Business Intelligence Lab

Written exam 11/9/2013

Deliver exercises within 4 h from start time

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

Exercise 1 (8 pts). For a given customer c, the day of the week with the highest sales to c is called the *most valuable* day of c. Write a Java program ValuableDay.java using JDBC that reads sales from the `sales_fact_1998` table from the *foodmart* database, and then it outputs the pairs c, d where c is a `customer_id` and d is its most valuable day (taken from the `the_day` attribute from `time_by_day`). The usage of the GROUP BY clause in SQL queries to perform computation at server side is not permitted.

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

Exercise 2 (8 pts). Develop a SSIS package reading `sales_fact_1998` from the *foodmart* database, with the purpose of writing on a text file all pairs c, d where c is a `customer_id` and d is its most valuable day (taken from the `the_day` attribute from `time_by_day`). The usage of the GROUP BY clause 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: BIDS/SSDT solution.

Exercise 3 (8 pts). Write a MDX query to answer the following question on the Sales cube of the `ruggieri_foodmart` OLAP database:

- for each customer city, the name and the total sales of the customer with the highest sales in that city.

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

Exercise 4 (8 pts). Using association rules in Weka, find (at least 3) pairs A, B of itemsets such that

$$1.5 \cdot supp(A) \leq conf(B \rightarrow A) \leq conf(A \rightarrow B)$$

on the dataset provided by the teacher. Justify your approach.

What to deliver: a PowerPoint file with screenshots of Weka explorer, description of the steps of the analysis.

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

Results and oral exam. Results will be published on-line by tomorrow morning. Oral exams will start tomorrow afternoon at teacher office.