Analytics in Action

Divyanshu Vats & Brett Martin

“In god we trust; all others bring data” - W. Edwards Deming

Course #: B8146 sec. 1: Analytics in Action

Room #: URI 332 (subject to change, please check Canvas)

Max Capacity: 24 MBAs, 12 ME

E-mail: brettlucasmartin@gmail.com / vats.div@gmail.com

Class Schedule: Tuesdays - Full Term

Class Hours: 5:45:00 PM to 9:00:00 PM

Office Hours: See below

Teaching Assistants: Raymond Li (yl3551@columbia.edu) & Raphael Ruscassie (raphael.ruscassie@columbia.edu)

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. Students 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 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

Example Company Projects (Subject to Change)

Ecommerce/Telemedicine: Work with a fast growing ecommerce company to better understand the connection between onsite user behavior and purchasing decisions. When are users willing to spend more? How long are users using the product? Can we predict when a customer needs new product (to preemptively market to them and capture spend)? Can we predict reorders? Can we build a model for who will reorder and why (and why not so that we can proactively drive retention)?

Financial Services: Work with a large bank to look at "people like you" and compare their income and spending patterns to yours...but instead of simply looking at current similar or "baseline" users, focus on consumers who were like you 3 years ago, but now are better off. What did they change to improve their state? Did they spend less, or spend more but increase their income? Did they take on debt (i.e., make payments to mortgage or debt companies) or did they change their location? This project aims to create an approach that would illustrate how "people like you" achieved some measure of financial success, by looking at their financial data.

SaaS/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.

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 PRE/CO-REQUISITES

○ Managerial Statistics
○ Business Analytics
○ Python Programming (co-requisite)

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 identifying repeat customers 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.

To ensure regular progress toward the successful completion of student projects, the students will be required to meet their companies on a monthly basis, present demos in-class, present a demo to the company, and deliver a final report that could potentially be used as a blog article by the company.

**TENTATIVE COURSE TIMELINE**

One-third of the course will have in-class lectures, one-third of the class will consist of hand-on “office hours” where students will work on projects with the instructors, and the final one-third of the course will be company visits and presentations.

**Week #1 (Sept 5, 2017)**
- Lecture: Introduction and project selection
  - Before the class begins, instructors will form teams based on student profiles.
  - During class, the teams will rank their preference for projects
- Homework #1
  - (Group) Answer analytical questions about the project data (designed by the instructors for each project) (5 points)
  - (Individual) Reading material on data formats and related questions (5 points)

**Week #2 (Sept 12, 2017)**
- Office hours and meet the companies
- Homework #2
  - (Group) Report on company visit and preliminary project plan (5 points)

**Week #3 (Sept 19, 2017)**
- Lecture: Data-driven problem solving using Instacart data
  - [https://tech.instacart.com/3-million-instacart-orders-open-sourced-d40d29ead6f2](https://tech.instacart.com/3-million-instacart-orders-open-sourced-d40d29ead6f2)
- Homework #3
  - (Group) Analysis of how the project will improve the established KPI set forth by the company. (5 points)

**Week #4 (Sept 26, 2017)**
- Office hours and prep for demo

**Week #5 (Oct 3, 2017)**
- Midterm demo
  - (Group) 15-20 minute presentation covering project overview, project data analysis, preliminary results, and timeline. (15 points)
- Homework #4
  - (Individual) Give detailed feedback on your classmates’ Demos (5 points)
Week #6 (Oct 10, 2017)
- Office hours and meet the companies

Week #7 (Oct 31, 2017)
- Lecture: Example and tools/tips for telling effective data stories
  - We will use the Instacart and various plotting libraries in python to understand the data and tell useful and informative data stories. The content from this lecture will be useful for students to tell data stories for their own projects.
- Homework #5
  - (Individual) Answer visualization based questions for data presented in class (5 points)

Week #8 (Nov 14, 2017)
- Guest Lecture: Guthrie Collin (Senior Manager, Technical Products at Amazon) on TFIDF
- (Group) A Data story based on project data (5 points)

Week #9 (Nov 21, 2017)
- Lecture: The Business of Data Science
- Homework #6
  - (Individual) Research a company started in the last 5 years and discuss how they use data and analytics for business growth. (5 points)

Week #10 (Nov 28, 2017)
- Guest Lecture: Paul Fisher (CTO at Phosphorus) on Data Science Implementation in Production Environments

Week #11 (Dec 5, 2017)
- In-class Final Demo
  - (Group) 15-20 min presentation to the class (15 points)

Week #12 (week of Dec 11, 2017)
- (Group) On-company location Final Demo
- Homework #7
  - (Group) Final report that could be used by the companies as a blog article. (10 points)

A Note on Office Hours: These office hours are designed to be scheduled meetings between the student groups and the instructors. The goal of these meetings is to get updates from the students on their progress, answer any questions about the projects, both technical and non-technical, and get students ready and prepared for their meetings with the companies. These meetings are flexible in location and could be held off-site or at the company location. Note that we will have regular weekly office hours in addition to these scheduled office hours.

ASSIGNMENTS/METHOD OF EVALUATION

Grading for this course is based on a combination of class participation (20%) as well as individual (20%) and group assignments (60%). Class participation will be measured by attendance in class as well as during scheduled office hours.

Participation

- 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.
Individual and Group Assignments

Assignments are designed to help guide groups down the path to the successful completion of the company project. For assignments that require coding, we expect all code and relevant data to be submitted as well. We have given a high level overview of the assignments. We will give out more details on each individual assignment when the class starts.

REQUIRED COURSE MATERIALS

- No textbooks required. From time to time, we may share relevant links via Canvas.
- Bring a laptop to class. Most of class we’ll ask you to put your laptops away so that we can focus on the discussion, but there may be exercises that require a laptop with internet access. Bring your charger.

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.

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.

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.

Divyanshu (Div) Vats is a Quantitative Researcher at Two Sigma Investments. Before joining Two Sigma, he worked at Sailthru as a Senior Data Scientist, where he designed and developed data-driven products using advanced Machine
Learning to enable marketers to optimize customer lifetime revenue and customer engagement. Divyanshu got his PhD from Carnegie Mellon University and undergraduate degree from The University of Texas at Austin.

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.