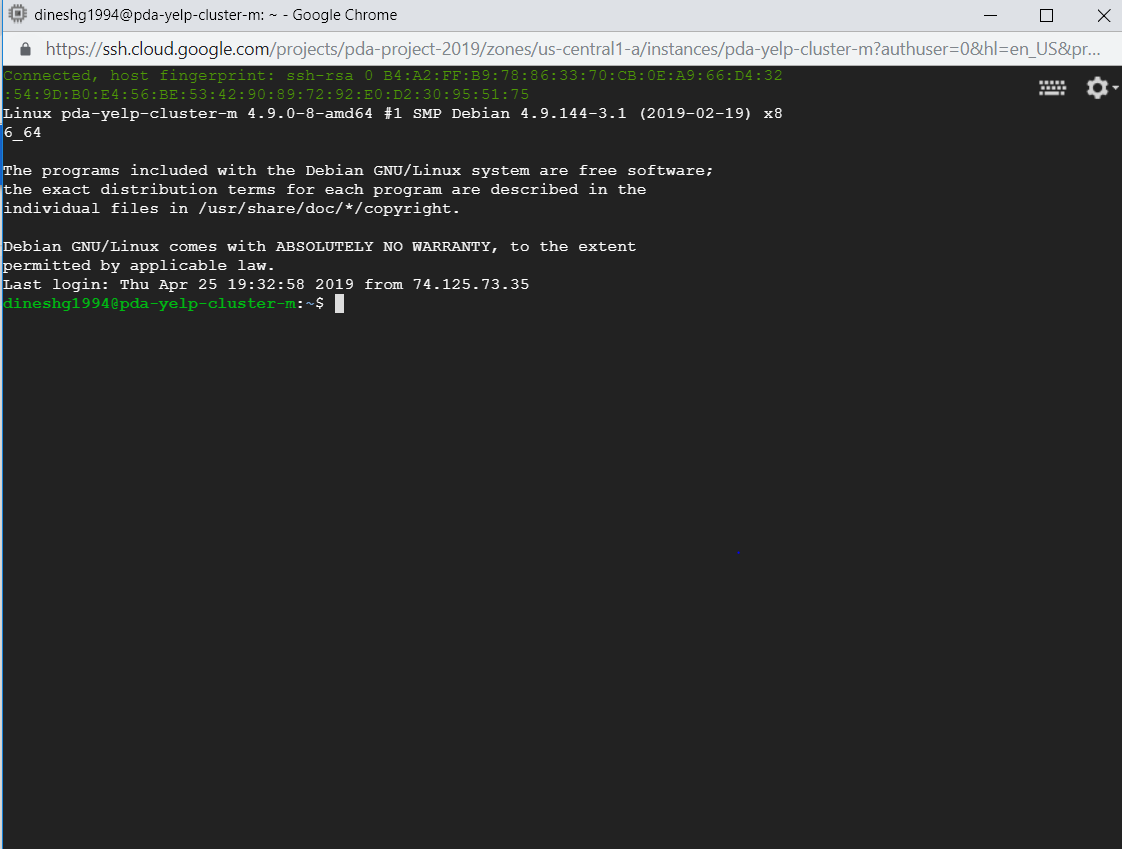
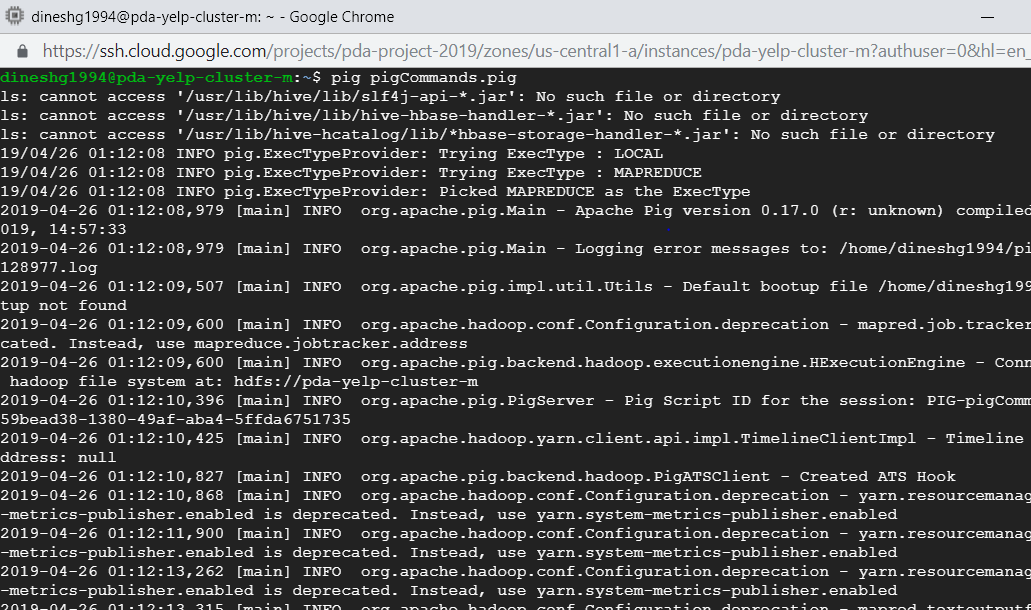
Task 1 in pig: Find average stars grouped by city and category and also sorted categories in ascending and stars in descending order.

