

**Cloud Computing
Big Data Analysis Project**

**Submitted by
Sai Sahaja Madhyahnapu
(M12467281)**

The project will perform Data analysis and processing using Map reduce processing in python. It will use the data from GHNC Daily: <ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily/readme.txt>. It has records of data from 2000 to 2018 which contains maximum and minimum temperatures recorded with station ID's.

In particular, the processing obtains the following results:

- Average TMIN, TMAX for each year excluding abnormalities or missing data
- Maximum TMAX, Minimum TMIN for each year excluding abnormalities or missing data
- 5 hottest, 5 coldest weather stations for each year excluding abnormalities or missing data
- Hottest and coldest day and corresponding weather stations in the entire dataset

I chose MapReduce because it is scalable, flexible, fast and have parallel processing. And also, it is easy to exclude abnormalities in MapReduce. I run the mapper and reducer files in Hadoop cluster. The obtained output is saved in a csv file and exported it to Excel and have done the analysis.

Results:

Average TMAX and TMIN:

Year	Average TMAX	Average TMIN
2000	175.588292	44.30704639
2001	178.7506964	47.92369897
2002	177.2364563	46.55436225
2003	177.0952437	48.62549307
2004	174.9770195	49.55504364
2005	177.9061243	49.9807912
2006	181.089305	51.120869
2007	178.4365732	49.22216519
2008	170.3911742	40.93179589
2009	169.1661447	43.70164404
2010	171.781345	47.47453654
2011	172.7793752	46.29123543
2012	184.4661432	53.87463503
2013	168.028435	43.53508188
2014	169.2431074	44.31857444
2015	178.0581972	54.04334792
2016	179.8717209	55.58451407
2017	176.5666926	52.69286723
2018	78.86609582	-37.35880684

Maximum TMAX and Minimum TMIN:

Year	Max TMAX	Min TMIN
2000	522	-578
2001	528	-528
2002	533	-472
2003	533	-500
2004	517	-533
2005	539	-550
2006	528	-528
2007	556	-539
2008	528	-578
2009	533	-556
2010	517	-533
2011	511	-517
2012	537	-544
2013	539	-528
2014	522	-500
2015	556	-528
2016	539	-469
2017	528	-520
2018	400	-528

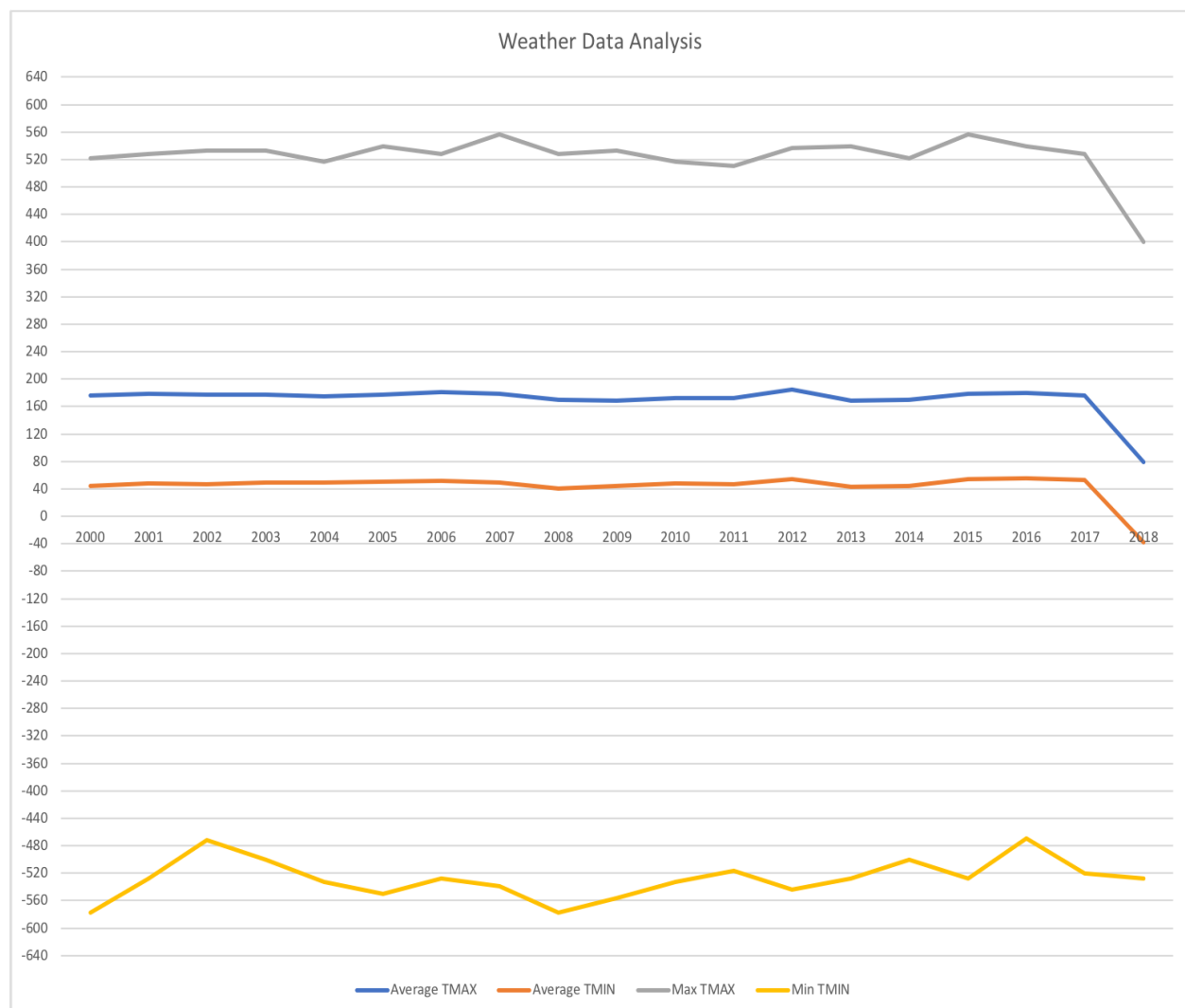
5 Hottest and 5 Coldest stations:

Year	Hottest Weather Stations	Coldest weather stations
2000	['USC00042319', 'USW00003104', 'USR0000AHAV', 'USC00021050', 'USC00024761']	['USC00508140', 'USW00026440', 'USC00502350', 'USC00505644', 'USC00501684']
2001	['USC00042319', 'USR0000AHAV', 'USC00024829', 'USC00021050', 'USR0000CM0J']	['USR0000ABCA', 'USC00508470', 'USS0051R01S', 'USW00026508', 'USC00504567']
2002	['USR0000CM0J', 'USC00042319', 'USC00265846', 'USC00044297', 'USR0000CCOR']	['USR0000CM0J', 'USC00042319', 'USC00265846', 'USC00044297', 'USR0000CCOR']
2003	['USC00021050', 'USR0000CMEA', 'USR0000AHAV', 'USC00029376', 'USC00042319']	['USS0050S01S', 'USC00501492', 'USC00509869', 'USS0051R01S', 'USW00026533']
2004	['USC00042319', 'USW00053139', 'USC00261371', 'USC00021050', 'USC00024761']	['USC00501684', 'USC00502568', 'USS0045010S', 'USW00026412', 'USC00503009']
2005	['USW00023179', 'USC00042319', 'USR0000CM0J', 'USW00053139', 'USC00029376']	['USW00026412', 'USC00509869', 'USC00509313', 'USC00502568', 'USC00501684']
2006	['USR0000CM0J', 'USC00044259', 'USW00053139', 'USC00042319', 'USC00029376']	['USS0051R01S', 'USC00501492', 'USR0000ACHL', 'USR0000ABEV', 'USR0000ASEL']
2007	['USW00053139', 'USR0000AHAV', 'USR0000CSAW', 'USC00021050', 'USC00042319']	['USR0000ACHL', 'USC00502607', 'USW00026422', 'USC00501684', 'USS0045R01S']
2008	['USC00261371', 'USC00042319', 'USC00024761', 'USC00044297', 'USW00053139']	['USS0045R01S', 'USR0000ACHL', 'USR0000ABEV', 'USC00501684', 'USC00501492']
2009	['USW00053139', 'USR0000CSQU', 'USR0000CBUU', 'USC00042319', 'USR0000CM0J']	['USC00502101', 'USC00501684', 'USC00509313', 'USR0000AEAG', 'USW00026412']
2010	['USC00021050', 'USR0000AHAV', 'USR0000CBUU', 'USC00042319', 'USR0000CM0J']	['USC00501684', 'USR0000ACHL', 'USS0045R01S', 'USS0051R01S', 'USC00502101']
2011	['USC00042319', 'USR0000CCAA', 'USW00053139', 'USR0000CBUU', 'USC00022782']	['USC00509869', 'USC00501492', 'USS0051R01S', 'USC00501684', 'USR0000ANOR']
2012	['USW00053139', 'USR0000AHAV', 'USC00021050', 'USS0005N23S', 'USC00042319']	['USC00504210', 'USC00503165', 'USS0051R01S', 'USS0045R01S', 'USC00503212']
2013	['USW00053139', 'USW00004134', 'USR0000AHAV', 'USC00044297', 'USC00042319']	['USS0051R01S', 'USS0045004S', 'USS0045R01S', 'USC00502339', 'USC00501684']
2014	['USW00053139', 'USC00042319', 'USC00028396', 'USW00003125', 'USS00009504S']	['USC00501684', 'USR0000ABCK', 'USS0048V01S', 'USS0051R01S', 'USS0045R01S']
2015	['USW00053139', 'USR0000CBUU', 'USW00003104', 'USC00045502', 'USC00042319']	['USS0041P07S', 'USC00509314', 'USW00026533', 'USC00501684', 'USS0045R01S']
2016	['USC00042319', 'USW00023179', 'USC00040924', 'USW00053139', 'USR0000CBEV']	['USR0000ALOS', 'USC00501684', 'USR0000ACHL', 'USS0045R01S', 'USS0041P07S']
2017	['USW00053139', 'USC00021050', 'USW00023179', 'USC00046386', 'USC00042319']	['USR0000ASLC', 'USS0051R01S', 'USS0045004S', 'USW00026529', 'USR0000ALIV']
2018	['USC000417624', 'USC000411524', 'USR0000CCAA', 'USW000012959', 'USR0000TFAL']	['USW000096406', 'USR0000AKAV', 'USR0000ANOR', 'USW00026529', 'USC00502339']

