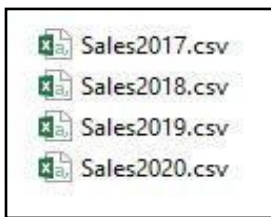


Laboratorio 0–Data Model (assieme in classe)

1. In questo Laboratorio inizieremo dal portare i dati all'interno di power query per trasformarli poi e vedremo come arrivare alla costruzione di un data model efficiente
2. Creare un nuovo file con Power BI Desktop

IMPORTAZIONE DEI FATTI

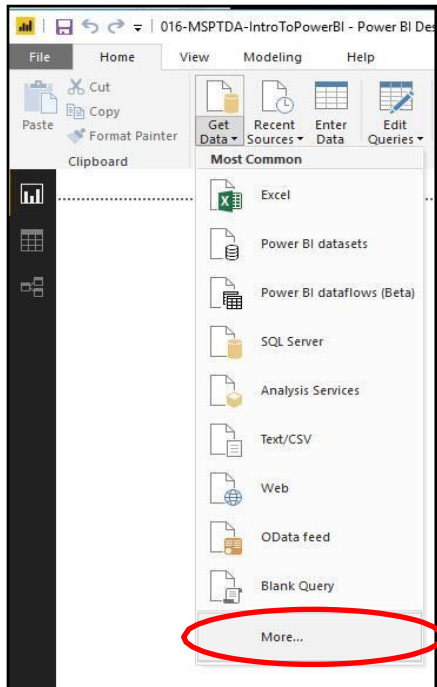
3. I dati della tabella dei **fatti** che dobbiamo importare provengono da più CSV (Valori separate da virgola) che si trovano nella cartella **dati**.
4. Dobbiamo importare queste tabelle e aggiungerle in un'unica tabella dei fatti.
5. I file che dobbiamo importare e aggiungere sono:



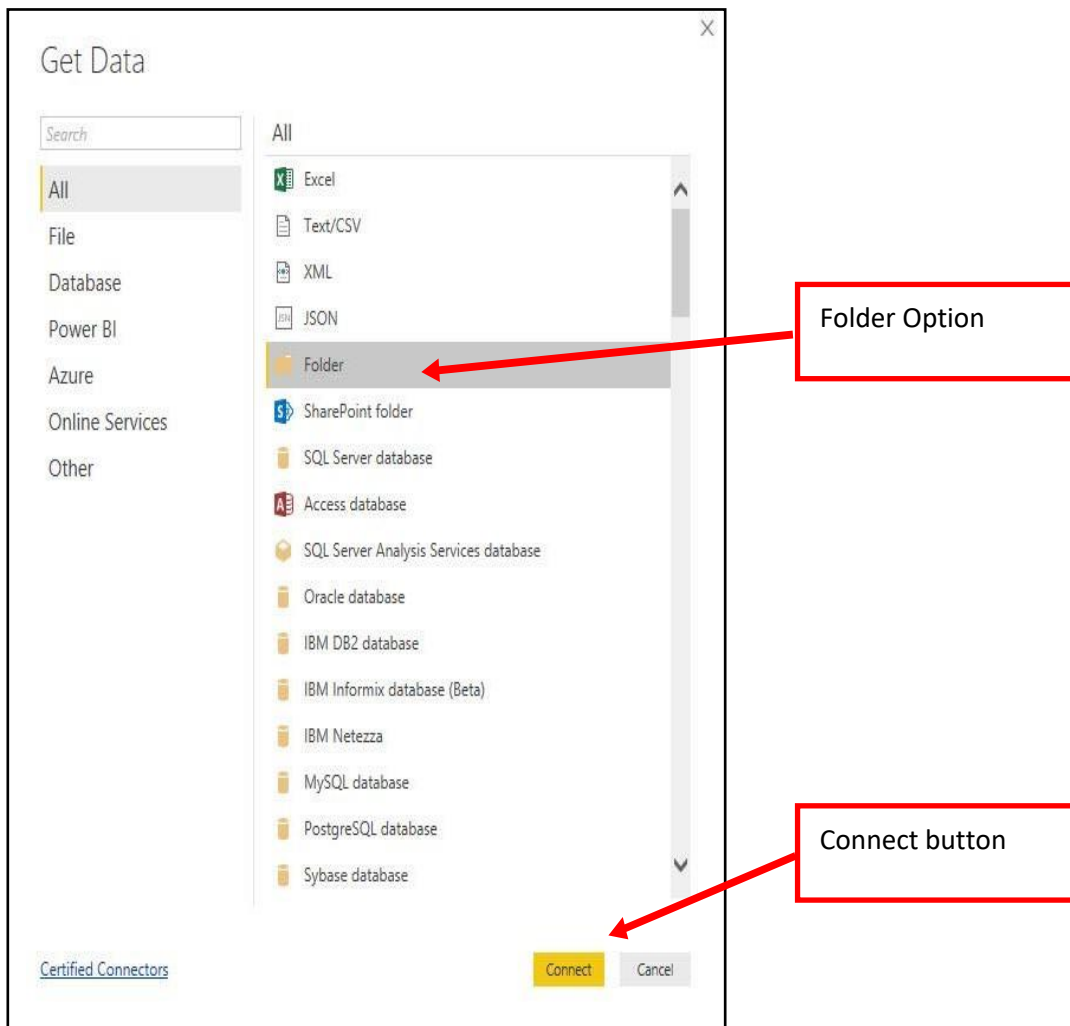
6. I singoli file CSV si presentano così:

	A	B	C	D	E	F	G
1	ISO Date	ProductID	SalesRep	UnitsSold	Discount	COGSTotal	
2	20171228	1	2	7	0.055	131.0582355	
3	20170118	7	4	95	0.4375	675.4277925	
4	20171012	12	2	72	0.4375	708.2377406	
5	20171212	1	1	4	0	74.89042027	
6	20171109	6	15	6	0.055	60.91124705	
7	20170124	7	1	72	0.4375	511.9031691	
8	20171119	10	6	8	0.055	76.829026	
257498	20170419	1	2	72	0.4375	1348.027565	
257499	20170428	2	22	96	0.4375	953.3091909	
257500							

7. Per iniziare la nostra importazione e trasformazione dei file CSV in un singolo set di dati corretti
Dalla ribbon Home, quindi nel gruppo "Extrnal data", facciamo clic sulla freccia a discesa "Get data", quindi facciamo clic su "More...."



