

(4)

MB-303(BA2)

using data visualization techniques, for the following points :

- Sales percentage/proportion of products sold
- Outliers in a sales dataset of products
- To depict the relationship between two products and the trends between them

Now which plot you will use and why to create reports ?

MB-303(BA2)

180

H

Roll No.

MB-303(BA2)

M. B. A. (THIRD SEMESTER)

END SEMESTER

EXAMINATION, Jan., 2023

DATA SCIENCE USING R

Time : Three Hours

Maximum Marks : 100

Note : (i) This question paper contains two Sections—Section A and Section B.

(ii) Both Sections are compulsory.

(iii) Answer any *two* sub-questions among (a), (b) and (c) in each main question of Section A. Each sub-question carries 10 marks.

(iv) Section B consisting of a case study is compulsory. Section B is of 20 marks.

P. T. O.

Section—A

1. (a) Define Data Science. Discuss the role of Python in Data Science. (CO1)
- (b) Outline various applications of data science and the role of big data in data science. (CO3)
- (c) Assess the basic skills, roles and responsibilities of a Data Scientist. (CO4)
2. (a) Recommend various functions of the DPLYR package to manage the dataframe. (CO4)
- (b) Describe the Hypothesis test and also discuss the level of confidence and significance. (CO1)
- (c) Demonstrate Descriptive, Diagnostic, Predictive and Prescriptive analysis with an example. (CO2)
3. (a) What is the use of CRAN in RStudio ? Discuss 3 packages and their use in RStudio. (CO2)

- (b) How to construct clusters using cluster analysis ? Discuss hierarchical and k-means clustering. (CO5)
- (c) Discuss the concept of multiple regression Analysis. Also, explain the concept of multicollinearity. (CO1)
4. (a) Discuss T, Z and ANOVA tests. How will you test the hypothesis using these tests in R ? (CO1 & CO3)
- (b) What is the use of Data Visualization ? Discuss different plots with their advantages. (CO2)
- (c) Create data structures : Vectors, Matrix and Data Frame in R with suitable examples. (CO5)

Section—B

5. **Case Study :** (20 Marks) (CO5)

Caselet/Numerical

In an FMCG company sales have been recorded for the last financial year, now being as a data scientist you have to depict reports