Columbia Business School

Analytics in Action; Fall 2022

Daniel Guetta & Brett Martin

"In God we trust; all others bring data" - W. Edwards Deming

Course #:	DROM B8146
Room #:	TBA
Max Capacity:	Roughly 24 MBA, roughly 12 engineers
Instructors:	Daniel Guetta; guetta@gsb.columbia.edu Brett Martin; blm2150@columbia.edu
Class Schedule:	Tuesdays, 6pm-9:15pm
Office Hours:	See below
Teaching Assistants:	Elliot Shin Oblander; EOblander23@gsb.columbia.edu Second TA TBD
Course prerequisites (MBA):	Managerial statistics Business analytics Basic Python Qualification Exam (see gsb.columbia.edu/courses/python)
Highly recommended (MBA):	Business Analytics 2

Columbia Business School

COURSE DESCRIPTION AND LEARNING OBJECTIVES

Companies like Google, Amazon, Microsoft, and Facebook have led the way in developing data-driven applications that have transformed our everyday lives. Based on the success of these data-driven pioneers, business leaders across all industries now realize the need to more effectively harness their own data to improve business operations and decision making. Managers who can effectively transform raw data into actionable insights will not only predict the future but control it.

This course provides students the opportunity to learn business analytics and data science by working on a set of company-sponsored applied projects. Student teams of 5-6 people, with 3-4 MBA students and 1-2 engineering (SEAS) students, will work hand in hand with the instructors and company representatives to achieve company goals through the practical application of data analytics. For example, students may be tasked with translating an e-commerce company's website activity into a data-driven marketing campaign or building a tool to predict which of a SaaS company's accounts are most likely to churn. The list of sponsoring companies spans large firms from financial services, cosmetics, media, and smaller NYC startups. Companies provide the data, faculty provides guidance on best practices, and your team will provide the answers.

Throughout this course, students will execute on a data-driven project to:
- Use tools and ideas from Business Analytics and Python's analysis environment to solve interesting and exciting business problems
- Learn how to formulate relevant business questions that can be answered using data
- Understand the various steps of data preparation like data cleaning and feature extraction
- Break down a complex data problem into multiple smaller, solvable problems
- Evaluate the effectiveness of a solution through statistical testing
- Learn how to iterate on a solution to continually improve it
- Learn how to measure improvement using Key Performance Indicators (KPIs)
- Learn how to collaborate meaningfully with multiple stakeholders
- Communicate results to both technical and non-technical audiences

TARGET AUDIENCE

This course is targeted at students that have a demonstrated interest in analytics, data science, and computer programming but seek the opportunity to implement their interests in a real-world environment.

EXAMPLE COMPANY PROJECTS FROM PAST YEARS

Media: Work with a large media company to build a model to accurately predict the KPIs of a given post, such as the total impressions, views, and engagements. What combination of factors has the greatest impact on the number of impressions? How do the average number of impressions changed over time? Can the model capture this change?

Mobility: Work with a large fleet management firm to proactively identify customers likely to return their vehicle late. Building pricing models to understand the potential effects of various late fees on churn.

Columbia Business School

Software as a Service/Artificial Intelligence: Work with a fast-growing SaaS company to implement a predictive analytics solution to minimize customer service costs. Identify and automate redundant workflows using data-driven algorithms. Build mechanisms to efficiently escalate service requests to humans when automated solutions are deemed likely to fail.

Financial Services: Work with a specialty finance company to build a new credit scoring model that incorporates social signals as well as financial ones. Test the model against real customers in real-time to generate millions of dollars in loans.

Retail/Data Analytics: Work with a large data provider to identify early indicators of retail closures. Identify and illustrate the potential for weather, location, and demographic trends to predict retail headwinds.

COURSE ROADMAP

The course is designed as a guided workshop between student teams and companies, supported by practical lectures from industry experts. Lectures will demonstrate proven approaches to solving common analytics problems like customizing prices for different customers, identifying churn, and identifying items most likely to be bought together. These lectures will also review tools like basic statistics, data analysis in python, statistical modeling, and data visualization. Through these lectures, students will get ideas for their own projects and learn the key skills needed for implementation.

