



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/141673/>

Version: Published Version

---

**Book:**

Lovelace, R, Nowosad, J and Muenchow, J (2019) Geocomputation with R. The R Series. Chapman and Hall/CRC Press. ISBN: 9781138304512.

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# Geocomputation with R

*Robin Lovelace, Jakub Nowosad, Jannes Muenchow*

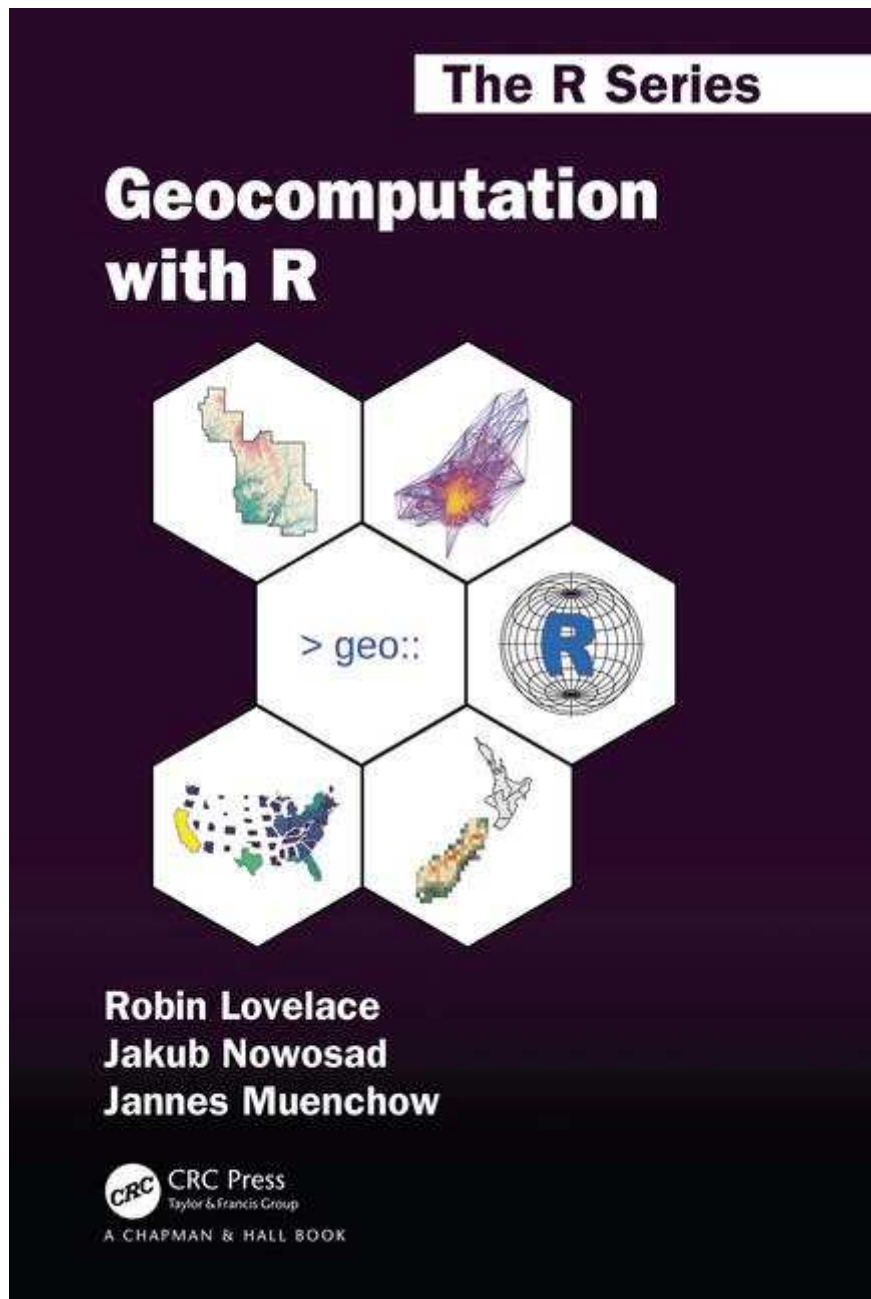
2019-03-10

## Welcome

---

This is the online home of *Geocomputation with R*, a book on geographic data analysis, visualization and modeling.

**Note:** This book has now been published by CRC Press in the [R Series](#). You can buy the book from [CRC Press](#), [Wordery](#), or [Amazon](#).



Inspired by [bookdown](#) and the Free and Open Source Software for Geospatial (FOSS4G) movement, this book is open source. This ensures its contents are reproducible and publicly accessible for people worldwide.

