



## Table of contents

1	Module Overview . . . . .	1
1.1	Weekly timetable . . . . .	2
1.2	Office Hours . . . . .	2
1.3	Releases . . . . .	2
1.4	Schedule . . . . .	2
1.5	Feedback . . . . .	3
1.6	Computer and Software . . . . .	3
1.7	Data . . . . .	3
1.8	Bibliography and other Resources . . . . .	5
1.9	Accessibility . . . . .	5
2	Textbooks . . . . .	5
3	Other Resources . . . . .	7
4	Assessment . . . . .	7
5	Learning Outcomes . . . . .	8

This document is still a draft version and subject to change.

# 1 | Module Overview

This module has two principal aims:

1. Introduce you to secondary data acquisition, management and analysis in the social sciences
2. Prepare you to attend further statistical training and make use of statistics in future research projects, be they academic (such as master's or PhD dissertations) or not.

The module consists of nine weeks of teaching. Seminars are designed as discussions on theory and concepts, as well as 'hands-on' computer workshops, giving you direct experience of exploring and analysing data in the R environment. R is a free and open software that you easily download and install on your computer. The module can help you work with different kinds of data, such as individual surveys, country-level aggregate data, corporate statistics, administrative records and data collected online. However, for practical reasons, in the module, we will focus on a selection of individual-level survey datasets, relevant to the social sciences at large. Skills acquired with such surveys may easily be transferred to other data.

This module does not require any prior knowledge of mathematics or statistics. You only need to understand the importance of statistics for empirical social sciences, and show willingness to learn about them. The focus will be less on mathematical equations than on the selection and application of a range of methods and tools to concrete scientific social data. You will be invited to use skills you have been developing in other domains, such as logical and causal reasoning, research design, reading and writing on this module. Statistical reasoning and computer languages are best acquired on the basis of a good command of natural language (in this case, English).

If you already have some knowledge or practice of statistics, please be aware that this module is designed for a range of levels, including beginners. If you already have experience with statistics, you may skim or skip the parts that you are confident with, and spend more time on more advanced readings and weekly "Going further" exercises. Meanwhile, if you are a beginner, you may rather focus on understanding the lecture, reading on the aspect of the module where you lack background (either statistics, coding, or social scientific reasoning), and doing the "Core" exercises.

To succeed on this module, it is not enough to attend, listen and review the exercises' solutions. You need to engage actively with all the module content: follow lectures and seminars, complete the worksheets, do relevant readings, and take quizzes.

## 1.1 Weekly timetable

Weeks 1-5 and 7-10 for lecture, seminar, plus optional individual meetings at the module director's office hours. Week 6, a.k.a. "reading week", is not taught.

### Lecture

All the lectures will take place on Monday, 10:00-11:00, in [S0.19](#).

### Seminars

- Group 1: Monday, 11:00-13:00, [S0.09](#)
- Group 2: Tuesday, 12:00-14:00, [OC1.02](#)

Please stick to the group you have been registered. If you wish to switch, please write to [paispg@warwick.ac.uk](mailto:paispg@warwick.ac.uk).

## 1.2 Office Hours

Office Hours are for questions and issues that have a personal dimension, such as illness, learning difficulty and reasonable adjustments, and any kind of question about the module's content that are not answered in class or in the module material.

My office hours take place Tuesdays, 9:45-11:45 in [D1.08](#). Please go to [my webpage](#) to book an appointment.

## 1.3 Releases

- Slides will be released on Moodle on Friday of the preceding week at 12 noon (latest).
- The material for all seminars is available on the modules [Seminar Companion](#) which is hosted on my GitHub page. All exercises and solutions are available without time restrictions, but please do take note of the disclaimer at the top of the [Downloads page](#).

## 1.4 Schedule

Students are expected to attend all the sessions, as they build up from one to the next in a cumulative way. The material provided online enables you to catch up on your own if you miss a session.