Time in each class will be set aside to give you time to work with your teams with the assistance of the instructors and TAs, though you should, of course, also meet with your teams and companies as necessary outside class times. In addition, each class will comprise a combination of
- lectures on "Analytics in Action" - relevant examples of analytics in the news,
- guest speakers from industry, and
- in the earlier part of the course, lectures on specific analytics-related topics such as visualization, predictive analytics, and data cleaning.

Additional office hours with the instructors and/or TAs can be arranged by appointment.

To ensure regular progress toward the successful completion of student projects, students will be required to submit a progress report, every week, to the course instructors and to their companies. In addition, students will be required to present their findings to the rest of the class at least twice - once for a midterm presentation halfway through the semester, and once for a final presentation at the end of the course. In addition to the final presentation, students will be expected to deliver a final report that could potentially be used as a blog article by the company.

REQUIRED COURSE MATERIALS

No textbook is required; all relevant materials and links will be posted to Canvas.

CLASSROOM NORMS AND EXPECTATIONS

- **Class will start on time.** Show up five minutes early for the good seats.

- **Stow your computer and cell phones during lectures and presentations unless otherwise instructed*.** If you have any questions, let us know. If the class is moving too quickly or slowly, please let us know. We and your fellow students appreciate your courtesy.
- **Bring a notebook and a pen for notes.** Since cell phones and computers are prohibited during class, you'll need some way of taking notes.
- **Be brave.** You stand to lose more by toeing the line than by genuinely putting yourself out there, as long as you think before you speak.
- **Be courteous.** Please be respectful and professional toward your fellow classmates. We wholeheartedly encourage vigorous debate but please don't be nasty, aggressive, or condescending. If you feel uncomfortable in class for any reason (the material, other students, me, you, anything), please confidentially email us or the TA. We will do everything I can to make our class as hospitable as possible but we can only do so if we are aware of the situation.

*Pending COVID circumstances situation…

INSTRUCTORS

Brett Martin is a Brooklyn-based entrepreneur and investor. He is the managing director of Charge Ventures, an early-stage venture capital fund. Previously, he co-founded @GetSwitch, the easiest way for passive job seekers to find their next job on their mobile phones, and @Sonar, a popular location-based mobile app that leveraged social networks to connect millions of people in the physical world. Before that, Brett built K2 Media Labs, a New York-based seed stage mobile incubator, launched Vice Magazine's web presence (VBS.tv), and worked on Wall Street as a senior equity research associate at Thomas Weisel Partners. He graduated from Dartmouth College with an A.B. in Economics.

Daniel Guetta is the Director of the Business Analytics Initiative jointly led by the Columbia Business School and Columbia Engineering, and a Lecturer in Discipline at the Columbia Business School. He teaches classes in business analytics, including data science, pricing, and supply chain management. Prior to joining the faculty at Columbia, he was a Deployed Computational Engineer and Engagement Manager at Palantir Technologies, where he worked with clients in fields from finance to pharmaceuticals to help them solve their hardest problems using data. He completed his undergraduate studies in physics and mathematics at Cambridge and MIT, and holds a Ph.D. in Operations Research from the Columbia Business School.

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.

Students seeking accommodation in the classroom may obtain information on the services offered by Columbia University's Office of Disability Services online at www.health.columbia.edu/docs/services/ods/index.html or by contacting (212) 854-2388.

Columbia Business School

TENTATIVE COURSE TIMELINE

Please bring a laptop to every class. The first part of each class will comprise a presentation from one or both of the instructors, followed by office hours, in which the instructors will work with each group individually on their projects.

