Crime Mapping in R

Reka Solymosi and Juanjo Medina

2021-02-03

Contents

Prelude

0.1 Introduction

This workbook contains the lab materials for our Crime Mapping module in Department of Criminology at the University of Manchester. This module is an optional unit open to 3rd year undergraduate and postgraduate students.

It makes use of R, as it is a free, open source tool, that has tremendous community support, and great versatility in mapping applications. You can find more details about the advantages of R for geospatial work here

Crime Mapping introduces students to the concepts of spatial data analysis. The aim is to familiarise students with basic concepts of GIS, and get acquainted with spatial statistics to be able to talk about data about crime, policing, and criminal justice topics situated in the places they occur. Details can be found in the Syllabus.

0.2 Disclaimer

Please beware that:

- In making these notes, while we briefly cover some concepts, students are expected to do the weekly reading, and attend the weekly lectures, as well as participate in lab disucssions to receive a complete course experience. These notes are *not* intended to be a stand-alone reference or textbook, rather a set of exercises to gain hands-on practice with the concepts introduced during the course.
- These pages are the content of the university course mentioned above. They are meant to (very gently) introduce students to the concept of spatial data analysis, and cover descriptive statsitics and the key concepts required to build an understanding of quantitative data analysis in crime research.

- The handouts below use, among other data sets, dara from the UK data service such as the Crime Survey for England and Wales that is available under a Open Government Licence. This dataset is designed to be a learning resource and should not be used for research purposes or the production of summary statistics.

Chapter 1

A first lesson about R

In this lesson, you will be introduced to the programming language, `R`. After installing the related software and getting a basic idea of the R Studio interface, you will learn three `R` basics: operators, objects, and packages.

Unlike other statistical software like SPSS and STATA, `R` is a free, open-source software for performing statistical analysis and data visualization. In addition, `R` offers more analytical solutions, flexibility, and customization than these commonly used statistical software, and its popularity has increased substantially over the years.

We learn `R` because we hope that this is an important tool that you will continue to use in future. As it is free and has a community feel to it where anyone can create and upload new techniques, the idea is that you can use `R` long after this course. Even if data analysis is not in the future for you, learning how to conduct and interpret statistical output is a good skill to have – much of our knowledge of the world includes statistics, so understanding the numbers and how they were derived are advantages. `R` uses a language called **object-oriented programming**, and though it may seem daunting at first, practice makes familiarity. Also, you can impress your friends with all your coding.

1.1 Install R & R Studio

As `R` and `R Studio` are free software, you should be able to install these on your own machines at home. You may be working with different IT, so there are different options to allow a successful install. Our first activity will be to decide what approach to working with `R` and `R Studio` will be best for you.

1.1.1 Activity 1: Identifying your operating system

In this activity, you need to answer a question about your computer/IT that you will be working with for this class. That question is:

- **What is your operating system**? Operating system refers to the software that your computer is running to deliver the basic functions. You may have, for example:
  - *Windows or Linux* - if you have these, you are most likely going to have an easy time installing `R` and `R Studio`, so you should give the installation instructions below a try
  - *Apple* - if you have a Mac, there are some extra steps to install `R` and `R Studio`. Specifically, there will be an additional programme to download called Xcode, and additional steps to take.
  - *Chromebook* - Installing `R` and `R Studio` on a Chromebook involves installing Linux. Like with a Mac, there are some additional steps you will need to take, and some patience.

In your group google sheets, write down which operating system you have. This will guide which directions to follow later.

1.1.2 Activity 2: Install R & R Studio

1.1.2.1 Some notes specific to your operating system.

Before you move on to the installation steps, look at your answer from Activity 1, and read or watch the advice specific to your operating system:

- Windows: click here for instructions
- Chromebook: read the tutorial here
- Mac, follow the guidance in the video here and then, you will also need to install command line tools, for that you can watch another video here
- Linux: for ubuntu see the video here and if you have questions, let the teaching team know!

Once you have watched or read the instructions for your relevant operating system, you are now ready to actually have a go at downloading the software for yourself. Before you start, write in the google doc any questions or concerns, and once ready, install!

