**Session 2**

**Assignment 2.8**

Student Name: Karthik K

Course: Big Data Hadoop & Spark Training

# Problem Statement

1. Fetch date and temperature from temperature\_data where zip HIVE Commands is greater than 300000 and less than 399999.
2. Calculate maximum temperature corresponding to every year from temperature\_data table.
3. Calculate maximum temperature from temperature\_data table corresponding to those years which have at least 2 entries in the table.
4. Create a view on the top of last query, name it temperature\_data\_vw.
5. Export contents from temperature\_data\_vw to a file in local file system, such that each file is '|' delimited.

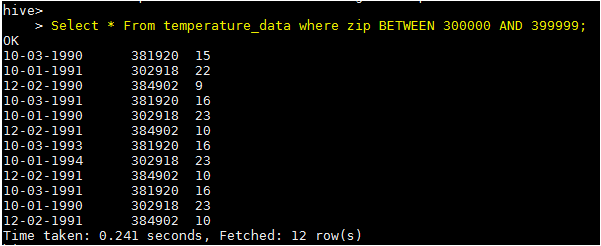
# Task1

Fetch date and temperature from **temperature\_data** where **zip** is greater than 300000 and less than 399999.

### HIVE Commands

***hive (custom) >Select \* From temperature\_data where zip BETWEEN 300000 AND 399999;***

### Output

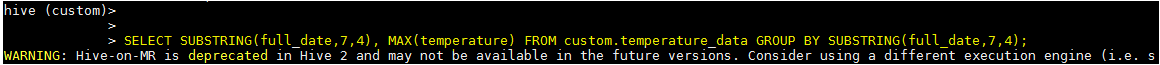


# Task2

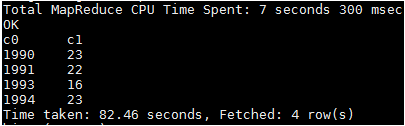
Calculate maximum temperature corresponding to every year from **temperature\_data** table.

### HIVE Commands

***hive (custom) > SELECT SUBSTRING(full\_date,7,4), MAX(temperature) FROM custom.temperature\_data GROUP BY SUBSTRING(full\_date,7,4);***



### Output



# TASK 3

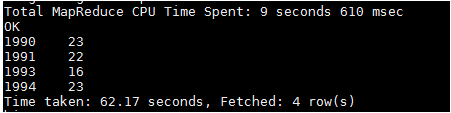
Calculate maximum temperature from **temperature\_data** table corresponding to those years which have at least **2** entries in the table.

### HIVE Commands

***hive(custom)>SELECT full\_date, MAX(t1.temperature) as temperature FROM (SELECT SUBSTRING(full\_date,7,4) full\_date, temperature FROM temperature\_data)t1 GROUP BY full\_date HAVING count(t1.full\_date)>=2;***



### Output

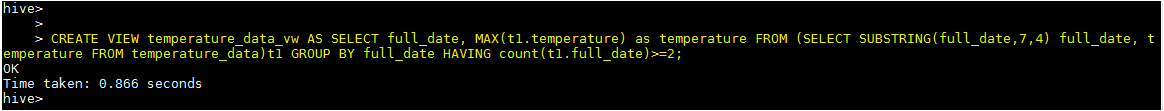


# TASK 4

Create a view on the top of last query, name it **temperature\_data\_vw.**

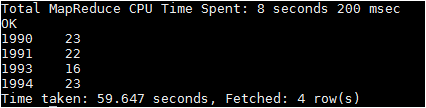
### HIVE Commands

***CREATE VIEW temperature\_data\_vw AS SELECT full\_date, MAX(t1.temperature) as temperature FROM (SELECT SUBSTRING(full\_date,7,4) full\_date, temperature FROM temperature\_data)t1 GROUP BY full\_date HAVING count(t1.full\_date)>=2;***





### Output



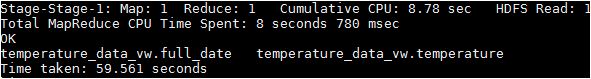
# TASK 5

Export contents from **temperature\_data\_vw** to a file in local file system, such that each file is '|' delimited.

### HIVE Commands

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





### Output

***cat /home/acadgild/hadoop/temperature\_data\_vw.txt/\****

