# MA 326

## Analytics for Strategy

### *Lies, Damned Lies, and Statistics*

Most managers lack real world experience working with data, above and beyond the basics of spreadsheet analysis. As a result, managers either perform inadequate evaluations of the data themselves, or turn the data over to statisticians who may lack the relevant institutional knowledge. Ideally, data-driven evaluation should combine both skills (data analysis and business expertise). This course is intended to help you obtain skills with data analysis. Bearing this in mind, this course has several goals:

- To familiarize you with the process of assembling and analyzing a data set
- To introduce the concept of an "experiment" and teach you how to distinguish correlation from causation
- To reinforce statistical techniques learned in your DS class and also introduce new statistical techniques
- To increase your comfort level working with data
- To enhance your ability to work with Stata
- To enable you to assess research and perform convincing research of your own.
- To help you distinguish convincing research from damnable lies.

### *Pedagogy*

This is a *clinical* course. You may be familiar with the educational process in other clinical fields, such as medicine. A famous dictum in medical education is "see one, do one, teach one." We will adhere to the same philosophy. I will give several lectures during the quarter, but you will spend most of your time performing and discussing empirical research. You will design and implement three research projects that I have chosen to highlight a range of key research issues. You will also be asked to read and discuss several empirical research papers.

The course is divided into four modules. The first module consists of lectures that build rapidly from basic to advanced research methods. During the first module you will perform a simple empirical project using real world data. The centerpiece of each remaining module is an extensive empirical analysis. Working in small groups, you will answer specific questions by developing and implementing an empirical model. We will discuss your modeling ideas during in-class "lab sessions." You should also feel free to visit me in my office (Jacobs 622); best time to come is before 3pm on M-Th. E-mail ahead of time to make sure I am available.

Each of the last three project modules will end with group presentations, during which we will discuss methods for presenting empirical results.

I expect you to perform all analyses in Stata, complemented by Excel or other graphics-friendly software when appropriate. You will learn many advanced Stata skills, taking you well beyond the prepackaged commands used in a basics statistics class.

***A Note about the Readings***

Most of the readings are from *Practical Regression*, which is a compilation of Technical Notes published by Kellogg. You can find the book posted to Canvas.

## Course Schedule

**Session 1 (September 20) Introduction and How to Perform Convincing Research**

Reading: *Practical Regression: Ten Steps towards Convincing Research* (pp.2-10).
Technical Note: *Looking at Data*
*A Quick One*: Begin work on For-profit Hospital project (Due Date: September 29)
Note: If you are unfamiliar with Stata, you should perform the capm tutorial. All necessary information and data files may be found at Canvas.

**Session 2-3 (September 22 and September 27) Regression Basics/Building Models**

Reading: *Practical Regression*: *Regression Basics* (pp. 11-24)
Reading: *Practical Regression: Building Your Model* (pp. 25-33)
Reading: *Practical Regression: Omitted Variable Bias* (pp. 34-39).
Technical Note: *Interpreting Slope Dummies*
Technical Note: *Stata Log File*
Technical Note: *Preparing Tables*
Technical Note: *The Partial-F test*
Technical Note: *Creating a do-file*
In-class exercise with dummies, interactions, and multiple slope dummies: *Autosales*

**Session 4 (September 29) Heteroskedasticity**

Nonprofit Hospital project due
*A Tasty One*: Begin work on Yogurt Project (due date: October 18)
Reading: *Practical Regression: Heteroskedasticity* (pp. 49-63)
Technical Note: *Clustering Standard Errors*

**Session 5 (October 4) Log Models**

Reading: *Practical Regression: Log versus Linear Specification* (pp. 64-76)
In-class Box/Cox exercise using yogurtsmall data

**Session 6 (October 6) Fixed Effect Models**

Reading: *Practical Regression: Fixed Effect Models* (pp. 40-48)
In-class Fixed Effects exercise using TCE data

**Session 7 (October 11) Yogurt Lab Session**

Technical Note: *Influential Observations*
Lab Session – Discuss ongoing work on Yogurt project

**Session 8 (October 13) Fixed Effect/Panel Data in Action**

Technical Note: *ANOVA*
Research Paper Discussion: Henderson, R. and I. Cockburn, 1994, "Measuring Competence? Exploring Firm Effects in Pharmaceutical Research" *Strategic Management Journal* 15: 63-84. (Feel free to skip the following parts of the paper: the section entitled "Specification of the Econometric Model"; the discussion of Table 4 on page 77; the Appendix.)

