

Advanced Power Query & Data Modelling - p2

- Merge Queries [Customer & Sales].
- Append Queries [Sales & Sales - p2].
- What is Data Model.
- Dimension & Fact Table.
- Star / Snowflake Schema.
- Relationships and Cardinality.
- Filter Flow - Downstream.
- Single/ Bi-directional Filters.
- Active / Inactive Relationship.

FACT & DIMENSION TABLES

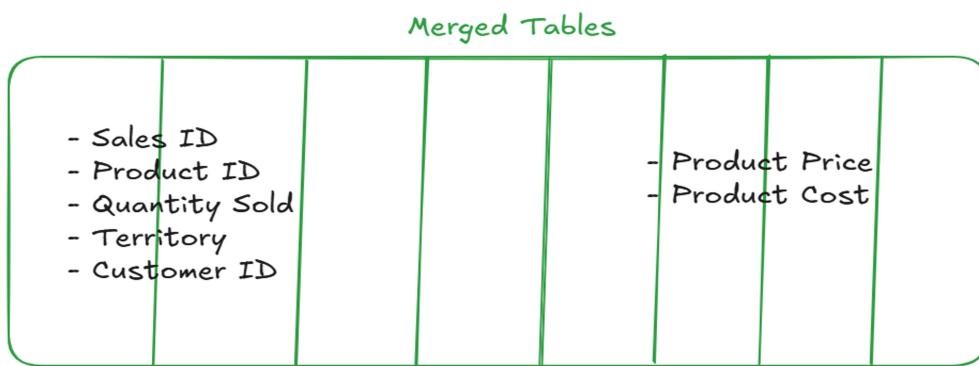
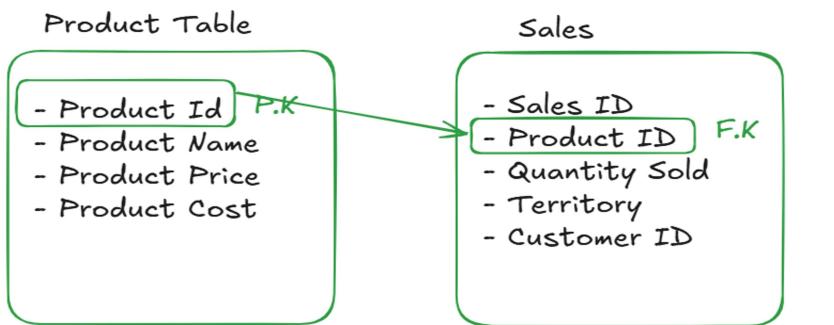
Data models generally contain two types of tables: fact ("data") tables, and dimension ("lookup") tables:

- Fact tables contain numerical values or metrics used for summarization (sales, orders, transactions, pageviews, etc.)
- Dimension tables contain descriptive attributes used for filtering or grouping (products, customers, dates, stores, etc.)

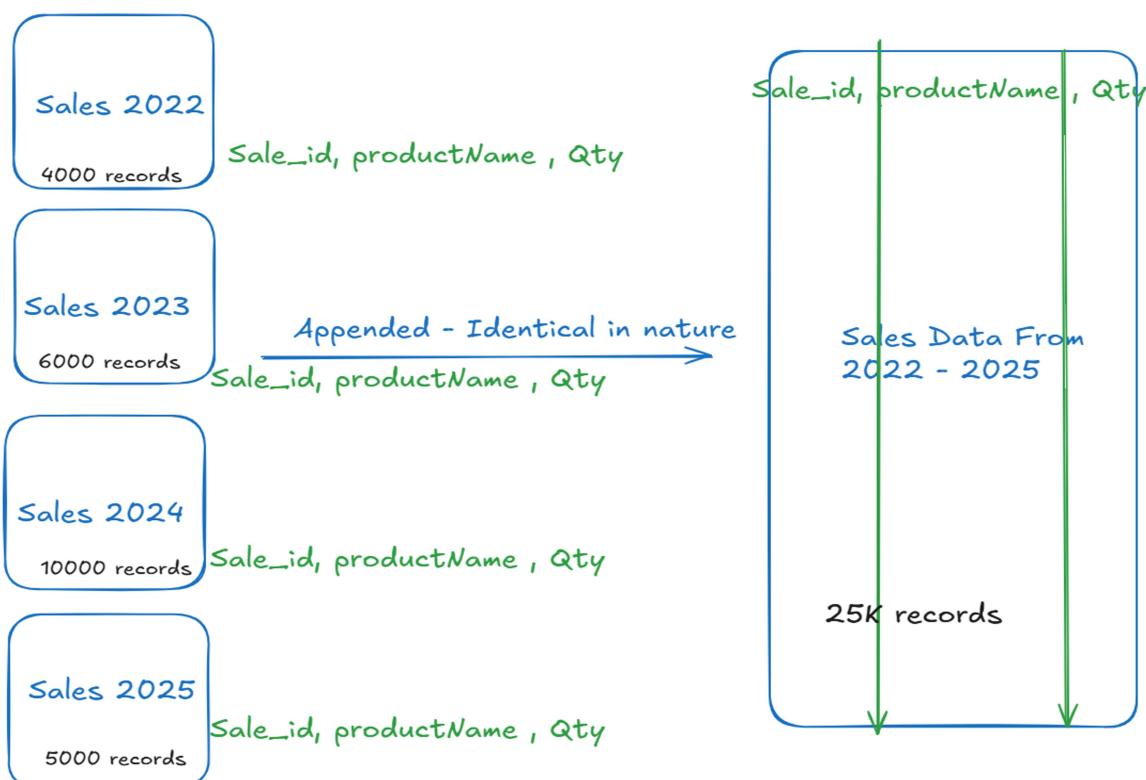


Merge Queries

Your final table will be merged in a horizontal Direction



Appended Queries → It connects multiple identical table vertically.



