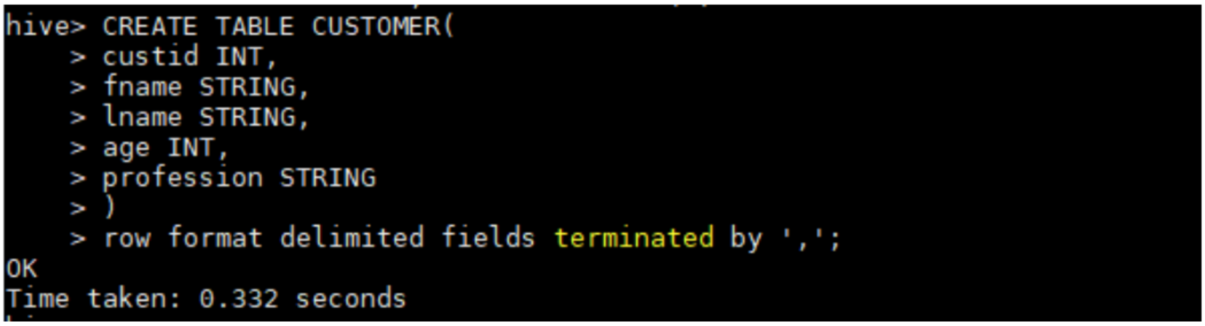
To-do

Case Study Description Let us take up the CUSTOMER and TRANSACTIONS table we have created in the Let's Do Together section. Let us solve the following use cases using these tables :-

1. Find out the number of transaction done by each customer (These should be take up in module 8 itself)

Step:1 Start The Hive Shell

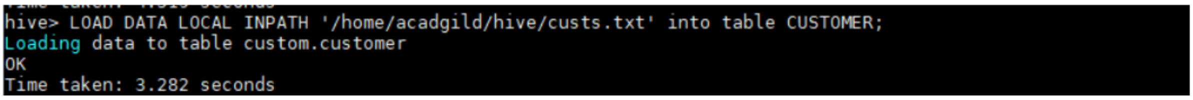
Step 2: Create Customer Table



```
hive> CREATE TABLE CUSTOMER(
    > custid INT,
    > fname STRING,
    > lname STRING,
    > age INT,
    > profession STRING
    > )
    > row format delimited fields terminated by ',';
OK
Time taken: 0.332 seconds
```

Step 3: We have inserted data into customer table from local file custs.txt as shown below

LOAD DATA LOCAL INPATH '/home/acadgild/hive/custs.txt' into table CUSTOMER;



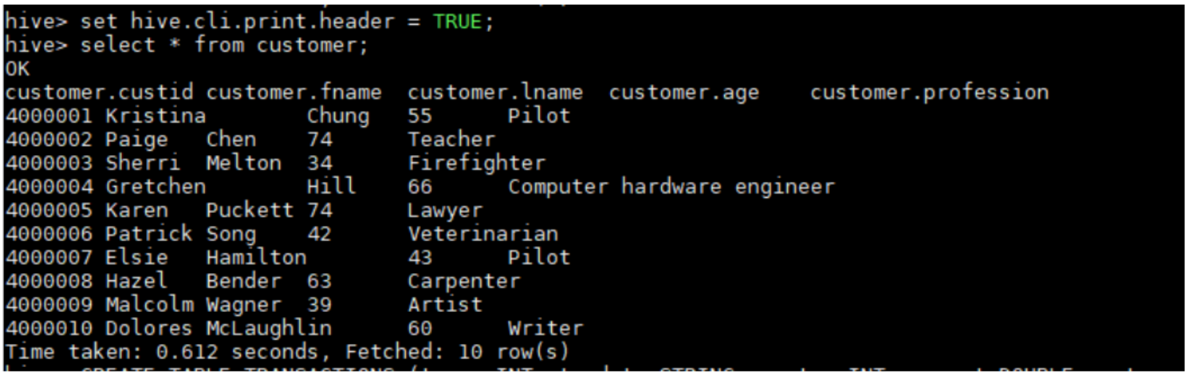
```
hive> LOAD DATA LOCAL INPATH '/home/acadgild/hive/custs.txt' into table CUSTOMER;
Loading data to table custom.customer
OK
Time taken: 3.282 seconds
```

Step 4: After inserting data into customer table, we are verifying that data is present in customer table by fetching rows by using query :

select * from customer;