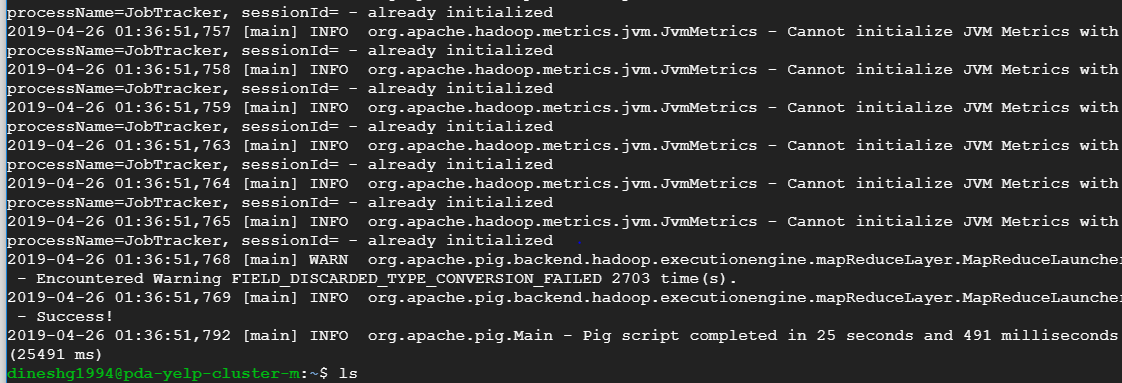
1. Cluster with anaconda and Jupiter, Bucket creation already done in gcp setup
2. SSH the cluster