You to Everyone 21:14

CN Sales - Part 2.xlsx 10.52 KB

You to Everyone 21:15

CN Sales Detail.xlsx 9.54 KB

	Customer ID	Customer Region	Customer Age	Sales
1	21	North	10	6,532.00
2	22	North	23	7,958.00
3	23	South	56	1,012.00
4	24	East	38	5,816.00
5	25	East	22	5,651.00
6	26	West	22	3,215.00
7	27	West	58	8,314.00
8	28	East	11	3,575.00
9	29	East	43	7,459.00
10	30	North	45	1,607.00
11	31	North	14	6,078.00
12	32	South	27	9,065.00
13	33	East	57	4,423.00
14	34	East	17	850.00
15	35	East	53	3,701.00
16	36	East	59	5,873.00
17	37	North	49	4,385.00
18	38	North	33	9,561.00
19	39	South	38	2,322.00
20	40	East	59	4,881.00
21	41	North	18	6,154.00
22	42	North	36	9,355.00

Merge Queries

Merge

Select a table and matching columns to create a merged table.

Sales

Customer ID	Customer Region	Customer Age	Sales
1	North	49	309.38
2	Unknown	10	57.71
3	North	47	686.95
4	South	11	1,619.19
5	East	17	1,005.90

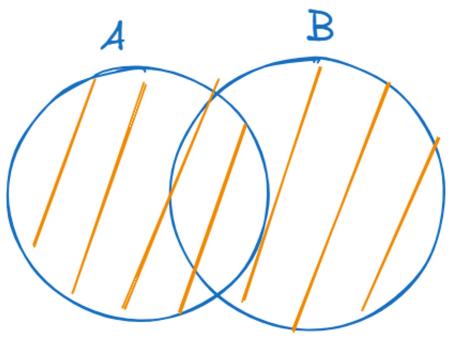
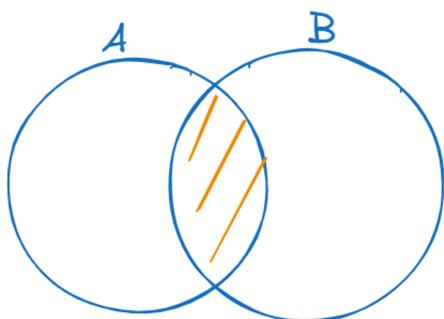
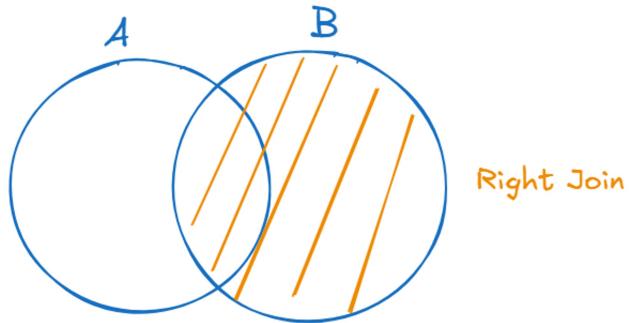
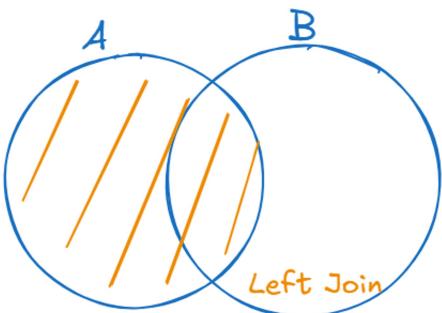
Customers

Customer ID	Customer Name
1	Elvin Catron
2	Gustano Primm
3	Lacresha Whitty
4	Octavio Ricci
5	Jacinda Moffett

Join Kind

- Left Outer (all from first, matching from second)
- Left Outer (all from first, matching from second)
- Right Outer (all from second, matching from first)
- Full Outer (all rows from both)
- Inner (only matching rows)
- Left Anti (rows only in first)
- Right Anti (rows only in second)

OK Cancel



Inner Join

Full Join - $A \cup B$

Left Join

Customer ID	Customer Region	Customer Age	Sales
1	North	49	309.38
2	Unknown	10	57.71
3	North	47	686.95
4	South	11	1,619.19
5	East	17	1,005.90

Customer ID	Customer Name
1	Elvin Catron
2	Gustano Primm
3	Lacresha Whitty
4	Octavio Ricci
5	Jacinda Moffett

- 6 Krishna Madan
- 7 Abhishek Behal
- 8 Naman Arora

After Join the 2 Table using Left Join

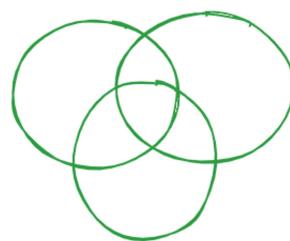
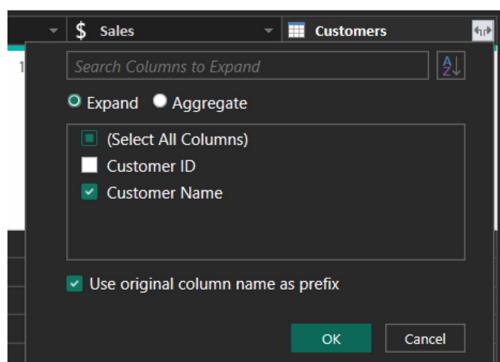
Customer ID	Customer Region	Customer Age	Sales	Customer Name
1	North	49	309.38	Elvin Catron
2	Unknown	10	57.71	Gustano Primm
3	North	47	686.95	Lacresha Whitty
4	South	11	1,619.19	Octavio Ricci
5	East	17	1,005.90	Jacinda Moffett

