

```
Activities Terminal Mon 23:22
hadoop@IT-SSL023: ~/hadoop
File Edit View Search Terminal Help
hadoop@IT-SSL023:~/hadoop$ ls
AddTwoMatrix.class  etc          k_means      KMeans.java  LICENSE.txt  matrix.class  pig-0.3.0  sbin
bin                 include      KMeans       lib          logs         matrix.java  Point.class share
centroid.txt        input.txt    k_means.java libexec       matrix       NOTICE.txt  README.txt  wordcount.java
hadoop@IT-SSL023:~/hadoop$ javac KMeans.java
hadoop@IT-SSL023:~/hadoop$ jar -cvf KMeans.jar -c /home/hadoop/hadoop/KMeans/ .
```

```
Activities Terminal Mon 23:23
hadoop@IT-SSL023: ~/hadoop
File Edit View Search Terminal Help
hadoop@IT-SSL023:~/hadoop$ ls
AddTwoMatrix.class  include      'KMeans$AvgMapper.class'  k_means.java  LICENSE.txt  matrix.java  README.txt
bin                 input.txt    'KMeans$AvgReducer.class' KMeans.java   logs         NOTICE.txt  sbin
centroid.txt        k_means     KMeans.class              lib           matrix       pig-0.3.0    share
etc                 KMeans      KMeans.jar                libexec       matrix.class Point.class  wordcount.java
hadoop@IT-SSL023:~/hadoop$ hdfs dfs -mkdir /EX8
hadoop@IT-SSL023:~/hadoop$ hdfs dfs -put input.txt /EX8
hadoop@IT-SSL023:~/hadoop$ hdfs dfs -put centroid.txt /EX8
hadoop@IT-SSL023:~/hadoop$
```

```
Activities Terminal Mon 23:24
hadoop@IT-SSL023: ~/hadoop

File Edit View Search Terminal Help

hadoop@IT-SSL023:~/hadoop$ ls
AddTwoMatrix.class  include  'KMeans$AvgMapper.class'  k_means.java  LICENSE.txt  matrix.java  README.txt
bin                 input.txt  'KMeans$AvgReducer.class'  KMeans.java   logs        NOTICE.txt  sbin
centroid.txt        k_means   KMeans.class               lib           matrix      pig-0.3.0    share
etc                 KMeans    KMeans.jar                 libexec       matrix.class  Point.class  wordcount.java

hadoop@IT-SSL023:~/hadoop$ hdfs dfs -mkdir /EX8
hadoop@IT-SSL023:~/hadoop$ hdfs dfs -put input.txt /EX8
hadoop@IT-SSL023:~/hadoop$ hdfs dfs -put centroid.txt /EX8
hadoop@IT-SSL023:~/hadoop$ hadoop jar KMeans.jar KMeans /EX8/input.txt /EX8/centroid.txt /output
```

```
Activities Terminal Mon 23:25
hadoop@IT-SSL023: ~/hadoop

File Edit View Search Terminal Help

hadoop@IT-SSL023:~/hadoop$ ls
AddTwoMatrix.class  include  'KMeans$AvgMapper.class'  k_means.java  LICENSE.txt  matrix.java  README.txt
bin                 input.txt  'KMeans$AvgReducer.class'  KMeans.java   logs        NOTICE.txt  sbin
centroid.txt        k_means   KMeans.class               lib           matrix      pig-0.3.0    share
etc                 KMeans    KMeans.jar                 libexec       matrix.class  Point.class  wordcount.java

hadoop@IT-SSL023:~/hadoop$ hdfs dfs -mkdir /EX8
hadoop@IT-SSL023:~/hadoop$ hdfs dfs -put input.txt /EX8
hadoop@IT-SSL023:~/hadoop$ hdfs dfs -put centroid.txt /EX8
hadoop@IT-SSL023:~/hadoop$ hadoop jar KMeans.jar KMeans /EX8/input.txt /EX8/centroid.txt /Output
23/02/20 23:24:56 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
23/02/20 23:24:56 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
23/02/20 23:24:58 INFO input.FileInputFormat: Total input paths to process : 1
23/02/20 23:24:59 INFO mapreduce.JobSubmitter: number of splits:1
23/02/20 23:25:00 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1676912324879_0007
23/02/20 23:25:00 INFO impl.YarnClientImpl: Submitted application application_1676912324879_0007
23/02/20 23:25:00 INFO mapreduce.Job: The url to track the job: http://IT-SSL023:8088/proxy/application_1676912324879_0007/
23/02/20 23:25:00 INFO mapreduce.Job: Running job: job_1676912324879_0007
```

```
Mon 23:25
hadoop@IT-SSL023: ~/hadoop

Total time spent by all reduce tasks (ms)=1636
Total vcore-milliseconds taken by all map tasks=1582
Total vcore-milliseconds taken by all reduce tasks=1636
Total megabyte-milliseconds taken by all map tasks=1619968
Total megabyte-milliseconds taken by all reduce tasks=1675264

Map-Reduce Framework
  Map input records=450
  Map output records=450
  Map output bytes=14400
  Map output materialized bytes=15306
  Input split bytes=100
  Combine input records=0
  Combine output records=0
  Reduce input groups=3
  Reduce shuffle bytes=15306
  Reduce input records=450
  Reduce output records=3
  Spilled Records=900
  Shuffled Maps =1
  Failed Shuffles=0
  Merged Map outputs=1
  GC time elapsed (ms)=66
  CPU time spent (ms)=1010
  Physical memory (bytes) snapshot=441778176
  Virtual memory (bytes) snapshot=3843379200
  Total committed heap usage (bytes)=311427072

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=5400
File Output Format Counters
  Bytes Written=109
hadoop@IT-SSL023:~/hadoop$
```

```
Mon 23:25
hadoop@IT-SSL023: ~/hadoop

Total megabyte-milliseconds taken by all reduce tasks=1675264

Map-Reduce Framework
  Map input records=450
  Map output records=450
  Map output bytes=14400
  Map output materialized bytes=15306
  Input split bytes=100
  Combine input records=0
  Combine output records=0
  Reduce input groups=3
  Reduce shuffle bytes=15306
  Reduce input records=450
  Reduce output records=3
  Spilled Records=900
  Shuffled Maps =1
  Failed Shuffles=0
  Merged Map outputs=1
  GC time elapsed (ms)=66
  CPU time spent (ms)=1010
  Physical memory (bytes) snapshot=441778176
  Virtual memory (bytes) snapshot=3843379200
  Total committed heap usage (bytes)=311427072

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=5400
File Output Format Counters
  Bytes Written=109
hadoop@IT-SSL023:~/hadoop$ hdfs dfs -cat /Output/part*0000
1.8943768115942032 5.360985507246374
4.590894230769228 6.605105769230767
5.004184115523467 3.468642599277982
hadoop@IT-SSL023:~/hadoop$
```