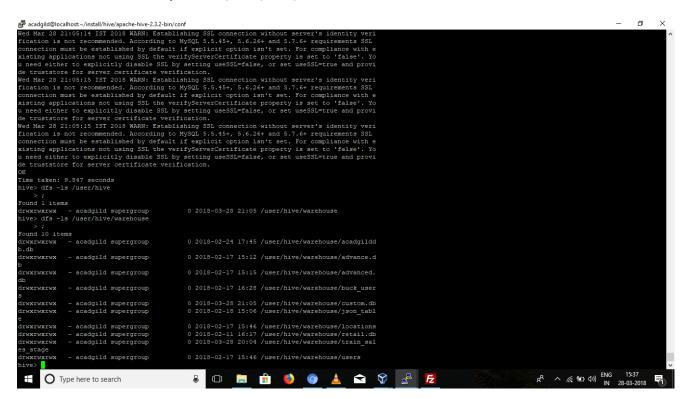# Assignment

Task1:

Command used:　　a) create database custom;　→　database created is stored as directory in hive and metadata in mysql.

　　　　　　　　　　b) dfs –ls /user/hive/warehouse　　→　check database
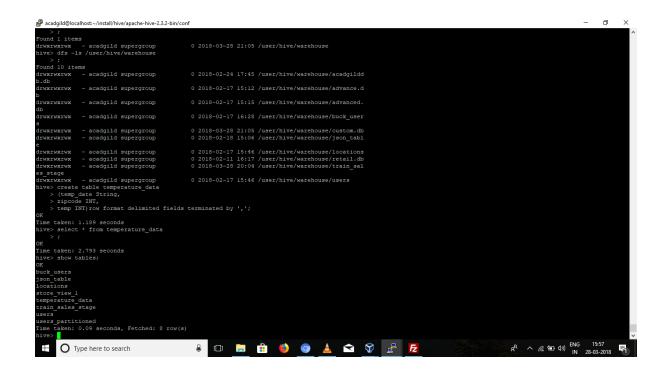

c) create table temperature_data →
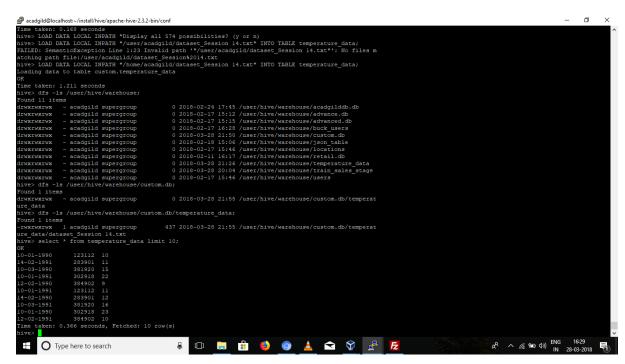
create table temperature_data

(temp_data String,

Zipcode INT,

Temp INT) row format delimited fields terminated by ',';

```
    > ;
Found 1 items
drwxrwxrwx   - acadgild supergroup          0 2018-03-28 21:05 /user/hive/warehouse
hive> dfs -ls /user/hive/warehouse
    > ;
Found 10 items
drwxrwxrwx   - acadgild supergroup          0 2018-02-24 17:45 /user/hive/warehouse/acadgildd
b.db
drwxrwxrwx   - acadgild supergroup          0 2018-02-17 15:12 /user/hive/warehouse/advance.d
b
drwxrwxrwx   - acadgild supergroup          0 2018-02-17 15:15 /user/hive/warehouse/advanced.
db
drwxrwxrwx   - acadgild supergroup          0 2018-02-17 16:28 /user/hive/warehouse/buck_user
s
drwxrwxrwx   - acadgild supergroup          0 2018-03-28 21:05 /user/hive/warehouse/custom.db
drwxrwxrwx   - acadgild supergroup          0 2018-02-18 15:06 /user/hive/warehouse/json_tabl
e
drwxrwxrwx   - acadgild supergroup          0 2018-02-17 15:46 /user/hive/warehouse/locations
drwxrwxrwx   - acadgild supergroup          0 2018-02-11 16:17 /user/hive/warehouse/retail.db
drwxrwxrwx   - acadgild supergroup          0 2018-03-28 20:04 /user/hive/warehouse/train_sal
es_stage
drwxrwxrwx   - acadgild supergroup          0 2018-02-17 15:46 /user/hive/warehouse/users
hive> create table temperature_data
    > (temp_date String,
    > zipcode INT,
    > temp INT)row format delimited fields terminated by ',';
OK
Time taken: 1.189 seconds
hive> select * from temperature_data
    > ;
OK
Time taken: 2.793 seconds
hive> show tables;
OK
buck_users
json_table
locations
store_view_1
temperature_data
train_sales_stage
users
users_partitioned
Time taken: 0.09 seconds, Fetched: 8 row(s)
hive>
```

