



Edoardo Cavaliere

Analisi dei dati con Power BI Desktop

Descrizione Progetto

L'obiettivo del progetto è la creazione di una Dashboard a partire da più fogli Excel.

I dati di input* contengono le classiche informazioni di un Database Gestionale, come l'elenco delle fatture, dei clienti, dei fornitori, etc.

Le tabelle di origine contengono dimensioni di analisi come la regione di provenienza dei clienti, la data di fatturazione e misure come l'importo della fattura.

Analisi del file di input

I dati di input derivano da un file Excel contenente due fogli:

- nel foglio «Fatture» sono presenti 8 colonne
- nel foglio «Database» sono presenti le informazioni su Clienti, Fornitori e Corrieri, separati da delle colonne vuote

	A	B	C	D	E	F	G	H
1	IdFattura	IdCliente	IdFornitore	DataFattura	DataTerminePagamento	DataPagamento	Importo	IdCorriere
2	1	1	15	12/08/2018	12/11/2018	21/10/2018	2,7	3
3	2	39	1	23/07/2018	23/10/2018	21/10/2018	5,2	1
4	3	10	1	25/08/2018	25/11/2018	23/11/2018	4,3	1
5	4	10	2	03/12/2018	03/03/2019	03/03/2019	5,1	3
6	5	29	2	15/11/2018	15/02/2019	13/02/2019	2,3	1
7	6	15	2	26/07/2018	26/10/2018	24/10/2018	8,5	3
8	7	35	2	28/01/2018	28/04/2018	28/04/2018	8,8	1
9	8	21	2	27/03/2018	27/06/2018	30/06/2018	1,3	1
10	9	40	3	27/06/2018	27/09/2018	25/09/2018	5,7	1
11	10	35	2	28/11/2018	28/02/2019	26/02/2019	9,5	2
12	11	32	1	11/05/2018	11/08/2018	09/08/2018	2,7	2
13	12	22	2	11/08/2018	11/11/2018	09/11/2018	10,8	1
14	13	7	2	09/09/2018	09/12/2018	08/12/2018	3	3
15	14	13	1	22/01/2018	22/04/2018	22/04/2018	7,1	3
16	15	22	3	12/09/2018	12/12/2018	30/12/2018	6,1	1
17	16	10	1	20/10/2018	20/01/2019	18/01/2019	1,3	1
18	17	22	2	27/09/2018	27/12/2018	26/12/2018	2,1	2
19	18	39	1	07/09/2018	07/12/2018	06/12/2018	4,7	2
20	19	40	3	07/12/2018	07/03/2019	07/03/2019	8,2	1

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	IdCliente	Nome	DataNascita	Regione		IdCorriere	Denominazione	Nazione	Regione		IdFornitore	Denominazione	Regione
2	1	Nicola	01/01/2010	Lazio			1 Corriere Iba	Italia	Lombardia		1	Fornitore Ava	Lombardia
3	2	Giovanni	01/03/1976	Lazio			2 Corriere Cic	Italia	Lazio		2	Fornitore Bcb	Lazio
4	3	Marco	01/04/1980	Lazio			3 Corriere Dais	Italia	Calabria		3	Fornitore Asa	Calabria
5	4	Giovanna	01/05/1977	Lazio									
6	5	Alice	01/06/1969	Sicilia									
7	6	Fabrizio	01/07/1996	Sicilia									
8	7	Irene	01/08/1990	Sicilia									
9	8	Maria	01/09/1999	Sicilia									

Import dei dati con Power Query– pt.1

L'import e aggiornamento dei dati può essere gestito con Power Query.

In tutti i casi si è posta attenzione a scegliere correttamente i **tipi** delle colonne.

In particolare, per gestire il foglio Database contenente più informazioni, sono stati utilizzati i passaggi «**scegli colonne**» e «**rimuovi righe vuote**».

