

3. Progettazione Data Warehouse

Il mercato della musica digitale è in forte espansione e le piattaforme Internet che offrono la possibilità di scaricare musica legale cresce di giorno in giorno. Ogni piattaforma memorizza localmente tutte le informazioni sui download dei brani musicali richiesti dai propri clienti. Per poter valutare l'efficienza di questo nuovo mercato e capire in quale direzione può essere ampliato, vi è la necessità di progettare un data warehouse che offra una visione d'insieme su questo nuovo mercato. In particolare, il data warehouse deve essere progettato per analizzare:

- il tempo medio di download per brano musicale,
- il numero medio di download giornaliero,
- il ricavo giornaliero medio per brano musicale.

L'analisi deve essere effettuata in funzione di:

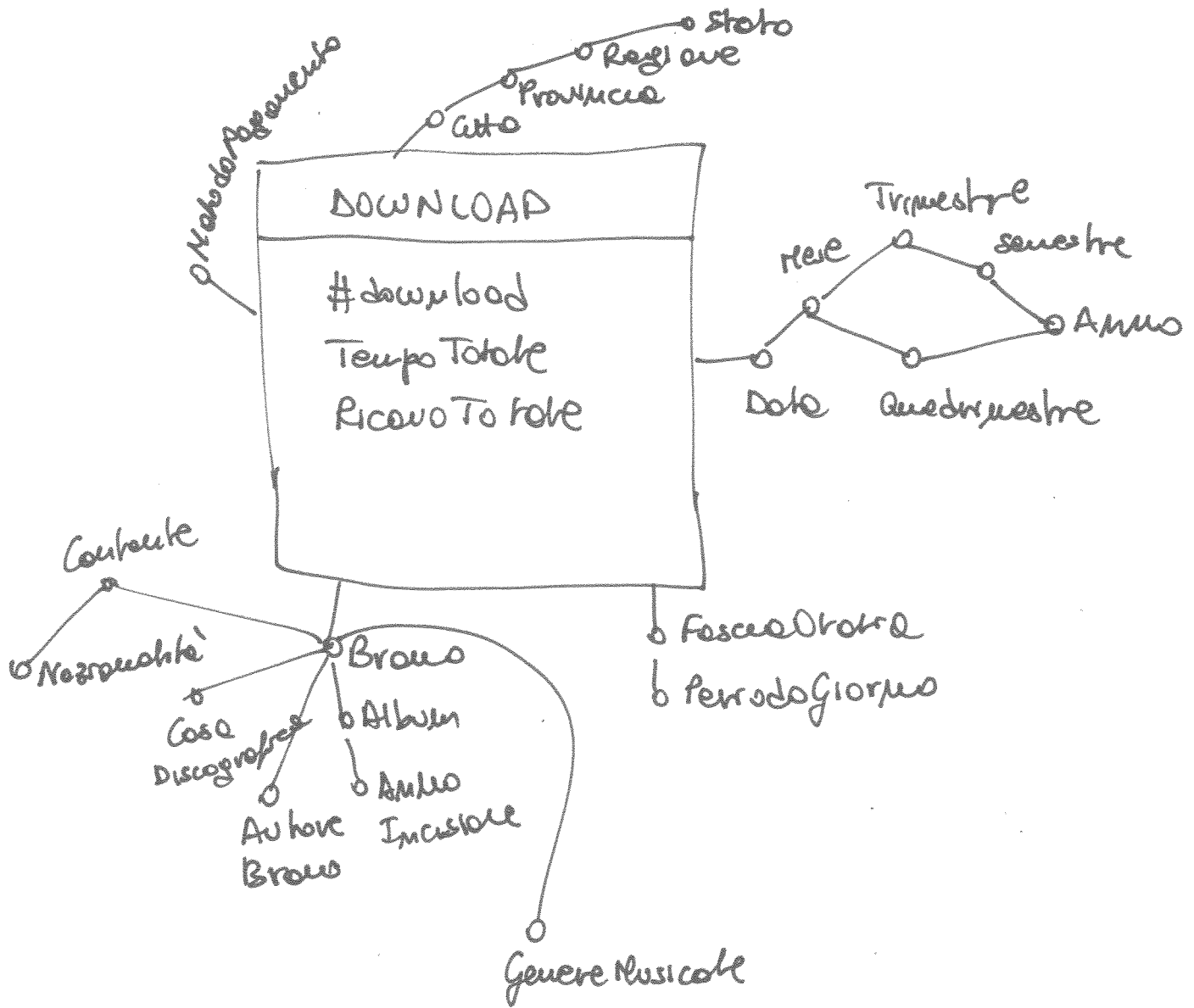
- data, mese, trimestre, quadrimestre, semestre e anno in cui è stato effettuato il download,
- fascia oraria (8:00-12:00, 12:00-16:00, 16:00-20:00, etc) e periodo del giorno (a.m., p.m.) in cui è stato effettuato il download,
- brano musicale, casa discografica che ha curato la registrazione del brano, album in cui il brano è stato pubblicato, anno di incisione dell'album,
- autore del brano musicale,
- genere musicale del brano,
- cantante/band e rispettiva nazionalità,
- città, provincia, regione e stato dell'utente che ha effettuato il download,
- metodo di pagamento (carta di credito, carta prepagata, etc.).

