

Week 2 - Introduction to R

Andrew Stewart

Andrew.Stewart@manchester.ac.uk



@ajstewart_lang



<https://github.com/ajstewartlang>

Week	Topic
1	Introduction, Open Science, and Power
2	Introduction to R
3	Data Wrangling and Visualisation
4	General Linear Model - Regression
5	General Linear Model - Regression
6	No Timetabled Lecture - Reading Week
7	Consolidation Lab
8	General Linear Model - ANOVA
9	General Linear Model - ANOVA
10	Tidy Thursday Data Wrangling & Visualisation Challenge
11	Reproducing your Computational Environment using Binder
12	Dynamic, Reproducible Presentations Using xaringan

Semester 1 Assignments

Data wrangling and visualisation – Due December 5th

ANOVA/ANCOVA – Due January 17th

Last week...

- We learned about Open Science and how research in Psychology is changing as a result of the replication crisis across the biosciences.
- Good scientists can engage in questionable research practices (QRPs) without being aware of it...
- *p*-hacking, HARK-ing, low power, publication bias etc. have all contributed to the current crisis.
- Psychological research is at the start of a renaissance period.

- We looked at the issue of Power and noted that under-powered experiments (with too few participants) can be misleading because:
 - They increase the probability of a Type II error (failing to detect an effect even though one is present).
 - Increase the probability of a Type I error (thinking we have an effect when we don't - this is often due to sampling error).
- We also looked at *p*-values and realised that *p*-values on their own tell us very little. We also need to know about *effect sizes*...

By the end of today you will...

- Understand the basics of how to interact with the RStudio environment.
- Understand the structure of data frames (and how you can start to manipulate them), how graphing functions like `ggplot` work with data frames, and how to plot some basic graphs.
- This will give you the foundation for tidying your data, visualising your data, and building models (e.g., AN(C)OVA (GLM), moderation and mediation in regression, and linear mixed models).

Why R?

- R allows you to engage in reproducible research.
- Statistical packages in R reflect the latest advances in the fields of Statistics and Data Science.
- Advanced models such as Linear Mixed Models (LMMs) are easy to build in R - many of the best journals now require LMMs to be used as they are more powerful and flexible than classical techniques such as ANOVA.
- Lots of amazing data visualisation packages available.
- In many institutions, the next generation of academics (M-level and PhD students, post-docs etc.) are learning R skills as part of their training.
- Plays an important role in Open Science.

What role can R play in Open Science?

- R scripts are easy to share allowing for reproducibility and easy public sharing of data and code.
- R is free, open source software that is much more flexible and powerful than SPSS.
- There is an active R community continuously updating statistical tests and packages that run in R.
- As R is a programming language, it forces you to know your data.

R vs. SPSS

“SPSS is like a bus - easy to use for the standard things, but very frustrating if you want to do something that is not already pre-programmed.

R is a 4-wheel drive off-roader, with a bike on the back, a kayak on top, good walking and running shoes in the passenger seat, and mountain climbing and spelunking gear in the back.

R can take you anywhere you want to go if you take time to learn how to use the equipment, but that is going to take longer than learning where the bus stops are in SPSS.” (Greg Snow, 2010, stackoverflow.com).

In meme form...

If statistics programs/languages were cars...

Image credit Darren Dahly @statsepi



O'REILLY®



R for Data Science

VISUALIZE, MODEL, TRANSFORM, TIDY, AND IMPORT DATA

Hadley Wickham &
Garrett Grolemund

Available electronically
for free at:

<http://r4ds.had.co.nz>

The R Series

Advanced R



Hadley Wickham

 CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Available electronically for
free at:

<https://adv-r.hadley.nz>

“Hadley Wickham, the Man Who Revolutionized R”



Chief Scientist at RStudio, author of key R packages incl. `ggplot2`, `tidyverse`, `dplyr` - all components of the tidyverse.

R User Groups...

R user groups exist all over the world with many in the UK (incl. Manchester)...

You can even buy the t-shirt...

UK

England

- Birmingham: [BRUM](#)
- Canterbury: [CanterburyR](#)
- Cambridge: [CambR](#)
- Coventry: [Warwick R User Group](#); [@WarwickRUG](#)
- Exeter: [Exeter R Users Group](#)
- London: [LondonR](#)
- Manchester: [ManchesterR](#)
- Newcastle: [R North East](#); [@RstatsNE](#)
- Nottingham: [Nottingham R User group](#)
- Oxford: [R user group Oxford](#); [@rusersoxford](#)
- Sheffield: [SheffieldR](#); [@Sheffield_R_](#)



RStudio
\$17



magrittr
\$17



Shiny
\$17



I like the way you R
\$17

...and conferences

WHO PLAN VENUE SPONSORS

R

rOpenSci Unconf

May 25 - 26 2017 • Los Angeles • CA



useR!2017
BRUSSELS

04.07.2017 - 07.07.2017



NEWS ABOUT REGISTRATION PROGRAM VENUE FAQ EXTRA



FOLLOW US ON

CONFERENCE BROCHURE

NEWS

LAST CALL

Only 15 spots left for the useR!2017 Conference in Brussels next week! Be quick! Go to <https://user2017.brussels/> to register! ...

NEWS

Conference Brochure online!

The useR!2017 Conference Brochure is online now! Take a look and find an answer to all your practical questions. A printed conference brochure will be ...

NEWS

useR!2017 App live now!

We are very proud to announce that our useR!2017 app is live! Discover the full schedule, speakers and the venue on the go! Download the app here: iOS ...

NEW YORK R CONFERENCE SPEAKERS AGENDA SPONSORS VIDEOS ▾

NYR

NEW YORK R CONFERENCE

BROUGHT TO YOU BY [LANDER ANALYTICS](#) AND [WORK-BENCH](#)

useR!

useR! 2018

THE CONFERENCE FOR USERS OF R
JULY 10-13, 2018.
BRISBANE, AUSTRALIA.



HOME PROGRAMME ▾ REGISTRATION ▾ CODE OF CONDUCT NEWS TRAVEL ▾ FAQ CONTACT

R skills are in high demand...

February 11, 2014

R skills attract the highest salaries

Two recent salary surveys have shown that [R language](#) skills attract median salaries in excess of \$110,000 in the United States.

In the [2014 Dice Tech Salary Survey](#) of over 17,000 technology professionals, the highest-paid IT skill was R programming. While [big-data skills in general featured strongly](#) in the top tier, having R at the top of the list reflects the strong demand for skills to make sense of, and extract value from big data.

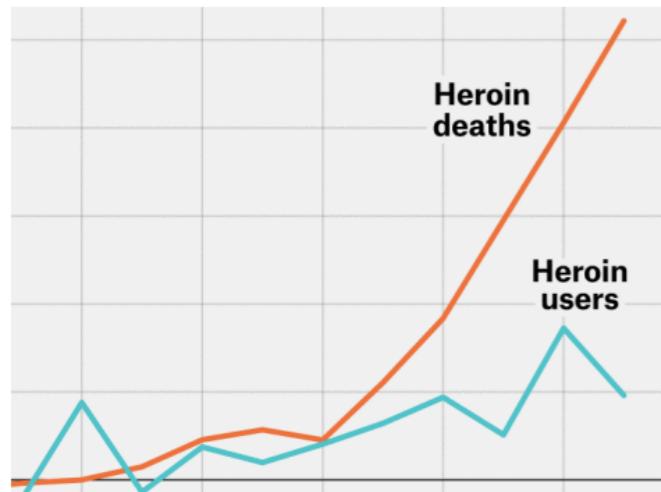
AVERAGE SALARY FOR High Paying Skills and Experience		
SKILL	2013	YR/YR CHANGE
R	\$ 115,531	n/a
NoSQL	\$ 114,796	1.6%
MapReduce	\$ 114,396	n/a
PMBok	\$ 112,382	1.3%
Cassandra	\$ 112,382	n/a
Omnigraffle	\$ 111,039	0.3%
Pig	\$ 109,561	n/a
SOA (Service Oriented Architecture)	\$ 108,997	-0.5%
Hadoop	\$ 108,669	-5.6%
Mongo DB	\$ 107,825	-0.4%

Similarly, the recent [O'Reilly Data Scientist Survey](#) also found R skills amongst those that pay in the \$110,000-\$125,000 range (albeit amongst a much smaller and specialized sample of respondents).

and R is widely used by a number of organisations (incl. the ONS)...

FiveThirtyEight

Politics Sports Science & Health Economics Culture



OPIOIDS
Data On Drug Use Is Disappearing Just When We Need It Most

By Kathryn Casteel

FEATURES

THE LATEST

JUN. 30 Why Republicans Might Be Forced To Oppose Tax Cuts

JUN. 29 Data On Drug Use Is Disappearing Just When We Need It Most

JUN. 28 The Health Care System Is Leaving The Southern Black Belt Behind

JUN. 26 The New CBO Report On Health Insurance Didn't Do Republicans Any Favors

JUN. 26



Most recent: Politics podcast

INTERACTIVES

35 Years Of American Death



YouGov UK

TAKE PART

SEE RESULTS

SOLUTIONS

Login

+ Join

How the YouGov model for the 2017 General Election works



By Douglas Rivers is a professor of political science at Stanford University and Chief Scientist at YouGov PLC.

In General Election 2017, Politics

On May 31, 2017, 6 a.m.

Share



YouGov ElectionCentre

The 2017 UK General Election

Doug Rivers, YouGov's chief scientist, sets out how YouGov's 2017 General Election model works

