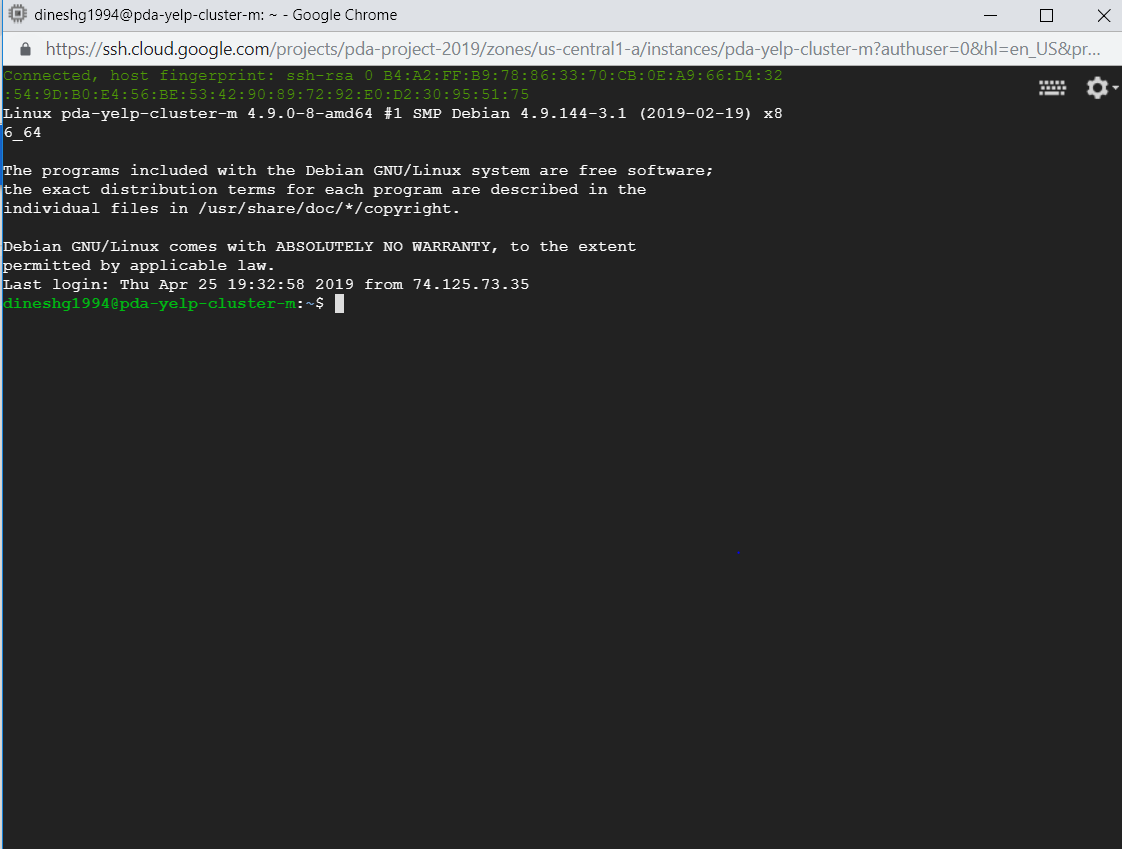
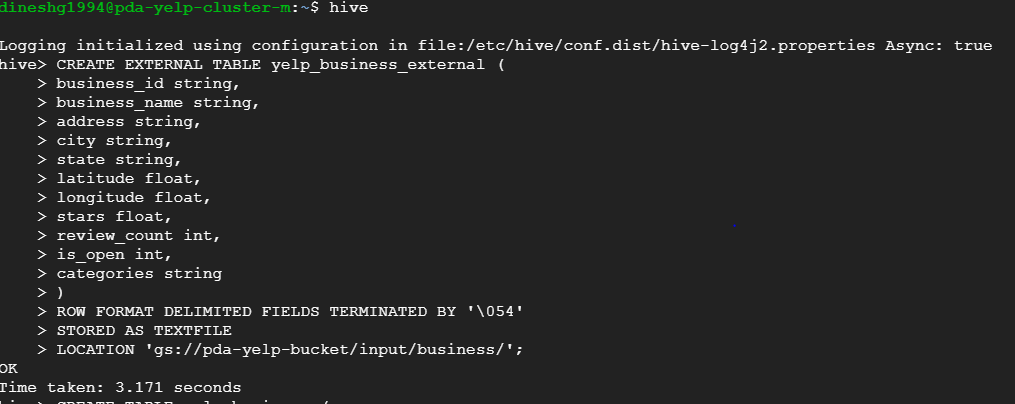
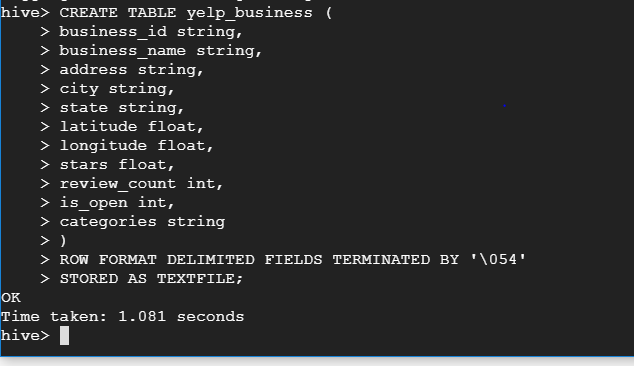
Task 1 in Hive: Top 15 business with most reviews.

