

Open Source Contribution and the Effectiveness of Public Work

Amy Yang
Senior Data Scientist
 @ayanalytics

Inspired by [David Robinson's keynote](#) at Rstudio :: conf 2019



Hadley Wickham



@hadleywickham



Some advice on managing your online presence if you're applying for a job or scholarship, based on my experiences reviewing rstudio::conf scholarship applications (results out soon):

1:25 PM · Oct 31, 2019 · [Twitter Web App](#)

311 Retweets **1.1K** Likes

- Polished GitHub page
- Pin repos that represent strong work
- Share public data analysis challenges
- Make a personal website

 Hadley Wickham  @hadleywickham · Oct 31

Replies to [@hadleywickham](#)

1. Spend some time polishing your GitHub home page. Make sure you have a custom image — if you don't want your face on in the internet, pick some image that represents you

3

11

117



 Hadley Wickham  @hadleywickham · Oct 31

2. Pin a few repos that represent particularly strong work. Make sure those repos have concise but descriptive titles, and that you've added a readme that briefly describes the contents. Include a plot/image if you have one

1

8

88



 Hadley Wickham  @hadleywickham · Oct 31

3. If you can't share your usual work, do a couple of tidy tuesday's ([github.com/rfordatascience.../tidytuesday](https://github.com/rfordatascience/tidytuesday)) or try out other public data analysis challenges



rfordatascience/tidytuesday

Official repo for the #tidytuesday project. Contribute to rfordatascience/tidytuesday development by ...

 [github.com](#)

2

16

209



 Hadley Wickham  @hadleywickham · Oct 31

4. If you want to do more, make a nice personal website — a good place to start is blogdown + academic theme. [@apreshill](#) has a great tutorial to get you started:



Edward Visel @alistaire · Oct 31

Doing these things is how I got my job.



Hadley Wickham ✅ @hadleywickham · Oct 31

Some advice on managing your online presence if you're applying for a job or scholarship, based on my experiences reviewing rstudio::conf scholarship applications (results out soon):

[Show this thread](#)



Edward Visel

Data Scientist at Uptake

Education



Pomona College

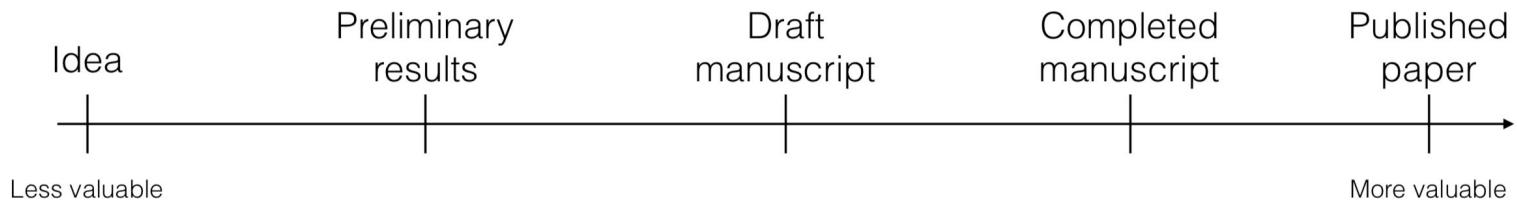
BA, Philosophy

2005 – 2009

- Completed extensive coursework in philosophy, French, economics, and politics.



Goals in grad school:



How I should have been thinking of them:

Anything still
on your computer

(Data, code, results,
draft, finished paper)

Anything out
in the world

(Paper, preprint, product,
blog post, open source,
tweet)