After Join the 2 Table using Right Join

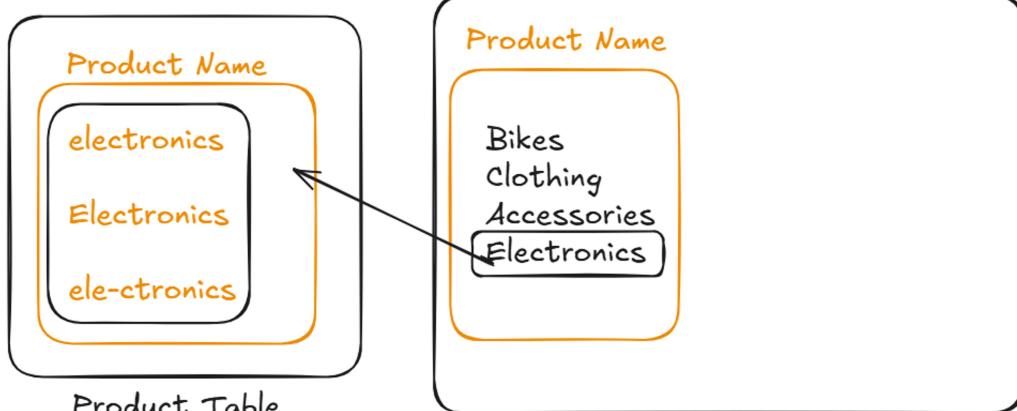
Customer ID	Customer Name	Customer Region	Customer Age	Sales
1	Elvin Catron	North	49	309.38
2	Gustano Primm	Unknown	10	57.71
3	Lacresha Whitty	North	47	686.95
4	Octavio Ricci	South	11	1,619.19
5	Jacinda Moffett	East	17	1,005.90
6	Krishna Madan	null	null	null
7	Abhishek Behal	null	null	null
8	Naman Arora	null	null	null

Customer ID	Customer Region	Customer Age	Sales	Customers
Valid	100%	Valid	100%	Valid 100%
Error	0%	Error	0%	Error 0%
Empty	0%	Empty	0%	Empty 0%
20 distinct, 20 unique	5 distinct, 2 unique	16 distinct, 12 unique	20 distinct, 20 unique	20 distinct, 20 unique
1 North		49	309.38	Table
2 Unknown		10	57.71	Table
3 North		47	686.95	Table
4 South		11	1,619.19	Table
5 East		17	1,005.90	Table
6 Unknown		23	174.65	Table
7 Unknown		21	825.00	Table
8 Unknown		27	250.00	Table
9 Unknown		44	9.03	Table
10 Unknown		10	255.84	Table
11 North		29	151.24	Table
12 South		29	251.72	Table
13 South		42	1,139.43	Table
14 North		12	575.36	Table
15 Unknown		45	18.06	Table
16 Unknown		30	299.85	Table
17 Unknown		35	54.89	Table
18 West		50	479.04	Table
19 Unknown		42	1,879.06	Table

Expand Aggregate



$\sum_{i=1}^n$ Customer ID	A _B	C Customer Region	$\sum_{i=1}^n$ Customer Age	\$ Sales	A _B	C Customers.Customer Name
Valid 100%	Valid 100%	100%	Valid 100%	100%	Valid 100%	100%
Error 0%	Error 0%	0%	Error 0%	0%	Error 0%	0%
Empty 0%	Empty 0%	0%	Empty 0%	0%	Empty 0%	0%
20 distinct, 20 unique	5 distinct, 2 unique	16 distinct, 12 unique	20 distinct, 20 unique	20 distinct, 20 unique		
1 North		49		309.88	Elvin Catron	
2 Unknown		10		57.71	Gustano Primmm	
3 North		47		686.95	Lacresha Whitty	
4 South		11		1,619.19	Octavio Ricci	
5 East		17		1,005.90	Jacintha Moffett	
6 Unknown		23		174.65	Betty scruggs	
7 Unknown		21		825.00	Era Gassner	
8 Unknown		27		250.00	Laticia Crimi	
9 Unknown		44		9.03	Alberta Fabela	
10 Unknown		10		255.84	Taylor Pogue	
11 North		29		151.24	Griselda Mariscal	
12 South		29		251.72	Sylvester Willbourn	
13 South		42		1,139.43	Ricky Cheshire	
14 North		12		575.36	Karisa Oquwndo	
15 Unknown		45		18.06	Debbie Orone	
16 Unknown		30		299.85	Loralee Widell	
17 Unknown		35		54.89	Season Viers	
18 West		50		479.04	Juan Luker	
19 Unknown		42		1,879.06	Shan texeria	



