Assignment 6.1 ( Hive Introduction)

Step1: The following hive command is used to create database named custom

CREATE DATABASE custom;

Step2: The following command is used to show database custom created successfully

SHOW DATABASES;

Step3:

The table temperature\_data is created with schema

date type string, zip\_code string, temperature float

Using the command below:

CREATE TABLE IF NOT EXISTS temperature\_data

(

date string,

zip\_code string,

temperature float

)

ROW FORMAT DELIMITED FIELDS TERMINATED BY ',';

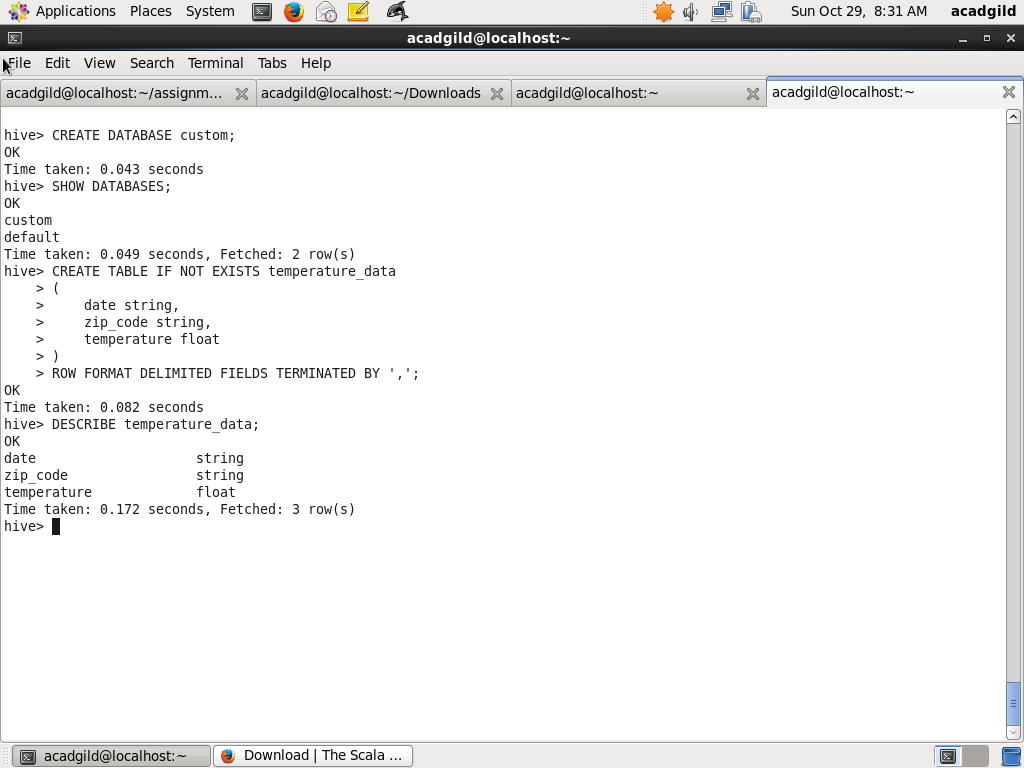
Note: The dataset given has date in the format dd-MM-yyyy as we have data like 02-14-1991, 14 can only be day (not month). But data is displayed, it will be displayed in MM-dd-yyyy

Step4:

Use the following command to check that table temperature\_data is created

DESCRIBE temperature\_data;

Screenshot of Step1 to Step4 is as below



Step5:

Use the following command to switch to newly created database custom

USE custom;

Step6:

Use the following LOAD DATA LOCAL INPATH command to load data from given dataset dataset\_Session 14.txt

LOAD DATA LOCAL INPATH '/home/acadgild/assignment\_6.1/dataset\_Session 14.txt' OVERWRITE INTO TABLE temperature\_data;

Step7:

Use the following SELECT command to display all the records of table temperatute\_data. Note date will be displayed in MM-dd-yyyy format

SELECT from\_unixtime(UNIX\_TIMESTAMP(date, 'dd-MM-yyyy'),'MM-dd-yyyy'), zip\_code, temperature FROM temperature\_data;

Screenshot of Steps 5 to 7 is as below:

