

Big Data Analytics Lab

PMDS507P

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

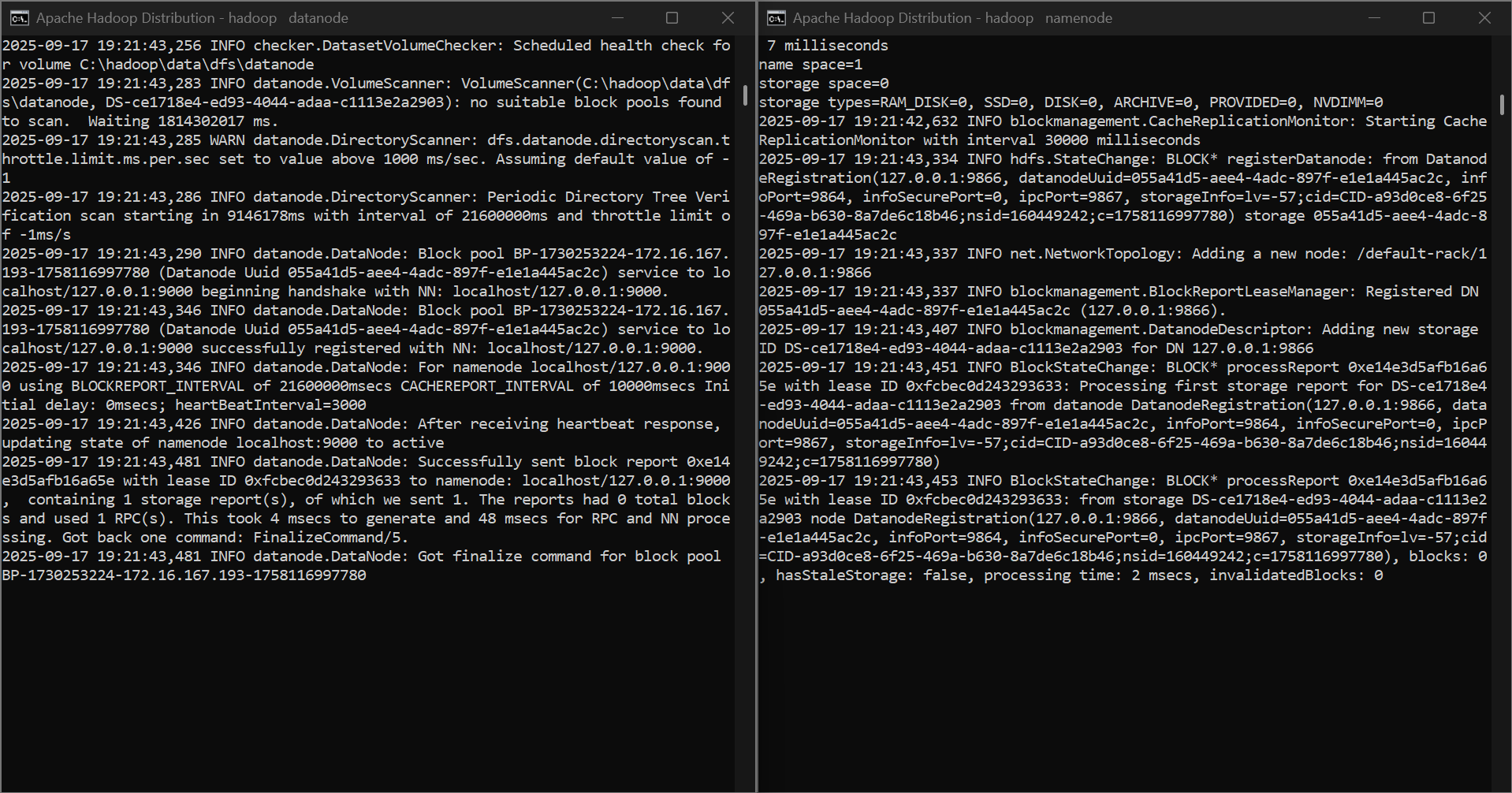
Slot: L29+L30

Digital Assignment 2

