

EXPERIMENT-7

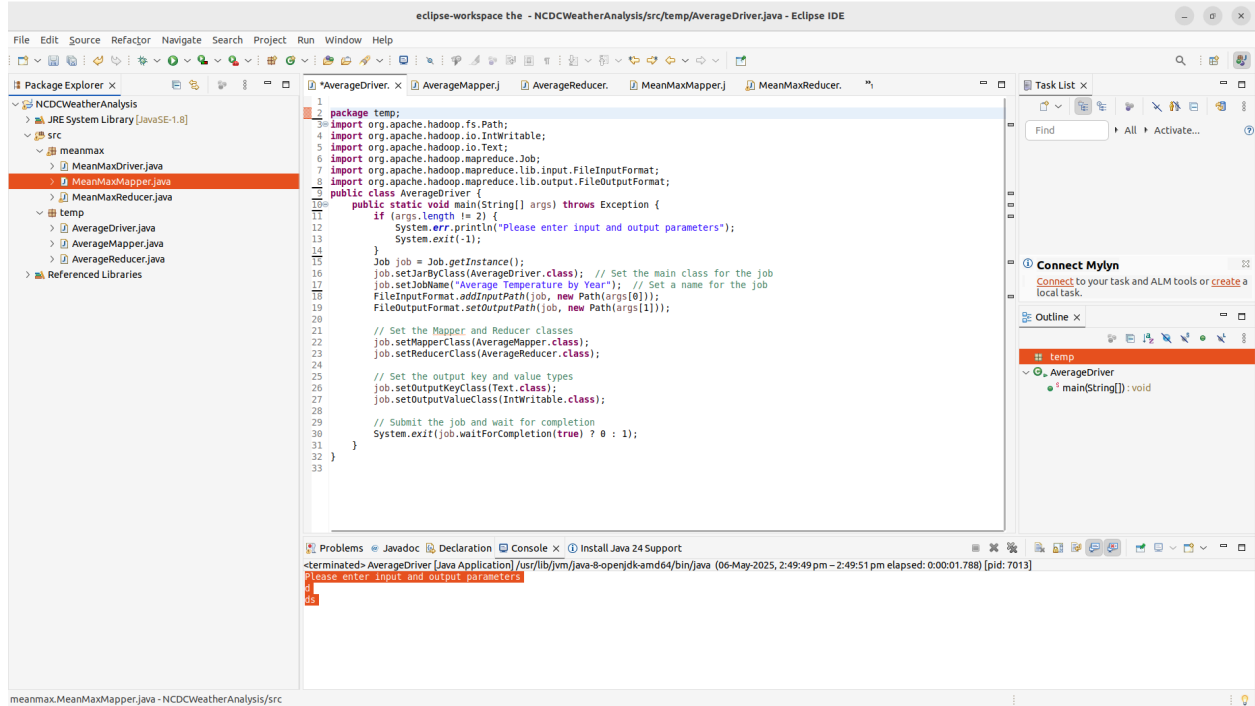
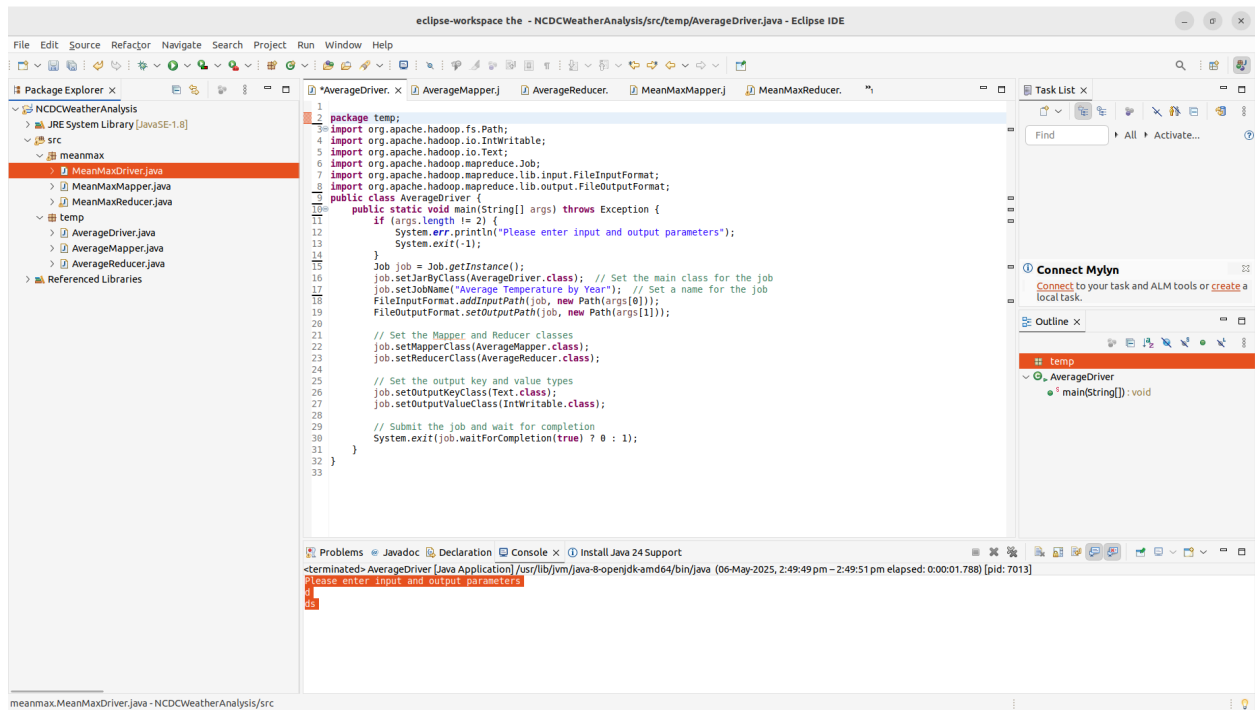
From the following link extract the weather data

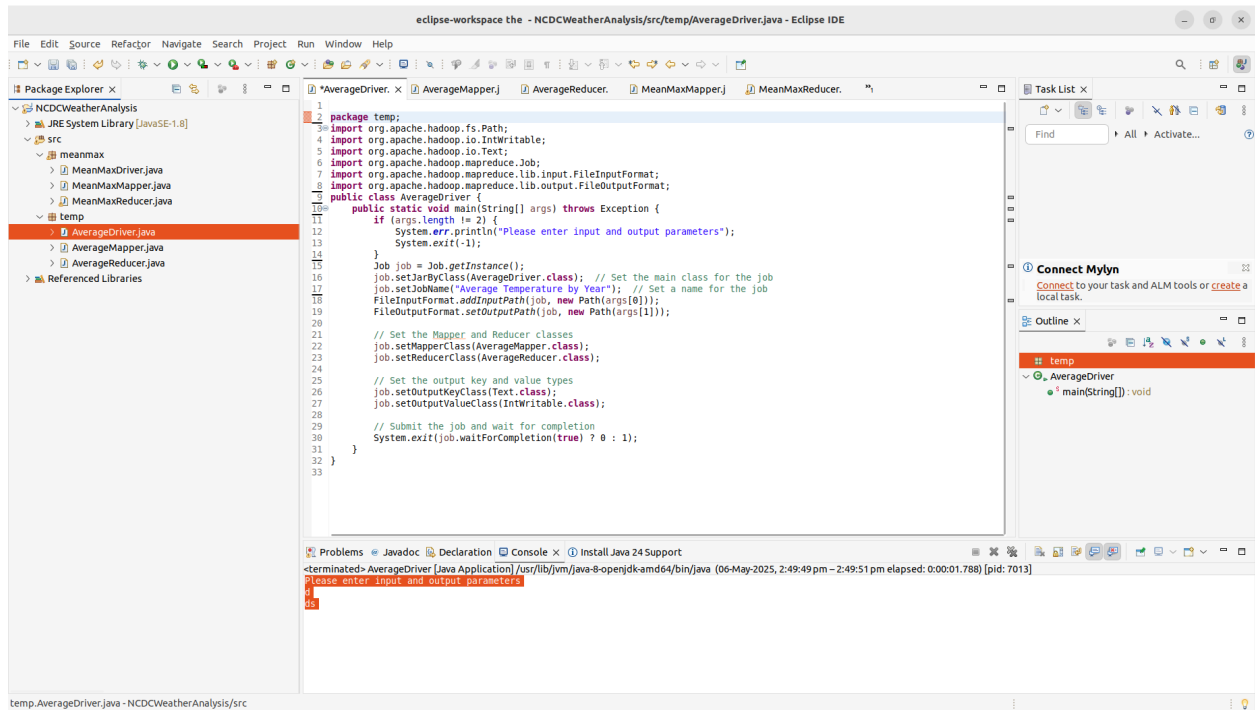
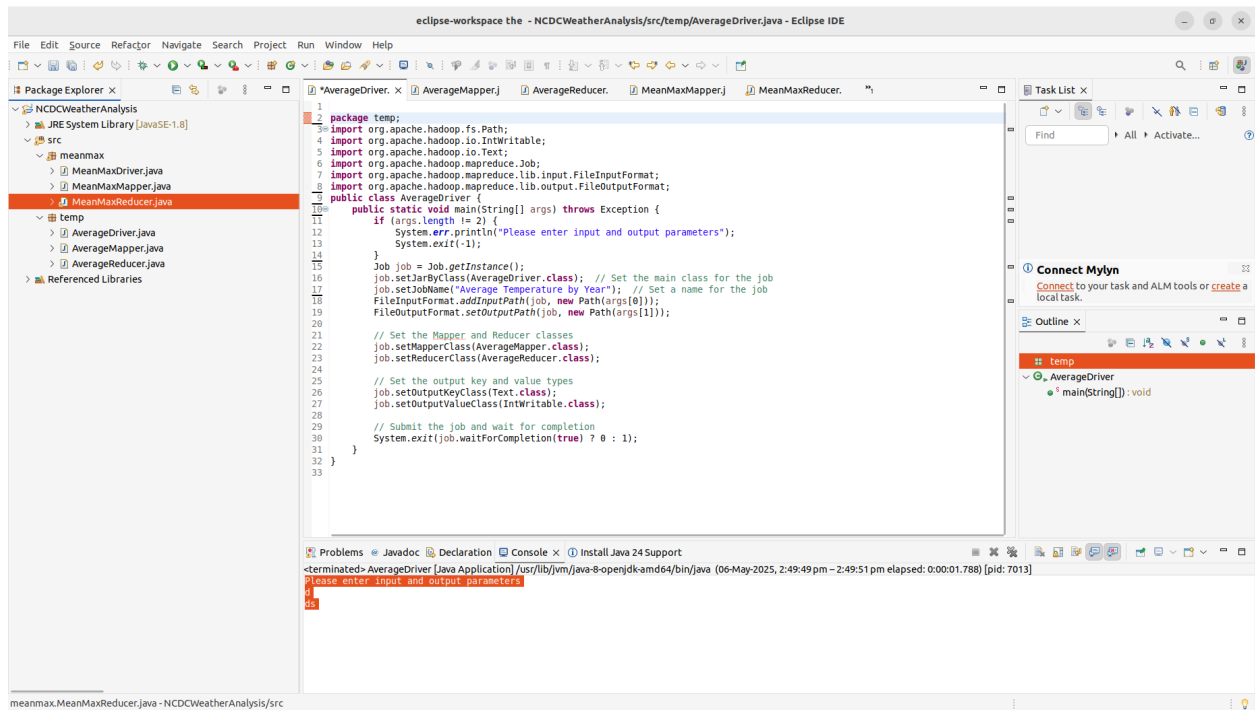
<https://github.com/tomwhite/hadoop-book/tree/master/input/ncdc/all>

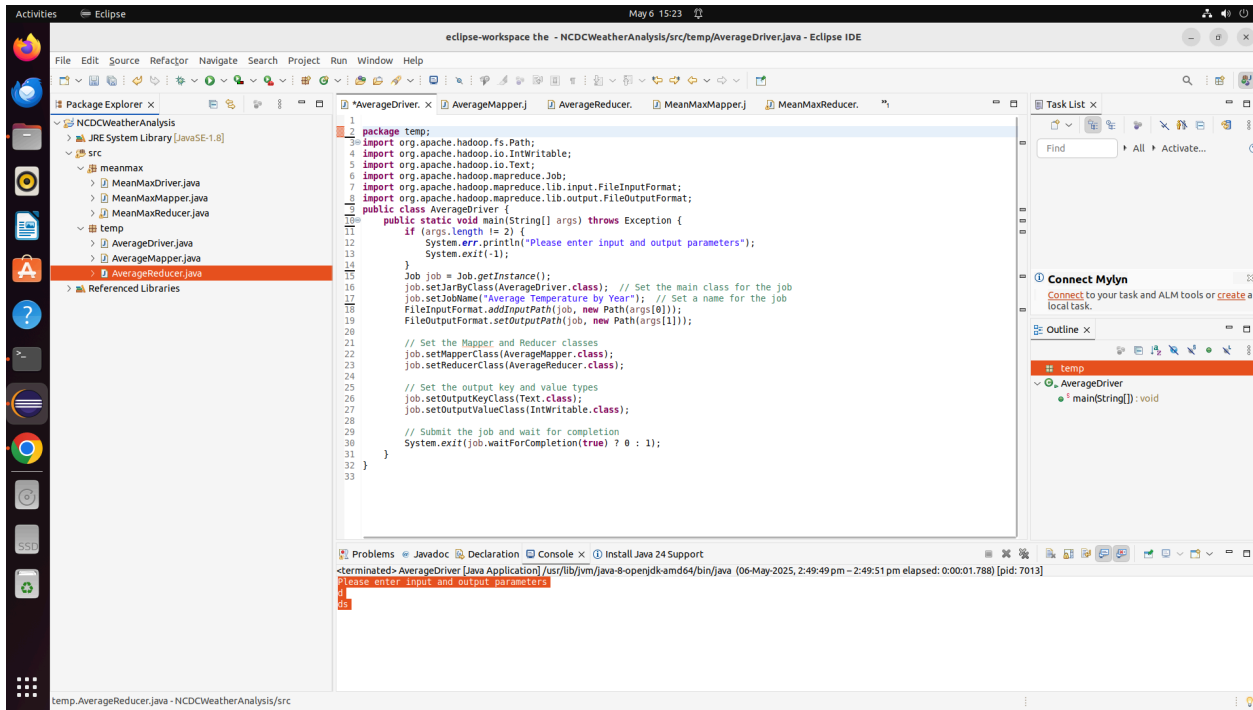
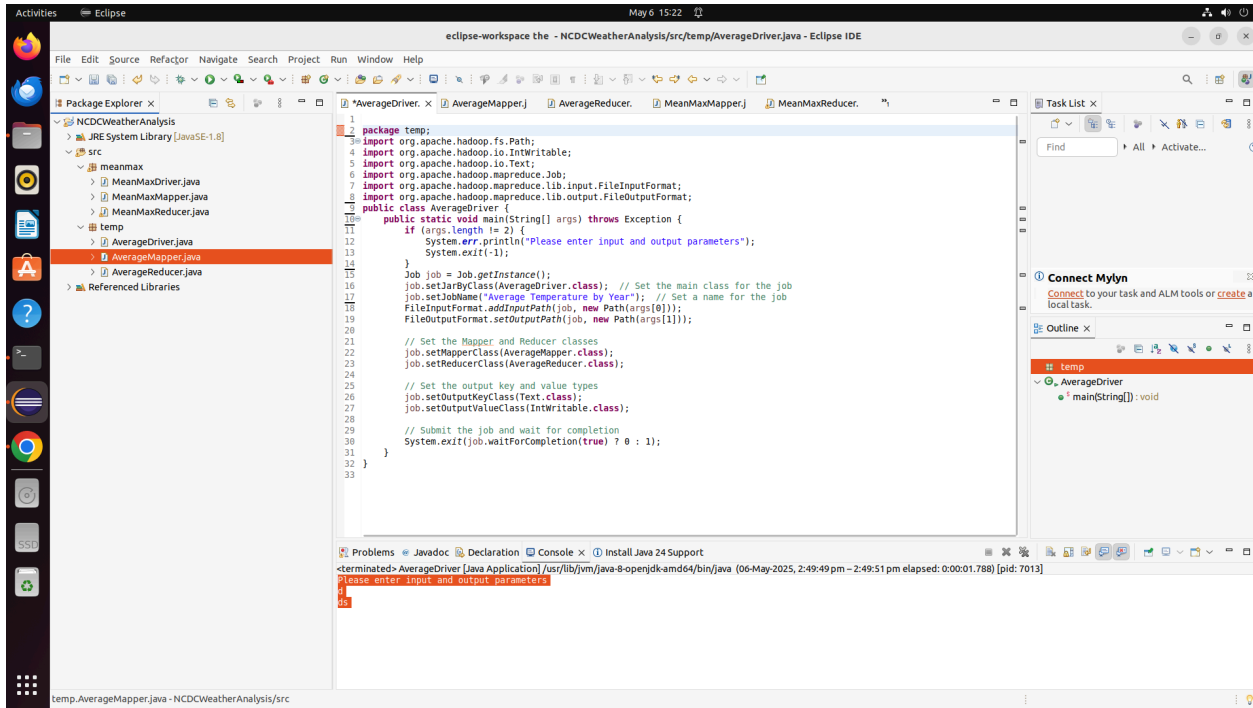
a) Create a MapReduce program to find average temperature for each year from NCDC data set.

b) find the mean max temperature for every month.

Code with screenshots:







```
Activities Terminal May 6 15:25
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC:~$ hdfs dfs -mkdir /ncdc_input
mkdir: call from bmscscse-HP-Elite-Tower-600-G9-Desktop-PC/127.0.0.1 to localhost:9000 failed on connection exception: java.net.ConnectException: Connection refused; For more details see: http://wiki.apa
che.org/hadoop/ConnectionRefused
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC:~$ start-dfs.sh
start-yarn.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [bmscscse-HP-Elite-Tower-600-G9-Desktop-PC]
Starting resourcemanager
Starting nodemanagers
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC:~$ jps
16032 SecondaryNameNode
15745 DataNode
15556 NameNode
7093 org.eclipse.equinox.launcher_1.6.1000.v20250227-1734.jar
16504 NodeManager
16937 Jps
16318 ResourceManager
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC:~$ hdfs dfs -mkdir /ncdc_input
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC:~$ git clone https://github.com/tomwhite/hadoop-book.git
fatal: destination path 'hadoop-book' already exists and is not an empty directory.
