



# **DATAMODELLERING VOOR SNELLERE POWER BI DASHBOARDS**

Peter van den Bos





powered by:

**GET  
RESPONSIVE**





# Peter van den Bos

**Data Modeling & DAX Specialist**  
**Freelance Trainer & Speaker**  
**Founder @ DutchBiGeek**

 [dutchbigeek.nl](https://dutchbigeek.nl)  
 [linkedin.com/in/petervandenbos](https://linkedin.com/in/petervandenbos)  
 [github.com/PetervandenBos/](https://github.com/PetervandenBos/)  
 [sessionize.com/peter-van-den-bos](https://sessionize.com/peter-van-den-bos)  
 [youtube.com/@DutchBiGeek](https://youtube.com/@DutchBiGeek)  
 [peter@dutchbigeek.nl](mailto:peter@dutchbigeek.nl)



**Geek by design**



*Because performance isn't optional*





# Waarom is het datamodel belangrijk?

- Het datamodel vormt de basis van Power BI.
- Een goed ontworpen datamodel verbetert de prestaties en versnelt query's.
- 80% van de prestatieproblemen zijn gerelateerd aan het datamodel.
- Een beter datamodel maakt DAX-query's eenvoudiger en efficiënter.
- Vermindert data-redundantie en minimaliseert fouten.

# VertiPaq-compressie

UnitCost ▼	StoreName ▼
€ 10,00	Contoso North America Online Store
€ 11,00	Contoso Europe Online Store
€ 14,00	Contoso Europe Online Store
€ 18,00	Contoso Europe Online Store
€ 15,00	Contoso North America Online Store
€ 14,00	Contoso North America Online Store
€ 19,00	Contoso Asia Online Store
€ 14,00	Contoso North America Online Store
€ 17,00	Contoso Europe Online Store
€ 19,00	Contoso Asia Online Store



# Dictionary Encoding

Vervangt unieke tekstwaarden door numerieke sleutels

Bespaart geheugen: tekst → geheel getal

Waarden worden één keer opgeslagen in een woordenboek en via sleutels benaderd

Vooraf effectief bij kolommen met een lage cardinaliteit

StoreName	StoreName.ID
Contoso North America Online Store	1
Contoso Europe Online Store	2
Contoso Europe Online Store	2
Contoso Europe Online Store	2
Contoso North America Online Store	1
Contoso North America Online Store	1
Contoso Asia Online Store	3
Contoso North America Online Store	1
Contoso Europe Online Store	2
Contoso Asia Online Store	3

# Run-Length Encoding (RLE)

Comprimeert opeenvolgende herhalingen van dezelfde waarde

Slaat op als (waarde + aantal)

Werkt alleen goed wanneer de kolom gesorteerd is

Wordt alleen toegepast als het daadwerkelijk geheugen bespaart

StoreName.ID	StoreName	StoreName.ID	Count	Rows
1	1	1	0	4
2	1	2	4	4
2	1	3	7	2
2	1			
1	2			
1	2			
3	2			
1	2			
2	3			
3	3			

