**Project Proposal**

**Supplier and Market Place Data Consolidation (ETL Project)**

**Names of Group Members**

Samira Hajimohammadzadeh

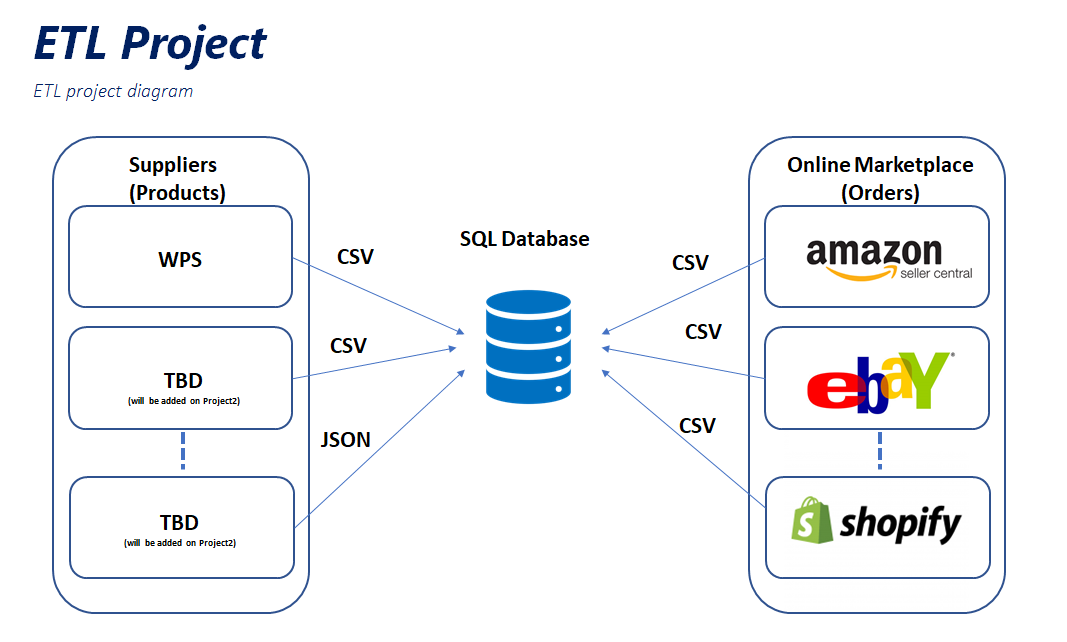
Shadi Salama

Warren Vince Sy

Shadi Askarzadeh

**Synopsis of Project:**

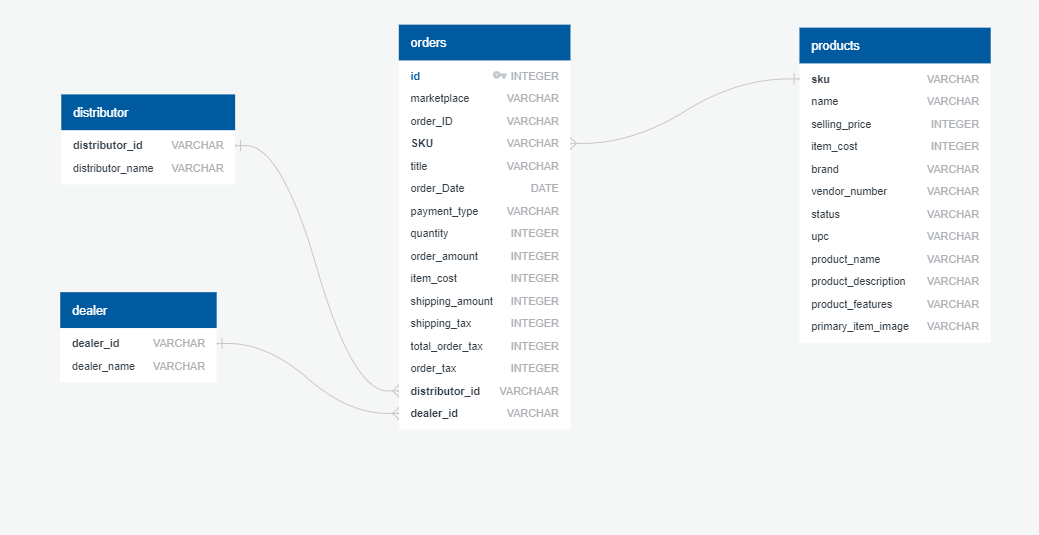
Online Sellers have no single view of their Data. All the sales data are in silo for every market place they sell. On top of that supplier data is on its own Silo as well. Goal for this project is to consolidate data from different market places and suppliers so that the seller/customer can have a single view of their data



