

BALLARI INSTITUTE OF TECHNOLOGY & MANAGEMENT

(Autonomous Institute under Visvesvaraya Technological University, Belagavi)

USN

--	--	--	--	--	--	--	--	--	--

Course Code

2	1	M	B	A	D	A	4	0	1
---	---	---	---	---	---	---	---	---	---

Fourth Semester MBA Degree Examinations, Sept/Oct 2023

R PROGRAMMING FOR MANAGERS

Duration: 3 hrs

Max. Marks: 100

Note: 1. Answer any FOUR full questions from Question No. 1 to 7.

2. Question No. 8 is compulsory

3. Missing data, if any, may be suitably assumed

<u>Q. No</u>	<u>Question</u>	<u>Marks</u>	<u>(RBTL:CO:PO)</u>
1.	a. Apply the Hadoop Distributed File system features in the design of Bigdata.	03	(4 : 1 : 1)
	b. Apply the various system roles for the HDFS components.	07	(4 : 1 : 1)
	c. Analyze the working of the MapReduce model.	10	(3 : 1 : 1)
2.	a. Apply the various HDFS user commands in Hadoop Framework.	03	(4 : 1 : 1)
	b. Apply the safe mode and rack awareness concepts in the Hadoop Distributed File System.	07	(4 : 1 : 1)
	c. Analyze the Apache Hadoop Parallel MapReduce data flow for the sample input text shown below: see spot run run spot run see the cat	10	(3 : 1 : 1)
3.	a. Outline the sources and types of data for a data warehouse?	03	(2 : 2 : 2)
	b. Explain any two Business Intelligence (BI) applications for various sectors.	07	(2 : 2 : 2)
	c. Outline any 8 design considerations for a data warehouse.	10	(3 : 2 : 2)
4.	a. What is a confusion matrix? Explain.	03	(1 : 2 : 2)
	b. Explain with diagram CRISP-DM data mining cycle.	07	(2 : 2 : 2)
	c. List and summarize the various charts use for data visualization.	10	(2 : 2 : 2)
5.	a. What are the different forms of data types and how to test the data type in R? Give one example for each.	03	(2 : 3 : 3)
	b. Explain the following R objects. a) vector b) data frame c) matrix d) list	07	(2 : 3 : 3)
	c. Develop an R program to find the roots of a quadratic equation $a_2x^2 + a_1x + a_0 = 0$	10	(3 : 3 : 3)
6.	a. List and Outline the syntax for the following data manipulation functions: select() b) arrange() c) filter()	03	(2 : 4 : 4)

- b. Develop an R Program to demonstrate the filter () using the logical operators considering a sample student performance dataset . **07** (3 :4 :4)

A gender	A race/ethni...	A parental le...	A lunch	A test prepa...	# math score	# reading sc...
female	group B	bachelor's degree	standard	none	72	72
female	group C	some college	standard	completed	69	90
female	group B	master's degree	standard	none	90	95
male	group A	associate's degree	free/reduced	none	47	57
male	group C	some college	standard	none	76	78
female	group B	associate's degree	standard	none	71	83
female	group B	some college	standard	completed	88	95
male	group B	some college	free/reduced	none	48	43
male	group D	high school	free/reduced	completed	64	64
female	group B	high school	free/reduced	none	38	60
male	group C	associate's degree	standard	none	58	54
male	group D	associate's degree	standard	none	48	52
female	group B	high school	standard	none	65	81
male	group A	some college	standard	completed	78	72
female	group A	master's degree	standard	none	50	53

- c. Develop a R Program to demonstrate the use of **10** (3 :4 :4)
- slice () to select the rows, to select the minimum and maximum values.
 - mutate () to modify and to add new columns.
7. a. List and outline the syntax for any of three different ways in which data can be imported into the R program. **03** (2 :4 :4)
- b. Explain with syntax and sample code to read from the following csv file using read. csv () and read_csv (). Also sort the results by countries losing most matches. **07** (2 :4 :4)

	Y..Teams	Wins	Lose	Points
1	India	5	2	10
2	South Africa	3	4	6
3	West Indies	1	6	2
4	England	2	4	4
5	Australia	4	2	8
6	New Zealand	2	5	4

- c. Demonstrate atleast five charts available in R using appropriate sample code for the above-mentioned student dataset. (As in question 6 (b)) **10** (2 :5 :5)
8. **Case Study (Compulsory)**
- a. Illustrate on any of case studies to highlight the application of Big Data **10** (2 :1 :1)
- b. What are the key characteristics of Business Analytics with R. Extend the use of R in helping organizations in the complicated process of data analysis. **10** (2 :3 :3)
