

Business Analytics Syllabus

Course Description

Business analytics refers to the ways in which enterprises such as businesses, non-profits, and governments can use data to gain insights and make better decisions. Business analytics is applied in operations, marketing, finance, and strategic planning among other functions. The ability to use data effectively to drive rapid, precise and profitable decisions has been a critical strategic advantage for companies as diverse as WalMart, Google, Capital One, and Disney. For example, Capital One uses sophisticated analytic capabilities to match credit card offerings to customers more accurately than their competition. WalMart uses analytics to monitor and update its inventory in a way that allows it to serve its customers at an exceptionally low cost. In addition, many current and recent startups such as Palantir and Splunk are based on the application of analytics to large data bases. With the increasing availability of broad and deep sources of information — so-called “Big Data” — business analytics are becoming an even more critical capability for enterprises of all types and all sizes.

In this course, you will learn to identify, evaluate, and capture business analytic opportunities that create value. Toward this end, you will learn basic analytic methods and analyze case studies on organizations that successfully deployed these techniques. In the first part of the course, we focus on how to use data to develop insights and predictive capabilities using machine learning, data mining and forecasting techniques. In the second part, we focus on the use of optimization to support decision-making in the presence of a large number of alternatives and business constraints. Finally, throughout the course, we explore the challenges that can arise in implementing analytical approaches within an organization.

The course emphasizes that business analytics is not a theoretical discipline: these techniques are only interesting and important to the extent that they can be used to provide real insights and improve the speed, reliability, and quality of decisions. The concepts learned in this class should help you identify opportunities in which business analytics can be used to improve performance and support important decisions. It should make you alert to the ways that analytics can be used — and misused — within an organization.

We have three goals in this course. The first is to help you think critically about data and the analyses based on those data — whether conducted by you or someone else. The second is to enable you to identify opportunities for creating value using business analytics. The third is to help you estimate the value created using business analytics to address an opportunity. Business analytics is an integral part of modern management — this course should provide you with the foundation you need to understand and apply these methods to drive value.

Course Materials

There is no required textbook for the class. There will be cases, articles as well as slides that we will distribute in each class, and post on canvas.

Requirements and Grading

There will be three homework assignments to hand in, preparation assignments for sessions, and a final. In addition, there will be periodic “concept checks” — short quizzes to be done individually to check your understanding of the material. If you understand the material, it should take you no more than 10 minutes to do each concept check.

Your grade will be determined as follows:

- Final exam: 50%
- Participation: 10%
- Concept checks: 10%
- Homework: 30%

You are expected to come to class prepared, and ready to discuss the pre-class reading, case or assignment questions. In many instances, we will ask you to provide a brief summary of some of your answers to pre-class assignments in an online system (a link to that effect will be provided on Canvas).

There will be one concept check quiz after each lecture (from Lecture 2 to Lecture 10). Quizzes will usually be due at the start of the week following the lecture (the exact due dates are provided in Canvas). In other words, two concept check quizzes will often be due at the same time. Each quiz is worth 10 points. If you score over 60 points in total, you will receive the full 10% credit for the concept check quizzes towards your final grade. (If not, you will receive percentage credit of $(\text{Total score})/6$.)

Homework

Homework assignments will need to be submitted on Canvas. There is one individual assignment (Type B) and two group assignments (Type A, to be done by learning teams).

- Homework 1 (individual assignment) is due (i.e., should receive a Canvas time stamp) before the start of session 4.
- Homework 2 (group assignment) is due (i.e., should receive a Canvas time stamp) before the start of session 7.
- Homework 3 (group assignment) is due (i.e., should receive a Canvas time stamp) before the start of session 11.

Software

This course will require the use of Windows Excel and we will provide a business analytics excel add-in. This add-in only runs under Windows, so Mac users will need to be able to boot into Windows. *There is no license associated with this add-in so you may keep it after the course and continue to use it.* The add-in extends the functionality of Excel to provide several additional analytic methods such as logistic regression, k-nearest neighbors, ROC curves and Monte Carlo simulation. It will allow you to employ all the analytic methods we learn in an Excel environment.


Syllabus at a Glance

Session 1: Introduction

Predictive Analytics

- **Session 2:** Predicting outcomes I / lending analytics
- **Session 3:** Predicting outcomes II / recommendation analytics
- **Session 4:** Quality of predictions I / healthcare analytics
- **Session 5:** Quality of predictions II / financial analytics
- **Session 6:** Predictions and skill versus luck / sports analytics

Prescriptive Analytics

- **Session 7:** Testing / retail analytics
- **Session 8:** Simulating the future / pension analytics 
- **Session 9:** Optimizing complex decisions / salesforce analytics
- **Session 10:** Optimizing with multiple objectives / portfolio analytics

Implementation

- **Session 11:** Decision-support systems -from concept to deployment- / supply chain analytics
 - **Session 12:** Review
-

Detailed Session Plan

- **Session 1:** Introduction

In this class, we introduce business analytics and some main levers that firms use to capture value through it.