**Datasets Used:**

For this ETL project, we will be extracting data from different market places (ebay, amazon, shopify) and suppliers (WPS) and saving it on a single database using relational tables.

**Database Architecture: Relational Database**



We created four tables in the database:

Orders Table – It will store all the order data from the different market places. We also added two additional columns (distributor\_id and dealer\_id) which are foreign keys to two masters tables. We added these tables as placeholders to be able to accommodate new dealers and distributors.

Products Table – It will store all the products the sellers are selling on the market place.

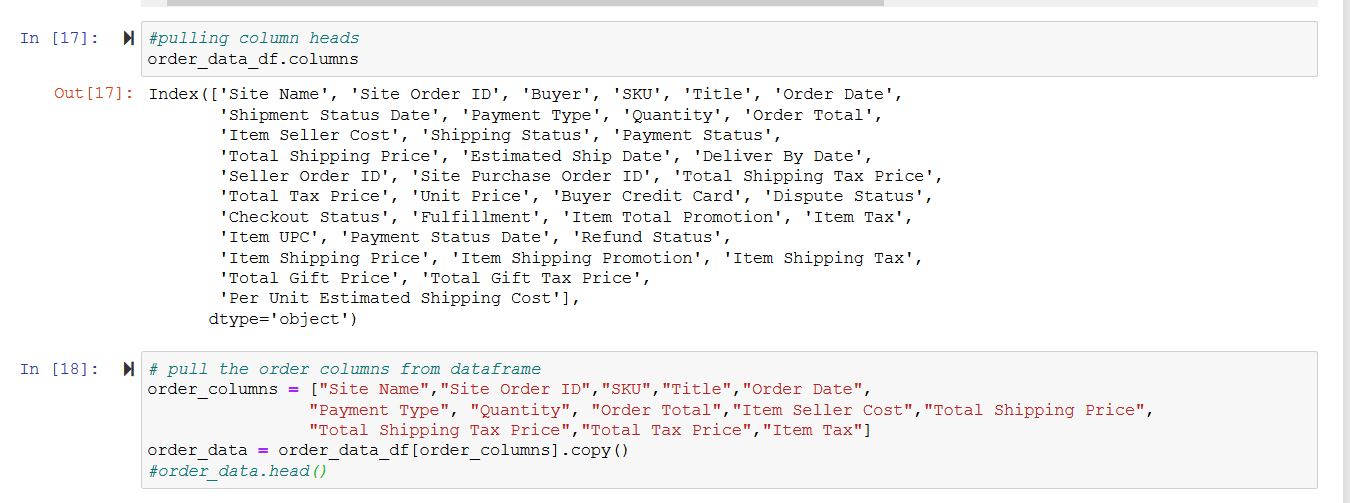
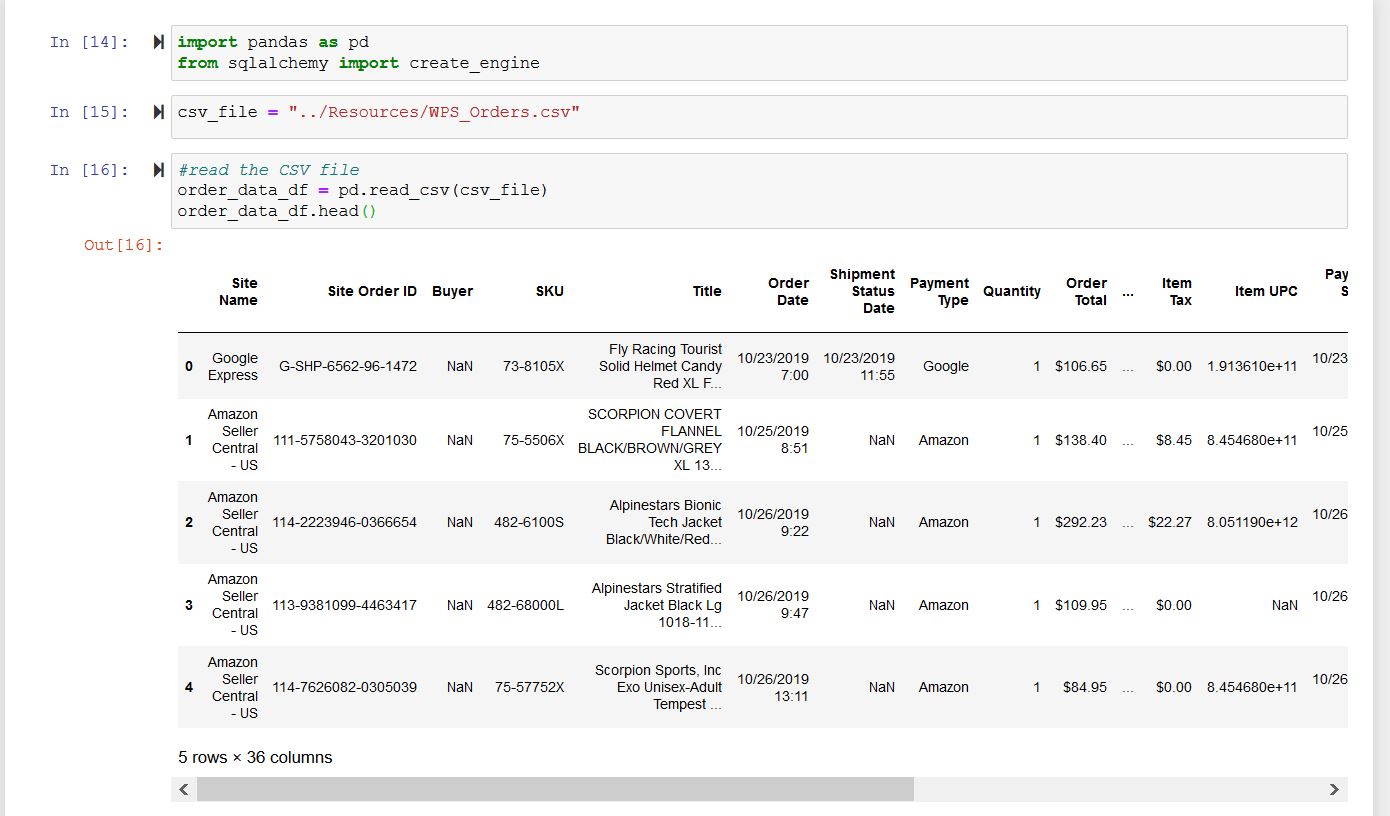
Dealer Table – It will store all Dealer Data. We will add more dealer related information on project 2.

Distributor Table - It will store all Distributor Data. We will add more distributor related information on project 2.

