Power BI – Data Transformations

April 2018

Insights & Data

Capgemini

AGENDA

Data Connectivity

- Files
- Database
- Azure
- Online Services
- Other

Data Transformations

- Choose Columns
- Remove Columns
- Keep Rows
- Remove Rows
- Split Column
- Group By
- Use First Row as Headers
- Replace Values
- Combine
- Transpose
- Reverse Rows
- Count Rows
- Rename

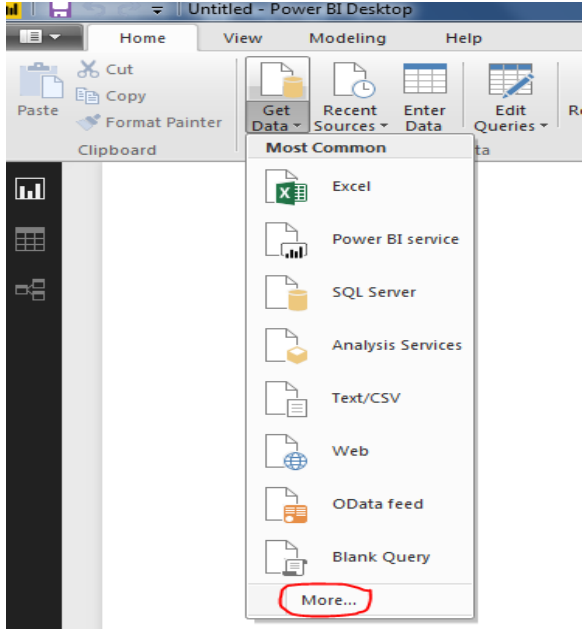
Contd.. Data Transformations

- Pivot Column
- Unpivot Columns
- Move
- Convert to List
- Fill
- Transform
- Change Type
- Duplicate Column
- Format (Lowercase, Uppercase, Capitalize Each Word, Add Prefix etc...)
- Extract (First Characters, Last Characters, Length, Range etc...)
- Statistics (Sum, Minimum, Maximum, Median, Average etc...)

- Standard (Add, Subtract, Multiply, Divide etc...)
- Scientific (Power, Square Root, Absolute Value etc…)
- Rounding (Round Up, Round Down)
- Information (Is Even, Is Odd and Sign)
- Trigonometry (Sine, Cosine, Tangent etc…)

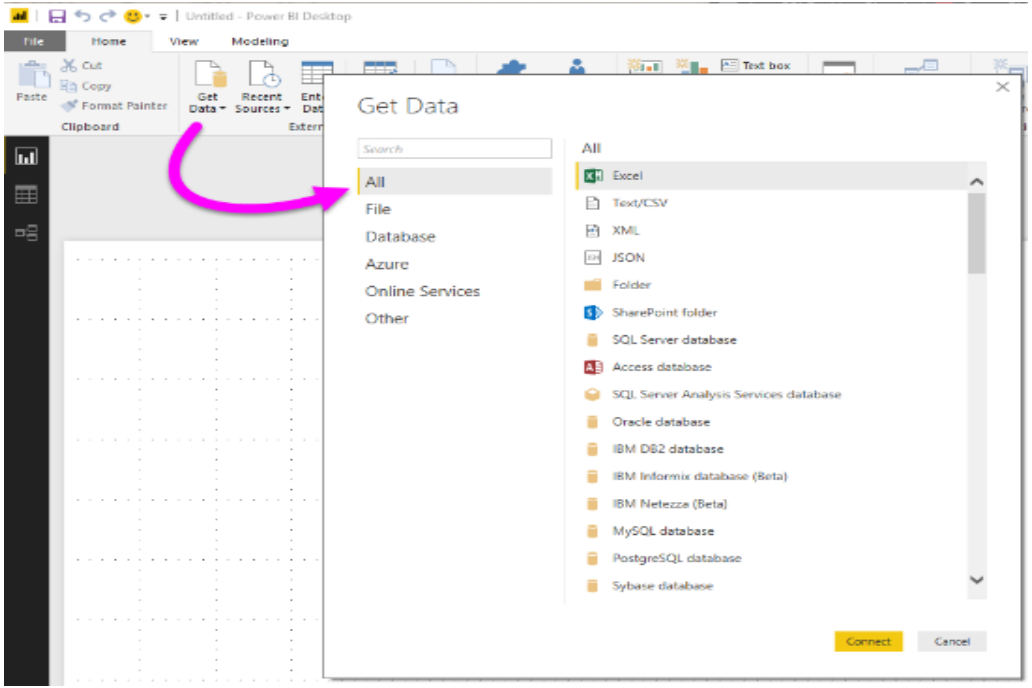
Data Connectivity

- With Power BI Desktop, you can connect to data from many different sources.
- To connect to data, select **Get Data** from the **Home** ribbon. Selecting the down arrow, or the **Get Data** text on the button, shows the **Most Common** data types menu. Please refer below screenshot.



Data Connectivity cont.…

- Selecting **More...** from the **Most Common** menu displays the **Get Data** window.

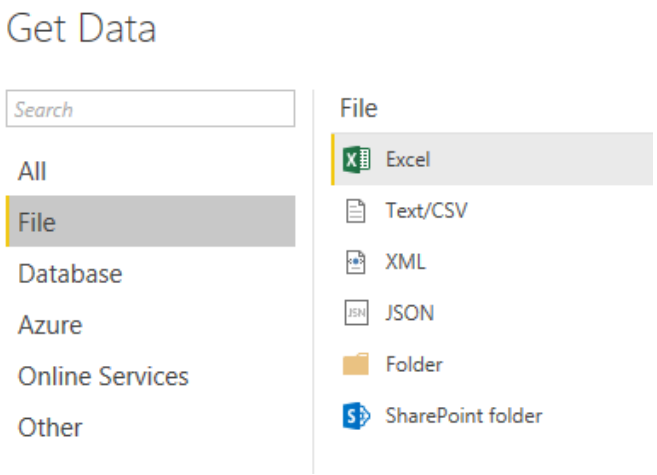


Data Connectivity cont.…

- **Data Sources** are organized in the following categories.
  - All
  - File
  - Database
  - Azure
  - Online Services
  - Other

- The **All** category includes all data connection types from all categories.
- The File category provides the following data connections:
  - Excel
  - Text/CSV
  - XML
  - JSON
  - Folder
  - SharePoint Folder

Data Connectivity cont….



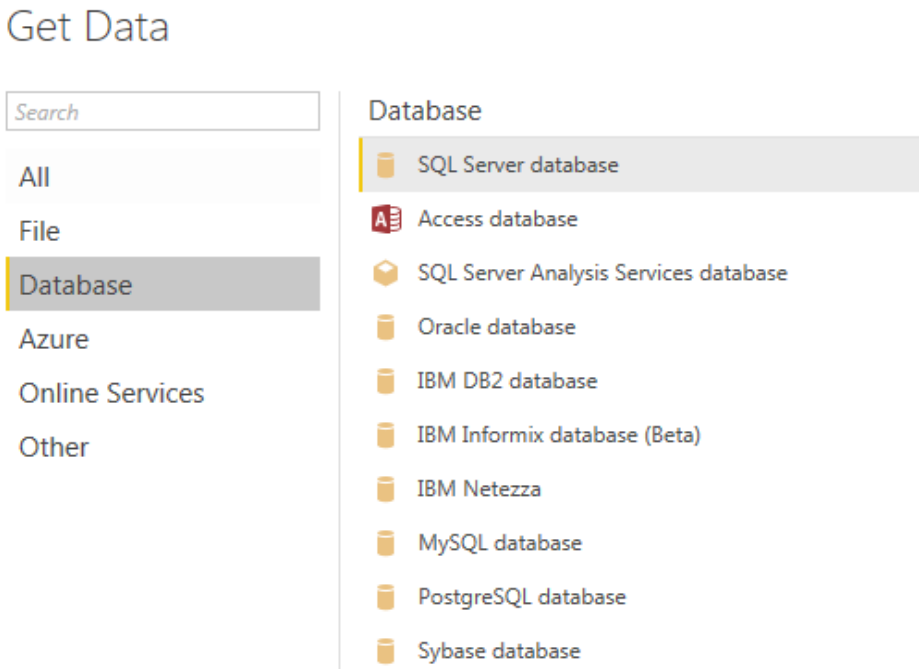
- The **Database** category provides the following data connections.
  - SQL Server Database
  - Access Database
  - SQL Server Analysis Services Database

Data Connectivity cont.…

- Oracle Database
- IBM DB2 Database
- IBM Informix database (Beta)
- IBM Netezza
- MySQL Database
- PostgreSQL Database
- Sybase Database
- Teradata Database
- SAP HANA Database
- SAP Business Warehouse Application Server
- SAP Business Warehouse Message Server (Beta)
- Amazon Redshift
- Impala
- Google BigQuery (Beta)

Data Connectivity cont.…

- The following Image shows the **Get Data** window for **Database**.

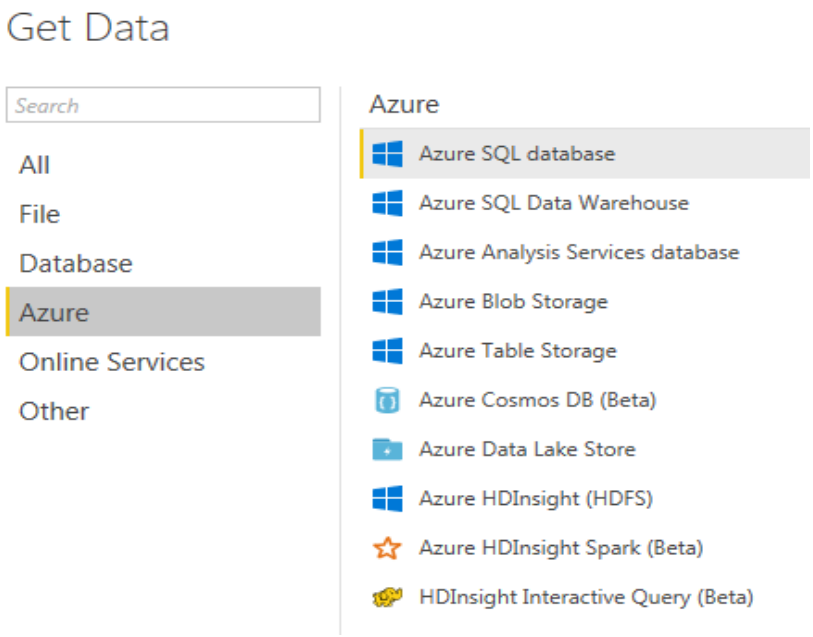


Data Connectivity cont.…

- The **Azure** category provides the following data connections.
  - Azure SQL Database
  - Azure SQL Data Warehouse
  - Azure Analysis Services database
  - Azure Blob Storage
  - Azure Table Storage
  - Azure Cosmos DB (Beta)
  - Azure Data Lake Store
  - Azure HDInsight (HDFS)
  - Azure HDInsight Spark (Beta)
  - HDInsight Interactive Query (Beta)

Data Connectivity cont.…

- The following image shows the **Get Data** window for **Azure**..

