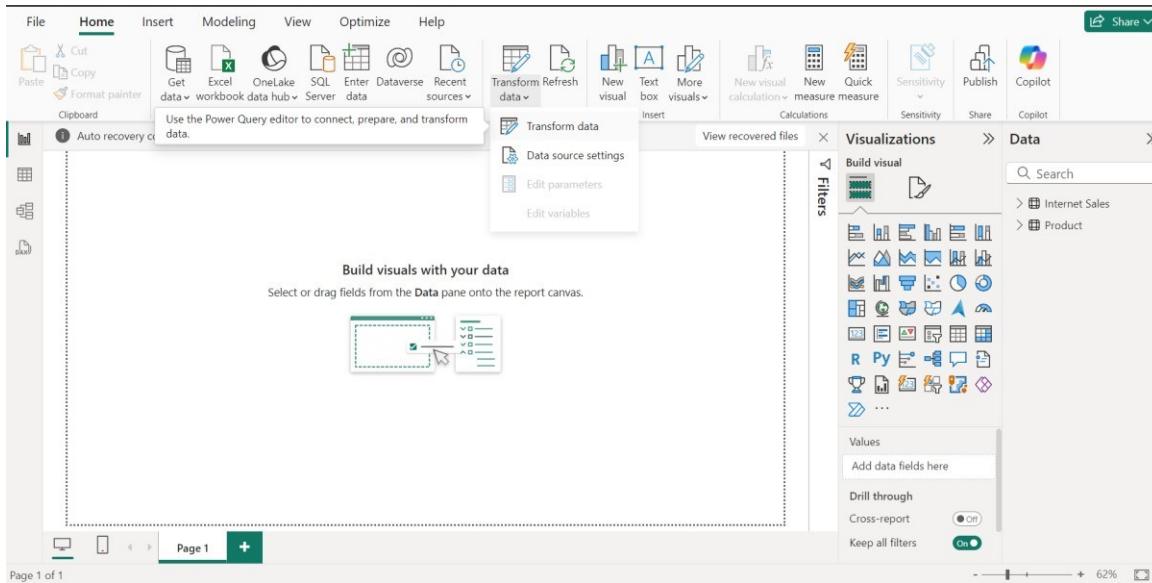


## Task 1: Understanding Power Query Editor (Data Cleaning & Transformation)

### 1. Import the dataset into Power BI using Power Query Editor.

Get data>load data>Transform data



### 2. Perform the following transformations:

- Remove missing values & duplicates.

## ○ After removing Duplicates

The screenshot shows the Power Query Editor interface. The main area displays a table with columns: ProductKey, OrderDateKey, DueDateKey, ShipDateKey, CustomerKey, and a final column with values ranging from 21 to 13. The 'APPLIED STEPS' pane on the right shows a step named 'Removed Duplicates'. The status bar at the bottom indicates 'PREVIEW DOWNLOADED AT 11:07 AM'.

## ○ Remove null values

The screenshot shows the Power Query Editor interface. The main area displays a table with columns: ProductKey, OrderDateKey, DueDateKey, ShipDateKey, CustomerKey, and a final column with values ranging from 19 to 1. The 'APPLIED STEPS' pane on the right shows a step named 'Filtered Rows1'. The status bar at the bottom indicates 'PREVIEW DOWNLOADED AT 20:25'.

## ○ Change data types

The screenshot shows the Microsoft Power BI Data Editor interface. A context menu is open over a column named 'CustomerKey' in a table named 'Internet Sales'. The 'Data Type' option is selected, opening a dropdown menu with various data type options like Decimal Number, Fixed decimal number, Whole Number, Percentage, Date, Time, etc. The 'Text' option is currently highlighted. The main workspace displays the 'Internet Sales' table with columns: eDateKey, ShipDateKey, CustomerKey, and PromotionKey. The 'CustomerKey' column contains values like 21768, 28389, 25863, etc. The bottom right corner of the screen shows a status bar with 'PREVIEW DOWNLOADED AT 18:34' and system icons.

## ○ After Changing Datatype

This screenshot shows the same Microsoft Power BI Data Editor interface after changing the data type of the 'CustomerKey' column. The 'Change Data Type' dialog is no longer visible, and the 'CustomerKey' column is now displayed as a percentage value, such as 2176800.00%, 2838900.00%, etc. The rest of the table and the interface remain the same, with the 'Internet Sales' table and its columns visible in the main workspace. The status bar at the bottom right shows 'PREVIEW DOWNLOADED AT 18:44'.

