

atomcamp

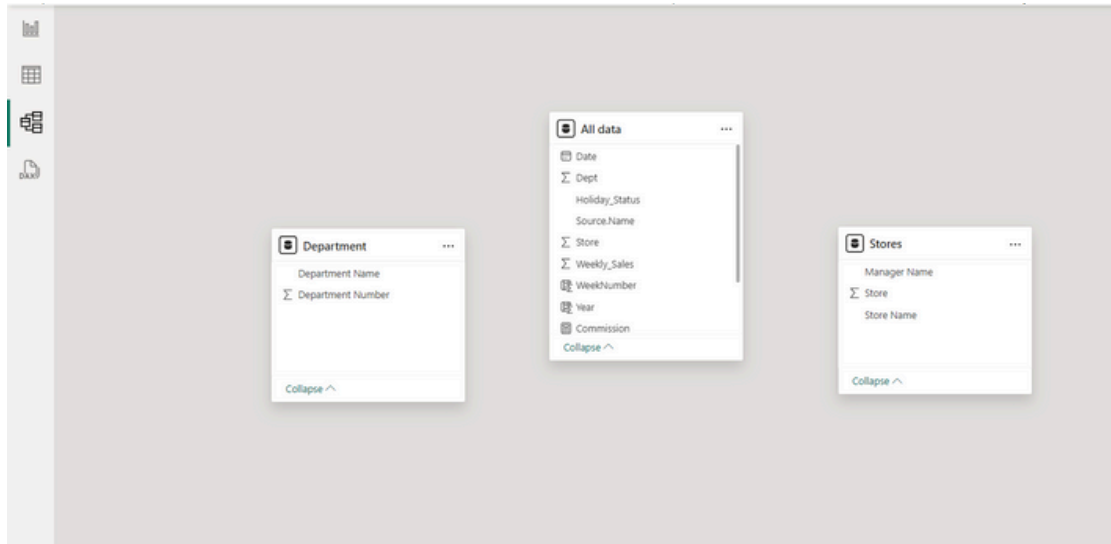
MICROSOFT POWER BI MANUAL FOR DESKTOP

Data Modeling



Creating a data model

What is a data model?



This IS NOT a data model

- This is a collection of independent tables, which share no connections or relationships
- If you tried to visualize Current Sale, this is what you'd get

Department Name	CurrentSale
Aquatics	4,448,332,866.70
Arts & Crafts	4,448,332,866.70
Astronomy	4,448,332,866.70
Automotive	4,448,332,866.70
Baby Essentials	4,448,332,866.70
Bakery	4,448,332,866.70
Beauty & Cosmetics	4,448,332,866.70
Bedding & Bath	4,448,332,866.70
Beverages	4,448,332,866.70
Bicycles & Accessories	4,448,332,866.70
Books & Magazines	4,448,332,866.70
Business Books	4,448,332,866.70
Camping & Outdoors	4,448,332,866.70
Cleaning Supplies	4,448,332,866.70
Collectibles	4,448,332,866.70
Commercial Kitchen	4,448,332,866.70
Dairy & Eggs	4,448,332,866.70
Total	4,448,332,866.70

Database Normalization

The tables are connected via relationships, based on a common field • Now Current Sales can be filtered using fields from the Lookup table



Department Name	CurrentSale
Aquatics	67,400,285.51
Arts & Crafts	21,773,750.43
Astronomy	66,565,374.40
Automotive	13,924,659.09
Baby Essentials	111,569,458.03
Bakery	53,932,127.36
Beauty & Cosmetics	128,912,417.47
Bedding & Bath	130,038,975.66
Beverages	139.06
Bicycles & Accessories	2,352,575.63
Books & Magazines	38,897,216.08
Business Books	197,663,588.55
Camping & Outdoors	19,446,709.84
Cleaning Supplies	6,182.87
Collectibles	195,216,073.03
Commercial Kitchen	299,461,722.85
Dairy & Eggs	5,236,160.15
Deli	10,211,053.43
Eco-Friendly Goods	31,292,100.42
Educational Supplies	19,747,611.13
Electronics	20,229,881.83
Total	4,448,332,866.70

Now this is a data model

Database normalization

Normalization

Normalization is the process of organizing the tables and columns in a relational database to reduce redundancy and preserve data integrity. It's commonly used to:

- Eliminate redundant data to decrease table sizes and improve processing speed & efficiency
- Minimize errors and anomalies from data modifications (inserting, updating or deleting records)
- Simplify queries and structure the database for meaningful analysis

Table: customers (1,004 rows) Column: Customer ID (1,000 distinct values)

