

Data Visualization Tools and Software

Aim : Prepare Data in Power BI Desktop : Connect and transform Raw data

Task:

1. Install Power-BI
2. Prepare data in Power-BI

In this lab, we learn how to:

Open Power BI Desktop

Set Power BI Desktop options

Types of Data Connectors;

- Connect to source data (Excel/CSV/etc.)
- Load data

Preview source data

Explore Transform Data options

- Clean data
- Remove/Replace Null values, duplicate values, missing value
- Merge / split columns
- Add columns, remove columns
- Change data types

Preview of Data

Dataset named as “99Bikers_Raw_data”

A1	transaction_id	product_id	customer_id	transaction_date	online_order	order_status	brand	product_line	product_class	product_size	list_price	standa
1	1	2	2950	2/25/2017	FALSE	Approved	Solex	Standard	medium	medium	71.49	
3	2	3	3120	5/21/2017	TRUE	Approved	Trek Bicycles	Standard	medium	large	2091.47	NA
4	2	3	3120	5/21/2017	TRUE	Approved	Trek Bicycles	Standard	medium	large		NA
5	3	37	402	10/16/2017	FALSE	Approved	OHM Cycles	Standard	low	medium	1793.43	
6	3	37	402	10/16/2017	FALSE	Approved	OHM Cycles	Standard	low	medium	1793.43	
7	4	88	3135	8/31/2017	FALSE	Approved	Norco Bicycles	Standard	medium	medium	1198.46	
8	5	78	787	10/1/2017	TRUE	Approved	Giant Bicycles	Standard	medium	large	1765.3	
9	6	25	2339	3/6/2017	TRUE	Approved	Giant Bicycles	Road	medium	medium	1538.99	
10	7	22	1542	4/21/2017	TRUE	Approved	WeareA2B	Standard	medium	medium	60.34	
11	8	15	2459	7/15/2017	FALSE	Approved	WeareA2B	Standard	medium	medium		
12	9	67	1305	8/10/2017	FALSE	Approved	Solex	Standard	medium	large		
13	10	12	3262	8/30/2017	TRUE	Approved	WeareA2B	Standard	medium	medium	1231.15	
14	10	12	3262	8/30/2017	TRUE	Approved	WeareA2B	Standard	medium	medium		
15	11	5	1986	1/17/2017	FALSE	Approved	Trek Bicycles	Mountain	low	medium	574.64	
16	12	61	2783	1/5/2017	TRUE	Approved	OHM Cycles	Standard	low	medium	71.16	
17	13	35	1243	2/26/2017	TRUE	Approved	Trek Bicycles	Standard	low	medium	1057.51	
18	14	16	2717	9/10/2017	FALSE	Approved	Norco Bicycles	Standard	high	small	1661.92	
19	15	12	247	6/11/2017	FALSE	Approved	Giant Bicycles	Standard	medium	large	1765.3	

Fig.-Transactions

A1	first_name	last_name	gender	past_3_years_DOB	job_title	job_industry_cate	wealth_segment	deceased_indicator	owns_car	tenure	address
2	Chickie	Brister	Male	86	1957-07-12	General Manager	Manufacturing	Mass Customer	N	Yes	14 45 Shopko
3	Morly	Genery	Male	69	1970-03-22	Structural Engineer	Property	Mass Customer	N	No	16 14 McCorm
4	Ardelia	Forrester	Female	10	1974-08-28	Senior Cost Accou	Financial Services	Affluent Customer	N	No	10 5 Colorad
5	Lucine	Stutt	Female	64	1979-01-28	Account Represen	Manufacturing	Affluent Customer	N	Yes	5 207 Annar
6	Melinda	Hadlee	Female	34	1965-09-21	Financial Analyst	Financial Services	Affluent Customer	N	No	19 115 Monta
7	Drudi	Brandli	Female	39	1951-04-29	Assistant Media Pl	Entertainment	High Net Worth	N	Yes	22 89105 Pear
8	Rutledge	Hallt	Male	23	1976-10-06	Compensation An	Financial Services	Mass Customer	N	No	8 7 Nevada C
9	Nancie	Vian	Female	74	1972-12-27	Human Resources	Retail	Mass Customer	N	Yes	10 85 Carioca
10	Duff	Karlowicz	Male	50	1972-04-28	Speech Pathologis	Manufacturing	Mass Customer	N	Yes	5 717 West C
11	Barthel	Docket	Male	72	1985-08-02	Accounting Assista	IT	Mass Customer	N	Yes	17 80 Scofield
12	Rockwell	Matson	Male	94	1995-01-01	Programmer Anal	Retail	High Net Worth	N	No	3 3682 Crow
13	Wheeler	Winward	Male	48	1999-08-30	Environmental Sci	Manufacturing	Mass Customer	N	No	10 3 Golden L
14	Olag		Male	60	1990-05-13	Human Resources	Telecommunicat	Mass Customer	N	No	9 0484 North
15	Melba	Spellacy	Female	38	1976-12-09	VP Marketing	Health	Mass Customer	N	No	4 0591 Anzir
16	Mandie	Fearles	Female	32	1964-04-19	Clinical Specialist	Health	Mass Customer	N	No	10 39 Kedzie I
17	Dukie	Swire	Male	88	1954-03-31		Manufacturing	Affluent Customer	N	Yes	5 64 Granby
18	Marcelia	Monkleigh	Female	61	1993-08-22	Associate Profess	Manufacturing	Mass Customer	N	Yes	4 610 Swallo
19	Winnifred	Beswetherick	Female	83	1976-06-08	Actuary	Financial Services	Mass Customer	N	No	14 61 4th Stre