### 3. Split full names into first and last names.

Product

EnglishProductName

FrenchProductName

StandardCost

APPLIED STEPS

Changed Type

- Delimiter-space > Left most delimiter

Product

EnglishProductName.1

EnglishProductName.2

SpanishProductName

FrenchProductName

APPLIED STEPS

Changed Type

#### 4. Merge tables based on keys (e.g., Orders & Customers).

Screenshot of Power BI Desktop showing the 'Merge Queries' dialog open over a query preview.

**Merge Queries** dialog:

- Source: Product SubCategory
- Promoted Headers: None
- Changed Type: None
- Merged Queries: Product Category1

Query Settings:

- Name: Product SubCategory
- All Properties

Applied Steps:

- Source: Product SubCategory
- Promoted Headers: None
- Changed Type: None
- Merged Queries: Product Category1

Preview:

	EnglishProductSubcategoryName	SpanishProductSubcategoryName	FrenchProductSubcategoryName	ProductCategoryKey
1	Mountain Bikes	Bicicleta de montaña	VTT	1
2	Road Bikes	Bicicleta de carretera	Vélo de route	2
3	Touring Bikes	Bicicleta de paseo	Vélo de randonnée	3
4	Handlebars	Barra	Barre d'appui	4

Column profiling based on top 1000 rows.

Screenshot of Power BI Desktop showing the 'Merged Queries' dialog open over a query preview.

**Merged Queries** dialog:

- Source: Product Category
- Promoted Headers: None
- Changed Type: None
- Merged Queries: Product Category1

Query Settings:

- Name: Product SubCategory
- All Properties

Applied Steps:

- Source: Product SubCategory
- Promoted Headers: None
- Changed Type: None
- Merged Queries: Product Category1

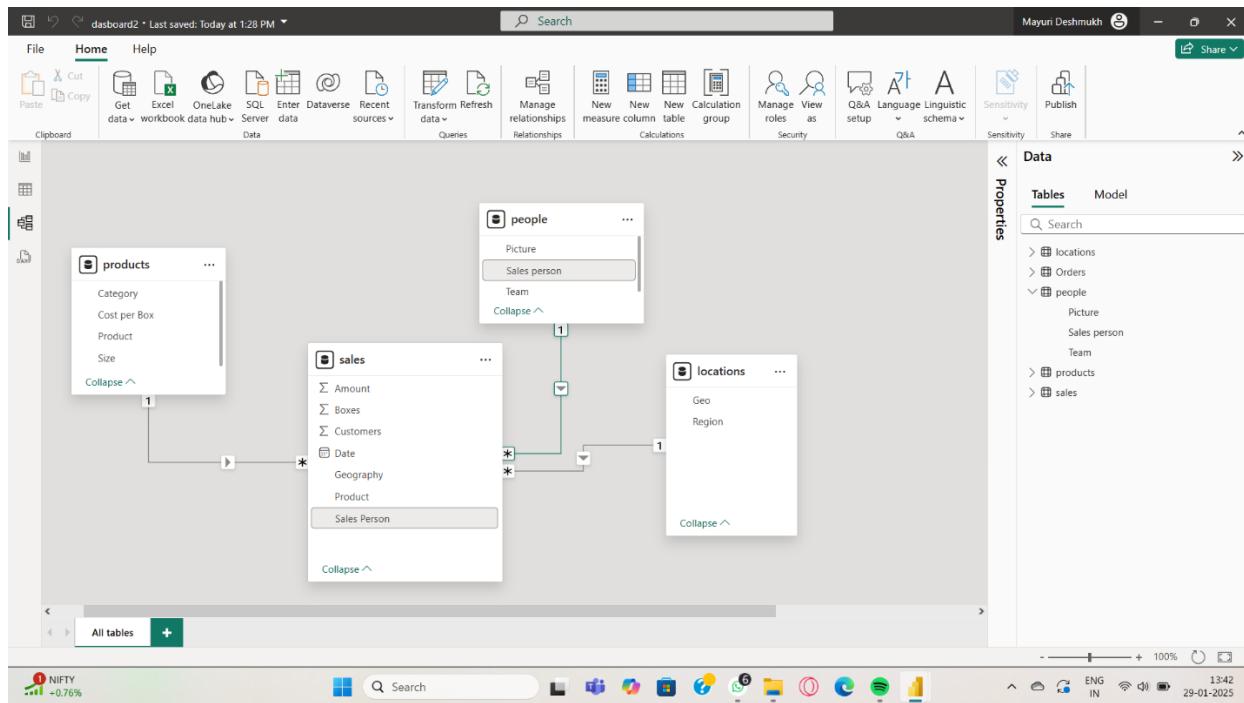
Preview:

	ProductCategoryKey	ProductCategoryAlternateKey	EnglishProductCategoryName	SpanishProductCategoryName	FrenchProductCategoryName
1	1	Bikes	Bicicleta	Vélo	

Column profiling based on top 1000 rows.

