# Assignment 6.2

**Problem Statement**

**Enhance the Map Reduce program of Task 8 (refer session 6, assignment 1) to use multiple reducers for sorting. The driver should accept three additional values: the minimum units sold, the maximum units sold and number of reducers to use. Use units sold as key and company as value. Write a custom partitioner to divide the keys on the basis of range. Take minimum to be 0 and maximum to be 10. Divide them across 2 reducers.**

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

*$ hadoop jar mapreduce.jar assignment62.Television /user/acadgild/hadoop/assignment6/output61a /user/acadgild/hadoop/assignment6/output62*

```
[acadgild@localhost assignment6]$ hadoop jar mapreduce.jar assignment62.Television /user/acadgild/hadoop/assignment6/output61a /user/acadgild/hadoop/assignment6/output62
17/08/28 18:02:02 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
17/08/28 18:02:03 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
17/08/28 18:02:04 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 18:02:04 INFO input.FileInputFormat: Total input paths to process : 1
17/08/28 18:02:04 INFO mapreduce.JobSubmitter: number of splits:1
17/08/28 18:02:05 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1503918153486_0011
17/08/28 18:02:05 INFO impl.YarnClientImpl: Submitted application application_1503918153486_0011
17/08/28 18:02:05 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1503918153486_0011/
17/08/28 18:02:05 INFO mapreduce.Job: Running job: job_1503918153486_0011
17/08/28 18:02:12 INFO mapreduce.Job: Job job_1503918153486_0011 running in uber mode : false
17/08/28 18:02:12 INFO mapreduce.Job:  map 0% reduce 0%
17/08/28 18:02:17 INFO mapreduce.Job:  map 100% reduce 0%
17/08/28 18:02:25 INFO mapreduce.Job:  map 100% reduce 50%
17/08/28 18:02:26 INFO mapreduce.Job:  map 100% reduce 100%
17/08/28 18:02:26 INFO mapreduce.Job: Job job_1503918153486_0011 completed successfully
17/08/28 18:02:26 INFO mapreduce.Job: Counters: 49
```

**Step 4: Check the output directory of the program:**
**$ hadoop fs -ls /user/acadgild/hadoop/assignment6/output62/**

```
[acadgild@localhost assignment6]$ hadoop fs -ls /user/acadgild/hadoop/assignment6/output62/
17/08/28 18:04:19 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 3 items
-rw-r--r--   1 acadgild supergroup          0 2017-08-28 18:02 /user/acadgild/hadoop/assignment6/output62/_SUCCESS
-rw-r--r--   1 acadgild supergroup         33 2017-08-28 18:02 /user/acadgild/hadoop/assignment6/output62/part-r-00000
-rw-r--r--   1 acadgild supergroup         10 2017-08-28 18:02 /user/acadgild/hadoop/assignment6/output62/part-r-00001
[acadgild@localhost assignment6]$
```

**Above directory shows two reduce files since we have chosen 2 reducer task as mentioned in partition class.**
/user/acadgild/hadoop/assignment6/output62/part-r-00000
/user/acadgild/hadoop/assignment6/output62/part-r-00001

**Step 5: check the files output:**

**Reducer 0 gets the keys which are less than 5.**

**Reducer1 gets the keys greater than 5.**

*$ hadoop fs -cat /user/acadgild/hadoop/assignment6/output62/part-r-00000*

*$ hadoop fs -cat /user/acadgild/hadoop/assignment6/output62/part-r-00001*

```
[acadgild@localhost assignment6]$ hadoop fs -cat /user/acadgild/hadoop/assignment6/output62/part-r-00000
17/08/28 18:08:07 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin
1       NA
1       Akai
2       Zen
3       Lava
4       Onida
[acadgild@localhost assignment6]$ hadoop fs -cat /user/acadgild/hadoop/assignment6/output62/part-r-00001
17/08/28 18:08:15 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin
7       Samsung
[acadgild@localhost assignment6]$
```

Thus we get the final output of the problem statement.