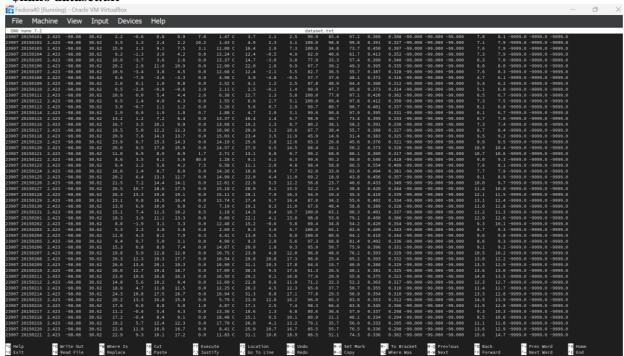
ROLL NUMBER: 210701108

EXP NO: 3 MAP REDUCE PROGRAM TO PROCESS A WEATHER DATASET

\$cd DA-Lab \$mkdir exp3 \$cd exp3

\$nano dataset.txt



\$nano mapper.py



\$nano reducer.py



ROLL NUMBER: 210701108

\$start-all.sh

```
Karthickragav@fedora:-$ start-all.sh
MARNING: Attempting to start all Apache Hadoop daemons as karthickragav in 10 seconds.
MARNING: This is not a recommended production deployment configuration.
MARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [fedora]
Starting resourcemanager
Starting nodemanagers
```

\$ jps

```
Karthi daragang fedora: $ jps
4916 Jps
4491 NodeManager
3669 NameNode
4809 SecondaryNameNode
4947 ResourceManager
3775 DataNode
```

\$hdfs dfs -mkdir/exp3

\$hdfs dfs -copyFromLocal ~/DA-Lab/exp3/dataset.txt /exp2

```
    Karthickragav@fedora:-/dalab/exp3$ hdfs dfs -ls /weatherdata

    Found 2 items
    79568 2024-09-02 10:36 /weatherdata/dataset.txt

    -rw-r---- 1 karthickragav supergroup
    0 2024-09-02 11:03 /weatherdata/dataset.txt

    drwxr-xr-x - karthickragav supergroup
    0 2024-09-02 11:03 /weatherdata/output
```

\$chmod 777 mapper.py 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

ROLL NUMBER: 210701108

```
$hdfs dfs -cat /exp3/output/*
karthickragav@fedora:~/dalab/exp3$ hdfs dfs -cat /weatherdata/output/*
01
         26.5
02
         26.6
03
         29.1
04
         30.8
05
         31.1
06
         33.6
07
         38.5
98
         40.2
09
         36.5
10
         36.9
11
         27.6
         25.9
```