Hottest Day and Coldest Day:

Day	Date	Station ID
Hottest Day	20071013	USR0000CSAW
Coldest day	20000101	USC00501684

Analysis of obtained results from given data in graph format:



Link to Github:

<https://github.uc.edu/madhyasa/Cloud-Project-2>

<https://github.uc.edu/madhyasa/Cloud-Project-2/blob/master/README.md>

Appendix:

Mapper.py

```
#!/usr/bin/env python

import sys
import csv

file = csv.reader(sys.stdin)
for line in file:
    if line[5] == "" and line[4] != "P" and line[6] != "" and (line[2] == 'TMAX' or line[2] ==
'TMIN') and line[3] != -9999:
        print '%s\t%s\t%s\t%s' % (line[1], line[0], line[2], line[3])
```

Reducer.py

```
#!/usr/bin/env python

from operator import itemgetter
import sys

tmax = 0
tmin = 0
maxtmx = 0
mintmn = 999999
tmaxct = 0
tminct = 0
currentyear = None
year = None
element = None
value = 0
date = None
stationid = None
hottestweatherstationsids = []
coldestweatherstationsids = []
hottestweatherstationsvalues = []
coldestweatherstationsvalues = []
hottestvalue = 0
coldestvalue = 999999
hottestday = None
coldestday = None
hottestdaystationid = None
coldestdaystationid = None
```

```

for line in sys.stdin:

    line = line.strip()

    date, stationid, element, value = line.split('\t', 3)
    year = str(date)[:4]

    try:
        value = int(value)
    except ValueError:
        continue

    if currentyear is None:
        currentyear = year

    if year == currentyear:
        if element == "TMAX":
            tmax += value
            if value > maxtmx:
                maxtmx = value
            tmaxct += 1

        elif element == "TMIN":
            tmin += value
            if value < mintmn:
                mintmn = value
            tminct += 1

        if value > hottestvalue:
            hottestvalue = value
            hottestdaystationid = stationid
            hottestday = date

        if value < coldestvalue:
            coldestvalue = value
            coldestdaystationid = stationid
            coldestday = date

    if len(hottestweatherstationsids) < 5:
        hottestweatherstationsids.append(stationid)
        hottestweatherstationsvalues.append(value)
    else:
        minhottestvalue = min(hottestweatherstationsvalues)
        if minhottestvalue < value and stationid not in hottestweatherstationsids:

```

```

        for idx, i in enumerate(hottestweatherstationsvalues):
            if i == minhottestvalue:
                hottestweatherstationsvalues[idx] = value
                hottestweatherstationsids[idx] = stationid
                break

    if len(coldestweatherstationsids) < 5:
        coldestweatherstationsids.append(stationid)
        coldestweatherstationsvalues.append(value)
    else:
        maxcoldestvalue = max(coldestweatherstationsvalues)
        if maxcoldestvalue > value and stationid not in coldestweatherstationsids:
            for idx, i in enumerate(coldestweatherstationsvalues):
                if i == maxcoldestvalue:
                    coldestweatherstationsvalues[idx] = value
                    coldestweatherstationsids[idx] = stationid
                    break
    else:
        print '%s' % currentyear
        print 'Average TMAX\t%s\t' % (tmax * 1.0 / tmaxct)
        print 'Average TMIN\t%s\t' % (tmin * 1.0 / tminct)
        print 'Max TMAX\t%s\t' % maxtmx
        print 'Min TMIN\t%s\t' % mintmn
        print 'Hottest Weather Stations\t%s' % hottestweatherstationsids
        print 'Coldest Weather Stations\t%s' % coldestweatherstationsids

    currentyear = year
    tmax = 0
    tmin = 0
    maxtmx = 0
    mintmn = 999999
    tmaxct = 0
    tminct = 0
    hottestweatherstationsids = []
    hottestweatherstationsvalues = []
    coldestweatherstationsids = []
    coldestweatherstationsvalues = []
    if element == "TMAX":
        tmax = value
        maxtmx = value
        tmaxct = 1

    elif element == "TMIN":
        tmin = value
        mintmn = value
        tminct = 1

```

```

if value < coldestvalue:
    coldestvalue = value
    coldestdaystationid = stationid
    coldestday = date

if value > hottestvalue:
    hottestvalue = value
    hottestdaystationid = stationid
    hottestday = date

if year == currentyear:
    print '%s' % currentyear
    print '%s\tAverage TMAX\t%s\t' % (currentyear, (tmax * 1.0 / tmaxct))
    print '%s\tAverage TMIN\t%s\t' % (currentyear, (tmin * 1.0 / tminct))
    print '%s\tMax TMAX\t%s\t' % (currentyear, maxtmx)
    print '%s\tMin TMIN\t%s\t' % (currentyear, mintmn)
    print '%s\tHottest Weather Stations\t%s' % (currentyear, hottestweatherstationsids)
    print '%s\tColdest Weather Stations\t%s' % (currentyear, coldestweatherstationsids)
    print '\n'
    print 'Hottest Day\t%s\t - StationID\t%s' % (hottestday, hottestdaystationid)
    print 'Coldest Day\t%s\t - StationID\t%s' % (coldestday, coldestdaystationid)

```

Logfile:

```

madhyasa@hadoop-gate-0:~/CloudPratice$ hadoop jar /usr/hdp/2.6.3.0-235/hadoop-
mapreduce/hadoop-streaming-2.7.3.2.6.3.0-235.jar -file map.py -mapper map.py -file red.py -
reducer red.py -input /user/tatavag/PIIweather -output result4.csv
18/04/24 21:42:39 WARN streaming.StreamJob: -file option is deprecated, please use generic
option -files instead.
packageJobJar: [map.py, red.py] [/usr/hdp/2.6.3.0-235/hadoop-mapreduce/hadoop-streaming-
2.7.3.2.6.3.0-235.jar] /tmp/streamjob4163282495771342433.jar tmpDir=null
18/04/24 21:42:40 INFO client.RMProxy: Connecting to ResourceManager at hadoop2-0-
0.csccloud.ceas.uc.edu/192.168.2.20:8050
18/04/24 21:42:40 INFO client.AHSPProxy: Connecting to Application History server at hadoop2-
0-0.csccloud.ceas.uc.edu/192.168.2.20:10200
18/04/24 21:42:40 INFO client.RMProxy: Connecting to ResourceManager at hadoop2-0-
0.csccloud.ceas.uc.edu/192.168.2.20:8050
18/04/24 21:42:40 INFO client.AHSPProxy: Connecting to Application History server at hadoop2-
0-0.csccloud.ceas.uc.edu/192.168.2.20:10200
18/04/24 21:42:41 INFO mapred.FileInputFormat: Total input paths to process : 19
18/04/24 21:42:41 INFO mapreduce.JobSubmitter: number of splits:37