We discuss the first case, Nomis, exploring the early stages of the analytics life cycle of a company aiming to capture value through pricing analytics.

Case: Nomis Solutions (Part A)

- **Session 2:** Predicting outcomes I / lending analytics

In this session, we study a common tool to leverage the data for prediction purposes: logistic regression.

Case: Nomis Solutions (Part B)

- **Session 3:** Predicting outcomes II / recommendation analytics

In this session, we discuss the possibilities offered by big data and the possibilities for mass customization of services. We introduce recommendation systems that companies such as Pandora, Netflix and Amazon are using to predict user tastes based on the information they collect. We introduce nearest neighbors methods as an alternative to regression or logistic regression.

Case: Pandora

- **Session 4:** Quality of Predictions I / Healthcare Analytics

In this session, we discuss how to assess the quality of predictions and quantify the errors one may make when predicting 0/1 outcomes. In particular, we introduce the notion of ROC curves and detail how to account for such errors when making economic trade-offs.

To illustrate these concepts, we focus on the task of predicting re-admissions of patients in hospitals, and the implications for preventive treatment measures.

Case: Tahoe Healthcare

- **Session 5:** Quality of predictions II / financial analytics

In this session, we discuss the notion of in-sample and out-of-sample predictions, and how the latter is key to properly assess the quality of predictions.

We discuss these concepts in the context of the design of trading strategies.

- **Session 6:** Predictions and skill versus luck / sports analytics

In this session, we discuss how to disentangle skill versus luck when attempting to make predictions about future performance of, e.g., sports players or investors. We introduce the notion of shrinkage estimators for such purposes.

- **Session 7:** Testing / retail analytics

In this session, we transition to the prescriptive analytics module, and in particular, the evaluation of the quality of decisions and the optimization of these.

We introduce a tool, Difference in Differences, to test the quality of changes in decisions in complex environments in the absence of perfect experiments.

We illustrate these concepts through the evaluation of a major new initiative at a retail chain and through the evaluation of the ROI of online advertising campaigns.

Case: Home and Kitchen

- **Session 8:** Simulating the future / pension analytics

In this session, we introduce simulation, a tool to evaluate decisions in the presence of uncertainty. We illustrate its use in the evaluation of pension liabilities at GM.

- **Session 9:** Optimizing complex decisions / salesforce analytics

In this session, we introduce how to formalize the optimization of a large number of decisions while accounting for different kinds of physical and business constraints. We discuss more specifically the assignment of sales representatives to doctors in the context of pharma retailing.

Case: Pharma detailing

- **Session 10:** Optimizing with multiple objectives / portfolio analytics

In this session, we discuss decision-making in the presence of multiple objectives. We introduce the notion of efficient frontier and discuss the pitfalls to keep in mind while attempting to construct such frontiers in, e.g., portfolio management.

We illustrate the use of optimization in portfolio management and the value this may add to, e.g., the more basic trading strategies discussed in Session 5.

Pre-class reading and assignment: Portfolio optimization

- **Session 11:** Optimizing with multiple objectives / portfolio analytics

Recently, Zara introduced new algorithms to dispatch products in its chain of stores. We will analyze the models underlying its decision support system. We will then discuss approaches to evaluate such a new decision support system. This session will bring together many of the concepts developed during the course.

Case: Zara

- **Session 12:** Review

Other Resources

Optional Recommended Readings

While there is no required textbook for the class and there is no definitive and comprehensive reference on the subject, there are several books we recommend in case you would like to learn more and follow up beyond what we will cover in class.

- For data mining and predictive analytics, the following are introductory textbooks:

Data Science for Business, Provost and Fawcett: O'Reilly

Data Mining for Business Intelligence, Concepts, Techniques and Applications, Shmueli, Patel, and Bruce: Wiley

For Excel modeling and optimization, the following is a good textbook:

Management Science: The Art of Modeling with Spreadsheets, Powell and Baker: Wiley

The three books above are on reserve at the Watson library.

- For scientific retail management, the following is an interesting account of various approaches that may be adopted to leverage analytics:

The New Science of Retailing, Fisher and Raman: Harvard Business Press

- An interesting recent industry report on big data and the role that analytics may play in industry is:

Big data: The next frontier for innovation, competition, and productivity

and can be found at:

http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation

Optional Software

- Rattle: this is an open source R-based data analysis tool.

<http://rattle.togaware.com/>

An introductory book to Rattle

Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery, Williams: Springer

is on reserve at the Watson library.

- Orange: this is an open source data-mining tool which is intuitive and workflow oriented.

<http://orange.biolab.si/>

- Tableau: this is a data visualization tool. A trial version may be downloaded at

<http://www.tableausoftware.com/products/desktop/download>