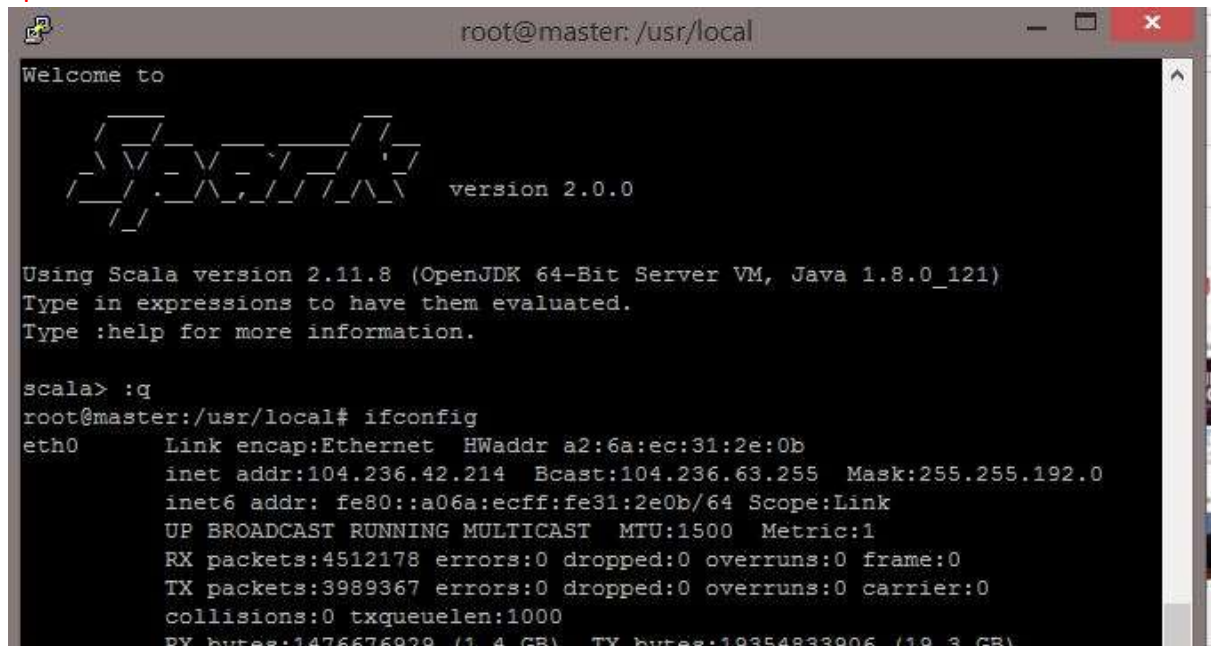


Mini Project 2

1- Spark Shell

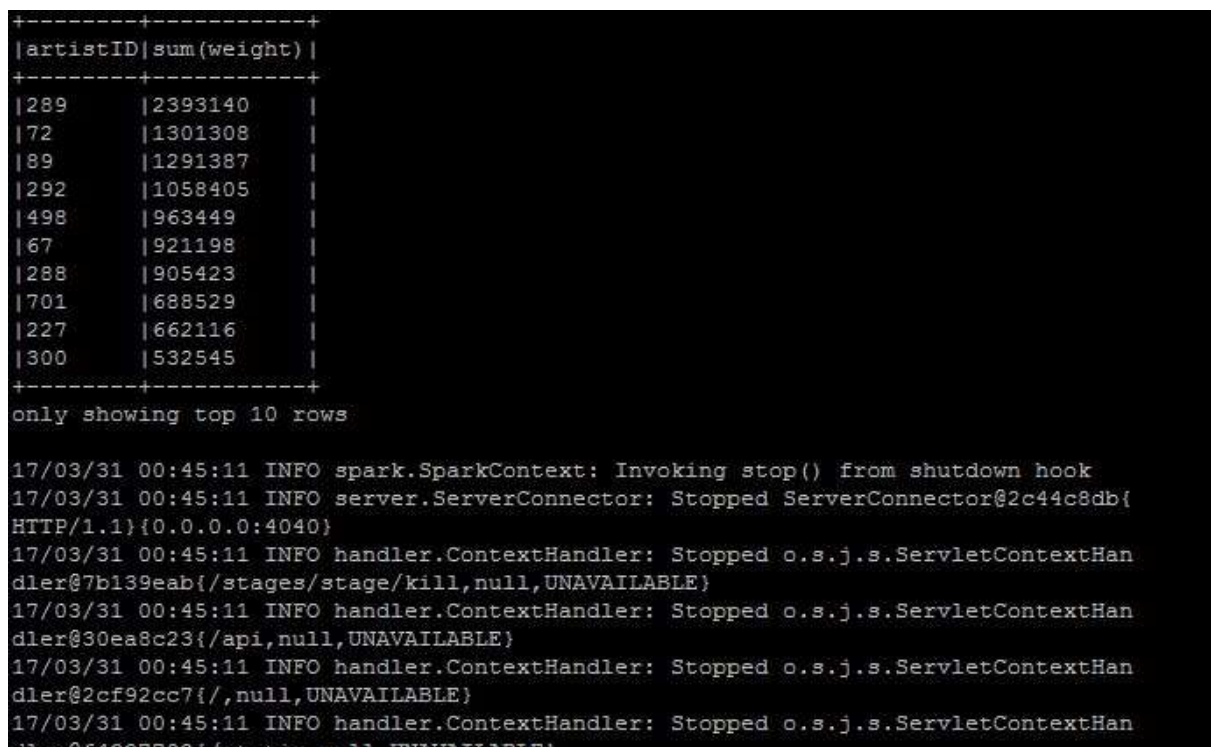


```
root@master: /usr/local
Welcome to
  ____  ____  ____  ____  ____  ____  ____  ____  ____  ____
 /_ _\/_ _\/_ _\/_ _\/_ _\/_ _\/_ _\/_ _\/_ _\/_ _\/_ _\
version 2.0.0

Using Scala version 2.11.8 (OpenJDK 64-Bit Server VM, Java 1.8.0_121)
Type in expressions to have them evaluated.
Type :help for more information.

scala> :q
root@master:/usr/local# ifconfig
eth0      Link encap:Ethernet  HWaddr a2:6a:ec:31:2e:0b
          inet addr:104.236.42.214  Bcast:104.236.63.255  Mask:255.255.192.0
          inet6 addr: fe80::a06a:ecff:fe31:2e0b/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:4512178  errors:0  dropped:0  overruns:0  frame:0
          TX packets:3989367  errors:0  dropped:0  overruns:0  carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:1476676929 (1.4 GB)  TX bytes:19354833906 (19.3 GB)
```

2- Task 2



```
+-----+-----+
|artistID|sum(weight)|
+-----+-----+
|289     |2393140    |
|72      |1301308    |
|89      |1291387    |
|292     |1058405    |
|498     |963449     |
|67      |921198     |
|288     |905423     |
|701     |688529     |
|227     |662116     |
|300     |532545     |
+-----+-----+
only showing top 10 rows

17/03/31 00:45:11 INFO spark.SparkContext: Invoking stop() from shutdown hook
17/03/31 00:45:11 INFO server.ServerConnector: Stopped ServerConnector@2c44c8db{
HTTP/1.1}{0.0.0.0:4040}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@7b139eab{/stages/stage/kill,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@30ea8c23{/api,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@2cf92cc7{/null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@64422702{/static,null,UNAVAILABLE}
```

```
|288      |905423      |
|701      |688529      |
|227      |662116      |
|300      |532545      |
+-----+-----+
only showing top 10 rows

17/03/31 00:45:11 INFO spark.SparkContext: Invoking stop() from shutdown hook
17/03/31 00:45:11 INFO server.ServerConnector: Stopped ServerConnector@2c44c8db{
