Columbia Business School

The Analytics Advantage: Case Studies in Big-Data and Analytics in Industry

> **Note**: the first lecture of this class will involves a panel discussion with various industry guest speakers (see details below). This first class will take place in Geffen 120 (Cooperman Commons) on March 21st at 2pm, <u>for all sections</u>. Attendance exemptions for this first lecture will be granted only to those with class conflicts.

This is the final class in the CBS Business Analytics sequence and focuses on how analytics have transformed industries from healthcare to finance. This class will build upon concepts you encountered in the core and in follow-on analytics classes. The only pre-requisites for this class are core Statistics and Business Analytics.

Course Description

Business analytics refers to the ways in which enterprises such as businesses, non-profits, and governments 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. In addition, many current and recent startups are based on the application of analytics to large databases. 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.

This class will focus on how analytics have generated value in a broad range of industries. Each class will be taught by a different faculty member with specific subject matter expertise and will focus on one specific industry and on how it has been transformed through the use of analytics.

To reflect the importance of this connection to industry, our first session will involve a panel discussion with guests from industry, who will tell us about their data-driven work within their organizations.

Following this first class, each of the remaining five sessions will focus on a specific industry including one or more of the following:

1. Background on the history and importance of the problem in that industry, and the impact analytics have had in that respect

2. A primer on relevant technologies and methods

3. A discussion of the inherent complexities involved in implementing an analytical solution in a messy, fast-moving environment

4. Current, state-of-the-art approaches to the problem

In each session, the focus will be on how analytics can be used to drive value; each of the points above, whether about data, technology, or analytics, will converge on that topic
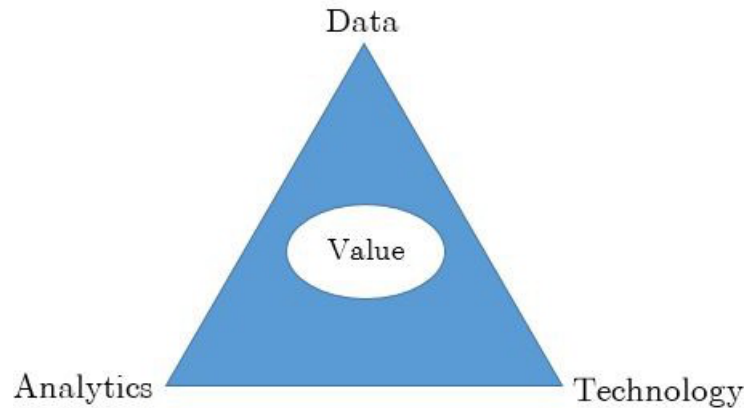
As in the core, this class will involve 'hands on' work using the Business Analytics add-in and other tools as appropriate.

Connection to the Core & Analytics Curriculum

You were introduced to the fundamentals of business analytics in your core 'Business Analytics' class and given an opportunity to deepen and broaden your exposure to these topics in various elective classes (including Business Analytics 2, Applied Regression Analysis, Data Analytics in Python, and Modern Econometrics for Business, among others).

These classes were anchored in cases, but their focus was on understanding the analytics techniques and the way value could be extracted using these techniques. As a result we generally picked one very narrow problem to work on without a broad focus on the industry itself. 'The Analytics Advantage' takes the opposite perspective and focuses on the industry first, looking at a range of ways analytics can be used in that industry.

Many students will choose to take this class after having taken one of our more advanced analytics electives, but those interested in broad applications without a deep dive into any specific technique may find this class attractive as a follow-on to the core.

Course Materials

There is no required textbook for the class. There will be slides for each class that will be posted on canvas.

Requirements and Grading

This class will be assessed by means of data-based exercises that will be set either before or after each class. These exercises will give you an opportunity to either prepare for the upcoming class or extend the examples you have studied in the previous class.

There will be five assignments, one for each class. You grade for the four best assignments you

attempt will be used; thus, you are free to drop one of the assignments.
See Canvas for assignment details and due dates.
Grades will be determined as follows:

- 70%: assignments
- 30%: attendance and participation

You are expected to come to class prepared, on time, and ready to discuss any pre-class readings and articles assigned for that class.

