```
Last login: Sat Nov 14 14:38:26 on ttys000
(base) lejiahu@b-10-27-95-32 ~ % ssh -Y lh2514@dumbo.es.its.nyu.edu

Password:
Last login: Sat Nov 14 14:27:55 2020 from b-10-27-95-32.dynapool.vpn.nyu.edu
Rocks 6.2 (SideWinder)
Profile built 16:54 24-Aug-2015

Kickstarted 16:58 24-Aug-2015
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
   THIS NYU-OPERATED FACILITY MAY BE ACCESSED ONLY BY AUTHORIZED USERS.
                  UNAUTHORIZED PERSONS ... TYPE 'exit' NOW
  See https://wikis.nyu.edu/display/NYUHPC/NYU+HPC+Access+Policy for details
  of the access policy. Questions regarding the access policy should be
  directed (by e-mail) to security@nyu.edu or (by phone) to 212-998-3333
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                WELCOME TO
*****************************************************************************
     ___           ___           ___           ___           ___
    /  /::\        /__/\         /__/\         /  /::\        /  /\
   /  /:/\:\       \  \:\        |  |::\       /  /:/\:\      /  /::\
  /  /:/  \:\       \  \:\       |  |:|:\     /  /:/~/::\    /  /:/\:\
 /__/:/ \__\:|  ___  \  \:\    __|__|:|\:\   /__/:/ /:/\:|  /  /:/  \:\
 \  \:\ /  /:/ /__/\  \__\:\  /__/::::| \:\  \  \:\/:/~/:/ /__/:/ \__\:\
  \  \:\  /:/  \  \:\ /  /:/  \  \:\~~\__\/   \  \::/ /:/  \  \:\ /  /:/
   \  \:\/:/    \  \:\  /:/    \  \:\          \  \:\/:/    \  \:\  /:/
    \  \::/      \  \:\/:/      \  \:\          \  \::/      \  \:\/:/
     \__\/        \  \::/        \  \:\          \__\/        \  \::/
                   \__\/          \__\/                       \__\/

*****************************************************************************
            https://wikis.nyu.edu/display/NYUHPC/Clusters+-+Dumbo


*****************************************************************************
[lh2514@login-2-1 ~]$ vi IntSumReducer.java
[lh2514@login-2-1 ~]$ vi LineCountMapper.java
[lh2514@login-2-1 ~]$ vi LinesCountRunner.java
[lh2514@login-2-1 ~]$ javac -classpath "$(yarn classpath)" -d . LineCountMapper.java
[lh2514@login-2-1 ~]$ javac -classpath "$(yarn classpath)" -d . IntSumReducer.java
[lh2514@login-2-1 ~]$ javac -classpath "$(yarn classpath)":. -d . LinesCountRunner.java
[lh2514@login-2-1 ~]$ jar -cvf LinesCount.jar *.class
added manifest
adding: Airbnb.class(in = 1428) (out= 789)(deflated 44%)
adding: AirbnbMapper.class(in = 2269) (out= 1047)(deflated 53%)
adding: AirbnbReducer.class(in = 1694) (out= 714)(deflated 57%)
adding: IntSumReducer.class(in = 1694) (out= 714)(deflated 57%)
adding: LineCountMapper.class(in = 1478) (out= 596)(deflated 59%)
adding: LinesCountRunner.class(in = 1430) (out= 793)(deflated 44%)
adding: Main.class(in = 1461) (out= 872)(deflated 40%)
adding: SelectFieldsMapper.class(in = 255) (out= 198)(deflated 22%)
adding: SelectFieldsReducer.class(in = 258) (out= 196)(deflated 24%)
adding: SelectFieldsRunner.class(in = 1819) (out= 964)(deflated 47%)
adding: WordCount.class(in = 1440) (out= 789)(deflated 45%)
adding: WordCountMapper.class(in = 2096) (out= 962)(deflated 54%)
adding: WordCountReducer.class(in = 1700) (out= 717)(deflated 57%)
[lh2514@login-2-1 ~]$ ls
Airbnb.class              myNewDir
Airbnb.jar                NYC_Mapper.py
Airbnb.java               NYC_Reducer.py
AirbnbMapper.class        NYPD_Complaint_Data_Historic.csv
AirbnbMapper.java         output_11
AirbnbReducer.class       output_hw3.txt
AirbnbReducer.java        part-r-00000
CleanMapper.py            project
CleanReducer.py           SelectFieldsMapper.class
com                       SelectFieldsMapper.java
data.csv                  SelectFieldsReducer.class
hw1                       SelectFieldsReducer.java
input.csv                 SelectFieldsRunner.class
