



Power BI

Giorno 2

Agenda

Cosa vedremo

- Modellare i dati in Power BI Desktop
- Star Schema



Agenda

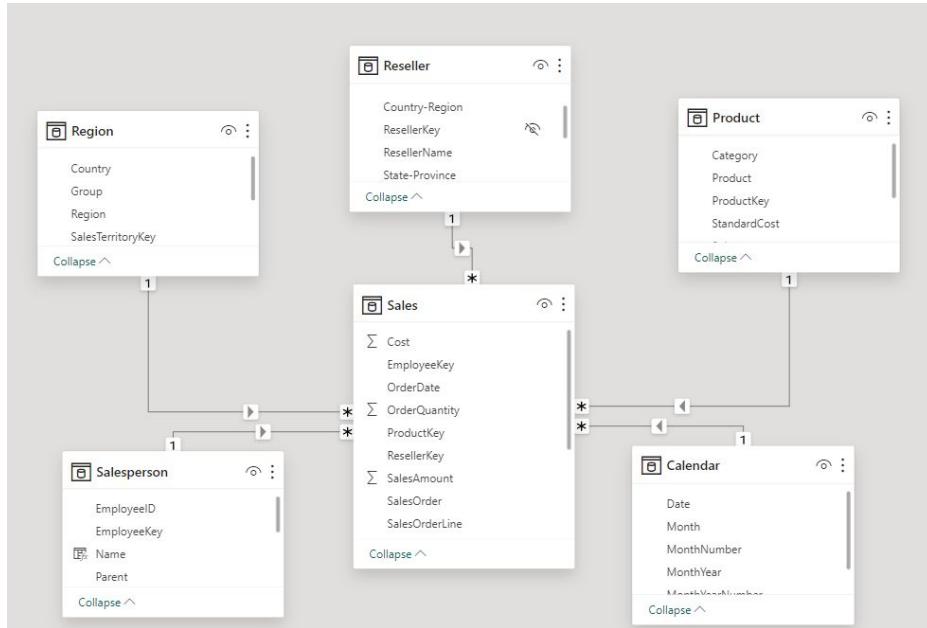
Obiettivi di apprendimento

- apprendere i concetti di modellazione dati
- conoscere il comportamento delle relazioni in PBI



Modellare i dati in Power BI Desktop

Modello dati



La progettazione di un modello dati è l'attività più critica di sviluppo.

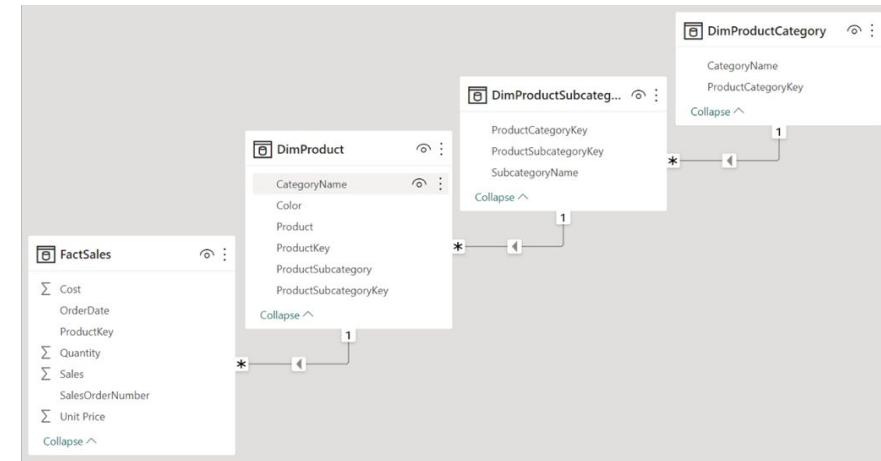
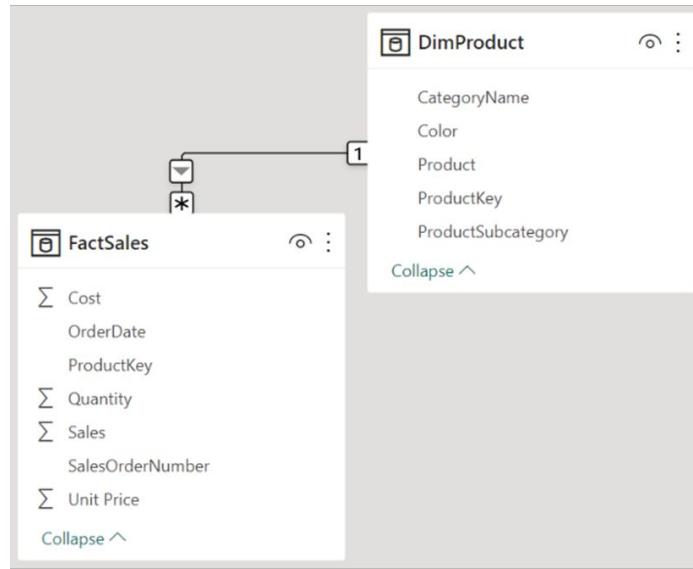
Il modello dati impatta:

