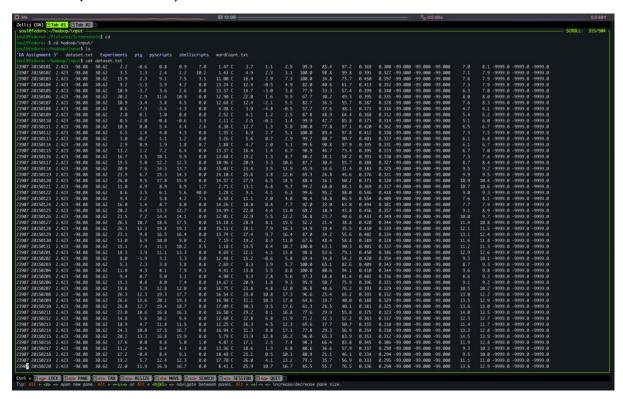
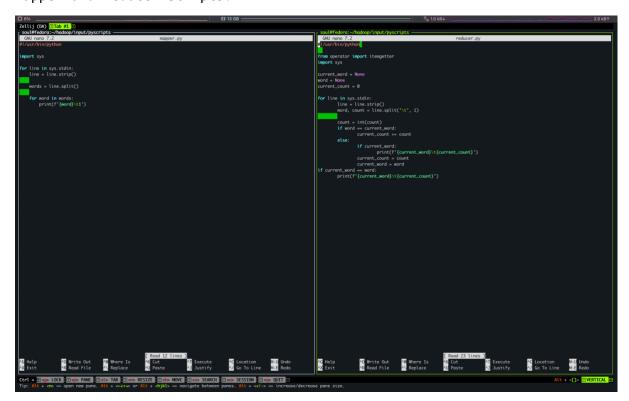
EXP NO: 3

Map Reduce Program To Process A Weather Dataset

Dataset (dataset.txt) :



Mapper and Reducer Scripts:



\$hadoop jar \$HADOOP_STREAMING -input /exp3/dataset.txt -output
/exp3/output -mapper ~/DA-Lab/exp3/mapper.py -reducer ~/DALab/exp3/reducer.py

```
Hadoop jar $HADOOP_STREAMING -input /exp3/dataset.txt -output /exp3/output -mapper -/DA-Lab/exp3/mapper.py -reducer -/DA-Lab/exp3/reducer.py
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings-on -Dswing.aatext-true
2024-09-21 00:13:19,993 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
packageJobJar: //tmp/hadoop-unjar38305904447873820999] [] /tmp/streamjob2158010624070613243.jar tmpDir=null
2024-09-21 00:13:20,918 INFO client.DefaultNoHARWFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.08032
2024-09-21 00:13:27,216 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/hadoop/.staging/job_1726891437845_0001
2024-09-21 00:13:28,202 INFO mapreduce.JobSubmitter: number of splits:2
2024-09-21 00:13:28,613 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1726891437845_0001
2024-09-21 00:13:29,203 INFO conf.Configuration: resource-types.xml not found
2024-09-21 00:13:29,203 INFO consource.ResourceUtils: Unable to find 'resource-types.xml'.
2024-09-21 00:13:29,895 INFO impl.YarnClientImpl: Submitted application application_1726891437845_0001
2024-09-21 00:13:29,998 INFO mapreduce.Job: Neurling job: job_1726891437845_0001
2024-09-21 00:13:29,998 INFO mapreduce.Job: Neurling job: job_1726891437845_0001
2024-09-21 00:13:29,999 INFO mapreduce.Job: Neurling job: job_1726891437845_0001
2024-09-21 00:13:29,999 INFO mapreduce.Job: nmp 0 reduce 0%
2024-09-21 00:13:29,999 INFO mapreduce.Job: nmp 0 reduce 0%
2024-09-21 00:13:29,999 INFO mapreduce.Job: nmp 0 reduce 0%
2024-09-21 00:13:29,999 INFO mapreduce.Job: nmp 100% reduce 0%
2024-09-21 00:14:01,012 INFO mapreduce.Job: nmp 100% reduce 100%
2024-09-21 00:14:01,012 INFO mapreduce
```

\$hdfs dfs -cat /exp3/output/*

```
-$ ./hdfs dfs -cat /exp3/output/*
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-21 00:15:38,966 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...
01
         26.5
02
         26.6
          30.8
          33.6
         38.5
08
09
         40.2
         36.5
10
         36.9
11
12
          27.6
```

Result:

Thus the dataset is processed and analysed using Hadoop Map-Reduce.