Course Description

This course introduces the fundamental ideas of data science and the data-analytical thought process, and illustrates various ways of applying data science in the business context. Through the above, it facilitates developing an understanding of potential routes for extraction of business information, insight, and value from data.

The course utilizes business-domain examples to illustrate the key ideas and thereby also introduce some of the most commonly used methods in the field. Through these,
students learn how data science methods can be employed to support the business decision-making process.

class 1
- introduction: data-analytic thinking
- from business problems to data science solutions - part I

class 2
- from business problems and data science solutions - part II
- introduction to predictive modeling - from correlation to supervised segmentation - part I

class 3
- introduction to predictive modeling - from correlation to supervised segmentation - part II

class 4
- fitting a model to data

class 5
- issues in modeling: overfitting (variance), underfitting (bias), and their avoidance

class 6
- finding similarities - neighbors and clusters

class 7
- decision-analytic thinking I - what is a good model?

class 8
- visualizing model performance

class 9
- probabilities and evidence - probabilistic modeling approaches

class 10
- extracting information from textual data - representing and mining text

class 11
- decision-analytic thinking II - toward analytic engineering

class 12
- addendum - a brief mention of some data science techniques that weren't discussed in the course

conclusion - self-study
Course Goals

- Develop a solid understanding of the core ideas behind data science and the data-analytical thought process, and acquaint students with some of the most commonly used data-science methods and the process of applying data science tools in the business context.

- Develop a basic understanding of the manner in which data science can be applied for the extraction of business information, insight, and value from data.

- Allow a smooth transition into the more technical courses (in years 2 and 3) that rely on notions presented here.

Grading

Final exam: 70%; home assignments: 30%. One must pass the exam (i.e. obtain a minimum grade of 60) to pass the course.

Most classes have associated home assignments. In total, there are 14 assignments (divided into 12 installments). Timely submission is required in order for an assignment to be accepted. Other than justifiable reasons (e.g. military reserve service), as specified in the Reichman University bylaws, late submissions will be disqualified and graded 0. The home assignment grade is the average (arithmetic mean) of ALL 14 assignments. To be entitled to take the exam, each participant is required to submit a minimum of 12 of the 14 assignments. (Keep in mind that unsubmitted assignments will be graded 0 and taken into account in the calculation of the overall home assignments grade.) Further details can be found on the course moodle page.

NOTE: Class attendance is mandatory. Exemptions will be given only in accordance with Reichman University's bylaws. Students who do not comply will not be allowed to take the exam.

Learning Outcomes

Upon successfully completing the course, students will have acquired a basic understanding of many of the main methods and tools used by data science. In addition, students will have developed an understanding of the types of tasks these tools can be used for within the business realm and the types of business questions they may help address. Students will have also acquired skills to enable them to assess the potential...
benefits of applying data science tools in various business contexts.

Lecturer Office Hours

Officially: Thursdays, 14:30 to 16 (by appointment).

Unofficially: If we're on campus, I'm at my office nearly daily and am happy to meet any time, as long as I don't have prior commitments. If we're off-campus, I can be reached by mail. Zoom meetings will be scheduled for issues that can be better handled that way.

Additional Notes

The classes focus is on addressing selected topics from the reading material and answering related questions. Students are required to read the relevant material for each class in advance (i.e., before the class starts).

Reading List

All reading material is available on the course Moodle page

Students are required to read the relevant material for each class in advance (i.e., before the class starts). The class itself focuses on addressing selected topics from the reading material and answering related questions.

The course is based on Data Science for Business by Foster Provost and Tom Fawcett. The book may serve as a handy reference and for enrichment.