**Session 9 (October 18) Presentations of Yogurt Projects**

Presentation of Yogurt Projects (Projects due)
Technical Note: Best Practices in Quantitative Data Presentation

**Session 10 (October 20) Discrete Dependent Variables: the Logit Model**

Reading: *Practical Regression: Maximum Likelihood Estimation* (MLE) (pp. 77-80)
Reading: *Practical Regression: Discrete Dependent Variables* (pp. 81-100)
In-class Logit exercise with *ISP_Rural* data
A Healthy One: Begin work on Report Card project (due date: November 3)

**Session 11 (October 25) Other Discrete Dependent Variables Models**

Reading: *Practical Regression: Discrete Dependent Variables* (pp. 81-100)
*In-class Poisson and Ordered Probit exercises with *ISP* data

**Session 12 (October 27) The Identification Revolution I – RCTs**

Recommended Reading: Manzi, J., *Uncontrolled* New York: Basic Books, 2012 (This book provides an opinionated but highly valuable discussion of techniques used by business for data-driven analyses.)

**Session 13 (November 1) More on RCTs**

Read the following online threads:
http://blogs.worldbank.org/psd/what-will-be-the-next-victim-of-randomized-control-trials
https://nyudri.wordpress.com/initiatives/deaton-v-banerjee/

Research paper discussion: Ashraf N. et al. 2014, "No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery" *Journal of Public Economics*

**Session 14 (November 3) Presentations of Report Card Projects**

In addition to these presentations, we will discuss special topics in data analytics. The exact topics and reading assignments will depend on issues raised during the quarter.

**Session 15 (November 8) The Identification Revolution II – Instrumental Variables and DID**

Reading: *Practical Regression: Causality and Instrumental Variables* (pp: 101-104)

**Session 16 (November 10) Exam**

**Sessions 17 and 18 (November 15 and 17) Meetings with Professor Dranove**

There are no class sessions on these two days. Instead, all student groups will meet with Professor Dranove to give final status reports on their projects. Meetings are projected to last 30 minutes. A meeting schedule will be posted and groups will be randomly assigned to meeting slots.

**Sessions 19 and 20 (November 29 and December 1) Presentations of Final Projects**

***Final Project***

You should work in a group of your choosing, consisting of 4-6 students. Your group may design your own final project, subject to my approval. I also have a number of data sets that you can use for a final project, but I strongly prefer that you seek out your own data. All projects must include some regression analysis, though this need not be the main focus of the project.

All topics for final projects must be approved by October 20. Late approvals will receive an automatic point deduction.

Most groups will have an opportunity to present their final projects during the tenth week of the quarter.

***Grading***

Your course grade will be broken down as follows:

| | |
|---|---|
| Nonprofit Hospital | 3% |
| Yogurt | 15% |
| Report Cards | 15% |
| Final Project | 18% |
| Exam | 33% |
| Class Participation | 16% |

***The Class Participation Grade***

Class participation promotes the collaborative nature of empirical research. Be sure to listen to your classmates as they discuss their work, especially during the lab sessions and project presentations. If you can help others become stronger researchers through your questions and comments, then you will receive a strong participation grade. If you say nothing, then you are not helping others and you will receive a low participation score.

Be sure to come to the research paper discussions with more than a passing knowledge of the assigned papers. (I will provide some questions to help guide you through these papers.)