**Project Proposal**

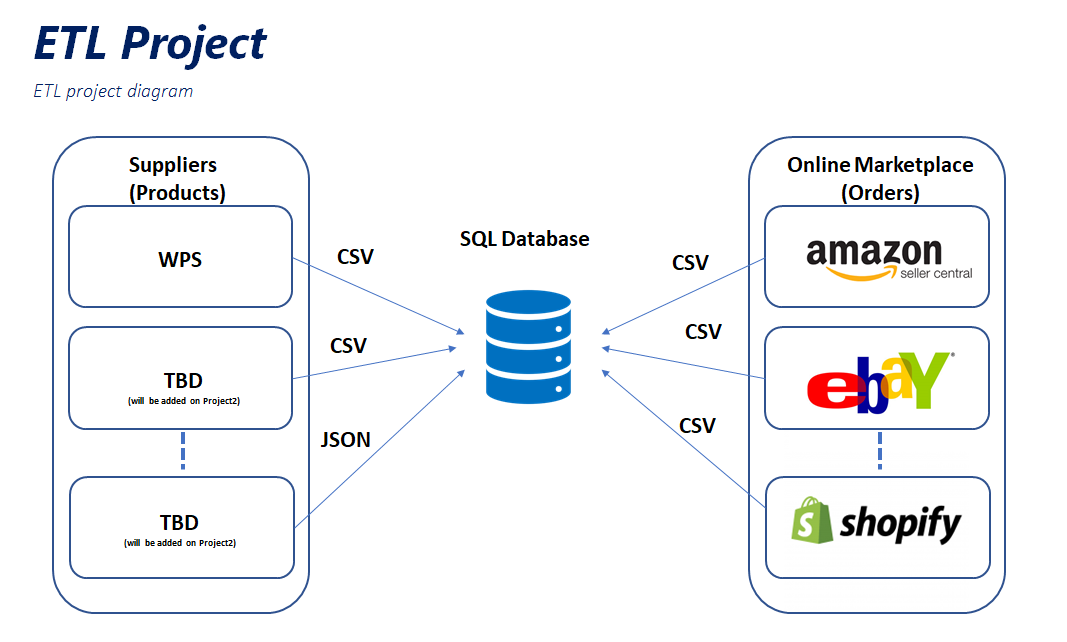
**Supplier and Market Place Data Consolidation (ETL Project)**

**Names of Group Members**

* Samira Hajimohammadzadeh
* Shadi Askarzadeh
* Shadi Salama
* Warren Vince Sy

**Synopsis of Project:**

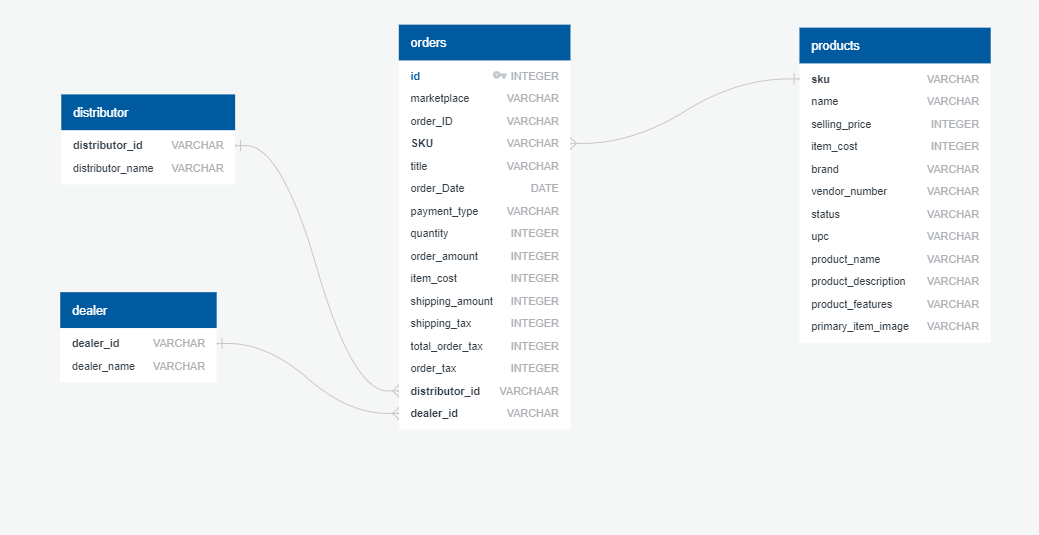
Online Sellers have no single view of their Data. All the sales data are in silo for every marketplace they sell. On top of that supplier data is on its own Silo as well. Goal for this project is to consolidate data from different marketplaces and suppliers so that the seller/customer can have a single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

