Social Media and Web Analytics

Course Introduction

Lachlan Deer
Tilburg University
Updated: 2022-04-07


Goedendag!

Learning Goals for this Class

- Understand the logistical structure of the course
- Meet your Instructors
- Develop an understanding what's needed in a Social Media Analytics toolkit
- Understand the topic coverage in this course
- Explain how you will be assessed

Logistics

Class Structure

- Hybrid-ish, *Mainly in person*
- **Lectures**
  - Recorded Lectures - Discussion of papers & marketing concepts
    - *Not* in-person
  - Live Computer Lectures - Methods intros, biweekly
    - *In person, on campus*
- **Lab Sections** (*on campus*)
  - Week 1: Intro to Networks & Getting Data via Twitter API
  - From Week 2:
    - Group Presentations (2×30 mins)
    - Review of Previous Lab Assignment (30 mins)

Where to Find Stuff

- Course website is your # 1 resource
  - Canvas only used for group allocation, grades
- Discussion / Chat: Slack
  - Group Chat across different 'channels'
- Email (if you must): tisem.social.media@gmail.com
  - *Not* our personal email addresses
- Office Hours
  - Sign up using links
  - Check the Syllabus for instructions

Textbooks, Readings

- **No Textbook** for the class
  - Become outdated fast as social media changes rapidly
  - Some emphasise too much of the hype, remainder are too dry
  - I do like Aral's *'Hype Machine'* though ... but its audience is not quite this class
- **My job**: filter through the sea of information and teach what is useful and relevant
  - No hype, hopefully not too dry
- **Your job**: read over assigned papers, learn the skills
  - I've curated course content to (hopefully) deliver the most insight with the least technical details
  - Readings shouldn't be 'hype', occasionally dry
  - Coding skills acquired should be easy to transfer to a different problem
- Lecture slides will be on course website as PDFs

Coding, Maths and Stats

This is an **analytics** class (it's even in the course name!)

- We will develop an analytics toolkit
    - Mainly 'code based', occasionally we'll need some maths and stats
    - ... I've tried to minimise unnecessary maths / stats
- Coding backgrounds among students is heterogeneous
    - Principle: Leave no-one behind
    - Instructors are here to help catch you up and provide further guidance as needed
    - $\Longrightarrow$ work with us, even when it is hard
    - Be willing to learn from each other, form study groups, use the Slack chat etc

... More on coding stuff *later*

Things to Keep in Mind

- This course is **quite new**
  - We look at *current* state of the art *knowledge*
  - There will likely be bugs ...
  - ... We'll work through them together
- Mix of students at various levels, from various backgrounds
  - Learn from other's unique perspectives
- I (awkwardly) straddle the line between marketing & economics
  - I care about 'mechanisms' and explaining why things happen
  - I care about more than only the 4P's
  - I'm not a 'data scientist' or computer scientist
  - $\Longrightarrow$ pragmatism towards analysis that asks interesting questions

Meet the Instructors

Lachlan: Background

- Born and raised in Australia
- **New to the Netherlands**: second year here
  - Currently live in Rotterdam
  - Journey to the the Netherlands: London $\longrightarrow$ Zurich $\longrightarrow$ the Bay Area $\longrightarrow$ Zurich $\longrightarrow$ Chicago
  - Still learning the ins and outs of the Netherlands *and* the university
  - ... am have yet to explore outside of Rotterdam!
- Now: Assistant Prof @ Tilburg University

Lachlan: Areas of Expertise

- **Substantive:**
  - Quantifying the impact of Twitter on demand for new products (Movies)
  - Quantifying relationship between advertising and social media discussions
  - Understanding the role of Social Media Networks in Political Revolutions (Arab Spring)
- **Methods**:
  - Working with 'big' data
  - Text-as-data
  - Network analytics
  - Econometrics / statistics

Bottom Line: *analysing social media data to answer economic and marketing questions*

Lachlan: Areas of Expertise

Why is all this relevant?

- I've used social media data throughout my career
  - From constructing simple tables and figures to convince firms to update their decisions and strategies
  - ... to quite elaborate statistical / econometric models
    - that appeal to other academics

I believe in the value of what I am teaching

Gijs: Background

- Dutch, Esch Based (very small, Boxtel more familiar)
  - But studied at Tilburg and been working here for almost a year
- Lecturer in the Marketing Department
  - Might know me from:
    - Customer Analytics
    - ACCO (pre-master)
    - Bachelor's / Master's Thesis

Gijs: Interests

- Academics:
  - Education, marketing research
  - Data driven decision making
  - Sport Analytics
- Real World:
  - Sunday League Football
  - Fitness

Social Media Analytics:
A possibly biased perspective

What is Social Media Analytics?

Social Media Analytics is the application of statistical methods to understand behaviour on social media websites to make business decisions

- It's generally *empirical*, sometimes theoretical (i.e. mathematical)

What kinds of empirical analysis are of interest to us as marketers?

- Descriptive Analysis
- Causal Analysis
- Predictive Analysis

Descriptive Analysis

Descriptive Analysis: summarise characteristics of a data set

- What does the data look like?
  - Means, standard deviations distribution of data
  - Results are (stylised) *facts*

Examples:

- How are users who discuss the US election connected on Twitter?
- What topics are discussed on Yelp Reviews?
- Are discussions on Reddit about Albert Heijn different from those on Twitter?

Causal Analysis

Causal Analysis: Does A lead to B?

- Might also care about the mechanism of how it happens

Examples:

- Do Facebook ads increase product purchases?
- Does product adoption by influencers increase demand?
- Do tweets by TV studios increase the number of viewers of their show?

We'll spend a lot of time in this space

Predictive Analysis

Predictive Analysis: How can I best predict an outcome?