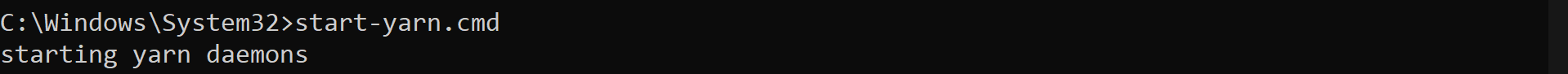
# 

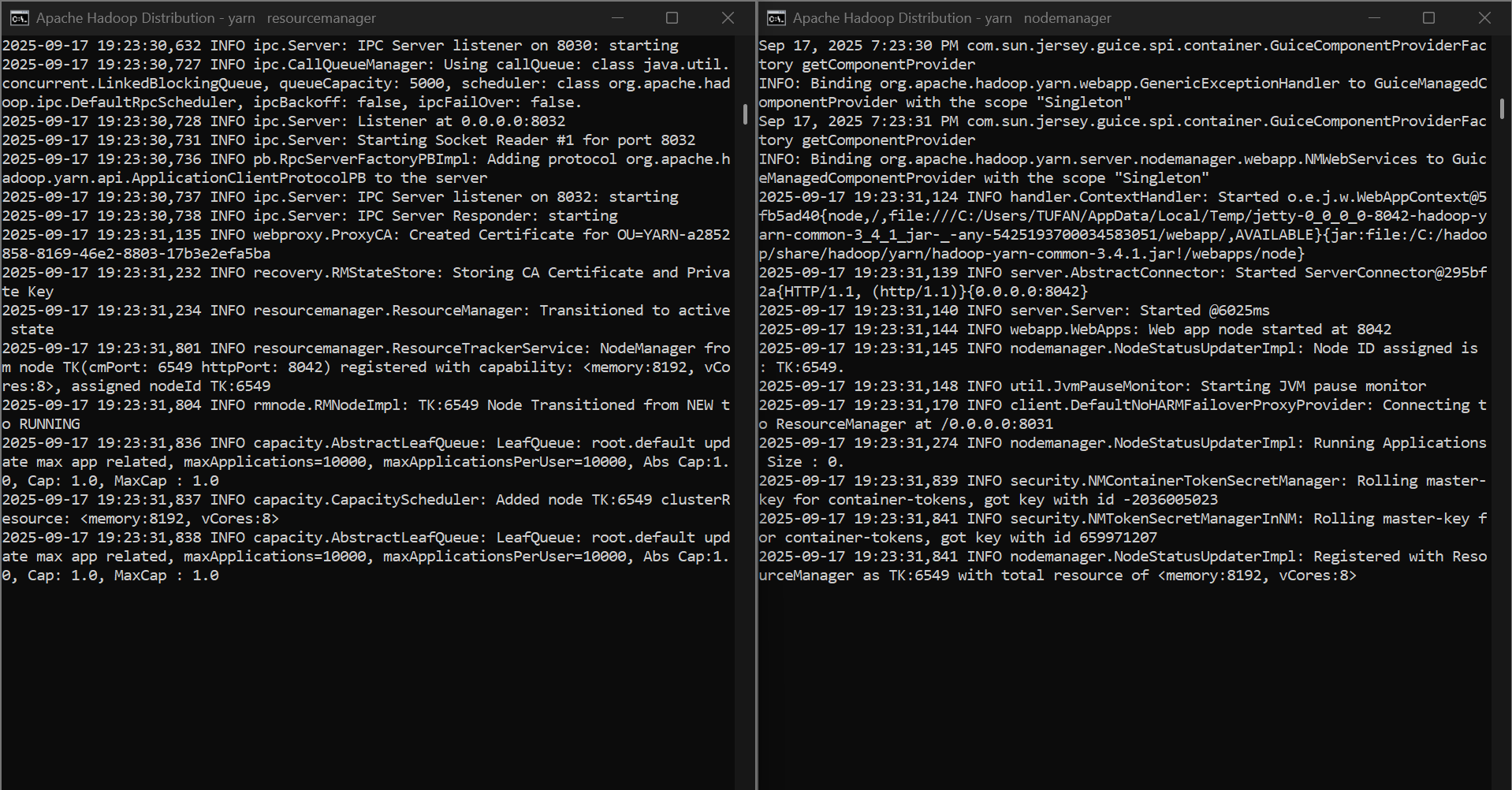
**Step 1**: Go to Command Prompt and type ([start-dfs.](http://start-dfs.sh)cmd)





