

Business Intelligence

Appello del 20 Luglio 2016

Descrizione del problema

La VVFF è una rinomata azienda di produzione di tessuti in lana pregiata che ha diversi stabilimenti sparsi in tutta Europa. Ogni stabilimento dispone di diversi macchinari utilizzati per la lavorazione della lana fino al prodotto finito (denominati tessuti) destinati alla vendita.

I tessuti prodotti dall'azienda sono ricavati dalla lavorazione di diversi titoli, così come più titoli possono realizzare più tessuti. I titoli sono i diversi fili utilizzati durante la lavorazione. Ad esempio, il tessuto "Flanella Pettinata" è realizzato utilizzando i titoli "Filo Grezzo" e "Filo Pettinato".

Inoltre ciascun tessuto è ottenuto utilizzando delle percentuali (variabili) di diversi titoli. I tessuti sono inoltre suddivisi in linee (fantasia, unita e classica) e in famiglie.

Per la realizzazione della produzione l'azienda dispone di diverse risorse, ciascuna associata ad un unico centro di costo. Ciascun centro di costo presta servizio per una particolare area operativa aziendale di interesse (e.g., Area Campioneria, Area Disegnatori, Tintoria, etc). Ogni area fa riferimento alla rispettiva macroarea di lavorazione. Inoltre per disporre di strumenti di controllo, l'azienda VVFF vorrebbe poter analizzare la propria produzione attraverso diversi scenari, per poter effettuare un'efficiente analisi degli scostamenti tra lo scenario di budget e quello consuntivo.

Attualmente ogni stabilimento ha una base di dati indipendente per gestire la propria attività economica, ma l'azienda VVFF è interessata a progettare un data warehouse che offra una visione di insieme che consenta di analizzare:

1. la produzione media mensile dei diversi stabilimenti;
2. gli incassi derivanti dalla vendita dei tessuti ai diversi clienti sparsi per il mondo.

Per il punto (1) l'azienda è interessata ad analizzare la quantità di metri di tessuto prodotti, il costo medio per metri di tessuto prodotti e il valore totale della produzione in funzione di:

- mese, mese dell'anno, bimestre, trimestre, quadrimestre e anno;
- titolo (e.g., Filo grezzo, Filo pettinato, etc), tessuto (e.g., Flanella pettinata, Tela Tradizionale, Raso Lana Cashmere, etc);
- percentuale di titolo variabile per un tessuto (inteso come la percentuale di titolo utilizzata per un determinato tessuto);
- famiglia e linea del tessuto;
- categoria del titolo;
- stabilimento;
- caratteristiche dello stabilimento in termini di attrezzature presenti: presenza della caldaia, del pettine, dell'orditoio, etc;
- città, provincia, regione, e stato in cui è sito lo stabilimento;
- tipo di scenario;
- risorsa, centro di costo, area, macroarea.

Per il punto (2) l'azienda è interessata ad analizzare l'incasso medio mensile, l'incasso per posizione geografica del cliente e i metri di tessuto venduti mediamente per cliente in funzione di

- città, provincia, regione, stato, continente in cui è sito il cliente;
- tessuto, famiglia e linea del tessuto;
- tipo di pagamento;
- città, provincia, e stato in cui è sito lo stabilimento;
- mese, bimestre, trimestre, semestre e anno.

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

- a) **Separatamente per lo scenario di budget e consuntivo, per ogni tessuto e trimestre visualizzare**
- **la quantità totale di metri di tessuto prodotti,**
 - **il costo medio per metri di tessuto prodotti,**
 - **il valore totale della produzione,**
 - **la percentuale di metri di tessuto prodotti rispetto al totale di metri di tessuto della stessa famiglia,**
 - **la percentuale di metri di tessuto prodotti rispetto al totale di metri di tessuto prodotti da tutti gli stabilimenti siti nella stessa regione.**

Effettuare l'analisi per ogni stabilimento.

- b) **Separatamente per le linee fantasia, unita e classica, visualizzare per ogni città dello stabilimento e semestre,**
- **gli stabilimenti dove sono stati prodotti la quantità massima di tessuto,**
 - **la quantità prodotta con il relativo costo e valore totale della produzione.**

Ordinare e assegnare un rank in ordine decrescente di quantità di tessuto prodotto.

- c) **Per ogni stato in cui è sito almeno un cliente e per ogni anno, visualizzare**
- **l'incasso medio mensile,**
 - **l'incasso cumulativo al trascorrere degli anni**
 - **i metri di tessuto venduti mediamente per cliente**

Effettuare l'analisi separatamente per linea di tessuti. Ordinare e assegnare un rank in ordine decrescente di metri di tessuto venduti mediamente per cliente, separatamente per ogni anno.

- d) Considerando soltanto lo scenario di budget, per ogni coppia (anno, titolo), visualizzare i metri totali di tessuto prodotti e il corrispondente costo totale. Effettuare l'analisi separatamente per ciascun stabilimento.
- e) Per ogni coppia (scenario, anno), visualizzare il valore totale prodotto, considerando gli stabilimenti con solo l'orditoio e il pettine.
- f) Per ogni stabilimento, visualizzare i metri di tessuto prodotti e il valore totale della produzione, separatamente per ogni mese.
- g) Considerando soltanto i clienti europei, selezionare per ogni coppia (stato, tessuto), l'incasso totale e il numero di clienti, separatamente per ogni anno.
- h) Per la linea fantasia, visualizzare per ogni bimestre e per ogni città dello stabilimento, i metri totali venduti, l'incasso totale e il numero medio di metri di tessuto venduti per cliente.
- i) Per ogni mese dell'anno 2015, visualizzare il numero totale di clienti che hanno effettuato un acquisto, separatamente per ogni tipo di pagamento.

Progettazione

1. Il data warehouse conterrà le informazioni relative agli anni 2009-2015. Al fine di una corretta realizzazione del data warehouse si ipotizzino le cardinalità dei diversi attributi dimensionali.
2. Esprimere le interrogazioni frequenti **(a)**, **(b)** e **(c)** delle specifiche del problema utilizzando il linguaggio SQL esteso.
3. Considerando le caratteristiche del data warehouse realizzato e le 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.