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 Amazon, 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, recent developments in AI and Large Language Models open many new possibilities, and 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 capture 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 capturing value using business analytics. The third is to help you estimate the value to be captured 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 and slides, that will all be posted on canvas. We will also provide a printed book of cases (and preclass assignments), whereas articles and slides will be individually distributed in class.

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.

There will be one concept check quiz after each lecture (from Lecture 2 to Lecture 10). The exact due dates are provided in Canvas. 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 (Total score)/6.)

The concept check quizzes are assignment of type B[1].

Homeworks

Homework assignments will need to be submitted on Canvas. There is one individual assignment (Type B[1]) 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 3.

- 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 Excel and we will provide a business analytics excel add-in. We have developed the add-in to extend the functionality of Excel providing 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. **There is no license associated with this add-in so you may keep it after the course and continue to use it as you see fit.**

We provide use cases with examples of the various functions associated with the add-in.

Attendance Policy

Students are required to attend each class. Students should reach out to the Office of Student Affairs (OSA) by using Core Absence Form on this course's Canvas page regarding excused absences (for religious observances; personal, medical, and family emergencies; military service; court appearances such as jury duty). Unexcused absences will affect your course grade as follows:

- Students that miss more than **33%** of their classes (unexcused absences) will at most receive a **P for the course grade**

- Students that miss more than **50%** of their classes (unexcused absences) will receive an **F for the course grade**

In addition to the effect on your final course grade, absences may also affect your final exam grades as follows:

- Students that miss the exam for an **excused** reason but are unable to take the exam within the stated make-up period will receive a **zero for the final exam grade**

- Students that miss the exam without notifying OSA (**unexcused**), will receive an **F for the course grade**

This course will use PollEverywhere as a tool to increase in-class student engagement. PollEverywhere will also be used to confirm student attendance and participation records. If a student is absent from class and is allowed to attend a different cluster's class meeting, it is at the discretion of the professor to count responses to any polls presented during this time as attendance or participation.

Responding to a poll when not present in the classroom is a violation of the Honor Code.

Inclusion, Accommodations, and Support for Students

At Columbia Business School, we believe that diversity strengthens any community or business model and brings it greater success. Columbia Business School is committed to providing all students with the equal opportunity to thrive in the classroom by providing a learning, living, and working environment free from discrimination, harassment, and bias on the basis of gender, sexual orientation, race, ethnicity, socioeconomic status, or ability.

Columbia Business School will make reasonable accommodations for persons with documented disabilities. Students are encouraged to contact the Columbia University's Office of Disability Services for information about registration. Students seeking accommodation in the classroom may obtain information on the services offered by Columbia University's Office of Disability Services online at https://health.columbia.edu/content/disability-services or by contacting (212) 854-2388.

Columbia Business School is committed to maintaining a safe environment for students, staff and faculty. Because of this commitment and because of federal and state regulations, we must advise you that if you tell any of your instructors about sexual harassment or gender-based misconduct involving a member of the campus community, your instructor is required to report this information to a Title IX Coordinator. They will treat this information as private, but will need to follow up with you and possibly look into the matter. Counseling and Psychological Services, the Office of the University Chaplain, and the Ombuds Office for Gender-Based Misconduct are confidential resources available for students, staff and faculty. "Gender-based misconduct" includes sexual assault, stalking, sexual harassment, dating violence, domestic violence, sexual exploitation, and gender-based harassment. For more information, see https://sexualrespect.columbia.edu/university-policy.

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

Deployment of Analytics and Frontiers

- **Session 11:** Algorithms and Accountability
- **Session 12:** Large Language Models and new Frontier

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.

  Case: GM Pension Liability Evaluation

- **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 of kinds of physical and business constraints. We discuss more specifically the assignment of salesforce representatives to target customers.

- **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:** Algorithms and Accountability

  In this class, we will discuss pitfalls to look for when deploying Analytics across applications, discussing challenges such as bias and algorithmic fairness.

- **Session 12:** Large Language Models and new frontier

  In this class, we will cover recent developments in Artificial Intelligence, and in particular Large Language Models (LLMs). We will discuss the path that lead to the current state of LLMs, the underlying predictive tasks they are solving, and the steps towards deployment in specialized tasks.

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.

– A general overview book of Data Science across applications is

Data Science in Context, Spector, Norvig, Wiggins, and Wing: Cambridge University Press.

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

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

– For more advanced machine learning tools, the following textbooks summarize and illustrate many methods

Business Data Science, Taddy: McGraw Hill

The Elements of Statistical Learning, Hastie, Tibshirani and Friedman
available here: https://web.stanford.edu/~hastie/ElemStatLearn/