Assignment 14.1

4. Associated Data Files

https://drive.google.com/file/d/0Bxr27gVaXO5sa0JBamZXdkpYUFk/view?usp=sharing

5. Problem Statement

**Create a database named 'custom'.**

**Create a table named temperature\_data inside custom having below fields:**

**1. date (mm-dd-yyyy) format**

**2. zip code**

**3. temperature**

**The table will be loaded from comma-delimited file.**

**Load the dataset.txt (which is ',' delimited) in the table**

CREATE TABLE temperature\_data(

date Date,

zipcode String,

temperature INT

)

COMMENT ""

ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' STORED AS TEXTFILE;

//creating temp table to get data into it so that we can do a date conversion later

CREATE TABLE temperature\_data\_t(

date string,

zipcode String,

temperature INT

)

COMMENT ""

ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' STORED AS TEXTFILE;

**LOAD DATA INPATH '/user/ec2-user/hive\_input\_files/temperaturedata.txt' INTO TABLE temperature\_data\_t;**

**insert overwrite table temperature\_data select FROM\_UNIXTIME(unix\_timestamp(date ,'dd-MM-yyyy'), 'yyyy-MM-dd'),zipcode,temperature from temperature\_data\_t;**