QTM 520: Applied I

Contact Hours: 150 minutes of class meeting time, one 50-minute lab, plus 3 hours of regular out-of-class work required as preparation for in-class work per week throughout a full semester

Credit Hours: 4

Prerequisites: None

Instructor:	Xxx
Semester	Fall 20xx
Meeting Time and Place:	xxx
Office:	xxx
Office Hours:	xxx
Email/Contact:	xxx
Course Website:	xxx
TA:	xxx

COURSE OBJECTIVES

Procedures for regression analysis for descriptive and causal inference; Applied, and formal foundations for regression and more advanced methods treated later in the masters. Foundations of statistical hypothesis testing via linear regression models. Omitted variable bias, multicollinearity, and heteroskedasticity. Generalized Linear Methods, Principal component analysis, Multidimensional scaling, Exploratory factor analysis, Confirmatory factor analysis, Cluster analysis, panel data methods, machine learning techniques

The course is designed to be taken simultaneously with QTM 510 and QTM 530 where students will also learn a logical and intuitive introduction to quantitative analysis from first principles and how to use a statistical software program to organize and analyze data.

CLASS REQUIREMENTS

Grades will be based on

- homework assignments (25%)
- midterm exam (25%)
- final exam (30%)
- data analysis project (20%)

HOMEWORK

The homework assignments will consist of analytical problems and data analysis with and without computer work. There will be approximately 6 homework assignments. The homework must be word processed with tables and figures together in the document. If you are using R markdown, make sure to knit it and submit as pdf or html (pdf is preferred) file. Each assignment will have its due date indicated and should be submitted through Canvas. Usually, assignments are due before the class meeting (e.g., if class is Tuesday 11:30 am, the assignment is due by 11:29am on Tuesday morning). Any assignment submitted after the due date/time will earn 0 points. To accommodate unexpected circumstances, your lowest homework grade will be automatically dropped at the end of the semester. Working together on the homework assignments is encouraged, but you must write your own solutions. This includes the computer assignments. It is highly recommended that you make your solo effort on all the problems before consulting others.

EXAMS

The midterm and the final will be in-class exams. **No collaboration is allowed on the exams**.

DATA ANALYSIS PROJECT

The final term project is a group-based work. You will make a team of 4-5 students. 2-3 papers will be provided. Your team must choose one paper and replicate it. The details of the project will be discussed after the Fall break.

LAB SESSIONS

Lab sessions will be run by TA. The purpose of the lab sessions is to learn the R programing and data analysis skills useful for regression and other empirical analysis. Students must bring a laptop to these sessions.

HONOR CODE

All students enrolled at Emory are expected to abide by the Emory College Honor Code. Any type of academic misconduct is not allowed which includes 1) receiving or giving information about the content or conduct of an ex- amination knowing that the release of such information is not allowed and 2) plagiarizing, whether intentionally or unintentionally, in any assignment. For the activities that are considered to be academically dishonest, refer to the Honor Code: http://catalog.college.emory.edu/academic/policies-regulations/honor-code.html.

DISABILITY ACCOMMODATIONS

If you are seeking classroom accommodations or academic adjustments under the Americans with Disabilities Act, you are required to register with Office of Accessibility Services (http://accessibility.emory.edu). To receive academic accommodations for this class, please obtain the relevant letter and meet with me at the beginning of the semester. Students are expected to give two weeks notice of the need for accommodations.

REQUIRED TEXTBOOK

- [W] Jeffrey M. Wooldridge, *Introductory Econometrics: A Modern Approach*, 7th Edition, Cengage Learning

SUPPLEMENTAL TEXTBOOKS

- [ISL] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. Introduction to Statistical Learning. http://faculty.marshall.usc.edu/gareth-james/ISL/
- [HR] Hernan, Miguel and Robis, James, *Causal Inference* : http://www.hsph.harvard.edu/faculty/miguel-hernan/causal- inference-book/
- [MD] Chester Ismay and Albert Y. Kim, *Modern Dive* http://moderndive.com/index.html
- [KI] Kosuke Imai, 2017, *Quantitative Social Science: An Introduction*, Princeton University Press
- [GH] Gelman, Andrew and Hill, Jennifer, 2007, *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge University Press.
- [MHE] Joshua D. Angrist and Jorn-Steffen Pischke, *Mostly Harmless Econometrics*, Princeton University Press
- ~~[GG] Alan Gerber and Donald Green. Field Experiments. W.W. Norton~~
- Kennedy, Peter, *A Guide to Econometrics*, 5*th* Edition, Blackwell.