Import dei dati con Power Query – pt.2

fx = Table.PromoteHeaders(Database_Sheet, [PromoteAllScalars=true])

	ABC 123	IdCliente	ABC 123	Nome	ABC 123	DataNascita	ABC 123	Regione	ABC 123	Column5
		Validi 100%		Validi 100%		Validi 100%		Validi 100%		Validi 0%
		Errore 0%		Errore 0%		Errore 0%		Errore 0%		Errore 0%
		Vuoto 0%		Vuoto 0%		Vuoto 0%		Vuoto 0%		Vuoto 100%
1			1	Nicola		01/01/2010		Lazio		
2			2	Giovanni		01/03/1976		Lazio		
3			3	Marco		01/04/1980		Lazio		
4			4	Giovanna		01/05/1977		Lazio		
5			5	Alice		01/06/1969		Sicilia		
6			6	Fabrizio		01/07/1996		Sicilia		
7			7	Irene		01/08/1990		Sicilia		
8			8	Maria		01/09/1999		Sicilia		
9			9	Grazie		01/10/1990		Sicilia		
10			10	Giovanni		01/11/1971		Toscana		
11			11	Maria		01/12/1996		Toscana		
12			12	Giuseppe		01/01/1976		Toscana		
13			13	Francesco		01/02/1999		Toscana		
14			14	Mike		01/03/1996		Toscana		

fx = Table.SelectColumns("#Intestazioni alzate di livello",{"IdFornitore", "Denominazione_2",

	ABC 123	IdFornitore	ABC 123	Denominazione_2	ABC 123	Regione_3
		Validi 7%		Validi 7%		Validi 7%
		Errore 0%		Errore 0%		Errore 0%
		Vuoto 93%		Vuoto 93%		Vuoto 93%
1			1	Fornitore Ava		Lombardia
2			2	Fornitore Bcb		Lazio
3			3	Fornitore Asa		Calabria
4				null		null
5				null		null
6				null		null
7				null		null
8				null		null
9				null		null
10				null		null
11				null		null
12				null		null

Import dei dati con Power Query– pt.3

Power Query Editor showing the transformation of data from a source table to a target table.

Top Panel (Source Table): The source table is named "Intestazioni alzate di livello". It has columns: IdFornitore, Denominazione_2, and Regione_3. The data is filtered to show rows where the status is "Validi" (Valid).

Bottom Panel (Target Table): The target table is named "Modificato tipo". It has columns: IdFornitore, DenominazioneFornitore, and RegioneFornitore. The data is filtered to show rows where the status is "Validi" (Valid).

Formulas:

- Top Panel: `= Table.SelectColumns("#Intestazioni alzate di livello",{"IdFornitore", "Denominazione_2", "Regione_3"})`
- Bottom Panel: `= Table.SelectRows("#Modificato tipo", each not List.IsEmpty(List.RemoveMatchingItems(`

Query Settings (Impostazioni query):

- Top Panel: The "PASSAGGI APPLICATI" (Applied Steps) list shows "Rimosse altre colonne" (Remove other columns) as the active step.
- Bottom Panel: The "PASSAGGI APPLICATI" list shows "Righe vuote rimosse" (Remove empty rows) as the active step.

A green arrow points from the "Rimosse altre colonne" step in the top panel to the "Righe vuote rimosse" step in the bottom panel, indicating the flow of the transformation process.

Arricchiamo i dati con il linguaggio DAX – pt.1

Tramite il linguaggio DAX possiamo arricchire le informazioni a disposizione.

Ad esempio è possibile combinare le colonne *DataTerminePagamento* e *DataPagamento* per classificare le fatture in base all'esito del pagamento.

Utilizziamo le **formule DATEDITT, IF, OR, BLANK** facendo attenzione anche alla gestione dei dati mancanti.

