

# Analytics in Action

*Daniel Guetta & Brett Martin*

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

<b>Course #:</b>	B8146 sec. 1: Analytics in Action
<b>Room #:</b>	WJW 416
<b>Max Capacity:</b>	24 MBAs, 12 ME
<b>Instructors:</b>	Daniel Guetta; <a href="mailto:guetta@gsb.columbia.edu">guetta@gsb.columbia.edu</a> Brett Martin; <a href="mailto:blm2150@columbia.edu">blm2150@columbia.edu</a>
<b>Class Schedule:</b>	Tuesdays, 6pm - 9pm Full Term
<b>Office Hours:</b>	See below
<b>Teaching Assistants:</b>	TBD
<b>Course pre-requisites:</b>	Managerial statistics Business analytics Python programming

## 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 have greatest impact on 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.

**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.

At least an hour of 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 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 questions, let us know. If 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 towing 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.

## **INSTRUCTORS**

**Brett Martin** is 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 PhD 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](http://www.health.columbia.edu/docs/services/ods/index.html) or by contacting (212) 854-2388.

**TENTATIVE COURSE TIMELINE**

Please bring a laptop to every class.

<p><b>Week 1</b> <i>Tuesday Sept 4th</i></p> <p><i>Deliverable:</i> list of project/team preferences due by end of class</p>	<p><b>Introduction &amp; Project Selection</b></p> <ul style="list-style-type: none"> <li>● <i>Pre-class assignment</i> <ul style="list-style-type: none"> <li>○ Install Anaconda</li> </ul> </li> <li>● <i>In-class agenda</i> <ul style="list-style-type: none"> <li>○ Ensure everyone has a working Python environment; introduce Jupyter notebook</li> <li>○ Presentations from each partner company on their projects</li> </ul> </li> </ul> <p>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</p>
<p><b>Week 2</b> <i>Tuesday Sept 11th</i></p> <p><i>Deliverable:</i> email to company, with instructors and TAs cc'ed, due by end of class</p>	<p><b>Data Cleaning</b></p> <ul style="list-style-type: none"> <li>● <i>Pre-class assignment</i> <ul style="list-style-type: none"> <li>○ Read “<i>Dynamic Pricing for Online Marketplace Sales</i>” case</li> </ul> </li> <li>● <i>In-class agenda</i> <ul style="list-style-type: none"> <li>○ Analytics in Action</li> <li>○ Data cleaning in Python</li> <li>○ Teams to meet for the first time, and establish initial contact with companies</li> <li>○ Office hours</li> </ul> </li> </ul>
<p><b>Week 3</b> <i>Tuesday Sept 18th</i></p> <p><i>Deliverable:</i> weekly stand-up to be emailed to company, with instructors and TAs cc'ed, by end of class</p>	<p><b>Data Visualization</b></p> <ul style="list-style-type: none"> <li>● <i>Pre-class assignment</i> <ul style="list-style-type: none"> <li>○</li> </ul> </li> <li>● <i>In-class agenda</i> <ul style="list-style-type: none"> <li>○ Analytics in Action</li> <li>○ Data visualization in Python</li> <li>○ Office hours</li> </ul> </li> </ul>
<p><b>Week 4</b> <i>Tuesday Sept 25th</i></p> <p><i>Deliverable:</i> weekly stand-up to be emailed to company, with instructors and TAs cc'ed, by end of class</p>	<p><b>Predictive Analytics</b></p> <ul style="list-style-type: none"> <li>● <i>Pre-class assignment</i> <ul style="list-style-type: none"> <li>○ Read “<i>Data-Driven Strategies for Peer-to-Peer Lending</i>” case</li> </ul> </li> <li>● <i>In-class agenda</i> <ul style="list-style-type: none"> <li>○ Analytics in Action</li> <li>○ Predictive analytics in Python</li> <li>○ Office hours</li> </ul> </li> </ul>