Fig.- NewCustomerList

A	B	C	D	E	F	G	H	I	J	K	
1	customer_id	first_name	last_name	gender	past_3_years_bike_DOB	job_title	job_industry_categ	wealth_segment	deceased_indicator	default	owns_a_bike
2	1	Laraine	Medendorp	F	93	1953-10-12	Executive Secretary	Health	Mass Customer	N	"
3	2	Eli	Bockman	Male	81	1980-12-16	Administrative Offic	Financial Services	Mass Customer	N	<script>alert('hi')</s
4	3	Arlin	Dearle	Male	61	1954-01-20	Recruiting Manager	Property	Mass Customer	N	1-Feb
5	4	Talbot		Male	33	1961-10-03	IT	Mass Customer	N	0 { ... } > \${\$0} !No	Yes
6	5	Sheila-kathryn	Calton	Female	56	1977-05-13	Senior Editor	n/a	High Net Worth	N	NIL
7	6	Curr	Duckhouse	Male	35	1966-09-16	Retail	Financial Services	Affluent Customer	N	89 8 8 8
8	7	Fina	Merali	Female	6	1976-02-23	n/a	Media Manager I	Affluent Customer	N	89 8 8 8
9	8	Rod	Inder	Male	31	1962-03-30	Business Systems	Agribusiness	Mass Customer	N	(-89 8 8 8)
10	9	Mala	Lind	Female	97	1973-03-10	Media Manager I	Financial Services	Affluent Customer	N	0D
11	10	Fiorenze	Birdall	Female	49	1968-10-11	Senior Quality Engg	Manufacturing	Mass Customer	N	89 8 8 8
12	11	Uriah	Visatt	Male	99	1954-04-30	Property	Financial Services	Mass Customer	N	89 8 8 8 E
13	12	Sawyer	Flatman	Male	58	1994-07-21	Nuclear Power Eng	High Net Worth	Mass Customer	N	nil
14	13	Gabriele	Norcross	Male	38	1955-02-15	Developer I	Manufacturing	Affluent Customer	N	-1.00E+02
15	14	Rayshell	Kitterman	Female	85	1983-03-25	Account Executive	Financial Services	High Net Worth	N	89 8 8 8
16	15	Erroll	Radage	Male	91	2000-07-13	Junior Executive	Agribusiness	Affluent Customer	N	8
17	16	Harlin	Parr	Male	38	1977-02-27	Media Manager IV	n/a	Mass Customer	N	1.00E+96
18	17	Heath	Faraday	Male	57	1962-03-19	Sales Associate	n/a	Affluent Customer	N	1957 8 8 8 8 8 8 8 8 8 8 8
19	18	Marjie	Neasham	Female	79	1967-07-06	Professor	n/a	Affluent Customer	N	100 8 8 8 8 8 8 8 8 8 8 8

Fig.- CustomerDemographic

Lab Assignment No: 2 Data Transferring

dv_assignment - Google Drive

Vivek_23070149030_DV_Assignment

99Bikers_Raw_data.xlsx - Google Sheets

docs.google.com/spreadsheets/d/1AOUFqJjdAUairTdMIZsYbf-JtSN3VCw/edit#gid=360167927

vivek.gurve.mtech2... vivek.gurve.mtech2... SIT-Classroom colab.google

99Bikers_Raw_data.xlsx

File Edit View Insert Format Data Tools Help

Share

A1 | customer_id

	A	B	C	D	E	F	G	H	I	J	K
1	customer_id	address	postcode	state	country	property_valuation					
2	1	060 Morning Avenue	2016	New South Wales	Australia	10					
3	2	6 Meadow Vale Court	2153	New South Wales	Australia	10					
4	4	0 Holy Cross Court	4211	QLD	Australia	9					
5	5	17979 Del Mar Point	2448	New South Wales	Australia	4					
6	6	9 Oakridge Court	3216	VIC	Australia	9					
7	7	4 Delaware Trail	2210	New South Wales	Australia	9					
8	8	49 Londonderry Lane	2650	New South Wales	Australia	4					
9	9	97736 7th Trail	2023	New South Wales	Australia	12					
10	11	93405 Ludington Park	3044	VIC	Australia	8					
11	12	44339 Golden Leaf Alley	4557	QLD	Australia	4					
12	13	2 Sutherland Street	3799	VIC	Australia	6					
13	14	9 McBride Trail	2760	New South Wales	Australia	8					
14	15	9861 New Castle Avenue	2428	New South Wales	Australia	9					
15	16	52 Moland Street	3331	VIC	Australia	4					
16	17	82391 Kensington Lane	3058	VIC	Australia	9					
17	18	092 2nd Alley	2135	New South Wales	Australia	12					
18	19	59 Spaight Circle	2233	New South Wales	Australia	9					
19	20	032 Bartelt Crossing	2444	New South Wales	Australia	8					

Transactions NewCustomerList CustomerDemographic CustomerAddress

Type here to search

Fig.- CustomerAddress

Data Transformation

Power BI Power Query Editor can be used to edit or transform data files before they get loaded into the Power BI dashboard. The Query Editor serves as an intermediate data container that allows you to modify data by choosing columns and rows, pivoting and unpivoting columns, splitting columns and rows, etc.

Power BI Desktop has three views:

- Report view – use queries that you create to build compelling visualizations, arranged as we want them to appear, and with multiple pages, that you can share with others.
- Data view – See the data in your report in data model format, where we can add measures, create new columns, and manage relationships.
- Model view – Get a graphical representation of the relationships that are established in your data model, and manage or modify them as needed.

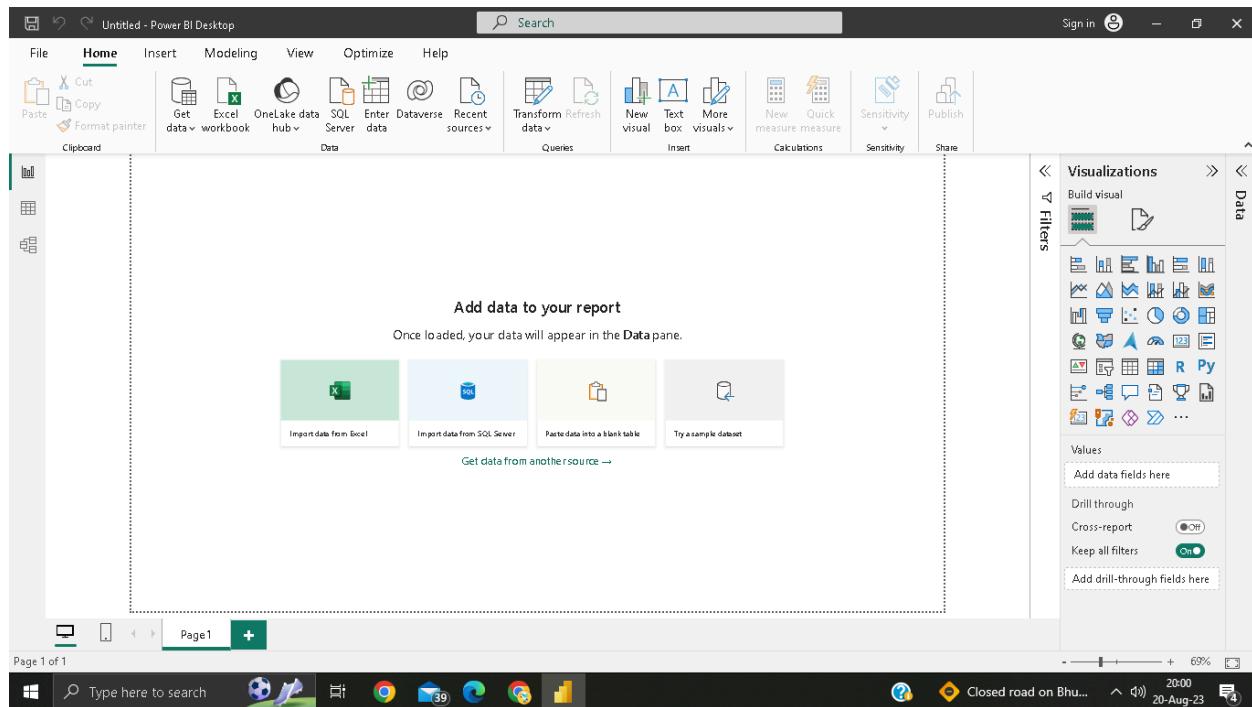


Fig.- Power BI Desktop

Get Data

Select the Text/CSV option in Get Data. This action launches a local file browser where you can select your text file. Select Open to open the file. From the Navigator, you can either transform the data in the Power Query Editor by selecting Transform Data, or load the data by selecting Load.

