

rstudio::conf
from RStudio

Phrasing

communicating data science through gifs, tweets, & classic misdirection

Mara Averick
tidyverse developer advocate
@dataandme

<http://bit.ly/phrasing-ds>

PHRASING

communicating data science through gifs,
tweets, & classic misdirection

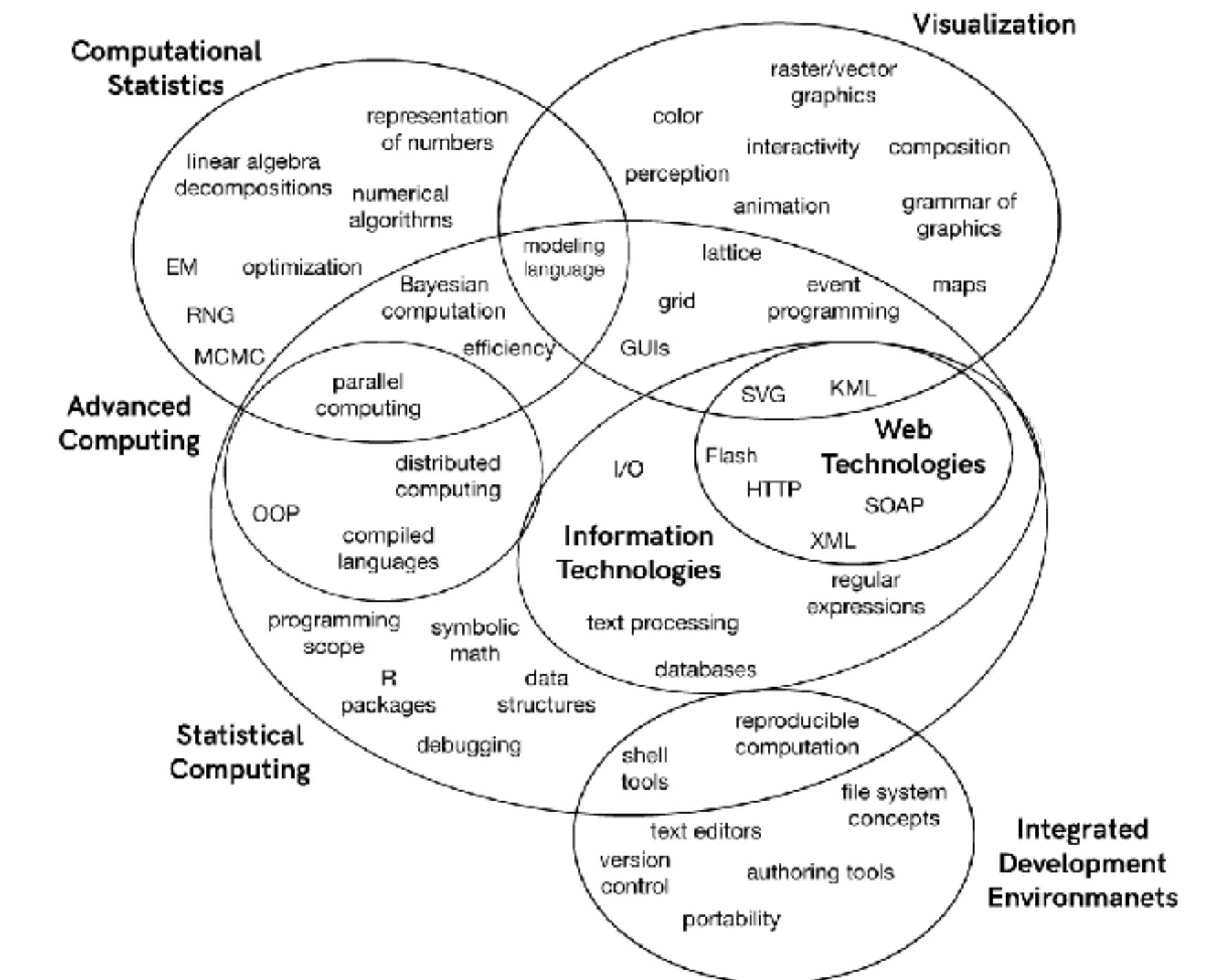
PHRASING





Mara Averick

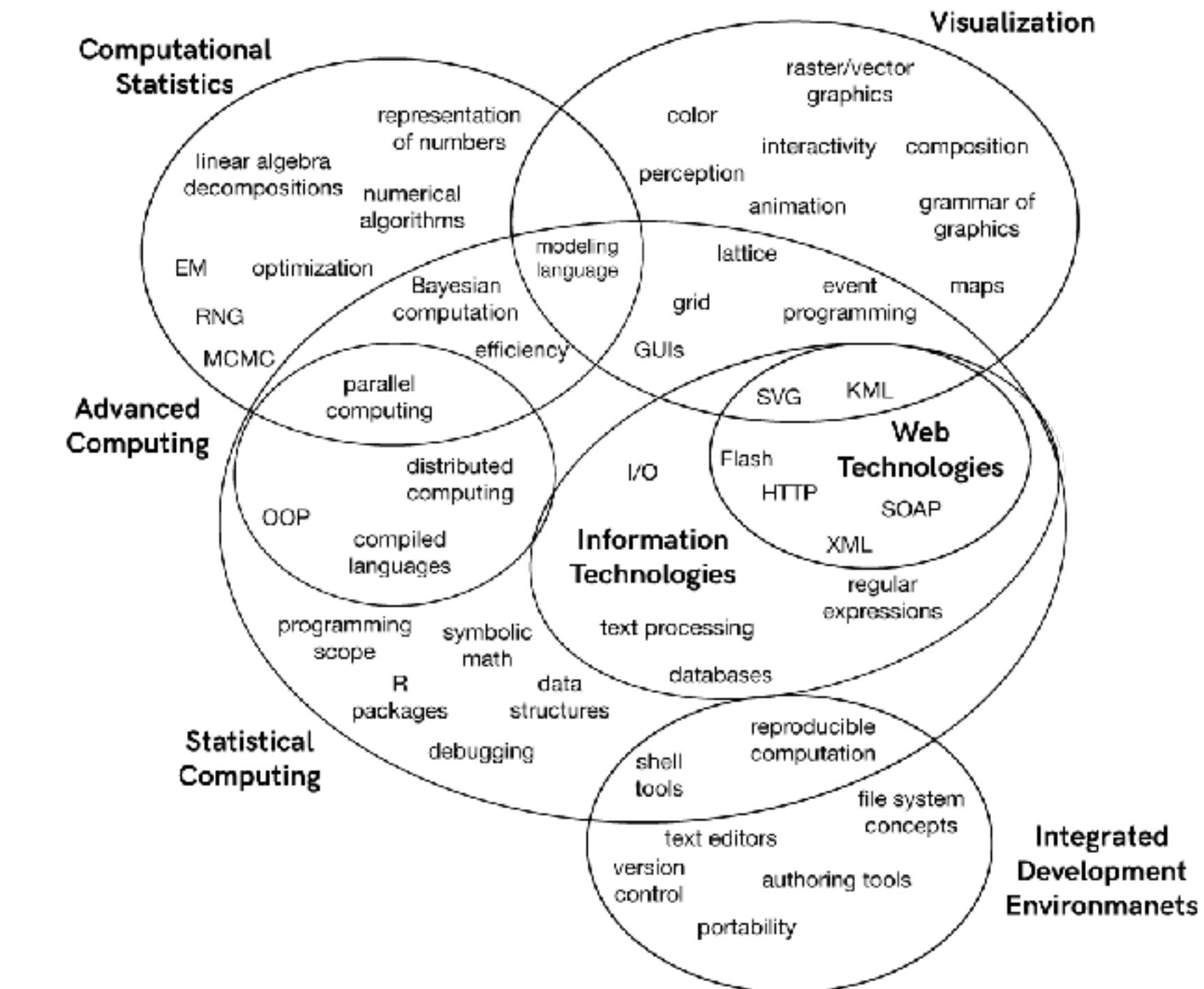
TIDYVERSE DEV ADVOCATE, RSTUDIO





Mara Averick

TIDYVERSE DEV ADVOCATE, RSTUDIO