Week 1	Research Design & Statistical Concepts
Week 2	Descriptive Statistics and Graphs
Week 3	Foundations of Statistical Inference
Week 4	Confidence Intervals and Significance Tests
Week 5	Introduction to Causality
Week 6	Reading Week
Week 7	Bivariate Methods
Week 8	Bivariate Regression
Week 9	Multivariate Regression I
Week 10	Multivariate Regression II

## 1.5 Feedback

Detailed feedback will be provided throughout the module seminars, especially collective feedback on weekly exercises, and individual written (plus possibly verbal) feedback on the final examination.

You are encouraged to post any questions about the content of the module on the Moodle forum: experience shows that there is great benefit for all students to share. Send an email to me for short questions, or book an in-person appointment during my weekly office hours for longer questions.

## 1.6 Computer and Software

We will be using R and R-Studio on this module. It is beneficial to everybody to learn R, whatever your previous experience with alternatives, and whatever your professional prospects.

- These programs are free to use for all Warwick students and easy to install: please allow between 5 and 20 minutes (depending on your experience) to install them on your own. Please do so as soon as you can, and no later than week 1. First, install the latest version of R from [www.r-project.org](http://www.r-project.org), then the latest (free) version of RStudio Desktop on [rstudio.com](http://rstudio.com). You can obtain MS Excel from the university's [IT Services](#). We will use this to create csv files.
- Seminars will take place in regular seminar rooms and not in computer labs. It is therefore imperative that you bring a functioning, fully charged laptop with you to the seminar. As some buildings date back to the 1960s in the university, you cannot rely on the availability of a plug socket.
- If you need help with IT, such as with your Windows/Macintosh operating system, software installation or connection issues, please visit the University Help Desk. I am unfortunately unable to handle these issues, beyond tailoring and using of the aforementioned module-specific software.
- The University network (such as OneDrive) can be used to store your PO91Q files and transfer easily between computers. This should be used instead of USB drives or email attachment. Please note that USB ports have been disabled throughout the university due to GDPR concerns.
- I strongly recommend you back up your work online in case your computer fails. You have free access to OneDrive during your time at Warwick. If you want to make independent arrangements, I can recommend Dropbox after decades of faithful and reliable service.

## 1.7 Data

As the focus of the module is on data analysis, and not data collection, we will use secondary data. This year exercises will make use of the six data sets below. Data sets are available for free for academic purpose, for example from the UK Data Service website.

- **The British Social Attitudes** (BSA 36, fieldwork 2018) asks over 3,000 people every year “what it’s like to live in Britain and how they think Britain is run. Since 1983 the survey has been tracking people’s changing social, political and moral attitudes. It informs the development of public policy and is an important barometer of public attitudes used by opinion leaders and social commentators.” (British Election Study, 2020)