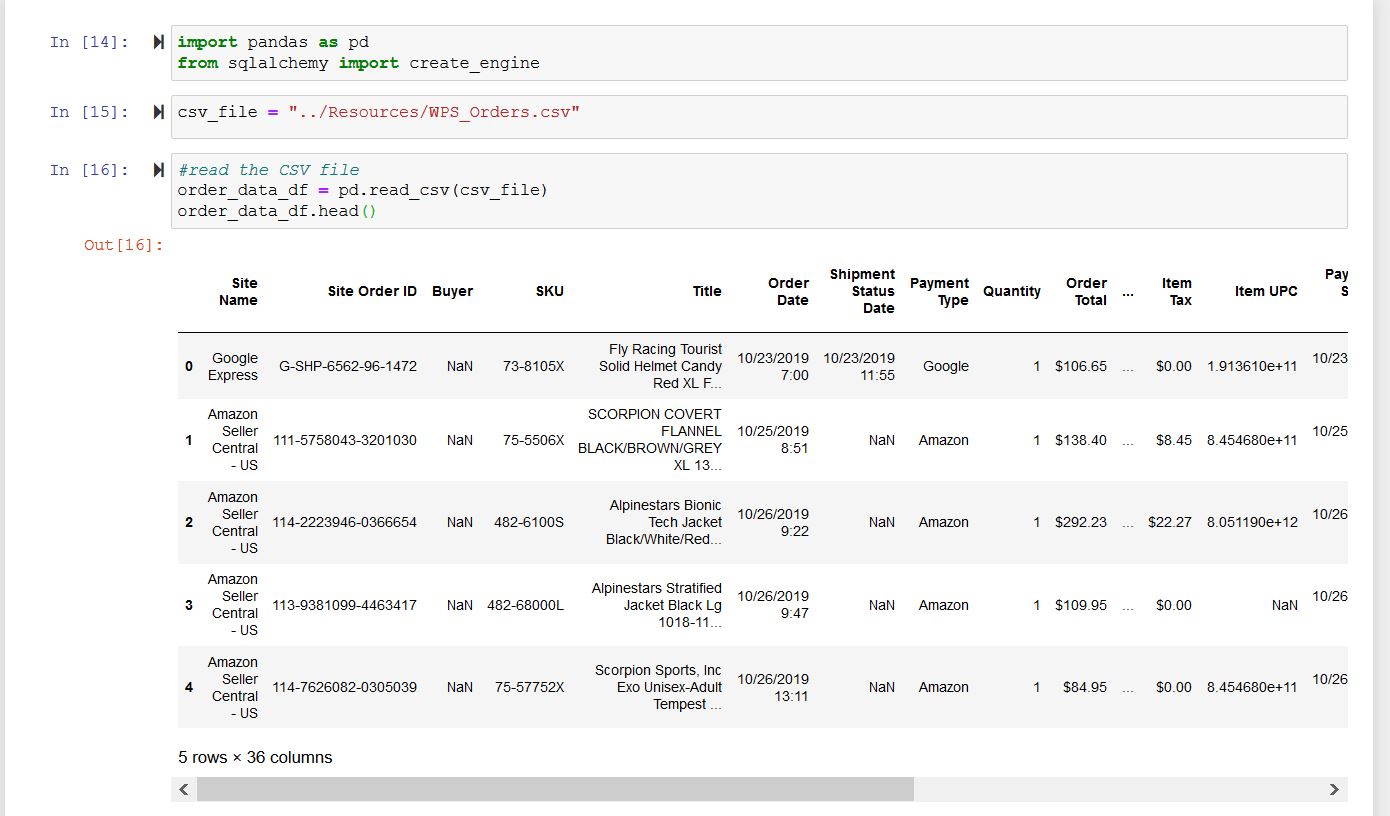
Dealer Table – It will store all Dealer Data. We will add more dealer related information on project 2.

