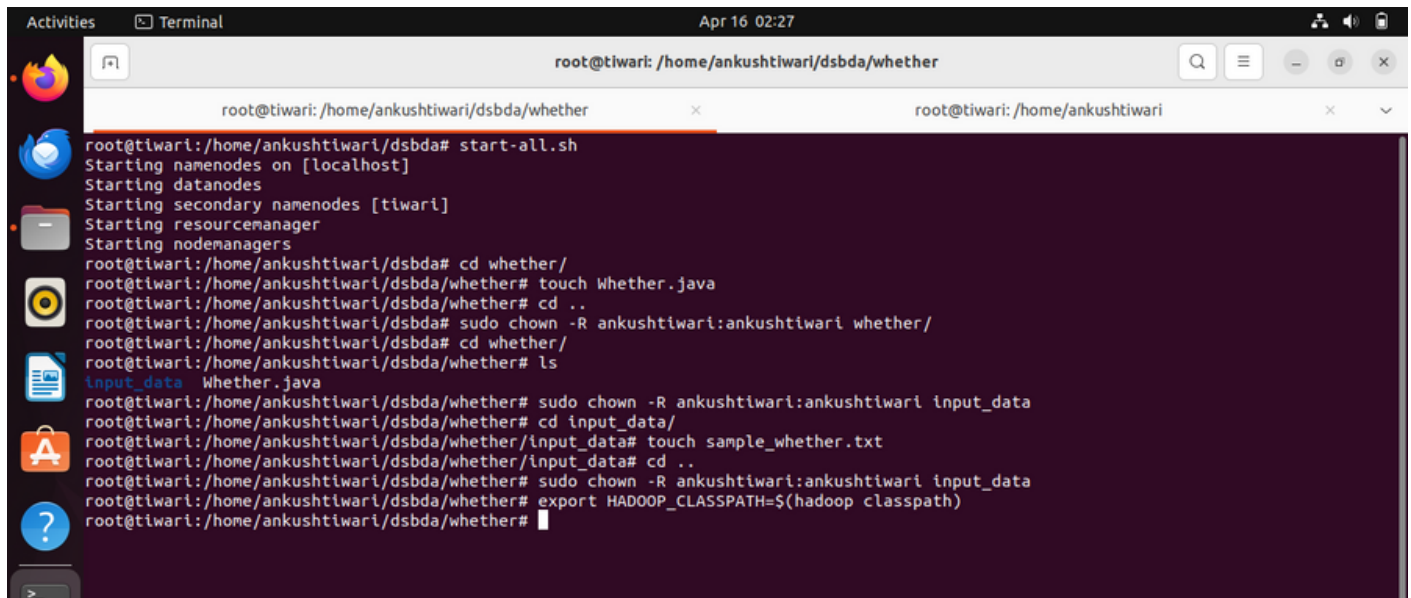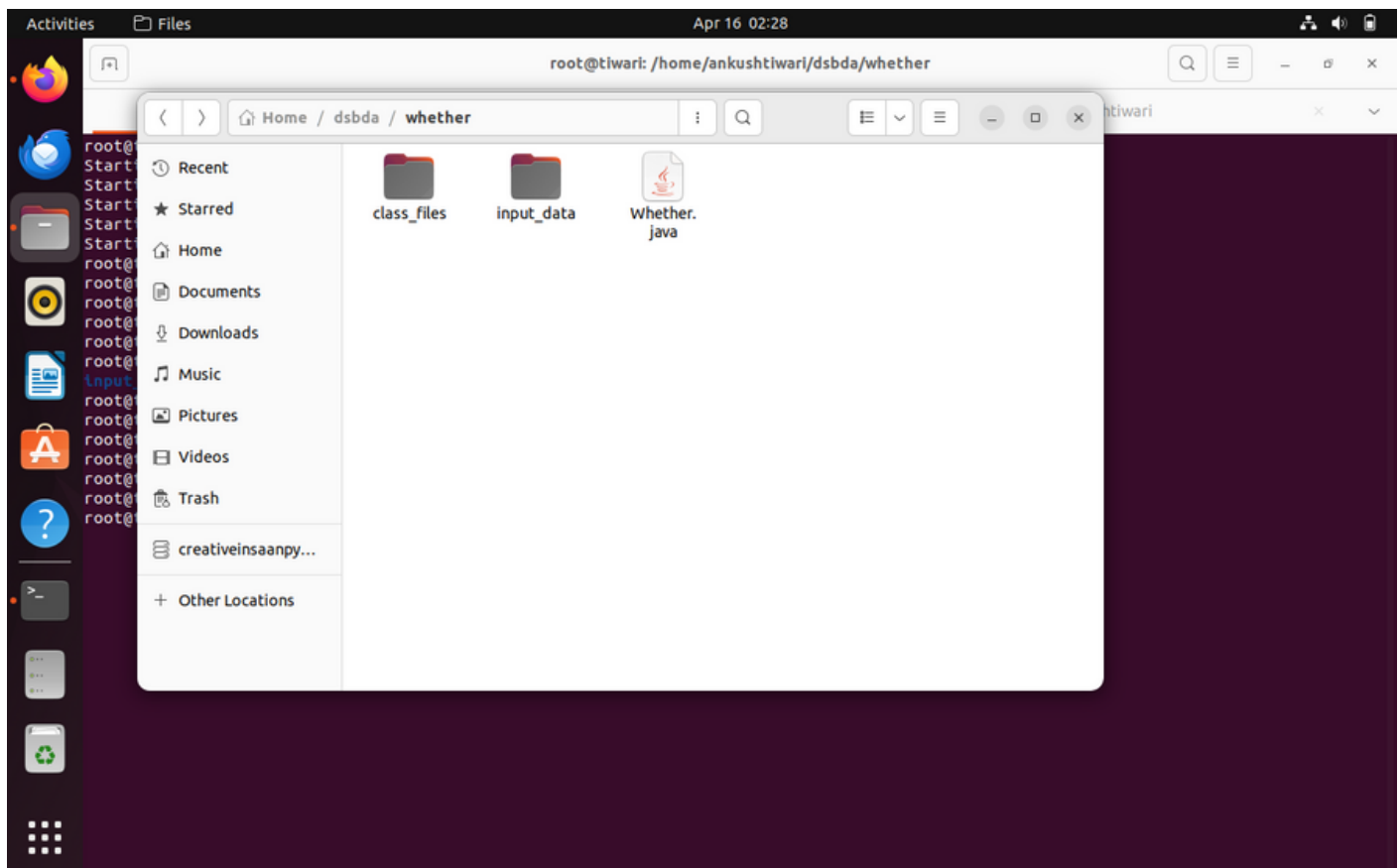**AIM: Locate dataset (e.g., sample_weather.txt) for working on weather data which reads the text input files and finds average for temperature, dew point and wind speed.**

```
root@tiwari:/home/ankushtiwari/dsbda/whether# export HADOOP_CLASSPATH=$(hadoop classpath)
root@tiwari:/home/ankushtiwari/dsbda/whether# hdfs dfs -mkdir /WeatherTut
root@tiwari:/home/ankushtiwari/dsbda/whether# hdfs dfs -mkdir /WeatherTut/Input
root@tiwari:/home/ankushtiwari/dsbda/whether# hdfs dfs -put input_data/sample_wheather.txt /WeatherTuT/Input
put: `/WeatherTuT/Input': No such file or directory: `hdfs://localhost:9000/WeatherTuT/Input'
root@tiwari:/home/ankushtiwari/dsbda/whether# hdfs dfs -put input_data/sample_wheather.txt /WeatherTut/Input
put: `input_data/sample_wheather.txt': No such file or directory
root@tiwari:/home/ankushtiwari/dsbda/whether# hdfs dfs -put '/home/ankushtiwari/dsbda/whether/input_data/sample_whether.txt' /Weathe
rTut/Input
root@tiwari:/home/ankushtiwari/dsbda/whether#
```