- **The European Social Survey** (ESS round 9, 2018) is “a large-scale, cross-national, and longitudinal survey research program on basic human values. It provides insights into the ideas, beliefs, preferences, attitudes, values and opinions of citizens all over Europe. It is a unique research project on how Europeans think about life, family, work, religion, politics and society”. (European Social Survey, 2019)
- **The World Values Survey** (WVS wave 6, 2010-12), “a global research project, explores people’s values and beliefs, how they change over time and what social and political impact they have. Since 1981 a worldwide network of social scientists have conducted representative national surveys as part of WVS in almost 100 countries. The WVS measures, monitors and analyses: support for democracy, tolerance of foreigners and ethnic minorities, support for gender equality, the role of religion and changing levels of religiosity, the impact of globalization, attitudes toward the environment, work, family, politics, national identity, culture, diversity, insecurity, and subjective well-being.” Wave 7 is in progress; you may use these data, but make sure the countries you need are available. (World Values Survey Association, 2025)
- **London Ward Atlas:** The ward profiles and ward atlas provide a range of demographic and related data for each ward in Greater London. They are designed to provide an overview of the population in these small areas by presenting a range of data on the population, diversity, households, life expectancy, housing, crime, benefits, land use, deprivation, and employment. (London Data Store, 2013)
- **World Development Indicators:** The World Development Indicators is a compilation of relevant, high-quality, and internationally comparable statistics about global development and the fight against poverty. The database contains 1,400 time series indicators for 217 economies and more than 40 country groups, with data for many indicators going back more than 50 years. (World Bank, n.d.)
- **Polity V:** The Polity5 dataset covers all major, independent states in the global system over the period 1800-2018 (i.e., states with a total population of 500,000 or more in the most recent year; currently 167 countries with Polity5 refinements completed for about half those countries). (Marshall & Gurr, 2020)
- **A Complete Dataset of Political Regimes, 1800–2020** provides a comprehensive cross-national dataset classifying regimes as democracies or dictatorships for all independent states from 1800 to 2020. It uses transparent coding rules based on contestation and participation to produce a binary coding of democracy and autocracy. The dataset enables long-run comparative analyses of democratization, regime change, and political development across nearly two centuries. (Miller et al., 2022)
- **Crime Survey for England and Wales, 2013-2014:** Unrestricted Access Teaching Dataset. The Crime Survey for England and Wales is an important monitor of the extent of crime in England and Wales. It is used by the Government to evaluate and develop crime reduction policies as well as providing vital information about the changing levels of crime over the last 30 years. (University of Manchester, Cathie Marsh Institute for Social Research (CMIST), UK Data Service, Office for National Statistics, 2019)

Each of these datasets is long (number of units) and wide (number of variables) enough to form the basis of in-depth analyses on many social and political topics. They differ with respect to their geographical scope and their thematic focus. Note that you may use two or more waves of the same survey for studying historical change or evolutions. Class exercises are limited to the last one or two waves, but older waves and cumulative datasets are available.

## 1.8 Bibliography and other Resources

- Detailed weekly references, with pages and links towards resources, are available on Talis (the university's electronic reading list repository) through Moodle. Most references are available at the University Library, either electronically or on paper.
- Lecture slides are fully referenced, and a list of references is provided at the end of each week's slide collection.
- The seminar companion is equally fully referenced, with a [full list of references](#) at the end.
- You can also download the module's [complete bibliography](#).
- All slides contain an interactive glossary that will help you learn the core concepts of the module. The full module glossary is available in the [Online Companion](#).
- Flashcards are provided for each week of the module. These contain the functions we encounter in R. They are integrated in the online companion, see [Week 1](#) for an example.

## 1.9 Accessibility

Considerable effort has been made to make all module materials as accessible as possible:

- Lecture Slides and documents
  - Have a grey background
  - Use the font [FS Me](#) for plain text
  - Use the font [Gintronic](#) for R code
  - Use the font [Fira Math](#) for equations
  - Use an accessible colour palette with high contrasts
- Seminar Companion
  - Uses the font [Lexend](#)
  - Uses the font [Recursive Mono](#) for R code
  - Uses an accessible colour palette with high contrasts

However, if you have any concerns about your ability to access the module content due to any kind of learning difference, disability, or other practical issues, please contact me and will do my best to accommodate. I am aiming to deliver material that is convenient for all participants to use, without exception.

## 2 | Textbooks

Textbooks are not only a matter of content, but also of fit between the writer's purpose and style, on one side, and the reader's background, intellectual profile and expectations on the other side. Therefore, you are encouraged to look at several of the textbooks below, starting with the table of contents and a sample chapter, before deciding which one will accompany you for the whole term or year. Also, some students may spend a lot of time in textbooks because they feel the need to read about statistics, while others might feel they make more progress by practising exercises and referring mainly to the slides for conceptual and theoretical matters. A balanced mix is probably a good recipe for progress, but this is just my experience.