Follow @YouGov on twitter and stay up to date with the latest news and results



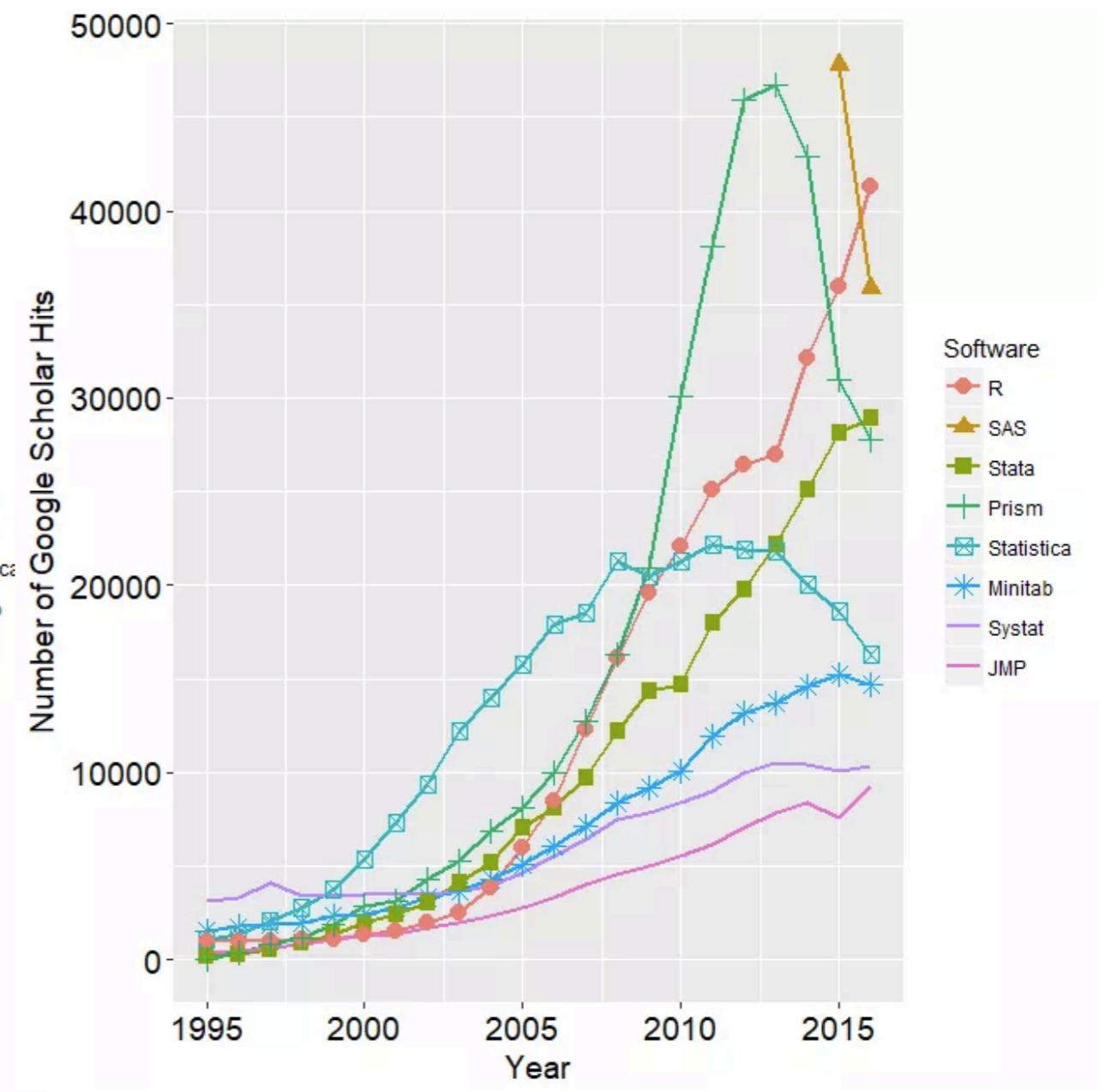
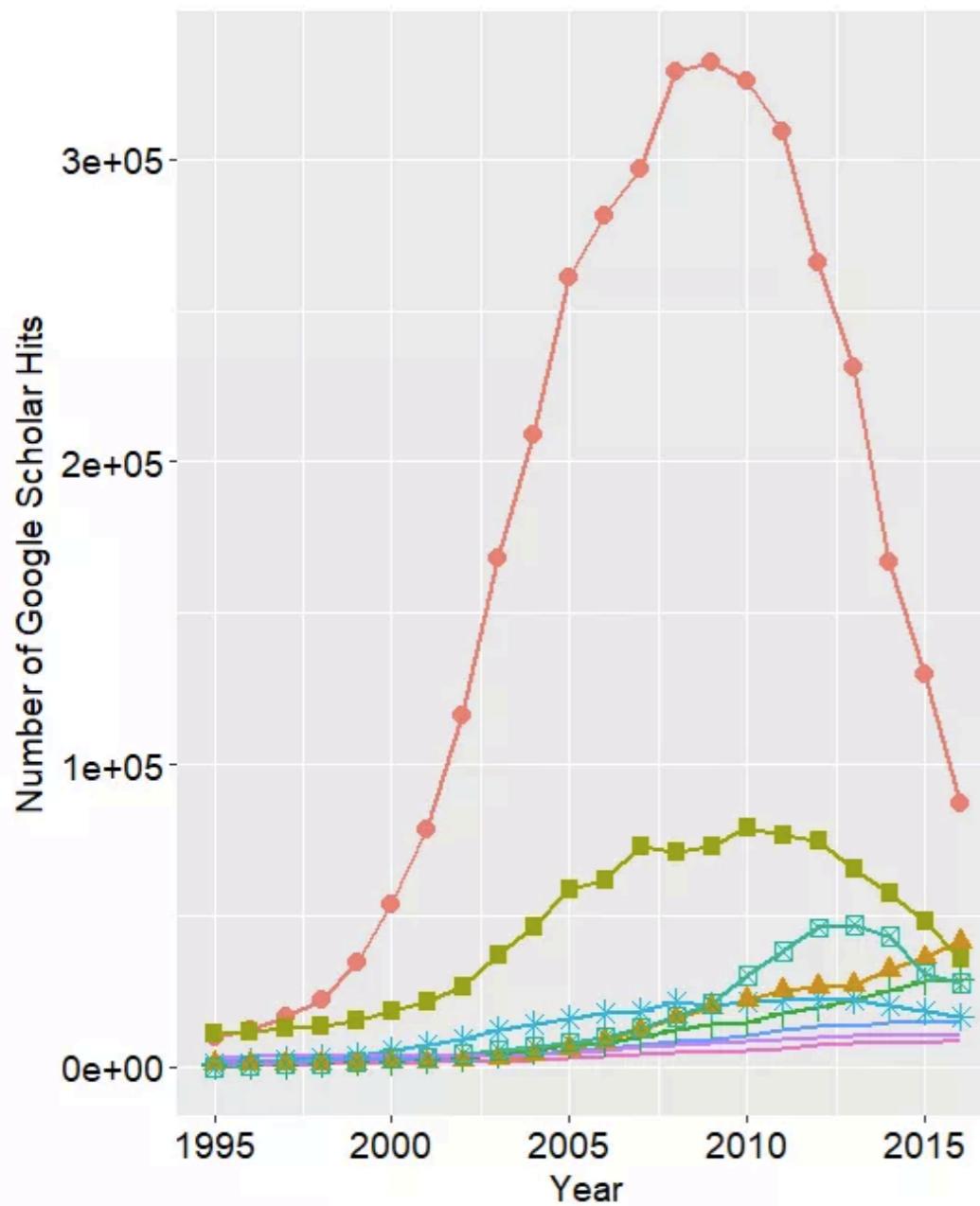
How did 2015 voters cast their ballot at the 2017 general election?

It wasn't just the kids that boosted Labour

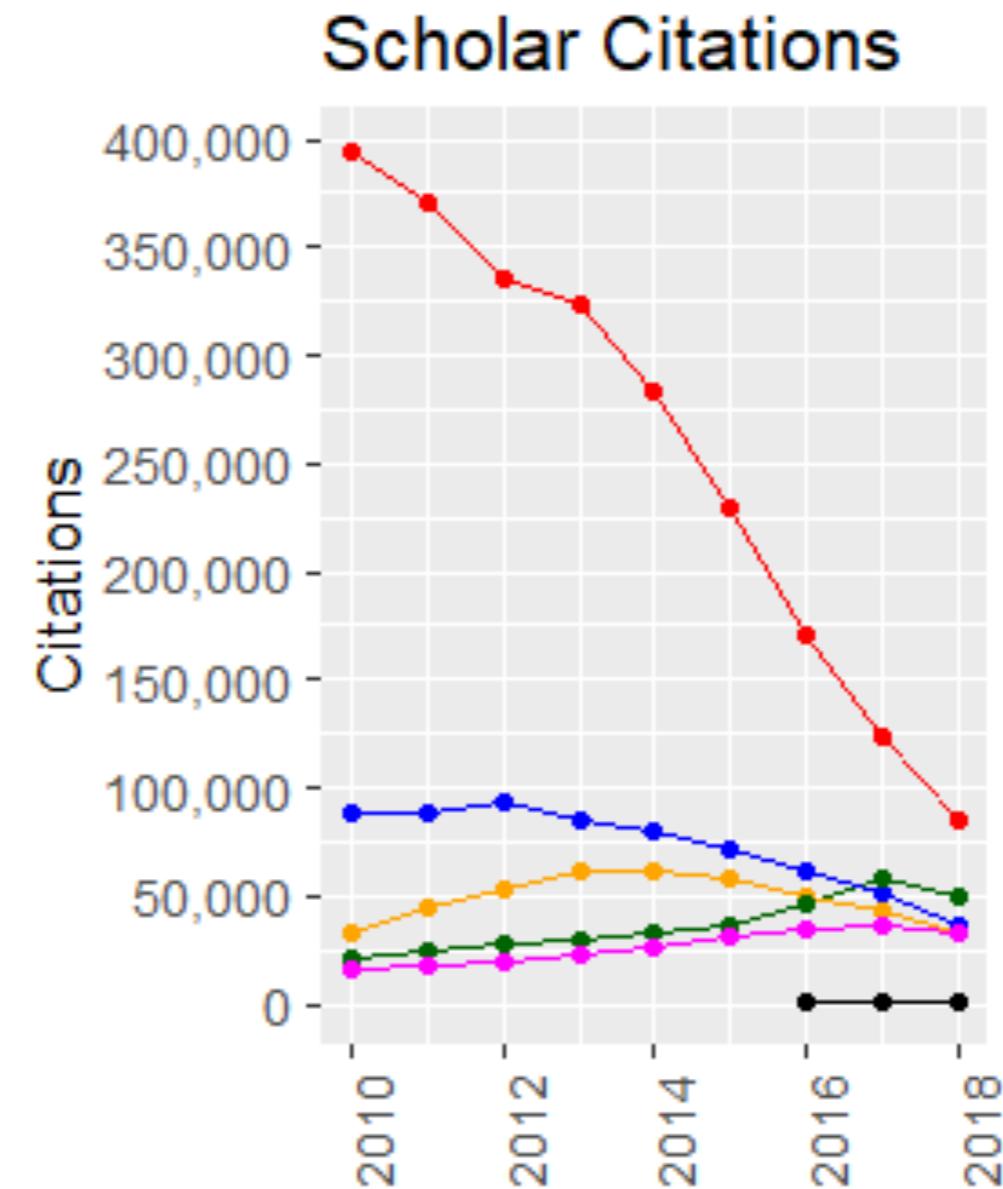
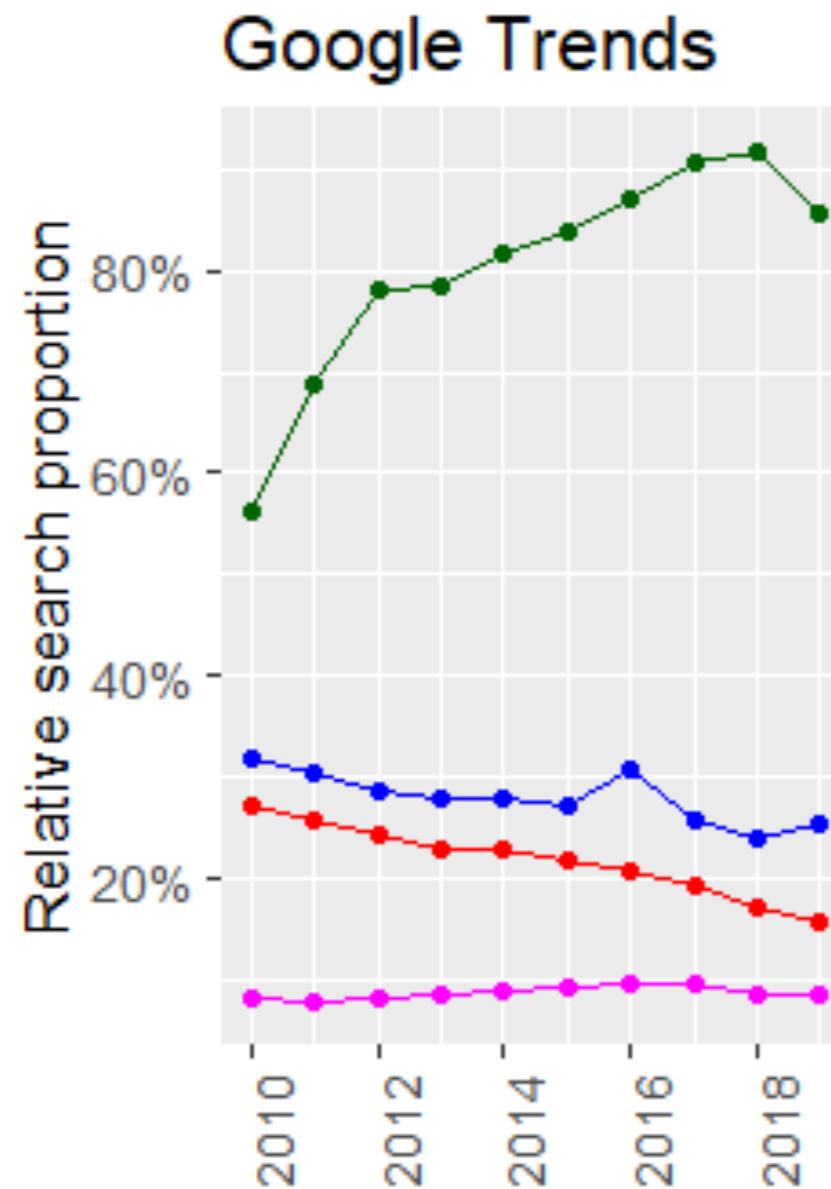
John Humphrys - A Government Adrift?

Theresa May is now almost as unpopular as pre-campaign Corbyn

and across academia...



Google Search Trends and citations in the academic literature.



software

- SPSS
- R
- SAS
- STATA
- JASP
- Prism

with Data Science a growing employment destination for Psychologists.



AMERICAN PSYCHOLOGICAL ASSOCIATION

ABOUT APA

TOPICS

PUBLICATIONS & DATABASES

PSYCHOLOGY HELP CENTER

NEWS & EVENTS

SCIENCE

[Home](#) // [gradPSYCH Magazine](#) // [January 2013 gradPSYCH](#) // Hot jobs: Big-data psychologists

CAREER CENTER

Hot jobs: Big-data psychologists

Wanted: Scientists who can help companies make sense of consumer and employee data.



By Rebecca Voelker

Print version: page 18

Every time you use an Internet search engine, sign up for a company "rewards" program or swipe your credit card, that information is saved and stored. The industries that amass these billions of bytes of data are increasingly hiring psychologists to help make sense of it, says Douglas Reynolds, PhD, president of APA's Div. 14 ([Society for Industrial and Organizational Psychology](#)).

"Big data, and what it means for business, is a hot topic right now," says Reynolds, who also serves as vice president of assessment technology at [Development Dimensions International Inc.](#), a human resources consulting firm. "We're in high demand."

