

Marketing Analytics

Wei Miao

2021-09-01

Contents

1	Preface	5
2	Syllabus: MSIN0094 Marketing Analytics	7
2.1	Instructor Information	7
2.2	Teaching Assistants	7
2.3	Course Objective	7
2.4	Module Structure Overview	8
2.5	Office Hours	8
2.6	Programming Language	9
2.7	Classroom Etiquette	9
2.8	Assignments and Grading policy	10
2.9	How to Make the Best Use of Case Studies	11
2.10	Copyright Issues	11
3	Detailed Class Descriptions	13
3.1	Week 1 (Thu, 07 Oct)	13
3.2	Week 2 (Thu, 14 Oct)	14
3.3	Week 3 (Thu, 21 Oct)	15
3.4	Week 4 (Thu, 28 Oct)	16
3.5	Week 5 (Thu, 04 Nov)	17
3.6	Week 6 (Thu, 11 Nov)	18
3.7	Week 7 (Thu, 18 Nov)	19
3.8	Week 8 (Thu, 25 Nov)	20
3.9	Week 9 (Thu, 02 Dec)	20
3.10	Week 10 (Thu, 09 Dec)	22
4	Lecture Notes	23

Chapter 1

Preface



This is the online supplement to the MSIN0094 Marketing Analytics Module at UCL School of Management.

Chapter 2

Syllabus: MSIN0094 Marketing Analytics

(Last update on 01 September, 2021)

2.1 Instructor Information

Module Leader: Dr. Wei Miao¹ (wei.miao@ucl.ac.uk)

Office: S3, Level 38, One Canada Square

Office hour:

- Appointment link
- Please refer to the office hour session for how to make office hour appointments. More details to be confirmed in Week 1's class.

2.2 Teaching Assistants

- Christopher Ogbunuzor (christopher.ogbunuzor.14@ucl.ac.uk)
- Henry Bellhouse (henry.bellhouse.18@ucl.ac.uk)

2.3 Course Objective

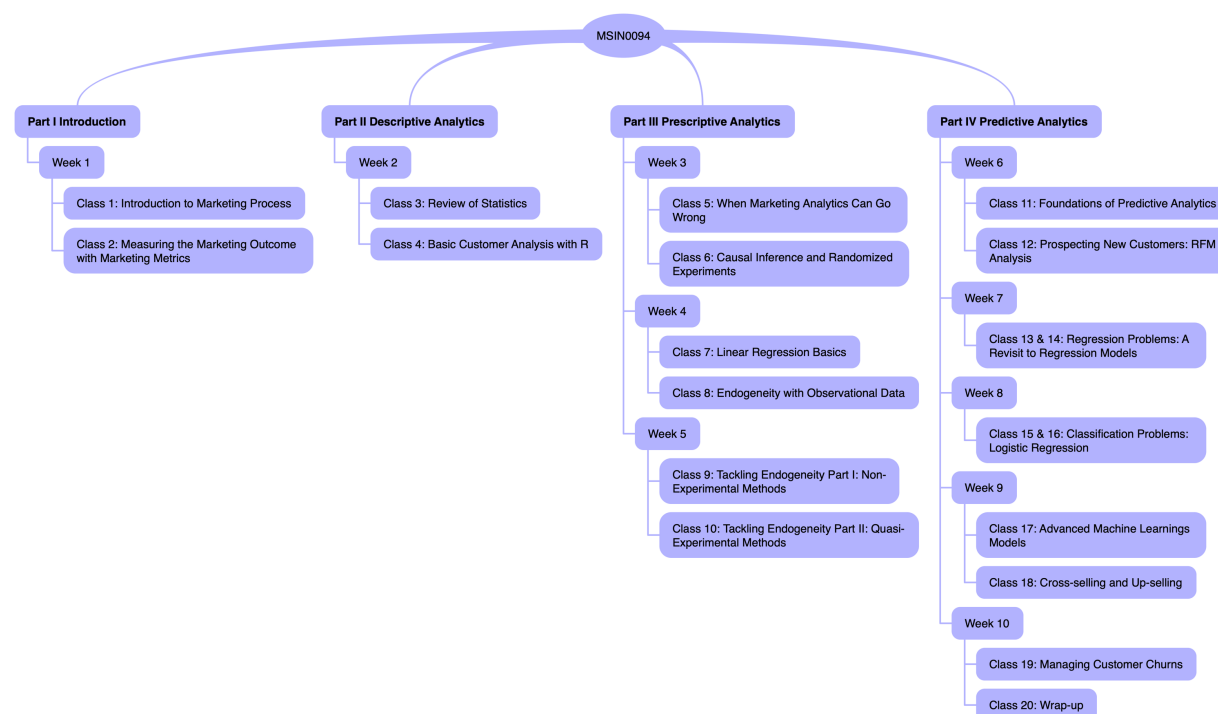
Marketing analytics addresses how to utilize analytics tools to better carry out marketing tactics and strategies. The digital age has fundamentally altered the way we collect, process, analyze, and disseminate data intelligence. Driven by advances in computing power, software, and other information technologies, the very nature of marketing analytics has evolved. In this module,

