

## Infomatica (SGI) - Esercizi SQL – Esercizio 5 (Foodmart parte 2)

Nei laboratori, troverete all'interno del DBMS MySQL il database foodmart (lo stesso dell'esercitazione precedente). Nei prossimi esercizi, vi sarà richiesto di analizzare le vendite del 1998, contenute nella tabella `sales_fact_1998`.

Le query seguenti servono per effettuare una valutazione delle attività di gestione della catena di supermercati. L'obiettivo della valutazione è identificare degli interventi per aumentare sia il fatturato sia il reddito dei supermercati. Dopo aver realizzato le query descritte nei punti seguenti, create ulteriori query liberamente, in modo da completare a vostro piacere l'attività di valutazione. Le indicazioni sui risultati sono riportate dopo l'ultimo punto

**Attività di marketing.** Le query seguenti permettono di analizzare le attività di marketing svolte

1. Ci sono dei prodotti venduti sottocosto nel 1998? I prodotti sottocosto sono quelli per cui il valore dell'attributo `store_cost` è maggiore del valore di `store_sales`
2. Individuate le promozioni che hanno fatto guadagnare di più nelle vendite del 1998. Per avere informazioni sulle promozioni guardate l'attributo `sales_fact_1998.promotion_id` e la tabella `promotion`. Visualizzate il guadagno totale generato e il nome di ogni promozione. Visualizzate le promozioni in ordine decrescente di guadagno totale (escludete dalla visualizzazione la non promozione indicata con il nome "No Promotion").

**Margine e redditività.** Le query seguenti permettono di analizzare come l'azienda costruisce utile e margine. Il margine è un indicatore definito con qs formula: **margine = (ricavo tot. – costo tot.) / ricavo tot.**

3. Individuate i 5 prodotti che offrono il margine maggiore (suggerimento: ordinate il risultato di una query e del risultato considerate solo le prime 5 tuple).
4. Individuate le 5 categorie di prodotti che offrono il margine maggiore. La categoria di un prodotto è identificata dall'attributo 'product\_class.product\_category'.
5. Individuate i 5 settori merceologici del supermercato che offrono il margine maggiore. Il settore merceologico è identificata dall'attributo 'product\_class.product\_department'.
6. Come il punto precedente, tuttavia visualizzate i `product_department` che offrono un margine superiore al valore di 0,6 (60%).

**Analisi del portafoglio vendite.** Le query seguenti permettono di analizzare le attività di vendita

7. In questo punto vi sarà richiesto di scrivere delle query che permettano di confrontare, per ogni supermercato della catena, le vendite del settore carni con il settore prodotti in scatola. Questo esercizio sarà suddiviso in sottopunti.
  - a. Individuate per ogni negozio il ricavo conseguito vendendo i prodotti del settore merceologico "Prodotti in scatola" (valore "Canned Products" dell'attributo `product_department` della tabella `product_class`). La query deve visualizzare l'id del supermercato e il ricavo conseguente alla vendita dei prodotti del settore "Prodotti in scatola". Per ogni supermercato deve apparire una sola riga.
  - b. Come la query precedente, analizzando però il settore carni (valore "Meat" dell'attributo `product_department` della tabella `product_class`)
  - c. [NB: per svolgere questo punto è necessario utilizzare le query annidate] Costruite una query che per ogni supermercato mostri: il valore delle vendite del settore "Prodotti in scatola"; il valore delle vendite del settore carni; il rapporto tra le vendite del settore carni e le vendite del settore "Prodotti in scatola". Suggerimento: per poter scrivere la query richiesta è necessario utilizzare le query scritte nei due punti precedenti inserendole come query annidate nella query da scrivere. L'esempio che segue mostra l'uso delle query annidate (o sottoquery):

```
select ...  
from tab_a, tab_b, (select store_id, sum(store value*unit sales) from sales_fact_1998 group by  
store_id) as tab_virtuale_c  
where tab_virtuale_c.store_id = tab_a.store_id ...
```

La query esterna contiene al suo interno una query (la parte sottolineata e in corsivo) il cui risultato viene usato come se fosse una tabella vera e propria e a cui si fa riferimento con il nome `tab_virtuale_c` nella query esterna. In generale, in una query SQL dopo la parola chiave `from` possono essere riportate una tabella reale oppure il risultato di una query (che in questo caso viene chiamato query annidata).

8. Determinate il ricavo totale suddiviso per "state" (attributo `store.store_state`) e mostrate gli stati in ordine decrescente di ricavo (ogni "state" deve apparire una sola volta). Per ogni riga del risultato, visualizzate anche il campo `store_country` oltre che il campo `store_state`.
9. Identificate i prodotti che hanno fatto guadagnare di più nel 1998 (visualizzate `product_id`, `product_name`, e totale del guadagno conseguito nel 1998, visualizzate i prodotti in ordine decrescente).

