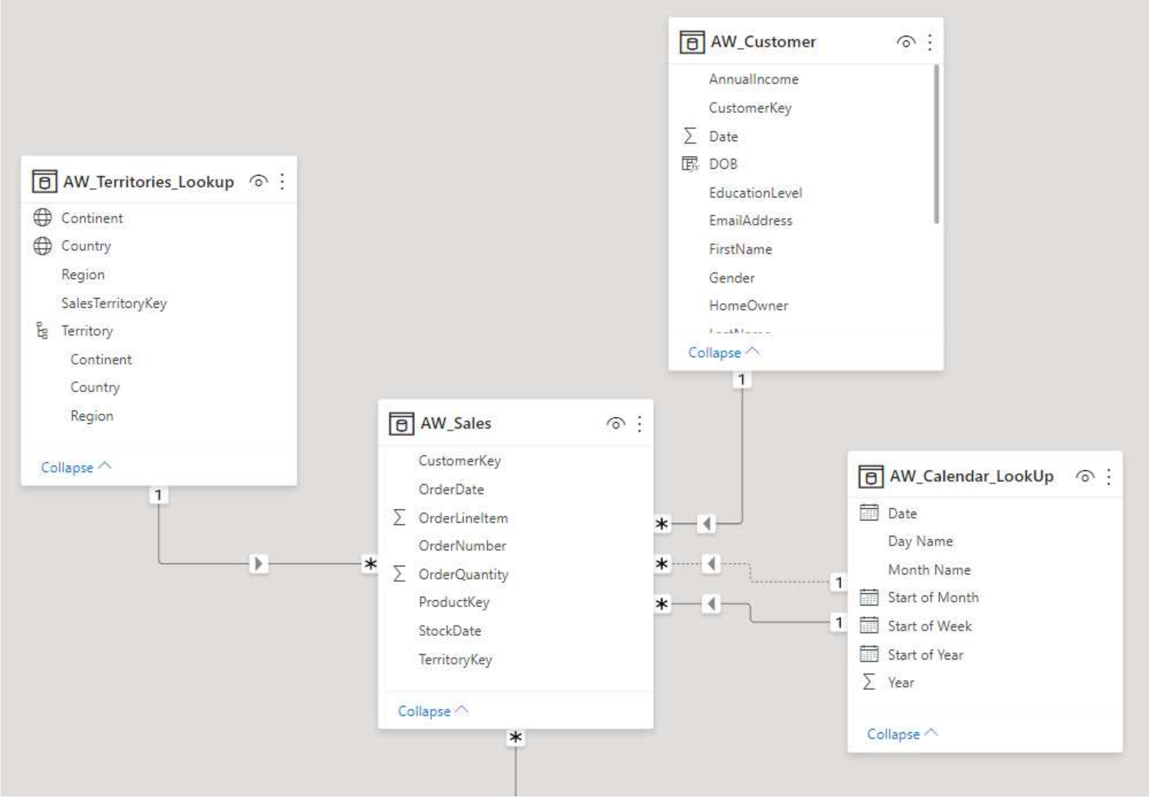
DATA MODELLING

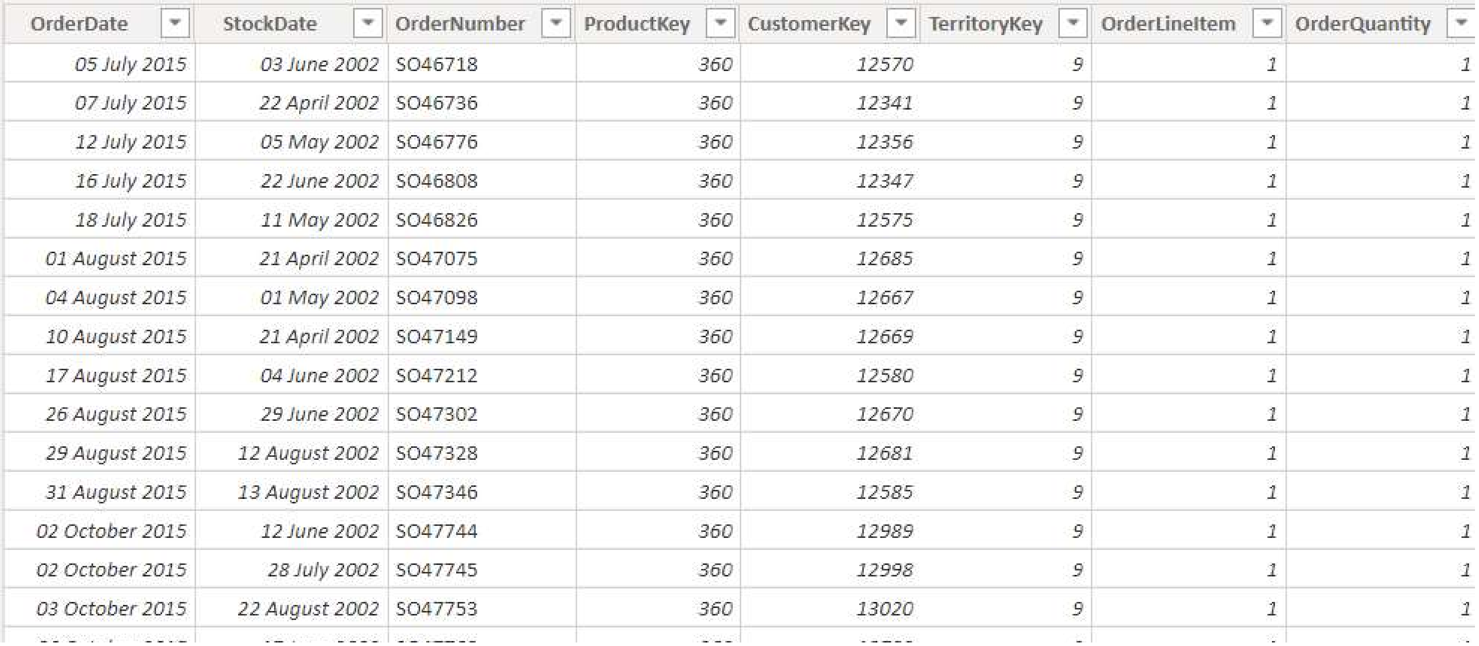
Data Models

- Tables which are connected to each other via relationships based on common fields.



Fact and Data Table

- Fact Table/Lookup Tables
  - Unique Values
- Data Tables
  - Multiple values



OrderDate	StockDate	OrderNumber	ProductKey	CustomerKey	TerritoryKey	OrderLineItem	OrderQuantity
05 July 2015	03 June 2002	SO46718	360	12570	9	1	1
07 July 2015	22 April 2002	SO46736	360	12341	9	1	1
12 July 2015	05 May 2002	SO46776	360	12356	9	1	1
16 July 2015	22 June 2002	SO46808	360	12347	9	1	1
18 July 2015	11 May 2002	SO46826	360	12575	9	1	1
01 August 2015	21 April 2002	SO47075	360	12685	9	1	1
04 August 2015	01 May 2002	SO47098	360	12667	9	1	1
10 August 2015	21 April 2002	SO47149	360	12669	9	1	1