Sales Table

Append Queries

5 Columns

Customers

Sales

Product Information

Sales detail

Sales - part 2

	Customer ID	Customer Region	Customer Age	Sales	Customer Name
1	1 North		49	309.38	Elvin Catron
2	2 Unknown		10	57.71	Gustano Primmm
3	3 North		47	686.95	Lacresha Whitty
4	4 South		11	1,619.19	Octavio Ricci
5	5 East		17	1,005.90	Jacinda Moffett
6	6 Unknown		23	174.65	Betty scruggs
7	7 Unknown		21	825.00	Era Gassner
8	8 Unknown		27	250.00	Laticia Crimi
9	9 Unknown		44	9.03	Alberta Fabela
10	10 Unknown		10	255.84	Taylor Pogue
11	11 North		29	151.24	Grisede Mariscal
12	12 South		29	251.72	Sylvester Willbourn
13	13 South		42	1,139.43	Ricky Cheshire
14	14 North		12	575.36	Karisa Oquwndo
15	15 Unknown		45	18.06	Debbie Orone
16	16 Unknown		30	299.85	Loralee Widell
17	17 Unknown		35	54.89	Season Viers
18	18 West		50	479.04	Juan Luker
19	19 Unknown		42	1,879.06	Shan texeria

Queries [5]

Customers

Sales

Product Information

Sales detail

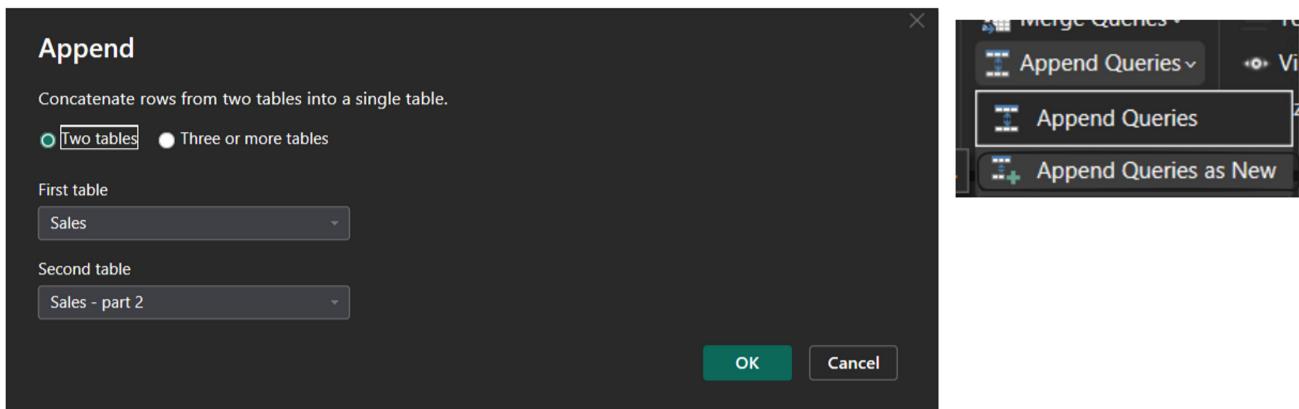
Sales - part 2

```
= Table.SelectRows(#"Changed Type", each [Customer ID] <> null and [Customer ID] <> "")
```

	Customer ID	Customer Region	Customer Age	Sales
1	21 North		10	6,532.00
2	22 North		23	7,958.00
3	23 South		56	1,012.00
4	24 East		38	5,816.00
5	25 East		22	5,651.00
6	26 West		22	3,215.00
7	27 West		58	8,314.00
8	28 East		11	3,575.00
9	29 East		43	7,459.00
10	30 North		45	1,607.00
11	31 North		14	6,078.00
12	32 South		27	9,065.00
13	33 East		57	4,423.00
14	34 East		17	850.00
15	35 East		53	3,701.00
16	36 East		59	5,873.00
17	37 North		49	4,385.00
18	38 North		33	9,561.00
19	39 South		38	2,322.00

Are they
both
Identical??

Disclaimer - In order to apply Append Query try to remove Customer Name from the Sales Data to make both table "Sales" & "Sales - p2" identical to apply append queries in vertical direction.



Queries [6]

	Customer ID	Customer Region	Customer Age	Sales
1	1 North	100%	49	309.38
2	2 Unknown	0%	10	57.71
3	3 North	0%	47	686.95
4	4 South	0%	11	1,619.19
5	5 East	0%	17	1,005.90
6	6 Unknown	0%	23	174.65
7	7 Unknown	0%	21	825.00
8	8 Unknown	0%	27	250.00
9	9 Unknown	0%	44	9.03
10	10 Unknown	0%	10	255.84
11	11 North	0%	29	151.24
12	12 South	0%	29	251.72
13	13 South	0%	42	1,139.43
14	14 North	0%	12	575.36
15	15 Unknown	0%	45	18.06
16	16 Unknown	0%	30	299.85
17	17 Unknown	0%	35	54.89
18	18 West	0%	50	479.04
19	19 Unknown	0%	42	1,879.06

What is Data Model

Independant Tables

The screenshot shows the Power BI Data Model view with five separate tables listed:

- Customers**: Contains Customer ID and Customer Name.
- Product Information**: Contains Customer ID, Order Category, Price per item, Product bought, and Product ID.
- Sales**: Contains Customer Age, Customer ID, Customer Region, and Sales.
- Sales - part 2**: Contains Customer Age, Customer ID, Customer Region, and Sales.
- Sales Record**: Contains Customer Age, Customer ID, Customer Region, and Sales.