## Task 2: Understanding Data & Data Modeling

1. Identify key tables: Fact (Sales, Orders) and Dimension (Customers, Products, Regions).



## Task 3: Implementing DAX Functions A) Text Functions Create calculated columns using the following Text DAX Functions:

### 1.Extract the first 3 letters of Product Category:

The screenshot shows the Power BI desktop interface with a table named 'Product'. A new calculated column 'Left' is being created using the formula `LEFT([EnglishProductName],3)`. The Data pane on the right shows various product categories like Adjustable Race, Bearing Ball, HeadsetBall Bearings, etc.

EnglishProductName	Left
Adjustable Race	Adj
Bearing Ball	Bear
HeadsetBall Bearings	Head
Crown Race	Cro
Decal 1	Dec
Decal 2	Dec
Flat Washer 1	Flat
Flat Washer 6	Flat
Flat Washer 2	Flat
Flat Washer 9	Flat
Flat Washer 4	Flat
Flat Washer 3	Flat
Flat Washer 8	Flat
Flat Washer 5	Flat
Flat Washer 7	Flat
Thin-Jam Hex Nut 9	Thin
Thin-Jam Hex Nut 10	Thin
Thin-Jam Hex Nut 1	Thin
Thin-Jam Hex Nut 2	Thin
Thin-Jam Hex Nut 15	Thin
Thin-Jam Hex Nut 16	Thin
Thin-Jam Hex Nut 5	Thin
Thin-Jam Hex Nut 6	Thin
Thin-Jam Hex Nut 3	Thin
Thin-Jam Hex Nut 4	Thin
Thin-Jam Hex Nut 13	Thin
Thin-Jam Hex Nut 14	Thin

### 2.Create a full name column from First & Last Names:

The screenshot shows the Power BI desktop interface with a table named 'Product'. A new calculated column 'Concatenate' is being created using the formula `CONCATENATE('Product'[EnglishProductName.1],'Product'[EnglishProductName.2])`. The Data pane on the right shows various product names.

EnglishProductName.1	EnglishProductName.2	Concatenate
Race	Adj	AdjustableRace
Bearing	Ball	BearingBall
Headset	Ball	HeadsetBall
Crown	Race	CrownRace
Decal	1	Decal1
Flat	Washer	FlatWasher
Washer	6	FlatWasher6
Washer	2	FlatWasher2
Washer	9	FlatWasher9
Washer	4	FlatWasher4
Washer	3	FlatWasher3
Washer	8	FlatWasher8
Washer	5	FlatWasher5
Washer	7	FlatWasher7
Pulley	Guide	GuidePulley
HL	Grip Tape	HLGrip Tape
Thin-Jam	Hex Nut 9	Thin-JamHex Nut 9
Thin-Jam	Hex Nut 10	Thin-JamHex Nut 10
Thin-Jam	Hex Nut 1	Thin-JamHex Nut 1
Thin-Jam	Hex Nut 2	Thin-JamHex Nut 2
Thin-Jam	Hex Nut 15	Thin-JamHex Nut 15
Thin-Jam	Hex Nut 16	Thin-JamHex Nut 16
Thin-Jam	Hex Nut 5	Thin-JamHex Nut 5
Thin-Jam	Hex Nut 6	Thin-JamHex Nut 6
Thin-Jam	Hex Nut 3	Thin-JamHex Nut 3

## B) Logical Functions Use Logical DAX Functions for analysis:

### 1. Discount Category: Create a calculated column to categorize discounts

Screenshot of Power BI desktop showing the creation of a calculated column named "And". The formula is `1 And = AND(Orders1[Discount]==0,Orders1[Profit]>3)`. The Data pane shows a table with columns: ID, Category, Sub-Category, Product Name, Sales, Quantity, Discount, Profit, and And. The And column contains values like True, False, and Null. The ribbon is at the top, and the status bar at the bottom indicates it's saved today at 7:05 PM.

