Advanced Quantitative Techniques
UAP 5494, Fall 2021
Tuesday 7:00 to 9:30 PM, Zoom

Contact Information

Instructor: Tom Sanchez, sanchezt@vt.edu
Offices: Virginia Tech Research Center, 900 N. Glebe Rd., Arlington, VA
TA: Trey Gordner, gg3@vt.edu
Office Hours: By appointment

Course Objectives

This course introduces a variety of intermediate quantitative research approaches in statistics, emphasizing discrete methods and multivariate analysis more generally. The objective is to expose students to these methods; including assumptions, techniques of application, and interpretation. More specifically, the course aims to help students through lectures, readings, student presentations, and exercises to develop a better understanding of:

- fundamentals of the R programming language for statistical analysis
- basic principles of sample design and inference
- the assumptions and limitations of ordinary least squares (OLS) regression models in planning, policy, and engineering analyses
- alternative approaches to analyzing categorical and other discretely-measured data that violate OLS assumptions
- techniques for incorporating fuller information into analysis through the application of panel, self-selection, and survival regression models
- data reduction and pattern recognition techniques to reduce the dimensionality of data

Learning Objectives

By the end of the course, each student will demonstrate the ability to:

- work with the R programming language
- draw inferences informed by statistical theory from samples
- recognize situations in statistical analysis where one or more of its assumptions are violated
- apply non-linear and other intermediate statistical methods to address planning, policy, or engineering questions ill-suited for OLS regression techniques
- critically assess the concepts, methods, and conclusions of an academic paper using intermediate statistical methods in a planning, policy, or engineering study
- frame a research question, collect data, conduct analysis using intermediate statistical methods, draw a conclusion on a planning, policy, engineering, or other research question, and prepare an article suitable for publication in an academic journal describing this effort

Words of Caution and (mostly) of Comfort

You should not take this course if you have not had a statistics class and you do not feel comfortable with basic statistical concepts of probability, hypothesis testing, and regression. It will be quite frustrating. However, we do not assume that anyone is a statistics expert. While labeled "advanced quantitative techniques," the course really represents an intermediate level of sophistication. We will see lots of numbers but little abstract math in the course, other than discussion of OLS basics. You do not need calculus or matrix algebra to understand the material. We will discuss the intuition behind techniques rather than the mathematical reasoning—using simple examples to talk about them—and rely mostly on applied, hands-on learning with the R programming language.

We will enforce the tenets of Virginia Tech's Graduate Honor Code, with all work subject to the code. For more information on the Graduate Honor Code, please refer to the GHS Constitution. We encourage students with special needs or circumstances to meet with me after the first class or as soon as possible. In all cases, please feel free to contact me should you have any questions or concerns about any course requirements.

Other Potentially Useful Course Materials

[insert links to online resources for both R and stats]

Assignments, Due Dates, and Grading

- Individual Problem Sets, **45%**: 3 sets of problems at 10% for first 3 sets, 15% for 4th set
- Readings and Class Participation, **10%**: Active participation in class activities, including 2 written questions prepared for **each** reading each week. I will call on students randomly during the discussion so be prepared!
- Reading Presentation & Write-Up, **15%**: One 2-3 page, single space, written synopsis and 15-minute oral presentation of a reading (sign up via Google docs at https://tinyurl.com/VTQuantReadings, first come-first serve). See below for format.
- Individual Research Paper, **30%**: Research Paper, with an oral presentation on December 7th and written paper due on December 8th. Note that the interim due date for the research question and data set. The paper should be abbreviated journal length (10-16 pages double-spaced) and likely will entail finding and using an existing and cleaned public use dataset. More details to follow.

All assigned work is due by the end of the day on the due date unless otherwise noted in this schedule. Written assignments will be _submitted online_*. If you turn in your work late, we will take off 10% of the grade for each 3 days that you submit it late (e.g.,* up to 3 days late it loses 10%, from 3 days to 6 days late it loses 20%, and so forth)*. If you take an incomplete for the class, you will lose 1 full course grade/month for each month (or portion of the month) after 11 December that you complete and submit* all *assigned work (e.g.,* A- would be changed to B- up until 11 January, a C- up until 11 February, failure after that).*

Format of Assignments

For written assignments—reading synopsis and individual paper—you must use the Chicago Manual of Style format. (See www.chicagomanualofstyle.org/tools_citationguide.html for simple guidance). Use the author-date/references style, NOT the notes-bibliography style. Be sure to properly reference all material obtained from the internet, including the access date. We recommend using EndNote or Zotero, which are available to all Virginia Tech students and can help you organize your references and notes. It also provides templates for doing reference formatting automatically.

