Social Media and Web Analytics

Lachlan Deer

Block 4 (Spring) 2021

E-mail: tisem.social.media@gmail.com
Web: tisem-digital-marketing.github.io/2021-smwa/
Learning Mode: Online
Dates: April 12 - June 4, 2021

Last Updated: 01 April, 2021

Motivation

"Some 73% of online adults now use a social networking site of some kind... [and] 42% of online adults now use multiple social networking sites," according to a recent Pew Research survey.[1] If a business needs to be where its customers are, then establishing a strong presence on social media is a critical component of a modern marketing strategy. The goal of this course is to develop an understanding of the social media landscape from the point of view of a marketing analyst. The course content will be unashamedly analytics heavy - utilizing modern statistical and mathematical modelling techniques to deliver managerially relevant conclusions from social media data To that end, we need to be able to deliver data-driven answers the following questions: How does one develop or maintain an online reputation or develop their brand online? How does one maintain a strong user community on social networking sites? How can one craft a potent social advertising strategy? What are the potential pitfalls? How does one quantitatively assess whether their online reputation is positive and/or improving, or whether their social marketing is effective?

To answer these questions, we must synthesize knowledge from a variety of disciplines. Sociologists have researched such concepts as homophily and social capital, which point to the underlying motivations in the relationships of individuals in a social network, computer scientists have uncovered surprising patterns in real-world social networks, and marketing experts have studied how network structure influences consumer demand, what makes ads go 'viral' and aspects of social advertising. This course will combine foundational material with examples from the social world to build a data-driven understanding of modern social media marketing.

The course is split into four modules:

1. **Network Patterns:** which seeks to describes and explain several common patterns found in real-world social networks,
2. **Branding and Community:** which explores the best practice methods for maintaining a strong brand online and managing the user community,
3. **Importance and Influence:** which discusses an individual's place in the network, and how memes, early adoption, word of mouth and other 'information cascades' propagate,

---

[1] Pew Research, 2019, Social Media Fact Sheet.

4. **Advertising and Marketing:** which focuses on viral marketing and social advertising techniques.

Target Audience

This course is designed for students in the Master of Marketing Analytics and Master of Marketing Management who have a strong interest in social media marketing and want to improve their quantitative and analytical skills.

In particular, I expect the typical student to have the following characteristics:

- Strong interest in understanding how social media and social networks effect consumer behavior, brand reputation and the actions of marketers.
- Strong interest in learning and applying statistical analysis and data science techniques to social media data, although with limited existing background (see Pre-reqs below).
- Strong interest in using quantitative results to develop management or marketing insights.
- Interest in learning how to write computer code and acquire "best practice" methods.
- Interest in learning new quantitative modelling techniques.
- Ability to work collaboratively.

Course Objectives

This course has five closely intertwined objectives:

1. Develop awareness of the challenges and opportunities social media and social networks present marketers.
2. Learning to use quantitative and statistical analysis to develop strategies for managers and marketers on social media platforms.
3. Understanding which modelling or statistical technique is best suited to a particular problem.
4. Develop the ability to understand the limitations of one's analysis.
5. Acquire necessary skills to work collaboratively in teams on analytical marketing problems.

Learning objectives for specific classes will be provided within the Course Notes.

Pedagogy

The COVID-19 pandemic has pushed the class online for this academic year.

The course will involve video lectures, self-guided tutorials, live discussions sessions and hands on assignments. The course strives to balance the practical aspects of social media analytics with a theoretical understanding of the concepts and frameworks involved. Each topic covered has a lecture/discussion portion as well as a practicum/exercise/demo portion. There are course projects in the form of group assignments and an individual final project where you will work on the developing your analytical toolkit and framing results around answering marketing questions.

The course structure takes advantage of an asynchronous format. This format lets you work through most of the course material in your own time, and aims to minimizes 'Zoom fatigue'. Each week, you will be expected to complete the following tasks:

- Read through the required readings
- Watch video lecture clips
- Work through the self-guided analytics tutorial
- Work collaboratively on group assignments (as needed)
- Work individually towards the course project's milestones.

Lecture material will be split into multiple (short) video clips. This allows you to work through topics at your own pace, and revisit specific sections multiple times relatively easily.