Install `R`:

1. Go to https://www.r-project.org/
2. Click the download R link under the *Getting Started* heading
3. You will be asked to select a Comprehensive R Archive Network (CRAN) mirror. Click the URL closest to your location
4. Click whichever download link is appropriate for your operating system (see above).
5. Then click the *install R for the first-time* link and proceed to install `R`

Install R Studio:

1. Go to https://rstudio.com/
2. Click the *Download* link
3. Scroll down and then click the DOWNLOAD button for the free version of `RStudio`
4. You will be taken to a website to download the free version of `RStudio` that is appropriate for your computer. Download and then install it.

1.1.2.2 Plan B: accessing R Studio remotely throught a web browser

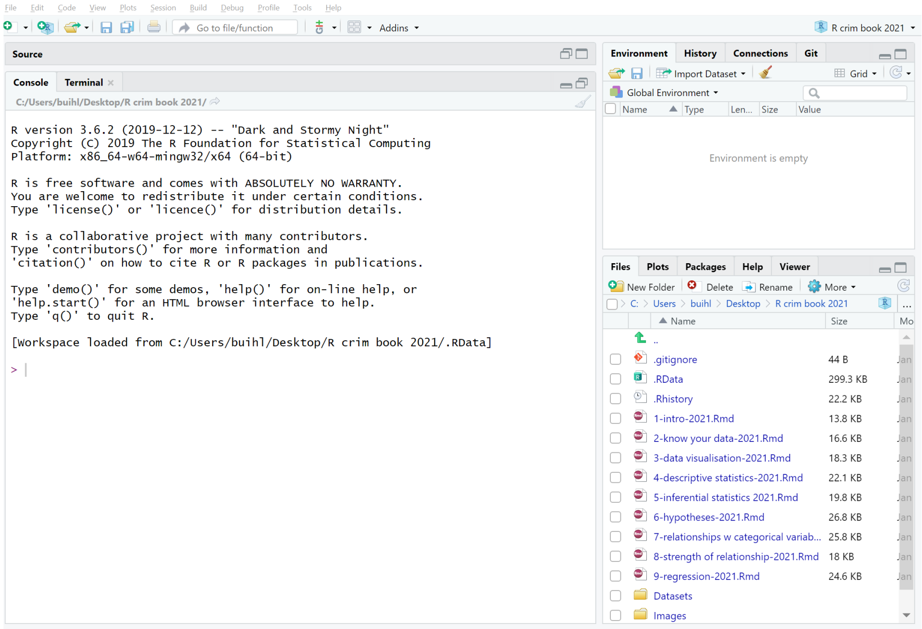
It might be that the above does not work and you find that there are some issues specific to your computer, or something just is not working. In that case, you have two options:

- *Option 1:* You can remotely access one of the university PCs from your browser (Firefox, Chrome, Safari, etc). You can find instructions how to do this here , and the university IT helpdesk can help you access this too. If you do this, you will be able to use the version of `RStudio` installed in the computer clusters.
- *Option 2*: You can access an online version of `R Studio`, which you can access through any web browser (Firefox, Chrome, Safari, etc). To do this, you visit https://rstudio.cloud/, click on ‘get started for free’, choose the free account, and click on ‘sign up’. Then you can always visit this website and log in to use `R Studio` in the cloud. Note that you should start a `New Project` and name it *Modelling Crime Data*, and then all your work will be saved in this project.

1.2 Open up and explore RStudio

In this session we will focus in developing basic familiarity with R Studio. You can use R without using R Studio, but R Studio is an app that makes it easier to work with R.

R Studio is what we call an IDE, an **integrated development environment**. It is a fancy way of saying that it is a cool interface designed to write programming code. Every time you open up R Studio you are in fact starting a R session. R Studio automatically runs R in the background. We will be interacting with R in this course unit via R Studio.



