SOCIOLOGY 514 – Analyzing Quantitative Data in Sociology
Monday, 10:00am-1:00pm

Professor Elizabeth Hirsh   Office: ANSO 2204
Email: ehirsh@mail.ubc.ca   Office hours: Thurs. 10-11am

___

DESCRIPTION

This course provides an examination of quantitative methods used in sociology and the social sciences, with a focus on linear modeling. We will begin with a review of bivariate regression and then cover the basics of multivariate linear regression, including estimation, interpretation, hypothesis testing, model assumptions, and model fit. We will review several tools for diagnosing violations of statistical assumptions and what to do when things go wrong, including dealing with outliers, collinearity, and weights. Finally, we will explore extensions of the linear regression model, such as logistic regression, generalized linear models, and hierarchical linear models (depending on time and students' interests). While statistical modeling is the focus of the course, we proceed with the assumption that models are only as good as the theoretical and substantive knowledge behind them. Thus in covering the material, we will spend considerable time discussing the link between substantive knowledge and data analysis.

REQUIREMENTS & EVALUATION

Class Meetings. We will meet once a week on Mondays from 10am to 1pm. The first half of class will be dedicated to new material, typically in lecture format. During the second half of class, we will transition to a "lab" format for demonstration of hands-on statistical modeling using software.

Readings. The course material will be drawn from multiple sources. There is one recommended textbook and several supplementary texts/readings that you may find helpful, depending on your previous experience with statistics and your goals for the course. The recommended textbook is:

- Agresti, Alan. *Statistical Methods for the Social Sciences*. Prentice Hall. 2017, 5th edition.

  This book provides an excellent introduction to descriptive statistics, statistical inference, and linear modeling, and assumes little prior knowledge.

There are several other texts on social statistics and regression analysis that you may find helpful. Some of the best include:

- John Fox. *Applied Regression Analysis and Generalized Linear Models*. Sage Publications. 2016, 3rd edition.

- Long, J. Scott. *Regression Models for Categorical and Limited Dependent Variables*. (Sage, 1997).

- Long, J. Scott and Jeremy Freese. *Regression Models for Categorical Dependent Variables Using Stata* (Stata Press, 2014, 3rd edition).

- Hox, Joop. Multilevel Analysis: Techniques and Applications. 3rd edition, 2017. Lawrence Erlbaum: New Jersey.

We will also read academic journal articles that provide examples of the various analytical techniques we cover in class.

Homework. Students will apply the statistical techniques covered in the course in weekly or biweekly homework assignments. These assignments will generally involve analyzing social science data sets using statistical software and discussing results. They will be geared toward application and interpretation rather than formal statistical derivations.

Software. Most of the labs and homework assignments will involve the use of statistical software. While students are free to use their preferred software package to complete homework assignments, I will demonstrate statistical techniques using Stata.
- Stata is available for purchase at student-discounted rates through UBC. For the purposes of this class, I suggest purchasing a Stata/SE license. For more information and purchasing online, see: https://www.stata.com/order/new/edu/gradplans/campus-gradplan/

I assume that students have some prior familiarity with this software package but will provide the Stata instruction necessary to complete homework assignments in class.

Research Project. Students will be required to complete a research project. The project will involve analyzing a dataset utilizing the techniques covered in class and presenting the findings in a paper. The goal of the research project is to practice quantitative data analysis in the context of a theoretically-inspired research project. More details on the research project will be distributed in class.

Grade Breakdown.

Homework assignments	60%
Research project	30%
Participation	10%

COURSE SCHEDULE & TOPICS
(more details posted on course Canvas page)

Part I. Linear Regression

Week 1. Introduction and Review

Week 2: Statistical Inference & t-tests

Week 3. Bivariate Regression

Week 4. Multivariate Regression

Week 5. Categorical Independent Variables and Interaction Effects

Week 6. Nonlinear Relationships

Part II. Model Building, Diagnostics, and Data Problems

Week 7. Modeling Building and Data Management

Week 8: Diagnosing and Dealing with Violations of OLS Assumptions

Week 9: Outliers, Influential Data, and Collinearity

Part III. Advanced Techniques: Extensions to the Linear Model

Week 10: Logistic Regression

Week 11: Generalized Linear Models

Week 12:   Multilevel Models

Week 13. Longitudinal Analysis