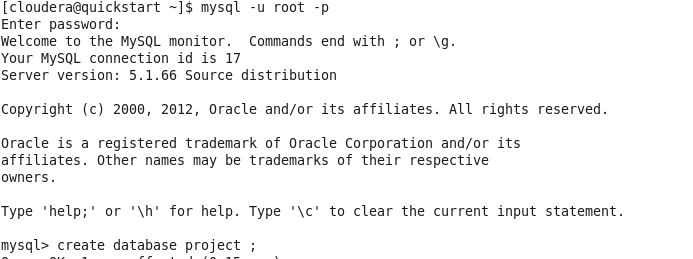
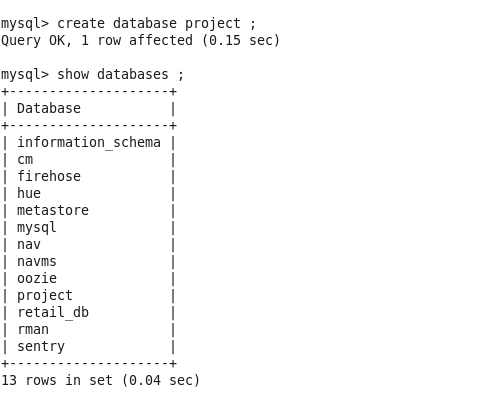
**Project Outcomes:**

**1.TaskOutcome:**

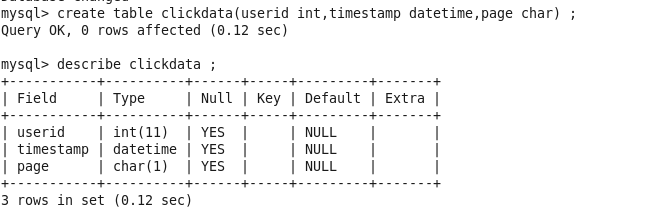
**🡺 start mysql in cloudera**



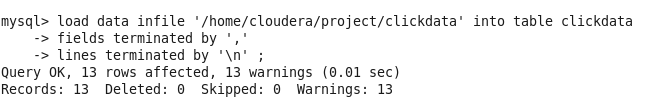
🡺**create a database** (database name is project)



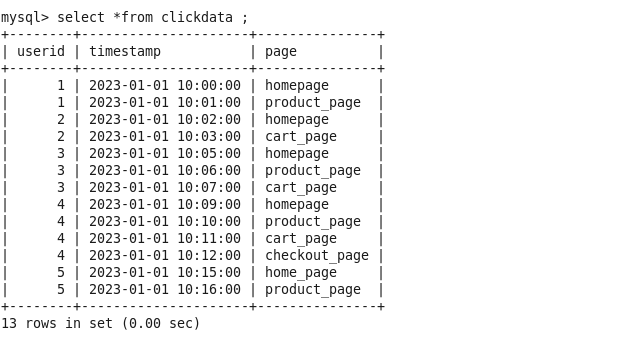
🡺**create a first table** (table name is click\_data)



🡺**put data in first table using load command**(click\_data table)