Customer ID	Customer Name	Email	Phone Number	Address Line 1	City	Country	Postcode	Loyalty Card
00246-15080-LE	Francesco Dressel			1 Fulton Road	Toledo	United States	43605	No
00256-19905-YG	Stanislaus Valsler		+353 (479) 865-9222	95 Southridge Alley	Castlebridge	Ireland	R14	No
00445-42781-KX	Faunie Brigham	fbrigham@	+353 (620) 657-2946	7246 Green Pass	Castlereagh	Ireland	F45	Yes
00539-42510-RV	Nickey Youles	nyoules2@	+353 (641) 846-7654	12461 Dryden Pass	Edgeworthstown	Ireland	E25	Yes
00841-75330-ZV	Rori Ollin		+1 (626) 704-3749	05889 Heath Place	Pasadena	United States	91199	Yes
00852-54571-WP	De Drewitt	ddrewitt@	+1 (571) 504-1175	6 Tomscot Hill	Alexandria	United States	22313	Yes
00886-35803-FG	Edeline Edney		+1 (205) 866-7629	43 Crest Line Road	Birmingham	United States	35263	No
00888-74814-UZ	Theresita Newbury	tnewbury@	+353 (803) 587-0026	79526 Bultman Lane	Clonskeagh	Ireland	D04	No
01035-70465-UO	Anselma Attwater	aattwater@	+1 (434) 821-8618	72 Maryland Terrace	Charlottesville	United States	22908	Yes
01282-28364-RZ	Koressa O'Geneay	kogeneay@	+1 (303) 637-0326	77 Rigney Hill	Aurora	United States	80045	No
01297-94364-XH	Deana Staite	dstaite3@	+1 (713) 478-3937	39 Dunning Hill	Houston	United States	77281	No
01304-59807-OB	Nico Hubert		+1 (646) 228-3492	027 Village Avenue	New York City	United States	10155	Yes
01338-83217-GV	Mord Meriet	mmeriet@	+1 (701) 560-2604	56000 Kedzie Alley	Grand Forks	United States	58207	No
01433-04270-AX	Felecia Dodgson		+1 (973) 434-8662	48053 8th Crossing	Newark	United States	7104	Yes
01474-63436-TP	Read Cutts	rcutts@	+1 (815) 758-8653	820 Reinke Pass	Rockford	United States	61105	No
01603-43789-TN	Brice Romera	bromera@	+353 (640) 110-9801	2311 Eastlawn Plaza	Foxrock	Ireland	D04	Yes
01811-60350-CU	Sarette Ducarel		+1 (914) 915-4328	04922 Colorado Street	Staten Island	United States	10310	No
01841-48191-NL	Cam Jewster	cjewster@	+1 (937) 925-7390	24010 Sunnyside Drive	Dayton	United States	45426	Yes
01881-40815-VO	Faith Powley	fpowley@	+1 (504) 873-5980	3 Talisman Hill	New Orleans	United States	70174	Yes
01927-46702-YT	Nelie Garnson		+44 (141) 640-7113	821 Annamark Park	Merton	United Kingdom	SW19	No
01932-87052-KO	Leonie Cullford	lcullford@	+1 (530) 998-9789	71 Sycamore Crossing	Chico	United States	95973	Yes

Models that aren't normalized contain redundant, duplicate data. Selection of the first column shows that there are some duplicates present in the data.

Facts and dimensions table

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.)

Product ID	Coffee Type	Roast Type	Size	Unit Price	Price per 100g	Profit
A-L-0.2	Ara	L	0.2	3.885	1.9425	0.34965
A-L-0.5	Ara	L	0.5	7.77	1.554	0.6993
A-L-1	Ara	L	1	12.95	1.295	1.1655
A-L-2.5	Ara	L	2.5	29.785	1.1914	2.68065
A-M-0.2	Ara	M	0.2	3.375	1.6875	0.30375
A-M-0.5	Ara	M	0.5	6.75	1.35	0.6075
A-M-1	Ara	M	1	11.25	1.125	1.0125
A-M-2.5	Ara	M	2.5	25.875	1.035	2.32875
A-D-0.2	Ara	D	0.2	2.985	1.4925	0.26865
A-D-0.5	Ara	D	0.5	5.97	1.194	0.5373
A-D-1	Ara	D	1	9.95	0.995	0.8955
A-D-2.5	Ara	D	2.5	22.885	0.9154	2.05965
R-L-0.2	Rob	L	0.2	3.585	1.7925	0.2151
R-L-0.5	Rob	L	0.5	7.17	1.434	0.4302
R-L-1	Rob	L	1	11.95	1.195	0.717
R-L-2.5	Rob	L	2.5	27.485	1.0994	1.6491
R-M-0.2	Rob	M	0.2	2.985	1.4925	0.1791
R-M-0.5	Rob	M	0.5	5.97	1.194	0.3582
R-M-1	Rob	M	1	9.95	0.995	0.597
R-M-2.5	Rob	M	2.5	22.885	0.9154	1.3731
R-D-0.2	Rob	D	0.2	2.685	1.3425	0.1611

products (48 rows)

This **Fact table** contains quantity values, along with product id and other fields

date	day_of_month	month	year	weekday	week_of_year	week_ending	month_name	quarter
1/1/1997	1	1	1997	Wednesday	1	1/5/1997	January	Q1
1/2/1997	2	1	1997	Thursday	1	1/5/1997	January	Q1
1/3/1997	3	1	1997	Friday	1	1/5/1997	January	Q1
1/4/1997	4	1	1997	Saturday	1	1/5/1997	January	Q1
1/5/1997	5	1	1997	Sunday	2	1/5/1997	January	Q1
1/6/1997	6	1	1997	Monday	2	1/12/1997	January	Q1

This **Calendar Lookup** table contains attributes about each date (month, year, quarter, etc.)

Primary Keys and foreign keys

Order ID	Order Date	Customer ID	Product ID
QEV-37451-860	Thursday, September 5, 2019	17670-51384-MA	R-M-1
KAC-83089-793	Thursday, July 15, 2021	23806-46781-OU	E-M-1
KAC-83089-793	Thursday, July 15, 2021	23806-46781-OU	R-L-2.5
ITA-87418-783	Friday, May 22, 2020	39396-12890-PE	R-D-2.5
QUQ-90580-772	Saturday, February 6, 2021	77634-13918-GJ	L-M-0.2
XWC-20610-167	Thursday, July 7, 2022	08350-81623-TF	E-D-0.2
ULR-52653-960	Monday, October 4, 2021	04152-34436-IE	L-L-2.5
HPI-42308-142	Saturday, April 11, 2020	06631-86963-XP	E-M-0.5
DJH-05202-380	Thursday, November 7, 2019	85589-17020-CX	E-M-2.5
VMW-26889-781	Friday, June 28, 2019	36078-91009-WU	A-L-0.2
XKB-41924-202	Friday, August 19, 2022	45089-52817-WN	L-D-0.5
YHV-68700-050	Wednesday, September 11, 2019	26333-67911-OL	L-L-2.5
ANM-16388-634	Friday, November 19, 2021	77343-52608-FF	L-L-0.2
GAZ-58626-277	Monday, January 4, 2021	69533-84907-FA	L-L-0.2
RPW-36123-215	Monday, July 19, 2021	80640-45811-LB	E-L-0.5
QLM-07145-668	Thursday, May 9, 2019	86437-17399-FK	E-D-0.2
EUH-08089-954	Sunday, December 6, 2020	29050-93691-TS	A-D-0.2
BLD-12227-251	Thursday, March 4, 2021	64395-74865-WF	A-M-0.5
FJQ-60035-234	Monday, April 5, 2021	08847-29858-HN	A-L-0.2
AJL-52941-018	Tuesday, May 5, 2020	55871-61935-MF	E-D-1

