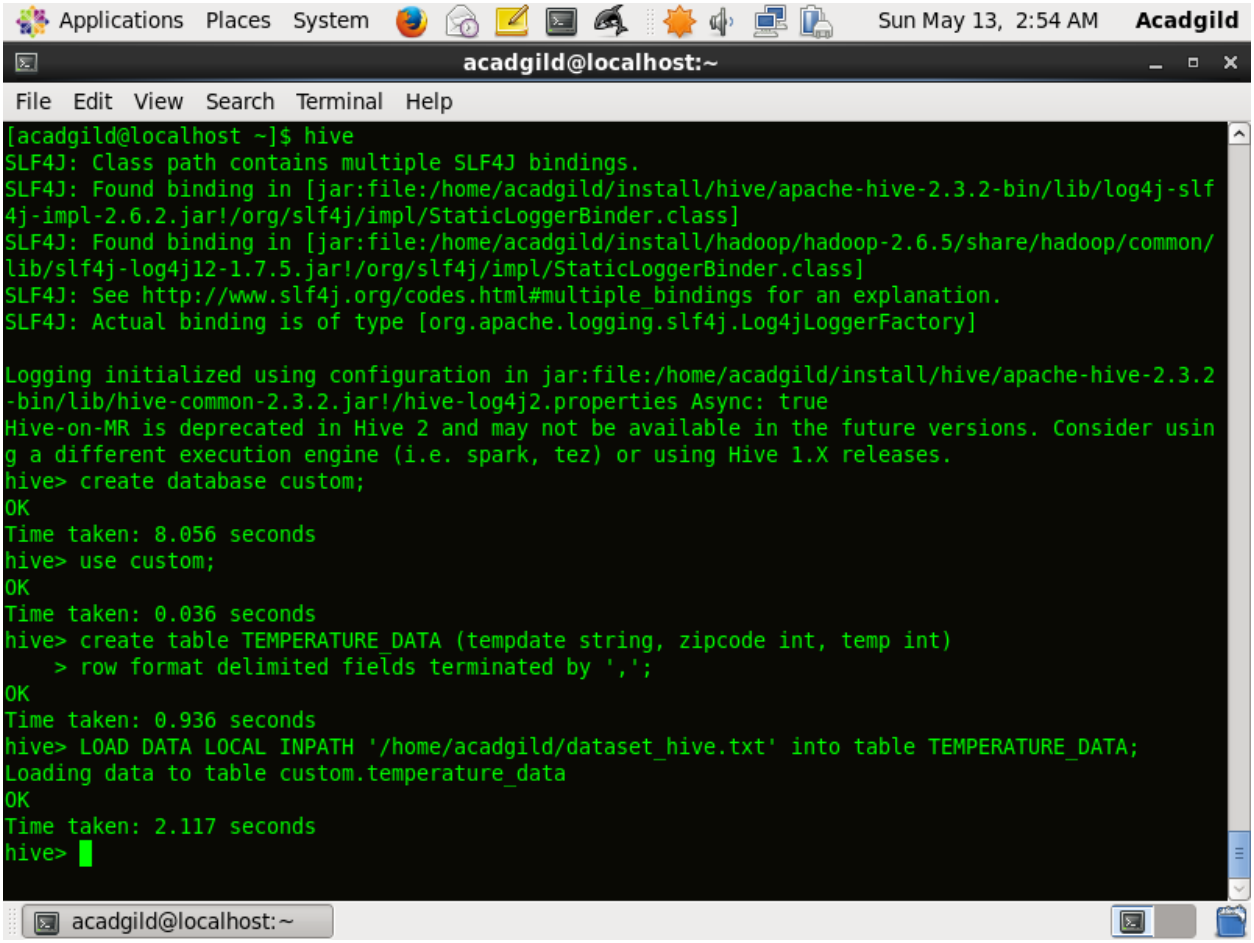
Creating and using a database:

Create database ‘custom’;
Use ‘custom’;

Creating the specified table:

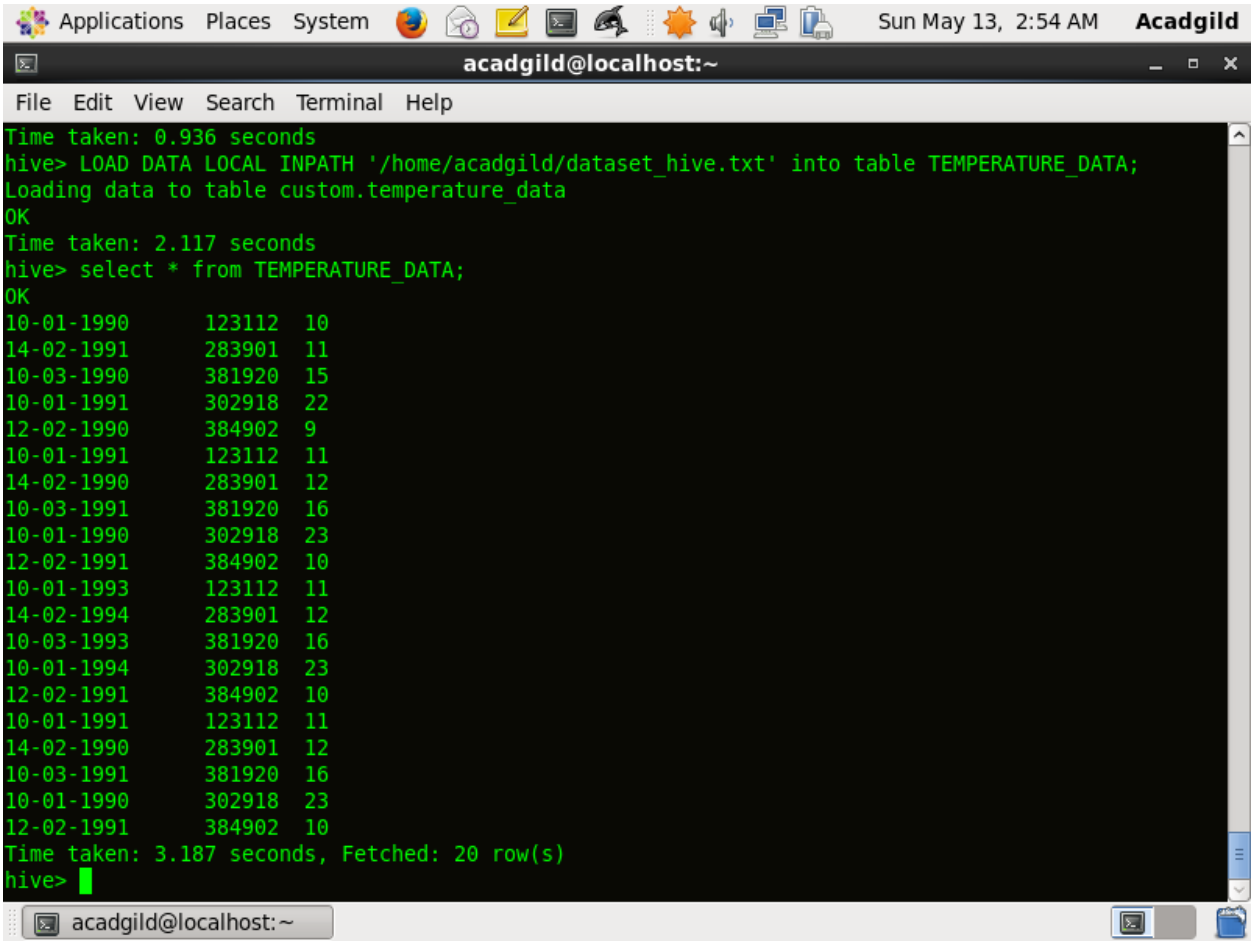
Create table temperature_data (tempdate string, zipcode int, temp int) row format delimited filelds terminated by ’,’;

Use command LOAD DATA LOCAL INPATH ‘<path>’ into <table-name>;



Check contents of the table after loading data by using command :