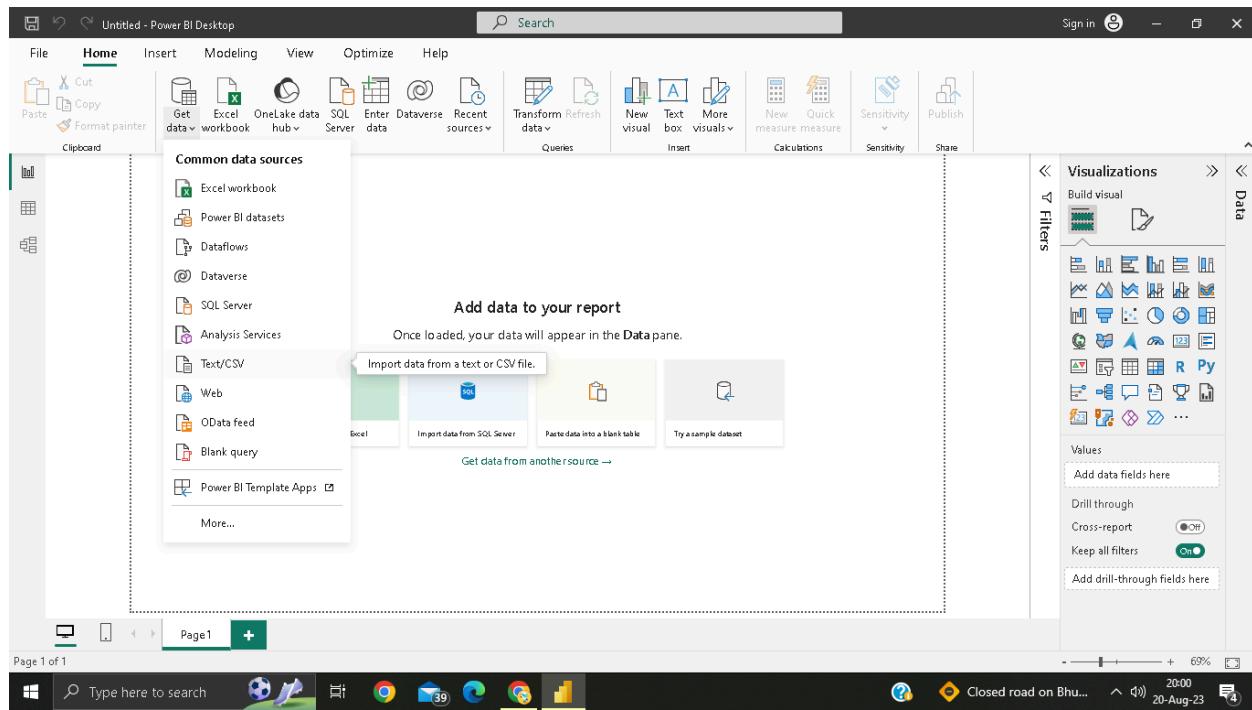


Fig.- Importing data from a CSV file

Navigator

The Navigator window displays a preview. At this point, you can edit the query before loading the table, by selecting Transform Data from the bottom of the window, or just load the table.

From the Navigator, select Transform Data to begin transforming the data in the Power Query Editor.

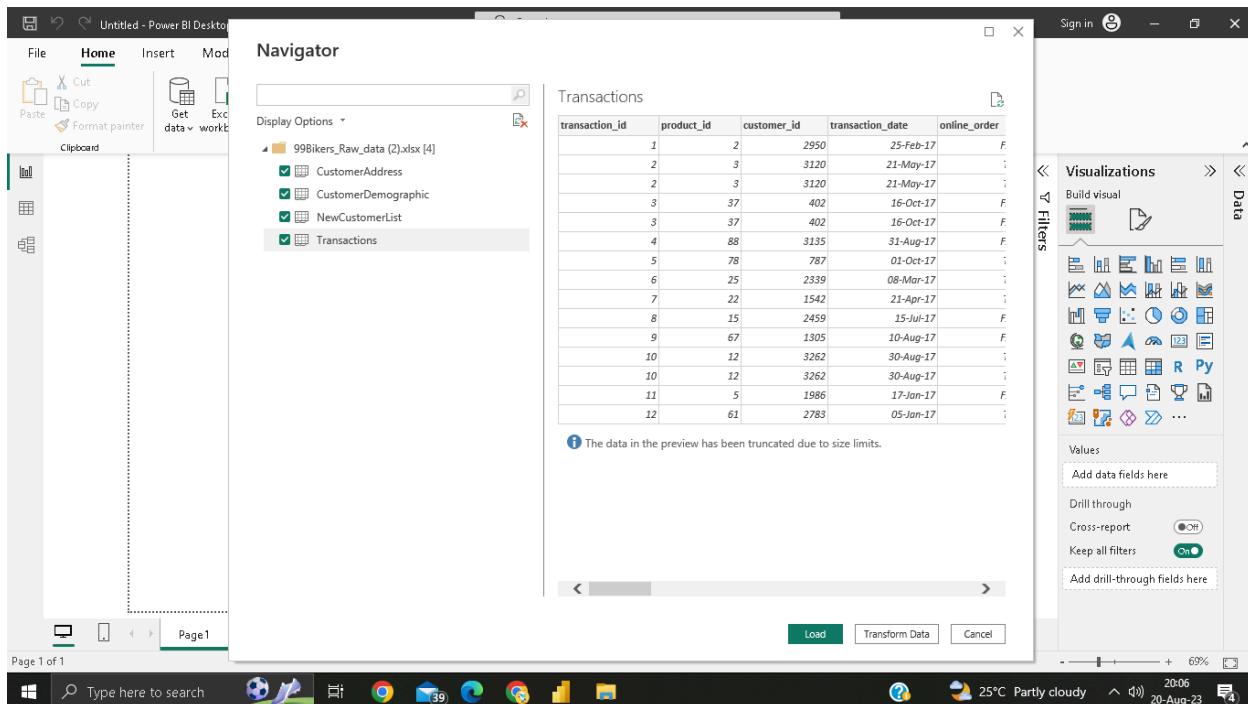


Fig.- Navigator from a CSV file

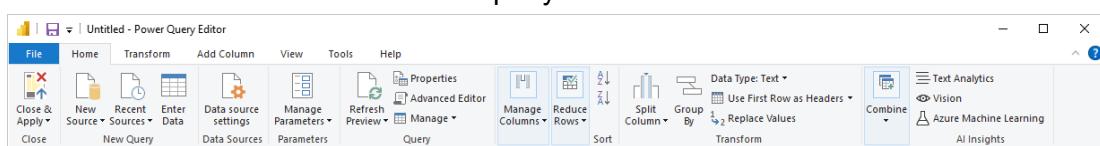
Power Query Editor

Use Power Query Editor to connect to one or many data sources, shape and transform the data to meet our needs, then load that model into Power BI Desktop.