10. Considerando le vendite del 1998, limitando la ricerca ai soli negozi di tipo supermarket (store.store\_type='Supermarket'), visualizzate una lista di supermercati (ogni supermarket deve apparire al massimo una sola volta) ordinata in base al rapporto (guadagno totale del supermarket) / (qt totale di oggetti venduti dal supermarket).

**Analisi varie.**

11. Visualizza l'utile generato dalle diverse tipologie di clientela. Le tipologie di clienti sono identificate dall'attributo customer.member\_card. Nella tabella risultato ci dovrà essere un solo valore per ogni tipologia di member\_card.
12. In questo esercizio saranno analizzati i margini dei settori merceologici nei diversi supermercati. L'esercizio è diviso in sottopunti.
- a. Visualizzate per ogni coppia <supermercato, settore merceologico> il margine corrispondente. I supermercati sono identificati dall'attributo store.store\_id mentre i settori merceologici sono identificati product\_class.product\_department. Nei risultati per ogni supermarket visualizzate anche la tipologia del supermarket (attributo store.store\_type)
  - b. Visualizzate per ogni coppia <tipologia supermarket, settore merceologico> il margine corrispondente. Questa query calcola il margine di un certo settore merceologico per tutti i supermercati di una stessa tipologia. A differenza del sottopunto precedente, il margine non va calcolato sul singolo supermarket ma sull'insieme di tutti i supermercati dello stesso tipo. Tenete presente che il margine non è una grandezza additiva (per calcolare il margine di un insieme di supermercati non si può sommare i margini dei singoli supermercati, ma occorre ricalcolare il margine sull'insieme di tutti i dati delle vendite).
  - c. [NB: per svolgere questo punto è necessario utilizzare le query annidate] Usando le due query precedenti come sottoquery, individuate quelle coppie di <supermercato, settore merceologico> per i quali il margine è inferiore al 96% del margine calcolato su <tipologia supermarket, settore merceologico> cioè il margine del corrispondente settore merceologico calcolato su tutti i supermercati della stessa tipologia. Per costruire la query dovete mettere in join i risultati delle query del sottopunto a) e del sottopunto b) e visualizzare solamente i dati dei supermercati che hanno un margine inferiore al 96% del margine dei supermercati della stessa categoria (quest'ultima informazione è presente nei risultati della query del punto b).
13. L'esercizio è diviso in sottopunti
- a. Visualizzate, per ogni product\_id, le quantità vendute nel primo trimestre del 1998. Le informazioni sul trimestre si trovano in time\_by\_day.quarter, il primo trimestre è 'Q1', il secondo trimestre è 'Q2', ...). Nel risultato potete limitarvi a visualizzare il product\_id, non serve il product\_name. Per ogni product\_id deve essere visualizzata una sola riga nella tabella risultato.
  - b. Visualizzate, per ogni product\_id, le quantità vendute in tutto il 1998. Nel risultato potete limitarvi a visualizzare il product\_id, non serve il product\_name. Per ogni product\_id deve essere visualizzata una sola riga nella tabella risultato.
  - c. [NB: per svolgere questo punto è necessario utilizzare le query annidate] Usando le due query precedenti come sottoquery, visualizzate, per ogni product\_id, la quantità di beni venduta nel 1° trimestre, la quantità di beni venduta in tutto l'anno, e il rapporto tra queste due quantità. Visualizzate i risultati in ordine crescente rispetto a product\_id. Nel risultato potete limitarvi a visualizzare il product\_id, non serve il product\_name.
14. L'esercizio è diviso in sottopunti
- a. Visualizzate, per ogni customer\_id, il numero medio di quantità acquistate per transazione. Una transazione corrisponde ad un singolo record della tabella sales\_fact\_1998. Nel risultato potete limitarvi a visualizzare il customer\_id, non servono i nomi e cognomi dei clienti. Per ogni customer\_id deve essere visualizzata una sola riga nella tabella risultato.
  - b. [NB: query annidate] Visualizzate, per ogni customer\_id il numero di transazioni effettuate (cioè il n.° di record di sales\_fact\_1998) in cui il cliente ha acquistato un numero di beni superiore alla quantità mediamente acquistata dal cliente nelle sue transazioni (in altre parole, ogni cliente avrà una media diversa). Si suggerisce di calcolare in una sottoquery le quantità medie per transazione dei clienti. Ordinate il risultato in ordine crescente rispetto al customer\_id.
  - c. [NB: query annidate] Arricchite la query del punto precedente visualizzando anche, per ogni cliente, il numero totale delle transazioni effettuate e il rapporto tra numero di transazioni in cui si è acquistata una quantità superiore alla media e numero di transazioni totali. Per farlo, si suggerisce di aggiungere alla query precedente una sottoquery che calcola, per ogni customer\_id, il numero totale di transazioni svolte. Ordinate il risultato in ordine crescente rispetto al customer\_id.

**Foodmart parte 2: risultati (senza le query).** Sono riportate solo le prime righe dei risultati. Attenzione, il . va interpretato come separatore tra la parte intera e la parte decimale.