Mara Averick

TIDYVERSE DEV ADVOCATE, RSTUDIO



alternative titles



alternative titles



sci•o•list

n • a superficial pretender to knowledge

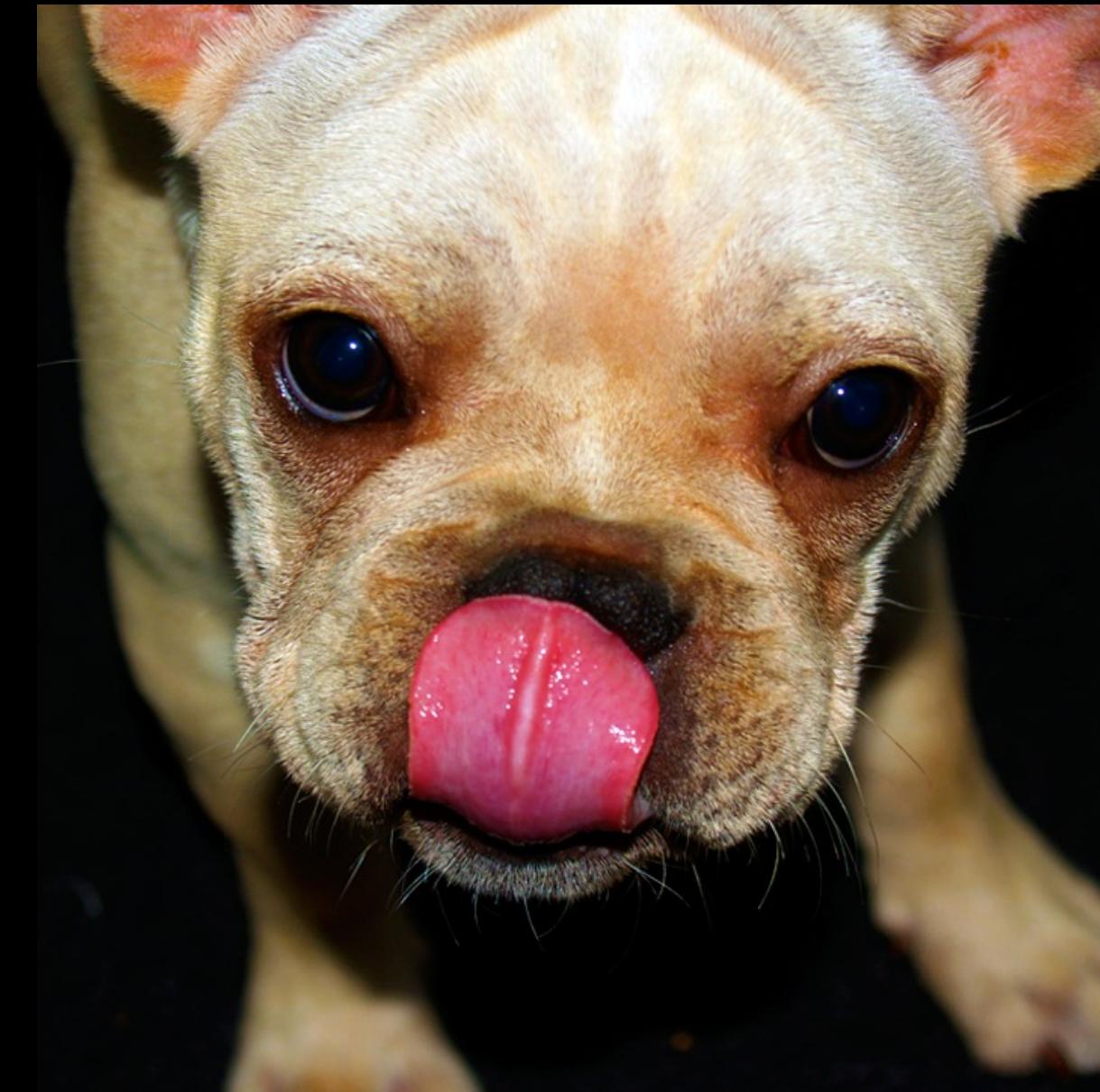
origins...

✓ Basketball



origins...

