SRM VALLIAMMAI ENGINEERING COLLEGE

(An Autonomous Institution)

SRM Nagar, Kattankulathur – 603 203

DEPARTMENT OF INFORMATION TECHNOLOGY

QUESTION BANK

I SEMESTER

1924102-DATA SCIENCE AND R ESSENTIALS

Regulation – 2019

Academic Year 2021 – 2022 (ODD SEMESTER)

Prepared by

Dr. S. SEKAR, Assistant Professor (Sel. G) / IT

SRM VALLIAMMAI ENGINEERING COLLEGE

(An Autonomous Institution)

SRM Nagar, Kattankulathur – 603 203.

DEPARTMENT OF INFORMATION TECHNOLOGY

QUESTION BANK

SUBJECT : 1924102-DATA SCIENCE AND R ESSENTIALS

SEM / YEAR : I Sem / I Year

UNIT I -INTRODUCTION

Introduction to core concepts and technologies: Introduction Terminology, data science process, data science toolkit, Types of data, Example applications. Data collection and management: Introduction, Sources of data, Data collection and APIs, Exploring and fixing data, Data storage and management, Using multiple data sources

PART – A

Q.No	Questions	BT Level	Competence
1.	What is Data Science?	BTL1	Remembering
2.	Who is a Data Scientist?	BTL1	Remembering
3.	Differentiate between structured and unstructured data.	BTL2	Understanding
4.	Mention the process in data preparation phase.	BTL1	Remembering
5.	Differentiate between Business Intelligence vs Data Science	BTL3	Applying
6.	List out the Open Source Big Data Tools.	BTL1	Remembering
7.	State the effective methods of data collection?	BTL3	Applying
8.	What you meant by data modelling?	BTL2	Understanding
9.	List out the phases of data science Life Cycle?	BTL4	Analyzing
10.	What is meant by data wrangling?	BTL1	Remembering
11.	What does a Data Scientist do?	BTL3	Applying
12.	What are the sources of the data?	BTL1	Remembering
13.	Difference between an API and a Library	BTL2	Understanding
14.	Discuss Primary Data and Secondary Data.	BTL2	Understanding
15.	Mention the various fields in computer science that utilize Data Science techniques.	BTL4	Analyzing
16.	Mention the categories of API in data science	BTL6	Creating
17.	List out the various components in data Science	BTL5	Evaluating
18.	Point out the basic elements of an API.	BTL5	Evaluating
19.	Identify the components of big data storage infrastructure.	BTL6	Creating
20.	List out the solutions for dealing the missing data.	BTL4	Analyzing

PART – B

1.	Explain the various phases of the data science life cycle?	(13)	BTL1	Remembering
2.	Explain in detail about open source data tool kits?	(13)	BTL1	Remembering
3.	(i)Explain the importance of Data Science? (ii) Describe the various applications of the data science?	(6) (7)	BTL1	Remembering

4.	(i)Differentiate between Business Intelligence vs Data Science (10) (ii) Discuss the Big data storage security strategies (3)	BTL1	Remembering
5.	Illustrate the various components of data science in detail? (13)	BTL2	Understanding
6.	Who is a Data Scientist? What are their roles and responsibilities? (13)	BTL2	Understanding
7.	Explain the type of data and Measurement Scales in detail? (13)	BTL2	Understanding
8.	(i)Discuss the Analytics Tools in detail (6) (ii)Discuss the Data Visualization tools in detail (7)	BTL3	Applying
9.	Discuss the different API should know for every data Scientist? (13)	BTL3	Applying
10	Summarize the various options to dealing with missing data (13)	BTL4	Analyzing
11.	(i)How big data storage compares to traditional enterprise storage (6) (ii) How machine learning affects big data storage (7)	BTL4	Analyzing
12.	Discuss the data storage management challenges and how to deal with them (13)	BTL4	Analyzing
13.	Explain briefly about the data science process (13)	BTL5	Evaluating
14.	Discuss the five important of sources of big data with an real time examples (13)	BTL6	Creating
	PART C		
1.	What if we could predict the occurrence of diabetes and take appropriate measures beforehand to prevent it? Diabetes Prevention(case study) (15)	BTL6	Creating
2.	Discuss the any five data science projects using API. (15)	BTL6	Creating
3.	Discuss about the various types of data in data science explain each type with example (15)	BTL5	Evaluating
4.	(i) Explain in detail the Top Tools for Data Scientists (8) (ii) Mention various Key features for Tools for Data Scientists (7)	BTL5	Evaluating