HTTP/1.1}{0.0.0.0:4040}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@7b139eab{/stages/stage/kill,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@30ea8c23{/api,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@2cf92cc7{/,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@64337702{/static,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@76c7beb3{/executors/threadDump/json,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@751d3241{/executors/threadDump,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@3f2ef586{/executors/json,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@756cf158{/executors,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@2f058b8a{/environment/json,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@6175619b{/environment,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@165b8a71{/storage/rdd/json,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@7fd4acee{/storage/rdd,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@5b970f7{/storage/json,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@1a1da881{/storage,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@119f1f2a{/stages/pool/json,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@1a1da881{/storage,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@119f1f2a{/stages/pool/json,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@3d08f3f5{/stages/pool,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@55795845{/stages/stage/json,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@4fe89c24{/stages/stage,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@5dbe30be{/stages/json,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@89c10b7{/stages,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@72bd06ca{/jobs/job/json,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@625e134e{/jobs/job,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@5cc126dc{/jobs/json,null,UNAVAILABLE}
17/03/31 00:45:11 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHan
dler@7e809b79{/jobs,null,UNAVAILABLE}
17/03/31 00:45:11 INFO ui.SparkUI: Stopped Spark web UI at http://104.236.42.214
:4040
17/03/31 00:45:11 INFO cluster.YarnClientSchedulerBackend: Interrupting monitor
thread
```


