19CMU34

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BSc DEGREE EXAMINATION MAY 2024
(Sixth Semester)

Branch – **COMPUTER SCIENCE**

DATA SCIENCE WITH R PROGRAMMING

Time: Three Hours Maximum: 50 Marks

SECTION-A (5 Marks)
Answer **ALL** questions
ALL questions carry **EQUAL** marks (5 x 1 = 5)

1. Which package in R is commonly used for data manipulation tasks, such as data cleaning and transformation?
   (i) ggplot (ii) dplyr
   (iii) tidyr (iv) lubridate

2. Which function in R is used to create interactive plots for data visualization?
   (i) plotly() (ii) ggplotly()
   (iii) shiny() (iv) visNetwork()

3. Which package in R is commonly used for building and training machine learning models?
   (i) dplyr (ii) ggplot
   (iii) caret (iv) tidyr

4. What does the 'hclust()' function in R perform?
   (i) Hierarchical Clustering (ii) Density-based Clustering
   (iii) Spectral Clustering (iv) K-Means Clustering

5. Which package in R is primarily used for creating static and interactive visualizations?
   (i) tidyr (ii) lubridate
   (iii) ggplot2 (iv) caret

SECTION - B (15 Marks)
Answer **ALL** Questions
ALL Questions Carry **EQUAL** Marks (5 x 3 = 15)

6 a. Narrate about reading data in R.
OR
b. Describe data frames in R with example.

7 a. Bring out the grammar of graphics.
OR
b. Describe Data types in R programming.

8 a. Outline about the supervised learning.
OR
b. Summarize on Naïve Bayes Algorithm.

9 a. Explain the hierarchical clustering in unsupervised learning.
OR
b. Describe dimensionality reduction in unsupervised learning.

10 a. Outline function operations in R.
OR
b. Sketch Roxygen in Advanced R programming.

Cont...

SECTION -C (30 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** Marks (5 x 6 = 30)

11 a. Elucidate data manipulation.

OR

b. Create a R program for calculator and explain.

12 a. Summarize graphics in visualizing data.

OR

b. Explain the control structures that are available in R programming.

13 a. Explain about the Linear Regression in R.

OR

b. Describe how to validate models in R Programming.

14 a Explain about the K-Means Clustering in detail.

OR

b. Narrate on Association Rules in Unsupervised learning.

15 a. Explain the vectors and vectorization functions in R.

OR

b. Explain how to create and build Packages in R Programming.

Z-Z-Z END