Maastricht University Leading in Learning!

COURSE BOOK

Data Analytics

SCI3051

Academic Year: 2017-2018

Course Period: 5

Bachelor year 3

Table of Contents

Introduction

The course Data Analytics shares many topics with the course Business Intelligence Case Studies that a compulsory course for the track Information Management and Business Intelligence in the master of International Business (IB). It is an elective for all master tracks in the master IB, and is available as elective for exchange students.

The course treats the theory and practice of Business Analytics. Methods for the analysis of data are presented, from current data mining toolboxes. We study how (and how not) to build predictive models to extract information from large databases and how to interpret the results. The thus discovered knowledge is used for intelligent decision making to make processes run more efficiently and to develop new services for the organizations that provide the data.

The course aims at getting hands-on experience in analysing managerial decision processes, based on available data from real-life cases. The course consists of applying up-to-date data mining techniques on real-life problems. These techniques will be implemented with modern software tools (Tableau and SPSS modeller). This course, we discuss cases from

1. Medtronic (medical technology) &
2. PureMinds (online marketing).

Prerequisites: SCI2033 Data Mining and recommended is SSC2061 Statistics 1.

Student Portal will be used as the primary means of communication between lecturer, tutors and students. The education method is a case based variant of PBL

Contact details:

Course coordinator: M. Peeters, m.peeters@maastrichtuniversity.nl

Lecturer(s): Guestspeakers &

Week 2-8: R. Jorge Almeida, rj.almeida@maastrichtuniversity.nl

Tutor(s): Week 1: S. van Hoesel, s.vanhoesel@maastrichtuniversity.nl

Week 2-8: M. Peeters & A. Abiad Monge, a.abiadmonge@maastrichtuniversity.nl

Learning objectives

Students will work on two real-life cases for which they have to visualize and analyse the data provided. Also, predictive models have to be built. They are supposed to find their own problems and have to find out themselves how to analyse the data to solve their problems. This involves the development of the following skills:

1. Be able to collect and analyse data.
2. Demonstrate up-to-date academic knowledge on data mining methods.
3. Be able to translate a problem from practice into a research question, and to develop a strategy to answer this question
4. Establish a meaningful link between a problem from practice and an academic approach
5. Be able to enrich an academic discussion of a practical problem by insights from related disciplines
6. Be able to propose solutions based on a well-grounded academic reasoning.
7. Learn to work in international tutorial groups in a PBL setting, inspire the discussion, and come to solutions

and the more generic skills

8. Be able to write an extensive report in a clear and well-structured way.
9. Demonstrate a professional attitude
10. Be able to present a complex matter in a precise, understandable way

Course structure

Description of the course

This course treats the theory and practice of Business Analytics. The data preparation process is discussed, and tools for the analysis of data, such as regression, clustering and decision trees. These methods will be applied during the tutorials.

The course consists of applying up-to-date data mining techniques on real-life problems. These techniques will be used with state-of-the-art software. We study how (and how not) to extract information from large databases with standard techniques from data mining to build prescriptive models and to interpret the results found in the models.

The cases are selected from business practices based on current topical developments of the various disciplines involved with data oriented decision making: financial, marketing, supply chain management etcetera. Selected companies will introduce these cases. Some companies involved in previous years are: VISA (London), Proctor & Gamble (Brussels), Xerox (Venray) and Smurfit-Kappa (Roermond).

For this course we have:

1. Medtronic: a multi-national in medical technology, services and solutions.
2. PureMinds: consultancy on online Marketing and web-analytics.

The course has one lecture per week on Mondays, and a group meeting on Thursdays. In the lectures, we provide the required theoretical knowledge and guest lectures in which we introduce the companies and the company-cases. These lectures are mandatory. Tutorials are used to discuss topics and explore the used software. In the first group meeting after a guest lecture, the data from the actual company and possible analysis topics are discussed. The second meeting is generally used to see where problems arise and discuss and solve these problems. In the subsequent meeting one or two groups present their findings (a group consists of three persons).

Literature

- Data Science for Business, What You Need to Know about Data Mining and Data-Analytic Thinking, by Foster Provost and Tom Fawcett, O'Reilly Media 2013. ISBN 978-1-4493-6132-7, EBook ISBN 978-1-4493-6131-0 (not compulsory).
- Other materials, i.e. articles, will be made available through Student Portal.