The online version of the book is hosted at [geocompr.robinlovelace.net](https://geocompr.robinlovelace.net) and kept up-to-date by [Travis](#), which provides information on its 'build status' as follows:

build passing

The version of the book you are reading now was built on 2019-03-10 and was built on [Travis](#).

## How to contribute?

**bookdown** makes editing a book as easy as editing a wiki, provided you have a GitHub account ([sign-up at github.com](https://github.com)). Once logged-in to GitHub, click on the 'edit me' icon highlighted with a red ellipse in the image below. This will take you to an editable version of the the source R Markdown file that generated the page you're on:



To raise an issue about the book's content (e.g. code not running) or make a feature request, check-out the [issue tracker](#).

## Reproducibility

To reproduce the code in the book, you need a recent version of R and up-to-date packages. These can be installed with the following command (which requires **devtools**):

```
devtools::install_github("geocompr/geocompr")
```

To build the book locally, clone or [download the geocompr repo](#), load R in root directory (e.g. by opening `geocompr.Rproj` in RStudio) and run the following lines:

```
bookdown::render_book("index.Rmd") # to build the book  
browseURL("_book/index.html") # to view it
```

## Supporting the project

If you find the book useful, please support it by:

- Recommending, [citing](#) or [linking-to](#) it
- 'Starring' the [geocompr GitHub repository](#)
- Reviewing it, e.g. on Amazon or [Goodreads](#)
- Buying a copy

Further details can be found at [github.com/Robinlovelace/geocompr](https://github.com/Robinlovelace/geocompr).



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

# Preface

---

## Who this book is for

This book is for people who want to analyze, visualize and model geographic data with open source software. It is based on R, a statistical programming language that has powerful data processing, visualization and geospatial capabilities. The book covers a wide range of topics and will be of interest to a wide range of people from many different backgrounds, especially:

- People who have learned spatial analysis skills using a desktop Geographic Information System (GIS) such as [QGIS](#), [ArcMap](#), [GRASS](#) or [SAGA](#), who want access to a powerful (geo)statistical and visualization programming language and the benefits of a command-line approach (Sherman 2008):

With the advent of ‘modern’ GIS software, most people want to point and click their way through life. That’s good, but there is a tremendous amount of flexibility and power waiting for you with the command line.

- Graduate students and researchers from fields specializing in geographic data including Geography, Remote Sensing, Planning, GIS and Geographic Data Science
- Academics and post-graduate students working on projects in fields including Geology, Regional Science, Biology and Ecology, Agricultural Sciences (precision farming), Archaeology, Epidemiology, Transport Modeling, and broadly defined Data Science which require the power and flexibility of R for their research
- Applied researchers and analysts in public, private or third-sector organizations who need the reproducibility, speed and flexibility of a command-line language such as R in applications dealing with spatial data as diverse as Urban and Transport Planning, Logistics, Geo-marketing (store location analysis) and Emergency Planning

The book is designed for intermediate-to-advanced R users interested in geocomputation and R beginners who have prior experience with geographic data. If you are new to both R and geographic data, do not be discouraged: we provide links to further materials and describe the nature of spatial data from a beginner’s perspective in Chapter 2 and in links provided below.

## How to read this book

The book is divided into three parts:

1. Part I: Foundations, aimed at getting you up-to-speed with geographic data in R.
2. Part II: Extensions, which covers advanced techniques.
3. Part III: Applications, to real-world problems.

The chapters get progressively harder in each so we recommend reading the book in order. A major barrier to geographical analysis in R is its steep learning curve. The chapters in Part I aim to address this by providing reproducible code on simple datasets that should ease the process of getting started.

An important aspect of the book from a teaching/learning perspective is the **exercises** at the end of each chapter. Completing these will develop your skills and equip you with the confidence needed to tackle a range of geospatial problems. Solutions to the exercises, and a number of extended examples, are provided on the book's supporting website, at [geocompr.github.io](https://geocompr.github.io).

Impatient readers are welcome to dive straight into the practical examples, starting in Chapter 2. However, we recommend reading about the wider context of *Geocomputation with R* in Chapter 1 first. If you are new to R, we also recommend learning more about the language before attempting to run the code chunks provided in each chapter (unless you're reading the book for an understanding of the concepts). Fortunately for R beginners R has a supportive community that has developed a wealth of resources that can help. We particularly recommend three tutorials: [R for Data Science](#) (Grolemund and Wickham 2016) and [Efficient R Programming](#) (Gillespie and Lovelace 2016), especially [Chapter 2](#) (on installing and setting-up R/RStudio) and [Chapter 10](#) (on learning to learn), and [An introduction to R](#) (Venables, Smith, and Team 2017). A good interactive tutorial is DataCamp's [Introduction to R](#).