Psychologists' ability to interpret numbers and human behavior makes them key members of many industry analytics teams, adds Suzie Weaver, PhD, a psychologist and senior analytic consultant with [Epsilon](#), a global marketing and analytics company. "Analytics is at the core of everything we do, whether it's in research, the academic sphere or the business world."

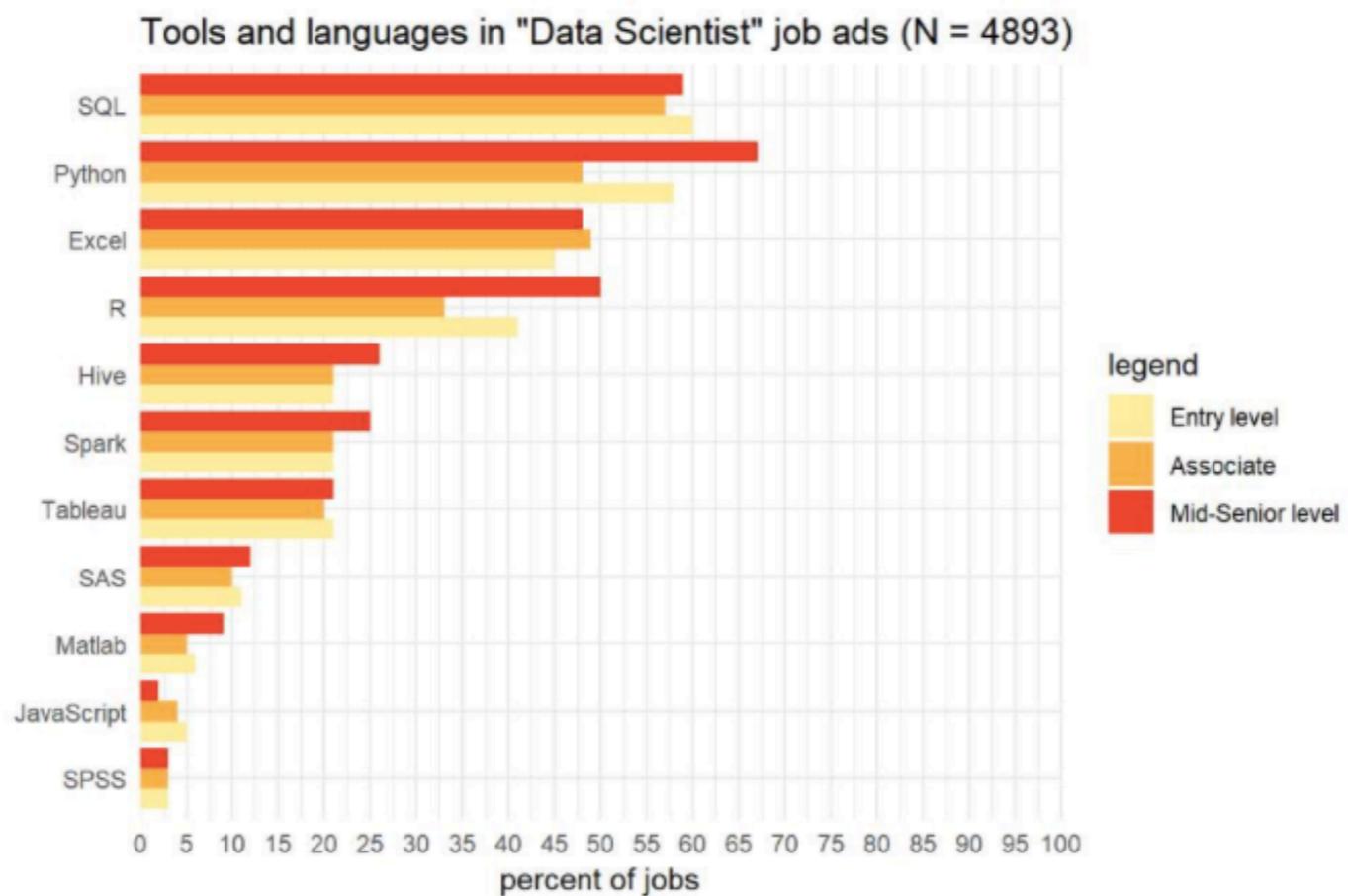
Tyler Burleigh
@tylerburleigh

▼

If your department teaches quantitative skills in SPSS, SAS, or Matlab, is it really preparing students for quantitative jobs outside of academia? [#datascientist](#) [#academia](#) [#rstats](#) [#python](#)

Data scientist jobs

4893 data scientist job ads were included in this analysis.



Jobs were scraped from LinkedIn in Sept 2019
Locations included NYC, SF, Seattle, Boston, and Toronto
Positions were Entry, Associate, and Senior levels



Nick Corriveau-Lecavalier
@CorriveauNick

#brainhackschool instructors be like: familiar with docker?

Me: 😬

Familiar with jupyter?

Me: 😬

Familiar with github?

Me: 😬

Familiar with binder?

Me: 😬

Familiar with python?

Me: 😬

It feels like I've spent these past 4 years of PhD on mars



7:05 PM · Aug 7, 2019 · Twitter for iPhone

How to create BBC style graphics

Load all the libraries you need

Install the bbplot package

How does the bbplot package work?

Save out your finished chart

Make a line chart

Make a multiple line chart

Make a bar chart

Make a stacked bar chart

Make a grouped bar chart

Make a dumbbell chart

Make a histogram

Make changes to the legend

Make changes to the axes

Add annotations

Work with small multiples

Do something else entirely

BBC Visual and Data Journalism cookbook for R graphics

Last updated: 2019-01-24

How to create BBC style graphics

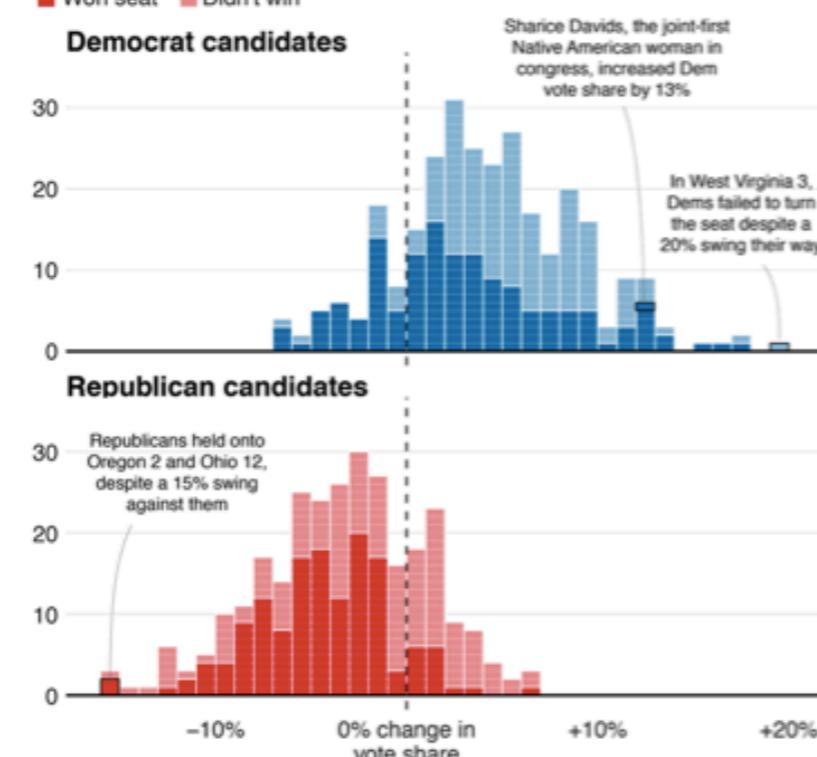
At the BBC data team, we have developed an R package and an R cookbook to make the process of creating publication-ready graphics in our in-house style using R's ggplot2 library a more reproducible process, as well as making it easier for people new to R to create graphics.

The cookbook below should hopefully help anyone who wants to make graphics like these:

Blue wave

■ Won seat ■ Didn't win

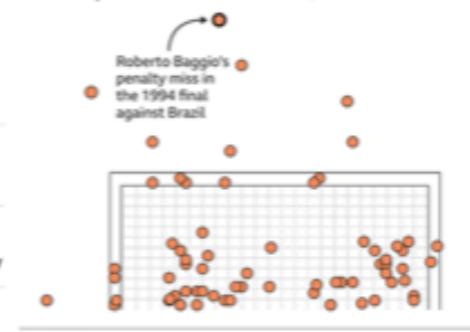
Democrat candidates



Source: AP, 19:01 ET

Where penalties are saved

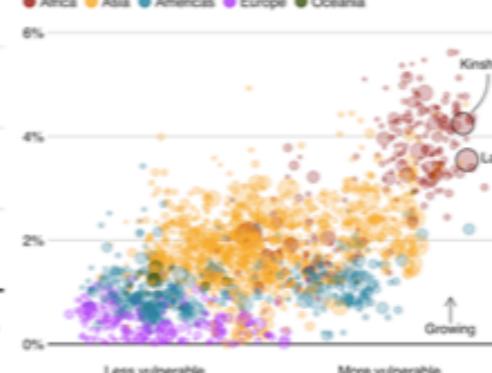
World Cup shootout misses and saves, 1982-2014



Fast-growing cities face worse climate risks

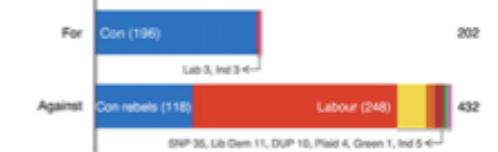
Population growth 2018-2035 over climate change vulnerability

■ Africa ■ Asia ■ Americas ■ Europe ■ Oceania



Source: Verisk Maplecroft. Circle size represents current population.

MPs rejected Theresa May's deal by 230 votes



Source: Commons Votes Services. Excludes 'Yellars', the Speaker and deputies

Earnings vary across unis even within subjects

Impact on men's earnings relative to the average degree



Source: Institute for Fiscal Studies

<https://medium.com/bbc-visual-and-data-journalism/how-the-bbc-visual-and-data-journalism-team-works-with-graphics-in-r-ed0b35693535>

Handy list of Psychology groups that teach R, plus links to course materials - list compiled by Andy Wills at Plymouth.

[View on GitHub](#)

rminr

Research Methods in R

Teaching Research Methods in R

This is a crowd-sourced list of uses of R to teach research methods in Psychology, and a link to Creative Commons teaching materials, where these are available. The year teaching in R was adopted at undergraduate and postgraduate level is also recorded, where known. Where there are no materials, but the organization's name has a link, this is a link to evidence that R is used.

If you'd like to add to this list, please submit a [pull request](#). Or, if you're not sure how to do that, just email me: andy@willslab.co.uk

Universities

University	Country	UG	PG	Link
Harrisburg University of Science and Technology	U.S.A.		2018	PG
Missouri State	U.S.A.		2017	PG
Nottingham Trent University	U.K.	2012	2010	
University of Edinburgh	U.K.	2018	2018	
University of Glasgow	U.K.	2015	2010	UG, PG
University of Lancaster	U.K.		2014	
University of Lincoln	U.K.		2018	PG
University of Manchester	U.K.		2018	PG
University of Plymouth	U.K.	2018 (Year 1) - 2020 (Year 3)	2017	UG, PG
University of Sussex	U.K.	2019		

<https://ajwills72.github.io/rminr/rminrinpsy.html>

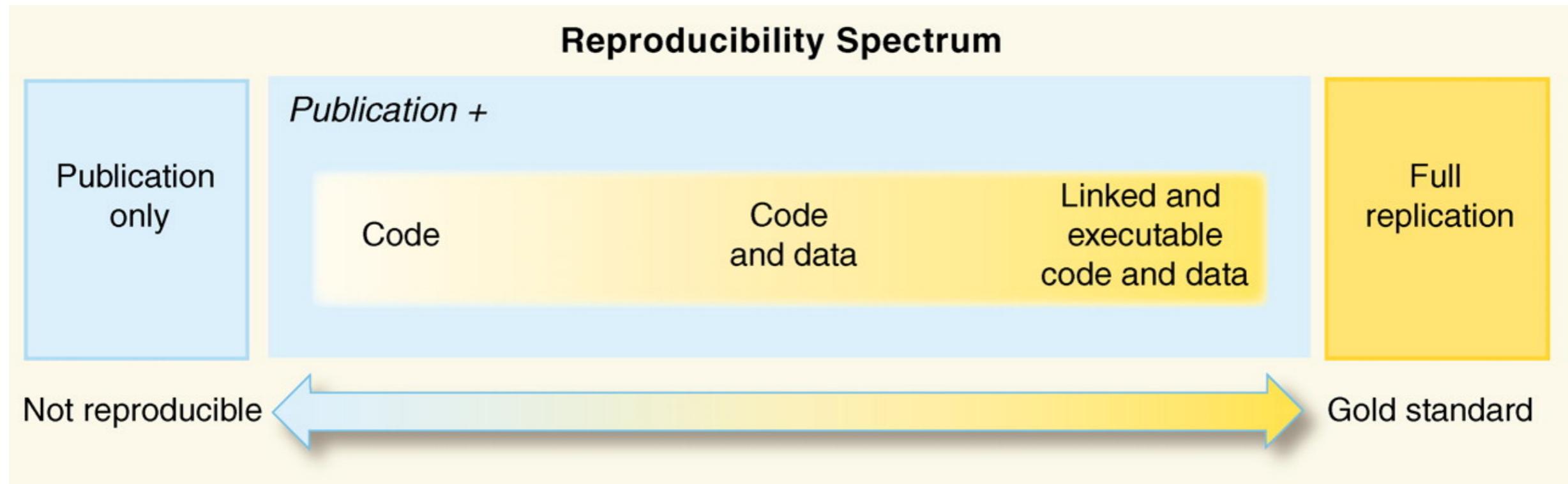
PERSPECTIVE

Reproducible Research in Computational Science

Roger D. Peng

[+ See all authors and affiliations](#)

