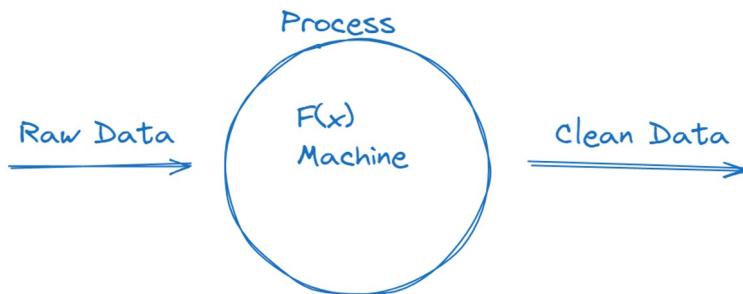


## Data Transformation - 1

---

→ Messy Data / Raw in format ----> Clean them up.  
And make it good for having insights.  
More Stories , patterns.



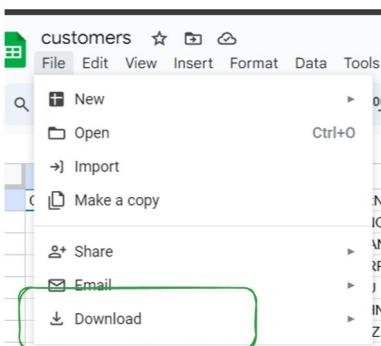
### Customer Table :

Perform the below mentioned operations in Power BI:

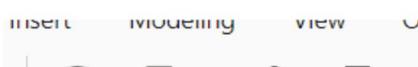
Customer Table:

- Check and modify the data type of each column in the customer table.
- Rename the modified dataset file as "d-customer".
- Eliminate duplicate dates from the birth date column in the customer table.
- Combine the columns "prefix, first name, and last name" and name it "Full Name".
- Sort the annual income column in descending order.

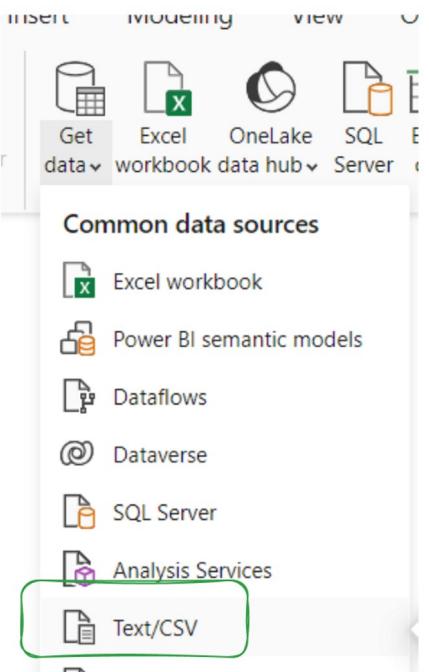
→ Open the csv file name Customer.csv.  
Go to file > Download > .csv format.



→ Next Step : Get Data > Import through the format you have chosen at the download . Either Excel Format / CSV format.



you have chosen to use the download . Excel / CSV formats.



→ Instead of Loading Data Directly. Transform the raw data for the cleaning process.

The screenshot shows the 'customers - customers.csv' file in the Power BI Transform Data dialog. The 'File Origin' dropdown shows '65001: Unicode (UTF-8)'. The 'Delimiter' dropdown is set to 'Comma'. The 'Data Type Detection' dropdown is set to 'Based on first 200 rows'. The main area displays the customer data with columns like CustomerKey, Prefix, FirstName, LastName, BirthDate, MaritalStatus, Gender, EmailAddress, AnnualIncome, and Total. At the bottom right, there are three buttons: 'Load' (green), 'Transform Data' (highlighted with a green rounded rectangle), and 'Cancel'.

→ Promote the headers :

The screenshot shows the 'Transform Data' editor for the 'customers - customers.csv' file. The 'Data type: Text' dropdown is open, and the 'Use First Row as Headers' checkbox is checked. The editor interface includes buttons for Choose Columns, Remove Columns, Keep Rows, Remove Rows, Split Column, Group By, Replace Values, and Transform. Below the editor, a preview table shows columns for FirstName, LastName, BirthDate, and MaritalStatus with their respective data types and validation statistics.

PCT Columns		PCT Columns		PCT Columns		PCT %
99%	Valid	100%	Valid	100%	Valid	V
0%	Error	0%	Error	0%	Error	E
< 1%	Empty	0%	Empty	0%	Empty	E
FirstName		LastName		BirthDate		Mar
JON		YANG		4/8/1966		M
EUGENE		HUANG		5/14/1965		S
RUBEN		TORRES		8/12/1965		M

→ Changing the data type :

Text Format Date - Locale Change type [Date]

Change Type with Locale

Change the data type and select the locale of origin.

Data Type: Text  
Locale: English (India)

OK Cancel

Input should be same to check the locale

mm / dd / yyyy

Text

\$90,000  
\$60,000  
\$60,000  
\$70.000

Currency - We only want numerical value without any special character.

0 1 2 3 4 5 6 7 8

\$ 1 1 1, 1 1 1

Suppose - "," will be included.

Extract Text Range

10 characters of length.

### Extract Text Range

Enter the index of the first character, and the number of characters to keep.

Starting Index	<b>1</b>
Number of Characters	<b>10</b>

10 characters of length.

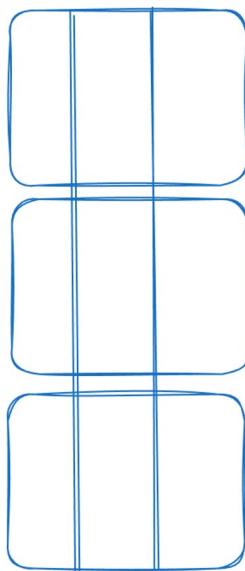
Sales Data - 2015 , 2016 , 2017.

OrderDate	StockDate	OrderNumber	ProductKey	CustomerKey	TerritoryKey	OrderLineItem	OrderQuantity
1/1/2015	9/21/2001	SO45080	332	14657	1	1	1
1/1/2015	12/5/2001	SO45079	312	29255	4	1	1
1/1/2015	10/29/2001	SO45082	350	11455	9	1	1

2015

2016 , 2017

Appended :  
Vertical Direction

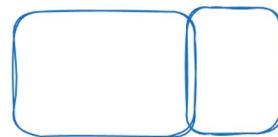


2015

2016

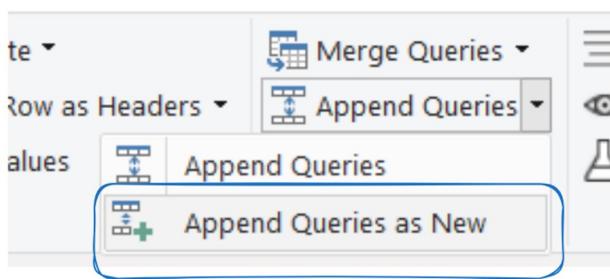
2016

Appending different files  
one below the other in  
vertical direction.



Appending

Horizontal Direction  
adding new column is merging



## Append

Concatenate rows from three or more tables into a single table.

Two tables  Three or more tables

sales 2015 - sales 2015 (Current)

d-customers

sales 2016 - sales 2016

sales 2017 - sales 2017

Add >>

sales 2015 - sales 2015 (Current)  
sales 2016 - sales 2016  
sales 2017 - sales 2017

current file to see 2015 date on top

OK

Cancel