- ✓ Basketball
- ✓ The Batpig
- ✓ @buffer



leaRning out loud

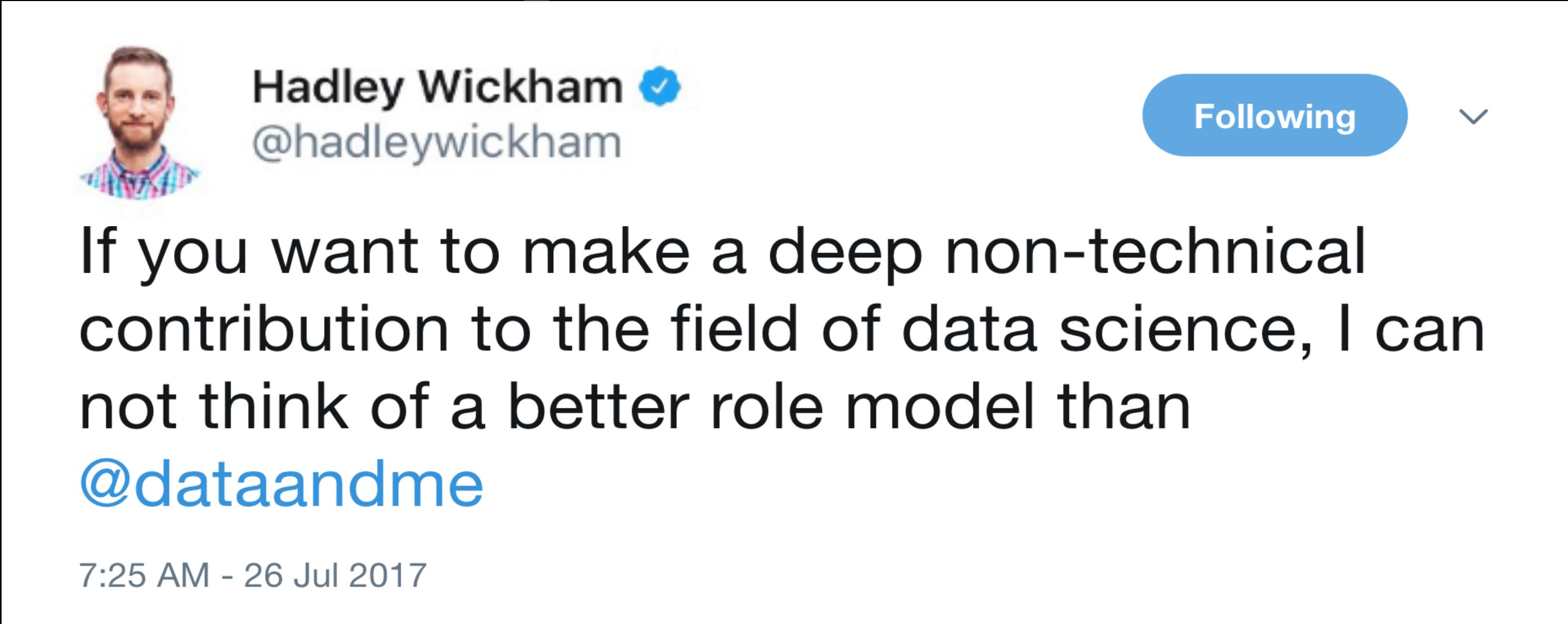
7



OMG I just learned a thing!! 🎉



but...



Hadley Wickham 
@hadleywickham

Following

If you want to make a deep non-technical contribution to the field of data science, I can not think of a better role model than
[@dataandme](#)

7:25 AM - 26 Jul 2017

but...

