

## Final Practical Exam

### Sample practical list for BDA SEM VII ( Computer Engineering, TSEC, Bandra)

- 1 HDFS commands ( for each practical 2-3 commands)
- 2 Use of Sqoop tool to transfer data between Hadoop and relational database servers.
- 3 Programming exercises in HBASE
- 4 Experiment on pig.
- 5 Create HIVE Database and Descriptive analytics(basic statistics).
- 6 Implement Bloom Filter using R Programming.
- 7 Implement FM algorithm using R Programming.

### Programming Using R

1. (A) HDFS Command  
  
(B) Write a R program to create a Data Frame which contain details of 5 employees and display summary of the data using R.
2. (A) HDFS Command  
  
(B) For anydataset visualize the following types of chart : [Scatterplot](#), [Bubble Chart](#), [Bar Chart](#) , [Dot Plots](#) ,[Histogram](#) ,[Box Plot](#) ,[Pie Chart](#)
- 3 (a) HDFS Command  
  
(B) Write the script in R to sort the values contained in the following vector in ascending order and descending order: (23, 45, 10, 34, 89, 20, 67, 99). Demonstrate the output using graph.
- 4 (A) HDFS Command

(B) The following table shows the number of units of different products sold on different days:

Product	Monday	Tuesday	Wednesday	Thursday	Friday
Bread	12	3	5	11	9
Milk	21	27	18	20	15
Cola Cans	10	1	33	6	12
Chocolate bars	6	7	4	13	12
Detergent	5	8	12	20	23

Create five sample numeric vectors from this data visualize data using R.

- 5 (a) HDFS Command  
  
(b) Consider the following data frame given below:

Subject	Class	Marks
1	1	56
2	2	75
3	1	48
4	2	69
5	1	84
6	2	53

(i) Create a subset of subject less than 4 by using subset () funcon and demonstrate the output.

(ii) Create a subject where the subject column is less than 3 and the class equals to 2 by using [] brackets and demonstrate the output using R

(iii) Visualize data

6 (a) HDFS Command

(b)The data analyst of Argon technology Mr. John needs to enter the salaries of 10 employees in R. The salaries of the employees are given in the following table:

Sr. No.	Name of employees	Salaries
1	Vivek	21000
2	Karan	55000
3	James	67000
4	Soham	50000
5	Renu	54000
6	Farah	40000
7	Hetal	30000
8	Mary	70000
9	Ganesh	20000
10	Krish	15000

i) Which R command will Mr. John use to enter these values demonstrate the output.

ii) Now Mr. John wants to add the salaries of 5 new employees in the existing table, which command he will use to join datasets with new values in R. Demonstrate the output.

(iii) Visualize the data using chart .

7 (a) HDFS Command

(b) Analyse and visualize churn modelling data using R.

8 (a) HDFS Command

(b) Analyse and visualize IRIS data using R.

9 (a) HDFS Command

(b) Analyse and visualize supermarket data using R.

10 (a) HDFS Command

(b) Analyse and visualize Loan data using R.

11. The following table shows the number of units of different products sold on different days

Product	Monday	Tuesday	Wednesday	Thursday	Friday
Bread	12	3	5	11	9
Milk	21	27	18	20	15
Cola Cans	10	1	33	6	12
Chocolate bars	6	7	4	13	12
Detergent	5	8	12	20	23

(i) Create five sample numeric vectors from this data.

(ii) Name and explain the operators used to form data subsets in R.

12. Which function is used to concatenate text values in R. Write a script to concatenate text and numerical values in R.

Text 1: Ram has scored

Text 2: 89

Text 3: marks

Text 4: in Mathematics

13. Consider the following data frame given below:

course	id	class	marks
1	11	1	56
2	12	2	75
3	13	1	48
4	14	2	69
5	15	1	84
6	16	2	53

i. Create a subset of course less than 3 by using [ ] brackets and demonstrate the output.

ii. Create a subset where the course column is less than 3 or the class equals to 2 by using subset () function and demonstrate the output.

i. Create a data frame from the following 4 vectors and demonstrate the output:

```
emp_id = c(1:5)
emp_name = c("Rick", "Dan", "Michelle", "Ryan", "Gary")
start_date = c("2012-01-01", "2013-09-23", "2014-11-15", "2014-05-11", "2015-03-27")
salary = c(60000, 45000, 75000, 84000, 20000)
```

- ii. Display structure and summary of the above data frame.
- iii. Extract the emp\_name and salary columns from the above data frame.
- iv. Extract the employee details whose salary is less than or equal to 60000.

15

- ii. Suppose you have two datasets A and B.  
Dataset A has the following data: 6 7 8 9.  
Dataset B has the following data: 1 2 4 5.  
Which function is used to combine the data from both datasets into dataset C.  
Demonstrate the function with the input values and write the output.

**Dr. Arti Deshpande**  
**TSEC, Bandra**