ANT4930/ANG6930:
R for Archaeological Data Analysis and Visualization
Spring 2020

Instructor:
Dr. Daniel Contreras
Office: B346 Turlington
Phone: 294-6880
email: daniel.contreras@ufl.edu
Office Hours: W 1:00p -3:00p and by appt

T 11:45a - 12:35p Th 10:40a - 12:35p
1208H TUR

Course Description:

This course introduces students to the basic quantitative methods required to describe and analyze quantitative archaeological data. Each week, students will be introduced to a new statistical technique and asked to use that technique in order to solve a small research problem. Because many of the challenges archaeological face in such research have to do with acquiring and managing data (especially legacy data), this course will emphasize not only statistical methods, but also 1) how to construct and critically evaluate arguments grounded in quantitative data, and 2) how to explore, analyze, and display data in the open-source R statistical environment.

Course structure:

This course combines lectures, discussions, and problem sets in a hybrid seminar-lab format to facilitate peer instruction and hands-on learning. Each week will be broken into three sections. The first section (Tues 11:45a – 12:35p) will follow a seminar format during which time a designated lead student will facilitate discussion of the previous week's assignment. In the second section (Th 10:40a – 11:40a), the instructor will review a new technique through lecture and demonstration. During the third section (Th 11:45a – 12:35p), students will be given time to begin applying the new technique to solve the lab assignments, in consultation with one another and the instructor.

You are not required to bring a laptop to class, but it is strongly recommended in order to work through demonstrations and labs with your peers and the instructor. If this presents a problem for you, please see me as soon as practical so that we can come up with a solution.

Participation:

All students must participate in weekly discussions (**10%** of final grade). Participation includes attendance, active listening, and constructive contributions to discussion.

Problem Sets:

All students enrolled in the course must complete *all* of the assigned problem sets (**30%** of final grade). Problem sets are due the Tuesday after they are assigned and must be uploaded to Canvas *before the start of class* on the day that they are due.

Problem set will be graded both on the completion of the necessary analyses and on *the quality of the associated argument*. Problem sets will generally emphasize questions rather than methods, and your answers should *employ* methods to address questions, rather than simply demonstrating methodological competence. You should strive to be clear about what procedures you have used, what the results demonstrate, and what you conclude. Drennan models this in the text and prompts you in the problem sets; see also Kintigh 2005 and Cowgill 2005.

All students are expected to complete their problem sets in R using the KnitR package to generate .pdf reports. Students will learn how to use KnitR in class and can also consult the online modules here: rmarkdown.rstudio.com/lesson-1. This is itself an important skill because clear reporting facilitates reproducible

research that can be easily shared and disseminated. In addition to completing the problem sets, students must be prepared to share their findings in class.

> Note: All students are invited, for any problem set, to instead use their own data and address their own questions, in which case a brief statement (½ page maximum), explaining which data are used and how they are analogous to the assigned problem set, should accompany the assignment.

Facilitation:

All students must facilitate a proportional share of the class discussions reporting on lab assignments (based on the total number of enrolled students; **20%** of final grade). In addition to walking everyone through their solution to the assigned problem set (focusing more on the quantitative logic [Drennan] than the R implementation [Phillips]), the facilitator should emphasize how the application of this particular statistical technique solved the specific research problem at hand, as well as outlining the limitations and appropriate uses such analysis. Relating this to your own research or adding analogous examples related to your own interests is encouraged but not required.

Office hours:

You are welcome in my office hours either individually or in groups, and may use that time either to ask specific questions or simply to work with the benefit of someone available to help you through roadblocks. You are in no way required to come, but please note that this should be considered part of the education available to you, not a last resort.

Term Project:

All students must complete a term project (in three parts throughout the semester, combining for **40%** of final grade). The project should address a clearly formulated question using original analyses based on the skills learned in class. Students must formulate a research question by Week 6 (**13 Feb; 5% of course grade**), identify the data and analyses needed to answer the question by Week 12 (**26 Mar; 10% of course grade**), and submit draft text of the content of their poster in Week 15 (**16 Apr, 5% of course grade**). Final projects will take the form of a poster, **due 1 May and comprising 20% of your course grade**.