```


18/04/24 21:42:41 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1524342268806_0173
18/04/24 21:42:42 INFO impl.YarnClientImpl: Submitted application application_1524342268806_0173
18/04/24 21:42:42 INFO mapreduce.Job: The url to track the job: http://hadoop2-0-cscloud.ceas.uc.edu:8088/proxy/application_1524342268806_0173/
18/04/24 21:42:42 INFO mapreduce.Job: Running job: job_1524342268806_0173
18/04/24 21:42:49 INFO mapreduce.Job: Job job_1524342268806_0173 running in uber mode : false
18/04/24 21:42:49 INFO mapreduce.Job: map 0% reduce 0%
18/04/24 21:43:00 INFO mapreduce.Job: map 1% reduce 0%
18/04/24 21:43:01 INFO mapreduce.Job: map 4% reduce 0%
18/04/24 21:43:02 INFO mapreduce.Job: map 6% reduce 0%
18/04/24 21:43:03 INFO mapreduce.Job: map 7% reduce 0%
18/04/24 21:43:04 INFO mapreduce.Job: map 12% reduce 0%
18/04/24 21:43:05 INFO mapreduce.Job: map 17% reduce 0%
18/04/24 21:43:06 INFO mapreduce.Job: map 20% reduce 0%
18/04/24 21:43:07 INFO mapreduce.Job: map 28% reduce 0%
18/04/24 21:43:08 INFO mapreduce.Job: map 31% reduce 0%
18/04/24 21:43:09 INFO mapreduce.Job: map 36% reduce 0%
18/04/24 21:43:10 INFO mapreduce.Job: map 44% reduce 0%
18/04/24 21:43:11 INFO mapreduce.Job: map 46% reduce 0%
18/04/24 21:43:12 INFO mapreduce.Job: map 52% reduce 0%
18/04/24 21:43:13 INFO mapreduce.Job: map 55% reduce 0%
18/04/24 21:43:14 INFO mapreduce.Job: map 56% reduce 0%
18/04/24 21:43:15 INFO mapreduce.Job: map 58% reduce 0%
18/04/24 21:43:16 INFO mapreduce.Job: map 61% reduce 0%
18/04/24 21:43:17 INFO mapreduce.Job: map 63% reduce 0%
18/04/24 21:43:18 INFO mapreduce.Job: map 65% reduce 0%
18/04/24 21:43:19 INFO mapreduce.Job: map 67% reduce 0%
18/04/24 21:43:20 INFO mapreduce.Job: map 70% reduce 0%
18/04/24 21:43:21 INFO mapreduce.Job: map 74% reduce 0%
18/04/24 21:43:22 INFO mapreduce.Job: map 74% reduce 11%
18/04/24 21:43:23 INFO mapreduce.Job: map 77% reduce 11%
18/04/24 21:43:24 INFO mapreduce.Job: map 79% reduce 11%
18/04/24 21:43:25 INFO mapreduce.Job: map 80% reduce 14%
18/04/24 21:43:26 INFO mapreduce.Job: map 81% reduce 14%
18/04/24 21:43:27 INFO mapreduce.Job: map 85% reduce 14%
18/04/24 21:43:28 INFO mapreduce.Job: map 85% reduce 15%
18/04/24 21:43:29 INFO mapreduce.Job: map 87% reduce 15%
18/04/24 21:43:37 INFO mapreduce.Job: map 87% reduce 18%
18/04/24 21:43:40 INFO mapreduce.Job: map 87% reduce 23%
18/04/24 21:43:45 INFO mapreduce.Job: map 88% reduce 23%
18/04/24 21:43:51 INFO mapreduce.Job: map 90% reduce 23%
18/04/24 21:43:53 INFO mapreduce.Job: map 90% reduce 25%
18/04/24 21:43:54 INFO mapreduce.Job: map 91% reduce 25%