Edwin Thoen
@edwin_thoen

Following

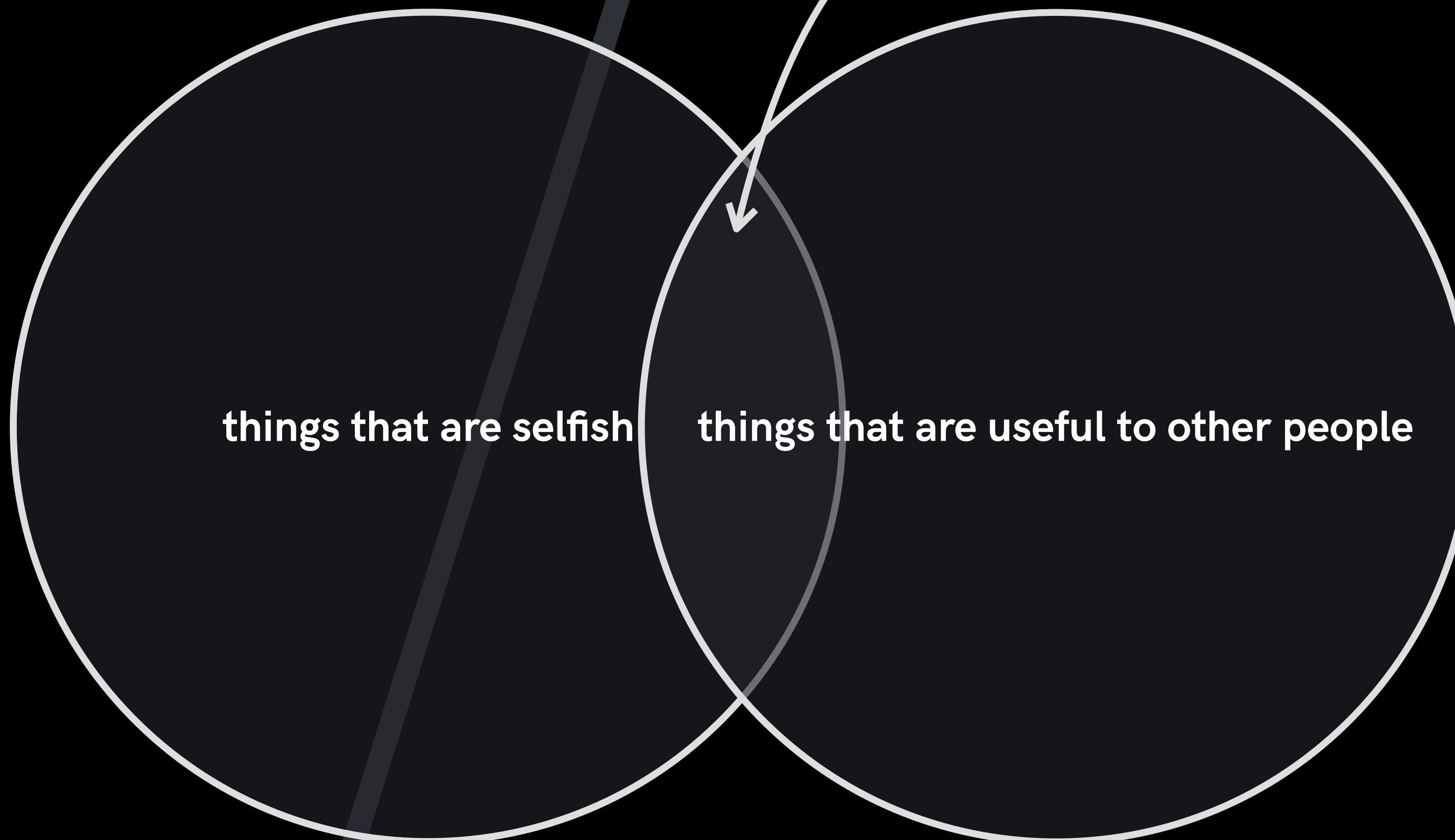
After 9 months as an R blogger I discovered the key to writing a succesfull blog: please [@dataandme!](#)

What determines a succesful blog post?

The figure is a histogram titled "What determines a successful blog post?". It compares two distributions on a horizontal axis labeled "Likes, retweets, pageviews, whatever" ranging from 0 to 500. A vertical line at approximately 350 is labeled "Tweeted about by Mara". The area to the left of this line is shaded red and represents the distribution for posts that were not tweeted about by Mara. The area to the right is shaded blue and represents the distribution for posts that were tweeted about by Mara. The blue distribution is shifted to the right, indicating higher engagement for posts that were tweeted about.

5:36 AM - 27 Aug 2017

but...





Hadley Wickham
@hadleywickham

Follow

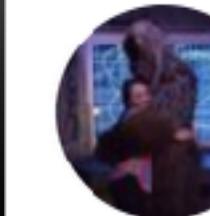


Replying to @juliasilge @arnicas @dataandme

I really didn't want to imply that Mara doesn't have awesome technical skills 😞 She does!

8:32 AM - 26 Jul 2017

8:35 AM - 26 Jul 2017



Lucy
@LucyStats

Follow



Replying to @hadleywickham @juliasilge and 2 others

I read as @dataandme has the ability to translate quite technical material to a non-tech audience, so if you're non-tech she's a 🔥 go to! 🎉🏆