1. Cluster with anaconda and Jupiter, Bucket creation already done in gcp setup
2. SSH the cluster

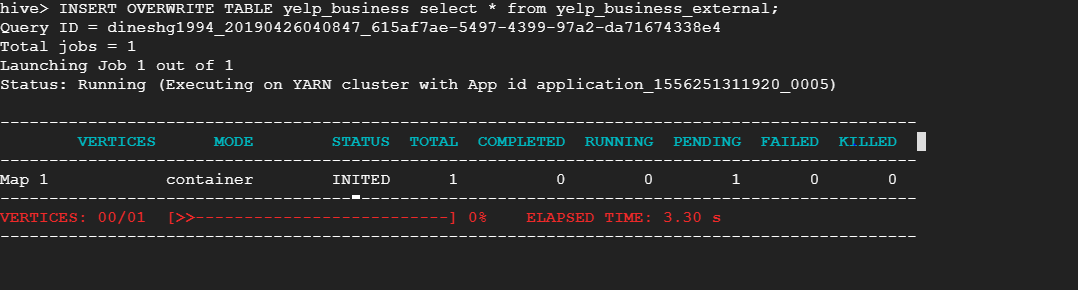


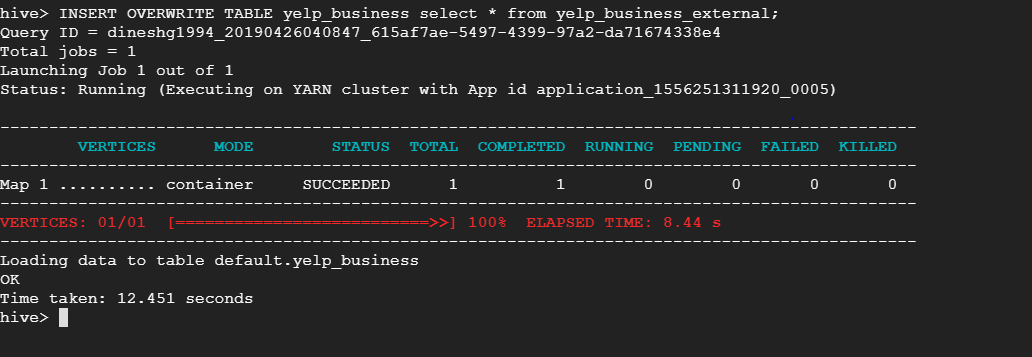
1. Table creation…first external table then hive table



