Course: IEOR 4575 Business Analytics for Operations Research
Lectures MW 2:40-3:55PM

Instructor
Prof. Guillermo Gallego

Office Hours
Tuesdays: 3-4pm
Office: CEPSR 822 (8th floor)

Textbooks and Learning Materials
There is no required textbook: all class materials will be available on our Courseworks website. However, some books are very useful if you want to learn more and deeper about data analytics. The best way to learn is by doing (especially with programming)


Optional Textbook 2 (solid primer, with theory and explanation): An Introduction to Statistical Learning with Application in R, by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani; Publisher: Springer (2013); ISBN-13: 978-1461471370;

Optional Textbook 3 (a great advanced text): Elements of Statistical Learning: Data Mining, Inference, and Prediction, by Trevor Hastie, Robert Tibshirani and Jerome Friedman, but it requires some mathematical sophistication and goes beyond the material we will be covering. The book is free at http://statweb.stanford.edu/~tibs/ElemStatLearn/index.html

Software:
• We require the R Statistical Software, which is powerful and free. R can be downloaded at the link below: http://www.cran.r-project.org/
• Rstudio is a free platform for both writing and running R, available at www.rstudio.org. Some students find it friendlier than basic R.
• The learning curve is very steep. Students can become proficient in a few weeks. Some manuals are very helpful to learn R, e.g., http://cran.r-project.org/manuals.html
• I provide limited software instruction, in-class demonstration, and code to accompany lectures and assignments. We do not assume that you have used R in a previous class. However, this is not a class on R. Like any language, R is only learned by doing. You should install R as soon as possible and familiarize yourself with basic operations.
• Additional resources: (a) Tutorials at data.princeton.edu/R are fantastic (and there are many others out there). (b) Youtube intros to R, e.g. the series from Google Developers.
Courseworks Site

A Courseworks site is set up for this course. Each student is expected to check the site throughout the semester as Courseworks will be the primary venue for outside classroom communications between the instructors and the students. Students can access the course site at https://courseworks.columbia.edu/welcome/

Business Analytics Course Description

This course prepares students to gather, describe, and analyze data, using advanced statistical tools to support operations, risk management, and response to disruptions. Analysis is done targeting economic and financial decisions in complex systems that involve multiple partners. Topics include: probability, statistics, hypothesis testing, experimentation, and forecasting. Prerequisite: Stat-IEOR 4150 or equivalent.

Business Analytics Course Overview

This course is about how to identify and capture value through analytics. We live in an information economy. Managers who understand how to leverage data through analytics will be better equipped to create and lead organizations than those who do not. This course will give you with that understanding.

Analytics is about information—turning data into action. Its value derives fundamentally from information gaps in the economic choices of consumers and firms. Analytics unlocks this hidden value.

Consider consumers. In economic theory, they make choices to maximize their total utility. That’s all fine, but to truly optimize one’s choices— to discover what’s best among the vast array of alternatives available is hard. Doing it well requires a lot of information and the ability to sort through that information to decide effectively. The result is that consumer decision-making is highly imperfect in practice and a great deal of potential value goes unrealized as a result.

Business analytics helps capture this latent value by improving our choices as consumers: Google’s search engine provides fast, highly relevant web content, enhancing the value of your time online.

Google Maps helps you uncover a faster route home. Pandora’s recommendation system makes it simple to find and enjoy the music you love; Netflix does the same for movies and television. Amazon makes it easy to discover great products to buy and streamlines the purchase transaction. TripAdvisor helps you create great vacations, while OkCupid finds you the perfect mate (well, at least a suitable one!). Each of these technologies enables us to capture value that would go unrealized without the help of data and analytics.

Firms benefit from analytics too. While economic theory suggests firms organize themselves to maximize the value they deliver, real companies frequently waste resources, fail to understand customer needs and hence engage in unproductive activities. Data and analytics help them overcome these inefficiencies. Capital One uses analytics to match credit card offers to customers more accurately than their competition. Walmart uses analytics to manage its inventory in a way that allows it to serve its customers reliably at exceptionally low cost. Cleveland Clinic uses analytics to provide more targeted treatment of patients and to fine-tune therapies that produce better health outcomes at lower costs.
Axioma uses analytics to construct portfolios that provide better risk-reward trade-offs for their clients. In each case, data and analytics helps these firms uncover new opportunities to focus their efforts on value-adding activities and hence increase the gap between the value they deliver and their cost to serve.