8. Nella finestra di dialogo **"Get Data"** seleziona l'opzione **"Folder"**, quindi fai clic sul pulsante **"Connect"**..



9. Nella finestra di dialogo Cartella, vai alla **cartella "Dati"** corretta che contiene i quattro file CSV iniziali, quindi fai clic su OK



10. Nella finestra di dialogo successiva, fai clic su "**Combine & Transform data**"

Content	Name	Extension	Date accessed	Date modified	Date created	Attributes	
Binary	Sales2017.csv	.csv	09/04/2020 14:44:24	21/11/2018 17:14:08	09/04/2020 14:44:24	Record	D:_GIT_M
Binary	Sales2018.csv	.csv	09/04/2020 14:44:24	21/11/2018 17:14:08	09/04/2020 14:44:24	Record	D:_GIT_M
Binary	Sales2019.csv	.csv	09/04/2020 14:44:24	21/11/2018 17:14:10	09/04/2020 14:44:24	Record	D:_GIT_M
Binary	Sales2020.csv	.csv	09/04/2020 14:44:24	21/11/2018 17:14:10	09/04/2020 14:44:24	Record	D:_GIT_M

Combine & Transform Data
Transform Data
Cancel

11. La finestra di dialogo successiva chiede quale Delimitatore utilizzare che per noi sarà la virgola. Dopo aver selezionato il delimitatore, fare clic sul pulsante OK.

Combine Files

Specify the settings for each file. [Learn more](#)

Example File: First file

File Origin: 1252: Western European (Windows)

Delimitator

Comma

Data Type Detection: Based on first 200 rows

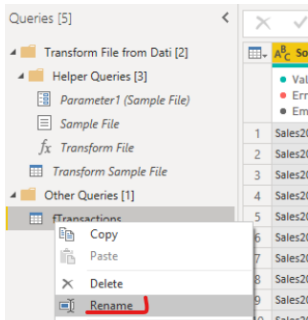
ISO Date	ProductID	SalesRep	UnitsSold	Revenue	ProfitTotal
20171228	1	2	7	0.055	131.0582355
20170118	7	4	95	0.4375	675.4277925
20171012	12	2	72	0.4375	708.2377406
20171212	1	1	4	0	74.89042027
20171109	6	15	6	0.055	60.91124705
20170124	7	1	72	0.4375	511.9031691
20171119	10	6	8	0.055	76.829026
20171223	1	2	84	0.4375	1572.698826
20171213	13	1	77	0.4375	1187.658002
20170610	6	17	108	0.475	1096.402447
20171025	1	3	60	0.4375	1123.356304
20170425	2	20	58	0.4375	575.9576362
20171015	11	11	108	0.475	2359.802639
20170510	2	4	96	0.4375	953.3091909
20171018	7	22	179	0.475	1272.648156
20170509	13	12	6	0.055	92.54477934
20171215	6	16	94	0.4375	954.2762037
20171016	1	7	12	0.375	149.1333622

☐ Skip files with errors
 OK
Cancel

12. Dopo aver selezionato il delimitatore e fare clic sul pulsante OK:

- Power Query crea una funzione Power Query ad-hoc che viene utilizzata per ogni file il contenuto
- Power Query unirà il risultato dell'estrazione da ciascun file in un unico singolo set di dati finale (tabella)

Rinominare il risultato in fTransactions



13. Le prime tre righe della tabella finale di **fTransactions** sono illustrate di seguito.

Ma dobbiamo eseguire delle attività di pulizia dei dati:

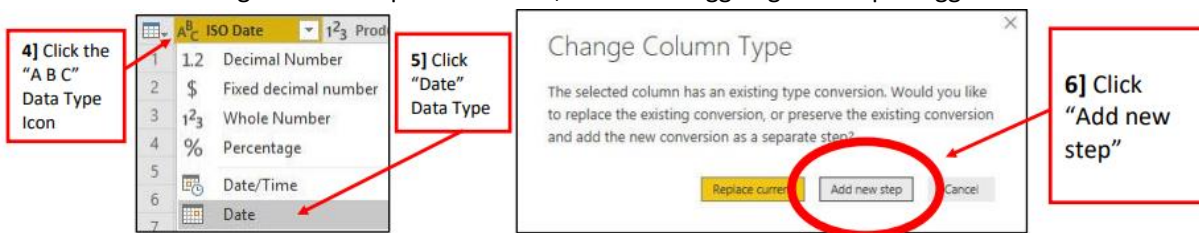
	Source.Name	ISO Date	ProductID	SalesRep	UnitsSold	Discount	COGSTotal
	Valid 100%	Valid 100%	Valid 100%	Valid 100%	Valid 100%	Valid 100%	Valid 100%
	Error 0%	Error 0%	Error 0%	Error 0%	Error 0%	Error 0%	Error 0%
	Empty 0%	Empty 0%	Empty 0%	Empty 0%	Empty 0%	Empty 0%	Empty 0%
1	Sales2017.csv	20171228	1	2	7	55	13
2	Sales2017.csv	20170118	7	4	95	4375	67
3	Sales2017.csv	20171012	12	2	72	4375	70

E' infatti necessario:

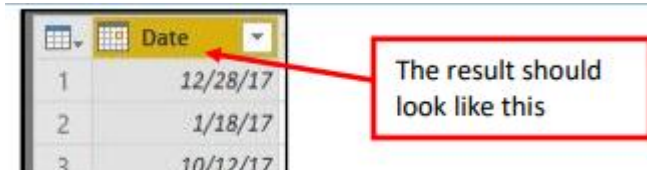
- Rimuovere la colonna Source.Name
- convertire la data ISO in una data corretta
- quindi arrotondare i numeri COGS (cost of goods sold)

14. Converti la **ISO Date** in data corretta:

- Fare clic sull'icona del tipo di dati "A B C" nell'intestazione della colonna della data ISO in alto a sinistra
- Dal menu a discesa, fai clic sul tipo di dati "Date"
- Nella finestra di dialogo "Cambia tipo di colonna", fai clic su "Aggiungi nuovo passaggio"

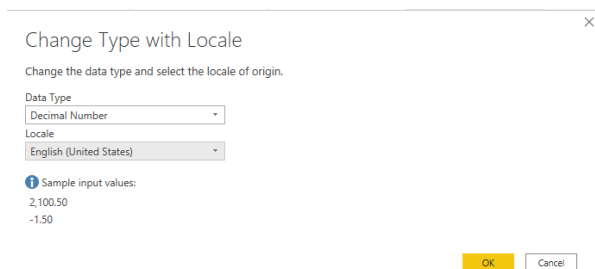


15. Fare doppio clic sul nome del campo **"ISO Date"**, digitare il nuovo nome del campo **"Date"**, quindi premere Invio.