8:37 AM - 26 Jul 2017

ex•o•ter•ic

adj. understandable by outsiders

or the general public



Lucy 
@LucyStats

Follow ▾

Replying to [@hadleywickham](#) [@juliasilge](#) and 2 others

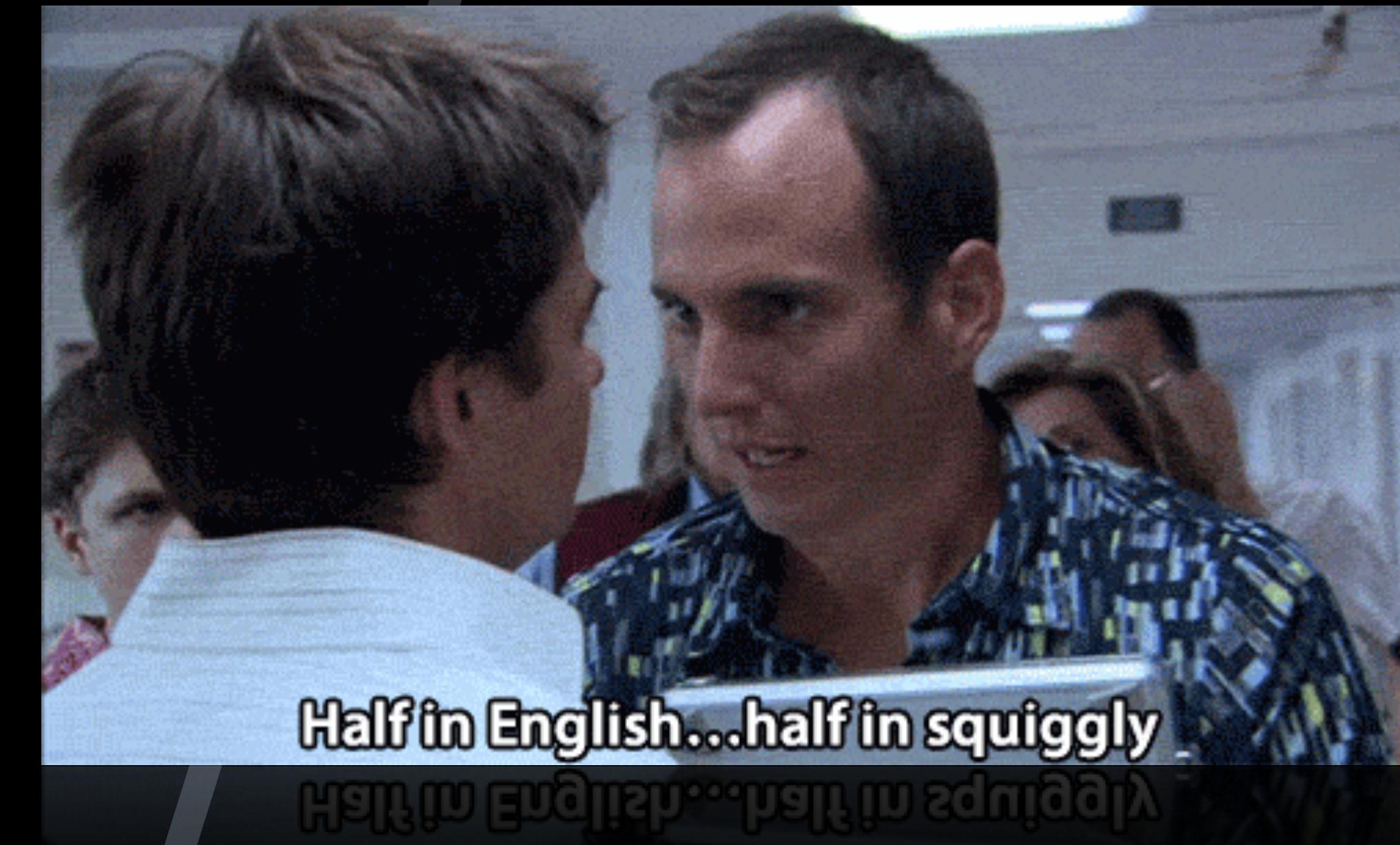
I read as [@dataandme](#) has the ability to translate quite technical material to a non-tech audience, so if you're non-tech she's a 🔥 go to! 🚀🏆

8:37 AM - 26 Jul 2017

ex•o•ter•ic

adj. *understandable by outsiders*

or the general public



a·po·ri·a

*n · patently insincere professings
(e.g. by a public speaker) of an
inability to know where to begin or
what to say*

hypotheses...