ID	Category	Sub-Category	Product Name	Sales	Quantity	Discount	Profit	And
10003479	Office Supplies	Storage	Eldon Base for stackable storage shelf, platinum	77.88	2	0	3.893999999999999	True
10003811	Office Supplies	Art	Newell 327	6.63	3	0	1.7901	False
10001246	Office Supplies	Art	Newell 317	5.88	2	0	1.7052	False
10000823	Office Supplies	Art	Newell 307	5.46	3	0	1.5288	False
10004456	Office Supplies	Art	Panasonic KP-4ABK Battery-Operated Pencil Sharpener	73.2	5	0	21.228	True
10002377	Office Supplies	Paper	Adams Telephone Message Book W/Dividers/Space For Phone Numbers, 5 1/4"X8 1/2", 200/Messages	22.72	4	0	10.224	True
10002005	Office Supplies	Paper	Xerox 225	45.36	7	0	21.7728	True
10002975	Office Supplies	Fasteners	Staples	11.34	3	0	5.2164	True
10003996	Office Supplies	Storage	Letter/Legal File Tote with Clear Snap-On Lid, Black Granite	80.3	5	0	20.878	True
10004477	Office Supplies	Paper	Xerox 1952	64.74	13	0	30.4278	True
10001963	Office Supplies	Storage	Tennsco Regal Shelving Units	405.64	4	0	12.1692	True
10001419	Office Supplies	Art	Newell 325	12.39	3	0	3.717	True
10000692	Office Supplies	Appliances	Fellowes Mighty 8 Compact Surge Protector	60.81	3	0	17.0268	True
10001457	Office Supplies	Paper	White GlueTop Scratch Pads	90.24	6	0	41.5104	True
10000675	Office Supplies	Storage	File Shuttle II and Handi-File, Black	305.01	9	0	76.2525	True
10000657	Office Supplies	Art	Newell 351	13.12	4	0	3.8048	True
10000575	Office Supplies	Art	Binney & Smith inkTank Desk Highlighter, Chisel Tip, Yellow, 12/Box	10.75	5	0	3.5475	True
10000134	Office Supplies	Fasteners	Advantus Push Pins, Aluminum Head	11.62	2	0	3.6022	True
10000552	Office Supplies	Paper	Xerox 200	12.96	2	0	6.2208	True
10001135	Office Supplies	Fastener	Brites Rubber Bands, 1 1/2 oz. Box	3.96	2	0	0.0792000000000002	False
10003602	Office Supplies	Art	Quartet Omega Colored Chalk, 12/Pack	11.68	2	0	5.4896	True
10004782	Office Supplies	Supplies	Elite 5" Scissors	16.9	2	0	5.07	True
10002301	Office Supplies	Supplies	Serrated Blade or Curved Handle Hand Letter Openers	6.28	2	0	0.0628000000000002	False
10000646	Office Supplies	Supplies	Premier Automatic Letter Opener	480.74	2	0	14.4222	True
10003442	Office Supplies	Storage	Eldon Portable Mobile Manager	141.4	5	0	38.178	True
10000482	Office Supplies	Paper	Snan-A-Way Black Print Carbonless Ruled Speed Letter, Triplicate	113.82	3	0	53.4954	True

### 2. High-Value Customers:

Screenshot of Power BI desktop showing the creation of a calculated column named "if". The formula is `1 if = IF(Orders1[Sales]>90,"high-value customer","low value customer")`. The Data pane shows a table with columns: Sub-Category, Product Name, Sales, Quantity, Discount, Profit, And, and if. The if column contains values like high-value customer and low value customer. The ribbon is at the top, and the status bar at the bottom indicates it's saved today at 7:05 PM.