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

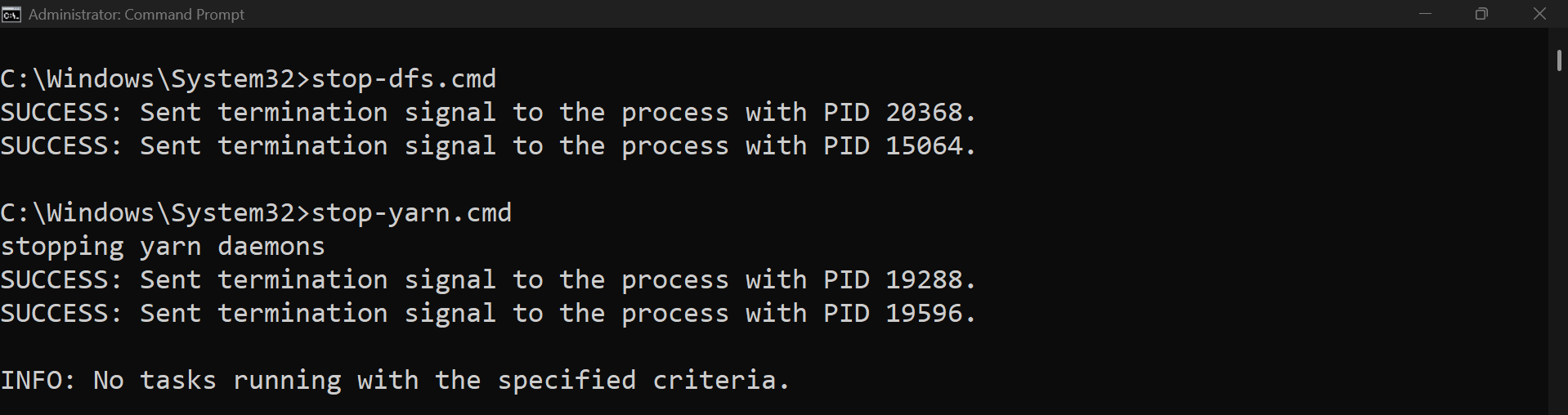




**Step 3:** Now type jps, which will display a list of all currently running Java processes on the local system.

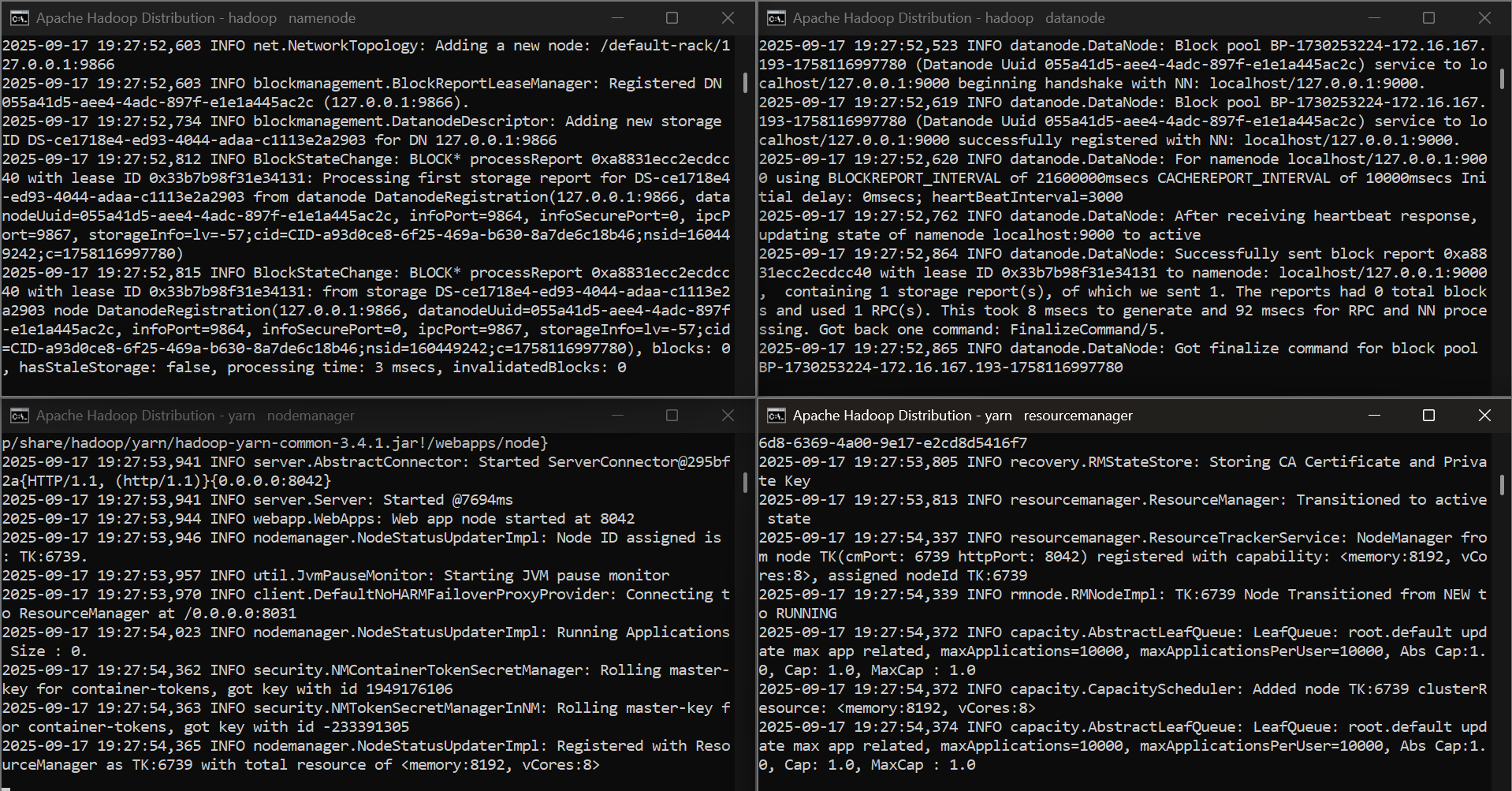


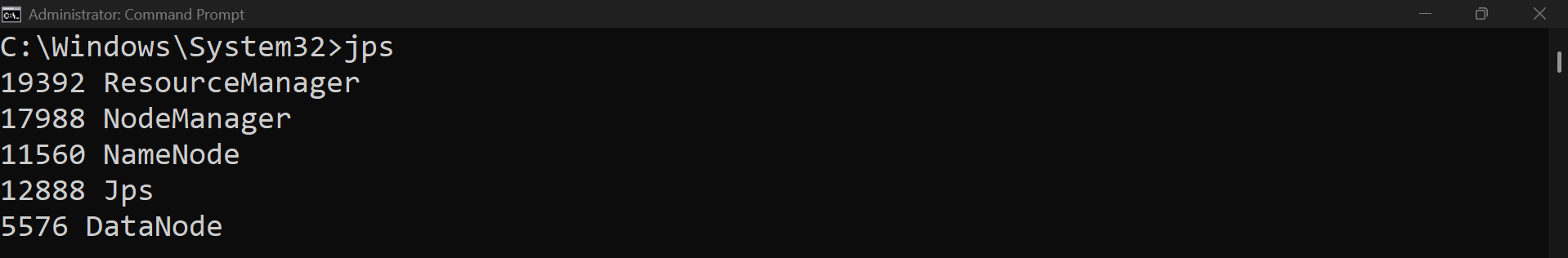
**Step 4:** Stop all the processes by typing (stop-dfs.cmd) and ([stop-yarn.](http://stop-yarn.sh)cmd)



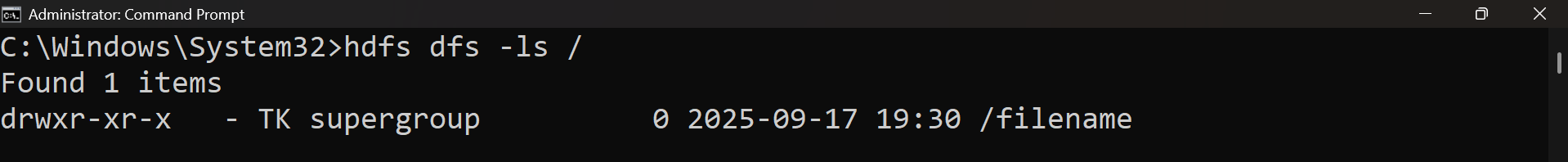
**Step 5:** Now try to open all the processes by a single command (start-all.cmd) and check if all the processes are running by the command jps