Textbooks do not all cover the same topics, in the same order, with the same importance. A fortiori, they do not all match the progress of the module. Some chapters I recommend in the reading list will match a class, others will cover more, or less, so you may have to find your way within a given chapter and identify what corresponds to what. But this is part of any academic reading.

If you enjoy working with a given textbook and you cannot get access to it in a format suitable to your needs, it is worth purchasing it. It may be important to “personalise” a book by writing on it, highlighting or annotating it. Chapters and pages I refer to will usually relate to the last edition, but you may use a previous edition. Unfortunately, publishers tend not to publish electronic versions of popular textbooks such as Stinerock. In these cases, the Library purchases more paper versions, but you may want to consider buying one of these for yourself, possibly second-hand.

Let me first cite some statistics and R references:

- **Field et al. (2012):** Good for those with absolutely no background in mathematics or statistics, with very detailed explanations and many illustrations
- **Fogarty (2023):** Best for beginners and quasi-beginners
- **Fox and Weisberg (2019):** Covers many aspects of regression analysis that cannot be covered in the module
- **Imai (2017):** A short and gentle introduction
- **Sheather (2009):** The most detailed on regression (weeks 8-10), with clearly presented equations. May be used without understanding all the equations
- **Stinerock (2022):** Very clear, with excellent recaps at the end of each chapter, good for beginners and intermediate-level student, more detailed than Field et al. and Fogarty
- **Teetor (2011):** Really for beginning with R
- **Long and Teetor (2019):** Goes further than the previous one, but remains very clear and user-friendly

The following textbooks do not deal, or not mainly, with R, but provide more detail on the statistical side:

- **Agresti (2018):** Some developments are well detailed, although sometimes too much technical detail is provided
- **Balnaves and Caputi (2001)**
- **Bartholomew et al. (2011):** Covers most families of methods, to be read by chapters or sections, yet with very good explanations and only the minimum of equations
- **de Vaus (2002):** A problem-based approach, with one concrete question per chapter
- **Fielding and Gilbert (2006)**
- **Gill (2006):** A mathematics refresher, especially useful if you have not touched maths since high school
- **Kellstedt and Whitten (2018):** With a disciplinary focus, but students in other disciplines may still find it very helpful
- **Kranzler and Anthony (2022):** I guess the title says it all
- **Marsh and Elliott (2008):** Quite basic, but it is important to have the basics right, before moving to more advanced readings
- **Sapsford and Jupp (2006)**
- **Tarling (2009)**



The following offer an introduction to social science methods beyond statistics. In particular if you do not have a background in social sciences, select relevant chapters in:

- **Burton (2000)**
- **Clark et al. (2021)**: A best-seller, comprehensive and with wise evaluations of the value of the different methods
- **Gilbert and Stoneman (2015)**
- **Pierce (2008)**

### 3 | Other Resources

In addition to textbooks, you are encouraged to read weekly one article with quantitative analyses in a journal of your discipline. Try and focus, more than the substance, on the data, methods and tools at play, and how they correspond to the module's weekly topics. As a matter of routine you could consult the most recent issues of a number of journals as they come into the library and establish for yourselves whether they contain pertinent articles. Search for applications of personal interest, as long as they include similar statistical developments as treated the corresponding week. Some journals tend to reduce quantitative analyses to a race to sophistication, which is sometimes sterile. I rather recommend journals that take a more open-minded approach to methods and publish articles that subordinate methods to research questions.

You could also consult on a regular basis one or more "reference" newspapers (e.g. Financial Times, Independent, Guardian, The Economist), blogs or other websites, where you will find considerable reference to quantitative studies. They provide plenty of examples of good and bad investigation.

If you are completely new to R and want a little more information on using the software itself, then you could consult my [Analysing Quantitative Data with R](#) webpage. You might also consider keeping a close eye on the [National Centre for Research Methods](#) materials, which provide introductions to more advanced topics.

