293 – Marketing Analytics

Syllabus - Fall 2018 (tentative)

Professor: Mike Palazzolo (mpalazzolo@ucdavis.edu)
Office hours: By appointment (in person or remote), or just stop by and see if I'm in my office (3308)

Course Description

"What's the counterfactual?"

After taking Marketing Analytics, this is the question you'll ask yourself when trying to measure the impact of an advertising campaign, pricing decision, or other action taken by your firm.

When attempting to measure the impact of a marketing action X (e.g., ad campaign) on an outcome Y (e.g., sales), the temptation is often to compare today to yesterday. We ask "Were our sales with the new ad campaign higher than our sales used to be with the old ad campaign?" But this is never the right question to ask, as many things change over time. For example, perhaps demand for our product is seasonal, or perhaps a competitor has entered the market since we began the new campaign. The right question to ask is "Were our sales with the new ad campaign higher than they would have been if we continued the old ad campaign, all else equal?"

The difference between the scenario we observe (today's sales with the new campaign) and the one we do not (today's sales with the old campaign—a counterfactual scenario) is the true impact of the new campaign on our sales. This course trains you to think in terms of counterfactuals, to run analyses in ways that better estimate the impact of past actions taken by a firm on consumer behavior, and to use corresponding insights to inform future action. In so doing, the course surveys different types of datasets and models that you are likely to encounter in the analytics world.

At the end of the day, no one marketing analytics course can provide experience with all analytical tools each student will use at their next job, as each person's work environment will differ. What a course *can* do is train each student how to *think* about marketing problems and data, to facilitate rapid emersion in their post-graduation analytics environment, and provide a foundation for making smart, data-driven decisions. That is the goal of this course.

Requirements

Prerequisites:	Stats Core (203A) or equivalent; Marketing Core (204) and 203B encouraged[1]
Software:	Microsoft Excel
Required Text:	Mastering 'Metrics by Angrist & Pischke
Optional Software:	R + R Studio (For extra credit portion of assignments)
Optional Texts:	An R Companion to Applied Regression by Fox and Weisberg Statistics for Management and Economics by Gerald & Keller (203A text) Introduction to Econometrics 3rd Edition by Stock & Watson

[1] Note that this is **not an introductory course** – it is an intermediate-to-advanced course that presumes some familiarity with statistics and some experience working with data. Having a strong understanding of linear regression by the end of session two is necessary. This is easier if you have already taken 203B. If you don't have experience with linear regression prior to the course, study Ch. 16-17 of the 203A text carefully.

Course Topics

Module	Wk	Topics	Readings: Required	Readings: Optional
The Basics	1	Causality and experimentation		
		Correlation versus causation A/B testing	MM Intro + Ch. 1 Canvas Articles	
	2	Regression		
		Linear regression review Empirical identification	MM Ch. 2	G&K Ch. 16-17, S&W Ch. 4-6
Working with market data	3	Causal effects with market data		
		"Control" groups in regression Difference-in-differences	MM Ch. 5 & 6	
	4	Incorporating Heterogeneity		
		Introduction to heterogeneity Heterogeneous regression		
Heterogeneous targeting	5	Heterogeneous response models I		
		Binary & discrete choice models Probabilities and expected values	Canvas Articles	Canvas Articles
	6	Heterogeneous response models II		
		Recommendation systems Regression discontinuity	Canvas Articles	
Consumer dynamics	7	Intertemporal dynamics		
		Purchase acceleration Rewards programs	Canvas Articles	Canvas Articles
	8	Dynamic pricing		
		Uber model Kayak model	Canvas Articles	
Fun stuff	9	Advanced topics		
		Auctions & dynamic bidding Consumer search	Canvas Articles	
	10	Final Projects		

General Teaching Philosophy

Because the primary objective of this course is to improve your thought process for tackling analytical questions, little emphasis is placed on memorization of substantive learnings; rather, emphasis is placed on developing your ability to think through data puzzles. This means that I will sometimes ask questions in class and on assignments that I don't necessarily expect you to be able to answer on a first pass, because the value of the exercise is in not encoding a specific answer to memory, but in the experience of trying to think through the problem and solve it. The homework assignments are therefore often lengthy and challenging.