3- Task 3:

Note: Time is in Seconds

Without using cached RDD

- a. How many hits were made to the website item “/assets/js/lowpro.js”?

293

```
4477550
/assets/js/lowpro.js 293
Timetaken77
17/03/31 01:25:00 INFO spark.SparkContext: Invoking stop() from shutdown hook
17/03/31 01:25:00 INFO server.ServerConnector: Stopped ServerConnector@6f0d554{HTTP/1.1}{0.0.0.0:4040}
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6326d182{/stages/stage/ks
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@475835b1{/api,null,UNAVAIL
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4fad94a7{/null,UNAVAILAB
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@36b0fcd5{/static,null,UN
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4eed46ee{/executors/threa
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@fd0e5b6{/executors/thread
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@52b56a3e{/executors/json,
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4c168660{/executors,null,
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@3c0fae6c{/environment/jsc
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@649725e3{/environment,nul
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@757d6814{/storage/rdd/jsc
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2c1b9e4b{/storage/rdd,nul
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@c7a975a{/storage/json,nul
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@743cb8e0{/storage,null,UN
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@235f4c10{/stages/pool/jsc
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@7fcbe147{/stages/pool,nul
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@44c79f32{/stages/stage/ja
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6256ac4f{/stages/stage,nul
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2a1edad4{/stages/json,nul
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4ebea12c{/stages,null,UN
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@1536602f{/jobs/job/json,r
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@25ddb5bb{/jobs/job,null,U
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@43b4fe19{/jobs/json,null,
17/03/31 01:25:00 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6e16b8b5{/jobs,null,UNAV
17/03/31 01:25:00 INFO ui.SparkUI: Stopped Spark web UI at http://104.236.42.214:4040
17/03/31 01:25:00 INFO cluster.YarnClientSchedulerBackend: Interrupting monitor thread
```

With cache:

```
/assets/js/lowpro.js 293
Timetaken70
17/03/31 02:04:03 INFO spark.SparkContext: Invoking stop() from shutdown hook
17/03/31 02:04:03 INFO server.ServerConnector: Stopped ServerConnector@5f1da3e{HTTP/1.1}{0.0.0.0:4040}
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6326d182{/stages/
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@475835b1{/api,nul
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4fad94a7{/null,U
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@36b0fcd5{/static,
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4eed46ee{/executo
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@fd0e5b6{/executo
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@52b56a3e{/executo
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4c168660{/executo
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@3c0fae6c{/environ
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@649725e3{/environ
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@757d6814{/storage
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2c1b9e4b{/storage
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@c7a975a{/storage/
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@743cb8e0{/storage
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@235f4c10{/stages/
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@7fcbe147{/stages/
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@44c79f32{/stages/
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6256ac4f{/stages/
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2a1edad4{/stages/
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4ebea12c{/stages,
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@1536602f{/jobs/jo
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@25ddb5bb{/jobs/jo
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@43b4fe19{/jobs/js
17/03/31 02:04:03 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6e16b8b5{/jobs,nu
17/03/31 02:04:03 INFO ui.SparkUI: Stopped Spark web UI at http://104.236.42.214:4040
```

