

# Database e Data Mining

## *Esercitazione N. 2*

L'esercitazione consiste nel progettare un data warehouse che permetta di gestire la problematica illustrata nei punti seguenti, mediante l'utilizzo della suite Pentaho Community Edition e relativi plug-in.

### **1. Descrizione del problema**

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, mentre lo schema logico è mostrato in Figura 2.

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

Ai fini dell'esercitazione si utilizzi l'apposito script per la creazione del data warehouse e il caricamento dei relativi dati.

*Durante l'esercitazione dovranno essere eseguite le seguenti attività:*

- *Creazione di cubi*
- *Creazione di report*
- *Creazione di dashboard*

## 2. Creazione di cubi

Utilizzando il data warehouse creato e alimentato mediante lo script fornito e tenendo in considerazione le query più frequentemente eseguite:

- creare uno o più cubi
- pubblicare i cubi creati su Pentaho BI User Console
- visualizzare e analizzare i 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:*

- *Utilizzare Pentaho schema workbench per costruire i cubi*
- *Utilizzare Pentaho BI User Console e Saiku Analytics per visualizzare ed interagire con i cubi*

## 3. Creazione di report

Utilizzando il data warehouse creato e alimentato mediante lo script fornito:

- Creare un report che visualizzi per ciascun **anno** una tabella con l'**incasso totale** di ogni **mese** per ciascuna **categoria di oggetto affittato**
- Creare un report che visualizzi per ciascun **anno** una tabella e un grafico a barre con l'**incasso totale** per ciascuna **categoria di oggetto affittato**
- Pubblicare i report realizzati su Pentaho BI User Console
- Visualizzare i report ed esportarli in formato PDF

*Durante l'esercitazione dovranno essere eseguite le seguenti attività:*

- *Utilizzare Pentaho report designer per realizzare i report*
- *Utilizzare Pentaho BI User Console per visualizzare ed esportare i report*

## 4. Creazione di dashboard

Utilizzando il data warehouse creato e alimentato mediante lo script fornito:

- Creare un dashboard che consenta di selezionare un **anno** di attività dello stabilimento balneare e per l'anno selezionato visualizzare un grafico a barre con il **numero di oggetti affittati** per ciascuna **categoria di oggetto**
- Creare un dashboard che consenta di selezionare un **anno** di attività dello stabilimento balneare e per l'anno selezionato visualizzare un grafico a barre con l' **incasso totale** per ciascuna **provincia**

*Durante l'esercitazione dovranno essere eseguite le seguenti attività:*

- *Utilizzare Community Dashboard Editor (CDE) per realizzare la dashboard*
- *Utilizzare Pentaho BI User Console per visualizzare la dashboard*

