

Experiment-3

Aim: To understand Data Modeling, create and manage tables and relationships in Power BI.

Data modeling is the process of creating a visual representation of either a whole information system or parts of it to communicate connections between data points and structures. Benefits of data modelling

Data modelling makes it easier for developers, data architects, business analysts, and other stakeholders to view and understand relationships among the data in a database or data warehouse. In addition, it can:

- Reduce errors in software and database development.
- Increase consistency in documentation and system design across the enterprise.
- Improve application and database performance.
- Ease data mapping throughout the organization.
- Improve communication between developers and business intelligence teams.
- Ease and speed the process of database design at the conceptual, logical and physical levels.

Understanding measures

In Power BI Desktop, measures are created and displayed in *Report View*, *Data View*, or *Model View*. Measures you create yourself appear in the **Fields** list with a calculator icon. You can name measures whatever you want, and add them to a new or existing visualization just like any other field.

For Data visualization first of all we need data so we import the dataset from our personal computer for analysis of data.

Power BI Desktop



Get data



Recent sources



Open other reports



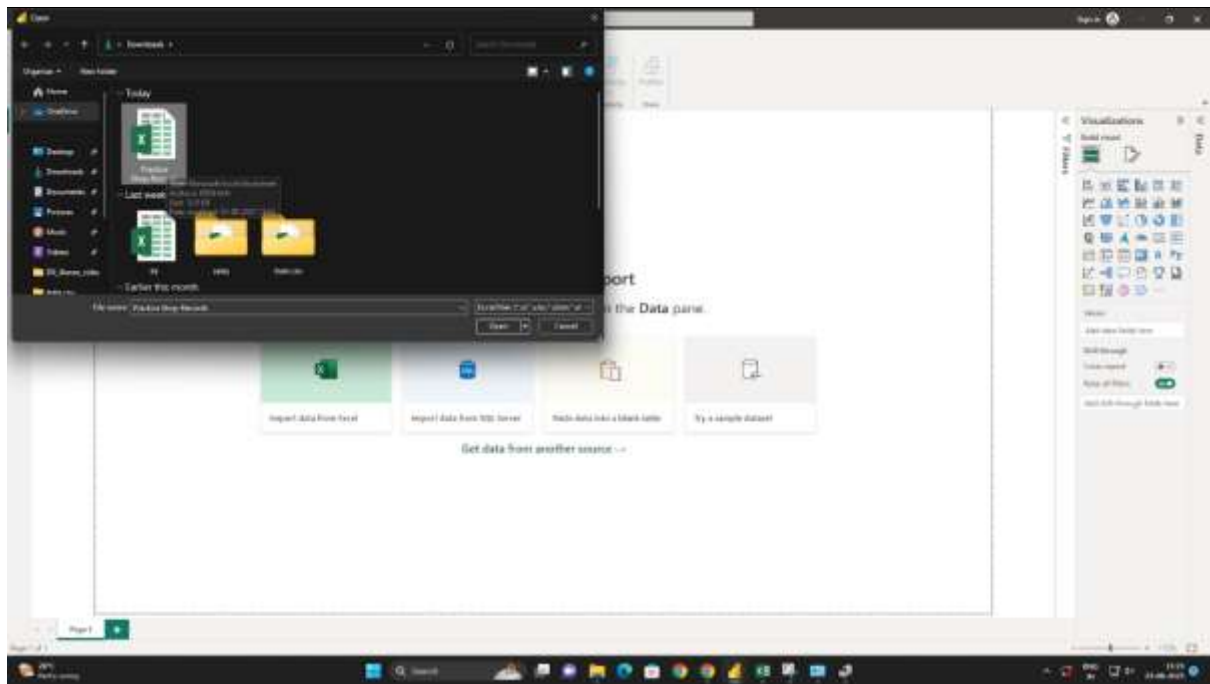
Collaborate and share

Sign in to publish your reports, access certified datasets, and share insights on organizational content in the Power BI service.

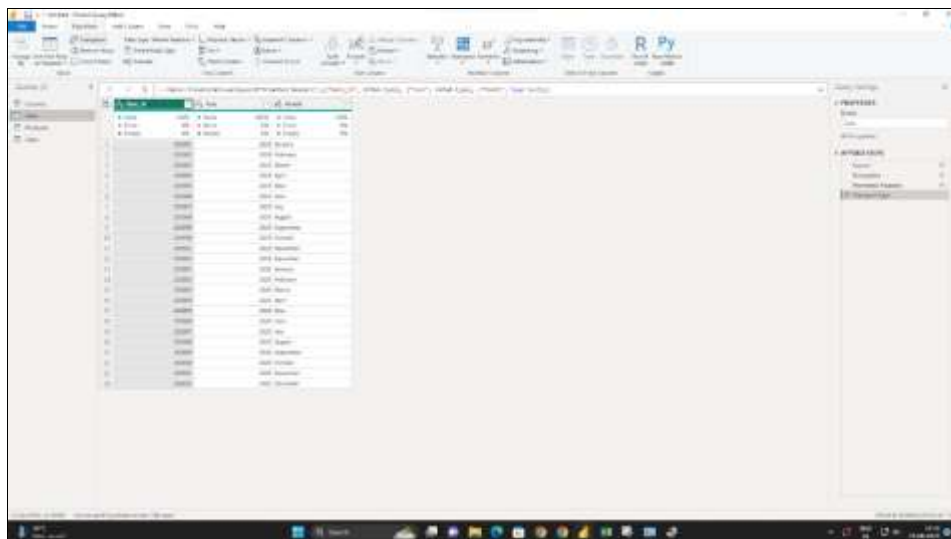
[Get started](#)

Want to buy a Power BI license? [Buy now](#)

Our dataset is in Excel file so we click on the excel workbook and connect..