# Value Encoding (Bit- Packing)

VertiPaq kiest het kleinste aantal bits  
per kolom

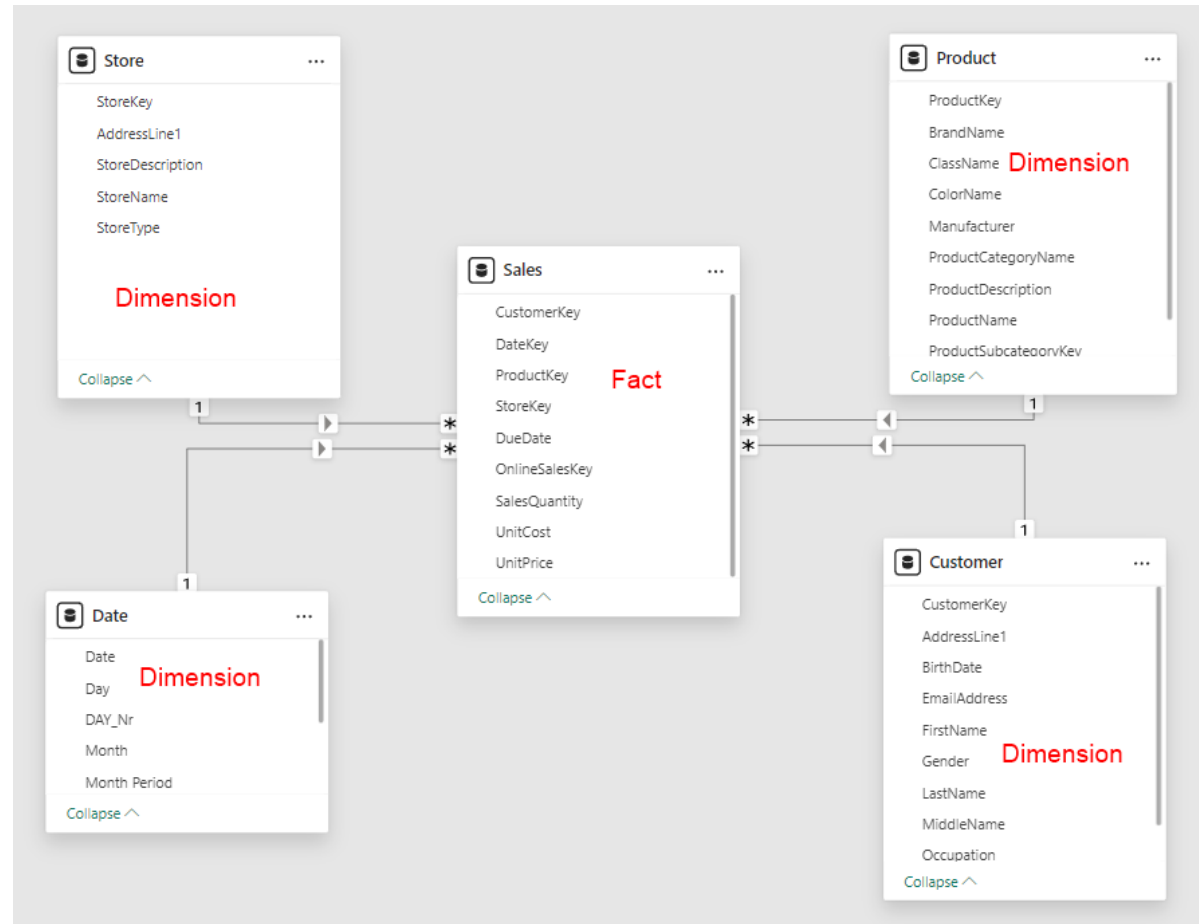
Kleinere waarde-interval = minder  
bits nodig