The query ribbon

The ribbon in Power Query Editor consists of four tabs: Home, Transform, Add Column, View, Tools, and Help.

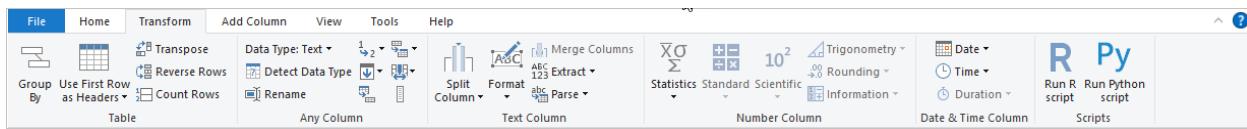
The Home tab contains the common query tasks.



Transform tab

The Transform tab provides access to common data transformation tasks, such as:

- Adding or removing columns
- Changing data types
- Splitting columns
- Other data-driven tasks



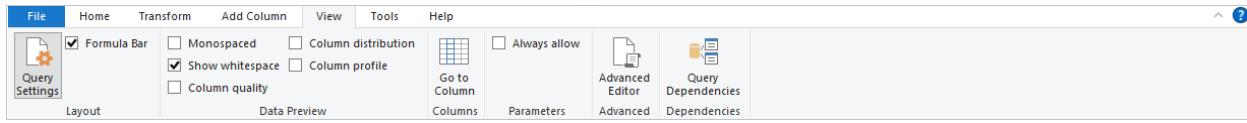
Add Column

The Add Column tab provides more tasks associated with adding a column, formatting column data, and adding custom columns. The following image shows the Add Column tab.



View tab

The View tab on the ribbon is used to toggle whether certain panes or windows are displayed. It's also used to display the Advanced Editor. The following image shows the View tab.



Extract

Creating a new column containing a specified number of characters from the start of each value in the selected column.

The screenshot shows the Power Query Editor interface. A query named 'NewCustomerList' is open. In the 'Transform' tab, the 'Format' section has the 'Extract' button selected. A tooltip explains: 'Create a new column containing a specified number of characters from the start of each value in this column.' Below it, a dropdown menu shows '1' selected. The formula bar contains the expression: `, {"first_name", type text}, {"last_name", type text}, {"gender", type text}`. The main table view shows columns for first name, last name, gender, and other details like past purchases and DOB. The 'APPLIED STEPS' pane on the right shows the 'Changed Type' step applied to the 'gender' column.

Fig.-Extract first character from gender column

This screenshot shows the 'Insert First Characters' dialog box overlaid on the Power Query Editor. The dialog asks 'Enter how many starting characters to keep.' with a 'Count' input field containing the value '1'. The 'OK' button is highlighted. The background shows the same 'NewCustomerList' query and 'Extract' step as the previous screenshot.

Fig.-Enter how many starting character to keep

Untitled - Power Query Editor

File Home Transform Add Column View Tools Help

Column From Custom Invoke Custom Examples Column Function Duplicate Column General

Merge Columns Extract Statistics Standard Scientific Rounding Trigonometry Date Time Duration From Text From Number From Date & Time AI Insights

Queries [4]

NewCustomerList

	Column21	Rank	Value	M or F	Column24	Column25
1	3125	1	1	M	null	
2	3125	1	1	M	null	
3	0.77	1	1	F	null	
4	1.575	4	4	F	null	
5	3625	4	4	F	null	
6	0.65	6	6	F	null	
7	1845	6	6	M	null	
8	7625	8	8	F	null	
9	3125	8	8	M	null	
10	3125	10	10	M	null	
11	.225	10	10	M	null	
12	.629	12	12	M	null	
13	.5875	13	13	M	null	
14	3125	14	14	F	null	
15	.442	14	14	F	null	
16	.5875	16	16	M	null	
17	.7625	17	17	F	null	
18	.0125	17	17	F	null	
19	.7625	19	19	F	null	
20	.075	19	19	F	null	
21	.5625	21	21	M	null	
22	.375	21	21	M	null	
23	.085	23	23	M	null	
24						

43 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 20:05

Query Settings

PROPERTIES

Name: NewCustomerList

All Properties

APPLIED STEPS

- Source
- Navigation
- Promoted Headers
- Changed Type
- Inserted First Characters
- Renamed Columns
- Removed Columns
- Reordered Columns

Fig.-M or F column extracted from gender column

Change Data Type

Converts the values from the Any data type to a data type based on the inspection of the values from each column.

The screenshot shows the Microsoft Power Query Editor interface. The top ribbon has tabs: File, Home, Transform, Add Column, View, Tools, Help. The 'Transform' tab is selected. The main area shows a table with columns: product_line, product_class, product_size, list_price, standard_cost, and product_first. A context menu is open over the 'list_price' column, showing options like Decimal Number, Fixed decimal number, Whole Number, Percentage, Date/Time, Date, Time, Duration, Text, True/False, and Binary. The 'list_price' column contains values such as 1.2, \$53.62, 248.82, etc. To the right, the 'Query Settings' pane shows 'Transactions' selected under 'Name'. The 'APPLIED STEPS' pane lists 'Changed Type' as the last step applied.

product_line	product_class	product_size	list_price	standard_cost	product_first
1 Standard	medium	medium	1.2	Decimal Number	\$53.62
2 Standard	medium	large		Fixed decimal number	
3 Standard	medium	large		Whole Number	
4 Standard	low	medium		Percentage	248.82
5 Standard	low	medium		Date/Time	248.82
6 Standard	medium	medium		Date	381.1
7 Standard	medium	large		Time	709.48
8 Road	medium	medium		Date/Time/Timezone	829.65
9 Standard	medium	medium		Duration	45.26
10 Standard	medium	medium		Text	13.44
11 Standard	medium	large		True/False	380.74
12 Standard	medium	medium		Binary	161.6
13 Standard	medium	medium		Using Locale...	161.6
14 Mountain	low	medium	1057.51	null	56.93
15 Standard	low	medium			154.4
16 Standard	low	medium	1661.92		1479.11
17 Standard	high	small			
18 Standard	medium	large	1765.3	null NA	
19 Standard	medium	large	1765.3	null NA	
20 Standard	medium	large	2091.47	388.92	
21 Standard	medium	medium	1555.58	818.01	
22 Standard	medium	small	1311.44	1167.18	
23 Standard	medium	medium	1292.84	13.44	
24					

