SOCIOLOGY 252 H1F:

INTERMEDIATE QUANTITATIVE METHODS IN SOCIOLOGY

Blair Wheaton

Department of Sociology

Fall, 2021

Office: Room 376
Department of Sociology
725 Spadina Avenue

Phone: 416-978-6317

E-mail: *blair.wheaton@utoronto.ca*

Web Site: Quercus

Place: SS 1088

Time: **Tuesday** 2-4

Tutorials: **Wednesday 4-6** or **Friday 12-2** (NOT weekly) FE36, Sociology, 725 Spadina.

T.A. ***Rachel Meiorin (rachel.meiorin@mail.utoronto.ca).***

Calendar Description

Provides students with the opportunity to develop an understanding of the logic of multivariate analysis by applying various strategies for the analysis of complex multivariate data.

Prerequisites:

SOC202H1** or equivalent.* ***Students without this/these prerequisite/s will be removed at any time discovered and without notice."

Overview

This course is a follow-up to a first methods course that includes some introduction to descriptive and inferential statistics and the logic of multivariate analysis. We begin by going back to go forward, reviewing some material on the fundamental building blocks of statistical theory and statistical inference.

We will look at tables to establish the nature of association and the logic of controlling for variables, and then introduce regression and correlation, followed by multiple regression. Multiple regression is a universal technique used in most disciplines that apply quantitative data: it is very flexible, and accommodates most styles of thinking and theorizing.

Near the end of the course, we will consider the extensions of this model that reflect its flexibility, including interaction effects between variables (intersections), nonlinear relationships, and models for categorical outcomes.

Where and How

In the first two weeks, this class will either meet online only, using Zoom, or in a hybrid format, with in-class participation for students who can come to class with a simulcast online for students who cannot.

After the first two weeks, the delivery mode may shift: we will either do all in-class, if all students are able to attend, or the hybrid model, if not.

We will be polling the class before the beginning of class to figure out what delivery method is best.

Required Work

Assignments

There will be two computer assignments, in which you will analyze data I provide. The first assignment will be on cross-classification and tables; the second will use multiple regression. I will teach the use of SAS and describe the data you can analyze in scheduled tutorials. There will be some minimal programming your group will do, depending on your choice of variables from the data provided, but I will post very specific templates for what you need to do, so that you can just edit these template programs with the specifics of the variables you choose.

We will not hold tutorials every week. Tutorials will be held on scheduled weeks in the course; these weeks are listed in the class schedule below. In general tutorials are held on weeks prior to due assignments and before tests. The purpose of tutorials is to introduce you to the software to be used in this class (SAS), the data to be used in class and for assignments, and to review assignment questions and test questions in the weeks before required work is due.

Because tutorials are not held every week as per the schedule below, both the TA and I will be available for individual meetings you schedule by email. This is to ask questions about class material or assignments.

Both assignments will be done in pairs. You will form pairs voluntarily. Grades will be given at the group level and will apply equally to both students.

Assignments are written up as short papers meant to analyze a specific research question, following the requirements of the question(s) in the assignment. Results from your computer analyses ***should*** be embedded into your assignment as tables. SAS output can be copied and pasted directly into your Word document.

Tests

There will be an in-class term test, and a final assessment during the scheduled final assessment period. The term test will focus on problems, including some calculation and/or interpretation. ***This test will be held in the tutorial room (FE36)***. The final is non-cumulative, and will include material only from the section on correlation and regression forward.

I will hand out some practice questions for tests.

In-Class Questions

There will be posted questions at the end of many classes, with the opportunity to answer the question in real time on Quercus. We will review answers to these questions at the end of each class. You will be given grades for being there and for attempting to answer the question that week. However, you only need to answer a question in 8 out of the last 10 weeks to get the full grade. Whether you answer the question correctly won't matter to the grade.

Software

This year, this class will use the free version of SAS University edition. SAS is the largest statistical software package in use around the world. You will be able to install SAS on your laptop, whether a PC or a Mac. But, in tutorials, you will be using the PC installed version of SAS, which looks very similar to the version you will install on your laptop.

I will post instructions about installing SAS.

Data

This year we will use a specific data set for assignments ---the 2015 General Social Survey for Canada. This is a long-term survey run every year by Statistics Canada on a representative sample of Canadians 15 and over. The sample size is 27,695.

We will provide derived versions of these data with already constructed variables, but we will also include the raw variables used to create the constructed variables so that you can create your own variables as well. This will be a choice you make: if you don't want to learn programming on the second assignment, you can use already constructed variables.

Due Dates and Weights for Required Work:

Given that the campus is open, physical copies of assignments will be handed in on the due date before 5 pm at the Department of Sociology, at 725 Spadina, in room 225 (200-level mailbox). If the campus is not open, assignments should be emailed to the TA by the same time and date as for physical assignments.