These are **primary keys (PK)** They uniquely identify each row of the table, and relate to foreign keys in fact tables

These are foreign keys (FK)
They contain multiple instances of each value, and relate to primary keys in dimension tables

Customer ID	Customer Name	Email	Phone Number
00246-15080-LE	Francesco Dressel		
00256-19905-YG	Stanislaus Valsler		+353 (479) 865-9222
00445-42781-KX	Faunie Brigham	fbrighamhg@blog.com	+353 (620) 657-2946
00539-42510-RY	Nickey Youles	nyoules2t@reference.com	+353 (641) 846-7654
00841-75330-ZV	Rori Ollin		+1 (626) 704-3749
00852-54571-WP	De Drewitt	ddrewittnf@mapquest.com	+1 (571) 504-1175
00886-35803-FG	Edeline Edney		+1 (205) 866-7629
00888-74814-UZ	Theresita Newbury	tnewburys@usda.gov	+353 (803) 587-0026
01035-70465-UO	Anselma Attwater	aattwater5u@wikia.com	+1 (434) 821-8618
01282-28364-RZ	Koressa O'Geneay	kogeneayrd@utexas.edu	+1 (303) 637-0326
01297-94364-XH	Deana Staite	dstaite3z@scientificamerican.com	+1 (713) 478-3937
01304-59807-OB	Nico Hubert		+1 (646) 228-3492
01338-83217-GV	Mord Meriet	mmeriet56@noaa.gov	+1 (701) 560-2604
01433-04270-AX	Felecia Dodgson		+1 (973) 434-8662
01474-63436-TP	Read Cutts	rcuttspsy@techcrunch.com	+1 (815) 758-8653
01603-43789-TN	Brice Romera	bromeramj@list-manage.com	+353 (640) 110-9801
01811-60350-CU	Sarette Ducarel		+1 (914) 915-4328
01841-48191-NL	Cam Jewster	cjewsterlu@moonfruit.com	+1 (937) 925-7390
01881-40815-VO	Faith Powley	fpowleybp@dyndns.org	+1 (504) 873-5980
01927-46702-YT	Nelie Garnson		+44 (141) 640-7113
01933-87053-KO	Leenie Gullford	lgullfordh0@vina.com	

Relationships VS Merged Tables

Table.Combine({"Changed Type", orders, products})

	Customer ID	Customer Name	Email	Phone Number	Address Line 1	City	Country	Postcode
1	17670-51384-MA	Aloisia Allner	aallner0@lulu.com	+1 (862) 817-0124	57999 Pepper Wood Alley	Paterson	United States	
2	73342-18763-UW	Piotr Bote	pbote1@yelp.com	+353 (913) 396-4653	2112 Ridgeway Hill	Crumlin	Ireland	D6W
3	21125-22134-PX	Jami Redholes	jredholes2@tmall.com	+1 (210) 986-6806	5214 Bartillon Park	San Antonio	United States	
4	71253-00052-RN	Dene Azema	dazema3@facebook.com	+1 (217) 418-0714	27 Maywood Place	Springfield	United States	
5	17670-51384-MA	Aloisia Allner	aallner0@lulu.com	+1 (862) 817-0124	57999 Pepper Wood Alley	Paterson	United States	
6	73342-18763-UW	Piotr Bote	pbote1@yelp.com	+353 (913) 396-4653	2112 Ridgeway Hill	Crumlin	Ireland	D6W
7	21125-22134-PX	Jami Redholes	jredholes2@tmall.com	+1 (210) 986-6806	5214 Bartillon Park	San Antonio	United States	
8	71253-00052-RN	Dene Azema	dazema3@facebook.com	+1 (217) 418-0714	27 Maywood Place	Springfield	United States	
9	23806-46781-OU	Christoffer O' Shea		353 (698) 362-9201	38980 Manitowish Junction	Cill Airne	Ireland	N41
10	86561-91660-RB	Beryle Cottier		+1 (570) 289-7473	2651 Stoughton Place	Scranton	United States	
11	65223-29612-CB	Shaylynn Lobe	slobe6@nifty.com	+1 (937) 954-4541	7005 Mariners Cove Place	Dayton	United States	
12	21134-81676-FR	Melvin Wharfe		+353 (507) 574-3034	7 Straubel Road	Kill	Ireland	P24
13	03396-68805-ZC	Guthrey Petracci	gpetracci8@livejournal.com	+1 (310) 868-1842	949 Paget Parkway	Los Angeles	United States	
14	61021-27840-ZN	Rodger Raven	rraven9@ed.gov	+1 (213) 263-0288	1 Reinke Avenue	Los Angeles	United States	
15	76239-90137-UQ	Ferrell Ferber	fferbera@businesswire.com	+1 (408) 383-5302	68 High Crossing Court	San Jose	United States	
16	49315-21985-BB	Duky Phizackerly	dphizackerlyb@utexas.edu	+1 (408) 533-6012	28643 Bluejay Crossing	San Jose	United States	
17	34136-36674-OM	Rosaleen Scholar	rscholarc@nyu.edu	+1 (804) 420-0420	80915 Montana Park	Richmond	United States	
18	39396-12890-PE	Terence Vanyutin	tvanyutin@wix.com		331 Bunting Hill	Migrate	United States	
19	95875-73336-RG	Patrice Trobe	ptrobee@wunderground.com	+1 (314) 240-7896	827 Declaration Plaza	Saint Louis	United States	
20	25473-43727-BY	Llywelyn Oscroft	loscroftf@ebay.co.uk		022 Roth Place	Philadelphia	United States	
21	99643-51048-IQ	Minni Alabaster	malabasterg@hexun.com	+1 (971) 483-6255	3 Charing Cross Trail	Portland	United States	