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

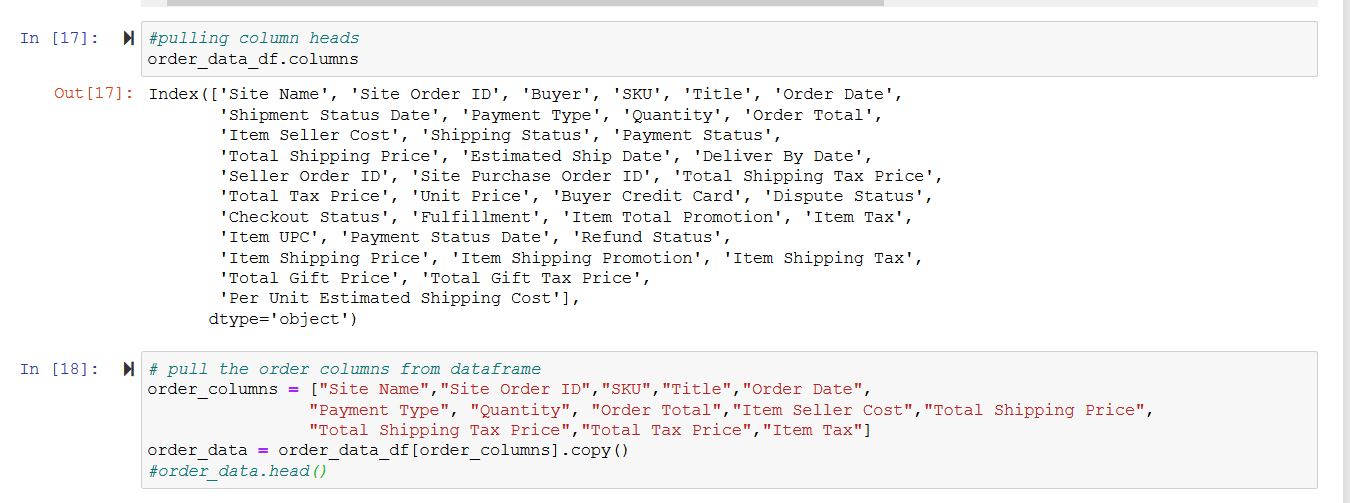
**Data Cleansing**

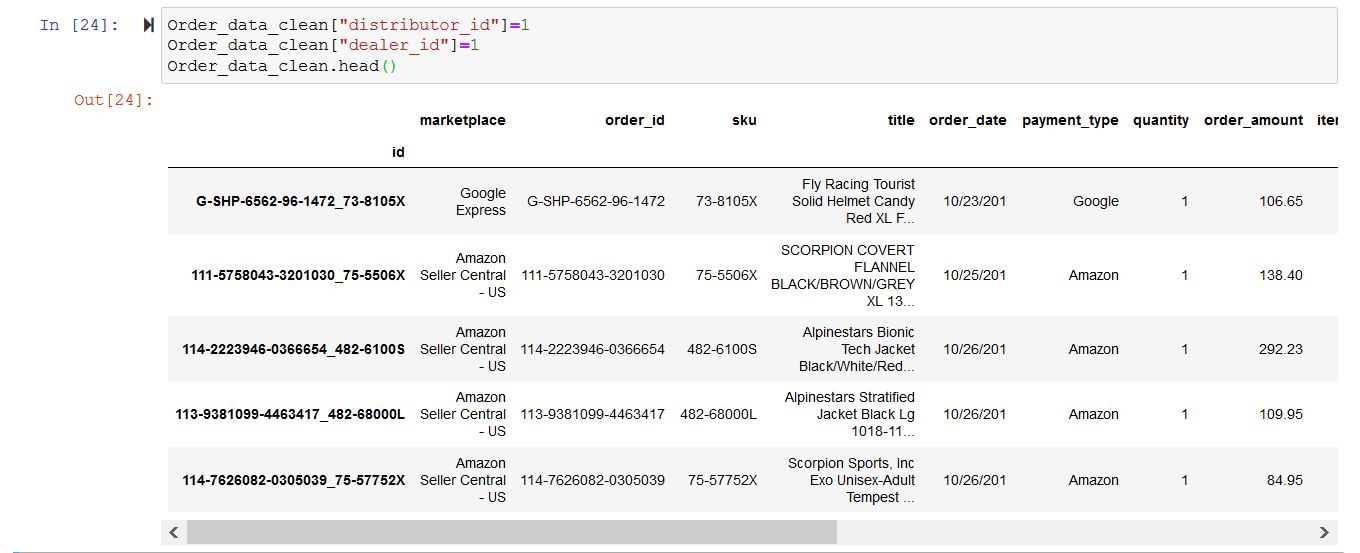