### 4 | Assessment

Unlike all other MA modules in PAIS, this module has an in-person exam in January. It will take place on campus at the beginning of Term 2. You will have to be physically present for this occasion, there are no exceptions. In the past, this exam has taken place on the very first day (Monday) of Term 2, so please make sure that you travel back to campus in time after the Christmas break.

The length of the exam is 3 hours. It consists of a compulsory section in which the questions relate to a case study in which I present some R output using real-world data. You then need to choose one of two or three optional sections. In the past, these have contained regression calculations by hand, an evaluation of true/false statements, and mini essays.

It will be clearly indicated on the exam paper how many points each (sub-)question is worth. You can achieve a maximum of 180 points in the exam. With 180 minutes of time available in the exam, the number of points thus translates directly to the number of minutes you have available for each question. The points will be converted to marks between 0 and 100 in the marking process.



All questions will draw on the content of the lectures, the seminars, and the items on the reading list marked “essential” (in other words, not the “recommended” items). It is, therefore, crucial for you to attend all lectures and seminars, and to stay on top of the reading throughout the module.

You will receive a formula collection with the exam paper, so there is no need to memorise these. You can use the weekly (non-assessed) quiz on Moodle to practise. There will be a mock exam as homework for the seminar in Week 10.

## 5 | Learning Outcomes

For the learning outcomes of this module, please refer to Table 1 on page 9.

(By the end of the module the student should be able to...)	Which teaching and learning methods enable students to achieve this learning outcome?	Which summative assessment method(s) will measure the achievement of this learning outcome?
<p>Subject Knowledge and Understanding</p> <ul style="list-style-type: none"> <li>• Acknowledge the strengths and weaknesses of quantitative methods</li> <li>• Be aware of and manipulate the fundamental notions of data management and statistics</li> <li>• Be aware of the role of statistical software and the best way to deal with them.</li> </ul>	<ul style="list-style-type: none"> <li>• Lectures and in class demonstrations (every week)</li> <li>• Core readings and discussions</li> <li>• Optional readings from list provided</li> <li>• Weekly formative tests</li> </ul>	<ul style="list-style-type: none"> <li>• Final Examination</li> </ul>
<p>Cognitive Skills</p> <ul style="list-style-type: none"> <li>• Have the ability to read quantitative studies, extract concepts and discuss the results</li> <li>• Be able to connect a research question with appropriate data, tools and research design</li> <li>• Assess and cite published literature</li> </ul>	<ul style="list-style-type: none"> <li>• Lectures and in class demonstrations (every week)</li> <li>• Home exercises with in-class correction</li> <li>• Compulsory readings and discussions</li> <li>• Optional readings from list provided</li> <li>• Weekly formative tests</li> </ul>	<ul style="list-style-type: none"> <li>• Final Examination</li> </ul>
<p>Key Skills; Subject-Specific/Professional Skills</p> <ul style="list-style-type: none"> <li>• Know how to plan a basic research project and convince about its relevance</li> <li>• Conduct quantitative research autonomously</li> <li>• Present results in writing in a rigorous manner.</li> </ul>	<ul style="list-style-type: none"> <li>• Home exercises with in-class correction</li> <li>• Optional readings from list provided</li> <li>• Weekly formative tests</li> </ul>	<ul style="list-style-type: none"> <li>• Final Examination</li> </ul>