Merging tables creates redundancy and often requires significantly more memory and processing power to analyze compared to a relational model with multiple small tables

The model view

Menu Ribbon

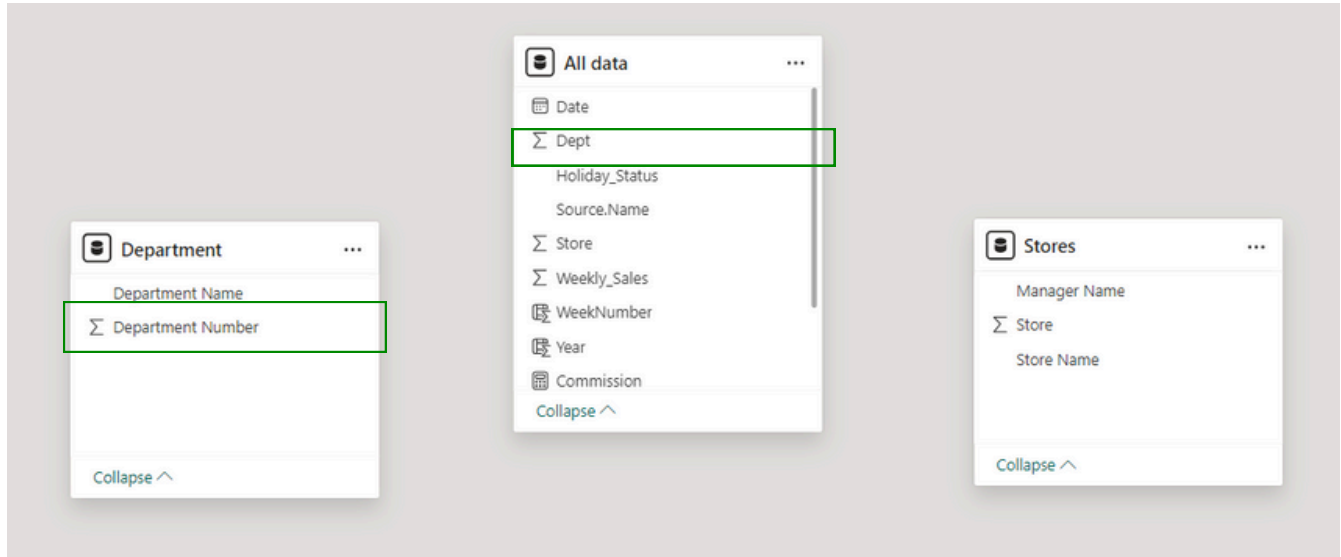
Model Canvas

Properties Pane

View Option

Creating Table Relationships

OPTION 1: Click and drag to connect primary and foreign keys within the **Model view**



OPTION 2: Add or detect relationships using the **Manage Relationships** dialog box

Edit relationship

Select tables and columns that are related.

From table
orders

Coffee Type	Country	Customer ID	Customer Na...	Email	Order Date	Order
null	null	17670-51384-...	null	null	Thursday, Sep...	QEV-
null	null	23806-46781-...	null	null	Thursday, July...	KAC-t
null	null	23806-46781-...	null	null	Thursday, July...	KAC-t

To table
customers

Address Line 1	City	Country	Customer ID	Customer Na...	Email	Loyali
2651 Stought...	Scranton	United States	86561-91660-...	Beryle Cottier	null	No
4 Mitchell Drive	Punta Gorda	United States	07591-92789-...	Culley Farris	null	Yes
95 Straubel Hill	Aurora	United States	51427-89175-...	Adrian Swaine	null	No

Cardinality
Many to many (**:*)

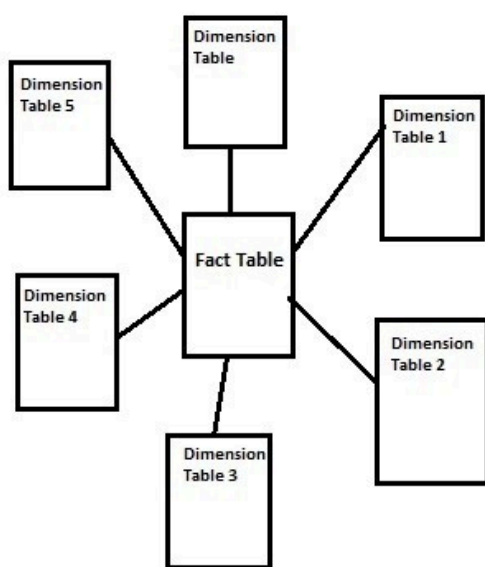
Cross-filter direction
Both

☒ Make this relationship active

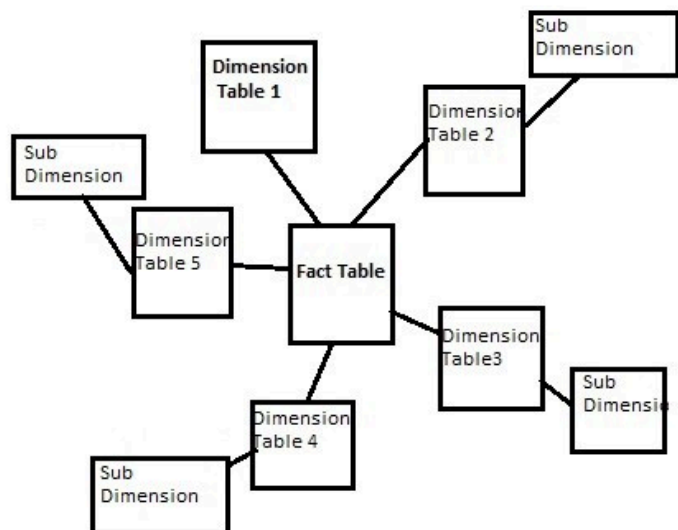
☐ Apply security filter in both directions

Editing tools allow you to activate or deactivate relationships and manage cardinality and filter direction

