the UCL MBA

The Data Analytics Advantage
MSIN0182: Syllabus

Module Description & Objectives

Data is increasingly critical across the organisation and across industries. Data is being used across a diverse set of activities, such as improving organisational design and hierarchy, increasing customer acquisition and retention, assessing market opportunities, complementing product offerings, and more. In the short to medium term, organisations will demand leaders are able to apply data and analytics across the organisation, build and deploy data resources and infrastructure, communicate with data teams, and expertly consume and contribute to the production of data analytics and output. The goal of The Data Analytics Advantage is to provide students with this "data mindset."

The aims of the module are as follows:

- Understand the foundations of probability and statistics
- Engage with three principal applications of analytics: descriptive, predictive, and causal
- Understand the fundamentals of data structures and management
- Build experience with and exposure to tools used in data analytics
- Understand the organisational and economic aspects of data and data management
- Consider different data approaches, the output and limitations of these approaches, and how to act on them
- Formulate an argument with appropriate use of data and evidence as support
- Understand various approaches to collecting data, and
- Develop a conception of mental models underlying data analytics including the data generating process, counterfactual thinking, uncertainty and Bayesian updating, and the nature of causal relationships

Method of Assessment

All marks for assessments will be given in Letter Grade Marks (A-F) as described in the UCL Academic Manual unless otherwise expressed in the module syllabus.

	% of Module Mark	Due
Quizzes	10%	Marked over the entire term
Class Interaction	30%	Marked each week
Individual Assignments	30%	Assignment 1 due 24 hours after Week 3 live session Assignment 2 due 24 hours after Week 8 live session
Team Assignment	30%	Due 48 hours after Week 10 live session

Quizzes – 10%: For the first nine weeks of the class, students will be asked to take a short quiz. There are 5 questions each week and quizzes will be graded each week, based on the number of correct answers (from 0 to 5 yields letter marks as follows: 5-A, 4-B, 3-C, 2-D, 1-F, 0-E). These quizzes are all multiple-choice and serve as checks on attention and the comprehension of the asynchronous content. Students are asked to complete the quizzes **no later than 24 hours prior to the start of the live sessions.**

Class Interaction – 30%: This graded component is linked to our module learning objective of enabling you to be able to critically discuss and debate module content with leaders from other cultures and industries. Class interaction is critical to adding context to the material we are focusing on during this module and central to the education aims of this degree programme. As such, students are expected to not only be present for live seminars but be prepared to engage with module leader and students on the weekly module material. This level of engagement requires preparation. There will be activities throughout the module that require you to demonstrate your preparedness before live seminars (prework). These are meant to help you digest the material but also to share your viewpoints with others in the class. There will thus be two components to this marking: your weekly preparation of module materials and your engagement in our weekly live seminars.

Marking will include both prework activities in the asynchronous material as well as live classroom interactions for a total grade mark each week. This mark (from 0 to 5 yields letter marks as follows: 5-A, 4-B, 3-C, 2-D, 1-F, 0-E) is a combination of these two engagement activities. Prework engagement is marked on being complete (2 points) before class with partial marks (1 point) being offered for incomplete responses or late submissions, and not completed being given no points. Remember, to communicate with your seminar leader if you are having any difficulties with the module material.

A prerequisite for classroom interactivity is being actively present. All 10 live seminars will be marked for Class Interaction. Missing a live seminar will contribute 0 points to your weekly (0-5 numerical points) mark. Being present (following accepted class etiquette where possible, such as having your camera feed live during class and respecting other class members) and engaged (responding to polls, following discussions, and responding to questions) are part of our standards of online class behaviour; even your non-verbals add to the dynamics of the class interaction. Being actively engaged during the live session at the minimum will contribute 1 point. Part of being an effective leader is to engage with developing other people. That is part of what we expect you to build and model in this programme. This is achieved by assisting others in your modules to gain a deeper view of and gaining insights into our focal material. To earn higher marks (2 or 3 points) you need to substantively add to the class discussion, by asking meaningful questions that are focused on the module material, answering seminar leader or student questions (even by chat), or actively participating in small group discussions. The intention is to encourage and reward students for adding value to the overall learning environment. Marks summed from prework and classroom engagement (0-5 points) will be posted to the LMS gradebook before the next live seminar to help give timely feedback. These weekly marks are intended to give you an indication as to the appropriate level of interaction you are displaying in the module.

Notice: Students are expected to attend each live seminar. However, we do understand that emergencies and conflicts do occur. If a student knows there may be a conflict, that student should contact their live seminar leader as soon as possible (preferably before missing the class). In the case where a student misses a live seminar, that student is responsible for the content discussed in the live seminar and is expected to watch the recording. To earn any participation marks for that missed seminar the student must submit a minimum of three substantive comments or questions to the module wall focused on the recorded seminar they missed. In the unlikely event that a student missed three or more seminars, that student may be disallowed from completing the module and would need to retake it at their own expense unless they have been granted an Extenuating Circumstances exemption. The key here is to be communicative. Let the seminar leader know early if you may need to miss a live seminar.