***Set up Orders table:***

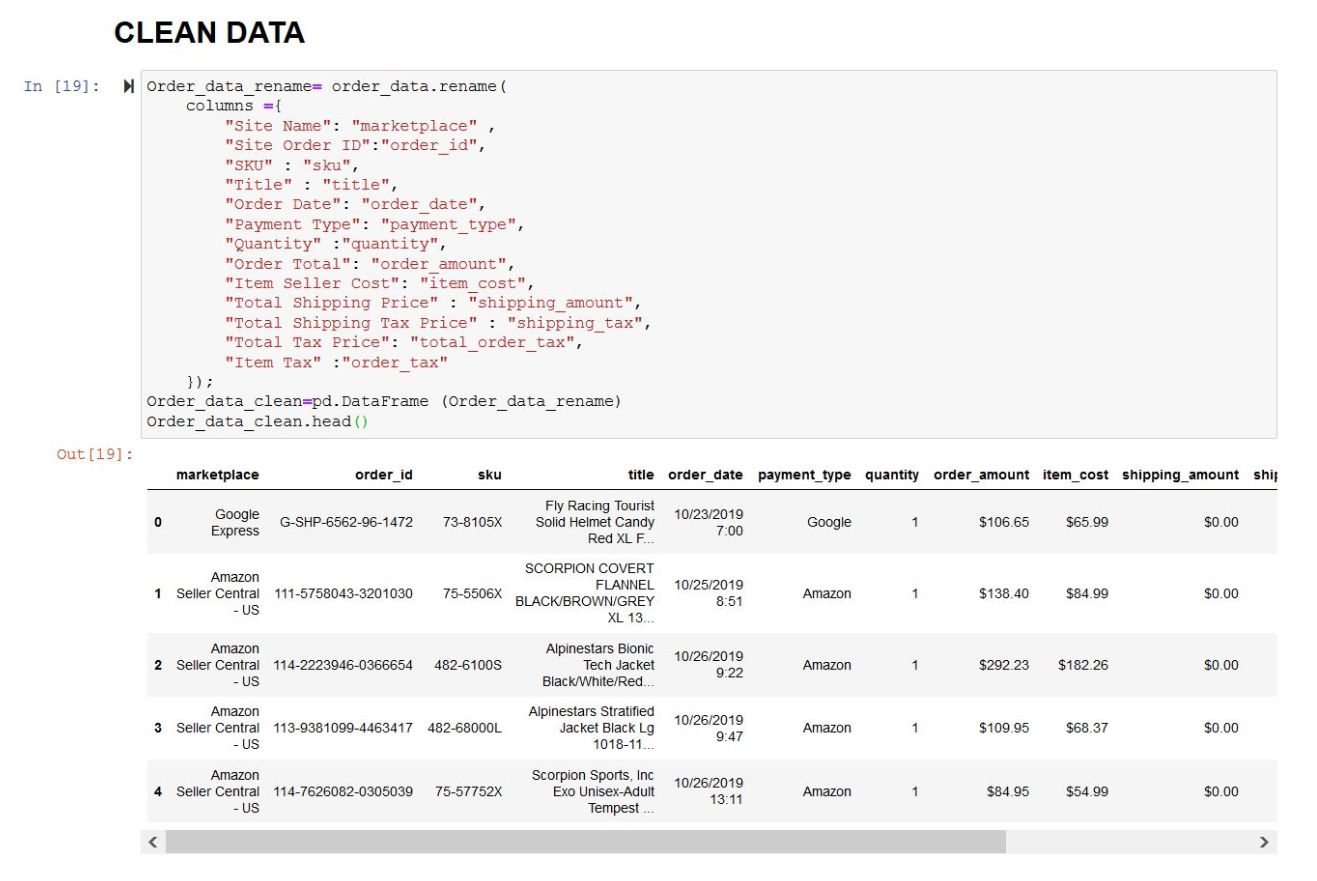
1. Extract Data from CSV file and convert it to Data Frame