Werkt na woordenboekcodering of  
op numerieke kolommen

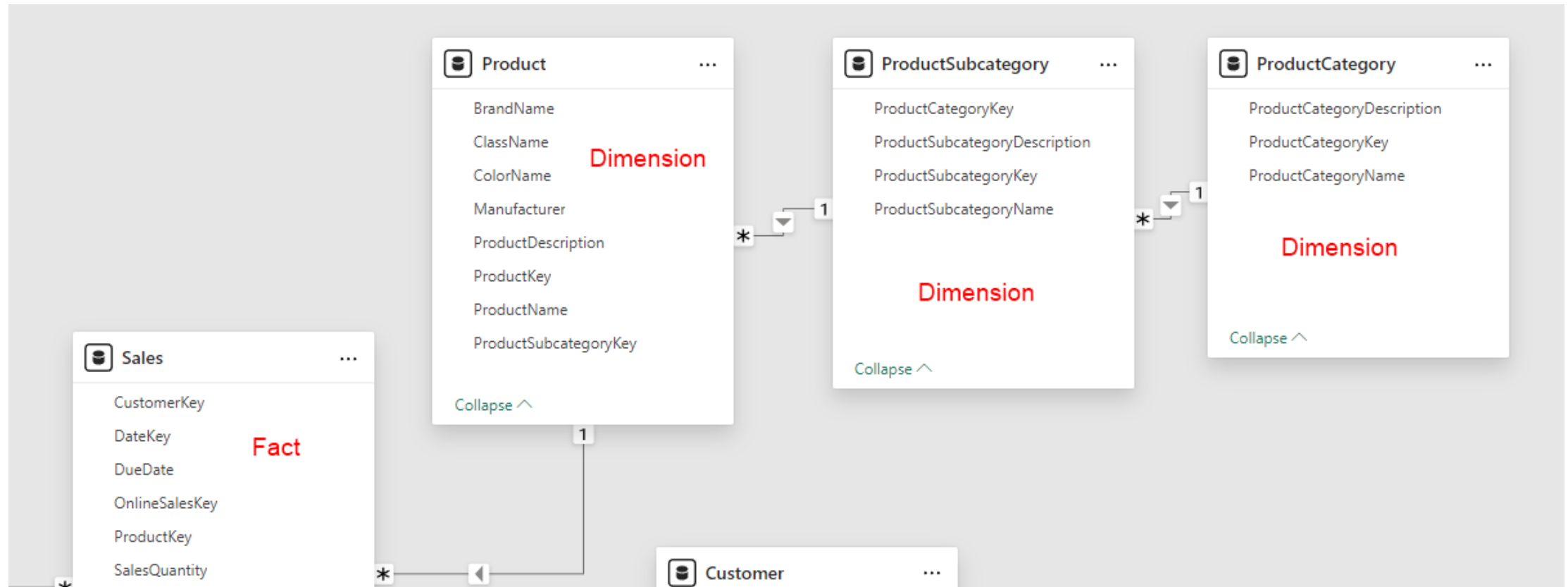
Original	Optimized	
UnitCost	UnitCost	
€ 10,00	0	+10
€ 11,00	1	+10
€ 14,00	4	+10
€ 18,00	8	+10
€ 15,00	5	+10
€ 14,00	4	+10
€ 19,00	9	+10
€ 14,00	4	+10
€ 17,00	7	+10
€ 19,00	9	+10



# Star Schema



# Snowflake-model





# Waaruit bestaat een geoptimaliseerd datamodel?

Het model bestaat uit facts en dimensions.

Een fact table bevat waarden die berekend kunnen worden, zoals:

- Omzet, aankoopmoment, verkochte producten, enzovoorts.

Een dimension table bevat waarden waarop gefilterd kan worden, zoals:

- Jaar, maand, fabrikant, klant, enzovoorts.



# Relaties en filtering

Er moet een relatie worden gemaakt tussen facts en dimensions.