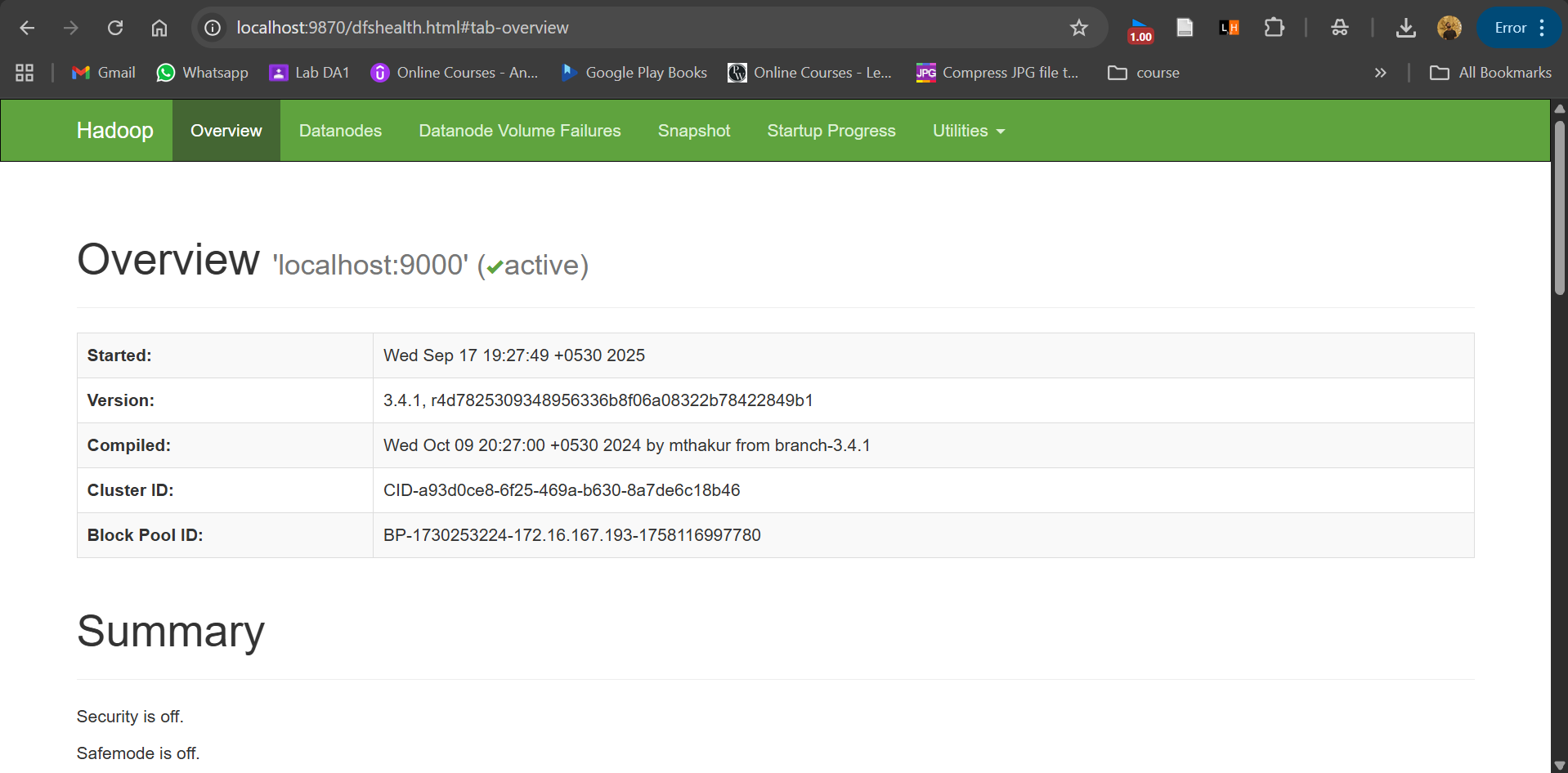


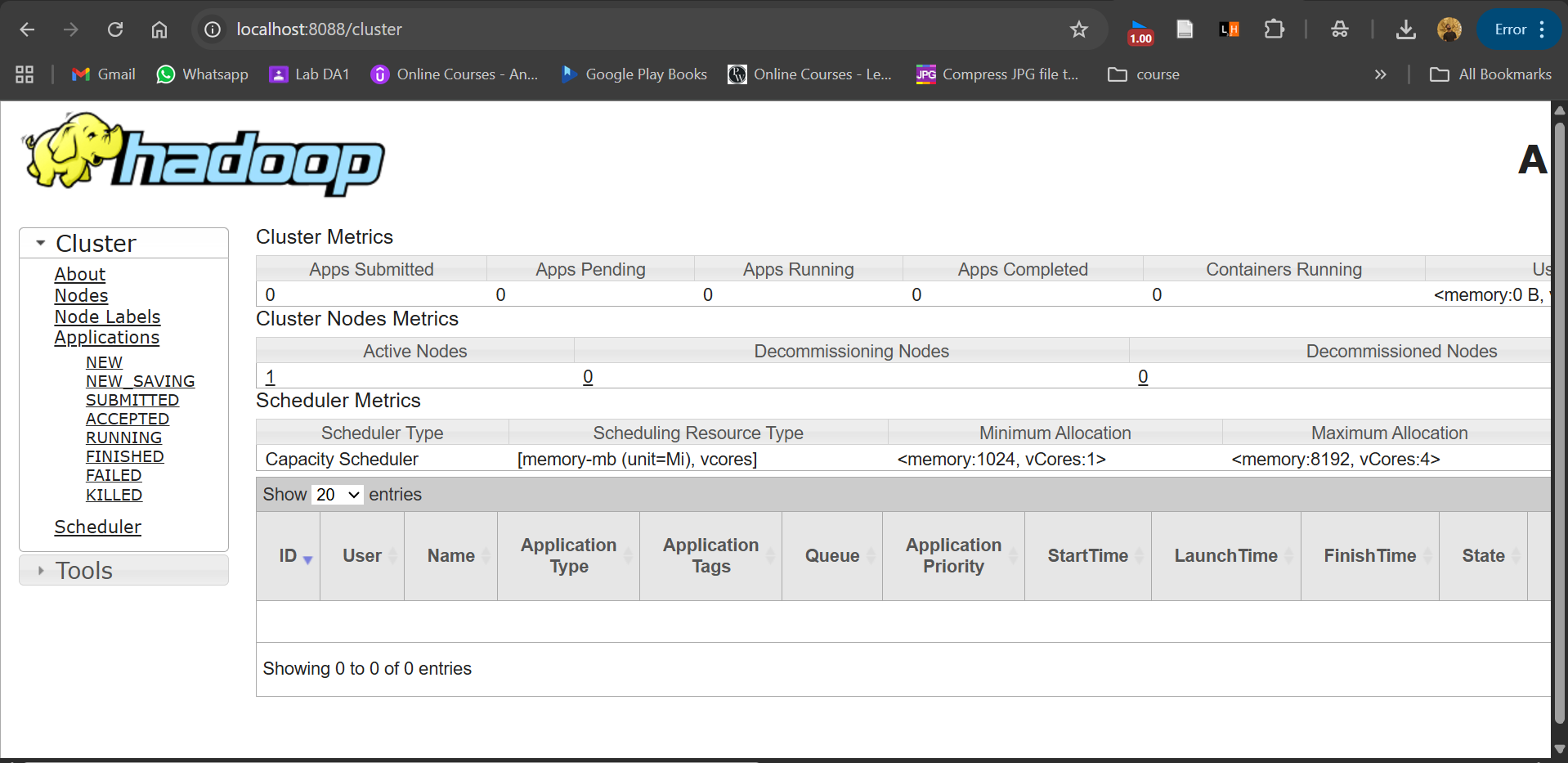


**Step 6:** List out all the root directories by