

JK Lakshmipat University Jaipur
INSTITUTE OF ENGINEERING AND TECHNOLOGY

End Term Examination, December 2025

B. Tech. (CSE)- Semester VII and V and BCA- Semester -V, 2025-26

Roll No.....

MA1107: Marketing Analytics using R, Python and
Tableau

Time: 2.5 hours

Max. Marks: 30

Instructions to students:

1. Do not write anything, other than your roll number, on the question paper.
2. Use of a normal calculator is permitted.
3. Attempt any three questions.

Q. 1 (a) A consumer electronics company wants to understand how customer journey mapping can help optimize its online store experience. Using the steps in marketing analytics, explain how the company should proceed from setting goals to supporting decisions. Illustrate each step with a relevant example. 4.0 (LO1)

(b) A consumer electronics company recently surveyed 40 dissatisfied customers to understand why they stopped purchasing its smart home devices. Customers were asked to share open-ended feedback about their dissatisfaction. After analyzing and grouping the comments, the following table was developed: 6.0 (LO1)

Category (Theme)	#of Comments	Total Comments	%	95% CI Low	95% CI High
Inferior product quality	20	40	50%	35%	65%
High prices	10	40	25%	14%	40%
Reliability of services	6	40	15%	6%	29%
Limited availability in stores	4	40	10%	3%	26%

In the light of the above results, answer the following questions:

- (i) Explain the process of converting qualitative comments into quantitative insights.
- (ii) Interpret the table:
 - Identify the top two problem areas based on percentage and confidence intervals.
 - Explain what the 95% confidence interval tells us about the reliability of each category's estimate.

(iii) As a marketing analyst, recommend two strategic actions the company should take to address customer concerns and improve satisfaction.

Q. 2 (a) You are working for an e-commerce platform trying to cluster users based on browsing and purchase behavior. Discuss how you would determine the optimal number of clusters (k) from both an analytical and business feasibility perspective. What risks arise if the chosen k is too small. 02.0 (LO3)

(b) A K-Means cluster analysis was conducted to segment bank customers based on their average balance, number of transactions, loan amount, and tenure (years). Results of the analysis were as under: 08.0 (LO3)

Table 1: Summary Diagnostics

Metric	Value
Number of Clusters	3
Number of Observations	30
Between-group Sum of Squares	10.80
Within-group Sum of Squares	4.20
Total Sum of Squares	15.00

Table 2: Cluster Summary

Cluster	Number of Customers	Avg. Balance	Avg. Transactions	Avg. Loan Amount	Avg. Tenure
1	8	0.85	0.78	0.62	0.70
2	10	0.25	0.30	0.15	0.28
3	12	0.60	0.58	0.50	0.66

In the light of the above results, answer the following questions

- Explain the difference between within-cluster homogeneity and between-cluster heterogeneity.
- Calculate within-group heterogeneity.
- Interpret the business meaning of each cluster.
- Suggest one marketing strategy per cluster.

Q. 3 (a) A farmer believes that the average yield of a new chili variety is greater than 12 tons per acre. To verify this claim, yield data (in tons per acre) were collected from 10 different farm plots as shown below: 06.0 (LO2)

Name of csv file- chili_yield_data

Plot_ID	Yield_ton_per_acre
Plot_1	11.5
Plot_2	12.8
Plot_3	13.2
Plot_4	12.7
Plot_5	11.9
Plot_6	13.0
Plot_7	12.5
Plot_8	12.6
Plot_9	13.1
Plot_10	12.4

The farmer wishes to statistically test the claim using a one-sample t-test at the 5% significance level.

Questions:

- Write the null and alternative hypotheses for this problem.
- Identify whether this is a left-tailed, right-tailed, or two-tailed test, and justify your choice.
- Write the R syntax to perform the hypothesis test using the given data.

(b) A retail chain claims that the average weekly sales at its new stores exceed ₹8.5 lakh. A random sample of 15 stores was selected to test this claim. After running the one-sample t-test, the analysis result is: 04.0
(LO2)

- One-Sample t-Test Results
- Sample Mean = 9.12
- Population Mean (μ_0) = 7.5
- $t = 2.18$, $df = 14$, $p = 0.046$
- 95% CI = [8.52, 9.72]

In the light of the above results, answer the following questions:

- Based on the p-value, should the company's claim be accepted at a 5% significance level?
- What does the confidence interval suggest about the average sales performance?

Q. 4 (a) The Department of Agriculture ran a regression to estimate how fertilizer usage (kg/acre) affects crop yield (quintals/acre). Results of the analysis were as under: 05.0
(LO4)

Regression Output:

Variable	Coefficients	Standard Error	t Stat	P-value
Intercept	8.25	1.40	5.89	0.0001

Fertilizer (kg/acre)	0.42	0.09	4.67	0.0007
-------------------------	------	------	------	--------

Dependent Variable: Crop Yield (quintals/acre)

$R^2 = 0.82$

In the light of the above results, answer the following questions:

- Write the estimated regression equation.
- What would be the expected crop yield when fertilizer usage is 100 kg/acre?
- What does the R^2 value imply about this model's explanatory power?

(b) Regression was run to predict hotel occupancy rate (%) using the following independent variables:

05.0
(LO4)

- PRICE = average room price (₹)
- RATING = average TripAdvisor rating (1–5 scale)
- LOCATION = location score (1–10 scale)

Results of the analysis were as under:

Table 1 – ANOVA

Source	df	SS	MS	F	Sig. F
Regression	3	22140.5	7380.17	8.72	0.001
Residual	26	22016.4	846.02		
Total	29	44156.9			

Table 2 – Regression Coefficients

Variable	Coefficients	Std. Error	t Stat	P-value
Intercept	22.8	9.5	2.4	0.024
PRICE	-0.09	0.03	-3.0	0.006
RATING	5.4	1.7	3.18	0.004
LOCATION	2.9	1.1	2.64	0.013

Dependent Variable: Hotel Occupancy Rate (%)

$R^2 = 0.73$

In the light of the above results, answer the following questions:

- Write the regression equation for predicting hotel occupancy rate.
- Based on the ANOVA results, should the null hypothesis (that all predictors jointly have no effect) be rejected at the 5% level? Explain.
- Calculate the predicted occupancy rate for a hotel with the following data: PRICE = ₹5,000, RATING = 4.2, LOCATION = 8.
- A second hotel has PRICE = ₹4,000, RATING = 3.5, LOCATION = 5. Compute its predicted occupancy rate and compare it with the previous result of question (iii).