input_hw3.txt             SelectFieldsRunner.java
IntSumReducer.class       smallWeather1.txt
IntSumReducer.java        WordCount.class
LineCountMapper.class     WordCount.jar
LineCountMapper.java      WordCount.java
LinesCount.jar            WordCountMapper.class
LinesCountRunner.class    WordCountMapper.class.save
LinesCountRunner.java     WordCountMapper.java
Main.class                WordCountReducer.class
Main.java                 WordCountReducer.java
myDir                     Word_output
[lh2514@login-2-1 ~]$ hdfs dfs -put NYPD_Complaint_Data_Historic.csv /user/lh2514/class1
put: `/user/lh2514/class1/NYPD_Complaint_Data_Historic.csv': File exists
[lh2514@login-2-1 ~]$ hadoop jar LinesCount.jar LinesCountRunner /user/lh2514/class1/NYPD_Complaint_Data_Historic.csv /user/lh2514/class1/output_NYC
20/11/14 14:43:51 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
20/11/14 14:43:51 INFO mapreduce.JobSubmitter: Cleaning up the staging area /user/lh2514/.staging/job_1604346392376_2024
20/11/14 14:43:51 WARN security.UserGroupInformation: PriviledgedActionException as:lh2514 (auth:SIMPLE) cause:org.apache.hadoop.mapreduce.lib.input.InvalidInputException: Input path does not e
xist: hdfs://dumbo/user/lh2514/NYPD_Complaint_Data_Historic.csv
Exception in thread "main" org.apache.hadoop.mapreduce.lib.input.InvalidInputException: Input path does not exist: hdfs://dumbo/user/lh2514/NYPD_Complaint_Data_Historic.csv
        at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.singleThreadedListStatus(FileInputFormat.java:323)
        at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.listStatus(FileInputFormat.java:265)
        at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.getSplits(FileInputFormat.java:387)
        at org.apache.hadoop.mapreduce.JobSubmitter.writeNewSplits(JobSubmitter.java:305)
        at org.apache.hadoop.mapreduce.JobSubmitter.writeSplits(JobSubmitter.java:322)
        at org.apache.hadoop.mapreduce.JobSubmitter.submitJobInternal(JobSubmitter.java:200)
        at org.apache.hadoop.mapreduce.Job$10.run(Job.java:1307)
        at org.apache.hadoop.mapreduce.Job$10.run(Job.java:1304)
        at java.security.AccessController.doPrivileged(Native Method)
        at javax.security.auth.Subject.doAs(Subject.java:422)
        at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1924)
        at org.apache.hadoop.mapreduce.Job.submit(Job.java:1304)
        at org.apache.hadoop.mapreduce.Job.waitForCompletion(Job.java:1325)
        at LinesCountRunner.main(LinesCountRunner.java:27)
        at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
        at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
        at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
        at java.lang.reflect.Method.invoke(Method.java:498)
        at org.apache.hadoop.util.RunJar.run(RunJar.java:226)
        at org.apache.hadoop.util.RunJar.main(RunJar.java:141)
[lh2514@login-2-1 ~]$ hadoop jar LinesCount.jar LinesCountRunner /user/lh2514/class1/NYPD_Complaint_Data_Historic.csv /user/lh2514/class1/output_NYC
20/11/14 14:44:55 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
20/11/14 14:44:55 INFO mapreduce.JobSubmitter: Cleaning up the staging area /user/lh2514/.staging/job_1604346392376_2025
20/11/14 14:44:55 WARN security.UserGroupInformation: PriviledgedActionException as:lh2514 (auth:SIMPLE) cause:org.apache.hadoop.mapreduce.lib.input.InvalidInputException: Input path does not e
xist: hdfs://dumbo/user/lh2514/NYPD_Complaint_Data_Historic.csv
Exception in thread "main" org.apache.hadoop.mapreduce.lib.input.InvalidInputException: Input path does not exist: hdfs://dumbo/user/lh2514/NYPD_Complaint_Data_Historic.csv