Less valuable

More valuable



The effectiveness of sharing





Aspiring

Junior

Senior

Blog

Tweet

Contribute to open source

Grow the community

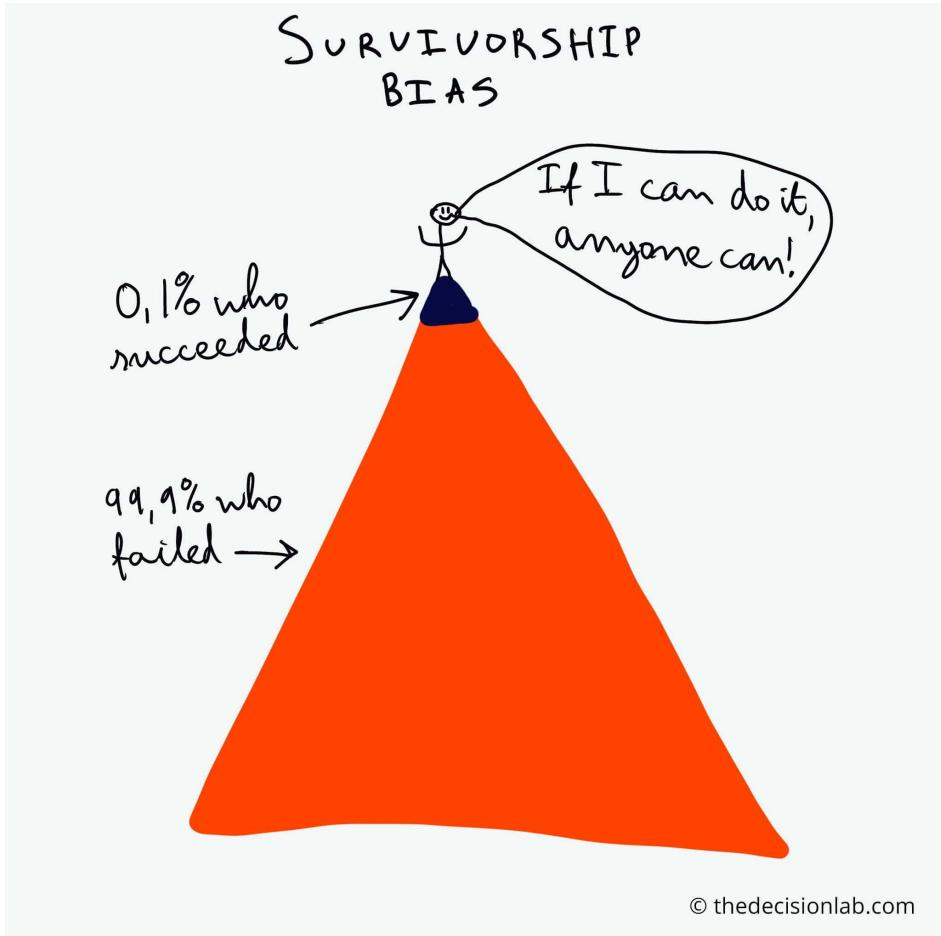
Give talks



Why spend time on public work?

- Advance your career
- Practice good habits
- Contribute to the community

Disclaimer 1





Disclaimer 2

Open Source Contribution and the
Effectiveness of Public Work

≠

Necessity



Aspiring

Junior

Senior

Blog



Why start a blog?

- Build a public portfolio



Edward Visel

Data Scientist at Uptake

About

<https://alistaire.rbind.io>



Field Organizer

Organizing for America - Virginia

Jun 2012 – Nov 2012 · 6 mos

Charlottesville, Virginia Area



Intern, Office of Congressman Juan Vargas

U.S. House of Representatives

Mar 2013 – Jul 2013 · 5 mos

Washington, D.C.



Regional Field Director

Democratic Congressional Campaign Committee

Jul 2014 – Nov 2014 · 5 mos

Rockford, Illinois Area



EDWARD VISEL

ODDS, ENDS, AND R CODE



RECENT POSTS

MATCH.ARG

JULY 4, 2019

[CONTINUE READING](#)

R | PROGRAMMING
 FUNCTIONS | UTILITY

MAPPING LEAVES

JANUARY 20, 2019



FIREWORKS

NOVEMBER 11, 2018

MAPPING LEAVES

RECURSIVE FUNCTIONAL PROGRAMMING

JANUARY 20, 2019

EDWARD VISEL

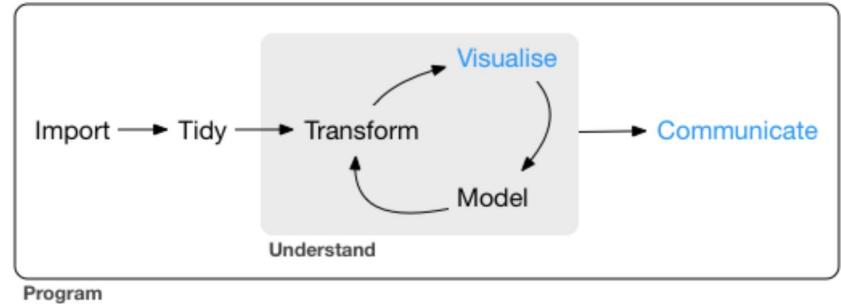
10 MINUTE READ

I love purrr.¹ Aside from its anonymous function notation, one of the functions that made me love the package was `at_depth`, which iterates across a list at a specified level of nesting. It has since been deprecated in favor of `modify_depth`, which is more powerful, but is significantly more finicky.



Why start a blog?

- Build a public portfolio
- Practice writing and visualization



"It doesn't matter how great your analysis is unless you can explain it to others: you need to **communicate** your results."

R for Data Science
Hadley Wickham & Garrett Grolemund



Why start a blog?

- Build a public portfolio
- Practice writing and visualization
- Teach/give advice in a way that

scales



David Robinson @drob · 8 Nov 2017

When you've written the same code 3 times, write a function

When you've given the same in-person advice 3 times, write a blog post



48



1.5K



4.2K



How can you start a blog?

blogdown: Creating Websites with R

Markdown



Please do not be misled by the word “blog” in the package name: blogdown is for general-purpose websites, and not only for blogs.

“What could I blog about?”

Data science/Data analysis project

kaggle  Search Competitions Datasets Notebooks Discussion Courses ...

General InClass Sort by: Grouped

All Categories Search competitions 

17 Active Competitions

Competition	Description	Prize Money	Team Count
 Deepfake Detection Challenge Identify videos with facial or voice manipulations Featured · Code Competition · 3 months to go · video data, online video	\$1,000,000 832 teams		
 2019 Data Science Bowl Uncover the factors to help measure how young children learn Featured · Code Competition · 20 days to go · video games, children, learning, education	\$160,000 2,785 teams		
 NFL Big Data Bowl How many yards will an NFL player gain after receiving a handoff? Featured · Code Competition · 4 days to go · american football, sports	\$75,000 2,038 teams		
 NFL 1st and Future - Analytics Can you investigate the relationship between the playing surface and the injury and performance of NFL ... Analytics · an hour to go · sports, tabular data	\$75,000		
 TensorFlow 2.0 Question Answering Identify the answers to real user questions about Wikipedia page content Featured · Code Competition · 20 days to go · text mining, text data	\$50,000 1,055 teams		
 Google QUEST Q&A Labeling Improving automated understanding of complex question answer content Featured · Code Competition · a month to go · text data, nlp	\$25,000 791 teams		
 Peking University/Baidu - Autonomous Driving Can you predict vehicle angle in different settings? Featured · 19 days to go · image processing, Image data	\$25,000 732 teams		



“What could I blog about?”