In addition, there will be a two "live" discussion sections each week, beginning in the second week of the class. These sections will run on the Zoom meeting platform, and links will be distributed on the course website. The format is as follows:

- Each session will last 60 - 75 minutes
- The schedule for discussion sections are available on the course website.
  - The numerous public holidays in Block 4 means that the section times move around week-to-week
- Each session will discuss 2-3 topics, topics *will differ between sessions*
  - Topics covered in week x will be based on week $x - 1$'s course material
  - Specific topics will be decided based on surveying student's interests each week, and will be announced in advance
- Instructors will aim to facilitate a discussion between students, rather than instructors leading the discussion
  - Instructors can end a session at their discretion if they feel students are not engaging in the discussion
- Discussion sections will *not* be recorded
- Attendance is optional

Because we are adapting this course to an online format, we reserve the right to change the design or content of these live sessions, depending on our experiences and your feedback.

All the videos, documents and (where appropriate) weblinks will be made available on the course website and via GitHub.

Prerequisites

To do well in this course it is suggested that you have passed the course Introduction to Marketing Research.

In addition, students will benefit from having completed other marketing courses (such as Customer Analytics and Price and Revenue Analytics) and other classes at TiSEM that deal with data and computing (i.e. the 'skills classes'). These are, however, not required.

Schedule

The following is a tentative course schedule. Check back regularly for updates.

Week	Marketing Topic	Analytics Topic	Due Dates
0	Course Introduction	Using Git and R	
1	Social Media Networks	Visualizing Network Data	
2	Online Reputation	Regression Review	
3	Community Structure	Text Analytics I	Project Milestone 1
4	Word of Mouth	Endogeneity: Causes & Solutions	Assignment 1
5	Influencers	Simulating Network Evolution	Project Milestone 2
6	Viral Marketing	Text Analytics II	
7	Social Advertising	TBA	Assignment 2 & Final Project

Readings

We will combine material from two books and recent academic papers. All readings from books are optional.

Books

The course uses readings from the following books:

- *Networks, Crowds and Markets: Reasoning About a Highly Connected World*, by David Easley and Jon Kleinberg
- *Networks: An Introduction* by Mark Newman

We will provide specific details about which chapters to review in advance of each week on the course website.

Papers

A list of required and optional readings from the academic literature will be provided in advance of each week on the course website.

Computing

This course is an analytics course, and as such we aim to use, understand and develop computational routines that help managers and marketers make better decisions. Large portions of the course will be *'hands-on'* and we will be using the statistical software `R` to demonstrate and apply methods and concepts in class.

All of the software requirements for this course are open-source and/or free. Please aim to have everything installed by the start of our first lecture.

We will be available for installation troubleshooting during the first week of the course. You can sign up for to get installation help via our office hour sign-up links (see below).

R & RStudio

You will be required to install `R` and `RStudio` to complete the Labs and group assignments. Installation instructions are available in the Installation Guide (they differ slightly by operating system):

After installing R and RStudio I suggest you get acquainted with the interface and basic data analysis commands before the first class. To facilitate the introduction, we've suggested some DataCamp modules for you to work through. The suggested modules can be found on the course webpage under 'Course Preparation'. Please look over these *before* the first class.

Improving Reproduciblity of our Analysis with Git and GitHub

In addition to helping you gain analytics skills using `R`, I want to help you improve the reproducibility of your computational analysis - i.e. I want you to document your code and the changes you make to it. To help achieve this goal we will use a version control software called `Git` and we will post our codes on a repository called `GitHub`.

- To install Git and create a Github account, follow the instructions in the Installation Guide.
- Also register for an education discount to get unlimited private repositories here.[2]

---

[2]GitHub recently announced unlimited free private repos for everyone. However, you are limited to three collaborators per private repo, so the education discount still makes sense.

As part of your preparation for the class, look over the DataCamp modules on "Version Control with Git" that we have highlighted on the course webpage under 'Course Preparation'. Please look over the material *before* the first class.

GitHub Classroom

I am going to run the course through GitHub Classroom. You will receive an email invitation to the course repo with instructions in due time, but for now it suffices to say that this is how we'll submit assignments and provide feedback, etc. Canvas will only be used to submit weekly quiz solutions, distribute grades, and registration of teams for group assignments.