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC:~$ cd ~/Desktop/
hdfs dfs -put * /ncdc_input
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC:~/hadoop-book/input/ncdc/all$ hdfs dfs -ls /ncdc_input
Found 2 items
-rw-r--r-- 1 hadoop supergroup 73867 2025-05-06 15:11 /ncdc_input/1901.gz
-rw-r--r-- 1 hadoop supergroup 74105 2025-05-06 15:11 /ncdc_input/1902.gz
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC:~/hadoop-book/input/ncdc/all$ hadoop jar AverageTemp.jar temp.AverageDriver /ncdc_input /output_avg
JAR does not exist or is not a normal file: /home/hadoop/hadoop-book/input/ncdc/all/AverageTemp.jar
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC:~/hadoop-book/input/ncdc/all$ hadoop jar AverageTemp.jar temp.AverageDriver /ncdc_input /output_avg
JAR does not exist or is not a normal file: /home/hadoop/Desktop/AverageTemp.jar
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC:~/hadoop-book/input/ncdc/all$ hadoop jar ACFFrogC11800KZ9jzVbDLx00_3-HyTtQ3oZcp_TewZ0Sd65Seh1Eqf_umh0I-d7NXMSbB49Hx4HgJzFkZ56Ak8MTMYxYCNBkewr7yOGfL17dL8J3C8KkdJx9GASRZJ3pZCQKyTrpgSUCSU.pdf
file1.txt
NCDCWeather
NCDCWeather.jar
sample.txt
Tejswini.jar
WordCount.jar
WordCount.jar
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC:~/hadoop-book$ hadoop jar NCDCWeather.jar temp.AverageDriver /ncdc_input /output_avg
2025-05-06 15:14:57.391 INFO Impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2025-05-06 15:14:57.428 INFO Impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2025-05-06 15:14:57.429 INFO Impl.MetricsSystemImpl: JobTracker metrics system started
2025-05-06 15:14:57.485 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool Interface and execute your application with ToolRunner to remedy this.
2025-05-06 15:14:57.537 INFO Input.FileInputFormat: Total input files to process : 2
2025-05-06 15:14:57.553 INFO mapreduce.JobSubmitter: number of splits:2
2025-05-06 15:14:57.607 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1944593155_0001
2025-05-06 15:14:57.607 INFO mapreduce.JobSubmitter: Executing with tokens: []
2025-05-06 15:14:57.656 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
2025-05-06 15:14:57.657 INFO mapreduce.Job: Running job: job_local1944593155_0001
2025-05-06 15:14:57.657 INFO mapreduce.LocalJobRunner: OutputCommitter set to config null
2025-05-06 15:14:57.661 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-05-06 15:14:57.661 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup_temporary folders under output directory:false, ignore cleanup failures: false
2025-05-06 15:14:57.661 INFO mapreduce.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter
```

```
Activities Terminal May 6 15:26
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop
WordCount.jar
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC:~/hadoop-book$ hadoop jar NCDCWeather.jar temp.AverageDriver /ncdc_input /output_avg
2025-05-06 15:14:57.391 INFO Impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2025-05-06 15:14:57.428 INFO Impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2025-05-06 15:14:57.429 INFO Impl.MetricsSystemImpl: JobTracker metrics system started
2025-05-06 15:14:57.485 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool Interface and execute your application with ToolRunner to remedy this.
2025-05-06 15:14:57.537 INFO Input.FileInputFormat: Total input files to process : 2
2025-05-06 15:14:57.553 INFO mapreduce.JobSubmitter: number of splits:2
2025-05-06 15:14:57.607 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1944593155_0001
2025-05-06 15:14:57.607 INFO mapreduce.JobSubmitter: Executing with tokens: []
2025-05-06 15:14:57.656 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
2025-05-06 15:14:57.657 INFO mapreduce.Job: Running job: job_local1944593155_0001
2025-05-06 15:14:57.657 INFO mapreduce.LocalJobRunner: OutputCommitter set in config null
2025-05-06 15:14:57.661 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-05-06 15:14:57.661 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup_temporary folders under output directory:false, ignore cleanup failures: false
2025-05-06 15:14:57.661 INFO mapreduce.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter
2025-05-06 15:14:57.692 INFO mapreduce.LocalJobRunner: Waiting for map tasks
2025-05-06 15:14:57.692 INFO mapreduce.LocalJobRunner: Starting task: attempt_local1944593155_0001_m_0000000_0
2025-05-06 15:14:57.701 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-05-06 15:14:57.701 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup_temporary folders under output directory:false, ignore cleanup failures: false
2025-05-06 15:14:57.707 INFO mapreduce.Task: Using ResourceCalculatorProcessFree: []
2025-05-06 15:14:57.708 INFO mapreduce.MapTask: Processing split: hdfs://localhost:9000/ncdc_input/1902.gz:0:74105
2025-05-06 15:14:57.739 INFO mapreduce.MapTask: (EQUATOR) 0 kvt 26214396(104857584)
2025-05-06 15:14:57.739 INFO mapreduce.MapTask: soft limit at 33886000
2025-05-06 15:14:57.739 INFO mapreduce.MapTask: bufstart = 0; bufvoid = 104857600
2025-05-06 15:14:57.739 INFO mapreduce.MapTask: kvstart = 26214396; length = 6553600
2025-05-06 15:14:57.741 INFO mapreduce.MapTask: Map output collector class = org.apache.hadoop.mapreduce.MapTask$MapOutputBuffer
2025-05-06 15:14:57.754 INFO zlib.ZlibFactory: Successfully loaded & initialized native-zlib library
2025-05-06 15:14:57.754 INFO compress.CodecPool: Got brand-new decompressor [.gz]
2025-05-06 15:14:57.815 INFO mapreduce.LocalJobRunner:
2025-05-06 15:14:57.816 INFO mapreduce.MapTask: Starting flush of map output
2025-05-06 15:14:57.816 INFO mapreduce.MapTask: Spilling map output
2025-05-06 15:14:57.816 INFO mapreduce.MapTask: bufstart = 0; bufend = 59085; bufvoid = 104857600
2025-05-06 15:14:57.822 INFO mapreduce.MapTask: kvstart = 26214396(104857584); kvend = 26188140(104752500); length = 26257/6553600
2025-05-06 15:14:57.826 INFO mapreduce.Task: Task:attempt_local1944593155_0001_m_0000000_0 is done. And is in the process of committing
2025-05-06 15:14:57.828 INFO mapreduce.LocalJobRunner: map
2025-05-06 15:14:57.829 INFO mapreduce.Task: Task 'attempt_local1944593155_0001_m_0000000_0' done.
