Big Data – Assignment 2

1. The Execution is as below:

```
Apoorvas-MacBook-Pro:q2 apoorvaajay$ cd ../q1
Apoorvas-MacBook-Pro:q1 apoorvaajay$ sbt package
[info] Set current project to Friend Recommender (in build file:/Users/apoorvaajay/Documents/UTD%20MS%20Sem4/Big%20Data/Assignment2/q1/)
[info] Updating {file:/Users/apoorvaajay/Documents/UTD%20MS%20Sem4/Big%20Data/Assignment2/q1/}q1...
[info] Resolving org.fusesource.jansi#jansi;1.4 ...
[info] Done updating.
[info] Compiling 1 Scala source to /Users/apoorvaajay/Documents/UTD MS Sem4/Big Data/Assignment2/q1/target/scala-2.10/classes...
[info] Packaging /Users/apoorvaajay/Documents/UTD MS Sem4/Big Data/Assignment2/q1/target/scala-2.10/friend-recommender_2.10-1.0.jar ...
[info] Done packaging.
[success] Total time: 10 s, completed Mar 14, 2016 10:52:06 PM
Apoorvas-MacBook-Pro:q1 apoorvaajay$
Apoorvas-MacBook-Pro:q1 apoorvaajay$ /Users/apoorvaajay/Documents/spark-1.6.0/bin/spark-submit --class org.assignment.scala.FriendRecommender --master local[4] target/scala-2.10/fr
riend-recommender_2.10-1.0.jar soc-LiveJournal1Adj.txt
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
16/03/14 22:53:28 INFO SparkContext: Running Spark version 1.6.0
16/03/14 22:53:28 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
16/03/14 22:53:28 INFO SecurityManager: Changing view acls to: apoorvaajay
16/03/14 22:53:28 INFO SecurityManager: Changing modify acls to: apoorvaajay
16/03/14 22:53:28 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: Set(apoorvaajay); users with modify permissions: S
et(apoorvaajay)
16/03/14 22:53:29 INFO Utils: Successfully started service 'sparkDriver' on port 57059.
16/03/14 22:53:29 INFO Slf4jLogger: Slf4jLogger started
16/03/14 22:53:29 INFO Remoting: Starting remoting
16/03/14 22:53:30 INFO Remoting: Remoting started; listening on addresses :[akka.tcp://sparkDriverActorSystem@192.168.0.6:57060]
16/03/14 22:53:30 INFO Utils: Successfully started service 'sparkDriverActorSystem' on port 57060.
16/03/14 22:53:30 INFO SparkEnv: Registering MapOutputTracker
16/03/14 22:53:30 INFO SparkEnv: Registering BlockManagerMaster
```

The recommendations for the users are as mentioned below :

924      439,2409,6995,11860,15416,43748,45881
8941     8943,8944,8940
8942     8939,8940,8943,8944
9019     9022,317,9023
9020     9021,9016,9017,9022,317,9023
9021     9020,9016,9017,9022,317,9023
9022     9019,9020,9021,317,9016,9017,9023
9990     13134,13478,13877,34299,34485,34642,37941,
9992     9987,9989,35667,9991
9993     36764, 44132, 10058,44088,36765,36909,10055,44068,44076,1000

2. The Execution is as below:

```
Apoorvas-MacBook-Pro:q2 apoorvaajay$ sbt package
[info] Set current project to Mutual Friend Finder (in build file:/Users/apoorvaajay/Documents/UTD%20MS%20Sem4/Big%20Data/Assignment2/q2/)
[info] Updating {file:/Users/apoorvaajay/Documents/UTD%20MS%20Sem4/Big%20Data/Assignment2/q2/}q2...
[info] Resolving org.fusesource.jansi#jansi;1.4 ...
[info] Done updating.
[info] Compiling 1 Scala source to /Users/apoorvaajay/Documents/UTD MS Sem4/Big Data/Assignment2/q2/target/scala-2.10/classes...
[info] Packaging /Users/apoorvaajay/Documents/UTD MS Sem4/Big Data/Assignment2/q2/target/scala-2.10/mutual-friend-finder_2.10-1.0.jar ...
[info] Done packaging.
[success] Total time: 10 s, completed Mar 14, 2016 10:47:18 PM
Apoorvas-MacBook-Pro:q2 apoorvaajay$ /Users/apoorvaajay/Documents/spark-1.6.0/bin/spark-submit --class org.assignment.scala.mutualFriends --master local[4] target/scala-2.10/mutua
l-friend-finder_2.10-1.0.jar soc-LiveJournal1Adj.txt 0 1
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
16/03/14 22:50:30 INFO SparkContext: Running Spark version 1.6.0
16/03/14 22:50:30 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
16/03/14 22:50:30 INFO SecurityManager: Changing view acls to: apoorvaajay
16/03/14 22:50:30 INFO SecurityManager: Changing modify acls to: apoorvaajay
16/03/14 22:50:30 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: Set(apoorvaajay); users with modify permissions: S
et(apoorvaajay)
16/03/14 22:50:31 INFO Utils: Successfully started service 'sparkDriver' on port 57036.
16/03/14 22:50:31 INFO Slf4jLogger: Slf4jLogger started
```

