293 – Marketing Analytics

Classroom Zoom Link: TBA

Course Description

There are two fundamental questions someone working in analytics must be able to answer.

"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 space, within the context of several different areas of marketing.

"What variation in your data answers your question?"

Variation in your data is key to answering any question. If you want to estimate how changes in price affect sales, you need to observe sales at multiple price points—that is, price must *vary* in your dataset. If you only ever observe one price, you cannot answer how price changing will affect sales. If counterfactuals are the conceptual foundation of this course (the "what"), empirical identification (the variation in your data that identifies the answer to a question) is the methodological foundation (the "how").

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.

Requirements

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

Difficulty Level

This is not an introductory course – it is an **intermediate-to-advanced** course that presumes solid 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 much experience with linear regression prior to the course, getting the required reading done early will be critical. Repeat reading may be necessary as well.

Course Topics

Module	Wk	Topics	Reading
Linear Regression Basics	1	Regression Basics	
		Methodology: Coefficient interpretation, model building	MM Ch. 2
		Application: **Pricing & Advertising**	G&K Ch. 16-17
	2	Heterogeneity	
		Methodology: Interaction and control variables	
		Application: **Social media analytics**	
Correlation vs Causation	3	Causality and experimentation	
		Methodology: Causation, experiments	MM Intro & Ch. 1
		Application: **Online A/B testing**	
	4	Causal inference with market data	
		Methodology: Difference-in-differences	MM Ch. 5 & 6
		Application: Various (focused more on methodology)	
Modeling Consumer Dynamics	5	Intertemporal dynamics	
		Methodology: Empirical identification, variable creation	
		Application: **Price promotion, reward programs**	
	6	Dynamic pricing	
		Methodology: Model-building & optimization	
		Application: **Ride-sharing & seat-booking platforms**	
Heterogeneity & Targeting	7	Dynamic auctions	
		Methodology: Auctions, experiments	
		Application: **Entertainment Venues**	
	8	Discrete choice models	
		Methodology: Logistic regression, logit choice model	
		Application: **Database marketing**	
Advanced Topics	9	Advanced topics	
		Guest Speakers	
		Possible topics: Big data analytics, healthcare analytics	
	10	Final Project Presentations	

We begin the course with the **foundations**: statistics (week 1), data manipulation and variable creation (week 2), and causal inference (week 3). In week 4, we combine all three foundations and learn how to accurately measure the causal effect of X on Y using market data.

In weeks 5 – 8, we build on the foundations from the first four weeks and explore more advanced forms of variable creation (e.g., modeling intertemporal consumer behavior), experimentation (e.g., running a dynamic bidding experiment), and statistical models (e.g., choice models).

Weeks 9 & 10 are elective weeks. You'll get to hear from industry connections on the analytics projects they are working on. You'll also learn from your classmates, who will be presenting their final projects.

Assignments & Grading

Assignments are always due at 8 am the day of class. The final case will be due during finals week, at 8 am on the day class is normally held. **Quizzes will be provided on Canvas 24 hours before the day of class** and must be completed before class begins.

Week #	Assignment Due	Type	Quiz?
1	*No assignment due*	*n/a*	*No*
2	Final Project: Topic Selection	Group	Yes
3	Homework 1: Regression Basics	Individual	*No*
4	Final Project: Proposal	Group	Yes
5	Homework 2: Steam A/B Testing & DiD	Individual	*No*
6	*No assignment due*	*n/a*	Yes
7	Case 1: Airline Dynamic Pricing	Group	*No*
8	Final Project: Preliminary Slide Deck	Group	*No*
9	*No assignment due*	*n/a*	*No*
10	Final Project: Presentation	Group	*No*
Finals	Case 2: Disneyland Targeted Marketing	Group	*No*

Assignments in this course are not like those in traditional stats courses. 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 about data. This means that I don't expect you to be able to answer all homework assignment questions perfectly on a first pass, because the value of the assignment is not retrieving a fact from memory, but the experience of trying to think through the problem.

Grading

********** **Teams should consist of 2-3 students** **********

Individual Problem Sets (20%): There will be **two problem sets which must be submitted individually**. You may work on the individual assignment with others, but must submit your own answers in your own words. Each is worth 10% of your grade, and will have extra credit opportunities pertaining to coding in R.

Three Quizzes (20%): There will be quizzes during weeks 2, 4, and 6. In total they will be worth 20% of your grade.

Team Cases (30%): There will be two team case assignments, each worth 15% of your grade.

Final Project (15%): 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.

In-class participation and exercises (15%): Attendance and active participation in class exercises will be worth 15% of your grade. You can miss one meeting without penalty (for part-time students, this means two classes on a given day). Missing part of more than three lectures (3 days for full-time students, 1.5 days for part-time students) without approval will result in a loss of all participation points. < This is almost never an issue, I'm just stating it here as a precaution.