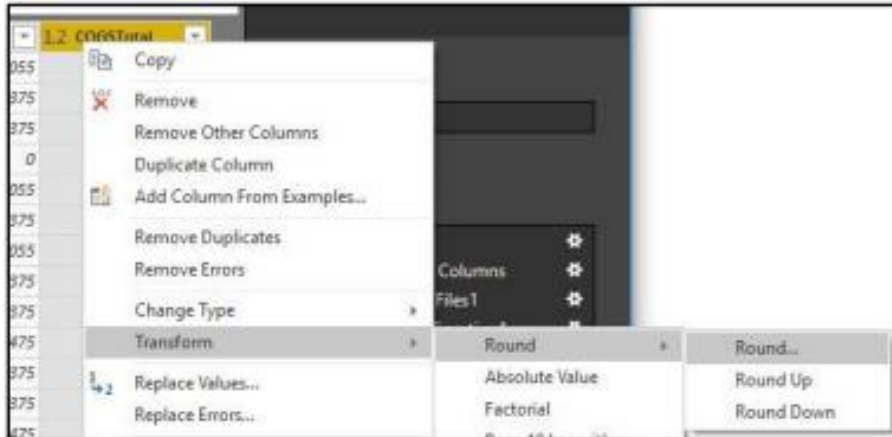


16. Campo COGS arrotondato: fare clic con il pulsante destro del mouse su Nome campo **COGSTotal**, selezionare **"Change Type"**, quindi fare clic su "Using Locale"

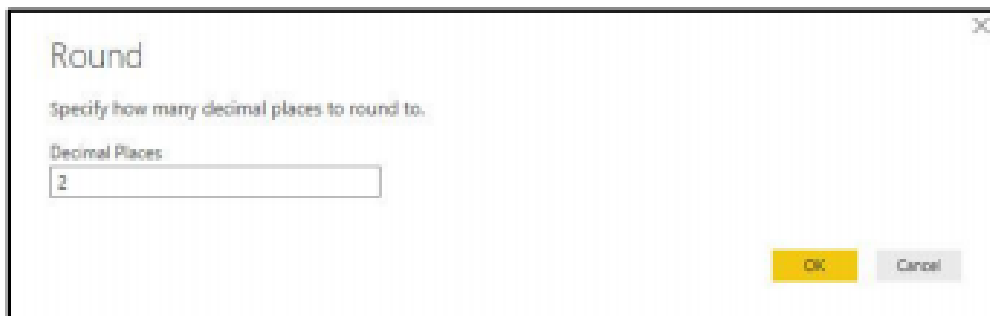
17. Dalla finestra di dialogo selezionare



18. Effettuare l'arrotondamento a due decimali usando "Transform" e scegliendo "Round".



19. Nella finestra di dialogo Round, **digitare 2**, quindi fare clic su OK

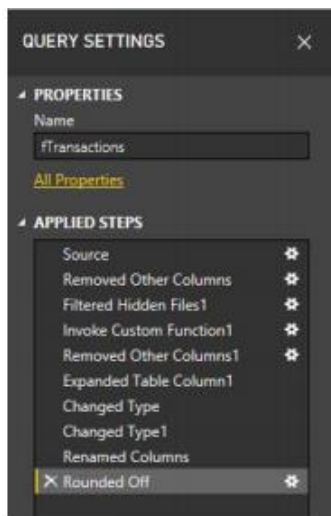


20. Ripetere gli step per il campo Discount

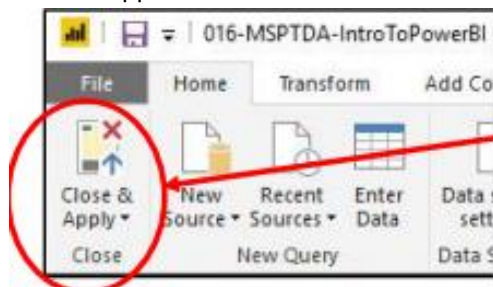
21. Assegnare il tipo **Whole Number** alle colonne

- ProductID
- SalesRep
- UnitsSold

22. Nell'Elenco dei passaggi applicati di Power Query, ecco i passaggi finali per la nostra importazione e trasformazione:

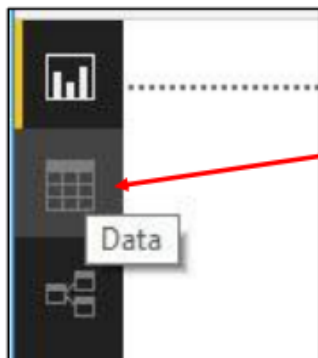


23. Chiudi e applica. Per caricare la tabella nel modello dati nella Ribbon Home.



Close & Apply:
Close = Close Power Query Editor.
Apply = Apply Applied Steps in Query.