2025-05-06 15:14:57.831 INFO mapreduce.Task: Final Counters for attempt_local1944593155_0001_m_0000000_0: Counters: 23
File System Counters
FILE: Number of bytes read=7873
FILE: Number of bytes written=719826
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=74105
HDFS: Number of bytes written=0
HDFS: Number of read operations=5
HDFS: Number of large read operations=0
HDFS: Number of write operations=1
HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
Map input records=6565
Map output records=6565
Map output bytes=59085
```

```
Activities Terminal May 6 15:26
hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop

Map-Reduce Framework
  Map input records=6565
  Map output records=6565
  Map output bytes=59085
  Map output materialized bytes=72221
  Input split bytes=105
  Combine input records=0
  Spilled records=6565
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=0
  Total committed heap usage (bytes)=526385152

File Input Format Counters
  Bytes Read=74105
2025-05-06 15:14:57,831 INFO mapred.LocalJobRunner: Finishing task: attempt_local1944593155_0001_m_000000_0
2025-05-06 15:14:57,831 INFO mapred.LocalJobRunner: Starting task: attempt_local1944593155_0001_m_000001_0
2025-05-06 15:14:57,832 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-05-06 15:14:57,832 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2025-05-06 15:14:57,832 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
2025-05-06 15:14:57,833 INFO mapred.MapTask: Processing split: hdfs://localhost:9000/ncdc_input/1901.gz:0:73867
2025-05-06 15:14:57,859 INFO mapred.MapTask: (EQUATOR) 0 kvl 26214396(104857584)
2025-05-06 15:14:57,859 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100
2025-05-06 15:14:57,859 INFO mapred.MapTask: sort.limit at 8386000
2025-05-06 15:14:57,859 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600
2025-05-06 15:14:57,859 INFO mapred.MapTask: kvstart = 26214396; length = 6553600
2025-05-06 15:14:57,860 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2025-05-06 15:14:57,887 INFO mapred.LocalJobRunner: map
2025-05-06 15:14:57,887 INFO mapred.MapTask: Starting flush of map output
2025-05-06 15:14:57,887 INFO mapred.MapTask: Spilling map output
2025-05-06 15:14:57,887 INFO mapred.MapTask: bufstart = 0; bufend = 59076; bufvoid = 104857600
2025-05-06 15:14:57,887 INFO mapred.MapTask: kvstart = 26214396(104857584); kvoid = 26188144(104752576); length = 26253/6553600
2025-05-06 15:14:57,890 INFO mapred.MapTask: Flushed spill 0
2025-05-06 15:14:57,892 INFO mapred.Task: Task:attempt_local1944593155_0001_m_000001_0 is done. And is in the process of committing
2025-05-06 15:14:57,894 INFO mapred.LocalJobRunner: map
2025-05-06 15:14:57,894 INFO mapred.Task: Task 'attempt_local1944593155_0001_m_000001_0' done.
2025-05-06 15:14:57,895 INFO mapred.Task: Final Counters for attempt_local1944593155_0001_m_000001_0: Counters: 23
File System Counters
  FILE: Number of bytes read=8102
  FILE: Number of bytes written=792068
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=147972
  HDFS: Number of bytes written=0
  HDFS: Number of read operations=7
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=1
  HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
  Map input records=6565
  Map output records=6564
  Map output bytes=59076
  Map output materialized bytes=72210
  Input split bytes=105
  Combine input records=0
  Spilled records=6564
```

```
Activities Terminal May 6 15:26
hadoop@bmsccse-HP-Elite-Tower-600-G9-Desktop-PC: ~/Desktop

Input split bytes=105
Combine input records=0
Spilled records=6564
Failed Shuffles=0
Merged Map outputs=0
GC time elapsed (ms)=0
Total committed heap usage (bytes)=526385152

File Input Format Counters
  Bytes Read=73867
2025-05-06 15:14:57,895 INFO mapred.LocalJobRunner: Finishing task: attempt_local1944593155_0001_m_000001_0
2025-05-06 15:14:57,895 INFO mapred.LocalJobRunner: map task executor complete.
