

# Business Intelligence

Politecnico di Torino

**Prova in itinere – 15 aprile 2014**

Cognome	
Nome	
Matricola	

## Descrizione del problema

La ditta CateBan è un'azienda italiana che offre il servizio di catering e banqueting sul territorio italiano. L'azienda ha un'esperienza decennale nel campo della ristorazione e della banchettistica, ed è formata da uno staff giovane, brillante ed entusiasta, in grado di organizzare diversi eventi con creatività e originalità. Gli eventi possono essere organizzati in diversi luoghi (e.g., Reggia Venaria, sala ricevimento Baratti&Milano) e nello stesso luogo possono essere organizzati eventi di diversa tipologia (e.g., matrimonio, compleanno, etc.). Ogni evento è caratterizzato da alcune informazioni sul tipo menù (vegano, vegetariano, misto, italiano, internazionale). Inoltre per ogni evento è noto il luogo in cui si svolge, il numero di persone invitate e il prezzo per persona. Il prodotto tra il prezzo e il numero di invitati all'evento determina il volume d'affari per la ditta CateBan associato all'evento. In ogni luogo (e.g., Reggia Venaria, sala ricevimento Baratti&Milano) possono essere organizzati più eventi nell'arco dello stesso mese. Ogni evento è gestito da un responsabile che ordina i prodotti alimentari necessari per organizzare l'evento. Per ogni prodotto, identificato da un codice univoco, sono noti la categoria merceologica, la marca, la tipologia e se è deteriorabile oppure no. Alla fine di ogni evento il responsabile compila un modulo in cui specifica per ogni prodotto la quantità utilizzata per l'evento e la quantità di scarto dell'evento. Un responsabile può gestire più eventi, anche di diverse tipologie (e.g., matrimonio, assegnazione premi).

Negli ultimi 8 anni, la ditta ho organizzato tantissimi eventi sul territorio e vuole analizzare:

1. il volume d'affari medio mensile, il volume d'affari medio per evento e il numero medio di persone per evento;
2. la quantità media di prodotti acquistati per evento, la quantità media di scarto per evento e il costo medio per evento.

Per effettuare tali analisi, è necessario creare un data warehouse in cui siano memorizzate tutte le informazioni di interesse.

La ditta CateBan è interessata ad analizzare il volume d'affari medio mensile, il volume d'affari medio per evento e il numero medio di persone per evento in funzione di:

- luogo in cui si svolge l'evento;
- tipologia del luogo in cui si svolge l'evento (e.g., villa, sala ricevimenti, dimora storica);
- città, provincia e regione del luogo in cui si svolge l'evento;
- bimestre, quadrimestre, anno in cui è stato organizzato l'evento;
- stagione in cui è stato organizzato l'evento;
- tipologia evento e tipo menù.

Inoltre la ditta CateBan è interessata ad analizzare per ogni mese, la quantità media di prodotti acquistati per evento, la quantità media di scarto per evento, il costo medio per evento in funzione di:

- prodotto, se è deteriorabile oppure no, categoria merceologica, tipologia di prodotto e marca;
- anno, trimestre, semestre, quadrimestre in cui è stato organizzato l'evento;
- responsabile dell'evento;
- città, provincia, regione in cui si svolge l'evento;
- tipologia evento e tipo menù.

Sono di seguito riportate alcune interrogazioni frequenti di interesse per la ditta CateBan.

- a) Considerando solo gli eventi del 2012 e 2013 con un menù vegano di cucina italiana, visualizzare per ogni provincia, le tipologie di luoghi associati al massimo volume d'affari medio per evento.
- b) Considerando solo le categorie merceologiche per le quali sono state acquistate più di 5 tipologie di prodotti per organizzare gli eventi nel 2013, visualizzare per ogni tipologia di prodotto la quantità media di prodotti acquistato per evento, la quantità media di scarto per evento, il costo medio per evento e la percentuale di scarto per tipologia di prodotto rispetto al totale di scarto della categoria merceologica di appartenenza. Si effettui l'analisi separatamente per ogni anno.
- c) Per ogni luogo in cui si svolge un evento, visualizzare per ogni quadrimestre il volume d'affari complessivo, il volume d'affari medio mensile, il numero medio di persone per evento, il volume d'affari quadrimestrale cumulativo dall'inizio dell'anno. Ordinare e assegnare un rank alla coppia (quadrimestre, luogo) per volume d'affari medio mensile (la prima posizione in graduatoria è assegnato alla coppia con il più alto volume d'affari medio mensile).
- d) Visualizzare per ogni tipologia di luogo in cui si svolge un evento e per ogni bimestre il volume d'affari totale e il numero medio di persone per evento.
- e) Relativamente al 2012 e 2013 e considerando solo gli eventi organizzati nel sud dell'Italia, per ogni categoria merceologica e per ogni trimestre, visualizzare la quantità media di prodotti acquistati per evento e la quantità media di scarto per evento.
- f) Considerando solo i menu vegani di cucina internazionale e gli eventi organizzati in Abruzzo, visualizzare per ogni tipologia luogo il numero totale di eventi organizzati, il relativo volume di affari e il numero medio di persone per evento.
- g) Visualizzare per ogni luogo in cui si svolge un evento e per ogni anno il volume d'affari totale, il volume d'affari medio per evento, il numero medio di persone per evento. Si effettui l'analisi separatamente per ogni tipologia evento.

## Progettazione

Il data warehouse conterrà le informazioni relative agli ultimi 8 anni. Al fine di una corretta realizzazione del data warehouse sono state fornite le seguenti informazioni:

- Prodotti: ~1000
- Categoria merceologica: ~10
- Tipologie prodotti: ~50
- Luoghi: ~800
- Città: ~100
- Tipologie eventi: ~10
- Tipologie luoghi: ~10

1. Progettare il data warehouse necessario per analizzare tutti le esigenze di analisi descritte in precedenza. Il data warehouse progettato deve inoltre permettere di rispondere in modo efficiente a **tutte** le interrogazioni frequenti presentate nelle specifiche del problema.
2. Esprimere le interrogazioni frequenti **(a), (b), (c)**, delle specifiche del problema utilizzando il linguaggio SQL esteso.
3. Considerando le caratteristiche del data warehouse realizzato e la cardinalità dei dati memorizzati nel data warehouse, decidere se e quali viste materializzate potrebbero essere definite al fine di ottimizzare i tempi di risposta delle interrogazioni proposte nelle specifiche del problema (considerare tutte le interrogazioni proposte e non solo quelle risolte in SQL al punto 2. Motivare le scelte fatte.
4. Decidere come gestire la dinamicità (variazioni) dei dati all'interno delle dimensioni. Motivare le scelte fatte.