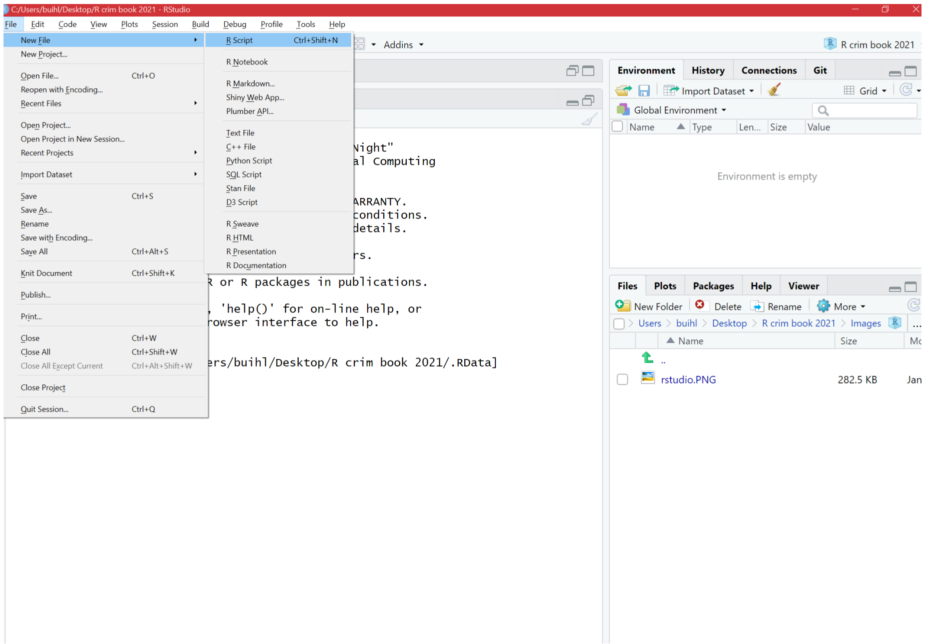
```
R version 3.6.2 (2019-12-12) -- "Dark and Stormy Night"
Copyright (C) 2019 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace loaded from C:/Users/buihl/Desktop/R crim book 2021/.RData]

> |
```

Name	Size	Mo
..		
.gitignore	44 B	Jan
.RData	299.3 KB	Jan
.Rhistory	22.2 KB	Jan
1-intro-2021.Rmd	13.8 KB	Jan
2-know your data-2021.Rmd	16.6 KB	Jan
3-data visualisation-2021.Rmd	18.3 KB	Jan
4-descriptive statistics-2021.Rmd	22.1 KB	Jan
5-inferential statistics 2021.Rmd	19.8 KB	Jan
6-hypotheses-2021.Rmd	26.8 KB	Jan
7-relationships w categorical variab...	25.8 KB	Jan
8-strength of relationship-2021.Rmd	18 KB	Jan
9-regression-2021.Rmd	24.6 KB	Jan
Datasets		
Images		

When you first open R Studio, you will see (as in the image above) that there are 3 main windows. The bigger one to your left is the console. If you read the text in the console you will see that R Studio is indeed opening R and you can see what version of R you are running. Since R is constantly being updated the version you installed is likely more recent that the one we used at time of writing.

