It is increasingly possible to use analysis of quantitative data to inform, make, and even automate marketing decisions. This is possible in the sense that the data is often available or economical to collect, but many firms lack the analytical and managerial expertise (and culture) to effectively use this data. The Internet industry has led in both the availability and use of behavioral data, so I will draw on my experiences there, but we will apply many course ideas to other industries.

We will cover quantitative approaches to: learning customer lifetime value, customer acquisition and viral growth, measuring preferences and demand, recommendations and personalization, advertising, up-selling and cross-selling, retention and churn, social influence and network effects, and discovering customer needs.

The goal of this course is to give students the expertise to initiate, participate in, manage, and evaluate marketing analytics efforts with substantial business impact. This requires knowledge of fundamental ideas at the intersection of statistics, machine learning, and human behavior, but also benefits from knowing specific widespread techniques. Impactful analytics can require a challenging balance of rigor, speed, repeatability, and interpretability.

In exercises and assignments, students will work with data sets — real and synthetic, of my selection and of theirs. The assigned data sets will include retail purchase micro-data, ridesharing data, data from randomized and non-randomized marketing campaigns, and data describing behavior on social media.

**Schedule of topics**
This is a draft schedule that is subject to change. Readings listed with a session should be read before that session. Even if available elsewhere, access these from Stellar or Study.net as some readings are annotated, including with guidance about what to read or skim.

1. **Introduction: Marketing decisions as interventions**
   Types of data relevant to marketing. Purpose-built data and data exhaust. Ways to use data, from generating ideas to automating decisions. Modeling interventions on your customers and potential customers. Gaps between data and marketing decisions.

2. **Learning customer lifetime value (with statistics/ML review)**

   James et al. (2013). *An Introduction to Statistical Learning*. [excerpt]

3. **Measuring needs and preferences**
4. Randomized experiments (with statistics review)
Designing and deploying randomized experiments (e.g., A/B tests) — online and offline. Statistical inference with minimal assumptions (randomization inference). Power to detect effects. Improving precision through design and analysis. Using exposure data. Examples from the Internet industry.


5. Marketing experiments


6. Recommendation and ranking systems
Basic recommendation systems, with matrix factorization and logistic regression. Applications in providing a personalized service, recommending the correct products, and advertising. Recommendations in viral growth and network formation (e.g., PYMK). What is the value of improving recommendations?

Patel, DJ. (2011). Building data science teams. [excerpt]

7. Predictive modeling and consumer choices (with statistics/ML review)
Understanding how customers are choosing among offerings. Conjoint analysis with millions of observations. Trees, forests, ensembles, and bootstraps.

Practical bootstraps for big data [lecture notes]

8. Adaptation, personalization, and targeting
The mere appearance of personalization. Finding and exploiting heterogeneous susceptibilities. Picking test markets.

Optional reading:

9. Learning during and after a product launch

Seward, Z.M. (2013). The first-ever hashtag, @-reply and retweet, as Twitter users invented them. Quartz.
Birthdays on Facebook [video from Udacity]
10. **Analytics and product design**
Repeated experimentation. Continual experimentation and contextual multi-armed bandit problems. Answering product design questions with data. Parameterizing products, with examples from consumer credit and Internet.

Yahoo Front Page case study.

11. **Special guest, TBA**

12. **Scaling up surveys (and other self-report data)**
Coping with or getting beyond small and biased samples. Adjusting biased samples. Examples from measuring customer attitudes in the Internet industry.

13. **Special guest, TBA**

14. **CLV revisited: Growth, retention, and engagement**

15. **Pre-empting churn and competition**
Interventions to prevent churn. Targeting and network effects. Examples from the telecommunications industry. Competitive intelligence in marketing. Examples from ride sharing.

16. **Network data and multiple units**
Using network data to learn otherwise observed consumer characteristics. Targeting with network data. Network interventions to prevent churn in telecom. Users, households, cookies, and other units.


17. **Data-informed marketing and company culture**

18. **Review**

**Guests**
We will have a couple of guests from industry, to be announced. This may require adjustments to the schedule above.

**Requirements**
This section provides an overview of the main requirements of the course. More details about the requirements — and the resources, such as data, needed to complete them — will be provided throughout the course.

**Class participation (20%)**
Attendance is required. Everyone’s learning can benefit from vibrant discussion that reflects being prepared for class and drawing on additional knowledge or experience. We will track
contributions to discussion during class. Productive contributions will be rewarded. Because of the work required for the assignments, all other readings have been carefully selected; read them before class.

Assignments (60%)
The will be three sets of problems given. In each set, some questions will be indicated as graded while others are required exercises graded only for completion. You can work on the assignments with other students but must turn in your own answers reflecting you actually doing data analysis and writing.

Some of the questions will involve working with a data set. Many of the questions will have you inform marketing decisions using a provided data set. Other questions will ask you to do some analysis of a data set of your choice. So before the first assignment, you should select and acquire one or two data sets that you could use for those questions; further guidelines will be provided. Don't use the same data set as someone you are collaborating with.

Potential topics for the assignments

Students can use any tools they like for the data analysis assignments, but we will provide some support for learning Tableau. Tableau is commercial software for exploratory data analysis (EDA), dashboards, etc., which is seeing substantial use in industry. It can be used in combination with other tools, such as SQL, R, and Excel. Students can use it without cost via a classroom license. If you instead want to learn to do EDA in R, the free online course that I created with Udacity and Facebook colleagues might be helpful: https://www.udacity.com/course/data-analysis-with-r--ud651.

Final exam (20%)
There will be a short final exam. Some questions will be similar to the assignments, but will not require computer-intensive data analysis. A few example questions will be provided in advance.

Further guidelines and policies
All the usual stuff will be added here.