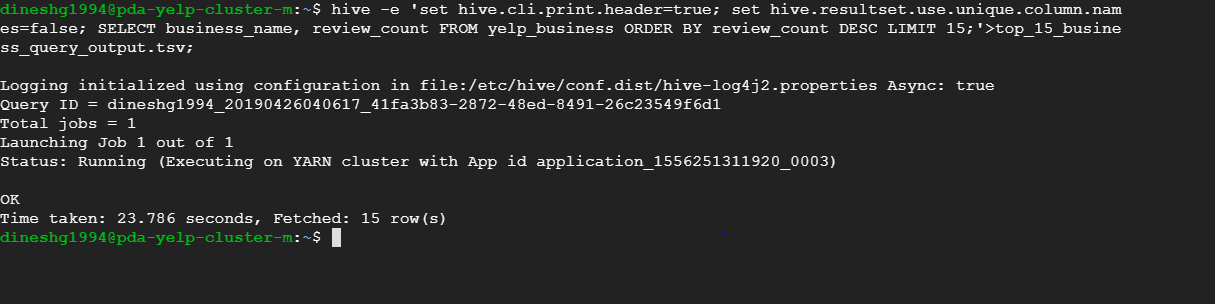


1. Inserting data from external table to hive table

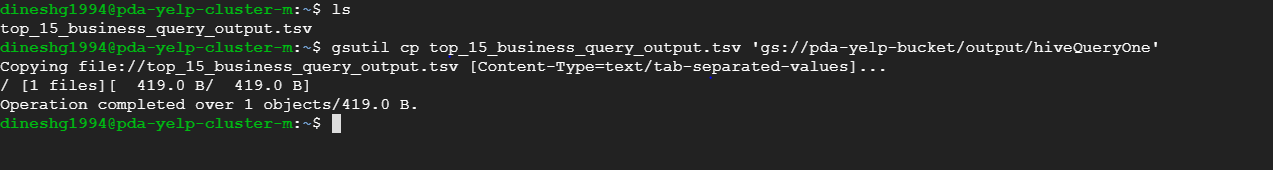




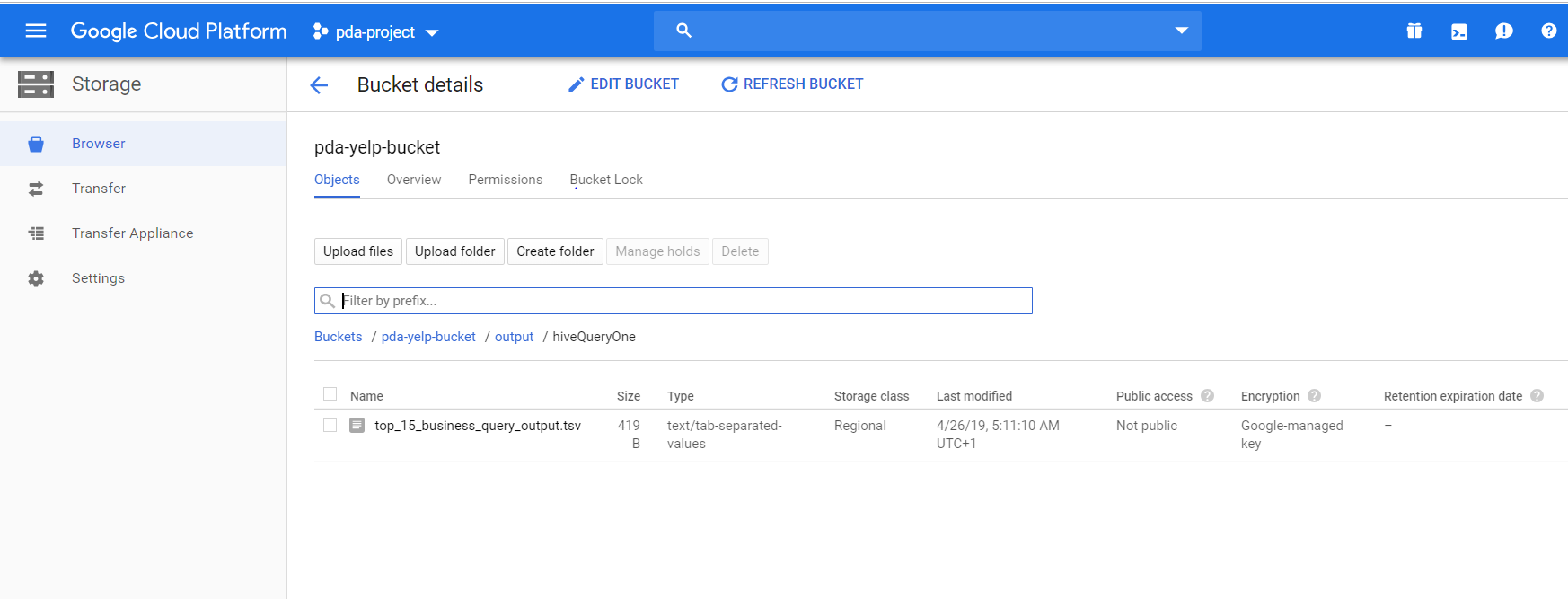
1. Executing command and saving it as csv file