Readings:

Required Texts:

- Drennan, Robert D. *Statistics for Archaeologists: A Common Sense Approach* (2nd Edition), Springer. ISBN 978-1-4419-0413-3
  - You may use a print or digital version of this text (available in the University Bookstore and elsewhere), but it must be the 2nd edition and reading assignments will reference page numbers from the print edition.
- Phillips, Nathaniel D. *YaRrr! The Pirate's Guide to R.* https://ndphillips.github.io/piratesguide.html
  - Freely available online, both as an online book and as a .pdf download. Reading assignments reference the Table of Contents in the online version (https://bookdown.org/ndphillips/YaRrr/).

Supplementary Readings:

- Additional readings, when noted in the syllabus, will be available on Canvas or at the URL listed.
- For each chapter of Drennan, notes on implementation in R will be available on Canvas.
  - I *highly* recommend re-typing this code yourself - so that you really understand what it's doing and how – rather than simply copying-and-pasting. You should also make extensive use of the '?' operator in R, which you can always use to decipher what inputs a function requires, what its defaults are, and what it does.
- Data for problem sets (when necessary) will be provided on Canvas.

- You will also find any number of other guides to using R available online. None of them are strictly necessary for this course, but all of them are your friends.

Required Software:

- *R* (https://cran.r-project.org/)
- *R Studio* (https://www.rstudio.com/)
  - Both R and R Studio are free and open-source, and available for Mac OSX, Windows, and Linux.
- *Tabula* (https://tabula.technology/)
  - Tabula is also free and open-source, and available for Mac OSX, Windows, and Linux. We will not use it much in this course, but being familiar with it is extremely useful.

See Phillips (2017) §2.1 for installation instructions for R and R Studio.

Readings should be completed *in preparation for* the Thursday class in the week that they are listed. Problem sets should be completed *in preparation for* the Tuesday class in the week that they are listed (they relate to the readings from the previous week).

Data:

Data required for problem sets will be available on Canvas or in the archdata R library.
Students are encouraged to use data related to their research for their final project; if your research has not (yet) produced appropriate data, then you should start looking for data for your project *early* in the semester (this might involve published or unpublished data; consult with the instructor if you don't know where/how to look).

Course Goals:

Students will gain basic understanding of common statistical methods, develop the tools to be critical consumers of published archaeological data analyses, and learn how to explore, visualize, and analyze their own data. Students will also become comfortable with manipulating data and de-mystify quantitative analyses.

Course Policies:

- Late/Missed Assignments:
  - Late problem sets will not be accepted unless you have missed class that week.
  - Grades on components of your final project will be reduced by 10% for each day late. Extensions may be granted in case of particular need, but *must* be requested and approved in advance.
- Digital Devices:
  - Using a laptop in class is expected in this course. You are expected to resist the temptation to use it for anything other than course purposes during class-time.

Grading:

Grades in this course will be based on:

- Participation (10%)
- Problem Sets (30%)
- Facilitation (20%)
- Final Project (40%): Proposal (5%), Data and Analysis (10%), Draft Text (5%), and Poster (20%)

Course Schedule:

Week	Dates	Topic(s)	Assignment
1	7/9 Jan	Introduction: Data Analysis, R, and You	Reading [9 Jan]: Phillips §1.3 – 1.5, §2, §3.1 – 3.2; Drennan Ch. 1; Cowgill 2001; Orton 1999
2	14/16 Jan	Batches of Numbers / Comparing Batches and Distributions	Reading / Problem Set (14 Jan): Drennan Ch.2 & 3; Phillips §4 – 6 / Phillips §4.5, § 5.4, § 6.6, Drennan Ch.1 Practice, Ch.2 Practice, & Ch.3 Practice Reading (16 Jan): Drennan Ch.4 & 5; Phillips §7
3	21/23 Jan	Categorical Data	Problem Set (21 Jan): Phillips §7.4; Drennan Ch. 4 Practice & Ch.5 Practice Reading (23 Jan): Drennan Ch. 6; Baxter 1995; Kintigh 2005; Cowgill 2005
4	28/30 Jan	Samples and Populations	Problem Set (28 Jan): Phillips §8.7; Drennan Ch.6 Practice Reading (30 Jan): Drennan Ch.7 & 8; Phillips §8
5	4/6 Feb	Confidence and Population Means	Problem Set (4 Feb): Phillips §9.9 & §10.6; Drennan Ch.7 Practice Reading (6 Feb): Drennan Ch.9; Phillips §9 – 10
6	11/13 Feb	Medians, Resampling, and Proportions	Problem Set (11 Feb): Phillips §11.10; Drennan Ch.9 Practice Reading (13 Feb): Drennan Ch.10 & 11; Phillips §11; Cleveland 1994:4-21 ***Final Project Proposal Due (13 Feb)***
7	18/20 Feb	Comparing Two Sample Means	Problem Set (18 Feb): Drennan Ch.10 & 11 Practice Reading (20 Feb): Drennan Ch.12; Phillips §12
8	25/27 Feb	Comparing >2 Sample Means	Problem Set (25 Feb): Drennan Ch.12 Practice Reading (27 Feb): Drennan Ch.13; Phillips §13.1 - 13.3; Cowgill 1977
	3/5 Mar	**No Class – Spring Break**	
9	10/12 Mar	Comparing Proportions	Problem Set (10 Mar): Phillips §13.6 Q1 & Q4; Drennan Ch.13 Practice Reading (12 Mar): Drennan Ch.14; Phillips §14
10	17/19 Mar	Relating Variables	Problem Set (17 Mar): Phillips §13.6 Q2, Q5, & Q6, §14.8; Drennan Ch.14 Practice Reading (19 Mar): Drennan Ch.15, Phillips §13.5
11	24/26 Mar	More Sampling	Problem Set (24 Mar): Drennan Ch.15 Practice Reading (26 Mar): Drennan Ch.17-20; Phillips §15 ***Final Project Data & Analyses Due (26 Mar)***
12	31 Mar/ 2 Apr	Multivariate Analysis: Similarities between Cases	Problem Set (31 Mar): Phillips 15.6; Problem Set at end of notes/code for Drennan Ch.17-20 Reading (2 Apr): Drennan Ch.11 & 22; Phillips §16
13	7/9 Apr	Multivariate Analysis: Cluster Analysis	Problem Set (7 Apr): Phillips §16.5; Problem Set at end of notes/code for Drennan Ch.22 Reading (9 Apr): Drennan Ch.25; Phillips §17; Homburg 2005 [DC conference]

14	14/16 Apr	Correspondence Analysis	Problem Set (14 Apr): Phillips 17.6; TBD Reading (16 Apr): Baxter & Cool 2010; Shennan Ch.13 ***Final Project Draft Text Due (16 Apr)***
15	21 Apr	MDS / PCA	Reading (21 Apr): Drennan Ch.23 & 24
	28 Apr	***Poster Session - Final Project Poster due*** (2:00p – 4:00p)	

List of Supplementary Readings (available on Canvas or at the listed URL):

- Amrhein, Valentin, Sander Greenland, and Blake McShane. 2019. Retire statistical significance. *Nature* 567:305-307.
- Baxter, M.J. 1995. My innate ontology may be insecure but I can spell 'cat'. Unpublished ms.
- Baxter, Michael J, and Hilary EM Cool. 2010 Correspondence Analysis in R for archaeologists: an educational account. *Archeologia e Calcolatori* 21: 211–228.
- Cleveland, William S. 1994. *The Elements of Graphing Data*. New Jersey: Hobart Press. pp.4-21.
- Cowgill, G.L. 1977. The Trouble with Significance Tests and What We Can Do about It *American Antiquity* 42:350–68.
- Cowgill, George L. 2001: Past, present and future of quantitative methods in United States archaeology. In Stancic, Z. and Veljanovski, T., (eds.), *Computer Archaeology for Understanding the Past: CAA 2000*. Oxford: Archaeopress, 35-40.
- Cowgill, G.L. 2005. Things to Remember About Statistics (Whatever Else You Forget) *SAA Archaeological Record* 5:35.
- Homburg, J.A. 2005. Tips for Improving the Quality of Your Poster Presentation *The SAA Archaeological Record* 5:22–23.
- Kintigh, K.W. 2005. Writing Archaeology: Analyses and Archaeological Argumentation *The SAA Archaeological Record* 5:33–35.
- Marwick, Ben (editor). 2017. *How To Do Archaeological Science Using R*. https://benmarwick.github.io/How-To-Do-Archaeological-Science-Using-R/
- Orton, Clive R. 1999: Plus ¸ca change? – 25 years of statistics in archaeology. In Dingwall, L., Exon, S., Gaffney, V., Laflin, S and van Leusen, M. (eds.), *Archaeology in the Age of the Internet: CAA97*. Oxford: Archaeopress, 25-34.
- Otárola-Castillo, Erik, and Melissa G Torquato. 2018 Bayesian Statistics in Archaeology. *Annual Review of Anthropology* 47: 435–453.
- Wasserstein, Ronald L., Allen L. Schirm, and Nicole A. Lazar. 2019. Moving to a World Beyond "p < 0.05". *The American Statistician* 73: 1–19.