¹Wei is pronounced as “way”, which stands for jade in Mandarin; Miao is my family name, which is exactly pronounced as “meow”, the sound of a kitty.

students will learn the scientific approach to real-world marketing problems with hands-on use of a combination of econometrics, statistics, machine learning, and programming, to collect, analyze, and act on customer intelligence.

Throughout the module, student will learn how to guide a company's marketing decisions along three dimensions: descriptive analytics, prescriptive analytics, and predictive analytics. Specifically, descriptive analytics involves using visualization tools and statistical tests to provide model-free insights into the customer intelligence; prescriptive analytics focuses on drawing causal inference via experimental, quasi-experimental, and non-experimental methods; predictive analytics shed light on how to prospect new customers, up-sell and cross-sell high value products, and mitigate customer churns, using appropriate machine learning models. At the end of the module, students will have the full capacity to carry out independent marketing research for their dissertation projects as well as lead a marketing analytics team in their future jobs.

2.4 Module Structure Overview



2.5 Office Hours

It is a good practice to utilize the Moodle forum as an interactive place to promote peer-to-peer learning. For any question, please first check the Moodle forum and see if the questions are already posted and answered by any of your fellow classmate there. If not, please post your questions in the Moodle forum. For any forum questions, we will first leave for your fellow classmates to provide potential answers. The teaching team will also monitor the forum and clarify any unanswered questions.

If your question is still unsolved, please make appointments for office hours. I am available to answer any questions you may have about past lectures, case studies, and any other queries related to marketing analytics during my office hours. I will confirm the detailed time slots and the link in the first class.

Please make sure of the following before/when you make your appointment:

- Please be careful when operating on the Excel form and kindly avoid overwriting/modifying other students' appointments.
- Please use the “comment” function in Excel form to briefly describe the questions you have so that other students sharing similar questions can drop in as well.
- Each session lasts for 15 minutes. If you expect your questions may take a longer time to solve, you can book two consecutive sessions. Each student is entitled to at most two sessions per week to ensure equal opportunities for all students.

2.6 Programming Language

We will use an open-source programming language R throughout the course. R is one of the most commonly used programming languages by data scientists, economists, and statisticians, and is sometimes called the “golden child” of data science. R is cross-platform compatible; with over 16,000 packages ready for use in CRAN (R's open-source repository), it has one of the richest ecosystems to perform data analytics tasks. More importantly, it's free of charge compared with other commercial software (it is good to have cost-effectiveness in mind as you are now business school students)!

In this module, I will cover the basics of R and then focus on how to perform data-cleaning tasks efficiently with R. If you have little prior exposure to R, **it is highly recommended that you start learning some R basics before Week 2 to avoid falling behind**. UCL library has rich digital resources of R textbooks, which can be downloaded free of charge using your UCL credentials.

