Columbia Business School
AT THE VERY CENTER OF BUSINESS™

MRKTB8633: Data Science for Marketing Managers
Spring 2023 A term, .5 Credits

Mondays 6:15 – 9:15 PM, Kravis TBD

Shawndra Hill	Course Assistant
Professor Office Location: 743	**TBD**
E-mail: sbh2146@columbia.edu	E-mail: TBD
Office Hours: 5pm Mondays or by appointment	

Communications from professor and teaching assistants about the course will take place through Canvas.
Students should make sure they regularly check for announcements and messaging notifications.

---

1. Course Overview

This course is for students who want to learn how to manage data scientists/data science projects in Marketing. This course connects real-world data on consumers and firms to decision-making and marketing management. The course will cover many real-world data driven marketing examples to illustrate applications of methods used in data science. The use of real-world examples and cases places these techniques in context and teaches students how to avoid the common pitfalls of data science management, emphasizing the proper application of data science techniques and pipelines. In addition, the course focusses on the unique requirements for managing data science teams and projects. This course covers the considerations that go into starting and completing successful data science projects in both small and large firms, focusing on the data science process of deploying models.

The goal of this course is three-fold. After taking this course you should:

Approach marketing and advertising problems data-analytically. Think carefully & systematically about whether & how data can improve business performance.

Be able to interact competently on the topic of data science. Know the basics of data science processes, algorithms, & systems well enough to interact with CTOs, expert data scientists, and business analysts. Be able to envision data-science opportunities.

Be able to manage data science projects in marketing: Learn how to build a strong team by understanding the different roles needed to support both large- and small-scale projects. Learn how to mitigate risks in data science product delivery and deliver impact.

Given recent advances in Large Language Models and Generative AI, we will sprinkle new and innovative applications to Marketing throughout the one week course.

It is important to note that while the use cases and motivating examples will be from **Marketing**, this course should be

relevant to anyone interested in managing data science projects and teams.

2. Instruction Method

Lectures

This is primarily a lecture-based course, but student participation is an essential part of the learning process. The course will explain with real-world examples the inner workings and uses of various data science techniques used in Marketing and Advertising. The primary emphasis is on understanding the various types of data science techniques, how to evaluate their results, and when and how to use them, and secondarily on the mechanics of how they work. Once a basic understanding of data science is established, we will focus on managing data science projects including building a team, what goes in to making go/no go decisions on projects, and investments in onsite/offsite infrastructure and data.

Quiz and Class Participation (25%)

Each class session, except the first, has materials you must read prior to class. We will have a very short quiz before (first 5-10 minutes) each class based on the readings. The quiz will open the night before class. The quiz grades will factor into your class participation grade. In addition, there will be opportunities to comment on content (questions, articles, etc.) on Canvas. Please pay attention (daily) to discussions.

Class participation is essential in order for you to get the maximum benefit from the course. Your grade will be based on attendance, effort and the content of your contributions to the class discussions. The evaluation of content will be based on the following: • Relevance: Are your comments clearly related to the readings and what is being discussed in class and to the comments of others? Do they move the discussion forward? • Fact-Based: Have you used specific data from readings, or from personal experience to support the assertions that you are making? If you feel that you are preparing well but that I am not calling on you enough, please let me know so that I can address the problem. In addition to voluntary participation, I often randomly select a few students to open and/or lead the case discussion. If you know things about the class topic, from previous experience, this can be very useful in class discussions. Please let me know about it in advance, and I will make a point to call on you at the appropriate time.

Learning by Assignments (25%)

There will be a total of 4 question assignments, each comprising a (multi-part) question. In addition, the assignments may involve hands-on work. You must turn in *all* question answers on the dates they are due. They will be graded and returned promptly. One of the assignments will be slightly longer to cover the crash course in data science material. From time to time there will be extra credit available for assignments for students who want to go deeper.

Data Science Project (40%)

As your final project, you will be required to complete a novel data science project proposal (~3 pages with presentation slides and links to supporting documents) to address a pressing issue or problem in the Marketing and Advertising industry right now for a company of your choosing. You will first pick a firm that is known to collect a significant amount of user or business data for Marketing and/or Advertising and then solve a Marketing Problem for that firm. You may pick your own, current firm. Some examples of Marketing Problems firms are facing today are: Cross Channel Advertising

Effectiveness Measurement, Signal Resilience for Target Marketing due to new Regulations (CCPA, GDPR), Utilizing Social Network Data for Ad Targeting. You will be given detailed guidelines for the structure of your final project on week 2. Your instructor will help you with some project area ideas to solve if needed, though you are encouraged to choose your own, which will need approval. You will work in groups of size 3-5 students (Depending on Class Size). We will need to keep the number of projects for the class to 10.

