

**NOTE: Attempt Any Two Questions from SECTION-I
SECTION-II is Compulsory**

SECTION-I

Attempt Any Two Questions

Q.1. (a) What is Multivariate Analysis? Describe in detail the classification of Multivariate Techniques. (10)

મલ્ટિવેરિયેટ એનાલિસિસ શું છે? મલ્ટિવેરિયેટ ટેકનિકોના વર્ગીકરણ વિગતવાર વર્ણન કરો.

(b) What is an outlier? How will you detect and handle outliers in the data? (5)

આઉટલેટર એટલે શું? મલ્ટિવેરિયેટ ડેટામાં તેમને ઓળખવા અને તેમને શોધી અને નિયંત્રિત કરી શકશો?

Q.2. Describe in detail the structured approach to Multivariate Model Building. (10)

મલ્ટિવેરિયેટ મોડેલ બિલ્ડિંગના સ્ટ્રક્ચર્ડ અભિગમનું વિગતવાર વર્ણન કરો.

(b) Answer Briefly Any Two :

કોઈપણ બે ના જવાબ સંક્ષિપ્તમાં આપો.

(1) Distinguish between Hierarchical Cluster Analysis and Non Hierarchical Cluster Analysis

હાયરાર્કિકલ ક્લસ્ટર એનાલિસિસ અને નોન હાયરાર્કિકલ ક્લસ્ટર એનાલિસિસ વચ્ચેનો તફાવત સમજાવો.

(P.T.O)

(2) What is model parsimony?

મોડેલ પાર્સમની એટલે શું?

(3) State the guidelines for Multivariate Analysis.

મલ્ટિવેરિએટ એનાલિસિસ માટે માર્ગદર્શિકા જણાવો.

(4) Explain briefly the Measurement Scales with one illustration for each.

ટૂંકમાં દરેક માટે એક ઉદાહરણ સાથે માપનના પ્રકારો સમજાવો.

Q.3. (a) What is Factor Analysis? How does it differ from Cluster Analysis? Explain the procedure of Factor Analysis.

પરિબળ વિશ્લેષણ શું છે? તે ક્લસ્ટર એનાલિસિસથી કેવી રીતે અલગ પડે છે?

પરિબળ વિશ્લેષણની પ્રક્રિયા સમજાવો.

(b) Distinguish between data missing at random (MAR) and missing completely at random (MCAR). Explain how each type impacts the analysis of the missing data.

રેન્ડમ (MAR) પર ગુમ થયેલ ડેટા અને રેન્ડમ કમ્પ્લીટી અટ મીસિંગ (MCAR) પર સંપૂર્ણ રીતે ગુમ થયેલ માહિતી વચ્ચેનો તફાવત સમજાવો કે તે કયા પ્રકારના અસરકારક વિશ્લેષણને અસર કરે છે.

Q.4. (a) What is Discriminant Analysis? How does it differ from Multiple Regression Analysis? Explain the procedure of Discriminant Analysis.

ભેદભાવ વિશ્લેષણ શું છે? તે મલ્ટીપલ રીગ્રેસન એનાલિસિસથી કેવી રીતે અલગ પડે છે?

ભેદભાવ વિશ્લેષણની પ્રક્રિયા સમજાવો.

(b) The manager of the Raymond's showroom in Bangalore wants to set up a model to screen the customers and classify them as "loyal" and "loyal" based on the following three independent variables:

બેંગ્લોરમાં રેમન્ડના શોરૂમના સંચાલક, ગ્રાહકોને તપાસવા માટે એક મોડેલ સ્થાપવા બનાવવા માટે અને તેમને નીચેના ત્રણ સ્વતંત્ર ચલિત્ત્વોના આધારે "વફાદાર" અથવા "બેવફા" તરીકે વર્ગીકૃત કરવા માંગે છે:

(1)FREQ: Frequency of purchase in a year

એક વર્ષમાં ખરીદીની આવર્તન

(2)AVGPURC: average purchase by customer in a year

એક વર્ષમાં ગ્રાહક દ્વારા સરેરાશ ખર્ચ

E 95-3

(3) YEARS: number of years since the customer has been purchasing from the Raymond showroom.

