2212N1054

Candidate's Seat No:____

IMBA in BI/IB/ APR Sem.-9 Examination
MDA
December-2025

Time : 2.30 Hours] **[Max.Marks : 70**

Instructions :(1) This paper contains **FIVE** questions.
(2) All questions are compulsory.
(3) Question No.**2, 3, 4** have internal options.
(4) Figures in the right side in parenthesis indicate marks.

Q.1 A smartphone company is launching a new flagship model and has collected a large dataset with many variables. The research team wants to apply multivariate analysis to achieve the following goals: **(14)**

1. Group similar customers or products into meaningful clusters based on several variables, without any prior group labels.
2. Test whether different demographic groups differ in their ratings of smartphone features such as battery life, camera, speed, and design.
3. Identify how consumers make trade-offs among features (screen size, storage, camera, price) to determine the most preferred feature combinations.
4. Examine whether advertising affects sales indirectly—by first increasing Brand Awareness, then improving Brand Attitude, which eventually boosts sales.

Task: For each of the four objectives:
- Identify the most appropriate multivariate technique.
- Classify it as a **dependence** or **interdependence** technique.
- Briefly justify your choice based on the role of variables, research purpose, and data characteristics.

Q.2 A market research team is conducting a customer satisfaction survey for a large retail chain. They use a 7-point Likert scale to measure satisfaction. After pilot testing, the team observes that the responses show unexpected patterns: **(14)**
- Some store locations consistently report higher satisfaction scores than others, even though their service operations are similar.
- Among customers surveyed at the same store, individual responses vary widely, even when customers report similar experiences in follow-up interviews.

Question:
Using the scenario above, explain the concepts of **Systematic Error** and **Random Error** in measurement. In your answer:
- Define each error clearly.

P.T.O

- Identify possible sources of each error in the survey context.
- Explain how each type of error can affect the interpretation of data.
- Suggest practical strategies the research team can use to reduce each type of error.

OR

Q.2 A large online food delivery platform wants to evaluate its customer experience performance for the upcoming financial year. **(14)**

The company specifically wants to estimate the percentage of customers who are highly satisfied with the delivery service. Previous pilot surveys suggest that approximately 70% of customers report being "highly satisfied." Management would like the estimate to have a precision of ±5 percentage points at a 95% confidence level.

Required:

- Explain the sample size determination approach based on precision rate and confidence level when estimating a population proportion.
- Using the information provided, compute the required sample size for estimating the proportion of highly satisfied customers.
- Briefly comment on why precision rate and confidence level are important when planning a survey-based study in business research.

Q.3 A company is preparing its dataset for multivariate analysis. During the initial review, the analytics team observes the following: **(14)**

- Some variables contain missing values.
- Several records appear duplicated.
- A few data points seem unusually high or low compared to the rest.
- Different units and formats have been used for the same variable across departments.

Before proceeding, the team wants to identify key data accuracy issues and select suitable graphical techniques to examine the dataset.

Required:

- Identify and explain any four issues affecting data accuracy illustrated in the scenario. (Explain briefly why each issue reduces the reliability of the dataset.) (8 marks)
- Recommend and justify any three graphical techniques that can help examine the dataset before multivariate analysis. (Explain what each technique reveals and why it is useful.) (6 marks)

OR

Q.3 Exploratory Factor Analysis (EFA) is widely used in multivariate data analysis to uncover underlying structures in observed variables. **(14)**

a) Explain the fundamental concept of factor analysis and discuss its importance in multivariate data analysis. (6 marks)

b) Describe the three major types of Exploratory Factor Analysis methods:
For each method, briefly explain its goal, how it works, and the type of variance it focuses on and Compare PCA, PAF, and MLFA. (8 marks)

Q.4 A researcher is modelling the probability of customer purchase (1 = Buy, 0 = Not Buy) using logistic regression. The model expresses the relationship between the predictors and the outcome through the logit transformation. **(14)**

Answer the following:

a) Explain why probabilities are first transformed into log-odds (logit) in logistic regression (7 marks)

b) Describe how the sigmoid (logistic) function converts the linear predictor (Z) back into a probability. (7 marks)

OR

Q.4 A researcher runs a multivariate analysis (such as MANOVA or Multiple Regression) and obtains statistically significant results. **(14)**
However, a senior statistician warns that "the results may not mean what you think they mean unless the core multivariate assumptions are met."

Using this remark as the context, explain how violations of each of the following assumptions can mislead the researcher, even if the output appears statistically significant:

For each assumption:
Briefly state what the assumption requires,
Describe one practical way of checking it, and
Explain how violating it can distort conclusions of multivariate analysis.

Q.5 All MCQs are compulsory, each contains 2 Marks **(14)**

1. Which of the following best explains why a researcher would use MANOVA instead of multiple ANOVAs?
   a) To increase the number of dependent variables tested independently
   b) To reduce Type I error when dependent variables are correlated
   c) To test interactions only between predictors
   d) To automatically perform discriminant analysis
2. A regression model has high R^2 but insignificant predictors. This likely indicates:
   a) Suppressor effect
   b) High multicollinearity
   c) Non-linearity in predictors
   d) Outliers only
3. The core purpose of conjoint analysis in marketing research is to:
   a) Predict which customer segments are most profitable
   b) Identify latent psychological traits behind preferences
   c) Decompose consumer choices to estimate attribute-level part-worth utilities
   d) Cluster consumers based on purchase behavior
4. In classical MDS, the primary goal is to:
   a) Predict classification accuracy based on similarity scores
   b) Represent similarity/dissimilarity data in low-dimensional spatial map
   c) Identify latent factors behind observed variables
   d) Reduce high-dimensional data using eigenvectors of covariance matrices only

P.T.O

5. Which of following is the FIRST step in screening multivariate data?
   a) Checking multicollinearity
   b) Assessing missing data
   c) Evaluating outliers
   d) Testing normality
6. Which of the following indicates problematic multicollinearity in multivariate analysis?
   a) Skewness value > 1
   b) VIF > 10
   c) Cronbach alpha < 0.7
   d) KMO < 0.5
7. A marketing researcher wants to test whether three age groups differ simultaneously on satisfaction, trust, and loyalty levels. Which method tests this overall effect?
   a) ANOVA
   b) MANOVA
   c) PCA
   d) Cluster Analysis

—X—