Star and snowflake schema



Star Schema



Snow Flake Schema

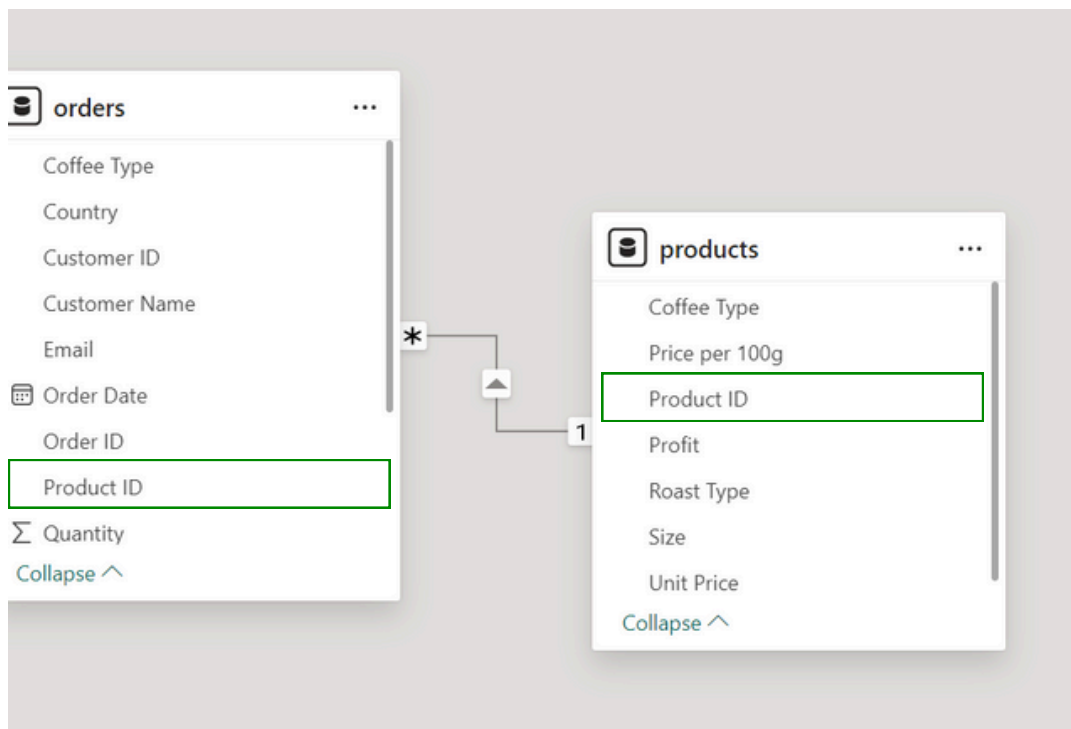
A **star schema** is the simplest and most common type of data model, characterized by a single fact table surrounded by related dimension tables

A **snowflake schema** is an extension of a star, and includes relationships between dimension tables and related sub-dimension tables

Relationship cardinality

Cardinality refers to the uniqueness of values in a column

- Ideally, all relationships in the data model should follow a one-to-many cardinality: one instance of each primary key, and many instances of each foreign key

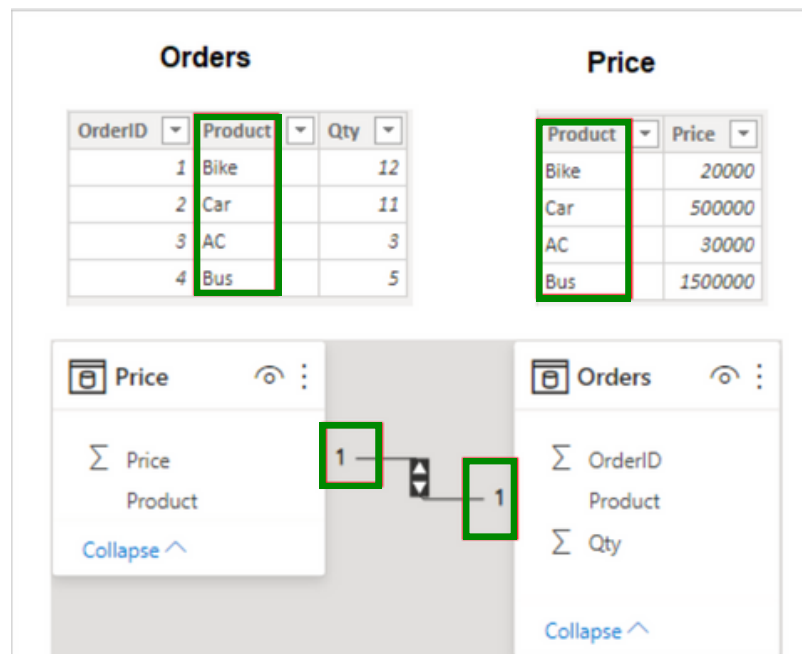


In this example there is only ONE instance of each Product ID in the Products table, since each row contains attributes of a single product. There are MANY instances of each Product ID in the orders table (noted by an asterisk *), since there are multiple sales for each product.

One-one cardinality

In a one-to-one relationship, the column in one table has only one instance of a particular value, and the other related table has only one instance of a particular value.

