



```
course = "Reproducible data processing and visualisation in R"
```

```
lesson_iteration = 1
```

```
lesson_title = "Installation, setup and fundamentals"
```

```
auth = "Ian Hussey"
```

```
dept = "Psychology of Digitalisation"
```

**[beginner] or [intermediate] R class
but not both**



Introduction

- 1 Installation and setup
- 2 Fundamentals
- 3 Reproducible reports
- 4 Loading, viewing, and saving data
- 5 The pipe and renaming
- 6 Data transformation I
- 7 Data transformation II
- 8 Working with strings and factors
- 9 Reshaping and pivots
- 10 Binding and joining
- 11 Structuring projects
- 12 Visualization
- 13 Reporting
- 14 Data privacy and anonymity
- 15 License and citation

Reproducible Data Processing and Visualization

</> Code

in R and tidyverse

AUTHOR

Ian Hussey

PUBLISHED

August 27, 2025

Table of contents

- [Introduction](#)
- [How to use this book](#)
- [Other learning resources](#)
- [Find me elsewhere](#)
- [Contributing](#)

Introduction

This [Open Source](#) eBook provides materials for the semester-long Master's seminar course "Reproducible Data Processing and Visualization in R" that I deliver at the University of Bern's Institute of Psychology.

How to use this book

This book is made up of individual Quarto (.qmd) workbooks. Many of the exercises are easiest to complete in your own local copy of these .qmd files.

I suggest that you [download the .zip file](#) of all the .qmd files and their supporting data and cheatsheets and use the eBook as a reference book. *If the .zip file does not contain the most recent versions, please contact me and I'll update it.* If I'm slow or you're feeling impatient, you can also [download the most recent versions from GitHub](#), along with the other files that create this eBook.

You can also copy and paste the code for any chapter directly from the website. Click the "</> Code" button on the top right of each page to see the full .qmd file's code. You can copy and paste this into a .qmd file. However, it's probably easier to download all the .qmd files and data as mentioned above.

Learning to code is a practice skill. Almost anyone can become competent in writing reproducible code for data processing and visualization with practice. More than anything else, completing this course requires that you practice in your own time, using not only the examples provided but also ones you create yourself. Take real data sets from your own studies, or the thousands available on the Open Science Framework ([osf.io](#)), or create simulated messy datasets with my R package [\[truffle\]](#) and practice.



Reproducible
Data Processing
and Visualization

in R and tidyverse

Ian Hussey

Ilias exercise: Installation and setup of R + RStudio

Problems?



Psychology of Digitalisation

Ian Hussey





Psychology of Digitalisation

Ian Hussey

Technical Skills

Beginner

R you Ready? Reproducible Data Preparation and Analyses in R [Beginner] (S/MS)

Intermediate

Reproducible Data Processing and Visualisation in R [Intermediate] (S/MS)

Advanced

Improving your Statistical Inferences through Monte Carlo Simulation Studies (S/MS)

Estimating the Credibility of Past Research (S/MS)

Transdisciplinary Data Science (L)

Substantive Knowledge

Meta-Science

Privacy and Security

Digitalisation

Psychology of Digitalisation: Everything is Computer (L)

Digitalisation of Psychology: Computer is Everything (L)

Psychology in Crisis? (S/MS)

Escape Rooms: Cybersecurity Education at School (S)

Influencer Marketing (S)

Usable Security (S)

Intervention & Consulting at Lufthansa (S)

u^b

Collegium generale

Kein Witz



**Interdisziplinäre Ringvorlesung
über Humor im HS25**

Mittwochs, 18.15 – 19.45 Uhr
Hauptgebäude, Auditorium maximum

KSL 505945-HS2025
3 ECTS



Why am I here?