- L'affidabilità del report
- La fruibilità del report

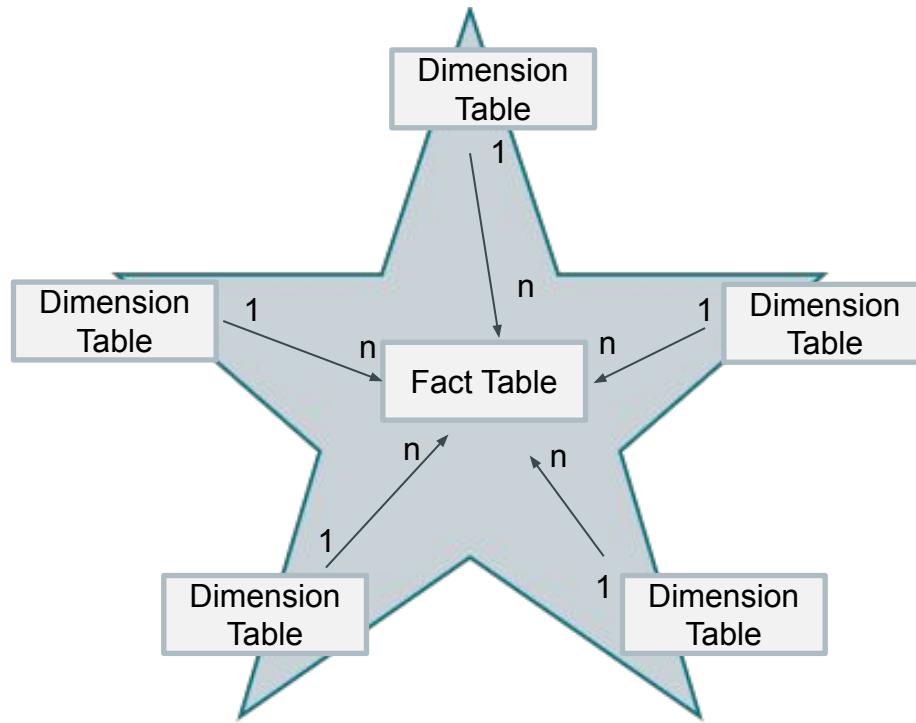
Un buon modello dati implica vantaggi in termini di:

- Qualità delle informazioni esposte
- Performance
- Aggregazioni (Roll-up / drill-down)
- Manutenibilità del report

Denormalization vs Normalization



Star schema



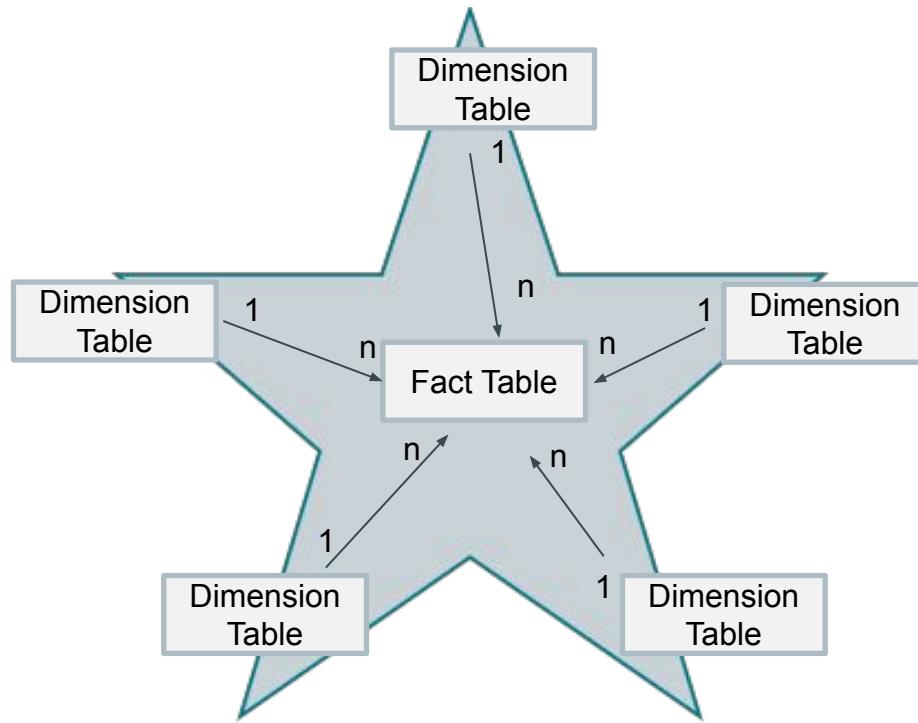
Lo **Star Schema** è un approccio logico alla modellazione dati maturo che ottimizza l'analisi dei dati.

In uno **Star Schema**, le tabelle sono classificate in fatti e dimensioni.

Fatto: è l'evento o l'insieme di osservazioni che bisogna misurare, quantificare o analizzare (vendite, acquisti, commesse, movimenti di magazzino, ...)

Dimensione: descrive il fatto; è una prospettiva di analisi rispetto la quale filtrare il fatto (prodotto, cliente, area geografica, calendario, ...)

Star schema

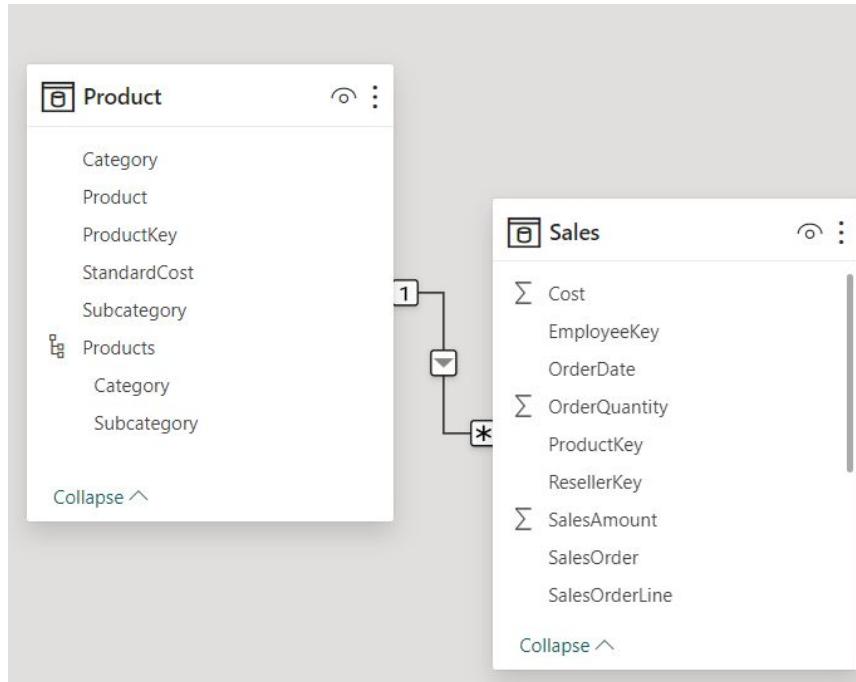


In uno **Star Schema**, per semplificare, il fatto supporta aggregazioni (è ciò che vogliamo aggregare) e la dimensione consente di applicare filtri (è ciò che utilizziamo per applicare i filtri).

