

Online Marketing Analytics

Syllabus

Spring 2009

Course Description:

The practice of marketing is changing. Due to increasing desktop computing power and companies amassing massive amounts of data, marketing decisions made by companies are becoming more and more data based. This holds in many sectors and in particular in internet marketing and retailing where the only interaction with the customer is in digital form. As a consequence, these “digital footprints” need to be analyzed very carefully in order to understand the customer’s preferences and needs.

In this course, we will study analytics for marketing decision makers. We will study a range of core analytical methods, and we will implement them using state-of-the art data mining software, applied to real marketing problems using real marketing data. At the core of this class is the application of analytics to online marketing. Students will learn about online marketing, they will create and manage their own online marketing campaigns and they will use analytics to monitor and adjust their campaigns. All of this will be accomplished within a world-wide online advertising competition, the Google Online Marketing Challenge. In that challenge, teams of students use real money to manage online advertising campaigns for real companies while competing against thousands of other student teams world-wide. This course is very hands-on and will have components of lectures, case discussion, data-driven projects and real-world campaigns.

Instructor:

Wolfgang Jank is associate professor of Decisions, Operations & Information Technologies at the Robert H. Smith School of Business, University of Maryland, and affiliated with the Center for Electronic Markets & Enterprises. He is interested in applying ideas from statistics and data mining to problems in electronic commerce, marketing, and operations management. Dr. Jank’s research has been published in the literature of statistics, data mining, information systems, and marketing. He has authored over fifty refereed articles and book chapters, and presented his work at national and international meetings. Dr. Jank received his Master’s degree from the Technical University of Aachen (Germany) and his PhD in Statistics from the University of Florida. After moving to the University of Maryland, he initiated, together with Dr. Shmueli, a new research area on Statistical Challenges in eCommerce. Dr. Jank is member of the American Statistical Society, the Institute of Mathematical Statistics, the European Network for Business and Industrial Statistics, the Association for Computing Machinery and INFORMS. He is past president of the University of Florida's chapter of the statistical honor society Mu Sigma Rho. Prof. Jank has been involved in a variety of consulting projects for private and public organizations, and he is advisory board member for several companies. Prof. Jank is teaching classes in data analytics in various programs (undergraduate, MBA, executive MBA and PhD) at the University of Maryland. He has received numerous awards including the top 15% teaching award for teaching MBA core classes.

- Contact info: wjank@rhsmith.umd.edu (e-mail).
- Website: <http://www.smith.umd.edu/faculty/wjank>

Course Pre- Requisites:

Students should have a basic understanding of statistics. I will assume that students have mastered a course in introductory statistics (e.g BUSI 630). Students should also have a basic knowledge of marketing. Basic software skills (particularly, for handling and manipulating data) will also be a plus. Since assignments are solved in teams, it is not essential that every student is a great master of every skill; but a portfolio of different skills across each team is definitely a plus. Moreover, while this course draws some ideas from the Data Mining course (BUDT 733), it is not essential that you have taken this course prior to this class.

Textbook:

There is no textbook for this course. Instead, there will be a variety of handouts and cases that students will prepare for each class. In addition, I recommend several books that cover material relevant to this class.

Relevant Books

The following is an (incomplete) list of books that cover material relevant to this course:

- Lattin, Carroll and Green “Analyzing Multivariate Data.” Duxbury/Thomson. (esp. Chapters 3, 8, 12 and 13)
- Hastie, Tibshirani and Friedman “The Elements of Statistical Learning” Springer. (esp. Chapters 3, 4, and 14)
- Berry and Linoff “Data Mining Techniques – For Marketing, Sales and Customer Relationship Management” Wiley. (esp. Chapters 5, 6, and 11 – but also read Chapters 17 and 18)
- Markov and Larose “Data Mining the Web” Wiley. (esp. Chapters 3 and 5)
- John, Whitaker and Johnson “Statistical Thinking in Business” Chapman and Hall. (esp. Chapters 3, 4, 8, and 9)

Software:

We will make use of the statistical software R. R is open source software and available from CRAN (<http://cran.r-project.org/>). CRAN hosts the basic software, add-on packages and a ton of additional reference material. You should spend a good amount of time *before the beginning of class* to check out all the resources available and to familiarize yourself with the software. I will also give a *brief* introduction to the main concepts during our first meeting.

