Advanced Quantitative Methods
Spring, 2011
R 6:00-7:50 PM

Professor: Richard Hendra, Ph.D.
rick.hendra@gmail.com

Important note: If you email me at hendr239@newschool.edu, please include your email address in the body of the email or it will typically take me a few days to get back to you. Emails to this account are forwarded to my gmail account without sender information. Sorry for this inconvenience.Gmail response time is typically w/in 1 day.

CONTACT INFORMATION

Email and the class "Blackboard" intranet site are heavily used in this class as a means of communicating, making announcements, and sharing documents. Students should check the class website frequently and make sure that their New School email address is up to date. Typically class materials will be posted by 9AM, the day of class (or earlier). I send an email out to let you know when materials are posted.

COURSE DESCRIPTION

In an information driven economy and an era of evidenced based decision making, quantitative skills are increasingly in demand. A firm understanding of data, quantitative methods, and computer driven analysis is an excellent means of enhancing your success in this labor market and improves the prospect that you can "make a difference" through research. This is a course about how to "do" statistics. This is not a math class and this class will only briefly touch on statistical theory. There are no hand calculations. We are here to learn about how to *use* statistics in SPSS to solve research and business problems and to improve decision making. This class picks up where quantitative methods left off. Students who work hard will leave the class with the ability to conduct their own serious quantitative research.

The main goal of the class is to broaden and deepen student's knowledge and ability to apply multivariate statistical techniques. Thus, after a quick statistics primer, we will start where you probably left off with quantitative methods: the multivariate regression model. We will elaborate the model and focus on how to detect and remedy violations to the many assumptions made by the model. In doing so, we will touch on many themes under the purview of econometrics. However, this is not an econometrics class. We will then discuss the special but common case of a dichotomous dependent variable. We will learn how to build clean and useful models for forecasting the probability of events (logistic regression). We will then spend a couple of weeks learning how to create statistical constructions called "factors" which concisely capture the key dimensions of multidimensional behaviors and outcomes such as "Happiness", "Depression", "Poverty." We will use this information to discuss how to create scales, and then learn how to use reliability analysis to assess the quality of these scales. In order to expose students to rapidly developing of data mining, we will briefly cover the key themes in cluster analysis and related techniques. Finally, the class will conclude with a deeper look

treatment of ANOVA analysis a powerful technique known as multilevel modeling. Students will learn that at many levels, ANOVA analysis is the same as regression analysis, although it is used for different purposes.

TEXTS/MATERIALS [TRY TO FIND DISCOUNTS!]

* Discovering Statistics Using SPSS. Andy Field **:** SAGE Publications. 3rd Edition only!! (January 23, 2009)

*Optional: students who aren't in the NSU computer labs often may opt to rent or rent SPSS. Available at the "**onthehub.com**" website. Search for SPSS for 6 month rental (49.00). Must get "standard" version (or higher). *The Base version won't cover everything we'll do in class!*

* Data sets will be posted on the class website. The textbook also has a very good companion website.

REQUIREMENTS

Deliverables
All work will be in SPSS. No hand calculators needed. Throughout the semester, students will be expected to execute and interpret statistical analyses in a series of problem sets. These will include short write-ups (2-3 pages) and output. Students should keep up with material throughout the semester. I don't want to clog up the schedule with too many quizzes, but there will be a cumulative final so it is critical to keep up. It's too much material to cram at the end. All assignments must be on time (in hardcopy) to get a grade.

Class schedule (will be adjusted as needed)

Week of	**Topic**	**Readings (all Chapters refer to Andy Field book)**	**Deliverable**
1/27	Review of basic statistics, research design, and data analysis concepts.	Ch 1,2, 5 http://www.nytimes.com/2011/01/11/science/11esp.html http://www.newyorker.com/reporting/2010/12/13/101213fa_fact_lehrer	
2/3	Multiple Regression Designs: the standard model	Ch 6,7 (Pages 197-225)	
2/10	Elaborating the Regression Model	Rest of Ch 7	Take home quiz. Due next week
2/17	Violations of Regression assumptions and how to redress them.	Online Regression Technical Notes: http://faculty.chass.ncsu.edu/garson/PA765/regress.htm Ch 3,4	Regression assignment distributed

		Regression critique: http://crab.rutgers.edu/~goertzel/mythsofmurder.htm	
2/24	Case study: regression in program evaluation.	Findings from New York Cities Conditional Cash Transfer Program http://www.mdrc.org/publications/549/execsum.pdf The core analytics of random assignment: http://www.mdrc.org/publications/437/full.pdf	
3/3	Logistic Regression	Ch 8 (through 294)	Final Regression model and write-up is due. Logistic Reg. assignment distributed.
3/10	Interpreting Logistic Regression and What can go wrong.	Rest of Ch 8	
3/17	Spring Break		
3/24	Logistic regression: advanced topics Possibly: multinomial logits	Online Logistic Regression Technical Notes: http://faculty.chass.ncsu.edu/garson/PA765/logistic.htm	Logistic regression model and write-up due.
3/31	Factor Analysis I: Principal Components	Ch 15	
4/7	Factor Analysis II: Building indexes and testing for reliability	Online Factor Analysis Technical Notes: http://faculty.chass.ncsu.edu/garson/PA765/factor.htm	Factor Analysis Assignment Distributed
4/14	Case study: Data mining to profile and target customers. -Clustering procedures	Online Data Mining Overview: http://www.thearling.com/text/dmtechniques/dmtechniques.htm	
4/21	*General Linear Model 1: One Way ANOVA* -Post hoc tests and multiple comparisons -Designed contrasts -Subgroup analysis	Ch 10 -Manzi on replication and multiple testing bias: http://www.city-journal.org/2010/20_3_social-science.html -When Is the Story in the Subgroups? Strategies for Interpreting and Reporting Intervention Effects on Subgroups http://www.mdrc.org/publications/551/full.pdf	Factor analysis and writeup due
4/28	General Linear Model 2: -Using covariates to improve precision -Using factorial ANOVA to measure interactions	"Contradicted and Initially Stronger Effects in Highly Cited Clinical Research" John P. A. Ioannidis, MD (I'll post this)	Quiz on Logistic, ANOVA, Factor

5/5	Data analysis topics: -Outlier detection and handling -Missing value detection and handling. -Subgroup analysis Survey topics: -Survey response analysis -Deriving weights -Fraud detection	Instructor Handouts	
5/12	Wrap-up/Final		Final exam

Grading Distribution:
Quizzes: 15%
Statistical Analyses: 50%
Final Exam: 35%