Science 02 Dec 2011;
Vol. 334, Issue 6060, pp. 1226-1227
DOI: 10.1126/science.1213847



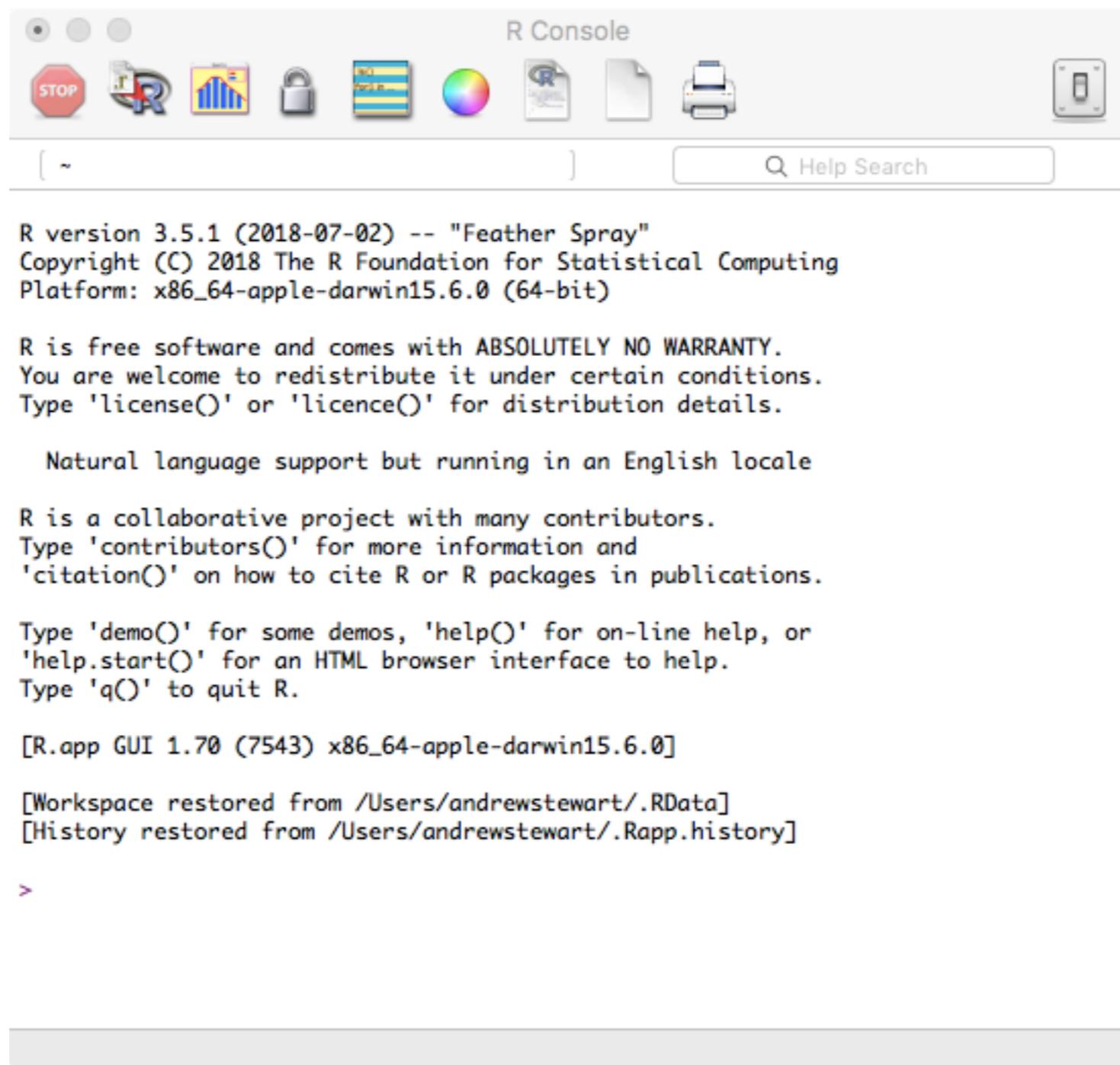
"You can't do data science in a GUI", Hadley Wickham

Starting R

- R is open source (i.e., free to download and add to). Main R site is:
- *www.r-project.org*
- From here you can download R (from one of the CRAN¹ mirrors for Windows, Mac and UNIX).
- R updates regularly (you need to update manually).

¹CRAN = *The Comprehensive R Archive Network*

- When you first load R it looks like this:



R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (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.

Natural language support but running in an English locale

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.

[R.app GUI 1.70 (7543) x86_64-apple-darwin15.6.0]

[Workspace restored from /Users/andrewstewart/.RData]
[History restored from /Users/andrewstewart/.Rapp.history]

>



- Rather than using the R interface, you should use the RStudio graphical interface.
- Download it from www.rstudio.com
- When you use RStudio, it looks like this:

RStudio File Edit Code View Plots Session Build Debug Tools Window Help

~/Desktop/Air Work/MRes 2016:17/R workshop MRes/R workshop directory/Workshop - RStudio

Addins

Workshop script2.R * DV

```

1 library(lmerTest)
2 library(lsmeans)
3 library(pbkrtest)
4
5 #Read in First Pass Data
6 FPs <- read.csv("~/Desktop/Air Work/R analyses/Indirect Request Expt/Experiment 1 - probability of success - Libby's data/FPs")
7
8 #this sets up the contrasts so that the intercept in the mixed LMM is the grand mean (i.e., the mean of all conditions)
9 my.coding <- matrix(c(.5, -.5))
10
11 contrasts(DV$Context)<-matrix(c(.5, -.5))
12 contrasts(DV$Sentence)<-matrix(c(.5, -.5))
13
14 #construct the models with crossed random effects for subjects and items for the pre-critical, critical and post-critical regions
15 model.full <- lmer(RT ~ Context*Sentence + (1+Context*Sentence | Subject) + (1+Context*Sentence | Item), data=DV, REML=TRUE)
16 model.null <- lmer(RT ~ (1+Context*Sentence | Subject) + (1+Context*Sentence | Item), data=DV, REML=TRUE)
17
18 summary(model.full)
19 lsmeans(model.full, pairwise~Context*Sentence, adjust="none")
20
21 model.full <- lmer(RT ~ Statement*Meaning*Probability + (1:Meaning*Probability | P_c) + (1:Meaning*Probability | Item), data=FPs)
22
23 (Top Level) ▾
  
```

Console ~ /Desktop/Air Work/MRes 2016:17/R workshop MRes/R workshop directory/Workshop /

The following object is masked from 'package:stats':

```

step
  
```

```

> library("lsmeans", lib.loc="/Library/Frameworks/R.framework/Versions/3.3/Resources/library")
Loading required package: estimability

Attaching package: 'lsmeans'

The following object is masked from 'package:lmerTest':
  
```

```

lsmeans
  
```

```

> model.full <- lmer(RT ~ Context*Sentence + (1+Context*Sentence | Subject) + (1+Context*Sentence | Item), data=DV, REML=TRUE)
  
```

Error in strsplit(keys, " * ") : non-character argument

Environment History

Import Dataset

Global Environment

Data

	DV	1680 obs. of 5 variables
my.coding	num [1:2, 1]	0.5 -0.5

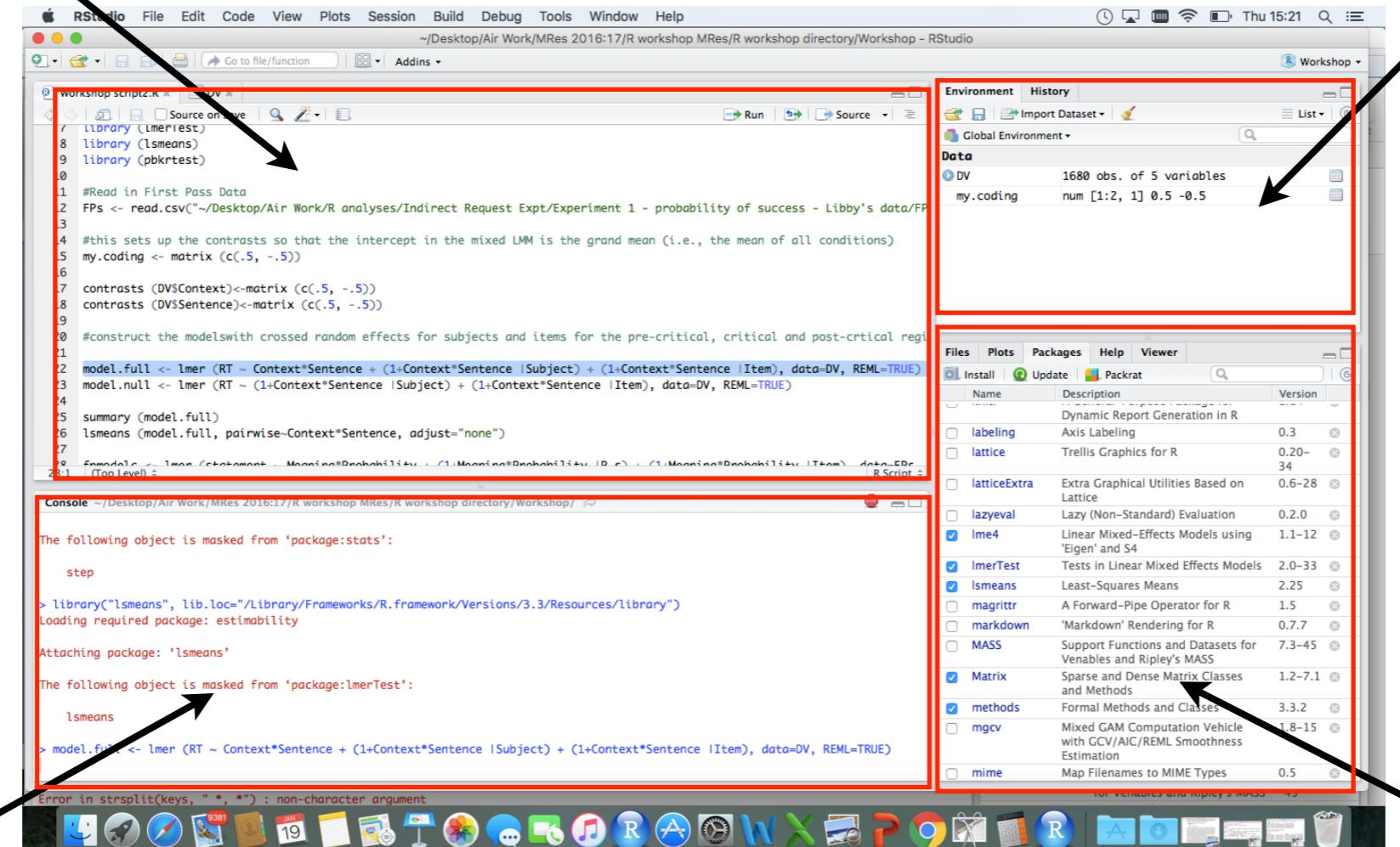
Files Plots Packages Help Viewer

Install Update Packrat

Name	Description	Version
lme4	Linear Mixed-Effects Models using 'Eigen' and S4	1.1-12
lmerTest	Tests in Linear Mixed Effects Models	2.0-33
lsmeans	Least-Squares Means	2.25
MASS	Support Functions and Datasets for Venables and Ripley's MASS	7.3-45
Matrix	Sparse and Dense Matrix Classes and Methods	1.2-7.1
methods	Formal Methods and Classes	3.3.2
mgcv	Mixed GAM Computation Vehicle with GCV/AIC/REML Smoothness Estimation	1.8-15
mime	Map Filenames to MIME Types	0.5

for Venables and Ripley's MASS 45

This is where you build your script and where data can be seen.