The following documents give a very detailed introduction and overview of the software:

- <http://cran.us.r-project.org/doc/manuals/R-intro.pdf>
- <http://cran.us.r-project.org/doc/manuals/R-data.pdf>
- <http://cran.us.r-project.org/doc/contrib/Farnsworth-EconometricsInR.pdf>
- <http://cran.us.r-project.org/doc/contrib/Owen-TheRGuide.pdf>
- <http://cran.r-project.org/doc/contrib/usingR.pdf>

You should go over these documents very carefully within the first weeks to understand the basic principles of R and to get started with the software.

R is primarily a command-line language. While usage of R is extremely straightforward, you may find a GUI environment even more convenient. The GUI can be obtained from the following link: <http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/>

The reasons why we use R in this course (and not any other statistical software) are

- R is free! This is in contrast to many other commercial packages that cost several hundred dollars per license.
- R is an open source project. As such, it grows at a pace much faster than any commercial package and, as a consequence, offers capabilities for data mining that is second to none.

Course Format:

Each class meeting consists of lectures, discussions and presentations of data-projects and case-analyses. During lectures, I will do most of the talking but there will be plenty of opportunity for you to contribute by asking thoughtful questions and adding personal insight. In fact, I expect students to contribute and interact continuously. We will use modern *clicker technology* which fosters the interaction between professor and students.

Other parts of our meeting will consist of presentations & discussions. Presentations are lead by teams of students (i.e. *you* will do most of the talking during that time). To that end, you (and your team) will prepare presentations on a particular topic. Each presentation will be approximately 15 minutes in length. Presentations should be prepared in PowerPoint. Teams will be chosen at random; i.e. not every team will present every single time. However, all teams are expected to prepare a presentation and submit their presentation to the instructor before class.

Presentations are followed by class discussion. In that discussion, all teams (including those that did not present) are expected to add their experience and insight on the topic.

There will be two types of presentations: some on data-driven projects, and others on case analyses. Data-driven projects require the application of ideas from statistics and data mining to solve real marketing problems. These projects will have a strong focus on real data and use software to manipulate and to extract intelligence from that data. During your presentation (and the ensuing discussion) we will discuss problems & challenges that arise from the data analysis and ideas & solutions to extract business knowledge from that data.

In the other type of presentation, the focus will be on case analyses. For those presentations, you will read and analyze a particular case. Cases focus on the use of marketing intelligence within a company. You will analyze these cases using specific questions as guidelines.

Overall, the format of this course aims at covering three fundamental instructional goals:

- Lectures (and associated readings) aim at conveying new concepts and ideas related to data analytics in the online marketing context.
- Data Projects (and associated discussions) aim at implementing these ideas on real data and real marketing problems. Most concepts in data analytics can only be truly understood when implemented using data and software. Using data analytics can be challenging, especially for the inexperienced user. Our discussions serve as feedback mechanism so students can learn from these challenges and solutions from everyone.
- Case Analyses (and associated discussions) aim at learning about the use of data analytics “in action.” To date, there are only very few firms that use analytics as part of their core strategy. And even if they do, the public hardly every knows about it because data analytics is still regarded as one of the last (and secret) competitive advantages. In that sense, case analyses serve as a vehicle to learn about successful examples of data analytics at selected companies.