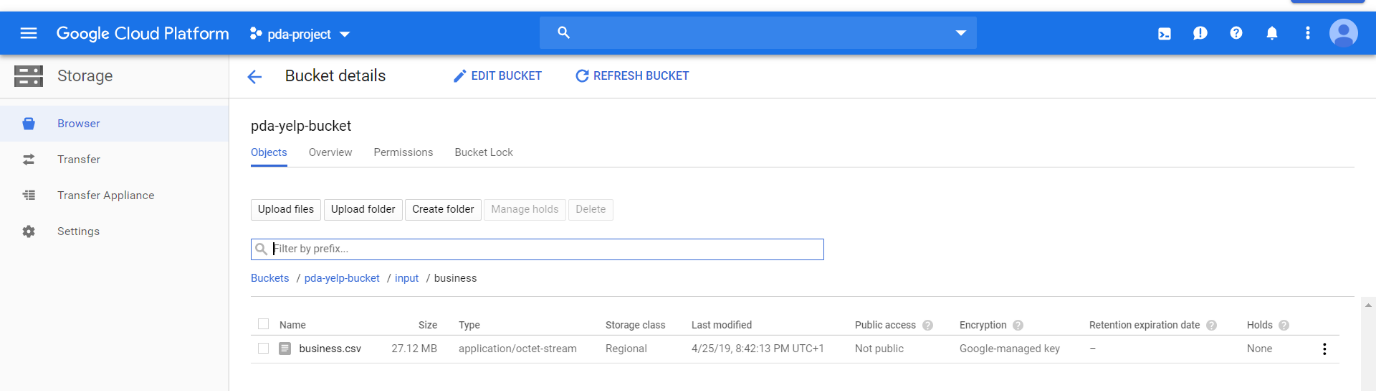
1. Run pig and execute pigCommands.pig



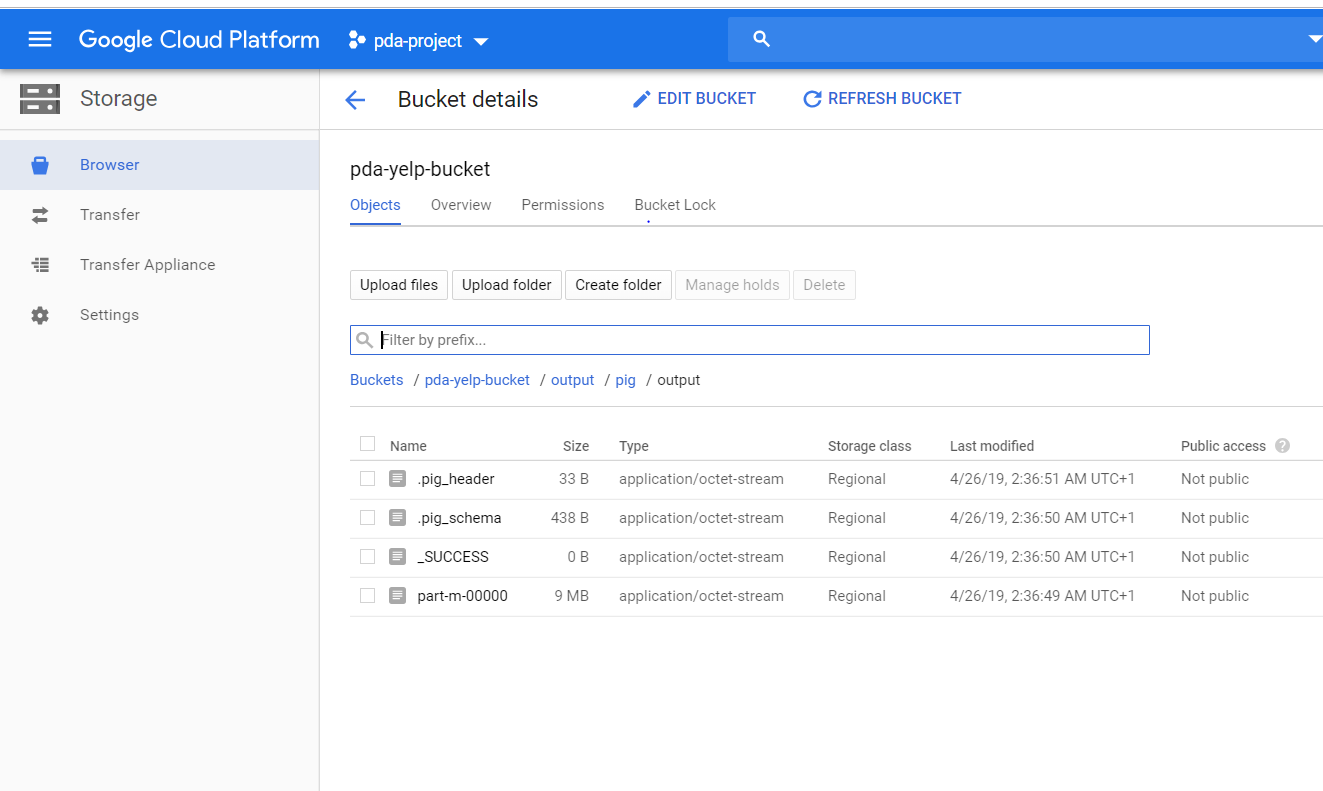
1. Successfully pig command executed



