HCMG 857 – HEALTH CARE DATA & ANALYTICS

Fall 2018 Q1 M: 3:00-6:00 pm Location: CPC Auditorium

Instructors: Matt Grennan, PhD grennan@wharton.upenn.edu
Hari Mahadevan, PhD MBA harimahadevan@gmail.com

Teaching Assistant: Stuart Craig stucraig@wharton.upenn.edu

Office Hours: by appointment

Course Description: In healthcare or anywhere else across science, or business, or sports, the importance of data and analytics ("D&A") is virtually unquestioned. That, however, does not mean that it needs no elucidation. In this course, we begin with a fundamental understanding of the state of D&A in healthcare and then move onto examples of its use in converting from business questions to implemented solutions.

We "sidestep" into the world of algorithms/machine-learning/AI and causal inference, but *our focus is on business applications of these tools to the available data in the healthcare industry*. As we discuss examples, we always seek to show how human creativity needs to be at the heart of the questions being probed. We highlight today's data universe in healthcare, the level of integration we have achieved, and the immensity of the remaining task, all with an eye to the business opportunities that exist now. We end with a showcasing of the art of the possible – in 2018 – and with (hopefully) a clear look ahead at what remains to be achieved.

At the end of this short course, students will:

1. Know the health care data landscape
2. Understand the "loop" that drives modern evidence-based businesses
3. Experience a deep dive into a real health care data analytics problem
4. Anticipate the business opportunities evolving in health care data analytics

Recommended Texts: The notes and handouts and class discussions will be sufficient. For additional reading, we recommend the following:

1. For "Data Science":
   Provost, Foster & Tom Fawcett (2013); *Data Science for Business*, O'Reilly.
   Annalyn Ng & Keneth Soo (2017); *Numsense! Data Science for the Layman: No Math Added,* Self-published
2. For "Causal Inference": Joshua D. Angrist & Jörn-Steffen Pischke (2014); *Mastering 'Metrics: The Path from Cause to Effect*, Princeton.
3. For background and futurism: Pedro Domingos (2015), *The Master Algorithm: How the Quest for the Ultimate Learning Machine will remake our world*, Basic Books.

Course Requirements: A curious mind, enthusiastic contribution to the class discussion (30% of your grade), a team-based problem set (20% of your grade), a team-based case write-up (20% of your grade), and a team-based final project (30% of your grade) form the bulk of preparation needed for the course.

A successful course will depend on your engagement with and contribution to the classroom discussion. We welcome questions, sharing of experiences, and an intellectual curiosity around anything discussed. Each lecture is designed to be part didactic and part conversation. Your professors and guest lecturers don't have all the answers but know how to ask good questions. We want the same from you; every student is expected to participate in all aspects of each lecture, case study discussion, as well as in the guest lectures.

The case write-up should not exceed 4 pages, excluding any supporting material—this page limit is purposefully short, so you will face hard decisions about what NOT to include. If you use a bunch of complex terminology or try to hit every potentially relevant point, you will run out of space. A single well-developed idea within an identified MECE (mutually exclusive, collectively exhaustive) set is better than ten vague ones.

Given the short length of this course, final projects will be limited to well thought-out "project proposals" to put the health care dataverse to use (as opposed to a full-fledged "take the project to completion", which we will reserve for future semester-long courses). By project proposal we mean: (1) a clearly defined question to be answered, (2) motivation for what we gain by knowing the answer to this question (think of what business problem/problems your project will solve for), (3) the data and toolkit that will be needed inclusive of analytical approach you propose, (4) anticipated challenges and the caveats they will introduce (or plans to overcome them), and (5) what insights and business implications you expect to furnish and (importantly) what additional lines of questioning the answer could reasonably be expected to lead to. The project should be of scope that a first attempt could be finished by a small team in a few weeks. Proposals should not exceed 10 pages.