root@tiwari: /home/ankushtiwari/dsbda/whether

root@tiwari: /home/ankushtiwari/dsbda/whether   ×   root@tiwari: /home/ankushtiwari

```
root@tiwari:/home/ankushtiwari/dsbda/whether# javac -classpath $(hadoop classpath) -d /home/ankushtiwari/dsbda/whether/class_files /
home/ankushtiwari/dsbda/whether/Weather.java
root@tiwari:/home/ankushtiwari/dsbda/whether#
```

Home / dsbda / whether / **class_files**

- Recent
- ★ Starred
- Home
- Documents
- Downloads
- Music
- Pictures
- Videos
- Trash
- creativeinsaanpy...
- + Other Locations

Weather.class   Weather$MapClass.class   Weather$Reduce.class

---

root@tiwari: /home/ankushtiwari/dsbda/whether

root@tiwari: /home/ankushtiwari/dsbda/whether   ×   root@tiwari: /home/ankushtiwari

```
root@tiwari:/home/ankushtiwari/dsbda/whether# jar -cvf weather.jar -C /home/ankushtiwari/dsbda/whether/class_files .
added manifest
adding: Weather$Reduce.class(in = 2786) (out= 1282)(deflated 53%)
adding: Weather.class(in = 3267) (out= 1693)(deflated 48%)
adding: Weather$MapClass.class(in = 2897) (out= 1376)(deflated 52%)
root@tiwari:/home/ankushtiwari/dsbda/whether#
```