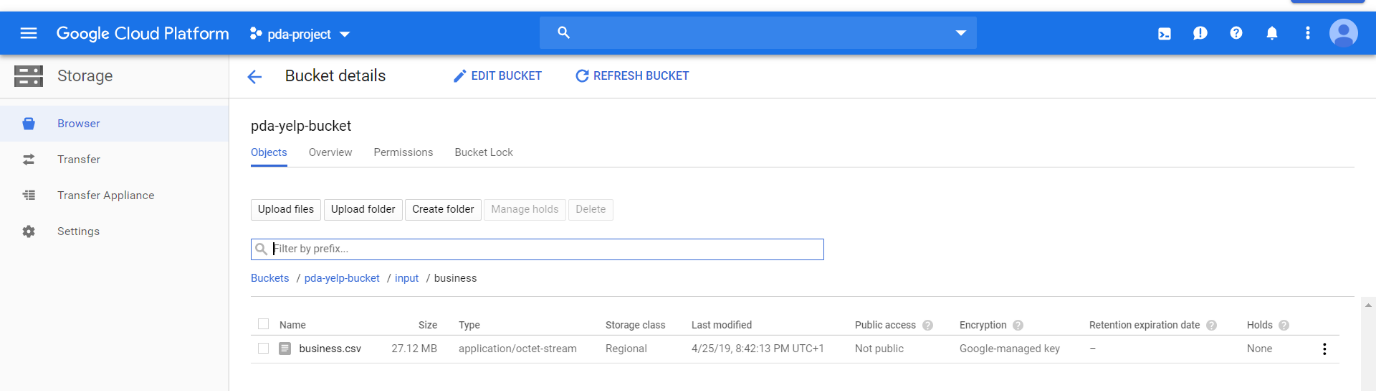
1. CSV file created and then Copying csv file from local to google file system