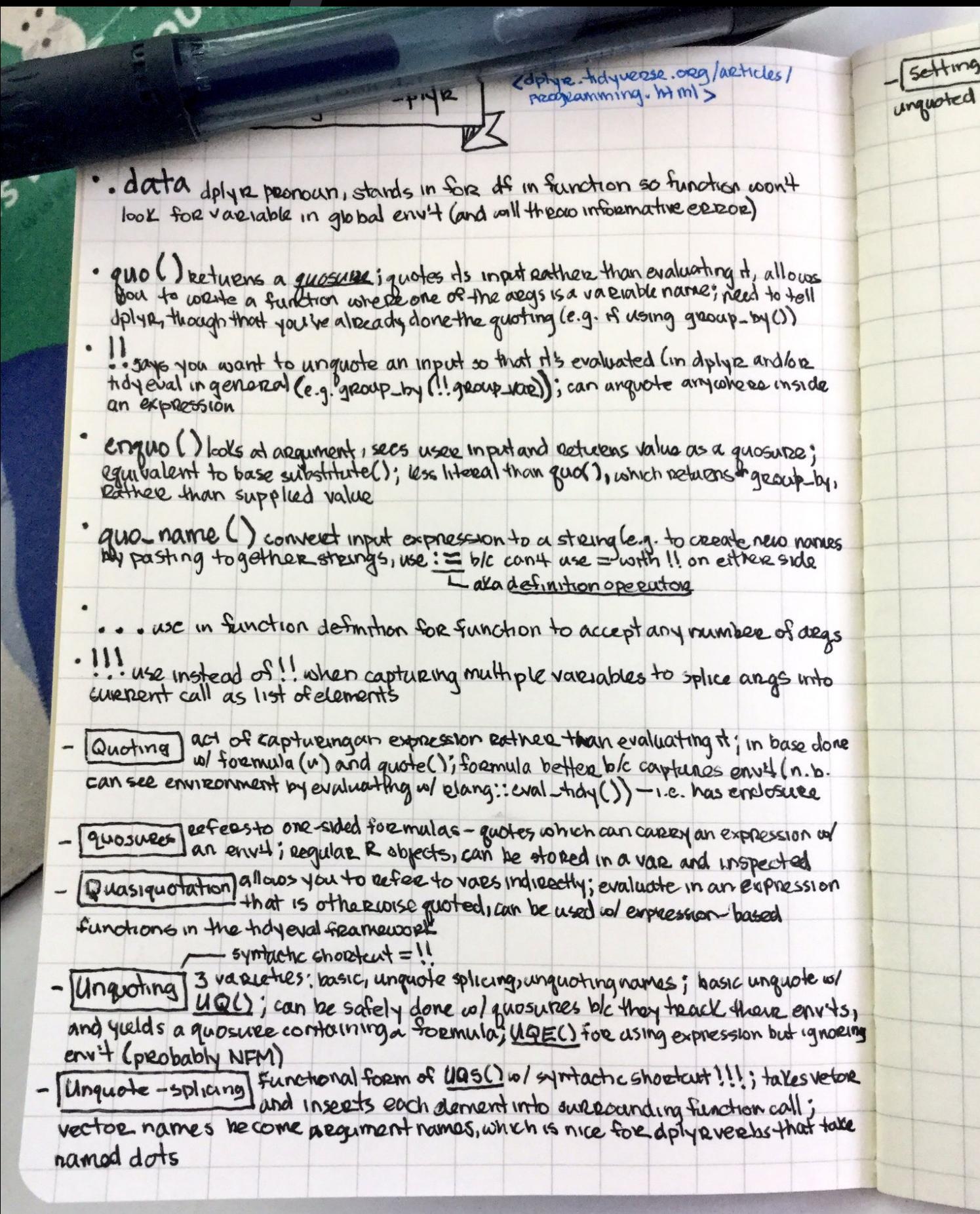


hypotheses...