UNIT II – Data Analysis

Data analysis: Introduction, Terminology and concepts, Introduction to statistics, Central tendencies and distributions, Variance, Distribution properties and arithmetic, Samples/CLT, Basic machine learning algorithms, Linear regression, SVM, Naïve Bayes. Data visualization Introduction, Types of data visualization

PART – A

Q.No	Questions	BT Level	Competence
1	What is data analysis with example?	BTL1	Remembering
2	Summarize the concepts of data cleaning.	BTL1	Remembering
3	Define data collection	BTL1	Remembering
4	What you meant by Descriptive statistics	BTL3	Applying
5	Write the types data analysis?	BTL1	Remembering
6	What is Data Mining?	BTL3	Remembering
7	Write the goal of business intelligence	BTL2	Understanding
8	Mention the main purpose of data analysis?	BTL2	Understanding
9	What is Support Vector Machine?	BTL2	Understanding
10	What is Predictive Analytics?	BTL2	Understanding
11	What is machine learning?	BTL1	Applying

12	Why is data analysis?	BTL1	Remembering
13	Interpret linear regression	BTL3	Analyzing
14	Point out the applications of Naïve bayes classifier	BTL4	Applying
15	Illustrate the supervised learning	BTL4	Analyzing
16	Express the Multiple linear regression	BTL4	Analyzing
17	Explain the mean,mode.	BTL5	Evaluating
18	State the data analysis with Excel?	BTL5	Evaluating
19	Give the different types of regression	BTL6	Creating
20	Differentiate regression and correlation	BTL6	Creating

PART – B

1	Explain in detail about the various data analysis techniques (13)	BTL1	Remembering
2	(i) Discuss about regression. List the types of regression. (7) (ii) List the purpose of using Regression Modelling in DataAnalysis. (6)	BTL1	Remembering
3	Discuss about mean medium, mode for discrete and continuous data with example(13)	BTL1	Remembering
4	Describe the different phases of Data Analysis Process. (13)	BTL1	Remembering
5	Discuss about the Descriptive statistics with an example. (13)	BTL2	Understanding
6	(i)What is the Central Limit Theorem and why is it important? (7) (ii) What is sampling? How many sampling methods do you know? (6)	BTL2	Understanding
7	(i)Explain the concepts of Support Vector Machine(SVM) algorithm (8) (ii)Write Pros and Cons associated with SVM (5)	BTL2	Understanding
8	Explain the concepts of linear regression in machine learning (13)	BTL3	Applying
9	Write the concepts of central limit theorem with example. (13)	BTL3	Applying
10	(i)What are the assumptions required for linear regression? (6) (ii) Distinguish between supervised and unsupervised learning. (7)	BTL4	Analyzing
11	What is data visualization and why is data visualization important? (13)	BTL4	Analyzing
12	(i) Give a short note of Data Analysis . (7) (ii) Discuss the importance of data analysis. (6)	BTL4	Analyzing
13	Explain in detail about various types of data visualization techniques with example (13)	BTL5	Evaluating
14	(i)Describe Naïve Bayes Classification. (7) (ii)Discuss about Variance and Standard deviance (6)	BTL6	Creating