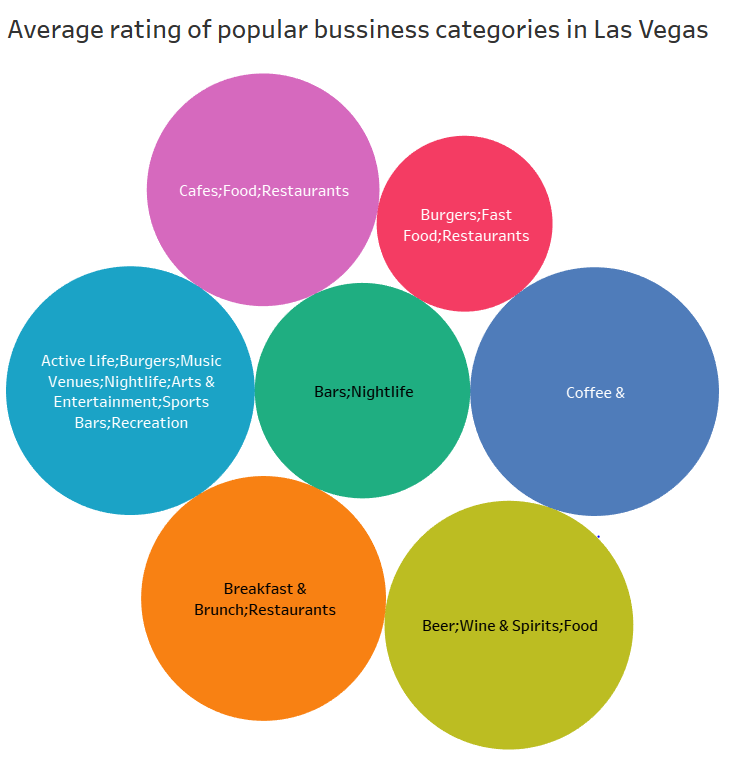
1. Input file on google file system

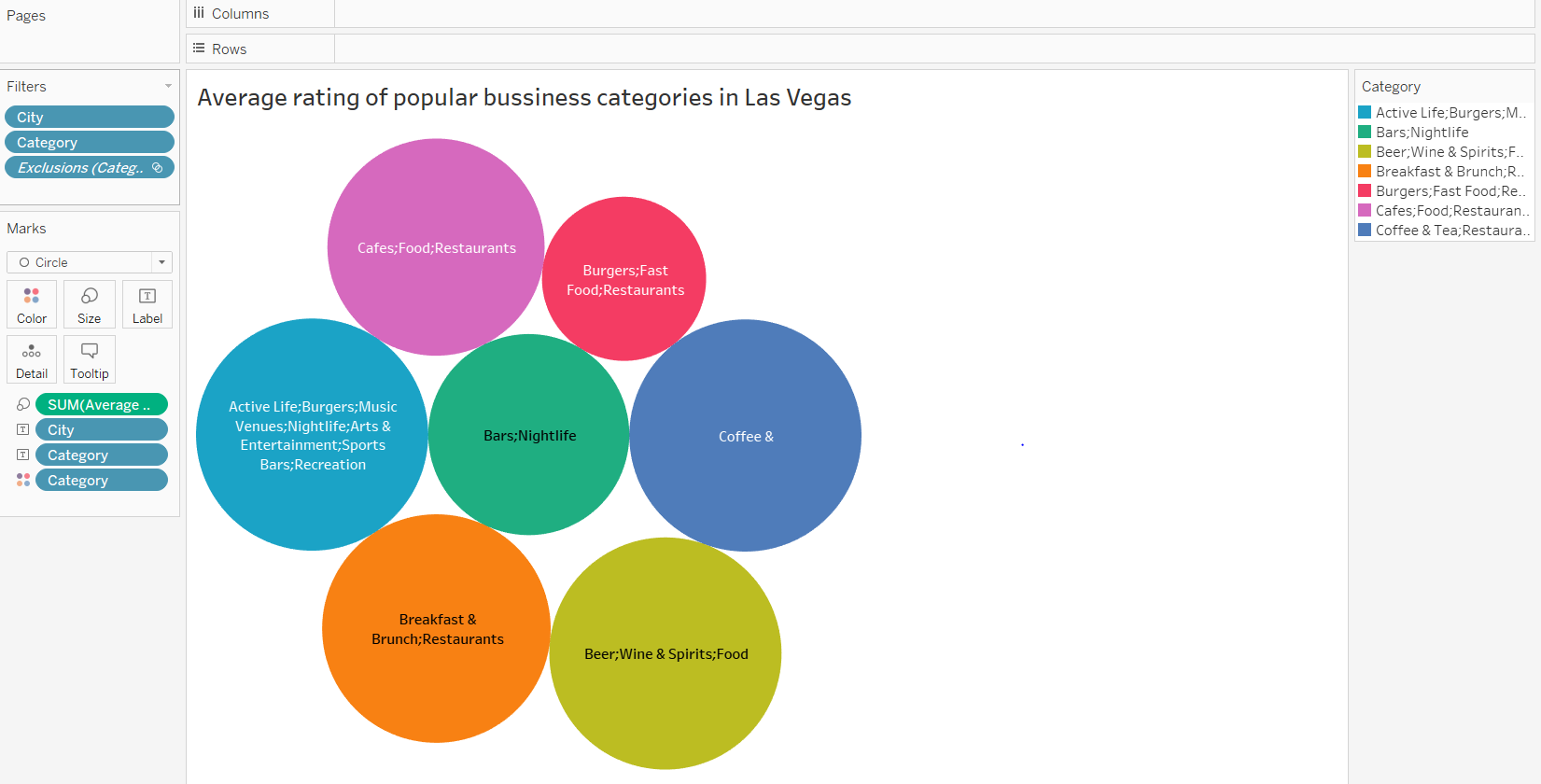


1. Pig output generated on google file system



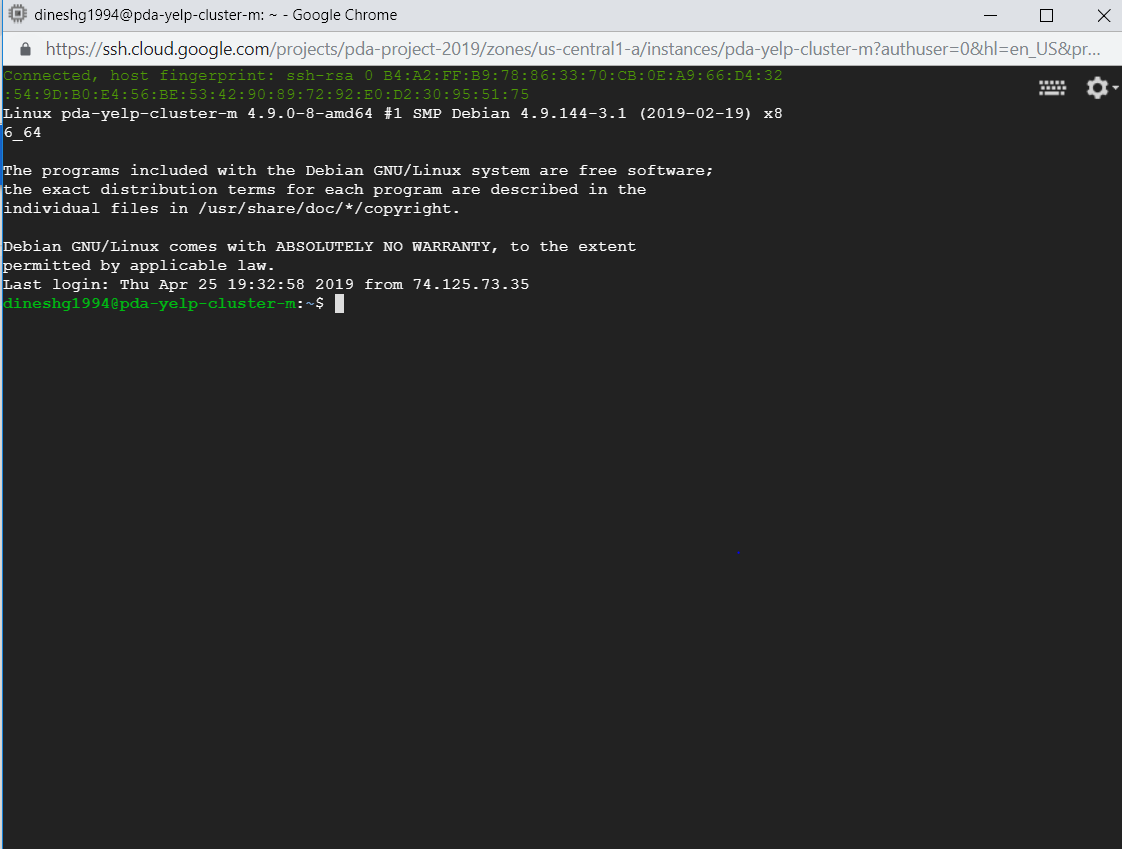
1. Steps to create big query table and connect big tableau with big query is followed same as in pyspark document task 1 steps 9-11.
2. Final visualisation