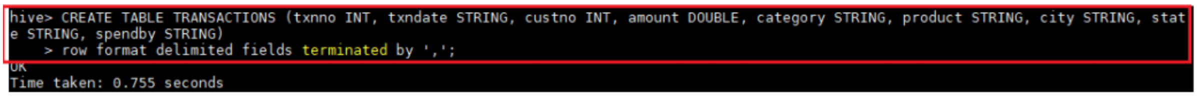
Before that, we have set column header to TRUE so that we can have column headers along with output by using command :

set hive.cli.print.header = TRUE;



```
hive> set hive.cli.print.header = TRUE;
hive> select * from customer;
OK
customer.custid customer.fname  customer.lname  customer.age    customer.profession
4000001 Kristina        Chung   55      Pilot
4000002 Paige   Chen    74      Teacher
4000003 Sherri  Melton  34      Firefighter
4000004 Gretchen        Hill    66      Computer hardware engineer
4000005 Karen   Puckett 74      Lawyer
4000006 Patrick Song    42      Veterinarian
4000007 Elsie   Hamilton        43      Pilot
4000008 Hazel   Bender  63      Carpenter
4000009 Malcolm Wagner  39      Artist
4000010 Dolores McLaughlin      60      Writer
Time taken: 0.612 seconds, Fetched: 10 row(s)
```

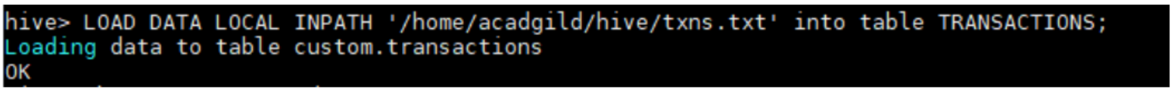
Now Create **TRANSACTIONS table**



```
hive> CREATE TABLE TRANSACTIONS (txnno INT, txndate STRING, custno INT, amount DOUBLE, category STRING, product STRING, city STRING, stat
e STRING, spendby STRING)
    > row format delimited fields terminated by ',';
OK
Time taken: 0.755 seconds
```

Insert Data into Transactions Table

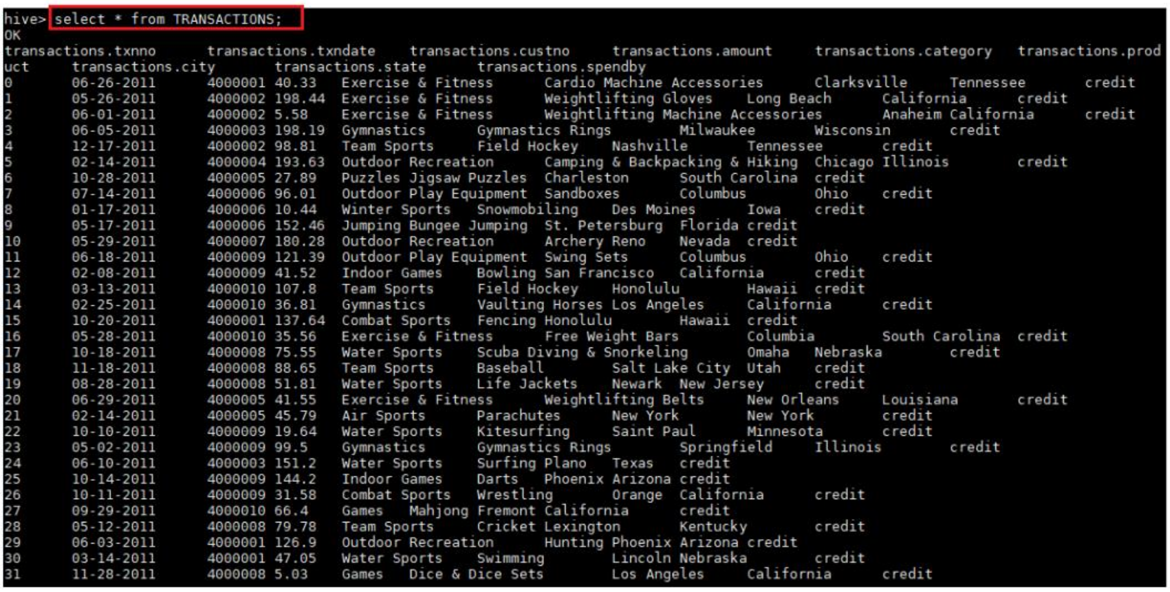
LOAD DATA LOCAL INPATH '/home/acadgild/hive/txns.txt' into table TRANSACTIONS;



```
hive> LOAD DATA LOCAL INPATH '/home/acadgild/hive/txns.txt' into table TRANSACTIONS;
Loading data to table custom.transactions
OK
```