Data Science Peer Project Review (10%)

You will be given 2 of your class projects to review from a quality, feasibility, (legal), and privacy perspective as a take home assignment. Your ability to critique data mining project proposals (as a manager) is one of the most important skills to take away from the course. This assignment will be done after final projects are handed in and you may opt out of this assignment to take a 10% loss on your overall grade.

3. Requirements and Grading

This is a lecture-style course, however student participation is extremely important. Students are required to be prepared and read the material or watch assigned videos **before** class. **Students are required to attend all sessions and discuss with the instructor any absence from class. 1 or more absences will result in a 0 for class participation (25% of your grade).** Attendance is important because every class builds on the last. In addition, if you missed the first class or two, you are still responsible for getting assignments in.

As discussed above, you will hand-in 4 (individual) write-ups to questions that will be assigned in class and will be posted on the class Canvas site. Answers should be well thought out but concise. The assignments will be based on the current day's readings and lecture and in some cases provide exercises to help you prepare for your final project. Points will be deducted for sloppy language and irrelevant discussion. You will submit all assignments on Canvas.

Note that you must put your name on your assignments to receive credit.

You will be assigned a longer assignment to cover the crash course in data science material. This assignment will be worth 33% of your assignment grade.

Late assignments

Turn in your assignments early if there is any uncertainty about your ability to turn it in on the due dates. **Assignments up to 1 day late will have their grade reduced by 50%.** After one day, late assignments will receive no credit (**no exceptions**).

There will be one team project (teams of 3-5 students) in which students will address business problems with data mining techniques. Students will hand in a pre-proposal (15% of project grade), a final writeup (accounts for 50% of project grade) and prepare a short class presentation of their work (30% of project grade). Students will also prepare a contribution report that details the contribution of team members including themselves (5%). A class discussion will follow the presentations. Details of the requirements for the project will be discussed the second day of class.

There will not be a final exam at the end of the course. The main deliverable in the class is the Data Science Project, including presentation.

The grade breakdown is as follows: 1. Pre-class Assignment Questions (4) including crash course in data science: 25 points 3. Data Science Project: 40 points 4. Class Participation (And Quizzes) 25 points. Peer Project Reviews 10 points.

Group Assignments: You may form you own group on Day 1. We will provide instructions for how to submit your group members on Canvas. Professors will assign groups for students not already in a group by Day 2.

Cell phone usage is not permitted in class. It is preferred that you use your CBS issued iPad to take notes and follow along. The lessons will be taught in person.

4. Teaching Materials

The following are reading materials for this course:

1. [Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking](#) Tom Fawcett and Foster Provost ([free same day in NY](#) if you haven't purchased already)

2. Supplemental readings will be provided as the class progresses.

*The required readings are subject to change based on student interest.

**The instructor will also post short videos from time to time covering real world examples of Data Science managers

Note there is no comprehensive book on the subject of this course. Therefore, the course readings will rely on a mix of publicly available research and practitioner content on the web.

Other optional reading

1. Managing Your Data Science Projects: Learn Salesmanship, Presentation, and Maintenance of Completed Models Robert de Graaf
2. Ethics and Data Science: https://www.oreilly.com/library/view/ethics-and-data/9781492043898/

3. **Course Schedule**

Session	Topic(s)	Readings	Due
1(A) Jan 22	**Fundamentals of Data Science for Marketing Part I:** Focus will be on K-means clustering and classification trees as illustrative examples of unsupervised and supervised learning. We will also discuss data, models, Applications, Big and Small Challenges	**Required** [Data Science in Marketing: A Comprehensive Guide (With Examples) - NoGood™: Growth Marketing Agency](#) [What Is MarTech and AdTech - And What's The Difference? | LinkedIn](#)	Assignment 0 (Personal Info Survey) +Team Members (3-5 students) Submitted by end of day

		[Customer Segmentation using Data Science (franciscojavierarceo.github.io)](#)

[What is Supervised Learning? | IBM](#)

[How Can Marketers Use ChatGPT? Here Are the Top 11 Uses. | Entrepreneur](#)

