

## First mid-term/Written exam 14/4/2015

**First mid-term: deliver solutions to ex. 1, 2, 3 within 3 h**

**Written exam: deliver solutions to ex. 1, 2, 3, 4, 5 within 4 h**

**Notice:** use your own SQL Server credentials (the lbi account is disabled)

**Exercise 1 (8 pts).** Consider the `foodmart` database. The *mean purchase per working day (MPD)* of a customer in a month of a given year is the total sales to that customer in that month divided by the number of non-Sundays days in that month. Write a Java program `MPD.java` which outputs, for every customer, month and year in which at least a purchase was done by the customer: the customer full name, the year, the month, the total sales to the customer, the number of non-Sundays days in that month, the MPD. The Java program can submit only SQL queries of the form “SELECT \* FROM table”.

**What to deliver:** `MPD.java`, `myJDBCdef.props` (with only the parameters needed for a test of the program).

**Exercise 2 (6 pts).** Develop a SSIS package that outputs on a CSV file the result of Ex. 1 ordered descending by MPD. The usage of GROUP BY / WHERE / ORDER BY clauses in SQL queries to perform computation at server side is not permitted. All the work must be done by the SSIS package.

**What to deliver:** SSDT solution.

**Exercise 3 (2 pts).** Write a single SQL query that solves the problem of Ex. 2 in a way as close as possible to the SSIS package you developed.

**What to deliver:** text file with SQL query and a brief comment about.

**Exercise 4 (8 pts).** Write a MDX query to answer the following question on the Sales cube of the `ruggieri_foodmart` OLAP database:

- for each store, the name and the total sales of the top-profit customers that sum up to 1000 dollars of profit.

**What to deliver:** MDX query and a brief comment about it, a PowerPoint file with the screenshot of the MDX query result.

**Exercise 5 (8 pts).** Describe how to compute frequent itemsets using SQL, possibly with CUBE or ROLLUP extensions. Show an example on the `census` table. Compare your approach with using Weka.

**What to deliver:** a PowerPoint file with the answer to the question and with screenshots of SQL Management Studio and Weka.

**How to deliver:** send an e-mail with a single <your surname>.zip file attached to `ruggieri@di.unipi.it`, including your name, surname, student ID, and computer IP address (<http://www.whatismyip.com>).

**Results and oral exam.** Results will be published on-line by this week. Oral exam dates (for the “Appello straordinario”) will be emailed to you.