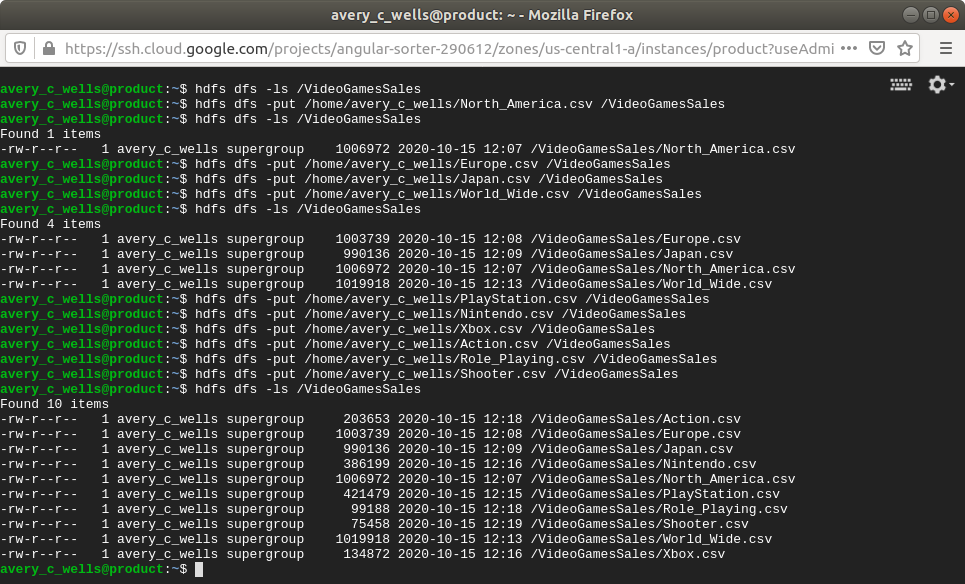
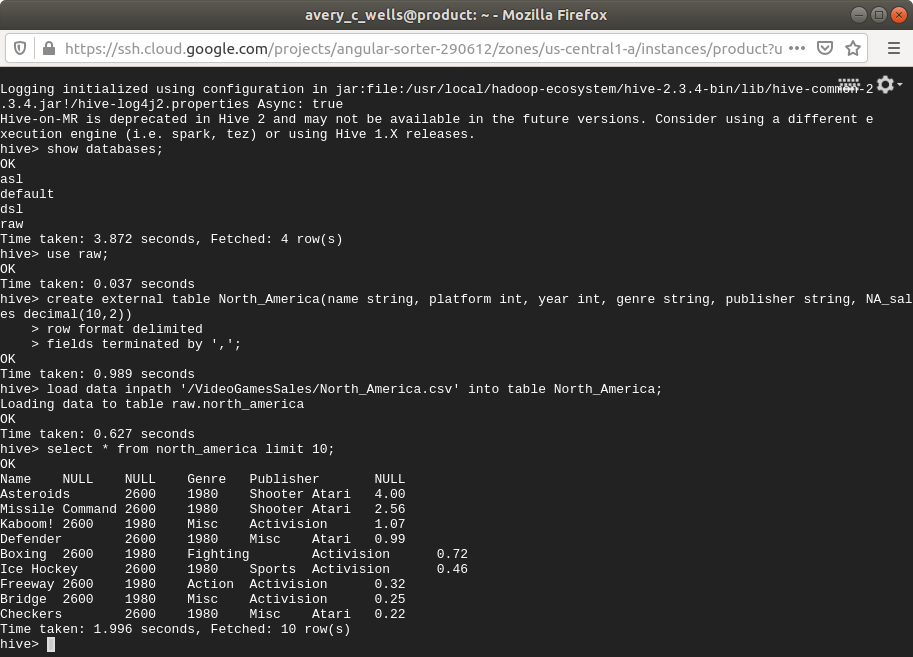
**Internal/External In Hive**

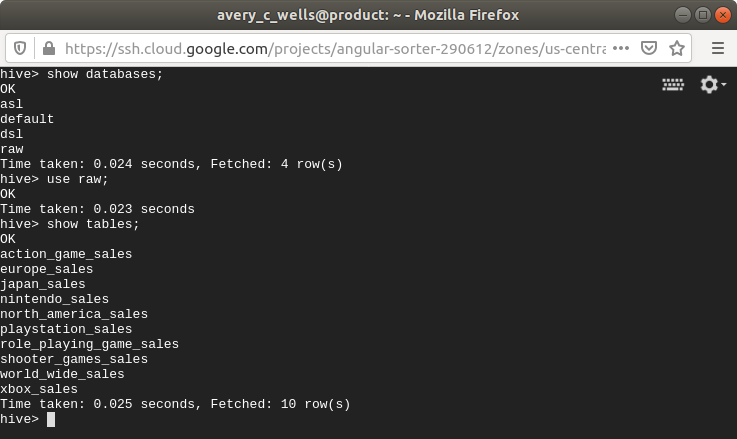
This process was done in the Google Cloud Platform using my production instances. Before getting to the internal and external tables in Hive, I had to place my csv files in HDFS. They are all in one folder “/VideoGamesSales” so it can be easier to look for them.



