GEOG 316/PLAN 351: Multivariate Statistics

Instructor: Dr. Chris Fletcher (chris.fletcher@uwaterloo.ca)

Lecture hours: 8:30-9:50am TuTh, PAS-1229

Labs on Tuesdays in EV2-1002A (Geddes Computing Lab – **entry code 1342**): 10:30-11:20am (section 101) and 11:30-12:20am (section 102).

Office hours: Wednesday 12:30-3:30pm in EV1-230

Teaching assistant: Bo Sun (b22sun@uwaterloo.ca)

COURSE DESCRIPTION: Specialists in the environmental sciences (e.g. geography, hydrology, ecology, atmospheric science) have to process, analyze and interpret a wide variety of data. Multivariate statistical techniques are particularly well suited for analysing diverse types of data (variables) simultaneously. This course will provide students with a foundation in the fundamentals of multivariate analysis, and practical experience of applying those techniques to problems in the environmental sciences.

COURSE LEARNING OBJECTIVES

By the end of the course students will be able to:

- Identify the most appropriate multivariate statistical techniques to use for data analysis in their academic discipline.
- Discriminate between the most common multivariate statistical techniques used in the environmental sciences.
- Apply these techniques to real data using the R statistical language.
- Interpret the output from R and communicate the results of these techniques.
- Develop confidence to explore beyond the techniques covered in this course.

PREREQUISITES: ENVS 278 (Advanced Environmental Research Methods).

RECOMMENDED READING:

Arnold, T., & Tilton, L. (2015). *Humanities Data in R*. Cham: Springer International Publishing. Free online text.

Everitt, B., & Hothorn, T. (2011). *An Introduction to Applied Multivariate Analysis with R*. New York, NY: Springer New York. Free online text.

Schumacker, R. E., 2016: *Using R with multivariate statistics*. Sage, Los Angeles, 383 pp.

WEB TUTORIALS AND RESOURCES:

https://www.kaggle.com/

http://trevorstephens.com/kaggle-titanic-tutorial/getting-started-with-r/

ADDITIONAL READING:

Last updated: Sep 6, 2017

Sharma, S. (1995). *Applied Multivariate Techniques* (Har/Dskt edition). New York: Wiley.

Peter J. A. Shaw, 2009. Multivariate Statistics for the Environmental Sciences, Wiley, ISBN: 978-0-340-80763-7, 248p.

McGarigal, K., Cushman, S., and S. Stafford, 2000. Multivariate Statistics for Wildlife and Ecology Research. Springer-Verlag: New York. 283 p.

Meyers, L.S., G. Gamst, and A.J. Guarino, 2006. Applied Multivariate Research: Design and Interpretation. Sage: Thousand Oaks. 722 p.

Shaw, P.J.A., 2003. Multivariate Statistics for Environmental Sciences. Arnold: London. 233 p.

Griffith, D.A. and Amrhein, C.G. 1997. Applied Multivariate Statistical Analysis for Geographers. Upper Saddle River, New Jersey: Prentice Hall, 800p.

COURSE WEB PAGE: Students registered in the course can access the course website via UW Learn (http://www.learn.uwaterloo.ca). The course website provides access to lecture presentations and selected research papers as well as lab assignments (and associated data). In addition, the course website supports announcements, discussion groups and e-mail. Please use the UW Learn e-mail tool for sending messages related to this course, but please carefully review the policy on email correspondence (below) before sending any email to the instruction team.

STUDENT ASSESSMENT:

Item	Sub-weights	Total weight
Assignments	A2=5%, A1, A3, A4=10%, A5=15%	50%
In-class tests	2 @ 25% each; non-cumulative	50%

ASSIGNMENTS: Assignments (PDF or .docx format, 12pt, single spaced) must be uploaded to the appropriate Dropbox on LEARN before **11:59pm** on each specified deadline date (see class schedule). Students will be provided with an answer template, and the data sources required to complete the assignments. You will apply the statistical analysis techniques learned in class to those data, and then interpret the results in the answer document. Show all mathematical working in your answer document, and include any graphs, tables and figures that are mentioned explicitly in the text. Upload your working R code (in a .R file) to the Dropbox in addition to your written answers.

