

Data Science e Tecnologie per le Basi di Dati

Homework 1

Una società italiana operante nel settore di ticketing per eventi musicali vuole analizzare gli incassi ed il numero di biglietti acquistati in funzione di data di acquisto, data e luogo di svolgimento dell'evento.

I biglietti di un evento sono caratterizzati dal nome del tour di un certo cantante/gruppo (e.g. "Evolve World Tour" del gruppo "Imagine Dragons"), data e luogo in cui l'evento si è svolto (e.g. 06/09/2018, "Milano Area Expo").

Un tour è caratterizzato da uno o più generi musicali. L'elenco dei possibili generi è noto a priori ed è di dimensioni limitate ("rock", "pop rock", "elettropop", "pop", "metal", ..., "jazz"). Ad esempio, il tour "Evolve World Tour" è associato ai generi "pop rock" e "elettropop".

Si vuole tener traccia della data di acquisto del biglietto e delle informazioni relative al suo acquisto, come il metodo di pagamento (carta di credito, carta prepagata, etc.). Un biglietto può essere acquistato attraverso tre modalità: online, presso una rivendita ufficiale o direttamente all'ingresso dell'evento.

L'azienda è interessata a svolgere analisi in funzione di:

- nome del tour e generi musicali,
- cantante/gruppo e rispettiva nazionalità,
- luogo dell'evento, città, provincia, regione (e.g., "Milano Area Expo", Milano),
- data (e.g. 06/09/2018), mese e anno di svolgimento dell'evento, indicazione se giorno feriale,
- data (e.g. 12/06/2018), mese, mese dell'anno, bimestre, trimestre, quadrimestre, semestre e anno in cui è stato effettuato l'acquisto,
- modalità di acquisto del biglietto (online, rivendita ufficiale, ingresso evento),
- metodo di pagamento (carta di credito, carta prepagata, etc.)

Homework Tasks

1. Progettare il data warehouse in modo da soddisfare le richieste descritte nelle specifiche del problema. Disegnare lo schema concettuale del datawarehouse e definire lo schema logico (tabelle dei fatti e delle dimensioni).
2. Esprimere le due interrogazioni seguenti utilizzando il linguaggio SQL esteso.
 - (a) Separatamente per ogni tipo di acquisto e per ogni mese di acquisto, analizzare:
 - i. l'incasso medio giornaliero
 - ii. l'incasso cumulativo dall'inizio dell'anno
 - iii. la percentuale di biglietti associati al tipo di acquisto considerato rispetto al totale del mese
 - (b) Per gli eventi che si sono verificati nel 2017, separatamente per ogni nazionalità cantante e per ogni città, analizzare:
 - i. l'incasso medio per biglietto,
 - ii. la percentuale di incasso rispetto al totale di incasso della provincia,
 - iii. la percentuale di incasso rispetto al totale di incasso regionale