

An Introduction to R

3rd – 5th December 2019

Alex Douglas

[@a.douglas@abdn.ac.uk](mailto:a.douglas@abdn.ac.uk)
[@Scedacity](https://twitter.com/Scedacity)

UNIVERSITY OF ABERDEEN




1

goals

- over the next three days.....
- introduce you to using R for data manipulation, graphics and statistics
- provide you with the practical skills necessary to explore and visualise your data
- provide you with the practical skills necessary to start your own data analysis
- BUT cannot teach you everything there is to know about R: that's down to you

2

course structure

- course split into 6 main components
 - introduction to R environment
 - basic R operations
 - dataframes
 - graphics and data exploration
 - basic statistics using R
 - R programming (optional)
- work through manual and associated exercises

3

what is R?

- environment for statistical computing, graphics and programming
- originally created by Ross Ihaka and Robert Gentleman (1996)
- currently maintained by international R-core development team
- very similar to S



4

R-Project and CRAN

- more information found at [www.r-project.com](http://www.r-project.org)
- download R from <http://cran.uk.r-project.org>

The R Project for Statistical Computing

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and Mac OS. To download R, please choose your preferred CRAN mirror.

If you have questions about R or how to download and install the software, or any of the license terms etc., please read our answers to frequently asked questions before you send us email.

News

- R version 3.4.0 (You Biased Darkness) has been released on Friday 2017-04-21.
- R version 3.3.3 (Great裹 R) has been released on Monday 2017-03-06.
- RwiR 2017 July 4 - 7, in Brussels has opened registration and more at <http://useR2017.brussels/>
- Tomas Kalvoda has joined the R core team.
- The R Foundation welcome five new ordinary members: Jennifer Bryan, Dianne Cook, Julie Josse, Tomas Kalvoda, and Balasubramanian Narasingham.
- The R Journal Volume 8/1 is available.
- The useR! 2017 conference will take place in Brussels, July 4 - 7, 2017.

R Foundation
 Nonprofit
 Board
 Members
 Donors
 Donate

5

why you shouldn't bother with R

relatively steep learning curve?
 it's command line driven
 you need to learn to 'speak' R



```
se.fnc <- function(x){ #start function for se
  std.x <- sd(x)           #calculate SD
  nos.x <- length(x)       #calculate number obs
  se.x <- std.x/(sqrt(nos.x))#calculate SE
  print(se.x)               # print SE
}                           #end of function
```

6

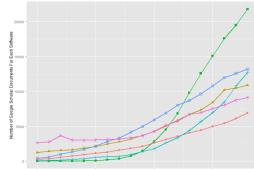
why you should learn 

it's free and platform independent

it's the software of choice for many students, academics, industries and charities worldwide

highly flexible and extensive

encourages you think about your analyses



7

why you should learn 

it allows you to keep an exact and reproducible record of your analyses*

excellent graphics facilities*

might get you that job/post-doc

R is a community not just a bit of software

8



9

why you should learn 

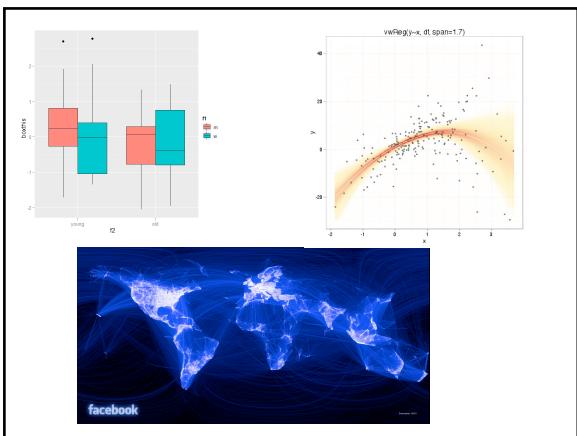
it allows you to keep an exact and reproducible record of your analyses

excellent graphics facilities*

might get you that job/post-doc

R is a community not just a bit of software

10



11

why you should learn 

it allows you to keep an exact and reproducible record of your analyses

excellent graphics facilities

might get you that job/post-doc

R is a community not just a bit of software

12

Using R - GUI

- fairly spartan GUI – very few menus or buttons

R is not menu driven

13

using R

- commands are typed (or sourced) into the console window at the > prompt

```
> 2+2
[1] 4
>
```

- R is object orientated. You can create variables and assign value(s) to them

```
> a<-2+2
> a
[1] 4
>
```

assign variables a value using the 'gets' operator

display variable value by typing variable name

14

using R

- once created, operations can be performed on variables

```
> a <- 2+2
> b <- 3*a
> a+b
[1] 10
```

- this is very powerful and flexible (as you will see)
- much of the functionality of R is enhanced by using variables called functions

15

using R

- functions contain a set of instructions that allow you to perform a specific task(s)
- you can use functions that are inbuilt in R (or R packages)

```
> numbers <- c(2,3,4,5,6)
> numbers
[1] 2 3 4 5 6
> mean(numbers)
[1] 4
> var(numbers)
[1] 2.5
```

concatenate

calculate the mean of variable numbers

calculate the variance of variable numbers

16

using R

- or write your own functions
- function to calculate standard error

```
> se.fnc <- function(x){ #start function for se
  std.x <- sd(x)           #calculate SD
  nos.x <- length(x)        #calculate number obs
  se.x <- std.x/(sqrt(nos.x))#calculate SE
  print(se.x)                #print SE
} #end of function
```