1. Clean data before loading to database









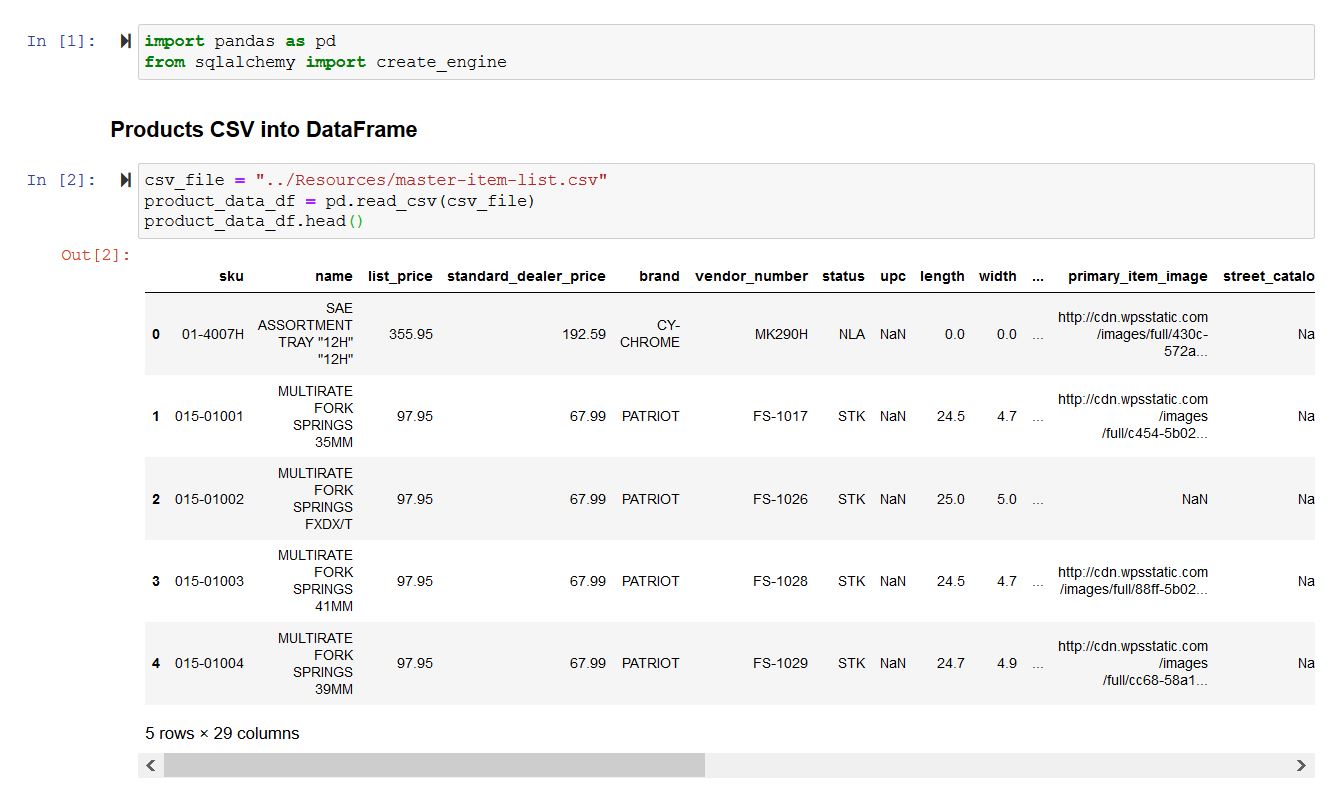


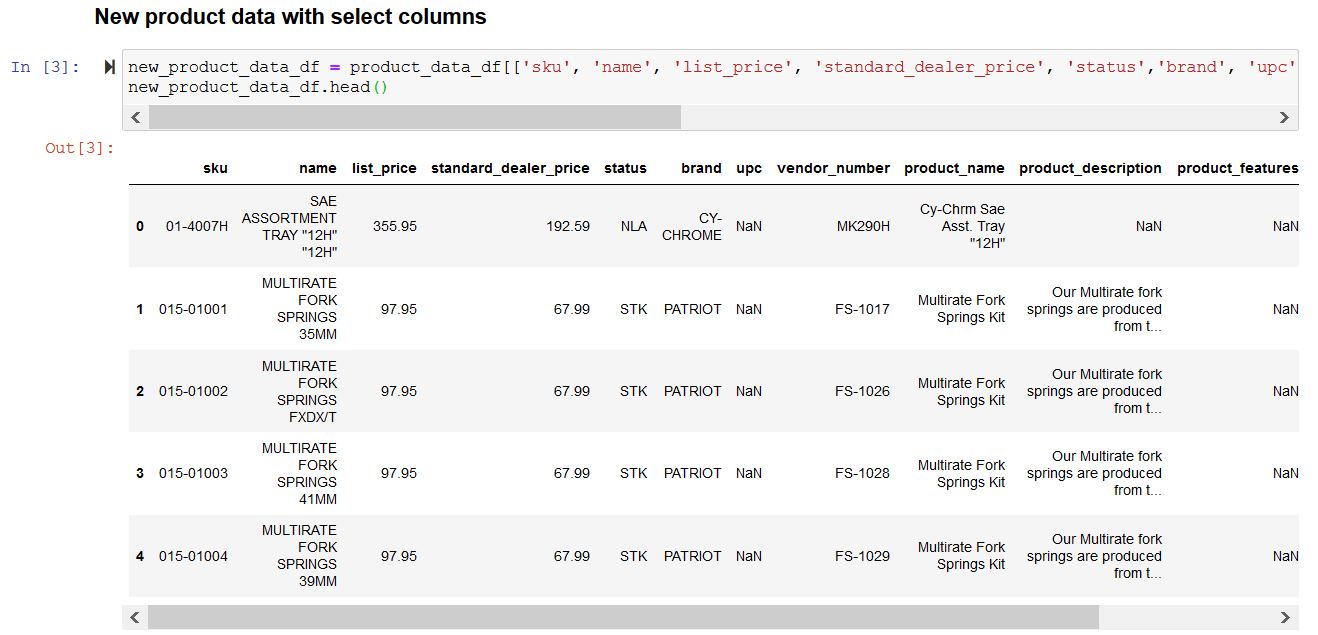
1. Load the data to SQL database. The tables type has been created in the pgAdmin in advance.