18/04/24 21:43:56 INFO mapreduce.Job: map 91% reduce 29%
18/04/24 21:44:02 INFO mapreduce.Job: map 92% reduce 29%
18/04/24 21:44:03 INFO mapreduce.Job: map 93% reduce 29%
18/04/24 21:44:05 INFO mapreduce.Job: map 94% reduce 29%
18/04/24 21:44:08 INFO mapreduce.Job: map 95% reduce 29%
18/04/24 21:44:27 INFO mapreduce.Job: map 97% reduce 29%
18/04/24 21:44:28 INFO mapreduce.Job: map 99% reduce 29%
18/04/24 21:44:48 INFO mapreduce.Job: map 100% reduce 29%
18/04/24 21:44:51 INFO mapreduce.Job: map 100% reduce 30%
18/04/24 21:44:57 INFO mapreduce.Job: map 100% reduce 31%
18/04/24 21:45:15 INFO mapreduce.Job: map 100% reduce 32%
18/04/24 21:45:24 INFO mapreduce.Job: map 100% reduce 33%
18/04/24 21:45:33 INFO mapreduce.Job: map 100% reduce 44%
18/04/24 21:45:36 INFO mapreduce.Job: map 100% reduce 67%
18/04/24 21:45:51 INFO mapreduce.Job: map 100% reduce 68%
18/04/24 21:46:10 INFO mapreduce.Job: map 100% reduce 69%
18/04/24 21:46:29 INFO mapreduce.Job: map 100% reduce 70%
18/04/24 21:46:47 INFO mapreduce.Job: map 100% reduce 71%
18/04/24 21:47:05 INFO mapreduce.Job: map 100% reduce 72%
18/04/24 21:47:23 INFO mapreduce.Job: map 100% reduce 73%
18/04/24 21:47:41 INFO mapreduce.Job: map 100% reduce 74%
18/04/24 21:47:56 INFO mapreduce.Job: map 100% reduce 75%
18/04/24 21:48:14 INFO mapreduce.Job: map 100% reduce 76%
18/04/24 21:48:32 INFO mapreduce.Job: map 100% reduce 77%
18/04/24 21:48:51 INFO mapreduce.Job: map 100% reduce 78%
18/04/24 21:49:09 INFO mapreduce.Job: map 100% reduce 79%
18/04/24 21:49:27 INFO mapreduce.Job: map 100% reduce 80%
18/04/24 21:49:45 INFO mapreduce.Job: map 100% reduce 81%
18/04/24 21:50:03 INFO mapreduce.Job: map 100% reduce 82%
18/04/24 21:50:21 INFO mapreduce.Job: map 100% reduce 83%
18/04/24 21:50:39 INFO mapreduce.Job: map 100% reduce 84%
18/04/24 21:50:57 INFO mapreduce.Job: map 100% reduce 85%
18/04/24 21:51:15 INFO mapreduce.Job: map 100% reduce 86%
18/04/24 21:51:35 INFO mapreduce.Job: map 100% reduce 87%
18/04/24 21:51:50 INFO mapreduce.Job: map 100% reduce 88%
18/04/24 21:52:08 INFO mapreduce.Job: map 100% reduce 89%
18/04/24 21:52:26 INFO mapreduce.Job: map 100% reduce 90%
18/04/24 21:52:44 INFO mapreduce.Job: map 100% reduce 91%
18/04/24 21:53:02 INFO mapreduce.Job: map 100% reduce 92%
18/04/24 21:53:20 INFO mapreduce.Job: map 100% reduce 93%
18/04/24 21:53:38 INFO mapreduce.Job: map 100% reduce 94%
18/04/24 21:53:57 INFO mapreduce.Job: map 100% reduce 95%
18/04/24 21:54:15 INFO mapreduce.Job: map 100% reduce 96%
18/04/24 21:54:33 INFO mapreduce.Job: map 100% reduce 97%
18/04/24 21:54:51 INFO mapreduce.Job: map 100% reduce 98%
18/04/24 21:55:09 INFO mapreduce.Job: map 100% reduce 99%

18/04/24 21:55:27 INFO mapreduce.Job: map 100% reduce 100%
18/04/24 21:55:34 INFO mapreduce.Job: Job job_1524342268806_0173 completed successfully
18/04/24 21:55:34 INFO mapreduce.Job: Counters: 51

File System Counters

FILE: Number of bytes read=3389252780
FILE: Number of bytes written=6784331842
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=4045895793
HDFS: Number of bytes written=5691
HDFS: Number of read operations=114
HDFS: Number of large read operations=0
HDFS: Number of write operations=2

Job Counters

Launched map tasks=37
Launched reduce tasks=1
Other local map tasks=9
Data-local map tasks=21
Rack-local map tasks=7
Total time spent by all maps in occupied slots (ms)=2263728
Total time spent by all reduces in occupied slots (ms)=1483580
Total time spent by all map tasks (ms)=1131864
Total time spent by all reduce tasks (ms)=741790
Total vcore-milliseconds taken by all map tasks=1131864
Total vcore-milliseconds taken by all reduce tasks=741790
Total megabyte-milliseconds taken by all map tasks=1738543104
Total megabyte-milliseconds taken by all reduce tasks=1519185920