Arricchiamo i dati con il linguaggio DAX – pt.2

GiorniRitardo =

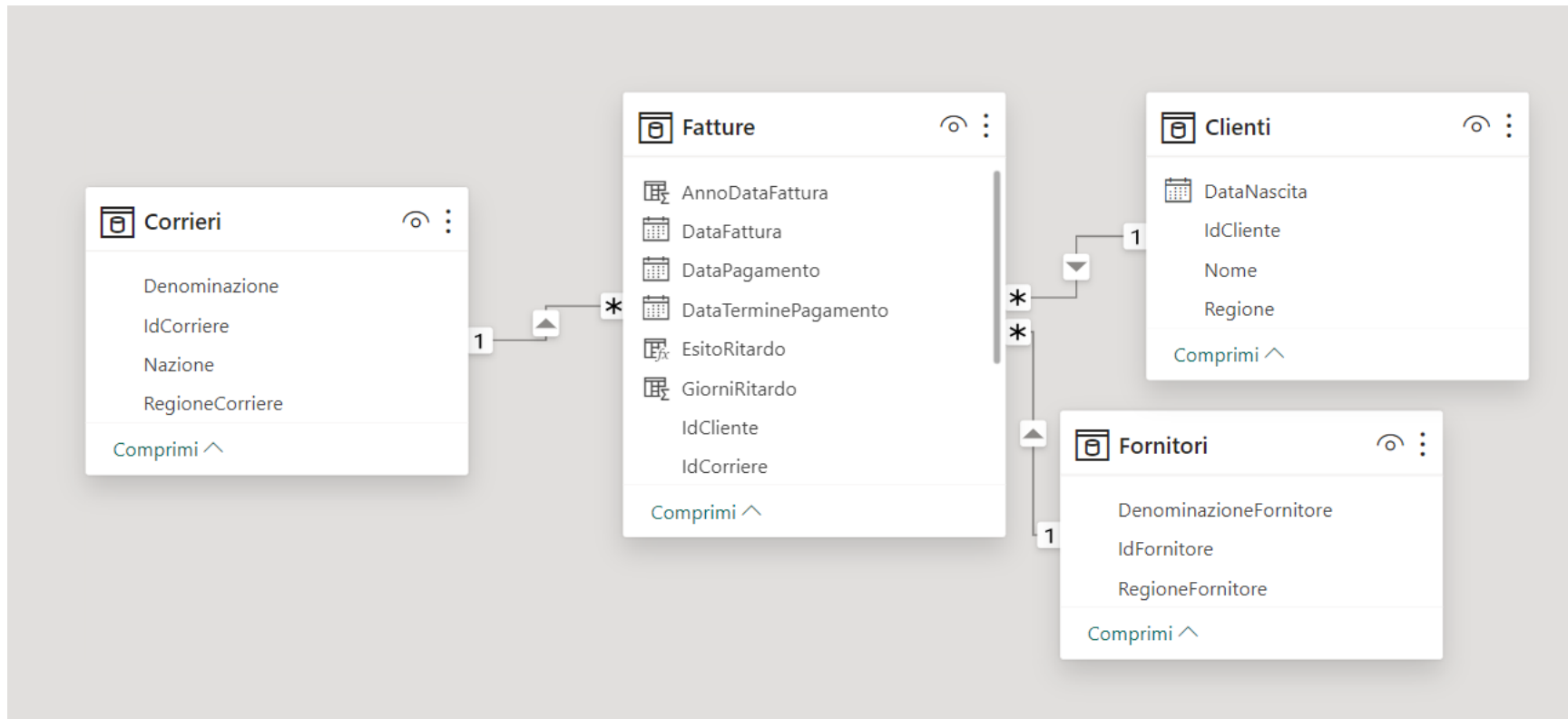
```
DATEDIFF(Fatture[DataTerminePagamento], Fatture[DataPagamento], DAY)
```

EsitoRitardo =

```
IF(OR(Fatture[DataPagamento]=blank(),  
      Fatture[DataTerminePagamento]=blank()),  
   "non disponibile",  
   IF(Fatture[GiorniRitardo]<0,  
       "in anticipo",  
       IF(Fatture[GiorniRitardo]=0  
           "in orario",  
           IF(Fatture[GiorniRitardo]<7,  
               "in extra ritardo", "in ritardo"  
           ))))
```


