**Assignemnet 1:**

**Objective** :

1. To create database named 'Custom'

2. create table named temperature\_data with the fields :

date (mm-dd-yyyy) format;

zipcode

temperature

3. load data from the dataset.txt (which is delimkited by ',') in to the table;

**solution**:

objective 1.

create database if not exists custom; // create db

show databases; // to verify the database created

use custom; // use the created db

objective 2 and 3 . As the objective is to load the date in theformat ('mm-dd-yyyy') , which i snot supported in hive date format directly (hive supports date in the format 'yyyy-mm-dd'). below process is followed:

-- create a table with the date column as string

create table temperature\_staging (date String , zipcode String , tem-perature String) row format delimited fields terminated by ',' ;

-- load the data in to the temperary table

load data local inpath '/tmp/dataset.txt' overwrite into table temperature\_staging;

-- create the final table with the date format

create table temperature\_data (date Date, zipcode String , temperature String) row format delimited fields terminated by ',' ;

-- load the data from the temp table to the final table using the unix date converter

insert into table temperature\_data select from\_unixtime(unix\_timestamp(date,'mm-dd-yyyy')),zipcode,temperature from temperature\_staging;