In uno **Star Schema** fatti e dimensioni sono generalmente denormalizzati.

Tra le dimensioni e i fatti si implementano relazioni uno-a-molti.

Relazioni



La relazione definisce un legame logico tra due tabelle.

Esistono diverse tipologie di relazioni:

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

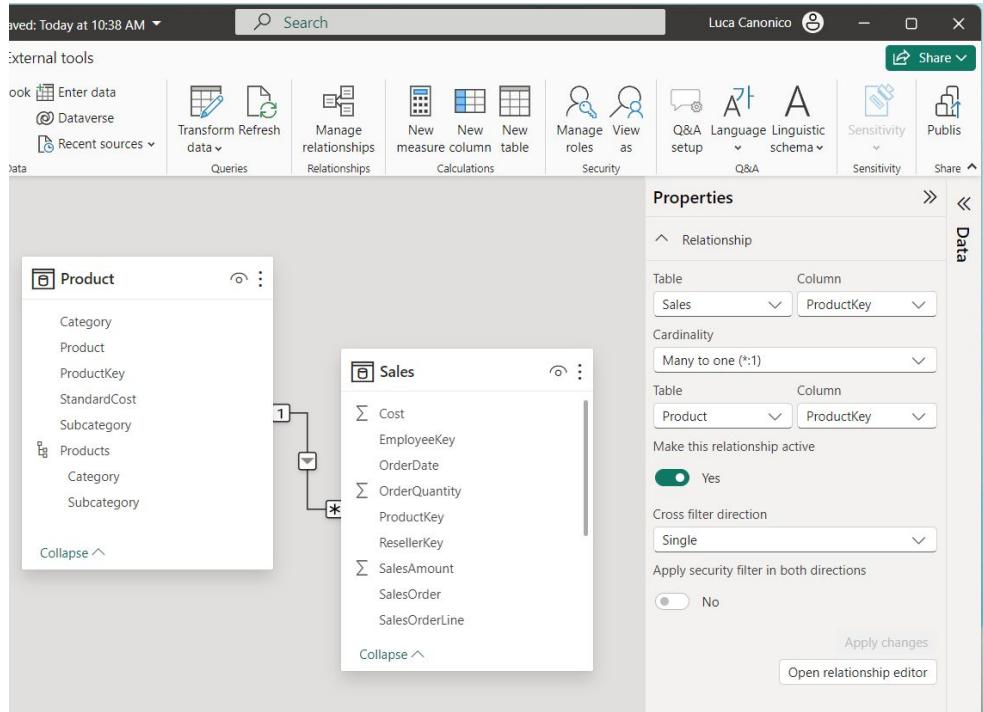
In Power BI è possibile implementare fisicamente anche relazioni:

- Many-to-many (*:*)

Le relazioni consentono ai filtri di propagarsi!

Il 'modo' con cui i filtri si propagano dipende dalla cardinalità!

Cardinalità



La cardinalità rappresenta l'univocità dei valori in una colonna.

In Power BI la cardinalità implica un concetto estremamente importante: la **direzione di propagazione dei filtri** (cross-filtering direction).

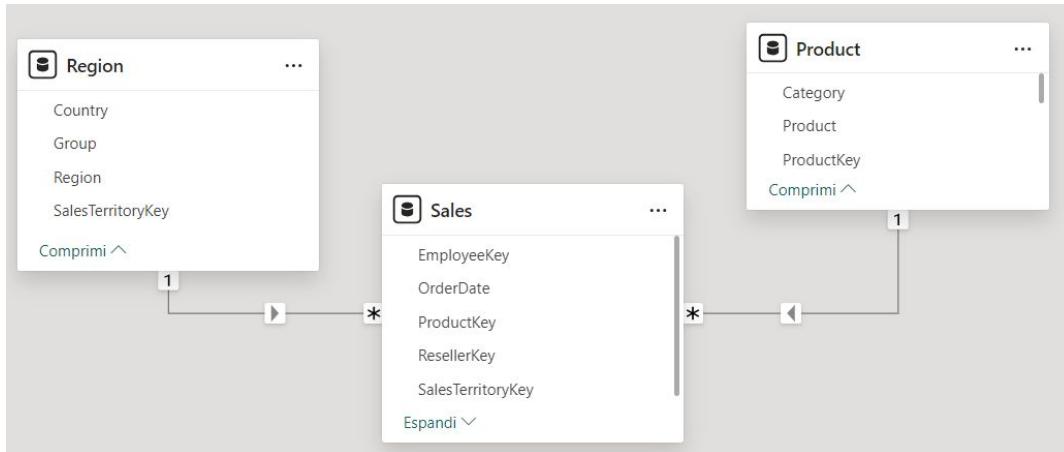
La direzione di propagazione dei filtri può essere:

- **Singola**, il filtro si propaga dal lato 1 al lato molti della relazione.
- **Doppia** (o bidirezionale), il filtro si propaga in entrambe le direzioni.

Per le relazioni one-to-one, l'unica opzione disponibile è la bi-direzionale.

Per le relazioni many-to-many, è possibile scegliere l'una o l'altra.

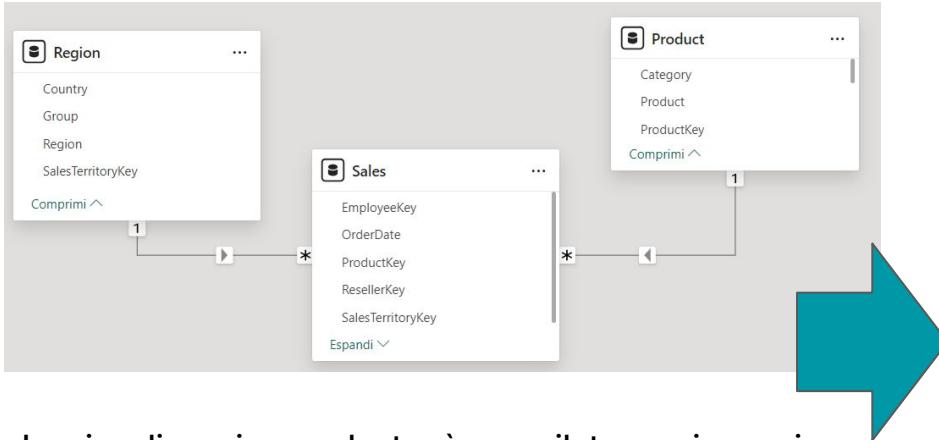
Direzione di propagazione dei filtri



Dato un modello logico come quello mostrato a sinistra, supponiamo di dover ottenere i ricavi totali e il numero di prodotti diversi venduti per ciascuna regione.