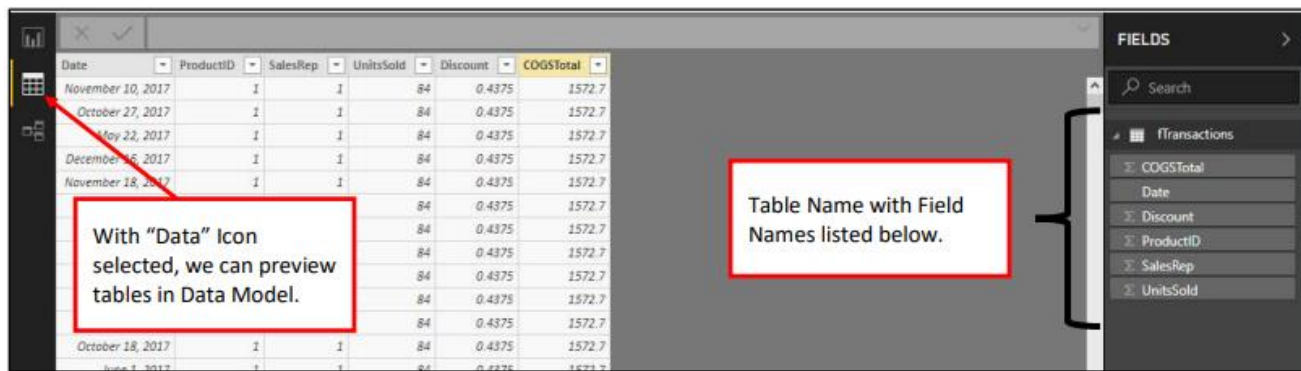
24. **Data Icon.** Nella finestra dell'Editor di Power Query, sull'estrema sinistra, fai clic sul pulsante icona "Data" (in realtà questa icona dovrebbe essere chiamata "Table icon").



In the Power BI Window, on the Far-Left Side, click the "Data" Icon Button.

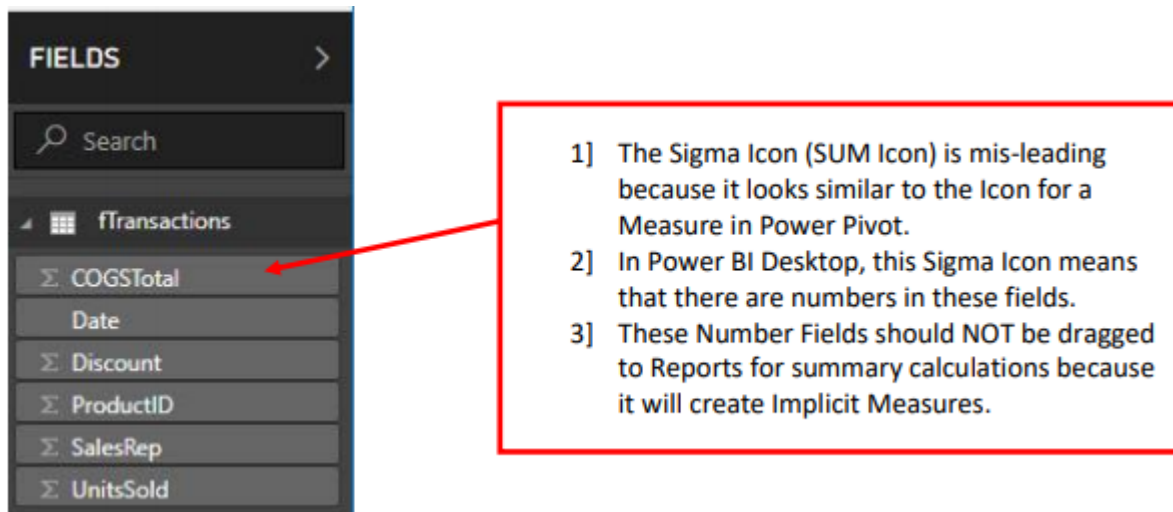
This Button should have been named "Tables" because this is the button we click to get a preview of the tables that are loaded to the Power BI - Data Model - Columnar Database.

25. Ecco come appare la tabella **fTransactions** caricata



26. Facciamo alcune considerazioni importanti sui campi Date e Numero nella tabella Power BI

a. Di seguito è riportata un'immagine dei nomi delle tabelle e dei campi **fTransactions**:



b. Dando un'occhiata ravvicinata al nome della tabella e ai nomi dei campi, notiamo:

i. Un'icona della tabella appare a sinistra del nome della tabella.

ii. Il campo della data non mostra un'icona a sinistra del nome del campo.

- I campi tipo date/datetime possono essere automaticamente promossi a gerarchia da Power BI. In realtà Power BI Desktop crea una tabella di data nascosta per ogni campo DATE, questo ha diverse controindicazioni per cui l'opzione va disabilitata
- Come vedremo è molto più efficiente crearne una nostra tabella delle date specifica con la nostra logica; e, soprattutto, creando una **singola** tabella Date e contrassegnandola come una tabella Date, evitiamo l'accumulo di più tabelle di date non necessarie.
- In seguito saranno nascosti tutti i campi di tipo date/datetime
- **Regola per i campi date/datetime dalle tabelle dei fatti: non vogliamo trascinare i campi data dalle tabelle dei fatti nelle tabelle o visualizzazioni**

iii. I campi che contengono numeri hanno un'icona **Sigma** (icona SUM) che indica che è attiva un'aggregazione di default.

- In seguito dovremo nascondere i campi del numero della tabella dei fatti dalla vista report
- **Regola per i campi numerici con icone Sigma: non vogliamo trascinare e rilasciare il campo con il numero dalle tabelle dei fatti in tabelle o visualizzazioni**

IMPORTAZIONE DELLE DIMENSIONI

1. Importa le tabelle delle dimensioni da un singolo file Excel. Nel file Excel con il nome "**0-Data Model.xlsx**", abbiamo queste tabelle delle dimensioni:
 - c. **dSalesRep** tabella Excel.
 - i. Collegheremo questa tabella dimensionale con la tabella dei fatti in una relazione uno-a-molti in modo da poter utilizzare i campi SalesRep e Region come condizioni / criteri / filtri per i nostri report.
 - d. **dProduct** tabella Excel.
 - i. La tabella dProduct contiene due chiavi esterne (**CategoryID** e **SupplierID**) che si collegano in una relazione uno-a-molti con il Tabelle **dCategory** e **dSupplier**. Dopo aver importato tutte e tre le tabelle, utilizzeremo Power Query per estrarre il campo Categoria dalla **dCategory** e campo del fornitore dalla tabella **dSupplier** alla tabella dProduct in modo da poter disporre di un **modello di dati adeguato allo schema a stella**.
 - ii. Collegheremo la tabella dimensionale **dProduct** con la tabella dei fatti in una relazione uno-a-molti in modo da poter usare vari campi come Condizioni / criteri / filtri per i nostri report, e quindi possiamo utilizzare le formule DAX dei prezzi al dettaglio per calcolare le entrate.
 - e. **dCategory** tabella Excel.
 - i. Utilizzando Power Query, uniremo questa tabella nella tabella **dProduct**.
 - f. **dSupplier** tabella Excel.
 - i. Utilizzando Power Query, uniremo questa tabella nella tabella **dProduct**.