Home / dsbda / **whether**

- Recent
- ★ Starred
- Home
- Documents
- Downloads
- Music
- Pictures
- Videos
- Trash
- creativeinsaanpy...
- + Other Locations

class_files   input_data   weather.jar   Weather.java

root@tiwari: /home/ankushtiwari/dsbda/whether

root@tiwari: /home/ankushtiwari/dsbda/whether ✕   root@tiwari: /home/ankushtiwari

```
root@tiwari:/home/ankushtiwari/dsbda/whether# cd ..
root@tiwari:/home/ankushtiwari/dsbda# sudo chown -R ankushtiwari:ankushtiwari whether/
root@tiwari:/home/ankushtiwari/dsbda# cd whether/
root@tiwari:/home/ankushtiwari/dsbda/whether# hadoop jar weather.jar Weather /WeatherTut/Input /WeatherTut/Output
2024-04-16 02:49:20,797 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2024-04-16 02:49:21,593 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2024-04-16 02:49:27,590 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/root/.stagin
g/job_1713214007656_0001
2024-04-16 02:49:31,741 INFO mapred.FileInputFormat: Total input files to process : 1
2024-04-16 02:49:32,080 INFO mapreduce.JobSubmitter: number of splits:2
2024-04-16 02:49:33,074 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1713214007656_0001
2024-04-16 02:49:33,075 INFO mapreduce.JobSubmitter: Executing with tokens: []
2024-04-16 02:49:33,723 INFO conf.Configuration: resource-types.xml not found
2024-04-16 02:49:33,724 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2024-04-16 02:49:45,260 INFO impl.YarnClientImpl: Submitted application application_1713214007656_0001
2024-04-16 02:49:45,490 INFO mapreduce.Job: The url to track the job: http://tiwari.myguest.virtualbox.org:8088/proxy/application_17
13214007656_0001/
2024-04-16 02:49:45,500 INFO mapreduce.Job: Running job: job_1713214007656_0001
2024-04-16 02:50:59,397 INFO mapreduce.Job: Job job_1713214007656_0001 running in uber mode : false
2024-04-16 02:50:59,405 INFO mapreduce.Job:  map 0% reduce 0%
2024-04-16 02:51:46,853 INFO mapreduce.Job:  map 100% reduce 0%
2024-04-16 02:52:03,475 INFO mapreduce.Job:  map 100% reduce 100%
2024-04-16 02:52:12,707 INFO mapreduce.Job: Job job_1713214007656_0001 completed successfully
2024-04-16 02:52:17,481 INFO mapreduce.Job: Counters: 55
        File System Counters
                FILE: Number of bytes read=348
                FILE: Number of bytes written=831122
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=1592
                HDFS: Number of bytes written=71
                HDFS: Number of read operations=11
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
                HDFS: Number of bytes read erasure-coded=0
        Job Counters
                Killed map tasks=1
```

root@tiwari: /home/ankushtiwari/dsbda/whether

root@tiwari: /home/ankushtiwari/dsbda/whether ✕   root@tiwari: /home/ankushtiwari

```
        Map-Reduce Framework
                Map input records=9
                Map output records=9
                Map output bytes=324
                Map output materialized bytes=354
                Input split bytes=218
                Combine input records=0
                Combine output records=0
                Reduce input groups=2
                Reduce shuffle bytes=354
                Reduce input records=9
                Reduce output records=2
                Spilled Records=18
                Shuffled Maps =2
                Failed Shuffles=0
                Merged Map outputs=2
                GC time elapsed (ms)=332
                CPU time spent (ms)=11660
                Physical memory (bytes) snapshot=798445568
                Virtual memory (bytes) snapshot=7694954496
                Total committed heap usage (bytes)=595591168
                Peak Map Physical memory (bytes)=298496000
                Peak Map Virtual memory (bytes)=2561880064
                Peak Reduce Physical memory (bytes)=216670208
                Peak Reduce Virtual memory (bytes)=2571194368
        Shuffle Errors
                BAD_ID=0
                CONNECTION=0
                IO_ERROR=0
                WRONG_LENGTH=0
                WRONG_MAP=0
                WRONG_REDUCE=0
        File Input Format Counters
                Bytes Read=1374
        File Output Format Counters
                Bytes Written=71
root@tiwari:/home/ankushtiwari/dsbda/whether#
```

