ETL - EXTRACT TRANSFORM AND LOAD

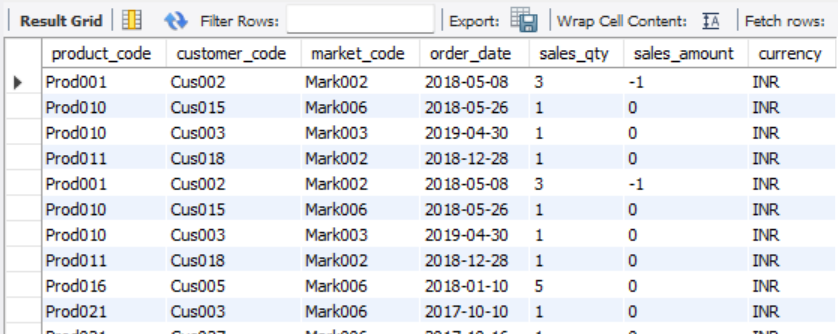
* Connect the database to Power BI
* Load the tables

DATA MODELING

* Check for the relatable fields
* Implement star schema (Manually connect relatable fields)
  + Fact Table - Transaction Table
  + Dimension Table - Customer Table, Market Table, Customer Table, Product Table.

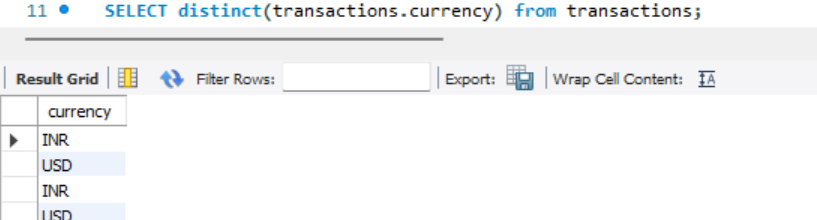
DATA CLEANING(Transformation)

* All the sales are running in India so we will filter the markets table as it has “NEW YORK” and “PARIS” in the “ZONE” column.
* There are few negative and zero values in Transactions table under ”sales\_amount” which are garbage.



* + Remove these values
* In Transactions table there are some “USD” values under “currency” remove them.
  + Create new converted currency column and convert all the currency to INR.
    - Create a custom column where if [currency]="USD" then 1 else null.
    - Then convert “USD” to “INR”.
      * **FORMULA:-**
      * = Table.AddColumn(#"Filtered Rows", "Normalised\_currency", each if[currency]="USD" then [sales\_amount]\*75 else [sales\_amount])
      * There is a another USD Format -"USD#(cr)" so convert into INR as well.

* + - * = Table.AddColumn(#"Filtered Rows", "Normalised\_currency", each if[currency]="USD" or [currency]=”USD#(cr)”then [sales\_amount]\*75 else [sales\_amount])
* There are some Duplicate values in Transactions table.



* Duplicates and messy data are :-
  + INR\r has more no. of sales than INR.
  + USD\r is a duplicate of USD.