# Assignment 6.3

**Problem Statement**
**Enhance the task 8 (refer session 6, assignment 1) to calculate the top 3 state-wise sales for each company. You may use multiple reducers for this activity**

**Step 1: Start Hadoop Daemons:**

*$ start-all.sh*

*$ jps*

```
[acadgild@localhost dataset]$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
17/08/28 16:35:28 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting namenodes on [localhost]
localhost: namenode running as process 2830. Stop it first.
localhost: datanode running as process 2931. Stop it first.
Starting secondary namenodes [0.0.0.0]
0.0.0.0: secondarynamenode running as process 3086. Stop it first.
17/08/28 16:35:35 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
starting yarn daemons
resourcemanager running as process 3263. Stop it first.
localhost: nodemanager running as process 3364. Stop it first.
[acadgild@localhost dataset]$ jps
2931 DataNode
3364 NodeManager
2830 NameNode
3086 SecondaryNameNode
3263 ResourceManager
4111 Jps
[acadgild@localhost dataset]$
```

**Step 2: Compile the program and get JAR file into Acadgild Sandbox:**

```
[acadgild@localhost dataset]$ cd assignment6/
[acadgild@localhost assignment6]$ ll
total 12
-rw-rw-r--. 1 acadgild acadgild 7180 Aug 28 16:37 mapreduce.jar
-rw-rw-r--. 1 acadgild acadgild  733 Aug 28 16:27 television.txt
[acadgild@localhost assignment6]$
```

**Step 3: Move dataset to hadoop directory:**

*$ hadoop fs -ls /user/acadgild/hadoop/assignment6*

```
[acadgild@localhost assignment6]$ hadoop fs -ls /user/acadgild/hadoop/assignment6/
17/08/28 18:25:35 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin
Found 4 items
drwxr-xr-x   - acadgild supergroup          0 2017-08-28 17:04 /user/acadgild/hadoop/assignment6/output61a
drwxr-xr-x   - acadgild supergroup          0 2017-08-28 17:12 /user/acadgild/hadoop/assignment6/output61b
drwxr-xr-x   - acadgild supergroup          0 2017-08-28 18:02 /user/acadgild/hadoop/assignment6/output62
-rw-r--r--   1 acadgild supergroup        733 2017-08-28 16:40 /user/acadgild/hadoop/assignment6/television.txt
[acadgild@localhost assignment6]$
```

**Step 4: run jar file with below command to get the desired result:**

*$ hadoop jar mapreduce.jar assignment63.Television63a /user/acadgild/hadoop/assignment6/television.txt /user/acadgild/hadoop/assignment6/output63a*

```
[acadgild@localhost assignment6]$ hadoop jar mapreduce.jar assignment63.Television63a /user/acadgild/hadoop/assignment6/television.txt /user/acadgild/hadoop/assignment6/output63a
17/08/28 18:27:09 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
17/08/28 18:27:10 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
17/08/28 18:27:10 WARN mapreduce.JobSubmitter: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
17/08/28 18:27:11 INFO input.FileInputFormat: Total input paths to process : 1
17/08/28 18:27:11 INFO mapreduce.JobSubmitter: number of splits:1
17/08/28 18:27:11 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1503918153486_0012
17/08/28 18:27:11 INFO impl.YarnClientImpl: Submitted application application_1503918153486_0012
17/08/28 18:27:11 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1503918153486_0012/
17/08/28 18:27:11 INFO mapreduce.Job: Running job: job_1503918153486_0012
17/08/28 18:27:17 INFO mapreduce.Job: Job job_1503918153486_0012 running in uber mode : false
17/08/28 18:27:17 INFO mapreduce.Job:  map 0% reduce 0%
17/08/28 18:27:22 INFO mapreduce.Job:  map 100% reduce 0%
17/08/28 18:27:29 INFO mapreduce.Job:  map 100% reduce 100%
17/08/28 18:27:29 INFO mapreduce.Job: Job job_1503918153486_0012 completed successfully
17/08/28 18:27:30 INFO mapreduce.Job: Counters: 49
```

Above class file is run to get the number of units sold by each company in the states.