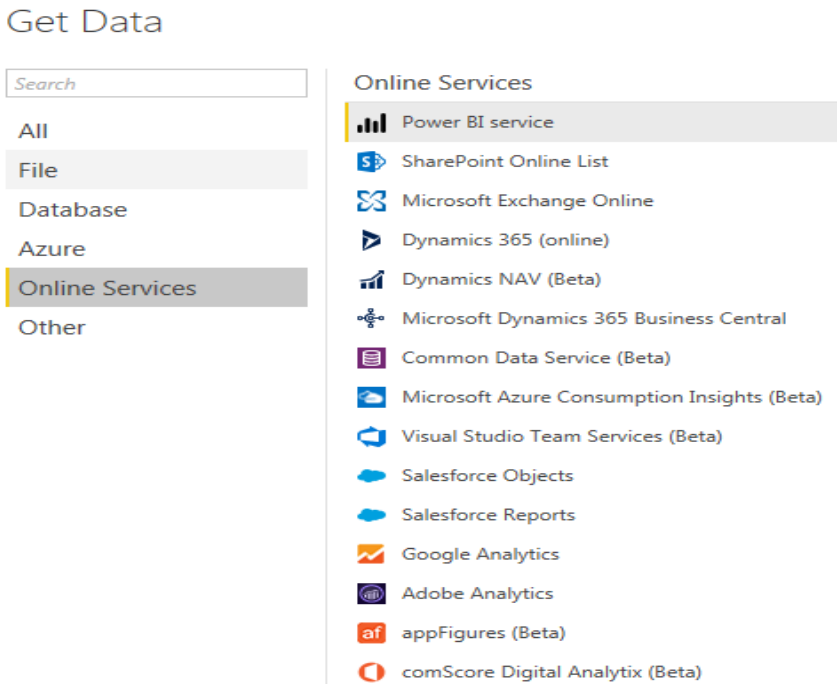


Data Connectivity cont.…

- The **Online Services** category provides the following data connections.
  - Power BI service
  - SharePoint Online List
  - Microsoft Exchange Online
  - Dynamics 365 (online)
  - Dynamics NAV (Beta)
  - Dynamics 365 for Financials (Beta)
  - Common Data Service (Beta)
  - Microsoft Azure Consumption Insights (Beta)
  - Visual Studio Team Services (Beta)
  - Salesforce Objects
  - Salesforce Reports
  - Google Analytics
  - Adobe Analytics
  - appFigures (Beta)

Data Connectivity cont.…

- The following image shows the **Get Data** window for **Online Services.**

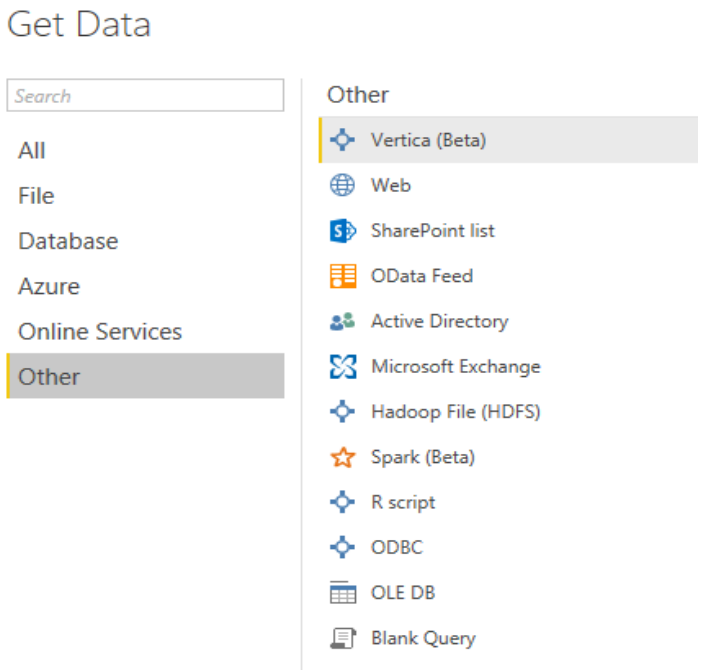


Data Connectivity cont.…

- The **Other** category provides the following data connections.
  - Vertica (Beta)
  - Kusto (Beta)
  - Web
  - SharePoint List
  - OData Feed
  - Active Directory
  - Microsoft Exchange
  - Hadoop File (HDFS)
  - Spark (Beta)
  - R Script
  - ODBC
  - OLE DB
  - Blank Query

Data Connectivity cont.…

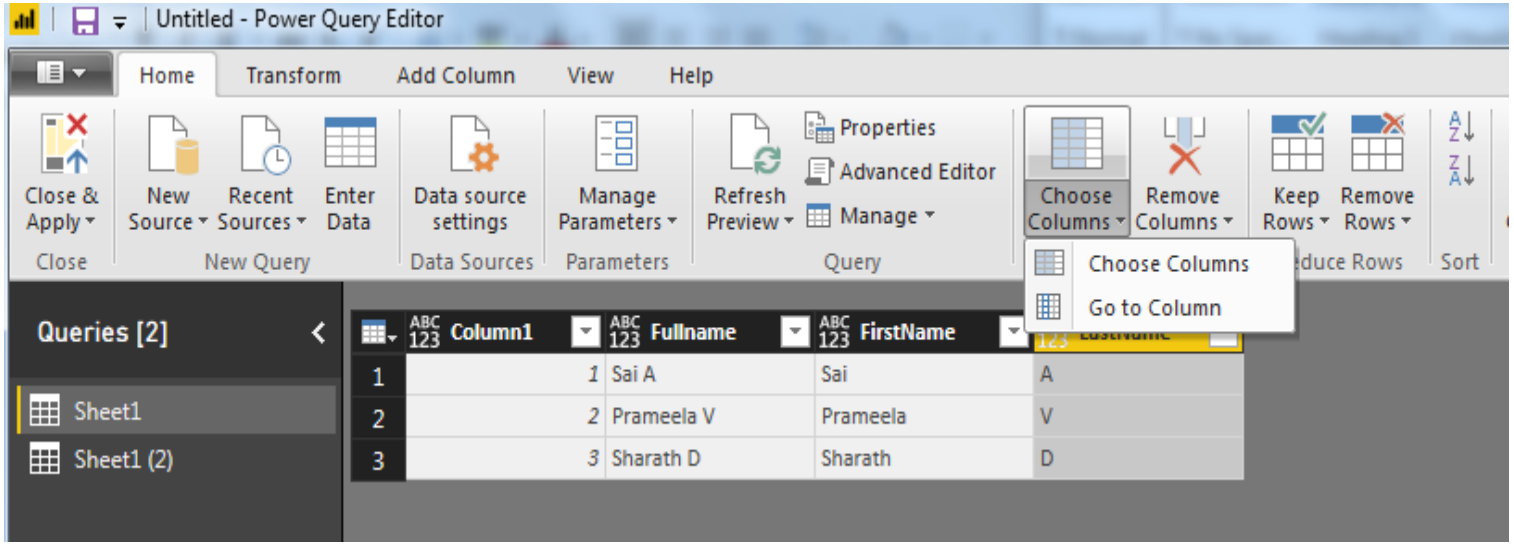
- The following image shows the **Get Data** window for **Other.**



Data Transformations:

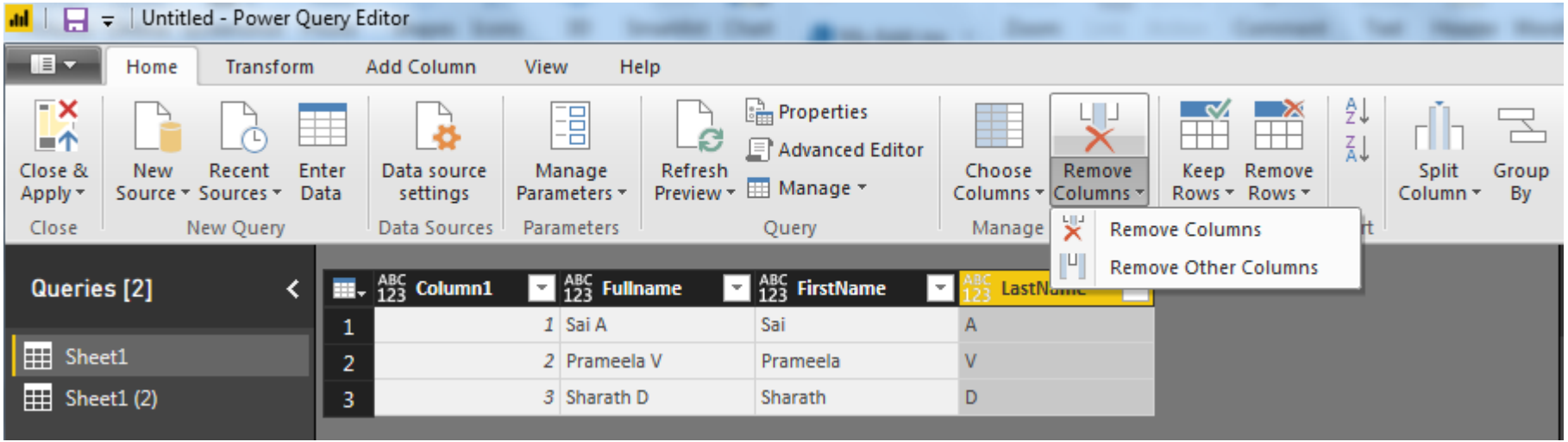
Choose Columns:

- It is used to keep the required columns in the report view.
- This can be used to select the necessary columns from the existing columns.



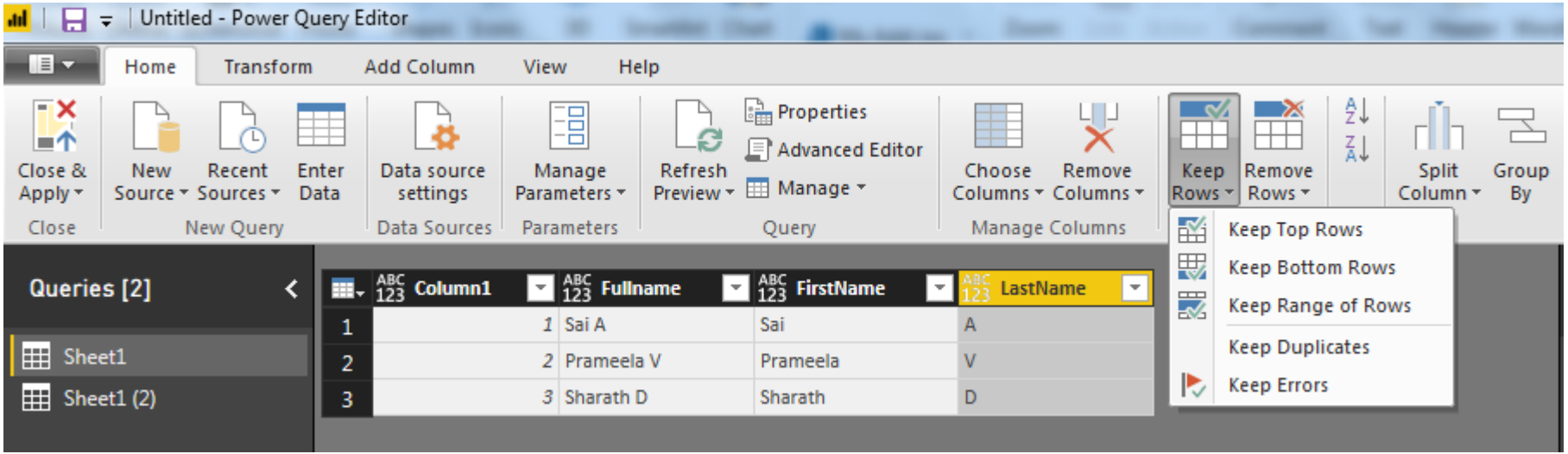
Remove Columns:

- It is used to remove the required columns in the report view.
- This can be used to remove the unnecessary columns from the existing columns.