The screenshot shows an Excel spreadsheet with the following tables:

SalesRepID	SalesRep	Region
1	Sioux Radcoolinator	NW
2	Tyrone Smith	NE
3	Chantel Zoya	SW
4	Chin Pham	SE
5	Diego Vasquez	MW
6	Vannessa Deloach	W
7	Shon Stein	NW
8	Toni Benton	NE
9	Ghislain Stidham	SW
10	Yoshiko Murillo	SE
11	Hoyt Potts	MW
12	Alysha Dewitt	W
13	Claudine Dupuis	NW
14	Shanta Spring	NE
15	Ramonita Babcock	SW
16	Janyce Betancourt	SE
17	Rhiannon Cathey	MW
18	Dominica Ordonez	W
19	Riana Burchfield	NW
20	Neida Ashe	NE
21	Marylouise Halverson	MW
22	Naoma Bloom	NW
23	Jolo Jones	W
24	Dean Washington	W
25	Kiki Lim	W

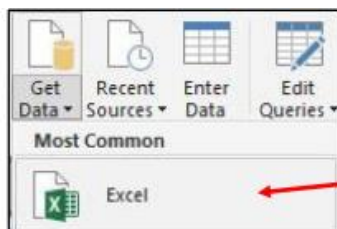
ProductID	Products	RetailPrice	CategoryID	SupplierID
1	Quad	43.95	3	GB
2	Yanaki	27.95	1	CO
3	Eagle	19.95	2	CC
4	Bellen	26.95	1	GB
5	Aspen	24.95	1	CO
6	Carlota	23.95	3	GB
7	Sunshine	19.95	4	GB
8	Sunset	22.95	4	GB
9	Beaut	35.95	5	CO
10	Kangaroo	25	2	CC
11	Elevate	48.95	5	GB
12	Flattop	25.95	2	CO
13	Vrang	12.95	4	GB
14	TriFly	21.95	3	GB
15	NaturalElbow	35	4	GB
16	LongRang	41	5	CC

CategoryID	Category
1	Beginner
2	Advanced
3	Freestyle
4	Competition
5	Long Distance

SupplierID	Supplier	City	State	E-mail
GB	Gel Boomerangs	Oakland	CA	gel@gel-boomerang.com
CO	Colorado Boomerangs	Gunnison	CO	Pollock@coloradoboomerang.com
CC	Channel Craft	Richland	WA	Dino@CC.com
GB	Darnell Booms	Burlington	VT	Darnell@Darnell.com

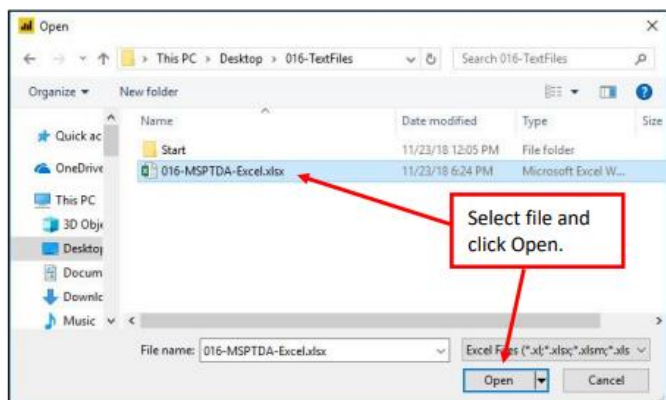
After we Import all four tables, We will use Power Query to Merge the two Snow Flake Dimension Tables into the dProducts Table.

2. Per importare le tabelle delle dimensioni dalla cartella di lavoro di Excel, andiamo alla scheda Ribbon Home di Power BI Desktop, quindi nel gruppo "External Data", facciamo clic sulla freccia a discesa Ottieni dati, quindi facciamo clic sull'opzione Excel (prima voce nell'elenco).



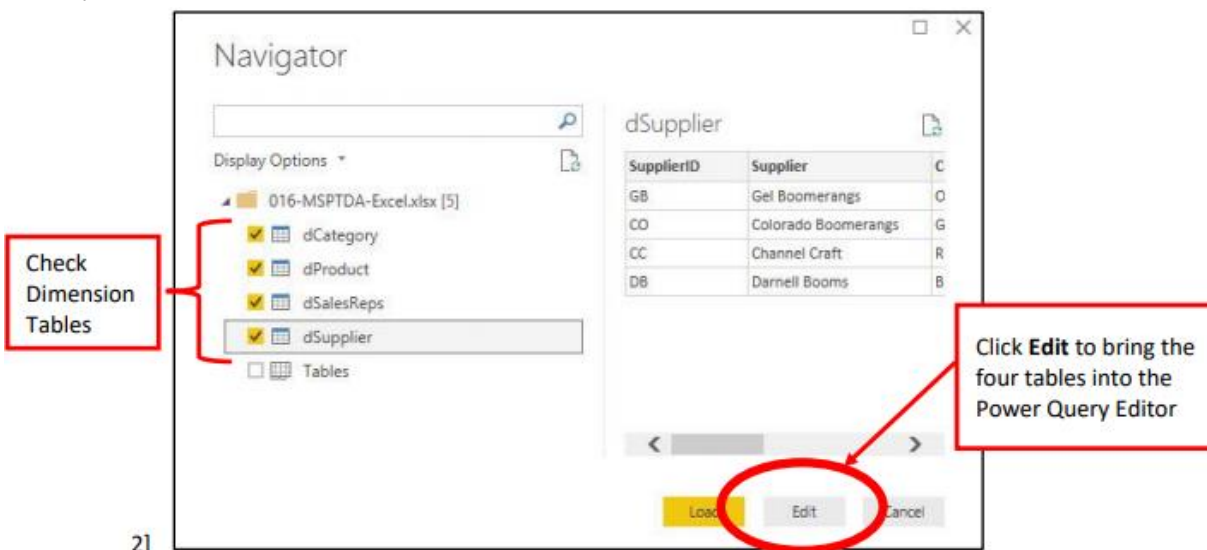
Import Excel Objects from Excel Workbook file.

