

Social media for data scientists

Karl Broman

Biostatistics & Medical Informatics, UW–Madison

kbroman.org

github.com/kbroman

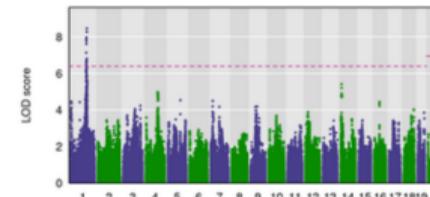
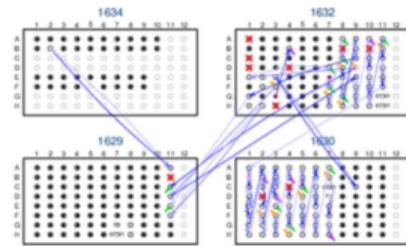
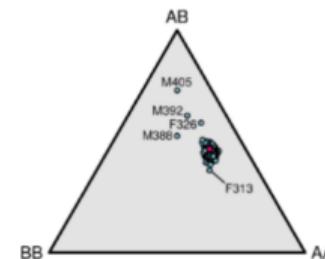
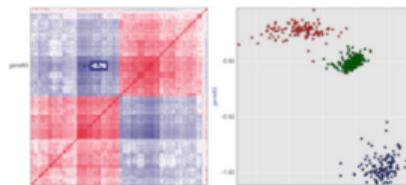
@kwbroman

Slides: kbroman.org/Talk_SocialMedia



Everyone should have a website

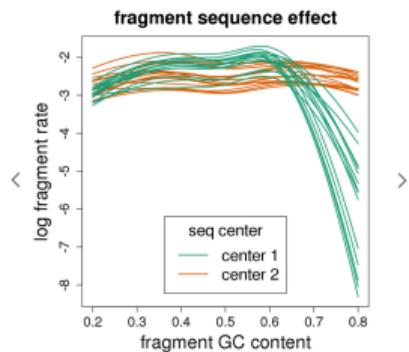
- ▶ How can I get ahold of that person I met?
- ▶ I'd like to see more of that cool thing I saw.
- ▶ Check out this cool talk I saw.
- ▶ What is X up to these days?
- ▶ We should invite X to give a talk.
- ▶ We should invite X to apply for this job.

[more figures](#)

Rafael Irizarry is Professor and Chair of the [Department of Data Sciences](#) at Dana-Farber Cancer Institute and Professor of Applied Statistics at Harvard. His research focuses on Genomics and he teaches several [Data Science courses](#).

This page provides information on his research and teaching activities. Blog posts are available at [Simply Statistics](#).

The best way to keep up with the latest news is to follow
[@rafalab](#).





ALEATORIC

BLOG ABOUT RESUME

LATEST POSTS

Sep, 2019

Sep 3, 2019 [announcing the distributions3 package](#)

Aug, 2019

Aug 31, 2019 [consistency and the linear probability model](#)

Jul, 2019

Jul 26, 2019 [an annotated bibliography on stochastic blockmodels](#)

Jun, 2019

Jun 7, 2019 [testing statistical software](#)

May, 2019

May 21, 2019 [type stable estimation](#)

[See more ...](#)



alexphayes.com

I am a research scientist in data curation at [Sage Bionetworks](#), where I help other researchers document and share their data. I've previously worked as an information manager at Washington State University and at the National Center for Ecological Analysis and Synthesis (NCEAS), where I combined data management with fieldwork at a remote Siberian lake.

I'm an enthusiastic R programmer, and I collect data visualizations gone beautifully wrong on a blog called [accidental aRt](#). When I'm not writing R code I can often be found in a ballet studio.



Follow me on [Twitter](#)



Hi! I'm Angela. I'm a data scientist, spatial thinker, and community builder.



I'm currently working as the R Spatial Advocate for the [Center for Spatial Data Science](#) at UChicago. Most recently, I did data science for transit at [TransLoc](#) and wrote a spatial econometrics thesis on the Detroit housing market.

In my day to day, I work closely with [Luc Anselin](#), leading [workshops](#) for researchers at our center and building out [spatial analysis resources](#) for applied social science research. I also [speak at conferences](#) about spatial data science and methods.

I founded [R-Ladies Chicago](#), which is the Chicago branch of a global organization whose mission is to promote gender diversity in the R community. We started in July 2017 and have expanded to 800+ members so far, putting on [20+ events](#) for the Chicago R community. Other organizations I'm affiliated with include [Chi Hack Night](#), [Software Carpentry](#), [rOpenSci](#), and [NASA Datanauts](#).

I'm currently in the process of organizing three events for the Chicago R community, including a [women's package workshop](#), an [unconference](#), and a [1-day conference](#). I'm always happy to chat about my experiences organizing these type of things (and more), so feel free to [schedule a chat](#) to meet in-person or online.

You can also find me on Twitter [@CivicAngela](#), where I post about #rstats, #rspatial, and #RLadies (and occasionally live-tweet a conference). Off the web, you can find me [doing a crossword](#), playing a game of ultimate frisbee, cooking something tasty, or traversing Chicago via bike, bus, or train.

Cora Allen-Coleman

Statistics PhD Candidate with 7 years of theoretical and applied statistics experience and 2 years of software development experience. Currently developing software in Julia for efficient search & model comparison in complex network space.

[View My GitHub Profile](#)



[resume](#) [cv](#) [linkedin](#) [mendeley](#)

Current Research

University of Wisconsin-Madison Department of Statistics

Developing software in Julia for efficient full likelihood-based network methods with Professor Cécile Ané, as part of [PhyloNetworks Julia Package](#)

University of Wisconsin-Madison Department of Biostatistics

Developing [ClusterRank R Package](#) for mixture model clustered ranking with Professor Ron Gangnon. Uses inverse-variance weighted ranking and mixture models to produce simultaneous complete, optimal rankings with clustering, error estimates, and visualizations.