1. Prodotti sottocosto: nessuno
2. Promozioni che hanno fatto guadagnare di più nelle vendite del 1998  
 'Weekend Markdown', '9848.02320000'  
 'Save-It Sale', '9567.95540000'  
 'Bag Stuffers', '7561.72810000'  
 'Money Savers', '7513.95520000'  
 'Double Down Sale', '7244.04940000'
3. Prodotti con margine maggiore (da sales\_fact\_1998)  
 1086 High Quality Paper Towels 0.618260368664  
 583 Walrus White Zinfandel Wine 0.618219844358  
 1153 Dual City Lox 0.617386172007  
 131 Faux Products Mint ... 0.6170093457941105  
 Mighty Good Monthly Auto Magazine 0.616990881459
4. 5 categorie di prodotti con margine maggiore (da sales\_fact\_1998)  
 Canned Sardines 0.604664648558  
 Carbonated Beverages 0.603163298714  
 Drinks 0.602574456253  
 Canned Tuna 0.602093418047  
 Kitchen Products 0.601710657683
5. 5 settori merceologici con margine maggiore (da sales\_fact\_1998)  
 Beverages 0.600962356385  
 Household 0.600517817892  
 Snacks 0.600422362200  
 Eggs 0.600201372713  
 Starchy Foods 0.600042831151
6. Product\_department che offrono un margine superiore al valore di 0,6 (60%) (da sales\_fact\_1998)  
 Beverages 0.600962356385  
 Household 0.600517817892  
 Snacks 0.600422362200  
 Eggs 0.600201372713  
 Starchy Foods 0.600042831151
7. Vendite carni vs vendite prodotti in scatola (solo punto c) (da sales\_fact\_1998)  

store_id,	vend_carne,	ven_scatol,	carne/scatol
1	1457.93000000	1015.09000000	1.436256883626
2	30.49000000	9.63000000	3.166147455867
3	961.25000000	815.24000000	1.179100632942
4	1038.92000000	1214.77000000	0.855240086601
5	68.49000000	18.23000000	3.756993965990
8. Ricavo per state (da sales\_fact\_1998) (ricavi in ordine decrescente)  

USA	WA	882323.57000000
USA	CA	510171.19000000
Mexico	Zacatecas	503675.41000000
USA	OR	429264.74000000
Canada	BC	329277.34000000
Mexico	DF	317063.20000000
Mexico	Yucatan	265264.47000000
Mexico	Veracruz	172919.87000000
Mexico	Guerrero	164537.21000000
Mexico	Jalisco	8006.61000000
9. Prodotti che hanno fatto guadagnare di più (da sales\_fact\_1998)  

product_id	product_name	gain
274	Great Pumpernickel Bread	3166.32140000
309	Super Extra Chunky Peanut Butter	2988.56340000
1180	Horatio Frosted Cookies	2971.02000000
175	High Top New Potatos	2969.12000000
1380	Hilltop Mint Mouthwash	2961.55820000
10. Supermercati: rapporto guadagno totale / qt oggetti venduti (da sales\_fact\_1998)  

24	4.312640339361
11	4.268505493549
15	4.245499745051
1	4.243960346164
16	4.237612303165

11. Utile generato dalle diverse tipologie di clientele (da sales\_fact\_1998)

Bronze	1198595.23050000
Normal	488982.83700000
Golden	279614.43460000
Silver	178904.91720000

12. Coppie <supermercato, settore merceologico> con margine inferiore del 4% (<96%) rispetto al margine del corrispondente settore calcolato su tutti i supermercati della stessa tipologia (da sales\_fact\_1998)

store_i,	store_type	prod_dep.	marg. sing.,	marg. Glob.
'2'	'Small Grocery'	'Eggs'	'0.573326823552'	'0.597476167896'
'2'	'Small Grocery'	'Seafood'	'0.562854152128'	'0.618038565109'
'5'	'Small Grocery'	'Eggs'	'0.571611518060'	'0.597476167896'
'9'	'Mid-Size Grocery'	'Carousel'	'0.574324621733'	'0.600637247569'
'18'	'Mid-Size Grocery'	'Meat'	'0.561517372187'	'0.588824870478'
'22'	'Small Grocery'	'Carousel'	'0.542876404494'	'0.612914298451'
'22'	'Small Grocery'	'Checkout'	'0.594654867257'	'0.623451892159'

13. Rapporto tra quantità vendute nel 1. trimestre (Q1) e quantità vendute in tutto l'anno. Viene visualizzato solo il risultato del punto c), che comprende i risultati dei punti precedenti

Product_id	qt_q1	qt_1998	rapporto
1	60.0000	148.0000	0.40540541
2	80.0000	275.0000	0.29090909
3	71.0000	266.0000	0.26691729

...

14. Rapporto tra il numero totale delle transazioni effettuate e il numero di transazioni in cui si è acquistata una quantità superiore alla media per transazione. Viene visualizzato solo il risultato del punto c), che comprende i risultati dei punti precedenti

customer_id	n_transazioni_mag	n_transazioni	indicatore
3	2	13	0.1538
6	4	6	0.6667
8	30	42	0.7143

...