



GEOG 5190

**Advanced Quantitative Techniques:
An Introduction to the Course**

The 4 Ground Rules for this Class

- 1. Do your work
- 2. Come to learn
- 3. Participate
- 4. Consider others

My goal here is to serve your needs well.
"I did not come to be served, but to serve"

Advanced Quantitative Techniques

- Hands-on introduction to powerful multivariate and spatial statistics

Advanced Quantitative Techniques

- Hands-on introduction to powerful multivariate and spatial statistics

Advanced Quantitative Techniques

- Hands-on introduction to powerful multivariate and spatial statistics

There is a basic concern reflected here that we'll come back to later in our class today

Advanced Quantitative Techniques

- **Q:** Why do we use statistics?
 - **"Statistics"** is the study of the collection, analysis, interpretation, presentation, and organization of data (Wikipedia)
 - As geographers and environmental scientists we have a lot of data, and our data are more complex than many people need to deal with
 - Always a spatial dimension, usually multiple variables, and often a time dimension as well

Advanced Quantitative Techniques

- **Q:** Why do we use statistics?
 - The methods we will study in this course will help us make sense of large and complex datasets
 - This course will emphasize the practical considerations associated with proper statistical practice

Your Instructor: Murray Rice

- **Education:** BS in Engineering Physics, MA & PhD in Geography
- **Background:** Proud Canadian 🇨🇦, Business Founder/Owner AltaVista Geographics
- **UNT:** Glad to be in my 13th Year at UNT 🌳
- **Family:** Wife (Sherri), Sons (Daniel, Timothy), Dogs (Maggie, Max) 🐕
- **Living life:** Committed Christian ✝️, traveler/explorer of new places (3 continents/45 US states/9 Canadian provinces so far) 🌍
- **Sports:** CFL Roughriders 🏈 • NHL Senators 🏒 • MLB Red Sox ⚾
- **Odd fact:** sons born in Red Sox championship years (2004, 2007) 🏆

Advanced Quantitative Techniques

- This is my dog, Maggie the vicious beagle



Advanced Quantitative Techniques

- This is the newest addition to my family, Max the cairn terrier/poodle



Advanced Quantitative Techniques

- **So tell me a little about you ...**
 - **Q:** what's your statistical background?
 - Courses taken? Where? How long ago?
 - Interests in statistics? Applications for statistical techniques?



Advanced Quantitative Techniques

- **Our class:** meeting Thursdays from Jan 18 to May 3, 6:00 to 8:50 pm
 - **Office hours:** Tuesdays, 10am to noon



Advanced Quantitative Techniques

Some people think that statistics professors think about statistics all the time.



Advanced Quantitative Techniques

- **Course goals**
 - **NOT** to get you to love statistics
 - **NOT** to get you fantastically excited about statistics ("let's do that again!")
 - **NOT** to make you a statistics expert
 - **NOT** to get you ready to "do statistics" perfectly
 - So what are we aiming at?

Advanced Quantitative Techniques

- **Course goals**
 - This is a practical "what are these methods, why would we use them, and how do I implement them" course
 - This is not an engineering or math course where we delve deeply into math equations and proofs

Advanced Quantitative Techniques

- **Course goals**

Aim: comfort with stats based on a foundational understanding of what each method does, and why you would use each one

Understanding & appreciation, not glee!

Please keep this in mind as we go through the semester

Advanced Quantitative Techniques

- **Course website**

- www.murrayrice.com (click on the teaching link, then click on the GEOG 5190 course link)
- Visit this site every week during the semester
- I post lecture and lab-related material here throughout the semester (slides from all lectures from last year are now up for your reference)

If we have winter weather, please visit the course website for updates and adjustments to our class schedule

Advanced Quantitative Techniques

Readings

- Your syllabus outlines readings for every non-lab week
- There is no formal textbook for the class, but I am providing two resources that you should read
 - 1. Readings packet:** available on CSAM's "class" network drive (read each week's readings before each class)
 - 2. Class slides:** available every non-lab week (posted on website; review after each class)

Advanced Quantitative Techniques

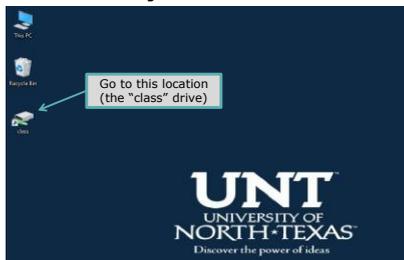
Readings

- One useful resource: a set of statistical readings we will refer back to throughout the semester

Mike Wulder's "A Practical Guide to the Use of Selected Multivariate Statistics"

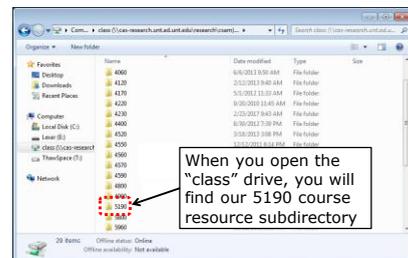
Accessible with the other readings on the CSAM "class" network drive

