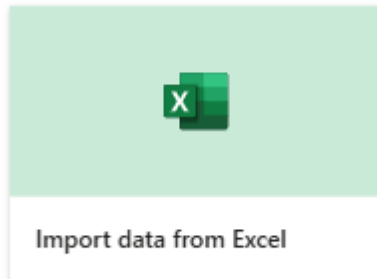


Section 1: Data Loading and Initial Cleanup

1. Load the dataset into Power BI.

- Import the dataset into Power BI Desktop using "Get Data".

The Excel file was imported using the "Get Data" option in Power BI Desktop. This made the dataset available for transformation and modeling



2. Open Power Query Editor.

- From the "Home" tab, click "Transform Data" to enter the Power Query Editor.

By selecting "Transform Data" from the Home tab, the dataset was opened in Power Query Editor, where all data cleaning and shaping operations were performed.

A screenshot of the Power Query Editor interface. The top ribbon shows the 'Home' tab with various options like 'Close & Apply', 'New Source', 'Recent Sources', 'Enter Data', 'Data source settings', 'Manage Parameters', 'Refresh Preview', 'Advanced Editor', 'Choose Columns', 'Remove Columns', 'Keep Rows', 'Remove Rows', 'Sort', 'Split Column', 'Group By', 'Data Type: Text', 'Use First Row as Headers', 'Replace Values', 'Merge Queries', 'Append Queries', and 'Combine Queries'. The main area displays a table with 7 columns: Product_ID, Sale_Date, Sales_Rep, Region, Sales_Amount, Quantity_Sold, and Product_Cate. The table contains 20 rows of data. The 'Sales_Rep' column has values like Bob, David, Charlie, Eve, and Alice. The 'Product_Cate' column has values like Furniture, Food, Clothing, and Electronics.

Product_ID	Sale_Date	Sales_Rep	Region	Sales_Amount	Quantity_Sold	Product_Cate
1052	2023-02-03	Bob	North	5053.97	18	Furniture
1093	2023-04-21	Bob	North	4384.02	17	Furniture
1015	2023-09-21	David	South	4631.23	30	Food
1072	2023-08-24	Bob	South	2167.94	39	Clothing
1061	2023-03-24	Charlie	East	3750.2	13	Electronics
1021	2023-02-11	Charlie	West	3761.15	32	Food
1083	2023-04-11	Bob	West	618.31	29	Furniture
1087	2023-01-06	Eve	South	7698.92	46	Furniture
1075	2023-06-29	David	South	4223.39	30	Furniture
1075	2023-10-09	Charlie	West	8239.58	18	Clothing
1088	2023-11-16	Eve	North	8518.45	13	Furniture
1100	2023-08-14	Bob	West	2198.74	43	Food
1024	2023-11-11	Eve	West	6607.8	21	Food
1003	2023-12-31	Alice	South	4775.59	30	Furniture
1022	2023-08-17	Charlie	South	8813.55	21	Food
1053	2023-10-16	Bob	North	2235.83	48	Furniture
1002	2023-05-30	David	North	6810.35	17	Furniture
1088	2023-10-04	Bob	East	6116.75	40	Electronics
1030	2023-07-17	David	West	3023.48	19	Clothing

Section 2: Data Cleaning Tasks

3. Remove unnecessary columns.

- Remove the Region_and_Sales_Rep column.

The column Region_and_Sales_Rep was removed because it was redundant or irrelevant to the analysis goals.

Column14	Removed Columns
Region_and_Sales_Rep	
North-Bob	
West-Bob	
South-David	
South-Bob	
East-Charlie	

4. Rename columns for better readability.

- Rename the following:
 - Product_ID → Product ID
 - Sale_Date → Sale Date
 - Sales_Rep → Sales Representative
 - Product_Category → Product Category
 - Payment_Method → Payment Method
 - Sales_Channel → Sales Channel
 - Customer_Type → Customer Type

Several columns were renamed to be more readable and user-friendly, aligning with reporting standards and easier interpretation. For example, Sale_Date became Sale Date, and Sales_Rep became Sales Representative.

