MKT 42370: Advanced Analytics & Big Data
Summer Trimester 2022/23

Time (Room):	Wednesday	10:00-13:00 (C201-GSB)
	Thursday	14:00-17:00 (C201-GSB)
	Friday	10:00-13:00 (C201-GSB)
Office Hours (Room):	Wednesday	13:30-14:30 (D010-GSB)
	Thursday	17:30-18:30 (D010-GSB)
	Friday	13:30-14:30 (D010-GSB)
Instructor:	David DeFranza	
Email:	david.defranza@ucd.ie	

Every day, an increasing variety of new information is created at a larger volume and faster velocity than ever before. This information presents incredible opportunities for businesses but contending with its size and speed poses considerable technical and strategic challenges. Since much of this newly generated data captures the thoughts, opinions, and behaviors of individual consumers, marketers play an important role in addressing these challenges. Indeed, businesses capable of extracting knowledge from large data sets can achieve a considerable competitive advantage, and marketers capable of facilitating this process will find they have an advantage in the job market. In this module, we will engage with business problems, specifically those which might be faced by marketers, using data analytic thinking. Along the way, we will discuss the fundamental principles guiding the extraction of knowledge from information, challenges posed by very large data sets (including "Big Data"), and some of the most common techniques and technologies used to mine such data.

Learning Outcomes

This module aims to prepare you for the challenges and opportunities today's marketers face when dealing with "Big Data" and extracting knowledge from large data sets more generally. At the end of this module, you should be able to:

1. Explain the unique characteristics of and challenges posed by Big Data
2. Summarize the industry standard data mining process within the context of a business problem
3. Identify an appropriate analysis method based on a description of the business problem and available data
4. Assess the performance of a model or analysis based on common diagnostic metrics

5. Develop a comprehensive data analysis proposal which addresses a specific business problem

Required Readings

This course will utilize a textbook, readings from which are required. The book is available from the library and new, used, and electronic editions are available from most book sellers. The book is:

Provost, F., & Fawcett, T. (2013) *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly.

- Publisher link: https://www.oreilly.com/library/view/data-science-for/9781449374273/
- UCD library link: https://tinyurl.com/b873bppp

Assessment

This module has three assessment components, each with a specific weighting and marks totally 100%. There will be no exam. The weighting assigned to and responsibility for (i.e., whether an assessment is to be completed by an individual or group) each component is shown below. Following, each assessment component is discussed in detail.

Assessment Component	Weighting	Individual / Group	Deadline
1. Online Quizzes	40%	Individual	12 June & 26 June
2. Class Participation	10%	Individual	23:59 after each lecture
3. Data Analysis Proposal	50%	Group	31 July

Online Quizzes

Online quizzes will test your knowledge of the course material. For each quiz you will have to answer between 20 and 30 multiple-choice type questions about the content covered in the preceding weeks. A specific time window during which the quizzes are available will be announced prior to the quiz day. All quizzes will be conducted through Brightspace.

Please note that quizzes will only be available for the time specified in the course and that failure to conduct the quiz will automatically result in zero marks. During that window you a limited amount of time to correctly answer as many questions as possible. Please note that you will answer a small subset taken from a large pool of questions, and that the questions will be randomly allocated to each student. This means that students are unlikely to see the same questions and/or questions in the same order. A set of sample questions with solutions will be provided before each quiz. Please note, however, that study guides are not comprehensive: Quizzes may cover any material discussed in class or in the assigned reading.

Class Participation

Students are, of course, encouraged to participate actively during lectures by asking and/or answering questions. Formal assessment of class participation will take place through the submission of regular "minute essays." After each lecture, an assignment will be made available on Brightspace. This assignment will consist of three questions: What was the most important or useful thing you learned in today's class; What question(s) do you have about the material covered in today's class; How might the material we discussed in today's class apply to your role as a digital marketer?

Responses will be due before 23:59 that day. You should be able to answer each question in fewer than 200 words written in only a few minutes. Reponses will be reviewed at the beginning of the following lecture.

