Marketing Information Systems - Database Marketing (F000880)

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

Credits 8.0   **Study time** 240 h

Course offerings and teaching methods in academic year 2025-2026

A (semester 1)	English	Gent	independent work
			group work
			seminar
			lecture

Lecturers in academic year 2025-2026

Offered in the following programmes in 2025-2026	**crdts**	**offering**

Teaching languages

English

Keywords

Information Systems, Data Warehousing, SQL, Big Data, Apache Spark/Spark SQL/Python, Machine Learning, Spark 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. "New data, new methods, and new skills - how to bring it all together?" is one of the research priorities 2016-2018 of the MSI (Marketing Science Institute) - THE leading marketing think tank worldwide.

Contents

1 Importance of information management in general: which developments are at the basis of the increased importance of information use in marketing, which businesses or functions are marketing information intensive?

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 automized data collection methods such as scanning and internet? How to handle the nosql evolutions?

3 Building a marketing database: which principles are at the basis of building a good marketing database? How to build the structure (Entity Relationship Diagrams)?

4 Querying marketing databases: SQL (structured query language) programming language (in casu: Oracle SQL and Hive/Apache Spark/Spark SQL/Python) with exercises on large existing marketing information systems.

5 Feature Engineering: constructing variables to be used different (predictive and prescriptive) models.

6 Machine learning using Spark ML pipelines.

7 Data Visualization is an important component in convincing management. PySpark supports nice visualization capabilities using e.g. Matplotlib. These will be demonstrated and used extensively.

8 Implementation/integration of MIS in the organisation: which traps are related to the process of implementing a MIS in the organisation, what are the principles of datawarehousing? Each of these topics will be treated in-depth based upon a mixture of interactive class discussions, real-life cases.

Initial competences

Introduction to Informatics. Basic knowledge of the CRISP-DM (data mining) model.

Final competences

1 Understanding the structure and set-up of a database.
2 Mastering the programming language SQL (and NoSQL Big Data) to achieve optimal data and information management.
3 Building applications based on these data structures in order to make adequate conclusions for complex marketing problems.
4 Providing correct communication towards both technical and non-technical professionals.
5 Using a variety of external data sources (including new forms such as social media data) in an optimal way.
6 Applying Big Data algorithms through Spark ML Pipelines on case studies in order to build complex marketing decision models.
7 Validating own research results with existing literature of international top journals.
8 Executing a company case study in an international and interdisciplinary team that includes different levels of experience.
9 Presenting professionally about an advanced problem and its solution.

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, Seminar, Lecture, Independent work

Extra information on the teaching methods

Ex cathedra or online sessions as well as active class discussions of the different techniques and models with interactive exercises in the PC room.

Study material

None

References

None

Course content-related study coaching

Numerous exercises are being solved during sessions. In addition, assignments (to be solved in teams) are handed out. Students will receive coaching in the process of solving the assignments and feedback afterwards (collectively, by team and individually). After tests about the programming language SQL students will receive collective feedback.

Assessment moments

continuous assessment

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

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

Examination methods in case of permanent assessment

Oral assessment, Skills test, Written assessment with open-ended questions, Peer and/or self 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

Written examination to investigate the extent to which students mastered the use of entitiy relationship diagrams, SQL query and NoSQL language, and the concept of Big Data.
Technical as well as managerial presentation of results of a Big Data assignment using Apache Spark (including Spark SQL, Spark ML Pipelines)

Calculation of the examination mark

30% on SQL exercises/exam and 70% on Big Data/Apache Spark/Spark SQL/Python group assignment potentially adjusted by peer assessment. Of the 70% group assignment score 30% is based on individual questions and 40% is based on group performance.
To pass, a student should pass both parts of the evaluation. If a student does not pass for both parts and the score is 10/20 or more, the score will be reduced to 8/20.
Teams will be randomly selected during class to present their solutions to assignments. A

maximum of two bonus points (of 20) can be earned in this way.