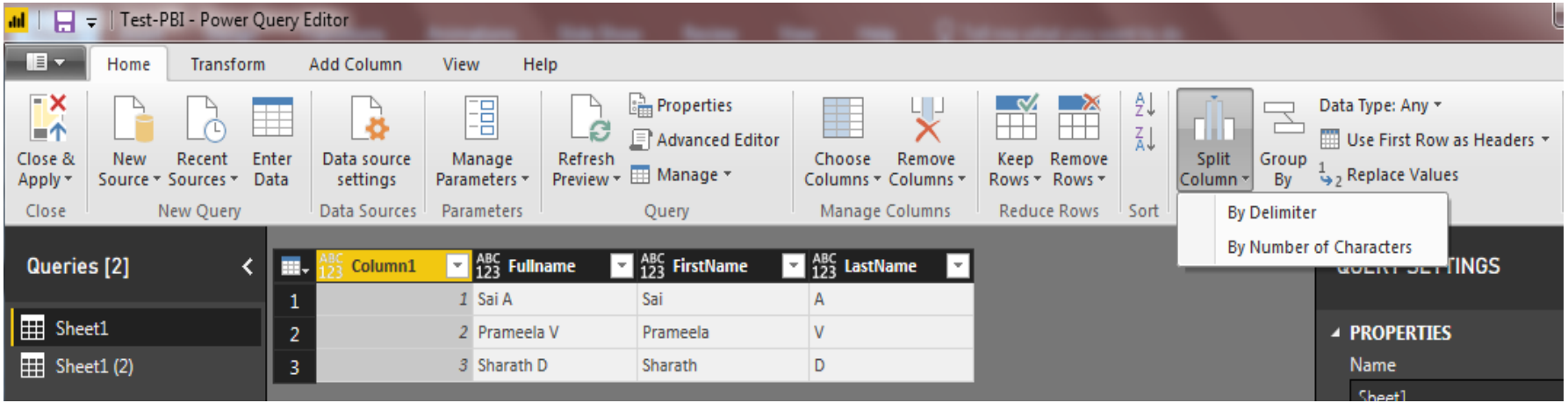
Keep Rows:

- It is used to keep the required rows in the report view.
- This can be used to keep the necessary columns from the existing columns.



Split Column:

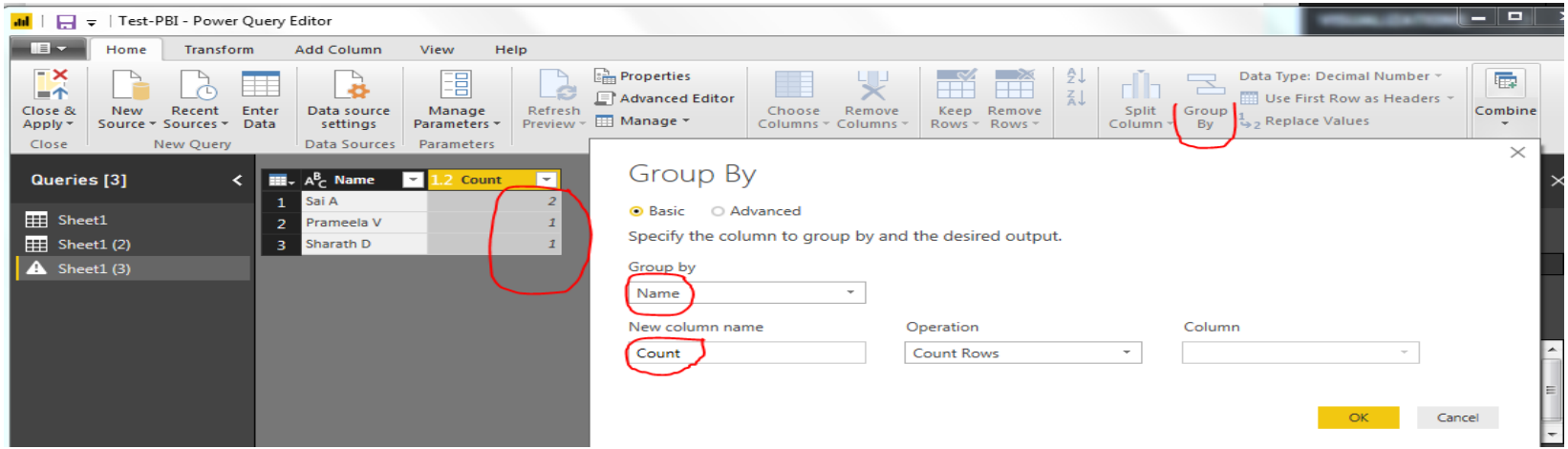
- It is used to split the column text value to different columns.
- A column of text can be split into multiple columns in two ways:
  - By delimiter
  - By a number of characters.



Group By:

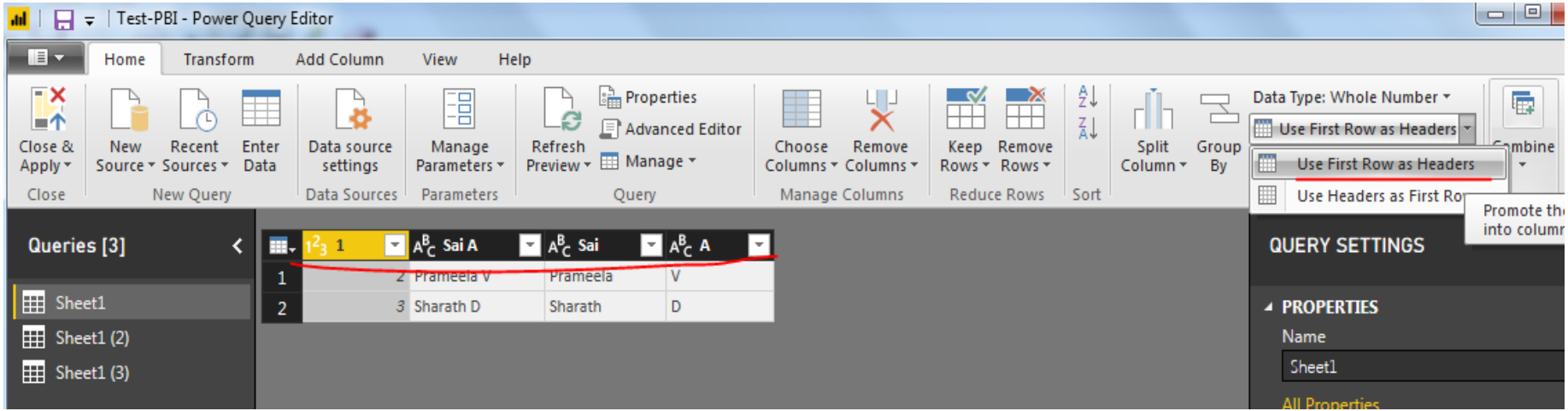
It is like our SQL Group By.

Group the values in multiple rows into a single value. This can be useful when summarizing the number of products offered, the total sales, or the count of students etc...



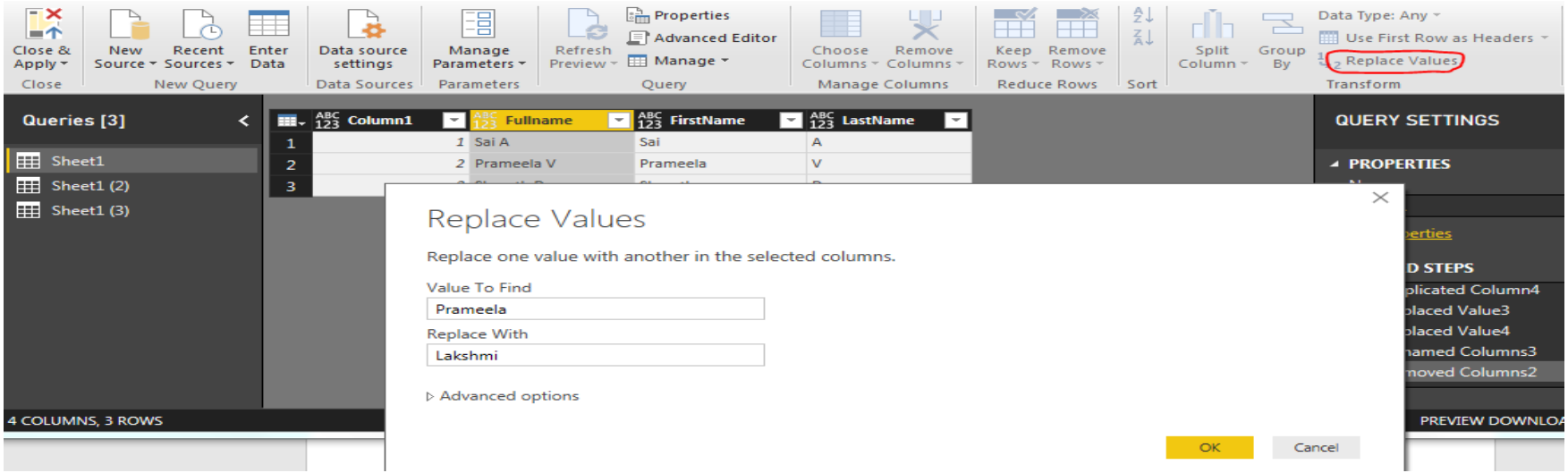
Use First Row as Headers:

- It is used to keep the First Row as Headers.
- When you want to make the first row as your header row, we need to use “Use First Row as Header” option



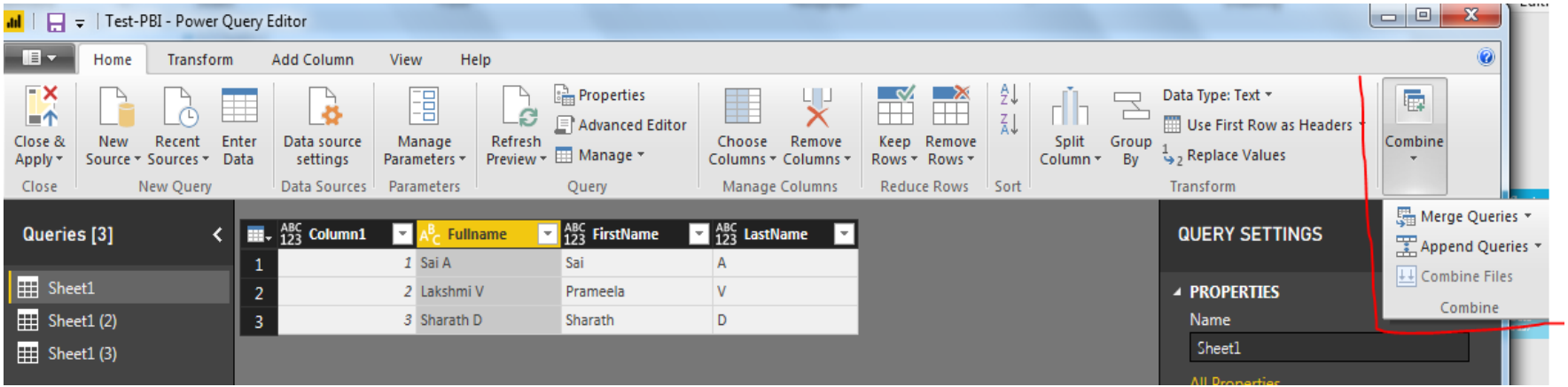
Replace Values:

- It is used to Replace one value with another value in the selected columns.
- This can be useful when you want to replace Null values with some default values like ‘0’ or ‘1’ etc..



