DATA SCIENCE

Delivery Time Analysis for an Ecommerce Company:

Question 1: Calculate Q1 and Q3

$Q1=0.25\times(n+1)$ Q1 = 0.25 \times (n+1) $Q1=0.25\times(n+1)$ $Q3=0.75\times(n+1)$ Q3 = 0.75 \times (n + 1) $Q3=0.75\times(n+1)$ ●Q1 = 4th value → 40 ●Q3 = 11th value → 75

Question 2: Find the IQR

$IQR=Q3-Q1=75-40=35$

Question 3: Detect Outliers Lower Bound=Q1−1.5×IQR {Lower Bound} = Q1 - 1.5 \times IQR Lower Bound=Q1−1.5×IQR Upper Bound=Q3+1.5×IQR {Upper Bound} = Q3 + 1.5 \times IQR Upper Bound=Q3+1.5×IQR

2.QUESTION-MEAN, MEDIAN, MODE:

Mean:

45+50+55+60+60+62+63+65+90+9510=6

5.5{45 + 50 + 55 + 60 + 60 + 62 + 63 + 65 + 90 + 95}{10} = 65.51045+50+55+60+60+62+63+65+90+95=65.5

- Median: Middle value = (60+62)/2=61(60 + 62) / 2 = 61(60+62)/2=61
- Mode: 60 (Occurs twice)

3. DETECT MULTICOLLINEARITY:

Calculate the Variance Inflation Factor (VIF). ○VIF > 10 indicates multicollinearity.

- Answer:

High VIF means the variables are correlated, impacting model accuracy

4. HYPOTHESIS TESTING:

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.

5. DETECTING OUTLIERS:

● Calculate the Interquartile Range (IQR).

●Step 2: Identify outliers using the formula: Outliers=(DataQ3+1.5×IQR)

●Outliers = (Data < Q1 - 1.5 \times IQR) (Data > Q3 + 1.5 \times IQR)

●Outliers=(DataQ3+1.5×IQR)

6. Understanding Customer Satisfaction:
Answer:

●Find the Mode to see the most common rating.

● Calculate the Mean and Median for further insights