Table 1: Learning Outcomes for PO91Q

## List of References

- Agresti, A. (2018). *Statistical Methods for the Social Sciences* (Fifth Edition). Harlow: Pearson.
- Balnaves, M., & Caputi, P. (2001). *Introduction to Quantitative Research Methods. An investigative Approach*. Thousand Oaks, CA: Sage.
- Bartholomew, D. J., Steele, F., Galbraith, J., & Moustaki, I. (2011). *Analysis of Multivariate Social Science Data* (Second Edition). London: Routledge.
- British Election Study. (2020). British Election Study. <https://www.britishelectionstudy.com/>
- Burton, D. (2000). *Research Training for Social Scientists: A Handbook for Postgraduate Researchers*. Thousand Oaks, CA: Sage.
- Clark, T., Foster, L., Sloan, L., & Bryman, A. (2021). *Bryman's Social Research Methods* (Sixth Edition). Oxford: Oxford University Press.
- de Vaus, D. (2002). *Analysing Social Science Data - 50 Key Problems in Data Analysis*. Thousand Oaks, CA: Sage.
- European Social Survey. (2019). Data. Round 9. <https://www.europeansocialsurvey.org/data>
- Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. Thousand Oaks, CA: Sage.
- Fielding, J., & Gilbert, N. (2006). *Understanding Social Statistics* (Second Edition). Thousand Oaks, CA: Sage.
- Fogarty, B. J. (2023). *Quantitative Social Science Data With R – An Introduction* (Second Edition). Thousand Oaks, CA: Sage.
- Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (Third Edition). Thousand Oaks, CA: Sage.
- Gilbert, N., & Stoneman, P. (Eds.). (2015). *Researching Social Life* (Fourth Edition). Thousand Oaks, CA: Sage.
- Gill, J. (2006). *Essential Mathematics for Political and Social Research*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/cbo9780511606656>
- Imai, K. (2017). *Quantitative Social Science: An Introduction*. Princeton: Princeton University Press.
- Kellstedt, P. M., & Whitten, G. D. (2018). *The Fundamentals of Political Science Research* (Third Edition). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108131704>
- Kranzler, J., & Anthony, C. J. (2022). *Statistics for the Terrified* (Seventh Edition). Lanham: Rowman; Littlefield.
- London Data Store. (2013). Ward Profiles and Atlas. <https://data.london.gov.uk/dataset/ward-profiles-and-atlas>
- Long, J., & Teeter, P. (2019). *R cookbook: proven recipes for data analysis, statistics, and graphics*. O'Reilly Media. <https://rc2e.com/>
- Marsh, C., & Elliott, J. (2008). *Exploring Data: An Introduction to Data Analysis for Social Scientists* (Second Edition). Cambridge: Polity Press.
- Marshall, M. G., & Gurr, T. R. (2020). Polity V Project: Political Regime Characteristics and Transitions, 1800-2018. <http://www.systemicpeace.org/inscrdata.html>

- Miller, M., Boix, C., & Rosato, S. (2022). *Boix-Miller-Rosato Dichotomous Coding of Democracy, 1800-2020*. <https://doi.org/10.7910/DVN/FENWWR>
- Pierce, R. (2008). *Research Methods in Politics*. Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9780857024589>
- Sapsford, R., & Jupp, V. (2006). *Statistics with R – A Beginner’s Guide* (Second Edition). Thousand Oaks, CA: Sage.
- Sheather, S. J. (2009). *A Modern Approach to Regression with R*. Berlin: Springer. <https://doi.org/10.1007/978-0-387-09608-7>
- Stinerock, R. (2022). *Statistics with R – A Beginner’s Guide* (Second Edition). Thousand Oaks, CA: Sage.
- Tarling, R. (2009). *Statistical Modelling for Social Researchers: Principles and Practice*. London: Routledge.
- Teetor, P. (2011). *25 Recipes for Getting Started with R*. O’Reilly Media.
- University of Manchester, Cathie Marsh Institute for Social Research (CMIST), UK Data Service, Office for National Statistics. (2019). Crime Survey for England and Wales, 2013-2014: Unrestricted Access Teaching Dataset. <https://doi.org/10.5255/UKDA-SN-8011-1>
- World Bank. (n.d.). World Development Indicators. <https://datacatalog.worldbank.org/dataset/world-development-indicators>
- World Values Survey Association. (2025). *World values survey*. <https://www.worldvaluessurvey.org/wvs.jsp>