Course Technology:

We will use modern clicker-technology for collecting feedback and checking progress (see e.g. http://www.news.com/New-for-back-to-school-Clickers/2100-1041_3-5819171.html or <http://www.oit.umd.edu/ITforUM/2005/Winter/clickers.html>). This technology will allow me to get your feedback in real-time. It will also allow you to perform reality checks relative to the entire class.

Class Deliverables:

Deliverables for this class consist of different components. Some components will require oral presentations and others written reports; above all, class participation will also be a major component. All components will enter the final grade; the precise weighing of each component is shown below.

Grading Policy:

| | |
|---|------|
| In-class <i>team</i> data project presentation | 25% |
| In-class <i>team</i> case analysis presentation | 25% |
| End-of-class <i>team</i> paper and presentation | 30% |
| <i>Individual</i> class participation | 20% |
| <hr/> | |
| Total | 100% |

In-class team data project presentation:

Three data-driven projects will be assigned over the course of this class. These projects cover different data-analytic concepts applied to real marketing problems. The goal is to implement ideas and methods learned in class (esp. during the lectures) on real data pertaining to real marketing problems.

Projects are to be solved in teams. Each project strongly correlates with the material covered in the previous class. For instance, as the first class will cover regression methods, the project due in the second class will relate to regression methods and its application to a particular marketing task.

Each project will relate data-driven decision making to a specific marketing function. Each project is complex and involves real decisions on real data. One of the learning objectives is to deal with real data. Real data can be “messy” (e.g. missing, unusual or duplicate observations) which complicates the knowledge extraction process. It will be your task, as a team, to deal with these data and their associated challenges, and derive actionable marketing decisions from them.

Results of the data project will be presented in class. Every class meeting, two teams will be selected at random to present their results. Each presentation will be followed by a class discussion. The goal is to receive immediate feedback on the assignments and to learn from the ideas (and also mistakes!) of your peers.

While only two teams will present their results, all teams are expected to solve the problem and prepare a presentation. All teams will email their presentation to the instructor before the start of class. All teams are expected to be prepared. In fact, every student is expected to be able to comment on the data modeling aspect and how it relates to the specific marketing task.

Presentations are to be prepared in PowerPoint.

In-class team case analysis presentation

Three case studies will be assigned over the course of this class. The cases cover different aspects of data-driven decision making for marketing in the context of a real business. The goal is to understand the “big picture”, that is, the use of data-analytics in the real world, to learn about success stories and also possible shortcomings of data-analytics.

Cases will be prepared in teams and discussed in class. Similar to the data projects, two teams will be chosen at random to present their results. That is, while all teams are expected to analyze the case and prepare a presentation, only two teams will present it. All teams will send their presentations to the instructor before the class starts. All other teams are expected to contribute to the class discussion.

For each case, you will receive a set of specific question which will guide your analysis. Please note that while cases typically contain a plethora of information about a firm (such as financial performance), our main focus will be on the information that pertains to the use of data-analytics in the context of online marketing. Moreover, while each case

contains valuable information about a firm's usage of marketing analytics, you are encouraged to do additional research (e.g. via the web) to find complementary information.

Presentations are expected to be prepared in PowerPoint.

End-of-class team paper, presentation & online marketing competition:

Designing, Executing and Mining an Online Marketing Campaign

Your end-of-class team paper and presentation involves applying your knowledge of marketing analytics to a real business. In that sense, it gives you an opportunity to apply the concepts learned in class to a real business scenario. In fact, you will design and execute an online marketing campaign, and you will analyze and optimize it using data-analytics. *And you will do all of this within a world-wide competition against thousands of other student teams!*

This project will last the entire duration of class. During the early parts, you will identify a business, familiarize yourself with its market and design a marketing campaign. Then, you will execute the campaign using real money, and you will monitor and adjust it using data-analytics. At the end of class, you will present your results; you will also write-up the key elements of your campaign in a semester paper.

This project should be executed in the following steps. Each step is to be performed in your team.