Fig.-Change list_price Data Type to Fixed decimal number

The screenshot shows the Microsoft Power Query Editor interface. The ribbon at the top includes File, Home, Transform, Add Column, View, Tools, and Help. The 'Transform' tab is selected. The 'Conditional Column' section is open, showing options like 'Column From Examples', 'Custom Function', 'Invoke Custom Function', 'General', 'Conditional Column', 'Index Column', 'Merge Columns', 'Format', 'Duplicate Column', 'From Text', 'From Number', 'From Date & Time', 'Text Analytics', 'Vision', and 'Azure Machine Learning'. A preview pane on the left lists four queries: CustomerDemographic, Transactions, NewCustomerList, and CustomerAddress. The main area displays a table with columns: product_line, product_class, product_size, list_price, standard_cost, and product. The 'list_price' column contains values like 71.49, 2,091.47, null, and Error. The 'standard_cost' column contains values like 53.62, 248.82, 248.82, and 381.10. The 'product' column contains values like 1,231.15, 161.60, 161.60, and null. The 'Transactions' query is currently selected. On the right, the 'Query Settings' pane shows 'PROPERTIES' for 'Transactions' and 'APPLIED STEPS' for 'Changed Type'. The status bar at the bottom indicates 33 COLUMNS, 999+ ROWS, and a preview download at 20:35 on 20-Aug-23.

Fig.- Changed Data Type of list_price

Column Quality

Shows the percentage of rows within a column that are valid, empty, or errors. From there we can hover over the summary box (or right-click on the column quality section) for additional prompts to correct any issues.

The screenshot shows the Power Query Editor interface with the 'Column quality' checkbox selected in the Formula Bar. The Data Preview pane displays a table with columns: product_class, product_size, list_price, standard_cost, and product_id. A tooltip above the preview area says 'Show column quality details in data preview.' The 'product_id' column has a tooltip indicating 'transformColumnTypes("#Promoted Headers", {"transaction_id", Int64.Type}, {"product_id", Int64.Type})'. The 'list_price' column has a tooltip indicating '\$ list_price'. The 'standard_cost' column has a tooltip indicating '\$ standard_cost'. The 'product_id' column also has a tooltip indicating 'ABC 123 product'. The Query Settings pane on the right shows the query name 'Transactions' and the applied step 'Changed Type'. The taskbar at the bottom shows various icons and the date '20-Aug-23'.

Fig.-Show column quality details in data preview

This screenshot shows the same Power Query Editor interface as above, but the 'transaction_id' column is selected in the Data Preview pane. The tooltip for this column now displays detailed column quality statistics: '99% Valid, 0% Error, < 1% Empty'. The Query Settings pane remains the same, showing the query name 'Transactions' and the applied step 'Changed Type'. The taskbar at the bottom shows the date '20-Aug-23'.

Fig.-Column quality details

Remove Duplicate

Use for removing duplicate rows and columns in Power BI. To do this, select the rows or columns that contain duplicates, then right-click and select the "Remove Duplicate Rows" or "Remove Duplicate Columns" option. This will create a new table with the duplicate rows or columns removed.

The screenshot shows the Power Query Editor interface. On the left, there's a navigation pane with 'Queries [4]' and several data sources like 'CustomerDemographic', 'Transactions', 'NewCustomerList', and 'CustomerAddress'. The main area displays a table with columns: transaction_id, product_id, customer_id, transaction_date, online_order, and order_status. A context menu is open over the first few rows of the transaction_id column, with 'Remove Duplicates' highlighted. To the right, the 'Query Settings' pane shows the query name is 'Transactions' and the applied steps include 'Promoted Headers' and 'Changed Type'. The status bar at the bottom indicates '33 COLUMNS, 999+ ROWS' and 'Column profiling based on top 1000 rows'.

Fig.-Remove duplicates in transaction_id

Remove Errors

To remove rows with errors in Power Query, first select the column that contains errors.

The screenshot shows the Power Query Editor interface. The 'list_price' column is selected, and a context menu is open, with 'Remove Errors' highlighted. The 'APPLIED STEPS' pane shows a step named 'Removed Duplicates'.

Fig.-list_price and standard_price column with Errors

From the drop-down menu, select Remove errors.

The screenshot shows the Power Query Editor interface. The 'list_price' column is selected, and a context menu is open, with 'Remove Errors' highlighted. The 'APPLIED STEPS' pane shows a step named 'Removed Duplicates'.

Fig.-Remove errors

Untitled - Power Query Editor

File Home Transform Add Column View Tools Help

Query Settings

Layout

Queries [4]

CustomerDemographic
Transactions
NewCustomerList
CustomerAddress

Data Preview

Columns Parameters Advanced Dependencies

Formula Bar

Monospaced Column distribution Always allow

Show whitespace Column profile

Go to Column Advanced Editor Query Dependencies

Advanced Editor Query Dependencies

Queries [4]

= Table.RemoveRowsWithErrors(#"Removed Errors", {"standard_cost"})

	A ^B product_class	A ^B product_size	\$ list_price	\$ standard_cost	ABC 123 product.first_sold_date
1	medium	medium	list_price	99% Valid	53.62
2	low	medium	997 (99%) Valid	0 (0%) Error	248.82
3	medium	medium	99% Valid	0% Error	381.10
4	medium	large	99% Valid	0% Error	709.48
5	medium	medium	99% Valid	0% Error	829.65
6	medium	medium	99% Valid	0% Error	45.26
7	medium	medium	null	13.44	
8	medium	large	null	380.74	
9	medium	medium	1,231.15	161.60	
10	low	medium	574.64	null	
11	low	medium	71.16	56.93	
12	low	medium	1,057.51	154.40	
13	high	small	1,661.92	1,479.11	
14	medium	large	1,765.30	null NA	
15	medium	large	2,091.47	388.92	
16	medium	medium	1,555.58	818.01	
17	medium	small	1,311.44	1,167.18	
18	medium	medium	1,292.84	13.44	
19	medium	medium	1,538.99	829.65	
20	medium	medium	499.53	388.72	
21					

33 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 20:05

Windows Type here to search 24°C Mostly clear 20-Aug-23

Fig.-list_price and standard_price column without Errors

Remove Empty

When you have blank values in your data table, you can use Power Query transformations to remove the row with blank values.

The screenshot shows the Power Query Editor interface. In the 'Data Preview' pane, there is a table with columns: product_class, product_size, list_price, standard_cost, and product_first_sold_date. The 'list_price' column contains several empty cells. A context menu is open over the third empty cell in the 'list_price' column, with the 'Remove Empty' option highlighted. The 'APPLIED STEPS' pane on the right shows a step named 'Removed Errors1'.

