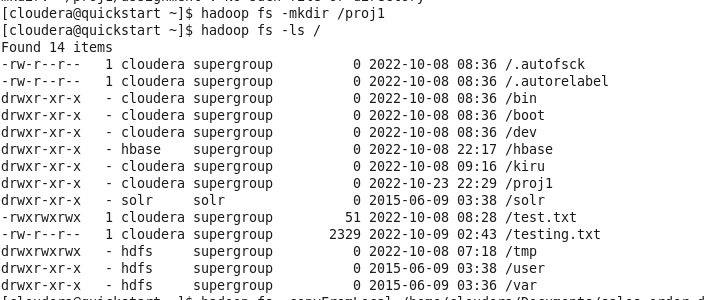
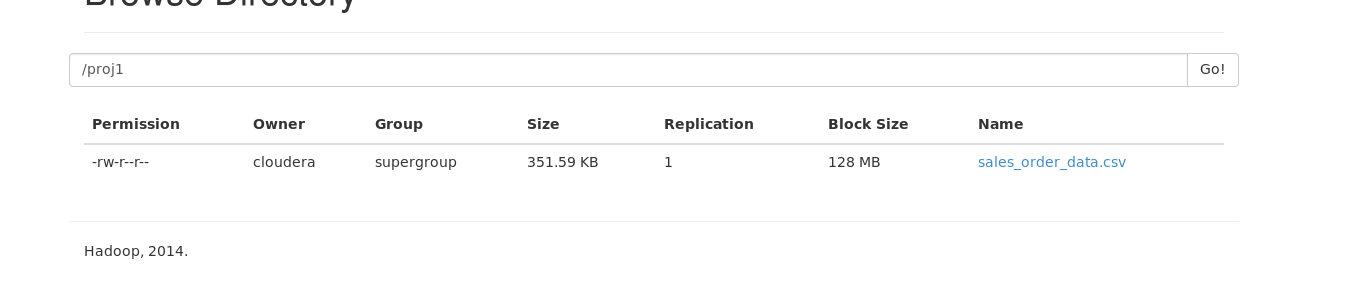
**1. Store raw data into hdfs location**

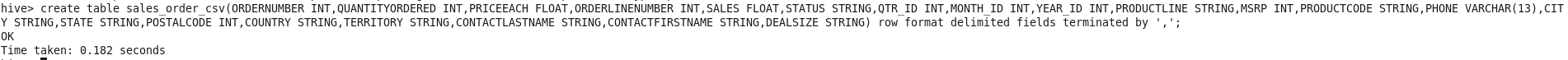


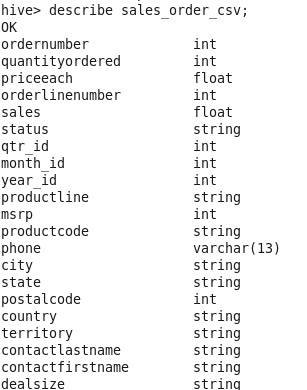




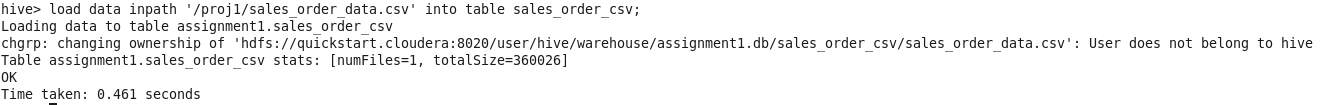


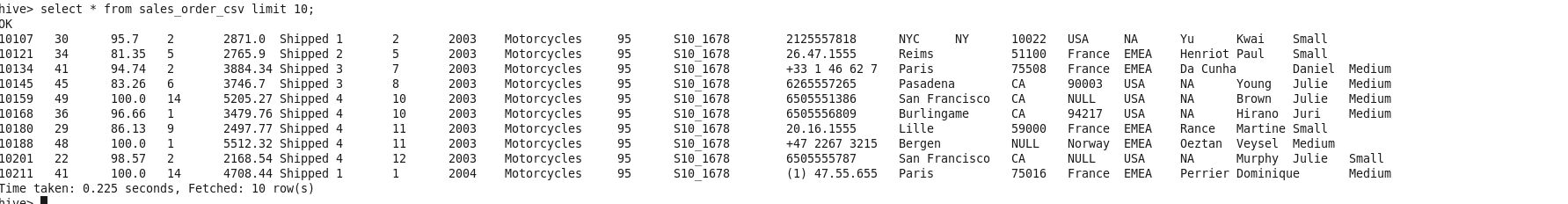
**2. Create a internal hive table "sales\_order\_csv" which will store csv data sales\_order\_csv .. make sure to skip header row while creating table**



