

# R Training for Beginners

Adam Rawles



# Contents

<b>1</b>	<b>Overview</b>	<b>5</b>
1.1	Welcome . . . . .	5
1.2	About Me . . . . .	5
<b>2</b>	<b>Introduction to R</b>	<b>7</b>
2.1	What is R and RStudio? . . . . .	7
2.2	Basic arithmetic operators . . . . .	7
2.3	Variable assignment . . . . .	7
2.4	Data types . . . . .	7
2.5	Data structures (and subsetting) . . . . .	8
2.6	Functions . . . . .	8
<b>3</b>	<b>Data Analysis in R</b>	<b>9</b>
3.1	Installing packages . . . . .	9
3.2	Loading data . . . . .	9
3.3	Cleaning data . . . . .	9
3.4	Summary statistics . . . . .	9
3.5	Graphs and plots . . . . .	10
<b>4</b>	<b>Programming in R</b>	<b>11</b>
4.1	User-defined functions . . . . .	11
4.2	For loops . . . . .	11
4.3	If/else statements . . . . .	11
<b>5</b>	<b>Tidyverse</b>	<b>13</b>
5.1	tidyr . . . . .	13
5.2	dplyr . . . . .	13
5.3	stringr . . . . .	13
5.4	ggplot2 . . . . .	13
<b>6</b>	<b>Exercises</b>	<b>15</b>



# Chapter 1

## Overview

### 1.1 Welcome

This book is a collection of training materials for an introduction to the R statistical computing programming language. Broken down into chapters, I've aimed to cover most of the basics. Alongside each chapter is a xaringan presentation. This can help for those looking to learn, but can also function as a first step for those looking to begin teaching R but who don't have the time to fully develop their own training modules. Hopefully, all the topics covered in the text version will be covered in the presentation and vice versa, so if you learn visually then you can rest easy knowing that you're not missing out! You can find a link to each presentation on the first page of each chapter.

This is a work in progress, and so I would greatly appreciate any feedback. Anything from typos to content suggestions, feel free to raise a GitHub issue if you feel something should be changed.

### 1.2 About Me

I began using R in my second year of university, during an internship looking at publication bias correction methods. I was under the tutorship of a member of staff who helped me immensely, but I must confess that I have never taken an official course in R, online or in person. I like to think, however, that this is not always a bad thing. Learning from the bottom up and struggling along the way is a fantastic way to acquire knowledge, and I believe instills a very important lesson:

You're not going to know everything there is to know about R. Ever. But that's okay.

I'm now 4 years into my R career, and I use R every day. With that in mind, I don't think there has ever been a day where I haven't referred to a tutorial, or Stack Overflow, or even just Googled the name of a function that I've used 1000 times before. There is a great repository of knowledge for R, and it's one of the things I love most about the R community. So please never feel as though you're an impostor in a world of R gurus. In reality, everyone else is just as lost as you. But if you keep ticking along and never feel that learning something new in R isn't worth your time, you'll end up doing some great things.

And in a roundabout way, that is part of the reason I decided to develop these materials. I don't pretend to be the ultimate R programmer, because I still know what it's like to learn something from the start. And everyone has to start somewhere. So I hope that I can help impart some of the lessons that I've learnt over the 4 years to anyone who's looking to learn R in a way that won't leave you feeling lost - and hopefully I can save some poor soul from the same headaches I suffered along the way.

The only final note I have before we start learning how to use R is another bit of advice:

Don't believe everything you learn in training.

Whilst this is probably a good thing to keep in mind for any type of training, I feel it's particularly relevant with R for two reasons. Firstly, when it comes to programming languages, lots of people have opinions. Some are true, most are not. Most things you read are a mix between fact and opinion, so take everything with a pinch of salt. For example, the developers of the `ggplot2` package are fervently against arbitrary second axes and so support for them in `ggplot2` is limited. I also share this view, but that doesn't mean that I'm right - read, learn, but question and make your own mind up. Secondly, R is a language that is still changing. For this reason, people may make statements relative to one version of R that aren't necessarily true in the future. Things have changed over the years, and so answers from a 10 year-old Stack Overflow question may not still be true when you come across them. A microcosmic version of this are some recent changes in the `tidyr` package. Historically, converting data from/to long and short formats was done using the `spread()` and `gather()` functions. However, in newer releases, these functions are deprecated in favour of `pivot_wider()` and `pivot_longer()`, which provide the same functionality but also some extra bits. The practical implication of this suggestion is don't always read one tutorial on a subject before you dive in.

## Chapter 2

# Introduction to R

Module 1 is a general overview of R and its basics. I personally use R Studio as my development environment of choice (more on development environments later on), so I've also included a very brief overview of R Studio.

The xaringan presentation for this module is [here](#).

### 2.1 What is R and RStudio?

filler

### 2.2 Basic arithmetic operators

filler

### 2.3 Variable assignment

filler

### 2.4 Data types

filler

## **2.5 Data structures (and subsetting)**

filler

## **2.6 Functions**



## Chapter 3

# Data Analysis in R

In Module 2, we'll look more specifically at how one might do some simple data analysis in R. For a more in-depth view, I would highly recommend Hadley's R4DS

The xaringan presentation for this module can be found [here](#).

### 3.1 Installing packages

filler

### 3.2 Loading data

filler

### 3.3 Cleaning data

filler

### 3.4 Summary statistics

filler

## 3.5 Graphs and plots

## Chapter 4

# Programming in R

In this Module, we'll look at some of the programming concepts and syntax in R.

The xaringan presentation for this module can be found [here](#).

### 4.1 User-defined functions

filler

### 4.2 For loops

filler

### 4.3 If/else statements

filler



## Chapter 5

# Tidyverse

In this module, we'll take a quick look over some of the packages in the tidyverse.  
The xaringan presentation for this module can be found [here](#).

### 5.1 tidyr

filler

### 5.2 dplyr

filler

### 5.3 stringr

filler

### 5.4 ggplot2

filler



## Chapter 6

# Exercises

Here are some interactive exercises covering the basic concepts we've looked at in the last 4 modules.

This exercises are hosted on a Shiny server [here](#)