3. Nella finestra di dialogo Apri, selezionare il file della cartella di lavoro di Excel denominato "**0-Data Model.xlsx**". Fai clic su Apri



Select file and click Open.

4. Nella finestra di dialogo Navigator, selezionare gli oggetti Excel desiderati. Per questo progetto scegli le quattro tabelle delle dimensioni. Quindi fare clic su Modifica per portare le quattro tabelle in Power Query Editor.
- g. Nota: ci sono più potenziali oggetti Excel che possono apparire in questa finestra di dialogo Navigator, tra cui Oggetti tabella Excel e oggetti foglio. Dovrai essere consapevole di quali oggetti sono presenti nell' Excel e selezionare gli oggetti che si desidera importare.



5. **Verificare** che le tabelle delle dimensioni siano state importate correttamente e verificare il tipo di dati corretto per ciascun campo. Nell'editor di Power Query, possiamo vedere che ci sono quattro nuove query elencate nel riquadro delle query a sinistra: dCategory, dProduct (quello selezionato nell'immagine), dSalesRep e dSupplier. A destra in Riquadro impostazioni query, possiamo vedere il nome della query e dei passaggi applicati per la query dProduct.

Uno per uno, **DEVI controllare ciascuna query per verificare che il nome della query sia adeguato e che i passaggi applicati**

abbiamo aggiunto **il tipo di dati corretto per ciascuna colonna di ciascuna tabella**. Per ogni nome di query, il nome tabella Excel ereditato è sufficiente. Ad esempio, "dProduct" è il nome della tabella Excel ereditato e questo è un buon nome per la query e per tabella risultante che verrà importata nel modello dati.

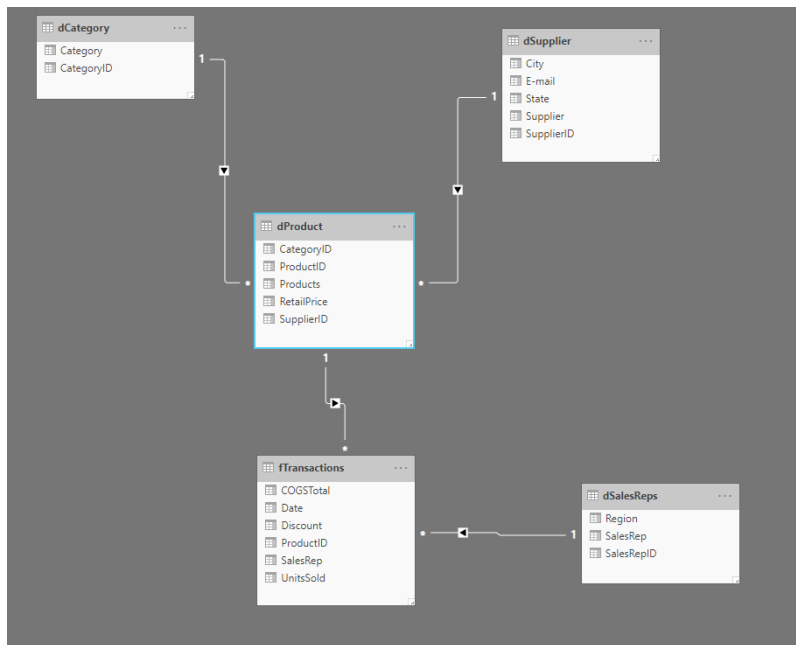
The screenshot displays the Power Query Editor window. The main area shows a table with 5 columns and 16 rows. The columns are ProductID, Products, RetailPrice, CategoryID, and SupplierID. The table contains data for various products like Quad, Yanaki, Eagle, Bellen, Aspen, Carlota, Sunshine, Sunset, Beut, Kangaroo, Elevate, Flattop, Vrang, TriFly, NaturalElbow, and LongRang.

On the left, the 'Queries [9]' pane shows a list of queries. A bracket groups four queries: fTransactions, dCategory, dProduct, and dSalesReps. A red box highlights this group with the text: "Four Individual Queries, one for each Excel Table, have been created".

On the right, the 'QUERY SETTINGS' pane is visible. It shows the 'PROPERTIES' section with the name 'dProduct'. Below it, the 'APPLIED STEPS' section lists 'Source', 'Navigation', and 'Changed Type'. A red box highlights the 'Changed Type' step with the text: "For each Table, Verify that Query Name and Applied Steps are correct.".

The bottom status bar indicates "5 COLUMNS, 16 ROWS" and "PREVIEW DOWNLOADED AT 12:42 PM".

UNIRE LE SNOW FLAKE DIMENSION DENTRO LA DPRODUCT



1. **Denormalizzare** le tabelle delle dimensioni snow flake per ricondurre il modello ad uno star schema.