I have seen there are four tables **Country,Date,Product,Sales** in this dataset . We select the tables what we want to do for analysis.



I have seen all the tables and done the data preprocessing . Why is this important ? Data preprocessing is an important step in the data mining process. It refers to the cleaning, transforming, and integrating of data in order to make it ready for analysis

File Home Transform Add Column View Tools Help

Transpose
 Reverse Rows
 Count Rows

Data Type: Text
 Detect Data Type
 Rename

Replace Values
 Fill
 Pivot Column

Unpivot Columns
 Move
 Convert to List

Split Column
 Format
 Parse

Merge Column
 Extract
 Parse

Table Any Column Text Column

Queries [4]

Country
Date
Products
Sales

= Table.Sort(#"Filtered Rows1",{{"Customer_Name", Order.Ascending}})

	Customer_ID	Customer_Name	Country_Code	Product_ID
1	127	A	IN	
2	179	A	IN	
3	153	A	IN	
4	101	A	IN	
5	128	B	FR	
6	154	B	FR	
7	102	B	FR	
8	180	B	FR	
9	103	C	US	
10	155	C	US	
11	181	C	US	
12	129	C	US	
13	156	D	IT	
14	182	D	IT	
15	104	D	IT	
16	130	D	IT	
17	131	E	MY	
18	157	E	MY	
19	105	E	MY	
20	183	E	MY	
21	184	F	UK	
22	132	F	UK	
23	106	F	UK	
24	158	F	UK	
25	107	G	AU	
26	185	G	AU	
27	159	G	AU	
28	133	G	AU	
29	186	H	TH	
30	134	H	TH	
31	160	H	TH	
32	108	H	TH	
33	187	I	NZ	
34	161	I	NZ	
35	135	I	NZ	
36	109	I	NZ	

File Home Transform Add Column View Tools Help

Group By Use First Row as Headers Count Rows

Transpose Reverse Rows

Data Type: Text Detect Data Type Rename

Replace Values Fill Pivot Column Convert to List

Unpivot Columns Move

Split Column Format Merge Columns Extract Parse

Statistics

Queries [4]

- Country
- Date
- Products
- Sales

fx = Table.Sort(#"Changed Type",{{"Month", Order.Ascending}})

	Date_ID	Year	Month
1		202	
2		201	
3		201	
4		202	
5		201	
6		202	
7		202	
8		201	
9		201	
10		202	
11		201	
12		202	
13		201	
14		202	
15		202	
16		201	
17		201	
18		202	
19		202	
20		201	
21		202	
22	201910	2019	October
23	202009	2020	September
24	201909	2019	September

Valid Error Empty

Sort Ascending Sort Descending Clear Sort

Clear Filter Remove Empty Text Filters

Search

- (Select All)
- April
- August
- December
- February
- January
- July
- June
- March
- May
- November
- October
- September

OK Cancel

The screenshot shows the Power Query Editor interface. The 'Queries' pane on the left lists 'Country', 'Date', 'Products', and 'Sales'. The main area displays a table with two columns: 'Country_Code' and 'Country'. The data is as follows:

Country_Code	Country
1	IN
2	FR
3	US
4	IT
5	MY
6	UK
7	AU
8	TH
9	NZ
10	RU
11	JP
12	KW

The 'Split Column' menu is open, showing options: 'By Delimiter', 'By Number of Characters', 'By Positions', 'By Lowercase to Uppercase', 'By Uppercase to Lowercase', 'By Digit to Non-Digit', and 'By Non-Digit to Digit'.

Arrange the country dataset in Alphabetical order.
And date and month in ascending order.

The screenshot shows the Power Query Editor interface with the 'Split Column by Delimiter' dialog box open. The dialog box has the following options:

- Select an origin delimiter:** A dropdown menu with 'Custom' selected.
- Split at:**
 - ☐ Left-most delimiter
 - ☐ Right-most delimiter
 - ☒ Each occurrence of the delimiter
- Advanced options:**
 - Quote Character:** A dropdown menu with 'None' selected.
 - ☐ Split using special characters

The 'OK' button is highlighted in green.

In country table we apply the delimiter to specify more clear data.
Also create a new column.

Home

Transform

Add Column

View

Tools

Help

Table

Use First Row as Headers

Count Rows

Transpose

Reverse Rows

Count Rows

Data Type: Text

Detect Data Type

Rename

Any Column

Replace Values

Fill

Pivot Column

Unpivot Columns

Move

Convert to List

Split Column

Format

Extract

Parse

Text Column

Merge Columns

Statistics

Standard

Scientific

Information

Number Column

fx

= Table.TransformColumnTypes(#"Split Column by Delimiter",{{"Column1.1", type text}, {"Column1.2", type text}})