the command (hdfs dfs -ls /)



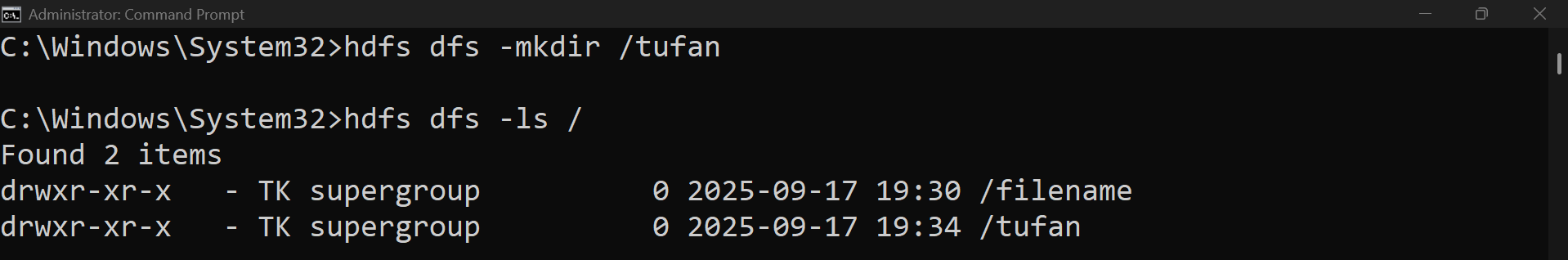
**Step 7:** Explore localhost:9870/ and localhost:8088/