Il data warehouse conterrà le informazioni relative agli anni 2007, 2008, 2009, 2010, 2011. Sono riportate di seguito alcune delle interrogazioni frequenti di interesse, alle quali il data warehouse deve poter rispondere in modo efficiente:

- (a) Considerando solo i download effettuati nel secondo trimestre dell'anno 2010 da utenti situati in Abruzzo, visualizzare, separatamente per ogni album, il numero totale di download e il tempo medio di download.
- (b) Per ogni autore di un brano musicale e per ogni fascia oraria, visualizzare il numero totale di download separatamente per ogni casa discografica. Ordinare e assegnare un rank in ordine crescente di numero totale di download.
- (c) Per ogni brano musicale di cantanti italiani e considerando solo i download effettuati nel 2011, visualizzare per ogni mese il ricavo complessivo, il ricavo giornaliero medio, il numero medio giornaliero di download, il numero totale di download cumulativo dall'inizio dell'anno, separatamente per ogni genere musicale.

Progettazione

- (a) (6 Punti) Progettare il data warehouse in modo da soddisfare le richieste descritte nelle specifiche del problema. Il data warehouse progettato deve inoltre permettere di rispondere in modo efficiente a tutte le interrogazioni frequenti indicate.
- (b) (8 Punti) Esprimere le interrogazioni frequenti (a) e (c) utilizzando il linguaggio SQL esteso.



TEMPO - I (CodT1, Data, Mese, Trimestre, Semestre, Quadrimestre, Anno)

TEMPO - II (CodT2, Fascicolo, PeriodoGiorno)

BRANO (CodB, Branco, Album, Anno Incisione, Genere Musicale, Autore Branco, Case Discografico, Contante, Notorieta)

MET-PAG (CodMP, MetodoPagamento)

WOGO (CodL, Città, Provincia, Regione, Stato)

DOWNLOAD (CodT1, CodT2, CodB, CodMP, CodL, #download, Tempo Totale, Ricavo Totale)

(b) SELECT Album
 SUM (#download)
 SUM (TempoTotale) / SUM (#download)

FROM DOWNLOAD D, TEMPO-I T, BRANO B,
 LUOGO L

WHERE D.CodT1 = T.CodT1 AND D.CodB = B.CodB
 AND D.CodL = L.CodL AND Anno = '2010'
 AND Trimestre = '2' AND Regione = 'Abruzzo'

GROUP BY Album

(c) SELECT Brano, GenereMusica, Hese
 SUM (RicavoTotale),
 SUM (RicavoTotale) / COUNT (DISTINCT Data),
 SUM (#download) / COUNT (DISTINCT Data),
 SUM (SUM (#download)) OVER
 (PARTITION BY Brano, GenereMusica
 ORDER BY Hese
 ROWS UNBOUNDED PRECEDING)

FROM DOWNLOAD D, TEMPO-I T, BRANO B

WHERE D.CodT1 = T.CodT1 AND D.CodB = B.CodB AND
 Anno = 2011 AND NomeMolice = 'Ithous'

GROUP BY Brano, Genere_Musica, Hese

Attributi che pilotano il GROUP BY