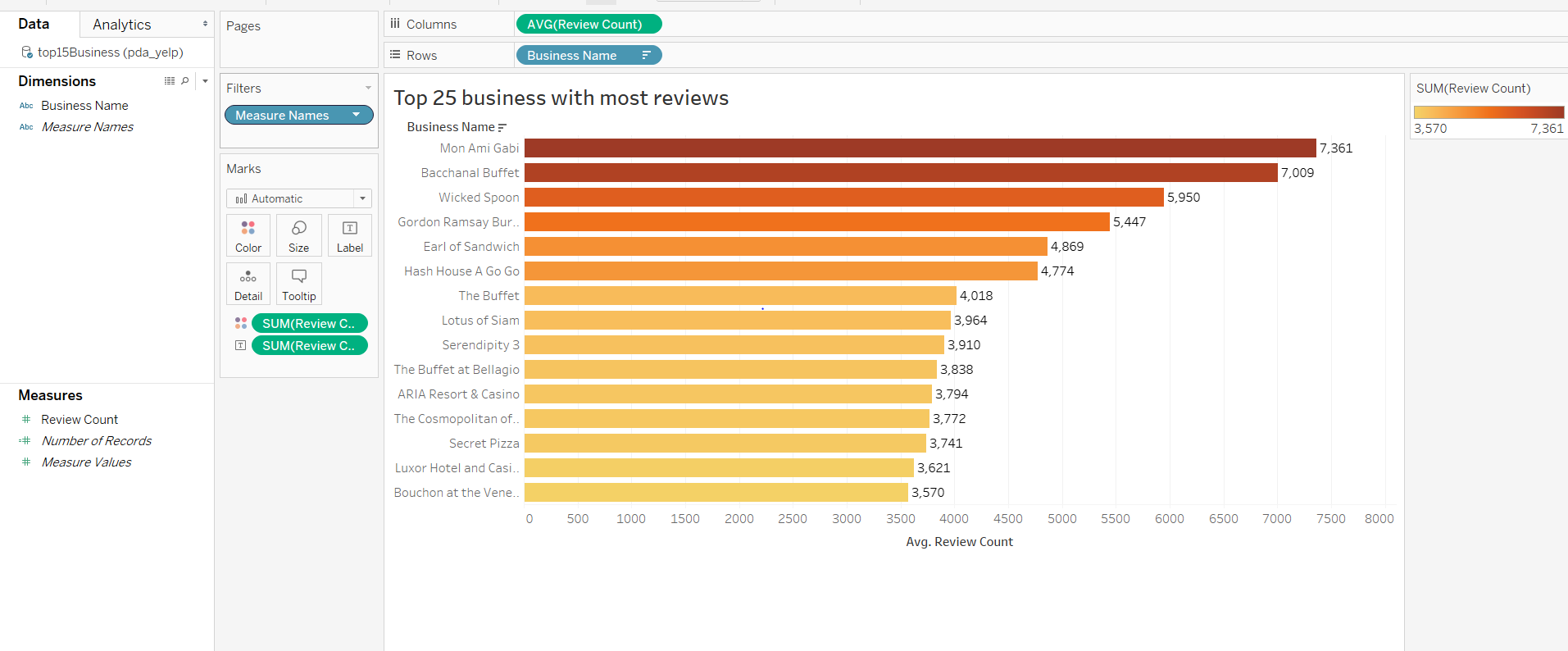
1. File moved into google file system

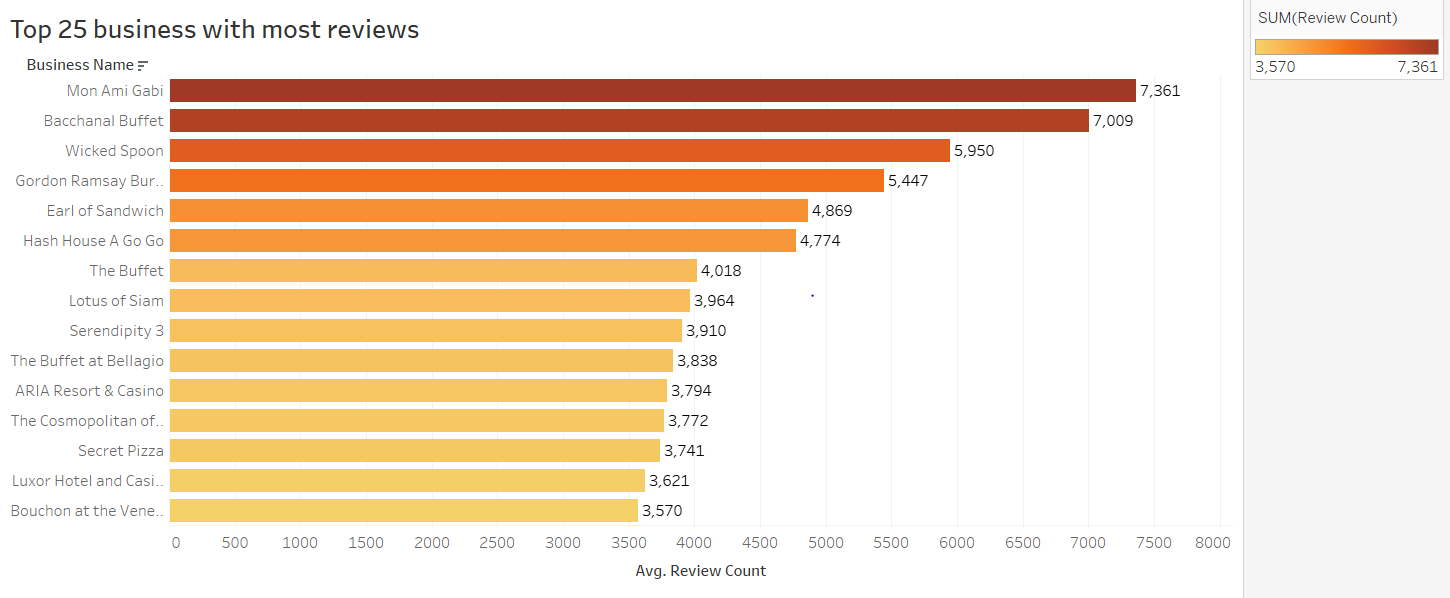


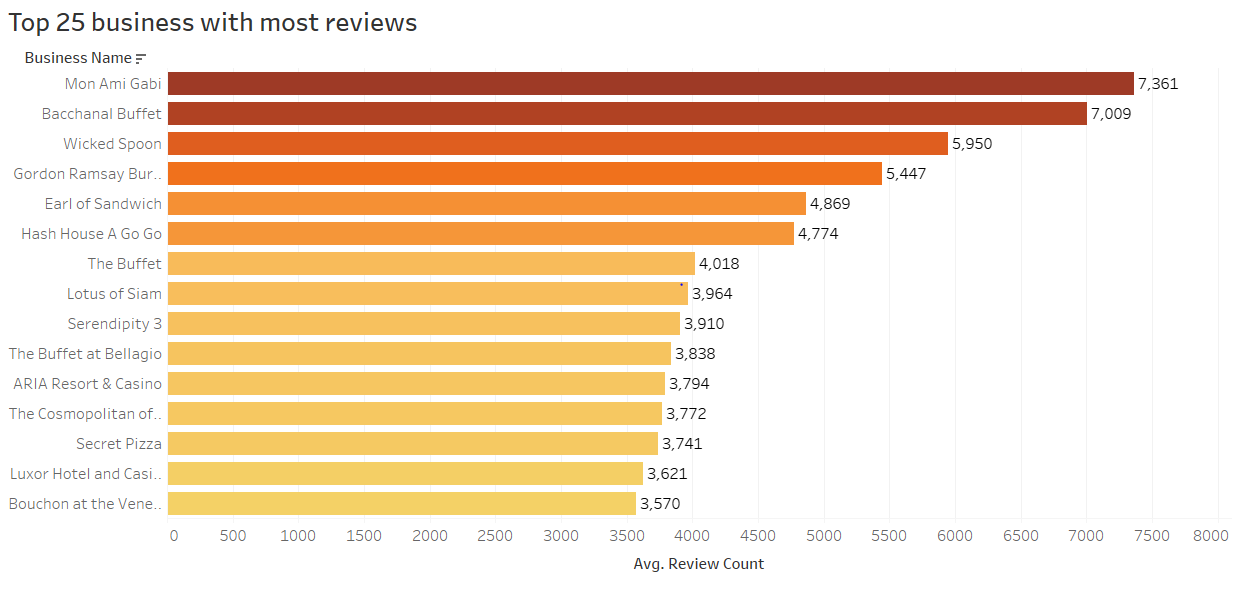
1. Input files in google file system



1. Steps to create big query table and connect big tableau with big query is followed same as in pyspark document task 1 steps 9-11.
2. Final visualisation

