Using R as a Research Tool.

Susan Johnston: Susan.Johnston@ed.ac.uk
Demonstrators: Gergana Dalaskova, John Godlee
Hat-Tips: Kyle Dexter, Coding Club

November 1, 2017

1 Introduction

1.1 What is R?

R began its life in New Zealand in 1993 as a language and environment for statistical computing and graphics. It is an interpreted programming language, meaning that rather than pointing and clicking, the user instead types in commands. It is **free** and works across all platforms.

1.2 Why use R?

> `“This is R. There is no if. Only how.”`
> `-- Simon ‘Yoda’ Blomberg, R-help (April 2005)`

Almost anything is possible in R. It is fast becoming the *lingua franca* of academic research and data science. It is used for:

- Processing and tidying data
- Statistical analyses
- Data visualistion (`ggplot`)
- Creating interactive web applications (`shiny`)
- Generating reports and presentations (`knitr`, `slidify`)
- Creating portable projects (RStudio Projects)

The analytical power of R lies in its many packages (11,172 at the time of writing). At least 300 of these are written for ecologists and evolutionary biologists. A list of packages are hosted on the Comprehensive R Archive Network (known as **CRAN**): `https://cran.r-project.org/`.

1.3 What we hope to achieve in this session.

Before beginning this course, we asked you to complete the

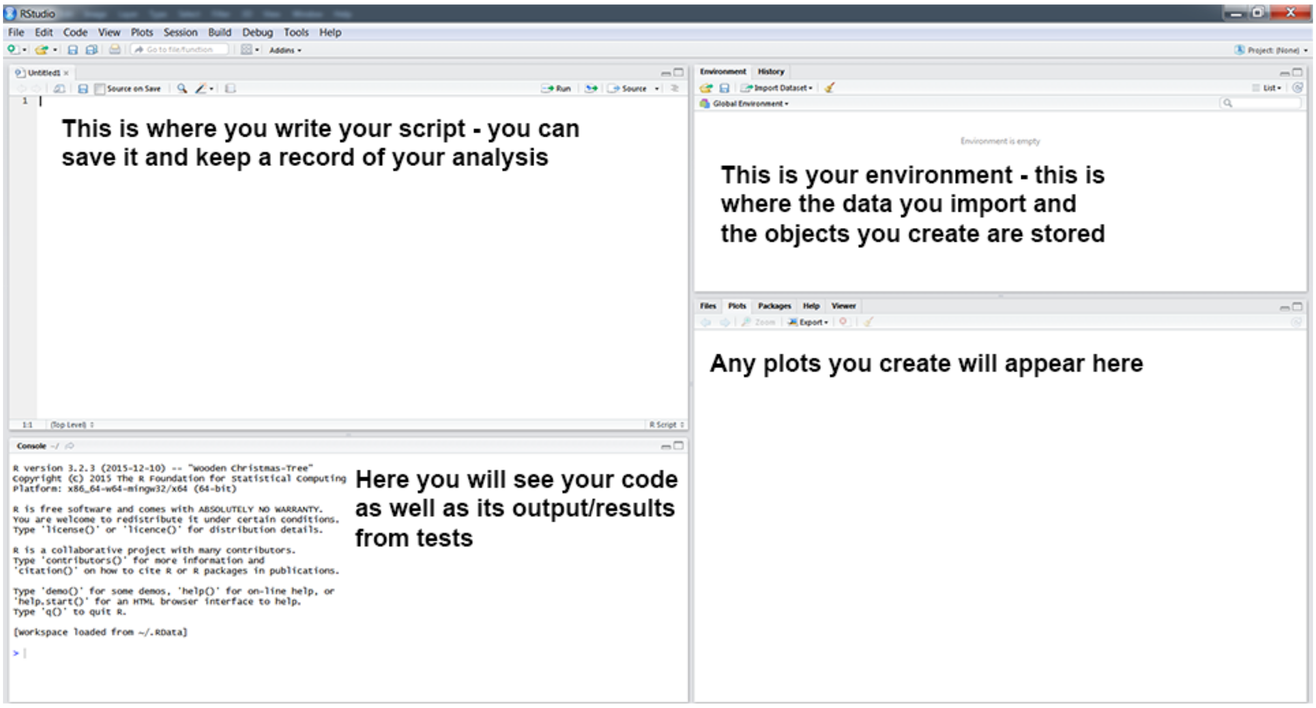
2 Getting Started: R and the RStudio Environment.

2.1 Installing R and RStudio.

R can be downloaded from the CRAN website. Whilst the CRAN download version provides a simple user interface, we recommend that R is run through the software RStudio. This is open-source, free, and available at `http://www.rstudio.com/`. [NB. It is important to install R first and RStudio second.]

2.2 Using RStudio.

Open a new R Script (`File > New File > R Script`) and save it. RStudio should look something like Figure 1.



This is where you write your script - you can save it and keep a record of your analysis

This is your environment - this is where the data you import and the objects you create are stored

Any plots you create will appear here

Here you will see your code as well as its output/results from tests

```
R version 3.2.3 (2015-12-10) -- "Wooden Christmas-Tree"
Copyright (C) 2015 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 ~/.RData]

>
```

Figure 1: The RStudio Environment. Taken from OurCodingClub.io

On the lower left is the *Console* pane - this is the engine of R. You can give instructions to R by directly typing at the prompt (>).

On the upper left is your new R Script - here, you can write commands and send them to the console by clicking “Run” or by typing `Ctrl-Enter` (or `Cmd-Enter`).

On the lower right, you can browse the packages installed on your machine, open files and search R Help. This pane will also show plots when we run them later in the practical.

Exercise 1.

Try running some basic commands directly in the console and from the R Script:

```
> 2+3

[1] 5

> 1:10

 [1]  1  2  3  4  5  6  7  8  9 10

> seq(1, 20, 4)

[1]  1  5  9 13 17

> mean(c(3, 6, 9, 3, 6, 7))

[1] 5.666667
```

Let’s assign a sequence of numbers to an object, x:

```
> x <- 1:10
> x

 [1]  1  2  3  4  5  6  7  8  9 10

> y <- seq(0, 4.5, 0.5)
> y

 [1] 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5
```

You can see that in the upper right pane, we can see this new objects x and y in the environment.

2.3 Finding Help with R.

The fastest way to find help in R is to search using `?`. For example:

```
> ?mean
```

should bring up a help page for the function `mean()` in the lower right corner. Typing two question marks:

```
> ??mean
> ??"standard error"
```

will search all help files and return a list of those that match.

Exercise 2.

1. Using only `?` and/or `??`, find a function for calculating the standard deviation. What is the standard deviation of `x`?
2. Using `?`, find the help file for the `sort()` function. Sort `x` and `y` in reverse order.

2.4 Troubleshooting and finding help outside of R.

- Coding Club Tutorials & Useful Links `https://ourcodingclub.github.io/`
- Stack Overflow `https://stackoverflow.com/`: Try searching with the tag [R]
- Google
- RStudio Cheatsheets `https://www.rstudio.com/resources/cheatsheets/`

3 Loading and Preparing Data.