However, my philosophy has always been to work as hard as I expect to students to work. To that end, I try to be available for meetings or to answer questions via e-mail as often as possible. I have had students meet with me every single week to go over the material from class. I certainly don't *expect* you to do that, but I want to make clear that I'm willing to help as much as possible. Note, however, that the students who have struggled with the course have often been those who did not have the time to get help, or felt uncomfortable asking for help. This course is more of a workshop than a class; to truly learn the material, you must be able and willing to work with your team and ask me for help when you struggle. The pre-class readings provide you with the foundation for what we cover in class, but don't cover everything. The text is effectively a pre-requisite to understand what is being taught in class. I'm pointing this out only to provide you with accurate expectations for the course, so that you're not caught off guard.

Changes from Fall 2017 offering

1) There is now a mandatory textbook.

Last year students mentioned that seeing material in class for the first time was challenging, and they would prefer to have some prep material. One challenge with this is that much of the material taught in this course is traditionally reserved for PhD-level courses (for reasons I don't fully understand), and the textbooks that might cover it are written in a manner you'd have difficulty reading if your only statistical background is 203A. However, I recently discovered a relatively new textbook that I think does a great job of communicating much of the core material from the first four weeks without a ton of equations or jargon. It's written by two prominent economics textbook authors whose other texts are regularly used in Masters- and PhD-level courses. I've also compiled some readings for topics covered later in the course.

2) R coding is now optional.

Each homework assignment will have an optional component for learning R. This component is extra credit, applied after the curve is set to raise your final letter grade. Off-line support for learning R will be provided (at office hours, via e-mail, and via documents on canvas). This change was made because the gap between students who wanted to learn R and those who didn't was too large to bridge in any other way, and I wanted to continue offering the opportunity to learn R without discouraging students from taking the course.

3) Anti-free-riding policy

In the past, I have offered teams the opportunity to fire team members that do not contribute, in order to prevent free-riders from taking advantage of their team members. The intention here was to **protect you** from others who might take a spot on your team and never show up, hoping you'll take the path of least resistance and do the work without them. Students have explained that there are social consequences to "firing" a team member. For this reason, I have made a few changes:

1) The first homework assignment is now an individual assignment.

2) Teams will be required to grade each member of the team at the end of the year.

3) If I determine a student has been free-riding, they will fail the course.

These policies are not aimed at students who may struggle with the material. **If you're still a little uncomfortable with statistics, do not let this discourage you**. You may work with others to figure out the answers to the individual assignment, but everyone must hand in their own submission. Additionally, I am always happy to help you with the assignments, and your classmates are generally happy to help those who make a genuine effort but struggle with the material. I'll do everything I can to help students succeed. This policy is literally only designed to stop the occasional (rare) student who might attempt to take advantage of you, by refusing to help with team assignments and hoping you will do all the work for them.

4) Changes to topics covered

Some of the more difficult assignments and topics (e.g., choice modeling) have been toned down. We'll spend more time on interpretation of linear regression results and problem/data structuring in the first two weeks to make sure no one feels left behind, and to make the work in later weeks easier. There are also only four homework assignments this year, down from five last year.

Grading & Assignment Schedule

Homework (60%): There will be four assignments, the first of which must be submitted individually. You may work on the individual assignment with others, but must submit your own answers in your own words. The last three assignments will be cases that you can submit in teams. Each is worth 15% of your grade, and each will have extra credit opportunities pertaining to coding in R. No late assignments are accepted.

Final Project (20%): Teams will also work on a quarter-long project of the choose-your-own-adventure variety, which will culminate in a 15-minute presentation on the last day of class. See below for details.

In-class participation and exercises (20%): Attendance and active participation in class exercises will be worth 20% of your grade. You can miss one class without penalty.

Final Project Proposal + Team Grade: You will be asked to submit a preliminary project proposal, and to grade each team member on their performance for your final project and homework via an online survey. You will not receive a grade for the course until you grade your team members' performance. Any team member reported to be free-riding will be investigated. **If I conclude the student was free-riding, they will fail the course**. You may also fire a team member for free-riding. Fired team members must work on their own.