## Why R?

Although R has a steep learning curve, the command-line approach advocated in this book can quickly pay off. As you'll learn in subsequent chapters, R is an effective tool for tackling a wide range of geographic data challenges. We expect that, with practice, R will become the program of choice in your geospatial toolbox for many applications. Typing and executing commands at the command-line is, in many cases, faster than pointing-and-clicking around the graphical user interface (GUI) a desktop GIS. For some applications such as Spatial Statistics and modeling R may be the *only* realistic way to get the work done.

As outlined in Section 1.2, there are many reasons for using R for geocomputation: R is well-suited to the interactive use required in many geographic data analysis workflows compared with other languages. R excels in the rapidly growing fields of Data Science (which includes data carpentry, statistical learning techniques and data visualization) and Big Data (via efficient interfaces to databases and distributed computing systems). Furthermore R enables a reproducible workflow: sharing scripts underlying your analysis will allow others to build-on your work. To ensure reproducibility in this book we have made its source code available at [github.com/Robinlovelace/geocompr](https://github.com/Robinlovelace/geocompr). There you will find script files in the `code/` folder that generate figures: when code generating a figure is not provided in the main text of the book, the name of the script file that generated it is provided in the caption (see for example the caption for Figure 12.2).

Other languages such as Python, Java and C++ can be used for geocomputation and there are excellent resources for learning geocomputation *without R*, as discussed in Section 1.3. None of these provide the unique combination of package ecosystem, statistical capabilities, visualization options, powerful IDEs offered by the R community. Furthermore, by teaching how to use one language (R) in depth, this book will equip you with the concepts and confidence needed to do geocomputation in other languages.

## Real-world impact

*Geocomputation with R* will equip you with knowledge and skills to tackle a wide range of issues, including those with scientific, societal and environmental implications, manifested in geographic data. As described in Section 1.1, geocomputation is not only about using computers to process geographic data: it is also about real-world impact. If you are interested in the wider context and motivations behind this book, read on; these are covered in Chapter 1.

## Acknowledgements

Many thanks to everyone who contributed directly and indirectly via the code hosting and collaboration site GitHub, including the following people who contributed direct via pull requests: [katygregg](#), [erstearns](#), [eyesofbambi](#), [tyluRp](#), [marcosci](#), [giocomai](#), [mdsumner](#), [rsbivand](#), [pat-s](#), [gisma](#), [ateucher](#), [annakrystalli](#), [gavinsimpson](#), [Henrik-P](#), [Himanshuteli](#), [yutannihilation](#), [jbixon13](#), [katiejolly](#), [layik](#), [mvl22](#), [nickbearman](#), [ganes1410](#), [richfitz](#), [SymbolixAU](#), [wdearden](#), [yihui](#), [chihinl](#). Special thanks to Marco Sciaini, who not only created

the front cover image, but also published the code that generated it (see `frontcover.R` in the book's GitHub repo). Dozens more people contributed online, by raising and commenting on issues, and by providing feedback via social media. The `#geocompr` hashtag will live on!

We would like to thank John Kimmel from CRC Press, who has worked with us over two years to take our ideas from an early book plan into production via four rounds of peer review. The reviewers deserve special mention here: their detailed feedback and expertise substantially improved the book's structure and content.

We thank Patrick Schratz and Alexander Brenning from the University of Jena for fruitful discussions on and input into Chapters 11 and 14. We thank Emmanuel Blondel from the Food and Agriculture Organization of the United Nations for expert input into the section on web services; Michael Sumner for critical input into many areas of the book, especially the discussion of algorithms in Chapter 10; Tim Appelhans and David Cooley for key contributions to the visualization chapter (Chapter 8); and Katy Gregg, who proofread every chapter and greatly improved the readability of the book.

Countless others could be mentioned who contributed in myriad ways. The final thank you is for all the software developers who make geocomputation with R possible. Edzer Pebesma (who created the `sf` package), Robert Hijmans (who created `raster`) and Roger Bivand (who laid the foundations for much R-spatial software) have made high performance geographic computing possible in R.



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

## References

Gillespie, Colin, and Robin Lovelace. 2016. *Efficient R Programming: A Practical Guide to Smarter Programming*. O'Reilly Media.

Grolemund, Garrett, and Hadley Wickham. 2016. *R for Data Science*. O'Reilly Media.

Sherman, Gary. 2008. *Desktop GIS: Mapping the Planet with Open Source Tools*. Pragmatic Bookshelf.

Venables, W.N., D.M. Smith, and R Core Team. 2017. *An Introduction to R. Notes on R: A Programming Environment for Data Analysis and Graphics*.