

Big Data Analytics Lab PMDS507P

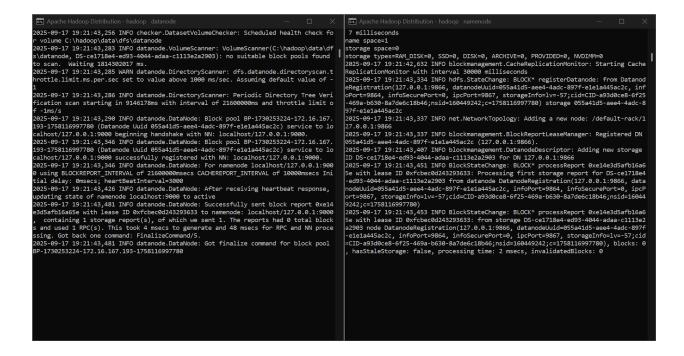
Name: **Tufan Kundu** Registration number: **24MDT0184**

Slot: L29+L30

Digital Assignment 2

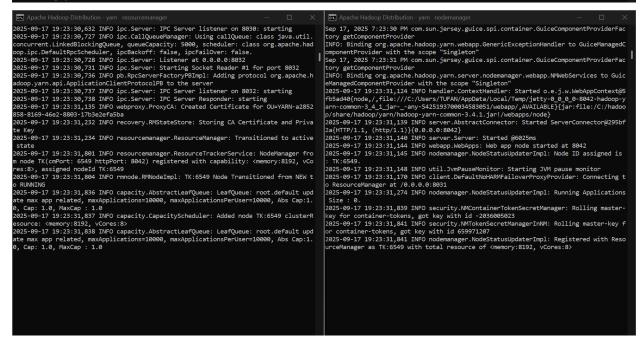
Step 1: Go to Command Prompt and type (start-dfs.cmd)





Step 2: Now type (start-yarn.cmd)

C:\Windows\System32>start-yarn.cmd
starting yarn daemons



Step 3: Now type jps, which will display a list of all currently running Java processes on the local system.

```
© Administrator: Command Prompt

C:\Windows\System32>jps
7712 ResourceManager
18356 Jps
12360 DataNode
6172 NodeManager
9596 NameNode
```

Step 4: Stop all the processes by typing (stop-dfs.cmd) and (<u>stop-yarn.</u>cmd)

```
C:\Windows\System32>stop-dfs.cmd
SUCCESS: Sent termination signal to the process with PID 20368.
SUCCESS: Sent termination signal to the process with PID 15064.

C:\Windows\System32>stop-yarn.cmd
stopping yarn daemons
SUCCESS: Sent termination signal to the process with PID 19288.
SUCCESS: Sent termination signal to the process with PID 19596.

INFO: No tasks running with the specified criteria.
```

Step 5: Now try to open all the processes by a single command (start-all.cmd) and check if all the processes are running by the command jps



```
© Administrator: Command Prompt

C:\Windows\System32>jps

19392 ResourceManager

17988 NodeManager

11560 NameNode

12888 Jps

5576 DataNode
```

Step 6: List out all the root directories by the command (hdfs dfs -ls /)

```
□ Administrator:Command Prompt

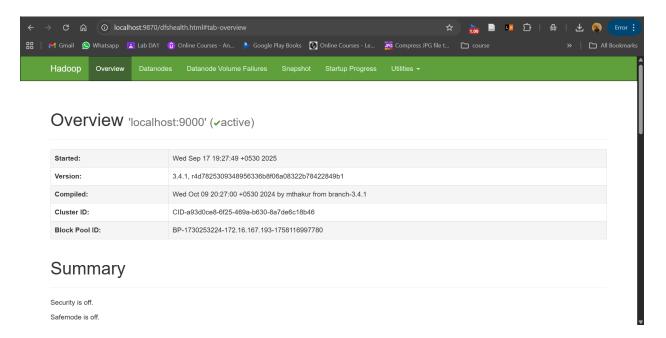
C:\Windows\System32>hdfs dfs -ls /

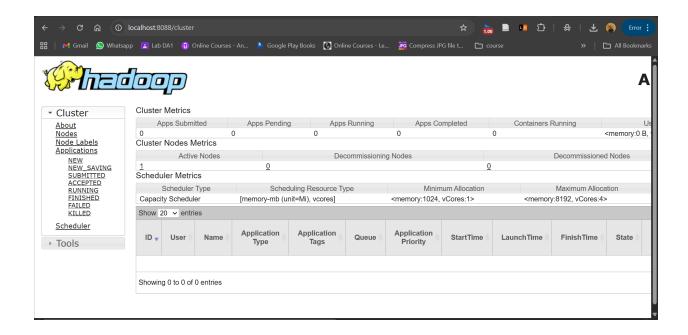
Found 1 items

drwxr-xr-x - TK supergroup

0 2025-09-17 19:30 /filename
```