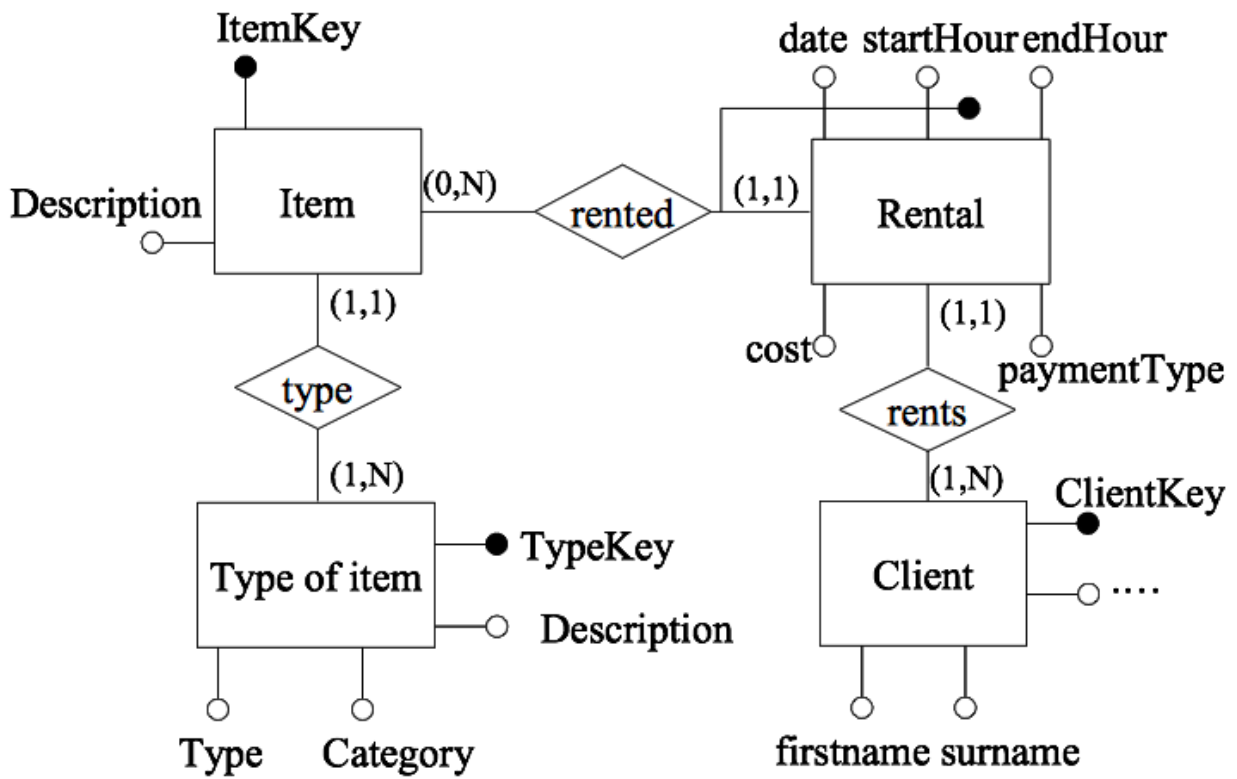


Figura 1. Modello concettuale della sorgente Stab1

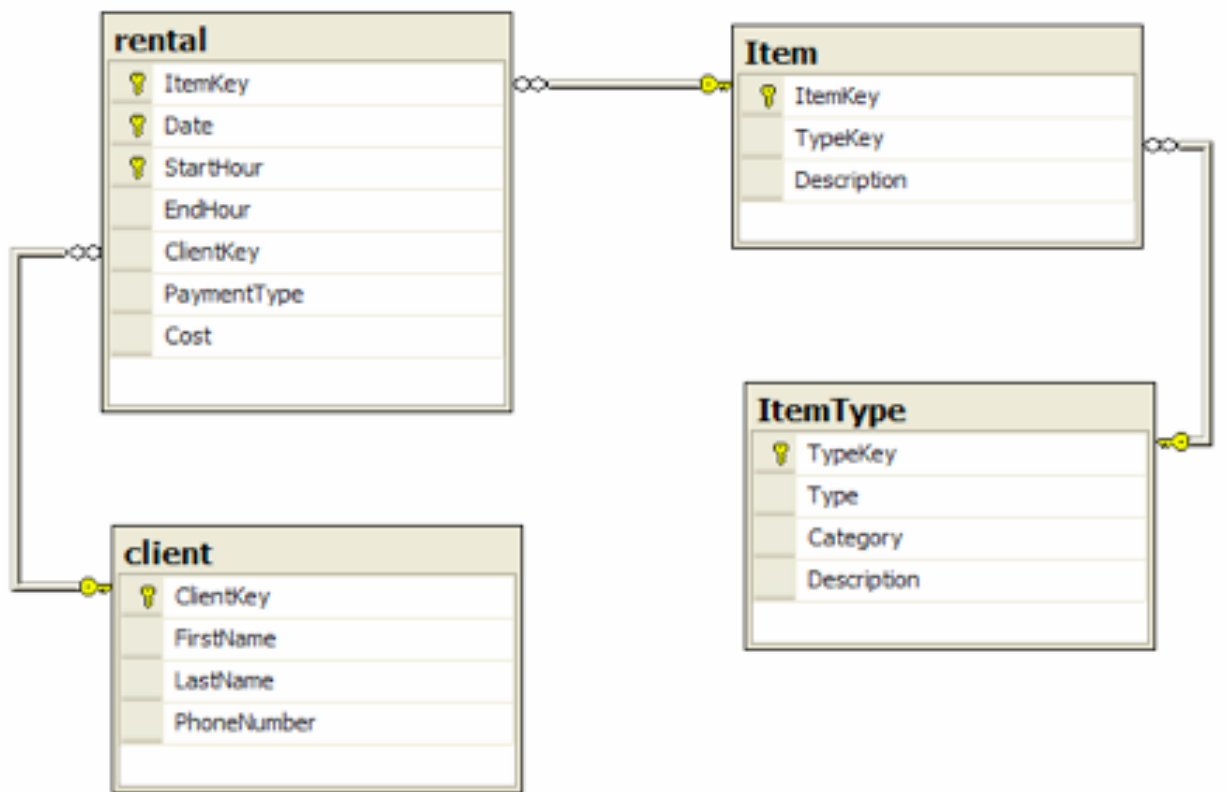


Figura 2 – Schema logico della sorgente Stab1