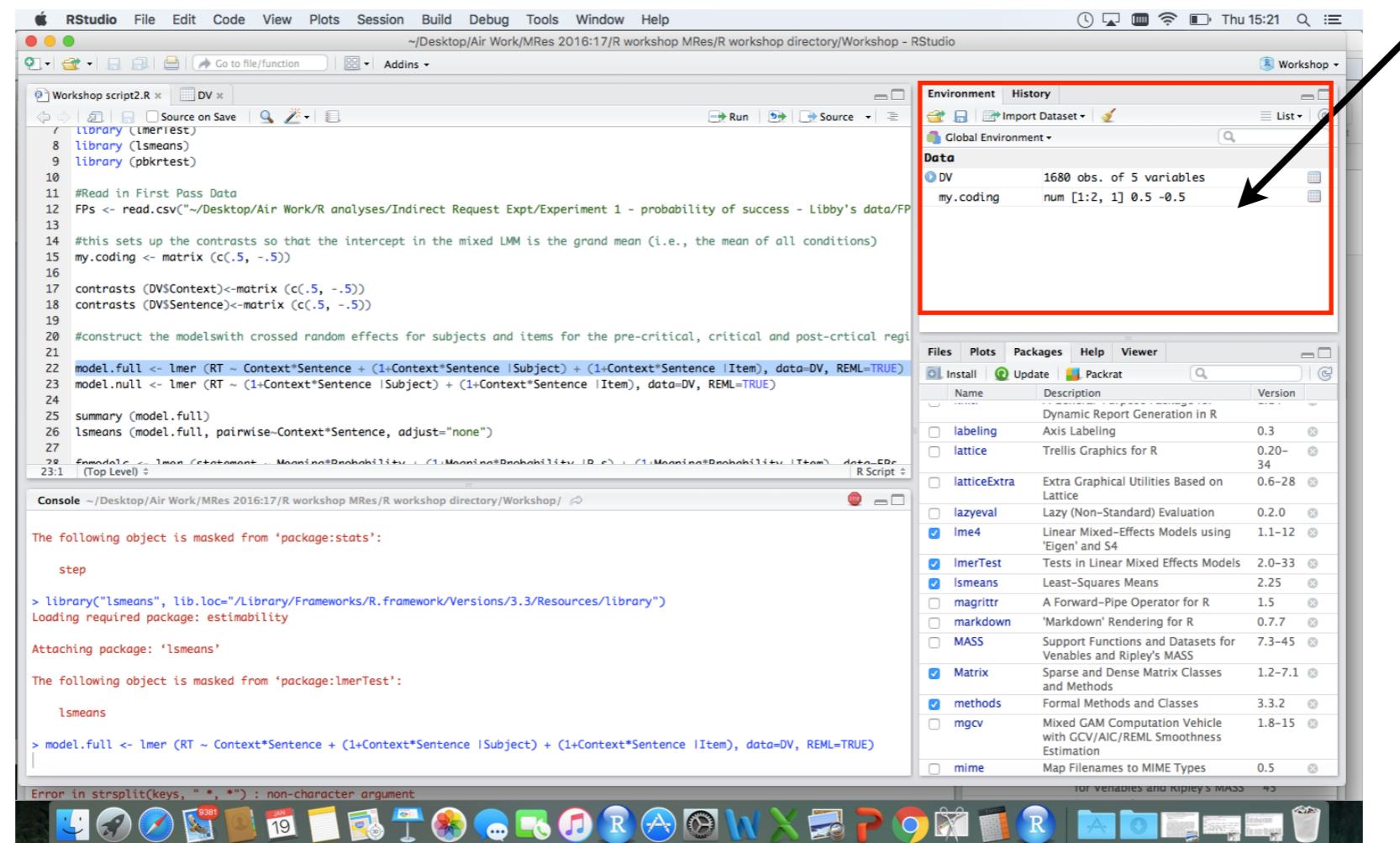


This is where you type commands.

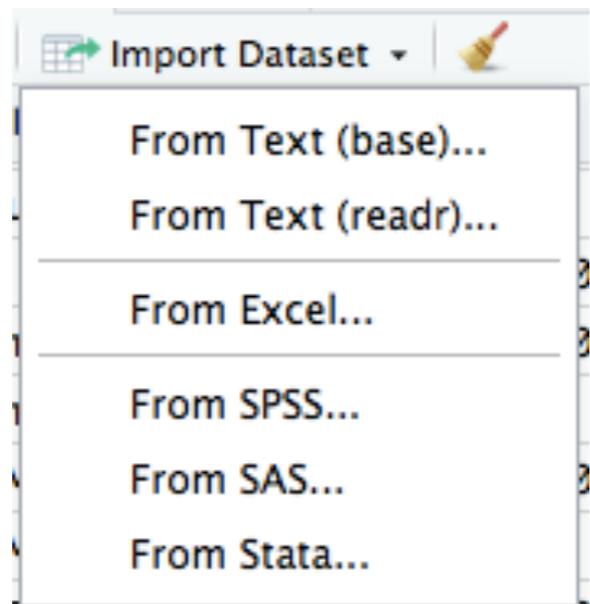
This contains information about your variables and open data sets.

This lists the packages you have loaded, and has tabs for help, graphs etc.

Here is where you import your data.



- Importing data (CSV (Text), Excel, SPSS, SAS and Stata format). CSV can be delimited by commas, tabs etc. Most of the time you'll use the `readr` import function...



Import Text Data

File/Url:

~/Desktop/Air Work/R analyses/R courses/R course/R Work Folder/priming data.txt

Data Preview:

Subject (integer) ▾	Priming Condition (character) ▾	RT (integer) ▾
1	LONG	500
2	LONG	551
3	LONG	479
4	LONG	561
5	LONG	522
6	SHORT	399
7	SHORT	423
8	SHORT	444
9	SHORT	410
10	SHORT	398

You can change the type of each variable by clicking on the down arrow.

You should change this to type Factor (LONG vs. SHORT).

Previewing first 50 entries.

Import Options:

Name: priming_data First Row as Names Delimiter: Tab Escape: None
Skip: 0 Trim Spaces Quotes: Default Comment: Default
 Open Data Viewer Locale: Configure... NA: Default

Code Preview:

```
library(readr)
priming_data <- read_delim("~/Desktop/Air Work/R analyses/R courses/R course/R Work Folder/priming data.txt",
  "\t", escape_double = FALSE, trim_ws = TRUE)
View(priming_data)
```

Can change here how the columns are delimited.

This is the code that corresponds to what you're getting RStudio to do.

- RStudio now looks like this:

The screenshot shows the RStudio interface with the following components:

- Data View:** Displays a data frame named "priming_data" with 10 observations and 3 variables: Subject, Priming Condition, and RT.
- Environment View:** Shows the global environment with "priming_data" listed.
- Console View:** Displays the R session history, including the command to read the data and the resulting error message.
- Packages View:** Shows a list of installed packages.

```

> 
> 
> 
> 
> 
> 
> 
> RTdata <- priming_data [c("Priming Condition", "RT")]
Error: object 'priming_data' not found
> library(readr)
> priming_data <- read_delim("~/Desktop/Air Work/R analyses/R courses/R course/R Work Folder/priming data.txt",
+   "\t", escape_double = FALSE, trim_ws = TRUE)
Parsed with column specification:
cols(
  Subject = col_integer(),
  `Priming Condition` = col_character(),
  RT = col_integer()
)
> View(priming_data)
> 
  
```

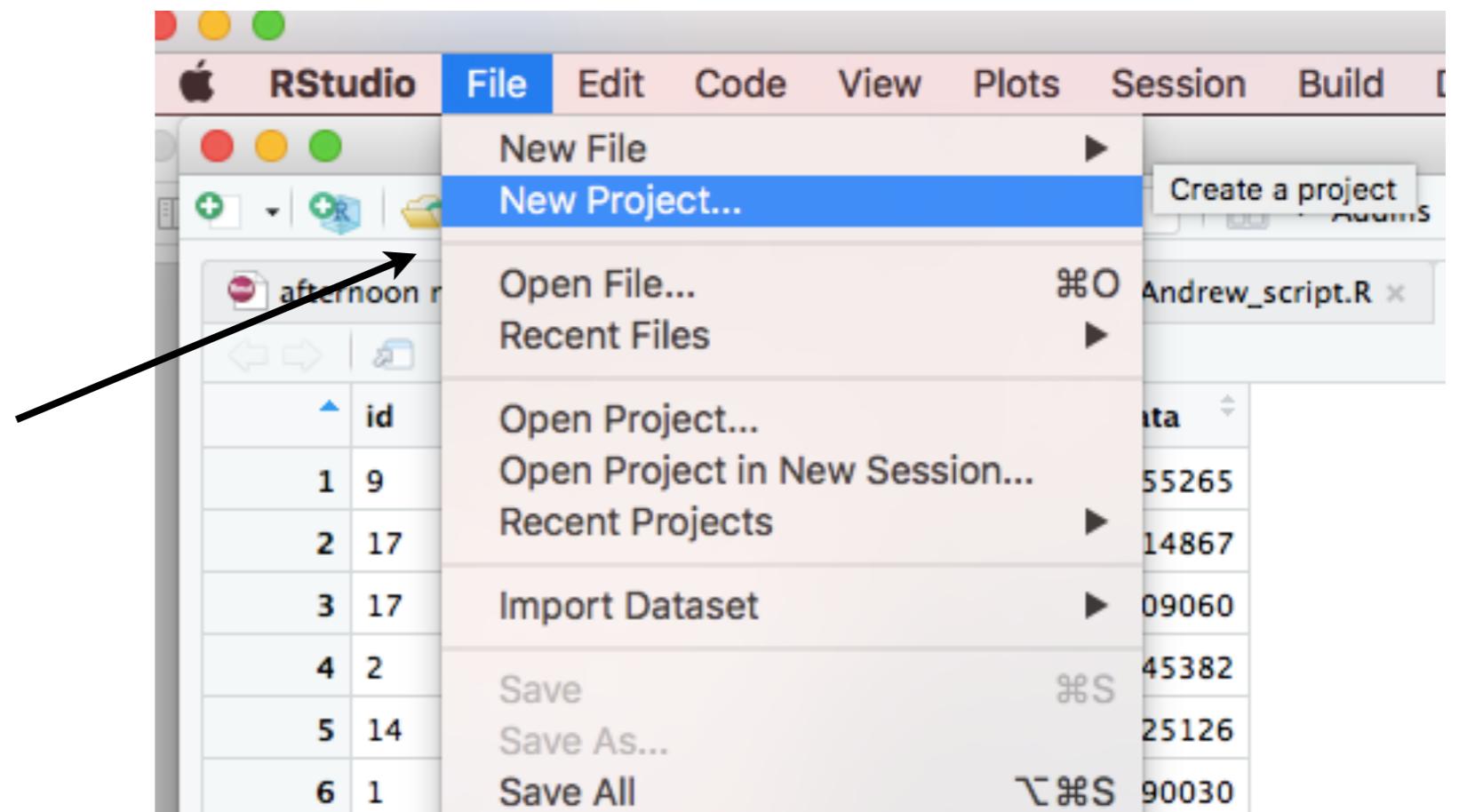
Subject	Priming Condition	RT
1	LONG	500
2	LONG	551
3	LONG	479
4	LONG	563
5	LONG	522
6	SHORT	399
7	SHORT	423
8	SHORT	444
9	SHORT	410
10	SHORT	398

Showing 1 to 10 of 10 entries

Name	Description	Version
labeling	Axis Labeling	0.3
lattice	Trellis Graphics for R	0.20-34
latticeExtra	Extra Graphical Utilities Based on Lattice	0.6-28
lazyeval	Lazy (Non-Standard) Evaluation	0.2.0
lme4	Linear Mixed-Effects Models using 'Eigen' and S4	1.1-12
lmerTest	Tests in Linear Mixed Effects Models	2.0-33
lsmeans	Least-Squares Means	2.25
magrittr	A Forward-Pipe Operator for R	1.5
markdown	'Markdown' Rendering for R	0.7.7
MASS	Support Functions and Datasets for Venables and Ripley's MASS	7.3-45
Matrix	Sparse and Dense Matrix Classes and Methods	1.2-7.1
methods	Formal Methods and Classes	3.3.2
mgcv	Mixed GAM Computation Vehicle with GCV/AIC/REML Smoothness Estimation	1.8-15
mime	Map Filenames to MIME Types	0.5
minqa	Derivative-free optimization algorithms by quadratic approximation	1.2.4
multcomp	Simultaneous Inference in General Parametric Models	1.4-6
munsell	Utilities for Using Munsell Colours	0.4.3

When Starting R

Always create a new project - this will keep all your files nice and organised.

