Databases
Project 2

Transportation network analysis.

You can use any airline company in the world.
The kinds of questions are generic and can be applied to any airline company with more than one hub (a hub is an airport where a large number of flights from the TAME group depart and arrive)

Example: You have been hired by the SAS group as a database designer to design a database that will help the board to know statistics about:

- Each TAME group flight:
  - the start and end airports?
  - the difference between the expected departure time and the real departure time?
  - the difference between the expected arrival time and the real arrival time?
  - the average time spent on the ground at the origin and the destination airports: i.e. from departure from the gate till take-off or from landing till complete stop?
  - the average length of the flight for each day of the week and each week in the year?
  - the time each SAS group plane spends on the ground between two consecutive flights for any given day and airport?
  - the number of passengers who have lost their flight transfer (either luggage only or passenger and luggage) after that flight?
  - The total amount of the flight tickets of the passengers who have lost their correspondence after that flight?
- Each TAME group ground employee:
  - which airport is he or she assigned to?
  - how many hours per week does he/she work for each given airport?
  - the number of flight delays he/she has been involved with?
  - the number of formal complaints about that employee?
- Each TAME group hub:
  - how many missed correspondences between flights of the same airline are there for any given day, week, or year?
  - how many missed transfers between flights of airlines of SAS group are there for any given day, week, or year?
  - the total of all flight departure delays?
  - the total of all flight arrival delays?

HASTA EL DOMINGO (ANDRÉS)

HASTA EL LUNES (LENIN)

HASTA EL MARTES (EMIL)

HASTA EL MIERCOLES (MAFER)

- the gates that have accepted all the airlines in the SAS group?
- the gate with the highest total time of flight delays?

You are given the following information on TAME EP's fleet and routes:
https://www.tame.com.ec/index.php/es/inicio/flota-de-tame
https://www.tame.com.ec/index.php/es/destinos/itinerarios.

You will make the following assumptions:

- each flight has one departure airport and one final destination airport through zero, one or several airport stops; however, each flight segment (between a take off and the next landing) has a different flight number;
- each flight is identified by the code of the airline and the flight number;
- each airport has its own three letters code (e.g. UIO is the code of Quito's Mariscal Sucre airport);
- each employee is identified by the personal identifier of the state in which he or she lives together with the name of that state;
- each passenger is identified by his/her passport number (which can be a string of numbers and/or letters) and the name of the state owning the passport.

Please ask any questions to me (consider yourself the database designer and me as the client company staff member).

Your objective is to build the database that will store all the relevant data on the TAME Group flights, employees, airports and gates and that will allow the board of governors to take some decisions based on the statistics mentioned above.

You will need to design the following milestones:

- data dictionary;
- entity relationship diagram;
- transformation of the entity-relationship diagram into a set of relations;
- SQL DDL code,
- normalization of the relations;
- SQL DML code (queries) and its optimization using relational algebra.
- Presentation of the project on the last day of class in the month of July.

The deadlines for the project are indicated in the "Schedule" on D2L.