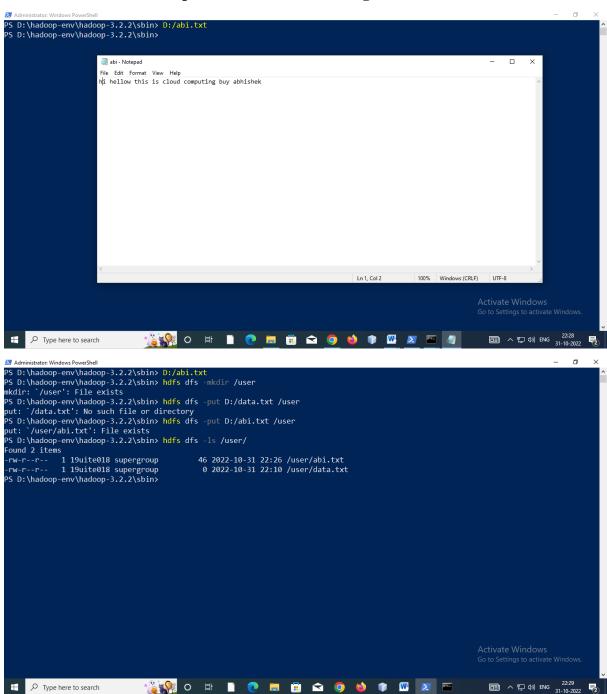Date : 15/10/22   Name : Abhishek S
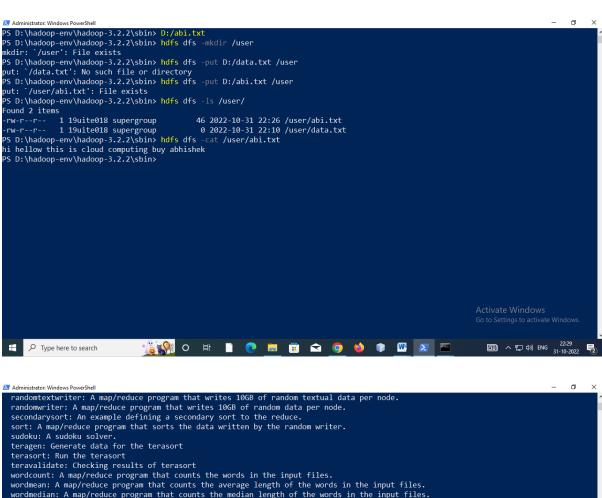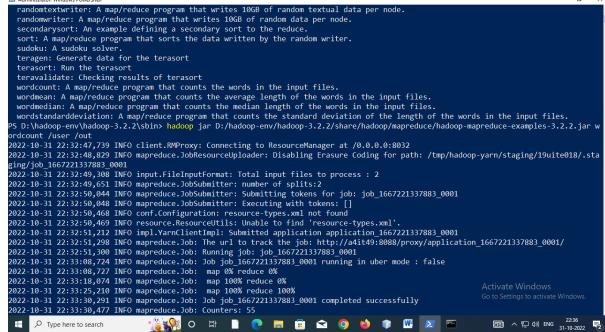
EX:10   Roll.No:19UITE018

                                                  Reg.No:920419205001

MapReduce (WordCount Program

```
PS D:\hadoop-env\hadoop-3.2.2\sbin> D:/abi.txt
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -mkdir /user
mkdir: `/user': File exists
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -put D:/data.txt /user
put: `/data.txt': No such file or directory
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -put D:/abi.txt /user
put: `/user/abi.txt': File exists
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -ls /user/
Found 2 items
-rw-r--r--   1 19uite018 supergroup         46 2022-10-31 22:26 /user/abi.txt
-rw-r--r--   1 19uite018 supergroup          0 2022-10-31 22:10 /user/data.txt
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -cat /user/abi.txt
hi hellow this is cloud computing buy abhishek
PS D:\hadoop-env\hadoop-3.2.2\sbin>
```

```
  randomtextwriter: A map/reduce program that writes 10GB of random textual data per node.
  randomwriter: A map/reduce program that writes 10GB of random data per node.
  secondarysort: An example defining a secondary sort to the reduce.
  sort: A map/reduce program that sorts the data written by the random writer.
  sudoku: A sudoku solver.
  teragen: Generate data for the terasort
  terasort: Run the terasort
  teravalidate: Checking results of terasort
  wordcount: A map/reduce program that counts the words in the input files.
  wordmean: A map/reduce program that counts the average length of the words in the input files.
  wordmedian: A map/reduce program that counts the median length of the words in the input files.
  wordstandarddeviation: A map/reduce program that counts the standard deviation of the length of the words in the input files.
PS D:\hadoop-env\hadoop-3.2.2\sbin> hadoop jar D:/hadoop-env/hadoop-3.2.2/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.2.jar wordcount /user /out
2022-10-31 22:32:47,739 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
2022-10-31 22:32:48,829 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/19uite018/.staging/job_1667221337883_0001
2022-10-31 22:32:49,308 INFO input.FileInputFormat: Total input files to process : 2
2022-10-31 22:32:49,651 INFO mapreduce.JobSubmitter: number of splits:2
2022-10-31 22:32:50,044 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1667221337883_0001
2022-10-31 22:32:50,048 INFO mapreduce.JobSubmitter: Executing with tokens: []
2022-10-31 22:32:50,468 INFO conf.Configuration: resource-types.xml not found
2022-10-31 22:32:50,469 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2022-10-31 22:32:51,212 INFO impl.YarnClientImpl: Submitted application application_1667221337883_0001
2022-10-31 22:32:51,298 INFO mapreduce.Job: The url to track the job: http://a4it49:8088/proxy/application_1667221337883_0001/
2022-10-31 22:32:51,300 INFO mapreduce.Job: Running job: job_1667221337883_0001
2022-10-31 22:33:08,724 INFO mapreduce.Job: Job job_1667221337883_0001 running in uber mode : false
2022-10-31 22:33:08,727 INFO mapreduce.Job:  map 0% reduce 0%
2022-10-31 22:33:18,074 INFO mapreduce.Job:  map 100% reduce 0%
2022-10-31 22:33:25,210 INFO mapreduce.Job:  map 100% reduce 100%
2022-10-31 22:33:30,291 INFO mapreduce.Job: Job job_1667221337883_0001 completed successfully
2022-10-31 22:33:30,477 INFO mapreduce.Job: Counters: 55
```