Applied Data Science Projects

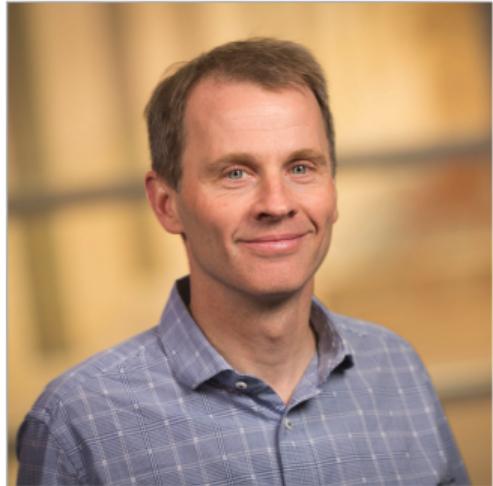
University of Wisconsin-Madison Department of Epidemiology

Lead analyst for a spatial analysis of breast cancer incidence and survival using generalized additive models, multiple imputation, and variable width smoothing techniques. *Paper In preparation*

University of Wisconsin-Madison Department of Medicine

Lead analyst for a study of UW's MD/PhD training program. Used a variety of multivariate generalized linear mixed effects models and creative visualizations to answer a diverse set of applied questions. *Paper under review at Academic Medicine*

coraallencoleman.github.io



Mark Craven

Professor

Department of Biostatistics and Medical Informatics

Department of Computer Sciences

University of Wisconsin-Madison

[Home](#)

Research

Group

Publications

Teaching

Contact

I am a professor in the Department of Biostatistics and Medical Informatics at the University of Wisconsin, and an affiliate faculty member in the Department of Computer Sciences. I am the Director of the Center for Predictive Computational Phenotyping, one of the NIH's Centers of Excellence for Big Data Computing. I am also the Director of the NIH/NLM-funded Computation and Informatics in Biology and Medicine (CIBM) Training Program, and a member of the Institute for Clinical and Translational Research, the Carbone Cancer Center, and the Center for Genomic Science Innovation.

I am also involved in the large and vibrant machine learning and quantitative biology research groups at the University of Wisconsin.

www.biostat.wisc.edu/~craven

What to include

- ▶ Who you are
- ▶ How to contact you
- ▶ Links to your other sites (e.g., GitHub and Twitter)
- ▶ What you have been up to
- ▶ Resume/CV

What to include

- ▶ Who you are
- ▶ How to contact you
- ▶ Links to your other sites (e.g., GitHub and Twitter)
- ▶ What you have been up to
- ▶ Resume/CV

Don't include: “under construction”

How to make a website

- ▶ html-based page at biostat.wisc.edu
- ▶ [Google sites](#)
- ▶ [Github pages](#) (uses [jekyll](#))
- ▶ [Blogdown](#) (uses [hugo](#))
- ▶ Create your own domain, e.g. at [Google domains](#)

How to make a website

- ▶ html-based page at biostat.wisc.edu
- ▶ Google sites
- ▶ Github pages (uses jekyll)
- ▶ Blogdown (uses hugo)
- ▶ Create your own domain, e.g. at Google domains

Does it look okay on a phone?

Blogs make the world better

Why blog?

- ▶ Practice writing
- ▶ Share small, useful things you've figured out
- ▶ Present side projects that could be really interesting to others
- ▶ Connect with people



the stupidest thing...



2019

[data analysis principles](#)

2019-08-15

[reproducibility guidelines](#)

2019-04-01

[omg, binder!](#)

2019-02-18

[writing text in Minecraft with R](#)

2019-02-11

[another new laptop](#)

2019-02-02

2018

[twitter break](#)

2018-10-05

[learning a new programming language](#)

2018-03-16

kbroman.org/blog

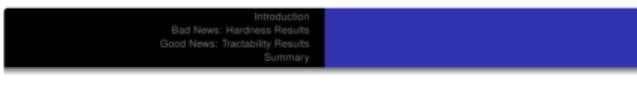
Better-looking LaTeX/Beamer slides

6 min read

2013-10-07

I like to use LaTeX to make slides for talks, largely because I prefer to write code (rather than use a mouse and menus) for control of things like colors and figure placement.

Most people that use LaTeX to make slides seem to use Beamer, but the resulting slides are usually a bit busy, like this:



On the Complexity of SNP Block Partitioning Under the Perfect Phylogeny Model

Jens Gramm¹ Tzvika Hartman² Till Nierhoff³
Roded Sharan⁴ Till Tantau⁵

¹Universität Tübingen, Germany

²Bar-Ilan University, Ramat-Gan, Israel

³International Computer Science Institute, Berkeley, USA

⁴Tel-Aviv University, Israel

⁵Universität zu Lübeck, Germany

Workshop on Algorithms in Bioinformatics, 2006

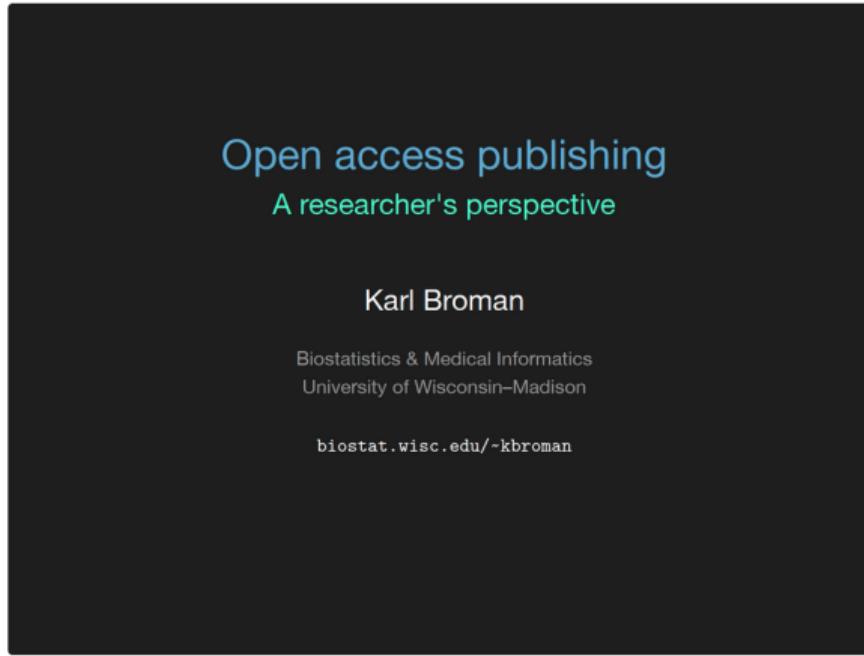


Better-looking LaTeX/Beamer slides

6 min read

2013-10-07

After several days work, I've finally figured out how to create LaTeX/Beamer-based slides that look like what I want:



In this post, I'll explain what I did.

