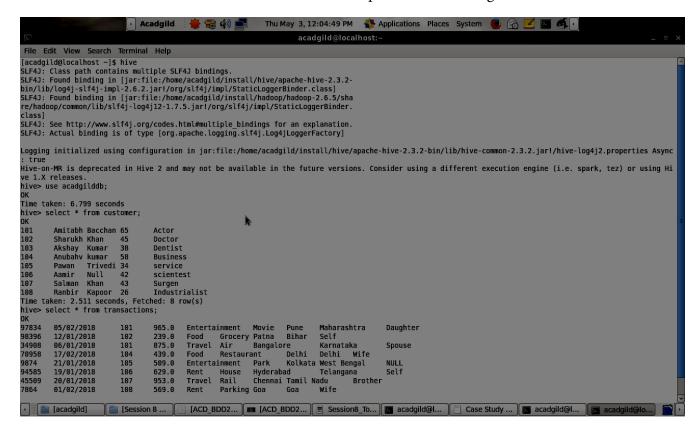# CASE STUDY - II HIVE

Created 2 Tables CUSTOMER , TRANSACTIONS as part of TODO assignment & loaded data for the same.

Q1: Find out the number of transaction done by each customer

I have achieved the same using below query, results shown in screenshot

Q2: Create a new table called TRANSACTIONS_COUNT. This table should have two fields - fname and count

I have created table as described & loaded data from TRASACNTIONS table into TRASACTION_COUNT table -- for the two coloumns FNAME,COUNT, result of table show below.

acadgild@localhost:~

File Edit View Search Terminal Help

```
Total jobs = 1
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hive/apache-hive-2.3.2-bin/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBind
er.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
2018-05-03 12:18:10     Starting to launch local task to process map join;       maximum memory = 518979584
2018-05-03 12:18:12     Dump the side-table for tag: 0 with group count: 8 into file: file:/tmp/acadgild/c9421d20-4085-4fe5-a6ce-b0c384857cfc/hive_2018-05-03
_12-18-00_979_3529013323770279398-1/-local-10003/HashTable-Stage-2/MapJoin-mapfile10--.hashtable
2018-05-03 12:18:12     Uploaded 1 File to: file:/tmp/acadgild/c9421d20-4085-4fe5-a6ce-b0c384857cfc/hive_2018-05-03_12-18-00_979_3529013323770279398-1/-local
-10003/HashTable-Stage-2/MapJoin-mapfile10--.hashtable (469 bytes)
2018-05-03 12:18:12     End of local task; Time Taken: 2.368 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1525322803121_0006, Tracking URL = http://localhost:8088/proxy/application_1525322803121_0006/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job  -kill job_1525322803121_0006
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
2018-05-03 12:18:22,698 Stage-2 map = 0%,  reduce = 0%
2018-05-03 12:18:30,460 Stage-2 map = 100%,  reduce = 0%, Cumulative CPU 1.92 sec
2018-05-03 12:18:38,300 Stage-2 map = 100%,  reduce = 100%, Cumulative CPU 4.43 sec
MapReduce Total cumulative CPU time: 4 seconds 430 msec
Ended Job = job_1525322803121_0006
Loading data to table acadgilddb.transaction_count
MapReduce Jobs Launched:
Stage-Stage-2: Map: 1  Reduce: 1   Cumulative CPU: 4.43 sec   HDFS Read: 13638 HDFS Write: 148 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 430 msec
OK
Time taken: 40.442 seconds
hive> 
```

‹ | 🖳 [acadgild] | 📁 [Session 8 ... | 📁 [ACD_BDD2... | ▬ [ACD_BDD2... | 🖳 acadgild@lo... | ☐ [Case Study... | 🖳 acadgild@lo... | 🖳 acadgild@l... | ☐ Untitled 1 - ... | 🖳 ›

---

acadgild@localhost:~

File Edit View Search Terminal Help

```
2018-05-03 12:18:12     Uploaded 1 File to: file:/tmp/acadgild/c9421d20-4085-4fe5-a6ce-b0c384857cfc/hive_2018-05-03_12-18-00_979_3529013323770279398-1/-local
-10003/HashTable-Stage-2/MapJoin-mapfile10--.hashtable (469 bytes)
2018-05-03 12:18:12     End of local task; Time Taken: 2.368 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1525322803121_0006, Tracking URL = http://localhost:8088/proxy/application_1525322803121_0006/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job  -kill job_1525322803121_0006
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
2018-05-03 12:18:22,698 Stage-2 map = 0%,  reduce = 0%
2018-05-03 12:18:30,460 Stage-2 map = 100%,  reduce = 0%, Cumulative CPU 1.92 sec
2018-05-03 12:18:38,300 Stage-2 map = 100%,  reduce = 100%, Cumulative CPU 4.43 sec
MapReduce Total cumulative CPU time: 4 seconds 430 msec
Ended Job = job_1525322803121_0006
Loading data to table acadgilddb.transaction_count
MapReduce Jobs Launched:
Stage-Stage-2: Map: 1  Reduce: 1   Cumulative CPU: 4.43 sec   HDFS Read: 13638 HDFS Write: 148 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 430 msec
OK
Time taken: 40.442 seconds
hive> select * from transaction_count;
OK
Aamir   1
Amitabh 2
Anubahv 1
Pawan   1
Ranbir  1
Salman  1
Sharukh 1
Time taken: 0.298 seconds, Fetched: 7 row(s)
hive> 
```

‹ | 🖳 [acadgild] | 📁 [Session 8 ... | 📁 [ACD_BDD2... | ▬ [ACD_BDD2... | 🖳 acadgild@lo... | ☐ [Case Study... | 🖳 acadgild@lo... | 🖳 acadgild@l... | ☐ Untitled 1 - ... | 🖳 ›