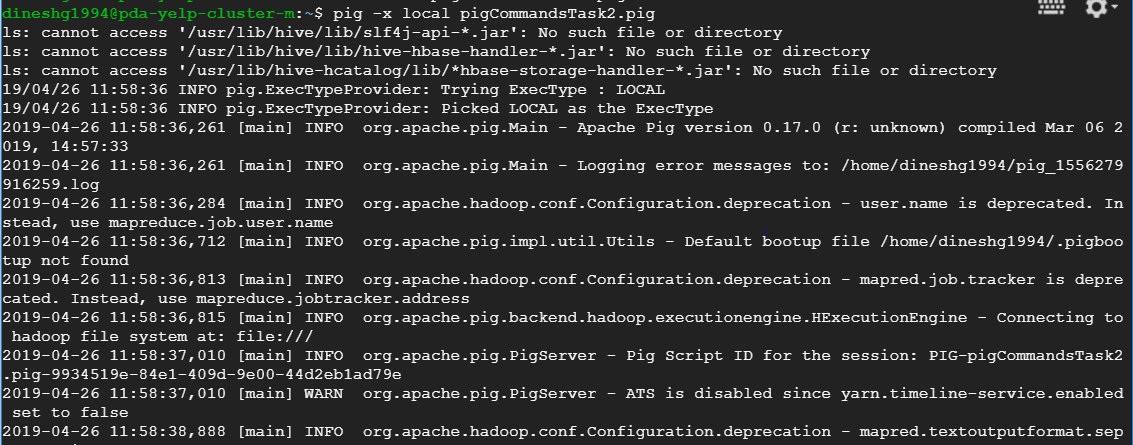


Task 2 in pig: Find average stars grouped by city and category and also sorted categories in ascending and stars in descending order.