Step 1: *Prepping for Online Marketing* -- Familiarize yourself with online marketing basics. Read the document "Marketing and Advertising using Google;" familiarize yourself with *Google AdWords* (<http://adwords.google.com/select/Login>) and *Google Analytics* (<http://www.google.com/analytics/>). This step should be completed *before* our class starts.

Step 2: *Collaborating with a Business* -- Identify a business for your online marketing campaign. Reach out to a business and obtain permission to run their online marketing campaign. Familiarize yourself with marketing basics of that business. You should obtain a basic understanding of

- **Company Background:** Obtain background information on your firm including industry, location, mission statement.
- **Segmentation Strategy:** Obtain information on target customer(s). If the target customer is a business, obtain information on geographic location, type of industry, company size, and product end-use. If the target customer is consumers, obtain information on socio-economic, demographic, benefits sought, and psychographic characteristics. Keep in mind that a company may have more than one group of target customers.

- **Product Strategy:** Obtain information on the product/service category(ies) offered. Obtain information on the brand names in the product mix; the image of the company and of the brand(s).
- **Pricing Strategy:** Obtain information on the position in the market (e.g., low priced, moderate, upscale) and any other relevant pricing information.
- **Communication Strategy:** Obtain information on the type of advertising messages used; the company's sales promotion programs. Obtain information on any customer relationship management programs in place.
- **Distribution Strategy:** Obtain information on the distribution channel through which the company's products are marketed.

You should also obtain permission to install Google Analytics and to set up a Google AdWords account.

Note: It will be your responsibility to identify, contact and interact with that business. Since our class meets over only a rather short period of time, you should identify that business **before our first class meeting**. Moreover, as we will compete in the Google Online Marketing Challenge, please make sure that the business conforms to the following basic rules (set forth by Google):

- The business should be a small- to medium-sized business
- The business must be new to AdWords. That is, the business only qualifies if it currently does not use Google AdWords campaigns. This is an important point since non-compliance will lead to disqualification from the challenge (and, Google's AdWords vouchers will not work if previous AdWords campaigns have been run for that business!).
- The business should give you permission to run their online marketing campaign. Ideally, it should also give you access to Google Analytics. This may be a tricky point since some businesses may not want Google Analytics installed in their HTML code; I would urge you to try as hard as possible for this point.

PLEASE CAREFULLY READ GOOGLE'S "SELECTING AND WORKING WITH A BUSINESS OR ORGANIZATION", ATTACHED AT THE END OF THIS SYLLABUS!

Step 3: *Background Check* -- Before starting the online marketing campaign, you want to understand organic web traffic. To that end, monitor *Google Analytics*. Monitor Google Analytics before starting your AdWords campaign. The goal is to better understand organic traffic to the target website. (Where do visitors typically come from? Which search words do they use? Are there markets that are under-covered?) This will also help you benchmark your campaign efforts (Does your campaign improve over organic traffic?), identify good search keywords, and relevant geographical areas for your campaign. You should monitor Google Analytics for at least one week in order to understand the organic traffic leading to your target web site.

You should also get background information on web traffic and associated costs for your particular business. To that end, use Google's Traffic Estimator (<https://adwords.google.com/select/TrafficEstimatorSandbox>) to learn about the costs for keywords and the estimated clicks per day. This will help you putting together a daily budget for your advertising campaign.

Step 4: Setting up the Campaign -- First, brainstorm (as a team) about the optimal combination of ad groups, the set-up of individual ad groups, keywords and maximum cost per click. You should also strategize about the maximum amount of money you are willing to spend every day to maximize the effectiveness of your campaign.

Step 5: Monitoring the Campaign -- After running your online marketing campaign for a week, analyze the results. Analyze the effectiveness of your ad groups and keywords. Compare the results with overall traffic using Google Analytics. Determine which combination of ad groups and keywords leads to the optimal campaign result.