Not assigned but recommended (also available on Canvas):

- Baxter, Mike J. 2014. *Landmark papers in quantitative archaeology – a commentary*. 1–16.
- Baxter, Mike J. 2015. *Notes on Quantitative Archaeology and R*.
- Baxter, Mike J, and Hilary E.M. Cool. 2016. *Basic Statistical Graphics for Archaeology with R: Life Beyond Excel.*
- Cool, Hilary EM, and Michael J Baxter. 1999 Peeling the onion: an approach to comparing vessel glass assemblages. *Journal of Roman Archaeology* 12: 72–100.
- Flannery, Kent V (editor) 1976. *The Early Mesoamerican Village*. Academic Press, New York. pp.131-136, 159-160.

University Policies:

Attendance Policy, Class Expectations, and Make-Up Policy

Requirements for class attendance and make-up exams, assignments, and other work in this course are consistent with university policies that can be found at: catalog.ufl.edu/UGRD/academic-regulations/attendance-policies/

Students Requiring Accommodations

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, https://www.dso.ufl.edu/drc) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester.

Course Evaluation

Students are expected to provide feedback on the quality of instruction in this course. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via ufl.bluera.com/ufl/. Summary results of these assessments are available to students at https://evaluations.ufl.edu/results/.

University Honesty Policy

UF students are bound by The Honor Pledge which states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code (https://www.dso.ufl.edu/sccr/process/student-conduct-honor-code/) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor.

Campus Resources:

Health and Wellness

U Matter, We Care: If you or someone you know is in distress, please contact umatter@ufl.edu, 352-392-1575, or visit umatter.ufl.edu/ to refer or report a concern and a team member will reach out to the student in distress.
Counseling and Wellness Center: Visit counseling.ufl.edu/ or call 352-392-1575 for information on crisis services as well as non-crisis services.
Student Health Care Center: Call 352-392-1161 for 24/7 information to help you find the care you need, or visit shcc.ufl.edu/.
University Police Department: Visit police.ufl.edu/ or call 352-392-1111 (or 9-1-1 for emergencies).
UF Health Shands Emergency Room / Trauma Center: For immediate medical care call 352-733-0111 or go to the emergency room at 1515 SW Archer Road, Gainesville, FL 32608; ufhealth.org/emergency-room-trauma-center.

Academic Resources

E-learning technical support: Contact the UF Computing Help Desk at 352-392-4357 or via e-mail athelpdesk@ufl.edu.

Career Connections Center: Reitz Union Suite 1300, 352-392-1601. Career assistance and counseling services career.ufl.edu/.
Library Support: cms.uflib.ufl.edu/ask various ways to receive assistance with respect to using the libraries or finding resources.
Teaching Center: Broward Hall, 352-392-2010 or to make an appointment 352-392-6420.
General study skills and tutoring: teachingcenter.ufl.edu/ Writing Studio: 2215 Turlington Hall, 352-846-1138.
Help brainstorming, formatting, and writing papers: writing.ufl.edu/writing-studio/
Student Complaints: sccr.dso.ufl.edu/policies/student-honor- code-student-conduct-code/