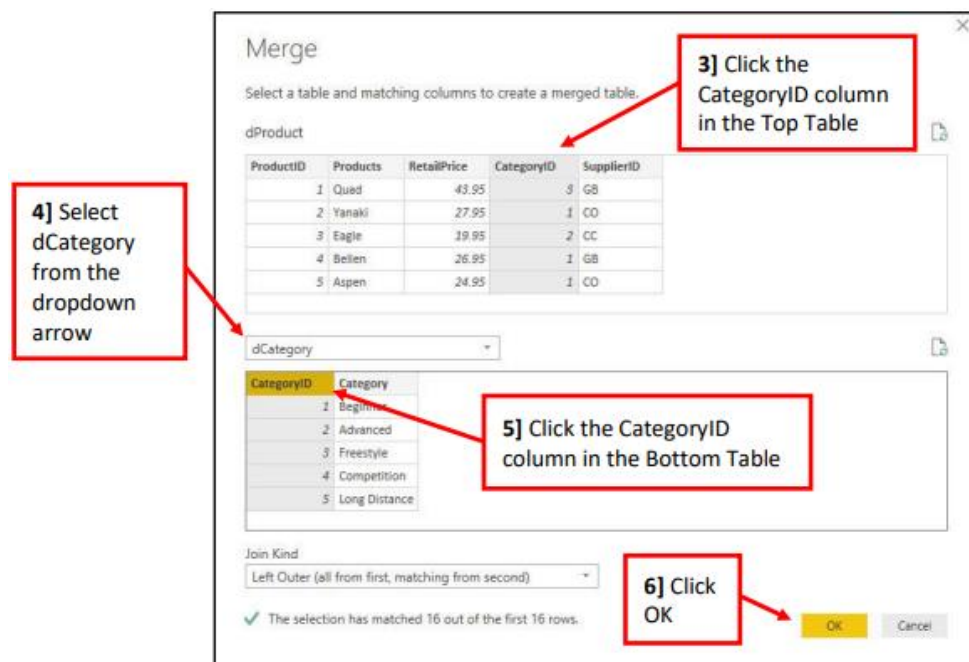
Attenersi alla seguente procedura per unire **dCategory** e **dSupplier** alla tabella **dproduct**:

- Selezionare la tabella **dProduct** nel riquadro delle query
- Nel gruppo Combina, nella Ribbon Home, fare clic sul menu a discesa **Merge** query, quindi da menu a discesa fai clic su **Merge query**.

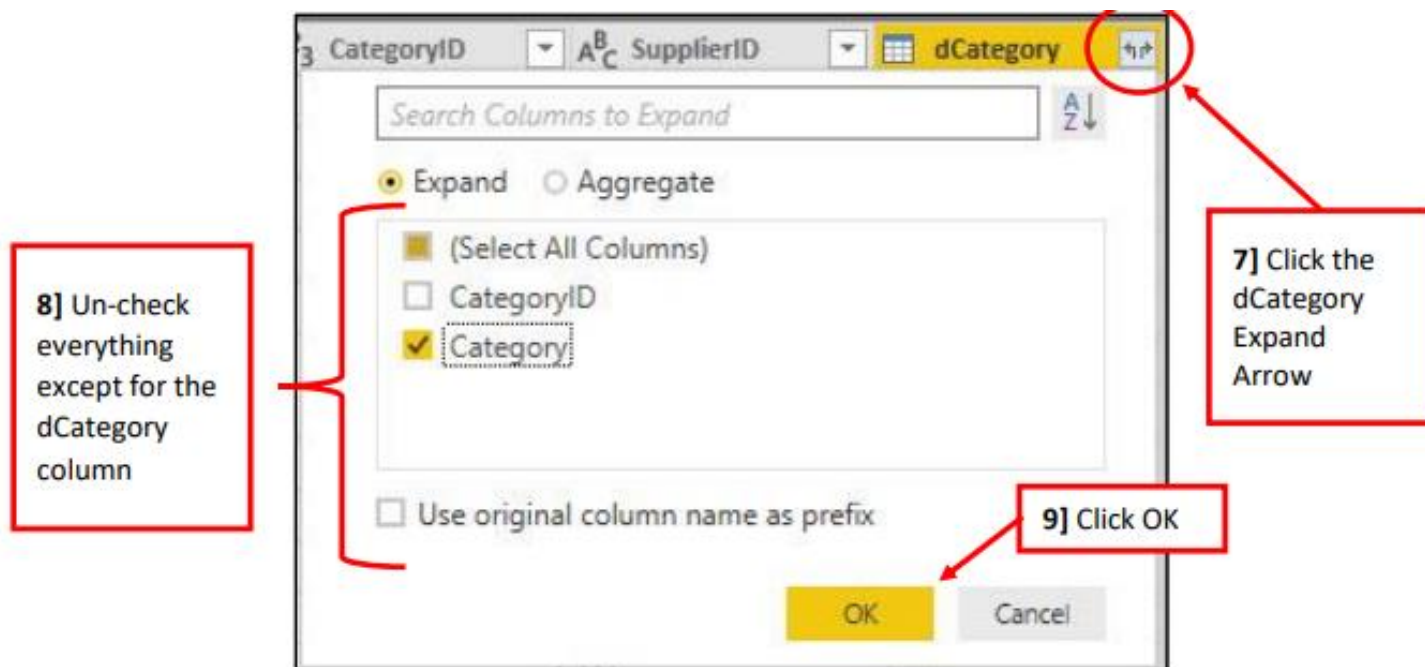
1] Select the dProduct table.

2] Click Merge Queries

- Nella finestra di dialogo di **Merge** fare clic sulla colonna **CategoryID** nella tabella superiore
- Per la tabella inferiore selezionare dCategoria dalla freccia a discesa.
- Fare clic sulla colonna **CategoryID** nella tabella inferiore.
- Fare clic su OK.



- Il processo di **Merge** ha aggiunto una nuova colonna denominata **dCategory**. Fare clic sulla freccia Espandi **dCategory**.
- Deseleziona tutto tranne la colonna **dCategory**.
- Fare clic su OK.



- Ripetiamo la stessa operazione con la tabella dSupplier sul campo SupplierID e recuperiamo solo la descrizione Supplier

Merge

Select a table and matching columns to create a merged table.

dProduct

ProductID	Products	RetailPrice	CategoryID	SupplierID
1	Quad	43.95	3	GB
2	Yanaki	27.95	1	CO
3	Eagle	19.95	2	CC
4	Bellen	26.95	1	GB
5	Aspen	24.95	1	CO

dSupplier

SupplierID	Supplier	City	State	E-mail
GB	Gel Boomerangs	Oakland	CA	gel@gel-boomerang.com
CO	Colorado Boomerangs	Gunnison	CO	Pollock@coloradoboomerang.com
CC	Channel Craft	Richland	WA	Dino@CC.com
DB	Darnell Booms	Burlington	VT	Darnell@Darnell.com

- Eliminiamo le chiavi usate per recuperare le informazioni quindi **CategoryID** e **SupplierID**. Poiché abbiamo solo bisogno delle tabelle **dCategory** e **dSupplier** per parte della nostra trasformazione Power Query, e poiché NON vogliamo caricare le due tabelle nel modello di dati, dobbiamo fare clic con il pulsante destro del mouse su ciascuna delle opzioni delle due query (dCategory & dSupplier) quindi deselezionare "Enabled Load".