        at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.singleThreadedListStatus(FileInputFormat.java:323)
        at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.listStatus(FileInputFormat.java:265)
        at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.getSplits(FileInputFormat.java:387)
        at org.apache.hadoop.mapreduce.JobSubmitter.writeNewSplits(JobSubmitter.java:305)
        at org.apache.hadoop.mapreduce.JobSubmitter.writeSplits(JobSubmitter.java:322)
        at org.apache.hadoop.mapreduce.JobSubmitter.submitJobInternal(JobSubmitter.java:200)
        at org.apache.hadoop.mapreduce.Job$10.run(Job.java:1307)
        at org.apache.hadoop.mapreduce.Job$10.run(Job.java:1304)
        at java.security.AccessController.doPrivileged(Native Method)
        at javax.security.auth.Subject.doAs(Subject.java:422)
        at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1924)
        at org.apache.hadoop.mapreduce.Job.submit(Job.java:1304)
```

```
        at org.apache.hadoop.mapreduce.Job.waitForCompletion(Job.java:1325)
        at LinesCountRunner.main(LinesCountRunner.java:27)
        at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
        at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
        at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
        at java.lang.reflect.Method.invoke(Method.java:498)
        at org.apache.hadoop.util.RunJar.run(RunJar.java:226)
        at org.apache.hadoop.util.RunJar.main(RunJar.java:141)
[lh2514@login-2-1 ~]$ hdfs dfs -ls
Found 10 items
drwxrwx---+  - lh2514 users          0 2020-11-13 19:00 .Trash
drwxr-xr-x+  - lh2514 users          0 2020-11-12 09:15 .sparkStaging
drwx------+  - lh2514 users          0 2020-11-14 14:44 .staging
-rw-r--r--+  3 lh2514 users       1324 2020-11-12 21:49 NYC_Mapper.py
-rw-rwxr--+  3 lh2514 users        314 2020-09-13 16:34 class
drwxrwxr-x+  - lh2514 users          0 2020-11-14 14:29 class1
drwxr-xr-x+  - lh2514 users          0 2020-10-22 16:47 hiveInput
drwxr-xr-x+  - lh2514 users          0 2020-10-22 15:37 impalaInput
drwxrwxr-x+  - lh2514 users          0 2020-09-08 18:05 myDir
drwxr-xr-x+  - lh2514 users          0 2020-11-13 18:22 project
[lh2514@login-2-1 ~]$ hdfs dfs -ls /user/lh2514/class1/
Found 17 items
drwxr-xr-x+  - lh2514 users          0 2020-10-12 01:27 /user/lh2514/class1/Airbnb_output
drwxr-xr-x+  - lh2514 users          0 2020-10-12 01:35 /user/lh2514/class1/Airbnb_output1
-rw-r--r--+  3 lh2514 users       1324 2020-11-12 00:53 /user/lh2514/class1/NYC_Mapper.py
-rw-r--r--+  3 lh2514 users       1037 2020-11-12 00:52 /user/lh2514/class1/NYC_Reducer.py
-rw-r--r--+  3 lh2514 users  265727269 2020-11-12 00:47 /user/lh2514/class1/NYPD_Complaint_Data_Historic.csv
-rw-r--r--+  3 lh2514 users        165 2020-11-14 14:29 /user/lh2514/class1/SelectFieldsMapper.java
-rw-r--r--+  3 lh2514 users        168 2020-11-14 14:28 /user/lh2514/class1/SelectFieldsReducer.java
-rw-r--r--+  3 lh2514 users       1851 2020-11-14 14:28 /user/lh2514/class1/SelectFieldsRunner.java
drwxr-xr-x+  - lh2514 users          0 2020-09-22 01:03 /user/lh2514/class1/Word_output
drwxr-xr-x+  - lh2514 users          0 2020-09-23 14:21 /user/lh2514/class1/Word_output1
-rw-r--r--+  3 lh2514 users    7077974 2020-10-12 01:26 /user/lh2514/class1/data.csv
-rw-r--r--+  3 lh2514 users  264678694 2020-11-05 01:14 /user/lh2514/class1/input.csv
-rw-rwxr--+  3 lh2514 users        314 2020-09-12 16:15 /user/lh2514/class1/input_hw2.txt
-rw-r--r--+  3 lh2514 users        289 2020-09-19 18:44 /user/lh2514/class1/input_hw3.txt
drwxr-xr-x+  - lh2514 users          0 2020-09-14 13:58 /user/lh2514/class1/output_hw2
drwxr-xr-x+  - lh2514 users          0 2020-11-12 01:45 /user/lh2514/class1/output_nyc
