**Task1:**

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

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

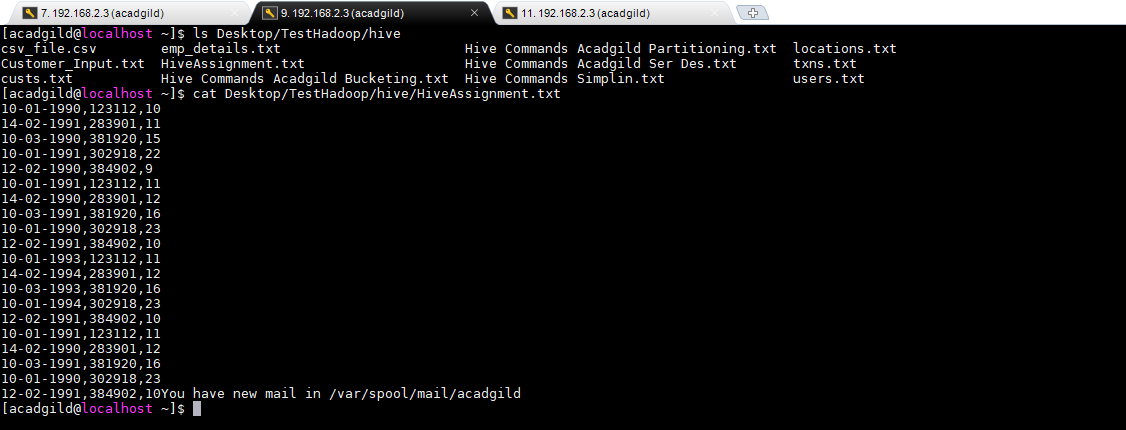
**2. zip code**

**3. temperature**

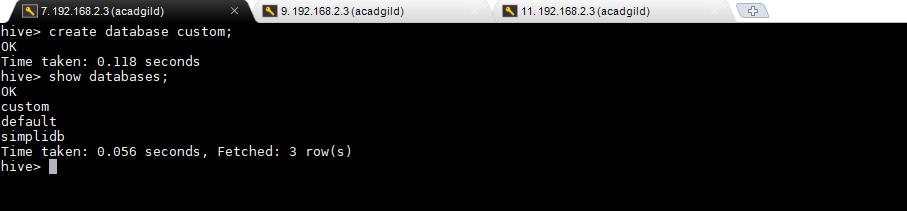
**The table will be loaded from comma-delimited file. Load the dataset.txt (which is ',' delimited) in the table.**

Solution:

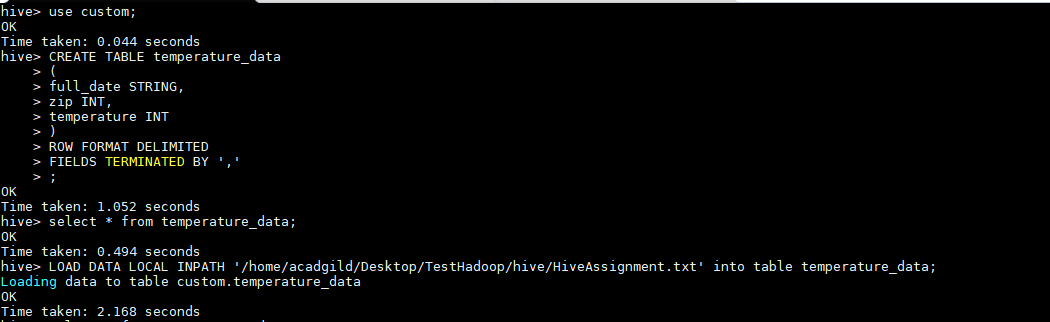
Input file is copied to the location - */home/acadgild/Desktop/TestHadoop/hive/HiveAssignment.txt*