Course Prerequisites: Pursuit of MBA – we expect most students to be 2nd years. We assume no prior knowledge of statistical tools/concepts beyond the level of the MBA Core, though of course further training in data science is a plus, and we welcome those with more advanced preparation. Where necessary, we cover the modern "big data" toolkit, but at a high level, with references to details for those unfamiliar. Our primary focus will remain on the data and applications in health care.

Course Outline:

I. "The Messy World of Healthcare Data" (September 10)

A quick-paced jump into the immensity and complexity of the world of healthcare data. An understanding of the types of companies in the various niches of this world and what they are individually/collectively collecting and attempting to achieve. Also, a look at M&A activity and what that might portend, as well as the role of governments and regulatory organizations. A look ahead into what the rest of the class will deliver across the mix of strategic insights and expected toolkit development. Why – and how – the use of data and its conversion into insights is in such a state of flux and what makes it fascinating!

Discussions:

- Intro to the Health Care Dataverse – Hari
- Intro to the Health Care Data *Analytics* space - Matt

Readings:

- Hsu, Nowak, and Smith (2017) *"Reframing Data and Analytics: Transforming Insights into Action" NEJM Catalyst.*
- Andrew Ng (2017) "*Uncovering The Real Value Of AI In Healthcare" Rock Health Podcast.*

II. "Data – It's a Business" (September 17)

We have learned what types of data exist, but what types of companies are collecting it and, most importantly, what are they doing with it? A live example from our guest lecturer(s) on highly current applications of insights developed and actions being taken today from the collection and analytics of large datasets. In some detail, what does the process of turning a business problem/question into an implemented solution actually entail?

Discussions:

- The Business of Health Care Data – Hari + Guests (Kepler, Inc)
- Case Study: Kyruus, Inc. Big Data's search for the killer app – Matt

Readings:

- *Kyruus, Inc. Big Data's search for the killer app,* HBS Case Study
- See *Canvas* for additional

III. "Is it all about Machines?" (September 24)

Let's demystify the buzzwords. What is Machine Learning and how is it different from AI? Where is this being used in healthcare? How? Lecture will continue to build on the business problem to implemented solution thread begun in Lecture 2 with an emphasis on the role of creativity in helping solve business issues (it isn't all about the "machines"!) and a turn doing using hypothesis-driven problem solving and data analytics from a complex, messy, large data set.

Discussions:

- ML and AI in Healthcare Today – Hari
- Application: ML for Estimating Causal Intervention Effects - Matt

Readings:

- Recommended texts
- Grennan, Matthew and Myers, Kyle and Swanson, Ashley Teres and Chatterji, Aaron, Physician-Industry Interactions: Persuasion and Welfare (July 17, 2018). Available at SSRN: https://ssrn.com/abstract=3216172
- See *Canvas* for additional

IV. "Perhaps it's all about the Analytics?" (October 1)

Real-life problem solution using hypothesis-driven problem solving and data analytics from a different complex, messy, large data set -- a case study and perspective from the field.

Discussions:

- Case Study: Intermountain Health; Precision Medicine – Hari
- Guest: Benson Hsu, MD, MBA, FAAP; Chief Medical Analytics Officer, Sanford Health https://www.linkedin.com/in/bensonhsumdmba/

Readings:

- *Intermountain Health; Precision Medicine,* HBS Case Study

V. "The Art of the Possible" (October 8)
Panel discussions by a set of industry practitioners across various facets of the healthcare world.

Discussions:

- David Lindsay, CEO and Co-founder, Oncora Medical https://www.linkedin.com/in/wdlindsay/
- Michael Ryan, SVP US Market Access & Government Affairs, Bristol Myers Squibb Pharmaceuticals https://www.linkedin.com/in/michael-ryan-38367b6a/
- Suzanne Sawyer, VP and CMO, Penn Medicine https://www.linkedin.com/in/suzsawyer/
- JT Treadwell, Managing Partner, STG Partners https://www.crunchbase.com/person/j-t-treadwell

VI. "It's Your Turn" (October 15)
Team presentations and discussions of project proposals.