WEEKLY CLASS SCHEDULE:

Date	Topic	Lab (Tue)/Assessment Activity
7-Sep	Course Organization and Overview; Review of ENVS278	
12-Sep	Review: Linear regression	Introduction to R

14-Sep	Logistic regression 1	
19-Sep	Logistic regression 2	A1: Logistic Regression [10%] due Friday Sep 22nd.
21-Sep	Introduction to multivariate analysis	
26-Sep	Graphical techniques for multivariate data	A2: Multivariate Techniques in R [5%] due Friday Sep 29th.
28-Sep	Inference for Multivariate Means	
3-Oct	Principal components analysis 1	A3: Principal components analysis [10%] due Friday Oct 13th.
5-Oct	Principal components analysis 2	
10 Oct	**Fall Study Break: No class or labs today**	
12-Oct	Principal components analysis 3	A3: continued. [Note: lab is on **Thursday**]
17-Oct	**Test 1** (held in class) [25%]	No Lab.
19-Oct	Multivariate ANOVA	
24-Oct	Classification I: linear discriminant analysis 1	A4: Classification [10%] due Friday Nov 3rd.
26-Oct	Classification I: linear discriminant analysis 2	
31-Oct	Classification II: cluster analysis 1	A4: continued.
2-Nov	Classification II: cluster analysis 2	
7-Nov	Introduction to Statistical learning models	A5: Machine Learning [15%] due Friday Nov 24th.
9-Nov	Random forest 1	
14-Nov	Random forest 2	A5: continued.
16-Nov	Support vector machines 1	
21-Nov	Support vector machines 2	A5: continued.
23-Nov	Support vector machines 3 / A5 help	
28-Nov	Review and prep for Test 2.	No lab.
30-Nov	**Test 2** (held in class) [25%]	

TESTS, CALCULATORS AND CRIB SHEETS: Two compulsory non-cumulative tests (held in class) will gauge student progress on material covered up to and including the class before the test. The tests will comprise multiple-choice and true-false questions, and long-answer questions based on the interpretation of output from R**.** Calculators are required for the tests and final exams but are *not* provided. Only calculators that are approved by the Math faculty are permitted in GEOG316 (see here). I strongly recommend buying one of these calculators at the start of term (for $12 - $15 at Walmart), and bringing it to every class.

Students are permitted to bring one page (double-sided, US Letter size 8.5" x 11") of hand-written study notes (aka a "crib sheet") with them into each test. Your crib sheet will be subject to inspection by exam proctors, and may be collected at the end of the test. Hand-written notes only: **photocopies are not permitted**.

Last updated: Sep 6, 2017

SOFTWARE & COMPUTING LABS: Access to the computing labs is restricted by passcode entry to those enrolled in the course. Food and/or drink are NOT permitted in the lab. The lab activities and assignments require using R software, which is available for students to use in the computing labs, or it can be downloaded for free. I strongly recommend downloading and using RStudio, both in labs and at home, instead of the standard R platform, because it is more user-friendly.

POLICY ON CELL PHONES: There is no valid academic reason to be using a cell phone during class. It is distracting to you, it is distracting to the instructor, and to the others around you. Therefore, during class, please keep cell phones in your bag/jacket on mute, rather than on your desk. Students often use their cell phone as a calculator, but this is not recommended because you will need an actual calculator during tests (see above).

BACKING UP YOUR DATA: Students are fully responsible for maintaining backups of any files and data you have modified. In computing the mantra is*: if it's not in two places it doesn`t exist.* Suitable options for backups include: networked drives; portable USB flash drives; external hard drives; laptops, or home desktop PCs; online "cloud" storage. No accommodation will be made for deadlines missed due to lost or corrupted data.

SUBMITTING WORK: Unless otherwise noted, all work should be submitted via UW Learn **in PDF format**. Please do not submit work in any format other than PDF (e.g., LibreOffice, Microsoft Word, Excel). Each assignment will have a specified due date and time on UW Learn.

LATE SUBMISSION: Any work submitted after the deadline **will not be graded** without a verifiable reason, for which official documentation is provided; for example, a notice or certificate of death in the event of bereavement, or a University Illness verification form.