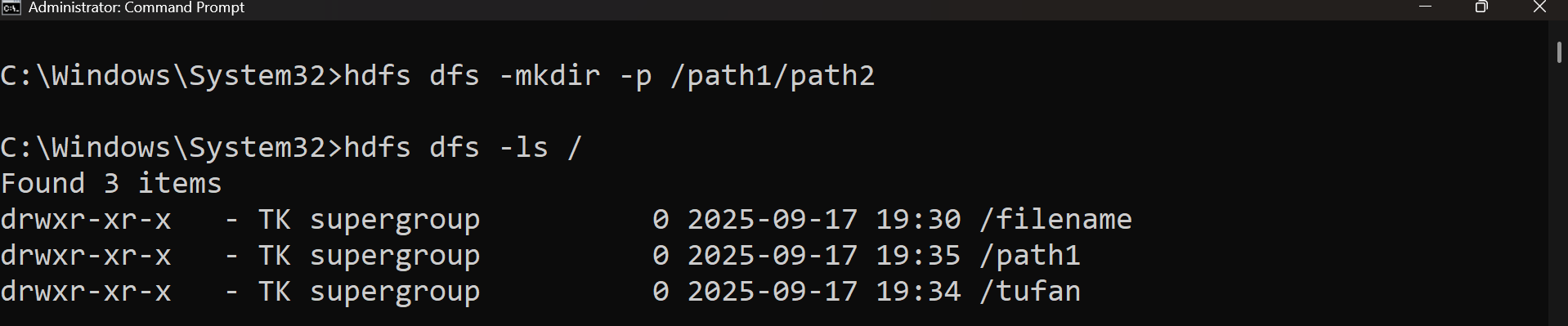
# Create Directory:

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



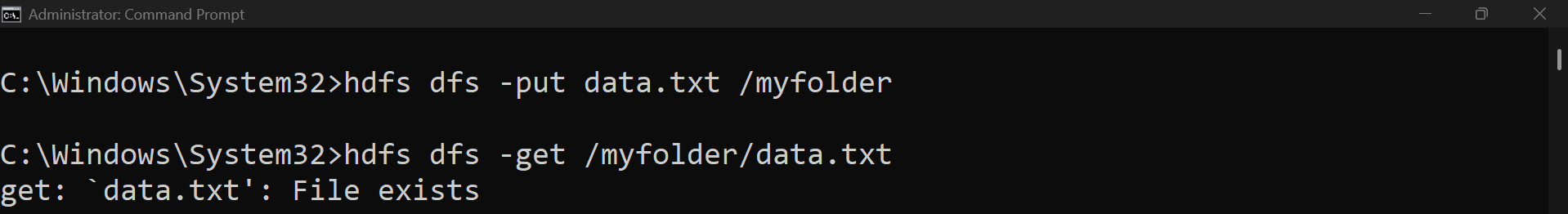
# Create multiple directories at once:

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



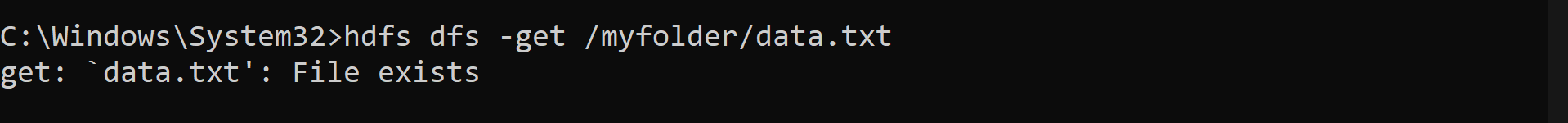
# Put (upload) a file to HDFS:

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



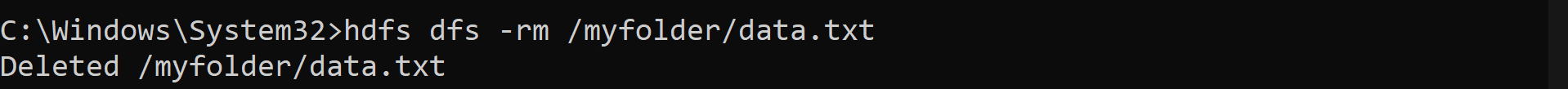
# Get(download) a file for HDFS:

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



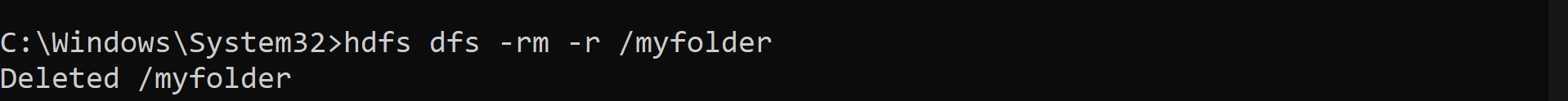
# Remove a file:

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

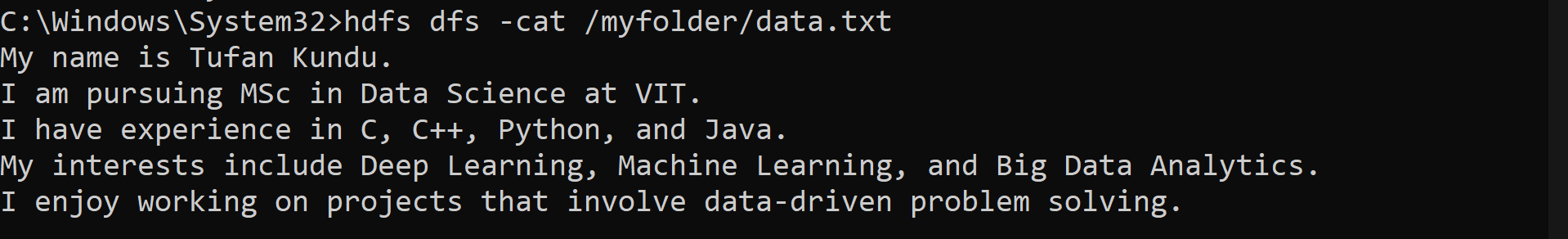


# Remove a file recursively:

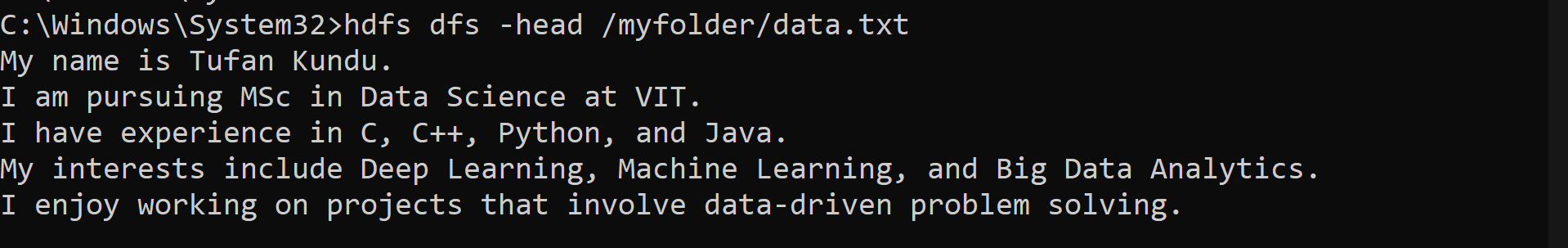
**Step 13:** To delete a directory, we can use (hdfs dfs -rm -r /myfolder) this will delete all the files/folders permanently



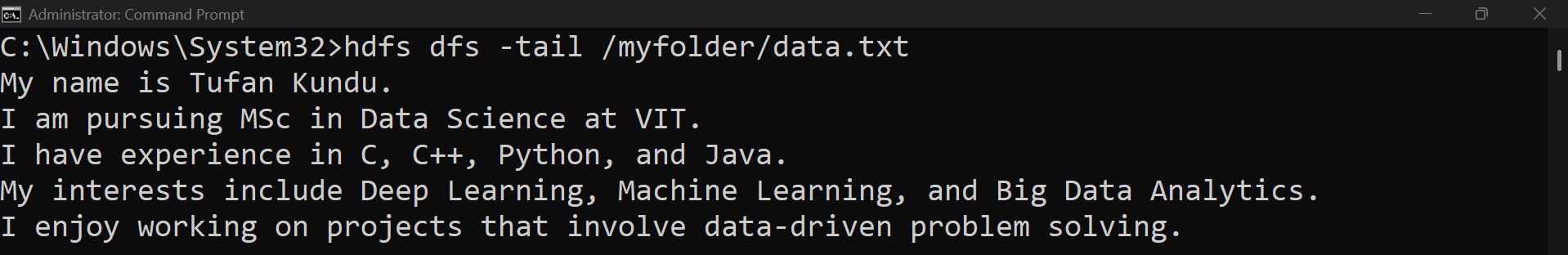
**Step 14:** For viewing file contents (hdfs dfs -cat /myfolder/data.txt)



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



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



**Step 17:** For copying files within hdfs (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)