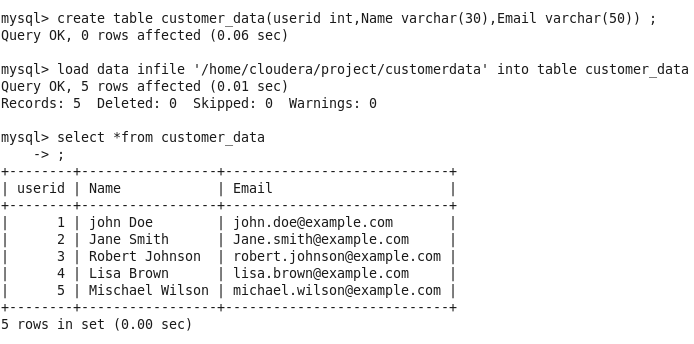


🡺**show the data first table using select statement**

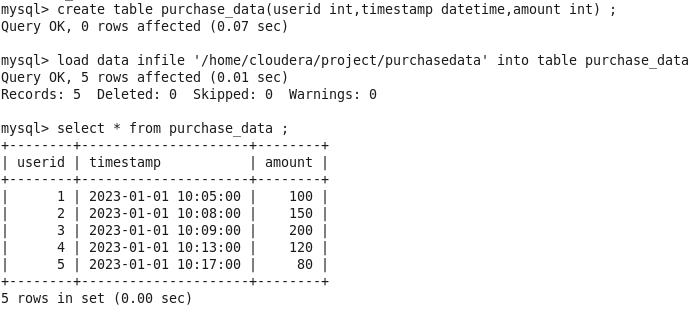


🡺**again doing same step in next table for customer\_data and purchase\_data:**

**For customer\_data table:**



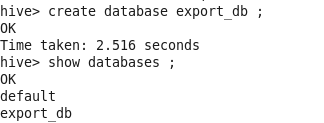
**For purchase\_data table:**

****

**2.TaskOutcomes:**

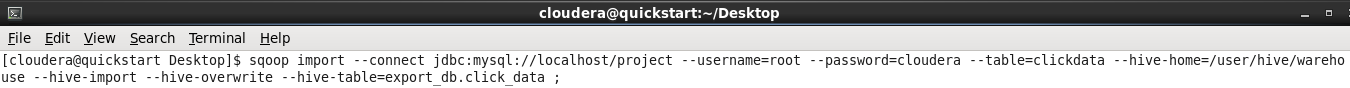
🡺**import data from mysql to hive using sqoop command:**

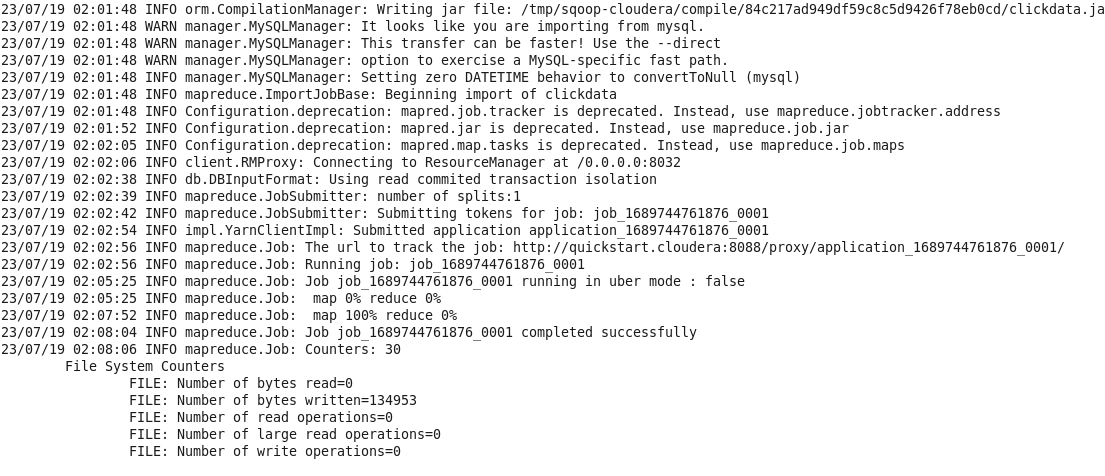
**🡺first create a database in hive and use this database and create a table:**

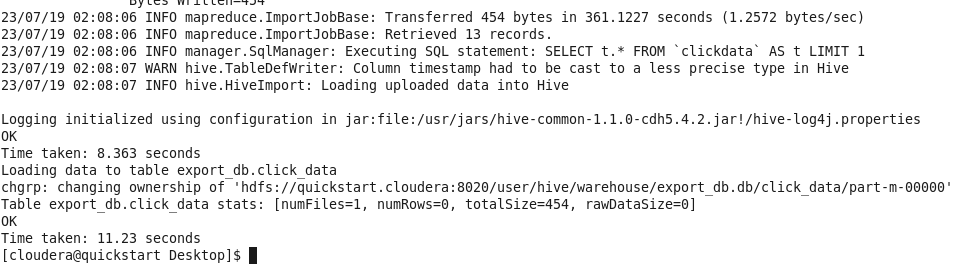
****

****

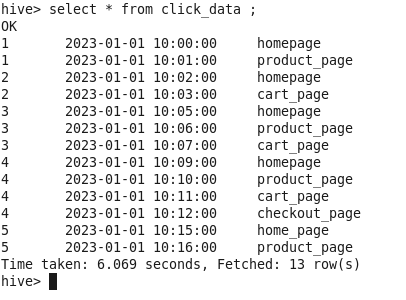
**🡺import mysql-table to hive using sqoop commands:**