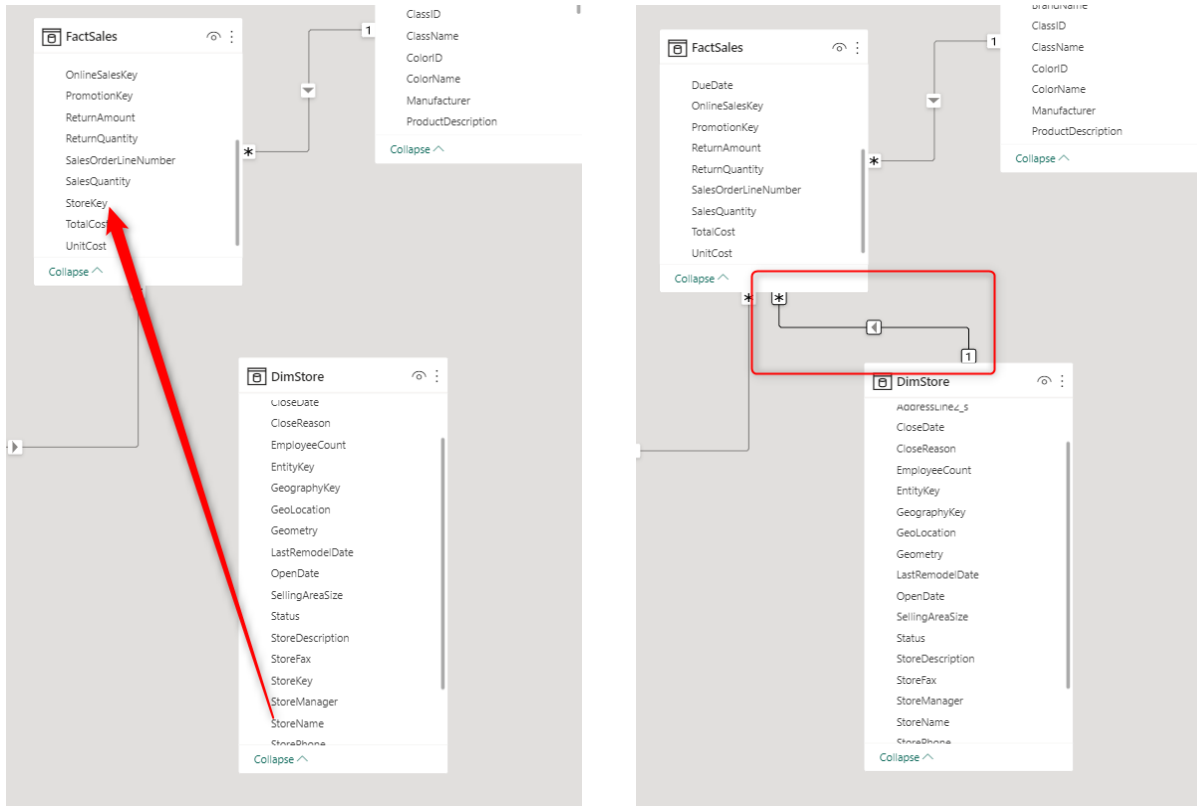
Mogelijke relaties zijn:

- One-to-many
- One-to-one
- Many-to-many

Daarnaast kan de filterrichting worden ingesteld op:

- Single
- Both

# Voorbeeld: Relaties en filtering



## Edit relationship

Select tables and columns that are related.

FactSales						
OnlineSalesKey	DateKey	StoreKey	ProductKey	PromotionKey	CurrencyKey	CustomerKey
20473555	2019-08-14 00:00:00.000	199	1670	1	1	1123
23353743	2020-07-23 00:00:00.000	199	1676	1	1	17042
20706373	2019-09-12 00:00:00.000	199	1676	1	1	7817

DimStore					
StoreKey	GeographyKey	StoreManager	StoreType	StoreName	StoreDescription
307	710	292	Online	Contoso Asia Online Store	Contoso Asia Online S
306	586	246	Online	Contoso Europe Online Store	Contoso Europe Onlin
199	800	212	Online	Contoso North America Online Store	Contoso Online Store

Cardinality

Many to one (\*:1)

Many to one (\*:1)

One to one (1:1)

One to many (1:\*)

Many to many (\*:\*)