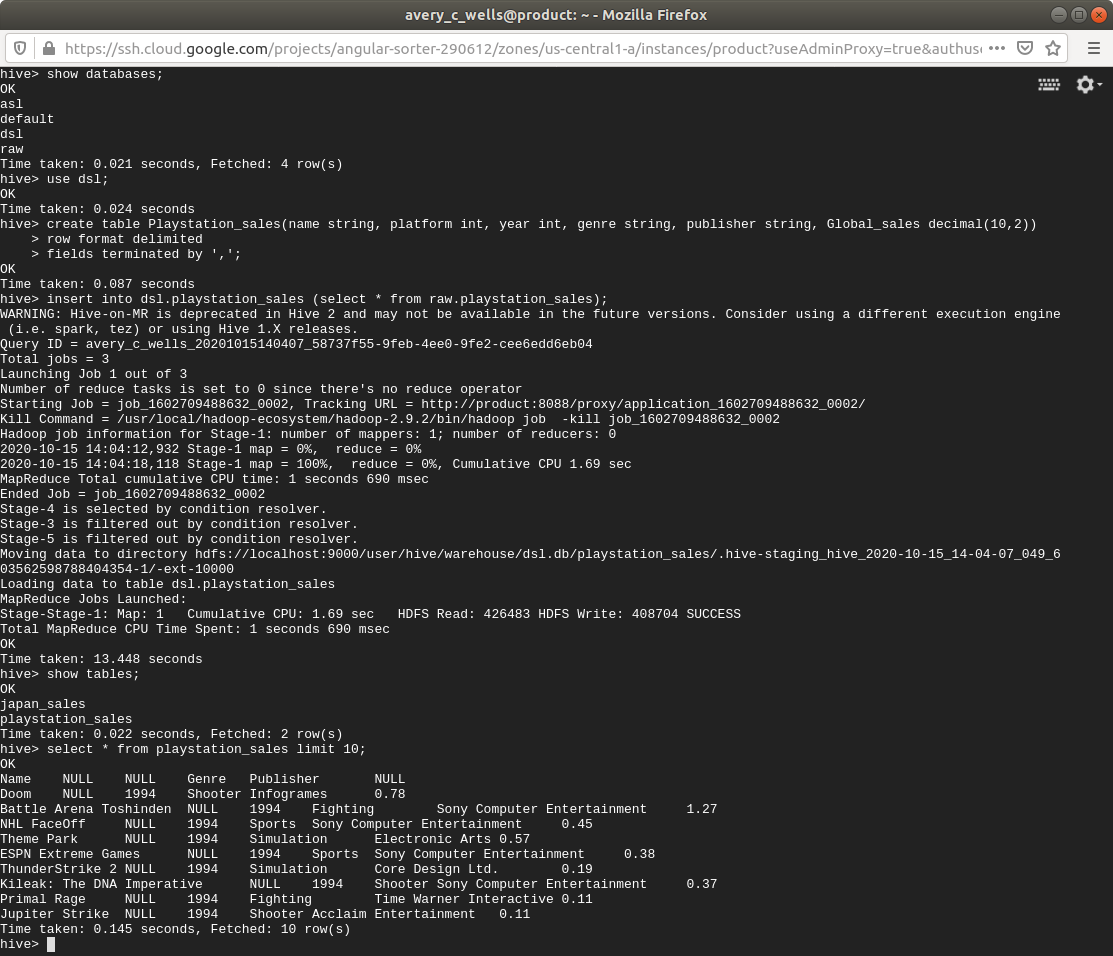
Next for me was to transfer those files/tables into Hive. I had to create an external table in the RAW database in Hive so all the tables can be in one place. Getting the right schema to match the table was important when moving tables to another database. Even though some of the tables had the same schema, some of them had them in different order so it was key to not misplace them.



I did rename all my tables so they can be described better.



Now it was time for the internal tables. Thee internal tables were to be placed in the DSL database. I only selected 5 of the tables from the RAW database. The command for creating an internal table were just the same as external except that there was no need to type “internal” in the command. Inserting the tables were also different from external. The command would be: “insert into dsl.table (select \* from raw.table);”.



Below, both the RAW and DSL databases are showing their respected tables.