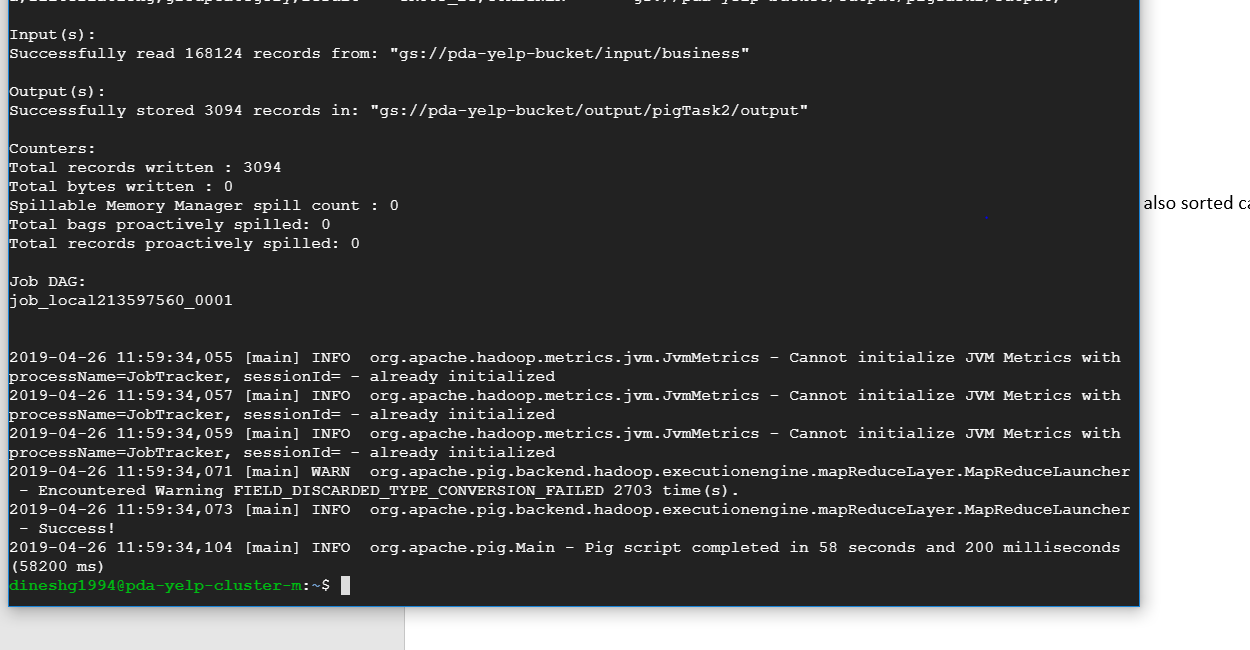
1. Cluster with anaconda and Jupiter, Bucket creation already done in gcp setup
2. SSH the cluster



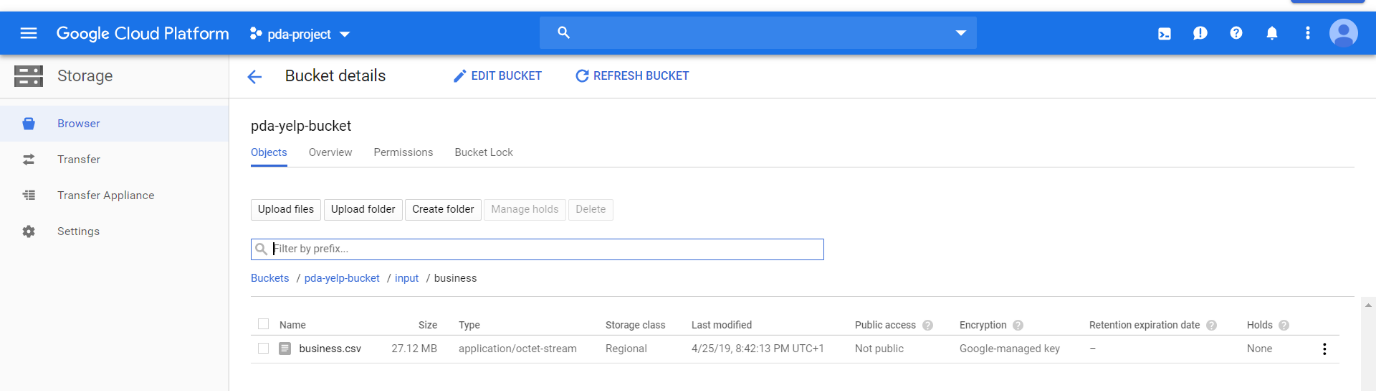
1. Run pig and execute pigCommandsTask2.pig