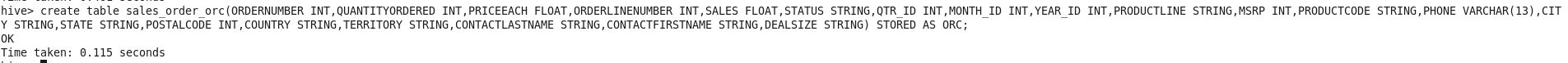


**3.Load data from hdfs path into "sales\_order\_csv"**

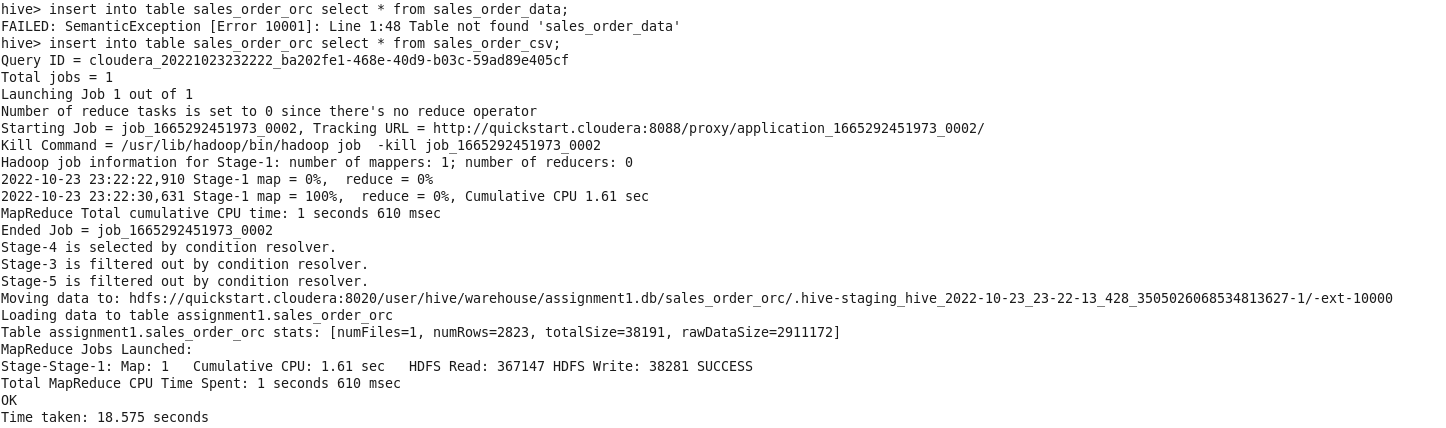


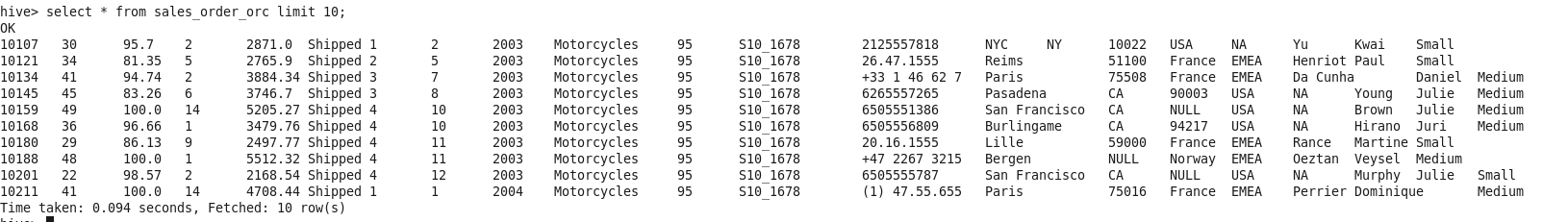


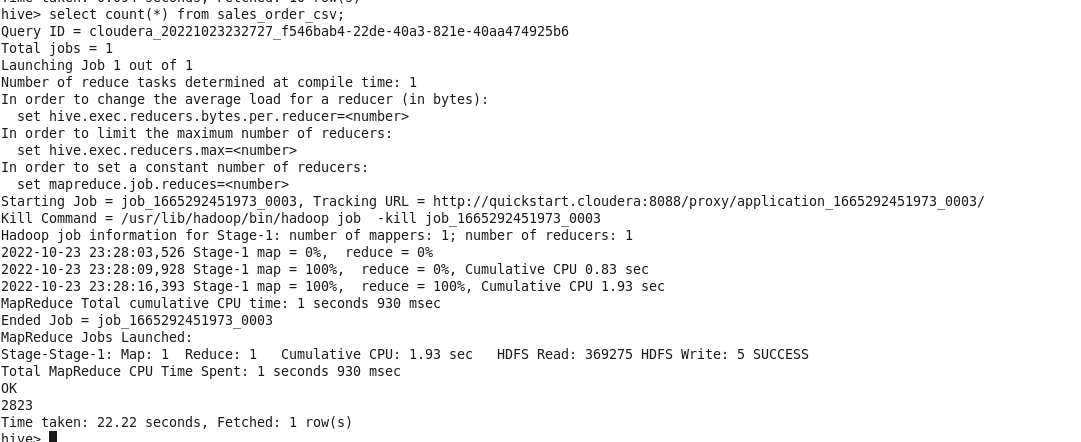
**4.Create an internal hive table which will store data in ORC format "sales\_order\_orc"**