Individual Assignments – 30%: This consists of two papers (1000 words each worth 15%). The first is a data summarisation and visualisation exercise and the second is an experiment proposal. Students will have the option of discussing individual assignments with a partner. The intention is to provide students with an "official" channel to provide and receive assistance on the assignments. However, the student's work and submission should be entirely her or his own. Detailed instructions for this assignment can be found on the LMS.

The individual assignments will be marked on a 25-point scale (see below for the grade implications) based on several factors:

- *Did you complete each of the requirements of the assignment?*
- *Did you discuss or use appropriate analytics techniques in an appropriate manner?*
- *How valid and sound are your interpretations of the results and your conclusions?*
- *How well did you communicate your ideas and arguments?*
- *How creative, novel, or insightful were your data, project proposal, and work?*

Team Assignment – 30%: This assignment will entail a data analytics project (3000 words) developed and submitted by a group of 3 to 4 students at the conclusion of the term. Students will self-select into groups for this paper and will utilise module techniques to analyse a dataset. There are very few constraints on what kind of project you do, what kind of data you use, or what kind of analysis you do.

Detailed instructions for this assignment can be found on the LMS. As this is a group project, substantial deviations in individual contributions may result in different marks among group members.

The assignment will be marked on a 25-point scale (see below for the grade implications) based on several factors:

- *Did you complete each of the requirements of the assignment?*
- *Did you discuss or use appropriate analytics techniques in an appropriate manner?*
- *How valid and sound are your interpretations of the results and your conclusions?*
- *How well did you communicate your ideas and arguments?*
- *How creative, novel, or insightful were your data, project proposal, and work?*

Scoring for Individual and Team Assignments: Each of the individual assignments and the team assignment will be assessed on a 25-point scale. Below is a table to translate the point score to your grade.

Grade	Individual and Team Assignment
A	20–25
B	17–19
C	13–16
D	10–12
E	7–9
F	6 points and below

Weekly Outline

This is a brief outline of weekly topics covered during the module. Also included are key readings assigned for the week and whether there is prework assigned during this week.

Required Purchases: No external purchases of books or materials are necessary for this module. Please see a list “Additional Resources” at the end of this syllabus for more readings on each topic.

Week 1: Descriptive Analytics

Summary	Summary statistics, data summarisation, and data visualisation
Readings	[Required] Review the different types of visualisations on Data to Viz and read through the caveats section. [Required] Watch the Tableau Getting Started (registration required) video and follow the video using your copy of Tableau to practice using the tool.

	[Optional] Tableau offers a wealth of additional training videos. If you find yourself stuck, there is a good chance what you need will be covered there.
Prework	Small-group activity output (one visualisation) to be submitted within 24 hours **after** the live session
Live Session	Data summarisation exercises and small-group visualisation activity
Assignments	Quiz 1 due 24 hours prior to the live session

Week 2: Data Wrangling

Summary	Data organisation and data collection
Readings	[Required] Read Parts 1–6 of Building Your First Extractor on the import.io help site. [Required] Watch the Advanced Pagination and Infinite Scroll video. [Required] Think of one or two examples of sites or pages that you would like to extract data from during the live session. Typically, pages that have structured data in a consistent format are the easiest to extract data. Examples include tables (like sports statistics or Wikipedia tables), product listings, product pages, or reviews. Feel free to explore though! [Optional] Import.io has a vast amount of resources in their help centre for you to refer to in the event you are stuck.
Prework	Small-group activity work (dataset) to be submitted within 24 hours **after** the live session
Live Session	Database exercise and small-group data collection activity
Assignments	Quiz 2 due 24 hours prior to the live session

Week 3: Survey Methods

Summary	Survey design and survey question design

Readings	[Required] We will be using the Office 365's Forms tool in class. Review how to create, collaborate on and share forms, and view results. Additional resources can be found in the help centre.
Prework	Small-group activity work (link to or survey) to be submitted within 24 hours **after** the live session
Live Session	Small-group survey design activity
Assignments	Quiz 3 due 24 hours prior to the live session **Individual assignment 1 due 24 hours after the live session**

Week 4: Probability and Statistical Inference

Summary	Foundations of probability, populations and samples, and hypothesis testing
Readings	[Optional] If you would like additional reference material on the topics we cover in this session, visit OpenIntro Statistics, Chapters 3–7.
Prework	Solution to assigned problems to be submitted 24 hours **before** the live session
Live Session	Probability and hypothesis testing exercises
Assignments	Quiz 4 due 24 hours prior to the live session

Week 5: Regression

Summary	Simple and multiple regression and correlation versus causation
Readings	[Optional] If you would like additional reference material on the topics we cover in this session, visit OpenIntro Statistics, Chapters 8–9.
Prework	Small-group activity work (regression output) to be submitted within 24 hours **after** the live session

Live Session	Regression exercises and small-group regression activity
Assignments	Quiz 5 due 24 hours prior to the live session