Assessment

(1) Attendance

Attendance is taken during the tutorial meetings and you have to be present in at least five of the tutorials. If you miss more than one of those meetings unexcused you will have to do an additional case in order to get a sufficient participation grade. Active participation can result in a bonus of 0,5 to your calculated grade.

Next, each group (3 students[1]) has to present one case. Plan your presentation to take about 30 minutes. Provide your tutor with a hand-out of the sheets. Presentations are required to pass the course but are not graded; feedback on them will be provided.

(2) Cases

We expect you to hand in a management report for each case via safe assign on Student Portal (The report should be to the point and all technical details (including the explanation of the approach, i.e. data cleaning/software input-output) are included as an appendix).

(3) Exam

In order to pass the course and set the final grade we discuss the cases with each group separately in an oral exam (note that the oral exam does not give extra points, but is mandatory to pass it in order to pass the course). The final grade is based on individual knowledge on the cases.

[1] Groups will be formed in the first tutorial. Presentations are assigned during tutorial 1.

Grading

To get familiar with the approach as well as with the grading we will use an example case. We discuss this case in detail, the approach and the usage of the software, and students can write a report that will be graded.

Each case will be graded with an integer and has a minimum requirement of 5. The final grade will be a weighted average of the two grades. For grading we use as reference base the structure as shown in Table 1.

If 1 case does not meet the requirements and in the oral exam you can convince the examiners that you are able to do a better job, you can get an additional case (resit). The grade for the additional case will replace the old one.

In order to pass the course you need a final grade of at least 5.5. If you do not succeed to get a sufficient grade and participate in the course in another block, you will have to start from scratch.

Table 1. Grading sheet

Case:	
1. Analysis	Specify what you do for who and why Result Completeness Methods applied correct Interpretation Presentation Conclusion What further analyses are possible? Creating/providing additional insights for other(s)
2. Additional	Use of theory Use of external data
3. Report structure	Executive summary Introduction problem Answer Discussion
4. Appendix	Data explanation & visualization Approach explained
Remarks report	Typos Use of graphs Writing/readability Layout

Fraud and Plagiarism

In order to protect the reputation of the degrees that you – as students – receive, instances of cheating or plagiarism are treated extremely seriously.

Fraud, including plagiarism, is understood as a student's act or failure to act that makes it partially or fully impossible to correctly assess his/her knowledge, insight and skills.

Plagiarism is understood as the presentation of one's own or other people's ideas or words without adequate reference to the source.

Any assignment is an individual piece of work, which means that plagiarism is strictly forbidden. Equally, the use of mobile phones, communication devices or any other information carrier (whether the phone or other device is turned on or off, used or not used, etc. is irrelevant) during an examination is also forbidden.

If the Board of Examiners concludes that anything has occurred in an examination that makes it partially or fully impossible to correctly assess his/her knowledge, insight and skills, they may impose a sanction in accordance with SBE's policy on fraud, including plagiarism.

More information can be found on Student Portal.

Complaints

If you do not agree with the examination or the results therefore, there are several procedures in place you can make use of. Please refer to Student Portal for more information.

Comment

Within five days after the examination date you can submit comments on the content and design of the examination (questions) to the course coordinator.

Appeal

For information regarding an appeal procedure, please read the information on Student portal carefully.

Course schedule, lectures & tutorials

	Lectures (Monday)	Group meetings (Thursday)
week 1		12 April: Business Intelligence
week 2	16 April Introduction course-theory Training Tableau	Introduction course-structure Introduction example case Tableau on example case
week 3	23 April: Introduce case 1 Methodology Data preparation	Tableau on example case SPSS modeller and data preparation Discussion case 1
week 4	30 April: Introduction data modelling Supervised data mining methods	Tutorial SPSS modeller on example case Discussion case 1
week 5	7 May: Introduction Case2 Supervised data mining methods	No Tutorial. Hand-in case 1 (9 May)
week 6	14 May: Unsupervised data mining methods	Presentation case 1 Tutorial SPSS modeller on example case Discussion case 2
week 7		Tutorial SPSS modeller on example case Discussion case 2
Week 8	28 May: Advanced BI	Presentation case 2 Oral exams