Combine:

- There are two primary ways of combining queries – merging and appending.
- When you have one or more columns that you'd like to add to another query, you merge the queries. Merge is kind of join (in sql terms).
- When you have additional rows of data that you'd like to add to an existing query, you append the query.
- One of the important thing is Append Queries will NOT remove duplicates.
- Append only appends data below each other.



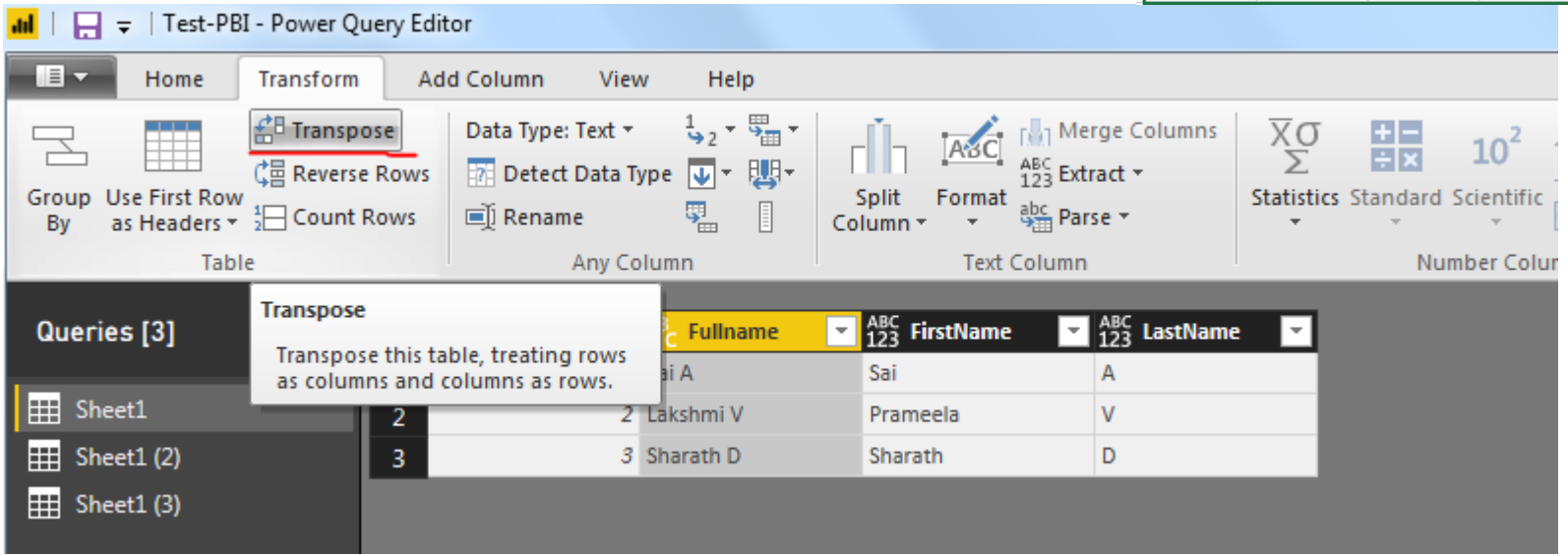
Transpose:

- Transpose the table, treating rows as columns and columns as rows.
- Sometimes you need to switch or rotate cells. We can do this by using the "Transpose".
- For example:

Jan	100
Feb	200
Mar	300
Apr	400

Transpose

Jan	Feb	Mar	Apr
100	200	300	400



Reverse Rows:

It is used to reverse the table's rows, hence the last rows are displayed first.

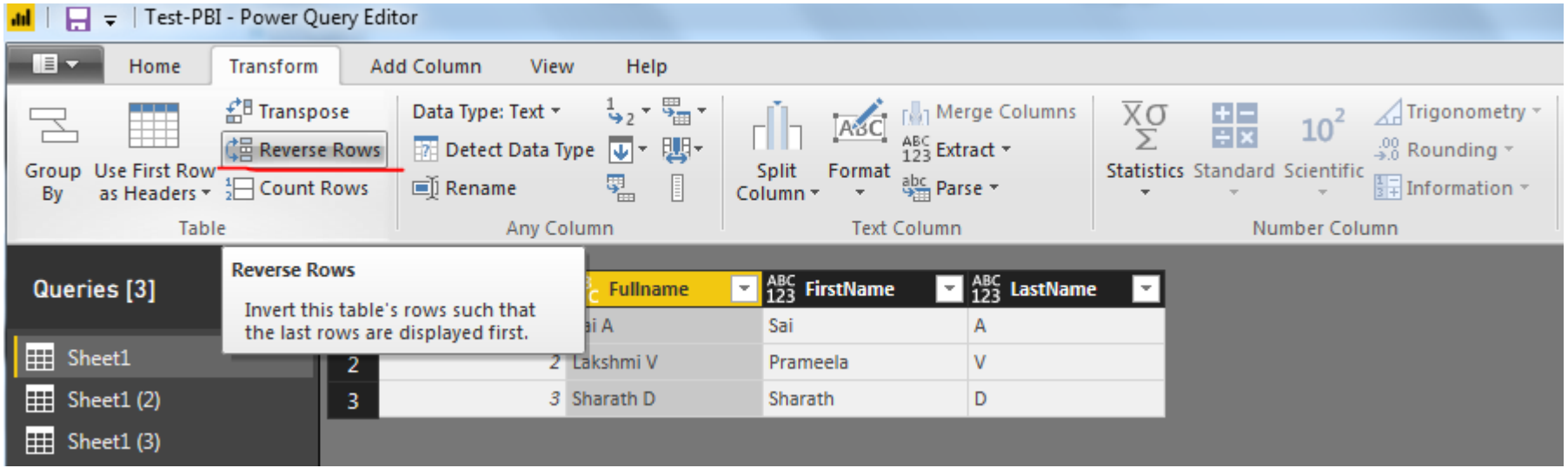
We can use this, when we want to display last rows as first.

For example:

Month	Qty
Jan	100
Feb	200
Mar	300

Reverse Rows

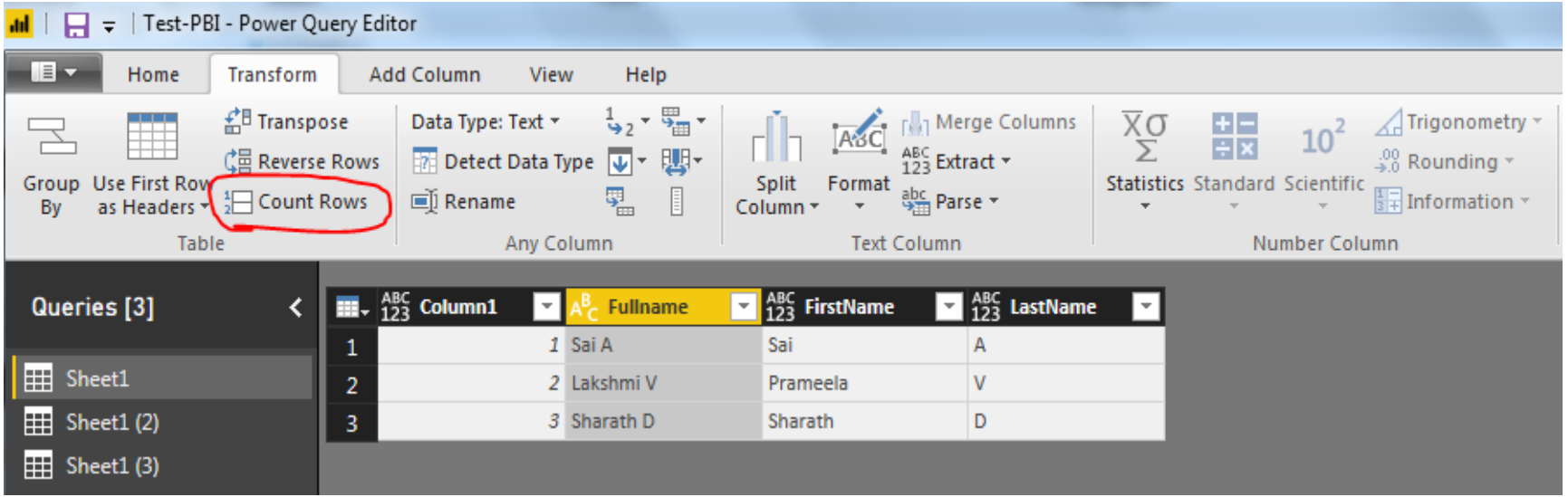
Month	Qty
Mar	300
Feb	200
Jan	100



Count Rows:

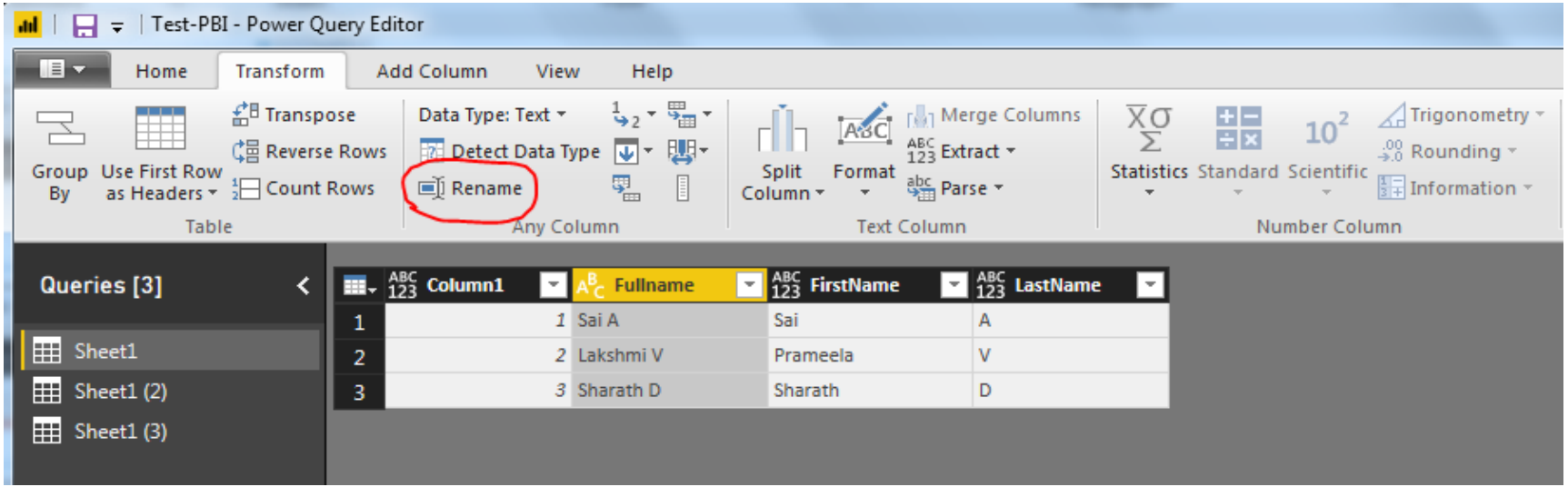
It is used to return the number of rows in the table.