17 August 2015	04 June 2002	SO47212	360	12580	9	1	1
26 August 2015	29 June 2002	SO47302	360	12670	9	1	1
29 August 2015	12 August 2002	SO47328	360	12681	9	1	1
31 August 2015	13 August 2002	SO47346	360	12585	9	1	1
02 October 2015	12 June 2002	SO47744	360	12989	9	1	1
02 October 2015	28 July 2002	SO47745	360	12998	9	1	1
03 October 2015	22 August 2002	SO47753	360	13020	9	1	1

Database Normalization

- Normalization is the process of organizing the tables and columns in a relational database to reduce redundancy and preserve data integrity.
- Tables should have distinct and specific data

Exercise

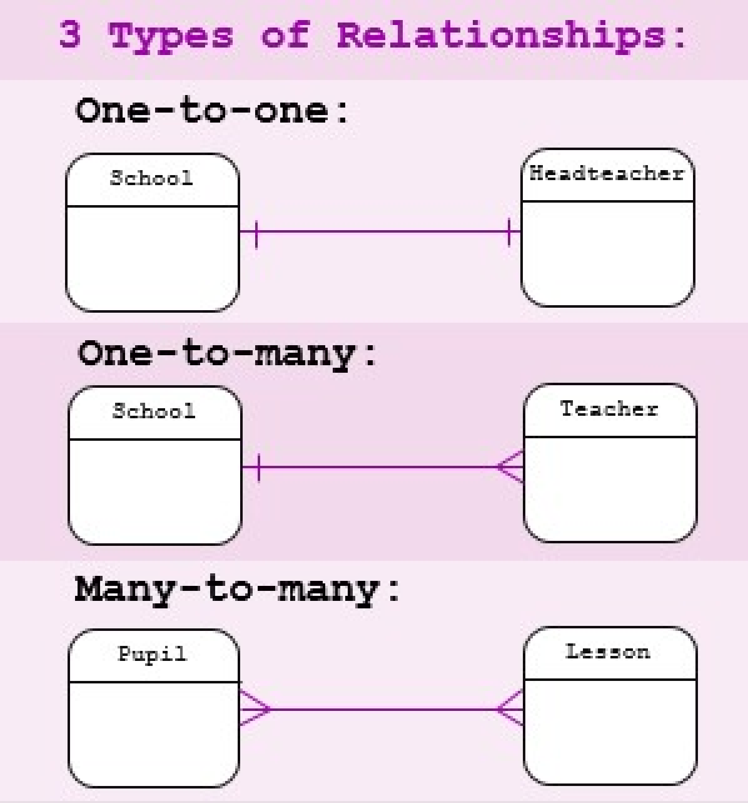
- Create tables from the Database shown following the rules of Normalization
  - Identify them as Fact or Data Table

