JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT
TEST -3 EXAMINATION- 2025
PhD – Sem I (For 256802)

COURSE CODE (CREDITS): 18P1WGE101 (3)

COURSE NAME: RMIQMCA

COURSE INSTRUCTORS: ASA

MAX. MARKS: 25

MAX. TIME: 2 Hours

Note: *(a) All questions are compulsory.*

(b) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems

Q.No	Question	CO	Marks
Q1	GreenLeaf Organics tested a new herbal shampoo on 5,000 customers to see if it improved satisfaction compared to the old product. The average satisfaction score increased only slightly—from 4.20 to 4.24—yet this tiny difference became statistically significant ($p < 0.01$) because of the very large sample size. Managers felt this improvement was too small to matter in real-world decision-making. At the same time, among customers with sensitive scalp, complaints dropped by 12%, which managers viewed as meaningful for a niche segment. However, this reduction was not statistically significant at the 0.05 level due to the smaller subgroup size. Thus, one finding is statistically significant but not practically relevant, while the other is practically important but not statistically significant. Using the GreenLeaf Organics case, explain the differences between statistical significance and practical significance. Is one a prerequisite for the other?		4
Q2	"Missing data can have significant impacts on any analysis, particularly those of a multivariate nature." Discuss the statement with identification of missing data and remedy for it.		4
Q3	FreshBasket is a national hypermarket chain operating in 40 cities with over 120 stores. After a decade of stable growth, the company is facing increasing competition from online retailers and new brick-and-mortar entrants. To strengthen its market position, FreshBasket launches a Customer Insights Initiative (CII) to better understand the factors influencing customer satisfaction, loyalty, and spending behaviour. The company collects data from 2,000 customers through loyalty card transactions and an in-store survey. The dataset contains: Demographic Variables (Age, Income, Family size, Education level, Occupation);		4+3 = 7

	Shopping Behaviour Variables (Monthly store visits, Average bill amount, Product category preferences (e.g., groceries, personal care, home products), Promotion sensitivity index, Time spent per visit); Attitudinal Variables (Perceived product quality, Store ambience rating, Staff courtesy, Perceived value for money, Satisfaction, Shopping enjoyment, Loyalty intention); Transactional Variables (Number of items per visit, Membership tier (Silver/Gold/Platinum), Coupon usage, Online vs offline purchase ratio) a) Identify the variables suitable for Dependance Technique Vs Independence Technique. Develop the suitable model also in both the cases. b) Discuss how the technique would change if any key dependent variable became: Categorical instead of metric OR Metric instead of categorical.			

Q4

Following is the result of a multivariate regression analysis with sample size 420:

Independent Variable	Coefficient (β)	Std. Error	p-value
Constant	18.42	2.9	<0.001
Self-Regulated Learning (SRL)	0.62	0.08	<0.001
Digital Competence (DC)	0.29	0.07	<0.001
Online Engagement (OE)	0.11	0.05	0.028
Digital Distraction (DD)	−0.47	0.09	<0.001
Motivation for Learning (ML)	0.18	0.06	0.003

$R^2 = 0.61$; Adusted $R^2 = 0.60$; F-stat = 131.5; p-value (model) = 0.001.

a) Which predictors align strongly with existing academic behaviour theories?
b) Why might Digital Distraction have a strong negative coefficient?
c) Is an $R^2 = 0.61$ theoretically plausible for behavioural research?
d) Do the significant predictors collectively make sense in the context of blended learning theory?

Marks: 1.5x4 = 6

Q5

"CFA is similar to EFA in some respects, but philosophically it is quite different". Discuss the statement with suitable example.

Marks: 4