Crayon colors in R

3 min read

Last night I was working on a [talk on creating effective graphs](#). Mostly, I needed to update the colors, as there'd been some gaudy ones in its previous form (e.g., slide 22).

I usually pick colors using the crayons in the [Mac Color Picker](#). But that has just 40 crayons, and I wanted more choices.

That led me to the [list of Crayola crayon colors on wikipedia](#). I wrote a ruby script to grab the color names and codes and added them to my R/broman package.

Use `brocolors("crayons")` to get the list of colors. For example, to get "Tickle Me Pink", use

```
library(broman)  
pink <- broman::caveman1[["Tickle_Me_Pink"]]
```

Use `plot_cravons()` to get the following summary plot of the colors:



Thoughts on statistical consulting

8 min read

2013-04-02

The [Statistics Department](#) at UW-Madison has a course on statistical consulting, offered each semester. I'm often asked to give a lecture, which I do in an informal way: summarizing my experiences and answering questions.

I thought it might be useful write my thoughts on statistical consulting here: why, how, and difficulties. This will be a bit rough, and long. I'll revert to bullet points, to be more compact (and because I'm lazy).

My situation

I'm an applied statistician at a university. I have no defined consulting responsibilities. Others at the university have a defined consulting role for a department or school, which pays part of their salary. And then there are statisticians in industry (some effectively doing statistical consulting within a company), or working for a consulting company or as a contractor; I don't know much about these things, but only about my own situation: I feel I should do my best to help colleagues who ask for my advice (though I *can* say no, at least in theory).

MongoDB with D3.js

6 min read

2015-06-22

I consider interactive data visualization to be the critical tool for exploration of high-dimensional data.

That's led me to spend a good amount of time in the last few years learning some new skills ([D3](#) and [CoffeeScript](#)) and developing some new tools, particularly the R package [R/qtlcharts](#), which provides interactive versions of the many data visualizations in [R/qtl](#), my [long-in-development](#) R package for mapping genetic loci (called quantitative trait loci, QTL) that underlie complex trait variation in experimental organisms.

[R/qtlcharts](#) is rough in spots, and while it works well for moderate-sized data sets, it can't well handle truly large-scale data, as it just dumps all of the data into the file viewed by a web browser.

For large-scale data, one needs to dynamically load slices of the data based on user interactions. It seems best to have a formal database behind the scenes. But I think I'm not unusual, among statisticians, in having almost no experience working with databases. My collaborators tend to keep things in Excel. Even for quite large problems, I keep things in flat files.

apply vs for

1 min read

2013-04-02

It's widely understood that, in R programming, one should avoid `for` loops and always try to use `apply`-type functions.

But this isn't entirely true. It may have been true for Splus, back in the day: As I recall, that had to do with the entire environment from each iteration being retained in memory.

Here's a simple example:

```
> x <- matrix(rnorm(4000*40000), ncol=4000)

> system.time({
+   mx <- rep(NA, nrow(x))
+   for(i in 1:nrow(x)) mx[i] <- max(x[i,])
+ })
  user  system elapsed
 3.719   0.446   4.164

> system.time(mx2 <- apply(x, 1, max))
  user  system elapsed
 5.548   1.783   7.333
```

There's a great commentary on this point by Uwe Ligges and John Fox in the May, 2008, issue of R News (see the "R help desk", starting on page 46, and note that R News is now the [R Journal](#)).

"The Role of Nature versus Nurture in Wealth and Other Economic Outcomes and Behaviors"

Posted by [Andrew](#) on 22 October 2019, 1:31 pm

Sandra Black, Paul, Devereux, Petter Lundborg, and Kaveh Majlesi [write](#):

Wealth is highly correlated between parents and their children; however, little is known about the extent to which these relationships are genetic or determined by environmental factors. We use administrative data on the net wealth of a large sample of Swedish adoptees merged with similar information for their biological and adoptive parents. Comparing the relationship between the wealth of adopted and biological parents and that of the adopted child, we find that, even prior to any inheritance, there is a substantial role for environment and a much smaller role for pre-birth factors and we find little evidence that nature/nurture interactions are important. When bequests are taken into account, the role of adoptive parental wealth becomes much stronger. Our findings suggest that wealth transmission is not primarily because children from wealthier families are inherently more talented or more able but that, even in relatively egalitarian Sweden, wealth begets wealth. We further build on the existing literature by providing a more comprehensive view of the role of nature and nurture on intergenerational mobility, looking at a wide range of different outcomes using a common sample and method. We find that environmental influences are relatively more important for wealth-related variables such as savings and investment decisions than for human capital. We conclude by studying consumption as an overall measure of welfare and find that, like wealth, it is more determined by environment than by biology.

Black is speaking at the sociology department [this Thursday](#). The topic seems important.

RECENT COMMENTS

> Anoneuoid on "Starting at the beginning again can be exhausting and stressful. But, opportunities are finally coming into focus . . ."

> Daniel Lakeland on "Starting at the beginning again can be exhausting and stressful. But, opportunities are finally coming into focus . . ."

> Anoneuoid on "Starting at the beginning again can be exhausting and stressful. But, opportunities are finally coming into focus . . ."

> Anoneuoid on "Starting at the beginning again can be

[Home](#)[About](#)[Archive](#)[Conferences](#)[Courses](#)[Interviews](#)[Contributing](#)[RSS](#)[Twitter](#)[GitHub](#)

© 2011 - 2017. All rights reserved.

Built with [blogdown](#) and
Hugo. Theme [Blackburn](#).