[A Beginner's Guide to Large Language Models | by Digitate | May, 2023 | Medium](#)

Background (Optional but strongly encouraged):
Data Science for Business
Chapters 2, 3, 6 | |
| 1 (B)
Jan 22 | (cont,) supervised/unsupervised learning and appropriate evaluation methods are used in Marketing

Applications will include: Audience Analysis, Target Marketing, Lead Scoring, Identity Resolution, Customer Personas/Segmentation, Text/Sentiment Analysis, Ad Ranking, Advertising Optimization | | |
| 2 (A)
Jan 29 | **Fundamentals of Data Science for Marketing Part II:** We will cover evaluation and what it takes to deploy models in the real world | [ROC Graphs: Notes and Practical Considerations for Data Mining Researchers (hp.com)](#)

Background (Optional but strongly encouraged):
Data Science for Business
Chapters 8 | Assignment 1 |
| 2 (B)
Jan 29 | **Fundamentals of Data Science for Marketing Part II:** We will cover field experiments and causal | [Run Field Experiments to Make Sense of Your Big Data (hbr.org)](#)

[A/B testing - Optimizely](#) | Assign students to groups who haven't |

	inference techniques used in Marketing and Advertising Applications will include: Brand Lift Measurement and others		already built a team
3 (A) Feb 5	**Connecting Data Science for Marketing with Project and Performance Management:** Data Science Process, Timelines and Milestones, Aligning on strategy, effective Presentation of Ideas	What Is OKR? The Goal-Setting System To Scale Up Your Business FourWeekMBA Ally.io	OKR Examples What Matters: Free OKR software & tools: Best ways to track goal setting -- How to Use OKRs for your AI Team	Assignment 2 +Pre-proposal Check in (Optional – can run idea by Prof)
3 (B) Feb 5	**Survey Design and Management (some time with your team)**		
4 (A) Feb 12	**Connecting Data Science for Marketing with Product and Company Impact:** Product Impact. Social Impact, Decision Making Impact, Understand Work	**Background (Optional but strongly encouraged):** **Data Science for Business** Chapters 13, 14, Appendix A	Assignment 3
4 (B) Feb 12	**Privacy, Policy, Ethics**	--[1803.09010] Datasheets for Datasets (arxiv.org) --What is a Walled Garden? And why it is the strategy of Google, Facebook and Amazon Ads platform?	by Pierre de Poulpiquet	mediarithmics what is?	Medium --The Ethics of Big Data	Socialnomics -- https://hannawallach.medium.com/big-data-machine-learning-and-the-social-sciences-927a8e20460d	
5 (A) Feb 19	**Building a Data Science Team:** Roles, Domain Experts, Complementary Skills, Interviewing, Onboarding, Paying Third Parties, Common Motivations of Data Science Team	Building Data Science Teams (datascienceassn.org) Building up a Data Science Team from Scratch	by Amadeus Magrabi	commercetools tech	Assignment 4

	Members, Supporting the Larger Organization **Communicating Results and Forming Partnerships from the Start**		
Feb 26 presentation in class	Proposal and Proposal Presentation Due		-- Proposal and Proposal Presentation (proposals TBD)
TBD	Final Projects Due 3 page write up and Presentation Deck (video optional)		Due TBD
TBD	Project Peer Reviews Due (Individual assignment)		(due on TBD)

CLASSROOM CULTURE

This course adheres to Columbia Core Culture. Therefore, you are expected to be present, prepared, and participating. This means: • Being on time and present for all parts of every session (attendance will be tracked) o Students who miss 1 or more sessions will receive 0% for class participation in the course. Students who miss 50%+ of classes (unexcused absences) will receive an F in the course. • Completing the readings and assignments before class • Being ready to respond to cold-calling • Participating in a meaningful way is expected and part of your grade • Being respectful of the professor and your fellow students (adhering to the Code of Conduct in the classroom)

NOTE:

- This version of the syllabus is as of November 14, 2023. Details may change with time.

Preparation Questions (be prepared to discuss these questions in class)

Will be updated by first day of class

Day 1

What are some examples of user data used in Marketing problems?

What are some examples of Data Driven Marketing problems?

What is supervised learning?

What is unsupervised learning?

What are different terms for dependent variable in prediction models?

Day 2

How do decision tress work?

How does K-means clustering work?

What are ROC curves and how are they used to evaluate classifier performance?

Day 3

Day 4

Day 5