Business Intelligence Lab

# Second mid-term 29/5/2014

**Deliver solutions to ex. 1, 2, 3 within 3 h**

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

**Exercise 1 (5 pts).** Write a MDX query to answer the following question on the Sales cube of the `ruggieri_foodmart` OLAP database:

- For every age-range, the percentage of store sales of top 10 customers in that age-range over the total sales to customers of that age range.

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

**Exercise 2 (3 pts).** Answer the question of Ex. 1 using an SQL query over the `foodmart` datawarehouse. The SQL query may use analytic functions.

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

**Exercise 3 (8 pts).** Consider the `census.arff` dataset splitted into 60% training and 40% test. The marketing department targets customers in the age ranges: 0-19, 20-39, 40-59, 60-79. They want to send 500 promotional discounts to customers in the test set having class `>50K`, with a proportional partition of the discounts among age ranges. Using Weka Explorer, or Weka Knowlegde Flow or Weka API at your choice, design a data mining process that supports those marketing objectives.

**What to deliver:** either a Weka knowledge flow `.kfml` file or a PowerPoint file with screenshots of Weka explorer or a Java program with Weka API calls, and a brief 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, 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 tomorrow, including the date and time for those who are admitted to the oral exam. Oral exams will start on Friday 6 June.