Simply Statistics

A statistics blog by Rafa Irizarry, Roger Peng, and Jeff Leek

You can replicate almost any plot with R

• Rafael Irizarry 2019/08/28

Although R is great for quickly turning data into plots, it is not widely used for making publication ready figures. But, with enough tinkering you can make almost any plot in R. For examples check out the [flowingdata](#) blog or the [Fundamentals of Data Visualization](#) book. Here I show five charts from the [lay press](#) that I use as examples in my data science courses. In the past I would show the originals, but I decided to replicate them in R to make it possible to generate class notes with just R code (there was a lot of googling involved).

[Read more »](#)

So You Want to Start a Podcast

• Roger Peng 2019/08/27

Podcasting has gotten quite a bit easier over the past 10 years, due in part to improvements to hardware and software. I wrote about both how I edit and record both of my podcasts about 2 years ago and, while not much has changed since then, I thought it might be helpful if I organized the information in a better way for people just starting out with a new podcast.

simplystatistics.org

"Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write." – H.G. Wells

October 23, 2019

Publicise it while it still works?

By Thomas Lumley

This is from *The Times*. In fact, per the BBC, it's the leading story [on the front page](#) of tomorrow's *Times*, above Brexit.

Doctors hail first drug to 'slow down' Alzheimer's

New treatment may reduce mental decline by 25%

Tom Whipple Science Editor
James Dean US Business Editor

Doctors have welcomed a "turning point" in the treatment of Alzheimer's disease after the first drug was found apparently able to slow the condition if it is caught in its earliest stages.

It would be the first effective treatment for Alzheimer's disease since dementia validating billions of pounds of research and decades of work.

But some experts said more information was needed and they were awaiting the opinion of US regulators, who are in talks with the drug's owner.

However, other experts said more information was needed and they were awaiting the opinion of US regulators, who are in talks with the drug's owner.

Anton Porsteinsson, director of the University of Massachusetts Alzheimer's program in New York state and lead investigator in the study, said it provided new hope for families. "There is now a drug that can help now, and the Alzheimer's disease community has been waiting for this moment."

Both the UK's main Alzheimer's charity and the Alzheimer's Research UK announced the news. Alzheimer's have waited a long time for a life-changing new treatment and this exciting announcement offers new hope that there is finally some progress.

Dr James Pickett, head of research at the Alzheimer's Society, said that while they were looking forward to seeing more data, the drug had "the potential



Recent comments

Matthew Rudling: Thank you for the response! I was pretty much on the fence with this one. Although I don't feel...

David Scott: Because their last game was nearly a year ago, when their ratings were very similar in my rating calculations (about...

Matthew Rudling: Why would South Africa beat Wales? In the last three matches Wales beat them.

Popular posts

Rugby World Cup Predictions 3

Rugby World Cup Predictions 4

Predictions for the RWC Quarter-Finals

Latest posts

statschat.org.nz



2019

[Some things I don't like about the Oxford-Munich Code...](#) 2019/10/01

[How to review a book](#) 2019/09/13

[Why isn't rimu tidy?](#) 2019/09/10

[\(What's up with the brackets?\)](#) 2019/09/10

[A package for multiple-response data](#) 2019/09/05

[Adding new functions to the survey package](#) 2019/07/16

[Denominator degrees of freedom in svyglm](#) 2019/06/26

[Wald, score, LRT: the picture](#) 2019/06/20

Julia Evans

ABOUT
TALKS
PROJECTS
TWITTER
GITHUB

FAVORITES ★ ZINES ★ RSS

Hey! I'm Julia. Welcome to my blog. Here's every post I've ever written, organized by category. Enjoy! You can subscribe to a [weekly digest of these blog posts](#).

I also publish computer zines at [Wizard Zines](#).

Most recent 10 posts

- Oct 21 2019 Print box set of my free zines
- Oct 3 2019 SQL queries don't start with SELECT
- Oct 1 2019 Zine revenue for 2019
- Sep 30 2019 Notes on building SQL exercises
- Sep 13 2019 Taking a year to explain computer things
- Sep 12 2019 New zine: HTTP: Learn your browser's language!
- Sep 6 2019 How to put an HTML page on the internet
- Sep 1 2019 How to write zines with simple tools
- Aug 30 2019 git exercises: navigate a repository
- Aug 27 2019 curl exercises

and all posts by category:

Cool computer tools / features / ideas

- Nov 2018 Tailwind: style your site without writing any CSS!
- Sep 2018 An awesome new Python profiler: py-spy!
- May 2018 Batch editing files with ed

jvns.ca



Stat Girl
Here I Come to Save Your Data!

ABOUT

Search ...

RECENT POSTS

[Baby, Baby, Baby Oh*](#)

[Why Tri?](#)



Of course I'm a good mother.
They're still alive,
aren't they?

your eCards
someecards.com

Baby, Baby, Baby Oh*

AUGUST 6, 2019 STATGIRLBLOG 1 COMMENT

My daughter recently turned 16. It's unbelievable that 16 years have passed since my life changed forever. It got me thinking about the challenges of being a parent, which led to thoughts about the challenges of being a mother – especially during graduate school.

When she was born, I was still a graduate student. My

statgirlblog.wordpress.com



Yihui's Blog

R · statistics · life · random thoughts

2019

2019-07-16 [Romain Lesur: The Most Wonderful Collaborator I Have Ever Worked With](#)

2019-06-28 [On TIOBE Index](#)

2019-05-22 [From RTFM to ITFM \(Improve The Fine Manual\)](#)

2019-05-01 [August 12, 1991 Review of S-PLUS Statistical Software](#)

2019-04-30 [Naming Software Packages with Common Words](#)

2019-03-06 [The Implementation of `yolo = TRUE` in xaringan via yolofy\(\)](#)

2019-02-20 [The Ultimate Infinite Moon Reader for xaringan Slides](#)

Concerns

- ▶ Putting yourself out there
- ▶ The web is forever
- ▶ People can be cruel or even evil, especially to women
- ▶ Feeling guilty for not posting

Twitter is not entirely bad

Why twitter?

