

# Course Content and Goals

5DV119 — Introduction to Database Management

Umeå University

Department of Computing Science

Stephen J. Hegner

`hegner@cs.umu.se`

`http://www.cs.umu.se/~hegner`

# Target Audience of the Course

- ▶ The course is oriented towards the needs of students in technical disciplines.
- ▶ It is not necessary to be a computer science major, but you should be in a program in the Faculty of Science and Technology (or pursuing other studies with a technical or formal orientation).
- ▶ Some background in programming (at least two courses), a knowledge of working within the Unix/Linux environment, and significant mathematical sophistication is expected.
- ▶ It is not oriented towards students in the humanities or social sciences.

## Target Audience of the Course — Levels

- ▶ The course has at least three distinct groups of participants:

Program	Year	CS background
Master in Technical CS (C-programmet)	3	11 courses
Bachelor in CS (Kandidatprogrammet)	2	8 courses
Other (Industrial Economics, Bioinformatics)	3-4	2 courses

- ▶ It is very difficult to design a course which is suitable for all groups.
- ▶ The course must be accessible to the students with the least background in CS.
- ▶ Students in Computer Science (particularly in the C-program) must understand that the course may not seem advanced enough or move quickly enough for their level and needs.

# Major Topics of the Course

- ▶ An introduction to the *relational model* of data, which is the one most widely used in practice (by far).
- ▶ A thorough introduction to *SQL*, the standard *query language* for the relational model, which provides basic access to databases.
- ▶ An introduction to *ODBC*, a means of using SQL from within a traditional programming language (*C* and *Python* in this course).
- ▶ An introduction to the *relational algebra* and *relational calculus*, formal query languages for the relational model.
- ▶ An introduction to the process of *normalization*, which helps to ensure that a relational schema is designed properly, avoiding so-called *anomalies* which can lead to inconsistent data.
- ▶ An introduction to the design of relational schemata using the ER (*entity-relationship*) approach.
- ▶ An introduction to security issues surrounding database applications.

# Goals of the Course

- ▶ A major goal is to provide a thorough introduction to the *use* of relational database-management systems (DBMSs).
- ▶ This includes the design and optimization of relational schemata for a given application.
- ▶ The course does not cover systems aspects of DBMSs, such as internal data organization, index structures, storage management, transaction management, recovery from failure, and query optimization.
- ▶ Those topics are covered in the followup course 5DV120, to be given during April-May.

# Systems to be Used

- ▶ Two popular *open-source* database-management systems will be used, *PostgreSQL* and *MySQL*.
- ▶ These are available on many platforms, but the final versions of all submitted work must run on the Linux installations of the department.
- ▶ Thus, while students are free to use whichever operating-systems they desire for development, the final submission must work under a specific Linux installation.
- ▶ A working knowledge of the GNU/Linux environment is thus essential.

# Important — Register for the Course ASAP

- ▶ In order to use *PostgreSQL* and *MySQL*, you must have a database account for each system.
- ▶ These database accounts will be generated in a few days, using the lists of accepted and registered students.
- ▶ Accepted students who are not registered *may* receive database accounts on the departmental systems, but students who are not at least accepted will not.
- ▶ These database accounts will be needed even for the first obligatory exercise.
- ▶ Register as soon as possible!