

Business Intelligence

Appello del 1 Luglio 2016

Descrizione del problema

La società GiocoVacanze gestisce diversi centri, distribuiti sul territorio nazionale italiano, che offrono un servizio ludico-educativo come supporto alle famiglie durante le vacanze scolastiche (e.g., pausa natalizia, pausa pasquale, periodo estivo, etc.). Ogni centro estivo dispone di spazi organizzati e attrezzati per consentire ai bambini (3-9 anni) e ai ragazzi (10-17 anni) che frequentano il centro opportunità educative, di socialità, di gioco e di comunicazione con i propri coetanei.

Le persone (bambini e ragazzi) che vogliono frequentare un centro possono iscriversi per un periodo minimo di una settimana (o per più settimane) e per un massimo di 10 ore giornaliere. Il centro offre diversi stimoli creativi coordinati da educatori qualificati. Il numero di educatori presenti in ogni settimana è proporzionale al numero di persone che sono iscritte. Durante la giornata vengono organizzate diverse attività che possono essere ripetute più volte nello stesso giorno e più volte nella stessa settimana.

Ogni attività è associata ad una tipologia specifica (es. torneo) ed è caratterizzata da un luogo in cui si svolge (es. campo da tennis, campo da volley, ecc.). Ogni attività è tipicamente caratterizzato da una o più edizioni (es. torneo di beach volley edizione 2015, torneo di tennis under 6 2016, ecc.). Data l'edizione è possibile identificare l'anno.

Alla fine di ogni settimana, le persone che hanno frequentato il centro possono esprimere un indice di gradimento delle diverse attività svolte assegnando un valore da 1 a 10 in modalità anonima e libera.

Attualmente ogni centro ha una base di dati indipendente per gestire al meglio il servizio, ma la società GiocoVacanze è interessata a progettare e realizzare un data warehouse che offra una visione d'insieme che consenta di analizzare:

1. gli incassi derivanti dalle attività organizzate nei diversi periodi dell'anno;
2. l'indice di gradimento delle attività proposte e qualità del servizio offerto.

Per il punto (1) la società GiocoVacanze è interessata ad analizzare l'incasso medio settimanale, l'incasso medio per persona iscritta e l'incasso medio per ora di attività in funzione di:

- settimana, settimana dell'anno, mese, bimestre, trimestre, quadrimestre, semestre e anno;
- fascia oraria (e.g., 8:00-12:00, 12:00-15:00, 15:00-18:00);
- fascia di età (e.g., 3-6 anni, 7-9 anni, 10-13 anni, 14-17 anni);
- categoria di persone che frequentano (bambini e ragazzi);
- caratteristiche del centro: tipologie di campi presenti (e.g., campo da tennis, campo da calcio, campo da volley), presenza di uno spazio esterno attrezzato, presenza della piscina;
- tipo di pagamento utilizzato dalle famiglie (e.g., bonifico, contrassegno, assegno);
- città, provincia e regione in cui è sito il centro;
- attività svolta, edizione, luogo, anno e tipologia dell'attività.

Per il punto (2) la società GiocoVacanze è interessata ad analizzare il numero medio di persone che frequentano per educatore, il valore medio di indice di gradimento per valutazione espressa, in funzione di:

- mese, mese dell'anno, bimestre, semestre, semestre dell'anno e anno;
- centro;
- città, provincia e regione in cui è sito il centro;
- caratteristiche del centro: tipologie di campi presenti (e.g., campo da tennis, campo da calcio, campo da volley), presenza di uno spazio esterno attrezzato, presenza della piscina;
- attività svolta, edizione, luogo, anno e tipologia dell'attività;
- categoria di persone che frequentano (bambini e ragazzi).

Sono di seguito riportate **alcune** delle interrogazioni frequenti di interesse:

- a) **Considerando i centri con solo la piscina all'aperto e il campo da volley, e le valutazioni espresse nel 2011 e 2012, visualizzare per ogni centro e attività:**
 - il valore medio di indice di gradimento per valutazione espressa,
 - il numero medio di educatori per persona,
 - il valore medio di indice di gradimento per valutazione espressa considerando tutte le attività della stessa tipologia,**Effettuare l'analisi separatamente per ogni trimestre. Ordinare e assegnare ad ogni record visualizzato una posizione all'interno di una classifica per numero medio di educatori per persona, separatamente per ogni attività. La posizione 1 va assegnata in corrispondenza del maggior numero medio di educatore per persona.**
- b) **Per le città in cui è sito almeno un centro in cui sono state svolte complessivamente 200 ore di attività nel 2012, visualizzare per ogni città e bimestre**
 - l'incasso medio per ore di attività, separatamente per fascia di età,
 - la percentuale di ore di attività per fascia di età rispetto al totale della categoria,
 - l'incasso cumulativo bimestrale, separatamente per fascia di età.
- c) **Separatamente per i bambini e i ragazzi, visualizzare per ogni tipologia di attività e anno, le attività con il maggior numero di ore svolte. Per tali attività visualizzare il numero di ore svolte, l'incasso totale e l'incasso medio per persona.**
- d) Per ogni coppia (fascia oraria, provincia), visualizzare l'incasso bimestrale, l'incasso medio bimestrale per persona e il numero totale di persone iscritte per ogni bimestre.
- e) Per ogni semestre e attività, visualizzare il numero medio di ore frequentate per persona, considerando solo i centri situati a Torino.
- f) Selezionare per ogni bimestre dell'anno 2015 e per ogni città, l'incasso medio per persona e il numero medio di ore frequentate per persona.
- g) Per ogni tipo attività e categoria di persone, visualizzare il numero totale di educatori e l'indice medio di gradimento fornito per valutazione espressa nel mese di Luglio 2015.
- h) Per ogni tipo attività e per ogni anno, visualizzare il numero medio di persone per educatore, separatamente per ogni categoria di persone.
- i) Per ogni tipo attività e per ogni semestre, visualizzare l'indice di gradimento medio per valutazione espressa, separatamente per ogni categoria di persone.

Progettazione

1. Il data warehouse conterrà le informazioni relative agli anni 2008-2015. Al fine di una corretta realizzazione del data warehouse si ipotizzino le cardinalità dei diversi attributi dimensionali.
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 potrebbe essere utile definire 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.