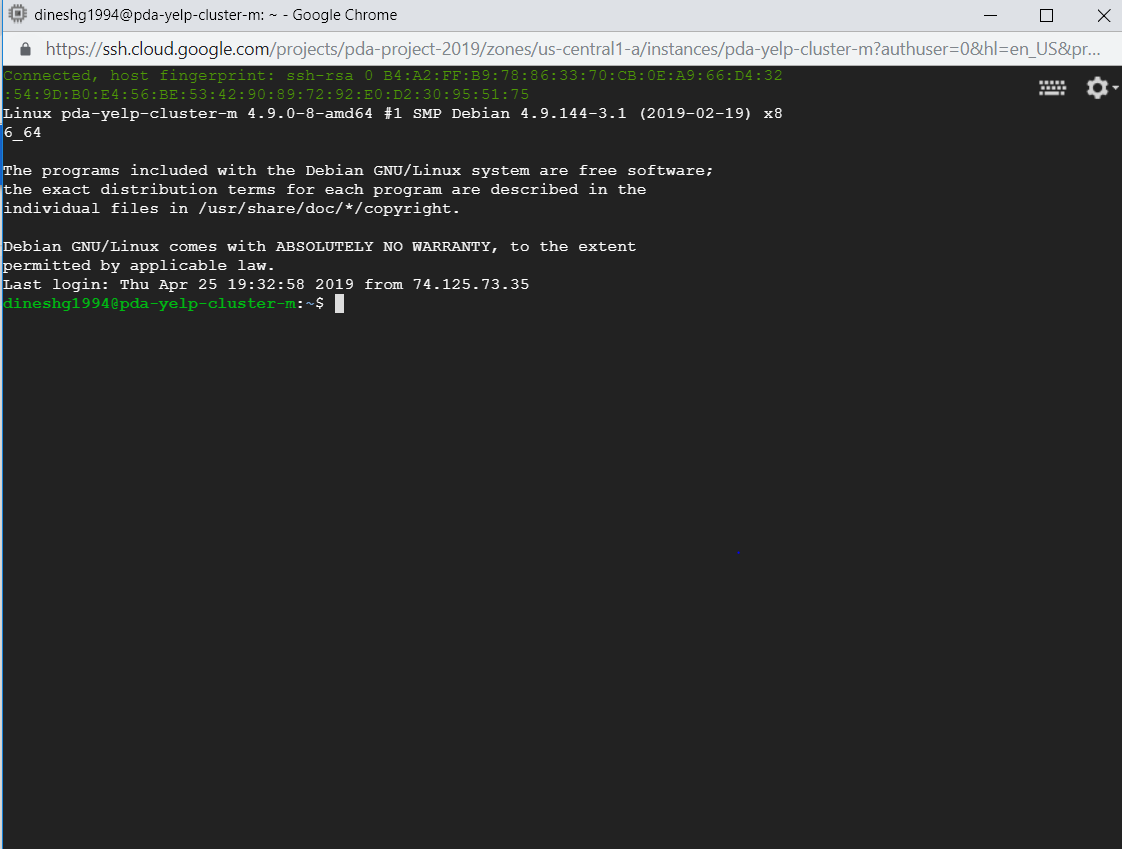




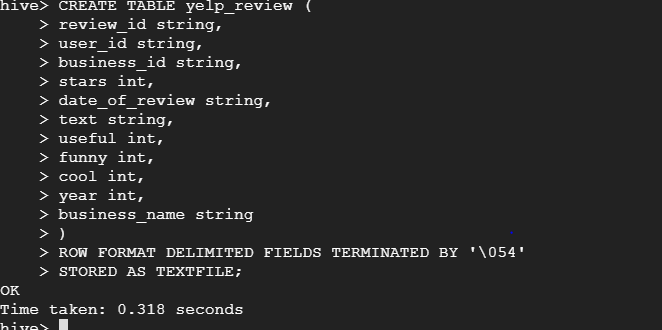


Task 2 in Hive: Top 15 coolest restaurants

1. Cluster with anaconda and Jupiter, Bucket creation already done in gcp setup
2. SSH the cluster

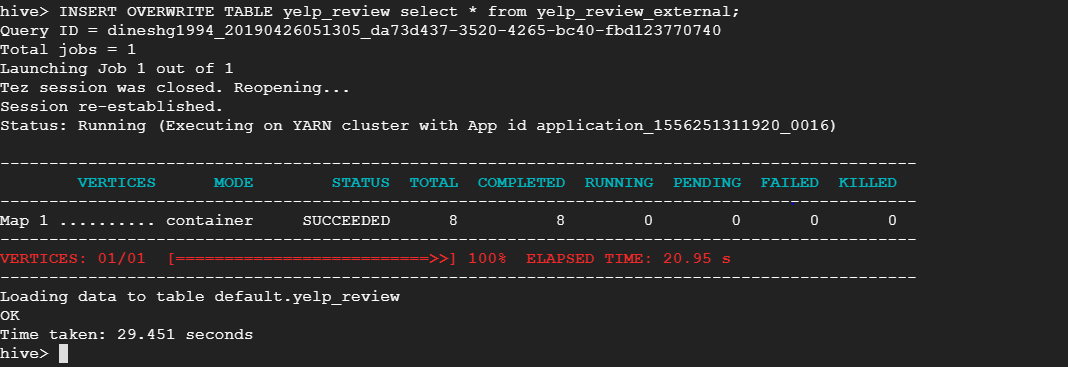


1. Table creation…first external table then hive table

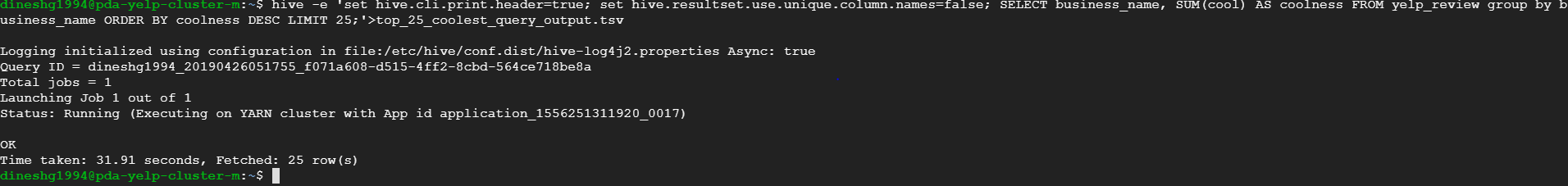