Primary and Foreign Key

- Primary Keys are unique and exist in Fact/Lookup tables
- Foreign keys may be duplicate and they exist in Data tables
- Primary and Foreign Keys are used to create relationships between the tables

Types of Relations

- One to many (used in Power BI)
- Many to many
- One to one
  - It's better to Merge



Creating Relationship

- Two Ways
  - Click and drag to another table on the key match
  - manage relationship button. More manual.
- Delete Relationship
- Edit Relationship
- Active vs Inactive Relationship
  - Multiple relationships from one table to another table are not allowed
  - One relationship can be active and shown by solid line
  - Another relationship will be inactive and shown by dotted line

Important - Relationships

- NO relationship created between two Data tables
  - It results in Many to Many relationship
- They are connected through Fact/Lookup Tables

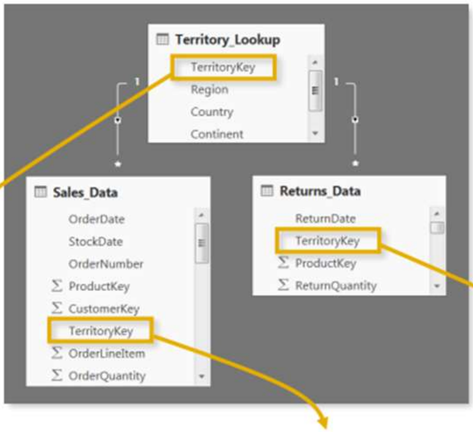
Filter Flow

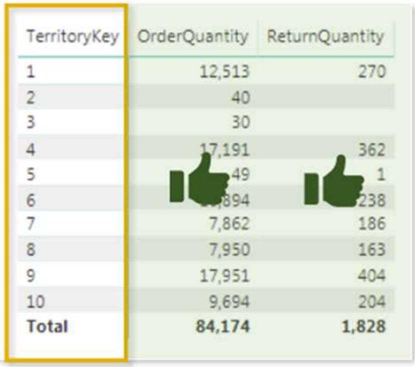
- When we create relationships between two tables we can see an arrow on the relationship lines. This indicates Filter Flow.
- Filter flow is always from one side to many side of the relationship.
- When we filter the table, the filter context is passed to all the downstream tables.
- Filters cannot flow upstream.

Filter Flow (Contd)

In this case, the only valid way filter both **Sales** and **Returns** data by Territory is to use the *TerritoryKey* field from the **Territory_Lookup** table, which is upstream and related to *both* data tables

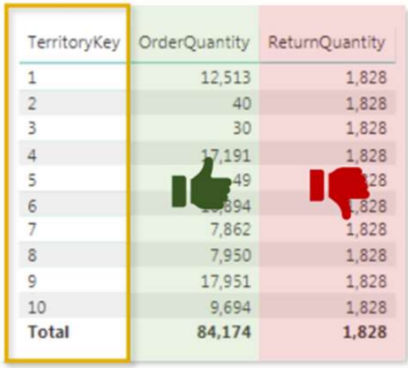
- Filtering using *TerritoryKey* from the **Sales** table yields incorrect **Returns** values, since the filter context ***cannot flow upstream*** to either one of the other tables
- Similarly, filtering using *TerritoryKey* from the **Returns** table yields incorrect **Sales** data; in addition, ***only territories that registered returns are visible in the table*** (even though they registered sales)





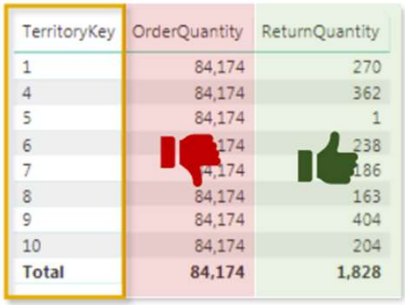
TerritoryKey	OrderQuantity	ReturnQuantity
1	12,513	270
2	40	
3	30	
4	17,191	362
5	49	1
6	894	238
7	7,862	186
8	7,950	163
9	17,951	404
10	9,694	204
Total	**84,174**	**1,828**

1) *Filtering using TerritoryKey from the **Territory_Lookup** table*



TerritoryKey	OrderQuantity	ReturnQuantity
1	12,513	1,828
2	40	1,828
3	30	1,828
4	17,191	1,828
5	49	28
6	894	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**

2) *Filtering using TerritoryKey from the **Sales_Data** table*



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

3) *Filtering using TerritoryKey from the **Returns_Data** table*

Professional School
ISO 9001 : 2000 Organisation

The END