Week 1
Tuesday, Sept 6th
 	
Deliverable: list of project/team preferences due by end of class
	Introduction & Project Selection
	● *Pre-class assignment*
	  ○ Install Anaconda
	● *In-class agenda*
	  ○ Presentations from each partner company on their projects
	
	At the end of class, you will be asked to rank projects in order of preference, and to indicate any preferences in re: what teammates you would like to work with. Based on this feedback, we will create teams and assign projects

Week 2
Tuesday, Sept 13
 	
Deliverable: weekly standup due Monday at midnight, emailed to instructors and TAs
	● *In-class agenda*
	  ○ Standups
	  ○ Office hours

Week 3
Tuesday, Sept 20
 	
Deliverable: weekly standup due Monday at midnight, emailed to instructors and TAs
	● *In-class agenda*
	  ○ Standups
	  ○ Office hours

Week 4
Tuesday, Sept 27
 	
Deliverable: weekly standup due Monday at midnight, emailed to instructors and TAs
	● *In-class agenda*
	  ○ Standups
	  ○ Office hours

Week 5
Tuesday, Oct 4
 	
Deliverable: weekly standup due Monday at midnight, emailed to instructors and TAs
	● *In-class agenda*
	  ○ Standups
	  ○ Office hours

Week 6 *Wednesday, October 11* *Deliverable*: weekly standup due Monday at midnight, emailed to instructors and TAs Begin preparing the mid-term demo	No class; use time for midterm presentation to companies
Tuesday, October 18 *Deliverable*: weekly standup due Monday at midnight, emailed to instructors and TAs. **Include feedback from company presentations**	**No class**: study week

Week 7
Tuesday, October 25
 	
Deliverable: no weekly standups required.
	Midterm demos
	● *Pre-class assignment*
	   ○ Prepare mid-term demos, covering project overview, data analysis, preliminary results, and timeline
	● *In-class agenda*
	   ○ Midterm demos
	
	You will be expected to provide feedback for every other team's midterm demo by the day after class.

Week 8
Tuesday, November 1
 	
Deliverable: weekly standup due Monday at midnight, emailed to instructors and TAs.
	● *In-class agenda*
	   ○ Standups
	   ○ Office hours

Tuesday, November 8 *Deliverable*: weekly standup due Monday at midnight, emailed to instructors and TAs.	**No class**: election day

Week 9
Tuesday, November 15
 	
Deliverable: weekly standup due Monday at midnight, emailed to instructors and TAs.
	● *In-class agenda*
	   ○ Standups
	   ○ Office hours

Begin preparing final presentations	

Tuesday, November 22
 	
No weekly standups required
	- *In-class agenda*
	    ○ Standups
	    ○ Office hours

Week 10
Wednesday, November 29
 	
Deliverable: no weekly standups required. Final demo in class
	Final Presentations
	- *Pre-class assignment*
	    ○ Prepare final presentation
	- *In-class agenda*
	    ○ Final presentations

Week 12 *Tuesday, December 6* *Deliverable*: no weekly standups required.	**No class**: final presentation to companies
Final deliverables: due by noon on Wednesday, December 13	

Columbia Business School

GRADING AND EVALUATION

Grading for this course will be based on the following four components:

Weekly standups with companies	20%	Standups are an essential part of the class, and writing a good standup is an essential skill. As such, 20% of your grade in the class will be awarded based on your standups. **Standups are limited to one, single-sided page, <u>including </u>all code, plots, etc...** **Standups will be graded out of 10 as follows** ● 0 : no standup, late standup, or standup > 1 page long ● 3 : rough, shoddy, or hard to read standup ● 10 : solid effort (most teams will get 10/10) ● 12 : (extra credit) outstanding work, we'd hire you Any standup that includes figures with missing or illegible axes will automatically score 3/10. You should store all of your standups in a google doc - the first page should be the most recent standup, the second page your next most recent, etc... Every week before Monday at midnight, you should email both instructors and TAs a link to this Google doc.
Midterm presentation and feedback on other teams'	10%	Your feedback to other teams will be essential and will help you hone your ability to evaluate analytics presented to you. 10% of your grade in the class will be awarded based on the feedback to give to the other teams' midterm presentations. If you make a good-faith effort at providing substantive feedback, you will score 10/10 in this category.
Final presentation	25%	A grading rubric for this component will be circulated separately.
Final blog post	25%	You will be required to produce a final blog post summarizing your work. A grading rubric for this component will be circulated separately.
Participation	20%	Class participation will be measured by attendance and contributions during class, and by pre-class preparation. Please come to class prepared and ready to participate actively. The success of this class depends on everyone's gracious sharing of their perspective, opinions, and experiences. We will prod, poke, and instigate debate with questions based on the readings, assignments, current events, and/or in-class discussion.