Customer Name	Sum of Sales
Alberta Fabela	\$5,18,809.73
Betty scruggs	\$5,18,809.73
Debbie Orone	\$5,18,809.73
Elvin Catron	\$5,18,809.73
Era Gassner	\$5,18,809.73
Gisede Mariscal	\$5,18,809.73
Gustano Primm	\$5,18,809.73
Jacinda Moffett	\$5,18,809.73
Juan Luker	\$5,18,809.73
Karisa Oquwndo	\$5,18,809.73
Lacresha Whitty	\$5,18,809.73
Laticia Crimi	\$5,18,809.73
Loralee Widell	\$5,18,809.73
Octavio Ricci	\$5,18,809.73
Ricky Cheshire	\$5,18,809.73
Season Viers	\$5,18,809.73
Shan texeria	\$5,18,809.73
Sylvester Willbourn	\$5,18,809.73
Taylor Pogue	\$5,18,809.73
Yee Bordeau	\$5,18,809.73
Total	\$5,18,809.73

Build section (Suggestions):

- Matrix
- ... (other chart types)

Data section:

- Customers** (selected):
 - Customer ID
 - Customer Name (checked)
- Product Information**
- Sales detail**
- Sales Record** (selected):
 - Customer Age
 - Customer ID
 - Customer Region
 - Sales (checked)

New relationship

Select tables and columns that are related.

From table: Customers

Customer ID	Customer Name
1	Elvin Catron
2	Gustano Primm
3	Lacresha Whitt...

To table: Sales Record

Customer Age	Customer ID	Customer Re...
49	1	North
10	2	Unknown
47	3	North

Cardinality: One to one (1:1)

Cross-filter direction: Both

Make this relationship active

Assume referential integrity

Save **Cancel**

Customers

- Customer ID
- Customer Name

Sales Record

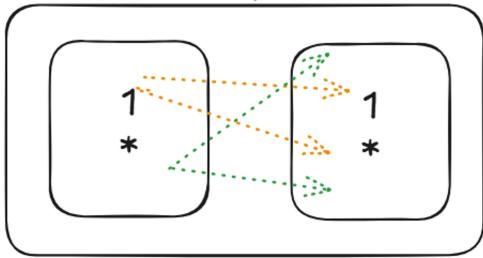
- Customer Age
- Customer ID
- Customer Region
- Sales

Cardinality: One to one (1:1)

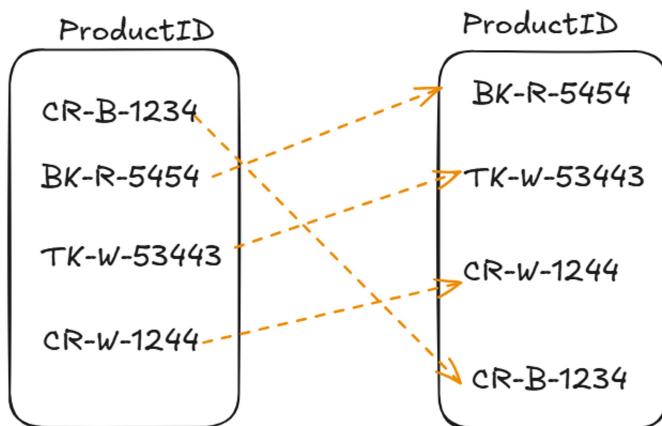
- Many to one (*:1)
- One to one (1:1)
- One to many (1:*)
- Many to many (*:*)

Customer Name	Sum of Sales
Alberta Fabela	\$5,08,681
Betty scruggs	\$9.03
Debbie Orone	\$174.65
Elvin Catron	\$18.06
Era Gassner	\$309.38
Grisede Mariscal	\$825
Gustano Primm	\$151.24
Jacinda Moffett	\$57.71
Juan Luker	\$1,005.9
Karisa Oquwndo	\$479.04
Lacresha Whitty	\$575.36
Laticia Crimi	\$686.95
Loralee Widell	\$250
Octavio Ricci	\$299.85
Ricky Cheshire	\$1,619.19
Season Viers	\$1,139.43
Shan texeria	\$54.89
Sylvester Willbourn	\$1,879.06
Taylor Pogue	\$255.84
Yee Bordeau	\$251.72
Total	\$5,18,809.73

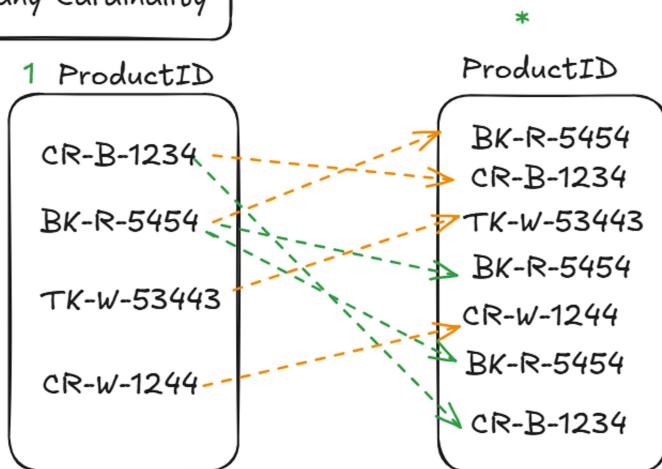
Why 4 Combination?