PART C

1.	.Explain the Central tendencies and distributions concepts and find the varience of first seven Natural Numbers?	BTL6	Creating
2.	Create a Regression Model for " happy people get many hours of sleep" using your own data and what kind of inferences it provides. (15)	BTL6	Creating
3.	Describe in detail, how Machine Learning Algorithms can be tested with necessary examples. (15)	BTL5	Evaluating
4.	Describe Learning Process & Explain Supervised Learning in detail, Give Necessary examples. (15)	BTL5	Evaluating

UNIT III - PYTHAN

Introduction to python- Data Types-Operators-input and output statements-flow control-string data type-list,

tuble, set data structure-Dictionary data structure-functions-modules-exception handling-file handling-python data base programme

PART – A

Q.No	Questions	BT Level	Competence
1.	Distinguish between string and list.	BTL1	Remembering
2.	List the various control flow structures	BTL1	Remembering
3.	List the standard data types in python.	BTL1	Remembering
4.	List the types of operators available in python	BTL1	Remembering
5.	Define Python list. How lists differ from Tuples.	BTL1	Remembering
6.	What are the list operations?	BTL1	Remembering
7.	Define Python Tuple.	BTL2	Understanding
8.	What are the different ways to create a list?	BTL2	Understanding
9.	What are the built-in functions that are used in Tuple?	BTL2	Understanding
10.	Point out the methods used in Tuples	BTL2	Understanding
11.	Define dictionary	BTL4	Analyzing
12.	Examine different set functions.	BTL3	Applying
13.	Write a few methods that are used in Python Lists	BTL3	Applying
14.	List the dictionary operations	BTL3	Applying
15.	Distinguish between files and modules	BTL4	Analyzing
16.	Examine the need for exceptions using an example	BTL4	Analyzing
17.	Give the features of python.	BTL5	Evaluating
18.	Difference between built-in exceptions and handling exception	BTL5	Evaluating
19.	Point out different modes of file opening.	BTL6	Creating
20.	Discover the format operator available in files.	BTL6	Creating

PART – B

1	Write short notes on types of operators in python with appropriate example (13)	BTL1	Remembering
2	List the different operators in python and estimate the precedence of execution. (13)	BTL1	Remembering
3	Describe the methods and operations of Dictionaries. (13)	BTL1	Remembering
4	Analyze the need for functions and explain with an example. (13)	BTL1	Remembering
5	What is a Python Tuple? What are the advantages of Tuple over list? (5) (ii)“Tuples are immutable”. Explain with example. (8)	BTL2	Understanding
6	Recollect the various dictionary operations and explain them with an example . (13)	BTL2	Understanding
7	Describe the following (i) Creating the list. (4) (ii) Accessing values in the lists. (3) (iii) Updating the list. (3) (iv) Deleting the list elements. (3)	BTL2	Understanding
8	(i)What are the accessing elements in a Tuple? Explain with suitable programs.(7) (ii) Explain how to return more than one value from a function with the help of a program. (6)	BTL3	Applying