Column1.1

Column1.2

Column2

Valid

100%

Valid

0%

Valid

100%

Error

0%

Error

0%

Error

0%

Empty

0%

Empty

100%

Empty

0%

1

Country_Code

Country

2

IN

India

3

FR

France

4

US

United States

5

IT

Italy

6

MY

Malaysia

7

UK

United Kingdom

8

AU

Australia

9

TH

Thailand

10

NZ

New Zealand

11

RU

Russia

12

JP

Japan

13

KW

Kuwait

File Home Insert Layout References Data Tools PowerPivot Help

Replace Values ▾ Unpivot Columns ▾
Move ▾
Convert to List
Any Column

Split Column ▾

Format ▾

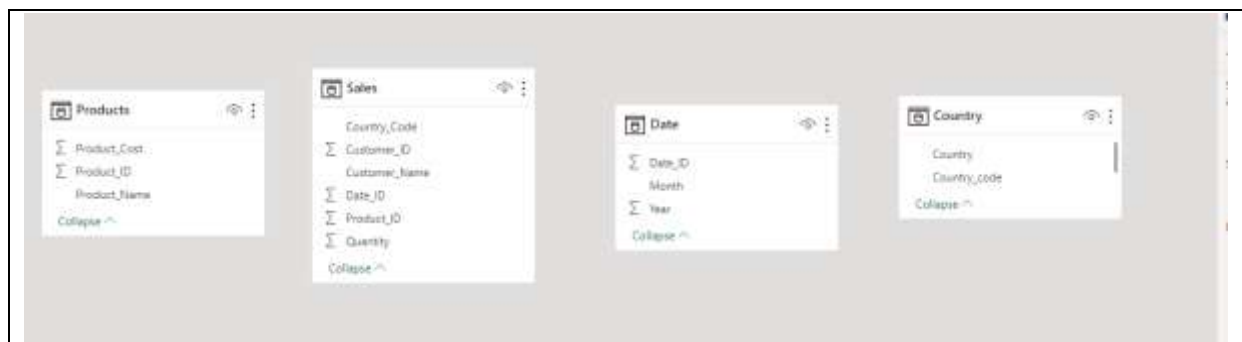
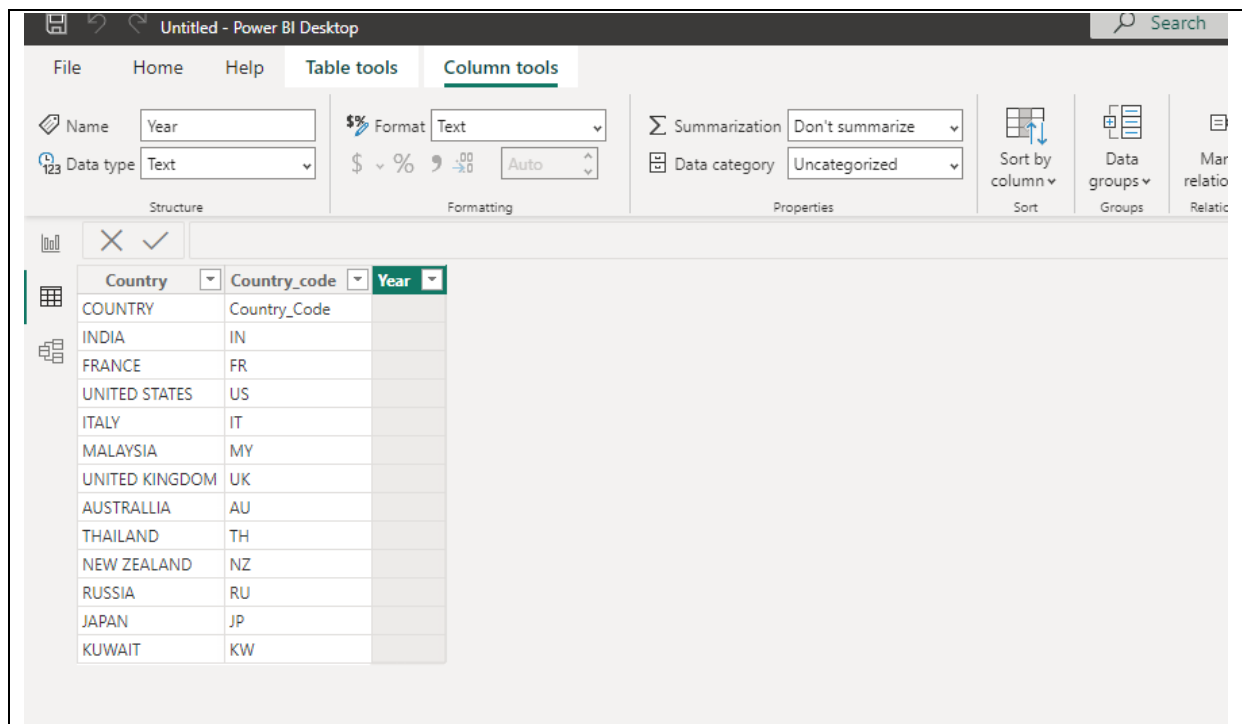
Merge Columns
Extract ▾
Parse ▾

Statistics ▾

11Down("#Changed Type1",{ "Column1.1"})

ABC Column1.2	ABC Column2
Valid 0%	Valid
Error 0%	Error
Empty 100%	Empty
	Country
	India
	France
	United States
	Italy
	Malaysia
	United Kingdom
	Australia
	Thailand
	New Zealand
	Russia
	Japan
	Kuwait

- lowercase
- UPPERCASE
- Capitalize Each Word
- Trim
- Clean
- Add Prefix
- Add Suffix



Now we do Data modelling in above screenshot there is no relationship between the tables, Check the relationship between those tables and draw a ER (Entity-Relationship) diagram.

Create relationship

Select tables and columns that are related.