Sub-Category	Product Name	Sales	Quantity	Discount	Profit	And	if
plies Storage	Eldon Base for stackable storage shelf, platinum	77.88	2	0	3.893999999999999	True	high-value customer
plies Art	Newell 327	6.63	3	0	1.7901	False	low value customer
plies Art	Newell 317	5.88	2	0	1.7052	False	low value customer
plies Art	Newell 307	5.46	3	0	1.5288	False	low value customer
plies Art	Panasonic KP-4ABK Battery-Operated Pencil Sharpener	73.2	5	0	21.228	True	low value customer
plies Paper	Adams Telephone Message Book W/Dividers/Space For Phone Numbers, 5 1/4"X8 1/2", 200/Messages	22.72	4	0	10.224	True	low value customer
plies Paper	Xerox 225	45.36	7	0	21.7728	True	low value customer
plies Fasteners	Staples	11.34	3	0	5.2164	True	low value customer
plies Storage	Letter/Legal File Tote with Clear Snap-On Lid, Black Granite	80.3	5	0	20.878	True	low value customer
plies Storage	Tennsco Regal Shelving Units	405.64	4	0	12.1692	True	high-value customer
plies Art	Newell 325	12.39	3	0	3.717	True	low value customer
plies Appliances	Fellowes Mighty 8 Compact Surge Protector	60.81	3	0	17.0268	True	low value customer
plies Paper	White GlueTop Scratch Pads	90.24	6	0	41.5104	True	high-value customer
plies Storage	File Shuttle II and Handi-File, Black	305.01	9	0	76.2525	True	high-value customer
plies Art	Newell 351	13.12	4	0	3.8048	True	low value customer
plies Art	Binney & Smith inkTank Desk Highlighter, Chisel Tip, Yellow, 12/Box	10.75	5	0	3.5475	True	low value customer
plies Fasteners	Advantus Push Pins, Aluminum Head	11.62	2	0	3.6022	True	low value customer
plies Paper	Xerox 200	12.96	2	0	6.2208	True	low value customer
plies Fasteners	Brites Rubber Bands, 1 1/2 oz. Box	3.96	2	0	0.0792000000000002	False	low value customer
plies Art	Quartet Omega Colored Chalk, 12/Pack	11.68	2	0	5.4896	True	low value customer
plies Supplies	Elite 5" Scissors	16.9	2	0	5.07	True	low value customer
plies Supplies	Serrated Blade or Curved Handle Hand Letter Openers	6.28	2	0	0.0628000000000002	False	low value customer
plies Storage	Premier Automatic Letter Opener	480.74	2	0	14.4222	True	high-value customer
plies Paper	Eldon Portable Mobile Manager	141.4	5	0	38.178	True	high-value customer
plies Paper	Snan-A-Way Black Print Carbonless Ruled Speed Letter, Triplicate	113.82	3	0	53.4954	True	high-value customer

## C) Time & Intelligence Functions Use Time Intelligence DAX Functions to analyze sales trends:

### 1. Calculate Total Sales Last Year:

Screenshot of Power BI desktop showing the 'Measure tools' ribbon tab selected. A DAX query is displayed in the main pane:

```

1 TotalSales2017 - 
2 CALCULATE(
3     SUM(olist_order_items_dataset[price]),
4     YEAR(olist_order_items_dataset[shipping_limit_date]) = 2017
5 )

```

The Data pane on the right shows the following relationships:

- locations
- olist\_order\_items\_dataset
  - $\Sigma$  freight\_value
  - order\_id
  - order\_item\_id
  - price
  - product\_id
  - seller\_id
- shipping\_limit\_date
- TotalSales2017

Table: olist\_order\_items\_dataset (1,12,650 rows) Column: TotalSales2017 (0 distinct values)

System status: 88°F Sunny

System status: ENG IN 13:48 29-01-2025

### 2. Find Year-to-Date (YTD) Sales:

Screenshot of Power BI desktop showing the 'Table tools' ribbon tab selected. A DAX query is displayed in the main pane:

```

1 YTDsales = TOTALYTD(SUM('olist_order_items_dataset'[price]), 'olist_order_items_dataset'[shipping_limit_date])

```

The Data pane on the right shows the following relationships:

- locations
- olist\_order\_items\_dataset
  - $\Sigma$  freight\_value
  - order\_id
  - order\_item\_id
  - price
  - product\_id
  - seller\_id
- shipping\_limit\_date
- YTDsales

Table: olist\_order\_items\_dataset (1,12,650 rows) Column: YTDsales (0 distinct values)