Map-Reduce Framework

Map input records=107556969
Map output records=106767613
Map output bytes=3175717524
Map output materialized bytes=3389252972
Input split bytes=4736
Combine input records=0
Combine output records=0
Reduce input groups=6675
Reduce shuffle bytes=3389252972
Reduce input records=106767613
Reduce output records=137
Spilled Records=213535226
Shuffled Maps =37
Failed Shuffles=0
Merged Map outputs=37
GC time elapsed (ms)=118858
CPU time spent (ms)=1822710

Physical memory (bytes) snapshot=45491015680
Virtual memory (bytes) snapshot=124659118080
Total committed heap usage (bytes)=45830635520

Shuffle Errors

BAD_ID=0
CONNECTION=0
IO_ERROR=0
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0

File Input Format Counters

Bytes Read=4045891057

File Output Format Counters

Bytes Written=5691

18/04/24 21:55:34 INFO streaming.StreamJob: Output directory: result4.csv

```
2000
Average TMAX      175.588292003
Average TMIN      44.3070463878
Max TMAX          522
Min TMIN          -578
Hottest Weather Stations ['USC00042319', 'USW00003104', 'USR0000AHAV', 'USC00021050', 'USC00024761']
Coldest Weather Stations ['USC00508140', 'USW00026440', 'USC00502350', 'USC00505644', 'USC00501684']
2001
Average TMAX      178.750696444
Average TMIN      47.9236989653
Max TMAX          528
Min TMIN          -528
Hottest Weather Stations ['USC00042319', 'USR0000AHAV', 'USC00024829', 'USC00021050', 'USR0000CMOJ']
Coldest Weather Stations ['USR0000ABCA', 'USC00508470', 'USS0051R01S', 'USW00026508', 'USC00504567']
2002
Average TMAX      177.236456277
Average TMIN      46.5543622476
Max TMAX          533
Min TMIN          -472
Hottest Weather Stations ['USR0000CMOJ', 'USC00042319', 'USC00265846', 'USC00044297', 'USR0000CCOR']
Coldest Weather Stations ['USR0000ABEV', 'USC00503212', 'USS0050S01S', 'USC00501684', 'USW00026533']
2003
Average TMAX      177.095243712
Average TMIN      48.6254930721
Max TMAX          533
Min TMIN          -500
Hottest Weather Stations ['USC00021050', 'USR0000CMEA', 'USR0000AHAV', 'USC00029376', 'USC00042319']
Coldest Weather Stations ['USS0050S01S', 'USC00501492', 'USC00509869', 'USS0051R01S', 'USW00026533']
2004
Average TMAX      174.977019472
Average TMIN      49.5550436384
Max TMAX          517
Min TMIN          -533
Hottest Weather Stations ['USC00042319', 'USW000053139', 'USC00261371', 'USC00021050', 'USC00024761']
Coldest Weather Stations ['USC00501684', 'USC00502568', 'USS0045010S', 'USW00026412', 'USC00503009']
2005
Average TMAX      177.906124277
Average TMIN      49.9807911967
Max TMAX          539
Min TMIN          -550
Hottest Weather Stations ['USW00023179', 'USC00042319', 'USR0000CMOJ', 'USW000053139', 'USC00029376']
Coldest Weather Stations ['USW00026412', 'USC00509869', 'USC00509313', 'USC00502568', 'USC00501684']
2006
Average TMAX      181.089305009
Average TMIN      51.1208690005
Max TMAX          528
Min TMIN          -528
Hottest Weather Stations ['USR0000CMOJ', 'USC00044259', 'USW000053139', 'USC00042319', 'USC00029376']
Coldest Weather Stations ['USS0051R01S', 'USC00501492', 'USR0000ACHL', 'USR0000ABEV', 'USR0000ASEL']
2007
Average TMAX      178.436573222
Average TMIN      49.2221651922
Max TMAX          556
Min TMIN          -539
Hottest Weather Stations ['USW000053139', 'USR0000AHAV', 'USR0000CSAW', 'USC00021050', 'USC00042319']
Coldest Weather Stations ['USR0000ACHL', 'USC00502607', 'USW00026422', 'USC00501684', 'USS0045R01S']
2008
Average TMAX      170.391174208
Average TMIN      40.9317958853
Max TMAX          528
Min TMIN          -578
Hottest Weather Stations ['USC00261371', 'USC00042319', 'USC00024761', 'USC00044297', 'USW000053139']