Both table contains only one instance of a value

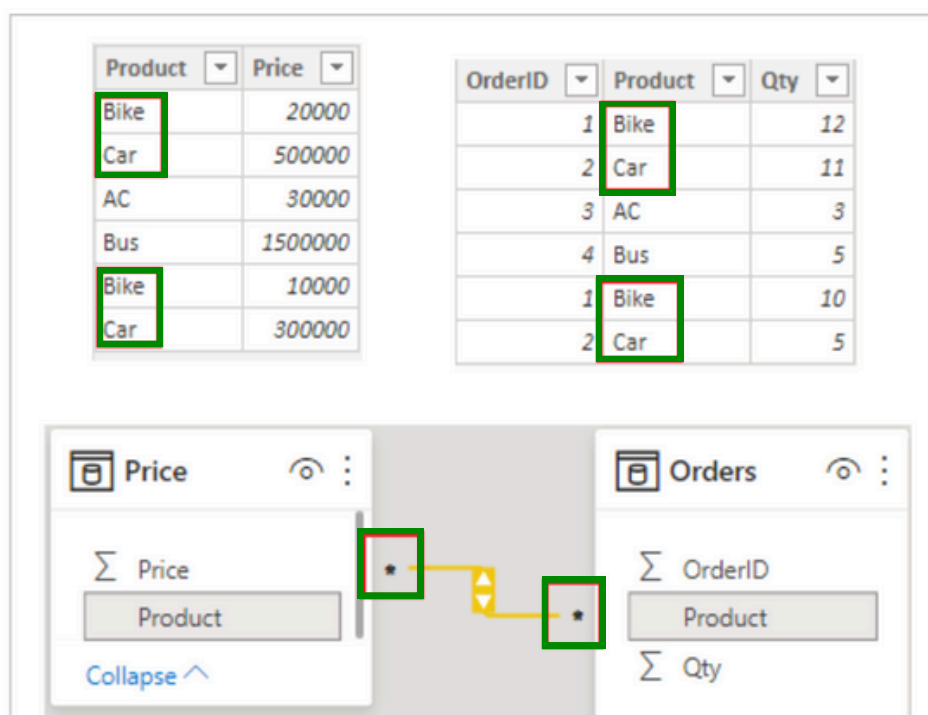


Many-many cardinality

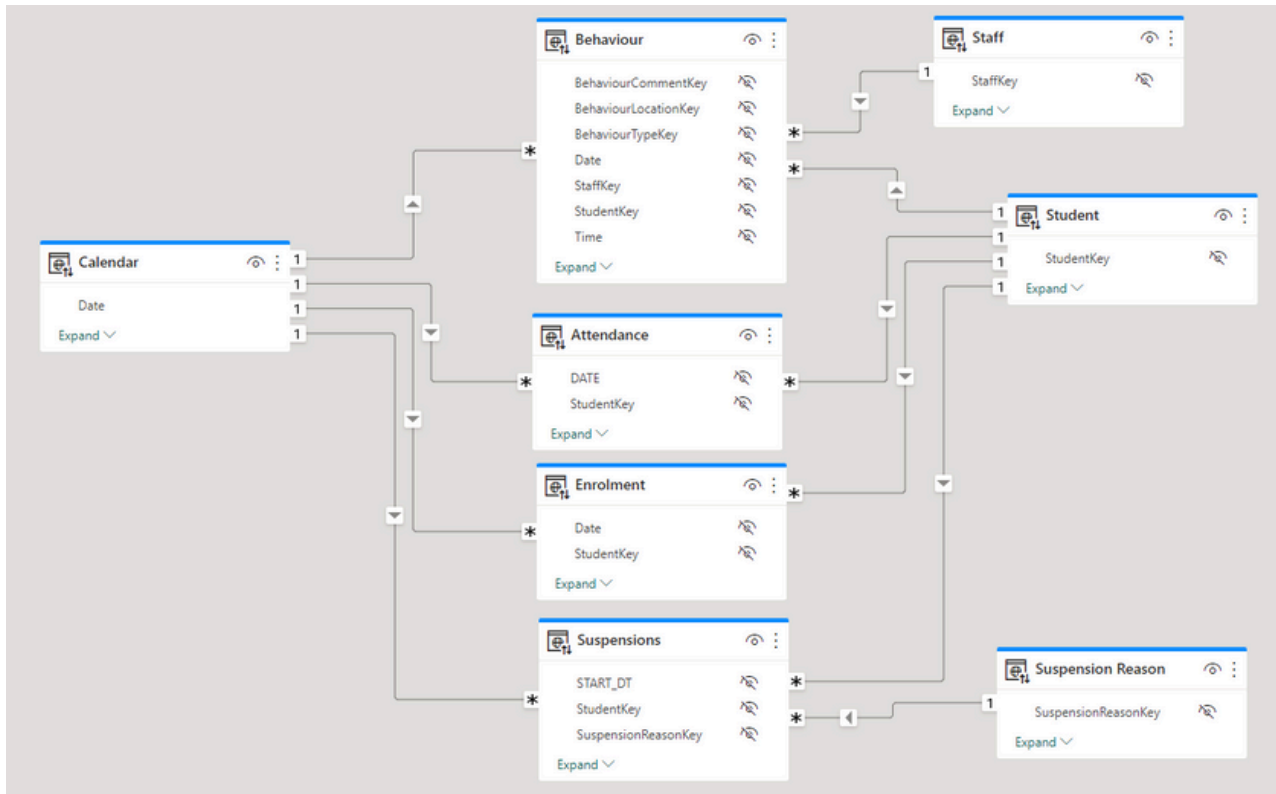
With composite models, you can establish a many-to-many relationship between tables, which removes requirements for unique values in tables.

Composite model: Allows two or more data connections, including DirectQuery connections or Import mode.

* : Both table contains more than one instance of a value.