Provisional due dates for required work are as follows:

Work	Date	Weights
1. First Assignment	Wednesday, October 20	20%
2. Term Test	Friday, October 29	20%
3. Second Assignment	Friday, December 3	20%
4. In-class ***participation*** on 8 of 10 questions in 10 weeks	Last 10 weeks of class	10%
5. Final Assessment	final assessment period	30%

Please note: Late assignments will be given a 10% reduction in the grade immediately. This means that the assignment will be given a weight equal to .90 of the assigned weight. This increases to 20% if the assignment is late more than 3 school days. Assignments will not be accepted if they are more than 7 school days late. Your grade will be zero on that part of the course.

Quercus

Quercus will be used in this course for three purposes: 1) I will post data, assignments, SAS examples, and course readings there; 2) I will post most lecture Power Point materials there – when they are presented in class; and 3) I will post in-class questions there that you can answer in real time at the end of class.

I will use Top Hat in class to ask questions. The class will answer the questions online, anonymously, and we will see the results live. This is important in this kind of class: issues in understanding material should be addressed at the time, rather than weeks later. I will also use Top Hat to post questions you will answer for homework exercises. These are short questions that should take only 10-15 minutes to answer in total.

Required Reading:

Required reading will occur in two forms in this class, and in sequence. We start with a set of notes I will post online, week by week, following the topics of the first four weeks. These notes cover the "review" part of the course, and are relevant over the first four weeks only.

After the first four weeks, you will be expected to read assigned sections from this textbook, recently published and available online:

Blair Wheaton and Marisa Young. 2021. ***Generalizing the Regression Model: Techniques for Longitudinal and Contextual Analysis. Sage: Thousand Oaks California.***

Please note: This text is for a range of second courses in statistics, and we will only be dealing with sections of Chapters 1, 2, 3, 4, and 5.

Optional Reading Online:

This course includes optional material that could be helpful as backup. None of these readings are required.

Barbara Illowsky and Susan Dean. Introductory Statistics. 2018. Rice University.

....***Also:*** Introductory Statistics - OpenStax

Larry D. Schroeder et al. Understanding Regression Analysis: An Introductory Guide. Beverley Hills, CA: Sage Publications. 1980. ***… or……***

Alan Sykes. An Introduction to Regression Analysis. Chicago Working Paper on Law and Economics.

Thomas Brambor et al. 2005. Understanding Interaction Models. Political Analysis 13: Pp. 1-11.

Chao-Ying Peng et al. 2010. An Introduction to Logistic Regression Analysis. Journal of Educational Research 96: Pp. 1-9.

The class schedule includes references to these online sources and to posted articles which will supplement the notes used in class. You should especially read introductory articles for multiple regression and logistic regression. They are intended as basic introductions for audiences who know nothing about these topics.

Student Accommodations

Please see me if you have a disability or other need that requires accommodation or classroom modification. I will be glad to help you in whatever way I can.

Missed Deadlines or Tests

Medical Issues: (NOTE: Because of Covid-19, students do NOT need to submit the usual documentation, i.e., medical notes or the Verification of Illness forms).

Students who miss the test, or are late in submitting an assignment for medical reasons, need to email the instructor (not the TA), and also declare their absence on the system (ACORN) on the day of the test or assignment due date.

Students who miss the test, or are late in submitting an assignment for other reasons, such as family or other personal reasons, should request their College Registrar to email the instructor.

Term Test

If you miss the term test, you must follow one of the procedures above to qualify for a make-up test. The T.A. will ***not*** run a make-up test separately for each individual. There will be one sitting arranged for all qualified students for a make-up test.

Academic Integrity

Students are expected to know and adhere to the University's principles of academic integrity. Any act of plagiarism or other unethical behavior will be addressed in accordance with University guidelines. Students should be aware that turning in an old paper, or large parts thereof, for credit in a second course, is considered an academic offense. Please see the "Code of Behaviour on Academic Matters" (http://www.governingcouncil.utoronto.ca/policies/behaveac.htm) for specific information on academic integrity at the U of T.

Appointments

I encourage you to email or phone me using the number on the first page and arrange an appointment at any time. I am in my office on most days. But, in general, I will be in my office from 4-5 on Tuesday after this class, if you want to stop by.

I will designate special office hours in the weeks before assignments are due – to be announced later.

Re-marking

We will use specific marking keys for both assignments and tests. Those keys define the universe of possible answers and possible variations in those answers. In a course such as this, the only issue that may come up is a mistake in applying the key to the answers in specific cases. If there is a mistake in an assignment or test you get back, you should see the T.A. ***within two weeks of your receipt of the test or assignment.*** In general, we will not consider work for re-grading after feedback on a later test or assignment, unless it is in this two week period.

Note that your grade after re-marking can remain the same, or increase, or decrease.