Creiamo la relazione tra le tabelle

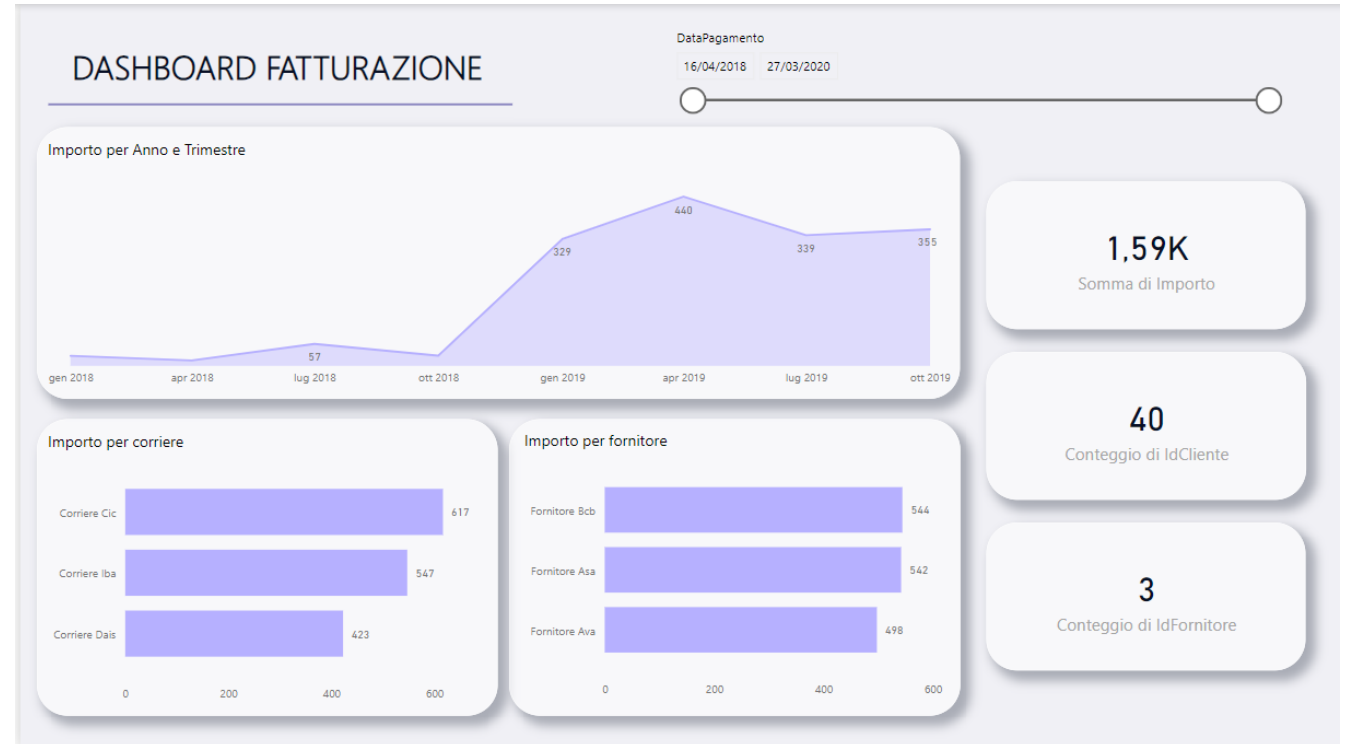
Utilizziamo le chiavi primarie e le chiavi esterne delle tabelle per creare tre relazioni **uno a molti**.



Creiamo la Dashboard – pt. 1

Iniziamo con una **HomePage** che riporti:

- un titolo
- alcune schede di riepilogo
- un grafico a linee che mostri l'andamento dell'importo al variare della data fattura
- due grafici a torta per mostrare come l'importo si suddivide per corrieri e fornitori
- un filtro dati sulla colonna della data fattura



Creiamo la Dashboard – pt. 2

Creiamo una scheda di dettaglio per i clienti con:

- un diagramma a barre verticali e una tabella per mostrare come l'importo delle fatture varia con le regioni dei clienti
- un grafico a torta per mostrare la provenienza dei clienti
- una matrice per confrontare al variare della regione gli importi del 2018 e gli importi del 2019
- un pulsante per tornare all'home page



Creiamo la Dashboard – pt. 3

Per confrontare l'importo delle fatture tra il 2018 e il 2019 ho creato due **measure**:

MisuraImporto2018 =

```
CALCULATE(SUM(Fatture[Importo]),  
FILTER(Fatture, Fatture[AnnoDataFattura] = 2018))
```

MisuraImporto2019 =

```
CALCULATE(SUM(Fatture[Importo]),  
FILTER(Fatture, Fatture[AnnoDataFattura] = 2019))
```

Risultato finale

