Expt-3:

Map Reduce program to process a weather dataset.

AIM:

To implement MapReduce program to process a weather dataset.

PROCEDURE:

1. Create Weather Dataset:

nano weather_data.txt

Example content:

20220101 30.5

20220102 29.8

2. Mapper Program (mapper.py):

```
#!/usr/bin/env python3
import sys
for line in sys.stdin:
   line = line.strip()
   month = line[4:6] # Extracting month
   temp = line[7:11] # Extracting temperature
   print(f'{month}\t{temp}')
```

3. Reducer Program (reducer.py):

```
#!/usr/bin/env python3
import sys

current_month = None
current_max_temp = -float('inf')

for line in sys.stdin:
    line = line.strip()
    month, temp = line.split('\t')
```

```
try:
                temp = float(temp)
              except ValueError:
                continue
              if current_month == month:
                current_max_temp = max(current_max_temp, temp)
              else:
                if current_month:
                  print(f'{current_month}\t{current_max_temp}')
                current\_month = month
                current_max_temp = temp
           if current_month == month:
              print(f'{current_month}\t{current_max_temp}')
4. Run the Program:
           hdfs dfs -mkdir /weatherdata
           hdfs dfs -copyFromLocal weather_data.txt /weatherdata
           hadoop jar $HADOOP_HOME/share/hadoop/tools/lib/hadoop-streaming-*.jar \
           -input /weatherdata/weather_data.txt \
           -output /weatherdata/output \
           -mapper mapper.py \
           -reducer reducer.py
5. Check Output:
```

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

OUTPUT:

```
vruthi@Ubuntu:~$ hdfs dfs -cat /weatherdata/output/part-00000
2024-09-23 01:34:15,049 WARN util.NativeCodeLoader: Unable to load native-hadoop library for
01
            -2.9
            9.3
02
03
            10.4
04
            15.7
05
            20.1
06
            28.3
07
            28.2
08
            28.4
```

RESULT:

Thus, the program for weather dataset using Map Reduce has been executed successfully.