For example: If you want to find out the Source file record count we can use this transformation.



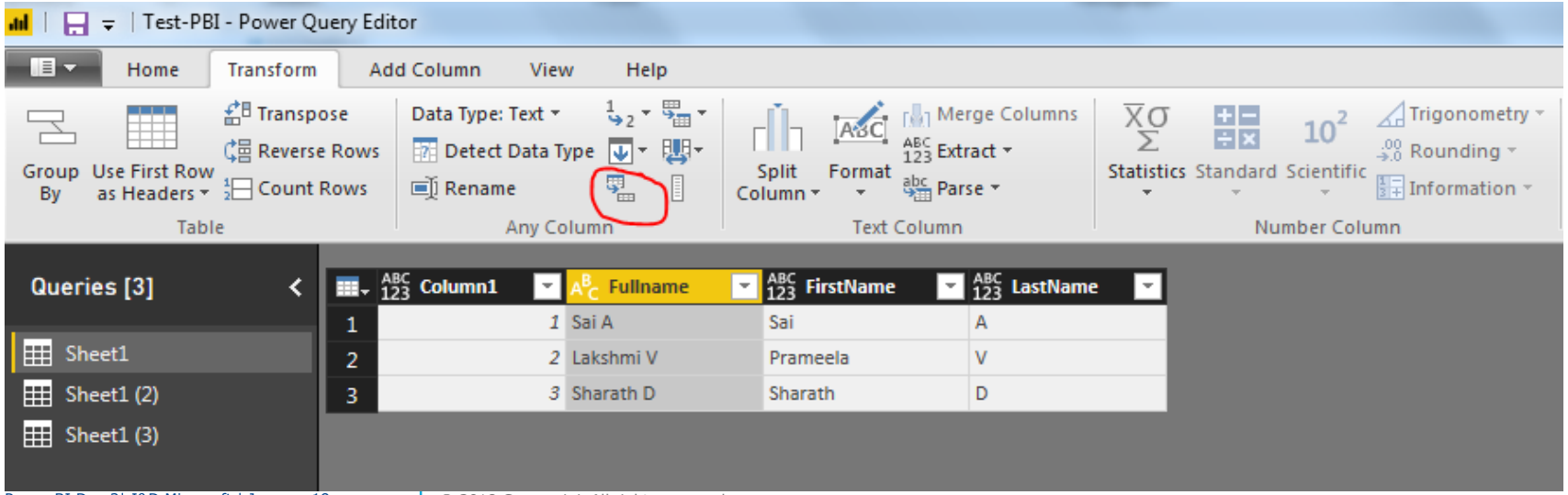
Rename:

- It is used to change the name of the currently selected column.
- There are a couple ways (at least) to rename a column. The following two approaches achieve the same result of changing the name of the column:
- In the **Query Editor**, **double-click** on a column, and enter the new name.
- In the **Query Editor**, you can also **right-click** on the column you want to rename, and select **Rename** from the menu that appears.



Pivot Column:

- With Power BI Desktop, you can pivot columns, and create a table that contains aggregated values for each unique value in a column.
- For example, if you need to know how many different products you have in each product category, you can quickly create a table the does precisely that.



Unpivot Columns:

Transform selected columns into attribute-value pairs where columns become rows.

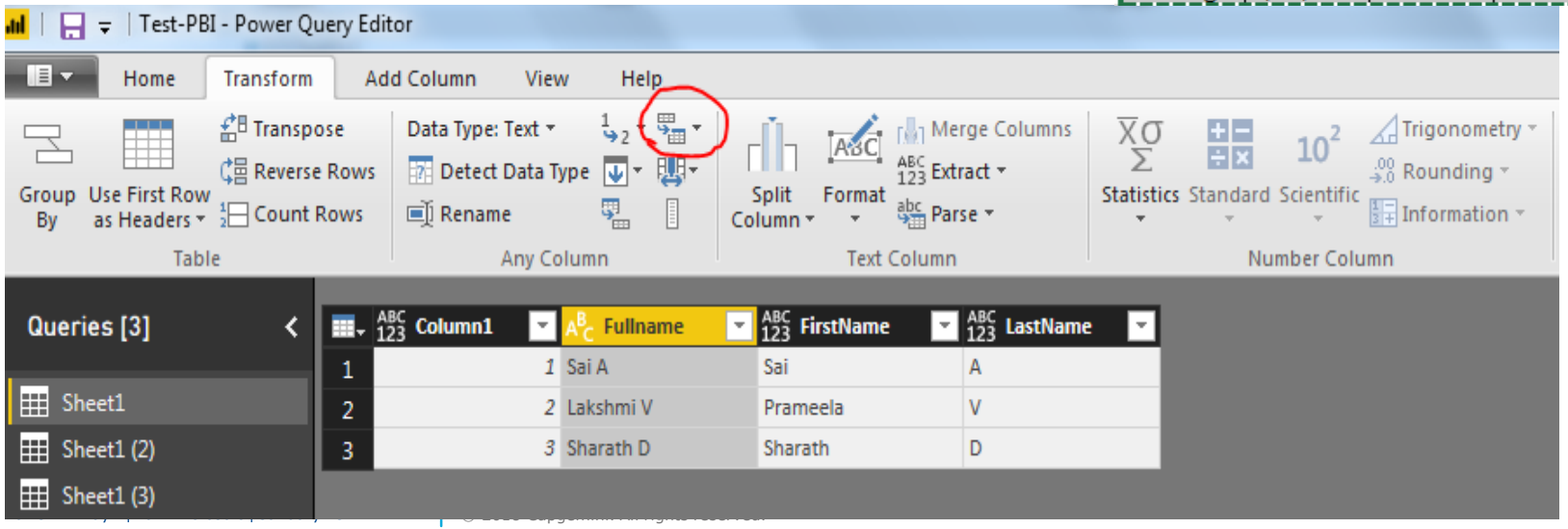
Unpivot Columns creates an **Attribute** column for each selected column heading and a **Value** column for each selected column cell value.

For example:

Category	1/31/2017	2/28/2017	3/31/2017
Beverages	969	399	300
Condimen	440	183	115

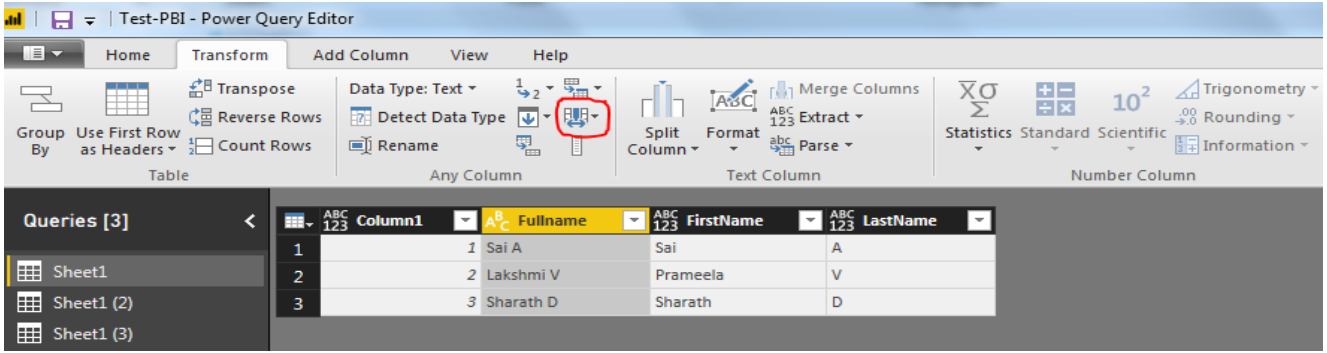
Category	Attribute	Value
Beverages	1/31/2017	969
Beverages	2/28/2017	399
Beverages	3/31/2017	300

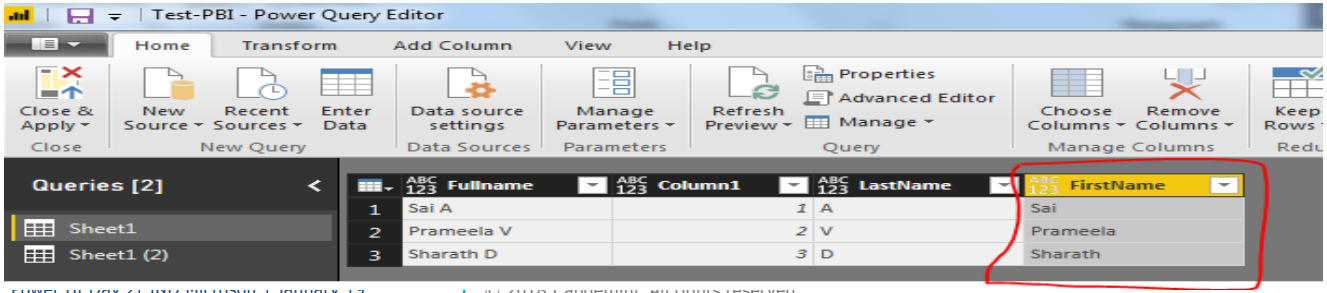
unpivot



Move:

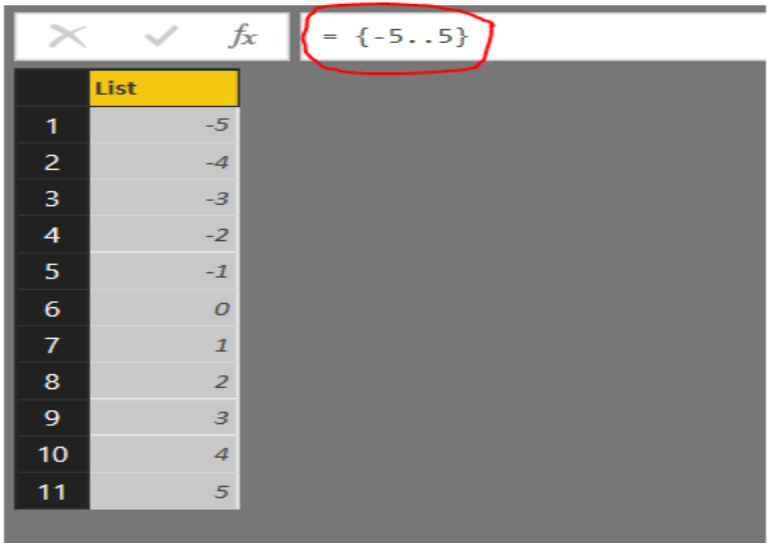
- Move selected columns to a different position in the table.
- We can adjust the columns by using this ‘Move’ transformation.
- By selecting the column we can adjust to left/right/To beginning/To End.
- For example: 2nd image shows “FirstName” column moved to Right.