- ▶ Publicize your stuff (blogs, papers, projects)
- ▶ Learn about new things
- ▶ Get help (e.g. with programming)
- ▶ Find a crew of people who share your interests
- ▶ Complain about stuff

Homework 4 - Karl Broman's Socks

 Karl Broman
@kwbroman



Following

That the 1st 11 socks in the laundry are each distinct suggests there are a lot more socks.

 Reply  Retweet  Favorite  More





David Robinson @drob · Jan 14, 2017
Best xaringan feature from @xieyihui:

If you set `yolo = TRUE`, it inserts a slide with a random picture of
[@kwbroman](#)

#studioconf

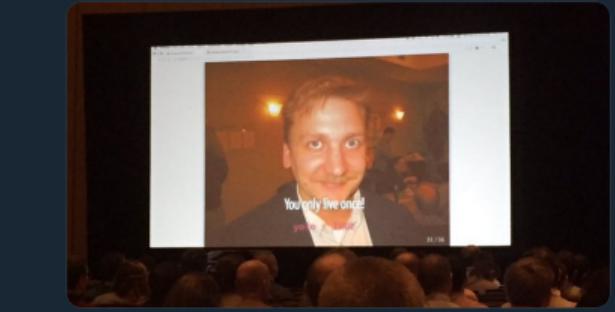




David Robinson @drob · Jan 14, 2017
Best xaringan feature from @xieyihui:

If you set `yolo = TRUE`, it inserts a slide with a random picture of
[@kwbroman](#)

#studioconf



Karandeep Singh
@kdpsinghlab

⬇️ output:

```
xaringan::moon_reader:  
yolo: true
```



Karl Broman @kwbroman · Oct 2

Replying to @kdpsinghlab @venkmurthy and 7 others
We need a Nature Onion twitter account.



Amelia MN

@AmeliaMN

Stats & data science professor at @UofStThomasMPL
@UCLA_MSAccelerate alumna. Info data visualization and open science, obsessed with R. Ironous is her.

Minneapolis, MN

amelia.mn



Colin Fay 🌟

@_ColinFay

I write code at @ThinkR_FR

Rennes, France

colinfay.me



Mara Averick

@tataandme

tidyverse 🐍 @studio, 🎵 hoop head, gnashgab, blatherlike, lesser % of @patigandie 🎩 🎤 shelter

Massachusetts

maraaverick.rbind.io



Hadley Wickham

@hadleywickham

R, data, visualisation, 📈, 📈, 🌈
Heihub

Houston, TX

hadley.nz



Hilary Parker

@hspster

Data Scientist @StitchFix, co-host of @NSTSEdevations #stats #rcatladies

San Francisco

hilaryparker.com



Karthik Ram

@mundata

Data Scientist at Berkeley Institute for Data Science @ics.berkeley.edu. Scientist at Berkeley Initiative for Global Change Biology. Co-founder + lead @ropensci

Yay Ana! 🎉 in transit

mundata.org



Jenny Bryan

@JennyBryan

Software engineer @studio, humaine #stats, adjunct prof @UBC where I created @STAT45, part of @ropensci

Vancouver, BC

jennybryan.org



Kara BOO! 🌸

@karawoo

Research scientist at @sagebio. Data curation, visualization, #stats, reproducibility, open science, ballet, shelter

karawoo.com



Lucy:assistant_prof(✓)

@LucyStats

stats assistant professor
@WakeForest 🌸 new mama 🌸
formerly @subrosal @wendy_bostic
@unc 🌸 co-founder @ladies_Nashville
心脏病, 🌸, disney, optimism, #stats
lucymcgowan.com



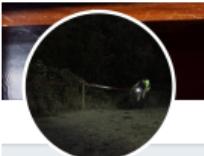
Mine CetinkayaRundel

@minebook

@EdTechHigh • @instudio • @DukeU, @OpenIntroOrg, Analyst, data sciences+stats ed., #dataviz, #rladies EDI, @datavizcat, cat videos = instant smiles. Shelter.

Edinburgh, Scotland

mine-ci.com



Scott

@scotttle

programmerologist • @ropensci co-founder

just@db

scottchamberlain.info



Susan Holmes

@SherlockHolmes

Statistician. Mother. Prof. Grandmthr. It is a panel mistake to theorize before one has data. Insensitivity one begins to twist facts to suit theories? Sherlock 🕵️

Stanford

susan@stanford.edu=susan/



Stephanie Hicks

@stephaniehicks

Mathematics, Engineering, Resistance, Activism, Research, Diversity, STEM advocacy, #H4UDataScience Lab, #Correspondent co-host, #rladiesBronx co-founder, mom, shusher

Baltimore, MD

stephaniehicks.com



Thomas Lin Pedersen

@thomaspl5

Visualisation and Data Science @studio.

What's about it on data-imagrati.com.

Generative Art. Support:

patreon.com/thomaspl5

paypal.me/thomaspl5

Copenhagen, Denmark

github.com/thomaspl5



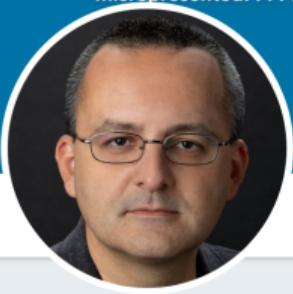
Thomas Lumley

@tstatm

Statistics, Computing, Genetics, Choral music. Product of more than one country. May contain nuts.

Auckland (mostly)

notstatschat.rbind.io



Alberto Cairo

@AlbertoCairo

Knight Chair at the U of Miami. NEW BOOK: 'How Charts Lie' (W.W. Norton). Previous: The Truthful Art (2016). Consultant and freelancer
albertocairo.com

⌚ USA, Brazil, Spain

🔗 thefunctionalart.com



Lynn Cherny

@arnicas

Data science, data vis, python, r, JS/D3.js, NLP, analytics; gender, science fiction, TV. Ex-researcher/academic with Stanford PhD.