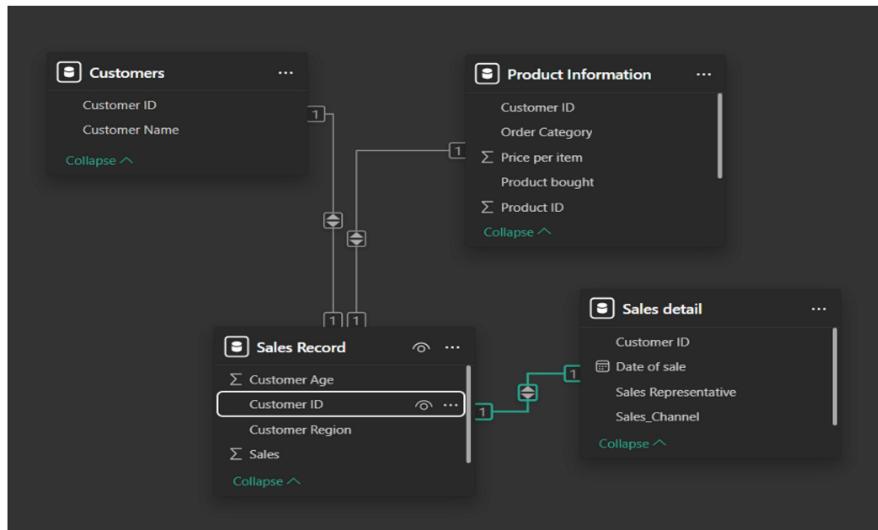
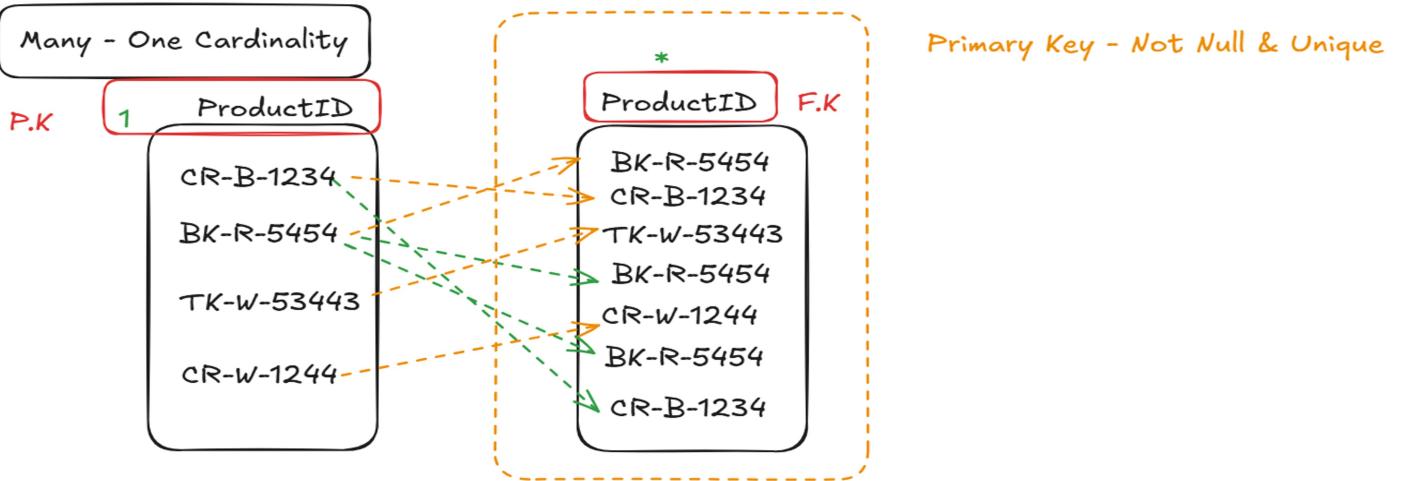