The output is as below:

```
16/03/14 22:50:34 INFO DAGScheduler: Job 3 finished: collect at mutualFriends.scala:12, took 0.1936
0       1       Set(5, 20)16/03/14 22:50:34 INFO SparkContext: Invoking stop() from shutdown hook
16/03/14 22:50:34 INFO SparkUI: Stopped Spark web UI at http://192.168.0.6:4041
```

3. The Execution is as below:

```
[Apoorvas-MacBook-Pro:q3 apoorvaajay$ sbt package
[info] Set current project to Mutual Friend Information (in build file:/Users/apoorvaajay/Documents/UTD%20MS%20Sem4/Big%20Data/Assignment2/q3/)
[info] Compiling 1 Scala source to /Users/apoorvaajay/Documents/UTD MS Sem4/Big Data/Assignment2/q3/target/scala-2.10/classes...
[info] Packaging /Users/apoorvaajay/Documents/UTD MS Sem4/Big Data/Assignment2/q3/target/scala-2.10/mutual-friend-information_2.10-1.0.jar ...
[info] Done packaging.
[success] Total time: 6 s, completed Mar 14, 2016 11:41:11 PM
[Apoorvas-MacBook-Pro:q3 apoorvaajay$
[Apoorvas-MacBook-Pro:q3 apoorvaajay$
[Apoorvas-MacBook-Pro:q3 apoorvaajay$ /Users/apoorvaajay/Documents/spark-1.6.0/bin/spark-submit --class org.assignment.scala.mutualFriendInfo --master local[4] target/scala-2.10/mu
tual-friend-information_2.10-1.0.jar soc-LiveJournal1Adj.txt userdata.txt 25 35
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
16/03/14 23:41:34 INFO SparkContext: Running Spark version 1.6.0
16/03/14 23:41:34 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
16/03/14 23:41:34 INFO SecurityManager: Changing view acls to: apoorvaajay
16/03/14 23:41:34 INFO SecurityManager: Changing modify acls to: apoorvaajay
```

The output is as below:

```
16/03/14 23:41:39 INFO ShuffleBlockFetcherIterator: Getting 2 non-empty blocks out of 2 blocks
16/03/14 23:41:39 INFO ShuffleBlockFetcherIterator: Getting 2 non-empty blocks out of 2 blocks
16/03/14 23:41:39 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 2 ms
16/03/14 23:41:39 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 2 ms
16/03/14 23:41:40 INFO BlockManagerInfo: Removed broadcast_8_piece0 on localhost:57359 in memory (size: 2.2 KB, free: 511.1 MB)
25      35      (Evangeline,45140)
16/03/14 23:41:40 INFO Executor: Finished task 3.0 in stage 7.0 (TID 17). 1165 bytes result sent to driver
16/03/14 23:41:40 INFO TaskSetManager: Finished task 3.0 in stage 7.0 (TID 17) in 1189 ms on localhost (1/4)
16/03/14 23:41:40 INFO Executor: Finished task 0.0 in stage 7.0 (TID 14). 1165 bytes result sent to driver
16/03/14 23:41:40 INFO TaskSetManager: Finished task 0.0 in stage 7.0 (TID 14) in 1195 ms on localhost (2/4)
```