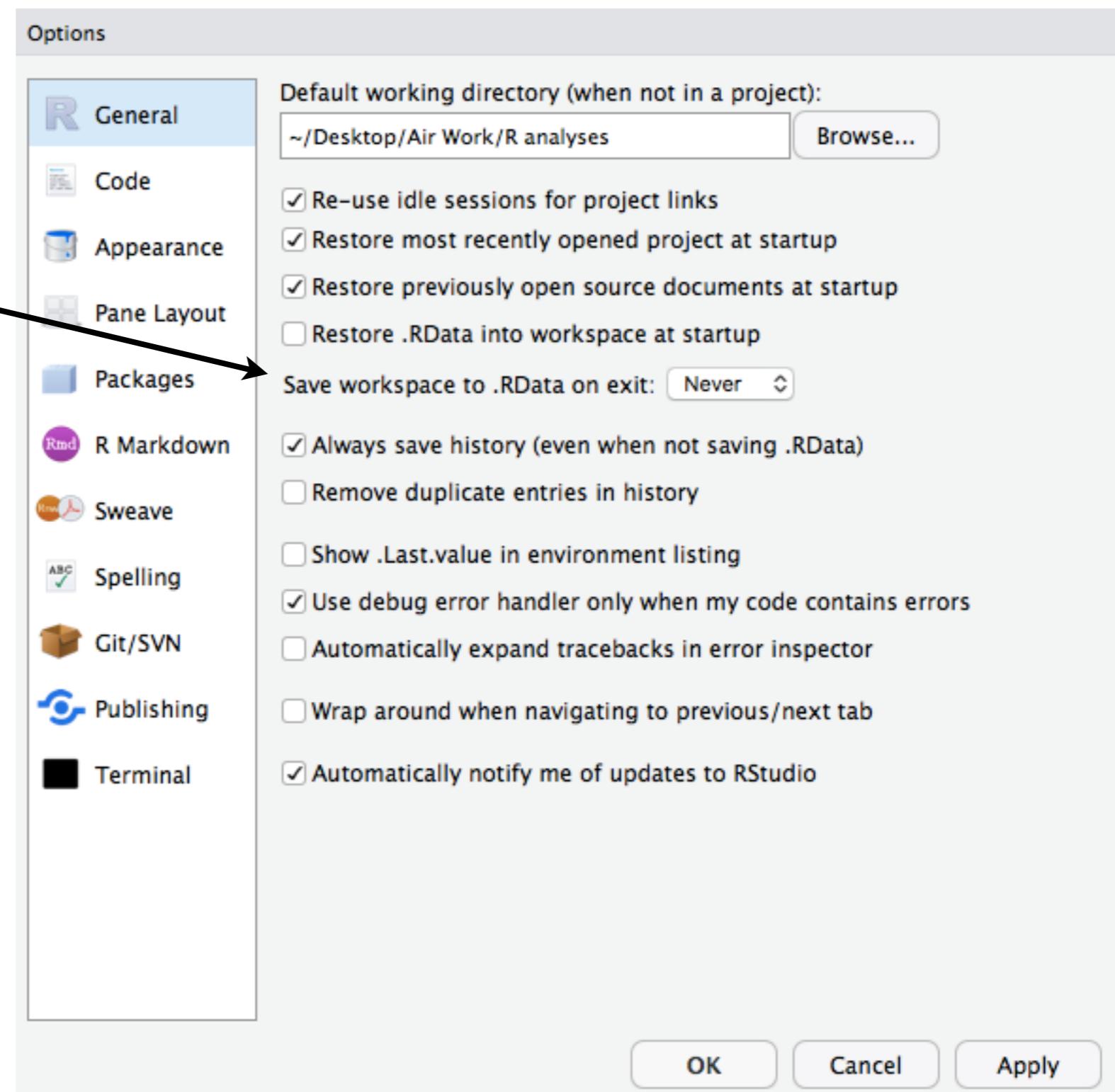


When you create an R Project in a folder, it's a good idea to keep that folder itself nicely organised:

```
name_of_project
| --name_of_project.Rproj
| --raw_data
|   |--data.csv
| --tidied_data
|   |--tidy_data.csv
| --R_scripts
|   |--data_tidy.R
|   |--data_vis.R
|   |--analysis.R
| --markdown_scripts
|   |--analysis.Rmd
| --output_reports
|   |--analysis.html
|   |--analysis.pdf
```

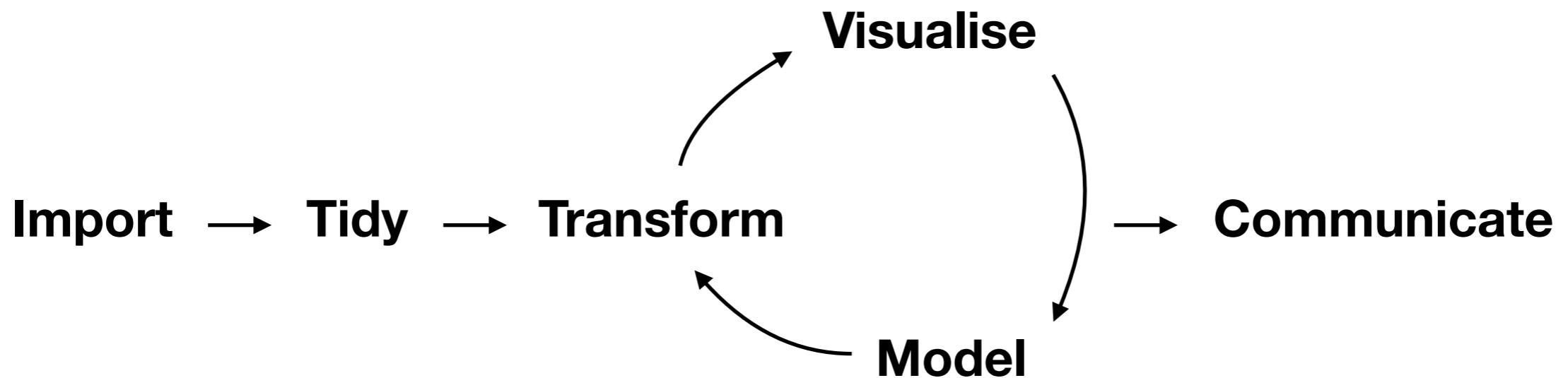
When Starting R

Under preferences,
make sure you
uncheck the saving
workspace
option...

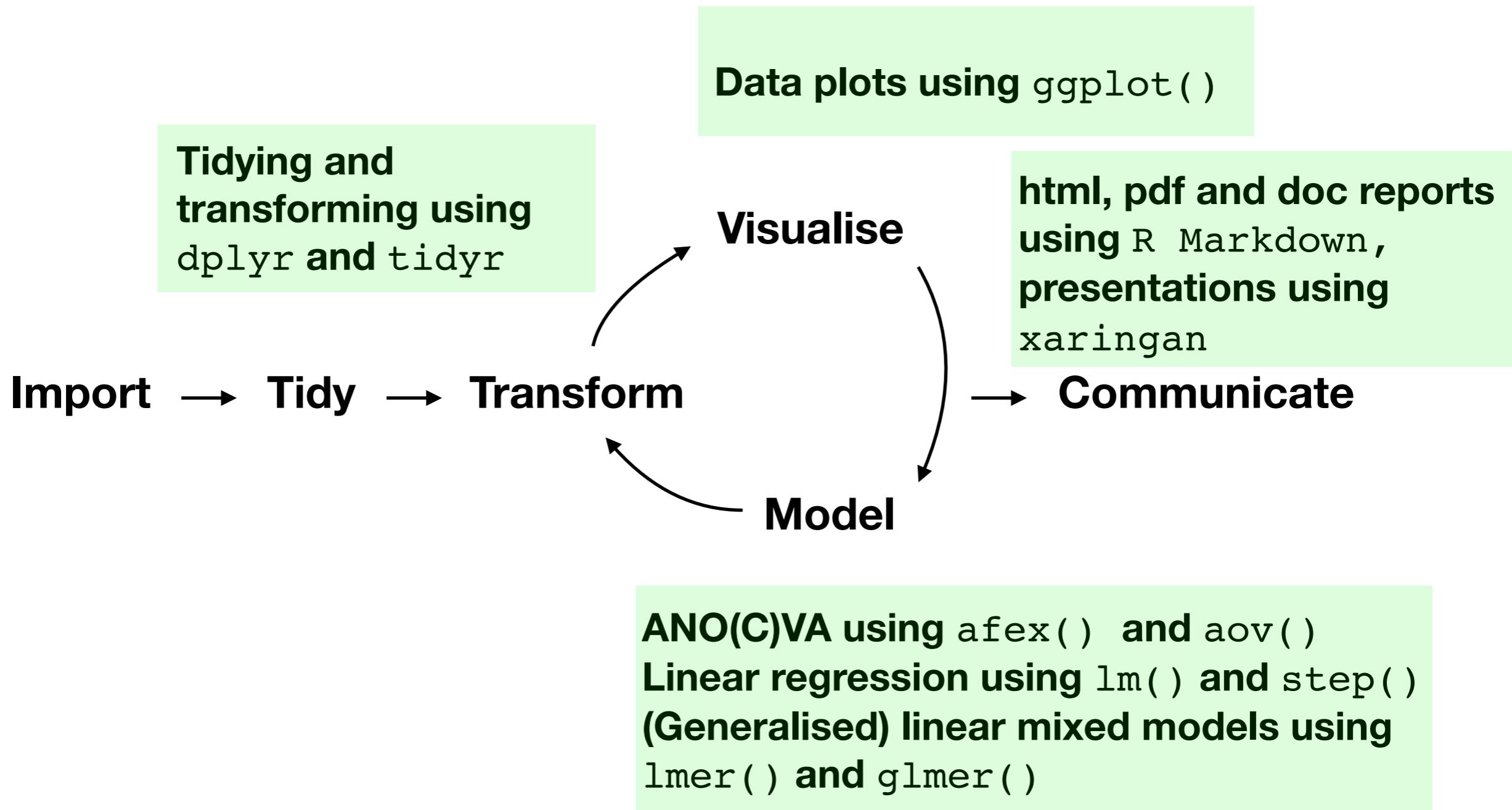


Analysis Workflow in the Tidyverse

(Garrett Grolemund and Hadley Wickham) - from Data to Write-up



Workflow



Packages in R

R has a number of functions already built in. These are part of “base R”. For much of what we do we need to load particular packages to let us use additional functions. We do this by:

```
> install.packages ("packagename")  
  
> library(packagename)
```

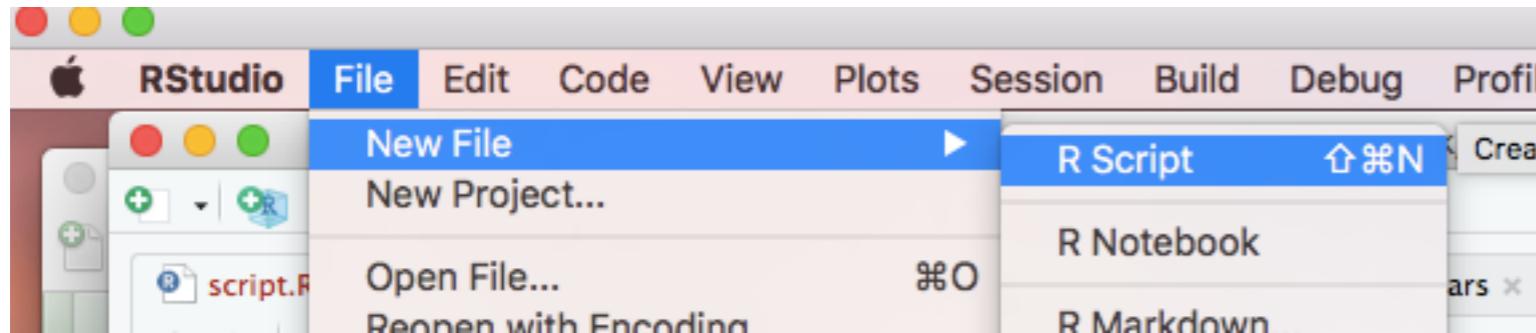
Every time you start R afresh, you need to load the packages you will use by using the `library` function. If you see notation like the following, it means we’re using a function (e.g., `summarise`) from within a package (e.g., `dplyr`)

```
dplyr::summarise
```

You don't need to re-invent the wheel...

- For any problem you want to solve, chances are others have had the same problem, solved it and have created an R package to do exactly what you want...
- One of the many great things about R is that you can add freely available packages to your library.
- ~15,000 R packages currently available
- The sites r-bloggers.com and rweekly.org are good places to find out about new packages.

