School of Social Policy and Practice
University of Pennsylvania
2018-19

NPLD 720: Data Analysis for Social Impact

Class time and locations:

Wednesday 6th February
10am-12pm (Zoom Lecture)
1pm-4pm (STATA computer Lab – Caster Building Computer Lab)

Thursday 7th February
10am-12pm (Zoom Lecture)
1pm-4pm (STATA computer Lab – Caster Building Computer Lab)

Wednesday 13th February
10am-12pm (Zoom Lecture)
1pm-4pm (STATA computer Lab – Caster Building Computer Lab)

Wednesday 20th March
1-4pm (Zoom Lecture)

Thursday 21st March
1-4pm (STATA computer Lab – Caster Building Computer Lab)

Thursday 28th March
10am-12pm (Campus Lecture)
1-4pm (STATA computer Lab – Caster Building Computer Lab)

Wednesday 3rd April (Matt on Campus)
10am-12pm (Campus Lecture)
1-4pm (STATA computer Lab – Caster Building Computer Lab)

Thursday 4th April (Matt on Campus)
10am-12pm (Campus Lecture)
1-4pm (STATA computer Lab – Caster Building Computer Lab)

Course Description

Practitioners, leaders and researchers need to engage with the latest cutting-edge research findings in their field. In this class you will develop an understanding of the quantitative methods that underpin social impact research, in an applied lab-based context. Theoretically, we will focus on developing your working statistical knowledge, and practically we will develop your data analysis skills by introducing you to a range of approaches for analyzing and handling large-scale secondary quantitative data that capture social impact. The substantive focus of the course will be on individual-level participation in the Non-profit Sector in activities such as volunteering and charitable giving.

This applied course covers the fundamental elements and approaches to handling and analyzing quantitative survey data. The emphasis is on developing an adequate understanding of basic theoretical statistical principles, descriptive and exploratory methods of analysis, graphical representation, operational procedures and interpretation of statistical results using STATA. The course will cover a wide range of statistical techniques from basic descriptive statistics to more advanced multivariate statistical techniques, such as OLS regression and logistic regression. You will also be introduced to a number of important topics, including theory testing and development; philosophy of science and research judgement; and replication in social impact research.

This course is an introduction to applied social impact research and is designed for those who want to engage with quantitative social impact research, but also those who wish to make their own original research contributions. No prior statistical knowledge or programming skills are required to enrol in the course.

Learning Outcomes

On completion of this course, students are expected to be able to:
- Demonstrate a sound understanding of the role of quantitative data analysis in social impact research.
- Understand basic statistical theory principles
- Develop important data management skills and basic programming skills (writing syntax)
- Report and present social impact research with close reference to theoretical frameworks.
- Understand the principles of some of the most frequently used statistical modeling methods such as ordinary least square (OLS) regression and logistic regression.
- Understand and interpret the quantitative outputs / findings, statistically and substantively.
- Write-up statistical results and present findings in a journal-acceptable format.

Readings

Any number of general statistics texts can be used for the class and will cover the general material. You may want to purchase the following title (recommended):

1. Agresti A and Finlay B. (2013). Statistical Methods for the Social Sciences. Pearson.

Further reading:
2. Acock AC. (2016). A Gentle Introduction to Stata. Stata Press.
3. Allison PD. (1999) Multiple Regression: A Primer. Pine Forge Press.
4. Long JS. (1997) Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks: Sage.
5. Long JS and Freese J. (2006) Regression Models for Categorical Dependent Variables Using Stata.
6. Wheelen C. (2014) Naked Statistics: Stripping the Dread from Data. Norton.

Other resources:

UCLA academic technology services: http://www.ats.ucla.edu/stat/
Statalist: http://www.statalist.org/

Course Grades

ATTENDANCE AND CLASS PARTICIPATION	20%
ASSIGNMENTS (3)	30%
FINAL PROJECT	50%
TOTAL	**100%**

Grade Scale

The following grading scale will be used for final grades:

98 - 100% =A+	80 - 82.9% = B-
93 - 97.9% =A	77 - 79.9% = C+
90 - 92.9% =A-	73 - 76.9% = C
87- 89.9% = B+	70 - 72.9% = C-
83- 86.9% = B	Below 70% = Failing

Attendance and Class Rules

Students are required to attend all lectures and labs. Laptops and cell phones are not permitted in class.

Policy on Scholastic Dishonesty

Students who violate University rules on scholastic dishonesty are subject to disciplinary penalties, including the possibility of failure in the course and/or dismissal from the University. Since such dishonesty harms the individual, all students, and the integrity of the University, policies on scholastic dishonesty will be strictly enforced. For further information, please visit Penn's Code of Academic Integrity website at http://www.upenn.edu/academicintegrity/

STATA assignments

There are three assignments during the course, accounting for 30% of your grade. These will be distributed at the end of labs 2, 4 and 5. These are to be submitted via Canvas and Turnitin one week later at 1pm.

Purchasing STATA

You do not need to purchase STATA as this software is available on campus and in our labs. However, If you would like to purchase your own copy of STATA you can do so. To do this you will need to go to the STATA website and purchase the IC gradplan version and select an 'annual' licence for $125. This can be done on one of the following websites:
https://www.stata.com/order/new/edu/gradplans/campus-gradplan/

https://www.stata.com/order/new/edu/gradplans/student-pricing/

Final Project

A 4000-word data analysis report that uses multivariate linear and/or logistic regression models in the general area of social impact research.

You should set out your research question(s) and explain how they are addressed using the statistical techniques learned on the course. You are free to choose any secondary data source you like (including those provided as part of the course), but if using your own data please check with me first. You should provide references both for the substantive topic, and for the research methods used. You may want to include a discussion of exploratory approaches before providing details on the regression analysis. In your report you need to demonstrate (i) your understanding of the statistical techniques used; (ii) your ability to apply these techniques in a meaningful way and; (iii) your ability to interpret your results in terms of their statistical significance and in the context of your research question(s).

Your final project is due at **1pm on Monday 29th April**

Late assignments / project

Late assignments/ projects will result in a deduction of one full letter grade for each day that it is late. The instructor may make exceptions in exceptional circumstances. Failing to plan ahead, unfortunately, is a common problem and not an exceptional one.

Class Schedule

Day 1: Introduction, Descriptive Statistics and Sampling
Reading: Agresti and Finlay (2013). CHAPTERS 1 - 3

Day 2: Probability Distributions, Confidence Intervals and Hypothesis Testing
Reading: Agresti and Finlay (2013). CHAPTERS 4-6

Day 3: Correlation and Bivariate Linear Regression
Reading: Agresti and Finlay (2013). CHAPTER 9

Day 4: Multiple Linear Regression
Reading: Agresti and Finlay (2013). CHAPTER 10 & 11

Day 5: Logistic Regression
Reading: Agresti and Finlay (2013). CHAPTER 15

Days 6-7: Statistical extensions and workshops