Fig.-list_price column 3 rows are empty

This screenshot is identical to the one above, showing the Power Query Editor with the 'list_price' column containing empty cells. A context menu is open over the third empty cell in the 'list_price' column, with the 'Remove Empty' option highlighted. The 'APPLIED STEPS' pane on the right shows a step named 'Removed Errors1'.

Fig.-list_price column remove empty

The screenshot shows the Power Query Editor interface with the following details:

- File**, **Home**, **Transform**, **Add Column**, **View**, **Tools**, **Help** menu items.
- Layout** tab selected in the ribbon.
- Query Settings** pane on the left.
- Data Preview** pane showing a table with columns: **product_class**, **product_size**, **list_price**, **standard_cost**, and **product_first_sold_date**.
- Advanced Editor** button.
- Dependencies** button.
- Query Settings** pane on the right with **Properties** and **Applied Steps** sections.
- Applied Steps** list includes: Source, Navigation, Promoted Headers, Changed Type, Removed Duplicates, Removed Errors, and **Removed Errors1**.
- Removed Errors1** step is expanded, showing a list of rows where the **list_price** column contains null or empty values.
- Bottom status bar: 33 COLUMNS, 999+ ROWS, PREVIEW DOWNLOADED AT 20:05.

Fig.-list_price column with no empty row

Replace values

In Power Query, you can replace one value with another value in a selected column. You can replace specific values or the whole value in a cell.

The screenshot shows the Power Query Editor interface with the following details:

- File**, **Home**, **Transform**, **Add Column**, **View**, **Tools**, **Help** menu items.
- Transform** tab selected in the ribbon.
- Table** context menu open, showing options like Transpose, Data Type, Replace Values, Unpivot Columns, etc.
- Replace Values** dialog box is open over the preview area.
- Replace Values** dialog: **list_price** column selected, **Replace existing values in the currently selected columns with the specified new value.**
- Replace Values** dialog: **list_price < null and [list_price] < ""**
- Applied Steps** list includes: Source, Navigation, Promoted Headers, Changed Type, Removed Duplicates, Removed Errors, and **Removed Errors1**.
- Removed Errors1** step is expanded, showing a list of rows where the **list_price** column contains null or empty values.
- Bottom status bar: 33 COLUMNS, 999+ ROWS, PREVIEW DOWNLOADED AT 20:05.

Fig.-Replace value in standard_cost column

The screenshot shows the Power Query Editor interface with the 'Replace Values' dialog open over a table query. The table has columns: list_price, standard_cost, product_first_sold_date, Column14, and Column15. The 'standard_cost' column contains many null values. The 'Replace Values' dialog has 'Value To Find' set to 'null' and 'Replace With' set to '0'. The 'APPLIED STEPS' pane shows the step 'Removed Errors1'.

Fig.-Replace value Null with 0 in standard_cost column

The screenshot shows the Power Query Editor interface with the table query now showing all 'standard_cost' values replaced by 0. The 'APPLIED STEPS' pane shows the step 'Replaced Value'.

Fig.- Null value Replaced with 0 in standard_cost column

Extract Component

Extract Month & Year from a Date, In the 'Date' dialog box, select the 'Month' options. This formula creates a new text column that displays the month

The screenshot shows the Power Query Editor interface. A context menu is open over the 'transaction_date' column, specifically at the top right of the column header. The menu path 'Month' -> 'Month' is highlighted. A tooltip explains: 'Extract the Month component from the Date/time values in the selected column.' The 'Month' option is also highlighted in the menu. The main table view shows transaction data with columns like transaction_id, product_id, customer_id, transaction_date, online_order, and order_status.

Fig.-Extract Month from transaction_date column

This screenshot shows the Power Query Editor after the transformation step. The 'transaction_date' column has been converted into a new column named 'Month'. The 'Month' column contains the extracted month names (e.g., Jan, Feb, Mar, etc.) corresponding to the dates in the original 'transaction_date' column. The 'Month' column is highlighted in the table. The 'APPLIED STEPS' pane on the right shows the step 'Extracted Month'.

Fig.-Extracted Month from transaction_date column

Extract Month & Year from a Date, In the 'Date' dialog box, select the 'Year' options. This formula creates a new text column that displays the Year

The screenshot shows the Power Query Editor interface. A context menu is open over the 'transaction_date' column, specifically at the cell containing '25-Feb-17'. The 'Extract' option is selected, and a submenu is displayed with the 'Year' option highlighted. The submenu also includes 'Month', 'Day', and other date-related options like 'Start of Max' and 'End of Max'. The main query editor window shows a table with columns: transaction_id, product_id, customer_id, transaction_date, and online_order. The transaction_date column contains various dates like '25-Feb-17', '16-Oct-17', etc.

Fig.-Extract Year from transaction_date column

This screenshot shows the Power Query Editor after the transformation. The table now includes an additional column, 'order_status', which appears to be a calculated column based on the original data. The 'transaction_date' column has been converted into a year column. The 'APPLIED STEPS' pane on the right side of the editor highlights the 'Extracted Year' step, indicating it was the last action taken. The main pane shows the same table structure as the previous screenshot, with the newly added 'order_status' column.

Fig.-Extracted Year from transaction_date column

Use Headers as First Row

To promote the first row to column headers, select Home > Use First Row As Headers. To demote column headers to the first row, select Home, select the arrow next to Use First Row As Headers, and then select Use Headers as First Row.

The screenshot shows the Microsoft Power Query Editor interface. The ribbon at the top has tabs for File, Home, Transform, Add Column, View, Tools, and Help. The Home tab is currently selected. In the ribbon, under the 'Transform' section, there is a dropdown menu labeled 'Use Headers' with two options: 'Use First Row as Headers' (selected) and 'Use Headers as First Row'. Below the ribbon, the main workspace displays a table with 33 columns and 999+ rows. The table has several columns with validation status (Valid, Error, Empty) and some columns containing 'null' values. On the right side of the screen, there is a 'PROPERTIES' pane showing the query name 'Transactions' and an 'APPLIED STEPS' pane listing various transformation steps. The bottom of the screen shows the Windows taskbar with icons for Start, Search, Task View, File Explorer, Edge, Mail, Google Chrome, and File Explorer, along with system status like weather (24°C Partly cloudy), date (20-Aug-23), and battery level.

Fig.-Use First Row As Headers

Group By

The Group By feature in Power BI summarizes data based on one or more aggregate functions. You can use the Group By feature to find the average, count, min, max, or any other aggregate value for one column, based on unique values in other columns.

