

Sistemi per la gestione di basi di dati

Esercitazione di laboratorio n. 6

SQL Server 2005

Durata: 2 settimane

L'esercitazione consiste nel progettare un data warehouse che permetta di gestire la problematica illustrata nei punti seguenti. Usare SQL Server per creare le tabelle e i cubi necessari per realizzare il data warehouse. Il server SQL Server a cui collegarsi è installato sulla macchina CCLIX3.

1. Progettazione del data warehouse

Una società che gestisce degli stabilimenti balneari italiani vuole analizzare gli incassi dei suoi stabilimenti nel corso degli anni. Attualmente ogni stabilimento ha una propria base di dati in cui memorizza giornalmente le informazioni relative agli oggetti affittati (ombrelloni, lettini, pedalò, ecc.) e il prezzo a cui sono stati affittati. Lo schema concettuale della base di dati di uno stabilimento balneare (sorgente dati Stab1) è mostrato in Figura 1.

La società vuole aggregare le informazioni provenienti dagli stabilimenti in un proprio data warehouse centralizzato. Il data warehouse deve essere progettato in modo tale da permettere di rispondere velocemente ad alcune interrogazioni effettuate frequentemente dai dirigenti della società.

Alcune delle informazioni a cui i dirigenti della società sono interessati sono le seguenti:

- incassi effettuati su base mensile in ogni provincia e in ogni regione
- incassi effettuati su base annua in ogni provincia e in ogni regione
- incassi effettuati su base trimestrale in ogni provincia e in ogni regione
- incassi effettuati su base quadrimestrale in ogni provincia e in ogni regione
- incassi effettuati su base mensile, trimestrale, quadrimestrale, annuale per ogni categoria di oggetti in funzione della provincia e della regione
- numero di oggetti affittati su base mensile, trimestrale, quadrimestrale, annuale in funzione della categoria di oggetto
- numero di oggetti affittati su base mensile, trimestrale, quadrimestrale, annuale in funzione della provincia e della regione.

La base di dati Stab1, presente sul server CCLIX3, è la base di dati di uno degli stabilimenti balneari.

Durante l'esercitazione dovranno essere eseguite le seguenti attività:

- ❖ Analizzare tale schema al fine di capire quali dati sono presenti nelle basi di dati dello stabilimento Stab1 (sorgente del data warehouse che deve essere creato).
- ❖ Progettare un data warehouse per la gestione delle problematiche illustrate in precedenza.
- ❖ Realizzare gli script necessari per la creazione delle tabelle del data warehouse. Utilizzare come nome del data warehouse il prefisso DW seguito dalla propria matricola (e.g., matricola s12345 nome del data warehouse DWs12345).

2. Caricamento dati

Realizzare le procedure ETL necessarie per effettuare il caricamento dei dati presenti nella base di dati Stab1 all'interno delle tabelle del data warehouse che avete creato al punto 1. Seguire come linea guida per la realizzazione delle procedure ETL la seguente descrizione dei macro blocchi del workflow.

- Creazione tabella temporanea e caricamento dati
 - creare una tabella temporanea *StagingTable* nella base di dati relazionale associata al data warehouse. Lo script di creazione sarà fornito durante l'esercitazione. Lo schema logico della *StagingTable* il seguente
StagingTable (month, year, province, category, numberRentedItems, netIncome)
 - caricare nella *StagingTable* i dati della sorgente (Stab1) in forma già aggregata (in funzione della granularità desiderata nel data warehouse). Ciò permette di avere i dati, già aggregati, in locale (nella base di dati relazionale associata al data warehouse). Nei passi successivi sarà possibile fare riferimento a tale tabella, più compatta di quella originale, senza dover accedere a Stab1.

- Caricamento dei dati nelle dimensioni
 - caricare i dati nelle tabelle associate alle dimensioni. I dati sorgenti sono ottenuti interrogando i dati presenti nella *StagingTable*.
- Caricamento dei dati nella tabella dei fatti
 - al termine del caricamento dei dati nelle tabelle associate alle dimensioni caricare i dati nella tabella dei fatti.

Durante l'esercitazione dovranno essere eseguite le seguenti attività:

- ❖ Usare Microsoft Visual Studio 2005 (Integration Services Project) per realizzare le procedure ETL necessarie per effettuare il caricamento dei dati presenti nella base di dati Stab1 all'interno del data warehouse che avete creato al punto 1.

3. Interrogazioni e creazione dei cubi

Utilizzare il data warehouse creato e alimentato ai punti 1 e 2 come sorgente per la fase di creazione dei cubi.

Sono di seguito elencate alcune delle interrogazioni che vengono frequentemente eseguite.

1. Calcolare per ogni coppia regione, trimestre l'incasso totale. Calcolare inoltre l'incasso totale per ogni regione e l'incasso totale per ogni trimestre.
2. Calcolare l'incasso totale per ogni coppia *[categoria oggetto affittato, mese]*. Calcolare inoltre l'incasso totale per categoria di oggetto (indipendentemente dal mese) e l'incasso totale per ogni mese (indipendentemente dalla categoria di oggetti affittati).
3. Calcolare il numero totale di oggetti affittati per ogni coppia *[categoria oggetto affittato, mese]*. Calcolare inoltre il numero totale di oggetti affittati per categoria di oggetto (indipendentemente dal mese) e il numero totale di oggetti affittati per ogni mese (indipendentemente dalla categoria di oggetti affittati).