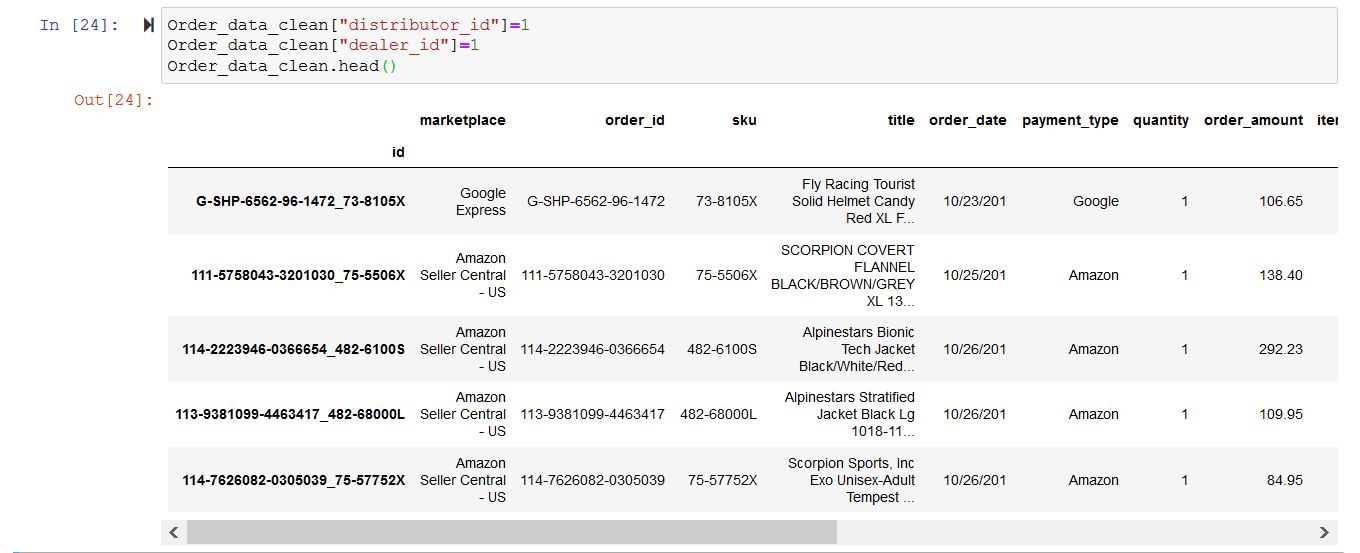
**Data Cleansing**