Data Analysis Proposal

Working in randomly assigned teams, you will be asked to develop a comprehensive data analysis plan addressing a specific business problem. The problem is detailed in the required case "Predicting Consumer Tastes with Big Data at Gap." In this case, you will learn that the new CEO of Gap has decided to upend the creative process at the clothing brand by replacing traditional creative directors with data analysts. As marketers, it will be your responsibility to predict fashion trends and develop new products using data and analytics. Your team must craft a plan capable of addressing this considerable challenge. Your analysis plan should follow the template provided by Provost and Fawcett (2013; see above) and submitted via Brightspace no later than 31 July, 2023. An opportunity for peer- and self-evaluation will be provided. The assignment will be graded in accordance with a rubric which will be provided in advance. Additional details will be provided in class and via Brightspace.

Assessment Criteria and Grade Descriptors

This module utilizes criterion referencing and UCD grade descriptors. Before attempting the assessments for this module, you are encouraged to review the grade descriptors. A copy of the UCD grade descriptors can be downloaded from:

https://www.ucd.ie/history/t4media/UCD%20Module%20Grade%20Descriptors-1.pdf

Protocol for submitting your assignments: All continuous assessment should be submitted electronically via Brightspace, by the deadline specified. Please do not email assignments directly to the teaching team, unless explicitly directed to do so.

Statement of Inclusion

This module strives to be a model of inclusion. We respect and value student diversity in all of the modules we offer. We aim to provide and promote equitable access and opportunity to all students regardless of disability, race, age, gender, sexuality or socio-economic status. Students are

encouraged to approach staff to discuss their learning needs. Any information disclosed will be treated confidentially.

University Policies

You should ensure you are familiar with the following UCD protocols:

- *Plagiarism and Academic Integrity:* UCD and the College of Business take academic integrity extremely seriously. All work must be your own, be completed specifically for this module and not have been submitted elsewhere. It should also be accompanied by a signed own work statement, such as the following:

  > *I declare that all materials included in this essay/report/project/dissertation is the end result of my own work and that due acknowledgement have been given in the bibliography and references to ALL sources be they printed, electronic or personal.*

  The university's plagiarism and academic integrity policy is available from: https://www.ucd.ie/secca/studentconduct/

- *Harvard Referencing Style:* UCD College of Business uses the Harvard style of referencing. The UCD library has developed some resources on avoiding plagiarism and on how to reference correctly using the Harvard style. These resources are available from: https://libguides.ucd.ie/academicintegrity

- *Assessment Submission Form:* When submitting a piece of assessment, you are asked to attach an assessment submission form. This form is available from: https://www.ucd.ie/t4cms/assessment%20submission%20form.pdf

- *Late Submission of Coursework:* This policy outlines the steps you should take where you know in advance that you will not be in a position to meet a submission deadline and the penalties imposed in such circumstances. See https://www.ucd.ie/t4cms/latesub_po.pdf

- *UCD Extenuating Circumstances policy:* If, during the course of this module, you encounter any serious unforeseen circumstances that are beyond your control and which prevent you from meeting the requirements of the module, you should consult this policy. A student guide to this policy is available from: https://www.ucd.ie/students/studentdesk/extenuatingcircumstances/

- *UCD Student Code:* The UCD Student Code establishes the University's regulations and expectations in respect of student behaviour and conduct. The Student Code is available from: https://www.ucd.ie/secca/studentconduct/

Module Topics

The schedule below outlines the planned themes, required readings, and tasks lecture by lecture. Updates and additions will be notified in class and on Brightspace. Required reading from the textbook, Provost and Fawcett (2013) is abbreviated, such that "PF Ch. 1" indicates "Chapter 1 in Provost and Fawcett, *Data Science for Business*."

Lecture	Date	Topic	Required Reading
1	31 May	Introduction: What is Big Data? What is Data Science?	PF Ch. 1
2	1 June	Managing data science	PF Ch. 2 & Ch. 13
3	2 June	Applying data analytic thinking to a business problem	PF Ch. 1 & Ch. 2
4	7 June	Building predictive models; Intro to tree-based models	PF Ch. 3
5	8 June	Implementing tree-based models	PF Ch. 3
6	9 June	Introduction to linear functions	PF Ch. 4
	12 June	**Online Quiz 1**	
7	14 June	Review and Data Analysis Proposal introduction	
8	15 June	Support vector machines and logistic regression	PF Ch. 4
9	16 June	Distance- and similarity-based models	PF Ch. 6
10	21 June	Evaluating model performance	PF Ch. 7 & Ch. 8
11	22 June	Reducing bias and enhancing model performance	PF Ch. 7 & Ch. 5
12	23 June	Applying data analysis techniques	
	26 June	**Online Quiz 2**	
	31 July	**Data Analysis Proposal Due**	