**Q.1] Perform following task on HDFS cluster using HDFS commands**

**1. Check Linux login details and list all the files available in the current directory**

**their size and date**

**2. List all files HDFS directory**

**3. Create and show new directory**

**4. Print the total size of directory**

**Answer -**

Run the following command on cmd

1. Start-all
2. Jps
3. Whoami
4. hadoop fs -ls <file:///C:\hadoop-3.3.0\share\hadoop\hdfs>
5. hadoop fs -mkdir <file:///C:/hadoop-3.3.0/share/hadoop/hdfs/newDR>
6. hadoop fs -ls file:///C:/hadoop-3.3.0/share\hadoop/hdfs

**Q.2] Write a program to demonstrate Map Reduce(any operation).**

1. Start-all
2. Jps
3. hdfs dfs -mkdir /input
4. hdfs dfs -put "C:\hadoop-3.3.0\data.txt" /input
5. hadoop jar "C:\hadoop-3.3.0\share\hadoop\mapreduce\hadoop-mapreduce-examples-3.3.0.jar" wordcount /input /output
6. hdfs dfs -cat /output/part-r-00000

**Q.3] Create a Hive Managed tables, employee and dept under database college**

CREATE DATABASE IF NOT EXISTS college;

USE college;

CREATE TABLE employee (

empcode INT,

empfname STRING,

emplname STRING,

job STRING,

manager STRING,

hiredate STRING,

salary INT,

commission INT

)

STORED AS TEXTFILE;

INSERT INTO TABLE employee

VALUES

(1, 'John', 'Doe', 'Manager', 'Alice', '2020-01-01', 60000, 5000),

(2, 'Jane', 'Smith', 'Engineer', 'John', '2021-03-10', 45000, 3000),

(3, 'Mike', 'Johnson', 'Analyst', 'John', '2019-07-05', 40000, 2500);

CREATE TABLE dept (

deptcode INT,

deptname STRING,

location STRING

)

STORED AS TEXTFILE;

INSERT INTO TABLE dept

VALUES

(10, 'HR', 'New York'),

(20, 'Engineering', 'San Francisco'),

(30, 'Sales', 'Chicago');

SHOW TABLES;

SELECT \* FROM employee;

SELECT \* FROM dept;

**Q.4] Write a script to load data of employee\_data.txt to Pig ( empcode INT, empfname**

**STRING, emplname STRING, job STRING, manager STRING, hiredate**

**STRING, salary INT, commission INT)**

1. create data.txt file in pig -> bin
2. hdfs dfs -mkdir /pig
3. hdfs dfs -put “C:/path\_to\_file\_in\_bin” /pig
4. hdfs dfs -cat hdfs://localhost:9000/pig/file\_name.txt
5. pig -x local
6. employee = LOAD 'hdfs://localhost:9000/pig/employee.txt' USING PigStorage(',') as (empcode:int,fname:chararray,lname:chararray,job:chararray,manager:chararray,hiredate:chararray,salary:int,commission:int);
7. dump employee;
8. ascending=order employee by fname asc;
9. dump ascending
10. split employee into x if empcode> 5, y if empcode > 2;
11. dump x
12. dump y

**Q.5] Perform following task on HDFS cluster using HDFS commands**

**a. Create new directory**

**b. Create file**

**c. Check the last modified time of current directory**

**d. HDFS Command to get help for any command on HDFS**

**e. Check the size of each file in directory of HDFS**

**f. Copy file from local system to Hadoop**

**g. Display the content of file in Hadoop file system**

**Answer-**

1. **start-all**
2. **jps**
3. **hdfs dfs -mkdir /new\_directory**
4. **hdfs dfs -ls /**
5. **hdfs dfs -put sample.txt /new\_directory/**
6. **hdfs dfs -ls /new\_directory**
7. **hdfs dfs -cat /new\_directory/sample.txt**
8. **hdfs dfs -ls /new\_directory**
9. **hdfs dfs -help ls**
10. **hdfs dfs -du -h /new\_directory**