Consider your passions and things you care about

Australia fires: A visual guide to the bushfire crisis

4 hours ago

Australia fires



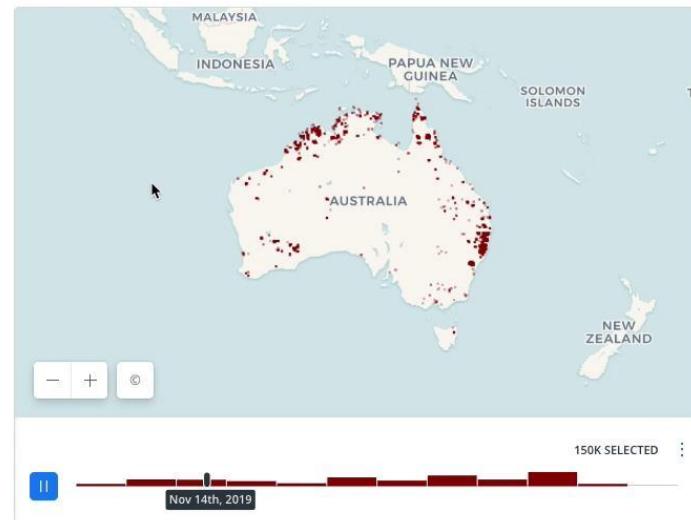
Record-breaking temperatures and months of severe drought have fuelled a series of massive bushfires across Australia.

Although recent cooler conditions have brought some respite, the states of New South Wales and Victoria are once again braced for the return of severe fire danger - with high temperatures and gusty winds forecast over the next few days.

Some 27 people have so far been killed - including four firefighters - and an estimated 10 million hectares (100,000 sq km or 15.6 million acres) of bush, forest and parks across Australia has burned.

Fire detections / hotspots

- In given week
- Previously





“What could I blog about?”

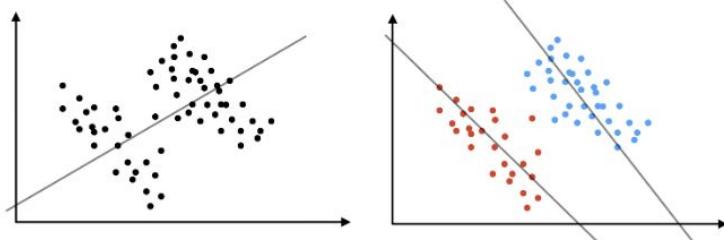
Teach a concept

Simpson's Paradox and Interpreting Data

The challenge of finding the right view through data



Tom Grigg [Follow](#)
Dec 9, 2018 · 7 min read



A visual example: the overall trend reverses when data is grouped by some colour-represented category.

Simpson's Paradox:

A trend or result that is present when data is put into groups that reverses or disappears when the data is combined.



“What could I blog about?”

Turn your papers into a blog post



Philip Guo



@pgbovine

every published research paper should ideally have an accompanying summary blog post. i've tried to do this for everything i've published post-Ph.D. (sometimes batching several papers into one post) & encourage my students to do the same. 10-100x more people will read blog posts.

11:14 AM · Jul 19, 2018 · [Twitter Web Client](#)

427 Retweets 1.8K Likes



Aspiring

Junior

Senior

Blog

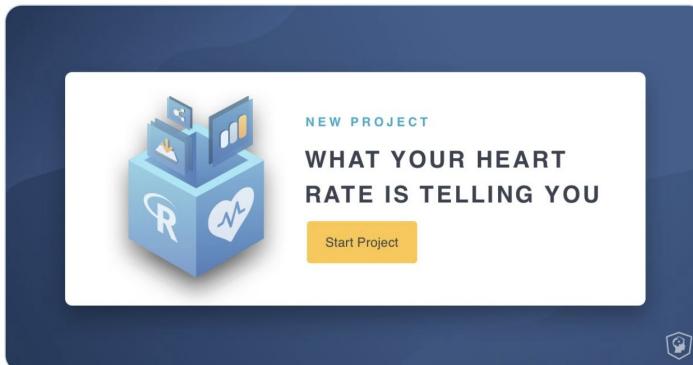
Tweet



Promote and discuss your work



New Project by [@ayanalytics](#): What Your Heart Rate Is Telling You! In this R project, you will examine the relationship between the maximum heart rate one can achieve during exercise and the likelihood of developing heart disease. datacamp.com/projects/445



James Lamb
 @_jameslamb

Do you like writing [#rstats](#) and [#python](#) code? Have you wanted to try [#opensource](#), but held back because it seems intimidating? I have a bunch of "good first issues" on a side project I'm passionate about, and I promise I'm a kind and patient reviewer: github.com/jameslamb/dopp...



Promote and discuss the work of others



Mara Averick
@dataandme

👉 guide 📚 now = 🔥!