9	(i) Explain with example of writing a file. (7) (ii) Discover syntax for reading from a file. (6)	BTL3	Applying
10	Discuss the different modes for opening a file and closing a file. (13)	BTL4	Analyzing
11	Describe in detail exception handling with sample program. (13)	BTL4	Analyzing
12	Explain with an example to copy the contents of one file to another. (13)	BTL4	Analyzing
13	i) Describe the use of try block and except block in python with syntax. (8) (ii) Describe with an example exceptions with arguments in python. (5)	BTL5	Evaluating
14	What is Dictionary? Explain Python dictionaries in detail discussing its operations and methods.. (13)	BTL6	Applying
PART C			
1.	Write a program to delete all the duplicate elements in a list. (15)	BTL6	Creating
2.	Evaluate the different operations like union, intersection and differences of list using Set and functions. (15)	BTL6	Creating
3.	Using a dictionary variable, estimate the frequency of each character in a sentence obtained from the user. (15)	BTL5	Evaluating
4.	Write a function that reads a file file1 and evaluates and displays number of words and vowels in the file. (15)	BTL5	Evaluating
UNIT IV - INTRODUCTION TO R PROGRAMMING			
What is R? –RStudio Overview - Arithmetic Operators - Logical Operations -Using Functions - Creating Variables - Numeric, Character and Logical Data - Vectors - Data Frames - Factors - Sorting Numeric, Character, and Factor Vectors - Special Values			
PART – A			
Q.No	**Questions**	**BT Level**	**Competence**
1.	What is a vector in R? Explain operations on vectors	BTL1	Remembering
2.	What are the Data Types in R?	BTL4	Analyzing
3.	List the differences between vector and list	BTL1	Remembering
4.	State the Arithmetic Operators in R	BTL1	Remembering
5.	Write any 3 math functions in R.	BTL1	Remembering
6.	Mention the various R Logical Operators	BTL1	Remembering
7.	Write syntax of if else in R.	BTL2	Understanding
8.	What is R?	BTL3	Applying
9.	Differentiate between vector, List, Matrix, and Data frame	BTL2	Understanding
10.	Explain initialize() function in R?	BTL2	Understanding
11.	Differentiate R while comparing with other programming languages?	BTL3	Applying
12.	Give any five features of R.	BTL2	Applying
13.	How to compare the two datasets in R	BTL5	Evaluating
14.	Differentiate between R and Python in terms of functionality?	BTL4	Analyzing
15.	What are the applications of R?	BTL4	Analyzing
16.	Write the purpose of cor() function in R	BTL3	Understanding
17.	Why do we use apply() function in R?	BTL5	Evaluating
18.	Differentiate b/w lapply and sapply.	BTL5	Evaluating
19.	Create a simple matrix with 3X3 size in R	BTL6	Creating

20.	Explain dnorm() function	BTL6	Creating
	PART – B		
1	Explain in detail about dataframe and arrays with example R code. (13)	BTL1	Remembering
2	Explain the procedure and concepts for reading data in R (13)	BTL1	Remembering
3	(i)Discuss the Data frames utilization in R (7) (ii) Describe the List function in R (6)	BTL1	Remembering
4	Demonstrate about (i) Input-Output functions (7) (ii) Assignment statement (6)	BTL1	Remembering
5	(i)Illustrate in detail the data types of R. (10) (ii) Define the window panes of RStudio GUI. (3)	BTL2	Understanding
6	Discuss the Data frames utilization in R with example (13)	BTL2	Understanding
7	Elaborate the following R objects. (i) vector (4) (ii)data frame (3) (iii)matrix (3) (iv)list (3)	BTL2	Understanding
8	Describe in detail about various logical operators in R programming. (13)	BTL5	Evaluating
9	Give example and define the following function of R (i) read() (3) (ii) head() (3) (iii) summary() (3) (iv) plot() (2) (v) summary() (2)	BTL3	Applying
10	List the inbuilt summary functions to apply on vectors. Create a vector and apply all functions on it. (13)	BTL4	Analyzing
11	What is a vector in R? Explain operations on vectors. (13)	BTL4	Analyzing
12	(i)What are the different ways to read the dataset? How to create and rename a variable in R? (7) (ii)What are the read write methods available in R and explain? (6)	BTL4	Analyzing
13	(i) Write about the following with example a)Mean b)Mode c)Median d)Cumulative Sum f)Cumulative Min (10) (ii) Write about data frame? Write about operations on data frame (3)	BTL6	Creating
14	Explain in details about factors in R Programming.	BTL3	Applying
	PART C		
1	What are the data frames? Write its significance in R-Language? (15)	BTL5	Evaluating
2	How to get system date in R? Generate sequence of previous and coming 10 dates from today in R. (15)	BTL6	Evaluating
3.	That is data frame and how to create a data frame using the following data: Height GPA (15) 66 3.80 62 3.78 63 3.88 70 3.72 74 3.69 Write an R program to prepare the inventory.	BTL5	Creating
4.	Create a data frame with a = c(1, 2, 3), b = c(4, 5, 6), c(7, 8, 9) and find the value of the following (i) How do I select the c (4, 5, 6)? (5) (ii) How do iselect the 1? (5) (iii) How do I select the 5? (5)	BTL6	Creating