Durante l'esercitazione dovranno essere eseguite le seguenti attività:

- ❖ Usare Microsoft Visual Studio 2005 (Analysis Services Project) per creare i cubi ritenuti utili per rispondere velocemente alle interrogazioni proposte.
- ❖ Provare a "navigare" i cubi creati tramite l'interfaccia grafica messa a disposizione da Microsoft Visual Studio 2005 (Analysis Services Project).

4. Uso dei cubi in Excel

Durante l'esercitazione dovranno essere eseguite le seguenti attività:

- ❖ Visualizzare i risultati di uno dei cubi creati all'interno di Excel utilizzando l'opzione Tabella Pivot.

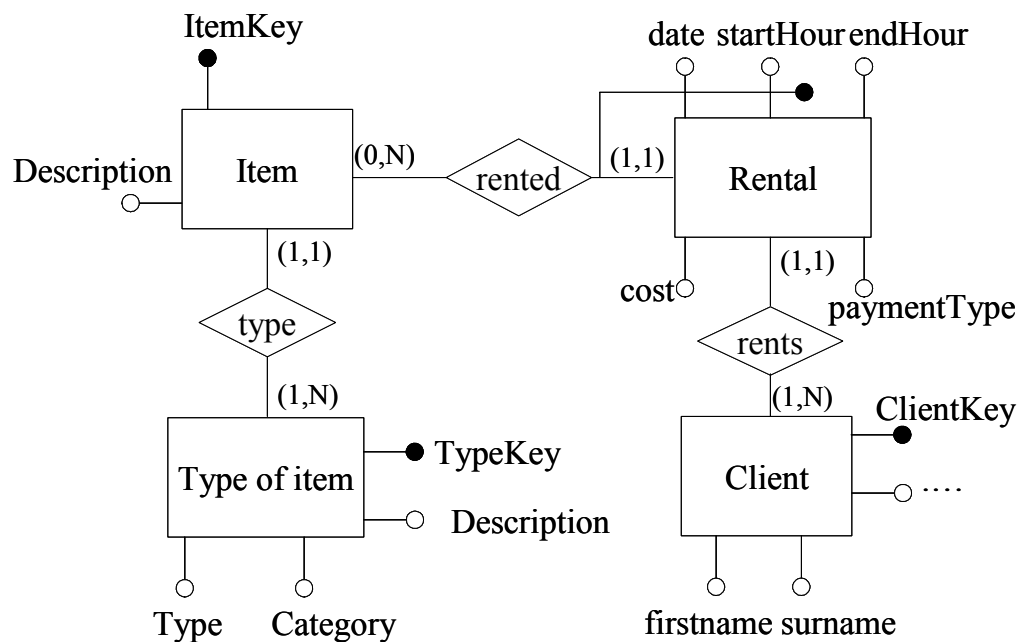


Figura 1. Modello concettuale della sorgente di dati Stab1

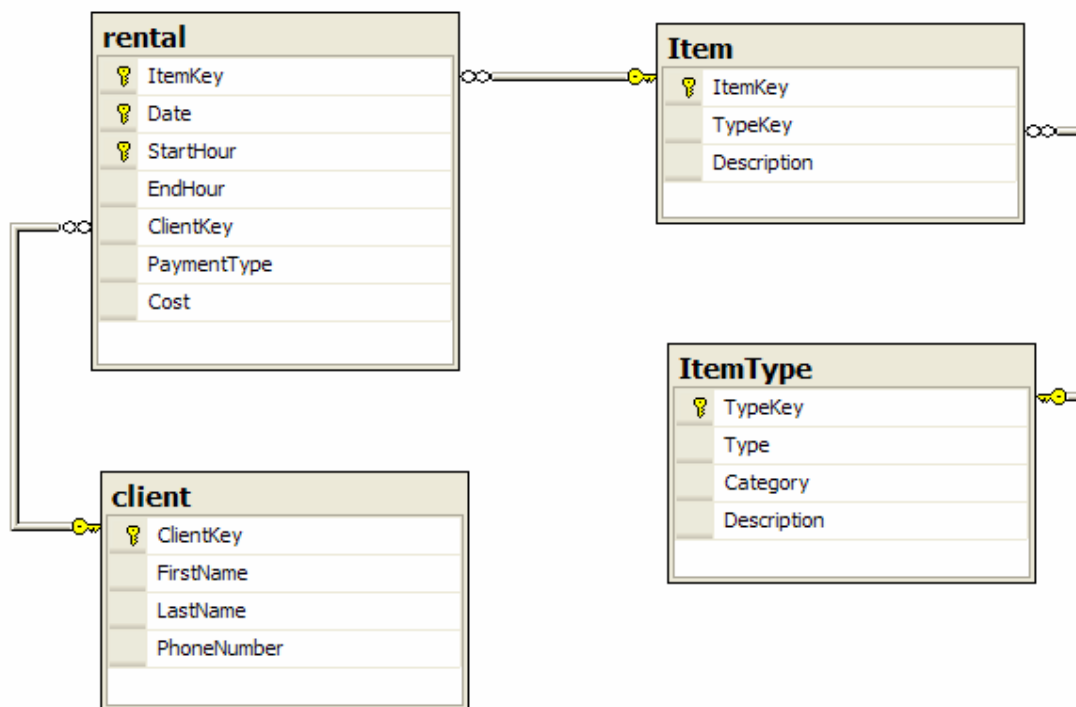


Figura 2 – Schema della base di dati Stab1 (sorgente dati presente in ogni stabilimento usata per alimentare il data warehouse)