"blogdown: Creating Websites with R Markdown"

buff.ly/2x8vo48 by @xieyihui @apreshill @ProQuesAsker
#rstats

blogdown: Creating Websites with R Markdown

Yihui Xie, Amber Thomas, Alison Presmanes Hill

@xieyihui @apreshill @ProQuesAsker

2017-09-18



The online version of this book is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

% <https://bookdown.org/yihui/blogdown/>

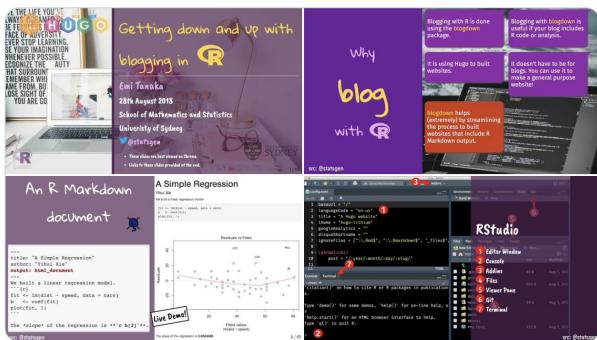


Mara Averick
@dataandme

🌈 super fun slides and great #blogdown guide by @statsgen:

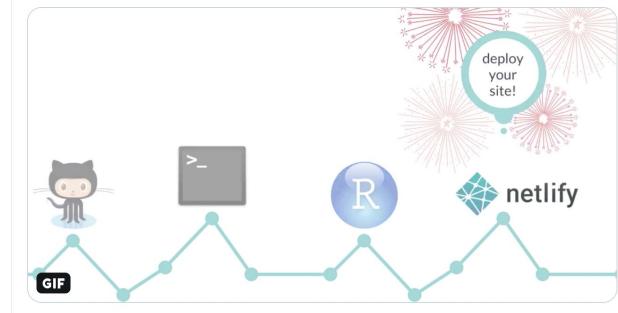
🎥 "Getting down and up with blogging in R"

buff.ly/2LygxHh #rstats

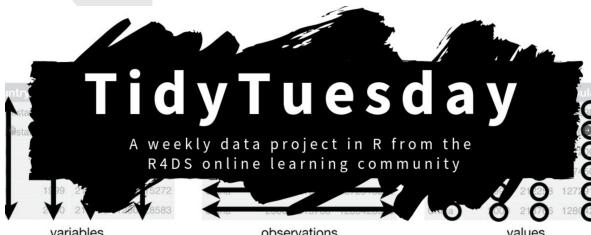


Mara Averick
@dataandme

🔥 step-by-step: "Up and running with blogdown" by @Apresmanes buff.ly/2tj3G3J via @rbindio #rstats #blogdown



Work with a community

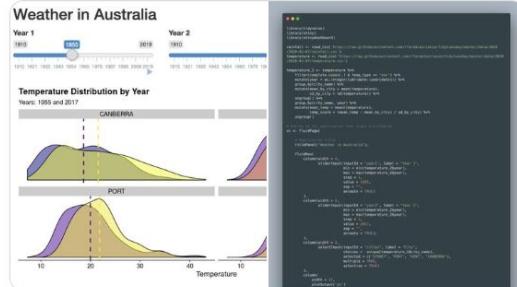


 **James Wade** @JamesHWade · Jan 8
I made a `#shiny` app for `#TidyTuesday` "Australia Fires" sub

Use it compare temperature by year for several Australian cities. I'm working on embedding it on my [#blogdown](#) site with help from [@rstudio](#) community.

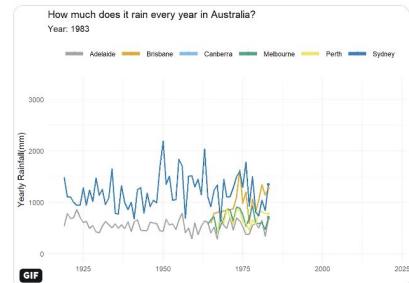
app: james-h-wade.shinyapps.io/TidyTuesday_20...

#rstats #r4ds #dataviz



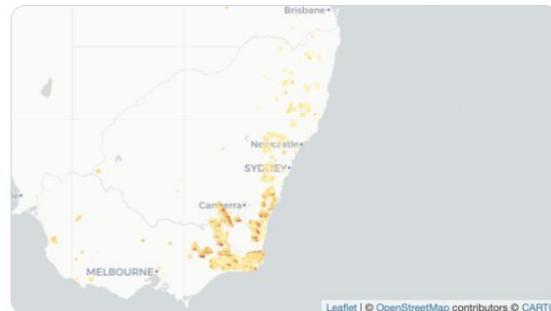
Adam Houlette

Going to try to do a lot of [#TidyTuesday](#) posts this year.
Starting off simple by looking at yearly rainfall in Australia
using [#gganimate](#). A little more detail at
ahoulette.com/2020/01/09/tid...



 Luisa Groher @luisagroher · Jan 8

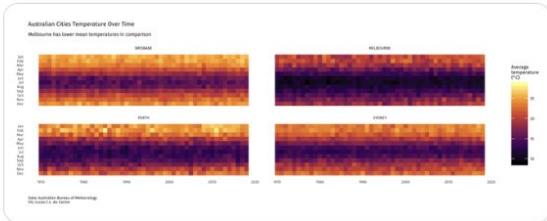
 I always wanted to participate in tidy tuesday and finally got around to it! This is fun. Hope it becomes a habit. Australian Wildfire data + leaflet
[#TidyTuesday](#) github.com/luisagroher/tidyTuesday



meu nome não é charles @comsemassim · Jan 11

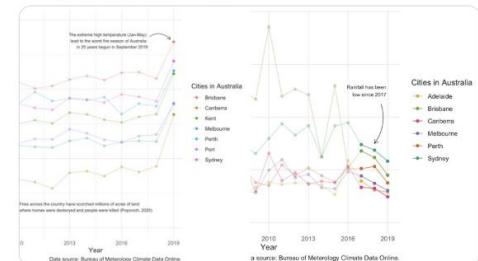
First **#TidyTuesday** submission ever, heavily inspired by [@htinkyawave](#)

Code: github.com/carmo-lucas/ti...