Evaluation

Weekly Quizzes (5 percent)

Each week there will be a quiz that recaps the course material from the current week. You must complete the quizzes by 5pm Central European Time on each Friday. We will allow you to re-try the quiz each week up to *three* times. We use your 5 best completed weekly quiz grades to compute your final grade. There is absolutely no make up for these quizzes.

Homework Assignments (40 percent)

There will be 2 group assignments worth equal weight. To pass the assignment, each group member needs to have made a substantial contribution to the project. The finished reports (& code where appropriate) must be submitted via GitHub no later than midnight of the due date (i.e. before 23:59 on the due date).

Group assignments **must** be done in groups. Students will be **randomly assigned** to groups before each assignment is released, i.e. the groups will change between assignments. Each group will consist of 4 students. Working with other student's you are unfamiliar with is an important skill to acquire, and mimics aspects of team work in the *real world*.

Further information on group assignments will be distributed closer to their issue.

Course Project (55 percent)

You must develop a project on a topic related to social media analytics that *has marketing implications* (broadly defined) that interests you. The final submission of the report has three features: (1) a written report of a max. 7 pages, (2) the underlying code & data to produce the report, and (3) a 5 minute video presentation of your research.

The project should have the following characteristics:

- There should be a clear motivation. You should be trying to solve a question that matters, either to society or business.
- You should do a thorough survey of the relevant literature. This includes articles from the popular press as well as from scholarly sources such as academic journals.
- Your report and video presentation should briefly organize this literature survey, and summarize your main insights and findings.

To assist you design your project with these characteristics, I have broken the project into three milestones:

- Milestone 1: Proposal and Data Collection
- Milestone 2: Literature review, and a motivating figure or table using the data, and

- Milestone 3: the final report, video presentation and underlying code and data.

Each milestone has different due dates *and* contributes to your final grade.

The final grade for your project is computed as:

$$\text{Project Grade} = 0.20 \times \text{Milestone 1} + 0.20 \times \text{Milestone 2} + 0.60 \times \text{Milestone 3}$$

The written part of each milestone must be based on the templates we will make available. These templates are designed to help you structure your submission to include the relevant information and do so in a concise manner.

Late Submission

Assignments and Projects must be submitted before the deadline. If you submit one assignment late, you will be penalized by 25 percent per calendar day. This penalty applies to an entire group where applicable. After the first late submission, you will get zero points on all subsequent late submissions.

Office Hours

We are here to help you through the course & want you meet your expectations and personal learning goals. To that end, the instructors have scheduled consultation hours each week for you to meet us and discuss aspects of the lecture material, lab sessions and assignments that you are struggling with. The office hours are:

Name:	Lachlan Deer	Hendrik de With
Office Location:	Zoom	Zoom
Consultation Hours:	Thursday, 10am - 12.30pm	Wednesday, 2pm - 4pm
Link to Book Slot:	Click Here	Click Here

If you plan on coming to our scheduled hours, please make a booking using the links provided. If you cannot make the scheduled office hours, talk to us on Slack and propose three times to meet and we will try and make it work. Office Hours are subject to a "fair and equitable use" policy.

Notes for scheduling with Lachlan:

- After you book a time slot, a calendar invite will be sent to your email automatically. You *must* accept this invite. Approx. 1 hour before our scheduled meeting a reminder with a Zoom link will be emailed to you.

Notes for scheduling with Hendrik:

- On the booking page, a link to Hendrik's Zoom meeting room is included. Copy down this link so you can join the scheduled meeting.

We expect you to come to your schedule slot prepared and having clear questions in mind that you want to ask. If either of us feel you have come unprepared, we reserve the right to end the appointment and ask you to come back at a future time when you are more prepared.

Rules of the Game

I have clear expectations of how we should all interact in class and in meetings. They are summarized by the "7 Be's" below:

Be Active

Learning via passive listening is an ineffective learning technique. Come to discussion classes prepared to ask questions, and to answer questions posed by me and others in class. This will lead to a much more enjoyable learning experience for us all. I would much rather have an engaging discussion on topics where we share mutual interest and cover less material than have Discussion Sections become a monologue.

