

Progetto di un data warehouse – BOZZA di Soluzione

Alberghi

Progettazione concettuale

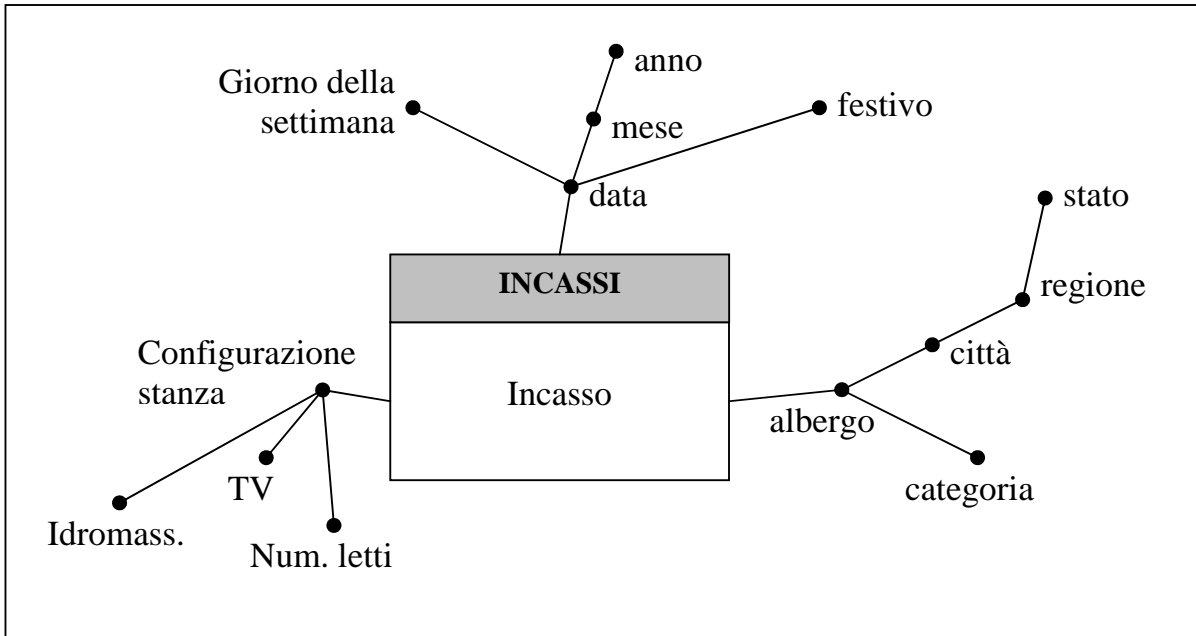


Figura 1 – Fatto Incassi

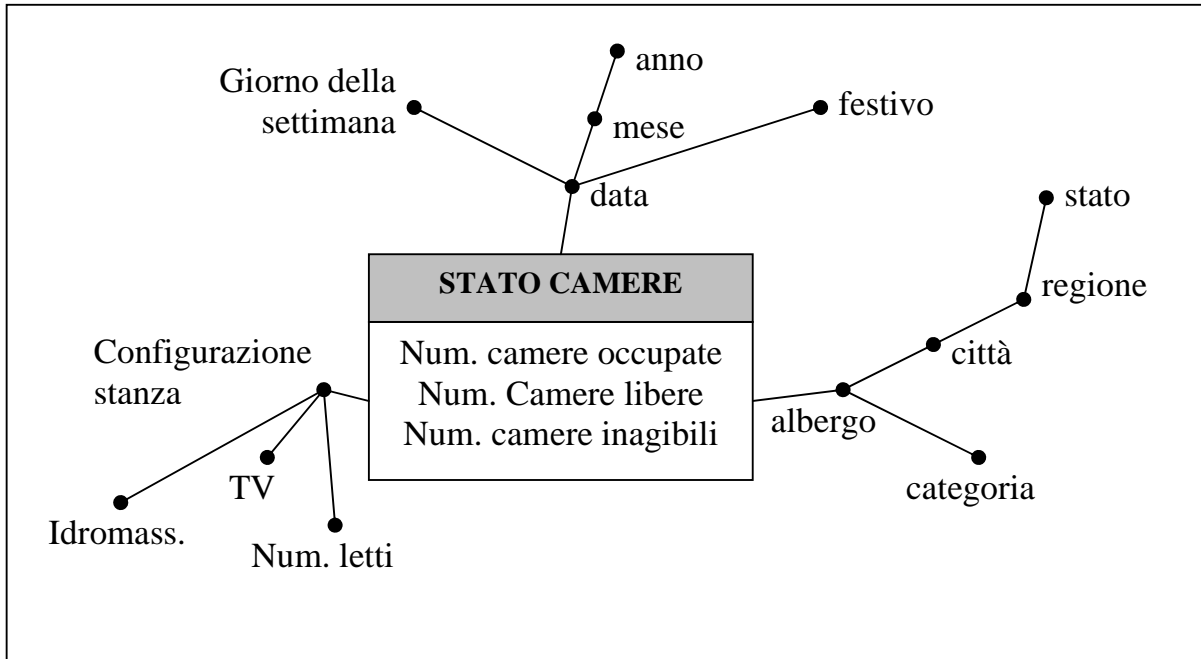


Figura 2 – Fatto Camere

Progettazione logica

Le dimensioni dei due fatti sono tutte identiche, quindi creo una sola tabella per ogni dimensione.