Writing Scripts in R



Ultimately you will be writing scripts that will allow yourself and others to recreate your analysis just by running the script - the code at the start of the script will load the packages your analysis needs, then load your data, tidy, transform and visualise your data, and then code for your model...

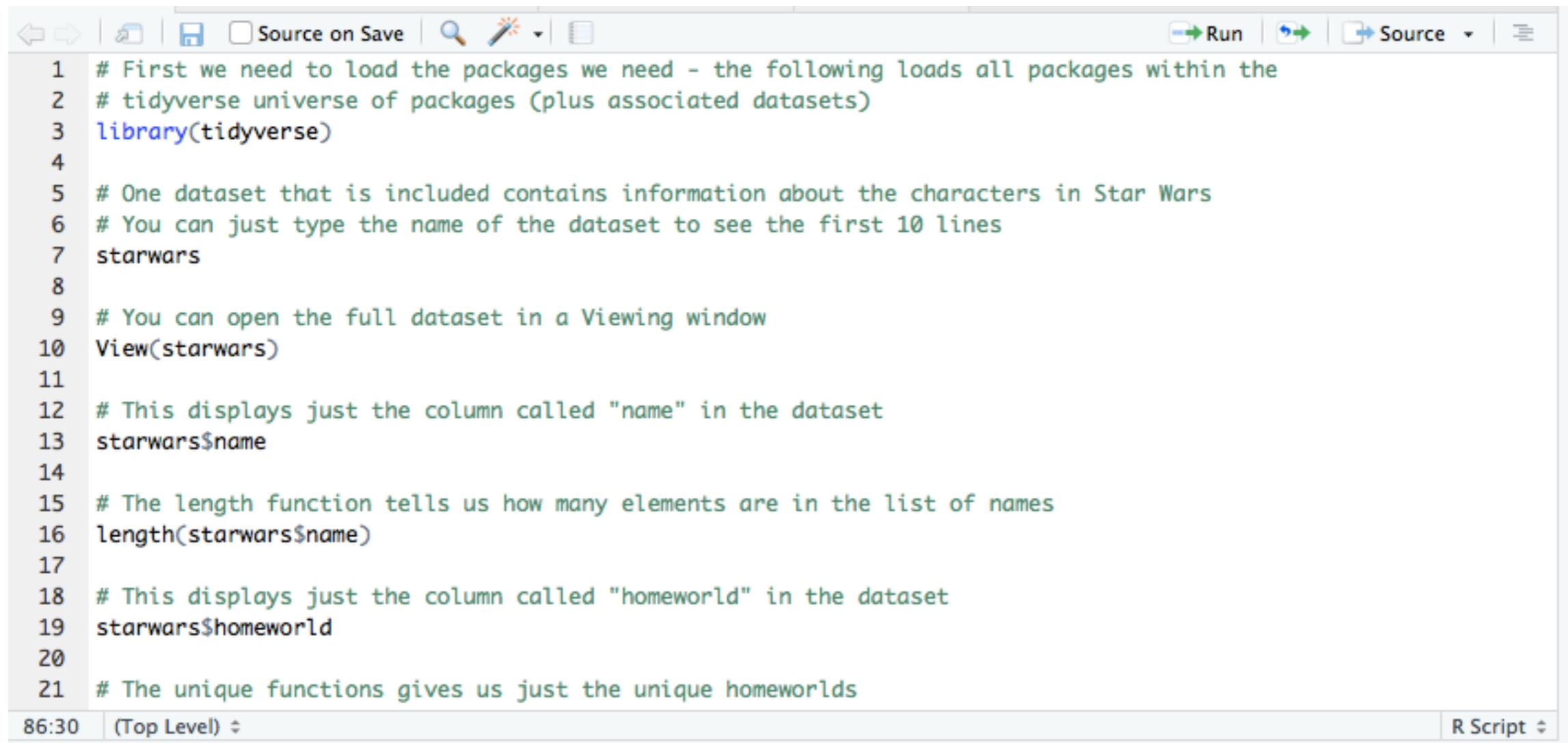
You should write your code as a script, and use the Console window for single command instructions relatively rarely...

Writing Scripts in R

Scripts should have lots of comments, indicated by a # symbol - this helps others read your code, and also yourself when you look back at it later.

Scripts can be saved and uploaded to (e.g.) OSF, GitHub, or submitted as an electronic supplement alongside your journal submission. Even when journals don't require you to adhere to Open Science practices, it's a good idea to make your code and data available during the reviewing process - and publicly available once your paper is published.

If you aren't sharing your data and code, you aren't engaged in generating reproducible or open research...

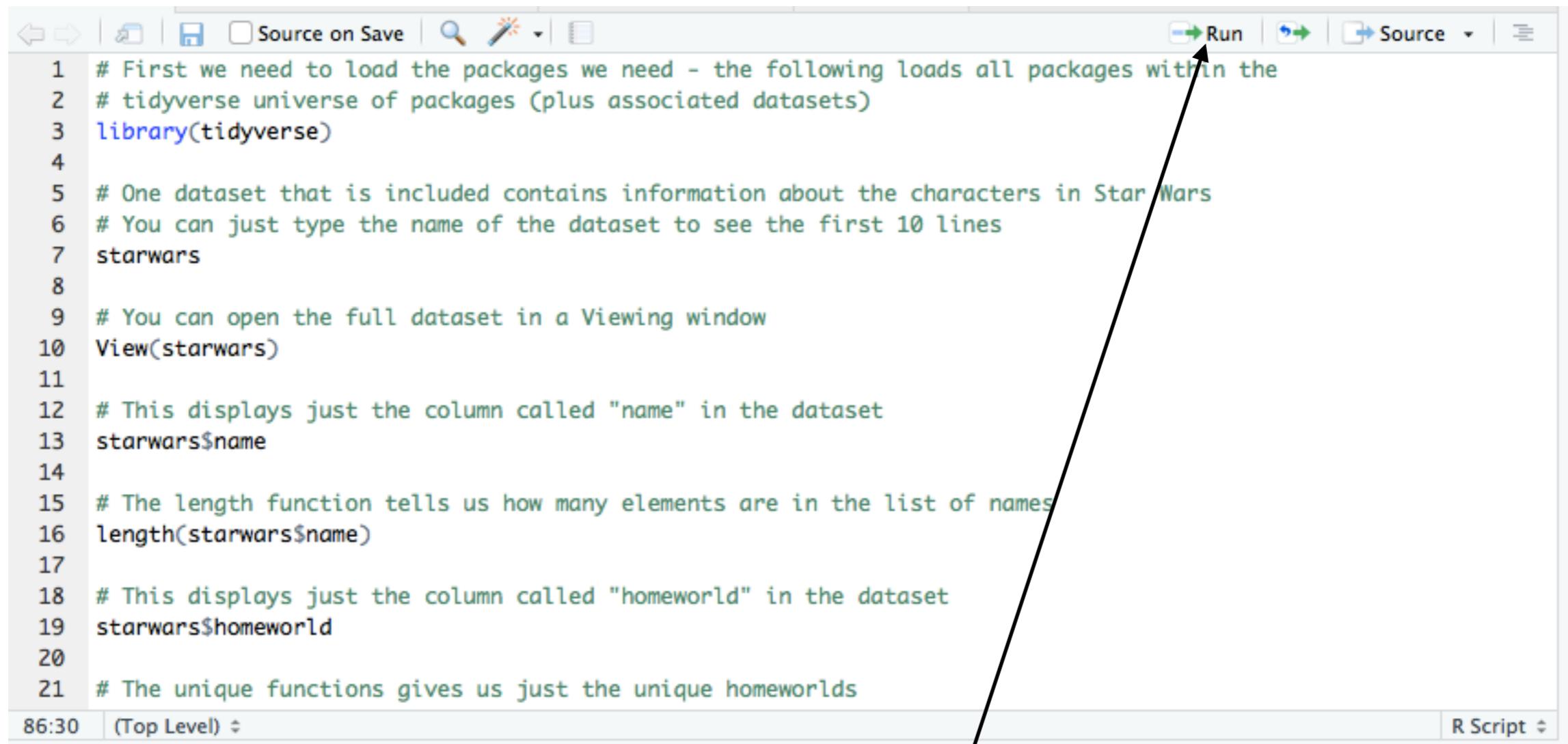


The screenshot shows the RStudio interface with an R script editor. The script contains numbered comments explaining various R commands:

```
1 # First we need to load the packages we need - the following loads all packages within the
2 # tidyverse universe of packages (plus associated datasets)
3 library(tidyverse)
4
5 # One dataset that is included contains information about the characters in Star Wars
6 # You can just type the name of the dataset to see the first 10 lines
7 starwars
8
9 # You can open the full dataset in a Viewing window
10 View(starwars)
11
12 # This displays just the column called "name" in the dataset
13 starwars$name
14
15 # The length function tells us how many elements are in the list of names
16 length(starwars$name)
17
18 # This displays just the column called "homeworld" in the dataset
19 starwars$homeworld
20
21 # The unique functions gives us just the unique homeworlds
```

The status bar at the bottom shows the time as 86:30, the project level as (Top Level), and the file type as R Script.

Lots of comments explaining what each section of code does, and lots of white space.



```
1 # First we need to load the packages we need - the following loads all packages within the
2 # tidyverse universe of packages (plus associated datasets)
3 library(tidyverse)
4
5 # One dataset that is included contains information about the characters in Star Wars
6 # You can just type the name of the dataset to see the first 10 lines
7 starwars
8
9 # You can open the full dataset in a Viewing window
10 View(starwars)
11
12 # This displays just the column called "name" in the dataset
13 starwars$name
14
15 # The length function tells us how many elements are in the list of names
16 length(starwars$name)
17
18 # This displays just the column called "homeworld" in the dataset
19 starwars$homeworld
20
21 # The unique functions gives us just the unique homeworlds
```

Once a script is written, you can run the entire script by highlighting it all (CMD-A in OSX) and clicking on ‘Run’, or you can highlight a section and run only that. CMD-Return will also Run the highlighted code.

Like this...

Style Guide

File names (both scripts and data files) should be meaningful:

regression_model.R - Good

somemodel.R - Bad

Variable names should be meaningful and in lowercase:

age - Good

A1 - Bad

expt1_data - Good

Experiment1datawithoutliersremoved - Bad

Style Guide

Place spaces around operators like + , - , = , <-

Whitespace used sensibly makes your code easier to read. Separate discrete sections of code in your script with a blank line.

When writing scripts, continue on a subsequent line rather than writing one very long line.

Comment, comment, comment...

```
# First we load our datafile.
```

<https://style.tidyverse.org>

R Style Guide Ⓜ ⚐ A ⌂ ⌂ ⓘ The tidyverse style guide

Welcome

I Analyses

1 Files

1.1 Names

1.2 Organisation

1.3 Internal structure

2 Syntax

2.1 Object names

2.2 Spacing

2.3 Argument names

2.4 Code blocks

2.5 Long lines

2.6 Assignment

2.7 Semicolons

2.8 Quotes

2.9 Comments

3 Functions

3.1 Naming

The tidyverse style guide

Hadley Wickham

Welcome

Good coding style is like correct punctuation: you can manage without it, but it sure makes things easier to read. This site describes the style used throughout the [tidyverse](#). It was derived from Google's original R Style Guide - but Google's [current guide](#) is derived from the tidyverse style guide.

