Department of Artificial Intelligence and Data Science

RAMAIAH
Institute of Technology

QUESTION BANK FOR IV SEMESTER
(Term: 15th April 2024-22nd July 2024)
Statistical Programming Laboratory (ADL 47)

Marks: 50   **Exam Hours: 03**

Credits: 0:0:1

	Part A
1	Suppose that queue <- c ("Steve", "Russell", "Alison", "Liam") and that queuerepresents a supermarket queue with Steve first in line. Using R expressions update the supermarketqueue as successively: a) Barry arrives; b) Steve is served; c) Pam talks her way to the front with one item; d) Barry gets impatient and leaves; e) Alison gets impatient and leaves. f) For the last case you should not assume that you know where in the queue, Alison is standing. g) Finally, using the function which(x), find the position of Russell in the queue.
2	Plot the graph of linear equation, Non-linear equation (square, cubic, square root, etc.), Trigonometric function and logarithmic functions
3	The performance of a student in 3rd semester CSE as given below {SUB1, SUB2, SUB3, SUB4, SUB5, SUB6} with score {91, 73, 65, 45, 54, 32}. a) Draw the Bar chart to indicate the performance of the student. b) Draw the pie chart to indicate the performance of the student. c) Find the average marks and discuss qualitatively about the performance in each subjects.
4	Write a R program to extract first 10 English letters in lower case and last 10 letters in upper case and extract letters between 22nd to 24th letters in upper case.
5	Write a R program to create an array of two 3x3 matrices each with 3 rows and 3 columns from two given two vectors. Print the second row of the second matrix of the array and the element in the 3rd row and 3rd column of the 1st matrix.
6	Write a R program to demonstrate the use of rbind() function to combine two or more data frames in vertically.
7	Write a R program to demonstrate various Apply functions(lapply and sapply)
8	Write a R Program to demonstrate debugging through traceback ().

	PART – B
1.	You are working as a cashier at a grocery store. Your task is to create a program that simulates the checkout process for a customer's shopping cart. The program should calculate the total cost of the items, including tax, and provide a detailed receipt. i. Define a list of products, each represented as a dictionary with keys: "name", "price", and "quantity". ii. Allow the cashier to input the products in the customer's shopping cart, including the name, price, and quantity of each item. iii. Calculate the subtotal (price * quantity) for each item and display a detailed receipt with product names, quantities, prices, and subtotals. iv. Calculate the total cost of the items in the cart before tax. v. Apply a tax rate (e.g., 8%) to the total cost to calculate the tax amount. vi. Calculate the final total cost by adding the tax amount to the total cost before tax.
2	You have been tasked with creating a program that calculates and assigns grades for students enrolled in multiple courses. The program will take input for the marks obtained by 10 students in 5 different courses, compute the total and average marks for each student, and assign corresponding grades based on their average performance. Declare constants for the number of students (num_students) and the number of courses (num_courses). Initialize an empty list to store student information. For each student: • Input the student's name. • Input marks for each of the 5 courses. • Calculate the total marks and average marks. • Determine the grade based on the average marks using a grading scale. • Display the student information, including their name, individual course marks, total marks, average marks, and the assigned grade.

3	You are developing an inventory management system for a small store. The system needs to handle inventory items and their quantities. Write a program that uses arrays to store inventory items and their quantities, and includes functions to add new items, update quantities, and display the inventory.
	
	 • Define an array to store inventory items.
	 • Define an array to store corresponding quantities.
	 • Implement functions to:
	     ○ Add a new item along with its quantity.
	     ○ Update the quantity of an existing item.
	     ○ Display the inventory items and quantities.
	     ○ Use the functions to manage the inventory and handle user interactions.

4	You are working as an educational analyst and need to analyze the performance of students in a school. You have data on student names, their scores in different subjects, and attendance. Write a program that uses data frames to manage and analyze student data, including calculating average scores, identifying students with low attendance, and generating a report. Create a data frame to store student information with columns: "Name", "Math_Score", "Science_Score", "History_Score", "Attendance". Implement functions to: • Calculate the average scores for each student. • Identify students with attendance below a certain threshold.

