Economic Data Analytics
Introduction to Data Management, Statistics, and Regression Methods for Decision Making
01:220:212

Instructor: Nora Paxton
Email: npaxton@economics.rutgers.edu
Class Hours: W 6:10-9:00
Classroom: Rutgers Academic Building 4450
Office hours: Wed. after class & by appointment

Introduction:

Organizations of all sorts—business, government, scientific, educational, non-profit and cultural—have 'customers' and need to make allocation decisions to operate effectively and efficiently. While certain specialized skills in big data analytics are strongly in demand in the current job market, many employers also look for individuals with general skills who are "trainable" in the specifics of a job. This lower-level course provides the tools and knowledge to use data to "size up" a situation or problem and to suggest alternative approaches or solutions based on available data. This course will demystify the process of data collection, visualization, analysis, and presentation. It will also show you how to work in the ubiquitous Microsoft Excel environment and how to do basic statistical analysis. In addition, you will be able to enhance your Excel skills by learning basic regression methods, which are econometric tools for estimating and explaining relationships among variables. These tools will allow you to provide important data-based decision support that organizations require.

Upon conclusion of the course, students will have an introduction to methods of economic data analysis and will be able to:

- Use spreadsheet software to collect, clean, transform, visualize, interpret and present data.
- Understand, conduct and interpret basic descriptive statistics including means, variability, and correlations and basic inferential statistics.
- Present results of data analysis for a non-specialist audience.

Prerequisites: Advanced algebra, placement into precalculus. Those intending to complete the Economic Data Analytics Minor should take precalculus (640:111, 640:115, or equivalent), as this is required for the introductory economics sequence, Introduction to Microeconomics (220:102) and Introduction to Macroeconomics (220:103).

Text: *Statistics for People Who (Think They) Hate Statistics*, Excel 2016 version, by Neil Salkind

Data analysis tools: Microsoft Excel. Access to personal or laptop computer with Excel installed is essential.

Evaluation: How is the grade determined?

- Class participation (10%)
- Homework and group project – (20%)
- Mid-term Exam (30%) - Data types and structures, Descriptive statistics
- Final Exam (40%) – CUMULATIVE **May 9th**
- *There is no extra credit given in this course*

Logistics and grade scale - Presentations and other class materials will be made available through SAKAI or email as needed. The syllabus on SAKAI will be the most current one and will supersede any previous versions. The grade scale used for the class is:

0-59.99 = F/ 60-69.99 = D/ 70-76.99 = C / 77-79.99 = C+/ 80-86.99 = B/ 87-89.99 = B+ / 90-100 = A

Late policy – All work must be submitted by the due date and time. Late submissions will be penalized with five percentage points per day after the due date and time. For example, if you get 90% on an assignment but it was handed in one day late, the grade will go down to 85%. Late submissions without penalty allowed only in cases of documented health or medical emergency. Missing assignments will get a zero grade.

Make-up exam policy – In cases of documented illness or personal emergency, email me prior to the exam. If a delay is granted, make up exams will be held by the economics department in **New Jersey Hall on Fridays from 12pm to 3pm in room SC 216**. Make up exams must be completed before exams are returned to students. After exams are returned, no more make ups will be allowed.

Academic integrity - Get familiar with the university's policy on academic integrity, it will be enforced in this class: http://academicintegrity.rutgers.edu/academic-integrity-policy/

Absence reporting - It is expected that students attend all sessions. Please note the dates of the Mid-Term, group projects, and the Final Exam (**May 9th**) to ensure you will be present on those key dates. However, if are going to miss more than one class due to illness or a family emergency, please use the University absence reporting website https://sims.rutgers.edu/ssra/ to indicate the date and reason for your absence. An email is automatically sent to me. Note, reporting your absence does not excuse from your course responsibilities.

Accommodations - If you need special accommodation due to disability, check the procedures and guidelines set by the Office of Disability Services: https://ods.rutgers.edu.

Class Topics

Class	Date	Topic	Readings
1	1/17	Introduction • What is data analysis and why do we care? • Excel is a powerful and ubiquitous tool • Review syllabus • Review class logistics Types of data • Time series • Cross sectional • Pooled data (panel/longitudinal data) Sources of data • Survey data • Administrative data	Chapter 1 Chapter 22 http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?_r=0 https://www.brookings.edu/wp-content/uploads/2016/06/04_obama_social_policy_haskins.pdf

		• Extant data from domestic and international agencies • Client data (sales, revenue, outputs, etc.)	
2	1/24	Methods of evaluation • Exploratory • Descriptive • Inferential Qualitative vs. quantitative data and analysis Exploring Excel - The Basics and Beyond *--In-class Excel exercises*	Appendix A (start here) Chapter 1.A Chapter 1.B https://ies.ed.gov/ncee/pubs/20174023/pdf/20174023.pdf https://ies.ed.gov/ncee/pubs/20104029/pdf/20104029.pdf (pages 1-22 and pages 41-51) https://www.acf.hhs.gov/opre/resource/the-role-of-social-networks-among-low-income-fathers-findings-from-the-pact-evaluation (pact RF Social Networks.pdf on Sakai)
3	1/31	How to prepare data for analysis *Group Challenge – create an analysis file*	Higher Education http://www.npr.org/sections/ed/2016/10/30/499200614/how-one-university-used-big-data-to-boost-graduation-rates https://www.nytimes.com/2017/02/02/education/edlife/will-you-graduate-ask-big-data.html?_r=0 Business: https://www.theatlantic.com/business/archive/2014/11/when-to-book-your-plane-ticket-a-guide/383146/
4	2/7	Descriptive Statistics • Averages • Median • Mode • Variability *Group exercise to illustrate descriptive statistics*	Chapter 2 Chapter 3 Public Health: https://www.cdc.gov/obesity/data/prevalence-maps.html Use statistics carefully:

		In-class Excel exercises	http://huff.to/2jxDzE7
5	2/14	**HOMEWORK #1 DUE BEFORE 6PM** Class discussion of Group Presentations • Overview of requirements • Brainstorm research questions • Brainstorm data sources Exploratory analysis • Using descriptive statistics • Plotting data • Scatter plots, histograms • Outliers Correlations • What is correlation? • How do we measure it? • Correlation does not equal causation Reliability and Validity *In-class Excel work*	Chapter 4 Chapter 5 Chapter 6 Come to class with ideas for presentations! Housing: https://www.washingtonpost.com/news/where-we-live/wp/2017/10/05/mortgage-rates-hold-steady-but-are-trending-higher/?utm_term=.ace75edf8314 Environment: http://www.motherjones.com/environment/2015/06/california-sinking-drought-ground-water
6	2/21	**Mid-Term – bring a calculator (separate from phone)** *Group time to select a topic.*	
7	2/28	Hypothesis testing Distributions • General definition • Normal distribution • Z score	Chapter 7 Chapter 8 Public Policy https://www.cbo.gov/system/files/115th-congress-2017-2018/costestimate/americanhealthcareact.pdf https://www.cbo.gov/sites/default/files/110th-congress-2007-2008/reports/10-31-healthinsurmodel.pdf Fashion: https://www.theguardian.com/technology/2014/jan/30/fashion-data-tool-editd-helps-asos-push-revenues-up-37

			Advocacy: https://www.oxfam.org/en/research/economy-994
8	3/7	Inferential statistics • Statistical significance • Significant vs meaningful • Determine what test to use Testing one sample • Z test • How to interpret	Chapter 9 Chapter 10 Defining and measuring poverty: http://www.irp.wisc.edu/faqs/faq2.htm https://www.cbpp.org/research/social-security/social-security-keeps-22-million-americans-out-of-poverty-a-state-by-state Finance: https://www.nytimes.com/2017/03/28/business/dealbook/blackrock-actively-managed-funds-computer-models.html
9	3/21	Q&A session on Group Presentations Testing two independent samples • T test • How to interpret Testing two dependent (related) samples • T tests again • How to interpret *In-class Excel work*	Chapter 11 Chapter 12 Significant vs. Meaningful http://www.healthnewsreview.org/2017/03/everolimus-cancer-drug-spin/ http://www.nydailynews.com/life-style/one-third-u-s-marriages-start-online-dating-study-article-1.1362743 Auto Industry https://datafloq.com/read/self-driving-cars-create-2-petabytes-data-annually/172 Sharing data https://news.stanford.edu/2017/10/02/psychologists-simplifying-brain-imaging-data-analysis/
10	3/28	Significance of correlation coefficient Linear Regression • What is a simple linear regression? • Dependent and independent variables • Using Excel data analysis	Chapter 15 Chapter 16 Predictive analytics Can an Algorithm Tell When Kids Are in Danger?

		• Assess quality of regression • Multiple regressions • Logistic regressions	Health care spending https://www.theatlantic.com/health/archive/2017/06/fixing-the-5-percent/532077/
11	4/4	**HOMEWORK #2 DUE BEFORE 6PM** Review Linear Regression • *In-class Excel work* Testing more than two samples • Analysis of variance • ANOVA Intro to factorial analysis	Chapter 13 Chapter 14 Zika (mathematical models, spatial analysis) https://www.cdc.gov/mmwr/volumes/66/wr/mm6613e1.htm?s_cid=mm6613e1_w http://www.pnas.org/content/114/22/E4334.full Labor studies https://www.bls.gov/news.release/empsit.nr0.htm
12	4/11	Q&A session on Final In-Class Presentations Other important statistical tests • chisq • F Test ANOVA (again)	Chapter 17 Chapter 18 Bring questions about Group Presentations LONG study! https://www.washingtonpost.com/news/inspired-life/wp/2016/03/02/harvard-researchers-discovered-the-one-thing-everyone-needs-for-happier-healthier-lives/?utm_term=.5a02975f2dde Kindergarten redshirting https://www.brookings.edu/blog/brown-center-chalkboard/2016/06/22/how-much-does-it-benefit-a-child-to-delay-kindergarten-entry-for-a-year/ New data! http://www.npr.org/sections/thetwo-way/2017/04/18/524553683/former-microsoft-ceo-launches-new-tool-for-finding-government-data
13	4/25	**Group Presentations**	
14	5/2	Introduction to Data Mining	Chapter 20

		Introduction to R Q&A session on student projects --work on projects in class	Intro to R http://r4ds.had.co.nz/index.html Conflict of interest: http://www.vox.com/2017/4/19/15350048/pubmed-publishing-conflicts-of-interest-funding-information Health care and big data: https://www.mapr.com/blog/reduce-costs-and-improve-health-care-with-big-data