

Marketing Analytics Weekly Arrangements

Wed, 21 Sep 2022

Preface: How to Use This Guide

Arrangements each week

We will have a 3-hour lecture each Thursday and I will aim to cover a new marketing analytics model each week. Whenever we learn a new technique (e.g., a new statistical model or a new analytics tool), the subsequent week will often start with some warm-up exercise and a workshop (with a case study) for you to review and practice the new technique learned in the previous week. This way, you would have time to digest what you've learned and can further reflect on your understanding of the technique by practicing your skills with a real-life application.

For instance, on Thursday 6 Oct (Week 1), I will first introduce the concepts of marketing and marketing process, and then will cover the concept of customer lifetime value (CLV) and how to compute CLV with R. In week 2, we will therefore begin with a case study that helps you practice your knowledge of CLV, so you can understand how to use CLV for better marketing decisions in your future projects/jobs.

In the remaining time of week 2, I will then introduce a new technique: the `dplyr` package in R, which helps us clean and manipulate datasets in R. Following a similar logic, then in week 3, we will start with a case study for you to practice the `dplyr` package. So on and so forth.

About the labels

- Essential: contents core to this week's materials. All pre-class preparations should be completed before class.
- Optional: supplemental readings for those interested in learning more

All materials, including the lecture slides and before-class readings will be released a few days before each Thursday's class.

Module Outline

Table 1: Module Outline

Week	Analytics/Methodology Topic	Substantive Topic	Case Study
1	R basics	Customer Lifetime Value	
2	R <code>dplyr</code>	Data Wrangling	CLSV
3	Unsupervised learning	Segmentation	Preliminary Customer Analysis
4	Supervised learning	Targeting	RFM Analysis
5	Randomized controlled trial	Customer churn management	Predicting Customer Churn at QWE Inc.
6	Linear regression I	Promotion analytics	A/B Testing at Vungle
7	Linear regression II	Pricing analytics I	Zalora: Data-Driven Pricing
8	Endogeneity and instrumental variable	Pricing analytics II	
9	Quasi-experiments: DiD and RDD	Search Engine Marketing	Evaluating the Impact of COVID-19 on Ride-sharing Market
10	Causal forest	Search Engine Marketing	The power of ranking

Induction Week: R Basics

! Pre-class preparation

- Finish reading “An introduction to R” (can be assessed in this [link](#)) Chapters 1, 2, and 3.
 - Please try to practice the codes in R along your reading.
 - Take a note of any questions you may have during your self-study. I will cover R basics in greater details in the induction week.
- What you will learn
 - An introduction to R basics
- After-class exercise
 - (essential) Finish data camp “Introduction to R” tutorial before Week 1 class. We will learn how to use R to compute customer lifetime value next week, so it’s very important that you are familiar with R basics before class.

Week 1: Module Introduction and Customer Lifetime Value

Module Introduction

- What you will learn
 - Concept of marketing and marketing process
 - An overview of the course topics and requirements
 - How marketing analytics can empower marketers in the digital era

Customer Profitability and Lifetime Value

- What you will learn
 - The concept of customer life cycle
 - The concept of break-even analysis
 - How to compute customer acquisition cost (CAC)
 - How to compute customer lifetime value (CLV) in R
- After-class reading
 - (optional) [Hubspot: How to compute CLV](#). This article introduces alternative ways to compute CLV, which are used in many companies.
 - (optional) [Important lessons for embracing customer lifetime value](#)

Week 2: Data Wrangling with R

Workshop: Customer Lifetime Social Value

! Pre-class preparation

- HBS 9-518-077: Customer Lifetime Social Value (CLSV).
 - Please carefully read the case background before class; we will be discussing the case in class

- What you will learn
 - How to apply CLV calculation in a real-life scenario
 - How to extend the concept of CLV to CLSV as needed
 - Discuss how CLV can be used by marketers to guide marketing decisions
- After-class reading
 - (optional) [The Dangerous Seduction of the Lifetime Value \(LTV\) Formula](#)

Data Wrangling with R: Part I

- What you will learn
 - How to use `filter`, `select`, `mutate`, and `arrange` for data manipulation with `dplyr` package in R
- After-class reading
 - (essential) Finish the practice exercise on data camp: [Data Manipulation with dplyr: 1.Transforming Data with dplyr](#)
 - (essential) [Cheatsheet for dplyr](#)
 - (optional) [Python Pandas vs. R Dplyr](#). In David's class, you will learn how to use Python Pandas for data wrangling. This cheat sheet provides a one-to-one comparison of the two libraries for your reference.

Week 3: Predictive Analytics for Customer Segmentation

Data Wrangling with R: Part II

! Pre-class preparation

- Case Study: Preliminary Customer Analysis
 - Please carefully read the case background before class; we will be discussing the case in class
- What you will learn
 - The usage of `group_by` and `join` for data aggregation and merge with `dplyr` package
 - How to use `dplyr` to conduct preliminary customer analyses
- After-class reading
 - (essential) Finish the practice exercise on data camp: [Data Manipulation with dplyr: 2.Aggregating Data](#)

Predictive Analytics: Unsupervised Learning

- What you will learn
 - Important concepts in predictive analytics
 - Concept of unsupervised learning
 - How to run K-means clustering in R

Week 4: Predictive Analytics for Customer Targeting

Workshop: Targeted Marketing with Predictive Analytics

! Pre-class preparation

- Case Study: Recency Frequency and Monetary Value (RFM) Analysis
 - Please carefully read the case background before class; we will be discussing the case in class
- What you will learn
 - The premise behind RFM analysis
 - How to implement an RFM campaign in R
 - How targeted marketing improves ROI of marketing campaigns

Decision Trees and Random Forest

- What you will learn
 - Intuition behind decision tree and random forest models
 - How to build random forest models in R
 - Application of random forest in targeted marketing
- After-class reading
 - (optional) Varian, Hal R. “Big data: New tricks for econometrics.” *Journal of Economic Perspectives* 28, no. 2 (2014): 3-28
 - (optional) Biau, Gérard, and Erwan Scornet. ‘A Random Forest Guided Tour’. *arXiv*, 18 November 2015. <http://arxiv.org/abs/1511.05741>.

