

## **For Starting Hadoop any time use Following Commands on Terminal**

- 1) ssh localhost
- 2) hadoop-3.3.1/bin/hdfs namenode -format
- 3) start-all.sh
- 4) Open the Browser and Type:- localhost:9870

## **Edit and save input file, mapper and reducer files**

gedit input.txt

gedit mapper.py

gedit reducer.py

## **check files**

ls

## **check contents**

cat input.txt

cat mapper.py

cat reducer.py

## **For simple output**

cat input.txt | python mapper.py

cat input.txt | python mapper.py | sort | python reducer.py

## **Access HDFS**

hdfs dfs -put /home/admin1/input.txt /

hdfs dfs -mkdir /input

hdfs dfs -put input.txt /input

hdfs dfs -ls /input

hdfs dfs -cat /input/input.txt

## **Run the program:**

```
hadoop jar /home/admin1/hadoop-3.3.1/share/hadoop/tools/lib/hadoop-streaming-3.3.1.jar -file  
/home/admin1/mapper.py -mapper mapper.py -file /home/admin1/reducer.py -reducer reducer.py -  
input /input/input.txt -output /output
```

See the output:

```
hdfs dfs -ls /output
```

```
hdfs dfs -cat /output/part-00000
```