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

$cd DA-Lab
$mkdir exp3
$cd exp3
$nano dataset.txt

$nano mapper.py


```python
#!/usr/bin/python3
import sys
for line in sys.stdin: # remove leading and trailing whitespace
        line = line.strip() # split the line into words
        words = line.split()
        month = line[10:12]
        daily_max = line[38:45]
        daily_max = daily_max.strip()
        for word in words:
                print ('%s\t%s' % (month ,daily_max))
```

$nano reducer.py


```python
#!/usr/bin/python3
from operator import itemgetter
import sys
current_month = None
current_max = 0
month = None
for line in sys.stdin:
        line = line.strip()
        month, daily_max = line.split('\t', 1)
        try:
                daily_max = float(daily_max)
        except ValueError:
                continue
        if current_month == month:
                if daily_max > current_max:
                        current_max = daily_max
        else:
                if current_month:
                        print ('%s\t%s' % (current_month, current_max))
                current_max = daily_max
                current_month = month
if current_month == month:
        print ('%s\t%s' % (current_month, current_max))
```

$start-all.sh

$ jps

$hdfs dfs -mkdir /exp3

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

```
┌──(hadoop㉿kali)-[~/hadoop/bin]
└─$ ./hdfs dfs -ls /exp3
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-21 00:11:13,818 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...
Found 3 items
-rw-r--r--   1 hadoop supergroup      79205 2024-08-29 10:50 /exp3/dataset.txt
drwxr-xr-x   - hadoop supergroup          0 2024-08-29 10:52 /exp3/new_output
drwxr-xr-x   - hadoop supergroup          0 2024-09-13 01:00 /exp3/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

$hdfs dfs -cat /exp3/output/*

```
  ┌──(hadoop@kali)-[~/hadoop/bin]
  └─$ ./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
03      29.1
04      30.8
05      31.1
06      33.6
07      38.5
08      40.2
09      36.5
10      36.9
11      27.6
12      25.9
```

```
  ┌──(hadoop@kali)-[~/hadoop/bin]
  └─$ ./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...
```