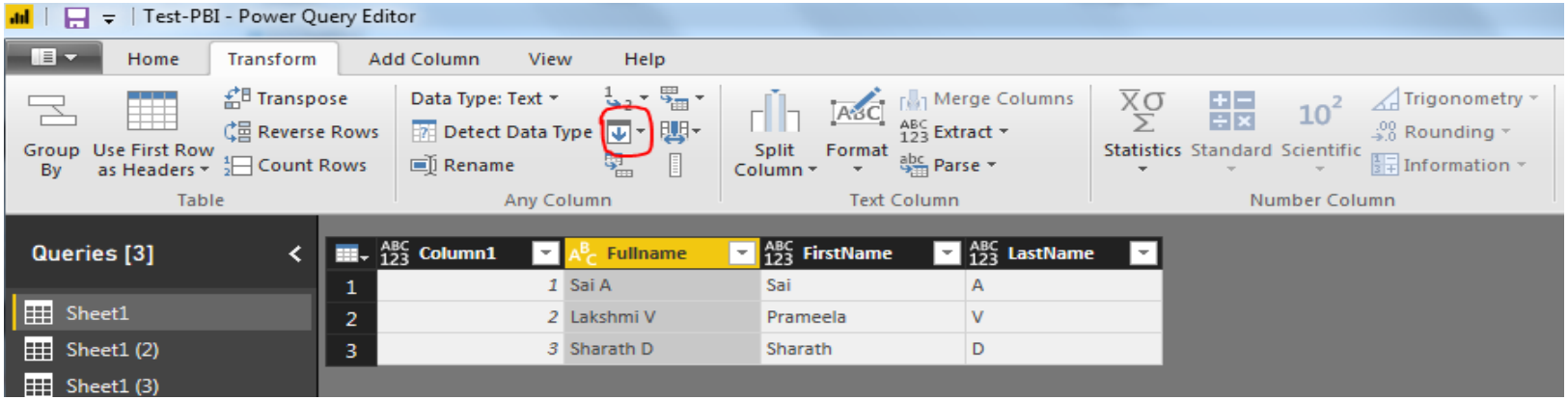
Convert to List:

- Converts the currently selected columns to a list.
- Disconnected lists of items can be useful for various purposes. For example, one can be used as a parameter table.
- We can also generate lists of characters and decimals, as well as non-continuous lists.



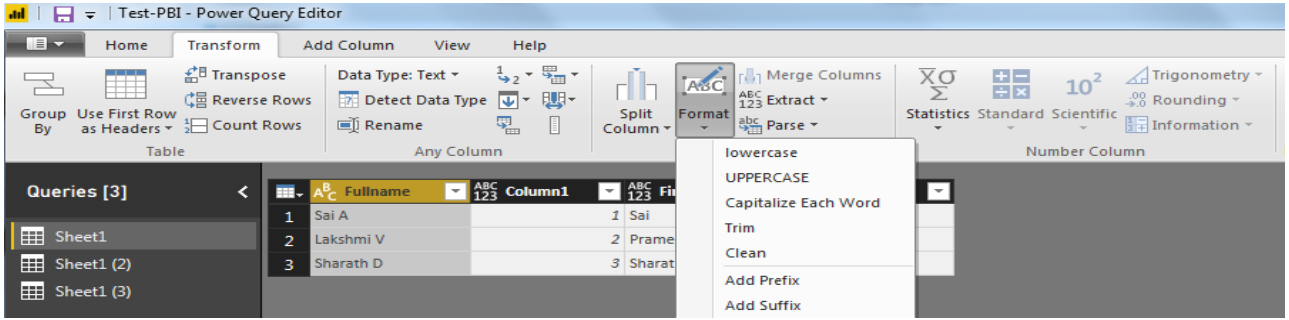
Fill:

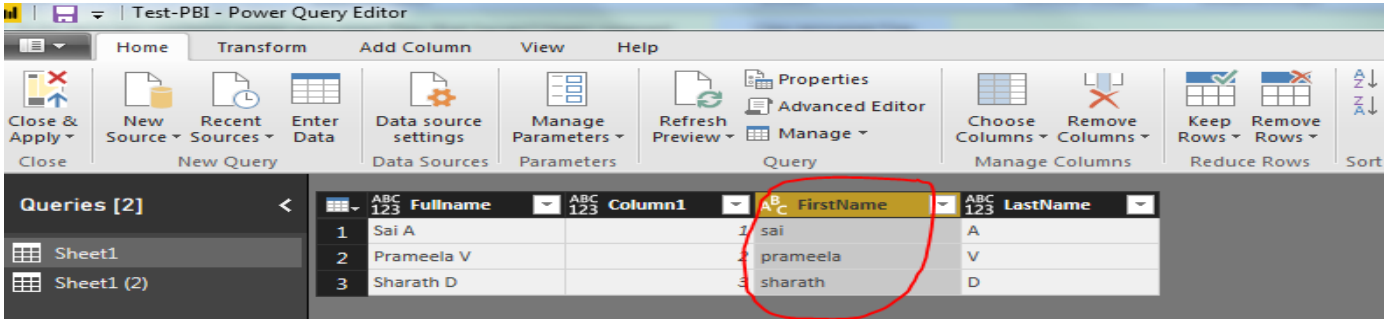
- Fill cell values to neighboring empty cells in the currently selected columns.
- This helps when your data contains cells or rows with blank (null) values, and you want to copy down the value from the cell above.
- The values will be filled down to replace the null values with the value above.
- The Fill function will NOT replace any existing values in the column. Once it hits a cell with a value, it will then look for the next blank cell and fill it with the value above.



Format:

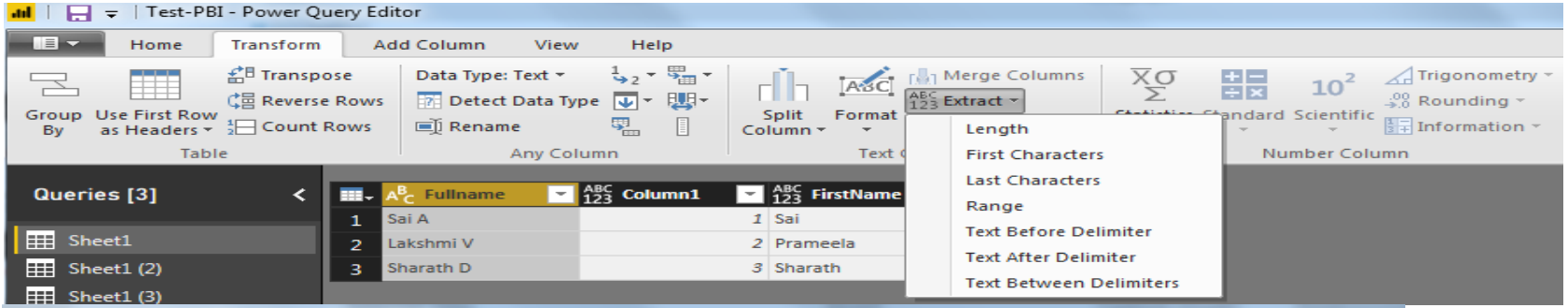
- Change casing of text or cleanse text.
- This can be used, when we want to change the particular column values to either lowercase / uppercase / trim etc.…
- For example: 2nd image shows the lowercase of Firstname column.

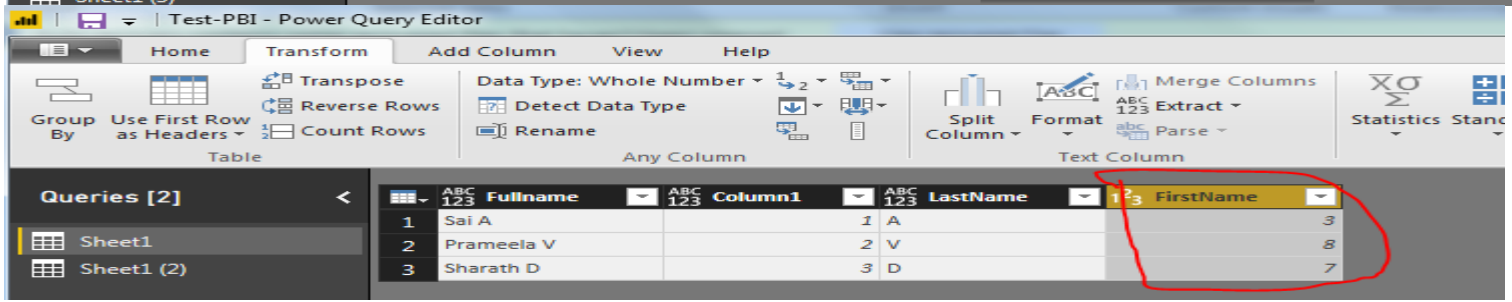




Extract:

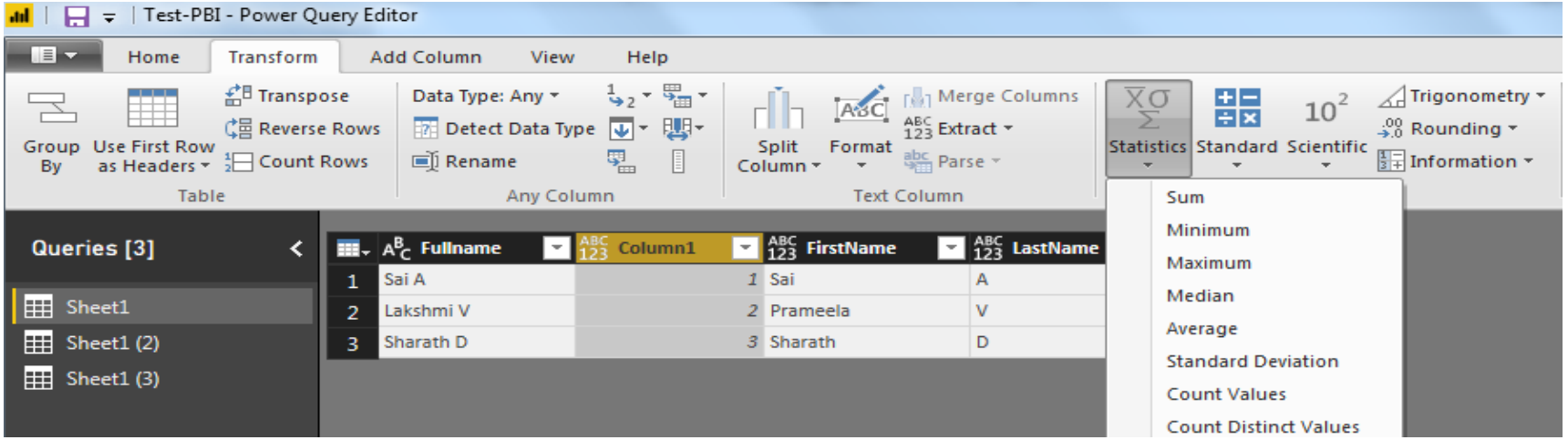
- It is used to extract characters from text values.
- Using this transformation we can find out the length of each row in a column. For example: 2nd image shows the length of "FirstName:
- Can extract range of characters in a selected column of a row.
- Can extract Frist character/Last character etc.…