Cross filter direction

Single

☐ Apply security filter in both directions

OK

Cancel



# Meerdere fact tables

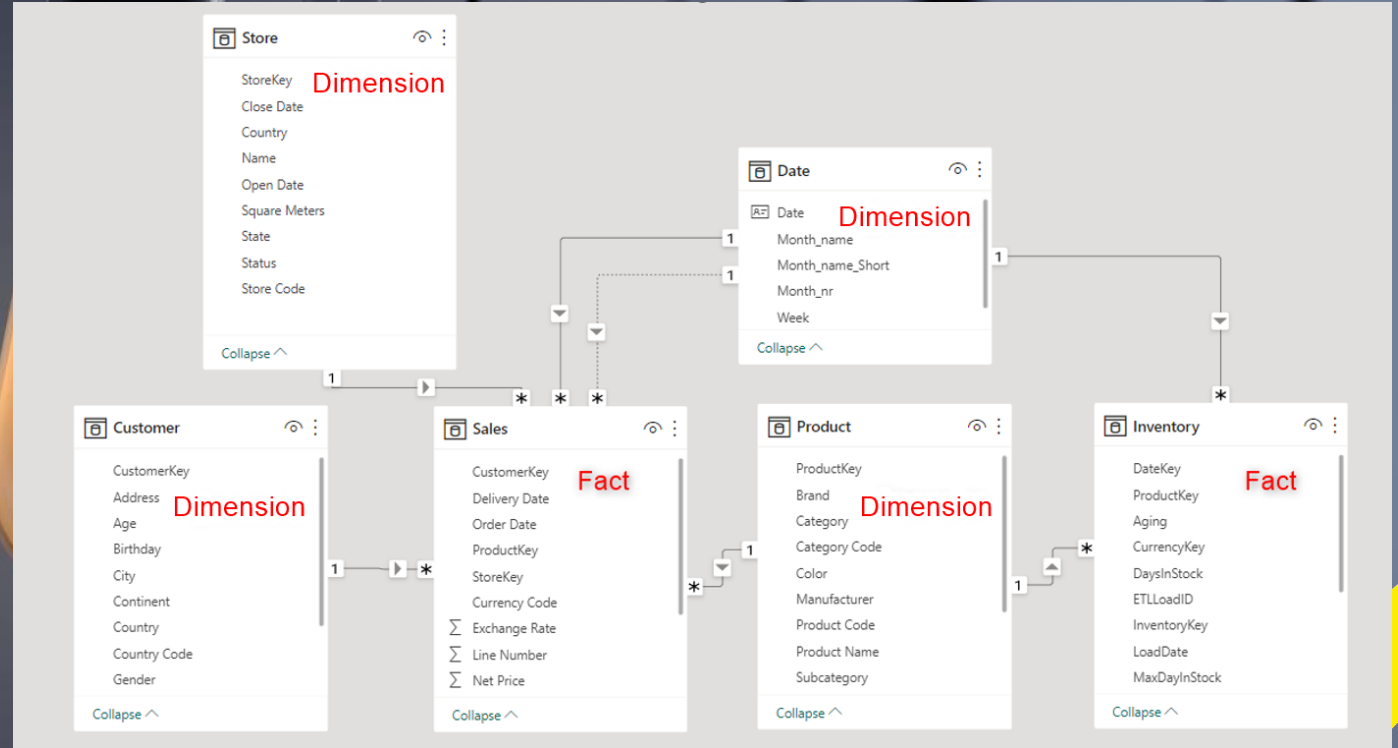
Het is mogelijk om meerdere fact tables in je model te gebruiken.

Deze fact tables hoeven niet veel gemeen te hebben

- zoals inventory en sales.

Het is echter noodzakelijk om enkele dimension tables te hebben die verbonden zijn met de fact tables

- date en product.