OUTPUT:

Screenshot 1 — Browser tabs:
You have successfully logged | Namenode information | DataNode Information | Mail - ABHISHEK.S - Outlook | Namenode information | +

localhost:9870/dfshealth.html#tab-overview

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities ▾

Browse the file system
Logs
Log Level
Metrics
Configuration
Process Thread Dump

Overview 'localhost:9820' (active)

Started:	Mon Oct 31 18:28:12 +0530 2022
Version:	3.2.2, r7a3bc90b05f257c8ace2f76d74264906f0f7a932
Compiled:	Sun Jan 03 14:56:00 +0530 2021 by hexiaoqiao from branch-3.2.2
Cluster ID:	CID-8e3bdf0a-913e-4e80-80b0-8c6fed996391
Block Pool ID:	BP-1203456125-192.168.56.1-1667220717770

Summary

Security is off.

Safemode is off.

18 files and directories, 5 blocks (5 replicated blocks, 0 erasure coded block groups) = 23 total filesystem object(s).

Heap Memory used 140.78 MB of 222 MB Heap Memory. Max Heap Memory is 889 MB.

Non Heap Memory used 66.87 MB of 68.59 MB Commited Non Heap Memory. Max Non Heap Memory is <unbounded>.

localhost:9870/explorer.html

Screenshot 2 — Browser tabs:
You have successfully logged | Namenode information | DataNode Information | Mail - ABHISHEK.S - Outlook | Browsing HDFS | +

localhost:9870/explorer.html#/

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities ▾

Browse Directory

| / | | Go! |

Show 25 entries   Search:

☐	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
☐	drwxr-xr-x	19uite018	supergroup	0 B	Oct 31 22:33	0	0 B	out	🗑
☐	drwx------	19uite018	supergroup	0 B	Oct 31 22:32	0	0 B	tmp	🗑
☐	drwxr-xr-x	19uite018	supergroup	0 B	Oct 31 22:26	0	0 B	user	🗑

Showing 1 to 3 of 3 entries

Previous 1 Next

Hadoop, 2021.

Hadoop   Overview   Datanodes   Datanode Volume Failures   Snapshot   Startup Progress   Utilities

Browse Directory

/out   Go!

Show 25 entries   Search:

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
	-rw-r--r--	19uite018	supergroup	0 B	Oct 31 22:33	1	128 MB	_SUCCESS	🗑
	-rw-r--r--	19uite018	supergroup	63 B	Oct 31 22:33	1	128 MB	part-r-00000	🗑

Showing 1 to 2 of 2 entries   Previous 1 Next

Hadoop, 2021.

Hadoop   Overview   Datanodes   Datanode Volume Failures   Snapshot   Startup Progress   Utilities

Browse Directory

File information - part-r-00000   ×

Download   Head the file (first 32K)   Tail the file (last 32K)

Block information — Block 0

Block ID: 1073741832
Block Pool ID: BP-1203456125-192.168.56.1-1667220717770
Generation Stamp: 1008
Size: 63
Availability:
- a4it49.KC.VNR

Close

/out   Go!

Show 25 entries   Search:

	Permission	Owner			Block Size	Name	
	-rw-r--r--	19uite01			MB	_SUCCESS	🗑
	-rw-r--r--	19uite01			MB	part-r-00000	🗑

Showing 1 to 2 of 2 entries   Previous 1 Next

Hadoop, 2021.