Connecting with multiple fact tables



There are four fact tables in this model

- Behaviour
- Attendance
- Enrolment
- Suspensions

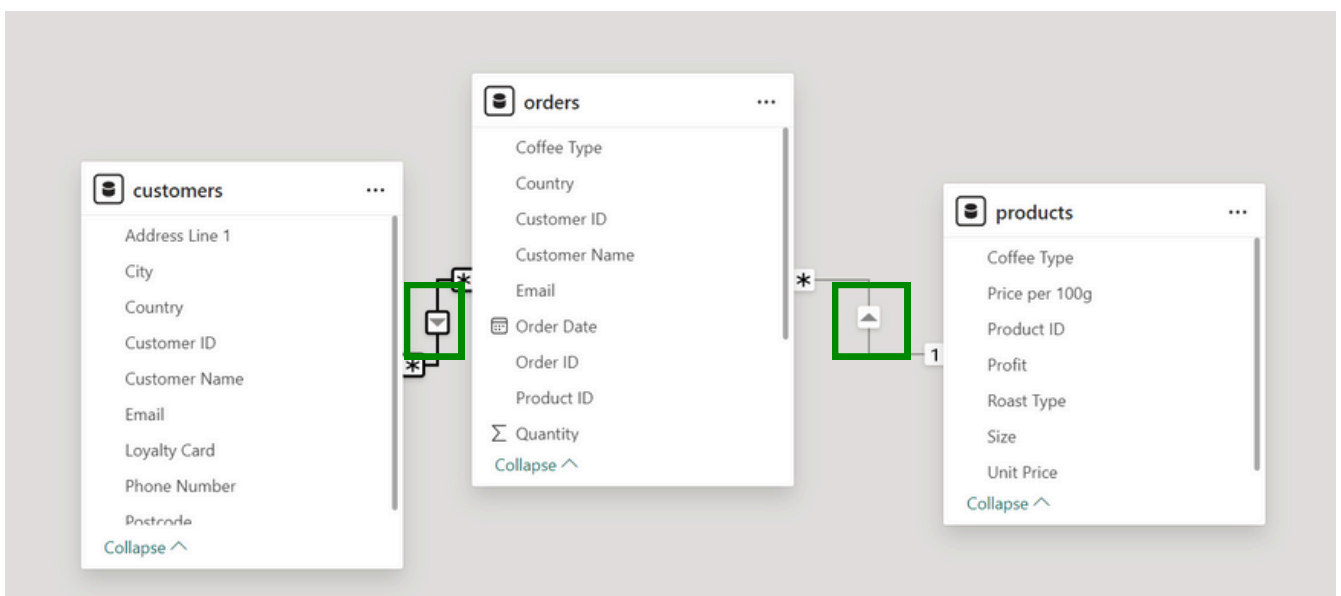
Connecting multiple fact tables in a data model is a common scenario in data warehousing and BI. This approach is often used to analyze different types of events or transactions that share common dimensions.

Filter Context and Flow

Here we have two data tables , connected to Territory Lookup

The arrows show the filter direction, and point from the one (1) side of the relationship to the many (*) side

- When you filter a table, that filter context is passed to any related “downstream” tables, following the arrow’s direction
- Filter context CANNOT flow “upstream”



Bi directional filter

Updating the cross-filter direction from Single to Both allows filter context to flow in either direction

The screenshot shows the Power BI interface with the relationship between customers and orders tables. The cross-filter direction is set to 'Both'. The Properties pane on the right shows the relationship details, including the cross-filter direction dropdown menu.

Properties

Relationship

Table	Column
orders	Customer ID

Cardinality: One to one (1:1)

Table	Column
customers	Customer ID

Make this relationship active: ☒ Yes

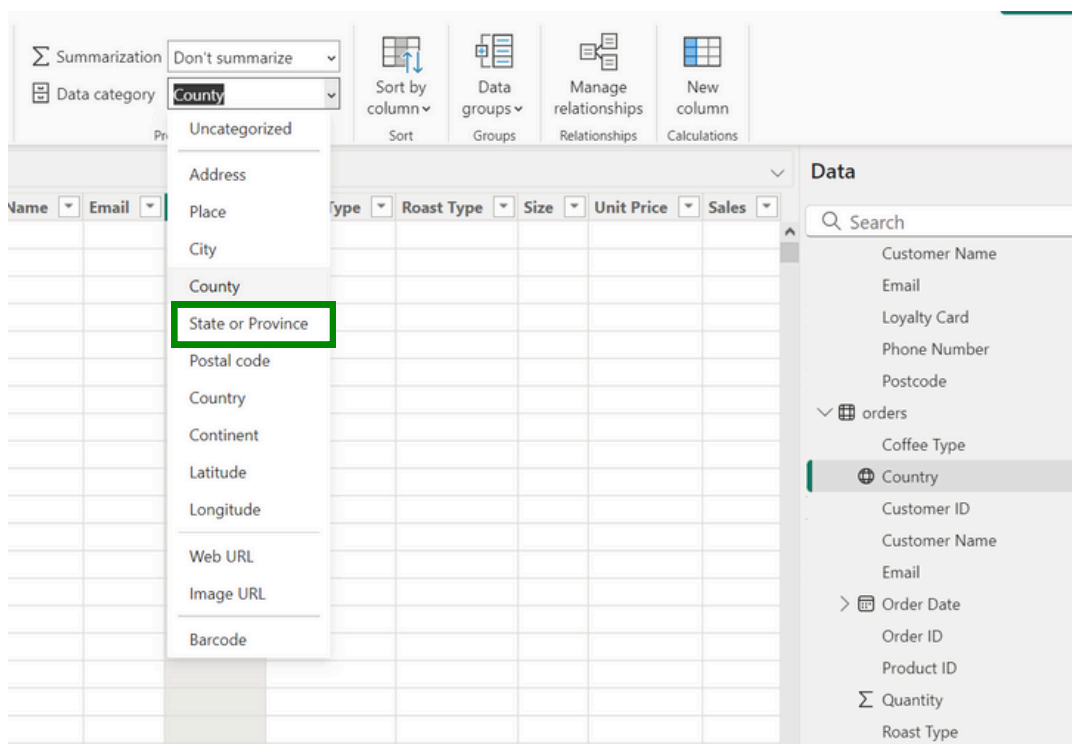
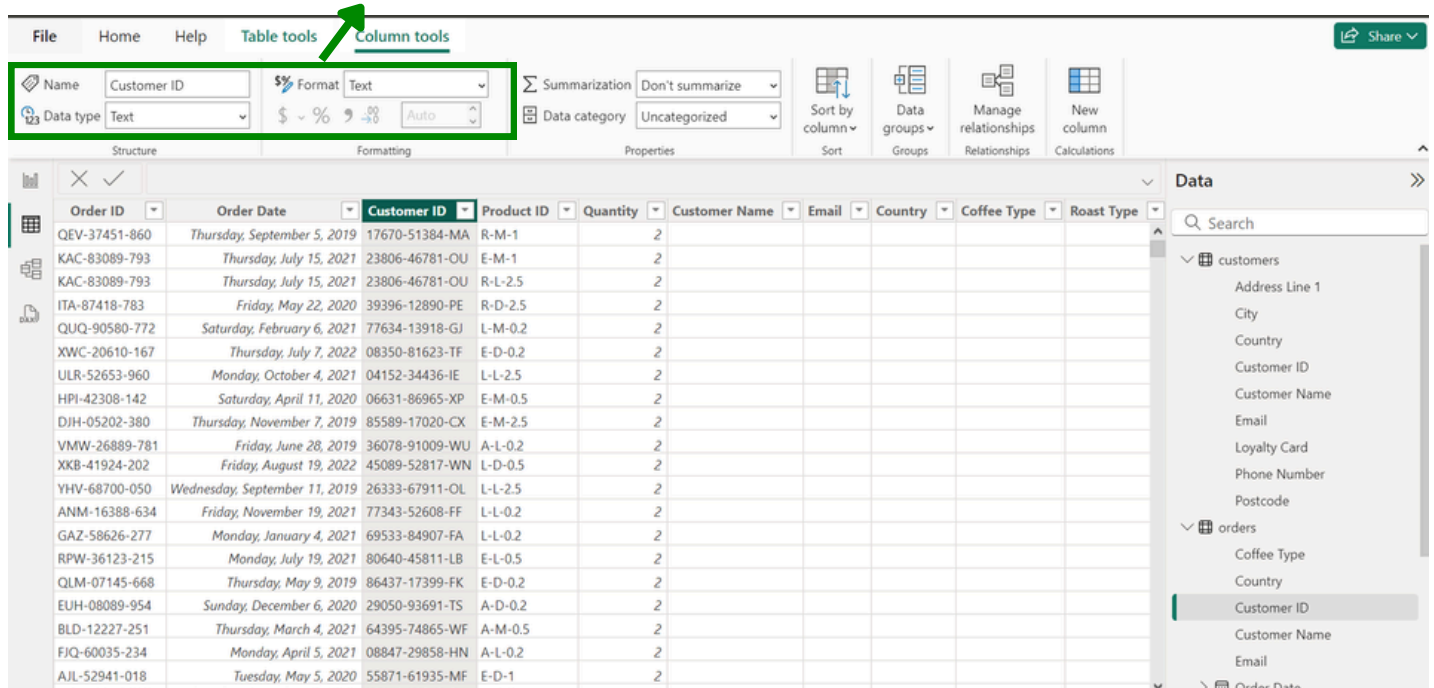
Cross-filter direction: **Both**