Thomas Herndon in 2013
Economics PhD student

He took a class where the professor asked students to examine the results of influential papers



He chose **Reinhart & Rogoff (2010) Growth in a Time of Debt**

“Economic growth slows dramatically when the size of a country's debt rises above 90% of Gross Domestic Product”

Reinhart & Rogoff's paper **changed the world**



Countries based economic ‘austerity’ policies on it

Ireland & Greece were sacrificed to stop the spread of the debt crisis to the rest of Europe

4 million Irish people took on 190 Billion CHF in debt

Slashed our healthcare, education, infrastructure,
- These still have not recovered

[Home](#) > [Lifestyle](#)

George Osborne plunged UK into austerity due to an 'error on a spreadsheet'

An academic paper produced by two Harvard economists and relied upon by Cameron and Osborne contained serious miscalculations, it has been claimed.

by **Jack Peat** — 2022-09-22 15:30 in [Lifestyle](#)[!\[\]\(735ceeed4e566aa93749bb6365185b00_img.jpg\) Facebook](#)[!\[\]\(67b99c4d964e257c5482ff1b2462b204_img.jpg\) Twitter](#)[!\[\]\(94480c799e843c3a4dcfaf8c99e6db79_img.jpg\) LinkedIn](#)[!\[\]\(7453c0f29ed3a7dcecf77fe714fbbf84_img.jpg\) Email](#)[!\[\]\(15d3dfb11951c9197b3fa51927099453_img.jpg\) WhatsApp](#)

George Osborne plunged the UK into austerity "all for nothing" due to an error on an Excel spreadsheet, according to a podcaster.

Hosts James Harkin, Andrew Hunter Murray, Anna Ptaszynski, and Dan Schreiber – from the No Such Thing As A Fish podcast – claim the UK reacted to incorrect data in an academic paper on economic growth and inflation.

Subscribe to our Newsletter

Email address:

Your e-mail address

[SIGN UP](#)[View our Privacy Policy and Terms & Conditions](#)

More from TLE

- ▶ Watchdog calls for ticket office closures to be delayed over machine upgrades.
- ▶ Richard Curtis uses Al Pacino's Any Given Sunday speech to inspire global action on climate change
- ▶ James Cleverly holidayed with well-connected Tory lobbyist in Gibraltar
- ▶ Eleveses: Has the Left Abandoned the Working Class?
- ▶ UK recession fears grow as economy contracts
- ▶ Labour calls on Sunak to block Liz Truss's soon-to-be-published honours list
- ▶ Mick Lynch suggests RMT union will not comply with new anti-strike legislation
- ▶ Russell Brand receives standing ovation at Wembley show following sexual assault

	B	C	I	J	K	L	M
2			Real GDP growth				
3			Debt/GDP				
4	Country	Coverage	30 or less	30 to 60	60 to 90	90 or above	30 or less
26			3.7	3.0	3.5	1.7	5.5
27	Minimum		1.6	0.3	1.3	-1.8	0.8
28	Maximum		5.4	4.9	10.2	3.6	13.3
29							
30	US	1946-2009	n.a.	3.4	3.3	-2.0	n.a.
31	UK	1946-2009	n.a.	2.4	2.5	2.4	n.a.
32	Sweden	1946-2009	3.6	2.9	2.7	n.a.	6.3
33	Spain	1946-2009	1.5	3.4	4.2	n.a.	9.9
34	Portugal	1952-2009	4.8	2.5	0.3	n.a.	7.9
35	New Zealand	1948-2009	2.5	2.9	3.9	-7.9	2.6
36	Netherlands	1956-2009	4.1	2.7	1.1	n.a.	6.4
37	Norway	1947-2009	3.4	5.1	n.a.	n.a.	5.4
38	Japan	1946-2009	7.0	4.0	1.0	0.7	7.0
39	Italy	1951-2009	5.4	2.1	1.8	1.0	5.6
40	Ireland	1948-2009	4.4	4.5	4.0	2.4	2.9
41	Greece	1970-2009	4.0	0.3	2.7	2.9	13.3
42	Germany	1946-2009	3.9	0.9	n.a.	n.a.	3.2
43	France	1949-2009	4.9	2.7	3.0	n.a.	5.2
44	Finland	1946-2009	3.8	2.4	5.5	n.a.	7.0
45	Denmark	1950-2009	3.5	1.7	2.4	n.a.	5.6
46	Canada	1951-2009	1.9	3.6	4.1	n.a.	2.2
47	Belgium	1947-2009	n.a.	4.2	3.1	2.6	n.a.
48	Austria	1948-2009	5.2	3.3	-3.8	n.a.	5.7
49	Australia	1951-2009	3.2	4.9	4.0	n.a.	5.9
50							
51			4.1	2.8	2.8	=AVERAGE(L30:L44)	

This course is intended to fill the gap
between other courses



Why are you here?