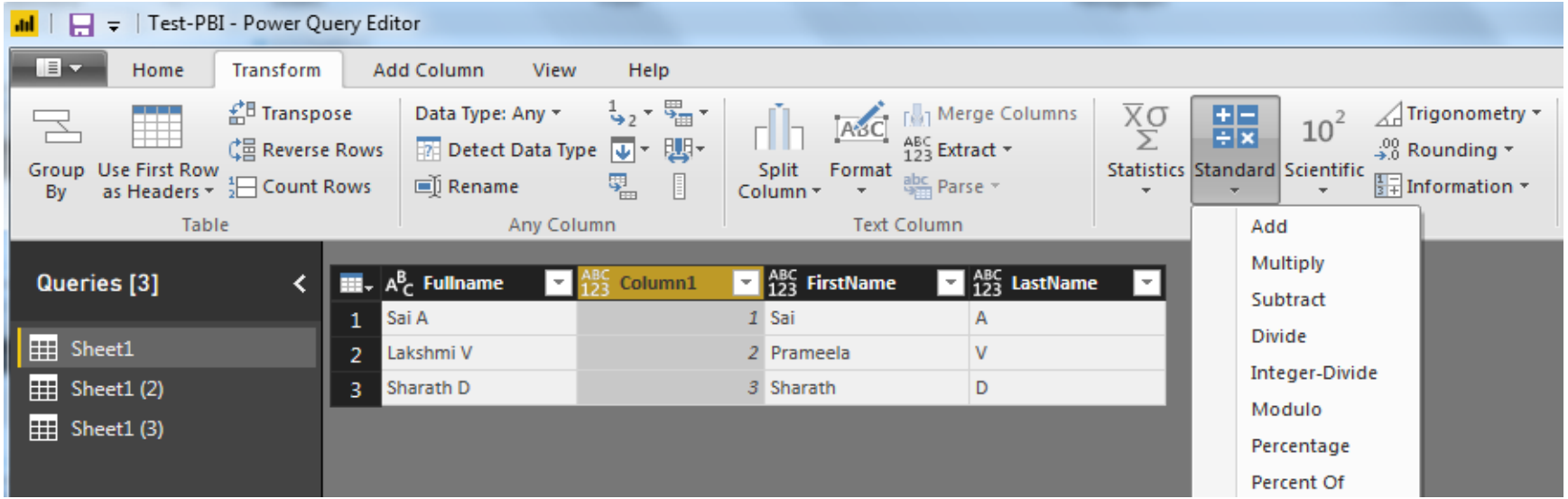
Statistics:

- It is used to perform statistical operations.
- Sum – It returns the sum of all the values in the currently selected column.
- Minimum - It returns the minimum of all the values in the currently selected column.
- Maximum - It returns the minimum of all the values in the currently selected column.



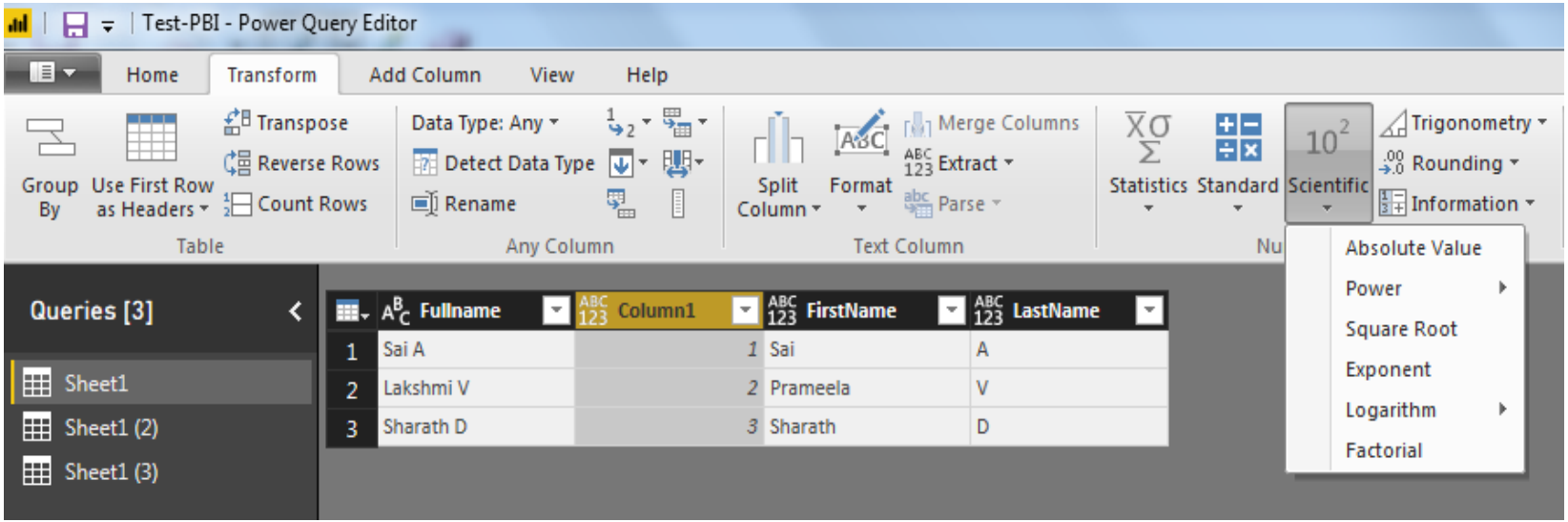
Standard:

- It is used to perform basic math operations.
- Useful to find out the % of profit / deriving threshold values etc.. by using the basic Math operations.



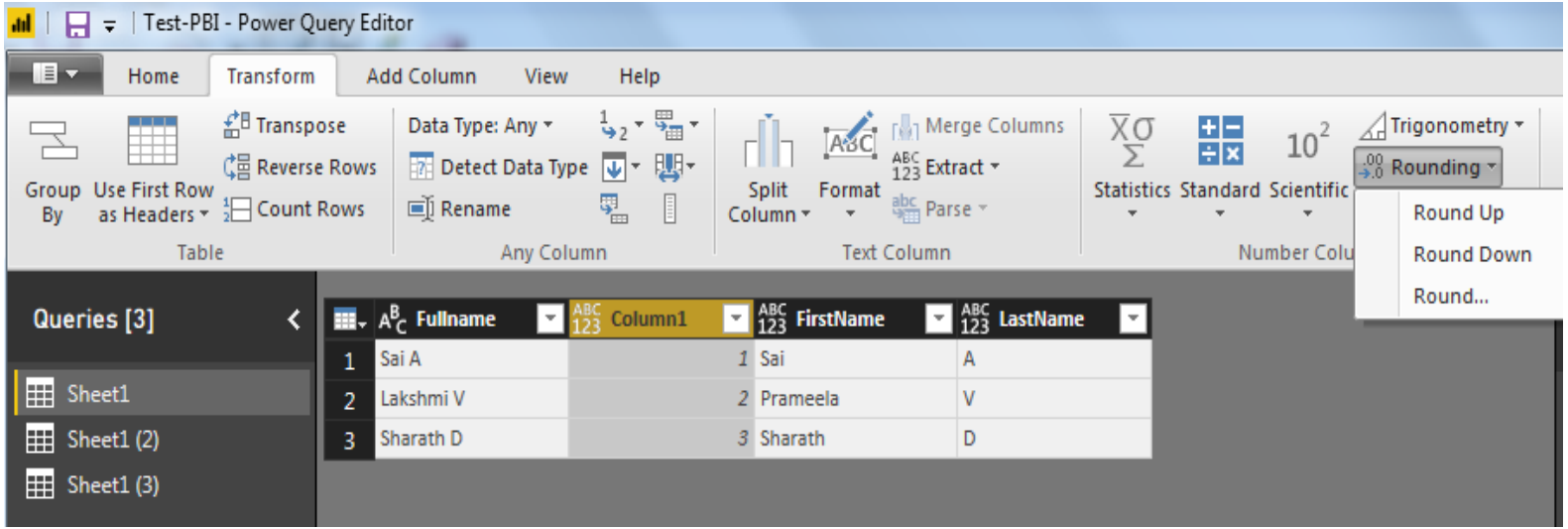
Scientific:

- It is used to perform scientific operations.
- This can be used to find out the Absolute Value / Square Root / Factorial etc.. of a selected column.



Rounding:

- It is used to perform rounding on number values.
- This can be used to rounding up the values to a nearest values.



Information:

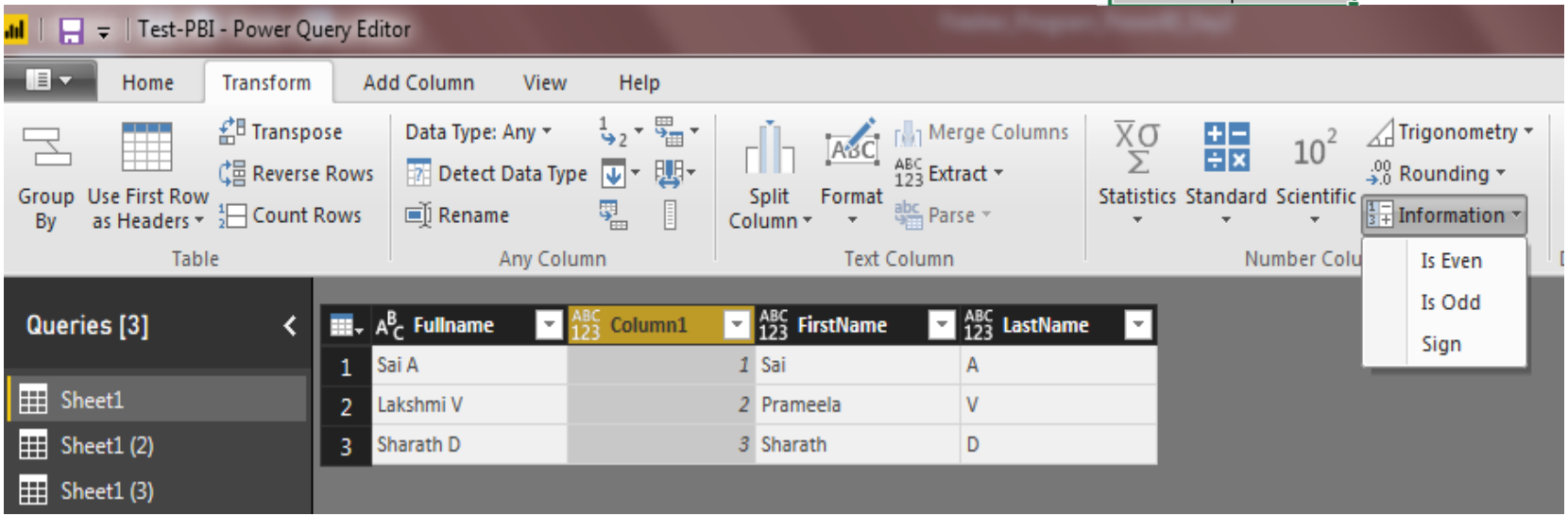
Extract information about a column.

