EXAM
Databases Modeling, DW modeling, SQL Queries

The project work can be done **by groups of 3-4 people**.

Exercise 1 (10)
Design the **conceptual and the logical model of a database** respecting the following requirements:

The database will register information about a library and the books loans. The loans processing system is based on the operational database with the following characteristics. The database must contain information about books characterized by a title, an author (or more authors), the publication year, the editorial company and by the people who borrowed that book. For each author of a book the information of interest is the name and surname, the nationality, the year of birth and the year of death, his or biography and the books authored. The database also must record information about people borrowing books. Each person is uniquely identified by a code and is characterized by a name, surname, the books borrowed. Given a loan of a book is interesting to maintain information about date of the loan and the date of return of the book and a flag indicating that the loan is concluded.

Exercise 2 (6)
Solve the following three queries on the database Sakila.
1. List of actors of English films
2. List of films with the title containing the letter "f" and released after 1958 or with a title starting with "T".
3. List of actors worked only in English film
4. List of actors surname and name such that the actors played only in movies of the category "Family"
5. List for each film with more that 10 rentals, the total rental amount and the average of the rental duration

Deliver a script file with text of query and SQL query

Exercise 3 (10)
Given the order processing system based on the operational database in Fig. 1. There is a general feeling among the directors and senior managers that the company is losing its market share. Within the past three months, two more companies have been formed and their presence in the market is already being felt. Also, recently, more customers than usual appear to be leaving the company and new customers are being attracted in fewer numbers than before.

Examples of business questions the directors of the company wish to ask of their data, in order to reach decisions about their future strategy, are:

1. Which bottle wines are increasing in popularity and which are decreasing, over the year 2014?
2. Which bottle wines are seasonal over the last 3 years?
3. Which Italian customers place the same bottle wine orders on regular basis over the last 3 years?
4. Are some bottle wines vintage more popular in different Italian customer regions?

5. Do Italian customers from different regions tend to purchase a particular class of wines over the year 2014?

Note that the business questions must usually be "**interpreted**" before they can be translated into an SQL query. For instance, the business question "**Do Italian customers from different regions tend to purchase a particular class of wines over the year 2014?**" can be interpreted as "**For the year 2014 and Italian customers, the total number of bottles ordered at more than twice the quantity as the average for all classes of wine, by customer region, by wine class name. The result is sorted by customer region and wine class name**".

With respect to the above business scenario, answer the following questions:

1. Rewrite all business questions in the following form: "For a data subset to use, the metrics to compute, by ..., ..., by .... The result is sorted by ...".

2. Design a conceptual schema for the data mart to support the business questions. Your schema should at least be able to satisfy the above mentioned analysis requirements. You may motivate other suitable attributes for the dimensions. Clearly specify the fact granularity. (**Deliver the schema of the conceptual design**)

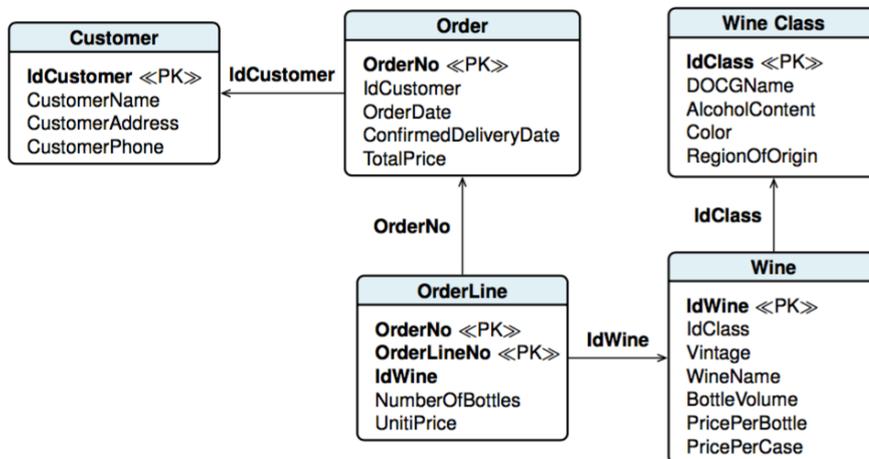3. Design the logical data mart design (**Deliver the schema of the logical design**)

Figure 1: Relational Database Schema

Exercise 4 (Analytic SQL) (6)
Answer the following business questions using SQL over the Foodmart database:

- Return a table where you list product family, store city and year, with the corresponding total sales and profit, and the percentage of profit over the total sales of the year.
- For every product family and customer country and gender return the customer with the highest store sales of the product family. If there are two or more customers with the same store sales, pick up any.
- Produce a report with information about the
  o profit and ratioP of each product family, store country, year and gender; the ratioP is between profit and the store sales in the gender category (Male or Female)
  o the profit of each subset of the above combination

Instructions for the delivery of the material of each exercise:

- **Ex.1:** you need to deliver
  1. a picture of the conceptual model (file name containing the students' names ex: **RossiVerdi-Ex1Conceprual.png**)
  2. a picture of the logical model (file name ex: **RossiVerdi-Ex1Logical.png**)
- **Ex.2:** you need to deliver: **file with text of query and SQL query** (file name ex: **RossiVerdi-Ex2.sql**)
- **Ex.3:** you need to deliver
  3. a picture of the conceptual model (file name: students names ex: **RossiVerdi-Ex3Conceprual.png**)
  4. a picture of the logical model (file name ex: **RossiVerdi-Ex3Logical.png**)
- **Ex.4:** you need to deliver: **file with text of query and SQL query** (file name ex: **RossiVerdi-Ex4.sql**)

All these files must be inserted in a folder with the name equal to the list of students' names and the folder must be zipped. At the end you need to get the zip file: **RossiVerdi.zip**

Send the zip file to the email: anna.monreale@unipi.it. The subject of the email must be**: [Exam Main2020]: names of students.** So for example: **[Exam Main2020]: Rossi, Verdi**