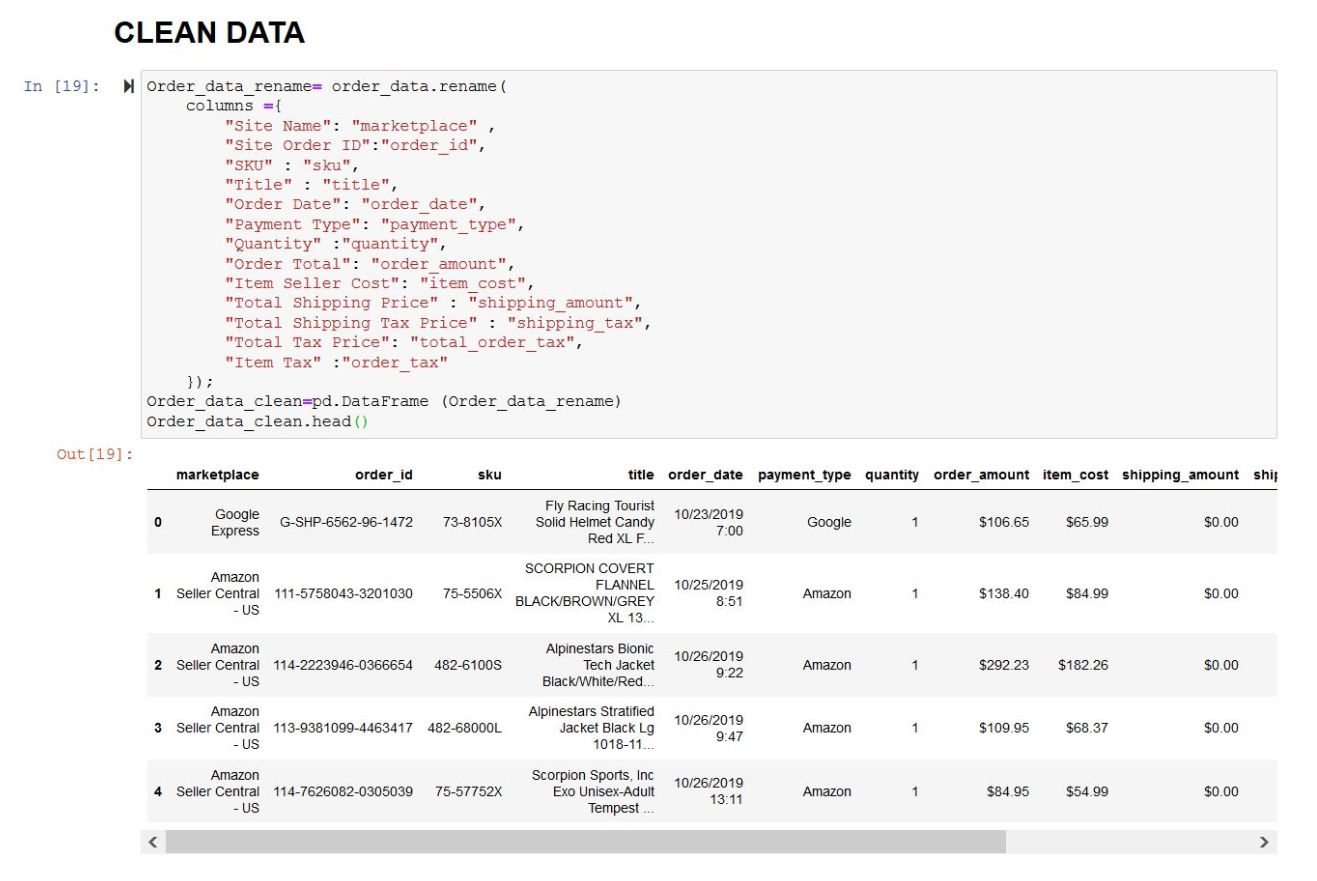
Order (Market Place File)