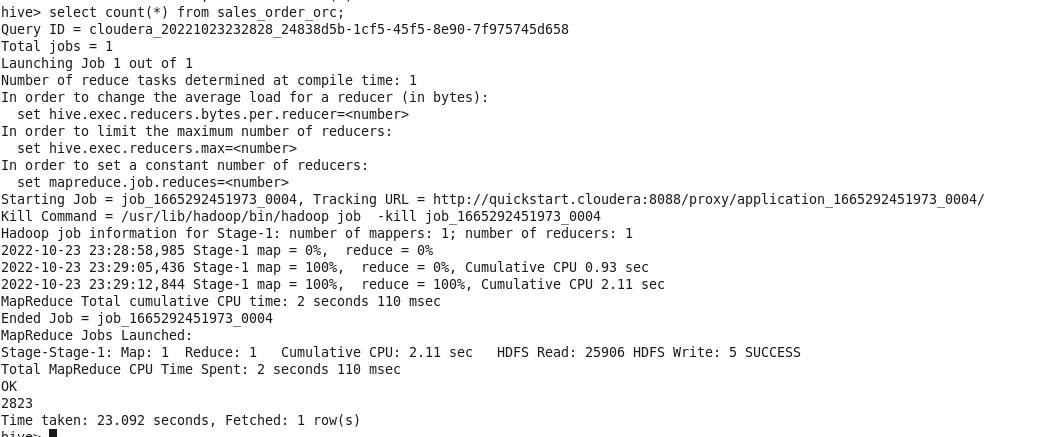


**5. Load data from "sales\_order\_csv" into "sales\_order\_orc"**

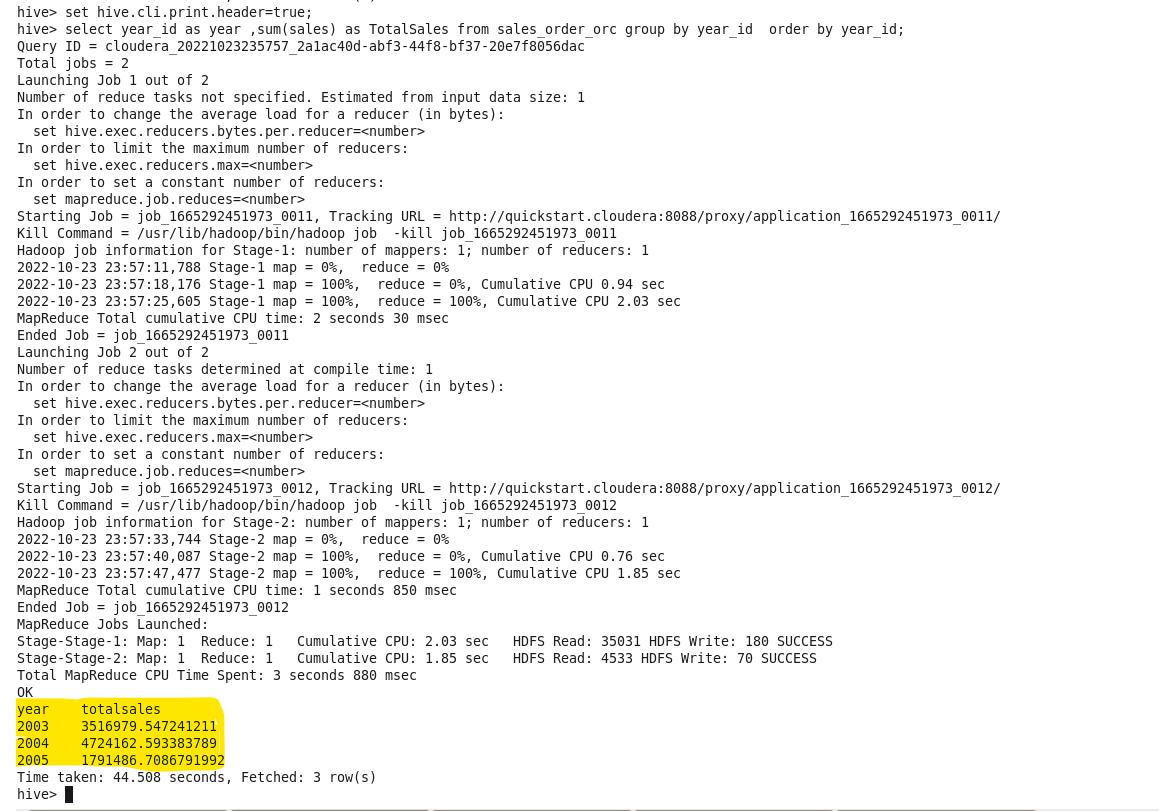




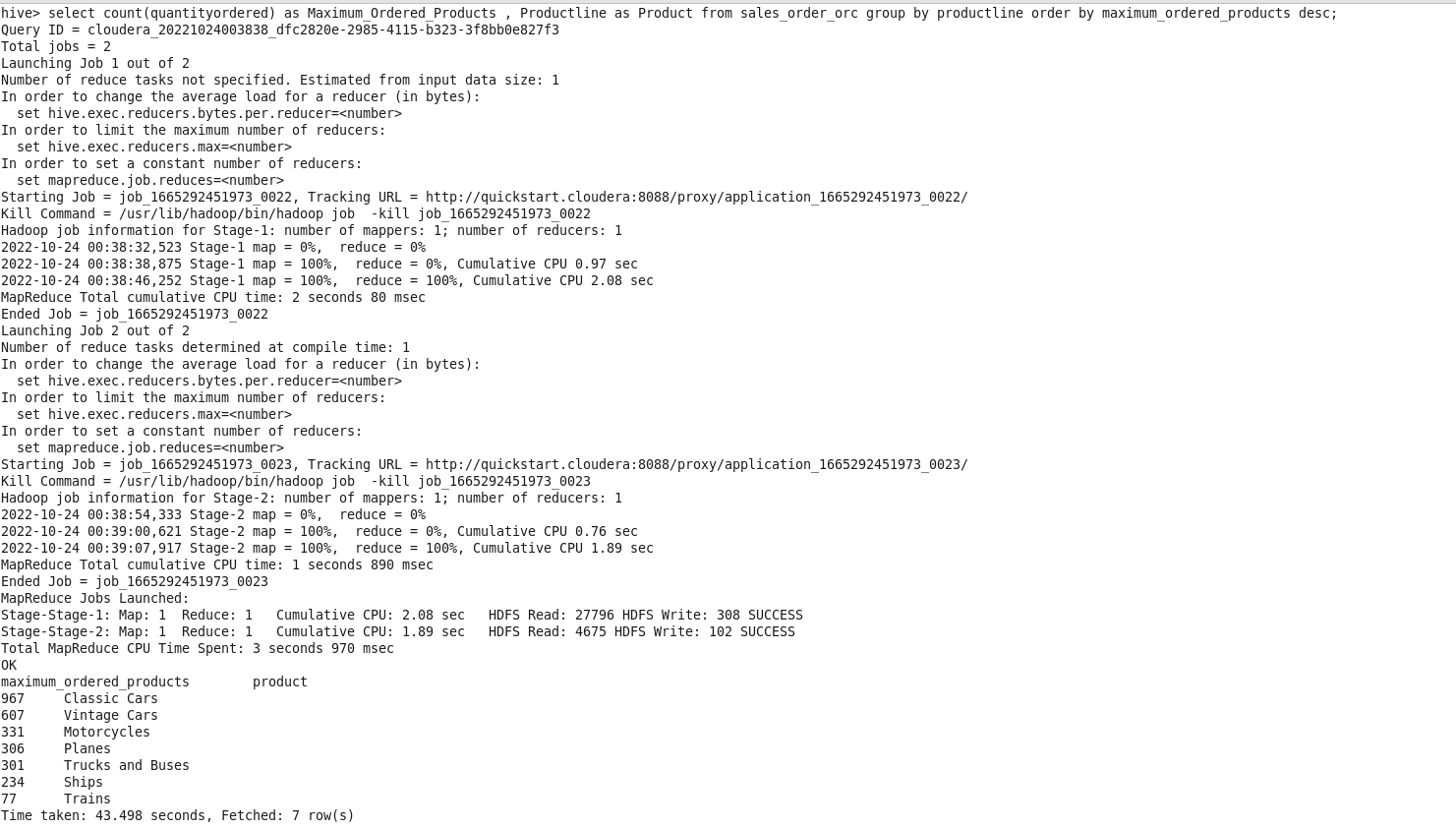




