

MGT-432

Data science for business

Younge Kenneth

Cursus	Sem.	Type
Energy Management and Sustainability	MA1, MA3	Opt.
Financial engineering	MA1, MA3	Opt.
Management, Technology and Entrepreneurship minor	H	Opt.
Managmt, tech et entr.	MA1, MA3	Opt.

Language of teaching	English
Credits	6
Withdrawal Session	Unauthorized Winter
Semester	Fall
Exam	During the semester
Workload	180h
Weeks	14
Hours	4 weekly
Courses	3 weekly
TP	1 weekly
Number of positions	100

It is not allowed to withdraw from this subject after the registration deadline.

Remark

MA3 only

Summary

Students will learn the basic concepts of Data Science so that they can make better business decisions. Students will also learn how to apply these concepts to real programming problems.

Content

This course introduces students to some of the programming tools used by data scientists to address real world business analytics problems. Accordingly, the course objectives are three fold: (1) to develop an understanding of how Data Science methods can support decision making in business environments; (2) to gain familiarity with how Data Science tools function through experience in addressing real-world problems and programming real-world solutions; (3) to evaluate the strengths and weaknesses of alternative approaches. The course is particularly applicable for students interested in working for, or learning about, data-driven companies.

Keywords

Data science; data analysis; business analytics; python; data-driven management

Learning Prerequisites**Required courses**

All students must have the following prerequisites:

Statistics: Prior to taking this course, all students must complete at least one course in statistics. You should have a basic understanding of descriptive statistics, the OLS linear regression model, and multiple regression.

General Programming: Prior to taking this course, all students must complete at least one course in a general computer programming language.

Python Programming: Prior to taking this course, all students must know (or quickly familiarize themselves with) the syntax and data structures of the **Python 3** programming language. There are numerous online tutorials and short courses for learning Python. We recommend one of the following options:

- The 7-hour Python tutorial at Kaggle: <https://www.kaggle.com/learn/python>. This course provides a fast

overview of Python, but it does assume basic knowledge of computing and programming languages. If this course is too advanced or moves too quickly, then you should select the next option.

- The JetBrains Academy track on Python: <https://hi.hyperskill.org> This is the best approach for an absolute beginner and it comes in interactive, bite-sized exercises. It also helps you track your progression of study with: <https://hyperskill.org/knowledge-map> However, this option may take up to 34 hours to complete.

If you do not know Python, and you cannot complete one of the two tutorials above by the end of the second week of class, then you should delay in taking this course until you have the necessary skills.

Important concepts to start the course

Descriptive statistics
The OLS linear regression model
Multiple regression

Learning Outcomes

By the end of the course, the student must be able to:

- Formulate prediction models
- Assess / Evaluate the performance of prediction models
- Describe their findings to others

Transversal skills

- Access and evaluate appropriate sources of information.
- Take feedback (critique) and respond in an appropriate manner.
- Plan and carry out activities in a way which makes optimal use of available time and other resources.
- Assess one's own level of skill acquisition, and plan their on-going learning goals.
- Assess progress against the plan, and adapt the plan as appropriate.
- Collect data.

Teaching methods

Weekly lectures, demonstrations, assignments, and exercises.

Expected student activities

Attending class regularly to both acquire content and to review problem sets and exercises. Take home, open-book exams will be given on the day of regularly scheduled class

Assessment methods

100 points	Qualifying Exam
500 points	Assignments
200 points	Business Use Case
200 points	Take home Final Exam

Resources

Virtual desktop infrastructure (VDI)

No

Bibliography

Textbook: "Data Science for Business" by Provost & Fawcett. (2013) Publisher: O'Reilly Media; ASIN: B017PNWLKQ

A list of additional readings will be distributed at the beginning of the course.

Ressources en bibliothèque

- [Data Science for Business / Provost](#)