- b. How many hits were made to the website item “/favicon.ico”?
66831

```
17/03/31 01:30:28 INFO Scheduler.DAGScheduler: Job 0 finished: collect at AccessLog2.java
/favicon.ico 66831
4411012
Timetaken: 73
17/03/31 01:30:28 INFO spark.SparkContext: Invoking stop() from shutdown hook
17/03/31 01:30:28 INFO server.ServerConnector: Stopped ServerConnector@24a11c1{HTTP/1.1}
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@632
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@475
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4fa
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@36b
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4ee
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@fd0
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@52b
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4c1
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@3c0
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@649
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@757
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2c1
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@c7a
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@743
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@235
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@7fc
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@44c
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@625
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2a1
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4eb
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@153
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@25d
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@43b
17/03/31 01:30:28 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6e1
17/03/31 01:30:28 INFO ui.SparkUI: Stopped Spark web UI at http://104.236.42.214:4040
17/03/31 01:30:28 INFO cluster.YarnClientSchedulerBackend: Interrupting monitor thread
```


- c. Which path in the website has been hit most? How many hits were made to the path?

```
17/03/31 01:46:09 INFO scheduler.DAGScheduler: Job 1 finished: take at AccessLog3.java:
117348 /assets/css/combined.css
Timetaken: 79
17/03/31 01:46:09 INFO spark.SparkContext: Invoking stop() from shutdown hook
17/03/31 01:46:09 INFO server.ServerConnector: Stopped ServerConnector@5b6813df{HTTP/1.
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@71
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@52
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@63
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@47
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4f
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@36
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4e
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@fd
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@52
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4c
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@3c
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@64
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@75
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2c
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@c7
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@74
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@23
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@7f
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@44
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@62
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2a
17/03/31 01:46:09 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4e
17/03/31 01:46:10 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@15
17/03/31 01:46:10 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@25
17/03/31 01:46:10 INFO ui.SparkUI: Stopped Spark web UI at http://104.236.42.214:4040
```

d. Which IP accesses the website most? How many accesses were made by it?

```
158613 10.216.113.172
Timetaken: 90
17/03/31 01:49:32 INFO spark.SparkContext: Invoking stop() from shutdown hook
17/03/31 01:49:32 INFO server.ServerConnector: Stopped ServerConnector@6a988392{HTTP/1.1
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@632
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@475
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4fa
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@36b
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4ee
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@fd0
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@52b
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4c1
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@3c0
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@649
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@757
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2c1
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@c7a
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@743
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@235
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@7fc
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@44c
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@625
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2a1
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4eb
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@153
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@25d
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@43b
17/03/31 01:49:32 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6e1
17/03/31 01:49:32 INFO ui.SparkUI: Stopped Spark web UI at http://104.236.42.214:4040
```

Task 3: part 2:

Loading the file every time to calculate counts

It took more time. I found the time difference of 15 seconds. I forgot to take the screen shot for the first time I ran it. When I tried it second time it is giving memory issue. So, I am unable to put screen shot for this one only.

Loading once and filtering in the middle

```
Timetaken: 77
17/03/31 02:53:08 INFO spark.SparkContext: Invoking stop() from shutdown hook
17/03/31 02:53:08 INFO server.ServerConnector: Stopped ServerConnector@6a988392{HTTP/1.1}{0.0.0.
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6326d182{/s
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@475835b1{/a
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4fad94a7{/f
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@36b0fcd5{/s
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4eed46ee{/e
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@fd0e5b6{/ex
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@52b56a3e{/e
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4c168660{/e
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@3c0fae6c{/e
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@649725e3{/e
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@757d6814{/s
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2c1b9e4b{/s
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@c7a975a{/st
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@743cb8e0{/s
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@235f4c10{/s
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@7fcbe147{/s
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@44c79f32{/s
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6256ac4f{/s
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2aledad4{/s
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4ebea12c{/s
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@1536602f{/j
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@25ddbbb{/j
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@43b4fe19{/j
17/03/31 02:53:08 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6e16b8b5{/j
17/03/31 02:53:08 INFO ui.SparkUI: Stopped Spark web UI at http://104.236.42.214:4040
```

Sai charan Talipineni
sat122@pitt.edu

Sai Rakesh Ghanta
sag163@pitt.edu

Note: Readme file is in zip folder.

References:

<http://spark.apache.org/docs/latest/programming-guide.html>

<http://spark.apache.org/examples.html>

<https://spark.apache.org/docs/0.9.1/java-programming-guide.html>

<http://www.javaworld.com/article/2972863/big-data/open-source-java-projects-apache-spark.html>

<https://www.simple-talk.com/cloud/cloud-data/start-big-data-apache-spark/>