📍 Lyon, France

🔗 ghostweather.com/bio.html



Nadieh Bremer

@NadiehBremer

Freelancing Data Visualization Designer. Graduated Astronomer. Love D3 | R | hexagons | travel | cats | paper art.
instagram.com/nadiehbremer

⌚ Amsterdam, The Netherlands

🔗 visualcinnamon.com



Angela Bassa 🎩🦉✍

@AngeBassa

I try to make sure we math on robot data right. Head of Data Science, Data Engineering, and Machine Learning @iRobot. Mathematician from @MIT. Opinions mine.

📍 Boston, MA

🔗 angelabassa.com/faq



Caitlin Boo-don 🎃

@beeonaposy

Data, traveling, and tacos. Cubs fan. Lead Data Scientist @onlinemeded. Co-organizer of @fadiesaustin, #pydata, #rstats, #NASADalanauts

📍 Austin, TX

🔗 caitlinhudon.com



**Casey Fiesler, PhD, JD,
geekD**

@cfiesler

Faculty in @CUInfoScience by way of @gtcomputing & @vanderbiltlaw. Studying social computing, law, ethics, fandom. Max level geek. Views/opinions my own!

📍 Boulder, CO

🔗 caseyfiesler.com



Greg Wilson

@gwwilson

Son of two immigrants. Married to an immigrant. Come a long way, got a long way to go.

🔗 third-bit.com



Hilary Mason ✨

@hmason

Working on something new. I ❤️ data and cheeseburgers.

📍 NYC

🔗 hilarymason.com



Megan Risdal 🧸

@MeganRisdal

Product Manager @StackOverflow. Former Product Lead @Kaggle. Xoogler @GCPCloud. Data/Science/Language. Weirdness. Minnesotan in Los Angeles. Tweets mine.

📍 Los Angeles

🔗 meg.dev



yo yehudi 🏳️🌈

@yoylehudi

I make software. Heart open*, sci-fi, cycling, running, veggies. 🇬🇧🇪🇺🇫🇷🇩🇪. Also @codeisscience & @intermineorg. @softwaresaved Fellow. Cat butler 🐱

📍 Cambridge, UK

🔗 yo-yehudi.com



Nancy Sims

@CopyrightLibn

Copyright Libn at U of MN. JD/MLS.
~50% tech/lib/acad law; ~40% other
leftish politics, ~10%
dogs/bugs/space/crafts. Any pronouns.
simsjd.com/copyrightlibn

⌚ Minneapolis, MN

🔗 lib.umn.edu/about/staff/nas...



Don Moynihan

@donmoyn

Professor at Georgetown, Visiting
Professor at Oxford & Aarhus. Immigrant.
Views my own. New book on
administrative burdens
russellsage.org/publications/a...

⌚ Washington DC

🔗 gufaculty360.georgetown.edu/s/contact/0033...



Kristin Briney

@KristinBriney

Biology librarian, ISO 8601 enthusiast,
data nerd, privacy advocate, tea drinker,
author of "Data Management for
Researchers": bit.ly/18ytZXL
[she/her/hers]

⌚ Los Angeles, CA

🔗 dataab initio.com



Hank Reichman

@HankReichman

Chair AAUP Committee A on Academic
Freedom and Tenure; Chair AAUP
Foundation. Author, *The Future of
Academic Freedom* (Johns Hopkins
University Press, 2019).

⌚ Albany, CA

🔗 academeblog.org/author/hankrei...



Hot Jupiter Wade

@TheGnarrator

Library Barbarian // Ambient Sound
Explorer // Hair Maven // Woman On The
Edge of Time // Migrainess //
She/Her/That Bitch

⌚ MKE

🔗 seadoubleyew.com



Nick Fleisher

@nickfleisher

Assoc prof @uwmling, pres @aaupwi.
Mostly syntax, semantics, WI politics,
higher ed. Views mine and possibly yours
too. Blog: languagepolitics.org

⌚ Milwaukee

🔗 people.uwm.edu/fleisher/

**Eva**

@evacide

Director of Cybersecurity @EFF / Tech advisor at @freedomofpress @callisto @CLTCBerkeley and more! My tweets are my own

eff.org

**Hannah Recht**

@hannah_recht

Data journalist at Bloomberg
@BBGVisualData. Policy, health, Census, R, gifs.

New York, NY

hrecht.github.io

**Lisa**

@meandertail

7% human, 1% caffeine, 1% EtOH, 1% capsaicin, 90% microbiome. These are my own views (my employer doesn't have opinions about The Bachelor).

Tallahassee, FL

**Laura Williams**

@MicroWavesSci

I am a microbiologist who uses genomics to study ecology and evolution of the littlest life. Assistant professor, Providence College

williamsmicrobegenomelab.wordpress.com

**Amy Lee**

@minisciencegeek

80% Microbiologist, 10% foodie/baker, 5% ceramic artist and 5% knitter. Loves data, food and art.

**Rani Molla**

@ranimolla

Data Editor @Recode / @voxdotcom making information visual. @Bfly / @business @WSJ @Gigaom @Columbiajourn alumna. Generalist.

Brooklyn

recode.net/authors/rani-m...

**Melissa A. Wilson**

@sexchrlab

Studying sex chromosomes, populations & evolution. Let's talk science. mathbionerd.blogspot.com

Tempe, AZ

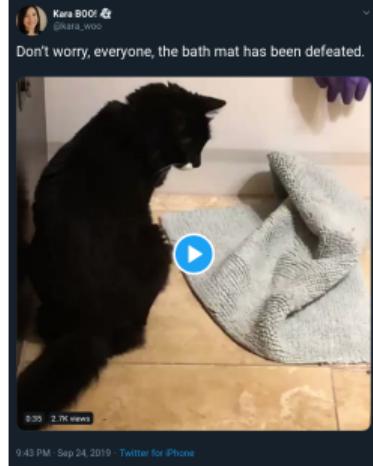
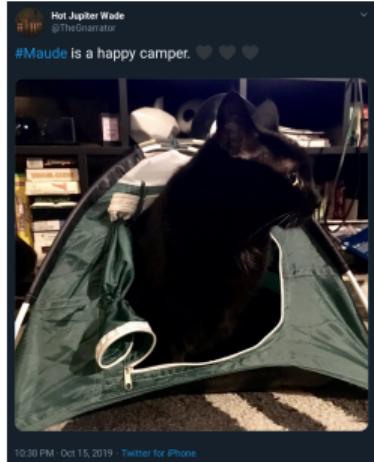
sexchrlab.org

**Myfanwy Johnston**

@Vooverb

Expect hilarious R code. #rstats

cats



cheese

Molly Beck ✅
@MollyBeck

"Bitch, I moved to the Midwest for the cheese curds."
@lizzo said in Madison on Thursday.



