

First mid-term/Written exam 4/11/2015

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

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

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

Exercise 1 (8 pts). Consider the `sales_fact` table of the `foodmart` database. The *deviation sales* in a day *time_id* is:

$$dev_sales(time_id) = sales(time_id) - avg_sales$$

where $sales(time_id)$ is the total sales of day *time_id*, and avg_sales is the average total sales of any day. We are interested in finding the interval $[time_id_1, time_id_2]$ such that the sum of deviations sales for the days in the interval:

$$\sum_{time_id \in [time_id_1, time_id_2]} dev_sales(time_id)$$

is maximum. Write a Java program `Deviation.java` which outputs such an interval and the sum of deviation sales. The Java program can submit only SQL queries of the form “SELECT * FROM table”.

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

Exercise 2 (8 pts). Develop a SSIS package that outputs on a CSV file the result of Ex. 1. 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 (8 pts). Write a SQL query with analytic functions or, at your choice, a MDX query that solves the problem of Ex. 1.

What to deliver: text file with SQL/MDX query and with a brief comment about, a screenshot of SQL Management Studio with query result.

Exercise 4 (8 pts). Consider the problem of sending an offer to customers for buying products from the *Seafood* department. Model the problem as a classification problem. Use SQL plus Weka Explorer, or Weka Knowledge Flow or Weka API at your choice for experimenting your solution.

What to deliver: screenshots of SQL Management Studio plus either a Weka knowledge flow `.kfm1` file or a PowerPoint file with screenshots of Weka explorer or a Java program with Weka API calls, and a description of the steps of the designed solution.

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.