

MGMT 298D - HEALTHCARE ANALYTICS

Spring 2020

INSTRUCTOR Professor Elisa Long
elisa.long@anderson.ucla.edu

TEACHING ASSISTANT Jenny Brewster
jennifer.brewster.2020@anderson.ucla.edu

CLASS TIMES Wednesdays, 1:00-3:50pm
Zoom link posted on CCLE

COURSE WEBSITE <https://ccle.ucla.edu/course/view/20S-MGMT298D-7>

COURSE DESCRIPTION

With US healthcare spending reaching 18% of GDP and the demand for health services continuing to increase, improvements in the quality and efficiency of healthcare delivery are needed. This course explores how analytics tools can be used to describe relationships, predict outcomes, and prescribe solutions to healthcare challenges. Classes will include a combination of lectures, data analysis and software demonstrations, group break-out sessions, and student-led presentations.

TARGET AUDIENCE

Students interested in consulting, data science, pharma/biotech, medical devices, or non-profits. Although we will be mostly discussing healthcare applications, the general methodologies apply across a range of industries and business sectors. No prior experience in healthcare is needed.

LEARNING OBJECTIVES

- Identify applications of healthcare data analytics and the trade-offs of various tools.
- Gain practical experience with developing quantitative tools and performing empirical analyses.
- Become familiar with working with data in R, and running linear and logistic regressions.
- Improve public speaking and presentation skills using remote technologies.

COURSE GRADING

- In class participation and lab submissions (30%)
- Three in-class discussions/debates (30%)
- Final project and group presentation (30%)
- Peer evaluations (10%)

CLASS ATTENDANCE

Classes will be held synchronously on Zoom, with recordings posted after each class. The Zoom link will be posted on CCLE; please sign in through CCLE using your UCLA logon.

* If you cannot attend any live classes due to a time-zone difference, please let me know.

GROUPS

Since we are all remote this year, I will assign groups (4-5 students each) to try to accommodate time-zone differences. This will be your primary team throughout the quarter. After the last day of class, each student will submit a short peer evaluation of your fellow teammates.

LABS

We will have 5 classes with labs in R Studio. I will give a short demonstration on the methodology and then groups will have a break-out session to work together on the analysis in R. Groups should upload their code and solutions to a few questions to CCLE after class.

* If your group is in a different time zone, you may work on the lab at a different time and post your files within 1-2 days of class completing.

IN-CLASS DEBATES

During classes 3, 6, and 9, we will hold a live discussion/debate on Zoom about a current healthcare topic or organization working in the area of healthcare analytics. I will assign topics and readings in advance, and each group should be prepared to discuss the merits for or against the particular topic.

FINAL PROJECT TOPICS (SUGGESTED)

Groups should apply one or more analytics tools from class to a real-world dataset, to test a particular hypothesis. Groups will present their results during class 10. Below are publicly available datasets, but you are welcome to explore a different dataset of your choice. You may consider merging this data (eg, California vaccination rates) with another dataset (eg, household incomes) to gain deeper insights.

1. National Health and Nutrition Examination Survey (NHANES)
<https://www.cdc.gov/nchs/nhanes/Default.aspx>
2. World Happiness Report
<https://worldhappiness.report/ed/2019/>
3. Centers for Medicare and Medicaid Services (CMS) physician referrals
<https://www.cms.gov/Regulations-and-Guidance/Legislation/FOIA/Referral-Data-FAQs.html>
4. California childhood vaccination rates by school, or disease cases by county
<https://data.chhs.ca.gov/dataset/school-immunizations-in-kindergarten-by-academic-year>
5. Humanitarian Data Exchange (global health, disease outbreaks, refugees, etc.)
<https://data.humdata.org/>

FREE ONLINE TEXTBOOK

“Statistical Inference via Data Science” by Ismay and Kim
<https://moderndive.com/index.html>

SOFTWARE & DATASETS

We will conduct in-class exercises using Excel and R Studio. Datasets will be posted on CCLE before class.
No prior experience with R is required.

First, download and install R here: <https://cloud.r-project.org/>

Second, download and install R Studio here: <https://www.rstudio.com/products/rstudio/download/>

HONOR CODE

The UCLA Anderson Honor Code will apply at all times, and I expect you to strictly adhere to this policy. If you have any questions or concerns, please feel free to speak with me.