10:55 PM · Oct 10, 2019 · Twitter for iPhone

Valerie Schremp Hahn ✅
@valeriehahn

My 9-year-old daughter has taken an old lip balm tube and filled it with cheese so she can eat it in class.



6:47 PM · Sep 17, 2019 · Twitter for iPhone

Karl Broman
@kwbroman

oh my god we're out of cheese. this is a disaster.



4:14 PM · Jun 5, 2019 · Tweetbot for iOS

Karl Broman
@kwbroman

omg there these huge trays full of cheese #rstudioconf



6:24 PM · Jan 16, 2019 · Tweetbot for iOS

software

Katie Jolly
@katiejolly3

I ❤️ WPA style National Park posters and often look for color schemes from them so I decided to make an `#rstats` package! I'll write a blog post with examples soon, but for now feel free to try them out and let me know what you think! 18 to choose from! 🎉

github.com/katiejolly/nat...



11:45 AM - Oct 22, 2019 · Twitter for Phone



Lisa Charlotte Root
@lisacroot

New article! I wrote down the Excel techniques & formulas I use the most to clean up data sets (so that I can use them for `#dataviz`). Which techniques are you using all the time? Which ones did I miss? Which ones should I add?

A1 Look at the data & understand it
A2 Rename your file
A3 Duplicate the data sheet(s)
A4 Save your source in an extra sheet
B1 Delete everything above the header
B2 Get rid of double-row headers
B3 Bring metrics in the header
B4 Check if the header content make sense
C1 Delete columns you don't need

How to prepare your data for analysis and charting in Excel & Google Sheets
The methods and Excel formulas you'll need most often to get your data ready for analysis or plug into Datawrapper.
blog.datarwrapper.de

7:02 AM - Oct 23, 2019 · Twitter Web App

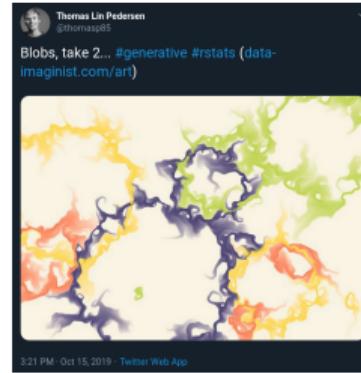
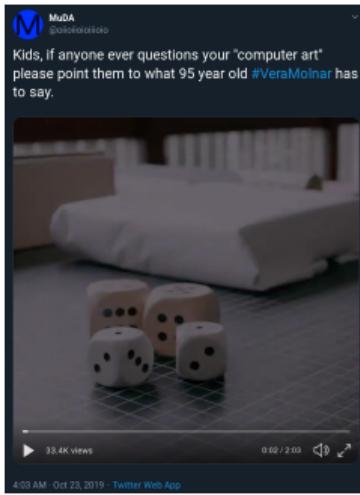
Hadley Wickham
@hadleywickham

The first five chapters of mastering-shiny.org are basically complete – this should be a great place to start if you've ever wanted to learn Shiny, and I'd love your feedback about what is confusing and what questions you still have `#rstats`

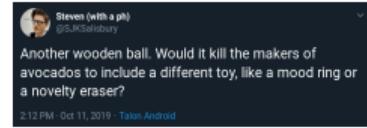
Mastering Shiny
The official guide to the Shiny web application framework for R.
mastering-shiny.org

9:32 AM - Oct 8, 2019 · Twitterbot for Mac

random stuff



more random stuff



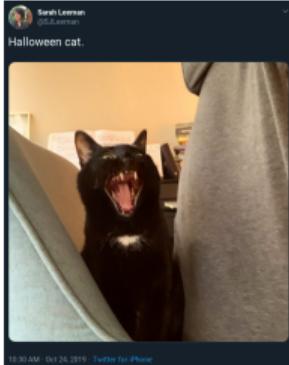
data vis

A screenshot of a tweet from Nadeem Bremer (@NadeemBremer). The tweet text is: "Learning #jupyter + #Python | Day 5 (after finally finishing to judge the IIB Awards): I made a tiny animation of my floating ball of gum from day 4! Somehow the colors became quite muted though. S And also still trying to find a good way to create a nice "background". Below the tweet is a thumbnail image showing a 3D rendering of a pink, translucent sphere with a blue play button icon in the center, set against a dark background.

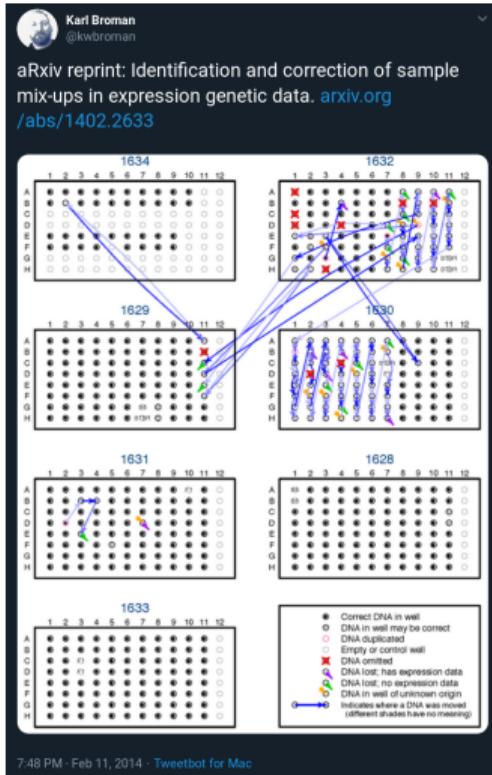
An infographic titled "Spell-O-Ly" featuring the subtitle "Hilarious Misspellings in Reddit Comments". It includes a list of misspellings such as "POSSIBLY", "CAPTIVELY", "GUNNAROLY", and "FAMOUSLY", each accompanied by a small cartoon character illustration. The infographic is attributed to "Hilary Bremner" and "infobeautyaward.com".