Apply changes

Open relationship editor

Data Formats and categories

Customize **data formats** from the Column tools menu in the **Data** view or the Properties pane in the **Model** view

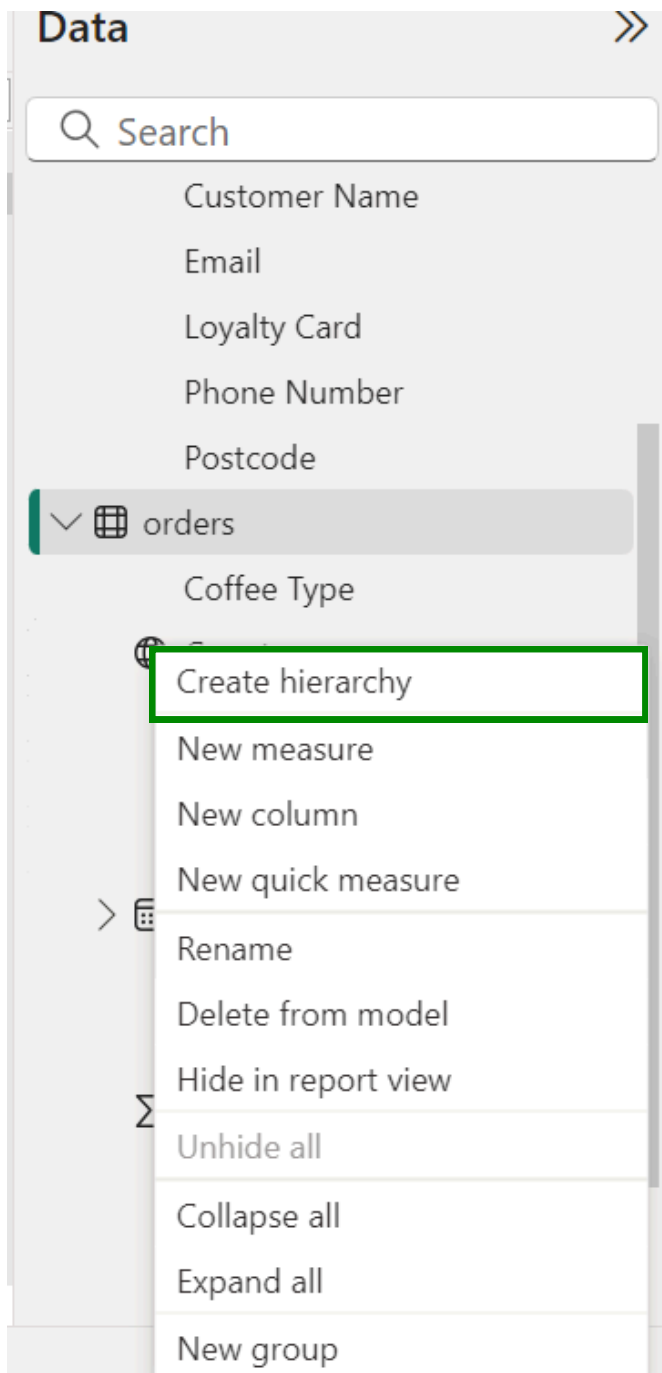


Assign data categories for geospatial fields, URLs or barcodes. This is commonly used to help Power BI map location-based fields like addresses, countries, cities, coordinates, zip codes, etc

Hierarchies

Hierarchies are groups of columns that reflect multiple levels of granularity

- For example, a Geography hierarchy might include **Country, State and City fields**
- Hierarchies are treated as a single item in tables and reports, allowing users to “drill up” and “drill down” through each level



In the Data pane, right-click a field and select Create hierarchy. You can keep on adding hierarchy till you reach your desired outcome.

For example in case of a territory you will add continent, then country, city and area and so on.