During (and even after) this module, whenever you run into any problems with R, Google (especially Stack Overflow) is always your best place to seek answers for most debugging issues; if you have any specific tasks in mind, the R Task Views page provides comprehensive guidance of which packages on CRAN are relevant for tasks related to a certain topic.

2.7 Classroom Etiquette

- **Punctuality.** Classes will start on time. Please be on time.
- **Class participation.** Please keep your mobile phones and laptops muted at all times unless otherwise instructed.

2.8 Assignments and Grading policy

Much of the learning during the course will happen with the help of individual assignments. Each assignment is carefully designed to examine the learning outcome of the topics taught during the previous weeks. Your final grade in this module will be based on 3 individual assignments with their associated weights as follows:

- 1st assignment, 30% weight, 1500 words, due on Oct 29
- 2nd assignment, 40% weight, 2000 words, due on Nov 26
- 3rd assignment, 30% weight, 1500 words, due on Dec 17

Detailed assignment descriptions will be given in due course.

2.8.1 Code of Conduct and Plagiarism

As all assignments in this module are **individual** assignments, you should **NOT** work with other students and the write-up should reflect your own work only. Please refrain from consulting classmates, friends, or seniors who have taken the course already. In the context of this module, it is acceptable to refer to concepts, frameworks, and analytical tools from the readings or class lectures with appropriate citation.

However, do not directly copy or paraphrase anything from other sources outside our classroom and present it as your own. It will be a violation of academic integrity if you base your assignments on solutions which you have found on the Internet or which you have obtained from others. You may risk failing this module if such a violation is found and investigated by the University. For detecting plagiarism, the School uses Turnitin. Turnitin checks your submission against 24+ billion web pages, 300+ million student papers, and 110,000+ publications and provides a similarity index. A particularly high similarity index will draw attention to the submission.

2.8.2 Word count

Please note, according to UCL's policy, you should **follow the word limit**. Otherwise, your mark may be reduced. Therefore, try to keep your answer concise and to the point.

2.8.3 Marks and Feedback

Your marks and feedback for each assignment will be usually released no later than 4 weeks after the submission date. Each assignment will be carefully marked according to the mark scheme by our teaching assistants, and there will be adequate second-marking procedures in the marking process to ensure the marking quality, as required by UCL. As a result, it is the university policy that students shouldn't question the academic judgement of markers, and **please refrain from emailing teaching assistants for any re-marking**.

2.9 How to Make the Best Use of Case Studies

To develop and test your skills in solving real-life marketing analytics problems, we will use many relevant case studies along the learning journey. Most case situations have been developed after careful research on actual situations faced by real companies and managers. The case authors have attempted to describe enough of the background and details of the situation in order to provide an adequate basis for class discussion.

Merely reading the case is not enough. To maximize the learning outcome, please follow the following advice for case preparations:

- After an initial reading to get a broad overview, go back and study the case thoroughly.
- Make notes wherever necessary and mark up the case, especially key information such as important numbers, to facilitate your understanding of the case background.
- Carefully think about the key problems. Always try to come up with alternative solutions to the problem and think about the pros and cons of each.

2.10 Copyright Issues

All of the materials in this module are copyrighted, either purchased by UCL, written by me, or owned by other people. Please refrain from uploading any materials in the module to other external websites or sharing any materials in the module with anyone who is not attending this class. A copyright violation may be investigated by the school.

Chapter 3

Detailed Class Descriptions

Students are expected to complete the required preparation work for each week's class, including reading case studies, preparing for case questions, etc. Detailed guidelines on how to prepare for each week are described below. Please follow the guidelines unless otherwise informed.

Note that the class schedules are tentative and are **subject to changes based on our progress**, which means this online supplement may be updated regularly. Therefore, It's a good practice to come back and check if there are any updates to the guideline each Wednesday before Thursday's class.