Product ID	Sale Date	Sales Representative	Region	Sales Amount	Quantity Sold	Product Category
1052	2/3/2023	Bob	North	5053.97	18	Furniture
1093	4/21/2023	Bob	West	4384.02	17	Furniture
1015	9/21/2023	David	South	4631.23	30	Food

Renamed Columns

5. Change column order to improve logical flow.

- Reorder the columns to:
 - Sale Date, Product ID, Region, Sales Representative, Product Category, Quantity Sold, Unit Price, Unit Cost, Sales Amount, Discount, Customer Type, Payment Method, Sales Channel

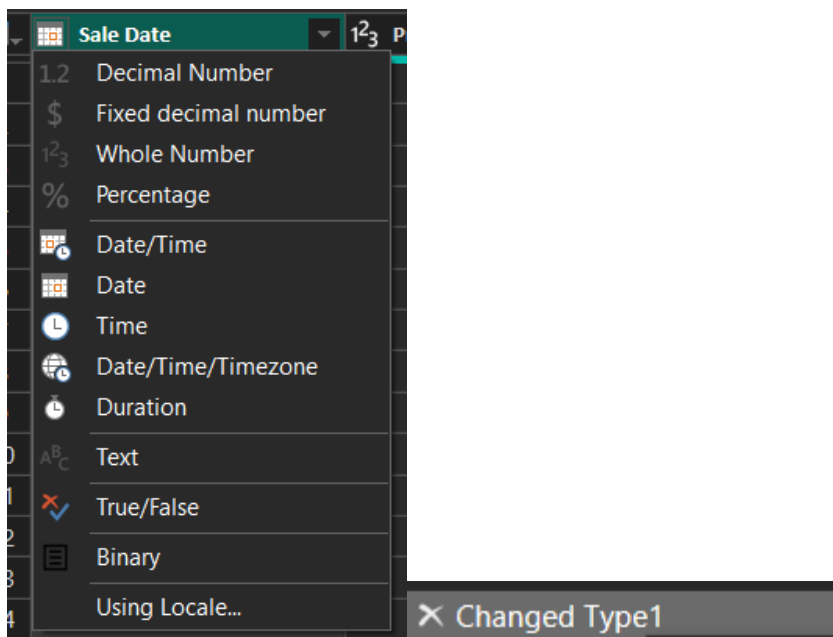
Columns were reordered to follow a logical sequence, starting with Sale Date and ending with Sales Channel, improving navigation and consistency during analysis.

= Table.ReorderColumns("#Renamed Columns",{"Sale Date", "Product ID", "Region", "Sales Representative", "Product Category", "Quantity Sold", "Unit Price", "Unit Cost"})							
Sale Date	Product ID	Region	Sales Representative	Product Category	Quantity Sold	Unit Price	Unit Cost
2/3/2023	1052	North	Bob	Furniture	18		



6. Correct data types.

- Ensure appropriate data types:
 - Sale Date: Date
 - Quantity Sold: Whole Number
 - Sales Amount, Unit Cost, Unit Price, Discount: Decimal Number
 - Categorical columns: Text



7. Format currency fields.

- Format Sales Amount, Unit Price, Unit Cost as currency.

Unit Price	Unit Cost	Sales Amount
₹ 3,007.47	₹ 2,565.30	₹ 5,238.42
₹ 4,186.98	₹ 3,851.45	₹ 7,751.92
₹ 4,061.04	₹ 3,741.30	₹ 7,724.57
₹ 4,361.70	₹ 4,138.41	₹ 8,090.84
₹ 2,147.14	₹ 1,833.95	₹ 9,631.41
₹ 3,108.30	₹ 2,873.53	₹ 6,773.89
₹ 1,955.56	₹ 1,531.20	₹ 7,952.11
₹ 4,784.28	₹ 4,646.55	₹ 5,842.88
₹ 1,709.71	₹ 1,281.65	₹ 6,966.82
₹ 1,482.88	₹ 1,434.20	₹ 6,581.04

8. Remove rows with null or blank Sales Representative values.

- Filter and remove any rows where the Sales Representative field is empty or null.

