Sistemas de Informação e Bases de Dados
Assignment 2 - Implementing the Database

October 2019

1 The Database Schema

The following relational model is a database schema for a veterinary hospital, inspired in what you modeled in Part 1 of the project.

```
person(VAT,name,address_street,address_city,address_zip)
RI: All persons are either clients, veterinary doctors, or assistants

phone_number(VAT,phone)
VAT: FK(person)

client(VAT)
VAT: FK(person)

veterinary(VAT,specialization,bio)
VAT: FK(person)

assistant(VAT)
VAT: FK(person)

species(name,desc)

generalization_species(name1,name2)
name1: FK(species)
name2: FK(species)

animal(name,VAT,species_name,colour,gender,birth_year,age)
VAT: FK(client)
species_name: FK(species)
RI: age is computed/derived from birth_year
```

```
consult(name,VAT_owner,date_timestamp,s,o,a,p,VAT_client,VAT_vet,weight)
name,VAT_owner: FK(animal)
VAT_client: FK(client)
VAT_vet: FK(veterinary)

participation(name,VAT_owner,date_timestamp,VAT_assistant)
name,VAT_owner,date_timestamp: FK(consult)
VAT_assistant: FK(assistant)

diagnosis_code(code,name)

consult_diagnosis(code,name,VAT_owner,date_timestamp)
name,VAT_owner,date_timestamp: FK(consult)
code: FK(diagnosis_code)

medication(name,lab,dosage)

prescription(code,name,VAT_owner,date_timestamp,name,lab,dosage,regime)
code,name,VAT_owner,date_timestamp: FK(consult_diagnosis)
name,lab,dosage: FK(medication)

indicator(name,reference_value,units,description)

procedure(name,VAT_owner,date_timestamp,num,description)
name,VAT_owner,date_timestamp: FK(consult)
RI: procedure cannot simultaneously be radiography, tests, and/or surgical

performed(name,VAT_owner,date_timestamp,num,VAT_assistant)
name,VAT_owner,date_timestamp,num: FK(procedure)
VAT_assistant: FK(assistant)

radiography(name,VAT_owner,date_timestamp,num,file)
name,VAT_owner,date_timestamp,num: FK(procedure)

test_procedure(name,VAT_owner,date_timestamp,num,type)
name,VAT_owner,date_timestamp,num: FK(procedure)
RI: type should be either "blood" or "urine"
RI: all tests are required to produce at least one indicator

produced_indicator(name,VAT_owner,date_timestamp,num,indicator_name,value)
name,VAT_owner,date,num: FK(test_procedure)
indicator_name: FK(indicator)
```

Part 2 of the project concerns with implementing a database with this schema, and designing SQL queries for answering some relevant information needs with basis on the stored data.

2 Expected Results

In this assignment, you are expected to provide the following results:

- For the relational model above, write the SQL instructions to create the database in our database server (i.e., MySQL on db.ist.utl.pt). You should choose the most appropriate SQL data types for each column.
- Write an SQL script to populate the tables with meaningful records of your choice, that you will design to ensure that we can validate the answers to the next questions.
- Write SQL queries for each of the following information needs:
    1. List the name, owner name, species, and age, for all animals that participated in consults with a veterinary doctor named *John Smith*.
    2. List the name of all indicators measured in *milligrams* and with a reference value above 100. The names should be presented together with the corresponding reference value, and sorted according to the reference value, in descending order.
    3. List the name, owner name, species and age for all animals with the most recent weight above 30 kilograms, and where the *objective* part of any SOAP note, associated to consults of that animal, mentions terms like *obesity* or *obese*.
    4. List the name, VAT and address of all clients of the hospital that are not owners of animals.
    5. For each possible diagnosis, list the number of distinct medication names that have been prescribed to treat that condition. Sort the results according to the number of distinct medication names, in ascending order.
    6. Present the average number of assistants, procedures, diagnostic codes, and prescriptions involved in consults from the year of 2017.
    7. For each animal sub-species of *dog*, present the name of the most common disease (i.e., the name associated to the most frequent diagnostic code for consults involving animals of that species).
    8. List the names of individuals that are simultaneously clients of the hospital (i.e., owners of animals or clients that have brought animals to consults) and employees of the hospital (i.e., veterinary doctors or assistants).
    9. List the names and addresses of clients that only own *birds* as their pets (i.e., the clients for whom all the owned animals contain the word *bird* as part of the species name).
- Suggest database indexes that could be used to improve the performance of the first two queries, from the list of information needs. Justify your choice, and provide SQL instructions for implementing the indexes.

- Write SQL instructions for each of the following changes in the database:

  1. Change the address of the client named *John Smith*, to a different city and street of your choice.
  2. Change the reference value of all indicators used in *blood tests* and measured in *milligrams*. The new reference values correspond to an increase in 10% from the old reference values.
  3. Delete the customer named *John Smith* from the database, removing also all the animals and all the consults (including the associated procedures, diagnosis and prescriptions) in which he was involved.
  4. Find the diagnostic code corresponding to *kidney failure*. Create also a new diagnosis code corresponding to *end-stage renal disease*. Change the diagnosis from *kidney failure* to *end-stage renal disease* for all diagnosed animals where a blood test shows a value above 1.0 for the indicator named *creatinine level*.

- Create views over the tables in the database model, corresponding to the following relational schema.

```
dim_date(date_timestamp,day,month,year)
RI: date_timestamp corresponds to a date existing in consults

dim_animal(animal_name,animal_vat,species,age)
animal_name,animal_vat: FK(animal)

facts_concults(name,vat,timestamp,num_procedures,num_medications)
name,vat: FK(dim_animal)
timestamp: FK(dim_date)
```

  Present the SQL code for creating each of the views corresponding to the tables in the previous model, so that the views list the corresponding records in the database (i.e., information on all the animals that had consults, together with the associated number of procedures and number of prescribed medications).

3 Submission Notes

A report for the 2nd assignment should be submitted to Fénix as a single PDF file, readable with a standard program such as Adobe Reader. The report should have one separate section for each of the aforementioned expected results.

The document cover page should mention the names, student numbers, and group number of its authors. Provide notes explaining the rationale behind non-trivial implementation decisions.