Week 6: Predictive Analytics

Summary	Extending regression to prediction, decision trees, limitations of prediction models
Readings	[Required] View the following videos from BigML's Education Videos: Introduction, Sources, Datasets, and Models I.
Prework	Small-group activity work (prediction model output) to be submitted within 24 hours **after** the live session
Live Session	Applications of prediction models exercises and small-group prediction model activity
Assignments	Quiz 6 due 24 hours prior to the live session

Week 7: Experiments

Summary	Experimental design and decision-making
Readings	None
Prework	Small-group activity work (experiment proposal slides) to be submitted within 24 hours **after** the live session
Live Session	Counterfactual exercises and small-group experiment proposal activity
Assignments	Quiz 7 due 24 hours prior to the live session

Week 8: Advanced Data Topics

Summary	Bayesian thinking, clustering, text analysis

Readings	[Required] View the following videos from BigML's Education Videos: Clusters and Topic Models.
Prework	Small-group activity work (clustering output) to be submitted within 24 hours **after** the live session
Live Session	Calibration of priors exercise and small-group clustering/text analysis activity
Assignments	Quiz 8 due 24 hours prior to the live session **Individual assignment 2 due 24 hours after the live session**

Week 9: Network Analytics

Summary	Network graphs, network and node statistics
Readings	None
Prework	Solution to assigned problems to be submitted 24 hours **before** the live session
Live Session	Network analysis discussion
Assignments	Quiz 9 due 24 hours prior to the live session

Week 10: Data, Ethics, and Organizations

Summary	Data-based businesses, data culture, data ethics
Readings	[Required] Any research you need to conduct to prepare your debate topic [Optional] Helping Organisations Navigate Ethical Concerns in Their Data Practices
Prework	Debate preparation notes to be submitted no later than 24 hours **before** the live session
Live Session	Data bias discussion and ethics debates
Assignments	**Team assignment; due 48 hours after the live session**

Module Expectations

This class covers a great deal of material. You will see each topic presented in multiple ways: concepts and foundations (sometimes with math), applications, and tools used to apply the methods. The intention is to provide multiple opportunities for you to grapple with the material, so you can build confidence at a personalised level. I expect each of you will work toward building intuition and understanding of the concepts. For some, it may mean delving further into the theory or tools. For others, it may mean considering concepts and applications. Regardless, I expect you to be active contributors in your groups and to our class. Every perspective is important.

Faculty Contact

Seminar faculty will let you know during the first live seminar, how it is best to contact them for individual attention. Many seminar leaders stay after live sessions to answer individual questions or schedule separate meetings with students. The Wall is also a great option for communicating general content questions or posting interesting material related to the module as it is a common space for faculty and students.

Additional Resources

This class will introduce you to a number of methods in data analytics. If you are interested in learning more, here is a list of resources to dig deeper into the material. Note the icon (‡) in Pricing indicates the resource is available from UCL. Clicking on the link will direct to the module reading list where you can access it.

Resource	Format	Pricing	Class Week	Tool
Show Me the Numbers	Book	Purchase	1	
Storytelling With Numbers: A Data Visualization Guide for Business Professionals	Book	Purchase(‡)	1	
Tableau eLearning	Site	Purchase	1	Tableau
Practical Tableau	Book	Subscription (‡)	1	Tableau
Database Design for Mere Mortals	Book	Subscription (‡)	2	
Damned Lies and Statistics	Book	Purchase(‡)	2, 4, 5	

How to Lie With Statistics	Book	Purchase	2, 4, 5	
Statistics: The Art and Science of Learning From Data	Site	Free	1, 4, 5	
OpenIntro Statistics	PDF	Free	1, 4, 5	
Cartoon Guide to Statistics	Book	Purchase	1, 4, 5, 7	
Seeing Theory	Site	Free	4,5	
Khan Academy Statistics and Probability	Videos	Free	4,5	
Python for Data Analysis	Book/ Videos	Subscription (‡)	2, 5, 6	Python
Survey Research Methods	Book	Purchase(‡)	3	
Data Smart	Book	Purchase(‡)	6, 8	Excel
Data Camp (Python skill tracks and career track)	Site	Free/ Subscription	1, 2, 4, 5, 6, 8, 9	Python
Impact Evaluation in Practice, Chapters 3 and 4	PDF	Free	7	
Running Randomized Evaluations	Book	Purchase(‡)	7	
Social Network Analysis for Startups	Book	Subscription (‡)	9	Python
Social and Economic Networks	Book	Purchase(‡)	9	
The Human Network	Book	Purchase	9	
Weapons of Math Destruction	Book	Purchase	10	
edX and Coursera Data Classes	Site	Free	All	Python, General

Tools

In this class, we will use a number of tools. Please refer to the LMS to make sure you are able to access each one.

Software	Availability	Description
Microsoft Excel	UCL student license	Tried and true
Tableau	Student license from site	Desktop data visualisation application
Import.io	Group license available	Online web scraping software
Micorsoft Forms	UCL student license	Online survey design tool
Google Colab	Free with Google account	Online environment to access Python code
BigML	Student license using UCL email	Online machine learning platform
Gephi	Open Source	Network analysis tool