

Data warehouse design exercise

Problem specifications

Design a data warehouse to manage the property rent adverts of the website www.SearchingYourHouse.com. It allows searching all the Italian cities to find rentals. People or agencies can publish property rent adverts on the website, specifying the property characteristics: city area (district), type of property (penthouse, attic, detached house, terraced house, etc.), monthly rent price, surface (number of square meters), number of rooms, number of bathrooms/toilets, floor, elevator presence, cellar presence, garage presence. Furthermore, the available furniture can be specified: tables, chairs, fridge, oven, heaters, beds, TV, etc.

Adverts are updated weekly. Every Monday the rented property adverts are removed and the new property adverts are inserted. The website users can add to their personal favorite list the properties they are interested in.

The website company wants to analyze the Italian rental situation. The analysis parameters are: the number of available properties to rent, the average monthly rent price for each property, and the average monthly rent price for each square meter of surface. These parameters must be available for each:

- week, month, 2-month period, trimester, 4-month period, semester, and year;
- city area (district), city, province, region, country area (Northern, Central, Southern Italy);
- presence of a university in the city; - type of property and available furniture; - number of rooms.

The website company wants to compute some statistics on the user favorite list of properties, thus evaluating the average number of users interested in a property for each:

- season (winter, spring, summer, and autumn) and year;
- city area (district), city, province, region, country area (Northern, Central, Southern Italy);
- presence of a university in the city;
- type of property and available furniture;
- the range of the monthly rent price (100-200 €, 200-300 €, etc.) and the range of the property surface (0-50 m², 50-100 m², etc.) - number of rooms.

The following are **some** of the frequent queries the website company is interested in:

- a) In 2004, including only the properties in cities where universities are present, select the average monthly rent cost for each city and month and the average monthly rent cost since the beginning of the year for each city and month.
- b) In September 2004, including only the properties in the province of Turin, select the total number of free properties for each city and week, the ratio of the total number of free properties for each city and week and the total number of free properties for the same week.
Give a rank according to the total number of free properties (rank 1st the highest number). Sort the data according to the rank.
- c) In summer 2005, including only the attics with bed, fridge, and table in Rome, for each city area (district) and for each range of the monthly rent price, select the average number of users per each property who have added the property in their favorite list, and select the average number of users per each property of the city area who have added the property in their favorite list. Order the results according to the city area and the average number of users.
- d) Including only the properties with bed and table in cities where universities are present, select the average monthly rent cost per property for each city, month, and year, select the average monthly rent cost per square meter for each city, month, and year, and select the average monthly rent cost per property of the city since the beginning of the year for each city, month, and year.
- e) Including only the properties in Piedmont (region), in September, October, and November 2004, select for each city the average monthly rent cost per property and the average monthly rent cost per property of the province in which the city is located.
- f) In 2004, including only the properties with bed and table in cities where universities are present, select the average monthly rent cost per property for each city and month, and the average monthly rent cost per square meter for each city and month.

Design

The data warehouse will store information of 2004 and 2005. The following cardinalities are known:

- Italian country areas: ~4, Provinces: ~100, Cities: ~8000
 - City areas (districts): ~10000, Cities with universities: ~100
 - Type of properties: ~5, Type of furniture: ~10, Rooms: ~1 to 5
 - Monthly rent cost ranges: ~10, Property surface ranges: ~10
1. Design the data warehouse to address the described issues and queries.
 2. Write the frequent queries using the extended SQL language.
 3. Decide whether and which materialized views and/or indexes are convenient.