Class Schedule and Readings

****Note: if necessary, and only if necessary, we will hold tutorials for students on Wednesdays if they cannot attend tutorials on Fridays. All scheduled tutorials noted here may therefore be held on the prior Wednesday, in addition to the schedule Friday.*

Date	*Day*	*Topic / Work*	*Readings*
September 14	Tuesday	**1. Overview** **Review of Descriptive Statistics**	**Class Notes: Basic Statistics Review sections:** **1. Measurement** **2. Sampling** **3. Descriptive Statistics** **Intro Statistics:** **1. Data and Sampling 1.2** **2. Measurement and Tables 1.3, 1.6.** **3. Descriptive Statistics 2.5 to 2.8**
September 21	Tuesday	**2. Probability and Inference**	**Class Notes: Review of Basic Statistics sections:** **4. Introduction to Probability** **5. Probability and Sampling Distributions** **6. More Sampling Distributions** **7. Tests of Hypotheses** **Intro Statistics:** **3. Probability 3.1 to 3.3, 3.5** **5. Continuous Random Variables 5.1 and 5.2** **6. The Normal Distribution 6.1 to 6.6** **7. The Central Limit Theorem 7.1 to 7.5** **9. Hypothesis Testing 9.1 to 9.4, 9.7 to 9.10**
September 28	Tuesday	**3. Cross-Classification: Studying Association in Tables**	**Class Notes: Cross-Classification Sections:** **1. Bivariate Associations** **2. Measures of Association in Tables** **Intro Statistics:** **3.4 Contingency Tables**
October 1	Friday	**Tutorial: Intro to SAS**	
October 5	Tuesday	**4. Multivariate Tables**	**Class Notes: Cross-Classification Section:** **3. Multivariate Tables**

Date	*Day*	*Topic / Work*	*Readings*
			Online: **https://onlinecourses.science.psu.edu/stat504/node/102** **Lesson 5.1, 5.2, 5.3**
October 8	Friday	**Tutorial: SAS Procedures, Class Data Assignment #1**	
October 12	Tuesday	**5. Introduction to Correlation and Regression**	**Wheaton and Young, Chapter 1 : Review of Correlation and Regression:** **1.1 to 1.6** **Quercus:** **Schroeder Understanding Regression Analysis: Chapter 1:. Linear Regression:** **Alan Sykes. An Introduction to Regression Analysis. Pp. 1-7.**
October 19	Tuesday	**6. Multiple Regression**	**Wheaton and Young, Chapter 1 : Review of Correlation and Regression:** **1.7-1.8** **Quercus:** **Schroeder Understanding Regression Analysis: Chapter 2: Multiple Regression** **Alan Sykes. An Introduction to Regression Analysis. Pp. 7-17.**
October 20	Wednesday	**Assignment #1 Due.**	
October 22	Friday	**Tutorial: Test Review**	***Tutorial room, FE36 Sociology***
October 26	Tuesday	**7. Interpretation of results: an Introduction to Models**	**Wheaton and Young, Chapter 6 : From Equations to Models:** **6.1 to 6.6, 6.9.**
October 29	Friday	**Term Test**	***Tutorial room, FE36 Sociology***
November 2	Tuesday	**8. Dummy Variables in Regression**	**Wheaton and Young, Chapter 1 : Dummy Variables sections** **1.9.** **Quercus:** **Schroeder Understanding Regression Analysis: Chapter 4: Dummy Variables**
November 8-12		**Reading Week**	
November 16	Tuesday	**9. Regression Extensions: Interactions I**	**Wheaton and Young, Chapter 2 : Interactions** **2.1. Continuous x Categorical Interactions**

Date	*Day*	*Topic / Work*	*Readings*
			Quercus: **Thomas Brambor et al. 2005. Understanding Interaction Models. Political Analysis 13: Pp. 1-11.**
November 19	Friday	**Tutorial: Assignment 2**	
November 23	Tuesday	**10. Regression Extensions: Interactions II**	**Wheaton and Young, Chapter 2 : Interactions** **2.2 to 2.4, 2.6. Generalizations of Interactions**
November 26	Friday	**Tutorial: Assignment 2**	
November 30	Tuesday	**11. Regression Extensions: Nonlinear Regression**	**Wheaton and Young, Chapter 3 : Nonlinear Regression** **3.1 to 3.4.1 (not 3.4.2), 3.7**
December 3	Friday	**Assignment #2 due;** **Tutorial: Final Exam review**	
December 7	Tuesday	**12. Models for Categorical Outcomes**	**Wheaton and Young, Chapter 4 : Logistic Regression** **4.1 to 4.4, 4.9** **Quercus::** **Chao-Ying Peng et al. 2010. An Introduction to Logistic Regression Analysis. Journal of Educational Research 96: Pp. 1-9.**