

## Written test 1/04/2019

**Deliver Part I (Ex.1 & Ex.2) or Part II (Ex.3 - Ex.5) within 2 h**

**Deliver solutions for all exercises within 4 h**

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

**Exercise 1 (8 pts).** Consider the database `foodmart`. For a given month and store, the `AverageGenderFoodRatio` is the ratio between the average sales of female customers and the average sales of male customers considering only sales of products belonging to the "Food" product family. Develop a python program `AvgFood.py` that computes the `AverageGenderFoodRatio` for every store and every month of every year and that produces a .csv file with the results. The python program can submit only SQL queries of the form "SELECT \* FROM table". The usage of PANDAS library is not permitted.

**What to deliver:** `AvgFood.py` and CSV file.

**Exercise 2 (8 pts).** Develop a SSIS package solving Exercise 1. No SQL query on data sources is allowed.

**What to deliver:** SSDT solution.

**Exercise 3 (8 pts). MDX DA FARE**

**What to deliver:** (1) Power Point file with the MDX queries and results and with a brief comment about them; (2) text file with the MDX queries.

**Exercise 4 (2 pts).** Answer the business question of Exercise 1 with **SQL** over the `foodmart` datawarehouse. Use analytic functions as needed.

**What to deliver:** (1) Power Point file with SQL queries and results and with a brief comment about them; (2) text file with SQL queries.

**Exercise 5 (6 pts).** Let  $A$  be the `AverageGenderFoodRatio` described in Exercise 1. Design a data mining approach predicting the value of  $A$  for a store, year and month number given only information available at the end of previous month.

**What to deliver:** screenshots of SQL Management Studio plus either a Weka knowledge flow .kfml file or a PowerPoint file with screenshots of Weka explorer (or Azure ML workflow and all the python scripts used) 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 **SUBJECT:LDS-Apr** with a single <your surname>.zip file attached to **annam@di.unipi.it** including your name, surname, student ID, and computer IP address (<http://www.whatismyip.com>).

**Results and oral exam.** Results will be emailed to the students shortly, including the date and time for those who are admitted to the oral exam.