

# Course Specifications

Valid as from the academic year 2018-2019

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 2018-2019

A (semester 1)	English	group work	5.0 h
		lecture	5.0 h
		seminar: practical PC room classes	37.5 h

Lecturers in academic year 2018-2019

Van den Poel, Dirk	EB23	lecturer-in-charge
--------------------	------	--------------------

Offered in the following programmes in 2018-2019

<a href="#">Master of Science in Statistical Data Analysis</a>	crdts	offering
	5	A

Teaching languages

English

Keywords

Information Systems, data Warehousing, SQL, Big Data, Apache Spark, Spark SQL, Python, Machine Learning, Apache MLlib, ML pipelines

Position of the course

The global objective of this course is to provide students with thorough theoretical as well as practical knowledge on the use and management of information. This knowledge can be of a strategic, a technical-analytical, as well as an organizational nature.

Contents

- 1 Importance of information management in general: which developments are at the basis of the increased importance of information use?
- 2 Data sources and data collection methods: which data sources are available to today's/tomorrow's data administrator, what is big data, how to deal with automatized data collection methods such as scanning and internet? How to handle the nosql evolutions?
- 3 Building a database: which principles are at the basis of building a good database? How to build the structure (Entity Relationship Diagrams)?
- 4 Querying databases: SQL (structured query language) programming language (in casu: Oracle SQL and Hive/Presto) with exercises on large existing information systems.
- 5 Implementation/integration of MIS in the organization: which traps are related to the process of implementing a MIS in the organization, what are the principles of data warehousing?

Each of these topics will be treated in-depth based upon a mixture of interactive class discussions, real-life cases.

Initial competences

Basic knowledge of programming.

Final competences

- 1 Have knowledge of concepts of the management of information.
- 2 Use effectively internal and external data sources.
- 3 Analyze the structure of databases.
- 4 Build applications to support queries in an efficient way.

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

Group work, lecture, seminar: practical PC room classes

#### Learning materials and price

- Oracle (TM) SQL en PL/SQL cursusmateriaal waaronder
- Oracle Database 10g: Introduction to SQL
- (wetenschappelijke) artikels
- Slides omtrent Apache Spark/Spark SQL/Python & Big Data
- Cases

Geraamde totaalprijs: 40 euro

#### References

#### Course content-related study coaching

#### Evaluation methods

continuous assessment

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

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

Examination methods in case of permanent evaluation

Assignment

Possibilities of retake in case of permanent evaluation

examination during the second examination period is possible

#### Calculation of the examination mark

30% on SQL exercises/exam and 70% on Big Data/Apache Spark/Spark SQL/Python assignment

To pass, a student should pass both parts of the evaluation. A student does not pass for both parts and the score is 10/20 or more, the score will be reduced to 9/20.