Multivariate Statistics for Behavioural Research:
PSYC-5151-2016F
Department of Psychology, Lakehead University

Instructor: Bruce Weaver, MSc, bweaver@lakeheadu.ca
Office: MS-2006
Phone: 807-346-7704

Office hours: By appointment
Class Times: Tuesdays & Thursdays, 10:00 to 11:30 AM (AT-3003)
Pre-requisites: PSYC-4111 or equivalent
Discussion Forum: https://groups.google.com/forum/#!forum/psyc-5151-2016f

Course Description

The university calendar course description for PSYC-5151 says:

> This course will focus on various advanced statistical techniques that can be used in psychological research. Topics will include multiple regression, factor analysis, cluster analysis, structural [equation] modeling, discriminant function analysis, multivariate analysis of variance, meta-analysis, canonical correlation, sequential analyses, and loglinear (*sic*) modeling.

My approach shall be to focus on visualization and interpretation of the results from various types of general and generalized linear models that are most commonly used in psychological research (e.g., OLS linear regression, ANOVA, multilevel models, logistic regression). Brief overviews of the other topics listed above (plus a few others I consider important) will be included by means of required readings with accompanying written assignments.

Course Objectives and Learning Outcomes

My objectives for this course are very much in keeping with the following statement by Michael Mitchell, the author of a textbook we shall be using (personal communication, 8-May-2016):

> When I think of teaching students, I think of training them for their future careers... it is like "boot camp" and when they leave boot camp (the class), we hope we gave them enough training and tools to survive in the battles they will encounter on their own.

Mitchell's *boot camp* metaphor is very apt in many ways. It suggests, among other things, that given the short duration of the course (12 weeks), the best approach is to provide students with a good grounding in the most important foundational types of statistical analysis: linear regression, ANOVA, multilevel models, logistic regression, the general linear model (GLM) and the generalized linear model (GzLM). Students with a good grounding in these types of models will be well-equipped to learn other more specialized techniques through self-study as the need arises in their own research programs.

Course Materials

1. **Required Textbooks & Articles**:
    a. Mitchell, Michael N. *Interpreting and Visualizing Regression Models Using Stata*. Stata Press, 2012.[1] (You will also need the datasets that accompany the book.)
    b. Norman, GR. & Streiner, DL. *PDQ Statistics (3rd Ed.)*. BC Decker Inc., 2003.
    c. PDFs of required articles will be made available.
2. **Recommended Readings**:[2]
    a. Abelson, Robert P. *Statistics as principled argument*. Erlbaum, Hillsdale, 1995.
    b. Aiken, Leona S., Stephen G. West, and Raymond R. Reno. *Multiple regression: Testing and interpreting interactions*. Sage, 1991.
3. **Required Software**: Students are required to purchase a license for **Stata/IC**.
    - **Please note that Stata is not installed in any of the university's computer labs, so you must purchase your own copy of it**.
    - Note that *Small Stata* cannot accommodate all of the datasets that accompany the required textbook. Therefore, you must purchase *Stata/IC*. I recommend the *perpetual license version* at a cost of $198 USD. However, less expensive limited term licenses are also available ($125 USD for 1 year, $75 USD for 6 months). See this page for all student pricing options.[3]
    - I recommend that you install Stata on a laptop computer that you can bring along to class, as some class time will be devoted to working through the examples in IVRMUS[4]. (Note that the single-user license agreement allows you to install Stata on up to 3 computers, with the understanding that you are the only user.)

See my Stata page for some online resources that I have found helpful.

Assessment

Description	Due	Value
Stata assignments	See below	50%
Written assignments	See below	50%
Please note that **I do not give extra work**. Final marks will be computed using only the components listed here.		

Submission of Assignments

[1] This book can be purchased from the Stata Bookstore. The paperback and e-book versions cost $58 and $52 USD respectively. Current prices at Amazon.ca are $87.56 and $78.19 for the paperback & Kindle editions respectively.
[2] Do not feel you must read these during the course. I do hope you will read them sometime during your graduate education though, as I think you'll find both of them very helpful. Both books are available through the Library.
[3] If there are any students who cannot afford a license for *Stata/IC*, I can make available smaller versions of any datasets that have more than 1200 cases (or observations, as Stata users call them) that will work with Small Stata (which can be purchased for as little as $38 USD). But note that these smaller files will contain random samples of the cases from the corresponding files Mitchell uses, and therefore, the results of analyses done with them will not match his results exactly.
[4] IVRMUS is short for *Interpreting and Visualizing Regression Models Using Stata*.

