GHENT UNIVERSITY

Course Specifications

Valid as from the academic year 2020-2021

Databases (C003803)

Due to Covid 19, the education and evaluation methods may vary from the information displayed in the schedules and course details. Any changes will be communicated on Ufora.

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

Credits 5.0 **Study time** 150 h **Contact hrs** 45.0 h

Course offerings and teaching methods in academic year 2020-2021

A (semester 2)	English	Gent	guided self-study	12.5 h
			online lecture	15.0 h
			online seminar: practical PC room classes	17.5 h

Lecturers in academic year 2020-2021

Bronselaer, Antoon	TW07	lecturer-in-charge

Offered in the following programmes in 2020-2021	**crdts**	**offering**
Master of Science in Statistical Data Analysis	5	A

Teaching languages

English

Keywords

Relational databases, SQL, Graph databases, Data warehouses, Data quality

Position of the course

The global objective of this course is to provide students with theoretical knowledge as well as practical usage of database technology. The main emphasis is on usage of existing databases (interpreting and reading schemas, retrieving data in an efficient manner, verifying quality of data...).

Contents

1 Relational databases: the relational model, constraints, relational algebra, the SQL query language, use of indices and query optimization
2 NoSQL databases: Document stores, Key-Value stores, Graph databases, property graph models, the Cypher query language
3 Data warehousing: dimensional models, ETL processes
4 Data Quality: measurement of data quality, edit rules in the Fellegi-Holt framework, the error localization problem, minimal set covers, branch-and-bound solutions

Initial competences

Basic knowledge of programming

Final competences

1 Understanding the relational model for databases and being able to use a relational database
2 Understanding the basics of NoSQL databases in general and being able to use a graph database
3 Understanding the basics of data warehouse models as an analytical tool
4 Understanding the basics of data quality measurement and being able to apply edit rules in practice

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

Guided self-study, online lecture, online seminar: practical PC room classes

Extra information on the teaching methods

SQL exercises will be done via the Qexr platform, which allows self-training of the SQL query language.
Important: because of COVID19, different didactical approaches can be used if this turns out to be necessary.

Learning materials and price

- Slides
- (Scientific) articles
- E-books
- Short videos

Estimated price: 20 euro

References

- S. Abiteboul, R. Hull, V. Vianu, Foundations of databases, Addison Wesley, 1995
- T. De Waal, J. Pannekoek, S. Scholtus, Handbook of Statistical Data Editing and Imputation, Wiley, 2011
- R. Kimball, M. Ross, The Data Warehouse Toolkit (3rd edition), 2013

Course content-related study coaching

Evaluation methods

end-of-term evaluation and continuous assessment

Examination methods in case of periodic evaluation during the first examination period

Written examination

Examination methods in case of periodic evaluation during the second examination period

Written examination

Examination methods in case of permanent evaluation

Report

Possibilities of retake in case of permanent evaluation

examination during the second examination period is possible

Calculation of the examination mark

The non-periodic evaluation will account for 70% of the total score, the periodic evaluation will account for 30% of the total score.
To pass, a student should pass both the periodic and non-periodic evaluation. In case a student fails one of both parts, the above mentioned weights are no longer valid. The total score will then be calculated as the minimum of the scores of both evaluations.