DELHI SCHOOL OF BUSINESS

By Vivekananda Institute of Professional Studies - TC

Delhi School of Business
PGDM / PGDM (FINTECH) Program
END-TERM EXAMINATION, JAN-FEB 2024
TERM – II (Batch: 2023-25)

Course Name	**R for Managers**	**Course Code**	
Duration	**2.5 Hours**	**Max. Marks**	**60**

Instructions:

1.You will use R Studio to answer the questions and use required packages.

2. Please write your Answers along with R code and the Output on your Answer sheets.

3. Please clearly write your Insights, Interpretations, Recommendations, Examples along with R code used on your Answer Sheets.

4. Please do not draw any charts/ graphs on Answer sheets but you need to mention the Insights, Interpretations of charts along with R code used for drawing those charts in your Answer Sheets.

5. All the six questions are compulsory and carry equal marks.

6. There is no negative marking. However, write comprehensive answers.

Questions:

- **1. Data Types and Structures (10 marks):**
  a. Write R code to demonstrate the creation of variables with different data types (numeric, character, logical, and factor) and the manipulation of various data structures, including vectors, matrices, lists, and data frames. For this question, either use inbuilt R datasets or create a sample dataset. Provide examples and explain the output. Discuss the importance of choosing the correct data type and structure in R programming. (10 marks)

- **2. Data Wrangling with ply (10 marks):**
  a. Using the dply package in R, write code to perform common data wrangling tasks on a given dataset. For this question, either use an inbuilt R dataset or create a sample dataset. Include operations such as filtering rows, selecting columns, creating new variables, and

summarizing data. Provide examples and discuss how the dplyr functions facilitate efficient and readable data manipulation in R. (10 marks)

- **3 .Statistical Analyses (10 marks):**
  a. Write R code to perform a statistical analysis, such as a t-test, z-test & proportion test, using a sample dataset. For this question, either use an inbuilt R dataset or create a sample dataset. Include loading the data, conducting the analysis, and interpreting the results. Discuss the practical significance of the analysis and how it can be applied to real-world scenarios. (10 marks)

- **4. Visualization and Exploratory Data Analysis (EDA) (10 marks):**
  a. Utilize R code to generate visualizations for exploratory data analysis, such as scatter plots, histograms, and boxplots. For this question, either use an inbuilt R dataset or create a sample dataset. Choose a dataset, load it into R, and create visualizations to uncover patterns and trends. Discuss the insights gained from the visualizations and the importance of EDA in data analysis. (10 marks)

- **5. Decision Making (10 marks)**

  a. A new detergent powder was recently developed by a multinational company. It was advertised in two States of India using two different media. In State A, it was advertised on TV only, while in State B, it was advertised on FM radio and local newspapers. The amount spent on advertisement was equal in both the States. After a month of the advertisement campaign, a survey was conducted to assess the awareness about this detergent in both the States. In State A, 600 people were selected randomly out of which 360 responded that they were aware about the new detergent powder. In State B, a random sample of 800 people was taken, out of which 620 indicated that they were aware about the new product.

  On the basis of this survey, can we conclude that the advertising media used in the two States were equally effective (or, there was no difference) in making the consumers aware about the new product? Test at 95% level of confidence.
  Please suggest your recommendation.
  (5 Marks)

b. A health club recently introduced a weight-reducing plan. From among a large number of people who participated in the plan, a random sample of 11 people was chosen and weights, (in kgs) were recorded both before and after completion of the plan. The information is given below. Using 5% level of significance, test whether the weight-reducing plan had been really effective in reducing the weight.

Person	1	2	3	4	5	6	7	8	9	10	11
Wt (before)	87.4	92.9	83.6	81.5	89.7	100.5	98.6	88.8	112.4	87.6	92.8
Wt (after)	85.4	88.3	84.7	81.2	83.3	94.6	90.1	87.2	104.6	88.4	91.7

Please elaborate your finding.
(5 Marks)

6. Business Insight Question:(10 marks)

a. Leveraging the hypothetical dataset containing information on customer age, purchase patterns, and satisfaction with services, employ appropriate visualization tools to extract insights. Additionally, conduct a statistical test to confirm if there is significant association between Customer Purchase Patterns & Satisfaction level.

What basic insights can be drawn from the data, and what simple recommendations would you provide to enhance customer satisfaction?

Hypothetical Dataset:
Consider the following simplified hypothetical dataset with columns:

- Customer_ID: Unique identifier for each customer.
- Age: Age of the customer.
- Purchase_Patterns: Frequency of purchases made by the customer (on a scale from 1 to 5, where 5 is frequent).
- Satisfaction: Customer satisfaction score (on a scale from 1 to 10, where 10 is highly satisfied).

```
# Generating simplified hypothetical data in R
num_customers <- 50
customer_data <- data.frame(
  Customer_ID = 1:num_customers,
  Age = sample(18:25, num_customers, replace = TRUE),
  Purchase_Patterns = sample(1:5, num_customers, replace = TRUE),
  Satisfaction = sample(5:10, num_customers, replace = TRUE)
)
```