One - One Cardinality



One - Many Cardinality

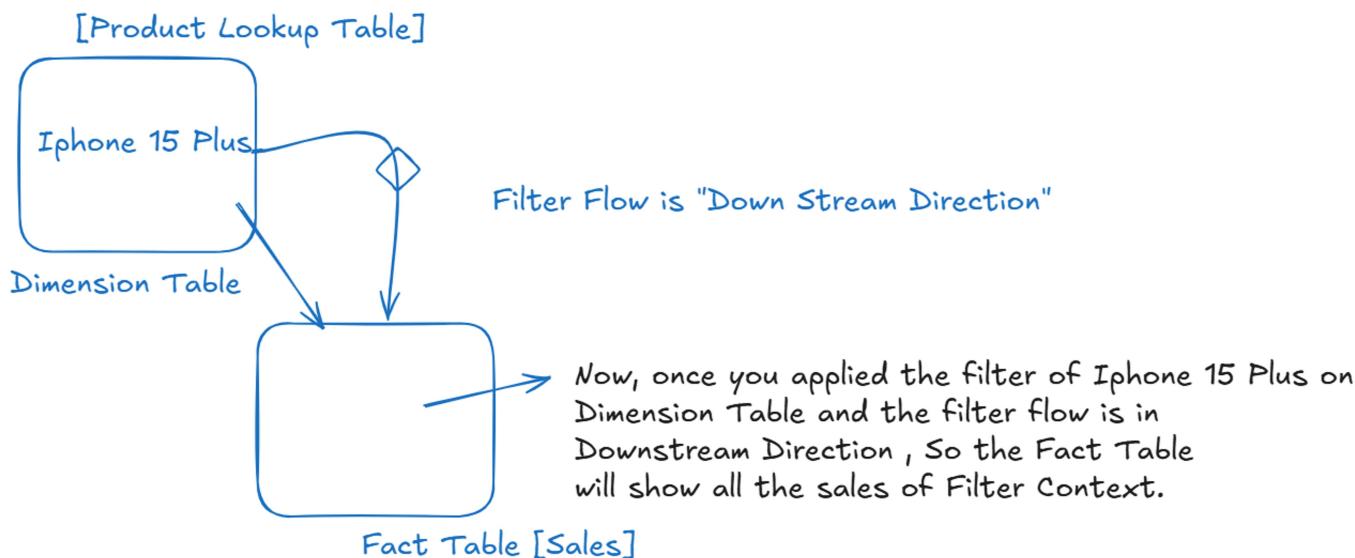
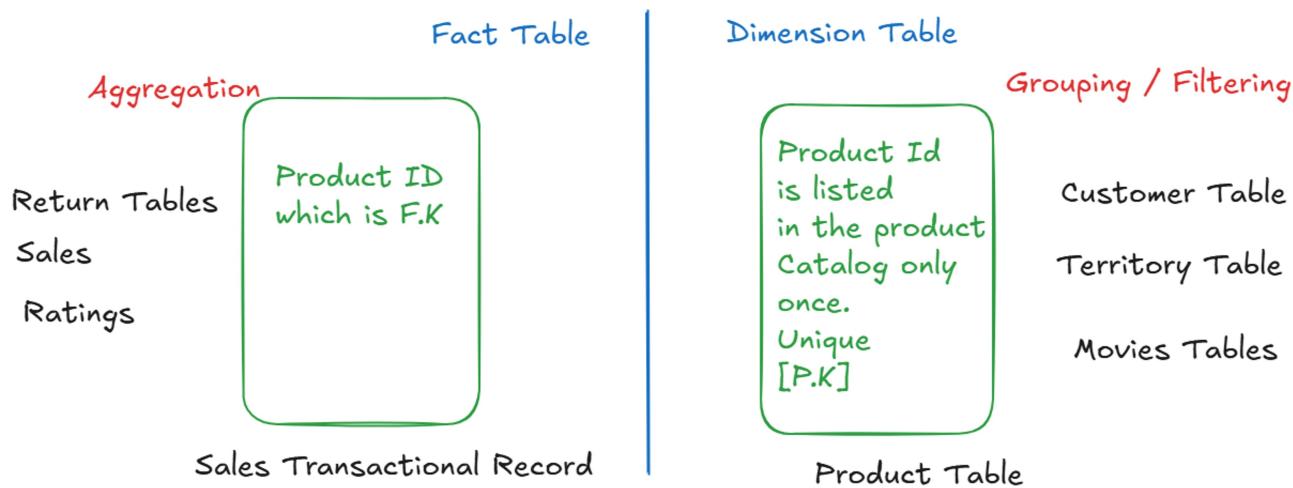




FACT & DIMENSION TABLES

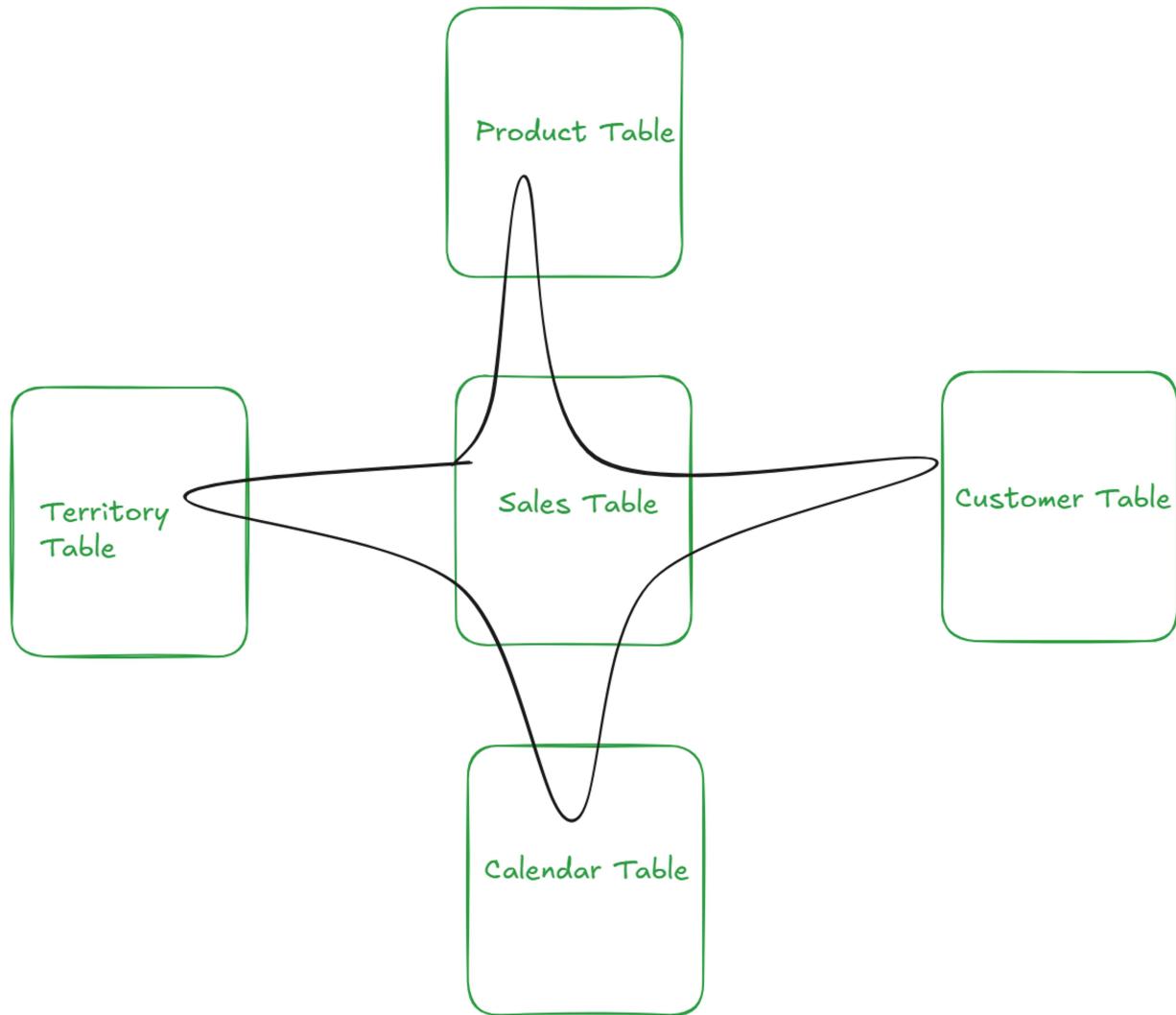
Data models generally contain two types of tables: fact ("data") tables, and dimension ("lookup") tables:

- Fact tables contain numerical values or metrics used for summarization (sales, orders, transactions, pageviews, etc.)
- Dimension tables contain descriptive attributes used for filtering or grouping (products, customers, dates, stores, etc.)



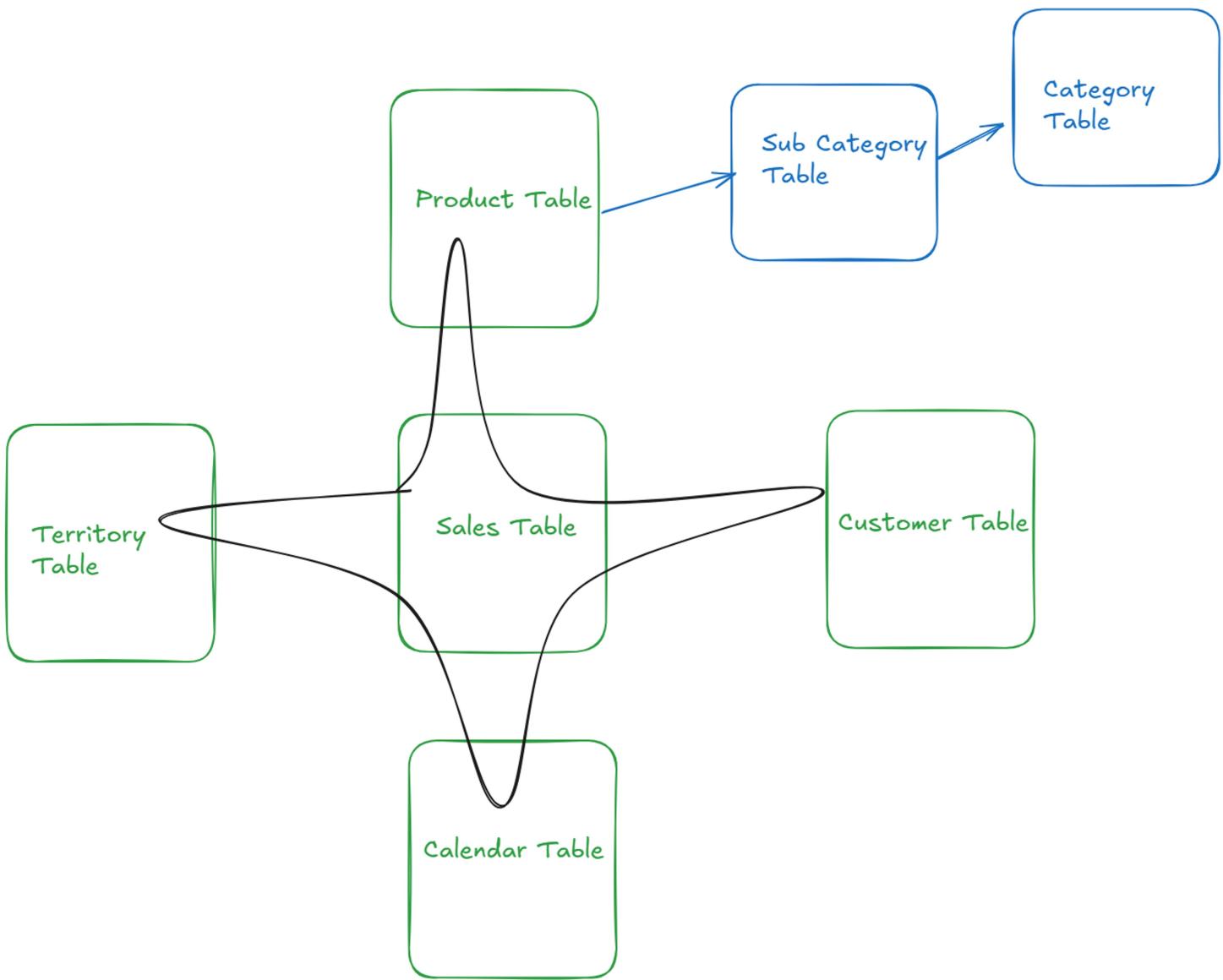
Star Schema

Star Schema is a relationship between atleast one Fact Table Connected with Multiple Dimension Table.



Snowflake Schema

Its an extended part of star schema where there is a sub-dimensional Table that won't directly connect with Fact table



Normalization

When you merge multiple Tables , you have to follow the rule of normalization. When you do merging of table using P.K & F.K, no duplicates values would be allowed.