**Step 5: check the intermediate output file:**

**We can check the intermediate output of the above step.**

*$ hadoop fs -ls /user/acadgild/hadoop/assignment6/output63a/*

*$ hadoop fs -cat /user/acadgild/hadoop/assignment6/output63a/**

```
[acadgild@localhost assignment6]$ hadoop fs -ls /user/acadgild/hadoop/assignment6/output63a/
17/08/28 18:29:06 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup          0 2017-08-28 18:27 /user/acadgild/hadoop/assignment6/output63a/_SUCCESS
-rw-r--r--   1 acadgild supergroup        165 2017-08-28 18:27 /user/acadgild/hadoop/assignment6/output63a/part-r-00000
[acadgild@localhost assignment6]$ hadoop fs -cat /user/acadgild/hadoop/assignment6/output63a/*
17/08/28 18:29:17 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Akai    Kerala  1
Lava    Assam   3
NA      Uttar Pradesh   1
Onida   Kerala  1
Onida   Uttar Pradesh   3
Samsung Kerala  1
Samsung Madhya Pradesh  3
Samsung Maharashtra     3
Zen     Maharashtra     2
[acadgild@localhost assignment6]$
```

**Step 6: Output of the step 4 will be given as input to the second class file:**

*$ hadoop jar mapreduce.jar assignment63.Television63b /user/acadgild/hadoop/assignment6/output63a/* /user/acadgild/hadoop/assignment6/output63b*

```
[acadgild@localhost assignment6]$ hadoop jar mapreduce.jar assignment63.Television63b /user/acadgild/hadoop/assignment6/output63a/* /user/acadgild/hadoop/assignment6/output63b
17/08/28 18:31:00 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
17/08/28 18:31:01 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
17/08/28 18:31:02 WARN mapreduce.JobSubmitter: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
17/08/28 18:31:02 INFO input.FileInputFormat: Total input paths to process : 1
17/08/28 18:31:02 INFO mapreduce.JobSubmitter: number of splits:1
17/08/28 18:31:02 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1503918153486_0013
17/08/28 18:31:02 INFO impl.YarnClientImpl: Submitted application application_1503918153486_0013
17/08/28 18:31:03 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1503918153486_0013/
17/08/28 18:31:03 INFO mapreduce.Job: Running job: job_1503918153486_0013
17/08/28 18:31:10 INFO mapreduce.Job: Job job_1503918153486_0013 running in uber mode : false
17/08/28 18:31:10 INFO mapreduce.Job:  map 0% reduce 0%
17/08/28 18:31:15 INFO mapreduce.Job:  map 100% reduce 0%
17/08/28 18:31:21 INFO mapreduce.Job:  map 100% reduce 100%
17/08/28 18:31:21 INFO mapreduce.Job: Job job_1503918153486_0013 completed successfully
17/08/28 18:31:21 INFO mapreduce.Job: Counters: 49
```

**Step 7: check the output generated by step 6:**

*$ hadoop fs -ls /user/acadgild/hadoop/assignment6/output63b/*

*$ hadoop fs -cat /user/acadgild/hadoop/assignment6/output63b/**

```
[acadgild@localhost assignment6]$ hadoop fs -ls /user/acadgild/hadoop/assignment6/output63b/
17/08/28 18:31:57 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup          0 2017-08-28 18:31 /user/acadgild/hadoop/assignment6/output63b/_SUCCESS
-rw-r--r--   1 acadgild supergroup        165 2017-08-28 18:31 /user/acadgild/hadoop/assignment6/output63b/part-r-00000
[acadgild@localhost assignment6]$ hadoop fs -cat /user/acadgild/hadoop/assignment6/output63b/*
17/08/28 18:32:08 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Akai    Kerala  1
Lava    Assam   3
NA      Uttar Pradesh   1
Onida   Uttar Pradesh   3
Onida   Kerala  1
Samsung Maharashtra     3
Samsung Madhya Pradesh  3
Samsung Kerala  1
Zen     Maharashtra     2
[acadgild@localhost assignment6]$
```

Thus we get the final output of the problem statement.