

Le aggregazioni in Power BI

di [Francesco Bergamaschi](#)

Pubblicato il 23 Marzo 2023

Vediamo **come gestire i report in Power Bi** quando le **tabelle sono troppo estese**: la **soluzione** canonica è costituita dalle **aggregazioni in Power BI**. Come funzionano?

Nel [precedente articolo](#) dal titolo, Luca ha identificato delle **strategie per evitare l'uso di DirectQuery nel caso di tabelle dei fatti molto estese** - tanto da non trovare **posto in RAM** e dunque che non possono avere la modalità di connessione preferita e più performante in Power BI: la modalità Import.

Quando la tabella è troppo grande

In dettaglio, Luca ha pensato che, nel caso, di una **tabella dei fatti troppo grande per l'Import**, le **opzioni di connessione** siano le seguenti:

- **DirectQuery**. In questo modo tutti i dati al massimo livello di dettaglio e di estensione temporale sarebbero trattati. Tuttavia, le prestazioni sarebbero molto più lente che in Import;
- Ridurre la granularità (cioè il livello di dettaglio) dei dati, in modo da importare tutto ma in modo meno dettagliato, ad esempio rinunciando ad avere il dettaglio di ogni scontrino e andando al livello giornaliero o mensile.
Poi bisognerà decidere cosa mostrare per ogni mese: le vendite totali mensili senza alcuna informazione dei prodotti o per ogni mese dettagliare i prodotti e le relative vendite? O le famiglie di prodotti e le relative vendite?



- Mantenere la granularità ma ridurre l'estensione temporale dei dati (ultimi tre anni? Ultimi cinque?)
- Un mix tra 2) e 3)

Supponiamo che solo la 1) sia possibile perché il cliente chiede di avere tutti i dettagli e tutta la storia dei dati (o perché, pur riducendo i dettagli e/o l'estensione temporali dei dati, le dimensioni della tabella restano oltre il limite di spazio disponibile in RAM per l'Import), **che fare per evitare di creare un report lento?**

La soluzione è costituita dalle aggregazioni in Power BI. Di che si tratta?

Si tratta di potere avere nel proprio modello dati in Power BI Desktop, oltre alla tabella dei fatti troppo grande per essere importata, una **seconda copia della stessa tabella dei fatti**, ma **aggregata**. La prima, per forza di cose, **sarà in DirectQuery**. La seconda, sarà **in Import**. Si veda l'immagine 1.

Immagine 1



Quindi, Luca conclude, **Power BI permette di avere tabelle con diverse modalità di connessione in uno stesso modello dati**, non male!

La tabella aggregata sarà, per forza di cose, **più piccola** rispetto alla tabella dei fatti in DirectQuery (altrimenti anche questa dovrebbe essere in DirectQuery).

Per renderla **più piccola**, sarà aggregata in Power Query soltanto per **data, cliente, territorio e prodotto** (di fatto collassando diverse righe di vendita dello stesso giorno in una singola riga).

Supponendo di avere circa 1.000 transazioni al giorno, questo vuol dire ridurre il numero di righe di circa 1.000 volte. Infine, **questa tabella aggregata sarà in Import (e non in DirectQuery)**.

In questo modo, questa tabella aggregata **sarà ancora collegabile con tutte le dimensioni del modello** come visibile in figura 1.

Non è detto che questa cosa sia sempre possibile: magari la tabella aggregata risulta ancora **troppo grande per l'Import** e, dunque, **bisognerà perdere qualche dettaglio in più**, per esempio il prodotto e il cliente e aggregare soltanto per territorio e data.

La tabella aggregata, in questo caso, **potrà essere connessa soltanto alle dimensioni SalesTerritory e Calendar**.

In questo caso **nulla vieta di generare una seconda tabella aggregata, in cui aggregare per Customer e Product** (Immagine 2). E se ne potrebbero creare anche tre o quattro o un numero qualunque.

Immagine 2



Come questa architettura renderà più veloce il report?

Succederà quanto segue: quando la visualizzazione di Power BI richieda un livello di dettaglio compatibile con quello delle tabelle aggregate, che sono in Import, Power BI userà quelle tabelle e così la performance sarà quella velocissima dell'Import.

Quando, invece, il livello di dettaglio richiesto è più spinto, non c'è alternativa ad usare la tabella in DirectQuery, con la risultante performance un po' degradata (si ricordi, tuttavia, che ciò permette di gestire un qualunque numero di righe in una tabella).

In questo modo si ottiene il meglio da entrambi i mondi (Import e DirectQuery) quando la tabella dei fatti può essere soltanto in DirectQuery!

Resta il forte suggerimento di evitare DirectQuery tutte le volte in cui non sia strettamente necessario, in modo da **usare Import nella stragrande maggioranza dei casi** e, così, avere sempre report molto performanti che danno all'utente finale un'esperienza piacevole di interazione con i report!

Al prossimo articolo!

A cura di Francesco Bergamaschi

Giovedì 23 marzo 2023

POWER B.I.

Controllo di gestione con la business intelligence per le aziende

PERCORSO ONLINE ON DEMAND

OLTRE 14 ORE DI LEZIONE DIVISE IN 5 MODULI TEMATICI

Introduzione al corso

Un percorso alla scoperta di **POWER BI**, l'**ecosistema di Microsoft per la Business Intelligence**.

Una **guida pratica e operativa** per apprendere come gestire e diventare autonomi in tutto il flusso di lavoro.

Passo dopo passo, partendo da zero, si esplorerà dettagliatamente sia la parte Desktop (per la **connessione ai dati**, la modellazione e la **creazione dell'analisi**) che la parte del **Cloud** (pubblicazione, condivisione, gestione delle aree di lavoro, implementazione della sicurezza e dei ruoli).

Il Corso prevede **oltre 14 ore formative** suddivise in 5 moduli tematici e può essere fruito *on demand*, dove e quando vuoi.



[SCOPRI DI PIU' >](#)