Extract Data from CSV file and convert it to Data Frame



Cleanse data before loading to database





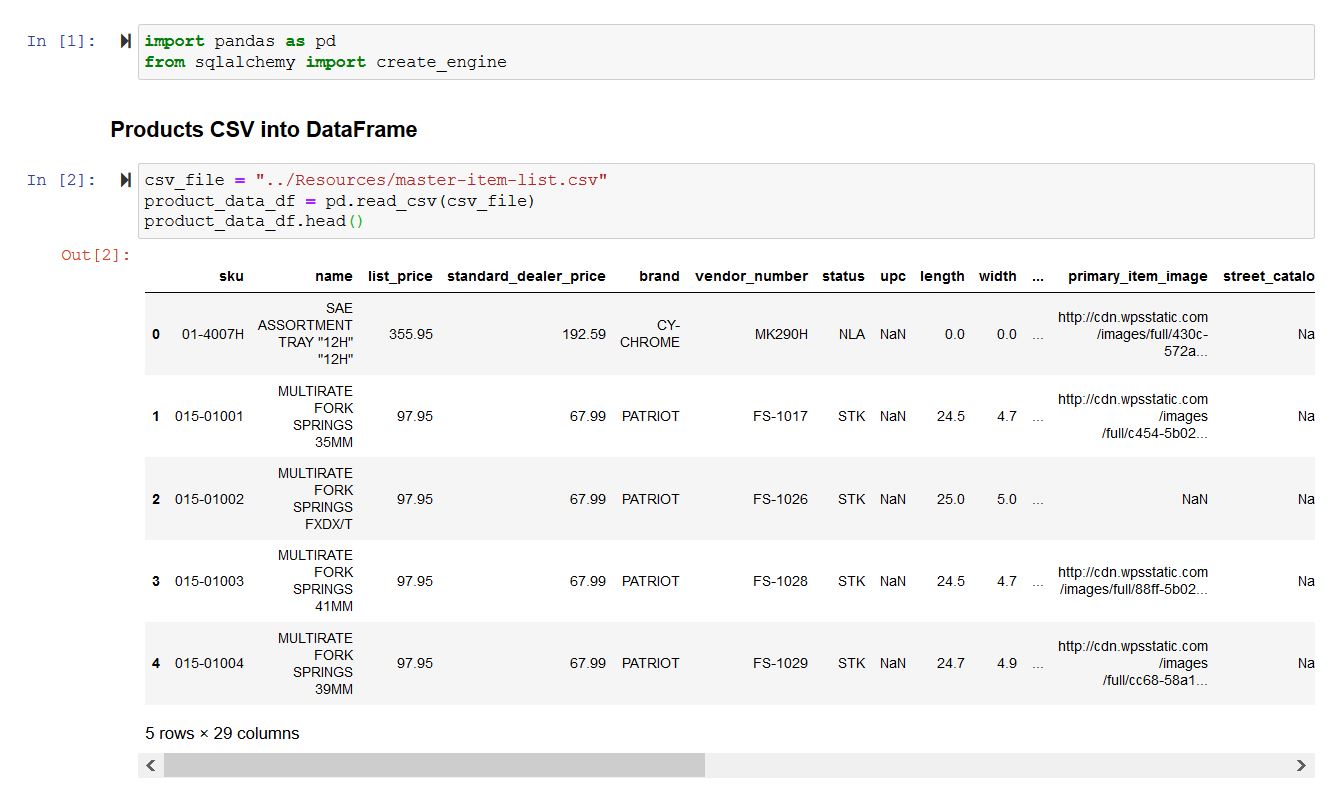


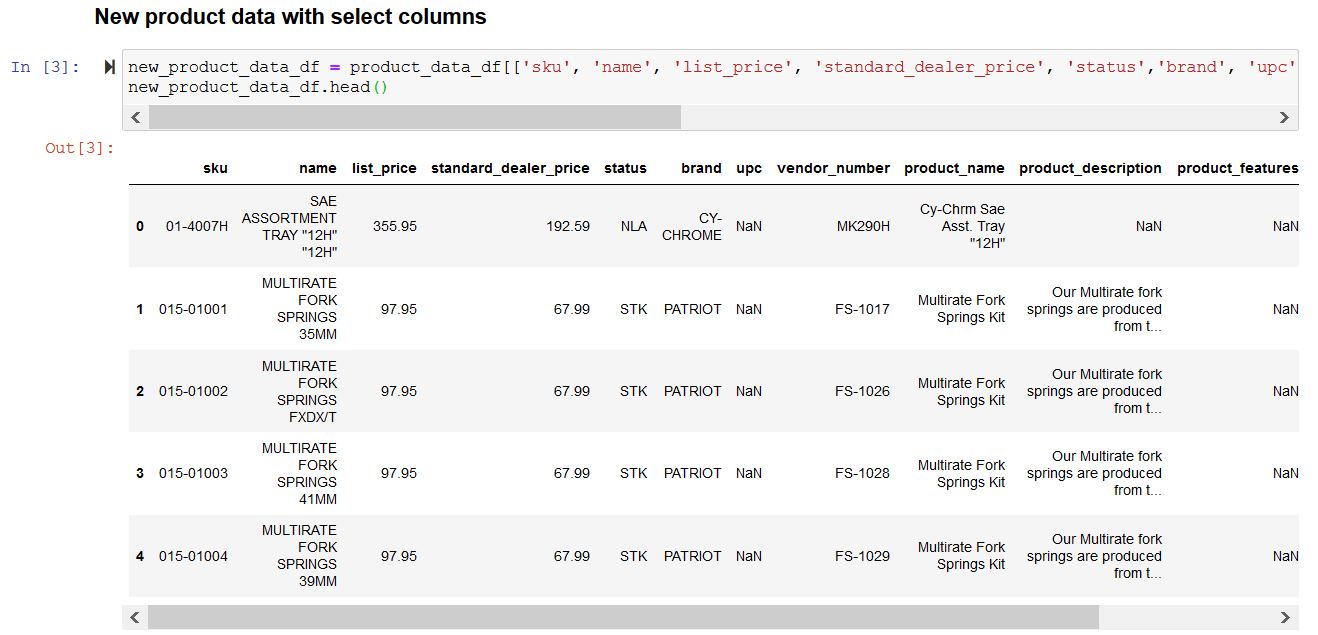
Load data to database