Using this transformation, we can find out each columns associated row value information

For example:

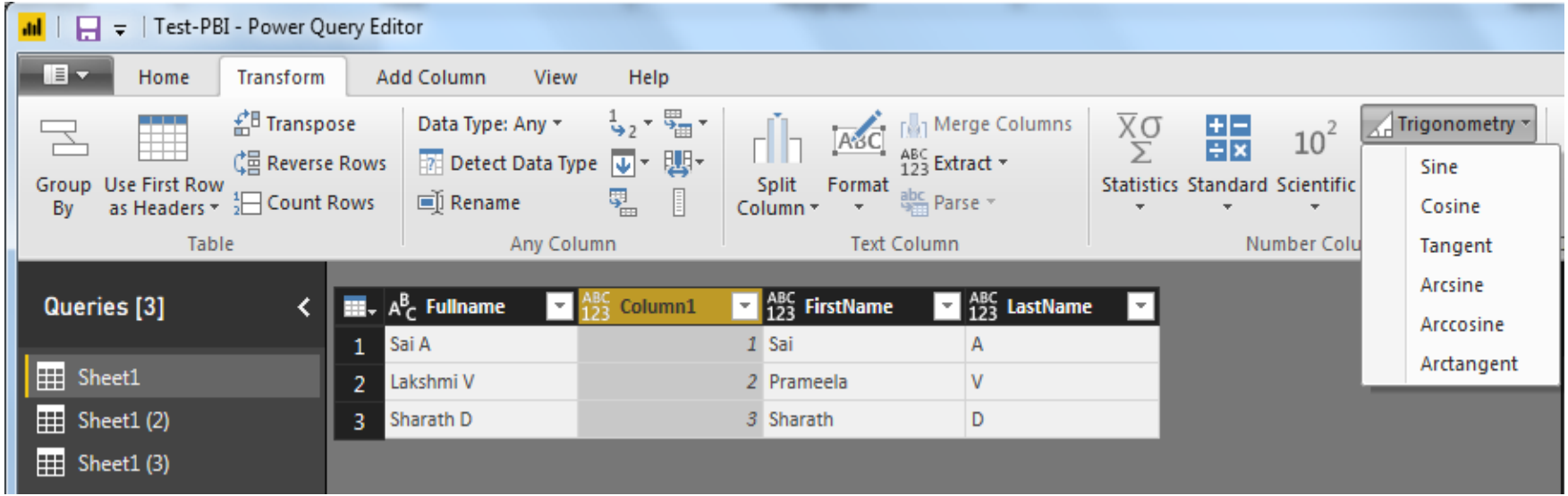
Month	Qty
Jan	100
Feb	200
Mar	300

Month	Qty
Jan	TRUE
Feb	TRUE
Mar	TRUE



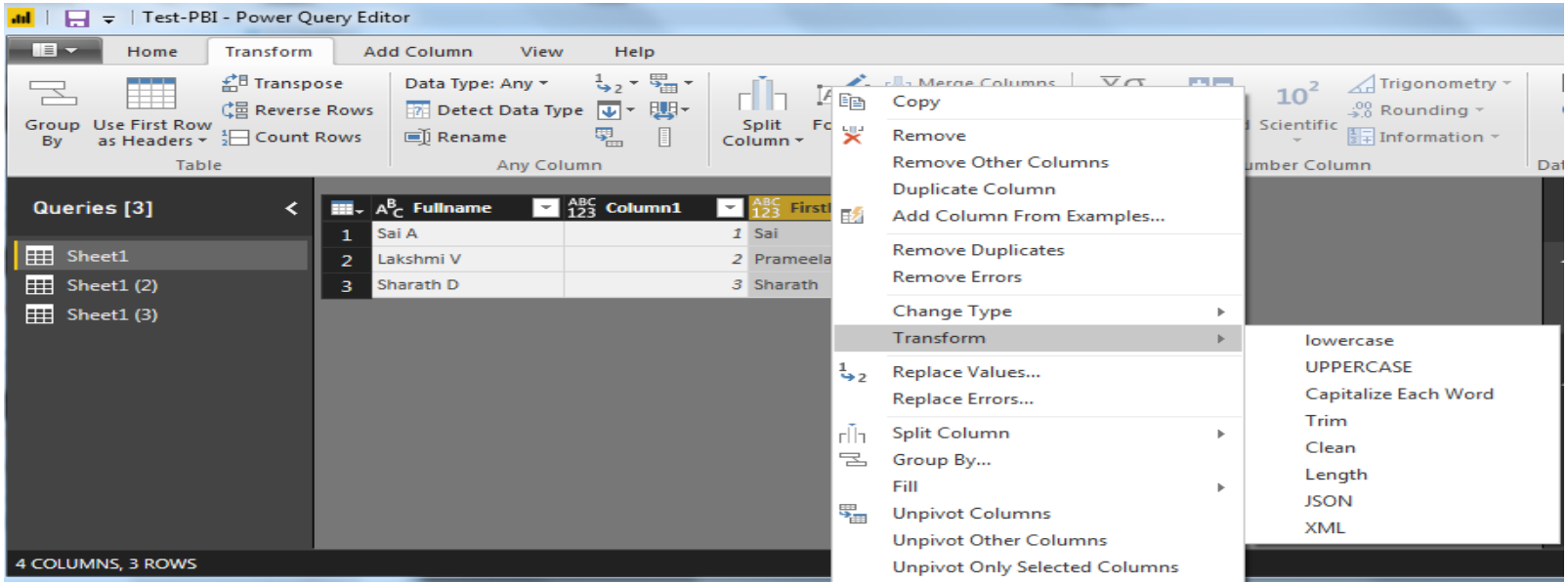
Trigonometry:

- It is used to perform trigonometric operations.
- If your data contains angular values ,you can find the sine, cosine values of it.
- It is helpful for mathematical data.



Transform:

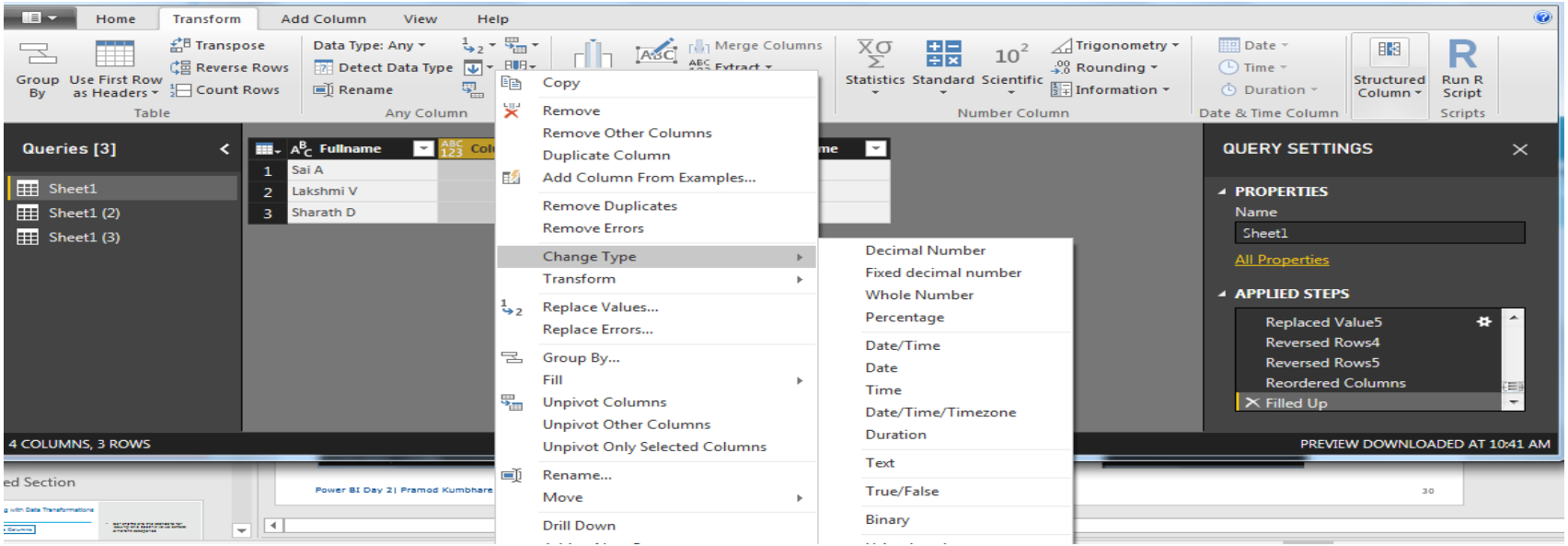
- Change casing of text or transform to selected format.
- This can be used, when we want to change the particular column values to either lowercase / uppercase / trim etc.…



Change Type:

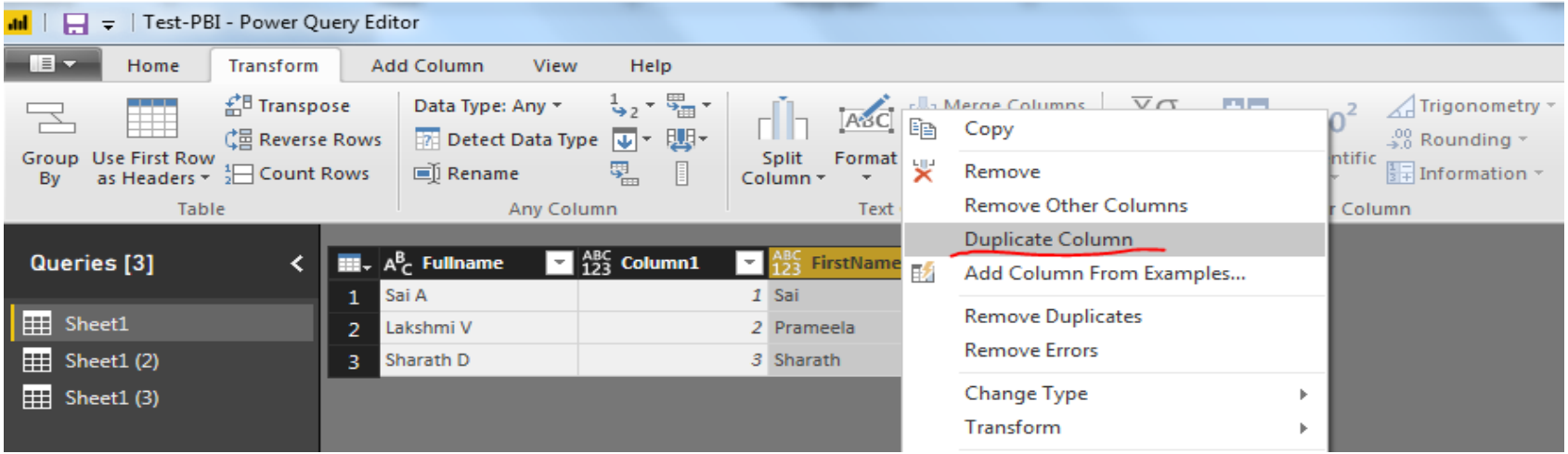
- It is used to change the data type of the selected column.
- We can use this transformation when we want to change the data type of a particular columns.

For example: We can change the numeric to Decimal/ixed Decimal etc..



Duplicate Column:

- It is used to duplicate the selected column.
- When we want to perform some transformations on the existing column, we can use the ‘Duplicate Column’.



AGENDA (A Look again)

Data Connectivity

- Files
- Database
- Azure
- Online Services
- Other

Data Transformations

- Choose Columns
- Remove Columns
- Keep Rows
- Remove Rows
- Split Column
- Group By
- Use First Row as Headers
- Replace Values
- Combine
- Transpose
- Reverse Rows
- Count Rows
- Rename

Contd.. Data Transformations

- Pivot Column
- Unpivot Columns
- Move
- Convert to List
- Fill
- Transform
- Change Type
- Duplicate Column
- Format (Lowercase, Uppercase, Capitalize Each Word, Add Prefix etc...)
- Extract (First Characters, Last Characters, Length, Range etc...)
- Statistics (Sum, Minimum, Maximum, Median, Average etc...)

- Standard (Add, Subtract, Multiply, Divide etc...)
- Scientific (Power, Square Root, Absolute Value etc…)
- Rounding (Round Up, Round Down)
- Information (Is Even, Is Odd and Sign)
- Trigonometry (Sine, Cosine, Tangent etc…)

Questions, Notes, Open Forum, Etc…

Thank You!

Capgemini