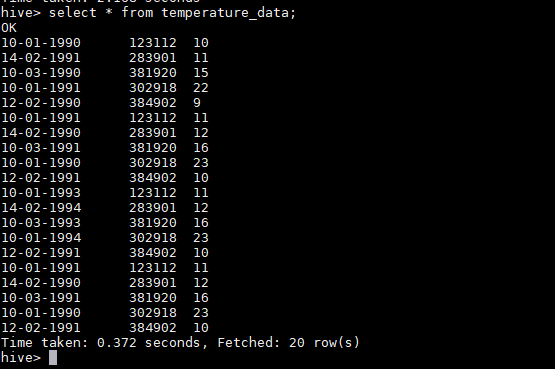


To create database ‘custom’ enter the HIVE shell and run the following command



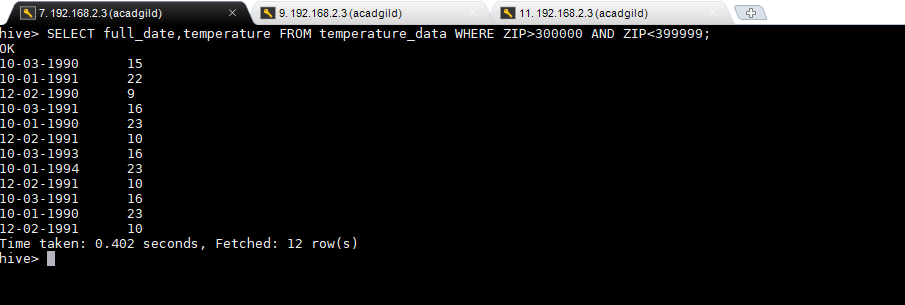
To create table ‘temperature\_data ‘ inside database ‘custom’





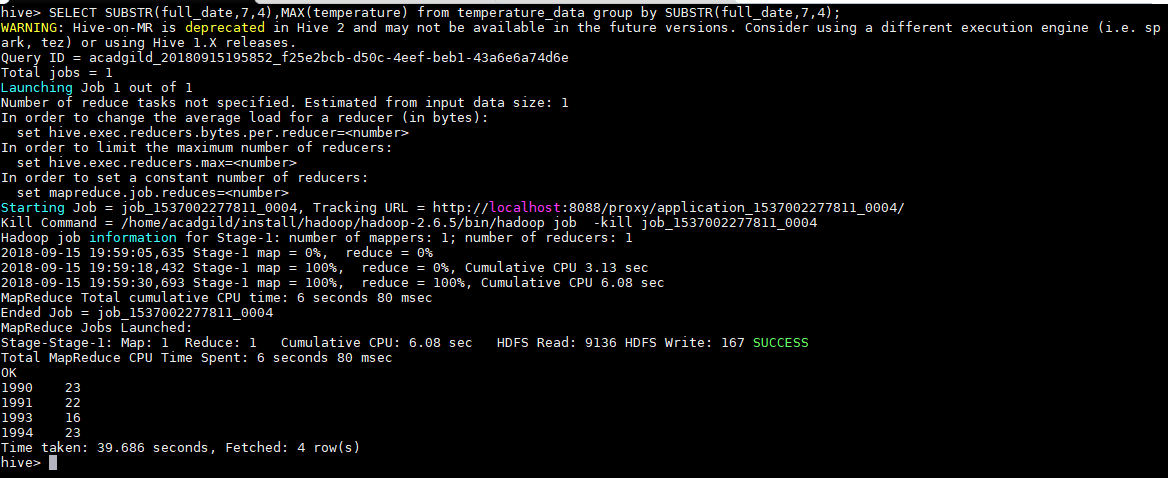
**Task 2(a): Fetch date and temperature from temperature\_data where zip code is greater than 300000 and less than 399999.**