2025-05-06 15:14:57,897 INFO mapred.LocalJobRunner: Waiting for reduce tasks
2025-05-06 15:14:57,897 INFO mapred.LocalJobRunner: Starting task: attempt_local1944593155_0001_r_000000_0
2025-05-06 15:14:57,904 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-05-06 15:14:57,904 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2025-05-06 15:14:57,904 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
2025-05-06 15:14:57,905 INFO mapred.ReduceTask: Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@8777f0e5
2025-05-06 15:14:57,906 WARN Impl.MetricsSystemImpl: JobTracker metrics system already initialized
2025-05-06 15:14:57,915 INFO reduce.MergeManagerImpl: MergerManager: memoryLimit=5827985408, maxSingleShuffleLimit=1456996352, mergeThreshold=3846470400, ioSortFactor=10, memToMemMergeOutputsThreshold=10
2025-05-06 15:14:57,916 INFO reduce.EventFetcher: attempt_local1944593155_0001_r_000000_0 Thread started: EventFetcher for fetching Map Completion Events
2025-05-06 15:14:57,927 INFO reduce.LocalFetcher: localFetcher#1 about to shuffle output of map attempt_local1944593155_0001_m_000001_0 decomp: 72206 len: 72210 to MEMORY
2025-05-06 15:14:57,928 INFO reduce.InMemoryMapOutput: Read 72206 bytes from map-output for attempt_local1944593155_0001_m_000001_0
2025-05-06 15:14:57,928 INFO reduce.MergeManagerImpl: closeInMemoryFile -> map-output of size: 72206, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->72206
2025-05-06 15:14:57,929 INFO reduce.LocalFetcher: localFetcher#1 about to shuffle output of map attempt_local1944593155_0001_m_000001_0 decomp: 72217 len: 72221 to MEMORY
2025-05-06 15:14:57,930 INFO reduce.InMemoryMapOutput: Read 72217 bytes from map-output for attempt_local1944593155_0001_m_000001_0
2025-05-06 15:14:57,930 INFO reduce.MergeManagerImpl: closeInMemoryFile -> map-output of size: 72217, inMemoryMapOutputs.size() -> 2, commitMemory -> 72206, usedMemory ->144423
2025-05-06 15:14:57,930 INFO reduce.EventFetcher: EventFetcher is interrupted.. Returning
2025-05-06 15:14:57,930 WARN io.ReadAheadPool: Failed readahead on file
EBADF: Bad file descriptor
at org.apache.hadoop.io.nativeio.NativeIO$POSIX.posix_fadvise(Native Method)
at org.apache.hadoop.io.nativeio.NativeIO$POSIX.posixFadviseIfPossible(NativeIO.java:419)
at org.apache.hadoop.io.nativeio.NativeIO$POSIX$CacheManipulator.posixFadviseIfPossible(NativeIO.java:296)
at org.apache.hadoop.io.ReadAheadPool$ReadAheadRequestImpl.run(ReadAheadPool.java:226)
at java.base/java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1128)
at java.base/java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:628)
at java.base/java.lang.Thread.run(Thread.java:829)
2025-05-06 15:14:57,930 INFO mapred.LocalJobRunner: 2 / 2 copied.
2025-05-06 15:14:57,931 INFO reduce.MergeManagerImpl: finalMerge called with 2 in-memory map-outputs and 0 on-disk map-outputs
2025-05-06 15:14:57,933 INFO mapred.Merger: Merging 2 sorted segments
2025-05-06 15:14:57,933 INFO mapred.Merger: Down to the last merge-pass, with 2 segments left of total size: 144409 bytes
2025-05-06 15:14:57,938 INFO reduce.MergeManagerImpl: Merged 2 segments, 144423 bytes to disk to satisfy reduce memory limit
2025-05-06 15:14:57,938 INFO reduce.MergeManagerImpl: Merging 1 files, 144425 bytes from disk
2025-05-06 15:14:57,938 INFO reduce.MergeManagerImpl: Merging 0 segments, 0 bytes from memory into reduce
2025-05-06 15:14:57,938 INFO mapred.Merger: Merging 1 sorted segments
2025-05-06 15:14:57,939 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 144414 bytes
2025-05-06 15:14:57,939 INFO mapred.LocalJobRunner: 2 / 2 copied.
2025-05-06 15:14:57,968 INFO Configuration.deprecation: mapred.skip.on is deprecated. Instead, use mapreduce.job.skiprecords
2025-05-06 15:14:58,022 INFO mapred.Task: Task:attempt_local1944593155_0001_r_000000_0 is done. And is in the process of committing
2025-05-06 15:14:58,023 INFO mapred.LocalJobRunner: 2 / 2 copied.
2025-05-06 15:14:58,023 INFO mapred.Task: Task attempt_local1944593155_0001_r_000000_0 is allowed to commit now
2025-05-06 15:14:58,038 INFO output.FileOutputCommitter: Saved output of task 'attempt_local1944593155_0001_r_000000_0' to hdfs://localhost:9000/output_avg
2025-05-06 15:14:58,038 INFO mapred.LocalJobRunner: reduce > reduce
2025-05-06 15:14:58,038 INFO mapred.Task: Task 'attempt_local1944593155_0001_r_000000_0' done.
2025-05-06 15:14:58,038 INFO mapred.Task: Final Counters for attempt_local1944593155_0001_r_000000_0: Counters: 30
File System Counters
  FILE: Number of bytes read=297022
```

```
Activities Terminal May 6 15:27
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC: ~/Desktop

2025-05-06 15:14:58,038 INFO mapred.Task: Final Counters for attempt_local1944593155_0001_r_000000_0: Counters: 30
File System Counters
  FILE: Number of bytes read=297022
  FILE: Number of bytes written=936493
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=147972
  HDFS: Number of bytes written=16
  HDFS: Number of read operations=12
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=3
  HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
  Combine input records=0
  Combine output records=0
  Reduce input groups=2
  Reduce shuffle bytes=144431
  Reduce input records=13129
  Reduce output records=2
  Spilled Records=13129
  Shuffled Maps =2
  Failed Shuffles=0
  Merged Map outputs=2
  GC time elapsed (ms)=2
  Total committed heap usage (bytes)=633339904
Shuffle Errors
  BAD_ID=0
  CONNECTION=0
  IO_ERROR=0
  WRONG_LENGTH=0
  WRONG_MAP=0
  WRONG_REDUCE=0
File Output Format Counters
  Bytes Written=16
2025-05-06 15:14:58,038 INFO mapred.LocalJobRunner: Finishing task: attempt_local1944593155_0001_r_000000_0
2025-05-06 15:14:58,039 INFO mapred.LocalJobRunner: reduce task executor complete.
2025-05-06 15:14:58,659 INFO mapreduce.Job: Job: job_local1944593155_0001 running in uber mode : false
2025-05-06 15:14:58,659 INFO mapreduce.Job: map 100% reduce 100%
2025-05-06 15:14:58,660 INFO mapreduce.Job: Job: job_local1944593155_0001 completed successfully
2025-05-06 15:14:58,664 INFO mapreduce.Job: Counters: 36
File System Counters
  FILE: Number of bytes read=312997
  FILE: Number of bytes written=2448387
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=370049
  HDFS: Number of bytes written=16
  HDFS: Number of read operations=24
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=5
  HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
  Map input records=13130
  Map output records=13129
  Reduce output records=12
  Spilled Records=26258
  Shuffled Maps =2
  Failed Shuffles=0
  Merged Map outputs=2
  GC time elapsed (ms)=3
  Total committed heap usage (bytes)=1686110288
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=147972
File Output Format Counters
  Bytes Written=72
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -ls /output_avg
Found 2 items
-rw-r--r-- 1 hadoop supergroup 0 2025-05-06 15:14 /output_avg/_SUCCESS
-rw-r--r-- 1 hadoop supergroup 16 2025-05-06 15:14 /output_avg/part-r-00000
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
1901 46
1902 21
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -ls /output_avg
Found 2 items
-rw-r--r-- 1 hadoop supergroup 0 2025-05-06 15:14 /output_avg/_SUCCESS
-rw-r--r-- 1 hadoop supergroup 16 2025-05-06 15:14 /output_avg/part-r-00000
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
1901 46
1902 21
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
01 4
02 1
03 6
04 34
05 89
06 143
07 182
08 172
09 123
10 73
11 21
12 3
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
1901 46
1902 21
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
1901 46
1902 21
```

```
Activities Terminal May 6 15:27
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC: ~/Desktop

Reduce output records=12
Spilled Records=26258
Shuffled Maps =2
Failed Shuffles=0
Merged Map outputs=2
GC time elapsed (ms)=3
Total committed heap usage (bytes)=1686110288
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=147972
File Output Format Counters
  Bytes Written=72
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -ls /output_avg
Found 2 items
-rw-r--r-- 1 hadoop supergroup 0 2025-05-06 15:14 /output_avg/_SUCCESS
-rw-r--r-- 1 hadoop supergroup 16 2025-05-06 15:14 /output_avg/part-r-00000
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
1901 46
1902 21
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -ls /output_avg
Found 2 items
-rw-r--r-- 1 hadoop supergroup 0 2025-05-06 15:14 /output_avg/_SUCCESS
-rw-r--r-- 1 hadoop supergroup 16 2025-05-06 15:14 /output_avg/part-r-00000
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
1901 46
1902 21
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
01 4
02 1
03 6
04 34
05 89
06 143
07 182
08 172
09 123
10 73
11 21
12 3
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
1901 46
1902 21
hadoop@bmsccse-HP-Ellite-Tower-600-G9-Desktop-PC:~/Desktop$ hdfs dfs -cat /output_avg/part-r-00000
1901 46
1902 21
```

