POLS0008 – 2023-2024
Introduction to Quantitative Research Methods

Lecturers:	Anwar Musah (AM) & Stephen Jivraj (SJ)
Seminar Leads:	Marta Fatica, Mikaella Mavrogeni & Ricardo Mellado Labbe
Teaching:	The module will run lectures and seminars. All lectures will be in-person on Tuesdays from 11:00am to 1:00pm. All seminars will be on Thursdays in 8 blocks each 1-hour from 09:00am to 02:00pm (Note that sessions between 11:00am to 02:00pm are in parallel).
Credits:	15 Credits
Assessment:	One 3,000 word essay (100%)
Essay Deadline:	To be confirmed
Attendance:	Attendance is required for all seminars and will be monitored. Please contact polsci.ug@ucl.ac.uk within 48 hours of an absence. Engagement with online materials will also be monitored.

USEFUL LINKS

UCL Timetable

Extenuating Circumstances
Essay Extensions
Penalties for Late Submission
Penalties for Overlength Essays
Essay Submission Information
Examinations
Plagiarism and TurnItIn
Plagiarism and Academic Writing - a Guide for Students (you will need to log in and enrol yourself on the page)

Content

This module introduces students to quantitative research methods in the social sciences by covering the basics of what is required for those considering a career involving data analytics. It assumes no knowledge of quantitative methods or statistical software. The module caters for students from diverse academic disciplines and adopts a practical hands-on approach to learning, with tutor supported computer tutorials. The module covers descriptive statistics, data visualisation, data access, probability, sampling, hypothesis testing, inferential statistics and ends with an introduction to linear regression. Students will be introduced to the R statistical software and work with data used in current academic research.

By the end of the module, students will be able to:

- identify and understand levels of measurement
- examine and visualise data using descriptive statistics
- use basic commands in R
- understand probability and statistical inference
- conduct basic statistical tests
- run and interpret simple linear regression

Lectures and tutorials
Each week there will be an introductory lecture followed by a computer tutorial. The lectures will introduce students to many of the ideas and issues relating to the various topics. The computer tutorials will provide an opportunity to implement the techniques covered in the lectures.

The link below is to a UCL managed server version of RStudio that will run all functions in the weekly tutorials. You are advised to use this version. It is currently only available on the UCL network, or through VPN or UCL@Desktop Anywhere. It will work on any device you can access a browser (e.g., desktop, laptop, tablet, smartphone, etc).

R Studio Server

Assessment
The module is assessed through the completion of one essay based on the secondary analysis of quantitative data. It accounts for 100% of the total marks on the module. The essay must be a **maximum of 3,000 words**. Please include the word count at the top of the essay and submit your essay using your candidate number as the filename. Please check the Department of Political Science essay submission checklist and penalties for late submission and exceeding word limits.

You will find useful guidance for writing and presenting essays on the Department of Political Science assessments Moodle page. These guidelines are designed to help you, and you should read them carefully and do your best to follow them. Good essays give clear and focused answers to the question asked, they have clear structures, and they will be adequately and appropriately referenced. They do not provide a vague and unstructured discussion of the topic. Plagiarism is taken extremely seriously and can disqualify you from the module (for details of what constitutes plagiarism see http://www.ucl.ac.uk/current-students/guidelines/plagiarism). If you are in doubt about any of this, ask the tutors.

Other non-assessed work
The computer tutorials will allow students to apply and test their knowledge of the material covered on the module and weekly exercises should be submitted for feedback from the module tutors. It is intended that students will complete weekly exercises outside of class.

Reading list

To gain a sufficient understanding of the concepts and techniques taught on this module, students will need to do background reading. No one book covers all the content on the module, and it is worth reading as widely as possible. Note also there are useful online resources, with some examples included below.

The core reading is covered by the following open-source e-book available:

- Çetinkaya-Rundel and Hardin 2021. Introduction to Modern Statistics. https://www.openintro.org/book/ims/

Please note many of the following texts and web-resources go beyond the level required for the module:

- Levin, Jack, James Fox, and David Forde. 2009. *Elementary Statistics in Social Research*, 11th edition. (International ed) Pearson/Allyn and Bacon.
- Agresti, Alan, and Barbara Finlay. 2009. *Statistical Methods for the Social Sciences*. New Jersey: Pearson Education. International edition, 4th (or 3rd) ed.
- Kellstedt, Paul M., and Guy D. Whitten. 2013. *The Fundamentals of Political Science Research*. Cambridge: Cambridge University Press.
- Elliott, J. and Marsh C. (2008) Exploring Data (2nd Edition) Polity Press
- Garner, Roberta. 2010. *The Joy of Stats: A Short Guide to Introductory Statistics in the Social Sciences*. Ontario: University of Toronto Press. (Especially the "Math Refresher" section)
- Salkind, Neil J. 2004. *Statistics for People Who Think They Hate Statistics*. London: Sage.
- Dalgard, Peter. 2008. Introductory Statistics with R (2nd edition).
- Field, A., Miles, J. and Field, Z. (2012) Discovering Statistics Using R. London: Sage.
- Zuur, A.F., Ieno, E.N., & Meesters, E.H.W.G., 2008, A Beginner's Guide to R.

Most of the reading is available in UCL library, although there are generally limited copies. Many items are also held in Senate House library.

Online resources

There are many web-based resources for the study of quantitative methods and the R statistical software. Students may find useful those listed below.