4. The Execution is as below:

```
[Apoorvas-MacBook-Pro:q3 apoorvaajay$ cd ../q4
[Apoorvas-MacBook-Pro:q4 apoorvaajay$ sbt package
[info] Set current project to Friend Average Age Finder (in build file:/Users/apoorvaajay/Documents/UTD%20MS%20Sem4/Big%20Data/Assignment2/q4/)
[info] Updating {file:/Users/apoorvaajay/Documents/UTD%20MS%20Sem4/Big%20Data/Assignment2/q4/}q4...
[info] Resolving org.fusesource.jansi#jansi;1.4 ...
[info] Done updating.
[info] Compiling 1 Scala source to /Users/apoorvaajay/Documents/UTD MS Sem4/Big Data/Assignment2/q4/target/scala-2.10/classes...
[info] Packaging /Users/apoorvaajay/Documents/UTD MS Sem4/Big Data/Assignment2/q4/target/scala-2.10/friend-average-age-finder_2.10-1.0.jar ...
[info] Done packaging.
[success] Total time: 10 s, completed Mar 14, 2016 11:42:26 PM
[Apoorvas-MacBook-Pro:q4 apoorvaajay$
[Apoorvas-MacBook-Pro:q4 apoorvaajay$ /Users/apoorvaajay/Documents/spark-1.6.0/bin/spark-submit --class org.assignment.scala.FriendsAvgAge --master local[4] target/scala-2.10/frien
d-average-age-finder_2.10-1.0.jar soc-LiveJournal1Adj.txt userdata.txt
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
16/03/14 23:44:03 INFO SparkContext: Running Spark version 1.6.0
16/03/14 23:44:03 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
16/03/14 23:44:04 INFO SecurityManager: Changing view acls to: apoorvaajay
16/03/14 23:44:04 INFO SecurityManager: Changing modify acls to: apoorvaajay
16/03/14 23:44:04 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: Set(apoorvaajay); users with modify permissions: S
et(apoorvaajay)
```

The output is as below:

```
6822, 518 Modoc Alley Boise Idaho Idaho 83703 US          86
17279, 2252 Oakridge Farm Lane Waukesha Wisconsin Wisconsin 53186 US     86
16359, 4771 Ritter Avenue Detroit Michigan Michigan 48226 US    86
33662, 4755 Oakway Lane Los Angeles California California 90017 US       86
2448, 3057 Elmwood Avenue Scottsdale Arizona Arizona 85259 US    86
22893, 2661 Counts Lane Cincinnati Kentucky Kentucky 45202 US    86
30469, 548 Camden Street Sparks Nevada Nevada 89431 US  86
24907, 1460 Rose Street Harvey Illinois Illinois 60426 US       86
19318, 4035 Canis Heights Drive Los Angeles California California 90007 US       86
2433, 4999 Bee Street Traverse City Michigan Michigan 49686 US  86
16/03/14 23:44:13 INFO Executor: Finished task 3.0 in stage 17.0 (TID 28). 1165 bytes result sent to driver
16/03/14 23:44:13 INFO TaskSetManager: Finished task 3.0 in stage 17.0 (TID 28) in 115 ms on localhost (1/4)
16/03/14 23:44:13 INFO Executor: Finished task 2.0 in stage 17.0 (TID 27). 1165 bytes result sent to driver
16/03/14 23:44:13 INFO TaskSetManager: Finished task 2.0 in stage 17.0 (TID 27) in 116 ms on localhost (2/4)
37554, 371 Blue Spruce Lane Hanover Maryland Maryland 21076 US  86
16/03/14 23:44:13 INFO Executor: Finished task 1.0 in stage 17.0 (TID 26). 1165 bytes result sent to driver
16/03/14 23:44:13 INFO TaskSetManager: Finished task 1.0 in stage 17.0 (TID 26) in 176 ms on localhost (3/4)
21605, 3810 Willis Avenue Deltona Florida Florida 32738 US      86
6820, 438 Adams Avenue Waldorf Maryland Maryland 20601 US       86
16368, 151 Harvest Lane Warrensburg Missouri Missouri 64093 US  86
30438, 2498 Heron Way Portland Oregon Oregon 97204 US    86
15352, 693 Terry Lane Orlando Florida Florida 32810 US  86
21616, 1024 Graystone Lakes Cedar Grove Georgia Georgia 31021 US        86
19309, 3301 Palmer Road Westerville Ohio Ohio 43081 US  86
2435, 4287 Locust Street Albany Georgia Georgia 31707 US        86
21609, 131 Jett Lane Los Angeles California California 90017 US 86
```