 **Shuhan Song** @song_shuhan · Jan 8
A late post for 01/07 #TidyTuesday A
there will get better ASAP.

Code: github.com/Shubhanstack/ti



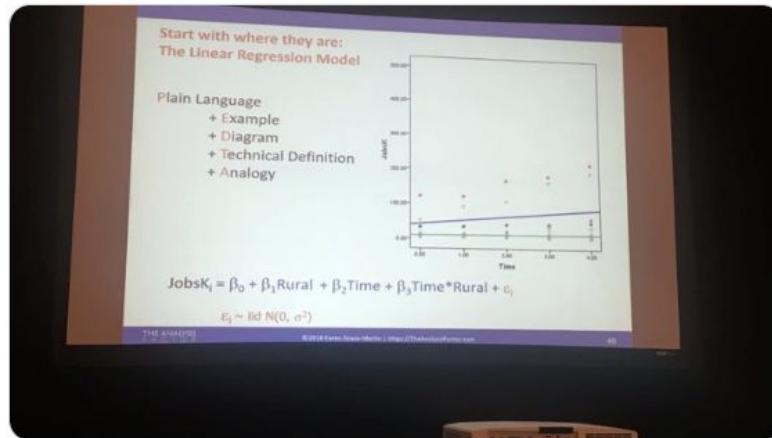
Share what you learned

- Conference/Talk/Journal club/Meetup etc.



Letisha Smith @LetishaAudrey · Feb 15, 2019

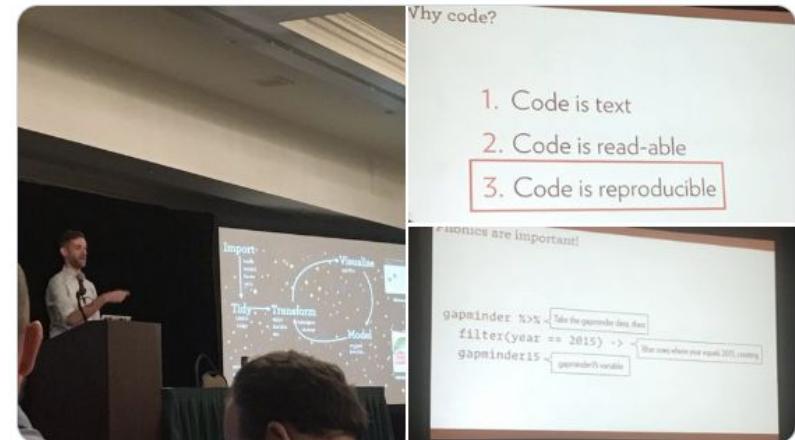
Very insightful talk from Karen Grace-Martin at [@analysis_factor](#) explaining complex concepts to non-statisticians. Her clarity in explanations gave me a deeper understanding of interactions, odds ratios, and multilevel models all in under than 30 minutes. #CSP2019



Letisha Smith @LetishaAudrey · Feb 15, 2019

Great #csp2019 kickoff with keynote from [@hadleywickham](#)! I learned some tips to reading r like:
%>% pipe translates to then &
-> assignment translates to create

Plus enjoyed hearing thoughts on why coding is important. #rstats





Aspiring

Junior

Senior

Blog

Tweet

Contribute to open source

```
> summary(lm(iris$Sepal.Length~iris$Sepal.Width))

Call:
lm(formula = iris$Sepal.Length ~ iris$Sepal.Width)

Residuals:
    Min      1Q  Median      3Q     Max 
-1.5561 -0.6333 -0.1120  0.5579  2.2226 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  6.5262    0.4789   13.63 <2e-16 ***  
iris$Sepal.Width -0.2234    0.1551   -1.44    0.152  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tidy_lm ← function(x) {
  coefs ← as.data.frame(coef(summary(x)))
  colnames(coefs) ← c("estimate", "std.error", "statistic", "p-value")
  coefs$term ← rownames(coefs)
  rownames(coefs) ← NULL
  coefs|
}
```



Individual contributor \longleftrightarrow Community

[tidymodels / broom](#)

Convert statistical analysis objects from R into tidy format <https://broom.tidyverse.org>

Code Issues Pull requests Actions Security Insights

1,020 commits 7 branches 0 packages 14 releases 110 contributors View license

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

alexphayes More NEWS Latest commit 3c922d5 on Dec 14, 2019

.github	New README and gam/mgcv doc linking	2 years ago
R	Remove insert_NAs	last month
data-raw	Move standardized tests and glossaries to modeltests package (#449)	2 years ago
inst	New `glance.aov()`, tidyverse DESCRIPTION, spelling fixes (#638)	11 months ago
man-roxygen	Disentangle tidy.lm() and ban use of lm tidiers by non-lm objects (#742)	last month
man	Remove confint_tidy() and col_name(). Closes #703.	last month
pkgdown/favicon	Attempt to automate pkgdown deploy	6 months ago
revdep	Very late merge of changes from 0.5.1 release (#620)	11 months ago
tests	Remove confint_tidy() and col_name(). Closes #703.	last month
vignettes	Remove quick functionality from tidy() and fix tests (#684)	6 months ago
.buildignore	Disentangle tidy.lm() and ban use of lm tidiers by non-lm objects (#742)	last month



David Robinson
@drob

Introducing broom: R package for turning model outputs into tidy data ([@hadleywickham](#) -style) for dplyr,tidyr
[#rstats](https://rstats.bit.ly/1IXtBbq) [bit.ly/1IXtBbq](https://rstats.bit.ly/1IXtBbq)

2:37 PM · Sep 11, 2014 · Twitter Web Client

24 Retweets 59 Likes

Individual contributor \longleftrightarrow Community

tidymodels / broom

Watch 52 Star 927 Fork 257

Code Issues 13 Pull requests 13 Actions Security Insights

First time contributing to tidymodels/broom?
If you know how to fix an issue, consider opening a pull request for it.
You can read this repository's contributing guidelines to learn how to open a good pull request.

