Group Projects Introduction

Aalto University

Databases for Data Science 2022

26 April, 2022

Project topics

- There are two topics available for the group project:
  1. dairy farm
  2. vaccine distribution.

- The formal descriptions of the project topics can be found on MyCourses.

- Decide with your group members which project you want to work on.

- Please glance through the topics before your exercise session.

Requirements: UML modelling

- You are first required to create an UML model for the database.

- Use the notation taught in the course, notational errors will affect your grade on the project.

- It is recommended to start working on the UML diagram already this week.

Requirements: Relational model

- Turn your UML model into a relational schema.

- Use the techniques taught in the course.

- Clearly list all the relations and underline key attributes.

- The relational model needs to be based on your UML model.

- It is recommended to do this task on Week 3 after the lecture on the topic.

Requirements: Functional dependencies, anomalies and BCNF

- List the non-trivial functional dependencies in your model.

- Analyse your model to find any anomalies or redundancies.

- Inspect whether or not the model is in BCNF.

- If it is not, make the appropriate changes to the relational model. Document the process of turning the model into BCNF. Submit both the original and the decomposed versions.

Documentation

The document should contain:

- The names and student numbers of your group members
- UML diagram for the model
- Relational schema of the converted UML
- Reasoning behind your design choices (recommended length 1-2 pages)
- Answers to the questions about functional dependencies, anomalies and BCNF

Deadlines and grading

- The project has three parts, with the following deadlines
  - **Monday, May 9th at 18:00**, Part 1 (modelling)
  - Monday, May 23th at 18:00, Part 2 (SQL implementation)
  - Monday, June 6th at 18:00, Part 3 (data analysis and corrections)
  - Friday, June 10th at 18:00, Project Deliverable Package

- The platform for group collaboration is Gitlab. **All parts must be submitted on A+.**

- For Part 1 to 3, each is worth 100 points. The Project Deliverable Package is worth 40 points. In total, you will get 340 points from the group project.

- For Parts 2 and 3, we offer 40 bonus points for groups that choose the advanced option of using PostgreSQL for defining and querying the database. More information about PostgreSQL can be found under the Platform and Tools section.