<http://www.anderson.ucla.edu/Documents/areas/adm/web/AndersonHonorCode.pdf>

ZOOM ETIQUETTE

- Have a working webcam and microphone. Earphones are encouraged to reduce feedback.
- Keep your video on, but mute your microphone (unless speaking) to avoid background noise.
- Please keep your attention on the live class discussion, and save the online shopping, Netflix, and laundry for after class.
- Use the Raise Hand and Chat features in Zoom, and I will do my best to monitor these.
- Please be professional and respectful of your fellow classmates, the TA, and instructor. We are all trying our best to make this new teaching experience as valuable—and enjoyable—as possible. Your patience is greatly appreciated!
- Some additional tips: <https://thewirecutter.com/blog/professional-video-call-from-home/>

FEEDBACK

Although this is my 12th year as a professor, this course is my first experience with remote instruction. I will try my absolute hardest to make the class a positive experience for everyone. In return, I ask for you to give me your undivided attention for 3 hours each week, and engage in class discussions, labs, and break-out sessions. If anything is unclear, or if you have constructive feedback, please let me know as soon as possible, so that I may update the course as we go along.

CLASS 1. THE CORONAVIRUS EPIDEMIC **APRIL 1**

Aim: To understand the trade-offs between data analytics and model building, and when each is appropriate.

Analytics tool: Mathematical modeling, optimization

Lab: Construct a Susceptible-Infected-Removed (SIR) epidemic model in Excel and estimate parameters from data.

Optional reading: Long, Nohdurft, Spinler (2018) "Spatial resource allocation for emerging epidemics: a comparison of greedy, myopic and dynamic policies", *MSOM*

CLASS 2. HEALTH INSURANCE **APRIL 8**

Aim: To examine how age, sex, BMI, smoking status relate to health costs. To discuss adverse selection in insurance markets with asymmetric information.

Analytics tool: Linear regression

Lab: Build a linear regression model in R using insurance claims data

CLASS 3. DISCUSSION: CURRENT ISSUES IN HEALTHCARE **APRIL 15**

Aim: To identify potential benefits and negative consequences of using data analytics in specific healthcare settings.

Topics: Groups will select 1 of the topics below and present one side (for or against) in a debate format

- Data privacy in genetic testing
- Artificial intelligence displacing healthcare jobs
- Racial biases in predictive algorithms

CLASS 4. SOCIOECONOMICS AND HEALTH BEHAVIOR (NHANES) **APRIL 22**

Aim: To explore a large dataset and test whether healthcare utilization differs by socioeconomic factors.

Analytics tool: Multiple linear regression

Lab: Merge multiple NHANES datasets and replicate a study on the relationship between income and exercise behavior.

Optional reading: Meltzer and Jena (2010) "The economics of intense exercise", *Journal of Health Economics*

CLASS 5. GLOBAL HEALTH (PEPFAR) **APRIL 29**

Aim: To illustrate how natural experiments can provide opportunities to evaluate policy interventions.

Analytics tool: Difference-in-differences

Lab: Develop a diff-in-diff analysis of country-level HIV outcomes before and after implementing PEPFAR.

Optional reading: Bendavid et al (2012) "HIV development assistance and adult mortality in Africa", *JAMA*

CLASS 6. DISCUSSION: HEALTHCARE ANALYTICS COMPANIES **MAY 6**

Aim: Groups select a current company working in the healthcare analytics space and prepare a 10-minute presentation to give on Zoom. Rather than just summarizing the firm's product/service, pick one aspect of the company and go deeper using tools from your other core courses. If you were hired by this company, what would you want to work on? Be creative!

Topics: Companies may be large or small, across a wide array of industries (e.g. electronic health records, precision medicine, diagnostics and therapeutics, telemedicine, wearable devices, etc).

Some ideas: <https://www.fastcompany.com/90457848/health-most-innovative-companies-2020>

CLASS 7. PREDICTING DISEASE (FRAMINGHAM STUDY) **MAY 13**

Aim: To introduce nonlinear regression models with binary outcome variables.

Analytics tool: Logistic regression

Lab: Build a logistic regression model in R to predict cardiovascular health outcomes.

CLASS 8. HOSPITAL PATIENT OUTCOMES **MAY 20**

Aim: To introduce causal estimation in healthcare.

Analytics tool: Regression discontinuity

Lab: Replicate a published study using regression discontinuity based on newborn birth weight and NICU assignment.

Optional reading: Doyle, Kowalski and Williams (2010) "Estimating marginal returns to medical care: Evidence from at-risk newborns", *Quarterly Journal of Economics*

CLASS 9. DISCUSSION: VALUING MEDICAL TECHNOLOGY **MAY 27**

Aim: To introduce cost-effectiveness analysis in healthcare and understand how U.S. payers decide what drugs and devices to cover.

Topics: TBA

CLASS 10. FINAL GROUP PRESENTATIONS **JUNE 3**

Groups should introduce the problem setting, describe the appropriate analytics tool, and present the analysis to the class via Zoom.