

Session 06: Hive Introduction

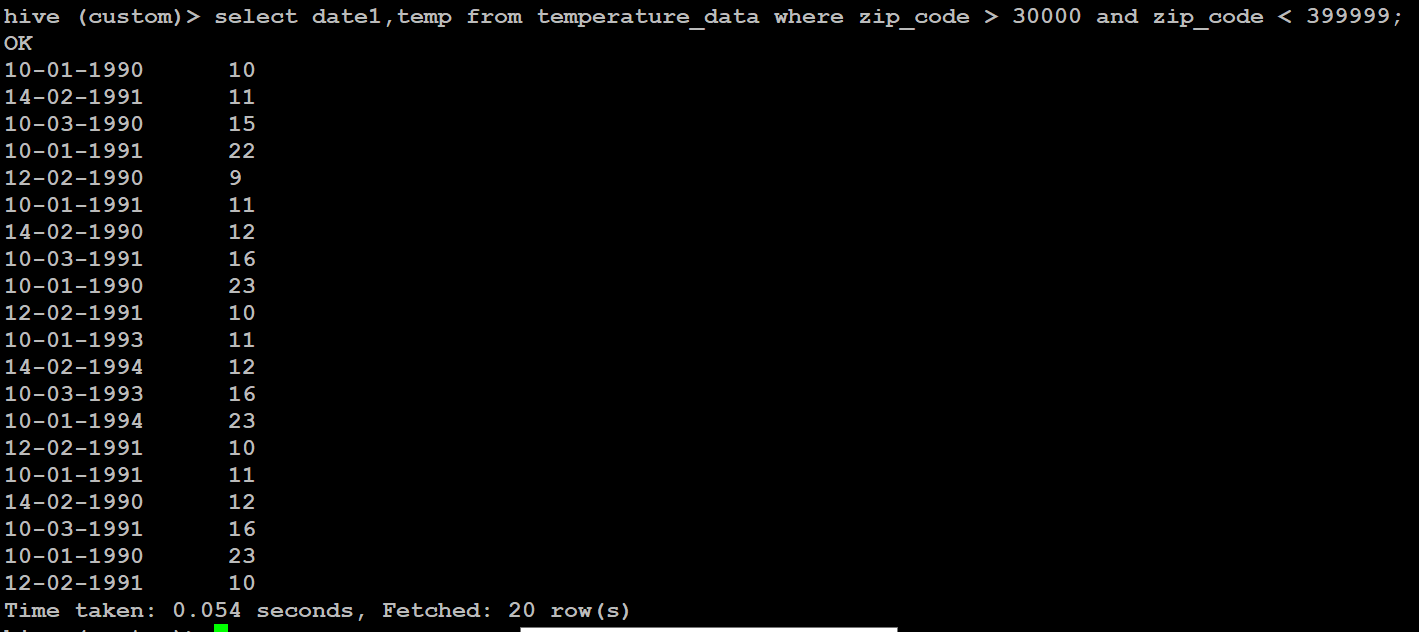
Assignment 2

*As from previous assignment we have already data in temperature\_data table*

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

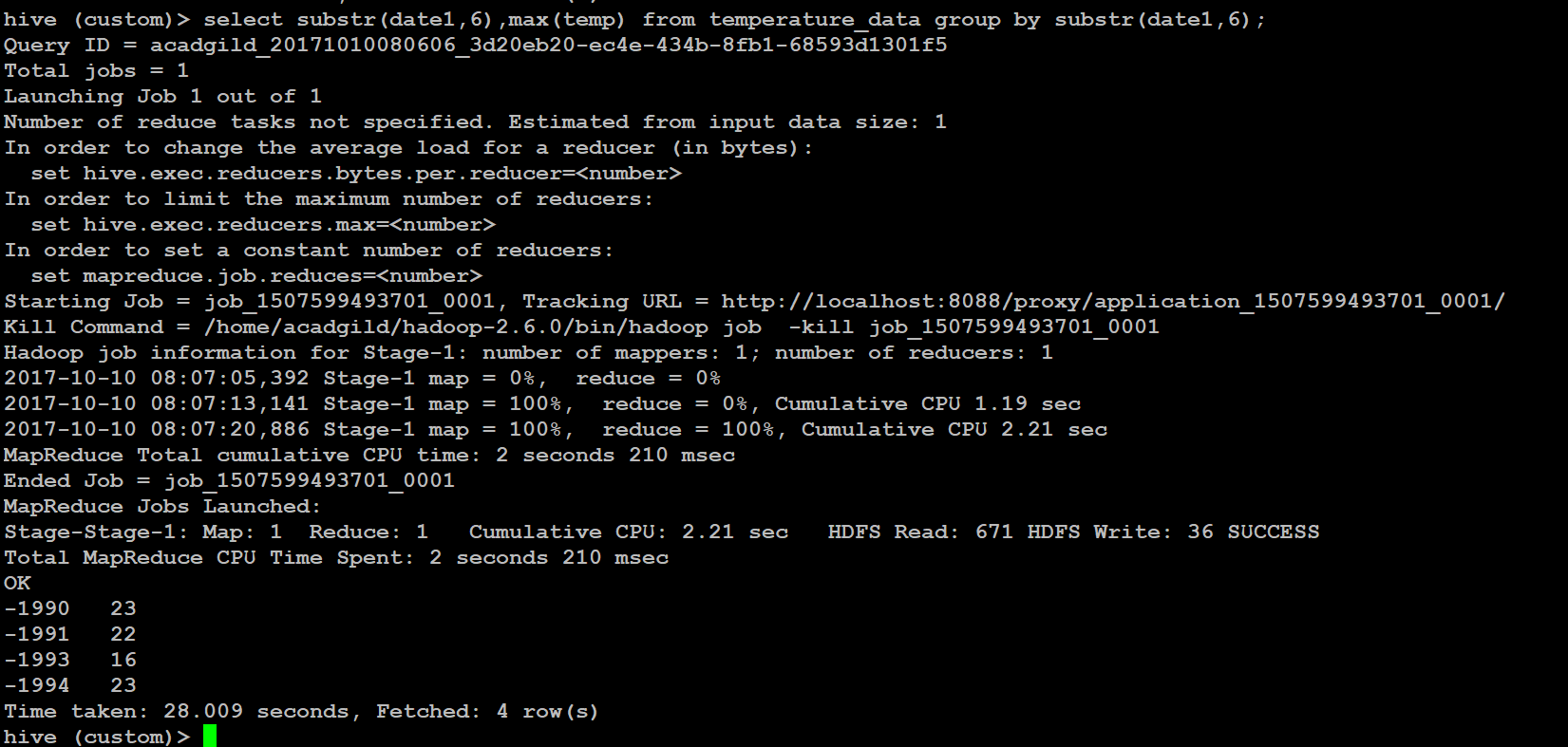
hive (custom)> select date1,temp from temperature\_data where zip\_code > 30000 and zip\_code < 399999;