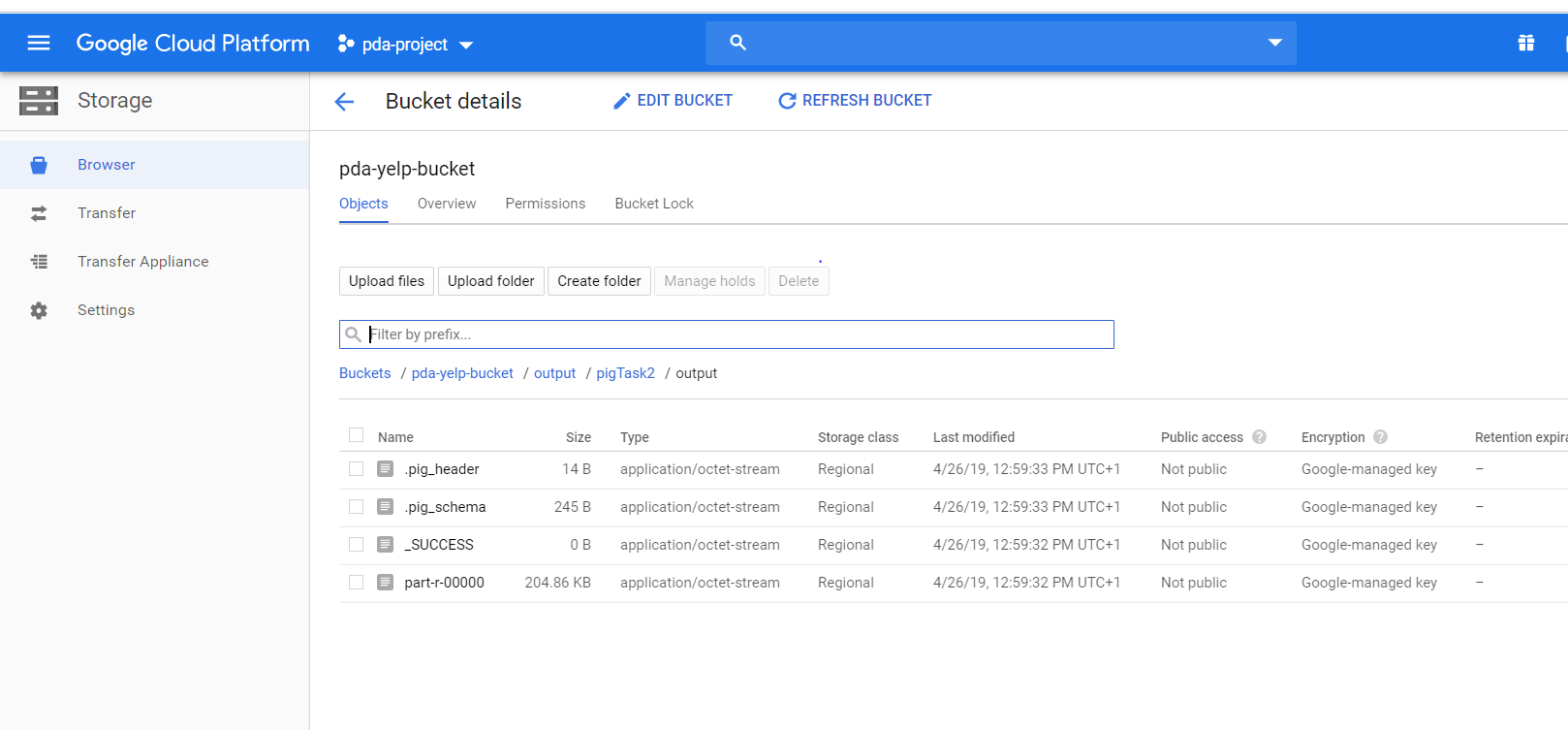
1. success



1. Input file on google file system



1. Output file on google file system



1. Final visualisation