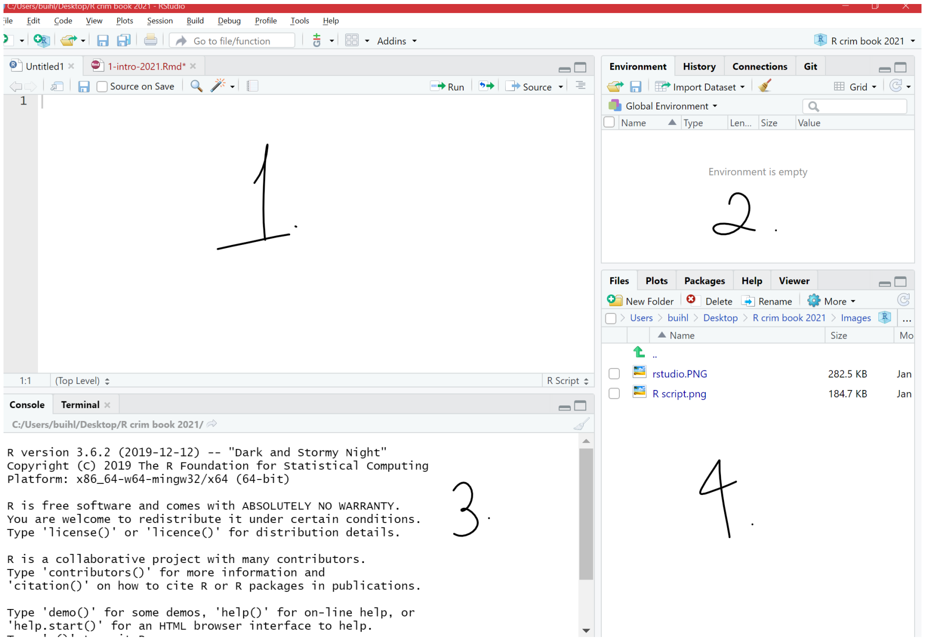
1.2.0.1 Activity 3: Opening up the script pane

The view in R Studio is structured so that you have 4 open windows in a regular session. However when you first open, you might be starting with only 3. To open the script pane (the one missing) click in the *File* drop down Menu, select *New File*, then *R Script*.



```
Night"
al Computing

ARRANTY.
conditions.
details.

rs.

R or R packages in publications.

, 'help()' for on-line help, or
rowser interface to help.

ers/buihl/Desktop/R crim book 2021/.RData]
```

You will now see the 4 window areas in display. On each of these areas you can shift between different views and panels. You can also use your mouse to re-size the different windows if that is convenient.


C:/Users/buihl/Desktop/R crim book 2021 - RStudio
ile Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
R crim book 2021
Untitled1
1-intro-2021.Rmd*
Source on Save
Run
Source
1.
1:1
(Top Level)
R Script
Console Terminal
C:/Users/buihl/Desktop/R crim book 2021/
3.
Environment History Connections Git
Import Dataset
Grid
Global Environment
Name Type Len... Size Value
Environment is empty
2.
Files Plots Packages Help Viewer
New Folder Delete Rename More
Users > buihl > Desktop > R crim book 2021 > Images
Name Size Mo
..
rstudio.PNG 282.5 KB Jan
R script.png 184.7 KB Jan
4.

```
R version 3.6.2 (2019-12-12) -- "Dark and Stormy Night"
Copyright (C) 2019 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
```

1.2.1 The four panes of R Studio

The purposes of the four panes in Figure above are the following:

1. **Script and data view**: where you type your programming code that tells `R` what you want to do. These are essentially instructions that you type and save as a **script**, so that you can return to it later to remember what you did and to share it with others so that they can reproduce what you did.

2. **Environment and history view**:

i) *2.1 Environment* tab- gives you the names of all the (data) objects that you have defined during your current `R` session, including number of observations and rows in those objects. We learn more about objects later.

ii) *2.2 History* tab- shows you a history of all the code you have previously evaluated in the main console. One of the key advantages of doing data analysis this way - with code versus with a point and click interface like Excel or ArcGIS is that you are producing a written record of every step you take in the analysis. First time around it will take you time to write these instructions, it may be slower than pointing and clicking. And unlike with pointing and clicking you need to know the "words" and "grammar" of this language.

3. **Main console**: this is considered `R's` heart, and it is where `R` evaluates the codes that you run. You can type your codes directly in the console, but for the sake of good habits, type them in the script and data view so you can save a record of them. Only type and run code from here if you want to debug or do some quick analysis.

4. **File directory, Plots, Packages, Help**:

i) *4.1 Files* tab- allows you to see the files in the folder that is currently set as your working directory.

ii) *4.2 Plots* tab- you will see any data visualizations that you produce here. You have not produced any yet, so it is empty now.

iii) *4.3 Packages* tab- you will see the packages that are currently available to install. We will explain what these are soon, but know that they are an essential feature when working with `R`.

iv) *4.4 Help* tab- you can access further information on the various packages.