Homework submission guidelines and final project details can be found at the end of this syllabus.

********** **There is no midterm or final exam** **********

Relationship of Marketing Analytics 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 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.

Course Policies

Anti-free-riding policy

The intention of this policy is to prevent free-riders from taking advantage of their team members. I want 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. The policy is as follows:

1) Teams may fire team members if they are not pulling their weight. Fired team members must work alone on all assignments.

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 (e.g., if they weren't fired because they worked on the first two team assignments but then bailed on the final project, hoping it was too late to fire them), **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**. 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.

Laptop policy

Laptops will be needed for some in-class exercises. However, when not doing those exercises, laptops must be closed. In the past I have had students request to have them open to take notes, but observational data tells me that laptops are more often a distraction than a note-taking device.

UC Davis and GSM Policies

Notice of the Code of Academic Conduct

Students are expected to conform with the code of academic conduct, which can be found here:

http://sja.ucdavis.edu/files/cac.pdf

Academic Affairs, as well as the Academic Director of the MBA Program, will be notified of any violations, and will take appropriate action.

Cross-Attendance Policy

Intention to take a make-up course in a location other than the one in which you are enrolled now follows a specific procedure. A request to attend must be made to your professor, with a justification provided.

If this request is made at least one week in advance of the section you wish to attend, you should CC the Instructional Operations contact for the location and get both their and the professor's permission. The contact for San Ramon is Sergio; for Sacramento it is Dedan. No Instructional Operations contact need be contacted for make-up courses in Davis, only your professor.

If the request is made less than one week in advance of the section you wish to attend, you must also CC the Senior Assistant Dean of Student Affairs, who must give first approval before the professor and Instructional Operations contact can approve.

Final Project

I. Content

You have two options for your final project's content.

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.

II. Presentation Format

You will have 15 minutes to present your final project. No more than 3 minutes should be spent on set-up/conclusion. Depending on what you're doing, your presentation format may vary, but a generally acceptable outline of a presentation looks like this:

[2 mins]	Introduction to question/problem of interest
[4 mins]	Introduction to dataset
[4 mins]	Description of how the data can be used to answer the question of interest. I.e., "What variation in the data answers the question?"
[4 mins]	Description of the statistical model used, how it answers the question, and what was found.
[1 min]	Conclusion

For some presentations covering academic papers, the statistical model will be too complicated for students to explain. For such presentations, place more emphasis on explaining how the data can answer the question of interest.

III. Intermediate Checkpoints

There will be three intermediate checkpoints for the final project. The first is merely you selecting your topic of interest, the second is your proposal, and the third is your preliminary slide deck (see the "Assignments" section of the syllabus for due dates). These checkpoints are designed to help you pace yourself, and to help me ensure I can provide support if you struggle (before it's too late). Expectations for each checkpoint will be posted on Canvas.

IV. Grading

The intention of the final project is to give you an opportunity to learn a little about at topic that is of interest to you, and then teach the rest of the class what you have learned. It is, in effect, a miniature independent study that feeds into a variety-filled lecture. Given that the intention of the project is not only for you to learn, but also for your classmates to learn, you will be graded on whether your presentation provides a meaningful learning opportunity for your classmates.

"Meaningful" here implies both that you should have a non-trivial amount of analytics content (e.g., don't just discuss the importance of branding) and that you've communicated that content clearly (don't just throw the results of an analyses up without explaining what you've done).

With that said, don't try to over-do it with content. Trying to present 20 slides in fifteen minutes will almost certainly leave you feeling rushed; it's better to present 15 minutes of material well than 20 minutes of material poorly. Not all students need to present, as long as everyone contributes to the project in some fashion.

Finally, your grade for your final presentation will be scaled based on your performance on the intermediate checkpoints. If you kept up with all of the checkpoints sufficiently, you will not lose points. If you did not, your final presentation grade will be docked 5% per checkpoint missed (or for which submissions were insufficient).

Homework Submission Guidelines

Report Structure:

Each homework assignment should be submitted as a report to the fictional individual tasking you with your analyses, per the homework prompt. Do not merely submit a list of numerical answers. Do not assume the grader will "know what you mean" because they have an answer key. Most of the points in an assignment will be based on how well you've explained what the numerical answers mean, and how you arrived at them. You should never submit something on canvas that you wouldn't be proud to submit to a recruiter or boss.

Submission Format:

All reports should be written in Word or similar software and saved as a PDF.

The text of this PDF should be no longer than four pages long.

Tables and figures should be included in an appendix at the end of the document (e.g., Excel output).

Extra Credit in R:

Write your code in R studio, save it, and send it (plus the source data file, in case you've made adjustments to the original file I sent you) to me zipped together so I can ensure it runs and generates the appropriate output.