The screenshot shows the Power Query Editor interface with the 'Transform' tab selected. A tooltip for the 'Group By' button indicates: "Group rows in this table based on the values in the currently selected columns." The 'Standard' column is selected. The formula bar shows the query: `= Table.TransformColumnTypes(#"Promoted Headers1", {{"1", Int64.Type}, {"2", Int64.Type}, {"3", Int64.Type}, {"4", Int64.Type}, {"5", Int64.Type}, {"6", Int64.Type}, {"7", Int64.Type}, {"8", Int64.Type}, {"9", Int64.Type}, {"10", Int64.Type}, {"11", Int64.Type}, {"12", Int64.Type}, {"13", Int64.Type}, {"14", Int64.Type}, {"15", Int64.Type}, {"16", Int64.Type}, {"17", Int64.Type}, {"18", Int64.Type}, {"19", Int64.Type}, {"20", Int64.Type}, {"21", Int64.Type}})`. The 'APPLIED STEPS' pane shows the 'Changed Type2' step, which is highlighted.

Category	Value	Count
Valid	100%	13912
Error	0%	3894
Empty	0%	418
Total	100%	12123

Fig.-Use group by on standard column

The screenshot shows the Power Query Editor interface with the 'Transform' tab selected. The 'Queries [4]' pane shows a new query named 'CustomerDemographic'. The formula bar shows the query: `= Table.Group(#"Changed Type2", {"Standard"}, {"Count": each Table.RowCount(_), Int64.Type})`. The 'APPLIED STEPS' pane shows the 'Grouped Rows' step, which is highlighted.

Category	Count
Standard	13912
Road	3894
Mountain	418
Touring	12123

Fig.-Grouped row of standard column

Reverse rows

With Power Query, it's possible to reverse the order of rows in a table.

The screenshot shows the Power Query Editor interface. A table named 'Transactions' is selected. In the 'Transform' ribbon, the 'Reverse Rows' button is highlighted. The 'APPLIED STEPS' pane shows the 'Promoted Headers1' step. The table data includes columns for ID, Approved, and Date. The 'Approved' column is highlighted in blue. The 'APPLIED STEPS' pane also lists other steps like 'Navigation', 'Changed Type', and 'Removed Duplicates'.

Fig.-Reverse rows of column 1

Move option

A common process when preparing data is to move columns in the dataset.

To accomplish this move, you can either select the Move option or drag and drop the column.

The screenshot shows the Power Query Editor interface. A table named 'Transactions' is selected. In the 'Transform' ribbon, the 'Move' button is highlighted. The 'APPLIED STEPS' pane shows the 'Reversed Rows' step. The table data includes columns for ID, Approved, and Date. The 'Approved' column is being moved to the right. The 'APPLIED STEPS' pane also lists other steps like 'Navigation', 'Changed Type', and 'Removed Duplicates'.

Fig.-Move selected column to right

Age

There is an Age menu option in the Power Query editor under Date. When you select a date column and use that Age option, it calculates the duration between the selected date and the current date in days.

The screenshot shows the Power Query Editor interface with a table query named "Transactions". The "Age" context menu is open over a date column, specifically the "Approved" column which contains dates like "2000-09-22". The menu includes options for "Date Only", "Parse", "Year", "Month", "Quarter", "Week", "Day", "Combine Date and Time", "Earliest", and "Latest". A tooltip for "Age" states: "Return the duration between the current local date and the values in the selected columns." Below the menu, a list of transformation steps is visible, including "Reversed Rows" which is currently selected.

Fig.-Calculates the duration between the selected date

This screenshot shows the same Power Query Editor session after the "Age" operation has been performed. The "Approved" column now contains duration values in days, such as "2158.00:00:00" and "2239.00:00:00". The "Calculated Age" step is highlighted in the "APPLIED STEPS" pane. The status bar at the bottom indicates the preview was downloaded at 20:05.

Fig.-Calculated duration between the selected date

Sort data

We can sort data to arrange in:

Ascending order (text is A-Z, number is smallest to largest, and date is oldest to newest).

Descending order (text is Z-A, number is highest to lowest, and date is newest to oldest).

The screenshot shows the Power Query Editor interface. The ribbon is visible with various tabs like File, Home, Transform, etc. In the center, there's a preview of a table with columns: job_title, job_industry_category, and wealth_se. The 'DOB' column is currently selected. On the left, the 'Queries [4]' pane shows a list of queries. On the right, the 'APPLIED STEPS' pane shows the history of steps taken, with 'Filtered Rows1' and 'Filtered Rows2' listed. The status bar at the bottom indicates 'PREVIEW DOWNLOADED AT 20:05'.

Fig.-Sort Ascending order of DOB column

This screenshot is identical to the one above, showing the Power Query Editor with the 'DOB' column sorted ascending. The preview shows the same range of dates from 08-Jun-38 to 28-Aug-39. The interface, applied steps, and status bar are all consistent with the first screenshot.

Fig.-Sort Ascending order of DOB

Save your work

When our query is where we want , select Close & Apply from Power Query Editor's File menu.
This action applies the changes and closes the editor.