To Access Reading Files on our CSAM network drive:



Spring 2018 Key Code to Enter ENV 340: 5279

To Access Reading Files on our CSAM network drive:



Advanced Quantitative Techniques

Grading

- Labs:** 5 sessions in total, worth 10% of your course grade each (50% in total)
- Participation:** class attendance & joining in on lecture discussions (10% in total)
- Term Project:** worth 40% in total, including
 - Proposal, due Mar. 8 (5%)
 - Report, due Apr. 26 (25%)
 - Presentation, times currently scheduled for Apr. 26 and May 3 (10%)

Advanced Quantitative Techniques

Labs

- We will devote five entire evenings to lab work (five separate labs)
- I will generally make instructions for each lab available on the course website well before lab night (minimum 3 days)
 - Please read the lab package online before you come to the lab class
- Labs are due at the beginning of class the Thursday following the lab session



Advanced Quantitative Techniques

- o **Labs**
 - You need to understand that I have high expectations for each of the lab reports you submit (and for the term project we will discuss in a minute)
 - Please see my "[writing expectations](#)" handout

Advanced Quantitative: Syllabus and Handouts

Please download and read the handout files below, listed by the week each will enter into our classroom conversation. Additional files may be added during the semester, so please periodically check this page. Each of these files contain important background information related to the course and one or more of the methods we will discuss in class.

Note, some links below direct you to articles on journal websites. If you are accessing this page from a UNT computer, you can download a full-text version of each of the articles.

Please feel free to contact me if you have questions about any of these documents.

Week 1: Introduction to the Course

- GEOG 5190 Syllabus
- GEOG 5190 Writing Expectations
- GEOG 5190 Performance Guidelines
- Online Conceptual Reading False Positives (introductory reading)
- Online Conceptual Reading Understanding Data
- Non-Taking Handout (from Stanford University's Center for Teaching and Learning)

Please review each of the week 1 documents linked here



Advanced Quantitative Techniques

- o **Labs**
 - Also, I should note my lab grading approach for this class
 - Everyone starts with an "8/10" grade, and your work either persuades me to increase or decrease it from there



Advanced Quantitative Techniques

- o **Term Project**
 - **Proposal:** an overview of the dataset and analysis you have planned
 - Key elements of the proposal (see your syllabus for more detail)
 - o **1. The dataset:** how obtained, why interesting
 - o **2. The analysis:** what method(s) you anticipate could be used in analyzing your data, and for what purpose



Advanced Quantitative Techniques

- o **Term Project**
 - **Key point:** the project analysis must focus on one of the quantitative techniques featured in this class
 - So, it is important that you do some reading early in the semester to get a basic idea of what each method entails (so you can make a good decision on which one to use)



Advanced Quantitative Techniques

- o **Term Project**
 - **Another possibility for this semester:** one project focus you can consider combines *education* and *analysis*
 - o For lab 2 (spatial stats) or lab 3 (multiple regression), you can create an entirely new lab for this course:
 - Lab Scenario, Database Selection, Lab Guidance & Questions Document

Please consult with me if you are considering doing this project option

Advanced Quantitative Techniques

Term Project

- I am aiming to release the 2018 versions of these two labs together so you can review well before the proposal due date (Mar 8)

Goal:
release
the 2018
versions
of both
labs by
Feb 12

5 (Feb 15)	Spatial Autocorrelation	IAB
6 (Feb 22)	Multiple Correlation and Regression <i>Reading:</i> Field, Sections 5.3 - 5.7 (pages 157-184); Walder, Section on Multiple Correlation and Regression. <i>Reading with Normality Handout</i> document (see syllabus and handouts page on course website)	
7 (Mar 1)	Multiple Regression	IAB

Advanced Quantitative Techniques

Term Project

- Paper:** a 12-22 page document (core text only, not including title page, table of contents, tables, maps, etc.: see guidelines in syllabus)
- Focus of paper is again on a **dataset** and its **analysis**
- You are responsible for finding and selecting your dataset

Advanced Quantitative Techniques

Term Project

- Presentation:** a 15 minute oral presentation, with up to 5 more minutes for questions
- Include some general interest and theoretical context ("why should anyone care?"), but the majority of your presentation should focus again on the **data** and your **analysis**

Advanced Quantitative Techniques

Term Project

- See the **syllabus** and **other handouts** for more information on my overall expectations for the project as a whole

Advanced Quantitative Techniques

Attendance

- See the syllabus for the course **attendance policy**, noting the **participation grade** and penalties for unexcused absences
- I will grant excused absences for personal and family emergencies, but it is **your** responsibility to come to me about this (preferably **before** the absence occurs)

What Methods Will We Study?

