**TASK-1 Assignment :**

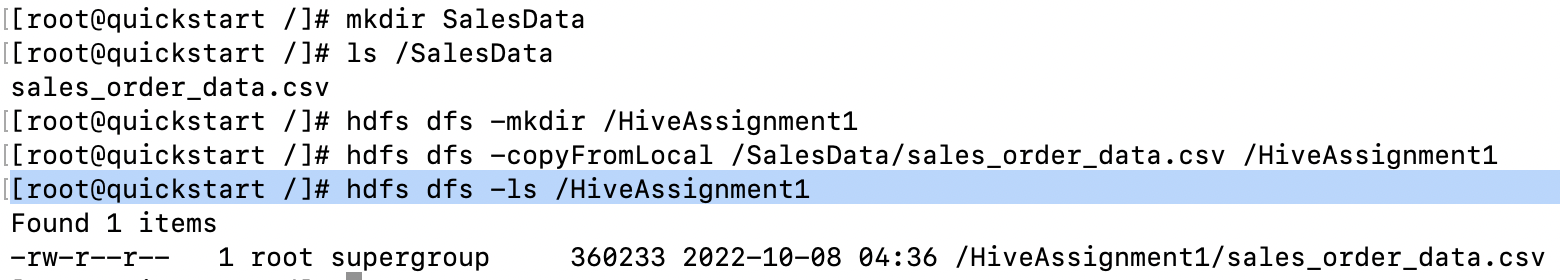
1. Download vehicle sales data 🡪 <https://github.com/shashank-mishra219/Hive-Class/blob/main/sales_order_data.csv>
2. Store raw data into HDFS location

docker cp sales\_order\_data.csv 322070b229b0:/SalesData/

hdfs dfs -mkdir /HiveAssignment1

hdfs dfs -copyFromLocal /SalesData/sales\_order\_data.csv /HiveAssignment1

hdfs dfs -ls /HiveAssignment1



1. 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.

Text

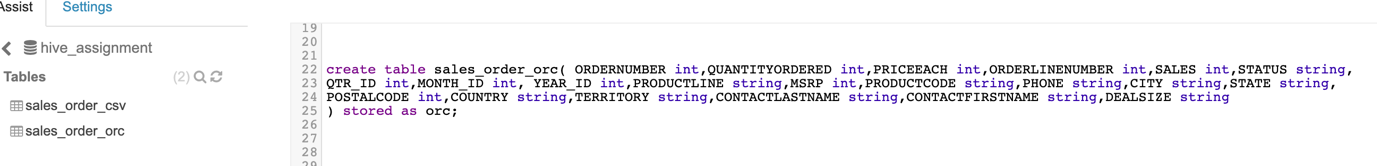
Description automatically generated with low confidence

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

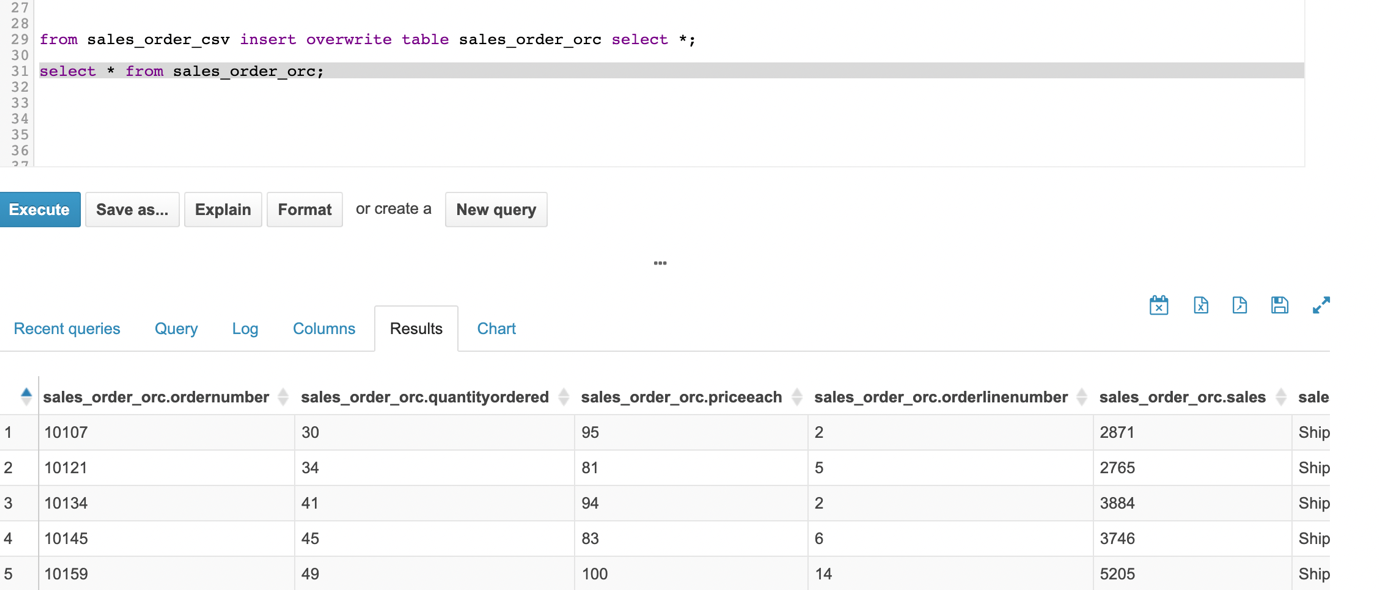
Graphical user interface, text, application, Teams

Description automatically generated

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



1. Load data from "sales\_order\_csv" into "sales\_order\_orc".



Perform below mentioned queries on "sales\_order\_orc" table :

1. Calculate total sales per year.

Graphical user interface, text, application

Description automatically generated

1. Find a product for which maximum orders were placed

Graphical user interface, text, application, email

Description automatically generated

1. Calculate the total sales for each quarter

Graphical user interface, table

Description automatically generated

1. In which quarter sales was minimum

Graphical user interface, text, application

Description automatically generated

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

Graphical user interface, text, application

Description automatically generated

1. Calculate quarterly sales for each city

Graphical user interface, table

Description automatically generated

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

Table

Description automatically generated