Coldest Weather Stations ['USS0045R01S', 'USR0000ACHL', 'USR0000ABEV', 'USC00501684', 'USC00501492']
```

```

2009
Average TMAX      169.166144748
Average TMIN      43.701644066
Max TMAX          533
Min TMIN          -556
Hottest Weather Stations      ['USW00053139', 'USR0000CSQU', 'USR0000CBUU', 'USC00042319', 'USR0000CMOJ']
Coldest Weather Stations      ['USC00502101', 'USC00501684', 'USC00509313', 'USR0000AEAG', 'USW00026412']
2010
Average TMAX      171.781345009
Average TMIN      47.4745365367
Max TMAX          517
Min TMIN          -533
Hottest Weather Stations      ['USC00021050', 'USR0000AHAV', 'USR0000CBUU', 'USC00042319', 'USR0000CMOJ']
Coldest Weather Stations      ['USC00501684', 'USR0000ACHL', 'USS0045R01S', 'USS0051R01S', 'USC00502101']
2011
Average TMAX      172.779375228
Average TMIN      46.29123543
Max TMAX          511
Min TMIN          -517
Hottest Weather Stations      ['USC00042319', 'USR0000CCAA', 'USW00053139', 'USR0000CBUU', 'USC00022782']
Coldest Weather Stations      ['USC00509869', 'USC00501492', 'USS0051R01S', 'USC00501684', 'USR0000ANOR']
2012
Average TMAX      184.466143175
Average TMIN      53.8746350331
Max TMAX          537
Min TMIN          -544
Hottest Weather Stations      ['USW00053139', 'USR0000AHAV', 'USC00021050', 'USS0005N23S', 'USC00042319']
Coldest Weather Stations      ['USC00504210', 'USC00503165', 'USS0051R01S', 'USS0045R01S', 'USC00503212']
2013
Average TMAX      168.028435001
Average TMIN      43.535081883
Max TMAX          539
Min TMIN          -528
Hottest Weather Stations      ['USW00053139', 'USW00004134', 'USR0000AHAV', 'USC00044297', 'USC00042319']
Coldest Weather Stations      ['USS0051R01S', 'USS0045004S', 'USS0045R01S', 'USC00502339', 'USC00501684']
2014
Average TMAX      169.243107388
Average TMIN      44.3185744368
Max TMAX          522
Min TMIN          -500
Hottest Weather Stations      ['USW00053139', 'USC00042319', 'USC00028396', 'USW00003125', 'USS0009S04S']
Coldest Weather Stations      ['USC00501684', 'USR0000ABCK', 'USS0048V01S', 'USS0051R01S', 'USS0045R01S']
2015
Average TMAX      178.058197184
Average TMIN      54.0433479213
Max TMAX          556
Min TMIN          -528
Hottest Weather Stations      ['USW00053139', 'USR0000CBUU', 'USW00003104', 'USC00045502', 'USC00042319']
Coldest Weather Stations      ['USS0041P07S', 'USC00509314', 'USW00026533', 'USC00501684', 'USS0045R01S']
2016
Average TMAX      179.871720866
Average TMIN      55.5845140688
Max TMAX          539
Min TMIN          -469
Hottest Weather Stations      ['USC00042319', 'USW00023179', 'USC00040924', 'USW00053139', 'USR0000CBEV']
Coldest Weather Stations      ['USR0000ALOS', 'USC00501684', 'USR0000ACHL', 'USS0045R01S', 'USS0041P07S']
2017
Average TMAX      176.566692563
Average TMIN      52.6928672336
Max TMAX          528
Min TMIN          -520
Hottest Weather Stations      ['USW00053139', 'USC00021050', 'USW00023179', 'USC00046386', 'USC00042319']
Coldest Weather Stations      ['USR0000ASLC', 'USS0051R01S', 'USS0045004S', 'USW00026529', 'USR0000ALIV']

```

```

2018
2018 Average TMAX      78.8660958193
2018 Average TMIN      -37.3588068434
2018 Max TMAX          400
2018 Min TMIN          -528
2018 Hottest Weather Stations      ['USC00417624', 'USC00411524', 'USR0000CCAA', 'USW00012959', 'USR0000TFAL']
2018 Coldest Weather Stations      ['USW00096406', 'USR0000AKAV', 'USR0000ANOR', 'USW00026529', 'USC00502339']

```

```

Hottest Day      20071013      - StationID      USR0000CSAW
Coldest Day      20000101      - StationID      USC00501684
madhyasa@hadoop-gate-0:~/CloudPratice$

```