Lab 8 Mean-Max Temperature using Map-Reduce

```
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [bmscecse-HP-Elite-Tower-800-G9-Desktop-PC]
Starting resourcemanager
Starting nodemanagers
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ jps
5922 NameNode
4503 org.eclipse.equinox.launcher_1.6.1000.v20250227-1734.jar
6807 NodeManager
6312 SecondaryNameNode
6058 DataNode
7226 Jps
6653 ResourceManager
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop fs -ls
Found 2 items
drwxr-xr-x   - hadoop supergroup          0 2025-04-29 15:04 op.txt
drwxr-xr-x   - hadoop supergroup          0 2025-04-29 15:11 out.txt
```

All Hadoop daemons (NameNode, DataNode, etc.) are started using start-all.sh on the local machine.

The jps command confirms active Hadoop services such as NameNode, DataNode, and ResourceManager are running.

The hadoop fs -ls command lists the contents of the HDFS root directory, showing two output folders: op.txt and out.txt.

```
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop fs -copyFromLocal /home/hadoop/Downloads/weather-data /uzairdir/wdata.txt
```

A local file weather-data.txt is copied to HDFS at /uzairdir/wdata.txt using the copyFromLocal command.

```
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop jar /home/hadoop/Desktop/Temp.jar AverageDriver /uzairdir/wdata.txt /uzair
dir/oxt
2025-05-06 15:16:20,395 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2025-05-06 15:16:20,440 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2025-05-06 15:16:20,440 INFO impl.MetricsSystemImpl: JobTracker metrics system started
2025-05-06 15:16:20,500 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool inte
rface and execute your application with ToolRunner to remedy this.
2025-05-06 15:16:20,553 INFO input.FileInputFormat: Total input files to process : 1
2025-05-06 15:16:20,583 INFO mapreduce.JobSubmitter: number of splits:1
2025-05-06 15:16:20,626 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1911472483_0001
2025-05-06 15:16:20,626 INFO mapreduce.JobSubmitter: Executing with tokens: []
2025-05-06 15:16:20,686 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
2025-05-06 15:16:20,686 INFO mapreduce.Job: Running job: job_local1911472483_0001
2025-05-06 15:16:20,686 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2025-05-06 15:16:20,689 INFO output.PathOutputCommitterFactory: No output committer factory defined, defaulting to FileOutputCommitte
rFactory
2025-05-06 15:16:20,689 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-05-06 15:16:20,689 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:f
alse, ignore cleanup failures: false
2025-05-06 15:16:20,690 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter
2025-05-06 15:16:20,729 INFO mapred.LocalJobRunner: Waiting for map tasks
2025-05-06 15:16:20,729 INFO mapred.LocalJobRunner: Starting task: attempt_local1911472483_0001_m_000000_0
2025-05-06 15:16:20,740 INFO output.PathOutputCommitterFactory: No output committer factory defined, defaulting to FileOutputCommitte
rFactory
2025-05-06 15:16:20,740 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-05-06 15:16:20,740 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:f
alse, ignore cleanup failures: false
2025-05-06 15:16:20,747 INFO mapred.Task:  Using ResourceCalculatorProcessTree : [ ]
2025-05-06 15:16:20,749 INFO mapred.MapTask: Processing split: hdfs://localhost:9000/uzairdir/wdata.txt:0+888190
2025-05-06 15:16:20,784 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)
2025-05-06 15:16:20,784 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100
2025-05-06 15:16:20,784 INFO mapred.MapTask: soft limit at 83886080
2025-05-06 15:16:20,784 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600
2025-05-06 15:16:20,784 INFO mapred.MapTask: kvstart = 26214396; length = 6553600
```

A MapReduce job is executed using the AverageDriver class to process wdata.txt and save results in oxt.

```
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop fs -cat /uzairdir/oxt/part-r-00000
1901    46
```

The output of the MapReduce job is viewed using hadoop fs -cat, showing results from the oxt/part-r-00000 file.