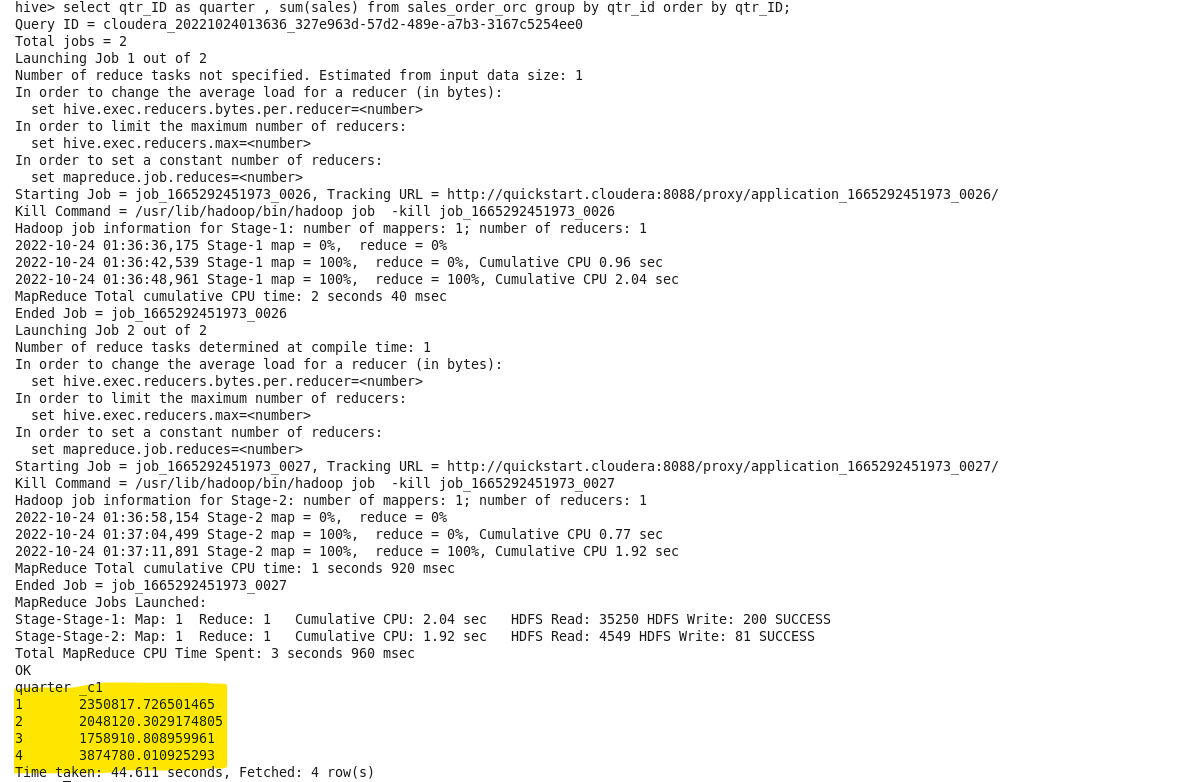
1. **Calculate total sales per year**



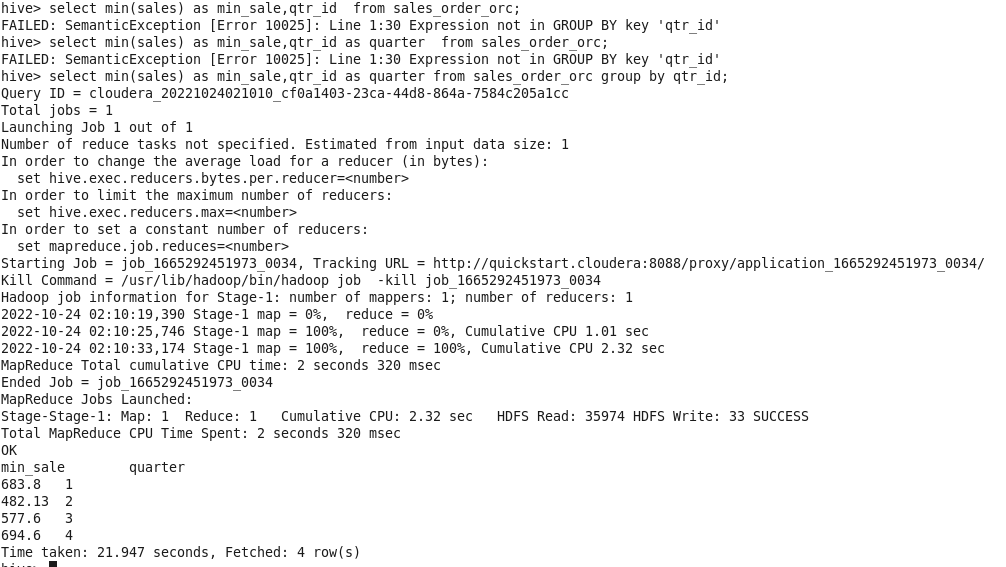
1. **Find a product for which maximum orders were placed**