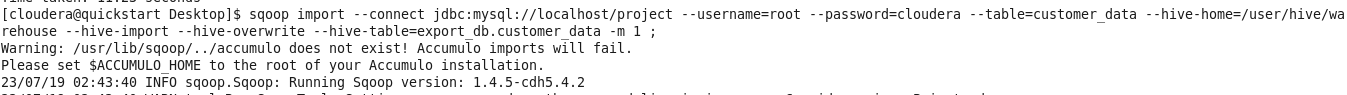


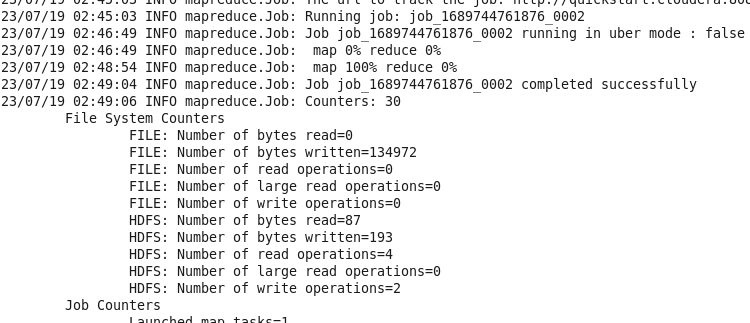
🡺**Go to hive check import status:**

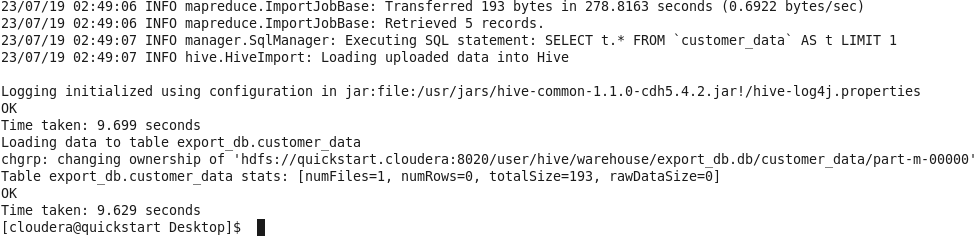


Success full import all rows are come in hive table this is first table import (clickdata) more two table are import mysql to hive

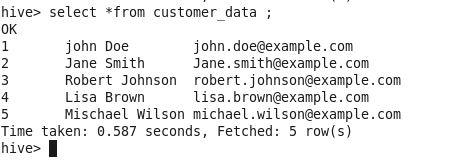
🡺**import customer\_data table into hive:**