Some explanations:

- **(before class)** means the reading material should be read and prepared before that week's class.
- **(after class)** means the reading material serves as a supplementary after-class reading to enhance your understanding of the week's lecture contents.
- **(optional)** means the reading material is technically challenging but can be handy when you need them. You can selectively read them based on your time availability.

3.1 Week 1 (Thu, 07 Oct)

3.1.1 Class 1: Introduction to Marketing Process

Class objectives:

- To provide an overview of the course topics and requirements
- To understand the typical marketing process in real-life examples

Readings:

- Supplementary reading: "Marketing Process" (after-class)

3.1.2 Class 2: Marketing Metrics: Customer Lifetime Value

Class objectives:

- Understand common marketing metrics
- Know how to compute customer lifetime value and how to use CLV to guide marketing decisions
- Apply CLV calculation in real-life business cases: Tesco & Vodafone

Readings:

- Assignment 1 Part I: Vodafone: Using CLV for Better Marketing Decisions (before-class, 15 min prep time)
- Supplementary reading: “Customer Lifetime Social Value” (after-class; optional)

Preparations:

- Important: Please read the Vodafone case carefully before Week 1’s class. Highlight any information (such as figures or key information) you feel may be relevant to solving the case. We will discuss the Vodafone case in class and solve the “in-class” questions which are listed at the end of the case.
- If you haven’t learned linear algebra before and if your time allows, I recommend you start reading “Review of basic linear algebra”, which is a prerequisite for Week 4.

Assignment 1:

- Assignment 1 examines your learning outcome from Week 1 to Week 4, including CLV; basic customer analysis; and causal inference. You now have access to Part I of Assignment 1. I suggest you start working on the part I now so that you leave yourselves enough time for the remaining parts.

3.2 Week 2 (Thu, 14 Oct)

3.2.1 Class 3: Review of Statistics

Class objectives:

- Understand the type of data and variables commonly used in marketing analytics
- Review of statistics
 - Statistical inference & hypothesis testing

- Statistical tests commonly used in marketing

Readings:

- “Tips for using statistics in marketing analytics” (after class)

Preparations:

- You should have installed R and RStudio on your laptops **before this class**

3.2.2 Class 4: Basic Customer Analysis with R**Class objectives:**

- Understand the usage and syntax of R data.table packages for data manipulation
- Able to apply basic statistics for basic customer analysis in business cases

Readings:

- Case study: “Basic Customer Analysis: Amazon Prime” (before class, 10 min prep time)
- Supplementary reading: data.table Cheat Sheet (after class)
- Assignment 1 Part II: “Basic Customer Analysis with Amazon Sales Data” (after class)

Preparations:

- Please read the case “Basic Customer Analysis: Amazon Prime”. There are quite a few questions at the end of the case. We will devote most of this class to hands-on exercise with R and data.table, answering the questions listed at the end of the Amazon Prime case.
 - We will use the dataset “Amazon Prime.csv” from Moodle. Remember to download the dataset to your hard disk before class.
- You should be able to solve the part II of assignment 1 after this week’s class. Get started now and remember not to procrastinate until the deadline!

3.3 Week 3 (Thu, 21 Oct)**3.3.1 Class 5: When Marketing Analytics Can Go Wrong****Class objectives:**

- Learn how to tell good analytics from bad ones

- Understand how to avoid interpreting statistical results incorrectly

Readings:

- Case study: “When Marketing Analytics Can Go Wrong: Case Studies” (before class, 20 min prep time)

Preparations:

- Read all case studies carefully before class
- Prepare and think over the questions at the end of each case for class discussion

3.3.2 Class 6: Causal Inference and Randomized Experiments**Class objectives:**

- Understand the Rubin causal inference framework
- Know the keys steps to run experiments
- Understand the pitfalls when running experiments

Readings:

- Case study: “Superb Trucks Lp: How to Conduct a Field Experiment” (before class, 20 min prep time)