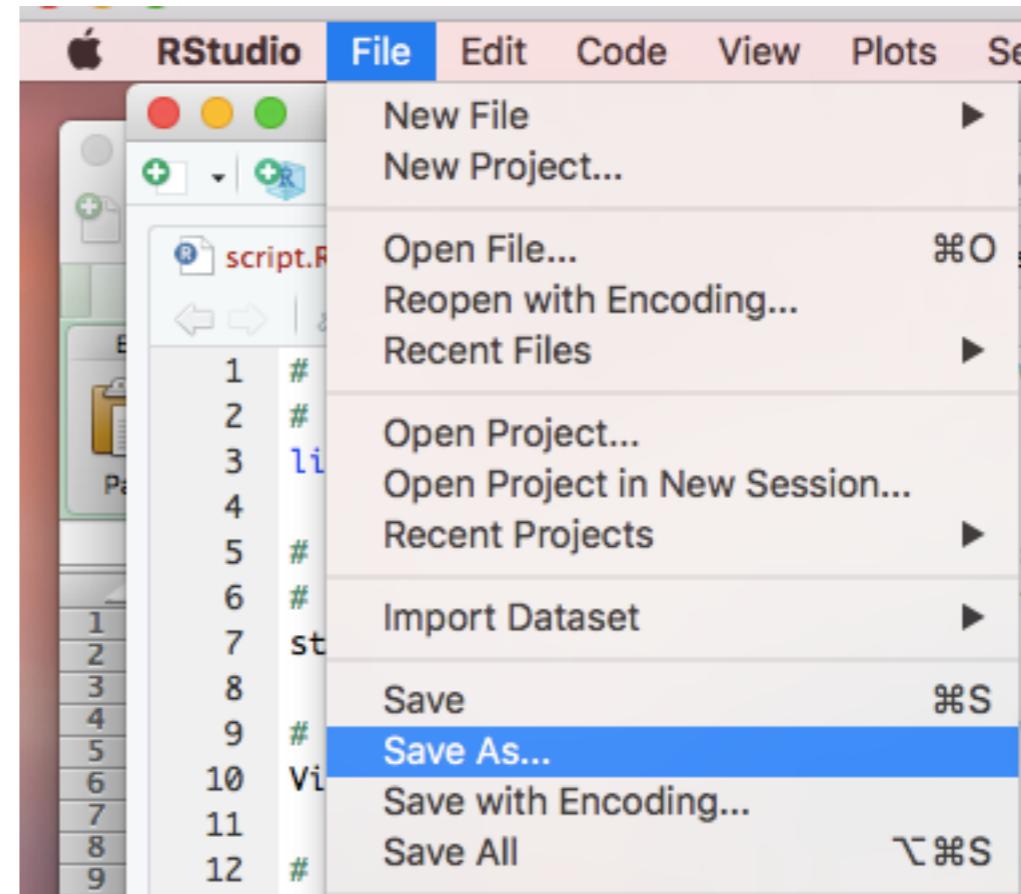
All style guides are fundamentally opinionated. Some decisions genuinely do make code easier to use (especially matching indenting to programming structure), but many decisions are arbitrary. The most important thing about a style guide is that it provides consistency, making code easier to write because you need to make fewer decisions.

Two R packages support this style guide:

- [styler](#) allows you to interactively restyle selected text, files, or entire projects. It includes an RStudio add-in, the easiest way to re-style existing code.



Once your
script is
finished, don't
forget to save
it...



Sharing your analysis

Journals are increasingly requesting analysis code and data to be published alongside the published paper (and sometimes at the review stage).

You can share your analysis and data even at the submission stage using something like GitHub or OSF.

Git is a powerful tool that allows for the version control in collaborative projects, and the sharing of code and projects.

You can set up a GitHub account, and install GitHub Desktop on your own computer.

Google Scholar Scopus jobs.ac.uk BBC News Chester Weather The Grauniad The Independent Google Maps Chester Weather Station GitHub

Search or jump to... / Pull requests Issues Marketplace Explore + 

 Overview Repositories 34 Projects 0 Stars 9 Followers 5 Following 1

Pinned Customize your pins

 Psychology_MRes_Stats_R_Course Slides for my MRes Stats Course  HTML  1 

 Affective-Theory-of-Mind-Inferences R code and data for Stewart, S.L.K., Schepman, A., Haigh, M., McHugh, R., & Stewart, A.J. (2019). Affective theory of mind inferences contextually influence the recognition of emotional facial expr...  R 

 Comprehension-of-indirect-requests-is-influenced-by-their-degree-of-imposition R code and data for Stewart, A.J., Le-luan, E., Yao, B., Wood, J., & Haigh, M., (2018). Comprehension of indirect requests is influenced by their degree of imposition. Discourse Processes, 55, 187-... 

 It's-hard-to-write-a-good-article R code and data for Stewart, A.J., Wood, J.S., Le-luan, E., Yao, B., & Haigh, M. (2018). "It's hard to write a good article." The online comprehension of excuses as indirect replies. Quarterly Jour... 

 Psychophysiology2017 Forked from RonanMcG/Psychophysiology2017 R scripts and dataset for McGarrigle, R.A., Dawes, P., Stewart, A.J., Kuchinsky, S.E., & Munro, K.J. (2017). Pupilometry reveals changes in physiological arousal during a sustained listening task....

 ajstewartlang.github.io



Google Scholar Scopus jobs.ac.uk BBC News Chester Weather The Grauniad The Independent Google Maps Chester Weather Station GitHub

Search or jump to... Pull requests Issues Marketplace Explore

ajstewartlang / Affective-Theory-of-Mind-Inferences Watch 0 Star 1 Fork 0

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

R code and data for Stewart, S.L.K., Schepman, A., Haigh, M., McHugh, R., & Stewart, A.J. (2019). Affective theory of mind inferences contextually influence the recognition of emotional facial expressions. Cognition & Emotion, 33, 272-287. Edit

Manage topics

15 commits 1 branch 0 releases 1 contributor

Branch: master New pull request Create new file Upload files Find File Clone or download

 ajstewartlang	Update README.txt	Latest commit e4b4ff4 28 days ago	
	Analysis with emotional word items excluded	typos corrected	2 years ago
	Analysis with emotional word items excluded_old	final changes	a year ago
	Both Expts comparison	final changes	a year ago
	Expt 1 Script and data	final changes	a year ago
	Expt 2 Script and data	final changes	a year ago
	README.txt	Update README.txt	28 days ago

Google Scholar Scopus jobs.ac.uk BBC News Chester Weather The Grauniad The Independent Google Maps Chester Weather Station GitHub

Search or jump to... Pull requests Issues Marketplace Explore

ajstewartlang / Affective-Theory-of-Mind-Inferences Watch 0 Star 1 Fork 0

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

Branch: master Affective-Theory-of-Mind-Inferences / Expt 1 Script and data / Create new file Upload files Find file History

		Latest commit dc6b924 on 4 Apr 2018
..		
AngerFearAcc.csv	Tidied up code with consistent labelling	2 years ago
AngerFearRT.csv	Tidied up code with consistent labelling	2 years ago
AngerFeargraphacc.csv	R code and data	2 years ago
AngerFeargraphdata.csv	R code and data	2 years ago
Expt1_AngerFear_script.R	final changes	a year ago
Expt1_AngerFear_script_final.R	final changes	a year ago
Expt1_AngerFear_script_with_exclusions_final.R	final changes	a year ago
Expt1_AngerFear_script_with_item_exclusions_fin...	final changes	a year ago

This repository Search Pull requests Issues Gist + ⚙

ajstewartlang / Affective-Theory-of-Mind-Inferences Watch 0 Star 0 Fork 0

Code Issues 0 Pull requests 0 Projects 0 Wiki Pulse Graphs Settings

Branch: master Find file Copy path

Affective-Theory-of-Mind-Inferences / Expt 1 Script and data / Expt1_AngerFear_script.R

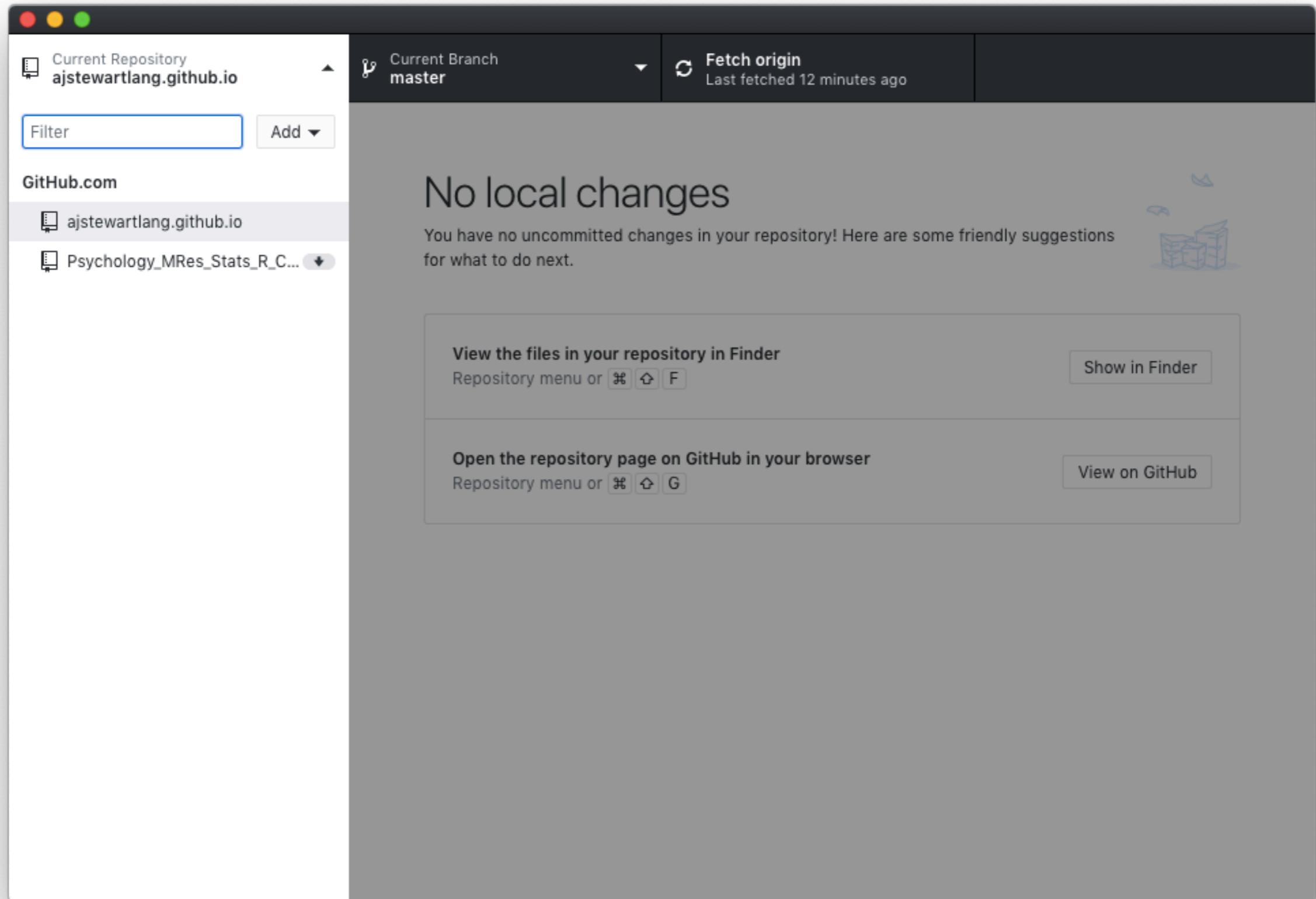
ajstewartlang Tidied up code with consistent labelling de7b584 3 hours ago

0 contributors

59 lines (46 sloc) | 3.17 KB Raw Blame History

```
1 library(lme4)
2 library(lmerTest)
3 library(lsmeans)
4 library(pbkrtest)
5 library(readr)
6 library(ggplot2)
7
8 #script for AngerFear RT and accuracy data analysis with arousal
9
10 #this is the analysis of the RT data
11 AngerFearRT <- read_csv("~/AngerFearRT.csv")
12
13 AngerFearRT$StoryEmotion <- as.factor(AngerFearRT$StoryEmotion)
14 AngerFearRT$FaceExpression <- as.factor(AngerFearRT$FaceExpression)
15
16 contrasts(AngerFearRT$StoryEmotion) <- matrix(c(.5, -.5))
17 contrasts(AngerFearRT$FaceExpression) <- matrix(c(.5, -.5))
18
19 #with Subject, Vignette, and Face as crossed random effects with arousal
20 #full model does not converge so need to drop interaction term from the random effects - in addition, Face random effect has only random in
21 modelRTAr1 <- lmer(RT ~ StoryEmotion*FaceExpression*Arousal + (1+StoryEmotion+FaceExpression|Subject) + (1+StoryEmotion+FaceExpression|Vig
22 summary(modelRTAr1)
23 modelRT <- lmer(RT ~ StoryEmotion*FaceExpression + (1+StoryEmotion+FaceExpression|Subject) + (1+StoryEmotion+FaceExpression|Vignette) + (1
24 anova(modelRTAr1, modelRT)
25
26 #difference between models not signif - arousal does not interact with effect so drop arousal from subsequent analysis
27
28 modelRTnull <- lmer(RT ~ (1+StoryEmotion+FaceExpression|Subject) + (1+StoryEmotion+FaceExpression|Vignette) + (1+FaceExpression|Face),
29 anova(modelRT, modelRTnull)
30 summary(modelRT)
```

This is GitHub Desktop



**Let's run
through a
script...**



TWENTIETH CENTURY-FOX Presents A LUCASFILM LTD. PRODUCTION **STAR WARS**
Starring **MARK HAMILL HARRISON FORD CARRIE FISHER**
PETER CUSHING
and
ALEC GUINNESS

Written and Directed by
GEORGE LUCAS

Produced by
GARY KURTZ JOHN WILLIAMS

PANAVISION® PRINTS BY DE LUXE® TECHNICOLOR®

Original Motion Picture Soundtrack on 20th Century Records and Tapes

Making Films Sound Better
DOLBY SYSTEM
Noise Reduction - High Fidelity

**STAR
WARS**