Detailed Class Plan

Each class will be taught by a faculty member in the Decision, Risk, and Operations division, based on their research and expertise in a specific industry. To ensure continuity and to link the topics in each class to each other, Angela Quintero and Nelson Fraiman will act as course leaders, introducing each class and handling any logistics that arise.

Classes will be as follows; order and exact content is subject to change.

Lecture 1, March 21st – Geffen 120, 2pm (all sections!)

Analytics in industry, panel discussion with guest speakers
- **Panelists:**

  - **Shahram Eballodahi**
    Ex Chief Data and AI Officer, Novartis
    Member, Data and Society Council, Data Science Institute, Columbia University
    https://www.linkedin.com/in/shahramebadollahi/
  - **Otis Jennings**
    Senior Director of Data Science, Elicit LLC
    Ex Senior research Scientist, Meta
    https://www.linkedin.com/in/otis-jennings-phd-06619b1/
  - **Mahmoud Naghshineh**
    Vice President, AI Impact & Partnerships, IBM Research
    https://www.linkedin.com/in/mahmoud-naghshineh-60091723/
  - **Margaret Pierson**
    Director, Data Science, Wayfair
    Ex Manager, research Science, Amazon
    https://www.linkedin.com/in/margaretpierson/
  - **Graham Poliner**
    SVP, Planning, Allocation, and Analytics at Five Below
    Ex SVP Analytics and Inventory Management, Macy's
    https://www.linkedin.com/in/graham-poliner-5877419/
  - **Ali Sadighian**
    Senior Manager, Uber
    Ex Senior Manager Supply Chain Optimization Technologies, Amazon
    https://www.linkedin.com/in/ali-sadighian-7b02b314/

- **Moderators:** Nelson Fraiman & Angela Quintero

Lecture 2, March 28th, Prof. Daniel Guetta

Digital Marketplaces and Ride Hailing: An Industry Based on Data

Ride hailing is an industry based on information technology and data. Mobile computing devices in the hands of every rider and driver – with broadband connectivity, graphical interfaces, GPS tracking and high-resolution mapping – have created an unprecedented capability to instrument and automate transportation. This has led to the creation of novel marketplaces for transportation that have fundamentally changed urban transportation throughout the world. In this session, we look at the role of analytics in enabling this innovation. Specifically, we will examine a fundamental spatial-temporal model of ride hailing service and how it is used to support product, strategic and operational decision making in the industry.

Lecture 3, April 4th, Prof. Paul Glasserman

Text Analytics in Finance: Signal from Chaos

Text analysis is increasingly used to extract predictive information from news, social media, corporate and legal documents. In this session, we will introduce two of the most widely used tools from natural language processing, sentiment analysis and topic modeling. We will discuss the main techniques used and highlight companies developing or applying these methods, with particular emphasis on using text analysis to forecast financial market returns.

Lecture 4, April 11th, Prof. Carri Chan

Analytics in Healthcare: When Regression Isn't Enough

There is great promise surrounding the use of Machine Learning in healthcare. We will introduce tools to evaluate the potential benefits of utilizing predictive analytics with a particular emphasis on healthcare delivery at hospitals. We will also discuss how to address challenges introduced by the need to provide guidance on how to go from predictive analytics to prescriptive analytics when randomized control trials are not always possible.

Lecture 5, April 18th, Prof. Ciamac Moallemi

Blockchain and Cryptocurrencies: Fundamentals, Approaches, and Future Outlook

The development of cryptocurrencies has garnered much attention as one of the most surprising fintech innovations in recent years. In this session, we will employ analytics to study these systems. We will begin by describing the enabling underlying technologies: decentralized blockchains and distributed ledgers. Building on this, we will analyze the basic economics of these systems and leverage simulation analytics to understand security and operational characteristics. We will conclude by considering the implications of our analysis for the long-term success of decentralized blockchains.

Lecture 6, April 25th, Prof. Assaf Zeevi

The Rise of Recommendation Engines: The Promise and Challenges

Recommendation engines (or recommender systems as they are often referred to) are one of the most successful and widely used applications of machine learning technologies in business. They are central to many areas, in particular in the burgeoning tech sector, including social networks, search engines, online retail, music and video streaming platforms, and content dissemination, to name but a few. In this session we will introduce the basic premise for recommendation engines, key working principles and relevant concepts, as well as some of the methods that pertain to their operation. We will illustrate these in several use cases, and discuss key challenges that present in these contexts.