STUDENT COLLABORATION: All assignments and tests are to be completed individually by each student. These pieces of work are expected to be the student's original work and to reflect her/his own thinking. Student collaboration in classroom exercises and during labs is permitted and encouraged, as long as all submitted work reflects each student's own thinking.

ACKNOWLEDGEMENT OF SOURCES: All sources used in the preparation of student work in this course must be acknowledged/cited in an appropriate way. I strongly recommend using free reference management software such as Zotero (www.zotero.org) or Mendeley (www.mendeley.com).

READABILITY AND CLARITY: Students are expected to present well organized and properly written work. The instruction team reserves the right not to grade any work submitted that does not meet commonly recognized standards of readability and clarity.

ILLNESS DURING TERM: Please refer to the University of Waterloo Policies regarding documentation and the management of requests for accommodation due to illness during the term. Illness verification forms are required for any student seeking accommodation for any course requirement missed due to an illness. Please refer to http://www.registrar.uwaterloo.ca/students/accom_illness.html for more information.

ACCOMMODATIONS DUE TO ILLNESS

MISSED DUE DATES AND TERM TESTS: If an assignment or midterm is missed because of illness, and all of the proper documentation is submitted on time, the weight of the missed assignments will be added to the final test. Assignment due dates will not be extended under any circumstances and term tests will not be written at a different time.

If the either test is missed for any reason, and all of the proper documentation is submitted on time, where sufficient other assignments and tests have been completed during the term a student's final grade may be calculated based on a reweighting of their other grades. If insufficient work is available to calculate a final grade, then the test(s) will be written the next time that this course is offered.

POLICY ON REGRADING ASSIGNMENTS:
If you notice an error in the assessment of your work please follow these steps:

1. Wait 72 hours after the assignment was returned before requesting a regrade
2. All requests for work to be regraded must be submitted to the instructor from your UWaterloo email account, in a message that fully describes the errors you believe were made. Verbal requests for regrades will not be accepted.
3. When writing your request, please follow the policy on student email (see below).
4. Be as specific as possible and list all relevant details, e.g., "*my marks were summed incorrectly for questions 1–5*".
5. If another student's work is used as an example or reason for an error in grading, both assignments will be subject to a regrade.
6. The entire assignment will be regraded, not the just the errors indicated in the written request. The resulting grade may increase or decrease depending on the result of the regrading.

ATTENDANCE: Attendance will not be taken at any lectures or labs. However, it is highly recommended that students attend all scheduled lectures and labs, since research shows very clearly that students who attend class earn higher grades.

POLICY ON EMAIL CORRESPONDENCE: Face-to-face contact time is available through office hours, during and after lectures, and in the tutorials. Discussion forums are available in LEARN for any inquiries that are of general interest (e.g., clarification of assignment instructions, or a concept introduced in class). Students are strongly

encouraged to attend office hours to discuss any issues related to the course, and email should only be used when none of these other options is appropriate. However, *if your question or concern cannot wait until the next lecture or office hour* then please remember these policies when sending email to the instructor or TAs:

- Always send emails from your University of Waterloo email account or from the email tool within UW LEARN. The instruction team reserves the right not to respond to emails sent from non-UW accounts.
- All emails should have the following subject line: "GEOG316: <<insert your subject here>>"
- The instructor must be copied on all course-related email communication with the TAs (the TAs are instructed not to respond to direct emails without the instructor cc'd).
- If your email includes an attachment, describe the contents of the attachment in the email.
- Be polite, respectful and professional.
- Proofread your email and use correct grammar and punctuation.
- Always use an appropriate greeting, and sign your full name.
- Allow the instructor or TA at least two business days to respond before sending the request again. Mark all urgent matters "URGENT" in the subject line.
- The instructor and TAs reserve the right to reply to you along with the entire class, if the question is deemed to be relevant to other students on the course. Alternatively, we may post the question and response in a discussion forum on LEARN. The questioner's identifying personal information will be removed from such announcements.