Filters is:pr is:open Labels 9 Milestones 3 New pull request

13 Open 325 Closed Author Label Projects Milestones Reviews Assignee Sort

- nest(data = -keep_col) ×**
#807 opened on Dec 18, 2019 by ChuliangXiao
- Merge in CRAN 0.5.3 release hotfixes ×**
#805 opened on Dec 14, 2019 by alexphghayes
- Add tidy, glance, augment methods for mlogit objects ×**
#794 opened on Dec 2, 2019 by gregmacfarlane
- Add tidiers for `fixest` models ×**
#793 opened on Nov 30, 2019 by karldw • Changes requested
- Improvements of consistency for post-hoc hypothesis test tidiers ×**
#788 opened on Nov 21, 2019 by crsh • Changes requested 9 of 9
- Fix issue 765 ×**
#770 opened on Oct 17, 2019 by jcfisher
- New tidier for pls::mvr object ✓**

- In one year, 200+ Pull requests from others
- In 2018, tidyverse dev team took over as full time maintainer

Steps to contribute

The screenshot shows a website with a header featuring a grey semi-circle graphic on the left. The main title is "How to contribute to open source". Below the title, there's a navigation bar with "Open Source Guides" on the left, and "About" and "Contribute" on the right. The main content area has a large heading "How to Contribute to Open Source" and a subtext: "Want to contribute to open source? A guide to making open source contributions, for first-timers and for veterans.". To the right of the main content is a sidebar with a "Table of Contents" section containing the following items:

- Table of Contents ▲
- Why contribute to open source?
- What it means to contribute
- Orienting yourself to a new project
- Finding a project to contribute to
- How to submit a contribution
- What happens after you submit a contribution

Opening an issue

You should usually open an issue in the following situations:

- Report an error you can't solve yourself
- Discuss a high-level topic or idea (for example, community, vision or policies)
- Propose a new feature or other project idea

Tips for communicating on issues:

- If you see an open issue that you want to tackle, comment on the issue to let people know you're on it. That way, people are less likely to duplicate your work.
- If an issue was opened a while ago, it's possible that it's being addressed somewhere else, or has already been resolved, so comment to ask for confirmation before starting work.
- If you opened an issue, but figured out the answer later on your own, comment on the issue to let people know, then close the issue. Even documenting that outcome is a contribution to the project.

Opening a pull request

You should usually open a pull request in the following situations:

- Submit trivial fixes (for example, a typo, a broken link or an obvious error)
- Start work on a contribution that was already asked for, or that you've already discussed, in an issue

A pull request doesn't have to represent finished work. It's usually better to open a pull request early on, so others can watch or give feedback on your progress. Just mark it as a "WIP" (Work in Progress) in the subject line. You can always add more commits later.

If the project is on GitHub, here's how to submit a pull request:

- Fork the repository and clone it locally. Connect your local to the original "upstream" repository by adding it as a remote. Pull in changes from "upstream" often so that you

Start small with beginner-friendly issue



alex hayes @ rstudio::conf
@alexphayes

looking to get involved but don't know how? check out our list of beginner friendly issues:

github.com/tidyverse/broom ...

if you see one that interests you, leave a comment and i'll walk you through making a pull request!

12:32 PM - 18 Jul 2018

12 Retweets 56 Likes



A screenshot of a GitHub repository page for 'tidyverse/broom'. The main heading is 'Want to contribute to tidyverse/broom?'. Below it, text says: 'If you have a bug or an idea, read the [contributing guidelines](#) before opening an issue. If you're ready to tackle some open issues, [we've collected some good first issues for you](#).'. A callout box is overlaid on the page, pointing to the 'Filter by label' dropdown menu.

A screenshot of the GitHub Issues page for 'tidyverse/broom' with the 'beginner-friendly' label selected. The sidebar shows a list of issues labeled as 'beginner-friendly':

- Error messages instead of output
- Incorrect v0.5.3 tag name
- Use as_tibble() instead of fix_data
- Improve documentation about def
- Implement a `tidy.Anova.mlm()` n
- Implement a tidy method for censl
- Implement a `tidy.cglm()` method
- Implement a `tidy.glmx()` method

The 'beginner-friendly' label is highlighted in the 'Filter by label' dropdown menu.



Contribute by creating packages

- Get other people to use your code
- Practice documentation, unit tests, creating user-friendly interfaces

Building R package from scratch

[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#) [Settings](#)

Branch: master • CSP2019-Rpackage / Building R package from scratch.pdf

amysheep update slides c100c13 on Feb 13, 2019

1 contributor

1.42 MB

Download History

Building R Package Tutorial

Amy Yang |Senior Data Scientist|UPTAKE
Twitter @ayyangalytics
CSP 2019



Pre-work

- R (version >3.5)
- Rstudio

You need the following packages

R packages

This is the book site for “R packages”. It was published with O'Reilly in April 2014.