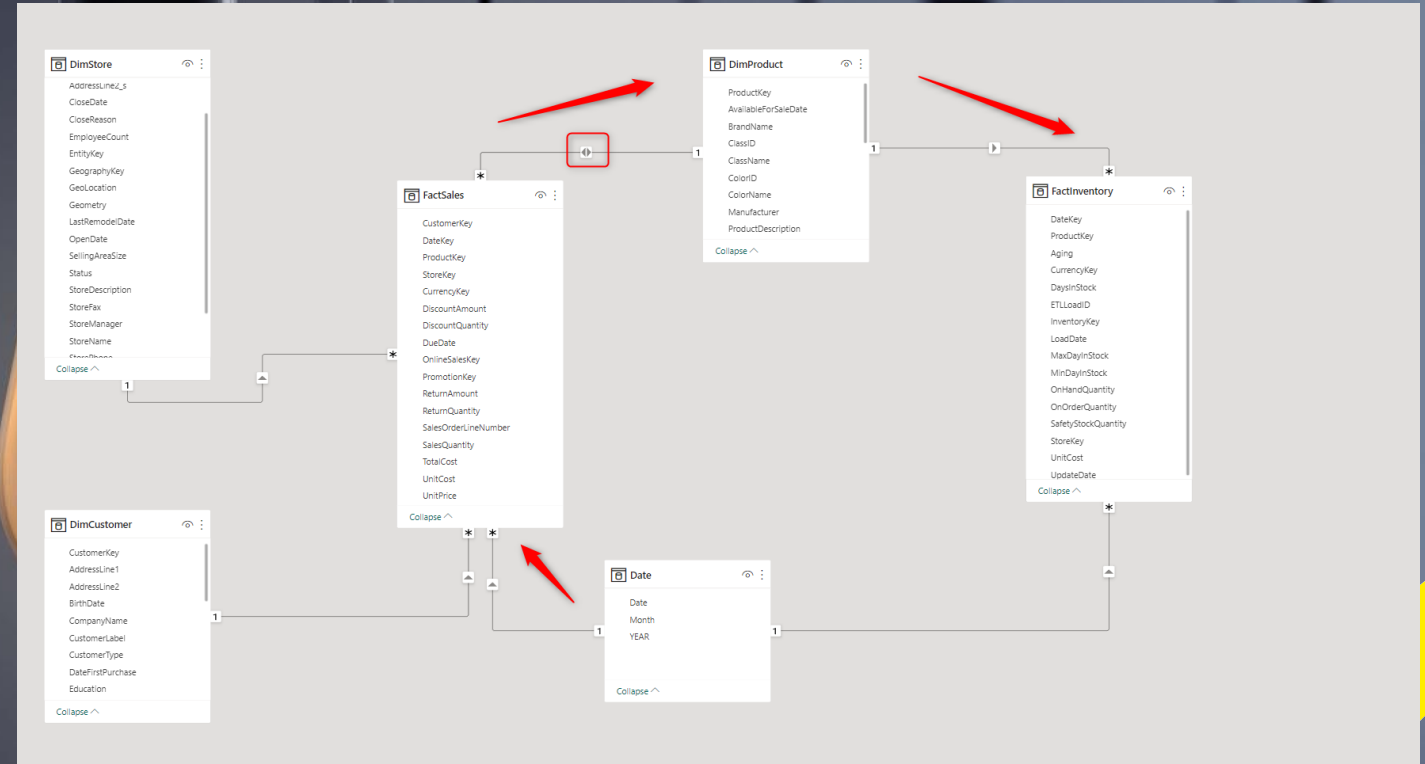


# Ambiguity

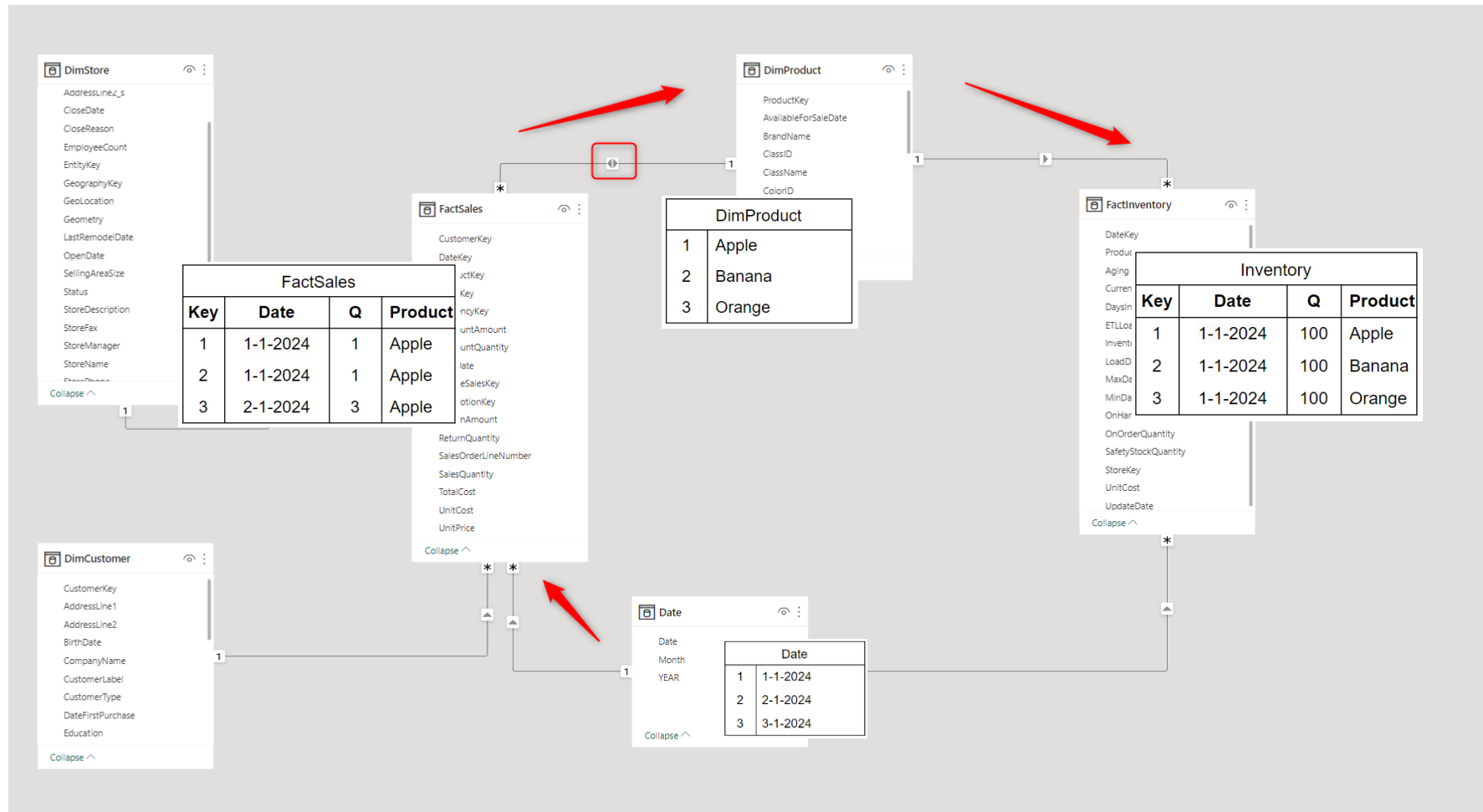
Wanneer een model meerdere fact tables bevat en bidirectionele relaties gebruikt, is er een risico op ambiguïteit.

Dit betekent dat het model niet weet welk pad het moet volgen bij het filteren.

Als het via meerdere tabellen kan filteren, kan het mogelijk onjuiste waarden weergeven.



# Voorbeeld ambiguity



# Many-to-many relaties

Many-to-many (M:N) relaties

- Voorbeeld: Books en Writers (CRM vs. ERP)

