Score: **92%**

No. of questions: **6**

Correct answer: **5**

Incorrect answer: **1**

Show incorrect attempt only

Question 1

1 Mark

Which of the statements given are correct?

Statement 1: Univariate analysis is for one variable
Statement 1: Bivariate analysis is for one variable
Statement 2: Univariate analysis is for two variable
Statement 4: Bivariate analysis is for two variable

A Statement I and II

B Statement I and IV

C Statement II and IV

D Statement I and III

Correct Answer: B. Statement I and IV

Univariate analysis refers to analyzing a single variable while bivariate analysis is for two variables.

Question 2

2 Marks

Which of the following is a plot used for bivariate analysis?

A	Scatter plot	✓
B	Histogram plot	
C	Both scatter and histogram are for bivariate	
D	Both scatter and histogram are for univariate	

Correct Answer: A. Scatter plot

Bivariate analysis is used to determine the relationship of one variable with another. Histogram gives a frequency distribution for a single variable while scatter plot is a plot between two variables.

Question 3

3 marks

Why do we need bivariate analysis?

A	It is used to find out the relationship between two variables.	✓
B	It is used to determine missing values in each individual variable.	
C	To determine the preprocessing steps required.	

Correct Answer: A. It is used to find out the relationship between two variables.

Bivariate analysis is used to determine the relationship of one variable with another.

Question 4

1 Mark

Which of the following would fall under data exploration?

A Imputing missing values

B Deleting duplicate rows and columns

C Creating histograms

D Both A and B

Correct Answer: C. Creating histograms

Imputing missing values and removing duplicates are preprocessing steps while creating histograms and other plots are a part of data exploration.

Question 5

2 Marks

Given below are 4 different orders of the machine learning life cycle. Select the right one

A Problem definition->data exploration->hypothesis generation->model building

B Problem definition ->hypothesis generation->data exploration->model building

C Problem definition ->data exploration->model building->hypothesis generation

D hypothesis generation->Problem definition ->data exploration->model building

Correct Answer: B. Problem definition ->hypothesis generation->data exploration->model building

The first step is problem definition. Then we perform hypothesis testing which is done before looking at the dataset and the variables. Then we explore the data and build a model.

Question 6

3 marks

Which of the following techniques can be used to deal with missing values?

A	Drop the rows having missing values
B	Replace missing with mean or mode
C	Use other columns to impute missing values
D	All of the above ✓

Correct Answer: D. All of the above

All the methods mentioned above can be used when the dataset has missing values.