

## Power BI - Data Modelling - p2

SumProduct SUMX

Order Quantity \* Product Price

Order Qty	Prod Price
2	100
3	200
5	400
2	100
1	400
3	300
6	200

Item Sales
200
600
2000
200
400
900
1200

Product

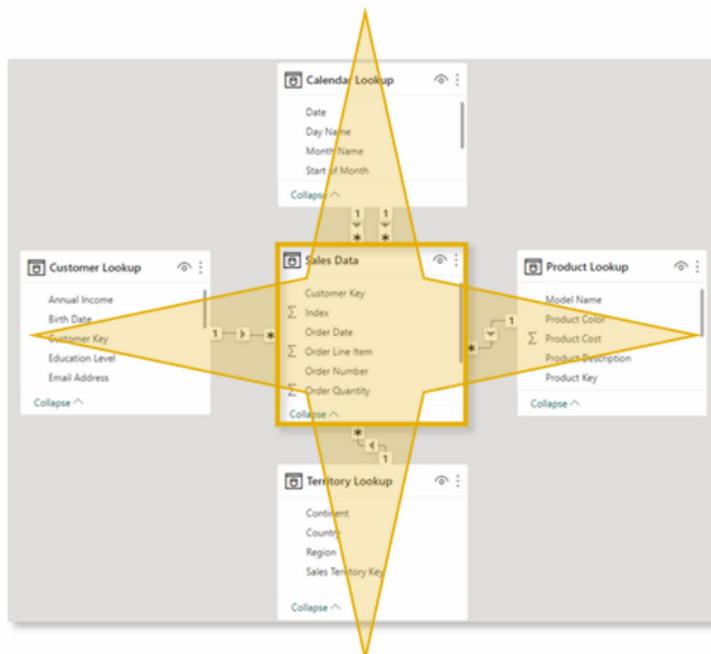
$$(2*100) + (3 * 200) + (5 * 400) + (2 * 100) + (1 * 400) + (3 * 300) + (6 * 200) = 5500$$

Revenue

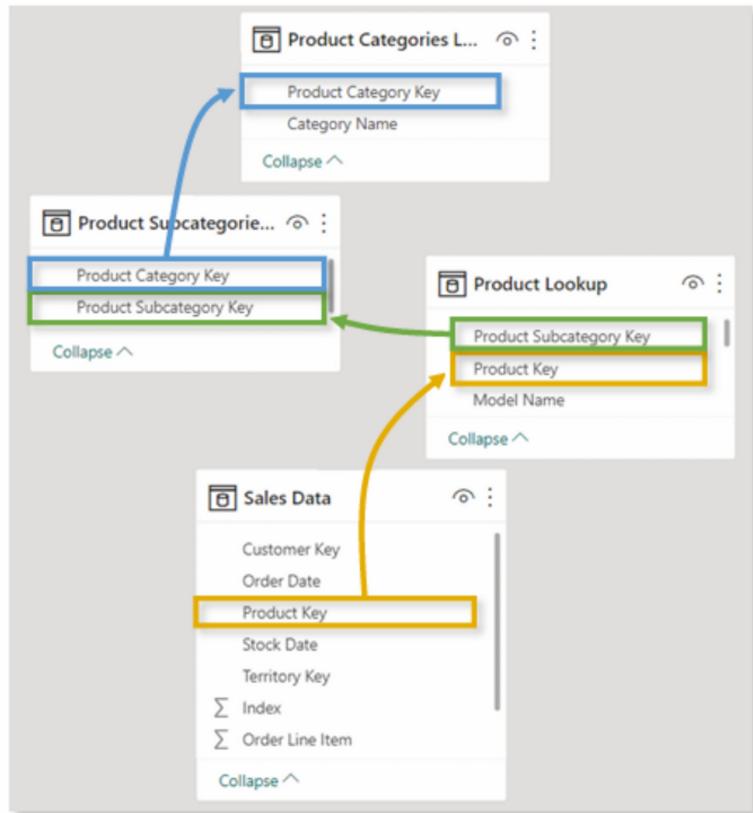
5500

Sum

## STAR & SNOWFLAKE SCHEMAS



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

Create all model view by dropping Fact & Dimension Table and add related tables around it.

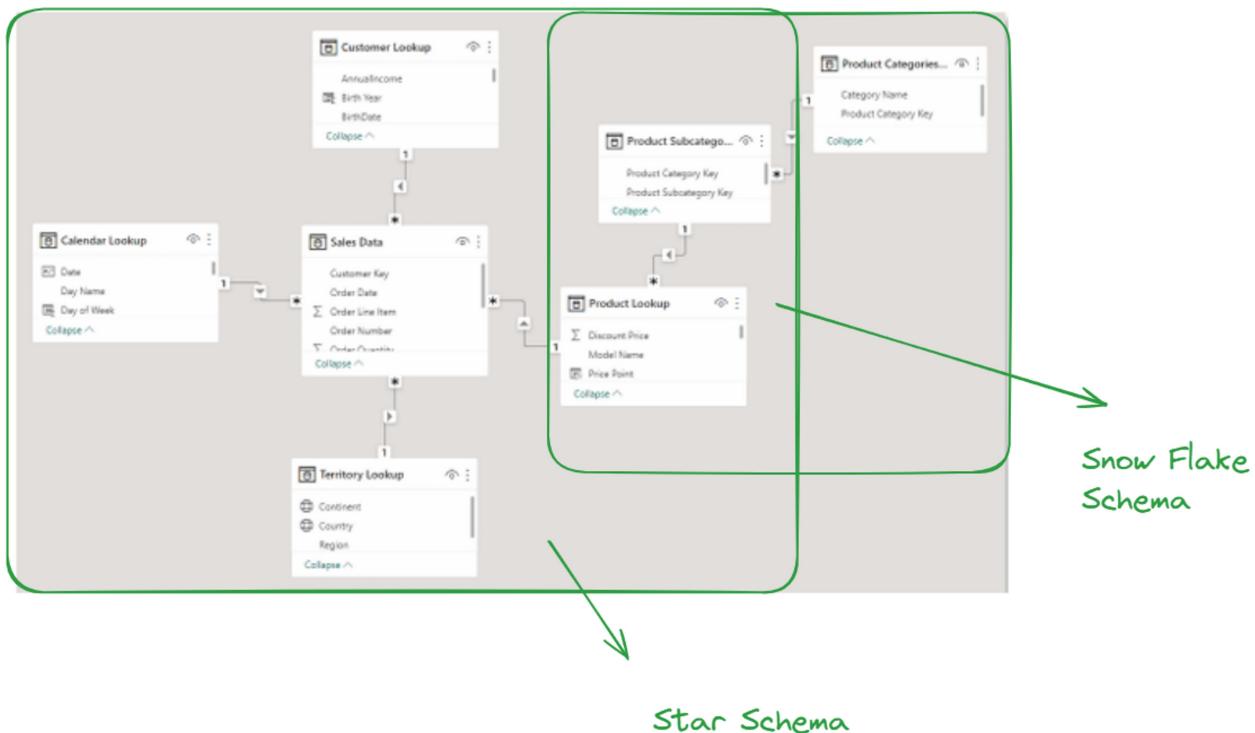
All tables	Sales Model View	Return Model View	Customer Model View	Calendar Model View	Product Model View	Territory Model View	<a href="#">+</a>
------------	------------------	-------------------	---------------------	---------------------	--------------------	----------------------	-------------------

Rename it well to understand.

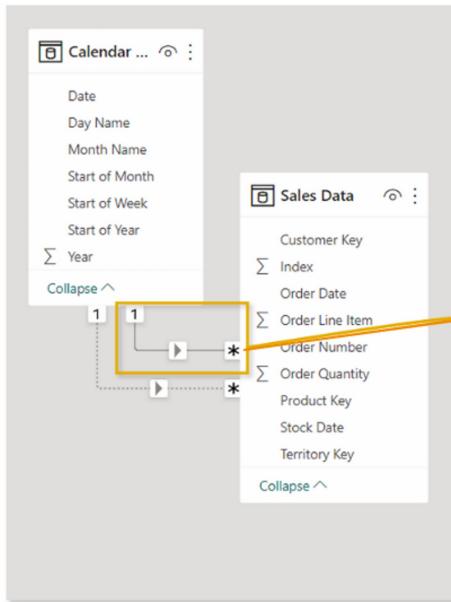
## ASSIGNMENT: TABLE RELATIONSHIPS

1. Delete all existing table relationships.
2. Create a star schema by creating relationships between the Sales, Calendar, Customer, Product and Territories tables.
3. Connect all three product tables (Product, Subcategory, Category) in a snowflake schema.
4. Use the matrix visual to confirm that you can filter Order Quantity values using fields from each dimension table.

Solution Preview



## ACTIVE & INACTIVE RELATIONSHIPS



Edit relationship

Select tables and columns that are related.

Sales Data							
Order Date	Stock Date	Order Number	Product Key	Customer Key	Territory Key	Order Line Item	C
7/3/2020	6/3/2020	SO46718	360	12570	9	1	
7/7/2020	4/22/2020	SO46736	360	12341	9	1	
7/12/2020	5/5/2020	SO46776	360	12356	9	1	

Calendar Lookup

Date	Day Name	Start of Week	Start of Month	Month Name	Start of Year	Year
1/1/2020	Wednesday	12/29/2019	1/1/2020	January	1/1/2020	2020
1/2/2020	Thursday	12/29/2019	1/1/2020	January	1/1/2020	2020
1/3/2020	Friday	12/29/2019	1/1/2020	January	1/1/2020	2020

Cardinality

Many to one (\*:1)

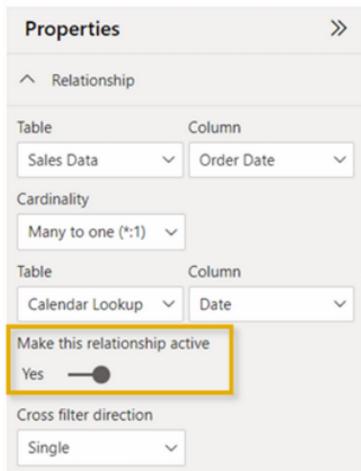
Cross filter direction

Single

Assume referential integrity

Make this relationship active

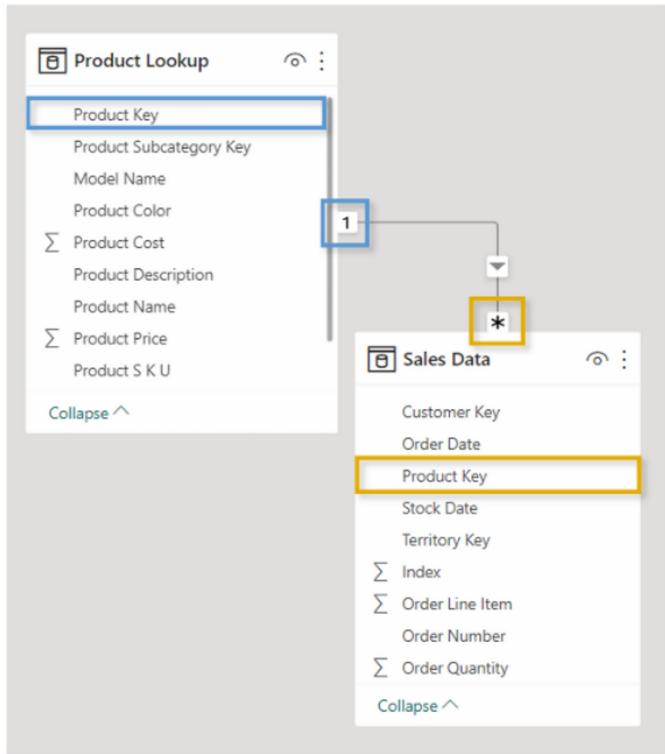
OK Cancel



The Sales Data table contains two date fields (Order Date & Stock Date), but there can only be one active relationship to the Date key in the Calendar table.

You can set relationships to active or inactive from either the Edit Relationships dialog box or the Properties (you must deactivate one before activating another)

## 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 Key in the Product table (noted by a "1"), since each row contains attributes of a single product (name, SKU, description, price, etc.)

There are MANY instances of each Product Key in the Sales table (noted by an asterisk \*), since there are multiple sales for each product

## ONE-TO-ONE CARDINALITY

Product Lookup

product_id	product_name	product_sku
4	Washington Cream Soda	64412155747
5	Washington Diet Soda	85561191439
7	Washington Diet Cola	20191444754
8	Washington Orange Juice	89770532250

Price Lookup

product_id	product_price
4	\$3.64
5	\$2.19
7	\$2.61
8	\$2.59

- Connecting the two tables above using `product_id` creates a one-to-one relationship, since each product ID only appears once in each table
- This isn't necessarily a "bad" relationship, but you can simplify the model by merging the tables into a single, valid dimension table

product_id	product_name	product_sku	product_price
4	Washington Cream Soda	64412155747	\$3.64
5	Washington Diet Soda	85561191439	\$2.19
7	Washington Diet Cola	20191444754	\$2.61
8	Washington Orange Juice	89770532250	\$2.59

NOTE: this still respects the rules of normalization, since all rows are unique and capture product-specific attributes

## MANY-TO-MANY CARDINALITY

Product Lookup

product_id	product_name	product_sku
4	Washington Cream Soda	64412155747
4	Washington Diet Cream Soda	81727382373
5	Washington Diet Soda	85561191439
7	Washington Diet Cola	20191444754
8	Washington Orange Juice	89770532250

Sales

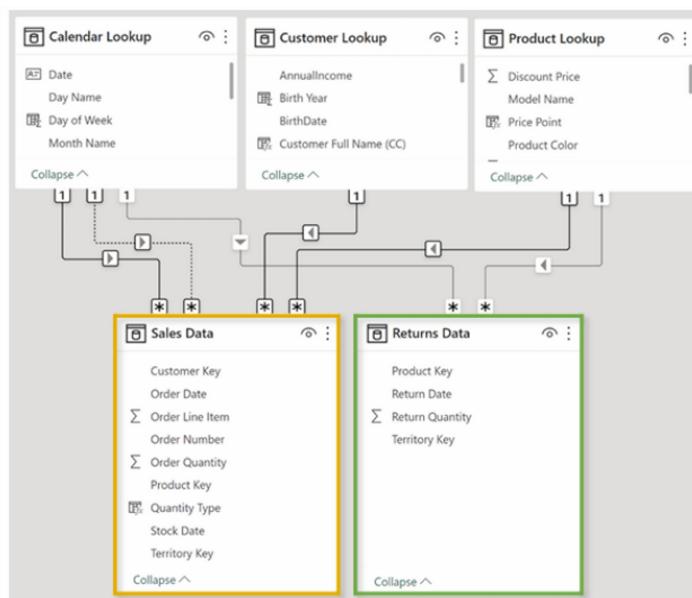
date	product_id	transactions
1/1/2017	4	12
1/2/2017	4	9
1/3/2017	4	11
1/1/2017	5	16
1/2/2017	5	19
1/1/2017	7	11



This relationship has cardinality Many-Many. This should only be used if it is expected that neither column (`product_id` and `product_id`) contains unique values, and that the significantly different behavior of Many-many relationships is understood. [Learn more](#)

- If we try to connect the tables above using `product_id`, we'll get a many-to-many relationship warning since there are multiple instances of `product_id` in both tables.
- Even if we force this relationship, how would we know which product was actually sold on each date – Cream Soda or Diet Cream Soda?

## CONNECTING MULTIPLE FACT TABLES



This model contains two fact tables:  
Sales Data and Returns Data

- Since there is no primary/foreign key relationship, we can't connect them directly to each other.
- But we can connect each fact table to related lookups, which allows us to filter both sales and returns data using fields from any shared lookup tables.
- We can view orders and returns by product since both tables relate to Product Lookup, but we can't view returns by customer since no relationship exists.

Generally speaking, fact tables should connect through shared dimension tables, not directly to each other

TerritoryKey	Sum of OrderQuantity	Sum of ReturnQuantity
1	12513	270
2	40	
3	30	
4	17191	362
5	49	1
6	10894	238
7	7862	186
8	7950	163
9	17951	404
10	9694	204
Total	84174	1828

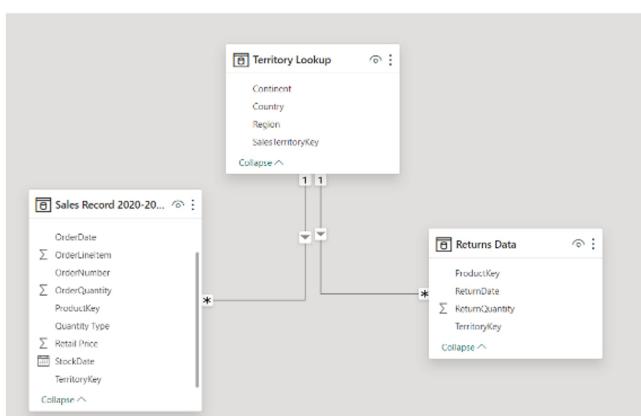
Territory Lookup

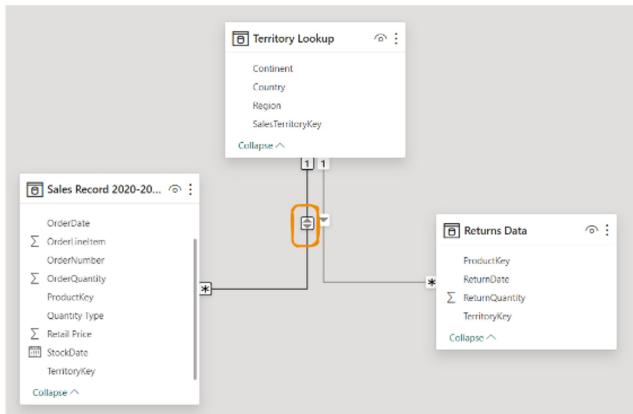
TerritoryKey	Sum of OrderQuantity	Sum of ReturnQuantity
1	12513	1828
2	40	1828
3	30	1828
4	17191	1828
5	49	1828
6	10894	1828
7	7862	1828
8	7950	1828
9	17951	1828
10	9694	1828
Total	84174	1828

Sales Record

TerritoryKey	Sum of OrderQuantity	Sum of ReturnQuantity
1	84174	270
4	84174	362
5	84174	1
6	84174	238
7	84174	186
8	84174	163
9	84174	404
10	84174	204
Total	84174	1828

Return Record





**Sales Record 2020-20...**

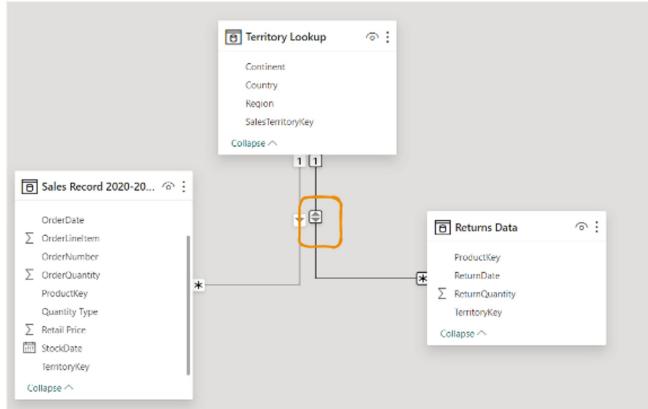
TerritoryKey	Sum of OrderQuantity	Sum of ReturnQuantity
1	12513	270
2	40	
3	30	
4	17191	362
5	49	1
6	10894	238
7	7862	186
8	7950	163
9	17951	404
10	9694	204
<b>Total</b>	<b>84174</b>	<b>1828</b>

**Returns Data**

TerritoryKey	Sum of OrderQuantity	Sum of ReturnQuantity
1	12513	270
2	40	
3	30	
4	17191	362
5	49	1
6	10894	238
7	7862	186
8	7950	163
9	17951	404
10	9694	204
<b>Total</b>	<b>84174</b>	<b>1828</b>

**Territory Lookup**

Continent	Country	Region	SalesTerritoryKey
-----------	---------	--------	-------------------



**Sales Record 2020-20...**

TerritoryKey	Sum of OrderQuantity	Sum of ReturnQuantity
1	12513	270
2	40	
3	30	
4	17191	362
5	49	1
6	10894	238
7	7862	186
8	7950	163
9	17951	404
10	9694	204
<b>Total</b>	<b>84174</b>	<b>1828</b>

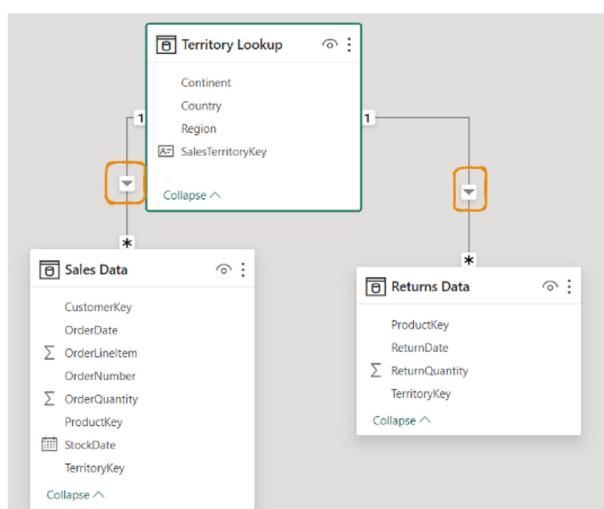
**Returns Data**

TerritoryKey	Sum of OrderQuantity	Sum of ReturnQuantity
1	12513	1828
2	40	1828
3	30	1828
4	17191	1828
5	49	1828
6	10894	1828
7	7862	1828
8	7950	1828
9	17951	1828
10	9694	1828
<b>Total</b>	<b>84174</b>	<b>1828</b>

**Territory Lookup**

Continent	Country	Region	SalesTerritoryKey
-----------	---------	--------	-------------------

## FILTER CONTEXT & FLOW

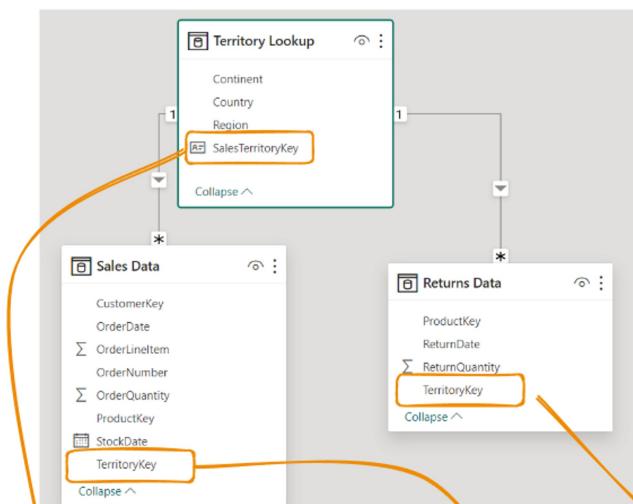


Here we have two data tables (Sales Data and Returns Data), connected to Territory Lookup

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"

**PRO TIP:** Arrange lookup tables above fact tables in your model as a visual reminder that filters always flow downstream



In this model, the only way to filter both Sales and Returns data by Territory is to use the Territory Key from the lookup table, which is upstream and related to both fact tables.

- Filtering using Territory Key from the Sales table yields incorrect Returns values, since the filter context can't flow to any other table

- Filtering using Territory Key from the Returns table yields incorrect Sales values, and is limited to territories that exist in the returns table

TerritoryKey	OrderQuantity	ReturnQuantity
1	12,513	270
2	40	
3	30	
4	17,191	362
5	49	1
6	10,94	38
7	7,862	186
8	7,950	163
9	17,951	404
10	9,694	204
Total	84,174	1,828

TerritoryKey	OrderQuantity	ReturnQuantity
1	12,513	1,828
2	40	1,828
3	30	1,828
4	17,191	1,828
5	49	1,828
6	10,94	1,828
7	7,862	1,828
8	7,950	1,828
9	17,951	1,828
10	9,694	1,828
Total	84,174	1,828

TerritoryKey	OrderQuantity	ReturnQuantity
1	84,174	270
4	84,174	362
5	84,174	1
6	84,174	238
7	84,174	186
8	84,174	163
9	84,174	404
10	84,174	204
Total	84,174	1,828

Filtering by Returns Data[Territory Key]

Filtering by Territory Lookup[Territory Key]

Filtering by Sales Data[Territory Key]

## BI-DIRECTIONAL FILTERS

### Edit relationship

Select tables and columns that are related.

Sales Data							
OrderDate	StockDate	OrderNumber	ProductKey	CustomerKey	TerritoryKey	OrderLineItem	Order
05 July 2020	03 June 2020	SO46718	360	12570	9	1	
07 July 2020	22 April 2020	SO46736	360	12341	9	1	
12 July 2020	05 May 2020	SO46776	360	12356	9	1	

Territory Lookup			
SalesTerritoryKey	Region	Country	Continent
1	Northwest	United States	North America
2	Northeast	United States	North America
3	Central	United States	North America

Cardinality

Many to one (\*:1)

Make this relationship active

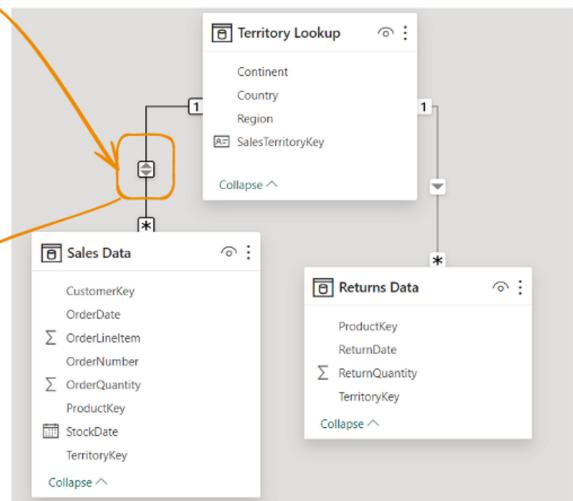
Assume referential integrity

Cross filter direction

Both

Apply security filter in both directions

OK Cancel



Properties

Relationship

Table Column

Sales Data TerritoryKey

Cardinality

Many to one (\*:1)

Table Column

Territory Lookup SalesTerritoryKey

Make this relationship active

Yes

Cross-filter direction

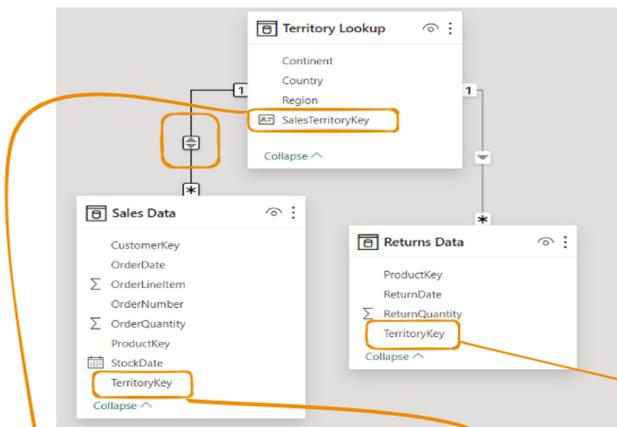
Both

Apply security filter in both directions

No

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

- In this example, filters applied to the Sales table can pass up to the Territory Lookup table, then down to Returns



With two-way cross-filtering enabled between Sales and Territory, we now see correct values using Territory Key from either table

- Filter context can now pass up to the Territory Lookup table, then downstream to Returns

- However, we still see incorrect values when filtering using Territory Key from the Returns table, since the filter context is isolated to that single table

TerritoryKey	OrderQuantity	ReturnQuantity
1	12,513	270
2	40	
3	30	
4	17,191	362
5	49	1
6	894	238
7	7,862	186
8	7,950	163
9	17,951	404
10	9,694	204
Total	84,174	1,828

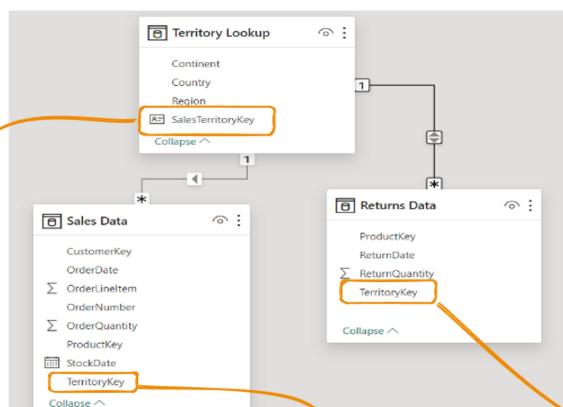
Filtering by Territory Lookup[Territory Key]

TerritoryKey	OrderQuantity	ReturnQuantity
1	12,513	270
2	40	
3	30	
4	17,191	362
5	49	1
6	894	238
7	7,862	186
8	7,950	163
9	17,951	404
10	9,694	204
Total	84,174	1,828

Filtering by Sales Data[Territory Key]

TerritoryKey	OrderQuantity	ReturnQuantity
1	84,174	270
4	84,174	362
5	84,174	1
6	174	238
7	174	186
8	84,174	163
9	84,174	404
10	84,174	204
Total	84,174	1,828

Filtering by Returns Data[Territory Key]



In this case, we've enabled two-way cross-filtering between the Returns and Territory tables

- As expected, we now see incorrect values when filtering using Territory Key from the Sales table, since the filter context is isolated to that single table

- While the values appear to be correct when filtering using Territory Key from the Returns table, we're missing sales data from any territories that didn't appear in the returns table (specifically Territories 2 & 3)

TerritoryKey	OrderQuantity	ReturnQuantity
1	12,513	270
2	40	
3	30	
4	17,191	362
5	49	1
6	894	238
7	7,862	186
8	7,950	163
9	17,951	404
10	9,694	204
Total	84,174	1,828

Filtering by Territory Lookup[Territory Key]

TerritoryKey	OrderQuantity	ReturnQuantity
1	12,513	1,828
2	40	1,828
3	30	1,828
4	17,191	1,828
5	49	1,828
6	894	1,828
7	7,862	1,828
8	7,950	1,828
9	17,951	1,828
10	9,694	1,828
Total	84,174	1,828

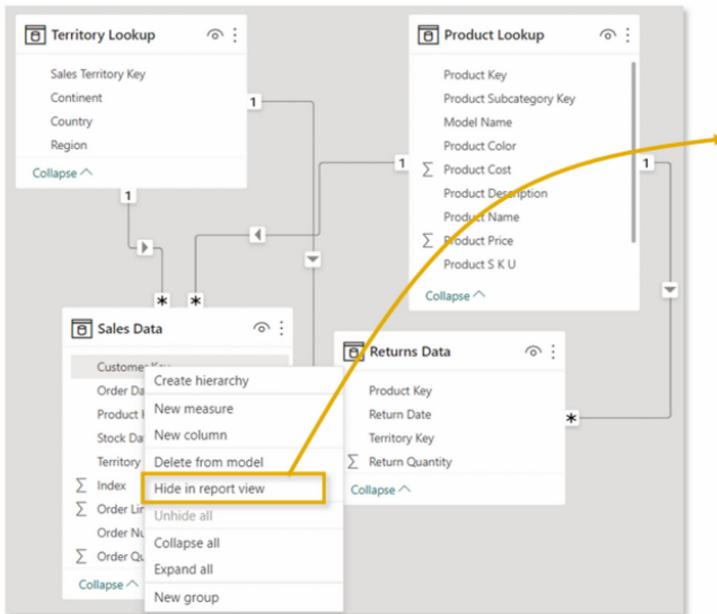
Filtering by Sales Data[Territory Key]

TerritoryKey	OrderQuantity	ReturnQuantity
1	12,513	270
4	17,191	362
5	49	1
6	10,894	238
7	862	186
8	7,950	163
9	17,951	404
10	9,694	204
Total	84,174	1,828

Filtering by Returns Data[Territory Key]

Territories 2 & 3 don't exist in the Returns table, so they aren't included in the filter context that passes to Territory Lookup and Sales

## HIDING FIELDS



Hide in Report View makes fields inaccessible from the Report tab, but still available in Data and Model views

- This can be controlled by right-clicking a field in the Data or Model view, or by selecting "Is hidden" in the Properties pane

- This is commonly used to prevent users from filtering using invalid fields, reduce clutter, or to hide irrelevant metrics from view

PRO TIP: Hide the foreign keys in fact tables to force users to filter using primary keys in dimension tables

