

Case Study on Statistics

1. Dataset: Monthly temperatures in a city (in °C): 12, 15, 14, 16, 18, 20, 22, 19, 17, 15, 13, 10

Questions:

- Calculate the mean temperature.
- Calculate the median temperature.
- Find the mode of the temperature data.
- Range and Variance

2. Dataset: Heights of students (in cm): 150, 160, 165, 170, 155, 160, 162, 168

Questions:

- Calculate the range of the heights.
- Calculate the variance of the heights.
- Standard Deviation

3. Dataset: Weekly expenses (in \$): 200, 250, 300, 220, 270, 290, 310

Questions:

- Calculate the standard deviation of the weekly expenses.
- Medium Level
- Covariance

4. Dataset: Advertising Spending (in \$1000s): 10, 20, 15, 25, 30; Sales (in \$1000s): 100, 150, 120, 180, 200

Questions:

- Calculate the covariance between advertising spending and sales.
- Correlation

5. Dataset: Study Hours: 2, 4, 6, 8, 10; Exam Scores: 55, 60, 65, 70, 75

Questions:

- a) Calculate the correlation coefficient between study hours and exam scores.
- b) Interpret the result.
- c) Quartiles and IQR

6. Dataset: Monthly savings (in \$): 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100

Questions:

- a) Calculate the first quartile (Q1) and third quartile (Q3).
- b) Calculate the interquartile range (IQR).
- c) Detecting Outliers

7. Dataset: Ages of employees: 22, 25, 28, 30, 35, 40, 45, 50, 55, 100

Questions:

- a) Identify any outliers using the IQR method.
- b) Hard Level
- c) Weighted Mean

8. Dataset: Grades (A, B, C, D, F) with corresponding weights: A=4, B=3, C=2, D=1, F=0.
Student grades: A, B, A, C, B, A, D, B, C, A

Questions:

- a) Calculate the weighted mean of the student grades.
- b) Regression Analysis

9. Dataset: Advertising Spending (in \$1000s): 10, 20, 15, 25, 30; Sales (in \$1000s): 100, 150, 120, 180, 200

Questions:

- a) Perform a simple linear regression analysis to find the relationship between advertising spending and sales.
- b) Interpret the slope and intercept of the regression line.
- c) ANOVA (Analysis of Variance)

10. Dataset: Test scores of students from three different classes:

Class A: 85, 88, 90, 92, 95

Class B: 78, 82, 85, 88, 90

Class C: 92, 95, 98, 100, 105

Questions:

- a) Perform an ANOVA test to determine if there are any significant differences between the mean scores of the three classes.

11. Dataset: Daily stock prices for 10 days: 100, 102, 101, 103, 105, 107, 109, 110, 108, 107

Questions:

- a) Calculate the 3-day moving average for the stock prices.
- b) Time Series Analysis

12. Dataset: Monthly sales data for a year: 1200, 1300, 1250, 1400, 1350, 1500, 1600, 1550, 1700, 1650, 1800, 1750

Questions:

- a) Plot the sales data on a line chart.
- b) Identify any trends or seasonality in the data.
- c) Chi-Square Test

13. Dataset: Observed frequencies of customer preferences for three products:

Product A: 50, 55, 60

Product B: 30, 25, 35

Product C: 20, 20, 25

Questions:

- a) Perform a chi-square test to determine if there is a significant difference in customer preferences among the three products.
- b) Hypothesis Testing

14. Dataset: Sample weights (in kg) of a new diet program: 70, 68, 72, 71, 69, 73, 72, 74, 70, 71

Questions:

- a) Test the hypothesis that the mean weight after the diet program is equal to 70 kg at a 5% significance level.
- b) Z-Scores

15. Dataset: Annual salaries of employees (in \$1000s): 40, 42, 45, 47, 50, 52, 55, 57, 60, 62

Questions:

- a) Calculate the z-scores for each salary.
- b) Interpret the z-scores.
- c) Normal Distribution

16. Dataset: Scores of students in a standardized test: 85, 87, 90, 92, 95, 97, 100, 102, 105, 107

Questions:

- a) Create a histogram of the test scores.
- b) Fit a normal distribution to the data and calculate the mean and standard deviation.
- c) P-Values

17. Dataset: Sample data of a new drug's effect on blood pressure reduction (in mmHg): 10, 12, 14, 15, 13, 16, 18, 17, 19, 20

Questions:

- a) Perform a t-test to determine if the new drug has a significant effect on blood pressure reduction.
- b) Report the p-value and interpret the results.

c) Confidence Intervals

18. Dataset: Monthly income of individuals (in \$1000s): 3.5, 4.0, 4.2, 4.5, 4.7, 5.0, 5.3, 5.5, 5.7, 6.0

Questions:

- a) Calculate the 95% confidence interval for the mean monthly income.
- b) Box Plot

19. Dataset: Daily temperatures (in °C) over two weeks: 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38

Questions:

- a) Create a box plot of the daily temperatures.
- b) Identify any outliers.
- c) Scatter Plot and Trend Line

20. Dataset: Advertising Spending (in \$1000s): 10, 20, 30, 40, 50; Sales (in \$1000s): 100, 150, 200, 250, 300

Questions:

- a) Create a scatter plot of advertising spending vs. sales.
- b) Add a trend line and display the equation of the line.
- c) Exponential Smoothing

21. Dataset: Monthly sales data for a year: 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260

Questions:

- a) Apply exponential smoothing to the sales data.
- b) Plot the original data and the smoothed data.
- c) Multiple Regression

22. Dataset: Advertising Spending (in \$1000s): 10, 20, 30, 40, 50; Sales (in \$1000s): 100, 150, 200, 250, 300; Online Marketing Spending (in \$1000s): 5, 10, 15, 20, 25

Questions:

- a) Perform a multiple regression analysis with advertising spending and online marketing spending as independent variables and sales as the dependent variable.
 - b) Interpret the coefficients and the R-squared value.
23. An auto company decided to introduce a new six cylinder car whose mean petrol consumption is claimed to be lower than that of the existing auto engine. It was found that the mean petrol consumption for the 50 cars was 10 km per litre with a standard deviation of 3.5 km per litre. Test at 5% level of significance, whether the claim of the new car petrol consumption is 9.5 km per litre on the average is acceptable.
24. A manufacturer of ball pens claims that a certain pen he manufactures has a mean writing life of 400 pages with a standard deviation of 20 pages. A purchasing agent selects a sample of 100 pens and puts them for test. The mean writing life for the sample was 390 pages. Should the purchasing agent reject the manufactures claim at 1% level?
25. The mean weekly sales of soap bars in departmental stores were 146.3 bars per store. After an advertising campaign the mean weekly sales in 400 stores for a typical week increased to 153.7 and showed a standard deviation of 17.2. Was the advertising campaign successful?