**Command**: *SELECT full\_date,temperature FROM temperature\_data WHERE ZIP>300000 AND ZIP<399999;*



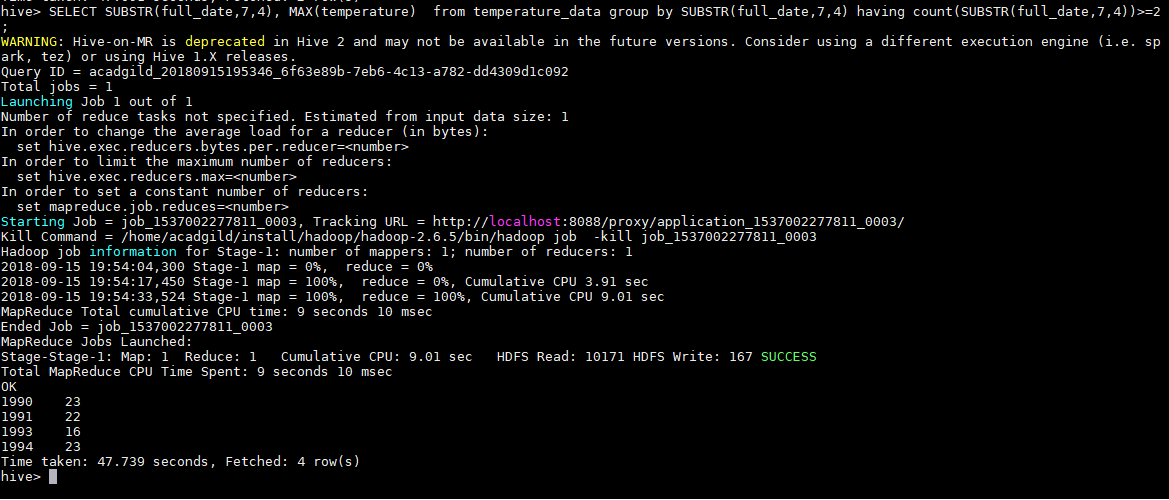
**Task 2(b): Calculate maximum temperature corresponding to every year from temperature\_data table**

**Command**: *SELECT SUBSTR(full\_date,7,4),MAX(temperature) from temperature\_data group by SUBSTR(full\_date,7,4);*

****

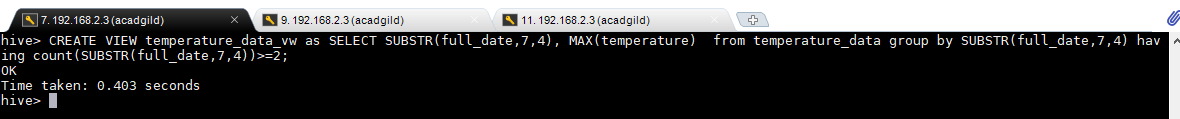
**Task 2(c): Calculate maximum temperature from temperature\_data table corresponding to those years which have at least 2 entries in the table.**

**Command**: *SELECT SUBSTR(full\_date,7,4), MAX(temperature) from temperature\_data group by SUBSTR(full\_date,7,4) having count(SUBSTR(full\_date,7,4))>=2;*

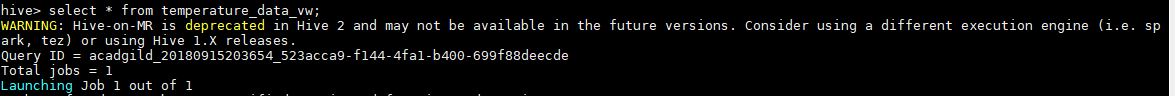
****

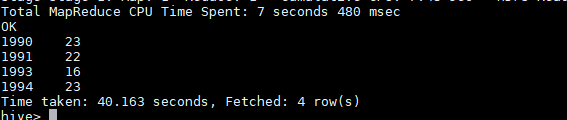
**Task 2(d) : Create a view on the top of last query, name it temperature\_data\_vw.**

**Command**: *CREATE VIEW temperature\_data\_vw as SELECT SUBSTR(full\_date,7,4), MAX(temperature) from temperature\_data group by SUBSTR(full\_date,7,4) having count(SUBSTR(full\_date,7,4))>=2;*

****

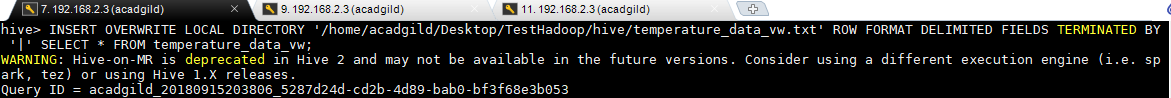
**To check the records in the view**

****

****

**Task 2(e): Export contents from temperature\_data\_vw to a file in local file system, such that each file is '|' delimited.**

Command : *INSERT OVERWRITE LOCAL DIRECTORY '/home/acadgild/Desktop/TestHadoop/hive/temperature\_data\_vw.txt' ROW FORMAT DELIMITED FIELDS TERMINATED BY '|' SELECT \* FROM temperature\_data\_vw;*



Output file path : */home/acadgild/Desktop/TestHadoop/hive/temperature\_data\_vw.txt'*

