



Data warehouse

Progettazione di un data warehouse

Tania Cerquitelli

Esercizio: Made in Italy

Nell'ambito della valorizzazione delle eccellenze del Made in Italy, il Governo italiano vuole analizzare le esportazioni di vini italiani all'estero, per capire, ad esempio, quali sono i vini più richiesti, da quali zone d'Italia provengono e quali sono i principali Stati esteri importatori.

Il Governo ha raccolto dalle aziende vinicole italiane l'elenco degli ordini di acquisto. Ogni ordine di acquisto è caratterizzato da data dell'ordine, importo in euro, quantità in litri, tipo di vino, tipo di confezionamento (es. bottiglia, cartone, damigiana, ecc.), numero di partita IVA dell'azienda vinicola di provenienza, stato estero di destinazione.

Gli analisti, a partire dagli ordini di acquisto, vogliono creare un data warehouse che sintetizzi i dati raccolti per rispondere ad un serie di interrogazioni.

In fase di realizzazione, gli analisti integrano i dati degli ordini con dei dati di contesto in loro possesso, necessari per svolgere le analisi di interesse. In particolare, ad ogni tipo di vino viene associato l'elenco delle certificazioni di qualità (DOC, DOP, o DOCG) che possiede. Per esempio: il tipo di vino "Barolo" è sempre associato alla certificazione DOCG, il tipo di vino "Nebbiolo" è sempre associato alle certificazioni DOP e DOCG, ecc. Ogni certificazione può essere associata a più tipi di vino. Alcuni tipi di vino possono essere senza certificazioni.

Esercizio: Made in Italy

Ad ogni ordine è, inoltre, associata la sua dimensione: sono considerati piccoli gli ordini fino a 100 litri, medi quelli fino a 1000 litri, e grandi quelli di quantità superiore. Inoltre è possibile risalire dal numero di partita IVA ai dati anagrafici completi (denominazione, indirizzo, provincia, regione, area geografica della regione).

Infine, per ogni azienda vinicola, Le interrogazioni cui sono interessati gli analisti riguardano la quantità a totale di vino esportata (in litri), e il prezzo medio per litro, al variare di:

- mese, bimestre, trimestre, quadrimestre, semestre e anno;
- provincia, regione e area geografica (Nord, Centro, Sud) dell'azienda vinicola;
- tipo di vino
- certificazioni di qualità
- tipo di confezionamento
- stato estero di destinazione, e relativo continente
- dimensione dell'ordine (piccolo, medio, grande)

Il data warehouse tratterà le informazioni relative agli anni dal 2010 al 2020

Esercizio: Made in Italy: interrogazione frequente

Considerando solo i vini che hanno la certificazione ``DOC'' esportati in Asia, selezionare il prezzo medio al litro per ogni anno, la percentuale di litri esportati nell'anno rispetto al totale di tutti gli anni, e il totale cumulativo annuale di litri esportati. Si effettui l'analisi separatamente per ogni tipo di confezionamento.

Esercizio: Made in Italy: interrogazione frequente

Per ogni regione, selezionare il prezzo medio al litro, il numero medio di litri di vino esportati per provincia (es. il Piemonte ha 8 province, quindi il totale dei litri esportati dal Piemonte deve essere diviso per 8) e la percentuale di litri di vino esportati da ogni regione rispetto al totale dell'area geografica di appartenenza (Nord, Centro, Sud). Si effettui l'analisi separatamente per ogni anno..