drwxr-xr-x+  - lh2514 users          0 2020-11-12 01:08 /user/lh2514/class1/python_code
[lh2514@login-2-1 ~]$ hadoop jar LinesCount.jar LinesCountRunner /user/lh2514/class1/NYPD_Complaint_Data_Historic.csv /user/lh2514/class1/output_NYC
20/11/14 14:46:51 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
20/11/14 14:46:52 INFO mapreduce.JobSubmitter: Cleaning up the staging area /user/lh2514/.staging/job_1604346392376_2027
20/11/14 14:46:52 WARN security.UserGroupInformation: PriviledgedActionException as:lh2514 (auth:SIMPLE) cause:org.apache.hadoop.mapreduce.lib.input.InvalidInputException: Input path does not e
xist: hdfs://dumbo/user/lh2514/NYPD_Complaint_Data_Historic.csv
Exception in thread "main" org.apache.hadoop.mapreduce.lib.input.InvalidInputException: Input path does not exist: hdfs://dumbo/user/lh2514/NYPD_Complaint_Data_Historic.csv
        at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.singleThreadedListStatus(FileInputFormat.java:323)
        at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.listStatus(FileInputFormat.java:265)
        at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.getSplits(FileInputFormat.java:387)
        at org.apache.hadoop.mapreduce.JobSubmitter.writeNewSplits(JobSubmitter.java:305)
        at org.apache.hadoop.mapreduce.JobSubmitter.writeSplits(JobSubmitter.java:322)
        at org.apache.hadoop.mapreduce.JobSubmitter.submitJobInternal(JobSubmitter.java:200)
        at org.apache.hadoop.mapreduce.Job$10.run(Job.java:1307)
        at org.apache.hadoop.mapreduce.Job$10.run(Job.java:1304)
        at java.security.AccessController.doPrivileged(Native Method)
        at javax.security.auth.Subject.doAs(Subject.java:422)
        at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1924)
        at org.apache.hadoop.mapreduce.Job.submit(Job.java:1304)
        at org.apache.hadoop.mapreduce.Job.waitForCompletion(Job.java:1325)
        at LinesCountRunner.main(LinesCountRunner.java:27)
        at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
        at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
        at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
        at java.lang.reflect.Method.invoke(Method.java:498)
        at org.apache.hadoop.util.RunJar.run(RunJar.java:226)
        at org.apache.hadoop.util.RunJar.main(RunJar.java:141)
[lh2514@login-2-1 ~]$ hadoop jar LinesCount.jar LinesCountRunner /user/lh2514/class1/input.csv /user/lh2514/class1/output_NYC
20/11/14 14:47:26 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
20/11/14 14:47:27 INFO mapreduce.JobSubmitter: Cleaning up the staging area /user/lh2514/.staging/job_1604346392376_2028
20/11/14 14:47:27 WARN security.UserGroupInformation: PriviledgedActionException as:lh2514 (auth:SIMPLE) cause:org.apache.hadoop.mapreduce.lib.input.InvalidInputException: Input path does not e
xist: hdfs://dumbo/user/lh2514/NYPD_Complaint_Data_Historic.csv
Exception in thread "main" org.apache.hadoop.mapreduce.lib.input.InvalidInputException: Input path does not exist: hdfs://dumbo/user/lh2514/NYPD_Complaint_Data_Historic.csv
        at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.singleThreadedListStatus(FileInputFormat.java:323)
        at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.listStatus(FileInputFormat.java:265)
        at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.getSplits(FileInputFormat.java:387)
        at org.apache.hadoop.mapreduce.JobSubmitter.writeNewSplits(JobSubmitter.java:305)
        at org.apache.hadoop.mapreduce.JobSubmitter.writeSplits(JobSubmitter.java:322)
        at org.apache.hadoop.mapreduce.JobSubmitter.submitJobInternal(JobSubmitter.java:200)
        at org.apache.hadoop.mapreduce.Job$10.run(Job.java:1307)