UW / Faculty of Environment Policies

Academic Integrity: In order to maintain a culture of academic integrity, members of the University of Waterloo community are expected to promote honesty, trust, fairness, respect and responsibility. The University's guiding principles on academic integrity can be found here: http://uwaterloo.ca/academicintegrity

ENV students are strongly encouraged to review the material provided by the university's Academic Integrity office specifically for students: http://uwaterloo.ca/academicintegrity/Students/index.html

Students are also expected to know what constitutes academic integrity, to avoid committing academic offenses, and to take responsibility for their actions. Student who are unsure whether an action constitutes an offense, or who need help in learning how to avoid offenses (e.g., plagiarism, cheating) or about "rules" for group work/collaboration should seek guidance from the course professor, academic advisor, or the Undergraduate Associate Dean. Students may also visit this webpage: https://uwaterloo.ca/library/get-assignment-and-research-help/academic-integrity/academic- integrity-tutorial

Discipline: When misconduct has been found to have occurred, disciplinary penalties will be imposed under Policy 71 – Student Discipline. For information on categories of offenses and types of penalties, students should refer to Policy 71 - Student Discipline, https://uwaterloo.ca/secretariat-general-counsel/policies-procedures-guidelines/policy-71

Religious Observances: Student needs to inform the instructor at the beginning of term if special accommodation needs to be made for religious observances that are not otherwise accounted for in the scheduling of classes and assignments.

Grievance: Students who believe that they have been wrongfully or unjustly penalized have the right to grieve; refer to Policy #70, Student Grievance: https://uwaterloo.ca/secretariat-general-counsel/policies-procedures-guidelines/policy-70. When in doubt please contact the department's administrative assistant for further assistance.

Appeals: A decision made or penalty imposed under Policy 70 (Student Petitions and Grievances) (other than a petition) or Policy 71 (Student Discipline) may be appealed if there is a ground. A student who believes he/she has a ground for an appeal should refer to Policy 72 (Student Appeals) www.adm.uwaterloo.ca/infosec/Policies/policy72.htm.

Turnitin: Plagiarism detection software (Turnitin) **will not** be used to screen assignments on this course.

Unclaimed Work: Unclaimed assignments will be retained for at least one month after term grades become official in quest. After that time, they will be destroyed in compliance with UW's confidential shredding procedures.

Note for students with disabilities: AccessAbility Services, located in Needles Hall, Room 1401, collaborates with all academic departments to arrange appropriate accommodations for students with disabilities without compromising the academic integrity of the curriculum. If you require academic accommodations to lessen the impact of your disability, please register with AccessAbility Services at the beginning of each academic term.

Co-op interviews and class attendance: Co-op students are encouraged to try and choose interview time slots that result in the least amount of disruption to class schedules. When this is challenging, or not possible, a student may miss a portion of a class meeting for an interview. Instructors are asked for leniency in these situations; but, a co-op interview does not relieve the student of any requirements associated with that class meeting. When a co-op interview conflicts with an in-class evaluation mechanism (e.g., test, quiz, presentation, critique), class attendance takes precedence and the onus is on the student to reschedule the interview. CECA provides an interview conflict procedure to manage these situations. Students will be required to provide copies of their interview schedules (they may be printed from JobMine/WaterlooWorks) should there be a need to verify class absence due to co-op interviews.

Intellectual Property: Students should be aware that this course contains the intellectual property of their instructor, TA, and/or the University of Waterloo. Intellectual property includes items such as: Lecture content, spoken and written (and any audio/video recording thereof); Lecture handouts, presentations, and other materials prepared for the course (e.g., PowerPoint slides); Questions or solution sets from various types of assessments (e.g., assignments, quizzes, tests, final exams); and Work protected by copyright (e.g., any work authored by the instructor or TA or used by the instructor or TA with permission of the copyright owner).

Course materials and the intellectual property contained therein, are used to enhance a student's educational experience. However, sharing this intellectual property without the intellectual property owner's permission is a violation of intellectual property rights. For this reason, it is necessary to ask the instructor, TA and/or the University of Waterloo for permission before uploading and sharing the intellectual property of others online (e.g., to an online repository).

Permission from an instructor, TA or the University is also necessary before sharing the intellectual property of others from completed courses with students taking the same/similar courses in subsequent terms/years. In many cases, instructors might be happy to allow distribution of certain materials. However, doing so without expressed permission is considered a violation of intellectual property rights.

Please alert the instructor if you become aware of intellectual property belonging to others (past or present) circulating, either through the student body or online. The intellectual property rights owner deserves to know (and may have already given their consent).