All assignments are to be submitted electronically via the *Desire2Learn* (D2L) Dropbox. They are due at 11:59 PM (one minute before midnight) on their due dates.

Stata Assignments

This section describes several assignments that will require you to generate Stata DO files. (DO files in Stata are the equivalent of syntax files in SPSS.)

In all of your DO files, include log commands at the top and bottom to save the command syntax and output to text files that you can submit for grading. Here is an example from my DO file for Chapter 10 of IVRMUS with the log commands highlighted:

```
log using "C:\bw\LU\PSYC\5151\2016\Stata log files\IVRMUS-10", ///
replace text name(log10)

* ================================================
* File:  IVRMUS-10.do
* Name:  Bruce Weaver, bweaver@lakeheadu.ca
* Date:  13-Jul-2016
* ================================================

// Stata commands to perform Chapter 10 analyses here

log close log10
```

Assignment	Description/Instructions	Value	Due Date
S1	Write a Stata DO file for GSWS, Chapter 1. Submit your log file.	2%	See note
S2-S18	Write Stata DO files for IVRMUS, Chapters 2-18. Submit your log files.	34%	See note
S19	Data analysis assignment—instructions to be provided in a separate document.	14%	8-Dec-2016

NOTE: Assignments S1 to S18 are due 10 days after their start dates as shown in the class schedule. For example, the start date for assignment S1 is 6-Sep-2016; therefore, it is due on 16-Sep-2016. The start date for assignment S18 is 10-Nov-2016; therefore, the due date is 20-Nov-2016.

Assignments S1 to S18 consist largely of transcribing Mitchell's Stata commands from his book to DO files, and therefore could be done in a very mindless and purely mechanical fashion. But if you approach them that way, you will learn very little. I encourage you to read each chapter of IVRMUS carefully as you create your DO files. As you transcribe a command, think about what it instructs Stata to do, and what you expect the output to look like. I also encourage you to include comments that indicate the relevant page numbers in Mitchell's book, what the syntax does, and any other information that you think might be helpful to someone reading your DO file. (Bear in mind that that *someone* could be you yourself at a later date when you've forgotten what you were thinking!) Because you are essentially transcribing commands in S1 to S18, everyone will have essentially identical commands in their DO files. That is fine. I shall be looking for evidence of independent work in the comments that you insert, however. In other words, *I do not want to see two DO files that look identical, including the comments*.

Written Assignments

These assignments require you to read various chapters or articles and then summarize the main points. For W1 to W9, your summaries should be *no more than one page* (using a reasonable font size that will not make me squint). The main thing I'll be looking for is evidence that you have read (and understood) the key points in the assigned chapters and articles. For W10, which is more substantial, you should aim for 2-3 pages at most. You may discuss the readings with classmates, but your written summaries must be generated independently.

Assignment	Description/Instructions	Value	Due Date
W1	Read chapters 13-15 in *PDQ Statistics* and the article by Huberty & Morris (1989). Write a brief summary of chapters 14-15 (MANOVA & DFA) that incorporates relevant information from Huberty & Morris.	4%	See note
W2	Read chapter 16 (EFA) in *PDQ Statistics* and the article by Preacher & MacCallum (2003). Write a brief summary of chapter 16 that incorporates relevant information from Preacher & MacCallum.	4%	See note
W3	Write a summary of chapter 17 (Path Analysis & SEM) in *PDQ Statistics*.	4%	See note
W4	Write brief summaries of chapters 18-19 (Cluster Analysis & Canonical Correlation) in *PDQ Statistics*.	4%	See note
W5	Write a brief summary of the Streiner & Lin (1998) article on log-linear analysis.	4%	See note
W6	Write a brief summary of the Field & Gillett (2010) article on meta-analysis.	4%	See note
W7	Read both articles by Schulz & Grimes in volume 365 of *The Lancet* (2005) [first article; second article]. Write a brief summary of the second article (*Multiplicity in randomised trials II: subgroup and interim analyses*). (Note that this assignment concerns the *sequential analysis* topic included in the university calendar description for PSYC-5151.)	4%	See note
W8	Write a brief summary of the article by Babyak (2004).	4%	See note
W9	Write a brief summary of Hill (1965). (See Streiner & Norman, 2010, for further commentary on Hill's criteria.)	4%	See note
W10	Read the following resources: ○ Affirming the consequent ○ AC Resource 1 ○ AC Resource 2 ○ Confirmation bias ○ CB Resource 1 ○ CB Resource 2 – Wason's 4-card problem (original) ○ CB Resource 3 – Wason's 4-card problem (modified) ○ MacKinnon, Krull & Lockwood (2000) Bearing in mind Hill's (1965) criteria for inferring causation (from assignment W9) and drawing upon material from the required resources for this assignment, discuss the appropriateness of concluding that "M is a mediator of the association between X and Y" after one has obtained positive results from a mediation analysis using observational data. **HINT**: Think about the definition of statistical mediation, and the distinction between mediation and confounding. As this is a more substantial assignment than the others, you should aim for 2-3 pages.	14%	1-Dec-2016