Preparations:

- Read the case “Superb Trucks Lp” before class. We will demonstrate how to use field experiments to solve marketing problems with this

3.4 Week 4 (Thu, 28 Oct)**3.4.1 Class 7: Linear Regression Basics****Class objectives:**

- Understand the concept of “data generating process” and a “model”
- Understand linear regression models from different perspectives
- Understand how to interpret the regression coefficients

- Continuous variables and standardization
- Log transformation
- Categorical variables
- Interactions

Readings:

- Prerequisite reading: “Review of basic linear algebra” (before class, 60 min prep time)
- Supplementary reading: “Practical Regression: Regression Basics” (before and after class)

Preparations:

- This week, we will learn the OLS linear regression model. This is the simplest yet most powerful model for data analytics, and it is the cornerstone of the remaining contents of this module. Therefore, it is important that you are well-prepared and stay focused for the class.
- *Important:* Please finish the “Review of basic linear algebra” before class. This is important, as during the lecture, I will show you how to run regression using linear algebra.

3.4.2 Class 8: Endogeneity with Observational Data**Class objectives:**

- Understand the concept of selection bias
- Understand the common causes of endogeneity from observational data

Readings:

- Supplementary reading: “Practical Regression: Introduction to Endogeneity: Omitted Variable Bias” (before and after class)

3.5 Week 5 (Thu, 04 Nov)**3.5.1 Class 9: Tackling Endogeneity Part I: Non-Experimental Methods****Class objectives:**

- Understand the intuition behind the instrumental variable methods
 - Able to find appropriate instrumental variables to solve endogeneity using 2SLS method.
- Understand the intuition behind the fixed effects

Readings:

- Supplementary reading: “Practical Regression: Causality and Instrumental Variables” (before and after class)

Preparations:**3.5.2 Class 10: Tackling Endogeneity Part II: Quasi-Experimental Methods****Class objectives:**

- Understand the intuition and how to use difference-in-differences method
- Understand the intuition and how to use regression discontinuity method
- Learn real-life applications of quasi-experimental methods

3.6 Week 6 (Thu, 11 Nov)**3.6.1 Class 11: Foundations of Predictive Analytics****Class objectives:**

- Understand the difference between prescriptive analytics and predictive analytics
- Understand the bias-variance tradeoff in predictive analytics
- Understand the difference between regression problems and classification problems

Readings:

- Supplementary reading: “Assessing Prediction Accuracy of Machine Learning Models” (after class)
- Supplementary reading: Varian, Hal R. 2014. ‘Big Data: New Tricks for Econometrics’. *Journal of Economic Perspectives* 28 (2): 3–28. (after class)
 - page 6-7: General Considerations for Prediction
 - page 21-24: Econometrics and Machine Learning

3.6.2 Class 12: Prospecting New Customers: RFM Analysis

Class objectives:

- Understand how to use RFM analysis to prospect new customers and improve the effectiveness of marketing activities
 - The concept of RFM
 - The steps to do RFM
 - Break-even analysis of RFM
 - Important variants of RFM
- Understand the concept of “lift” and “gain”, and how to use the concepts to evaluate predictive models

Readings:

- Case study: “UberEat: Mobile App Push Messaging” (before class, 10 min prep. time)
- Assignment 2: “Amazon Prime: Using Predictive Analytics to Improve Marketing ROI” (after class)

Preparations:

- Read and prepare “UberEat: Mobile App Push Messaging” for class discussion; the dataset “UberEat.csv” is on Moodle. Note that, for you to better understand and be able to compare different predictive models, we will be using the same UberEat case study and dataset in the next few weeks. Each week, we will learn a new predictive model, and you will learn how to apply different models on the UberEat dataset, and compare their performance.
- It’s time to catch up with our old friend again in the assignment. After this class, you should be able to solve the Part I of Assignment 2 “Amazon Prime: Using Predictive Analytics to Improve Marketing ROI”. It’s good to start early as the second assignment accounts for 40% of the total grade.