<p><b>Week 5</b> <i>Tuesday October 2nd</i></p> <p><i>Deliverable:</i> weekly stand-up to be emailed to company, with instructors and TAs cc'ed, by end of class</p>	<p><b>Predictive Analytics; relationships among records</b></p> <ul style="list-style-type: none"> <li>● <i>Pre-class assignment</i> <ul style="list-style-type: none"> <li>○</li> </ul> </li> <li>● <i>In-class agenda</i> <ul style="list-style-type: none"> <li>○ Analytics in Action</li> <li>○ Mining relationships among records in Python</li> <li>○ Office hours</li> </ul> </li> </ul>
<p><b>Week 6</b> <i>Tuesday October 9th</i></p> <p><i>Deliverable:</i> weekly stand-up to be emailed to company, with instructors and TAs cc'ed, by end of class.</p> <p>Begin preparing mid-term demo</p>	<p><b>Predictive Analytics; time series data</b></p> <ul style="list-style-type: none"> <li>● <i>Pre-class assignment</i> <ul style="list-style-type: none"> <li>○ Read "<i>Predicting Sales at a Supermarket Chain</i>"</li> </ul> </li> <li>● <i>In-class agenda</i> <ul style="list-style-type: none"> <li>○ Analytics in Action</li> <li>○ Time series data</li> <li>○ Office hours</li> </ul> </li> </ul>
<p><i>Tuesday October 16th</i></p> <p><i>Deliverable:</i> weekly stand-up to be emailed to company, with instructors and TAs cc'ed, by end of day.</p>	<p><b>No class:</b> study week</p>
<p><i>Tuesday October 23rd</i></p> <p><i>Deliverable:</i> weekly stand-up to be emailed to company, with instructors and TAs cc'ed, by end of day.</p>	<p><b>No class:</b> fall break</p>
<p><b>Week 7</b> <i>Tuesday October 30th</i></p> <p><i>Deliverable:</i> by the day after class (<i>Wednesday October 31st</i>), paste your comments on each other team's project in the relevant class Google Doc.</p>	<p><b>Midterm demos</b></p> <ul style="list-style-type: none"> <li>● <i>Pre-class assignment</i> <ul style="list-style-type: none"> <li>○ Prepare mid-term demos, covering project overview, data analysis, preliminary results, and timeline</li> </ul> </li> <li>● <i>In-class agenda</i> <ul style="list-style-type: none"> <li>○ Midterm demos</li> </ul> </li> </ul> <p>You will be expected to provide feedback for every other team's midterm demo by the day after class. We will not be meeting on Week 8; you will use the next two weeks to update your midterm presentation in light of the feedback you receive, and to meet with your company and present it to them.</p>
<p><i>Tuesday November 6th</i></p>	<p><b>No class:</b> election day</p>

<p><b>Week 8</b> <i>Tuesday November 13th</i></p> <p><i>Deliverable:</i> by the day before next class (<i>Monday November 20th</i>), email company, with instructors and TAs cc'ed, with summary of feedback from midterm presentation.</p>	<p><b>No class:</b> midterm presentation to companies</p>
<p><b>Week 9</b> <i>Tuesday November 20th</i></p> <p><i>Deliverable:</i> weekly stand-up to be emailed to company, with instructors and TAs cc'ed, by end of class.</p>	<p><b>Guest Lecture</b></p> <ul style="list-style-type: none"> <li>● <i>Pre-class assignment</i> <ul style="list-style-type: none"> <li>○ N/A</li> </ul> </li> <li>● <i>In-class agenda</i> <ul style="list-style-type: none"> <li>○ Guest lecture</li> <li>○ Office hours</li> </ul> </li> </ul>
<p><b>Week 10</b> <i>Tuesday November 27th</i></p> <p><i>Deliverable:</i> weekly stand-up to be emailed to company, with instructors and TAs cc'ed, by end of class.</p> <p>Begin preparing final presentations</p>	<p><b>Guest Lecture</b></p> <ul style="list-style-type: none"> <li>● <i>Pre-class assignment</i> <ul style="list-style-type: none"> <li>○ N/A</li> </ul> </li> <li>● <i>In-class agenda</i> <ul style="list-style-type: none"> <li>○ Guest lecture</li> <li>○ Office hours</li> </ul> </li> </ul>
<p><b>Week 11</b> <i>Tuesday December 4th</i></p> <p><i>Deliverable:</i> weekly stand-up to be emailed to company, with instructors and TAs cc'ed, by end of class.</p>	<p><b>Final Presentations</b></p> <ul style="list-style-type: none"> <li>● <i>Pre-class assignment</i> <ul style="list-style-type: none"> <li>○ Prepare final presentation</li> </ul> </li> <li>● <i>In-class agenda</i> <ul style="list-style-type: none"> <li>○ Final presentations</li> </ul> </li> </ul>
<p><b>Week 12</b> <i>Tuesday December 11th</i></p> <p><i>Deliverable:</i> slides and final report due by the end of the week, Friday December 14th</p>	<p><b>No class:</b> final presentation to companies</p>

**GRADING AND EVALUATION**

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

Weekly standups with companies	20%	<p>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.</p> <p>This component will be graded on a 0/80/100 scale. As long as your team makes a good effort at standups, you will receive 80%. Exceptional work (no more than one or two teams) will receive 100%. If you fail to complete more than one standup, you will receive a 0.</p>
Midterm presentation and feedback on other teams'	10%	<p>You 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.</p> <p>This component will also be graded on a 0/80/100 scale.</p>
Final presentation and report	50%	<p>Your final presentation and report will form the majority of your grade. A grading rubric for this component will be circulated separately.</p>
Participation	20%	<p>Class participation will be measured by attendance and contributions during class, and by pre-class preparation.</p> <p>Please come to class prepared and ready to participate actively. The success of this class depends on everyone's gracious sharing of their perspective, opinion, and experience. We will prod, poke, and instigate debate with questions based on the readings, assignments, current events and/or in-class discussion.</p>