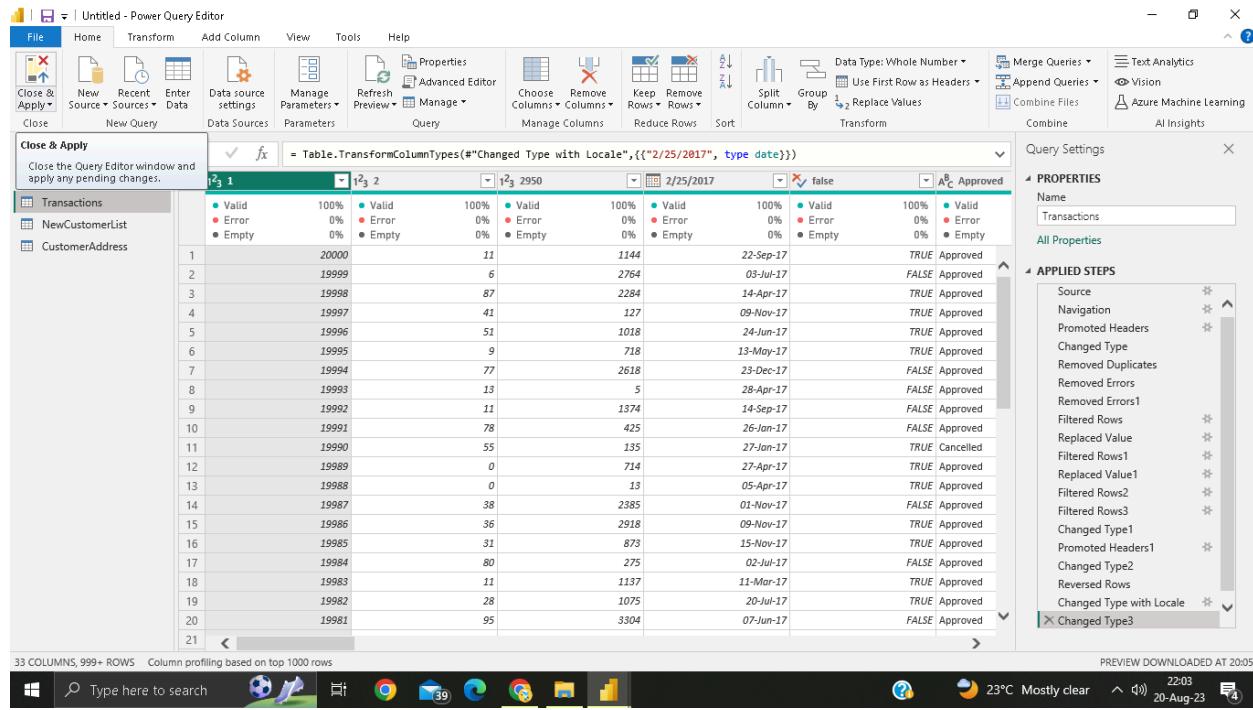


Fig.-Close the editor

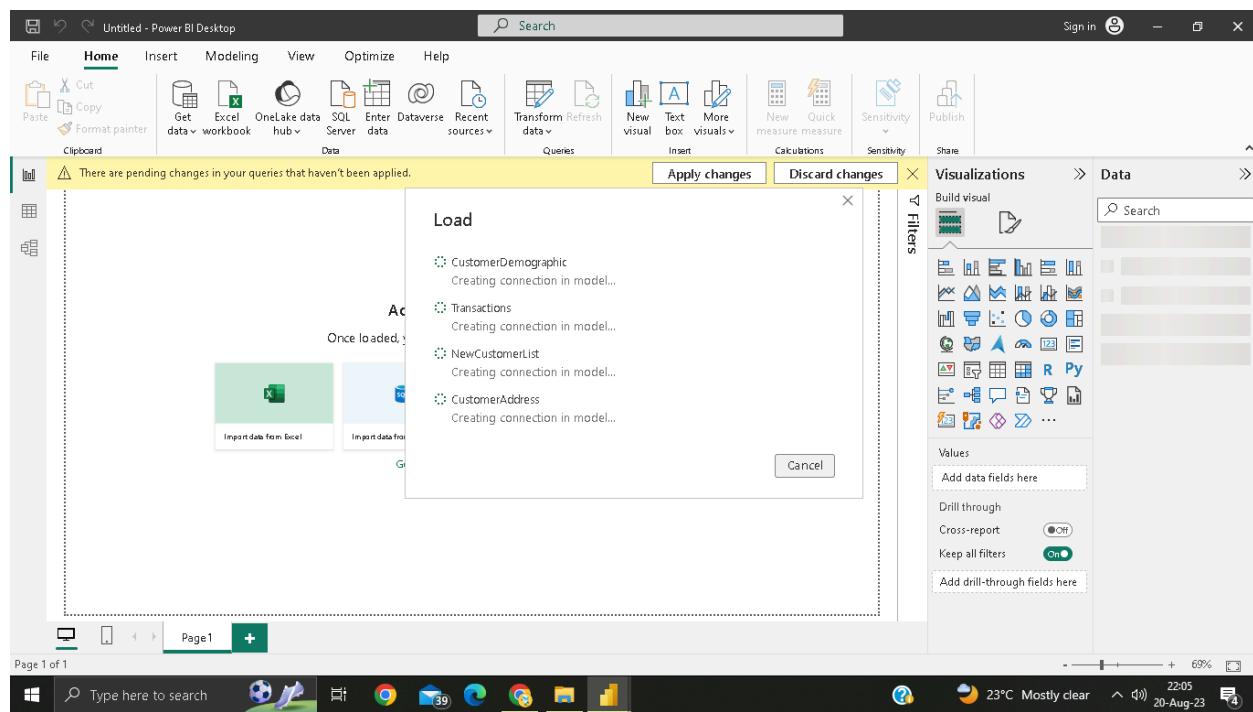


Fig.-Creating model

Model view

Get a graphical representation of the relationships that are established in our data model, and manage or modify them as needed.

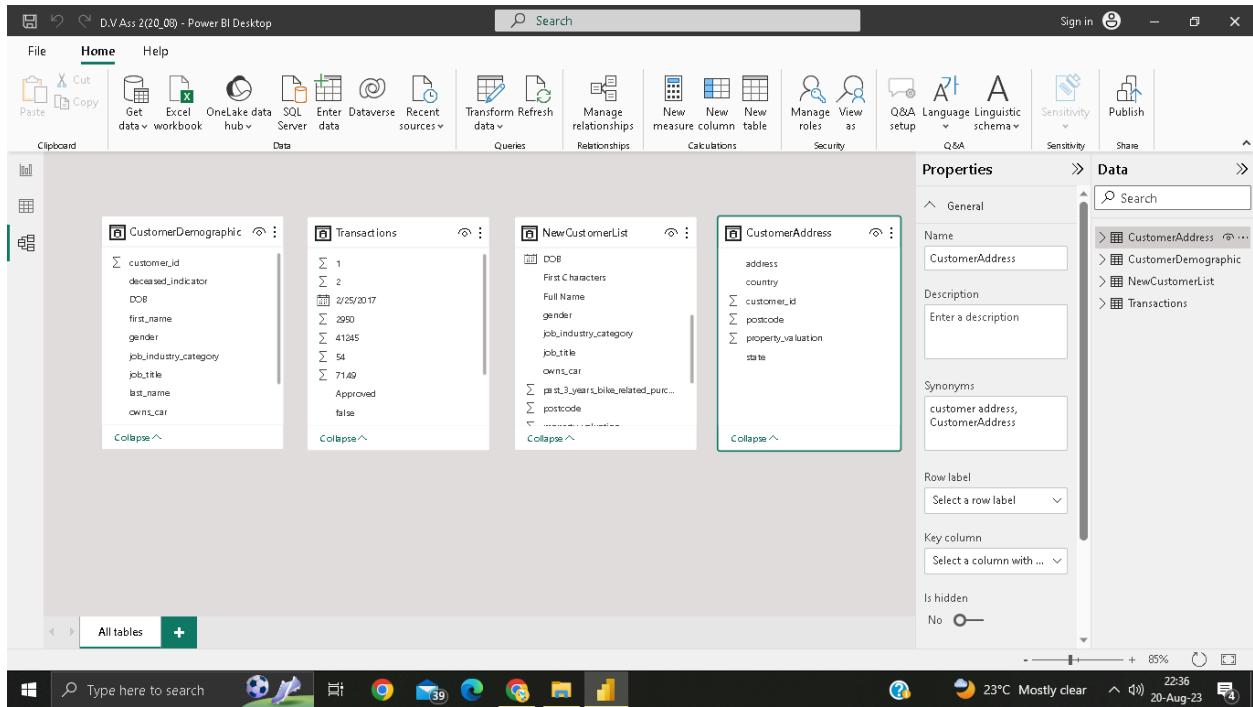


Fig.-Model view