NOTE: Assignments W1 to W9 are due 10 days after their start dates as shown in the class schedule. For example, the start date for assignment W7 is 27-Oct-2016; therefore, the due date is 6-Nov-2016.

Tentative Class Schedule

Date	Lectures & Labs	Start Date For: [1]
06-Sep-2016 (T)	Overview of the course	S1
08-Sep-2016 (Th)	Lab	S2; W1
13-Sep-2016 (T)	Lecture: Simple Linear Regression	S3
15-Sep-2016 (Th)	Lab	S4; W2
20-Sep-2016 (T)	Lecture: Simple Linear Regression	S5
22-Sep-2016 (Th)	Lab	S6; W3
27-Sep-2016 (T)	Lecture: Multiple Linear Regression	S7
29-Sep-2016 (Th)	Lab	S8; W4
04-Oct-2016 (T)	Lecture: Multiple Linear Regression	S9
06-Oct-2016 (Th)	Lab	S10; W5
11-Oct-2016 (T)	**Reading week—no classes**	
13-Oct-2016 (Th)		
18-Oct-2016 (T)	Lecture: Multiple Linear Regression	S11
20-Oct-2016 (Th)	Lab	S12; W6
25-Oct-2016 (T)	Lecture: Multiple Linear Regression	S13
27-Oct-2016 (Th)	Lab	S14; W7
01-Nov-2016 (T)	Lecture: Logistic Regression	S15
03-Nov-2016 (Th)	Lab	S16; W8
08-Nov-2016 (T)	Lecture: Multilevel Models	S17
10-Nov-2016 (Th)	Lab	S18; W9
15-Nov-2016 (T)	Lecture: GLM & GzLM	Work on S19 & W10
17-Nov-2016 (Th)	Lab	
22-Nov-2016 (T)	Lab	
24-Nov-2016 (Th)	Lab	
29-Nov-2016 (T)	Lab	
01-Dec-2016 (Th)	Lab	

1. All due dates can be seen by clicking on **Assignments** in D2L. Assignments S1-S18 and W1-W9 are due 10 days after their start dates. W10 is due on 1-Dec-2016, and S19 is due on 8-Dec-2016.

Tips to Help You Succeed in This Course

1. *Don't fall behind!*
2. Attend regularly and ask questions when things are unclear. (Review point #1).
3. Submit assignments earlier than the 10-day deadline if possible. (Review point #1).
4. Post questions to the class discussion forum.
5. Monitor the class discussion forum regularly—you might learn something from the questions that your classmates post.
6. Review the PowerPoint slides and read the assigned textbook chapters carefully.
7. Write down difficulties you have with the material (so that you can ask questions later).
8. Study with classmates—share notes, quiz each other, etc.
9. Complete all assessment items.
10. If problems arise, speak to the instructor as they arise, not at the end of the semester!

Course and University Policies

Course Policy on Late Submissions

Unless there are extenuating circumstances (i.e., documented medical or compassionate reasons) that have been discussed with the instructor in advance, **late submissions will be penalized 10% (of a perfect score) per calendar day late**. E.g., for an assignment that is scored out of 100, if your grade is 75, but is two days late, your penalized score will be 75 – 20 = 55. If your grade is 100, but you are 5 days late, your penalized score will be 100 – 50 = 50.

University Policy on Academic Dishonesty

The University takes a most serious view of offences against academic honesty such as plagiarism, cheating and impersonation. Penalties for dealing with such offences will be strictly enforced. A copy of the "Code of Student Behaviour and Disciplinary Procedures" including sections on plagiarism and other forms of misconduct may be obtained from the Office of the Registrar.