Packages are the fundamental units of reproducible R code. They include reusable code, data, and documentation. This section describes how to use them and sample data. In this section you'll learn how to 1

Table of contents ▾

Getting started
Introduction
Package structure
Package components
Code (R/)
Package metadata (DESCRIPTION)
Object documentation (man/)
Vignettes (vignettes/)
Testing (tests/)
Namespaces (NAMESPACE)
Data (data/)
Compiled code (src/)
Installed files (inst/)
Other components
Best practices
Git and GitHub
Checking
Release



Aspiring

Junior

Senior

Blog

Tweet

Contribute to open source

Grow the community



Start a group

Meetup

Part of R-Ladies – 181 groups

R-Ladies Chicago

📍 Chicago, IL

👤 1,050 members · Public group

👤 Organized by R-Ladies Global and 8 others



R-Ladies Chicago

Events:

- Tech talks
- Interactive data hackathons
- Social events
- Study groups

Check us out on Twitter:
@RLadiesChicago



R-Ladies Chicago
@RLadiesChicago

We're about to get started learning Shiny with #r spatial! #RLadies



Start a group



Chicago Python Users Group
chipy.slack.com

Channels ⊕

- # algorithms
- # announcements
- # at-chicagopython
- # call-for-papers
- # code_review
- # code-talk
- # community**
- # daily-mentorship
- # data
- # datascience
- # devops
- # evry1-luvs-sivg
- # general**
- # intro
- # jobs
- # mentorship-blogs
- # mentorship-program
- # projectnight
- # py-career-questions
- # pycon



Participate in public Q&A

stackoverflow Products Customers Use cases Search...
Home
PUBLIC
Stack Overflow Tags
Users Jobs
TEAMS What's this?
Free 30 Day Trial

Bayes Asked 4 yrs
R Studio Community

all categories all tags Latest Categories Top
Topic
1 h thr of 10 Lo the 1
Welcome to the RStudio Community!
meta
Welcome to community.rstudio.com — we're glad to have you! This welcome page will give you some advice on how to get the most out of the site if you're getting or giving help. We want this to be a friendly, inclusive com... [read more](#)

Different 10
[Bug] Selected text highlighting conflicts with pre-defined theme color RStudio IDE rstudio
Deep Learning with Keras and TensorFlow in R Workflow - rstudio::conf 2020 Conference Workshops
Can't get data to show up General rstudio
I can't get this plot to work General rstudio
change columns strings to numeric from a list General
renv::restore() in Docker fails to restore packages Package Management renv

ASA Amstat News ASA Community World of Statistics STAT/r@k STATS.org This is Statistics ASA Leader HUB
Join/Renew Code of Conduct Contact Us ASA Home Sign In

R Studio Community Sign Up Log In Search
R Studio Community Sign Up Log In

alistaire Edward Visel Sustainer Chicago, IL [alistaire.rbind.io](#)
A data scientist working in production systems at Uptake. Interests and experience in anomaly detection, politics, and education.

Company: Uptake
Joined Sep 14, '17 Last Post Jul 7, '19 Seen Jan 6 Views 555 Trust Level Sustainer Group Sustainers ...
Summary Activity Badges
STATS
645 days visited 4d read time 3m recent read time 695 topics viewed 4.9k posts read 285 ❤ given 3 topics created
256 posts created 620 ❤ received 30 ✎ solutions
TOP REPLIES
2017-09-16 · ❤ 19 Best practice to handle out-of-memory data
2017-09-14 · ❤ 16 R + Python in IDE
2018-03-10 · ❤ 14 Conditional pipelines
TOP TOPICS
2018-03-30 · ❤ 3 Vignette syntax highlighting in help viewer
2018-06-01 · ❤ 2 Drop shadows on geom_text
2017-09-27 · ❤ 2 Quoting garbles >
action of single regression r^2
and deployment Learn More »
let r_1^2 be the proportion of variance explained by the first variable in the regression model, and let r_2^2 be the proportion of variance explained by the second variable. Is it possible for $r_1^2 + r_2^2 < 1$? Perhaps as a function of $\text{Cov}(X_1, X_2)$? More details
surprised to find $r_{12}^2 > r_1^2 + r_2^2$. This turns out to be true if $\text{Cov}(X_1, X_2) = 0$. I assume the r_i^2 are additive in the sense that they represent the proportion of variance explained by each variable separately. I'm curious if one can quantify the size of the violation of the triangle inequality if X_1 and X_2 are correlated.

Organize a conference



ABOUT US

satRdays are free/cheap accessible R conferences being organised by the community for the community.

I love the R community! I want to help make it as awesome as possible, and that means getting as many people involved as we can. More people participating means more knowledge sharing and more community. I see these weekend, local, super low-cost conferences as a vital way of improving the accessibility of knowledge and growing the community.

STEPH LOCKE

Founder



Running the Hungarian R User Group was always such a good and rewarding experience, so it wasn't a question at all if I'm willing to help with the baby steps of the project and kicking off the first satRday conference. It was quite a lot of work, but the almost 200 registrations from 19 countries and 25 awesome speakers providing 12+ hours of fun and educating experience in a good company totally worthed it!



GERGELY DAROCZI

Main organizer of the
Budapest 2016 event



I had the privilege of being on the organizing committee for the 2nd satRday which took place in Cape Town. It was incredibly exciting for me to have the opportunity to meet with passionate R users and developers - both local and international - right on my doorstep. It was definitely one of my highlights for 2017!

JON CALDER

Co-organizer of
satRday Cape Town



The R community is a place where everybody is welcomed. The tool we use unites us, and the diversity of our backgrounds is its power. Joining the satRday Berlin team has been a heartwarming experience!

SONIA BOENDER

Co-organizer of
satRday Berlin





Aspiring

Junior

Senior

Blog

Tweet

Contribute to open source

Grow the community

Give talks



Embarrassingly Parallel R: Getting Started with Parallelization in R

Counterfactual Analysis of Cross-Sectional Data Using Quantile Process Regression

Statistical Image Processing for Machine Learning

Bayesian Analysis of Sparse Multivariate Matched Proportions Data



A talk isn't a tutorial or a resource,
It's a sales pitch

It's an opportunity to convince people that this is
something worth looking into



“What if I’m an
introvert?”