L'obiettivo è esporre per ciascuna regione le aggregazioni richieste.

Direzione di propagazione dei filtri



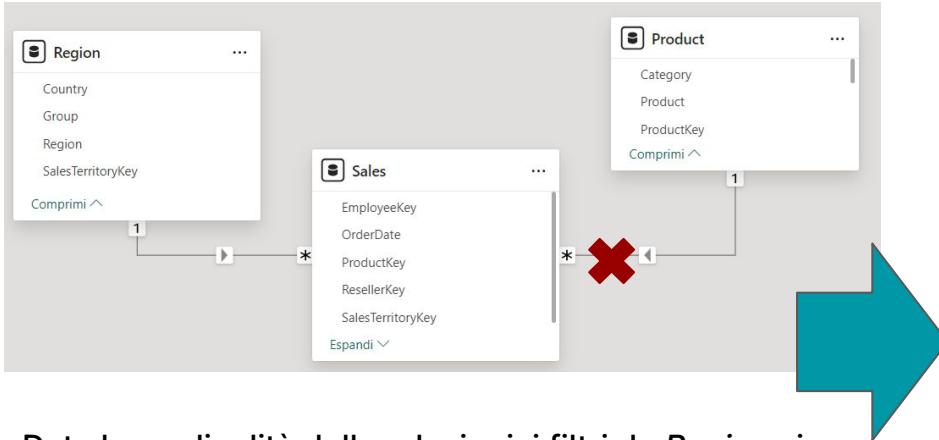
La visualizzazione a destra è compilata con i campi:
Region, SalesAmount e ProductKey rispettivamente di
Region, Sales e Product.

Il campo *Region* è un filtro implicito mentre *SalesAmount* e *ProductKey* sono due misure implicite in quanto campi aggregati (rispettivamente somma e conteggio).

Cosa desumi dalla visualizzazione? L'informazione esposta è corretta?

Region	Somma di SalesAmount	Conteggio di ProductKey
Australia	1.391.024,85	397
Canada	13.875.632,96	397
Central	7.633.386,86	397
France	4.527.839,69	397
Germany	1.877.743,39	397
Northeast	6.715.354,39	397
Northwest	12.004.822,45	397
Southeast	7.638.606,55	397
Southwest	18.001.116,10	397
United Kingdom	3.883.042,96	397
Totale	77.548.570,20	397

Direzione di propagazione dei filtri

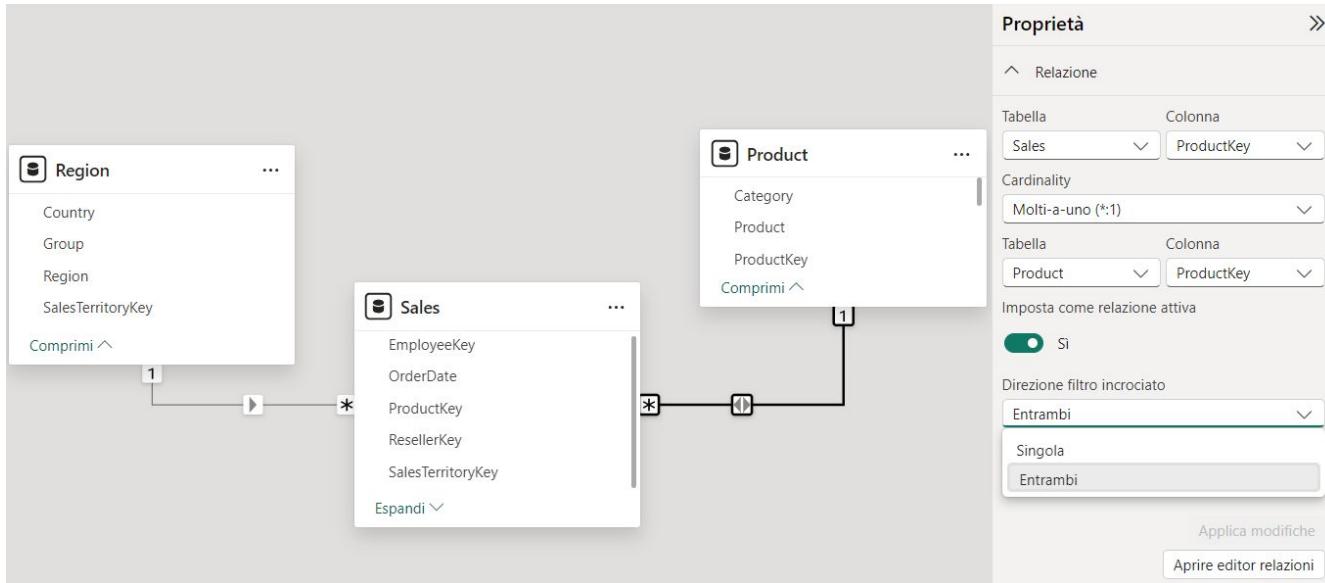


Data la cardinalità delle relazioni, i filtri da *Region* si propagano correttamente verso *Sales* ma non possono propagarsi verso *Product*.