Wollen wir uns duzen?

Course aims

How you will know it

Computational reproducibility of psychology research is low
e.g., 1 of 14 articles
(Crüwell et al., 2023; Hardwicke et al. 2020)

By the end of this class, you should be able to write more reproducible code than most psychology researchers

Course aims

How you will know it



© Sarah Andersen

Problem Based Learning

No slides

We move at the pace of the room

Asking questions is essential!

Please offer your broken code and
we'll fix it together

How is the seminar structured?

Organisation & assessment

**This seminar asks more from you
than some others**

**Hopefully it you will also learn
proportionally more**

Organisation & assessment



Assessment

Weekly unmarked exercises
3 larger graded assignments



Attendance

Weekly meetings in B202
Bring your laptop!
You can miss max 2 sessions
Plan ahead for weekly assignments



Communication

Slack wherever possible
Discussion & problem solving encouraged!
Email if necessary: ian.hussey@unibe.ch

Rough schedule

Weeks 1-7	<ul style="list-style-type: none">- Working through book and its exercises
Weeks 8-14	<ul style="list-style-type: none">- Larger assignments- Working through solutions together- Live-coding sessions and Q&As
Week 9 (?) November 10th	<p>Guest lecture by Dr Saloni Dattani on data visualisation</p> <p>Formerly <i>Our World in Data</i>, Currently <i>Works In Progress</i>, scientificdiscovery.dev, <i>Hard Drugs</i> podcast, and advisor at <i>Open Philanthropy</i></p>







Introduction

- 1 Installation and setup
- 2 Fundamentals
- 3 Reproducible reports
- 4 Loading, viewing, and saving data
- 5 The pipe and renaming
- 6 Data transformation I
- 7 Data transformation II
- 8 Working with strings and factors
- 9 Reshaping and pivots
- 10 Binding and joining
- 11 Structuring projects
- 12 Visualization
- 13 Reporting
- 14 Data privacy and anonymity
- 15 License and citation

Reproducible Data Processing and Visualization

in R and tidyverse

</> Code

Table of contents

[Introduction](#)

How to use this book

Other learning resources

Find me elsewhere

Contributing

AUTHOR

Ian Hussey

PUBLISHED

August 27, 2025

Introduction

This [Open Source](#) eBook provides materials for the semester-long Master's seminar course “Reproducible Data Processing and Visualization in R” that I deliver at the University of Bern’s Institute of Psychology.

How to use this book

This book is made up of individual Quarto (.qmd) workbooks. Many of the exercises are easiest to complete in your own local copy of these .qmd files.

I suggest that you [download the .zip file](#) of all the .qmd files and their supporting data and cheatsheets and use the eBook as a reference book. *If the .zip file does not contain the most recent versions, please contact me and I'll update it.* If I'm slow or you're feeling impatient, you can also [download the most recent versions from GitHub](#), along with the other files that create this eBook.

You can also copy and paste the code for any chapter directly from the website. Click the “</> Code” button on the top right of each page to see the full .qmd file’s code. You can copy and paste this into a .qmd file. However, it’s probably easier to download all the .qmd files and data as mentioned above.

Learning to code is a practice skill. Almost anyone can become competent in writing reproducible code for data processing and visualization with practice. More than anything else, completing this course requires that you practice in your own time, using not only the examples provided but also ones you create yourself. Take real data sets from your own studies, or the thousands available on the Open Science Framework ([osf.io](#)), or create simulated messy datasets with my R package [{truffle}](#) and practice.



Reproducible
Data Processing
and Visualization

in R and tidyverse

Ian Hussey

Agenda

Access book

- ianhussey.quarto.pub/reproducible-data-processing-and-visualization

Questions about chapter: installation and setup

Questions about chapter: Fundamentals

Do chapter: Reproducible reports