        at org.apache.hadoop.mapreduce.Job$10.run(Job.java:1304)
        at java.security.AccessController.doPrivileged(Native Method)
        at javax.security.auth.Subject.doAs(Subject.java:422)
        at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1924)
        at org.apache.hadoop.mapreduce.Job.submit(Job.java:1304)
        at org.apache.hadoop.mapreduce.Job.waitForCompletion(Job.java:1325)
        at LinesCountRunner.main(LinesCountRunner.java:27)
        at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
        at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
        at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
        at java.lang.reflect.Method.invoke(Method.java:498)
        at org.apache.hadoop.util.RunJar.run(RunJar.java:226)
        at org.apache.hadoop.util.RunJar.main(RunJar.java:141)
[lh2514@login-2-1 ~]$ vi LinesCountRunner.java
[lh2514@login-2-1 ~]$ javac -classpath "$(yarn classpath)":. -d . LinesCountRunner.java
[lh2514@login-2-1 ~]$ ar -cvf LinesCount.jar *.class
ar: no operation specified
[lh2514@login-2-1 ~]$ jar -cvf LinesCount.jar *.class
added manifest
adding: Airbnb.class(in = 1428) (out= 789)(deflated 44%)
adding: AirbnbMapper.class(in = 2269) (out= 1047)(deflated 53%)
adding: AirbnbReducer.class(in = 1694) (out= 714)(deflated 57%)
adding: IntSumReducer.class(in = 1694) (out= 714)(deflated 57%)
adding: LineCountMapper.class(in = 1478) (out= 596)(deflated 59%)
adding: LinesCountRunner.class(in = 1374) (out= 755)(deflated 45%)
adding: Main.class(in = 1461) (out= 872)(deflated 40%)
adding: SelectFieldsMapper.class(in = 255) (out= 198)(deflated 22%)
adding: SelectFieldsReducer.class(in = 258) (out= 196)(deflated 24%)
adding: SelectFieldsRunner.class(in = 1819) (out= 964)(deflated 47%)
adding: WordCount.class(in = 1440) (out= 789)(deflated 45%)
adding: WordCountMapper.class(in = 2096) (out= 962)(deflated 54%)
adding: WordCountReducer.class(in = 1700) (out= 717)(deflated 57%)
[lh2514@login-2-1 ~]$ hadoop jar LinesCount.jar LinesCountRunner /user/lh2514/class1/NYPD_Complaint_Data_Historic.csv /user/lh2514/class1/output_NYC
20/11/14 14:51:20 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
20/11/14 14:51:20 INFO input.FileInputFormat: Total input paths to process : 1
20/11/14 14:51:21 INFO mapreduce.JobSubmitter: number of splits:2
20/11/14 14:51:21 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1604346392376_2029
20/11/14 14:51:21 INFO impl.YarnClientImpl: Submitted application application_1604346392376_2029
20/11/14 14:51:21 INFO mapreduce.Job: The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1604346392376_2029/
20/11/14 14:51:21 INFO mapreduce.Job: Running job: job_1604346392376_2029
20/11/14 14:51:26 INFO mapreduce.Job: Job job_1604346392376_2029 running in uber mode : false
20/11/14 14:51:26 INFO mapreduce.Job:  map 0% reduce 0%
20/11/14 14:51:31 INFO mapreduce.Job:  map 100% reduce 0%
20/11/14 14:51:36 INFO mapreduce.Job:  map 100% reduce 97%
20/11/14 14:51:37 INFO mapreduce.Job:  map 100% reduce 100%
20/11/14 14:51:42 INFO mapreduce.Job: Job job_1604346392376_2029 completed successfully
20/11/14 14:51:42 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=897900
```

```
                FILE: Number of bytes written=7039288
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=265793063
                HDFS: Number of bytes written=20
                HDFS: Number of read operations=102
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=64
        Job Counters
                Launched map tasks=2
                Launched reduce tasks=32
                Data-local map tasks=2
                Total time spent by all maps in occupied slots (ms)=20912
                Total time spent by all reduces in occupied slots (ms)=426060
