

## Advanced Databases (E018441)

**Course size** *(nominal values; actual values may depend on programme)*

**Credits** 6.0

**Study time** 180 h

**Course offerings in academic year 2023-2024**

**Lecturers in academic year 2023-2024**

**Offered in the following programmes in 2023-2024**

**crdts**

**offering**

### Teaching languages

Dutch

### Keywords

data quality, data warehousing, NoSQL, imputation, normalisation, indexing

### Position of the course

This course is a specialization course in which theory and applications on databases are learned. A first part of the course deals with methods for measurement and improvement of data quality. Several algorithms for detection of errors are discussed. Also methods for systematic correction of these errors are discussed. Finally, the part on data quality also deals with database normal forms and their practical use. In a second part, non-relational database models are introduced (NoSQL) and we learn about the concept of data warehousing. Finally, we study elementary storage of data and faster querying of data (indexing).

### Contents

- Data quality: basic measurement
- Edit rules
- Functional dependencies and normal forms
- Repair of systematic and random errors
- Dynamic analysis of data quality
- Data deduplication
- Control digits
- Data warehousing: Concepts and techniques
- Primary and secondary file organization
- NoSQL solutions and distributed databases for management of Big data

### Initial competences

basic principles of (relational) databases

### Final competences

- 1 Understanding basic techniques for measurement of data quality.
- 2 Understanding basic techniques for improvement of data quality.
- 3 Understand the basic principles of physical database storage and indexing.
- 4 Being able to apply datawarehouse technology.
- 5 Knowing how 'Big' data can be managed via NoSQL and distributed database technology.

### Conditions for credit contract

Access to this course unit via a credit contract is determined after successful competences assessment

### Conditions for exam contract

This course unit cannot be taken via an exam contract

### Teaching methods

Seminar, Lecture

### Learning materials and price

- Syllabus (In English). Additional course material is available through the electronic learning environment Ufora
- Basic price: 20 EURO

### References

- R. Elmasri, S.B. Navathe, Fundamentals of Database Systems, Seventh Edition, Pearson Addison-Wesley, Boston USA, 2016 (ISBN: 9780133971330)
- T. de Waal, J. Pannekoek, S. Scholtus, Handbook of Statistical Data Editing and Imputation, Wiley, 2011 (ISBN: 978-0-470-54280-4)
- G. De Tré, Principes van databases, 2e editie. Pearson Education Benelux, 2017 (ISBN: 978-90-430-3580-4)

### Course content-related study coaching

All exercise courses and the practicum are supported by assistants.

### Assessment moments

end-of-term and continuous assessment

### Examination methods in case of periodic assessment during the first examination period

Written assessment

### Examination methods in case of periodic assessment during the second examination period

Written assessment

### Examination methods in case of permanent assessment

Assignment

### Possibilities of retake in case of permanent assessment

examination during the second examination period is possible in modified form

### Extra information on the examination methods

During examination period: Written open-book exam. If an "on campus" exam is not possible, exams will be organised via MS Teams.

Outside examination period: design exercise with report/design to be handed in half-way the course.

### Calculation of the examination mark

A weighted average of both evaluations is applied with the following weight:

- Periodic evaluation: 15/20

- Non-periodic evaluation: 5/20

Students can only pass this course if they obtain a minimum score of 10/20 for **both** parts of the evaluation. If students obtain less than 10/20 for at least one of the parts, the following rules apply:

- If one obtains an 8/20 or 9/20 for at least one part of the evaluation, one cannot pass the whole of the course. If the final score computed with the weighted average is 10/20 or higher, this will be reduced to the highest unsuccessful mark, which is 9/20.
- If one obtains less than 8/20 for at least one part of the evaluation, one cannot pass the course. If the final score computed with the weighted average is 8/20 or higher, this will be reduced to the highest non-deliverable mark, which is 7/20.

Partial exemption for the non-periodic evaluation is possible.

### Facilities for Working Students

- The non-periodic evaluation can be done at home and individually. Students have a server at their disposal.
- Refresher exercises and some SQL functionality needed during the non-periodic evaluation can be trained via the Qexr platform
- Short video recordings that summarize the course content are made available via Ufora.