ઘણા વર્ષોથી ગ્રાહક રેમન્ડ શોરૂમથી ખરીદી કરી રહ્યો છે

For this purpose he carried out Discriminant Analysis on the data of customers, whose output is given below. Interpret the output and write the conclusions: (17)

આ હેતુ માટે તેમણે 18 ગ્રાહકોના ડેટા પર ભેદભાવપૂર્ણ વિશ્લેષણ કાર્ય કર્યું જેનું આઉટપુટ નીચે આપેલ છે. આ આઉટપુટનો અર્થઘટન કરો અને અંતમાં નિષ્કર્ષ લખો:

Table:1

Discriminant Group	Classification Matrix	
	Percent Correct	G1 (predicted) G2 (predicted)
		Annual Income Threshold
G1 (observed)	88.8889	8
G2 (observed)	100	0
Total	94.4444	8

G1: Loyal વફાદાર

G2: Disloyal બેવફા

Table:2

Wilk's Lambda=0.3372559

(p-value =0.001272)

	Standardized Coefficient
Variable	Root 1
FREQ	-0.77738
AVGPURC	-0.73963
YEARS	-0.22719
Eigenval	1.965108
Cum. Prop	1.00000

SECTION-II

Multiple Choice Questions:

[10]

1. In one-way ANOVA, the null hypothesis may be tested by _____.

વન-વે અનોવા માં, નલ પૂર્વધારણા _____ દ્વારા પરીક્ષણ કરવામાં આવે છે.

- A. t statistic
- B. F statistic
- C. chi-square
- D. η^2

2. The _____ step of the data-preparation process consists of reviewing questionnaires to identify illegible, incomplete, inconsistent, or ambiguous data. ડેટા-તૈયારીની પ્રક્રિયાના _____ પગલામાં અયોગ્ય, અપૂર્ણ, અસંગત અથવા અસ્પષ્ટ પ્રતિભાવો ઓળખવા માટે પ્રશ્નાવલીઓની તપાસ કરવામાં આવે છે.

- A. coding
- B. data cleaning
- C. questionnaire checking
- D. editing

3. The general form of the multiple regression model is as follows, by which _____.

મલ્ટીપલ રીગ્રેસન મોડેલના સામાન્ય સ્વરૂપને અંકીત કરવામાં આવે છે, જે દ્વારા તેના _____ છે?

- A. $Y_i = \beta_0 + \beta_1 X_i + e_i$
- B. $\hat{Y} = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_k X_k$
- C. $\hat{Y} = a + b_1 X_1 + b_2 X_2$
- D. $\hat{Y}_i = a + b X_i$

4. _____ is the appropriate test statistic to use to determine the significance of the coefficient of determination in bivariate regression.

દ્વીચલીય રીગ્રેસનમાં નિશ્ચિતતાના ગુણાંકના મહત્વને નક્કી કરવા માટે _____ પરીક્ષણ આંકડા છે.

- A. F statistic
- B. T statistic
- C. Z statistic
- D. ω^2

5. D is the _____ in the model $D=b_0+ b_1 X_1+ b_2 X_2+ b_3 X_3+...+ b_k X_k$.

ડી એ મોડેલ $D=b_0+ b_1 X_1+ b_2 X_2+ b_3 X_3+...+ b_k X_k$ માં _____ છે.

- A. discriminant score
- B. disordinal interaction
- C. difference variable
- D. discriminant coefficients or weights

6. In discriminant analysis, the criterion or dependent variable is _____ and the predictor or independent variables are _____ in nature.

ભેદભાવ વિશ્લેષણમાં, માપદંડ અથવા આશ્રિત ચલ _____ છે અને આગાહી કરનાર અથવા સ્વતંત્ર ચલો _____ પ્રકૃતિમાં છે.

- A. interval; categorical
- B. ordinal; interval
- C. categorical; interval
- D. ordinal; categorical

7. In Factor Analysis, the amount of variance a variable shares with all other variables included in the factor analysis is referred to as _____.

પરિબળ વિશ્લેષણમાં, પરિબળ વિશ્લેષણમાં શામેલ અન્ય બધા ચલો સાથેના વેરિયેબલની ભિન્નતાની માત્રા _____ તરીકે ઓળખવામાં આવે છે.

- A. communality
- B. total variance
- C. shared variance
- D. percentage of variance

8. Which method of analysis does not classify variables as dependent or independent?

વિશ્લેષણની કઈ પદ્ધતિ ચલને આશ્રિત અથવા સ્વતંત્ર તરીકે વર્ગીકૃત કરતી નથી?

- A. regression analysis
- B. discriminant analysis
- C. analysis of variance
- D. factor analysis

9. _____ is a class of techniques used to classify objects or cases into relatively homogeneous groups.

વર્ગ પદાર્થો અથવા કેસોને પ્રમાણમાં એકરૂપ જૂથોમાં વર્ગીકૃત કરવા માટે _____ તકનીક વપરાય છે.

- A. Principal components analysis
- B. Cluster analysis
- C. Common factor analysis
- D. Discriminant analysis

E 95-6

10. In Cluster Analysis, the most commonly used measure of similarity is _____ or its square.

ક્લસ્ટર એનાલિસિસમાં, સમાનતાનો સૌથી વધુ ઉપયોગ થતો માપ _____ અથવા તેનો ચોરસ છે.

- A. Euclidean distance
- B. city-block distance
- C. Chebychev's distance
- D. Manhattan distance
