

Data Mining & Knowledge Extraction

Assignment 3

Deadline

Strict deadline **January 21st 2024 at 23:59**.

The deadline is definitive and it already accounts for vacation days.

No submission after the deadline will be considered.

Instructions

You are given an E/R diagram of a relational database of a big hotel chain, which can be found at the end of this document. The diagram describes a database storing information about all the hotels of the chain, their rooms, and the history of room reservations made by customers.

In particular, a reservation is made for a certain room by a certain customer, on a certain date. Each reservation determines the price of the room. Rooms have a number of beds, and hotels have a location and a number of stars (from 1 to 5).

Your goal is to design a data warehouse (actually just one data mart) that allows the hotel chain to analyze **room reservations**. In particular, the data mart must allow to analyze:

1. the number of reservations made
2. the income made by reservations

The above quantities must be known according to:

3. The date of the reservation at different levels: day, month, quarter, year
4. The size of the room reserved, expressed as small (#beds = 1), medium ($2 \leq \text{\#beds} \leq 4$), high (#beds ≥ 5)
5. The hotel location at different levels: city, region, country
6. The hotel quality (using the number of stars)
7. The customer location at the level of region, and country.

Design a Dimensional Fact Schema, using the attribute tree methodology, that implements the above 7 points.

Marking

The total mark of your assignment is 8 points, where you can get 1 point for each criterion above that your dimensional fact schema allows to implement (0 if not correct), plus 1 point for producing the correct **initial** (i.e., **without any grafting/pruning**) attribute tree (0 if not correct).

Delivery

- Go to the following Google Form: <https://forms.gle/Dpz9CEFDfEtYRr69>
- Upload the pdf file containing a picture of the attribute tree, a description of the eventual pruning/grafting operations you applied to the tree, and a picture of the final dimensional fact schema.
- You can retry your upload a maximum of 10 times (**every upload REPLACES the previous one**).

E/R Diagram

