Professor:
Dr. Laura E. RAILEANU

Assignement Nº 1

Data Warehousing and OLAP

A.

Briefly compare the following concepts. You may use an example to explain your point(s):

(a) Snowflake schema, fact constellation, starnet query model
(b) Data cleaning, data transformation, refresh
(c) Discovery-driven cube, multifeature cube, virtual warehouse

B.

Suppose that a data warehouse consists of the three dimensions time, doctor, and patient, and the two measures count and charge, where charge is the fee that a doctor charges a patient for a visit.

(a) Enumerate three classes of schemas that are popularly used for modelling data warehouses.

(b) Draw a schema diagram for the above data warehouse using one of the schema classes listed in (a).

(c) Starting with the base cuboid [day, doctor, patient], what specific OLAP operations should be performed in order to list the total fee collected by each doctor in 2008?

(d) To obtain the same list, write an SQL query assuming the data is stored in a relational database with the schema fee (day, month, year, doctor, hospital, patient, count, charge).

C.

Suppose that a data warehouse for *Big-University* consists of the following four dimensions: *student, course, semester*, and *instructor*, and two measures *count* and *avg grade*. When at the lowest conceptual level (e.g., for a given student, course, semester, and instructor combination), the *avg grade* measure stores the actual course grade of the student. At higher conceptual levels, *avg grade* stores the average grade for the given combination.

(a) Draw a *snowflake schema* diagram for the data warehouse.

(b) Starting with the base cuboid [student, course, semester, instructor], what specific *OLAP operations* (e.g., roll-up from *semester* to *year*) should one perform in order to list the average grade of *CS* courses for each *Big-University* student.

(c) If each dimension has five levels (including all), such as "*student* < *major* < *status* < *university* < all", how many cuboids will this cube contain (including the base and apex cuboids)?