See the inserted Data

select * from transactions;



```
hive> select * from TRANSACTIONS;
OK
transactions.txnno      transactions.txndate    transactions.custno     transactions.amount     transactions.category   transactions.prod
uct     transactions.city       transactions.state      transactions.spendby
0       06-26-2011      4000001 40.33   Exercise & Fitness      Cardio Machine Accessories      Clarksville     Tennessee       credit
1       05-26-2011      4000002 198.44  Exercise & Fitness      Weightlifting Gloves    Long Beach      California      credit
2       06-01-2011      4000002 5.58    Exercise & Fitness      Weightlifting Machine Accessories       Anaheim California      credit
3       06-05-2011      4000003 198.19  Gymnastics      Gymnastics Rings        Milwaukee       Wisconsin       credit
4       12-17-2011      4000002 98.81   Team Sports     Field Hockey    Nashville       Tennessee       credit
5       02-14-2011      4000004 193.63  Outdoor Recreation      Camping & Backpacking & Hiking  Chicago Illinois        credit
6       10-28-2011      4000005 27.89   Puzzles Jigsaw Puzzles  Charleston      South Carolina  credit
7       07-14-2011      4000006 96.01   Outdoor Play Equipment  Sandboxes       Columbus        Ohio    credit
8       01-17-2011      4000006 10.44   Winter Sports   Snowmobiling    Des Moines      Iowa    credit
9       05-17-2011      4000006 152.46  Jumping Bungee Jumping  St. Petersburg  Florida credit
10      05-29-2011      4000007 180.28  Outdoor Recreation      Archery Reno    Nevada  credit
11      06-18-2011      4000009 121.39  Outdoor Play Equipment  Swing Sets      Columbus        Ohio    credit
12      02-08-2011      4000009 41.52   Indoor Games    Bowling San Francisco   California      credit
13      03-13-2011      4000010 107.8   Team Sports     Field Hockey    Honolulu        Hawaii  credit
14      02-25-2011      4000010 36.81   Gymnastics      Vaulting Horses Los Angeles     California      credit
15      10-20-2011      4000001 137.64  Combat Sports   Fencing Honolulu        Hawaii  credit
16      05-28-2011      4000010 35.56   Exercise & Fitness      Free Weight Bars        Columbia        South Carolina  credit
17      10-18-2011      4000008 75.55   Water Sports    Scuba Diving & Snorkeling       Omaha   Nebraska        credit
18      11-18-2011      4000008 88.65   Team Sports     Baseball        Salt Lake City  Utah    credit
19      08-28-2011      4000008 51.81   Water Sports    Life Jackets    Newark  New Jersey      credit
20      06-29-2011      4000005 41.55   Exercise & Fitness      Weightlifting Belts     New Orleans     Louisiana       credit
21      02-14-2011      4000005 45.79   Air Sports      Parachutes      New York        New York        credit
22      10-10-2011      4000009 19.64   Water Sports    Kitesurfing     Saint Paul      Minnesota       credit
23      05-02-2011      4000009 99.5    Gymnastics      Gymnastics Rings        Springfield     Illinois        credit
24      06-10-2011      4000003 151.2   Water Sports    Surfing Plano   Texas   credit
25      10-14-2011      4000009 144.2   Indoor Games    Darts   Phoenix Arizona credit
26      10-11-2011      4000009 31.58   Combat Sports   Wrestling       Orange  California      credit
27      09-29-2011      4000010 66.4    Games   Mahjong Fremont California      credit
28      05-12-2011      4000008 79.78   Team Sports     Cricket Lexington       Kentucky        credit
29      06-03-2011      4000001 126.9   Outdoor Recreation      Hunting Phoenix Arizona credit
30      03-14-2011      4000001 47.05   Water Sports    Swimming        Lincoln Nebraska        credit
31      11-28-2011      4000008 5.03    Games   Dice & Dice Sets        Los Angeles     California      credit
```

2. Create a new table called TRANSACTIONS_COUNT. This table should have 3 fields - custid, fname and count. (Again to be done in module 8)

3. Now write a hive query in such a way that the query populates the data obtained in Step 1 above and populate the table in step 2 above. (This has to be done in module 9).

4. Now lets make the TRANSACTIONS_COUNT table Hbase complaint. In the sence, use Ser Des And Storate handler features of hive to change the TRANSACTIONS_COUNT table to be able to create a TRANSACTIONS table in Hbase. (This has to be done in module 10)

5. Now insert the data in TRANSACTIONS_COUNT table using the query in step 3 again, this should populate the Hbase TRANSACTIONS table automatically (This has to be done in module 10)

6. Now from the Hbase level, write the Hbase java API code to access and scan the TRANSACTIONS table data from java level.