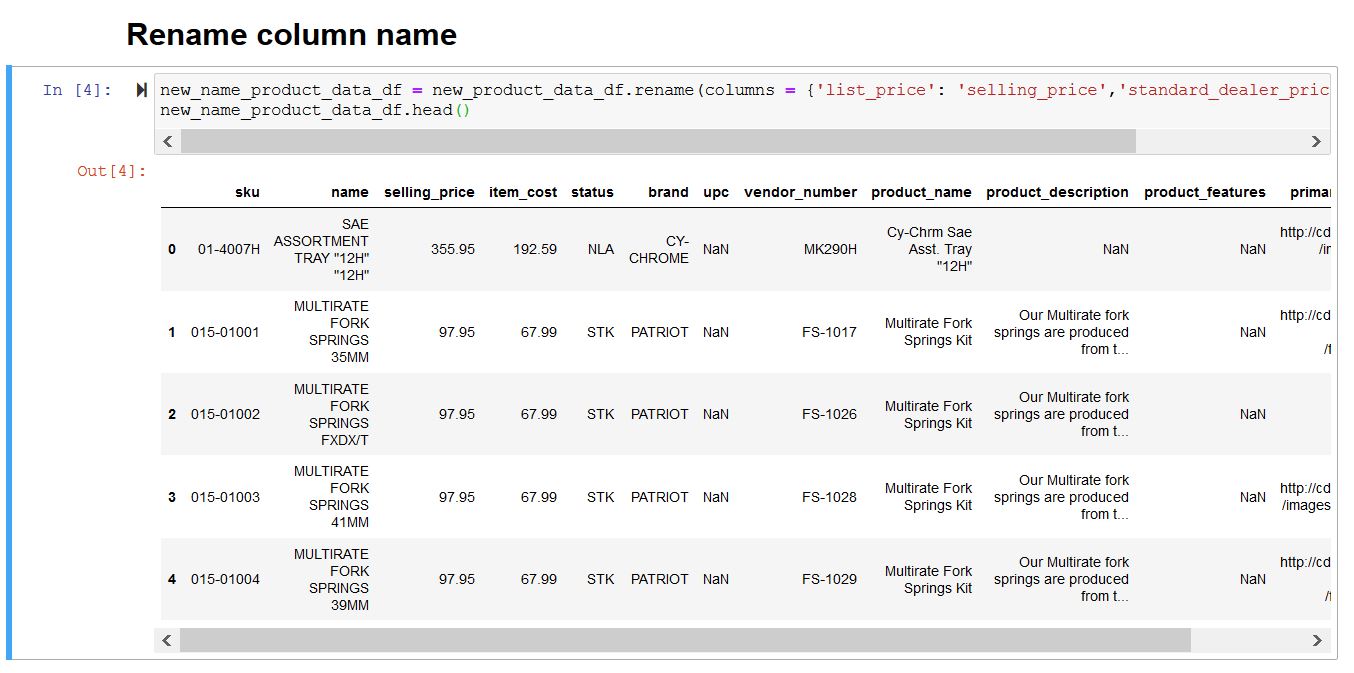
***Set up products table*** (Supplier File)***:***

1. Extract Data from CSV file and convert it to Data Frame





1. Clean data before loading to database



1. Load data to database