Browsing HDFS

localhost:9870/explorer.html#/WeatherTut/Output

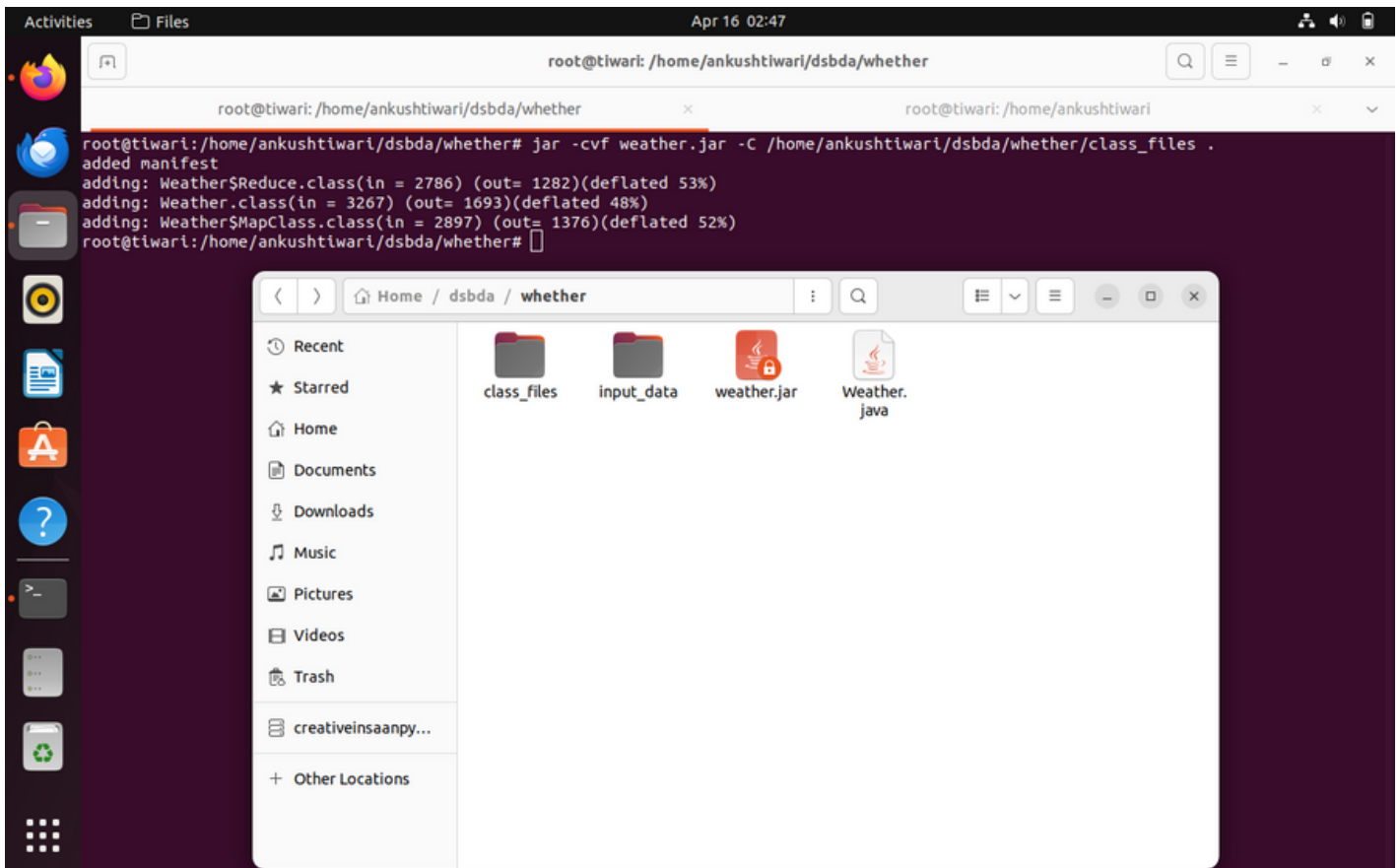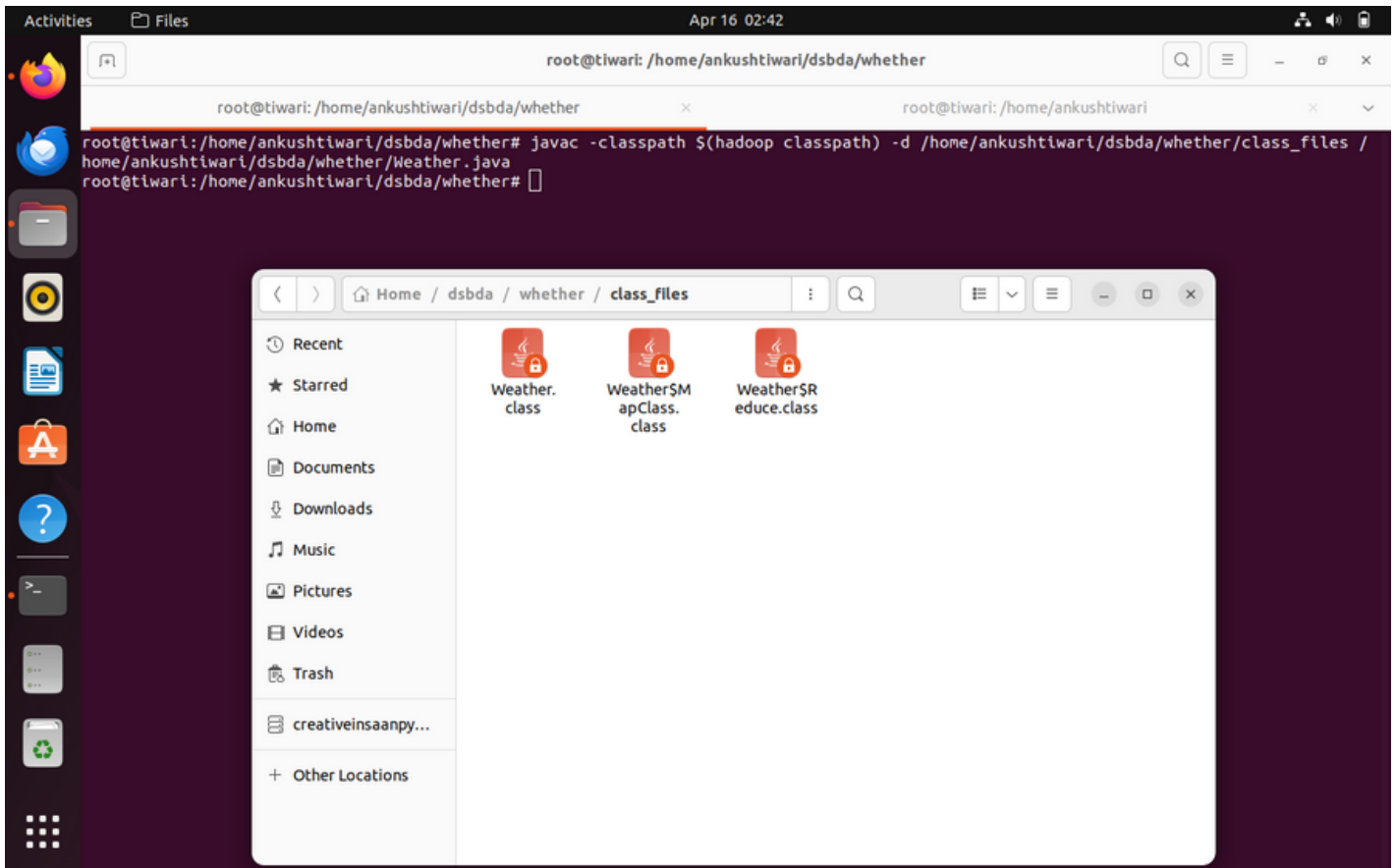Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

File information - part-00000

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

Block information -- Block 0

Block ID: 1073741848
Block Pool ID: BP-38197123-127.0.1.1-1713094056323
Generation Stamp: 1024
Size: 71
Availability:
- tiwari.myguest.virtualbox.org

File contents

```
690190_02_section1     56.03 33.0 10.7
690190_02_section4     54.884 33.0 10.7
```

Browse Directo

/WeatherTut/Output

Show 25 entries

Search:

	Permission	Owner		Block Size	Name
☐	-rw-r--r--	root		B MB	_SUCCESS
☐	-rw-r--r--	root		B MB	part-00000

Showing 1 to 2 of 2 entries

Previous 1 Next

Hadoop, 2023.

---

root@tiwari: /home/ankushtiwari/dsbda/whether

root@tiwari: /home/ankushtiwari/dsbda/whether × root@tiwari: /home/ankushtiwari

```
root@tiwari:/home/ankushtiwari/dsbda/whether# hdfs dfs -cat /WeatherTut/Output/part-00000
690190_02_section1     56.03 33.0 10.7
690190_02_section4     54.884 33.0 10.7
root@tiwari:/home/ankushtiwari/dsbda/whether#
```