The following rules shall govern the treatment of candidates who have been found guilty of attempting to obtain academic credit dishonestly:

a) The minimum penalty for a candidate found guilty of plagiarism, or of cheating on any part of a course will be a zero for the work concerned. A candidate found guilty of cheating on a formal examination or a test, or of serious or repeated plagiarism, or of unofficially obtaining a copy of an examination paper before the examination is scheduled to be written, will receive zero for the course and may be expelled from the University.
b) Students disciplined under the Code of Student Behaviour and Disciplinary Procedures may appeal their case through the Judicial Panel.

Note: "Plagiarism" shall be deemed to include:

1. Plagiarism of ideas as where an idea of an author or speaker is incorporated into the body of an assignment as though it were the writer's idea, i.e. no credit is given the person through referencing or footnoting or end noting.
2. Plagiarism of words occurs when phrases, sentences, tables or illustrations of an author or speaker are incorporated into the body of a writer's own, i.e. no quotations or indentations (depending on the format followed) are present but referencing or footnoting or end noting is given.
3. Plagiarism of ideas and words as where words and an idea(s) of an author or speaker are incorporated into the body of a written assignment as though they were the writer's own words and ideas, i.e. no quotations or indentations (depending on format followed) are present and no referencing or footnoting or end noting is given.

Accommodations

Lakehead University is committed to achieving full accessibility for persons with disabilities. Part of this commitment includes arranging academic accommodations for students with disabilities to ensure they have an equitable opportunity to participate in all of their academic activities. If you think you may need accommodations, you are strongly encouraged to contact Student Accessibility Services (SAS) and register as early as possible. For more information, please visit http://studentaccessibility.lakeheadu.ca.

Appendix: Hyperlinks Used in This Document

The original electronic version of this document contains several hyperlinks. Those hyperlinks are listed below, for the benefit of anyone who is using a (printed) hard copy of the document. They are listed in the order in which they appear in the document.

Description	Hyperlink
Calendar description for PSYC-4111	http://navigator.lakeheadu.ca/Catalog/ViewCatalog.aspx?pageid=viewcatalog&topicgroupid=16812&entitytype=CID&entitycode=Psychology+4111
Calendar description for PSYC-5151	http://navigator.lakeheadu.ca/Catalog/previouscals/2012-2013/pg134.html#25935
Stata Bookstore site for the required textbook, *IVRMUS*	http://www.stata.com/bookstore/interpreting-visualizing-regression-models/
Datasets for *IVRMUS*	http://www.stata-press.com/data/ivrm.html
Student pricing for Stata	http://www.stata.com/order/new/edu/gradplans/student-pricing/
Stata's single-user license agreement	http://www.stata.com/order/end-user-license-agreement/S14EULA.pdf
Some online Stata resources	https://sites.google.com/a/lakeheadu.ca/bweaver/Home/statistics/stata
Getting Started With Stata	http://www.stata.com/features/documentation/
Stata comments in DO files	http://www.stata.com/manuals14/pcomments.pdf
Huberty & Morris (1989)	https://wweb.uta.edu/management/Dr.Casper/Fall10/BSAD6314/Coursematerial/Huberty%20&%20Morris%201989%20-%20MANOVA%20vs.%20ANOVA.pdf
Preacher & MacCallum (2003)	http://www.quantpsy.org/pubs/preacher_maccallum_2003.pdf
Streiner & Lin (1998)	https://www.researchgate.net/publication/13480629_Life_after_chi-squared_An_introduction_to_log-linear_analysis
Field & Gillett (2010)	http://onlinelibrary.wiley.com/doi/10.1348/000711010X502733/pdf
Schulz & Grimes (2005a)	http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(05)66461-6/fulltext
Schulz & Grimes (2005b)	http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(05)66516-6/fulltext
Babyak (2004)	http://people.duke.edu/~mababyak/papers/babyakregression.pdf
Hill (1965)	https://www.edwardtufte.com/tufte/hill
Streiner & Norman (2010)	https://www.researchgate.net/publication/288380181_Causation
Affirming the consequent: Resource 1	http://www.skepdic.com/affirmingtheconsequent.html
Affirming the consequent: Resource 2	http://www.philosophy-index.com/logic/fallacies/affirming-consequent.php
Confirmation bias: Resource 1	http://skepdic.com/confirmbias.html
Wason 4-card problem (original)	http://www.allfunandgames.ca/brain/fourcard.shtml
Wason 4-card problem (modified)	https://www.youtube.com/watch?v=qNBzwwLiOUc
MacKinnon, Krull & Lockwood (2000)	http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2819361/