System status: 88°F Sunny

System status: ENG IN 13:51 29-01-2025

### 3. Calculate Month-over-Month Growth

Screenshot of Power BI desktop showing the 'Table tools' ribbon selected. A DAX query editor window is open, displaying the following DAX code:

```

1 PreviousMonthSales =
2 CALCULATE(
3     SUM(olist_order_items_dataset[price]),
4     DATEADD(olist_order_items_dataset[shipping_limit_date], -1, MONTH)
5 )
6

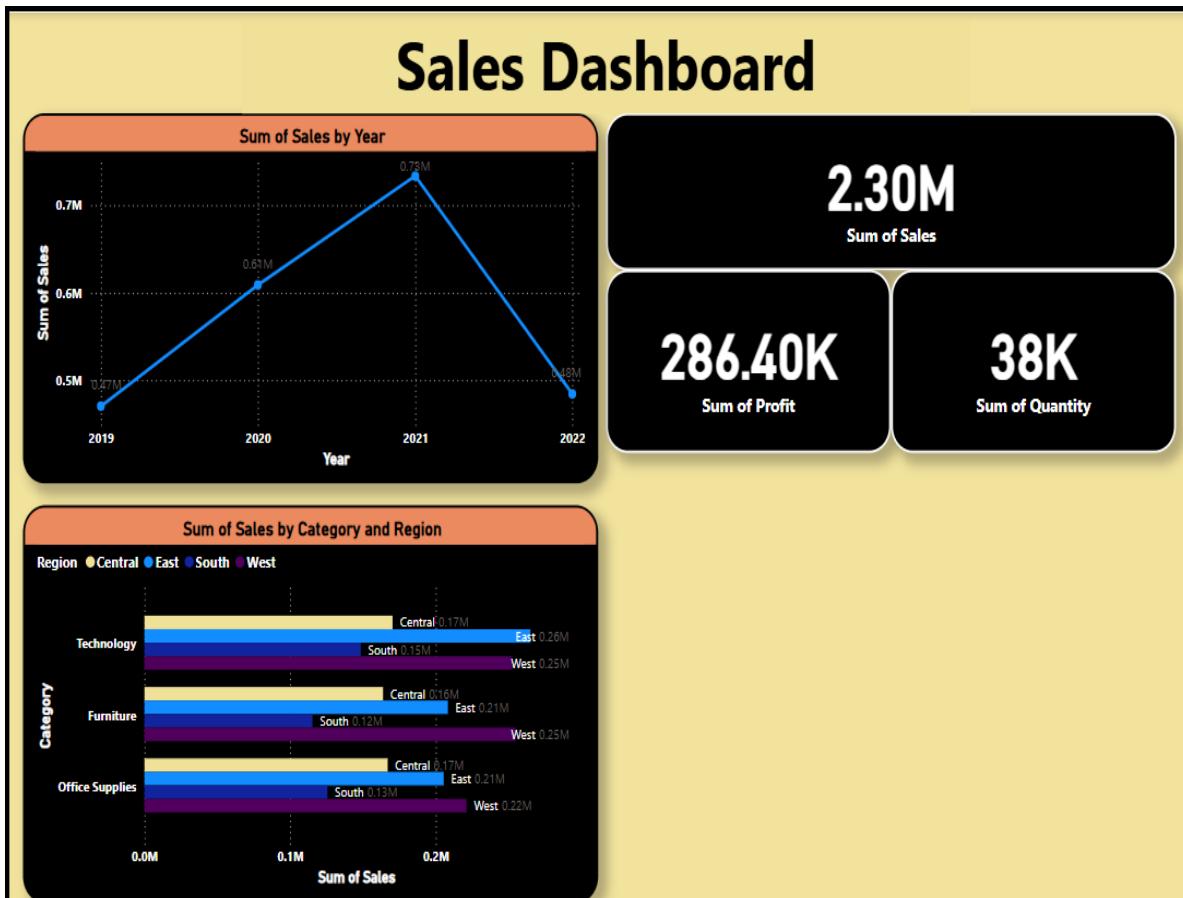
```

The Data pane on the right shows the structure of the 'olist\_order\_items\_dataset' table, including columns for freight\_value, order\_id, order\_item\_id, product\_id, seller\_id, shipping\_limit\_date, price, and freight\_value.

Below the table structure, the table preview shows approximately 12,650 rows of data. The status bar at the bottom indicates the following information: 88°F Sunny, ENG IN, 29-01-2025, and 13:53.

## Task 4: Data Visualization & Report Creation Create a Sales Dashboard:

- Total Sales, Profit, and Quantity (Cards)
- Sales Trends over time (Line Chart)
- Sales by Region & Product Category (Bar Chart)
- High-Value vs Regular Customers (Pie Chart)



## Create Customer Analysis Dashboard:

- Customer-wise sales & profits (Table)
- Top 10 Customers (Bar Chart)
- Discount Categories (Stacked Bar Chart)