Our aim is to teach you how they do it; how to identify and capture analytic value. Toward this end, you will learn about core analytic methods — and do so in the context of case studies of leading organizations that successfully deployed each method.

In the first part of the course, we focus on predictive analytics — the use of data together with machine learning, data mining and statistical forecasting techniques to construct generalizable predictions about future outcomes. Such predictions can be used to determine which patients a hospital should target for treatment or which stocks a high-frequency trader should buy and sell.

In the second part of the course, we focus on prescriptive analytics — how to use experimentation, simulation and optimization to improve business decisions. We will see how such techniques help decide what value to place on GM’s pension fund obligations or how a pharmaceutical company should best utilize its sales force.

Finally, throughout the course we examine the challenges involved in successfully implementing analytics within an organization and how to overcome them — in particular, how to validate a model and develop a delivery system to deploy it.

The course emphasizes that business analytics is not a theoretical discipline; it is a business practice. The techniques are only interesting and important to the extent that they can be used to provide actionable insights and improves important decisions. The concepts learned in this class will help you identify analytic opportunities and alert you to the ways that analytics can be used — and misused—within an organization.

We have three main 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 capturing value using business analytics. The third is to help you assess the potential value added by a given business analytics opportunity.

Business analytics is an integral part of modern management — this course will provide you with the knowledge and skills necessary to productively utilize business analytics throughout your career.

This is a course in statistics, machine learning, and data-driven analytics applied to business environments. This course is designed for students who wish to increase their capability to build, use and interpret data analysis models for business, health care and other quantitative management areas. This course prepares students to gather, describe, and analyze real-world data, use advanced analytical tools to provide scientific guidance in decision making. Students are supposed to have basic knowledge of calculus, probability and statistics and other quantitative background. Students should be comfortable with mathematical formulas and are willing to develop programming skills to analyze data.

Course topics include a review of basis statistical ideas, numerical and graphical methods for summarizing data, linear regression, logistic regression, model choice and false discovery rates, multinomial and binary regression, classification, decision trees, factor models, clustering, cross-validation, decision trees and other emerging data analytics methods. The course presents real-world examples where a significant competitive advantage has been obtained through large-scale data...
analysis. We learn both basic underlying concepts and practical computational skills, including techniques for analysis of distributed data. Examples include advertising, eCommerce, finance, health care, marketing, and revenue management. The ultimate goal is, of course, help to make better business decisions using advanced data analytics.

Attendance Policy

Attendance and class participation are part of each student’s course grade. Students are expected to attend all scheduled class sessions. Failure to attend class will result in an inability to achieve the objectives of the course. Excessive absence will result in loss of points for participation. Regular attendance and active participation are required for students to successfully complete the course.

Class participation is an important part of learning. If you have a question, it’s likely that others do as well. I encourage active participation, and course grades will take into account students who make particularly strong contributions.

Assignments

*Homework:* weekly individual homework assignments, due by the midnight of next class day. All homework assignment should be submitted through the Coursework links.

*Group Projects:* 2-4 students form a group and work on the projects as a team. Students can identify a company or a scenario, collect data, use techniques taught in class to study the data patterns or to predict future outcomes. Students are required to write a 4-6 page project report, and present in class using Power Point slides. Details will be available shortly.

*Final Exam:* the final exam is in-class Closed-book individual written exam.

Late submission including assignments, projects and exams will not be accepted.

Study Group (not required, but highly recommend)

Many students learn better and faster when working in a group, so I encourage collaborative learning. You can work together in a study group with 2-4 students, to discuss class materials, homework assignments and projects on a weekly basis. However, each student must individually write homework assignments to reflect their own understanding of the material. The study groups can be different from your project groups.

Evaluation and Grading

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attendance and participation in class discussion</td>
<td>10%</td>
</tr>
<tr>
<td>Homework</td>
<td>30%</td>
</tr>
<tr>
<td>Project</td>
<td>20%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>40%</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
</tr>
</tbody>
</table>
Important notes about grading policy:
The grade of A+ is reserved for those who demonstrate extraordinarily excellent performance. The grade of A/A- is awarded only for excellent performance. The grade for good performance in this course is a B+/B.