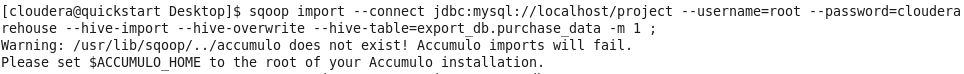


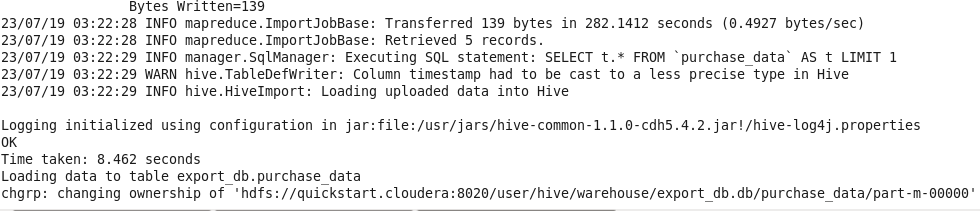


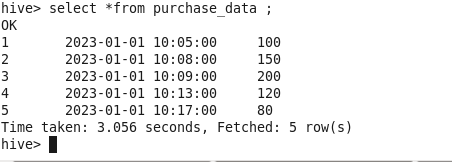
**See the results:**

****

**🡺import purchase\_data table into hive:**

****

****

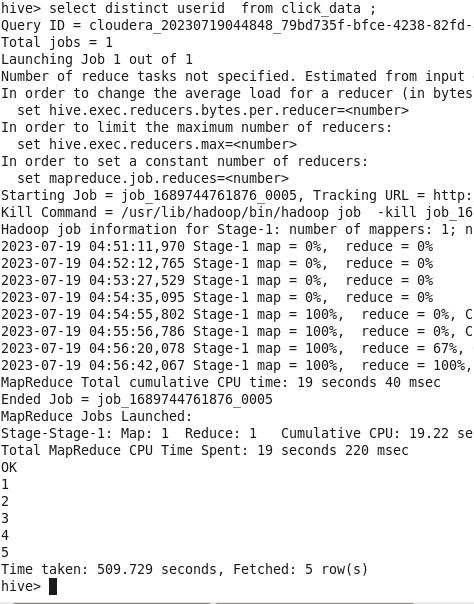
****

**Last table come in hive (purchase\_data table).**

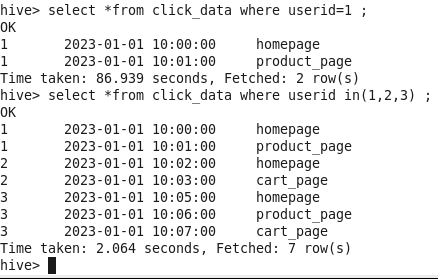
**3.TaskOutcomes:**

**1.Data Cleaning:**

🡺**Distinct element:**

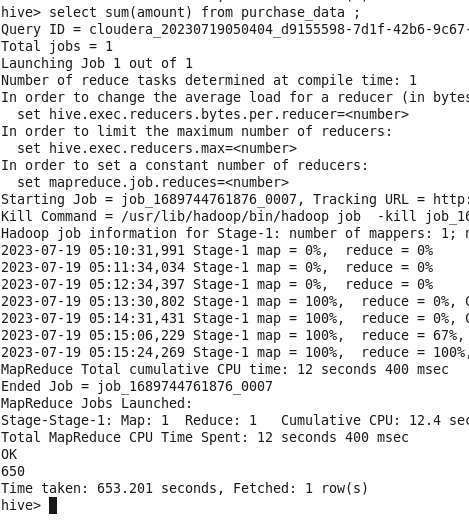


🡺**Data filter: use the where clause.**

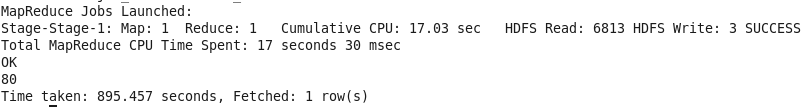
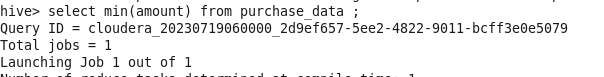


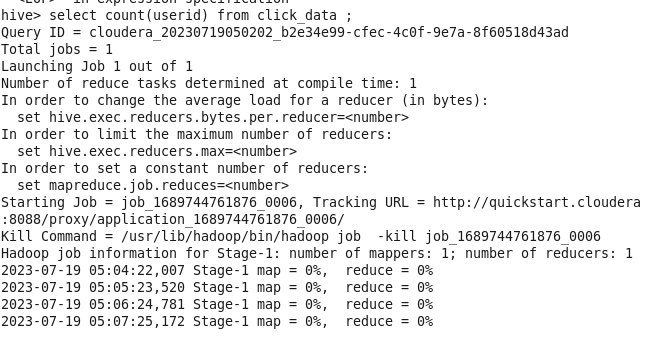
🡺**Data Aggregation :**

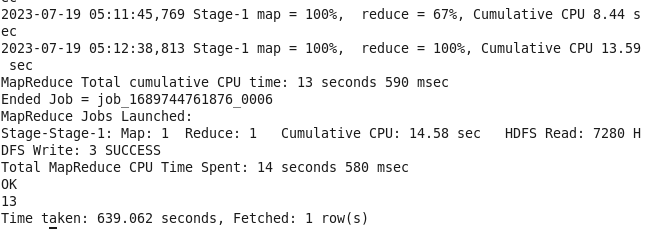
**Sum Function:**



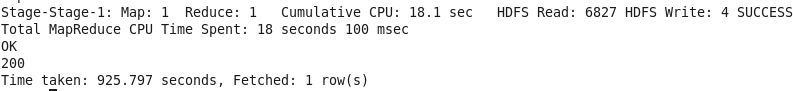
**Min Function:**

****

**Count:** 

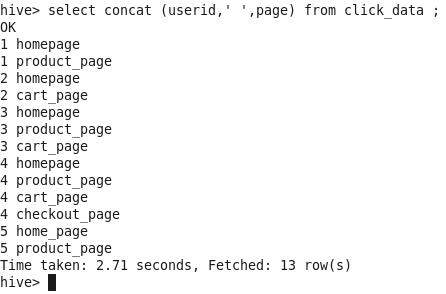


**Max Function:**

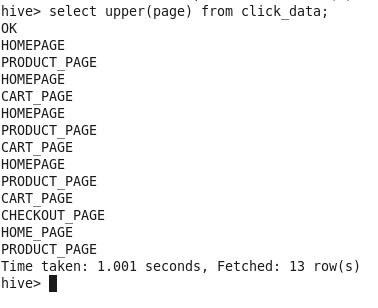
**C:\Users\HP\OneDrive\Desktop\boardinfinty-project\hadoop poto\30.PNG**

**3.Data Transformation:**

**Concat:**

****

**String Manipulation:**

****