

Marketing Information Systems - Database Marketing (F000880)

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

Credits	8.0	Study time	240 h	Contact hrs	80.0 h
---------	-----	------------	-------	-------------	--------

Course offerings and teaching methods in academic year 2018-2019

A (semester 1)	English	demonstration	2.5 h
		seminar: practical PC room classes	60.0 h
		group work	5.0 h
		lecture	7.5 h
		self-reliant study activities	5.0 h

Lecturers in academic year 2018-2019

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

Offered in the following programmes in 2018-2019

Master of Science in Marketing Analysis	crdts	offering
	8	A

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 automated 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

Demonstration, group work, lecture, self-reliant study activities, seminar: practical PC room classes

Extra information on the teaching methods

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

Learning materials and price

Oracle(TM) SQL course material including:

- Oracle Database 10g: Introduction to SQL
- (Scientific) papers
- Slides on Apache Spark/Spark SQL/Python and Big Data/Big Data Analytics in Spark ML Pipelines
- Cases

References

Van den Poel D. (2003), "Predicting Mail-Order Repeat Buying: Which Variables Matter?", Tijdschrift voor Economie en Management, 48 (3), p.371-404.

Ryza S., Laserson U. (2015), Advanced Analytics with Spark: Patterns for Learning from Data at Scale", O'Reilly.

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 and about Big Data & Hive/Apache Spark/Spark SQL/Python, students will receive collective as well as individual feedback & 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

Written examination with open questions, oral examination, assignment, skills test, peer assessment

Possibilities of retake in case of permanent evaluation

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 entity 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 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.