Step 7: Explore localhost:9870/ and localhost:8088/





Create Directory:

Step 8: Now create a new directory by command (hdfs dfs –mkdir /filename) and view the list of directories again

Create multiple directories at once:

Step 9: We can create multiple directories at once by the command (hdfs dfs -mkdir -p/path1/path2)

```
C:\Windows\System32>hdfs dfs -mkdir -p /path1/path2

C:\Windows\System32>hdfs dfs -ls /
Found 3 items
drwxr-xr-x - TK supergroup 0 2025-09-17 19:30 /filename
drwxr-xr-x - TK supergroup 0 2025-09-17 19:35 /path1
drwxr-xr-x - TK supergroup 0 2025-09-17 19:34 /tufan
```

Put (upload) a file to HDFS:

Step 10: Now create a text file and upload it into the new directory by command (hdfs dfs -put data.txt /myfolder)

```
C:\Windows\System32>hdfs dfs -put data.txt /myfolder

C:\Windows\System32>hdfs dfs -get /myfolder/data.txt
get: `data.txt': File exists
```

Get(download) a file for HDFS:

Step 11: To download a file from hdfs we can use the command (hdfs dfs -get /myfolder/data.txt)

```
C:\Windows\System32>hdfs dfs -get /myfolder/data.txt
get: `data.txt': File exists
```

Remove a file:

Step 12: To remove a file from any directory, we use the command (hdfs dfs -rm /myfolder/data.txt)

```
C:\Windows\System32>hdfs dfs -rm /myfolder/data.txt
Deleted /myfolder/data.txt
```

Remove a file recursively:

Step 13: To delete a directory, we can use (hdfs dfs -rm -r /myfolder) this will delete all the files/folders permanently

```
C:\Windows\System32>hdfs dfs -rm -r /myfolder
Deleted /myfolder
```

Step 14: For viewing file contents (hdfs dfs -cat /myfolder/data.txt)

```
C:\Windows\System32>hdfs dfs -cat /myfolder/data.txt
My name is Tufan Kundu.
I am pursuing MSc in Data Science at VIT.
I have experience in C, C++, Python, and Java.
My interests include Deep Learning, Machine Learning, and Big Data Analytics.
I enjoy working on projects that involve data-driven problem solving.
```

Step 15: To display the first few lines, we use (hdfs dfs -head /myfolder/data.txt)

```
C:\Windows\System32>hdfs dfs -head /myfolder/data.txt
My name is Tufan Kundu.
I am pursuing MSc in Data Science at VIT.
I have experience in C, C++, Python, and Java.
My interests include Deep Learning, Machine Learning, and Big Data Analytics.
I enjoy working on projects that involve data-driven problem solving.
```

Step 16: To display the last few lines, we use (hdfs dfs -tail /myfolder/data.txt)

```
C:\Windows\System32>hdfs dfs -tail /myfolder/data.txt

My name is Tufan Kundu.

I am pursuing MSc in Data Science at VIT.

I have experience in C, C++, Python, and Java.

My interests include Deep Learning, Machine Learning, and Big Data Analytics.

I enjoy working on projects that involve data-driven problem solving.
```

Step 17: For copying files within hdfs (hdfs dfs -cp /myfolder/data.txt /backup)

C:\Windows\System32>hdfs dfs -cp /myfolder/data.txt /backup

Step 18: For checking the disk usage (hdfs dfs -du -h /)

Step 19: To show the file checksum (hdfs dfs -checksum /myfolder/data.txt)