Products

Product_ID	Product_Name	Product_Cost
10	Smartphone	200
11	Tablet	270
12	Computer	470

Sales

Customer_ID	Customer_Name	Country_Code	Product_ID	Quantity	Date_ID
101	A	IN	10	800	201901
127	A	IN	10	800	201906
153	A	IN	10	800	201912

Cardinality

Cross filter direction

☐ Make this relationship active

☐ Assume referential integrity

☐ Apply security filter in both directions

⚠ There's already a relationship between these two columns.

OKCancel

Product_ID
Product_Name
Collapse

Customer_Name
Date_ID
Product_ID
Quantity
Collapse

Product_ID	Product_Name	Product_Cost
10	Smartphone	200
11	Tablet	270
12	Computer	470

Sales

Customer_ID	Customer_Name	Country_Code	Product_ID	Quantity	Date_ID
101	A	IN	10	800	201901
127	A	IN	10	800	201906
153	A	IN	10	800	201912

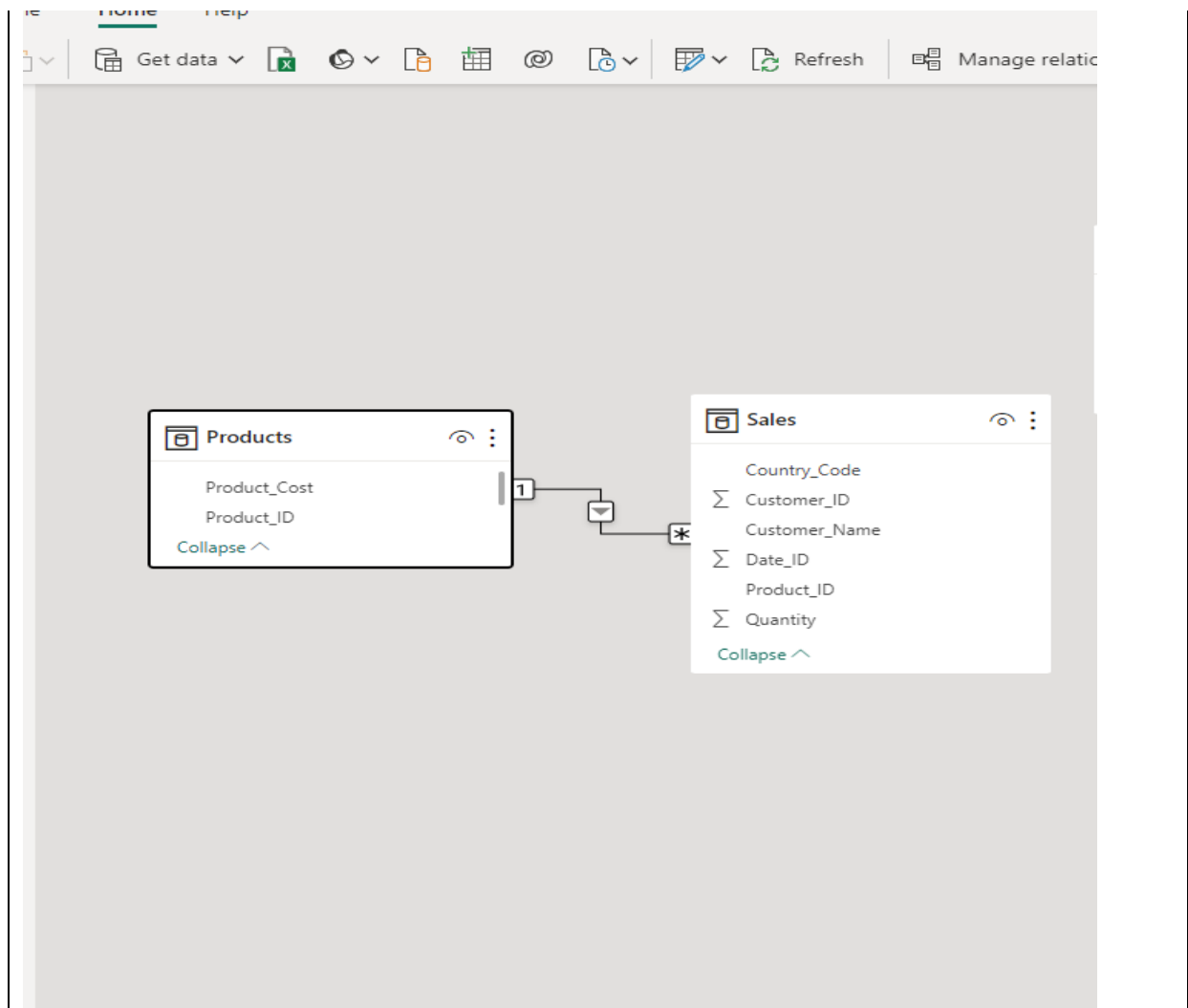
Cardinality

Cross filter direction

☒ Make this relationship active

☐ Assume referential integrity

☒ Apply security filter in both directions



We found that relationship between Products and Sales tables.

ry_Code

mer_ID

mer_Nam

D

ct_ID

ity

^

Create relationship

Select tables and columns that are related.

