



Preparatory Course on Data Analysis

Instructor: Dr. [Annabelle Wittels](#)

Register [online](#) by September 9, 2020
Please notify [Claudia Fanti](#) of your online enrolment.

Aim of this course

This short-course will provide you with an introduction to the most important concepts in statistics, to help you move on to more complex topics in October and to, more generally, equip you with the tools necessary to understand, use and question the use of simple statistics.

Most analysis for social science research today is conducted with the help of two programmes: R and Stata. You will get the chance as part of this short-course to learn the basics of one of these analysis programmes.

Structure of the course

There are six sessions. The first four will focus on grounding you in the basics of statistics. The last two will be focussed on data analysis in Stata and R. In the pre-course survey, you will be asked to choose to attend *either* the session on R or on Stata. All sessions will be conducted via the video-calling app Zoom.

No.	Date	Time	Topic	Assignment
1	Mo, Sept 14	13.00-16.00 CEST	Intro to What, Why and When Statistics	Take-Home exercise
2	Tue, Sept 15		Describing & Summarising Data	Take-Home exercise
3	Thurs, Sept 17		From Populations and Samples to Comparing Distributions	Take-Home exercise
4	Fr, Sept 18		Significance & Analysing Data	Coding task
5	Mon, Sept 21		Introduction to Stata	-
6	Tue, Sept 22		Introduction to R	-

What to prepare *ahead* of the course start

There are two things, you need to ensure before starting the course (Monday, September 14th). This will ensure that we all are off to a smooth start of the course.

1. Please install one of the two programmes *and check that they are working*. More detailed instructions on how to go about this, are provided below.
2. Answer the short survey circulated by Claudia Fanti. The survey will include a question on whether you want to complete the introduction to Stata or to R.

Course materials and readings

The course follows a structure borrowed loosely from Rowntree, D. (2018). *Statistics Without Tears: An Introduction for Non-Mathematics*. Penguin Books, Random House, UK.

Relevant sections and chapters are available as scanned PDFs on Sharepoint, the course platform (similar to Moodle). Alternatively, you can obtain a copy of the book from Amazon for about €10.

Readings on SharePoint are organised by the day on which you will be assigned the reading. If the folder says, for instance, Monday September 14th, you are supposed to read the materials on Monday, September 14th or at least before the next session on Tuesday, September 15th.

Data and software

The course will use the following types of software and platforms:

1. [Zoom](#)
2. [R](#)
3. [Stata](#)
4. [Slack](#)
5. [SharePoint](#) (access granted after course registration)

Please ensure that your programme of preference is installed and working ahead of the course start date. Guides on where and how to install each of these software packages is available via the linked URLs.

A note on installing Stata:

Different to R, Stata is a pay-for software. You will have to get in touch with the university ICT services, in order to install the software and obtain a license for your personal device.

Contact details for the relevant EUI ICT services are provided below.

Online Helpdesk:

<https://helpdesk.eui.eu/CustomerPortal/Account/Login?ReturnUrl=%2fCustomerPortal>

Email: portable.support@eui.eu

Phone: ext. 2030 (+39-055-4685-030) or 2020 (+39-055-4685-020)

Assignments and assessment

Each session will end with an explanation of a short take-home exercise that you will need to complete before the next session. A short description of the assignment will also be uploaded on SharePoint. You will be asked to submit your completed assignments via the course page on Sharepoint.

The Slack group will be there for you to discuss assignments, help each other with tasks and resources.

Sessions 2-4 will start with a brief recapitulation of the main points of the last session and offer time to flag any questions you had relating the take-home exercises.

Detailed schedule

#	Date	Time	Topic / Task
1	Mon, Sept 14	13.00-13.30	Introductions
		13.30-14.00	Course overview & questions
		14.00-14.15	Coffee break
		14.15-15.45	What and why statistics
		15.45-16.00	Explain take-home exercise & time for questions
2	Tue, Sept 15	13.00-13.30	Recap of the last session & discuss take-home task
		13.30-14.00	Describing a sample & Summarising Data Part I
		14.00-14.15	Coffee break
		14.15-15.45	Summarising Data Part II
		15.45-16.00	Explain take-home exercise & time for questions
3	Thurs Sept 17	13.00-13.30	Recap of the last session & discuss take-home task
		13.30-14.00	Populations and samples & Comparing distributions Part I
		14.00-14.15	Coffee break
		14.15-15.45	Comparing distributions Part II
		15.45-16.00	Explain take-home exercise & time for questions
4	Fri, Sept 18	13.00-13.30	Recap of the last session & discuss take-home task
		13.30-14.00	Significance & Data analysis Part I
		14.00-14.15	Coffee break
		14.15-15.45	Data analysis Part II
		15.45-16.00	Explain data exercises for R & Stata

5	Mon, Sept 21st	13.00-13.30	<i>Stata</i> — Go through take home exercise together
		13.30-14.00	<i>Stata</i> — Building blocks: variables and summary statistics
		14.00-14.15	Coffee break
		14.15-15.45	<i>Stata</i> — Simple data visualisation
		15.45-16.00	<i>Stata</i> — Time for questions
6	Tue, Sept 22nd	13.00-13.30	<i>R</i> — Go through take home exercise together
		13.30-14.00	<i>R</i> — Building blocks: variables and summary statistics
		14.00-14.15	Coffee break
		14.15-15.45	<i>R</i> — Simple data visualisation
		15.45-16.00	<i>R</i> — Time for questions