A screenshot of a Twitter post. The user @NadiehBremer has a profile picture of a woman with glasses and short hair. The tweet text is: "Really happy to have received this amazing looking 'Data Visualization Handbook' by @infomuotolu & @jhilden 😊". Below the tweet is a large image showing the front cover of the book 'Data visualization handbook' by 'INFO MUOTOLU & JHILDEN' and an open page displaying various data visualization examples like pie charts and bar graphs.

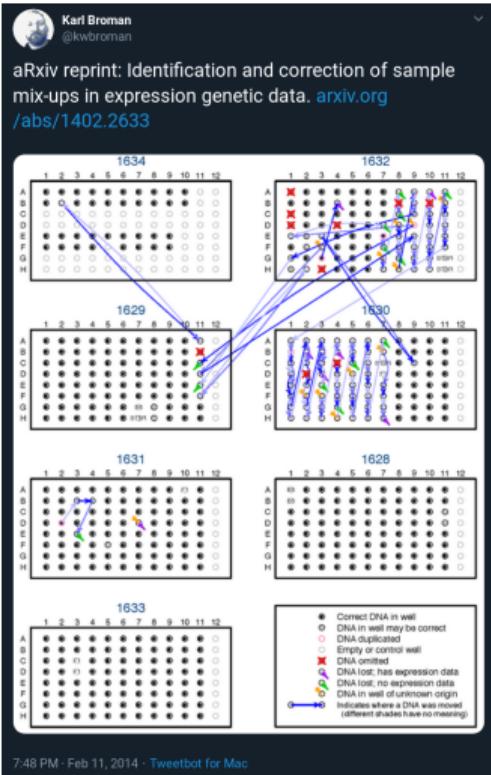
more cats



papers



papers

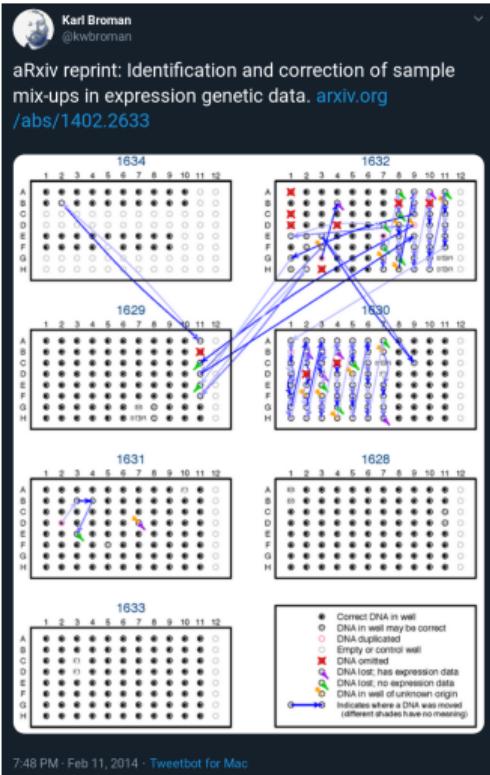


Replying to @yarbsalocin

@yarbsalocin @kwbroman Wasn't the exact same thing done 3 years ago by Westra et al.? bioinformatics.oxfordjournals.org/content/early/ ...

7:26 AM · Feb 12, 2014 · Twitter Web Client

papers



Tuuli Lappalainen
@tuuliel

Replying to @yarbsalocin

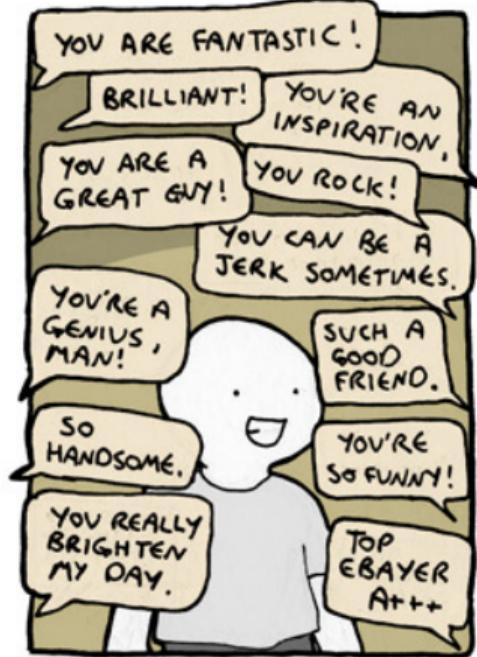
@yarbsalocin @kwbroman Wasn't the exact same thing done 3 years ago by Westra et al.? bioinformatics.oxfordjournals.org/content/early/ ...

7:26 AM · Feb 12, 2014 · Twitter Web Client

Karl Broman @kwbroman · Feb 12, 2014

Replying to @tuuliel

@tuuliel Yes; at least 4 other groups have done the same thing and are ref'd in my paper. My title includes "A case study." @yarbsalocin



WWW.DEATHBULGE.COM

Twitter challenges

- ▶ Controlling your time
- ▶ Keeping up
- ▶ People can be cruel or even evil, especially to women
- ▶ Feeling like a loser

Other sites

- ▶ GitHub
- ▶ Google scholar
- ▶ ORCID
- ▶ YouTube
- ▶ Facebook
- ▶ LinkedIn

Summary

- ▶ Create a personal website, preferably with a semi-permanent URL
- ▶ Consider writing an occasional blog post
- ▶ Consider participating on twitter
- ▶ List your website on the other sites; list the other sites on your website
- ▶ You don't need to include a picture, but consider at least making an “avatar”

Slides: kbroman.org/Talk_SocialMedia



kbroman.org

github.com/kbroman

@kwbroman