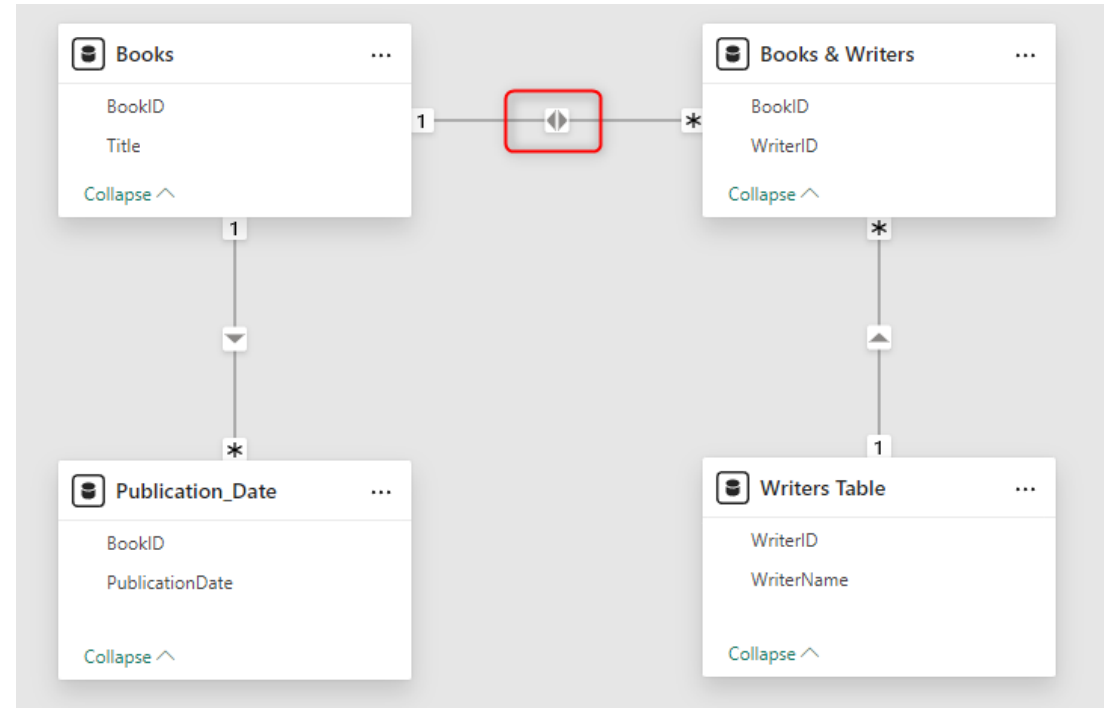
Hoe los je dit op?

Maak een bridge table:

- Maak een tabel die alle sleutels bevat.

Let op de richting van de cross-filter:

- Filtering werkt alleen van de 1 naar de \* kant.





# Belang van het Star Schema voor VertiPaq

Het Star schema bevordert de prestaties:

- Snellere query's door het scheiden van dimensions en facts.
- Vereenvoudigt data voor efficiënte VertiPaq-verwerking.

Goede datavoorbereiding is essentieel:

- Consistente dataformaten zorgen voor betere compressie en lager RAM-gebruik.

Kleine aanpassingen, grote winsten:

- Kleine veranderingen kunnen de prestaties aanzienlijk verbeteren.





# Het belang van het kiezen van de juiste gegevenstypen

Selecteer de juiste gegevenstypen om de modelprestaties te optimaliseren.

- Begrijp de gegevensvereisten en rapportbehoeften.

Power Query gebruikt standaard "Decimal" voor numerieke kolommen:

- Dit kan leiden tot onnodige decimalen.
- Evalueer of minder decimalen of gehele getallen voldoende zijn.

De juiste keuze van gegevenstype is cruciaal voor de VertiPaq-compressie.



# Optimaliseren van dataload voor compressie

Laad alleen de noodzakelijke data om optimale compressie te bereiken.

Verminder het aantal kolommen per tabel:

- Minder kolommen zorgen voor effectievere compressie.
- VertiPaq sorteertechniek: Slaat kolommen met de laagste cardinaliteit eerst op voor betere compressie.
- Meer kolommen, vooral met een hogere cardinaliteit, verminderen de efficiëntie van de compressie.

Beperk je tot ongeveer 15 kolommen per tabel om een efficiënt model te behouden.



# Best practices

- Gebruik altijd een star schema, of een snowflake schema indien nodig.
- Een fact table bevat waarden voor berekeningen.
- Een dimension table bevat waarden voor filtering.
- Vermijd het gebruik van bidirectionele relaties.
- Bevat alleen de data die je daadwerkelijk gebruikt.
- Bepaal van tevoren het detailniveau voor je tabellen.