Reading Presentation and Write-up Guidelines (15%)

Each person must select one of the readings in the schedule and prepare a written synopsis and an oral presentation (oral presentation date indicated in the Schedule of Topics). Sign up via Google docs at https://tinyurl.com/VTQuantReadings for a paper and date (type in your name for a paper and date, only one person per reading). The synopsis should take 2-3 pages, single-spaced (Times New Roman 12 pt. or equivalent, 1 inch margins). It should include a section for each of the following:

- Problem statement
- Literature review
- Research design and data analysis methodology
- Results and findings
- Interpretation and conclusions

You must post the synopsis on Canvas at least five days before you present it in class. This gives all other students (and me) a chance to read it before you present it. We expect that everyone will come to class having read and prepared to discuss the paper and synopsis.

Course Schedule (*We likely will revise this*)

Dates	Lecture Topic(s)	Reading	Assignment
Week 1 (8/24)	Intro to R and RStudio (Data Carpentry 1&2)	Watch Installing RStudio and follow along with the instructions.	
Week 2 (8/31)	Data Structures and Dataframes (DC 3)	Watch R Basics Read Ch 5: descriptive statistics	Import and explore Olympics data on your own: https://github.com/noellepablo/olympics-app https://github.com/JaikrishnaVS/Shalady_Jviswaka_FinalProject https://github.com/josephmbarnby/Olympics2021
Week 3 (9/7)	Data Wrangling (DC 4)	Watch Data Wrangling Read Ch 12: tidy data Read Ch 7.1-7.7: Data handling	
Week 4 (9/14)	Data Visualization (DC 5)	Watch Visualization Read Ch 6: Drawing graphics	
Week 5 (9/21)	Programming Basics (SWC 7, 10)	Watch Programming Basics in R Read Ch 15: Linear regression Watch Linear Models Pt 1 Watch Linear Regression in R	Problem Set 1 - R programming assignment
Week 6 (9/28)	Binomial logit	Read Lewis and Pitts, 2017 Read Logistic Regression Watch Logistic Regression playlist	
Week 7 (10/5)	Multinomial logit	Read Grigolon, et al 2017 Read Multinomial Regression Watch Multinomial Regression with R 1-4	Problem Set 2 - logit regression
Week 8 (10/12)	Instrumental variables	Read Sanchez, 2002 Read Instrumental Variables Watch An Intuitive Introduction to IV Watch Econometrics - IV Watch IV with R playlist	
Week 9 (10/19)	Factor/principal components analysis	Read Walter, Evans, and Atherwood 2016 Read Factor/principal component analysis Watch PCA, Step-by-Step Watch PCA in R Watch Factor Analysis in R	
Week 10 (10/26)	Time series	Read Phillips and Land, 2012 Read Times series analysis Watch Time Series Analysis with R 1-4 Watch ARIMA Modeling Using R	Problem Set 3 - IV, factor analysis, time series
Week 11 (11/2)	Cluster analysis	Read Stylidis 2018 Read Cluster Analysis in R	

		Watch Intro to Cluster Analysis with R	
Week 12 (11/9)	Network analysis	Read Sanchez and Hall, 2014 Watch Network Analysis Theory & Concept Watch Network Analysis Practice (igraph only)	Let's use Gephi for this
Week 13 (11/16)	Spatial analysis	TBA Read Spatial data analysis and mapping Watch Spatial data analysis in R	Problem Set 4 - cluster, network, spatial
Week 14 (11/23)	Thanksgiving		
Week 15 (11/30)	Presentation methods	TBA	
Week 16 (12/7)	Presentations	TBA	

Readings

Grigolon, Anna B., Aloys W. J. Borgers, Astrid D. A. M. Kemperman*, et al.* 2014. "Vacation Length Choice: A Dynamic Mixed Multinomial Logit Model." *Tourism Management* 41:158-167. doi: https://doi.org/10.1016/j.tourman.2013.09.002.

Lewis, Gregory B., and David W. Pitts. 2017. "Lgbt–Heterosexual Differences in Perceptions of Fair Treatment in the Federal Service." *American Review of Public Administration* 47 (5):574.

Phillips, Julie, and Kenneth C. Land. 2012. "The Link between Unemployment and Crime Rate Fluctuations: An Analysis at the County, State, and National Levels." *Social Science Research* 41 (3):681-94. doi: 10.1016/j.ssresearch.2012.01.001.

Sanchez, Thomas W. 2002. "The Impact of Public Transport on Us Metropolitan Wage Inequality." *Urban Studies* 39 (3):423-436. doi: 10.1080/00420980220112766.

Sanchez, Thomas W., and Ralph P. Hall. 2014. "Social Network Analysis of Trb Standing Committees." Transportation Research Board 93rd Annual Meeting, Washington, DC.

Stylidis, Dimitrios. 2018. "Residents' place image: a cluster analysis and its links to place attachment and support for tourism". *Journal of Sustainable Tourism*, 26(6), pp.1007-1026.

Walter, Rebecca, Aaron Evans, and Serge Atherwood. 2016. "Addressing the Affordable Housing Crisis for Vulnerable Renters: Insights from Broward County on an Affordable Housing Acquisition Tool." *Housing Policy Debate* 26 (1):123-149. doi: 10.1080/10511482.2014.1003190.