3.7 Week 7 (Thu, 18 Nov)

3.7.1 Class 13 & 14: Regression Problems: A Revisit to Regression Models

Class objectives:

- Review of simple linear regression models
- Learn the definition and application of step-wise and backward stepwise regression models.
- Learn the concept of L1 and L2 regularization in the context of linear regression models

- Know the pros and cons of Lasso and Ridge regressions; Understand when and how to apply the two regularized regression models

Readings:

- Supplementary reading: Varian, Hal R. 2014. ‘Big Data: New Tricks for Econometrics’. Journal of Economic Perspectives 28 (2): 3–28. (after class)
 - page 15: Variable Selection

3.8 Week 8 (Thu, 25 Nov)

3.8.1 Class 15 & 16: Classification Problems: Logistic Regression

Class objectives:

- Understand the mathematics and intuition behind the logistic regression
- Understand how to interpret logistic regression results
 - probability, odds, odds ratio, and log odds ratio
- Understand how to score customers using logistic regression
- Compare the predictive power of logistic regression against RFM analysis

Preparations:

- This week, I will demonstrate how to use logistic regression to solve the UberEat case. We will also compare the performance of logistic regression with that of RFM model.
- At this point we have covered everything you need to know to work on the second assignment.

3.9 Week 9 (Thu, 02 Dec)

3.9.1 Class 17: Advanced Machine Learnings Models

Class objectives:

- Learn the intuition behind advanced machine learning models and their application
 - Decision tree
 - Bagging and boosting models
 - Deep learning

- Learn how to use automatic machine learning tools to make our life easier

Readings:

- Supplementary reading: R-blog, [<https://www.r-bloggers.com/2020/04/automl-frameworks-in-r-python/>] (before and after class)
- Supplementary reading: Varian, Hal R. 2014. ‘Big Data: New Tricks for Econometrics’. Journal of Economic Perspectives 28 (2): 3–28. (after class)
 - page 7-15: Classification and Regression Trees

Preparations:

- Follow the instructions in <https://docs.h2o.ai/h2o/latest-stable/h2o-docs/downloading.html> and install h2o package on your laptop before the class.

3.9.2 Class 18: Cross-selling and Up-selling**Class objectives:**

- Understand the benefits of using customer purchase data for customer development
- Understand how to cross-sell and up-sell

Readings:

- Case study: “Cross-Selling Insurance Products Using Predictive Analytics” (before class)
- Supplementary reading: Knott, Hayes, and Neslin “Next-Product-to-Buy Models for Cross-Selling Applications,” Journal of Interactive Marketing 16(3) (2002) (after class)

Preparations:

- Please read the “Cross-Selling Insurance Products Using Predictive Analytics” before class. Since we have already learned all the tools in the previous weeks. Please try if you can develop you own predictive model using RFM and logistic regression, which we have covered in previous weeks.
- In the class, I will demonstrate how to use more sophisticated machine learning techniques to solve the case.

3.10 Week 10 (Thu, 09 Dec)

3.10.1 Class 19: Managing Customer Churns

Class objectives:

- Understand how to use attrition prediction for managing customer churn management

Readings:

- Case study: “Customer Churn Management Using Predictive Analytics” (before class)

Preparations:

- Read the case “Customer Churn Management Using Predictive Analytics” and prepare the questions for class discussion
- Since we have covered all the necessary predictive models in the previous weeks, please explore the dataset and develop at least one predictive model for managing customer churn before class. I will invite some of you to share your solutions in class.

3.10.2 Class 20: Module Wrap-up

Class objectives:

- To review and synthesize the module

Chapter 4

Lecture Notes

The lecture notes will be released one week before each class. Please do not circulate without permission from the module leader. To use the lecture notes outside MSIN0094, please contact the module leader (wei.miao@ucl.ac.uk)