

1.Flight delay analysis:

Percentile calculations:calculate 10th,25th,50th,(Median),75th and 90th-from this how flight delays are distributed at different levels

IQR calculation:

IQR=Q3-Q1 Detect outliers

Distribution plot:

Box plot-to visualize the spread,median,quartiles and potential outliers

Histogram-asses the distribution shape,skewness,frequency of delays

2.Employee salary analysis:

Central Tendency calculation

Calculate Mean,Median,Mode to understand the central value of the salaries

Skeweness check:

If Mean>Median,data is Right Skewed(higher salaries skew the mean)

If Mean<Median,data-left skewed(Lower salaries pull the mean down)

If Mean=Median,data is symmetrical

Best is use Median when there are outliers,as it better represents the central tendency in the presence of extreme values

3.Product Sales Analysis:

Frequency Distribution

Divide the sales date into intervals(eg 5 or 10 units) Count the sales within each interval to understand how sales are distributed.

Visualization:

Histogram:Displays the frequency distribution across sales intervals.

Bar Plot:shows trends in sales,helping to visualize changes over time

4.Student performance Analysis

1.Data preprocessing:

Handle missing values by imputing with the mean or median

Convert categorical values using label encoding or one=hot encoding

Descriptive Statistics:

Calculate mean,median,mode,and,standard deviation for each subject to understand the central tendency and dispersion.

3.Visualisation:

Use box plots to detect outliers

Create histograms to visualize the distribution of scores.

Plot scatter plots to check relationships between subjects.

Compute correlation coefficients to see how subject scores relate to each other.

Use aheatmap to visualize correlations

5.Clinical Trial for Diabetes Medication:

Hypothesis Test:

Ho(Null Hypothesis):No difference between medication and placebo.

H1(Alternative Hypothesis):Medication lowers blood suger more than the placebo

T-test

If $p<0.05$,reject Ho indicating medication is effective)

If $p>0.05$ fail to reject Ho(no significant difference)