****

1. **Calculate the total sales for each quarter**



1. **In which quarter sales was minimum**

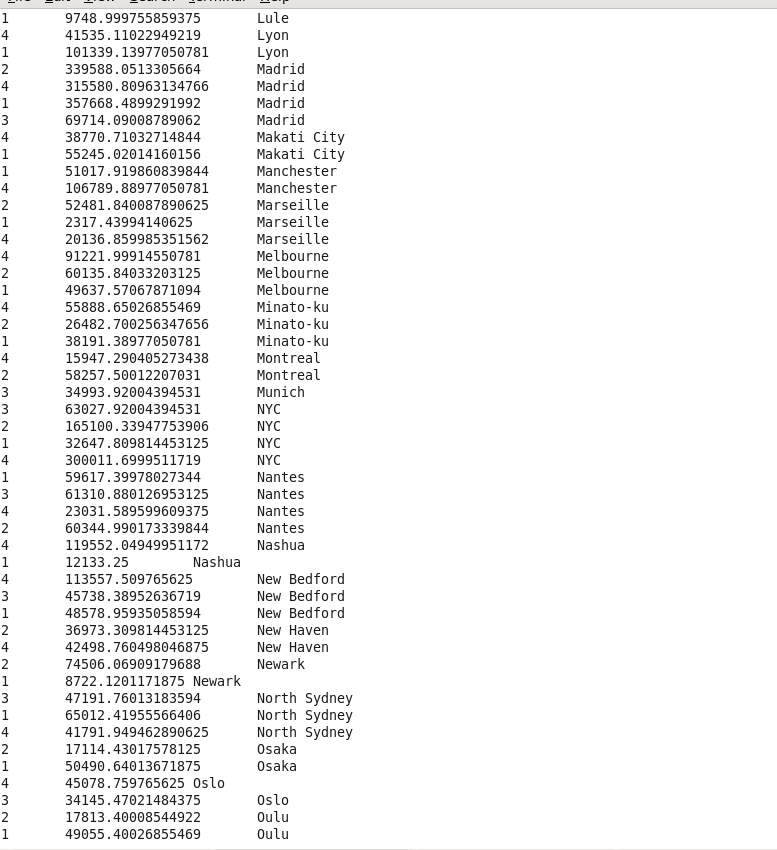


1. **In which country sales was maximum and in which country sales was minimum**



**F. Calculate quarterly sales for each city**







**g. Find a month for each year in which maximum number of quantities were sold**