Date

Date_ID	Year	Month
202004	2020	April
201904	2019	April
202008	2020	August

Sales

Customer_ID	Customer_Name	Country_Code	Product_ID	Quantity	Date_ID
101	A	IN	10	800	201901
127	A	IN	10	800	201906
153	A	IN	10	800	201912

Cardinality

One to many (1:*)

Cross filter direction

Single

☒ Make this relationship active

☐ Assume referential integrity

☐ Apply security filter in both directions

OK

Cancel

Products

Product_Cost

Product_ID

Collapse ^

1

*

Sales

Country_Code

Customer_ID

Customer_Name

Date_ID

Product_ID

Quantity

Collapse ^

*

*

Date

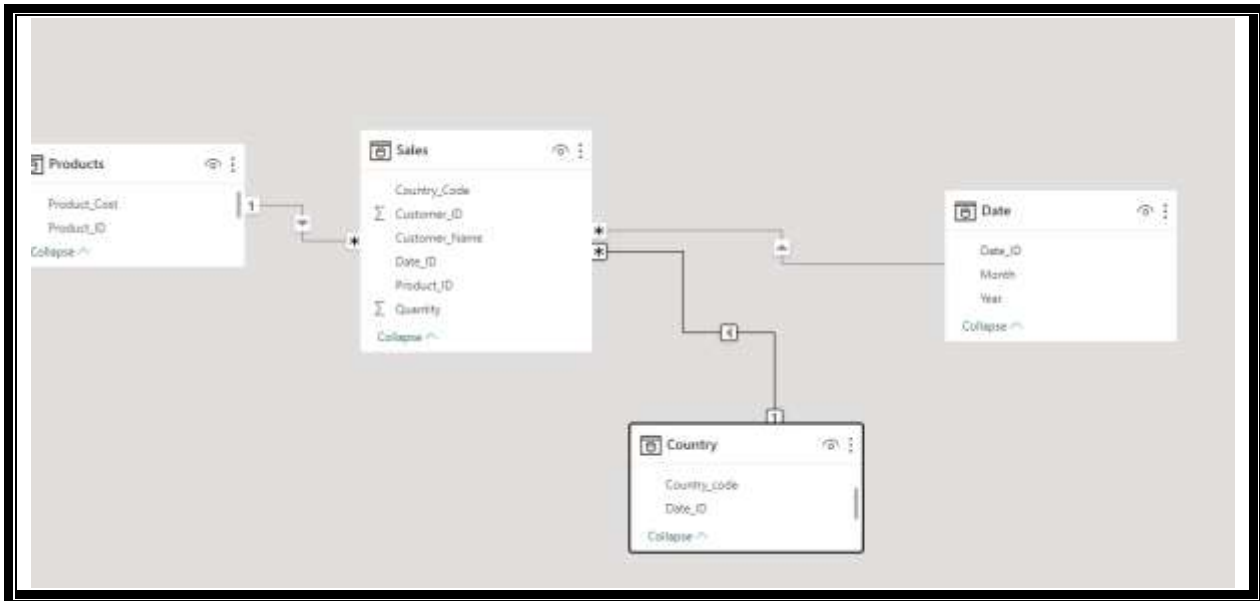
Date_ID

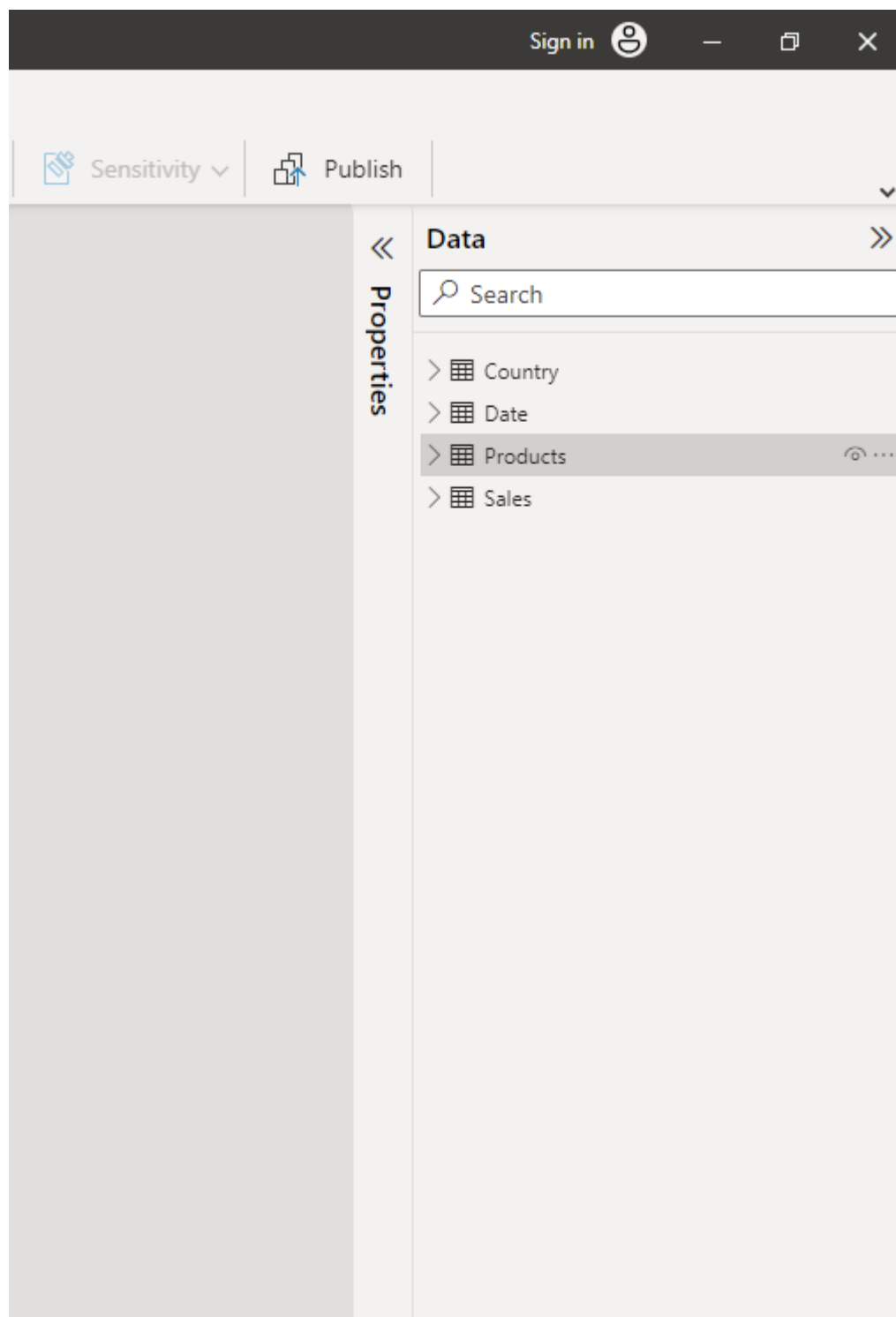
Month

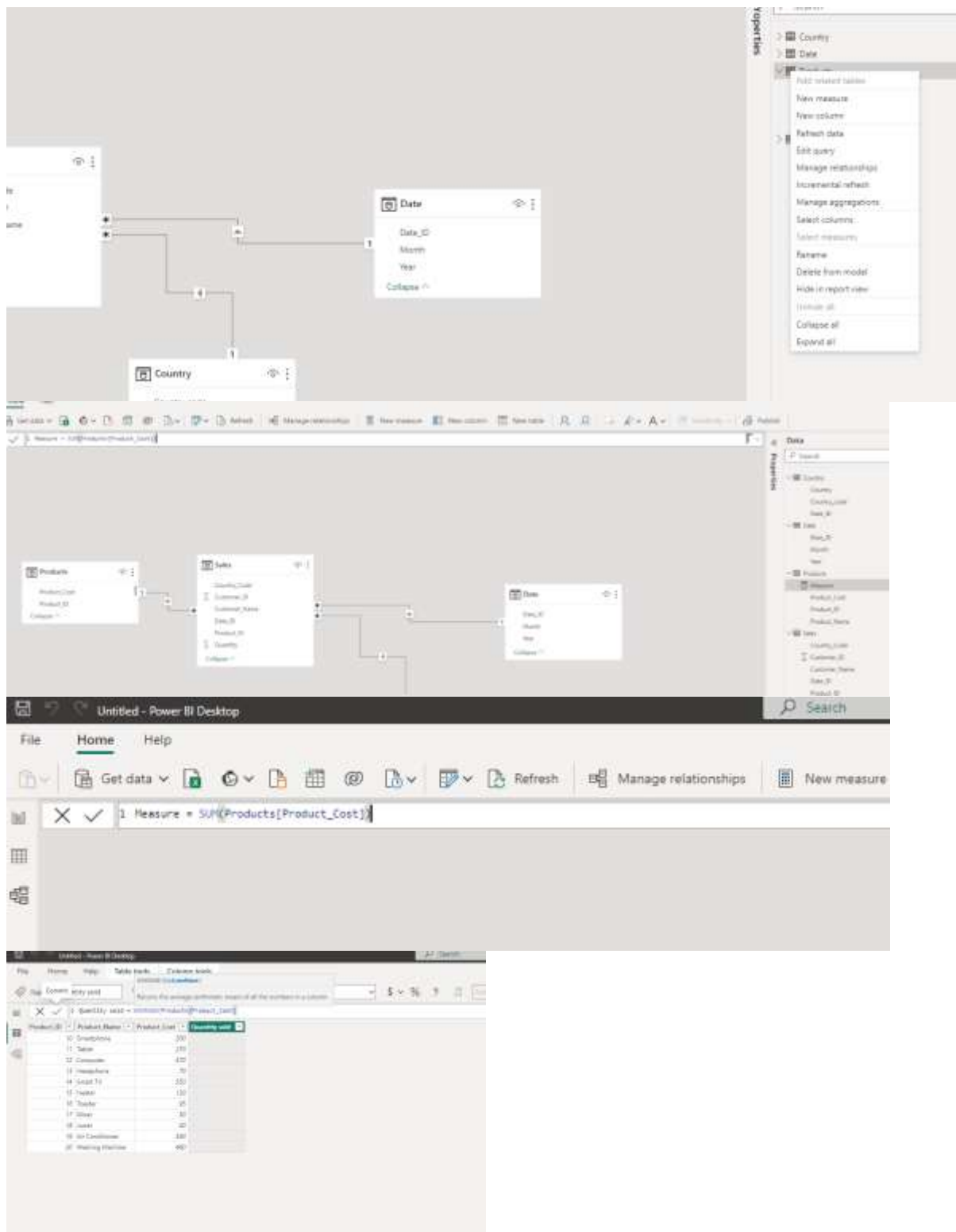
Year

Collapse ^

E-R Diagram is created. We found one to many relationships in those tables.







We create measures using DAX(Data Analysis Expression) for creating various expressions and filters.