UNIT V -DESCRIPTIVE STATISTICS IN R			
Statistical graphs-Statistical graphs-Iteration-Conditional statements-Data exploration and visualization-Data querying: SQL and R-Writing functions Reporting-Interactive reporting with Rmarkdown			
PART – A			
Q.No	**Questions**	**BT Level**	**Competence**
1.	Is R good for statistical analysis?Justify	BTL1	Remembering
2.	Mention the five descriptive statistics?	BTL1	Remembering
3.	Explain Pie chart in R.	BTL1	Remembering
4.	Give names of visualization packages.	BTL1	Remembering
5.	What are GGobi and iPlots?	BTL1	Remembering
6.	How do you use iteration in R?	BTL1	Remembering
7.	Differentiate SQL and No SQL databases.	BTL2	Understanding
8.	State the packages and function required to run SQL queries in R?	BTL2	Understanding
9.	How do I visualize data in R?	BTL5	Evaluating
10.	Differentiate between data exploration and data analysis?	BTL2	Understanding
11.	What is data exploration in R?	BTL3	Applying
12.	How do you write an IF condition in a For loop in R?	BTL3	Applying
13.	What is Graphical Data Analysis with R?	BTL3	Applying
14.	Why is R good for data visualization?	BTL4	Analyzing
15.	What is data exploration and visualization?	BTL4	Analyzing
16.	How do you write a function in R programming?	BTL4	Analyzing
17.	What is R Markdown used for?	BTL2	Understanding
18.	List out the function in R to calculates descriptive statistics?	BTL6	Evaluating
19.	What is an RMD file in R?	BTL5	Creating
20.	How do I connect SQL to R?	BTL6	Creating
1	Write the R script which include relevant packages and procedure to access .csv and .exl files. Elaborate with suitable example. (13)	BTL1	Remembering
2	Discuss about Descriptive Statistics in exploratory analysis (13)	BTL1	Remembering
3	Generalize the graphical analysis in data analysis? List the various plots in R and explain in detail. (13)	BTL2	Understanding
4	(i)Write about scatter plot and histograms with examples? Explain its importance? (7) (ii) How to plot multiple curves in same graph? Explain with example? (6)	BTL1	Remembering
5	List down at least 5 libraries in R that can be used for data Visualization. Explain three of them briefly. (13)	BTL2	Understanding
6	Describe the various plots in R to visualize the data and explain the purpose of each plot in detail. (13)	BTL1	Remembering
7	How to Output and Format A Table Of Data In R Markdown package (13)	BTL2	Understanding
8	Describe the types of flow control statements in R Programming? (13)	BTL4	Creating
9	(i) Explain in detail about Simple simple statistics functions in R Programming . (8)	BTL3	Applying

	(ii)Mention the steps to perform Exploratory Data Analysis (EDA) (5)		
10	What is Data Exploration? Explain about Data Exploration Tools? (13)	BTL4	Analyzing
11	Explain the process steps for data exploration and visualization with R? (13)	BTL6	Analyzing
12	Analyze the various R – Charts and Graphs with examples (13)	BTL4	Analyzing
13	Discuss about: (i) Data Visualization (6) (ii) Logical Operators. (7)	BTL5	Evaluating
14	Explain about the various statistical functions and math functions in R.(13)	BTL3	Applying
PART C			
1.	How to make multiple plots on to a single page layout in R Explain with Example. (15)	BTL5	Creating
2.	When will use a histogram and when will you use a bar chart in R ? Explain with an example by leveraging R package. (15)	BTL6	Creating
3.	How to Create Interactive Reports with R Markdown package. (15)	BTL5	Evaluating
4.	(i) How to write the own function in R programming with example. (8) (ii)What is the relation between data analysis and visualization? Explain.(7)	BTL6	Evaluating