d) load dataset →   load data local inpath "/user/acadgild/dataset-session 14.txt" into table temperature_data;


```
Time taken: 0.168 seconds
hive> LOAD DATA LOCAL INPATH "Display all 574 possibilities? (y or n)
hive> LOAD DATA LOCAL INPATH "/user/acadgild/dataset_Session 14.txt" INTO TABLE temperature_data;
FAILED: SemanticException Line 1:23 Invalid path '"/user/acadgild/dataset_Session 14.txt"': No files m
atching path file:/user/acadgild/dataset_Session%2014.txt
hive> LOAD DATA LOCAL INPATH "/home/acadgild/dataset_Session 14.txt" INTO TABLE temperature_data;
Loading data to table custom.temperature_data
OK
Time taken: 1.211 seconds
hive> dfs -ls /user/hive/warehouse;
Found 11 items
drwxrwxrwx   - acadgild supergroup          0 2018-02-24 17:45 /user/hive/warehouse/acadgilddb.db
drwxrwxrwx   - acadgild supergroup          0 2018-02-17 15:12 /user/hive/warehouse/advance.db
drwxrwxrwx   - acadgild supergroup          0 2018-02-17 15:15 /user/hive/warehouse/advanced.db
drwxrwxrwx   - acadgild supergroup          0 2018-02-17 16:28 /user/hive/warehouse/buck_users
drwxrwxrwx   - acadgild supergroup          0 2018-03-28 21:50 /user/hive/warehouse/custom.db
drwxrwxrwx   - acadgild supergroup          0 2018-02-18 15:06 /user/hive/warehouse/json_table
drwxrwxrwx   - acadgild supergroup          0 2018-02-17 15:46 /user/hive/warehouse/locations
drwxrwxrwx   - acadgild supergroup          0 2018-02-11 16:17 /user/hive/warehouse/retail.db
drwxrwxrwx   - acadgild supergroup          0 2018-03-28 21:26 /user/hive/warehouse/temperature_data
drwxrwxrwx   - acadgild supergroup          0 2018-03-28 20:04 /user/hive/warehouse/train_sales_stage
drwxrwxrwx   - acadgild supergroup          0 2018-02-17 15:46 /user/hive/warehouse/users
hive> dfs -ls /user/hive/warehouse/custom.db;
Found 1 items
drwxrwxrwx   - acadgild supergroup          0 2018-03-28 21:55 /user/hive/warehouse/custom.db/temperat
ure_data
hive> dfs -ls /user/hive/warehouse/custom.db/temperature_data;
Found 1 items
-rwxrwxrwx   1 acadgild supergroup        437 2018-03-28 21:55 /user/hive/warehouse/custom.db/temperat
ure_data/dataset_Session 14.txt
hive> select * from temperature_data limit 10;
OK
10-01-1990      123112  10
14-02-1991      283901  11
10-03-1990      381920  15
10-01-1991      302918  22
12-02-1990      384902  9
10-01-1991      123112  11
14-02-1990      283901  12
10-03-1991      381920  16
10-01-1990      302918  23
12-02-1991      384902  10
Time taken: 0.366 seconds, Fetched: 10 row(s)
hive>
```

Task2

a) Select * from temperature_data where zipcode>300000 and zipcode<399999