

Jamia Hamdard
School of Management and Business Studies
Department of Management
Odd Semester Examination - March 2023
B.Com (Hons)-I Semester
BCH103 : Business Statistics

Roll No...

Duration: 3 Hour

Max.Marks -60
(4X6=24)

Section A - Short Answer

Note: Attempt any Six Questions, All questions carry equal marks

- Q1: What is a 'statistic'?
- Q2: Differentiate between information and data using one simple example.
- Q3: Differentiate between quantitative and qualitative variables using one suitable example.
- Q4: Discuss the usefulness of pie charts. When they may be used to visualize the data?
- Q5: What is a Probability distribution?
- Q6: List the properties of a Normal Distribution.
- Q7: Differentiate between discrete and continuous variables using one suitable example.
- Q8: What is correlation?
- Q9: Discuss the usefulness of mode as a measure of central tendency.

Section B-Long Answer

(12X3=36)

Note: Attempt any Three Questions; All questions carry equal marks

Q1: You have collected data on the amount of time (in minutes) that 20 people spend exercising each week. The data is as follows:

25, 30, 40, 50, 20, 45, 35, 20, 30, 60, 25, 45, 55, 30, 50, 15, 20, 25, 35, 40.

Which visualization tool(s) would be most appropriate to use to show the distribution of exercise times in the dataset? Explain the reasoning for your choice.

Q2: You work for a company that sells a product, and you are interested in understanding the relationship between the price of the product and the number of units sold. You have collected the data on the price and number of units sold for the past 10 months as follows:

Price (Rs)	10	20	30	40	50	60	70	80	90	100
Units	100	80	60	50	40	30	25	20	15	10

- (a) Create a scatter plot to visualize the relationship between the price and the number of units sold. [4]
- (b) Based on the Scatter plot, what can you infer about the relationship between price and the number of units sold? [4]
- (c) Can you suggest any further analysis to better understand this relationship? Explain [4]

Q3: Suppose you are the manager of a company and you want to calculate the average salary of your employees. You have data on the salaries of your employees, but you know that some employees make significantly more than others. You want to calculate an average that reflects the overall salary level of your employees, but you don't want the average to be skewed by the high salaries of a few employees.

Salary (Rs.)

30,000	35,000	45,000	50,000	60,000	70,000	80,000	90,000	100,000	1,000,000
--------	--------	--------	--------	--------	--------	--------	--------	---------	-----------

- (a) Which type of average would be most appropriate to use in this case, and why? [4]
- (b) Show how you would calculate the average using the chosen method. [4]
- (c) Discuss any limitations or assumptions of your approach. [4]

Q4: The manager of a company wants to compare the salaries of two departments - the Sales department and the Marketing department. The salaries of the employees in both departments are shown below:

Sales department: [40, 50, 50, 60, 70, 80, 90, 100]

Marketing department: [30, 40, 50, 60, 70, 80, 90, 100, 110]

The manager wants to know the following based on the data provided above –

- (a) which department has a higher average salary, and
- (b) which department has a more spread-out distribution of salaries.

Q5: Suppose you're a researcher studying the distribution of income in a particular region. You collect income data from a random sample of 1000 individuals in the region, and you compute the skewness and kurtosis of the income distribution. The skewness is 1.5, and the excess kurtosis is 3.0.

- (a) Based on these numbers, can you conclude that the income distribution is highly skewed and leptokurtic? [4]
- (b) What does this tell you about the shape of the distribution? [4]
- (c) What are the implications for your research? [4]

Q6: Suppose you're a data analyst working for an e-commerce company, and you're interested in understanding the factors that influence customer spending on the company's website. You've collected data on a sample of 1000 customers, including their age, gender, income, and the total amount spent on the website over the past year.

You decide to run a multiple linear regression analysis with total spending as the dependent variable and age, gender, and income as independent variables. The regression equation you obtain is:

$$\text{Total Spending} = 250 + 2.5\text{Age} + 50\text{Gender} + 0.1*\text{Income}$$

Using this regression equation, answer the following questions:

- (a) What is the predicted total spending for a 30-year-old female customer with an income of Rs. 50,000?
 - (b) Interpret the coefficient for income. What does this tell you about the relationship between income and total spending?
 - (c) What does the R-squared value of 0.25 for the regression tell you about the overall fit of the model?
- *****