Untitled - Power BI Desktop

Home Help **Table tools** **Column tools**

Name Quantity sold Data type Decimal number Format

1 Quantity sold = AVERAGE(Product

Product_ID Product_Name Product_Cost

10	Smartphone	200
11	Tablet	270
12	Computer	470
13	Headphone	70
14	Smart TV	550
15	Heater	130
16	Toaster	95
17	Mixer	50
18	Juicer	40
19	Air Conditioner	540
20	Washing Machine	460

Whole number
Decimal number
Fixed decimal number
Date/time
Date
Time
Text
True/false
Binary

Untitled - Power BI Desktop

Home Help **Table tools** **Column tools**

Name Quantity sold Data type Whole number Format

1 Quantity sold = AVERAGE(Products[Product_Cost])

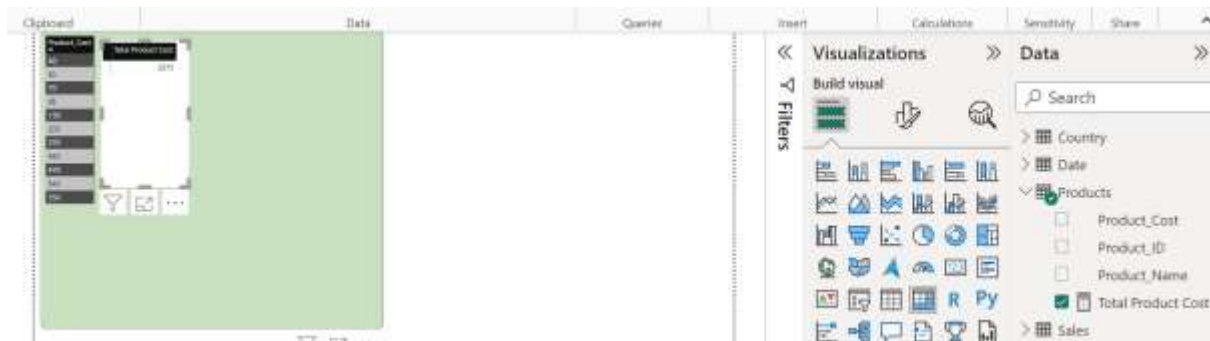
Product_ID	Product_Name	Product_Cost	Quantity sold
10	Smartphone	200	261
11	Tablet	270	261
12	Computer	470	261
13	Headphone	70	261
14	Smart TV	550	261
15	Heater	130	261
16	Toaster	95	261
17	Mixer	50	261
18	Juicer	40	261
19	Air Conditioner	540	261
20	Washing Machine	460	261

Home table Products \$ % 0

Structure Formatting

1 Total Product Cost = SUM([Product_Cost])

Product_ID	Product_Name	Product_Cost
10	Smartphone	200

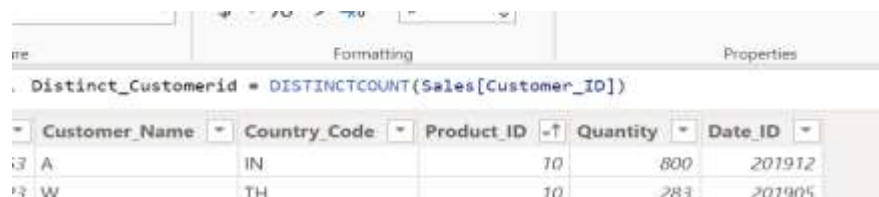


After using Sum dax expression we visualize the sum of total product in matrix form above screenshot shows the representation of matrix.



Total Quantity sold is sum of all quantity.

- We use DISTINCTCOUNT function counts the BLANK value. To skip the BLANK value, use the [DISTINCTCOUNTNOBLANK](#) function.



Data

Total Quantity sold	Distinct_Customerid
44536	104

Name:
 Format:
 Summarization:
 Data category:
 Sort by column:

Data type:
 \$ % 0.00 0

Structure Formatting Properties Sort

1 Product_Cost = RELATED(Products[Product_Cost])

Customer_ID	Customer_Name	Country_Code	Product_ID	Quantity	Date_ID	Product_Cost
101	A	IN	10	800	201901	200
102	B	FR	11	104	201901	270
103	C	US	12	300	201901	470

Formatting

Distinct_Productid = DISTINCTCOUNT(Products[Product_ID])

Product Name Product Cost

Distinct_Productid
11

TOTAL REVENUE = $\frac{[Total\ Quantity\ sold] * [Total\ Product\ Cost]}{[Total\ Quantity\ sold]}$

Queries

through its SQL endpoint.

9000000
1156000
3450000
2783000
1691100
1207000
5784000
5037000
4048000
1291000
4076000
11207400
Total 128041000