- using your new function

```
> se.fnc(numbers)
[1] 0.7071068
```

notice the use of comments

17

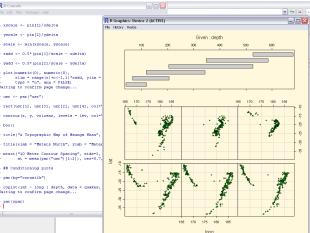
using R - syntax

- R is case sensitive 'A' is not the same as 'a'
- commands are generally separated by a new line, but you can also use a semicolon ;
- anything that follows the hash symbol (#) will be ignored by R. Use this to comment your code
- a series of commands can be grouped using braces { }
- don't worry too much about spaces
- previous commands can be recalled using ↑ and ↓ keys

18

using R

- some functions open other devices
 - i.e. many graphic functions open a plotting device
 - plots can be copied and pasted into word
 - or saved as a pdf, jpg etc
 - you can open more than one device



19

getting help in R

- R has extensive help facilities
- from within R the main method of getting help is to use the `help()` function

```
> help("plot")
or
?plot
to search for help

> help.search("plot")
or
??plot
```



20

getting help in R

- wealth of information on web
 - R-project web site
 - Rhelp mailing list with searchable archives
 - free pdfs from R-project website
 - www.Rseek.org
 - UoA Rusergroup mailing list

<https://mlisthost.abdn.ac.uk/mailman/listinfo/rusergroup>

21

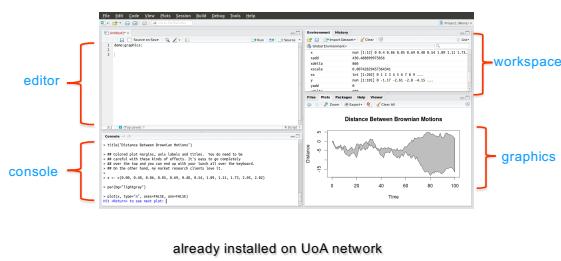
using a script editor: Tinn-R

- many script editors available (crimson, Emacs, WinEdt etc)
- Tinn-R is free and fully featured
- offers - syntax highlighting, R reference card, bracket matching, dynamic link with R for easy sourcing
- already installed on UoA network
- download from <http://nbcgib.uesc.br/lec/software/editores/tinn-r/en>

22

using an IDE: RStudio

- Using an Integrated Development Environment (IDE)
- RStudio (<http://www.rstudio.com/>)



23

keep scripting

- always use a script editor or IDE to write your R code rather than typing directly into console
- the more you use R the more complicated your code will become
- allows you to keep a permanent and reproducible record of your analyses

```
se.fnc <- function(x){
  std.x <- sd(x)
  nos.x <- length(x)
  se.x <- std.x/(sqrt(nos.x))
  print(se.x)
}
```

24

course survival guide (and beyond)

- keep a careful record of your code, analysis & plots
 - Tinn-R, Rstudio, MS word etc
- annotate your R code with plenty of comments (#)
- remember R has an extensive help facility
- ask plenty of questions
- start using R to explore and analyse your own data as soon as possible

25

course survival guide (and beyond)

- join the University of Aberdeen RUsergroup mailing list

<https://mlisthost.abdn.ac.uk/mailman/listinfo/rusergroup>

- invest in a good book (but see free pdfs first)

- have fun!

rusergroup October 2007 Archives by thread

• October 2007 messages sorted by [subject] [author] [date]
• Message to [list]
• List archive [Table of Contents](#)

• [Rmessage \(about inserting data and connecting single database\)](#) Douglas, Dr Alastair Mac, 23 Oct 2007 11:49:11+0100
• [Rmessage \(about rusergroup\)](#) Treloar, Jim, 20 Oct 2007 11:49:11+0100
• [Re: Rmessage \(about rusergroup\)](#) Douglas, Dr Alastair Mac, 20 Oct 2007 11:49:11+0100
• [Re: Rmessage \(about rusergroup, memory issue\)](#) Douglas, Dr Alastair Mac, 25 Oct 2007 13:33
• [Rmessage \(about inserting data into db\)](#) Gedde, L. E., 24 Oct 2007 13:33
• [Re: Rmessage \(about inserting data into db\)](#) Gedde, L. E., 24 Oct 2007 13:33
• [- Re: Rmessage \(about inserting data into db\)](#) Gedde, L. E., 24 Oct 2007 13:33
• [Re: Rmessage \(about inserting data into db\)](#) O'Brien, M. K., 24 Oct 2007 12:15:14+0100
• [Re: Rmessage \(about inserting data into db\)](#) Logie, C., 24 Oct 2007 11:52:39+0100
• [Re: Rmessage \(about inserting data into db\)](#) Logie, C., 24 Oct 2007 11:52:39+0100
• [Re: Rmessage \(about inserting data into db\)](#) Logie, C., 24 Oct 2007 11:52:39+0100
• [Re: Rmessage \(about inserting data into db\)](#) Treloar, Jim, 24 Oct 2007 11:52:39+0100
• [Re: Rmessage \(about inserting data into db\)](#) Douglas, Dr Alastair Mac, 24 Oct 2007 14:43:22+0100
• [Rmessage \(about inserting data into db\)](#) Treloar, Jim, 24 Oct 2007 14:43:22+0100
• [Rmessage \(about inserting data into db\)](#) Gedde, L. E., 24 Oct 2007 14:43:22+0100
• [Rmessage \(about inserting data into db\)](#) O'Brien, M. K., 24 Oct 2007 14:43:22+0100
• [Rmessage \(about inserting data into db\)](#) Treloar, Jim, 24 Oct 2007 14:43:22+0100

26

where to start?

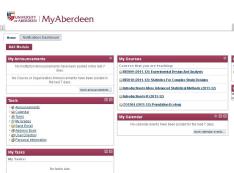
- work through sections of manual
- complete exercises for each section
- datafiles and exercises can be found at:

<https://abdn.blackboard.com>

Use your login ID and password

navigate to ‘An Introduction to R’

Click ‘Course Materials’



27