OK



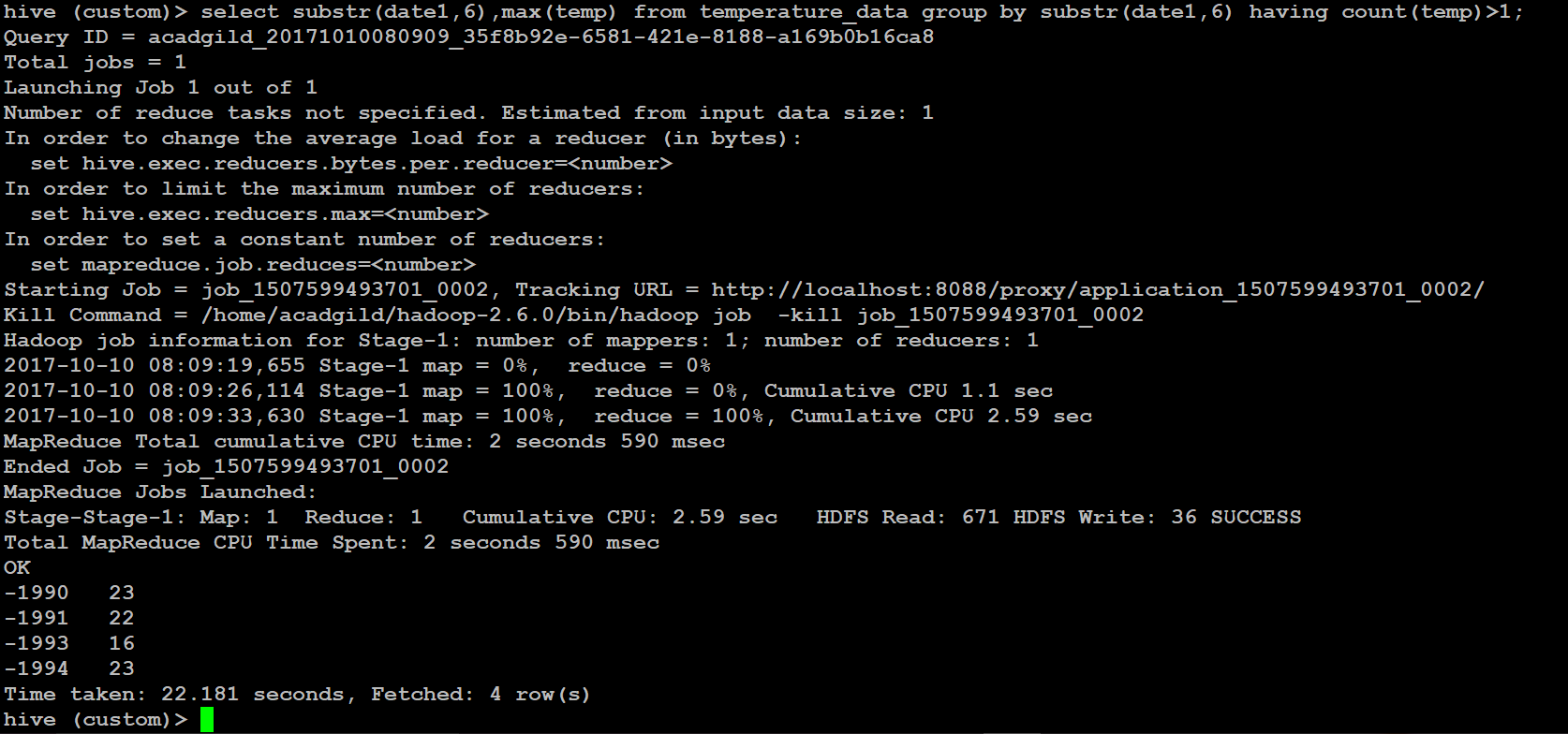
* *Calculate maximum temperature corresponding to every year from temperature\_data table*.

