2013
Department of Government

GV207-5-AU – "Political Analysis: Introduction to OLS"

Module Supervisor:

Dr. Christopher Claassen
Room 5.414
01206 873494
cclaas@essex.ac.uk
Office hours: Tuesday 3-5pm

Deadlines

Seven homework assignments: 9:45am on the day of the class immediately following the assignment of homework (typically a Thursday). There will be no homework due in the first week (week 1) and last week (week 10) of class. Homeworks will be marked within one week of the due date.

Replication Assignment: 9:45am on the last day of class (week 11). This report will be marked within 2 weeks.

Aims and Objectives

By the end of the module, our goals are to have you understand:

- analyse continuous data;
- conduct regression analyses of continuous data;
- conduct and interpret hypothesis tests;
- present data using graphics;
- use the R statistical framework to achieve the above;
- draw inferences from samples to populations;
- design a study to test a causal hypothesis; and
- understand, interpret and critically evaluate existing quantitative social science studies.

Content

This module teaches you how to test a question regarding the social and political world using existing quantitative data. It thus teaches you both how to analyse social science data, and how to draw inferences from that data so as to answer the question at hand. Most quantitative social science takes the form of a causal model, where an explanation is advanced – and tested – for some social phenomenon or behaviour. The module is oriented in this direction, so you will also learn how to – and how not to – test causal questions. Finally, a large part of the module concerns learning the software necessary to conduct data analysis. In particular, you will learn how to run statistical analyses in the R statistical language.

Key Skills

Communication: Writing clearly; interpreting data analysis.

Improving own learning and performance: Responding to feedback on assignments regarding data analysis and software usage.

Information technology: Learning the R statistical language; extracting data from online sources; producing reports with quantitative and graphical output.

Numeracy: understanding and conducting statistical analyses.

Problem solving: building a statistical model; using the R language to implement the statistical model and produce interpretable output.

Critical thinking: Linking a theoretical question to the supporting evidence through an understanding of what the evidence means and how it is produced.

Organisation

The module will comprise one two-hour seminar per week. The class time is Thursday 12-2pm. The module administrator is Sally West (Room 5B.316, telephone ext. 3011, email: sawest). The office is open 10am–1pm and 2–4pm, Monday to Friday and is closed on Wednesday afternoons.

Assessment and Workload

Assessment is by coursework only. There is, in other words, no exam. Marks will be assigned to two forms of coursework: weekly homework and a data analysis report.

Weekly homework (40%). Every week the class will be assigned a short homework assignment to familiarize yourself with R and the statistical methods we are learning. These will be due at 9:45am on the day of the next class.

Replication assignment (60%). Students are required to write a research report analyzing some political or social data of interest using OLS. The report should be around 5 pages, double-spaced. It should be accompanied by tables and figures showing results and diagnostics. These tables and figures are not included in the page length. The report will be due at 9:45am on the last day of class (week 25). More details will be provided in class.

All coursework (assignments and reports) must be submitted using the University FASer server. You can access on-line submission via your myEssex portal or via essex.ac.uk/e-learning/tools/faser/students. No coursework should be emailed to administrative or academic staff. You may find it helpful to look at the "most common asked questions regarding on-line submission" on the following webpage: essex.ac.uk/government/online_resources/troubleshoot.shtm

I encourage all students to hand in coursework in pdf or html format. I discourage handing in Microsoft Word documents.

Textbooks

In the first term, we will be using the following textbook:

> Daniel Kaplan. 2012. *Statistical Modeling A Fresh Approach*. Macalester College, St. Paul, MN, 2nd edition. The first few chapters are available online for free at: www.mosaic-web.org/go/StatisticalModeling/index.html.

A few chapters from the following book will also be used. We will continue to use this book in the next module:

> Andrew Gelman and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Other useful books, which are recommended but not required, are:

> Alan Angresti and Barbara Finlay. 2009. *Statistical Methods for the Social Sciences*, 4th Edition. Upper Saddle River, NJ: Prentice Hall.
>
> Larry Gonick and Woollcott Smith. 1992. *The Cartoon Guide to Statistics*. HarperCollins Publishers.

Software

We will be using the R statistical language (r-project.org). R is available for every computing platform, and most importantly, is free. If desired, you can download and install R on your own computer, or you can use the computers in the lab. I recommend using (and installing) the RStudio front-end for R, which is freely available at rstudio.com/ide/download/.

A useful handbook for R is the free online book by John Verzani (cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf), particularly if you want to learn more than the basics that we will cover in class. The rseek.org website is another good resource for problem-solving in R.

Module Outline

Week	Topic	Reading
W1	Introduction, data analysis, programming in R	Kaplan, chs. 1–3
W2	Models	Kaplan, ch. 6
W3	Formulas	Kaplan, ch. 7
W4	Fitting models	Kaplan, ch. 8
W5	Correlation	Kaplan, ch. 9
W6	Controls	Kaplan, ch. 10
W7	Confidence in models	Kaplan, chs 5 & 12
W8	Hypothesis tests	Kaplan, chs 13–15
W9	Assumptions and diagnostics	Gelman & Hill, chs 3–4
W10	Prediction and simulation	Gelman & Hill, ch. 7 (7.1–7.3)