Modern Data Management & Business Intelligence
Assignment #1 – Due Date: November 8th, 2022 – Groups of two (2) students

You can use either MySQL or Microsoft's SQL Server – For this assignment we suggest that you use MySQL. For the second assignment you will have to use SQL Server.

Send email to the TA (ssafras@gmail.com) if you are stuck, but first check Google!

Description of the Case:

Ministry of health wants to build a database for the analysis of medical prescriptions. A prescription is described by a unique id, the date and time of the prescription and a quantity (for the drug prescribed). A prescription is ordered by a doctor, who has a unique id, name and specialization. A doctor works in an area. An area has a unique id. a name and a mean income. A prescription is ordered for a patient, who is described by his/her SSN, a name, his/her phone number, date of birth and gender. Finally, a prescription is for a single drug. A drug has a unique id, a name, a description and a price (the price of a single unit).

Deliverables (in <u>one</u> word document):

1. (10%) Use the Entity-Relationship Diagram (ERD) to model entities, relationships, attributes, cardinalities, and all necessary constraints. Use any tool you want to draw the ERD.
2. (10%) Create the relational schema in MySQL/SQLServer and insert a few records into the tables to test your queries below. You will have to hand in the CREATE TABLE statements.
3. (60%) Write SQL code and test it to your data for the following queries

    a. Show the SSN and the name of male patients older than 30 years old that had at least one prescription ordered for them in 2021
    b. Show the SSN of female patients who had prescriptions ordered for them in 2021 for a total amount more than 1000€.
    c. For each area, show the area's id, the area's name and the total amount of drugs prescribed by doctors working in that area
    d. For each drug id and for each month of 2021, show the total amount of prescriptions
    e. Show for each doctor that lives in an area of mean income between 20.000 and 30.000, the doctor's id, doctor's name and his/her total amount of prescriptions.
    f. For each doctor's specialization, show the specialization and the total number of prescriptions ordered by doctors of that specialization in 2021.
    g. For each drug id, show the percentage change in the total amount of prescriptions for that drug in 2021 vs. 2020.
    h. For each drug id and for prescriptions of 2021, show the total amount of the drug's prescriptions ordered for male customers (2nd column), and the total amount of the drug's prescriptions ordered for female customers (2nd column).

4. (20%) Using the programming language of your choice (Python/R), write a program that connects to the database and then prints out a line for each prescription with the following information: the id, date and time of the prescription, the name and phone number of the patient, the name and specialization of the doctor who ordered the prescription, and the name of the drug, its unit price and the quantity of the prescription.