THEY SPEAK
IN MIND



Kevin Goldsmith

@KevinGoldsmith

Follow

introvert tip: One of the reasons I like speaking at conferences is because it makes networking a lot easier. People know who you are and have something to talk to you about.

Seems counterintuitive, but one of the best things for introverts is to become speakers!

2:01 PM - 14 Feb 2018



Emily Robinson

@robinson_es

Following



Best advice for reducing nerves before giving a data science talk: "just imagine that your audience uses p-values of .25 for significance"

5:14 PM - 27 Jul 2017

Copyright 2009 by Randy Glasbergen.
www.glasbergen.com



**"Remember, you'll be speaking to a group
of military generals. It's okay if you bomb."**



meetup



rstudio::conf

JSM



David Robinson @drob · 8 Nov 2017

When you've written the same code 3 times, write a function

When you've given the same in-person advice 3 times, write a blog post

48

1.5K

4.2K



Hadley Wickham ✅

@hadleywickham

▼

Replying to @drob

I'd say write a book but whatevs

8:08 PM - 8 Nov 2017



Aspiring

Junior

Senior

Blog

Tweet

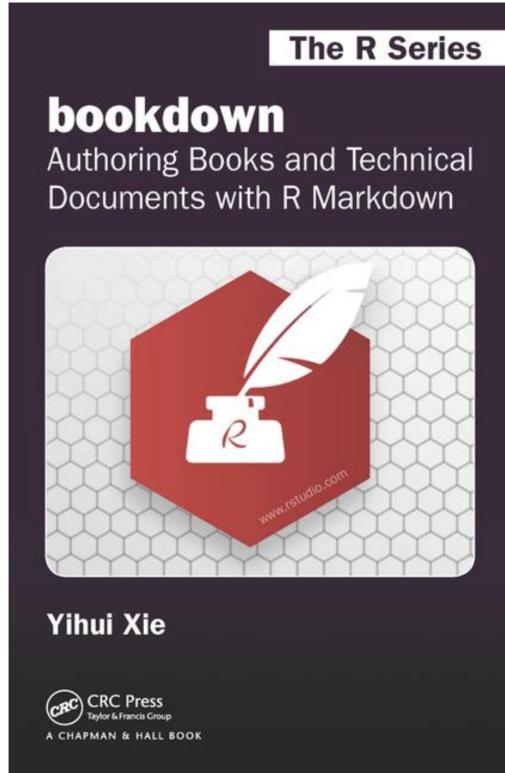
Contribute to open source

Grow the community

Give talks

Other

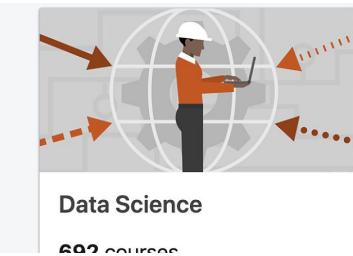
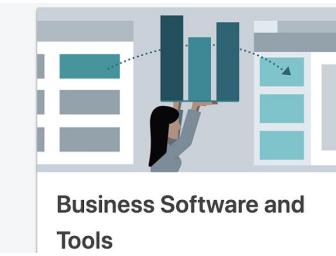
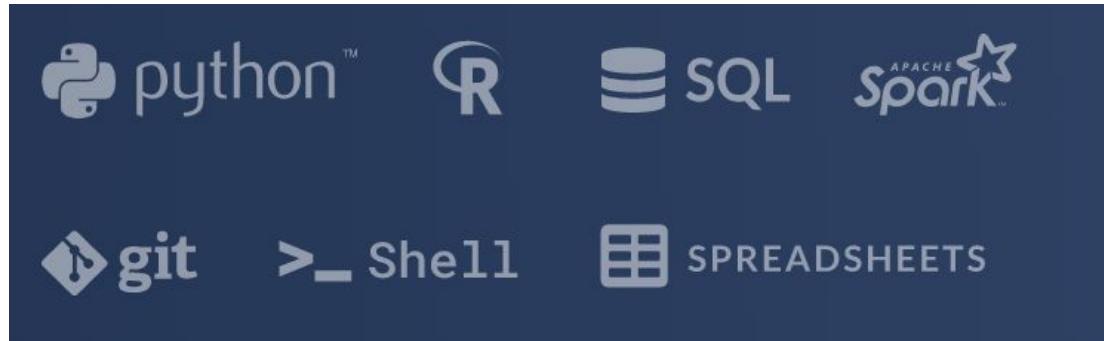
To write a book, you need **a good amount to say**
and **some practice in saying it**



[Bookdown by yihui xie](#)



Teach a course or workshop





About me

Biostatistician

Industry Data Scientist

Experimentation Data Scientist



R-Ladies

Blog/Tweet/Talk

Organizing conference/Building community

Acknowledgment

Hi David, I really liked your keynote on the effectiveness of public work and plan to give a talk on this same topic. Is that okay I borrow some slides of yours? If so, what's the best way to credit you during my talk?

Dec 31, 2019, 2:34 PM ✓

 Hi- of course you can. Just mention my talk somewhere in the talk. Is the format of the slides convenient for you?

Dec 31, 2019, 3:32 PM

 Hi Ed, Happy new year! I am preparing a talk on public work and open source and thought of you and how you used blog and other media to showcase your skills. Is that okay that I use you as an example in my talk?

Jan 1, 2020, 9:17 AM ✓

 Of course! Hope it goes well!

Jan 1, 2020, 3:28 PM



It's all about sharing