- When A occurs, so does B

Examples:

- Is this review posted by a real person or by a bot?
- How many retweets does Nike expect its next tweet to get?
- Who is a new Twitter user likely to follow?

Not the focus of this class

How to do Social Media Analytics

Social Media Analytics needs to combine tools from three areas:

1. Network Analytics
2. Text Analytics - 'Text-as-data'
3. Statistical / Econometric Methods

The exact mix of these depends on:

- **The question** you want to answer
  - Example: Can one deliver valuable insight by ignoring the network structure?
- **Personal taste**
  - ... I've increasingly started to value the network side of things lately
  - ... This view is not necessarily representative of all marketers

Good and Bad News ...

Good news: high quality social media analytics is incredibly useful

Why?

- Social media impacts a wide variety of industries
  - Media & entertainment, politics, health care, FMCG, fashion & beauty, etc
- It provides real answers to real problems in marketing and business strategy
  - And *people care about the answers*
- Being able to do (good) social media analytics ensures many (enjoyable) job prospects

Good and Bad News ...

Bad News: Its *hard*

- One needs to learn about networks, text analysis *and* statistics
- ... **and** be able to work on causal and predictive questions

(That seems like *a lot*...)

- And... learning the tools can be boring

Opinion: I think the upside far, far outweighs the bad.

Where We Are Going

Course Objectives

1. Explain and evaluate the challenges and opportunities social media and social networks present marketers.
2. Summarize state of the art knowledge from the academic marketing literature about social media's impact on marketing.
3. Provide intuitive explanations of statistical concepts from the areas of linear regression, causal inference, natural language processing and network analytics
4. Implement statistical analysis to analyze social media data using tools from linear regression, causal inference, natural language processing and network analytics
5. Interpret their own and other's statistical analysis of social media data
6. Prescribe Managerial and Marketing strategies to improve business performance based on analytical findings.
7. Appraise and critique the assumptions behind statistical analysis of social media data in a given setting and propose alternative methodologies to improve existing analysis

What We Will Cover

Four Blocks:

1. **Empirical Analysis of Patterns in Social Networks** (Week 1 & 6)
    - Structure of Social Media Networks
2. **Analysing Brand Reputation in Online Communities** (Weeks 3)
    - Online Reputation
    - Structure of Online Communities
3. **Quantifying the Importance of Influencers and Word of Mouth** (Weeks 4 & 5)
    - Word of Mouth
    - Influencers
4. **Measuring the Effectiveness of Social Media Advertising** (Weeks 6 & 7)
    - Viral Marketing
    - Social Advertising

Week 2? ... a deep dive into regression & causation

Building an Analytics Toolkit

1. **Network Analytics**
    - Summarise and plot network data
    - Identify communities within a network
2. **Text-as-Data**
    - Summarising Text
    - Measuring Sentiment
    - Identify topics being discussed
3. **Statistics/Econometrics**
    - Regression Modelling: OLS and extensions
    - Causal Inference

Building an Analytics Toolkit: Software

1. `R` - purpose built statistical software

    - Wide variety of statistical and graphical techniques
    - Can do most analysis marketers need to do
    - It's free and open source, and has a *friendly* user community

2. Git - Version Control

    - Track the changes to our code and writing systematically
    - Improves 'replicability'
    - *Highly valued* by employers in marketing analytics companies and in quantitative consulting

Your task before the Week 2:

- Install required software on your computer
- Coding Bootcamp - refresh your skills where required

Assessment

Group Assignments (30%)

- **2** group assignments - **each 15% of final grade**.
- Group Allocation: random + changes between assignments
- **Group Assignment 1**: Evaluating & Managing Online Reputation
- **Group Assignment 2**: Sentiment and Topic Analysis using Airline Tweets
- Assignment Structure:
  - Multiple Parts
  - Each part has multiple exercises
  - Mix of analytics and interpretation
  - Goal: walk through solving a problem
- **Submission via GitHub Classroom**
  - We'll introduce you to the details over next weeks

Group Presentations (20%)

- **2** group presentations & discussions - **each 10% of final grade**.
- Group Allocation: random + changes between assignments
- **Format**:
  - 2 groups per week starting Week 2
  - Presentation: 15 mins
  - Panel Q & A: approx 10 mins
    - Panel is other group presenting that week
  - Open Q & A: remaining time
  - Afterwards: Write a short individual reflection on how it went

Topics are available on the course website 2 weeks before a presentation

Final Exam (50%)

- Closed Book, Pencil/Pen and Paper
- 3 hours in length

Structure:

- PART A: True/False/Uncertain with an explanation (20%)
- PART B : Short answer questions (40%)
- PART C: An open ended / essay style question (40%)

Coverage:

- Everything from lectures, labs, and readings
- *Including* writing / explaining short code snippets

Lab Assignments (Ungraded)

Essential Notions:

- Learning by doing... especially for methods
- Mix of coding and conceptual
- *Invest your time in these, it will pay off*

Answers?

- Selected answers posted online
- Cover some material in Lab Section

Active Participation Bonus

Up to 1 Bonus Point (equiv. 10 percent of course grade)

- Want to encourage *active* & productive participation in classes
- How?
  - Engage in discussions
  - Ask questions
  - Participate in surprise short quizzes

Note: Max. course grade is still 10.


ANY
QUESTIONS?


Let's get
started!

License & Citation

Suggested Citation:

```
@misc{smwa2022_intro,
      title={"Social Media and Web Analytics: Course Introduction"},
      author={Lachlan Deer},
      year={2022},
      url = "https://tisem-digital-marketing.github.io/2022-smwa"
}
```

This course adheres to the principles of the Open Science Community of Tilburg University. This initiative advocates for transparency and accessibility in research and teaching to all levels of society and thus creating more accountability and impact.