Di conseguenza, la visualizzazione espone il conteggio dei *ProductKey* a prescindere dalle regioni (*Region* non filtra *Product*)

Region	Somma di SalesAmount	Conteggio di ProductKey
Australia	1.391.024,85	397
Canada	13.875.632,96	397
Central	7.633.386,86	397
France	4.527.839,69	397
Germany	1.877.743,39	397
Northeast	6.715.354,39	397
Northwest	12.004.822,45	397
Southeast	7.638.606,55	397
Southwest	18.001.116,10	397
United Kingdom	3.883.042,96	397
Totale	77.548.570,20	397

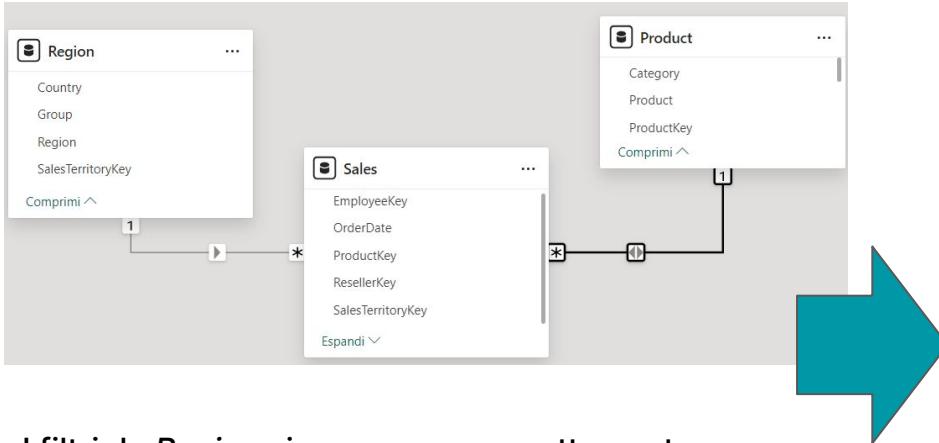
Direzione di propagazione dei filtri



La direzione di propagazione dei filtri puo' essere gestita come 'doppia'.

Quando la relazione è doppia i filtri si propagano in entrambe le direzioni.

Direzione di propagazione dei filtri



I filtri da *Region* si propagano correttamente verso *Sales* e viene filtrata di conseguenza anche *Product* (in quanto tra *Sales* e *Product* è stata definita una direzione di propagazione dei filtri a "doppio senso di marcia").

Cosa desumi dalla visualizzazione? è corretto?

Region	Somma di SalesAmount	Conteggio di ProductKey
Australia	1.391.024,85	117
Canada	13.875.632,96	329
Central	7.633.386,86	324
France	4.527.839,69	265
Germany	1.877.743,39	157
Northeast	6.715.354,39	319
Northwest	12.004.822,45	324
Southeast	7.638.606,55	329
Southwest	18.001.116,10	328
United Kingdom	3.883.042,96	273
Totale	77.548.570,20	397

Direzione di propagazione dei filtri

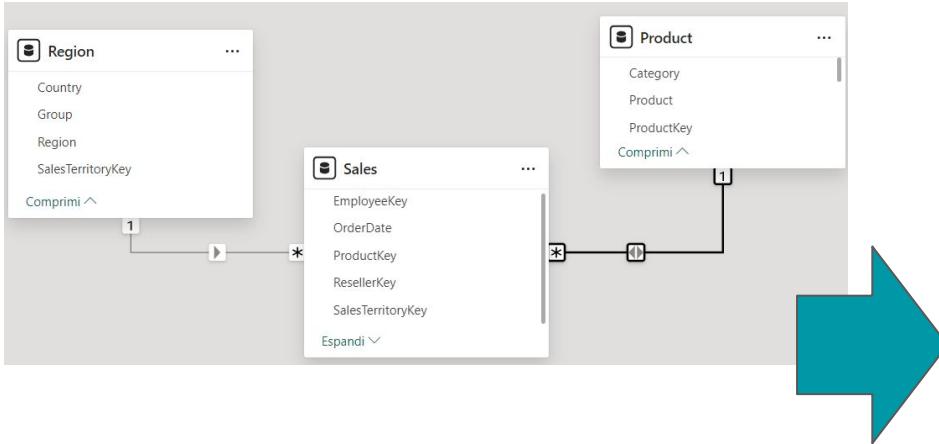
Utilizza la relazione doppia con parsimonia.

Quando la **Direzione filtro incrociato** è impostata in *Entrambi* le tabelle vengono considerate come se fossero una singola tabella (quando filtrate).



Le relazioni bi-direzionali degradano le prestazioni del report pertanto, utilizzale solo se necessario.

Direzione di propagazione dei filtri



... e se, per il calcolo del numero di prodotti diversi venduti per regione, utilizzassimo il campo *ProductKey* di *Product**?

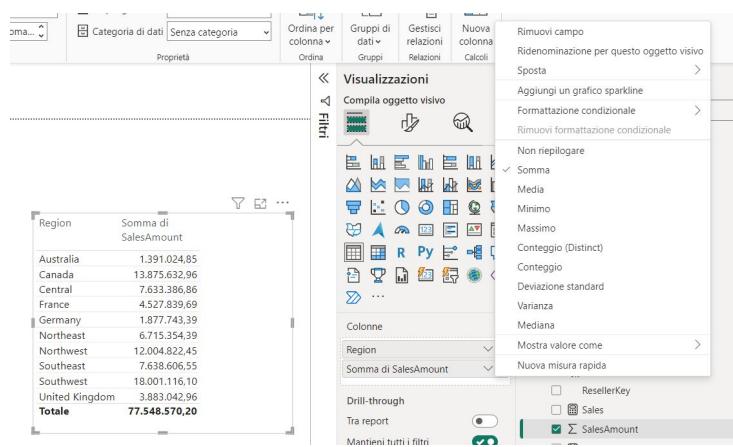
* definendo come modalità di aggregazione Conteggio (Distinct)

Region	Somma di SalesAmount	Conteggio di ProductKey (Product)	Conteggio di ProductKey (Sales)
Australia	1.391.024,85	117	117
Canada	13.875.632,96	329	329
Central	7.633.386,86	324	324
France	4.527.839,69	265	265
Germany	1.877.743,39	157	157
Northeast	6.715.354,39	319	319
Northwest	12.004.822,45	324	324
Southeast	7.638.606,55	329	329
Southwest	18.001.116,10	328	328
United Kingdom	3.883.042,96	273	273
Totale	77.548.570,20	397	334

Direzione di propagazione dei filtri

Ricorda che in Power BI il risultato di un'aggregazione (che sia una misura implicita o esplicita) dipende dai filtri applicati.