Matrix Dashboard in POWER BI

Clipboard

Data

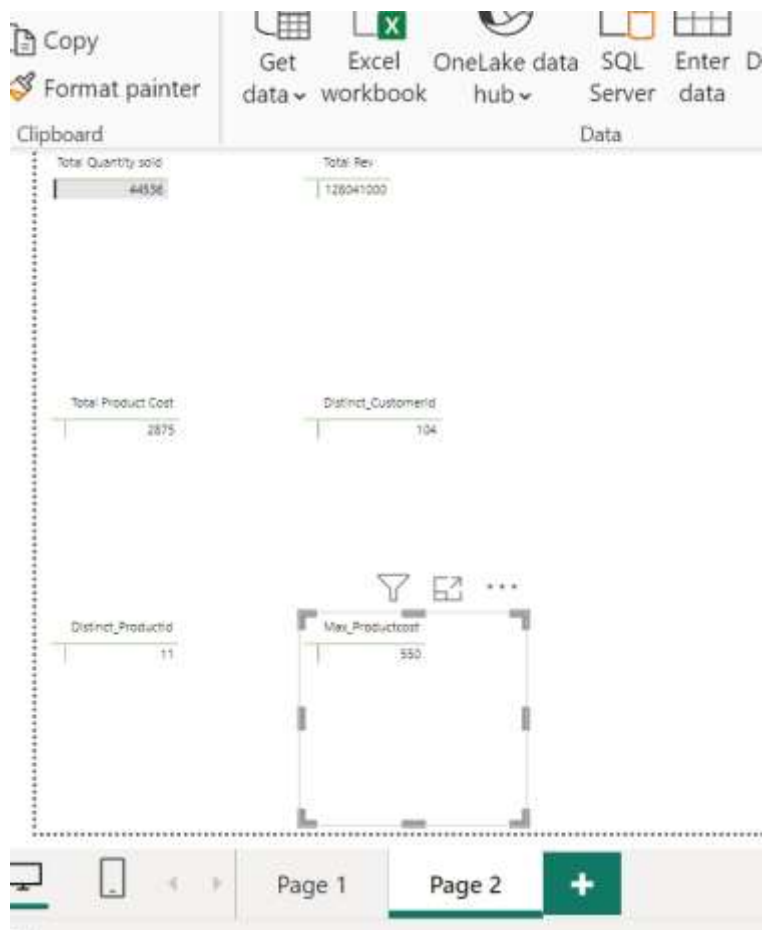
Queries

Product_Cost Total Product Cost

460

Total Quantity sold	Distinct_Customerid	Distinct_Productid
2888	8	1

Customer_Name	Total Quantity sold	Total Product Cost
C	780160	
	548320	
Total	1328480	



Create new Columns--

ata type

Whole number

\$ % 0

Data category

Uncategorized

Sort by column

Data groups

Structure

Formatting

Properties

Sort

Groups

X

✓

1 Total Cost = Sales[Quantity]*Sales[Product_Cost]

Customer_ID	Customer_Name	Country_Code	Product_ID	Quantity	Date_ID	Product_Cost	Total Cost
101	A	IN	10	800	201901	200	160000
102	B	FR	11	104	201901	270	28080
103	C	US	12	300	201901	470	141000
104	D	IT	13	240	201901	70	16800
105	E	MY	14	321	201901	550	176550
106	F	UK	15	105	201902	130	13650
107	G	AU	16	503	201902	95	47785
108	H	TH	17	438	201902	50	21900
109	I	NZ	18	352	201902	40	14080
110	J	IN	19	678	201902	540	366120
111	K	RU	20	424	201903	460	195040
112	L	IN	10	965	201903	200	193000

Power BI Desktop interface showing the DAX formula bar and a data table.

Measure: Sum_Totalcost

Format: Whole number

Data category: Uncategorized

Measure table: Sales

Structure: 1 Sum_Totalcost = SUM(Sales[Total Cost])

Customer_ID	Customer_Name	Country_Code	Product_ID	Quantity	Date_ID	Product_Cost	Total Cost
101	A	IN	10	800	201901	200	160000
102	B	FR	11	104	201901	270	28080
103	C	US	12	300	201901	470	141000
104	D	IT	13	240	201901	70	16800
105	E	MY	14	321	201901	550	176550
106	F	UK	15	105	201902	130	13650
107	G	AU	16	503	201902	95	47785

Measure: CostComparison

Format: Whole number

Structure: 1 CostComparison = If('Sales'[Total Cost]>=200000,"
2 IF('Sales'[Total Cost]]>=100000,"High",
3 IF('Sales'[Total Cost]]>=50000,"Average","Low"))



Why We Create Power Bi Dashboard-

Power BI report service, just like the desktop version, includes a variety of page formatting options, including visuals, shapes, and images, that can help your report stand out. One of the most efficient ways to identify and communicate insights is to **use Power BI to create visuals**.

format painter | data ▾ | workbook | hub ▾ | Server | data | sources ▾ | data ▾ | visual | box | vis

Clipboard | Data | Queries | Insert

104 Distinct Customer	2875 Total Product Cost	550 Max Product Cost	12M Sum Revenue	45K Total Quantity sold	11 Distinct Product	128M Total Rev
--------------------------	----------------------------	-------------------------	--------------------	----------------------------	------------------------	-------------------

Product Name | Product Cost | Total Quantity sold | Sum of Total Cost

Air Conditioner	540	4244	\$22,91,785
Computer	470	4334	\$19,90,920
Headphone	70	3928	\$2,67,960
Heater	130	1988	\$2,58,440
Juicer	40	2584	\$1,03,360
Killer	30	3452	\$1,72,600
Smart TV	550	3876	\$21,86,800
Smartphone	200	8192	\$16,38,400
Tablet	270	1732	\$10,07,840
Toaster	85	5416	\$5,14,520
Washing Machine	400	2888	\$13,26,400
Total		44136	\$1,57,60,880

Sum of Total Cost, Total Rev and Total Product Cost

Sum of Total Cost, Total Rev and Total Product Cost

Max Product Cost and Total Product Cost by Product Cost

The dashboard displays a table of product data with columns: Product Name, Product Cost, Total Quantity sold, and Sum of Total Cost. Below the table is a bar chart titled 'Sum of Total Cost, Total Rev and Total Product Cost' showing the distribution of these values. To the right is a pie chart titled 'Max Product Cost and Total Product Cost by Product Cost' showing the distribution of product costs.