	• Generate a report with student names, average scores, and attendance status. • Use the functions to analyse student performance and generate the report.
5	You are a data analyst at a retail company that sells products online. The company is interested in predicting sales for the upcoming months to better manage inventory and plan marketing strategies. As part of your role, you need to develop a program that utilizes time series analysis to forecast sales based on a historical sales dataset. Write an R program to forecast sales for the next three months using time series analysis techniques. The program should perform the following steps: • Load the required libraries, including the forecast package. • Create a data frame with two columns: Month and Sales. The Month column should contain a sequence of dates from January 2023 to June 2023 (inclusive), and the Sales column should contain the corresponding sales amounts (12000, 15000, 18000, 16000, 20000, 22000). • Convert the sales data into a time series object with a monthly frequency. • Fit an ARIMA (AutoRegressive Integrated Moving Average) model to the sales time series using the auto.arima() function. • Forecast sales for the next three months using the fitted ARIMA model and the forecast() function. • Display the forecasted sales results, including point forecasts and prediction intervals.
6	You are a data analyst working for an e-commerce company that specializes in selling a variety of products online. The company aims to analyze customer purchase data comprehensively to gain insights into customer behavior and spending patterns. Your goal is to develop a R program that performs an in-depth analysis of customer purchase data. You will calculate various statistical measures and generate visualizations to understand the distribution of purchase amounts among customers. Note: Load the necessary libraries, including the dplyr and ggplot2 packages. Given the example customer purchase data provided below, create a data frame named purchase_data with two columns: CustomerID and PurchaseAmount. Calculate and display the following statistical measures: • Mean (average) purchase amount • Median purchase amount • Standard deviation of purchase amounts • 1st quartile (25th percentile) of purchase amounts • 3rd quartile (75th percentile) of purchase amounts Create a histogram to visualize the distribution of purchase amounts using the ggplot2 package. Display the histogram with appropriate labels and titles. Example Customer Purchase Data: <table><tr><td>CustomerID</td><td>PurchaseAmount</td></tr><tr><td>101</td><td>150</td></tr><tr><td>102</td><td>200</td></tr><tr><td>103</td><td>120</td></tr><tr><td>104</td><td>300</td></tr><tr><td>105</td><td>80</td></tr></table>
7	Write an R program that generates two matrices, matrix_A and matrix_B, and conducts operations including element-wise addition, scalar multiplication, matrix transpose, and multiplication.

Department of Artificial Intelligence and Data Science

RAMAIAH Institute of Technology

8	You are a data analyst tasked with analyzing and visualizing a dataset. The dataset contains information about sales .
	 Write a program that performs data analysis and generates visualizations for a given dataset. The program should:
	   a) Load the necessary libraries (ggplot2).
	   b) Prepare example sales data
	   c) Perform the following tasks:
	     i. Create a scatter plot to visualize the relationship between scores and attendance percentages.
	     ii. Generate a bar plot to show the distribution of scores among different students.
	     iii. Create a line plot to display the trend of scores over time (assuming different students' scores were collected at different time intervals).
	     iv. Generate a histogram to visualize the distribution of scores.
	   d) Customize the appearance of each plot, such as color, labels, and titles.
	   e) Arrange the plots in a way that they are easy to compare and understand.
	   f) Provide appropriate titles for each plot and the axes.

9.	You are a data analyst at an e-commerce company that sells a variety of products online. The company has provided you with a dataset containing information about customer purchases. Your task is to perform a comprehensive data analysis to gain insights into customer behavior and spending patterns.
	 Dataset Description:
	   The dataset customer_purchases.csv contains the following columns:
	   CustomerID: Unique identifier for each customer.
	   PurchaseAmount: The amount spent by the customer on a purchase.
	 Problem Tasks:
	 You are required to perform the following tasks using R:
	 Task 1: Load the Dataset
	   Load the necessary libraries, including readr and dplyr.
	   Read the dataset customer_purchases.csv into a data frame named purchase_data.
	 Task 2: Data Summary
	   Calculate and display the total number of records in the dataset.
	   Calculate and display the total number of unique customers in the dataset.
	 Task 3: Calculate Statistical Measures
	   Calculate and display the mean, median and standard deviation for purchase amount.
	 Task 4: Customer Segmentation
	   Create a new column named Segment in the purchase_data data frame based on the following criteria:
	   "Low Spender" if the purchase amount is less than the median.
	   "High Spender" if the purchase amount is greater than or equal to the median.
	 Task 5: Visualize Data

Note:

Conduction and Result (both (a) & (b))	**: 35 Marks (Part A: 15 Marks, B: 20 Marks)**
Write-up	**: 8 Marks**
Viva	**: 7 Marks**
COP	**: -10 Marks**