I'd like to give you a quick run-through here of the statistical techniques we will cover this semester

- Note, there is a brief summary of these methods in your syllabus as well, and again note that last year's lecture slides are currently posted for your reference

What Methods Will We Study?

- **Spatial Autocorrelation**
 - Answers the question: does a given geographic distribution (points, areas) have a statistically-significant spatial pattern (clustered or uniform)?

Clustered Random Uniform

What Methods Will We Study?

- **Multiple correlation and regression**
 - Examine correlation and regression relationships among three or more variables (we will briefly review the simple meaning of “correlation” and “regression” before dealing with the multivariate case)

What Methods Will We Study?

- **Multiple correlation and regression**

What Methods Will We Study?

- **Spatial Pattern: Smoothing and Trend Surface Analysis**
 - Uncover and better represent overall spatial trends found on a map (uses some concepts developed in our multiple regression discussion)

What Methods Will We Study?

- **Spatial Pattern: Smoothing and Trend Surface Analysis**

Smoothing Example:

What Methods Will We Study?

- **Spatial Pattern: Smoothing and Trend Surface Analysis**

Smoothing Example:

What Methods Will We Study?

- **Spatial Pattern: Smoothing and Trend Surface Analysis**

Trend Surface Analysis Example:

TSA is really just multiple regression with location coordinates (longitude, latitude) as the independent variables

What Methods Will We Study?

- **Factor Analysis**
 - Simplify highly complex, multivariate situations by identifying major influences (particularly useful with datasets encompassing 10 or more variables)

Simple Example

Is there some common "factor" linking income and education? → We might say the two variables are really specific measures of "something more general": social status

Factor Analysis tries to uncover those "more general somethings" in a complex dataset

What Methods Will We Study?

- **Cluster Analysis**
 - A grouping or classification method that help us to identify groups of similar items in a long list of items

Simple Example

- You have a list of 100 cities and you want to identify groups of cities within the list: the cities that are most similar (e.g. Detroit and Cleveland, or Boston and Seattle)
- Cluster analysis would allow you to use a database for those cities (e.g. incomes, prominent industries, ethnic groups, etc. for each city) and objectively identify groups of similar cities

What Methods Will We Study?

- **Overall**
 - The methods selected for inclusion in this course include some of the most powerful methods geographers use
 - There are other useful methods around, but the ones in this course are the methods the UNT Geography faculty have selected as the ones you need to know

One Stats Detail

- **The one actual bit of "class content" I want to mention today: honest testing and the six steps of a formal statistical test**
 - You should have been exposed to these ideas (or something like them) before, but I want to run through this here as a starting point for this class
 - Proper statistical procedure is foundational for our work in this course

Class Discussion

- **Before we discuss the "six steps", let's come back to a theme related to the comic images I showed in opening today**
 - Do people misuse statistics? How? Why? How can we do better?
 - *Related discussion:* let's talk about the article reading for this first class

aps
Psychological Science

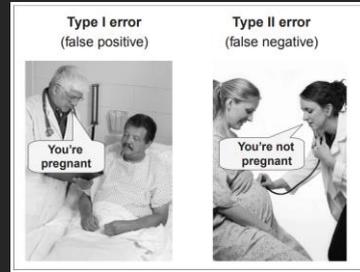
General Article

False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Joseph P. Simmons¹, Leif D. Nelson¹, and Uri Simonsohn¹
¹The Wharton School, University of Pennsylvania, and Thea School of Business, University of California, Berkeley

Psychological Science
2013, Vol. 24, No. 10, 1074–1079
© 2013 American Psychological Association
1074–1079, DOI: 10.1037/xap0000011
http://dx.doi.org/10.1037/xap0000011

Q: What is a false positive? Why is this a problem? What solutions does this article propose?



One Stats Detail

6 steps of a formal statistical test

1. Formulate hypotheses (H_0 and H_1)
2. Select a test statistic
3. Select your significance level and whether to use a one- or two-tailed test
4. Look up the critical value for acceptance/rejection of your H_0
5. Compute your test statistic from your data
6. Make a decision: accept/reject H_0 ?

How do these steps help with "honest testing"?

Class Discussion

- o I gave you one other reading to complete for today: "Understanding Data", a blog post by Joseph Kerski



Class Discussion

- o I gave you one other reading to complete for today: "Understanding Data", a blog post by Joseph Kerski
 - What is Kerski's main point? What do we need to build into our thinking to avoid problems?

"Encourage your data users ... to be critical of spatial data—knowing its source, who produced it, when and why it was produced, the scale at which it was produced, and its content."

Homework

- o 1. Do the reading for next week (see course schedule in the syllabus)
- o 2. Start looking at possible datasets and analytical methods you might use
 - See the course package for a list of the methods we will be examining (note that the package has a list and brief description of each method we will cover in this course)