Week 5: Causal Inference and RCT

Workshop: CSR with Machine Learning

! Pre-class preparation

- Case study: Predicting Customer Churn at QWE Inc.
 - Please carefully read the case background before class; we will be discussing the case in class
- What you will learn
 - How to use predictive models for Customer Relationship Management (CSR)

Causal Inference & Randomized Experiments

! Pre-class preparation

- This and next week, we will be conducting t-tests with R in class to estimate treatment effects from randomized experiments. If you're unfamiliar with t-tests, please go through this [Review of Statistics with R](#) before the next week's class.
- What you will learn
 - Concept of causal inference
 - Concept of Rubin's potential outcome framework
 - The intuition of why RCT guarantees causal inference
 - Steps to scientifically design and conduct an RCT
- After-class reading
 - [Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing](#). Cambridge University Press, 2020., by Ronny Kohavi (Airbnb, formerly Microsoft and Amazon), Diane Tang (Google), and Ya Xu (LinkedIn). This is a very practically oriented guide to experimentation, with many examples relevant to marketing and product management.

Week 6: Linear Regression

Workshop: A/B Testing at Vungle

! Pre-class preparation

- UVA-QA-0821: A/B Testing at Vungle
 - Please carefully read the case background before class; we will be discussing the case in class
- What you will learn
 - The business model of the advertising industry and mobile marketing industry
 - Design and analyze A-B testing to solve real-life marketing problems
- After-class reading
 - (recommended) [Test and learn: How a culture of experimentation can help grow your business](#)
 - (optional) [A Comparison of Approaches to Advertising Measurement: Evidence from Big Field Experiments at Facebook](#)
 - (optional): Berman, Ron, and Christophe Van den Bulte. 'False Discovery in A/B Testing'. Management Science, 30 December 2021

Linear Regression Model: Basics

! Pre-class preparation

- This week's content will involve quite some maths and will be likely challenging. It's highly recommended that you go through [this tutorial](#) to refresh your knowledge of linear algebra before class. Only read Sections 1, 2, 3.1 - 3.3, 3.6 - 3.7.

- What you will learn
 - The concept of Data Generating Process (DGP) and a model
 - The intuition behind coefficient estimation of linear regression models
 - Run linear regression models in R
 - How to interpret the regression coefficients and statistics
- After-class reading
 - (highly recommended) [Introduction to Econometrics with R](#), Chapters 4-7. These 4 chapters cover very detailed applied knowledge of linear regressions. Due to limited time, we cannot cover all contents in class, so it would be great if you can take time to go through these chapters thoroughly.

Week 7: Marketing Mix Modeling

Linear Regression Model: Advanced Topics

- What you will learn
 - How to interpret coefficients of log transformed variables
 - How to model non-linear relationship using linear regression
 - How to interpret the coefficients of categorical variables

Workshop: Marketing Mix Modelling with Linear Regression

! Pre-class preparation

- Case study: Zalora: Data-Driven Pricing Recommendations
 - Please carefully read the case background before class; we will be discussing the case in class

- What you will learn
 - Build regression-based marketing mix models to solve pricing and promotion optimization problems
- After-class reading
 - [A complete guide to Marketing Mix Modeling](#)

Week 8: Endogeneity and Instrumental Variables

Endogeneity and Panel Regression with Fixed Effects

- What you will learn
 - Endogeneity and its consequence in causal inference
 - Omitted variable bias
 - Reverse causality
 - Measurement error
 - fixed effects regression for panel data

Instrumental Variables

- What you will learn
 - Intuition of why instrumental variables solve endogeneity
 - The two requirements of a valid instrumental variable and how to find good instruments
 - Apply two-stage least square method to estimate the causal effects using instrumental variables
 - Application of instrumental variables in the marketing and business field
- After-class reading
 - (optional) [Econometrics with R: Instrumental Variables Regression](#)

Week 9: Quasi-Experimental Methods

Difference-in-Differences Design

- What you will learn
 - Concept of difference-in-differences (DiD) design
 - Estimation of causal effects using the DiD design
 - Application of DiD design in the marketing and business field

Regression Discontinuity Design

- What you will learn
 - Concept of regression discontinuity design (RDD)
 - Estimation of causal effects using the RDD design
 - Application of RDD designs in the marketing and business field
- After-class reading
 - (optional) Varian, Hal R. “Causal inference in economics and marketing.” Proceedings of the National Academy of Sciences 113, no. 27 (2016): 7310-7315.
 - (optional) [Econometrics with R: Quasi-experiments](#)

Week 10: Frontiers of Marketing Analytics

Workshop: Search Engine Marketing

! Pre-class preparation

- Case study: The Power of Ranking
 - Please carefully read the case background before class; we will be discussing the case in class
- What you will learn
 - What is search engine marketing (SEM)
 - How to apply data-driven analytics for optimal SEM
 - More applications of quasi-experimental method in Marketing

Frontiers of Marketing Analytics

- What you will learn
 - Causal machine learning with causal forest
 - Heterogeneous treatment effect estimation with causal forest in R
- After-class reading
 - [Estimate causal effects using ML](#) by Microsoft Research
 - Athey, Susan, and Stefan Wager. 'Estimating Treatment Effects with Causal Forests: An Application'. ArXiv:1902.07409 [Stat], 20 February 2019. <http://arxiv.org/abs/1902.07409>.