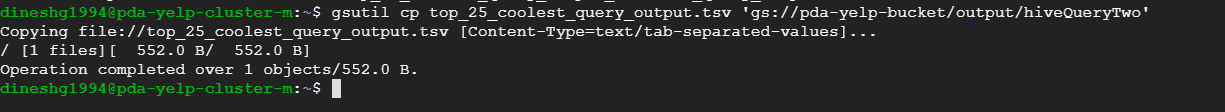
1. Inserting data from external table to hive table

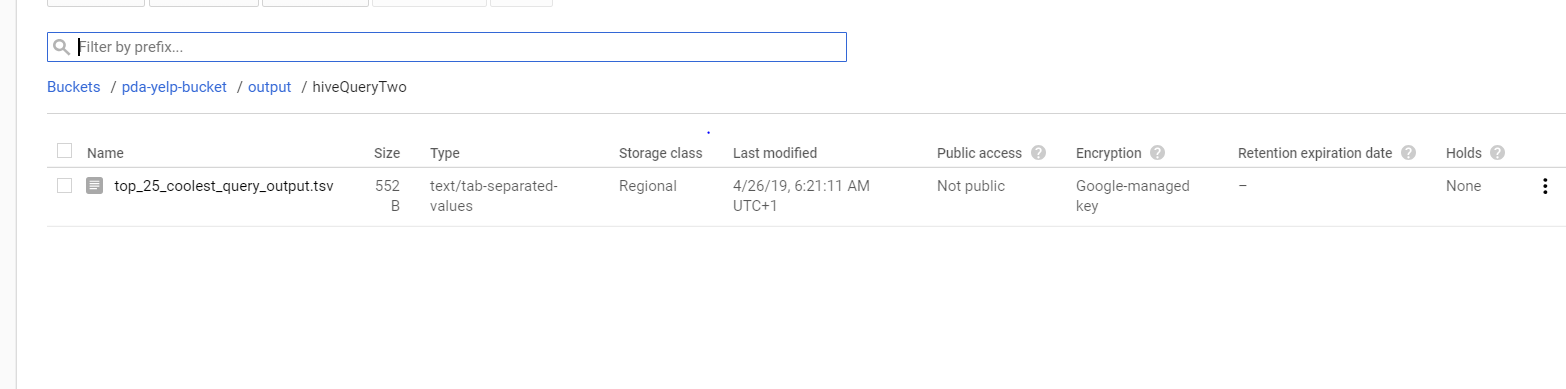


1. Executing command and saving it as csv file



1. Moving csv to google file system





1. Steps to create big query table and connect big tableau with big query is followed same as in pyspark document task 1 steps 9-11.
2. Final Visualisation