Teams should consist of 3 students

Assignment Schedule (all sections)

Assignment	Due Date
Form project teams & choose topic (team)	Class 2 (end of class)
Preliminary project proposal (team)	Class 3 (8 am)
HW 1: Regression & A/B Testing (solo)	Class 3 (8 am)
HW 2: Market Data Case (team)	Class 5 (8 am)
HW 3: Choice Modeling Case (team)	Class 7 (8 am)
HW 4: Dynamics Case (team)	Class 9 (8 am)
Final Project Presentation (team)	Class 10 (PPT @ 8 am)

Final Project

You have two options for your final project.

Option 1: Collect data from a firm (maybe one you are affiliated with) and examine a problem of interest to them. The firm must be willing to let you present your results to the class.

Option 2: Research a topic in analytics and teach the class what you have learned.

Below are a list of six topics you may be interested in pursuing for option 2.

- Measuring peer influence
- Measuring word of mouth in social media
- Text mining social media
- Freemium products and network effects
- Purchase intention and uncertainty
- Missing data and data fusion

A set of readings for each topic will be available on canvas—you will read these and consolidate into a presentation, so your fellow students can learn about the topic as well. These readings will be difficult—you are NOT expected to perfectly understand the math. Focus on the data and what the data can tell you about the topic of interest; complicated models are ultimately just a way of structuring what the data tells us. You may propose other topics you find interesting if you wish, so long as it is about analytics.

Final Project Proposal: Teams must write up a two-page preliminary project proposal. Examples of previous proposals will be posted online. As long as the proposal is sent on time and reasonably well thought-out you'll get full credit. The point of the proposal is just to begin a conversation with me, so that I can guide you over the course of the quarter.

Relationship to other courses

There are several courses pertaining to marketing analytics (and business analytics more generally) at the GSM. The faculty make an effort to ensure each is its own unique offering, without much redundancy. Below I briefly identify some (but not all) of the courses someone interested in Marketing Analytics might take, how those courses relate to this course (and each other), and how they differ.

0. Primary Statistics Courses

203A – Analysis for Managers ("Stats 1")

203B – Forecasting and Managerial Research Methods ("Stats 2")

285 – Times Series Analysis and Forecasting ("Stats 3")

1. Methodologically-Focused Courses

249 – Marketing Research

269 – Business Intelligence Technologies and Data Mining

293 – Marketing Analytics

2. Objective-Focused Courses

234 – Pricing

239 – Digital Marketing

243 – Customer Relationship Management

248 – Marketing Strategies

282 – Supply Chain Management

I separate courses other than the primary statistics courses (203A, 203B, 285) into two groups: (1) courses that have a broader approach pertaining to statistical methodology, and (2) courses that teach a specific set of marketing analytics pertaining to a specific objective (or set of related objectives) in depth. The first set of courses are more general and teach you how to think like a data scientist, while the second set are more likely to be directly related to specific job functions or a specific class of decisions that marketing managers need to make. For example, Pricing is included in the second bucket, as it tackles pricing-related analytics in depth.

The Marketing Analytics course is in bucket (1), and is most closely related to Marketing Research and Business Intelligence Technologies and Data Mining. The three courses cover both statistical methodology and data, with Marketing Research focusing more on the collection and analysis of primary data and BIT&DM focusing more on analysis of secondary data (especially "big data"). Marketing Analytics spends less time on the collection of data than Marketing Research, and less time on statistical methodologies that are useful for big data problems than BIT&DM. Instead, Marketing Analytics focuses on (a) how to recognize what questions a dataset allows you to answer and what it does not, and (b) how to most accurately answer a question of interest with that dataset.

Finally, I think it's worth noting that irrespective of what my intentions for the courses are, students have in practice often used my course in one of two ways: Either as "Stats 1.5," taking it before they take 203B if they feel they want more experience with statistical methodology before jumping into the second stats course, or as "Stats 2.5" (or "Stats 4"), taking it after 203B (or after 285) to help them further develop their skill at relating data to statistical models. Either approach is fine. This course will be easier if you have taken 203B first, but 203B will (I assume) be easier if you have taken Marketing Analytics first. Consequently, having taken 203B is recommended but not required for this course.