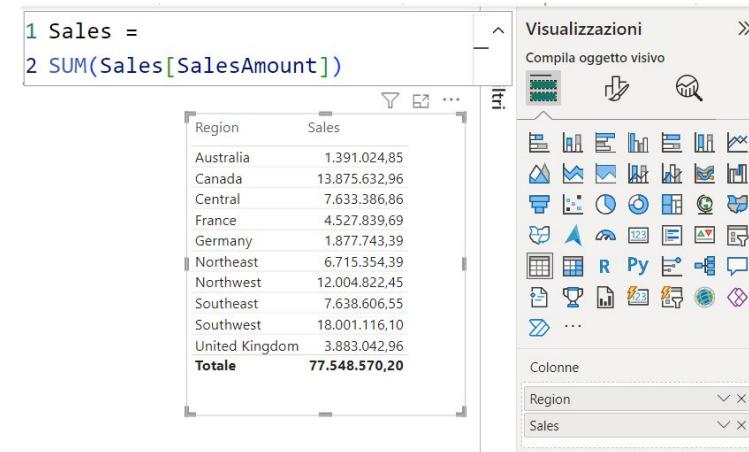
misura implicita: campo aggregato



The screenshot shows the Power BI Data View. A context menu is open over a table named 'Region'. The menu path 'Visualizzazioni > Somma di SalesAmount' has been selected. The menu lists various aggregation functions: Somma, Media, Minimo, Massimo, Conteggio (Distinct), Conteggio, Deviazione standard, Varianza, Mediana, and Mostra valore come. The 'Somma' option is highlighted. The table below shows the sum of SalesAmount for each region.

Region	Somma di SalesAmount
Australia	1.391.024,85
Canada	13.875.632,96
Central	7.633.386,86
France	4.527.839,69
Germany	1.877.743,39
Northeast	6.715.354,39
Northwest	12.004.822,45
Southeast	7.638.606,55
Southwest	18.001.116,10
United Kingdom	3.883.042,96
Totale	77.548.570,20

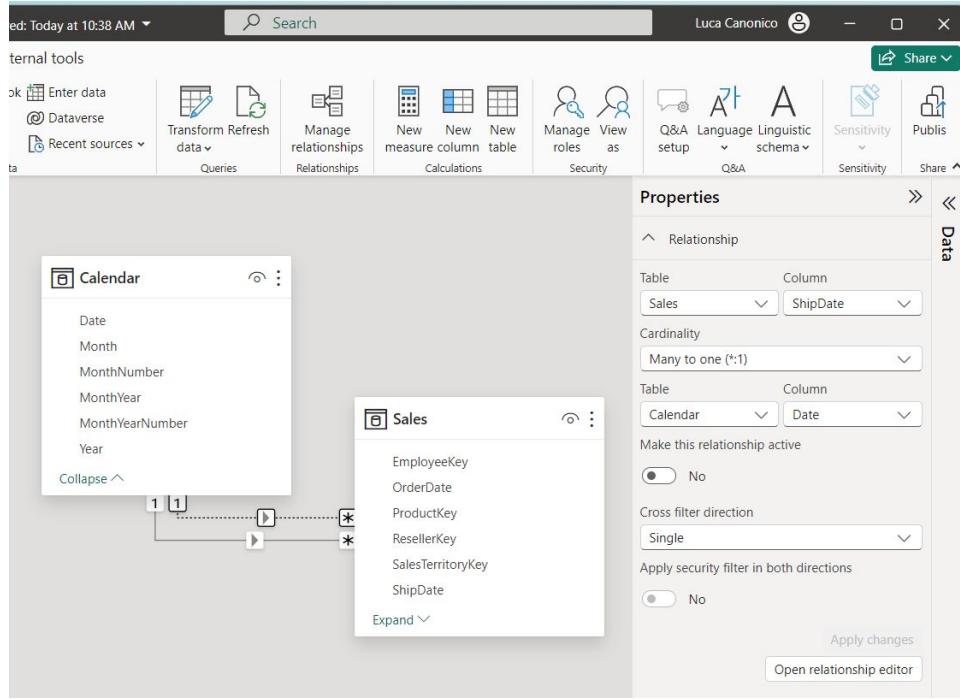
misura esplicita: aggregazione DAX



The screenshot shows the Power BI Data View. A context menu is open over a table named 'Region'. The menu path 'Visualizzazioni > SUM(Sales[SalesAmount])' has been selected. The table below shows the sales amount for each region, with a total row at the bottom.

Region	Sales
Australia	1.391.024,85
Canada	13.875.632,96
Central	7.633.386,86
France	4.527.839,69
Germany	1.877.743,39
Northeast	6.715.354,39
Northwest	12.004.822,45
Southeast	7.638.606,55
Southwest	18.001.116,10
United Kingdom	3.883.042,96
Totale	77.548.570,20

Role-playing Dimensions



È possibile definire una sola relazione attiva tra due tabelle di Power BI.

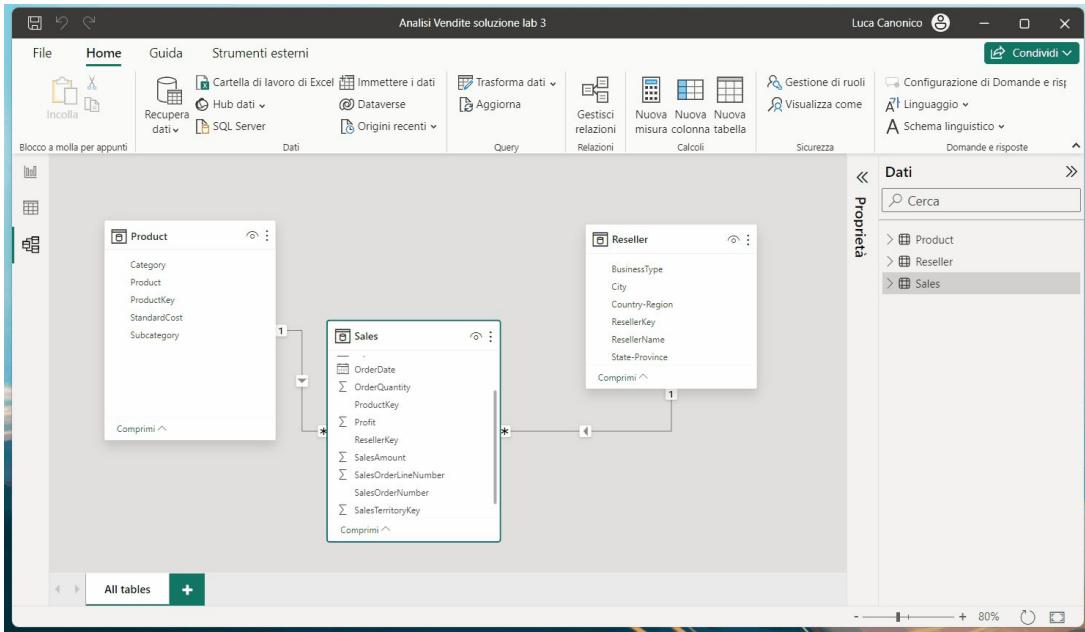
Eventuali altre relazioni sono inattive.