                Total time spent by all map tasks (ms)=5228
                Total time spent by all reduce tasks (ms)=71010
                Total vcore-milliseconds taken by all map tasks=5228
                Total vcore-milliseconds taken by all reduce tasks=71010
                Total megabyte-milliseconds taken by all map tasks=21413888
                Total megabyte-milliseconds taken by all reduce tasks=436285440
        Map-Reduce Framework
                Map input records=1048576
                Map output records=1048576
                Map output bytes=16777216
                Map output materialized bytes=898294
                Input split bytes=258
                Combine input records=0
                Combine output records=0
                Reduce input groups=1
                Reduce shuffle bytes=898294
                Reduce input records=1048576
                Reduce output records=1
                Spilled Records=2097152
                Shuffled Maps =64
                Failed Shuffles=0
                Merged Map outputs=64
                GC time elapsed (ms)=1642
                CPU time spent (ms)=37320
                Physical memory (bytes) snapshot=13810323456
                Virtual memory (bytes) snapshot=127167029248
                Total committed heap usage (bytes)=41371041792
        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=265792805
        File Output Format Counters
                Bytes Written=20
[lh2514@login-2-1 ~]$ hdfs dfs -ls /user/lh2514/class1/output_NYC
Found 33 items
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/_SUCCESS
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00000
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00001
-rw-r--r--+  3 lh2514 users         20 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00002
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00003
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00004
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00005
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00006
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00007
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00008
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00009
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00010
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00011
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00012
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00013
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00014
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00015
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00016
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00017
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00018
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00019
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00020
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00021
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00022
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00023
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00024
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00025
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00026
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00027
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00028
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00029
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00030
-rw-r--r--+  3 lh2514 users          0 2020-11-14 14:51 /user/lh2514/class1/output_NYC/part-r-00031
[lh2514@login-2-1 ~]$ hdfs dfs -cat /user/lh2514/class1/output_NYC/part-r-00002
Total Lines    1048576
[lh2514@login-2-1 ~]$
```