KPI's (Key Performance Indicators): There are several KPI's that you should monitor. First, the total impressions tell you how many times your ad has been shown. Total clicks tell you how many times users have clicked on your ads. The click-through-rate (CTR) equals the total clicks divided by the total number of impressions and measures the quality of the campaign (the larger, the better). You may also want to monitor your total costs, total number of campaigns, total ad groups, total number of ads and keywords.

At the end of class, every team will report on their campaign efforts and results. This will be in the form of in-class presentations (by every team) as well as a semester paper. The semester paper should be no more than 10 pages (including exhibits) and it should contain a) an overview of the firm, b) an overview of the campaign, c) campaign results, d) conclusion and recommendations for the firm.

Individual class participation and clickers:

Effective participation consists of not only responding to questions raised by the instructor but also asking thoughtful questions and responding to contributions from your fellow-students. Quality of participation is more important than quantity. However, you will not earn credit in this component, if you rarely speak in class. Quality of participation includes: Evidence of reading and prior analysis; Relevance of comments; Ability to listen and relate to input from other students; Ability to lead discussion into previously unexplored areas; Ability to admit error; Ability to intellectually interact with other students (and not just the instructor).

I will foster (and also measure) class participation using the clicker technology. Clickers allow one question to be answered simultaneously by all students. Clickers also allow for immediate feedback, thereby stimulating further discussion.

Class-by-class Schedule

Each class meeting will be 6 $\frac{3}{4}$ hours long; there are a total of 4 class meetings.

An outline of each meeting follows below; please note that the outline is tentative and subject to change.

I. Overview

| Class | Date | Time | Classroom |
|-------|------|--------------|-----------|
| 1 | 2/14 | 8:30 3:15 | TBD |
| 2 | 2/21 | 8:30 3:15 | TBD |
| 3 | 3/7 | 8:30 3:15 | TBD |
| 4 | 3/14 | 8:30 3:15 | TBD |

II. Detailed Outline

(Tentative and subject to change)

| Day1 | | | | |
|-------------|--------------|--|---|---------------------|
| Time | Type | Content | Reading | Deliverables |
| 8:30 | Introduction | Goals, Deliverables | | |
| 9:30 | | | | |
| | <i>Break</i> | | | |
| 9:45 | Lecture | Introduction to R/ Review of Forecasting with Regression | | |
| 10:45 | | | | |
| | <i>Break</i> | | | |
| 11:00 | Case | Online Advertising: History and Outlook | HBS: <i>How Media Choices change Online Advertising</i> | Presentation |
| 12:00 | Discussion | | | |
| | <i>Lunch</i> | | | |
| 1:00 | Case | Google's Online Advertising with AdWords | HBS: <i>Google Advertising</i> | Presentation |
| 2:00 | Discussion | | | |
| | <i>Break</i> | | | |
| 2:15 | Discussion/ | Google Challenge | | |
| 3:15 | Team Work | | | |

| Day 2 | | | | |
|--------------|--------------|--|-------------------------|---------------------|
| Time | Type | Content | Reading | Deliverables |
| 8:30 | Data Project | Conjoint Analysis and Product Optimization | | Presentation |
| 9:30 | Discussion | | | |
| | <i>Break</i> | | | |
| 9:45 | Lecture | Classification/Targeting: Overview | | |
| 10:45 | | | | |
| | <i>Break</i> | | | |
| 11:00 | Lecture | Targeted Marketing: An application to cereal preferences | | |
| 12:00 | | | | |
| | <i>Lunch</i> | | | |
| 1:00 | Case | Information-based Marketing, Online and Offline | HBS: <i>Capital One</i> | Presentation |
| 2:00 | Discussion | | | |
| | <i>Break</i> | | | |
| 2:15 | Discussion/ | Google Challenge | | |
| 3:15 | Team Work | | | |