Part I: The Basics of Simple and Multiple Regression for Cross-section Data

Week 1: Introduction

Topics:
- Syllabus
- Overview of empirical analysis
- Importance of controlling [Ethics component]
- Introductory/Review Sampling Activity
- Importance of uncertainty assessment [Ethics component]
- Introduction to data types

Lab Session: Introduction to GitHub Repository

Reading:
[W] Chap. 1

Week 2: The Simple Linear Regression Model

Topics:
- OLS estimator
- Interpretation
- Goodness of fit
- Units of measurement
- Understanding OLS estimator as a random variable
- Statistical properties of OLS estimator
- Estimator choice [Ethics component]

Reading:
[W] Chap. 2

Lab Session: Integrating R with Github

Weeks 3 and 4: Multiple Regression Analysis
- Importance of controlling and the consequence of omitted variables
- Interpretation of the estimated model
- Statistical Properties of OLS estimator
- Factors that affect the precision of an estimator & multicollinearity
- Qualitative Information (as a regressor) ([W] Chap. 7.1-7.4)
- FWL theorem
- Reporting standards [Ethics component]

Reading:
[W] Chap. 3 & Chap.7.1-7.4

Weeks 5 and 6: Inference on parameter estimates in a regression model

- Interval Estimate
- Null and Alternative hypothesis
- Size and Power
- P-value and Rejection region
- Confidence interval
- Linear and Nonlinear restrictions
- Reporting standards [Ethics component]

Reading:
[W] Appendix. C3 & C4
[W] Chapter 4

Week 7: Review and Midterm

Week 8: Review of Midterm & Diagnostics
- Non-normal distribution : Asymptotic properties of OLS ([W] Chap. 5)

Reading:
[W] Chap. 5

Week 8: Review of Midterm and Resampling for Descriptive Inference
- Example: Computing bootstrapped se of APE for the LDV Models

Reading:
[MD] Chap. 9

Week 9: Diagnostics (Cont'd)
- Non-normal distribution: Asymptotic properties of OLS ([W] Chap. 5)
- Nonlinearity: Functional form ([W] Chap. 6.2)
- Heteroskedasticity ([W] Chap 8.1-8.4) - Robust standard error, Test for heteroskedasticity, Weighted least squares estimators
- Outliers and Influential observations ([W] Chap. 9.5e)
- Reporting standards [Ethics component]

Reading:
[W] Chap. 5, Chap. 6.2, Chap. 8.1-8.4, Chap 9.5e

Week 10: Data Acquisition & Management

Part II: Limited Dependent Variables

Week 11: Limited Dependent Variables

- Linear probability model
- Logit and Probit regression model
- Multinomial regression
- Ordered logit
- Count Models
- Average Partial Effect (APE), Partial Effect at the Average (PEA)
- Reporting Standards [Ethics component]

Reading:
[W] Chap. 17.1

Part II: Causal Analysis

Week 9: Introduction of Causal analysis and Randomized controlled experiment

- Definition of Potential Outcomes Based Causal Inference
- Individual and Average Treatment Effects
- Estimation of Average Treatment Effects
- Non-monotonic Treatment Effects [Ethics component]

Optional Reading:
Angrist, Joshua D. and JornSteen Pischke (2010) *The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics*, Journal of Economic Perspectives

Marianne Bertrand and Sendhil Mullainathan (2004) *Are Emily and Greg More Employable than Lakisha and Jamal?*
A Field Experiment on Labor Market Discrimination, American Economic Review

Week 10: Randomized Control Experiments

- Unbiased estimators with completely randomized design
- Consistent estimators with completely randomized design
- Unbiased estimators with blocked randomized design
- Consistent estimators with blocked randomized design
- Standard errors and confidence intervals
- Natural experiments
- Human subjects research [Ethics component]

Reading:
[GG] Chs. 1- 4

[MHE] Chs. 1-2

Week 11: Randomized Control Experiments: Non-Compliance and Externality Validity

Topics:
- Noncompliance in experiments
- "Noncompliance" in natural experiments
- Constant effects models
- Hetergeneous effects models
- External Validity, Mechanisms and Theory [Ethics Component]
- Reporting standards for "broken" experiments [Ethics Component]

Reading:
[GG] Chs. 5-6
[MHE] Ch. 4

Week 12: Introduction to Causal Inference using Observational data
- The Assumption on No Unmeasured Confounding
- Choosing Conditioning Variables
- Matching

Reading:
[W] Chap. 15
[W] Chap. 13

 Optional Reading:
Daniel Treisman, *The Causes of Corruption: A Cross-National Study*, Journal of Public Economics, 2000

Weeks 13: Causal Inference Using Observational Data: Panel Data

Topics:
- Panel Data
- Diff-in-Diff
- Fixed and Random Effects
- Regression Discontinuity
- Reporting standards for observational studies [Ethics Component]
- Reporting standards for subpopulation effects [Ethics Component]

Reading:
[MHE] Chs. 3, 5

Weeks 14: Causal Inference Using Observational Data: Regression Discontinuity

Topics:
- Regression Discontinuity
- Local Average Treatment Effects versus Average Treatment Effects [Ethics Component]

Reading:
[MHE] Ch 6