Be "Hands-on"

The class is designed to be 'hands-on' in the sense that you will be expected to complete a substantial amount of coding exercises. We strongly believe the only way to learn programming is to do programming. Try to complete each 'Lab Session' be prepared to ask and answer questions about it during Discussion Sections. Codes that solve exercises will be posted *after* Discussion Sections.

Be Prepared

We expect you to dedicate an average of 20 hours per week of study time to this course. Use this time to look over required readings, re-examine course notes and codes and work on your assignments. Looking over required readings means that our discussions will have more substance, and (ideally) be more insightful.

Quick tip: If you find yourself not understanding part of the course content or not able to find bugs in your code, take a break and come back to it in an hour or so. Come talk to us if it remains unclear.

Be Professional

I expect you to treat us and your class members collegially. In addition, we have a set of Social Rules and a Code of Conduct that provides further details of acceptable and unacceptable behavior.

When writing an email, being professional also means using a meaningful subject line, identifying yourself (i.e. Regards, Lachlan at the end of the email), and keeping the message focused. We aim to reply to emails within 48 hours Monday - Friday, but there might be a slight delay if you write to us late Friday evening.

I strongly encourage you to provide feedback about both format and content during the course. We will provide survey links to collect this feedback at regular intervals. Be professional when filling these out.

Be Informal (but not too informal!)

Do not address us with "Mr.", "Dr.", or "Prof.". Please use our first name's (i.e. Lachlan and Hendrik). More formality than that makes me feel uncomfortable.

Also you can to add us on LinkedIn or Twitter - where we maintain professional accounts, but not Facebook.

Be International

This course is taught in English. You are obliged to speak English whenever there are international students present and all Slack messages on public channels must be in English. If in doubt, you need to communicate in English. As someone who has studied in in countries where I do not speak the native language - be assured it is *very much appreciated.*

Be Healthy (Physically & Mentally)

Physical and Mental health are important, and definitely more important than your coursework. Material posted online, the readings and talking with your classmates about what you missed is more than enough if you miss one or two classes.

On the mental health side, I know how challenging graduate study can be, especially when paired with one's 'non-study life.' *I strongly encourage you to prioritize mental health over **everything** else.* I'm happy to talk with you individually and share some common experiences and strategies, but I am not a professional. If you feel like you need to talk to someone, a good starting point might be Tilburg's Education Coordinators or the student psychologists at the University.

Be Fair & Honorable

In group assignments we expect all group members to actively and evenly contribute. If you are struggling with parts of a group assignment you are working on, talk to your group members and to us rather than withdrawing from conversations with your colleagues.

Being fair & honorable also means that you should not plagiarize others work. We stick closely with the School's Fraud and plagiarism policies which are outlined here.

FAQs

R looks cool, but I'm more familiar with Python/Julia/MatLab/SPSS/etc. Can I use that instead?

Short answer: No.

Longer answer: Look, I like and use a lot of those languages too, but I'm not changing my lecture notes or assignment templates for you. Plus, I really do think that R makes the most sense for quantitative marketing students looking to develop their data science skills. It already has all of the statistics and econometrics support, and is amazingly adaptable as a "glue" language to other programming languages and APIs. Learning multiple languages is never a bad idea in the long run, though.

I already have a BitBucket/GitLab/etc. account. Do I still have to use GitHub?

Since I'm running this course through GitHub Classroom, yes. But good for you! (Seriously... those are great platforms too and as an open-source advocate, I fully support a plurality of tools and software options.)

On that note, do you have any advice for running a course on GitHub Classroom?

I mostly followed this excellent tutorial by Jacob Fiksel. (Hat Tip to Grant McDermott at U Oregon for recommending this!)

This course looks interesting! Can I use/adapt your lecture notes for a similar course that I'm teaching at XYZ?

Sure. I've benefited greatly from other people making their teaching materials publicly available (and have tried my best to acknowledge them directly in the relevant sections of this course). I'm more than happy to pay it forward. I only ask two small favours. 1) Please let me know (email/Twitter) if you do use material from this course, or have found it useful in other ways. 2) A minor acknowledgment somewhere in your own syllabus or notes would be much appreciated.