10] Right-click each of the two Queries (dCategory & dSupplier) in the Query Pane on the left, then uncheck "Enable load"

- Di seguito è riportata un'immagine della tabella Transformed dProduct. A sinistra nel riquadro delle query, possiamo vedere che esiste una tabella **dei fatti (fTransactions)**, due Tabelle dimensioni (**dProduct & dSalesRep**) e **dCategory & dSupplier**, che sono **in corsivo** e **NON verranno caricati nel modello dati**. Sulla destra puoi vedere i passi applicati completati per la tabella dProduct.

016-MSPTDA-IntroToPowerBI - Power Query Editor

File Home Transform Add Column View Help

Close & Apply Close New Source Recent Sources Enter Data Data source settings Manage Parameters Refresh Preview Advanced Editor Choose Columns Remove Columns Keep Rows Remove Rows Sort Split Column Group By Data Type: Text Use First Row as Headers Replace Values Merge Queries Append Queries Combine Files Combine

Queries [9]

Transform File from... Sample Query [2] Sample File Sample File Para... Transform Sample... Transform File fra... Other Queries [5] fTransactions dCategory dProduct dSalesReps dSupplier

Table.ExpandTableColumn(#"Merged Queries1", "dSupplier", {"Supplier"}, {"Supplier"})

ProductID	Products	RetailPrice	CategoryID	SupplierID	Category	Supplier
1	Quad	43.95	3	GB	Freestyle	Gei Boomerangs
2	6 Carlots	23.95	3	GB	Freestyle	Gei Boomerangs
3	4 Bellen	26.95	1	GB	Beginner	Gei Boomerangs
4	2 Yanaki	27.95	1	CO	Beginner	Colorado Boomerangs
5	5 Aspen	24.95	1	CO	Beginner	Colorado Boomerangs
6	3 Eagle	19.95	2	CC	Advanced	Channel Craft
7	7 Sunshine	19.95	4	GB	Competition	Gei Boomerangs
8	8 Sources	22.95	4	GB	Competition	Gei Boomerangs
9		35.95	5	CO	Long Distance	Co
10		25	2	CC	Advanced	Co
11		48.95	5	GB	Long Distance	Gei Boomerangs
12		25.95	2	CO	Advanced	Co
13		12.95	4	DB	Competition	Darnell Booms
14		21.95	3	DB	Freestyle	Darnell Booms
15	15 NaturalElbow	35	4	DB	Competition	Darnell Booms
16	16 LongRang	41	5	CC	Long Distance	Channel Craft

7 COLUMNS, 16 ROWS

PREVIEW DOWNLOADED AT 2:28 PM

QUERY SETTINGS

PROPERTIES

Name dProduct

All Properties

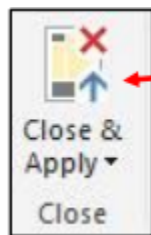
APPLIED STEPS

Source Navigation Changed Type Merged Queries Expanded dCategory Merged Queries1 Expanded dSupplier

dCategory & dSupplier are italicized and will NOT be loaded to the Data Model

Finished Steps for dProduct

3. Per chiudere e caricare le tabelle delle dimensioni nel database colonnare nel modello dati, fare clic sul pulsante Chiudi e applica



Clicking Close & Apply will apply Query changes, load tables to Data Model and close the Power Query window.

nel gruppo Close della ribbon Home

4. Visualizza tabelle nella finestra di Power BI Desktop:

- Per visualizzare le tre tabelle caricate nel Modello dati, fare clic sull'icona Dati nell'estrema sinistra nella finestra di Power BI Desktop.

Nel riquadro Campi a destra, è possibile selezionare le tabelle da visualizzare.

Clicking Data Icon (really should be names Tables Icon) allows you to see tables that have been loaded to the Data Model.

Click on Table Name to view table.

Click Expand triangle to reveal Fields below tables.

ProductID	Products	RetailPrice	CategoryID	SupplierID	Category	Supplier
1	Quad	43.95	3	GB	Freestyle	Gel Boomerangs
6	Carlota	23.95	3	GB	Freestyle	Gel Boomerangs
4	Bellen	26.95	1	GB	Beginner	Gel Boomerangs
2	Yanaki	27.95	1	CO	Beginner	Colorado Boomerangs
5	Aspen	24.95	1	CO	Beginner	Colorado Boomerangs
3	Eagle	19.95	2	CC	Advanced	Channel Craft
7	Sunshine	19.95	4	GB	Competition	Gel Boomerangs
8	Sunset	22.95	4	GB	Competition	Gel Boomerangs
9	Beaut	35.95	5	CO	Long Distance	Colorado Boomerangs
10	Kangaroo	25	2	CC	Advanced	Channel Craft
11	Elevate	48.95	5	GB	Long Distance	Gel Boomerangs
12	Flattop	25.95	2	CO	Advanced	Colorado Boomerangs
13	Vrang	12.95	4	DB	Competition	Darnell Booms
14	TriFiY	21.95	3	DB	Freestyle	Darnell Booms
15	NaturalElbow	35	4	DB	Competition	Darnell Booms
16	LongRang	41	5	CC	Long Distance	Channel Craft

TABLE: dProduct (16 rows)

5. Ci manca la dimensione del tempo per poter fare le analisi.

Vediamo la tecnica basata su PowerQuery / M

- Quindi entriamo in Power Query e apriamo una query vuota. Poi copiamo all'interno del designer il codice m che si trova nella cartella del Laboratorio **dCalendar.pq** che contiene una funzione che genera il calendario.
- Eseguiamo la funzione e creiamo il calendario.
- Poi facciamo chiudi e applica per collegare il calendario creato con le altre tabelle del modello dati.