hive (custom)> select substr(date1,6),max(temp) from temperature\_data group by substr(date1,6);



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

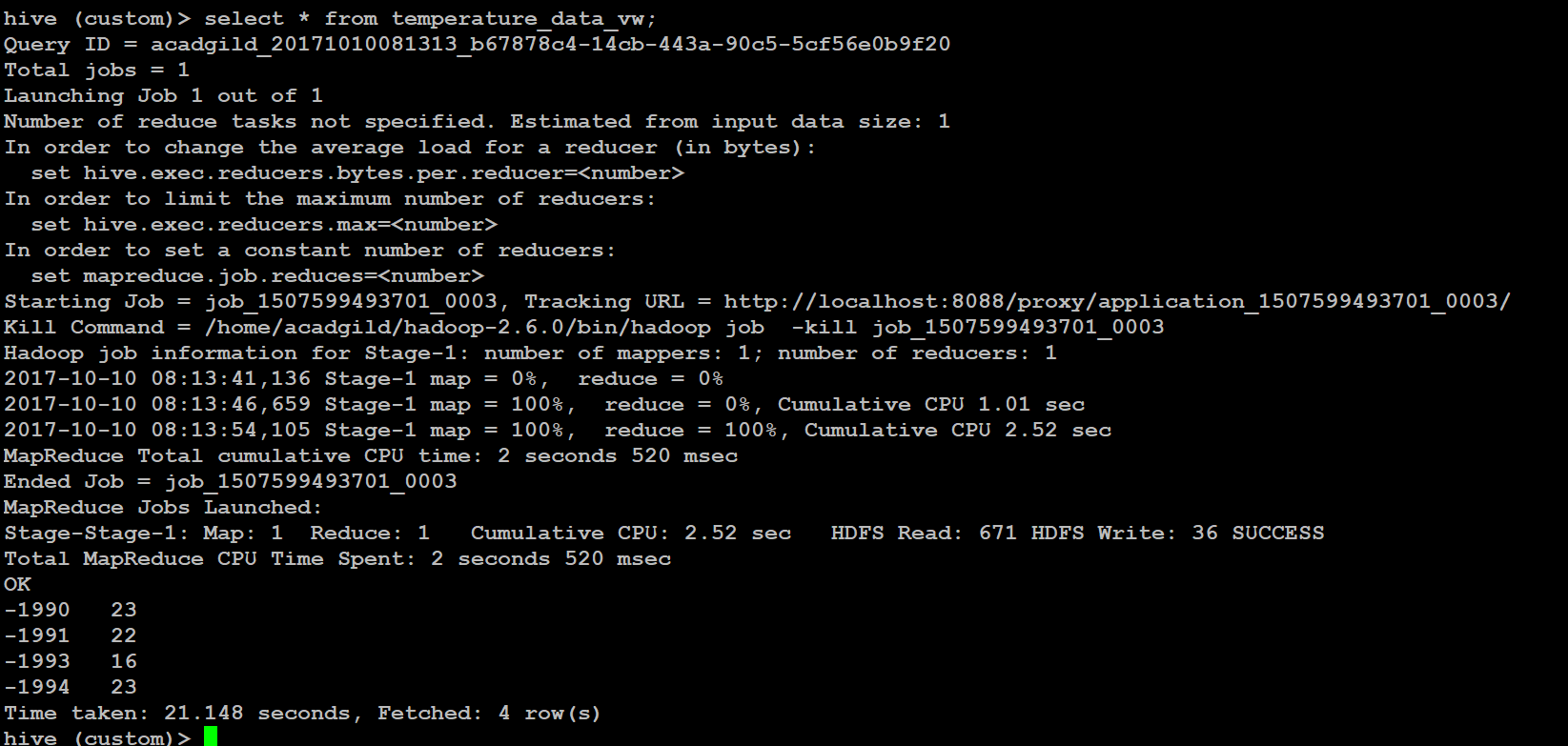
hive (custom)> select substr(date1,6),max(temp) from temperature\_data group by substr(date1,6) having count(temp)>1;



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

hive (custom)> create view temperature\_data\_vw select substr(date1,6),max(temp) from temperature\_data group by substr(date1,6) having count(temp)>1;

hive (custom)> select \* from temperature\_data\_vw;



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

hive (custom)> insert overwrite local directory '/home/acadgild/Desktop/hive\_out' row format delimited fields terminated by '|' select \* from temperature\_data\_vw;