Online interactive e-books covering statistical principles in the first half of the module
http://onlinestatbook.com/2/index.html
http://www.openintro.org/stat/

Statistical glossary - includes some good simple explanations of basic concepts used in module
http://www.stats.gla.ac.uk/steps/glossary/index.html

Statsoft electronic statistics textbook covering many of the techniques covered in the module
http://www.statsoft.com/textbook/stathome.html

R resources
UCLA Statistical Consulting Group introduction to R
http://www.ats.ucla.edu/stat/r/seminars/intro.htm

Neat websites with basic data analysis commands described
http://www.statmethods.net/index.html
http://www.cookbook-r.com

The R Guide to UK Data Service key UK Surveys
http://ukdataservice.ac.uk/media/398726/usingr.pdf

Producing simple graphics with R
http://www.harding.edu/fmccown/r/

Week 1 (AM)

Lecture: Understanding data
After providing an overview of the aims, learning objectives and practical arrangements for the module, the lecture moves on to take a brief look at the notion of variables and how they are measured, frequency distributions and ways of describing the central tendency and the dispersion of a variable.

Tutorial: Introduction to R
The first week provides an introduction to the statistical software R, which will be used in all subsequent tutorials. Students will learn how to download and install R, enter simple data into R, how to load data sets from other statistical packages into R and how to use the R environment for simple calculations.

Essential reading
Introduction to Modern Statistics Chapter 1

Week 2 (AM)

Lecture: Examining data 1
This session explores a survey dataset to identify and describe different types of variables using R. Various techniques for looking at variable distributions, including table and basic univariate graphs

Tutorial: Describing data I
The practical introduces univariate (one variable) descriptive statistics and graphs in R, including frequency tables, histograms and boxplots.

Essential reading
Introduction to Modern Statistics Chapter 4

Week 3 (AM)

Lecture: Examining data 2
This session explores a survey dataset to identify and describe different types of variables using R. Various techniques for looking at variable distributions, including tables, graphs and summary statistics are considered. We also address access to data and data management such as how to recode a variable.

Tutorial: Describing data II
This practical explores descriptive bivariate statistics and graphs using R, including two-way tables, bar charts, line graphs and scatterplots.

Essential reading
Introduction to Modern Statistics Chapter 5

Week 4 (AM)

Lecture: Sourcing Data
This session will be an opportunity to recap some of the core principles introduced in the previous 3 weeks before introducing you to sourcing of data from open-source websites. We will discuss the benefits and limitations of using secondary data from open-source platforms compared to primary data and discuss its ease for access as well as usefulness for research.

Tutorial: Analysing open-source data in R
This practical will introduce you to sourcing and preparing data for analysis from official sources with a strong focus on bringing together the techniques learnt over the past 3 weeks.

Essential reading
Introduction to Modern Statistics (see previous chapters i.e., 1, 4 & 5)

Week 5 (SJ)

Lecture: Probability and the normal distribution
The basics of probability theory are introduced as well as concepts of probability distributions, the normal curve, and sampling distributions of means.

Tutorial: Sampling distribution of the mean

This session introduces sampling distributions using a simple simulated data set in R.

Essential reading
Introduction to Modern Statistics Chapter 13

Week 6 (SJ)

Lecture: Confidence intervals and significance
The session explores the ability to generalise the findings from analysis of sample data to the wider population (inference). The theory of hypothesis testing, confidence intervals and statistical significance are introduced.

Tutorial: Testing for sampling error
This practical follows on from the previous week using the simulated data to derive confidence intervals.

Essential reading
Introduction to Modern Statistics Chapters 11-12, 14

Week 7 (SJ)

Lecture: Measures of difference (SJ)
This lecture considers statistical hypothesis tests for the difference in two means from paired and independent samples and the difference in two categorical variables.

Tutorial: Tests of significance
The session explores the use of appropriate use of test of significance for a mean.

Essential reading
Introduction to Modern Statistics Chapters 19-21

Week 8 (SJ)
Lecture: Exploring relationships between interval variables
This lecture starts with a look at graphical approaches of exploring the relationship between interval variables using scatterplots introduced in lecture 2 before moving on to look at measures of correlation and associated statistical tests and how to interpret them.

Tutorial: Exploring relationship between categorical variables
We cover how to test the statistical significance of two-way samples using a simple Chi Square test (use of the Cramers V test is also included as a measure of the strength of association between two variables).

Essential reading
Introduction to Modern Statistics Chapters 16-18

Week 9 (SJ)
Lecture: The simple linear regression model
This lecture introduces the theory and practice of simple linear regression and how to interpret the output. The components of a simple linear regression model are described and explained before looking at how to interpret the output and checking the assumptions of the method are met.

Tutorial: Correlation and linear regression
The session first covers how to run and interpret of two statistical tests of correlation (Pearson's r and Spearman's rho) and then how to design, run, and interpretation of a simple linear model.

Essential reading
Introduction to Modern Statistics Chapters 7-8

Week 10 (SJ)

Lecture: Assumptions of linear regression
The purpose of this lecture is to explore the basic assumptions underlying the multiple linear regression model such as collinearity, outliers/leverage and correlated residuals.

Tutorial: Testing the assumptions of multiple linear regression
This tutorial enables students to assess their linear regression models by testing for the basic assumptions underlying this statistical method.

Essential reading
Introduction to Modern Statistics Chapters 24-25