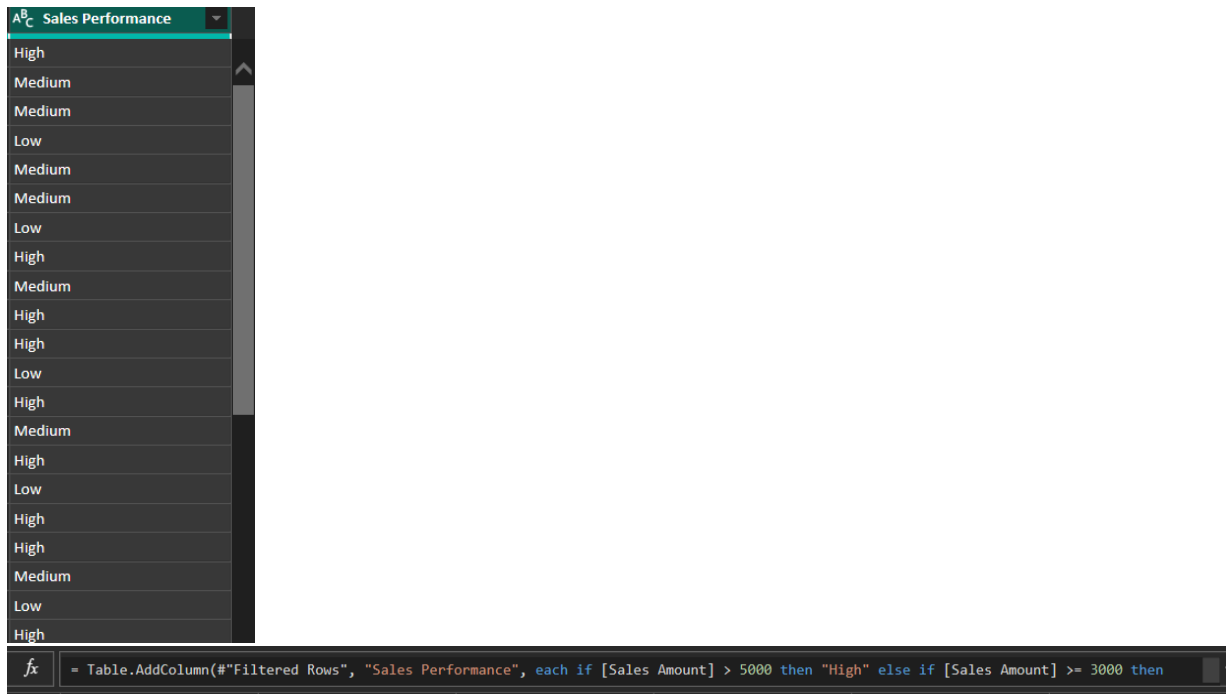
Here we don't have any null value so I'm not able to display it.

Sales Representative	Product Category	Quantity Sold	Unit Price
David			
David			
David			
David			
David			
David			
David			
David			
David			
David			
David			
David			
David			
David			
David			
David	Clothing	6	₹ 1,7
David	Electronics	45	₹ 4,
David	Electronics	42	₹ 5,

Section 3: Data Transformation

9. Add a new conditional column.

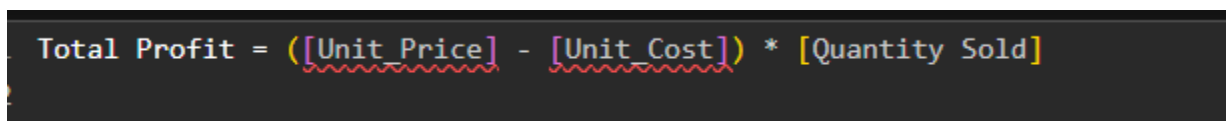
- Create a column Sales Performance:
 - "High" if Sales Amount > 5000
 - "Medium" if between 3000–5000
 - "Low" otherwise