- Dimensioni

ALBERGO (CodA, Albergo, Categoria, Città, Regione, Stato)

TEMPO (CodT, Data, Mese, Anno, Giorno_settimana, Festivo)

CONFIGURAZIONE_STANZA (CodC, Num_letti, TV, Vasca_idromassaggio)

I due fatti sono simili (stesse dimensioni con lo stesso livello di dettaglio) posso quindi creare una sola tabella dei fatti con tutte le misure.

- Fatti

GESTIONE_CAMERE (CodA, CodT, CodC, Incasso, Num_cam_occup, Num_cam_libere, Num_cam_inag);

Gesione dinamicità delle dimensioni

Dimensione ALBERGO

- Cambia la città, la regione o lo stato in cui si trova l'albergo.
E' un evento molto raro. Se succede, posso usare lo scenario temporale TIPO1 (*oggi per ieri*)
- Cambia la categoria dell'albergo
 - Uso TIPO1 (*oggi per ieri*) se mi interessa l'albergo (sovrascrivo la nuova categoria e pertanto tutti gli eventi, anche quelli passati, verranno interpretati in base all'attuale configurazione)
 - Uso TIPO2 (*oggi o ieri*) se mi interessa fare analisi precise in base alla categoria (il cambio di categoria porta alla "generazione" di un nuovo albergo/tupla nella dimensione ALBERGO)

Dimensione CONFIGURAZIONE_STANZA

- In caso di aggiunta di una nuova caratteristica per le camere aggiungo un nuovo attributo alla dimensione.

Query

C - Relativamente all'anno 2005, selezionare per ogni coppia (stato, mese) l'incasso mensile calcolato considerando solo gli alberghi a 4 stelle e l'incasso cumulativo da inizio anno (sempre considerando solo gli alberghi a 4 stelle).

```
SELECT stato, mese, SUM(incasso), SUM(SUM(incasso) OVER
(PARTITION BY stato ORDER BY mese ROWS UNBOUNDED PRECEDING)
FROM incassi I, tempo T, albergo A
WHERE
I.codT=T.CodT AND I.CodA=A.CodA AND
Categoria='4stelle' AND
Anno = 2005
GROUP BY stato, mese
```

E - Calcolare per ogni albergo l'incasso totale effettuato nell'anno 2005 considerando solo le camere dotate di TV e vasca idromassaggio

```
SELECT albergo, SUM(incasso)
FROM incassi I, tempo T, albergo A, configurazione_stanza C
WHERE
    I.codT=T.codT AND I.codA=A.codA AND I.CodC=C.codC AND
    Anno = 2005 AND
    Tv='si' AND
    Vasca_idromassaggio='si'
GROUP BY A.CodA, A.Albergo
```

Viste

Stima cardinalità dei fatti:

INCASSI= 500(alberghi) * 2x365 (2 anni) * 2^8 (configurazioni) = circa 93x10^6

INTERROGAZIONI:

- Cardinalità risultato = 40 (stati) x 12 (mesi) che è molto minore della cardinalità del fatto. Può essere utile quindi definire una vista materializzata associata all'interrogazione a). Meglio definire però una vista materializzata più generale in cui non ho nella WHERE la condizione sull'anno. L'anno lo uso come attributo di raggruppamento.

VISTA 1 - seleziona SUM(camere_libere), SUM(camere_occupate), SUM(camere_inagibili) e raggruppa per (stato,mese,anno).

Cardinalità VISTA 1 = 40(stati) * 12(mesi) * 2(anni)

- Posso usare la VISTA 1. Aggregando i dati in modo opportuno (x stato) ottengo il risultato della query.
- Cardinalità risultato = 40(stati) * 12 (mesi del 2005), molto minore della cardinalità del fatto. Utile definire una vista con la categoria.

VISTA 2 – seleziona SUM(incasso) e raggruppa per (stato,mese,anno,categoria)

cardinalità VISTA 2 – 40*12*2*7(categorie alberghi stimato)

Con questa vista posso rispondere, raggruppando adeguatamente, alla query c).

- Creo VISTA 3 dove raggruppo per (stato, anno, festivo)
- VISTA 4 raggruppo per (albergo, anno, tv, vasca)