| Day 3 | | | | |
|--------------|--------------|--|-------------------------|---------------------|
| Time | Type | Content | Reading | Deliverables |
| 8:30 | Data Project | Voter Targeting | | Presentation |
| 9:30 | Discussion | | | |
| | <i>Break</i> | | | |
| 9:45 | Lecture | Clustering/Segmentation: Overview | | |
| 10:45 | | | | |
| | <i>Break</i> | | | |
| 11:00 | Lecture | Market Segmentation: An application to the auto market | | |
| 12:00 | | | | |
| | <i>Lunch</i> | | | |
| 1:00 | Case | Mining Online Information | HBS: <i>Intelliseek</i> | Presentation |
| 2:00 | Discussion | | | |
| | <i>Break</i> | | | |
| 2:15 | Discussion/ | Google Challenge | | |
| 3:15 | Team Work | | | |

| Day 4 | | | | |
|--------------|--------------|--|--------------------------|---------------------|
| Time | Type | Content | Reading | Deliverables |
| 8:30 | Data Project | Segmenting Customer Preferences | | Presentation |
| 9:30 | Discussion | | | |
| | <i>Break</i> | | | |
| 9:45 | Lecture | Further Applications and Success Stories of Marketing Intelligence | | |
| 10:45 | | | | |
| | <i>Break</i> | | | |
| 11:00 | Case | Data & Privacy | HBS: <i>Choice Point</i> | Presentation |
| 12:00 | Discussion | | | |
| | <i>Lunch</i> | | | |
| 1:00 | Team | Google Challenge | | Presentation |
| 2:00 | Presentation | | | |
| | <i>Break</i> | | | |
| 2:15 | Team | Google Challenge | | Presentation |
| 3:15 | Presentation | | | |

Selecting and Working with a Business or Organization

One of the major aims of the Challenge is to give students practical, real-life experience as part of their studies. With this in mind, the best way to work with businesses is for the student teams to think of themselves as consultants, and the business as the client. Teams should aim to apply the same level of professionalism that a real-life consulting firm would.

Selecting the right business or organization is important. The right choice can range from one person focusing on their local region to a multi-office business servicing a few countries. A good idea is businesses between 1-100 employees.

Try to work with businesses relevant to the types of search queries that Google users conduct. A good example would be a traditional retail business, such as a home wares store, a vintage fashion store or a niche beauty store. You might want to try a few keyword searches for goods and services relevant to your potential business before making your final selection.

You should be aware that Google has content guidelines and will not run AdWords for sites promoting inappropriate items such as academic aids, some alcohol, bulk marketing, counterfeit designer goods or cigarettes. (Details available at <https://adwords.google.com/select/contentpolicy.html>)

Please note that the business or organization should not currently use AdWords in any capacity and not have had an active account within the last 6 months.

Tips on Businesses to Avoid

When searching for likely candidates, remember that you may compete against many companies who have large advertising budgets to spend on the same keywords you want to use. With this in mind, you might want to avoid businesses such as:

- Web Hosting
- Web Design Agencies
- Insurance Companies
- Mortgage Agencies
- Debt Consolidation Companies
- Multi-level Marketers – http://en.wikipedia.org/wiki/Multi-level_marketing
- Distributors
- Affiliate Companies – http://en.wikipedia.org/wiki/Affiliate_marketing

Also, please bear in mind that many big players in travel and finance have advertised and optimized their campaigns for years. They have significant online advertising budgets and experience. It might be tough for you to compete against them.

Please note: This isn't to say that AdWords isn't appropriate for these businesses – the point is that you have a limited budget and a limited timeframe so you might find it difficult to compete effectively using them as clients.

Finally, please note that the 'landing page quality score' of the website can affect your account performance. When selecting your business, to ensure their website is suitable, please read the landing page/website guidelines at <http://adwords.google.com/support/bin/answer.py?answer=46675&topic=9356>. For further information on websites that typically have poor landing page quality, please see <http://adwords.google.com/support/bin/answer.py?answer=66238>.