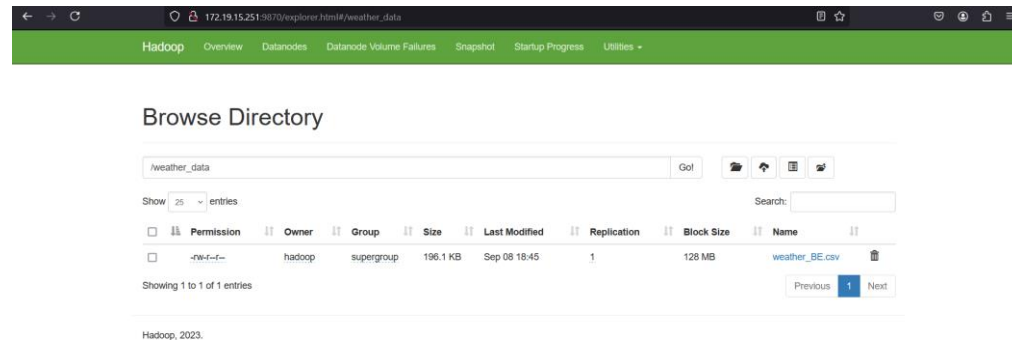


CIA-1 Hadoop Output :

```
hadoop@LAPTOP-MSOAV9AM:~$ hdfs dfs -mkdir /weather_data
```

```
hadoop@LAPTOP-MSOAV9AM:~$ hdfs dfs -put weather_BE.csv /weather_data/
```



```
hadoop@LAPTOP-MSOAV9AM:~/hadoop/bin$ hadoop jar /home/hadoop/hadoop/share/hadoop/tools/lib/hadoop-streaming-3.3.6.jar -input /weather_data/weather_BE.csv -output /weather_data/output -mapper /home/hadoop/mapper.py -reducer /home/hadoop/reducer.py -file /home/hadoop/mapper.py -file /home/hadoop/reducer.py
```

```
2024-09-08 23:05:26,230 INFO mapreduce.Job: map 100% reduce 100%
2024-09-08 23:05:26,231 INFO mapreduce.Job: Job job_local2010838308_0001 completed successfully
2024-09-08 23:05:26,241 INFO mapreduce.Job: Counters: 36
File System Counters
  FILE: Number of bytes read=292420
  FILE: Number of bytes written=1734522
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=401610
  HDFS: Number of bytes written=52
  HDFS: Number of read operations=15
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=4
  HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
  Map input records=6575
  Map output records=6562
  Map output bytes=131240
  Map output materialized bytes=144370
  Input split bytes=101
  Combine input records=0
  Combine output records=0
  Reduce input groups=1
  Reduce shuffle bytes=144370
  Reduce input records=6562
  Reduce output records=2
  Spilled Records=13124
  Shuffled Maps =1
  Failed Shuffles=0
  Merged Map outputs=1
  GC time elapsed (ms)=0
  Total committed heap usage (bytes)=587202560
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=200805
File Output Format Counters
  Bytes Written=52
2024-09-08 23:05:26,241 INFO streaming.StreamJob: Output directory: /weather_data/output
hadoop@LAPTOP-MSOAV9AM:~/hadoop/bin$
```

Browser interface showing the Hadoop file explorer for the directory `/weather_data/output`. The interface includes a navigation bar with links: Hadoop, Overview, Datanodes, Datanode Volume Failures, Snapshot, Startup Progress, and Utilities.

The main section is titled "Browse Directory" and shows a table of files in the directory `/weather_data/output`. The table has columns: Permission, Owner, Group, Size, Last Modified, Replication, Block Size, and Name.

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	hadoop	supergroup	0 B	Sep 08 23:05	1	128 MB	_SUCCESS
-rw-r--r--	hadoop	supergroup	52 B	Sep 08 23:05	1	128 MB	part-00000

Showing 1 to 2 of 2 entries

Navigation buttons: Previous, 1, Next

```
hadoop@LAPTOP-MSOAV9AM:~/hadoop/bin$ hdfs dfs -cat /weather_data/output/part-00000
Min Temperature: -11.17
Max Temperature: 29.67
```

Actual Temp using python(pandas):

```
[1] import pandas as pd
✓ 0.9s Python

[2] data = pd.read_csv('weather_BE.csv')
✓ 0.0s Python

[5] data['Temp'].max()
✓ 0.0s Python
... 29.67

[7] data['Temp'].min()
✓ 0.0s Python
... -11.17
```