The screenshot shows a data transformation interface. At the top, a dropdown menu is set to 'Sales Performance'. Below it, a list of values is displayed: High, Medium, Medium, Low, Medium, Medium, Low, High, Medium, High, High, Low, High, Medium, High, Low, High, High, Medium, Low, and High. At the bottom, a formula bar shows the logic for adding the column: `= Table.AddColumn("#Filtered Rows", "Sales Performance", each if [Sales Amount] > 5000 then "High" else if [Sales Amount] >= 3000 then`

10. Create a calculated column for Total Profit.

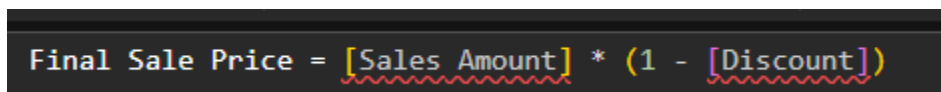
- Total Profit = (Unit Price - Unit Cost) * Quantity Sold



The screenshot shows a calculated column formula: `Total Profit = ([Unit Price] - [Unit Cost]) * [Quantity Sold]`. The fields [Unit Price], [Unit Cost], and [Quantity Sold] are highlighted with red wavy lines, indicating they are recognized as column names.

11. Create a column for Final Sale Price after discount.

- Formula: Final Sale Price = Sales Amount * (1 - Discount)

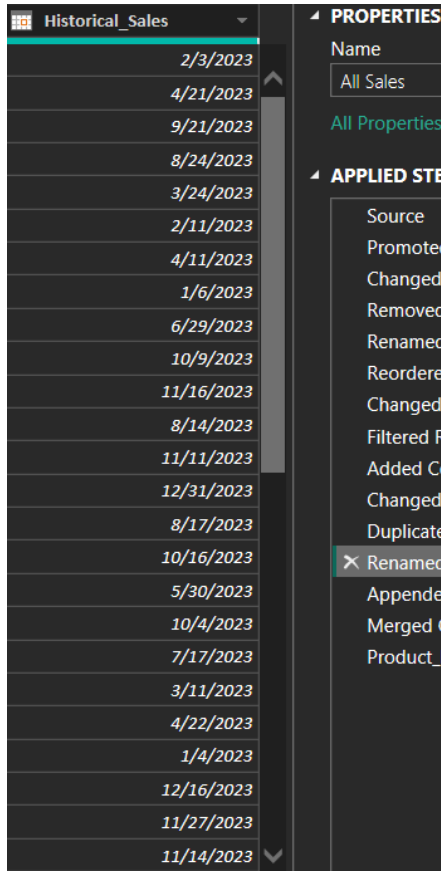


The screenshot shows a calculated column formula: `Final Sale Price = [Sales Amount] * (1 - [Discount])`. The fields [Sales Amount] and [Discount] are highlighted with red wavy lines, indicating they are recognized as column names.

Section 4: Appending & Merging

12. Duplicate the current query and name it Historical_Sales.

- Filter for records with Sale Date before 01-Jan-2023.

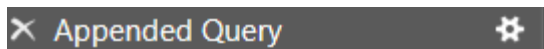


The screenshot shows a data table with the title 'Historical_Sales'. The table contains a single column of dates. The sidebar on the right has two sections: 'PROPERTIES' with a 'Name' field set to 'All Sales' and a link 'All Properties'; and 'APPLIED STE' (partially visible) with a list of actions including Source, Promote, Changed, Removed, Renamed, Reordered, Changed, Filtered, Added, Changed, Duplicate, and 'X Renamed' (which is highlighted with a red 'X').

Historical_Sales
2/3/2023
4/21/2023
9/21/2023
8/24/2023
3/24/2023
2/11/2023
4/11/2023
1/6/2023
6/29/2023
10/9/2023
11/16/2023
8/14/2023
11/11/2023
12/31/2023
8/17/2023
10/16/2023
5/30/2023
10/4/2023
7/17/2023
3/11/2023
4/22/2023
1/4/2023
12/16/2023
11/27/2023
11/14/2023

13. Append Historical_Sales to the original table.

- Create a new table All_Sales by appending both tables.



X Appended Query

Section 5: Final Steps

15. Apply all transformations.

- Click “Close & Apply” to load cleaned and transformed data back into Power BI for report building.

