|  |  |
| --- | --- |
| **Ex no : 3** | **ETL Process for a selected Data Source** |
| **Date :** |

**AIM**

To perform Extract Transform and Load (ETL) operations for a selected data source.

**PROCEDURE**

Step 1: Download and install the tableau-desktop 2023.2.1.

Step 2: Download and Install MySQL Server (Windows (x86, 64-bit), MSI Installer) from <https://dev.mysql.com/downloads/mysql/>

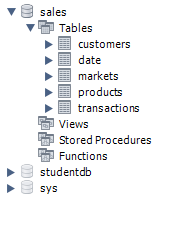
Step 3: Download and Install MySQL Workbench (Windows (x86, 64-bit), MSI Installer) from <https://dev.mysql.com/downloads/workbench/>

Step 4: Download and Install MySQL connecter (Windows (x86, 64-bit), MSI Installer) from <https://dev.mysql.com/downloads/connector/odbc/>

Step 5: Download the SQL DB which consists of customers, date, markets, products and transactions tables of sales DB from <https://github.com/codebasics/DataAnalysisProjects/blob/master/2_SalesInsightsTableau/db_dump.sql>

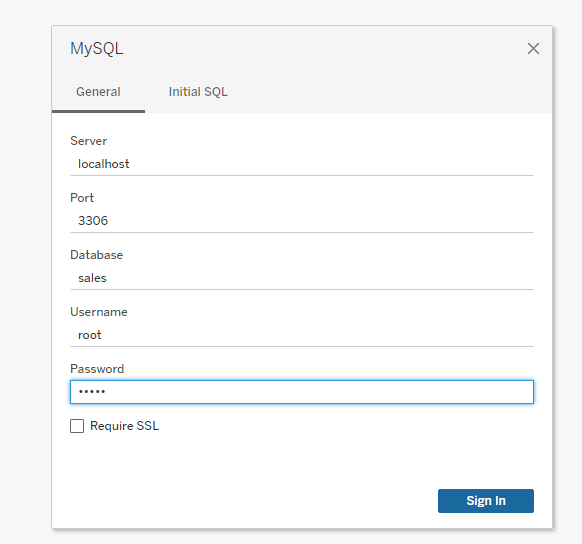
Step 6: Create mysql sales schema and import the downloaded db\_dump.sql to sales DB.

Step 7: You can find all the sales tables will be imported in the sales DB.



Step 8: Keep the Server on in MySQL and open the tableau desktop.

Step 9: After opening tableau desktop, connect to MySQL Server with sales DB by specifying Username and password as shown below.



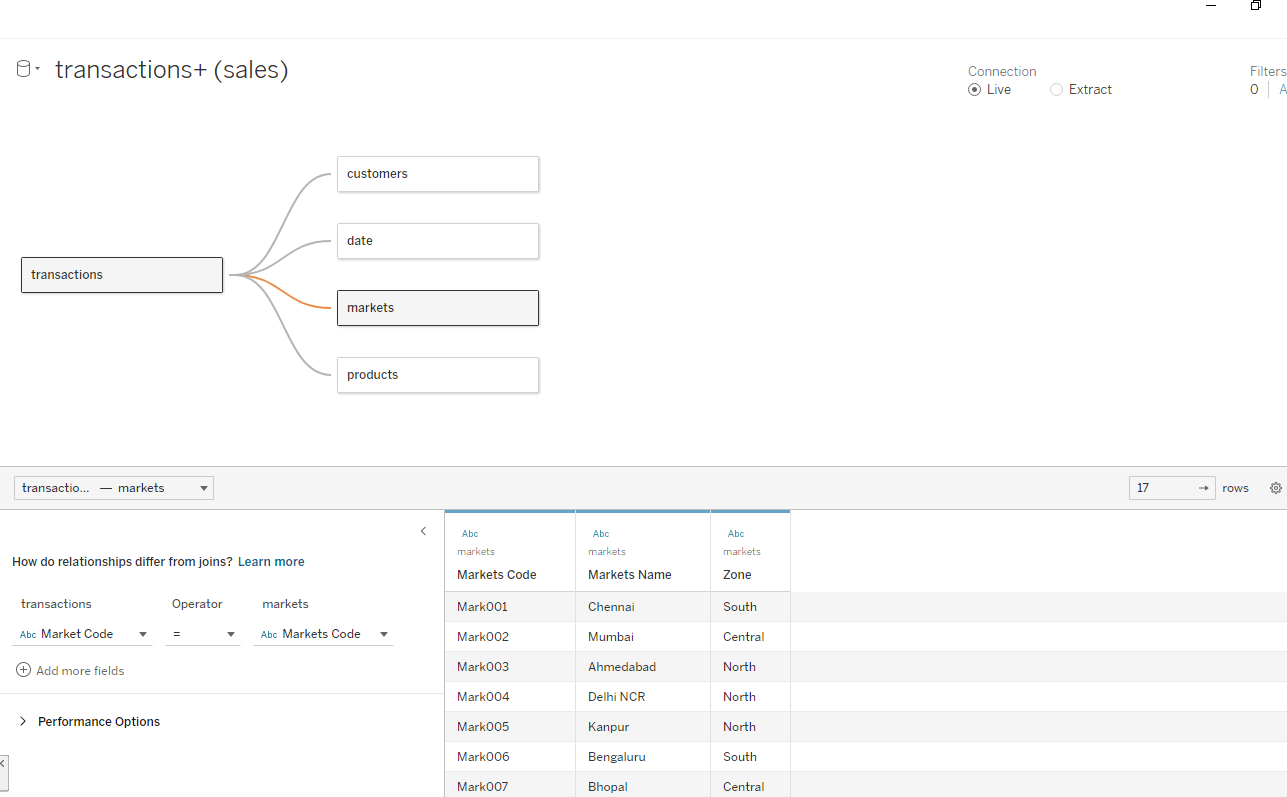
Step 10: After connecting, you can visualise all tables in the tableau, then drag the tables to form a star schema by specifying its relationships(keys), with fact and dimension table scenario as shown below in output.

Step 11: After extracting and connecting all the table data, perform transforming and loading process.

Step 12: Click on transaction table and inspect sales amount, we could able to find the sales amount is ‘-1’ for a data, so remove it by applying the filter, data->edit data source filters. Once filter is applied we can observe that ‘-1’ sales amount row will be deleted (As the sales amount would never be -1).

Step 13: Similarly, Inspect all the columns and transform the parameters like currency and sales amount and make necessary changes.

**OUTPUT**



**RESULT**

Thus, Extract Transform and Load (ETL) operations for a selected data source has been completed successfully.