



NATIONAL TECHNICAL UNIVERSITY OF ATHENS

SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

Data Science and Machine Learning Master's Programme

Εξόρυξη Γνώσης από Δεδομένα (Data-Driven Knowledge Extraction)

Fall 2023

Instructors: D. Tsoumakos, G. Alexandridis

Assignment #1

Question 1 – SQL:

- A. Construct an E-R diagram for a hospital: One or more doctors are associated to a patient during his/her stay at the hospital. Associate with each patient a log of the various medical tests/exams conducted to him/her by one or more of his doctors.

You are free to add attributes to each entity / relationship you define to include important pieces of information (primary keys, etc). Shortly explain any decisions you made (20%).

- B. Assume the following relational database schema on employees:

employee (employee-name, street, city)

works (employee-name, company-name, salary)

company (company-name, city)

manages (employee-name, manager-name)

where the primary keys are underlined. Give an expression **both** in Relational Algebra **and** SQL **for each** of the following queries:

- Find the names, street address, and cities of residence of all employees who work for "NT University" company and earn more than 22,000 per year. (20%)
- Find the names of all employees who live in the same city as the company for which they work. (20%)

Question 2 – Data Warehouses/OLAP:

A data warehouse consists of three dimensions: time, doctor, and patient, and one measure: charge (charge is the fee that a doctor received from the patient for the visit).

- Draw a star schema diagram for the above data warehouse. Assume and add useful attributes for each dimension table besides the necessary ones. (20%)
- 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 for 2022"? (20%)

Deliverable:

- This is an individual assignment.
- Your solutions must be uploaded to the helios class page in pdf format by the deadline. No late submissions or different file formats will be accepted.