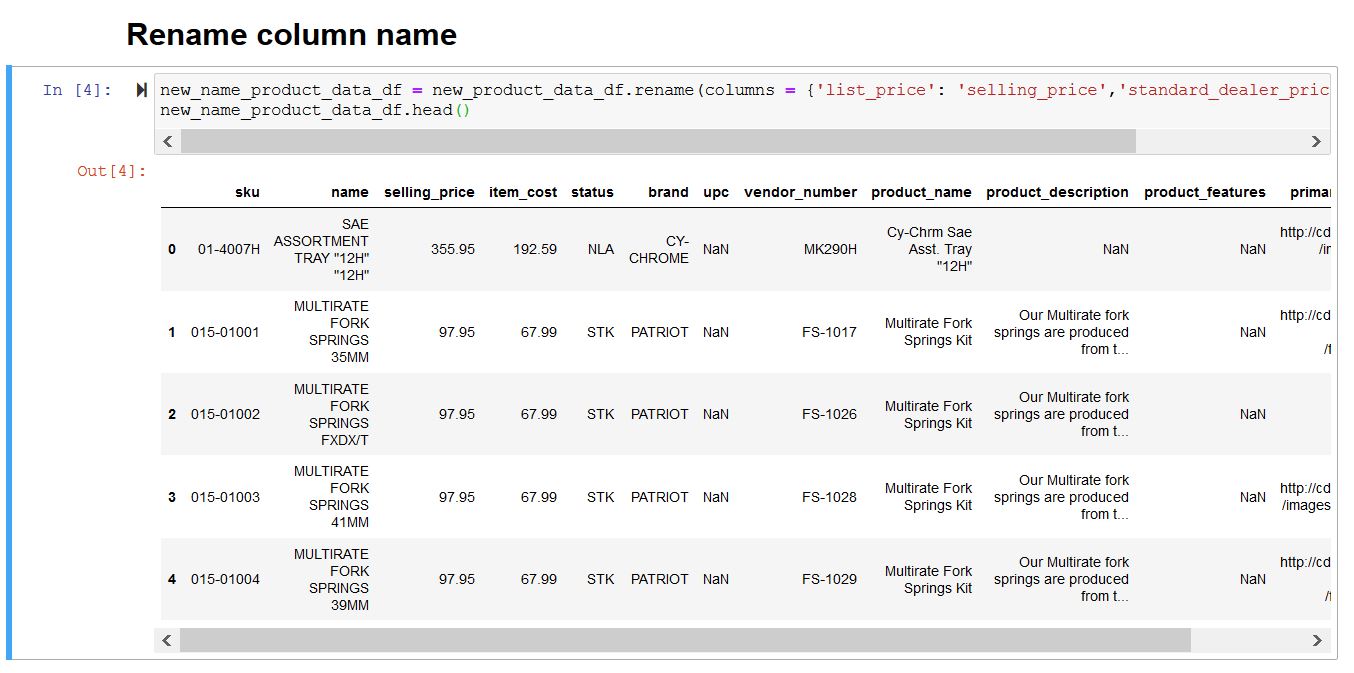
Products(Supplier File)

Extract Data from CSV file and convert it to Data Frame





Cleanse data before loading to database



Load data to database