Select * form temperature_data

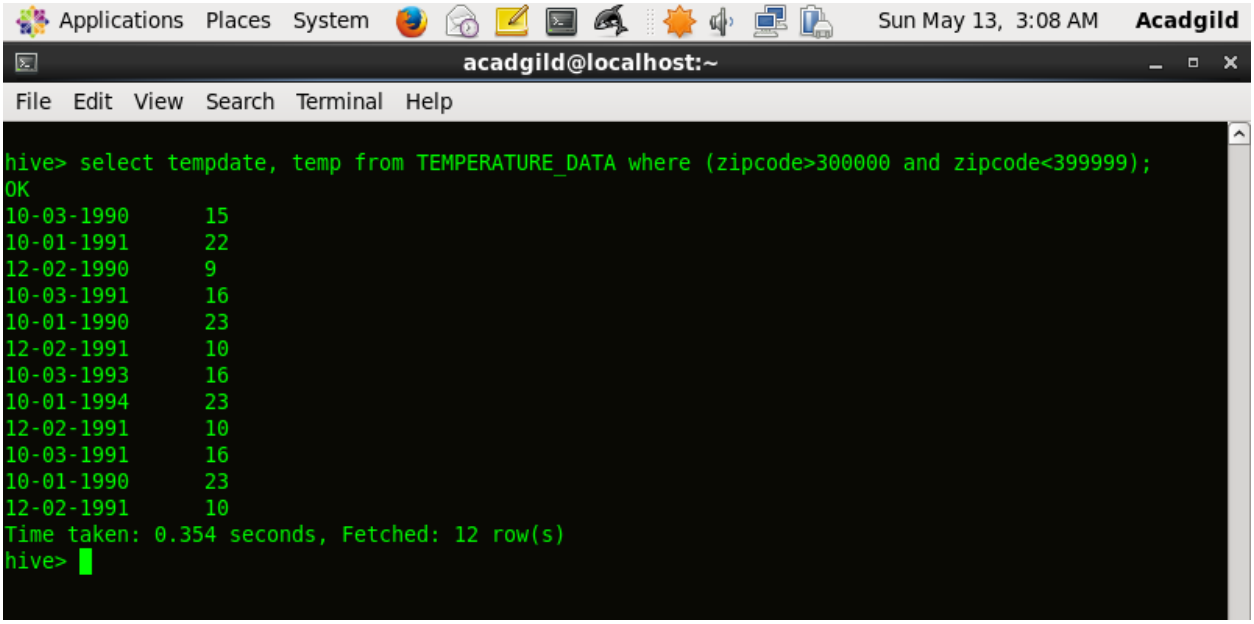


To Fetch date and temperature from temperature_data where zip code is greater than 300000 and less than 399999 use the where keyword to filter the records by using the condition as mentioned

```
hive> select tempdate, temp from TEMPERATURE_DATA where (zipcode>300000 and zipcode<399999);
```

I am selecting tempdate and temp because I want only those fields in the output.

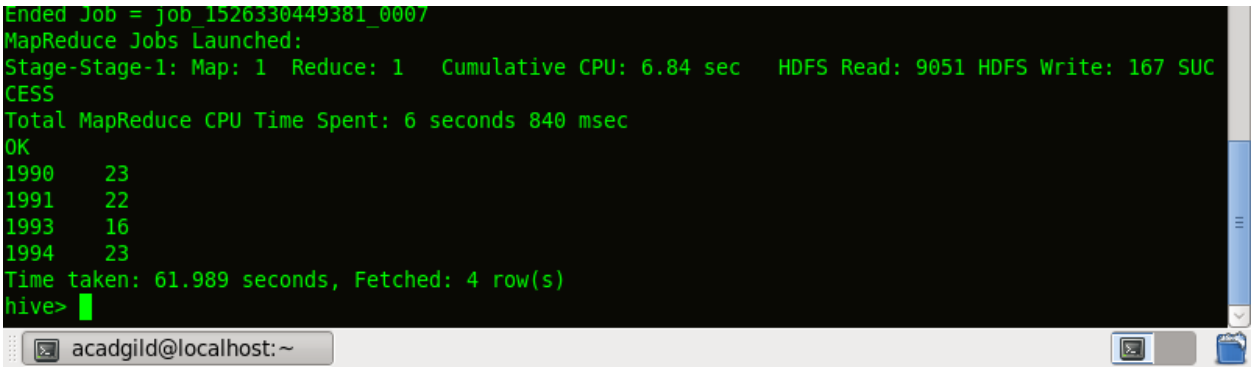
Output:



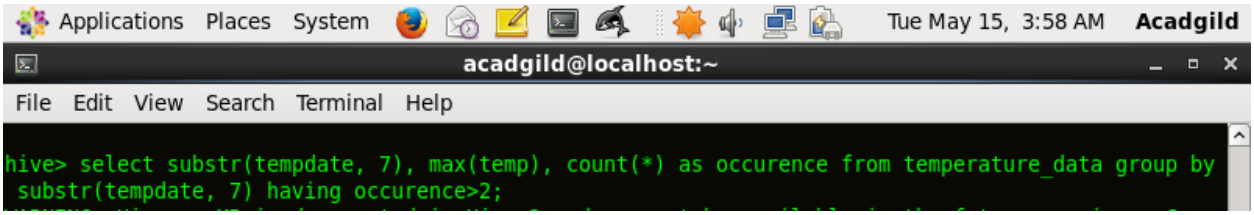
To find maximum temperature corresponding to every year from temperature_data table we have to group tempdate column and find the maximum temperature from each group.

We will use built in function substr() for returning a part of the tempdate string which will give the year and max() function to find the max vaue form temp column. Use the below command.

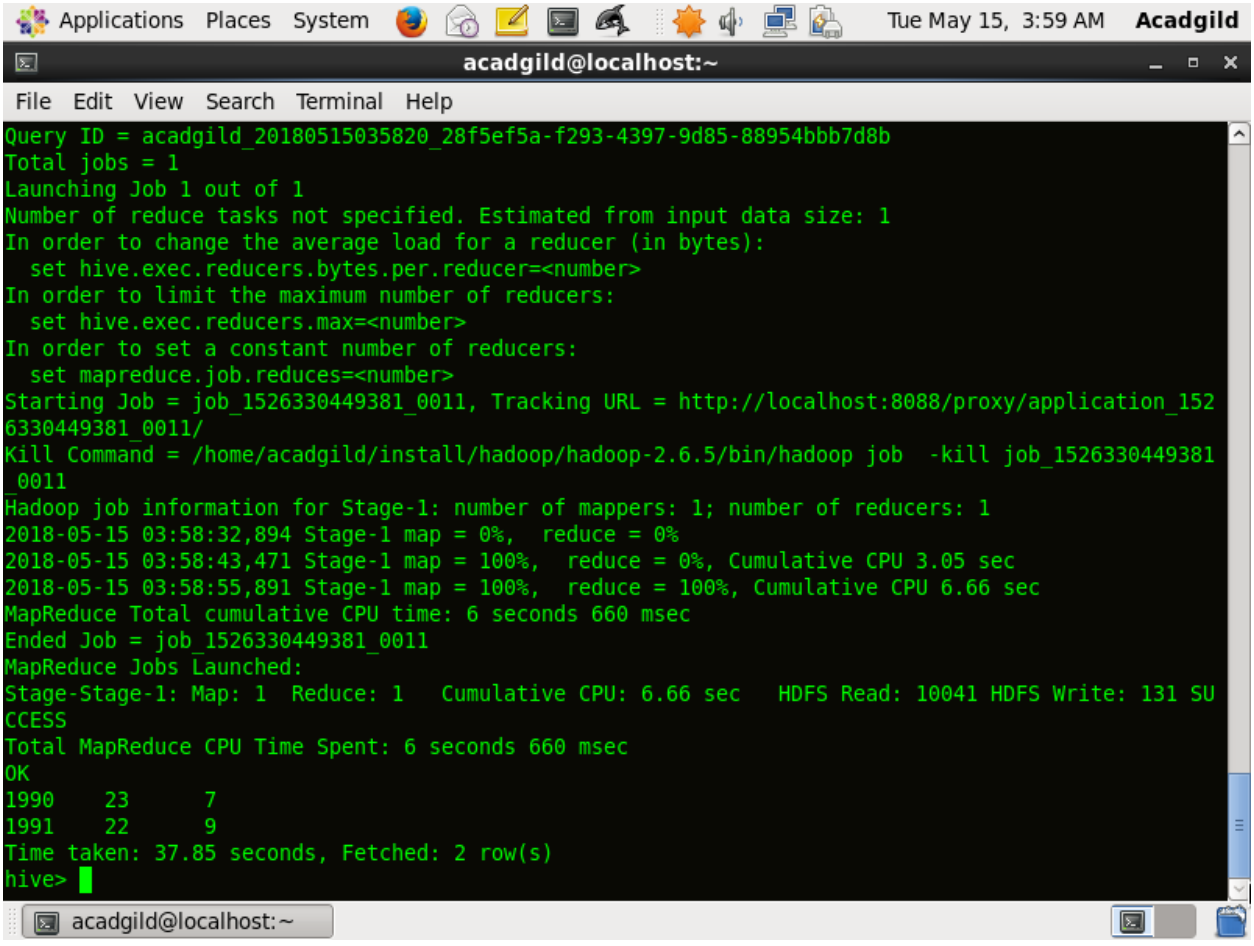
Output:



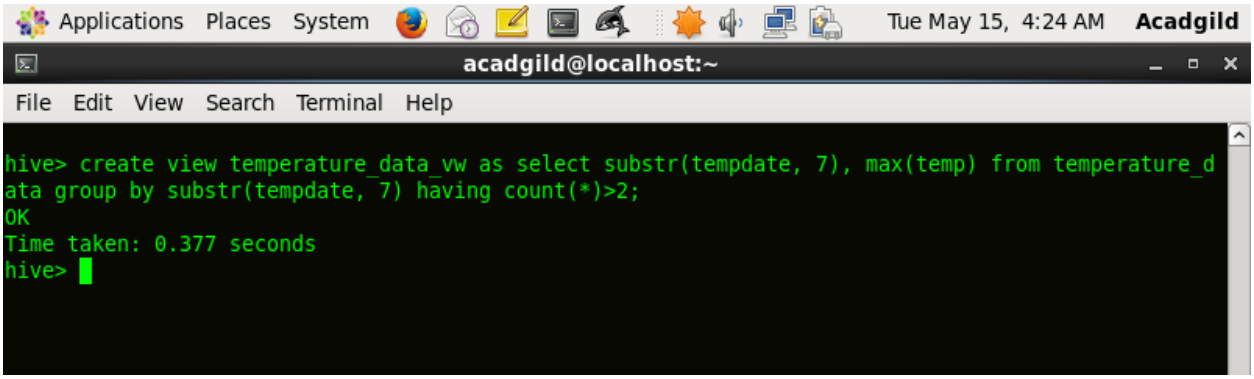
To Calculate maximum temperature from temperature_data table corresponding to those years which have at least 2 entries in the table, I add the having clause which is used to filter the grouped output. Here I assign the number of entries to occerence variable which will store the output of the coun() function and having condition will check whether the number of entries are more than 2.



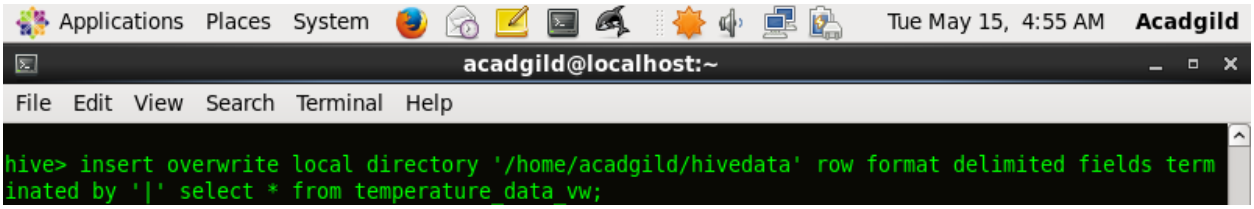
Output:



Create view <view name> as <query>
So temperature_data_vw is my view name and using my last query to store in the view.



To export content to a local file system with '|' delimiters use command :



Lets check the contents of our hivedata local directory: