

Power BI

Giorno 1

Agenda

Cosa vedremo

- ❑ Introduzione a Power BI
- ❑ Acquisizione, profilazione, preparazione e caricamento dei dati



Agenda

Obiettivi di apprendimento

- ❑ Descrivere Power BI e le sue parti fondamentali
- ❑ Acquisire e profilare i dati
- ❑ Scegliere la giusta modalità di storage dei dati

Introduzione a Power BI

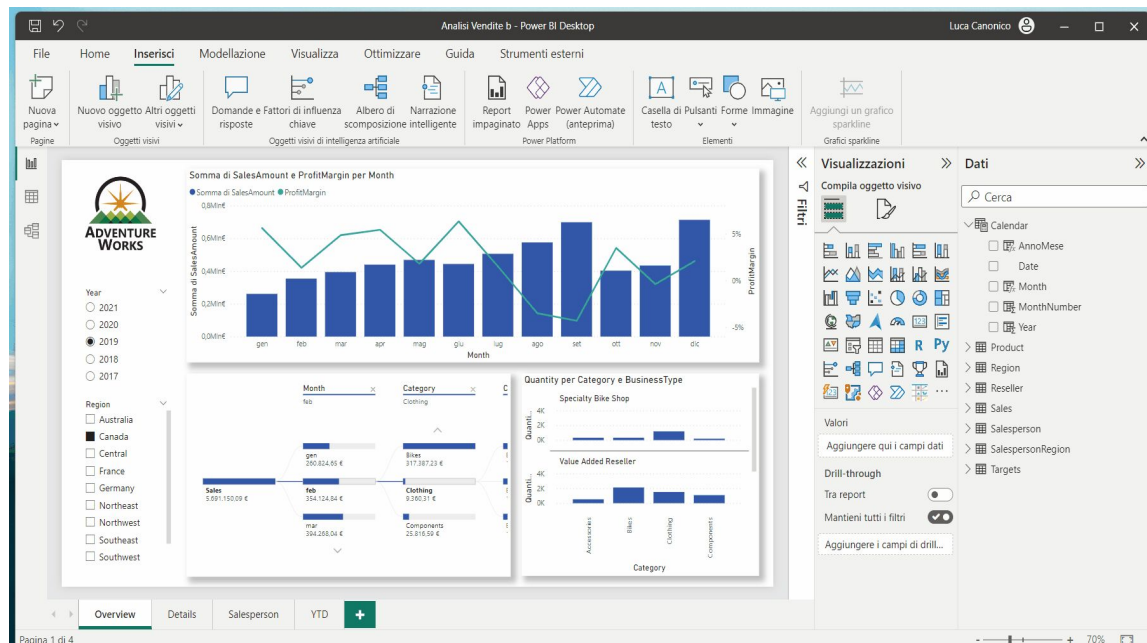
Business Intelligence

È un insieme di processi, metodologie e software a supporto del processo decisionale aziendale.

Lo scopo principale di una soluzione di BI è la trasformazione dei dati disponibili in informazioni, e quindi in conoscenza a supporto di un *decision maker*.

Il dato diventa un'informazione quando suggerisce un'azione. Quali sono gli aspetti sfidanti?

- La raccolta e la 'ristrutturazione' del dato
- Fare storytelling



Quali sono i task di un Data Analyst?

Comprensione dello scenario

È la fase di sviluppo durante la quale si raccolgono i requisiti, si individuano le sorgenti dati e si esplorano i dati per acquisire familiarità con il contesto di analisi...

Preparazione del dato

È un processo complesso che comprende l'analisi della qualità dei dati, la verifica dell'integrità dei dati, la correzione di errori, la ristrutturazione dei dati per ottimizzarne la leggibilità e le aggregazioni...

Costruzione di un modello logico

Costruire un modello logico implica la definizione delle relazioni tra tabelle...

Visualizzazione e Analisi del dato

Un report di BI deve consentire un'esplorazione efficace ed efficiente dello scenario di analisi...

Gestione della soluzione

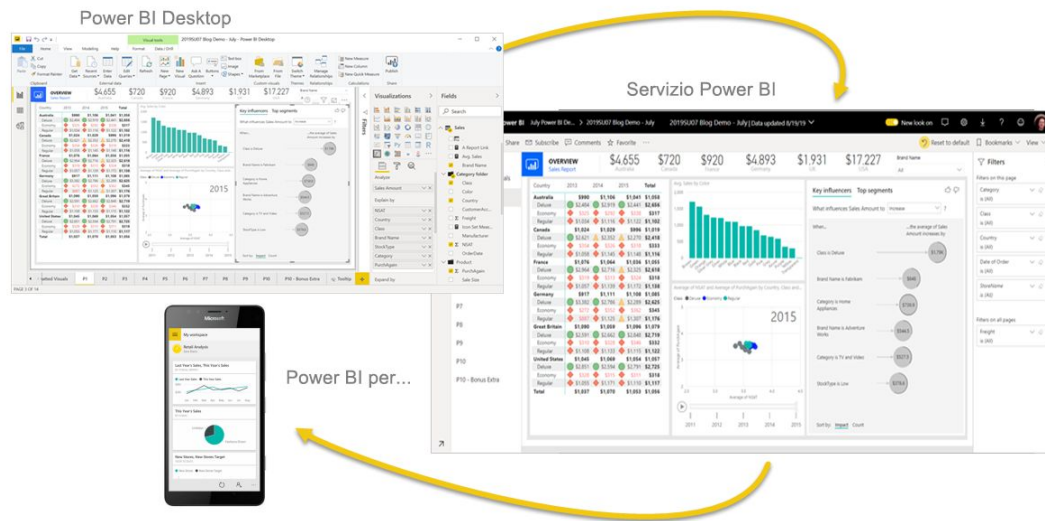
La soluzione di BI deve essere gestita in termini di aggiornamento dei dati sottostanti, distribuzione, accessi, ...

Che cos'è Power BI?

Power BI è una raccolta di servizi software, app e connettori che operano congiuntamente per trasformare origini dati non correlate in informazioni dettagliate, coerenti, visivamente coinvolgenti e interattive.

Gli elementi base di Power BI includono:

- Un'applicazione desktop di Windows, **Power BI Desktop**
- Un servizio SaaS (Software as Service), **Power BI Service**
- Un'applicazione per device mobile, **Power BI Mobile**



<https://learn.microsoft.com/it-it/power-bi/fundamentals/power-bi-overview>

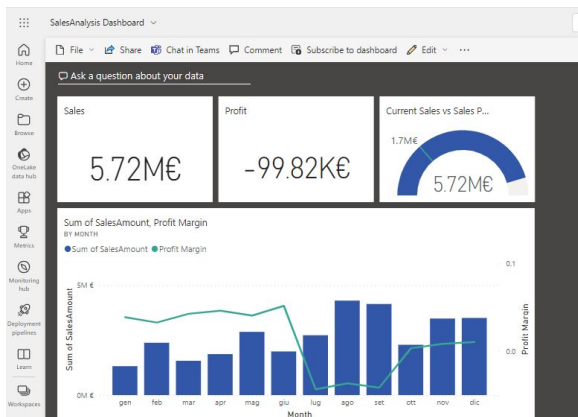
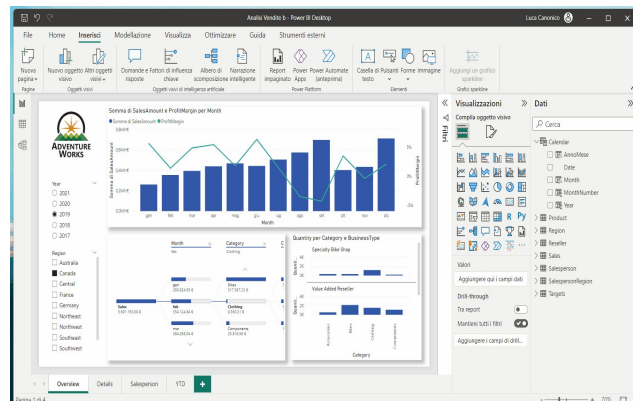
... un po' di nomenclatura!

Che cos'è una visualization?
È un singolo oggetto grafico
(line chart, pie chart, ...)

Che cos'è un report?
È un'insieme di visualization
organizzato in una o più pagine.
Il report è interattivo, dinamico,
filtrabile,...

Che cos'è una dashboard?
È un insieme selezionato di
visualization di più report riuniti
in un'unica pagina!

... definiamo **Tile** la singola
visualizzazione della dashboard.



La dashboard è statica
ed espone dati
tipicamente molto
aggregati.

Qual è il flusso tipico di attività in PBI?

1. Acquisizione dati in Power BI Desktop
2. Creazione report in Power BI Desktop
3. Pubblicazione del report in Power BI Service
4. Condivisione di report e dashboard

I dati possono provenire da sorgenti eterogenee:

- File csv
- Cartelle di lavoro Excel
- Database on-premise
- Database in cloud
- Liste Sharepoint
- ...



Esplora il report

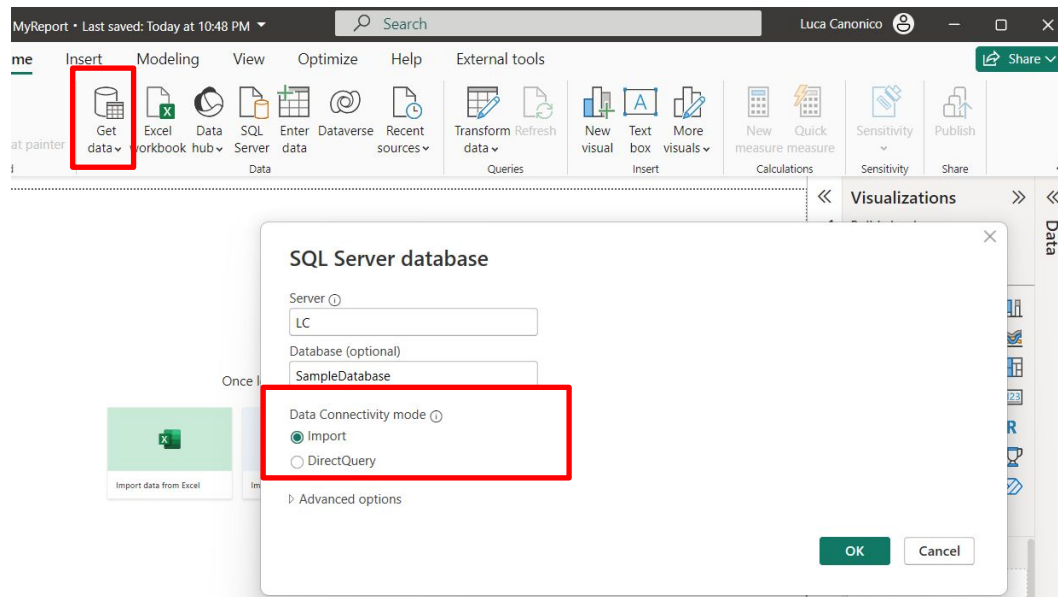
Modalità di storage del dato

Import

- Power BI importa effettivamente i dati creandone una copia locale.
- I dati sono memorizzati all'interno del file di Power BI (**pbix**).

DirectQuery

- Power BI non crea nessuna copia locale dei dati.
- I dati sono interrogati direttamente nella loro sorgente.



Modalità di storage del dato

Import

- Esperienza utente ottimizzata
- Sono garantite tutte le funzionalità di Power BI
- Aggiornamento dati manuale o schedulato
- Limiti sul volume dati (dipende dalla licenza)

DirectQuery

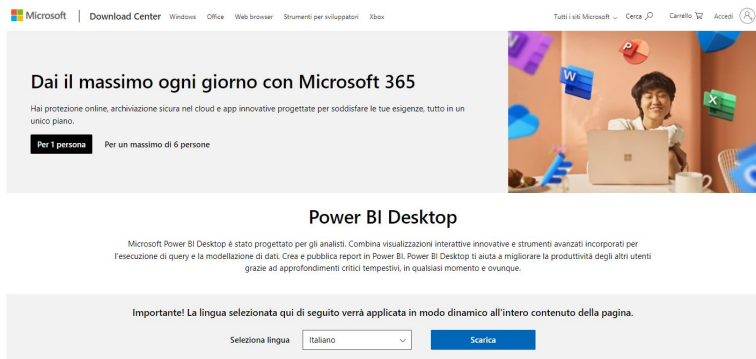
- L'esperienza utente può' essere degradata in termini di performance
- Non sono garantite tutte le funzionalità di Power BI (limiti nelle trasformazioni e funzioni DAX)
- Report real-time
- Utile per la gestione di grossi moli dati



Installazione di Power BI Desktop

Per scaricare il file eseguibile di Power BI Desktop dall'Area download, accedi alla pagina dedicata

Area download



Microsoft | Download Center | Windows | Office | Web browser | Strumenti per sviluppatori | Xbox

Tutti i siti Microsoft | Cerca | Canale | Accedi

Dai il massimo ogni giorno con Microsoft 365

Hai protezione online, archiviazione sicura nel cloud e app innovative progettate per soddisfare le tue esigenze, tutto in un unico piano.

Per 1 persona Per un massimo di 6 persone

Power BI Desktop

Microsoft Power BI Desktop è stato progettato per gli analisti. Combina visualizzazioni interattive innovative e strumenti avanzati incorporati per l'esecuzione di query e la modellazione di dati. Crea e pubblica report in Power BI. Power BI Desktop ti aiuta a migliorare la produttività degli altri utenti grazie ad approfondimenti critici tempestivi, in qualsiasi momento e ovunque.

Importante! La lingua selezionata qui di seguito verrà applicata in modo dinamico all'intero contenuto della pagina.

Seleziona lingua: Italiano Scarica

Specifica quindi il file di installazione a 32 bit o a 64 bit da scaricare.

Scegli il download che desideri

<input type="checkbox"/> Nome file	Dimensioni
<input type="checkbox"/> PBIDesktopSetup.exe	430.6 MB
<input checked="" type="checkbox"/> PBIDesktopSetup_x64.exe	479.0 MB

Scarica

Dimensioni totali: 479.0 MB

Viene richiesto di eseguire il file di installazione dopo aver completato il download.

Come ottenere Power BI Desktop

Check Point!

Quali sono i componenti principali di Power BI?

- ☐ SQL Server, PBI desktop, App Mobile
- ☐ PBI desktop, Power BI Service
- ☐ PBI desktop, Power BI Service, Power BI Mobile

Che cos'è un report di Power BI?

- ☐ Un grafico che espone un'informazione aggregata
- ☐ Un canvas che espone informazioni sintetiche e rilevanti
- ☐ Una collezione di visualization disposta in una o più pagine

Qual è il task più critico in fase di sviluppo?

- ☐ È la fase di pulizia e trasformazione dei dati
- ☐ È la fase di modellazione dei dati

Il dataset aziendale deve essere aggiornato ogni ora. È necessario sviluppare un report che esponga i dati con il minor ritardo possibile. Quale opzione di storage del dato soddisfa il requisito?

- ☐ Import mode
- ☐ DirectQuery mode

Power Query Editor

Recupero dati in Power BI Desktop

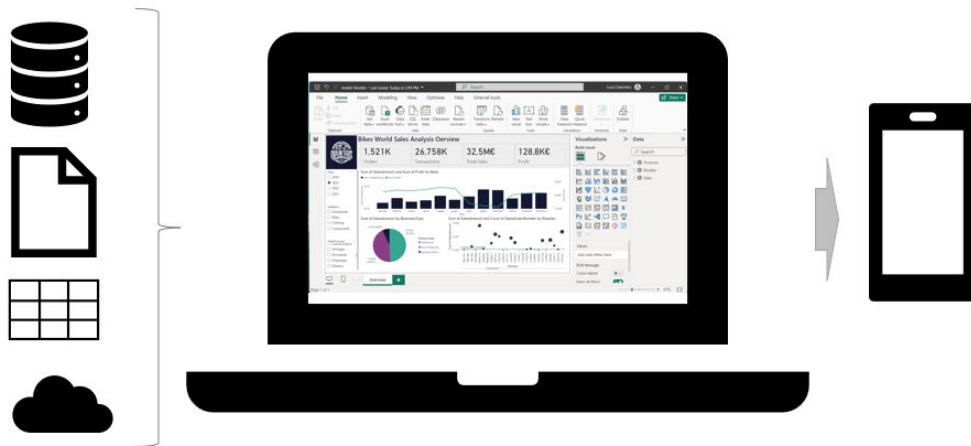
Il primo task nel processo di sviluppo è l'identificazione e il recupero dati.

È possibile recuperare dati da sorgenti on-premise (database, cartelle di lavoro Excel, csv, txt, ...), da sorgenti cloud (cloud database, sharepoint, ...).

Dopo la pulizia e la 'ristrutturazione' dei dati è possibile creare report in Power BI.

Il tool utile alla profilazione, pulizia e preparazione dei dati è il Power Query Editor.

Il Power Query Editor è uno strumento fondamentale utilizzabile come strumento di ETL in diversi prodotto/servizi Microsoft (Power Platform, Azure Data Factory, Fabric, ...)



Flat files, Excel

Le aziende molto spesso esportano e archiviano dati in file.

Tra i formati file più comuni ci sono i flat file (.csv e .txt).

Un'altra origine comune è rappresentata da cartelle di lavoro Excel (.xlsx).

Quali sono le criticità per questa tipologia di sorgente dati?

- 1) il percorso del file (file path)
- 2) modifiche strutturali al file sorgente (l'eliminazione o la ridenominazione di campi possono rompere il report)

ResellerSalesTargets.csv

File Origin: 1200: Unicode | Delimiter: Comma | Data Type Detection: Based on first 200 rows

Year	EmployeeID	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12
2017	61161660	-	-	-	-	-	-	200	400	600	400	800	800
2017	90836195	-	-	-	-	-	-	100	200	300	400	400	500
2017	112432117	-	-	-	-	-	-	500	1500	1000	1000	2200	1750
2017	139397894	-	-	-	-	-	-	100	200	300	300	300	500
2017	191644724	-	-	-	-	-	-	100	450	500	200	750	750
2017	234474252	-	-	-	-	-	-	100	200	300	300	300	300
2017	399771412	-	-	-	-	-	-	100	200	300	300	300	300
2017	502097814	-	-	-	-	-	-	500	1500	1000	1000	2200	1750
2017	615389812	-	-	-	-	-	-	100	100	100	200	200	200
2017	716374314	-	-	-	-	-	-	100	200	300	100	300	500
2017	841560125	-	-	-	-	-	-	100	500	400	400	1000	1000

Buttons: Extract Table Using Examples, Load, Transform Data, Cancel

Untitled - Power Query Editor

Queries [1] | ResellerSalesTargets

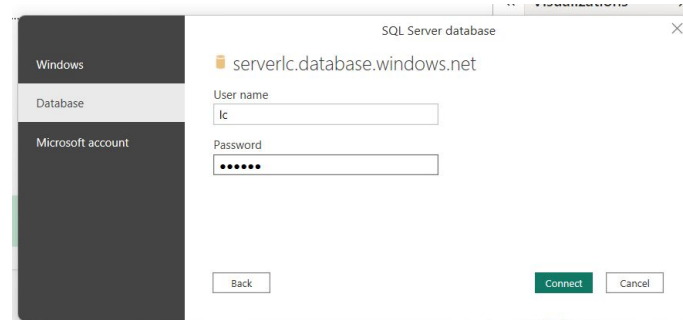
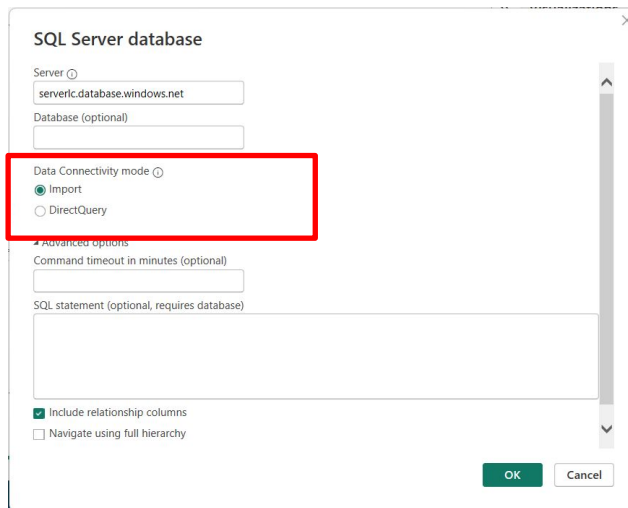
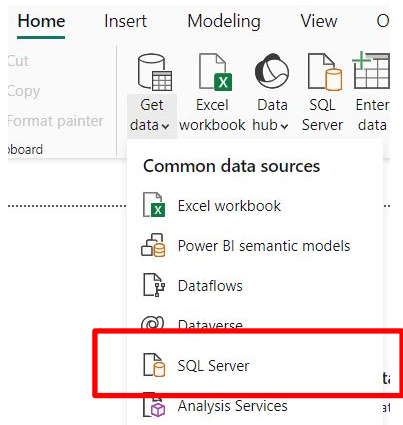
Formula Bar: = Csv.Document(File.Contents("C:\Users\LucaCanonica\Desktop\ResellerSalesTargets.csv"), [Delimiter=";", Columns=14, Encoding=1200, QuoteStyle=QuoteStyle.None])

Error Message: DataSourceError: Could not find file 'C:\Users\LucaCanonica\Desktop\ResellerSalesTargets.csv'. Details: C:\Users\LucaCanonica\Desktop\ResellerSalesTargets.csv

Query Settings: Name: ResellerSalesTargets

APPLIED STEPS: Source, Promoted Headers, Changed Type

Database relazionali

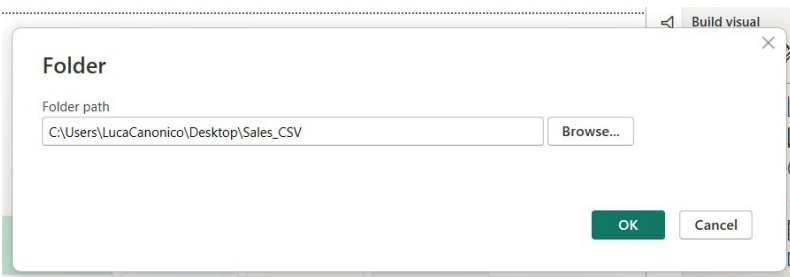


Con Power BI Desktop è possibile connettersi a diversi database relazionali (e non) che siano in cloud oppure on-premises. La connessione ad un database relazionale richiede la scelta della modalità di storage dei dati e la selezione delle tabelle o entità che bisogna acquisire. In alternativa, è possibile scrivere una query custom (in quest'ultimo caso è obbligatoria l'indicazione del database).

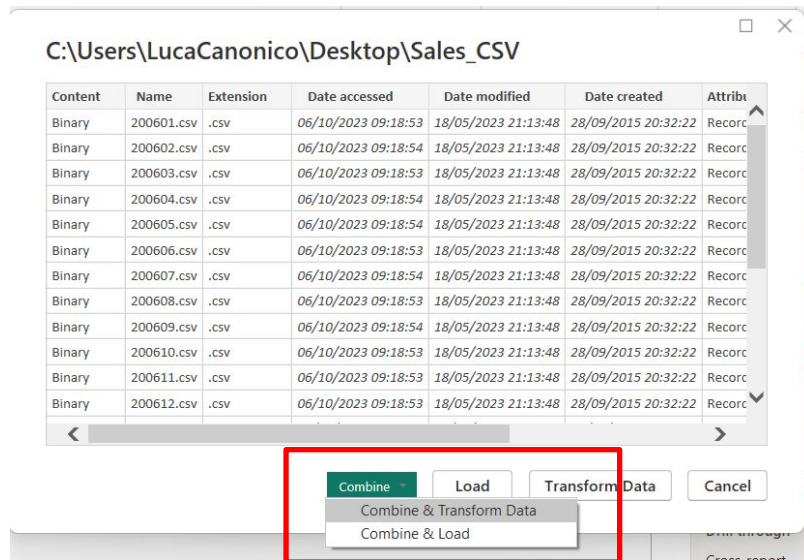
Folder

Un'altra opzione utile è la connessione ad un Folder.

La connessione ad un folder consente di eseguire un append dei file contenuti al suo interno.



Scenari tipici: estrazioni periodiche di file csv con la medesima struttura in uno specifico folder...



Profilazione campi

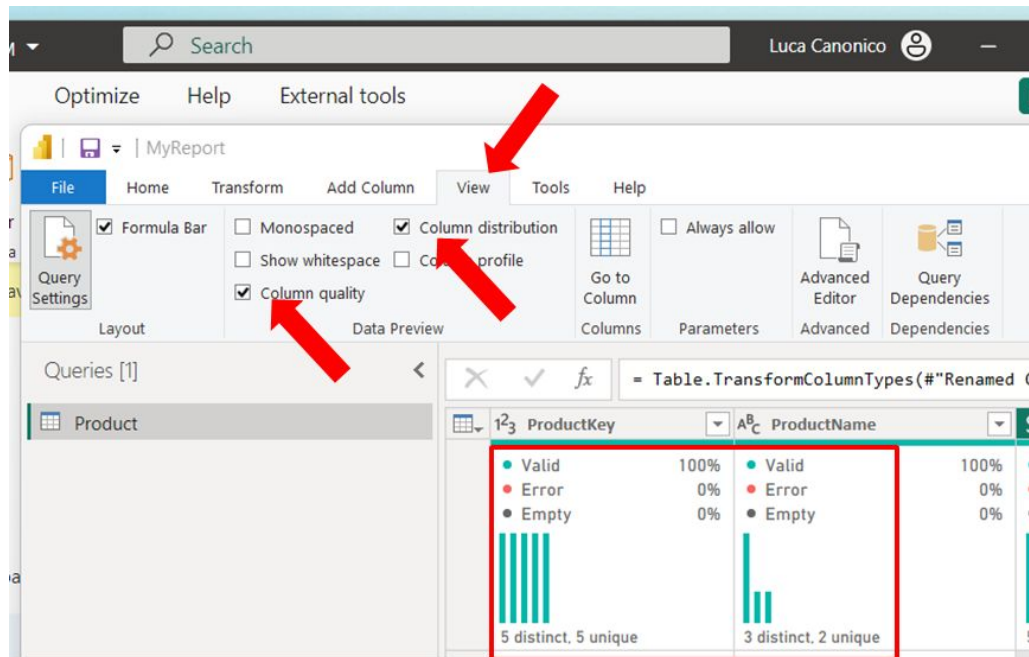
Column quality

Espone gli indicatori: Valid, Error, Empty

Column distribution

Espone la frequenza e la distribuzione dei valori per ciascuno dei campi della query (tabella) in esame. In particolare, fornisce il conteggio dei valori distinti (conteggio dei valori diversi) e dei valori univoci (conteggio dei valori che non hanno ridondanze).

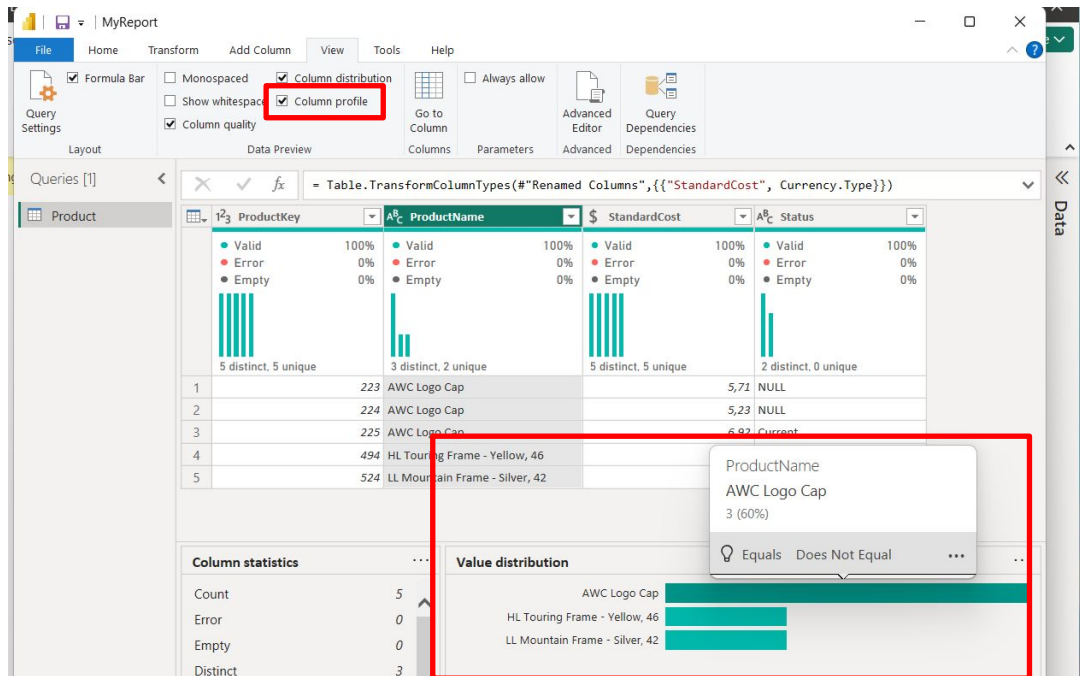
Cosa si desume valutando la profilazione nell'esempio?



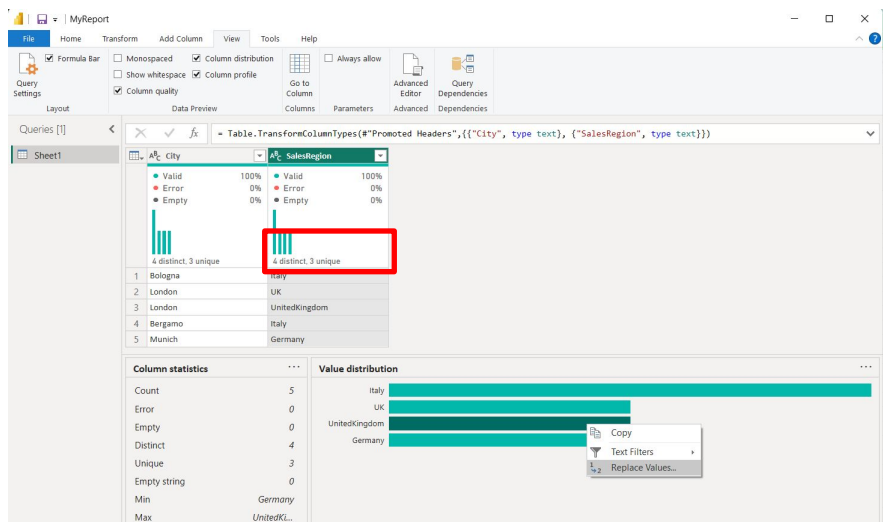
Profilazione campi

Dall'esempio si evince che:

- Ci sono 5 record
- Il campo *ProductKey* è una PK
- Il campo *ProductName* contiene 3 valori diversi (su un totale di 5 righe).
- Il campo *ProductName* contiene 2 valori univoci (su un totale di 5 righe). In altre parole, 2 valori su 5 non hanno duplicati (non sono ripetuti)
- Esiste 1 *ProductName* ridondante (conteggio distinct - conteggio unique = $3 - 2 = 1$)
- Il Column profile indica quante volte è ripetuto ciascun valore nella colonna selezionata



Profilazione campi e sostituzione valori errati



I dati devono essere consistenti!

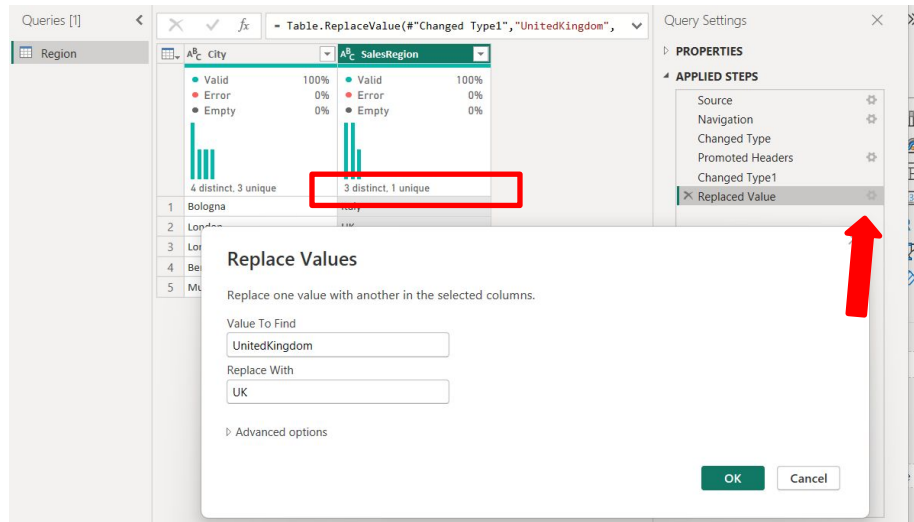


Table: Table.ReplaceValue(#"Changed Type1", "UnitedKingdom",

Replace Values

Replace one value with another in the selected columns.

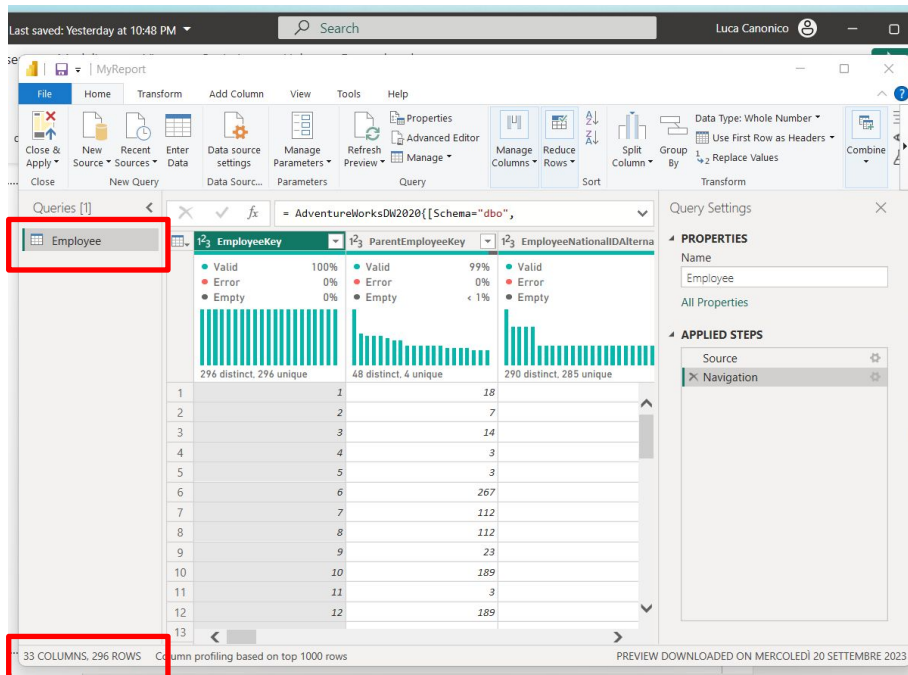
Value To Find
UnitedKingdom

Replace With
UK

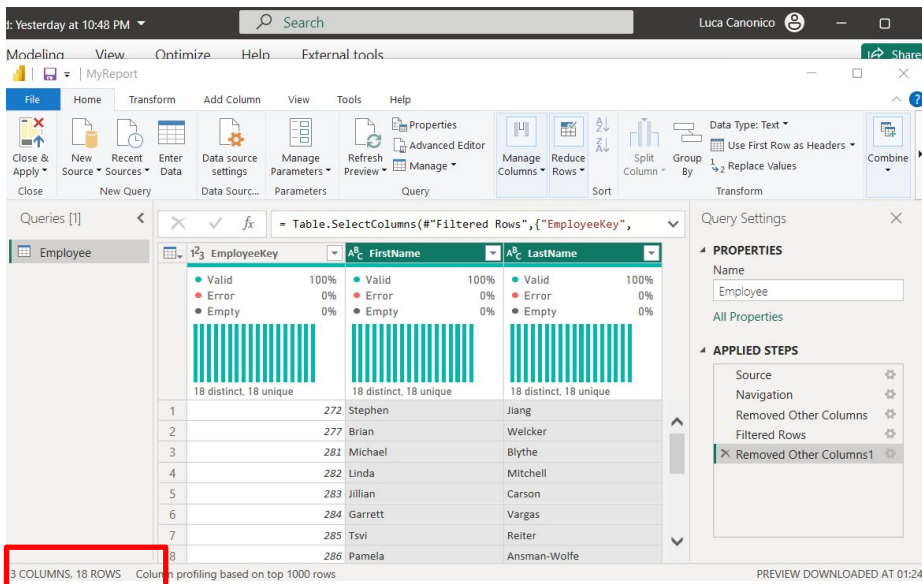
Advanced options

OK Cancel

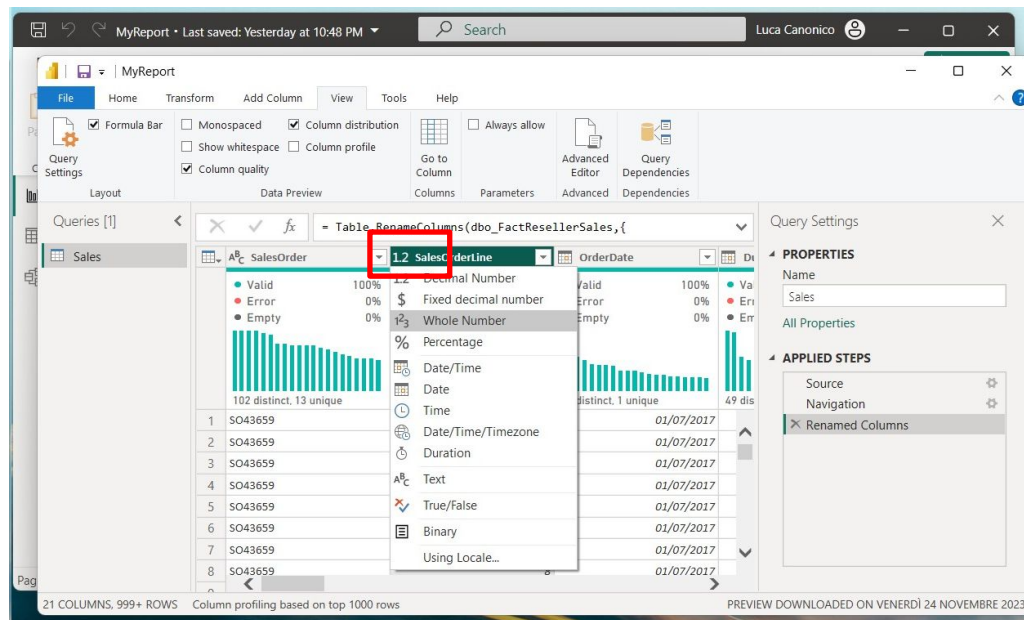
Rimuovere campi e filtrare righe



L'obiettivo è minimizzare il volume del dato coerentemente ai requisiti di analisi!



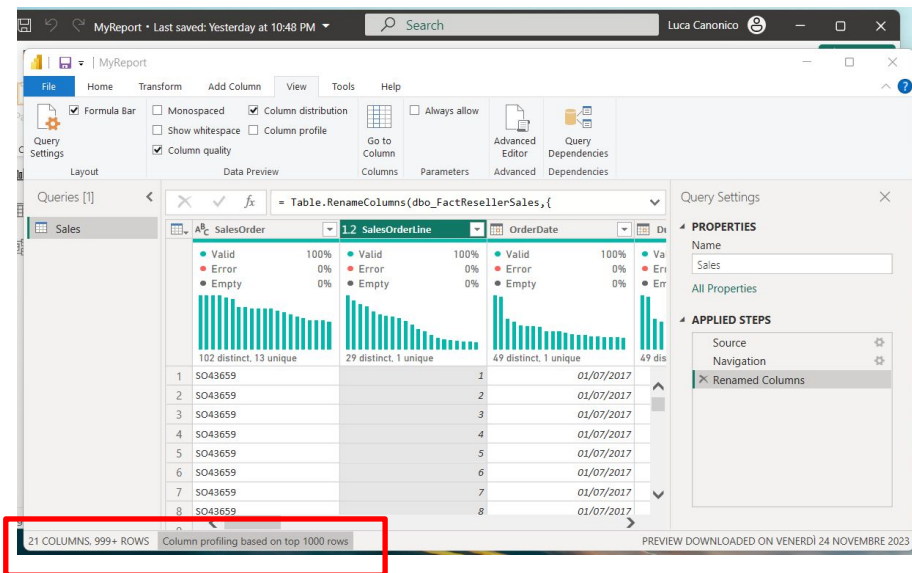
Rimuovere campi e filtrare righe



The screenshot shows the Epicode software interface. The 'Rename Columns' dialog box is open, and the '1.2 SalesOrderLine' column is selected. The 'Decimal Number' data type is highlighted in the list. The background shows a data table with columns like 'SalesOrder', 'SalesOrderLine', and 'OrderDate'.

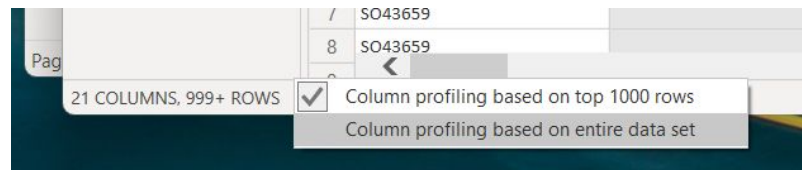
Assegnare il *data type* corretto a ciascun colonna coerentemente all'informazione descritta dalla colonna stessa!

Rimuovere campi e filtrare righe



In generale, fai attenzione al numero di record della tabella. Se il numero di record > 1000 abilita la profilazione sull'intero dataset in modo tale PBI profili i campi valutando l'intero set di dati!

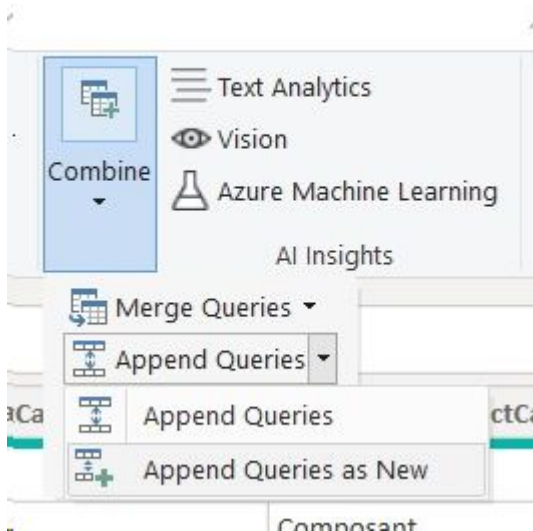
È limitato di default alle prime 1000 righe!!



Combinare più tabelle in un'unica tabella

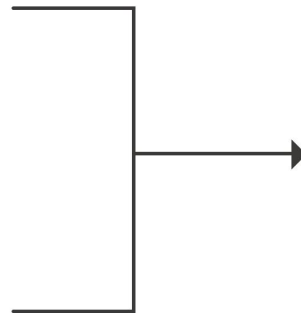
Append

È l'operazione con la quale è possibile 'accodare' i record di due o più tabelle.



A	B	C
1	1	1
2	2	2
3	3	3

A	B	D
4	4	4
5	5	5



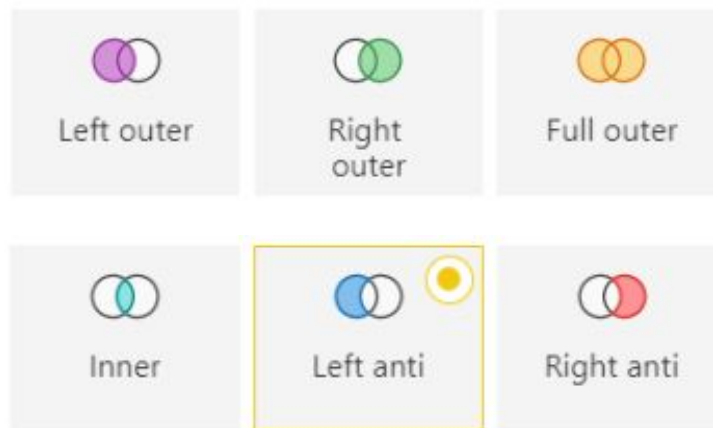
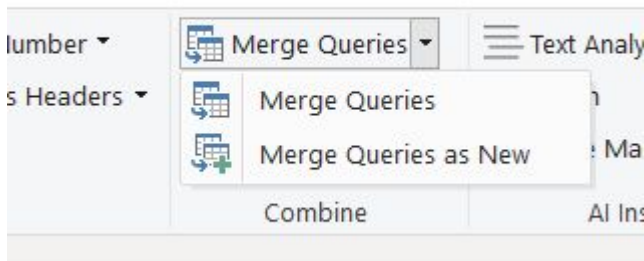
A	B	C	D
1	1	1	null
2	2	2	null
3	3	3	null
4	4	null	4
5	5	null	5

<https://learn.microsoft.com/en-us/power-query/append-queries>

Combinare più tabelle in un'unica tabella

Merge

È l'operazione con la quale è possibile combinare record di due tabelle in base alla corrispondenza dei valori di un campo in comune.



<https://learn.microsoft.com/it-it/power-query/merge-queries-overview>

Check Point!

Hai importato un tabella da un database SQL. Qual è il metodo più veloce per capire se ci sono campi con valori vuoti?

- ☐ Abilitare il column profile
- ☐ Abilitare il column quality
- ☐ Abilitare il column distribution

Hai importato una tabella di 2100 record. Bisogna verificare l'esistenza di campi chiave. Cosa fai?

- ☐ Abiliti il column profile e il column distribution
- ☐ Bisogna scrivere una query nella sorgente dati
- ☐ Nessuna delle precedenti

Hai importato due tabelle: Prodotti e CategorieProdotti. Vuoi ridurre il numero di tabelle per ottimizzare il modello. Quale operazioni esegui?

- ☐ Append
- ☐ Merge

**Let's take a
look!**



GRAZIE
Epicode