Una relazione può essere utilizzata 'forzosamente' utilizzando espressioni DAX (che vedremo più avanti!)

In questo scenario, la tabella **Calendar** potrebbe filtrare la tabella **Sales** sul campo **OrderDate** (è necessario analizzare le vendite in base alla data di registrazione degli ordini) oppure sul campo **ShipDate** (è necessario analizzare le vendite in base alla data di spedizione degli ordini).

Questo scenario verrà approfondito in seguito!

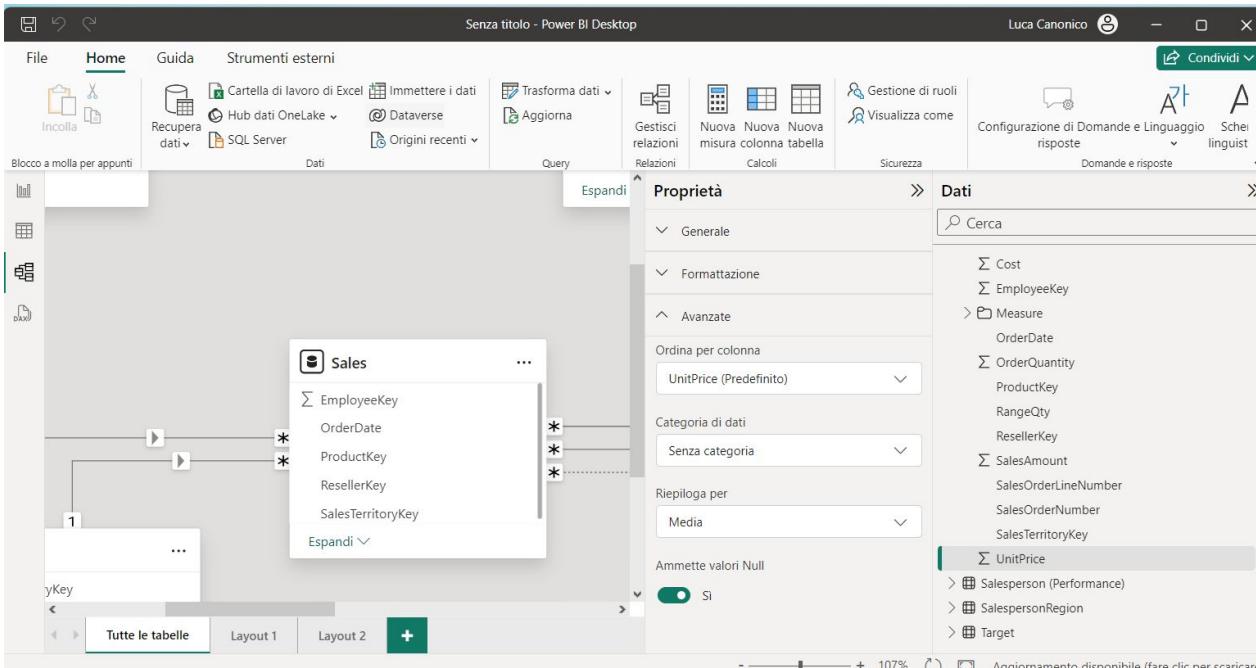
Proprietà di campi



Best Practice:

- definire correttamente la modalità di aggregazione dei campi utile. Altrimenti sui campi numerici (sono tutti quelli preceduti dal simbolo della sommatoria) viene applicata di default una somma oppure un conteggio (se il campo contiene valori progressivi).
- categorizzazione dei campi (utile soprattutto per le mappe)

Proprietà di campi: modalità di aggregazione dei campi



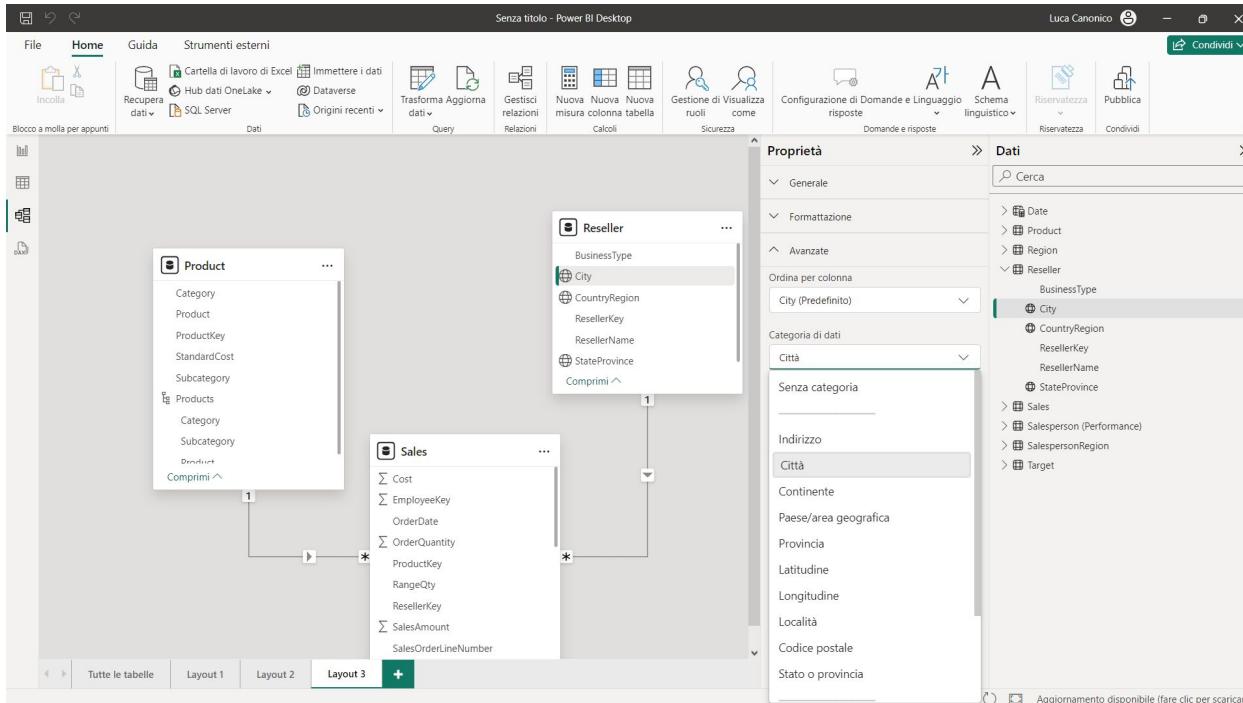
Best Practice:

- definire correttamente la modalità di aggregazione dei campi utile

Visualizzazione modello → (selezionato il campo di cui gestire le proprietà) → Proprietà → Avanzate → Riepiloga per → ...

In questo modo, i valori del campo verranno aggregati in base alla modalità di aggregazione indicata.

Proprietà di campi: modalità di aggregazione dei campi



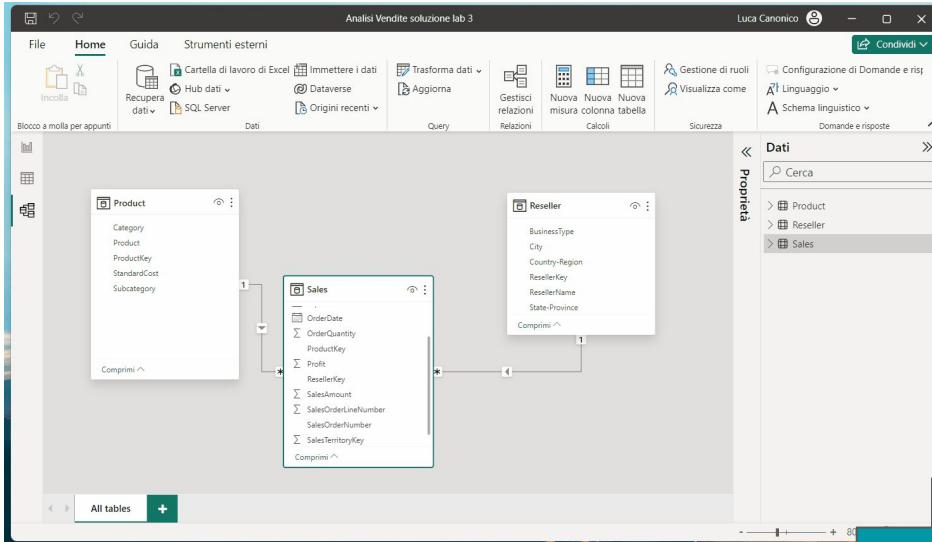
Best Practice:

- categorizzazione dei campi

Visualizzazione modello → (selezionato il campo di cui gestire le proprietà) → Proprietà → Avanzate → Categoria di dati → ...

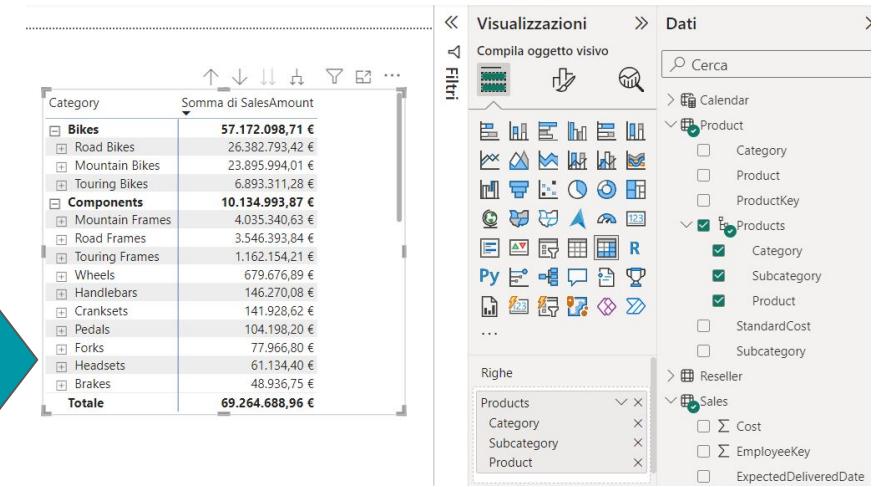
In Power BI Desktop è possibile classificare i dati geografici per una colonna, in modo che Power BI Desktop sappia come trattare i valori negli oggetti visivi in un report.

Gerarchia



Ogni dimensione contiene generalmente almeno una gerarchia!

Una gerarchia descrive i diversi livelli di dettaglio della dimensione rispetto i quali è possibile filtrare ed aggregare una metricali!



Category	Somma di SalesAmount
Bikes	57.172.098,71 €
Road Bikes	26.382.793,42 €
Mountain Bikes	23.895.994,01 €
Touring Bikes	6.893.311,28 €
Components	10.134.993,87 €
Mountain Frames	4.035.340,63 €
Road Frames	3.546.393,84 €
Touring Frames	1.162.154,21 €
Wheels	679.676,89 €
Handlebars	146.270,08 €
Cranks	141.928,62 €
Pedals	104.198,20 €
Forks	77.966,80 €
Headsets	61.134,40 €
Brakes	48.936,75 €
Totale	69.264.688,96 €

Check Point!

- Quali sono le due tipologie di tabelle in uno Star Schema?
- Qual è la differenza tra le due?
- Cosa si intende per cardinalità di una relazione?
- Quali sono alcune delle proprietà più utili da definire per campi e tabelle in PBI Desktop?



GRAZIE
EPCODE