Data Driven Dollars (B8696)  
How Businesses Transform Data Assets Into Revenues  
Spring 2021 Syllabus

GENERAL INFORMATION

Professor: Yuval Ariav (yuval.ariav@gmail.com)
Room: TBA
Time: TBA
Dates: TBA

Overview: Data assets represent a new and increasingly important corporate asset class. Data Driven Dollars is a course about these assets and the unique business models that they drive. The course teaches MBAs the theoretical and practical aspects of developing and managing data assets, as well as leveraging them to drive revenue in both startup and corporate environments.

As humanity becomes progressively more dependent on digital processes, more elements of everyday life and business reality are captured in digital form: consumers’ location data, financial history and browsing habits, as well as company history and internal workflows across the corporate landscape.

This exponential increase in “reality mining” has given rise to a new breed of corporate revenue streams and opportunities, where data is the product, or at the very least an indispensable component of it. Due to the unique processes needed to support the collection, cleaning, storage, processing, and presentation of data, building and running businesses around data assets is a type of expertise in and of itself.

Objectives: This course teaches students about what it means to build and run such a business. It provides the basic knowledge required to evaluate, direct, and manage data asset related processes, and the vocabulary to enable effective conversations and collaboration across and within organizations. It discusses the different data monetization strategies and further provides a framework for evaluating if and how data assets can be leveraged to create additional revenue opportunities for the business.
**Format**

Data Driven Dollars will be a 6-class half-term course.

Most classes will have both a theoretical and a practical element to them. The theory part will consist of a frontal presentation and/or discussion outlining the premise of the class’ topic, defining terminology and discussing challenges, and opportunities. The practice part will provide the students with concrete, real-life example(s) of the theory through presentations by external speakers, exercises, or projects.

Additionally, most classes will start with a 10-15-minute *data to revenue brainstorming* session. These sessions will begin with a presentation of a real-life dataset that is monetizable in a non-trivial way. The class will then ideate on how the data can be used to generate revenue, after which an actual monetization model will be presented. The goal of these sessions is to broaden students’ horizons about the business applications of data.

**Grading**

50% Class Participation  
50% Assignments (individual + group projects)

**Participation**

This course adheres to the Columbia Core Culture, and students are expected to be Present, Prepared, and Participating.

Class Attendance is assumed, and active, constructive participation is expected. Much of the time in classes will be spent in open discussion and conversation, and high-quality contribution and engagement will go toward the final grade.

**COURSE PLAN**

**Class 1**

**Background - The Data Revolution**

After a brief introduction of the professor and the course, we’ll dive into an overview of the last two decades with respect to humanity’s data generation, processing, storage, and learning capacity. This will be followed by discussing the history of the data-rich industries and the evolution of data-related business workflows, as well as methodologies for data and reality mining.

We’ll discuss examples of companies generating new types of digital data and how they make money, and talk about social data and its impacts on consumer behavior and business opportunities. We’ll proceed by defining what data assets are and aren’t, validate our definition by looking at real-
life examples, and talk about the rate of data asset development across different industries.

Class 2

The Building Blocks of Data Assets

In this class, we’ll learn about how modern data operations happen. First, we’ll cover the key workflows involving modern data assets:

- Collection
- Aggregation
- Sanitation and distillation
- Enrichment
- Siloing, compliance and protection
- Processing
- Reporting and presentation

Then, we’ll talk about the challenges presented by large data assets and the constraints they present to scale, e.g. volume, access frequency, broad data vs. deep data, etc.

We’ll cover modern data operations frameworks, such as Hadoop and Kafka, and learn about how they fit into the modern data operations architecture.

Class 3

Big Data and Machine Intelligence

Big Data, Machine Learning, and Artificial Intelligence are arguably the most (over)used buzzwords in the tech innovation ecosystem in recent years.

In this class, we’ll strive to provide a hype-less definition of what Big Data is, dispelling some of the preconceptions surrounding it. We’ll then cover some real-world examples from leading companies, and discuss the scope and limits of this technology.

We’ll then move on to Machine Intelligence, and talk about example implementations, from simple rule-based engines to complex cutting-edge algorithms, discussing the pros and cons of each approach. We’ll go through an illustration of a very simple Machine Learning algorithm to start developing an intuition into how these algorithms work.
Class 4  
**Outbound Data Monetization Strategies**
In this class, we’ll build on our understanding of how data operations work from the previous classes to discuss the different strategies used by companies to monetize their data assets.

We’ll start by reviewing valuation methods for other intangible assets, and proceed with an overview of a framework that can be used to value corporate data assets.

We will then cover the main outbound data monetization strategies, including the Data Licensing Model and the Oracle Model.

At the end of this class we will talk briefly about the scope of the group assignment that will be presented during the last class.

Class 5  
**Inbound Data Monetization Strategies**
Picking up from Class 4, we’ll discuss inbound data monetization strategies, including the Premium Service Model, the Platform/Broker Model, and the Advertising Model.

We’ll go through case studies of each of these models and understand why these models were chosen, and how the characteristics of the underlying data assets impact performance and execution.

Class 6  
**Starting and Running a Data Asset-Rich Business**
In the final class of this course we’ll discuss frequent challenges of starting and running companies that are data-rich.

We’ll talk about the cold start problem and how it impacts new businesses, including examples and the ways to overcome it.

We’ll then talk about what first mover advantage means for data-heavy companies, and what competitive advantages look and behave like in businesses that are data-driven.

We’ll briefly cover legal and regulatory frameworks that govern data assets, how companies may be impacted by them, and how the operational implications they bring about.

---

Columbia Business School
Finally, each group will present its assignment to the class, followed by a brief discussion.