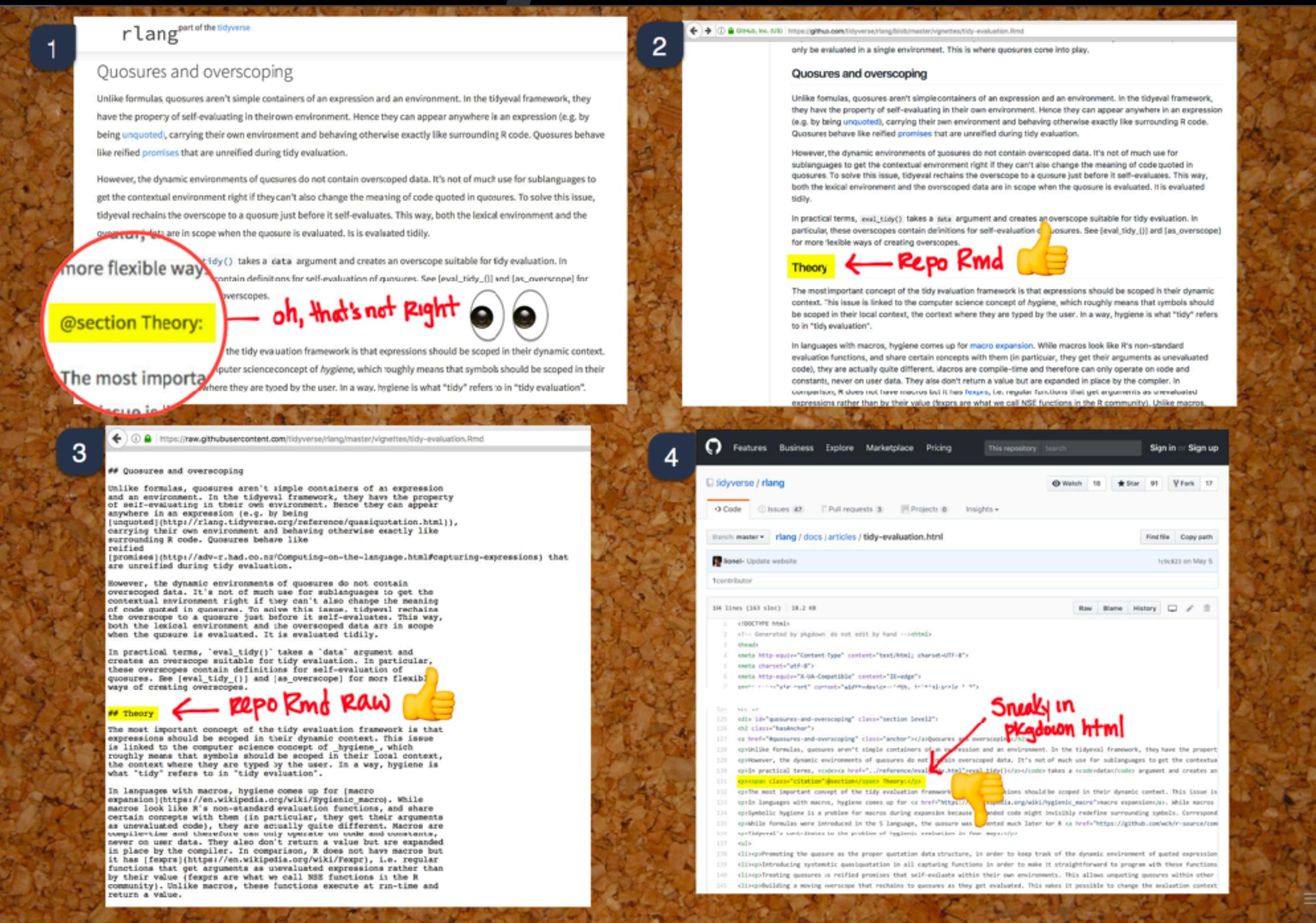


We gotta find a way to make this fit into the hole for this

untidy efforts...



untidy efforts...



other people's content...



other people's content...

↑
not necessarily new

Mara Averick
@dataandme

Follow

ICYMI, it's not breaking news, but @hspter's R-📦-✍️ing guide = 🔥: "Writing an R package from scratch" buff.ly/2wuSwg0 #rstats

Step 0: Packages you will need

The packages you will need to create a package are devtools and roxygen2. I am having you download the development version of the roxygen2 package.

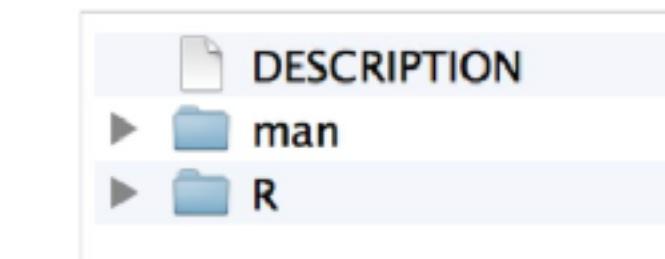
```
1 | install.packages("devtools")
2 | library("devtools")
3 | devtools::install_github("klutometis/roxygen")
4 | library(roxygen2)
```

Step 1: Create your package directory

You are going to create a directory with the bare minimum folders of R packages. I am going to make a cat-themed package as an illustration.

```
1 | setwd("parent_directory")
2 | create("cats")
```

you look in your parent directory, you will now have a folder called cats, and in it you will have two folders and one file called DESCRIPTION.



You should edit the DESCRIPTION file to include all of your contact information, etc.

Step 2: Add documentation

We didn't generate the most interesting docs to start. Let's make better – it's part of the package request that makes writing an R package an interesting challenge. That is that it needs to be well-documented. This is where the roxygen2 package comes in. It has a lot of great features for package documentation. The details can be found in the manual.

```
# A file 'DESCRIPTION'
# This function allows you to express your list of cats.
# You can use 'man' or 'R' to define the contents of these...
# Manually edit the 'DESCRIPTION' file to add the contents below and save it.
# Please add documentation
# If you need to add documentation from your annotations earlier, make sure to do
# "roxygen2" in your code, like it is in the code above.
# These tools are very useful.
# DESCRIPTION
# man
# R
```

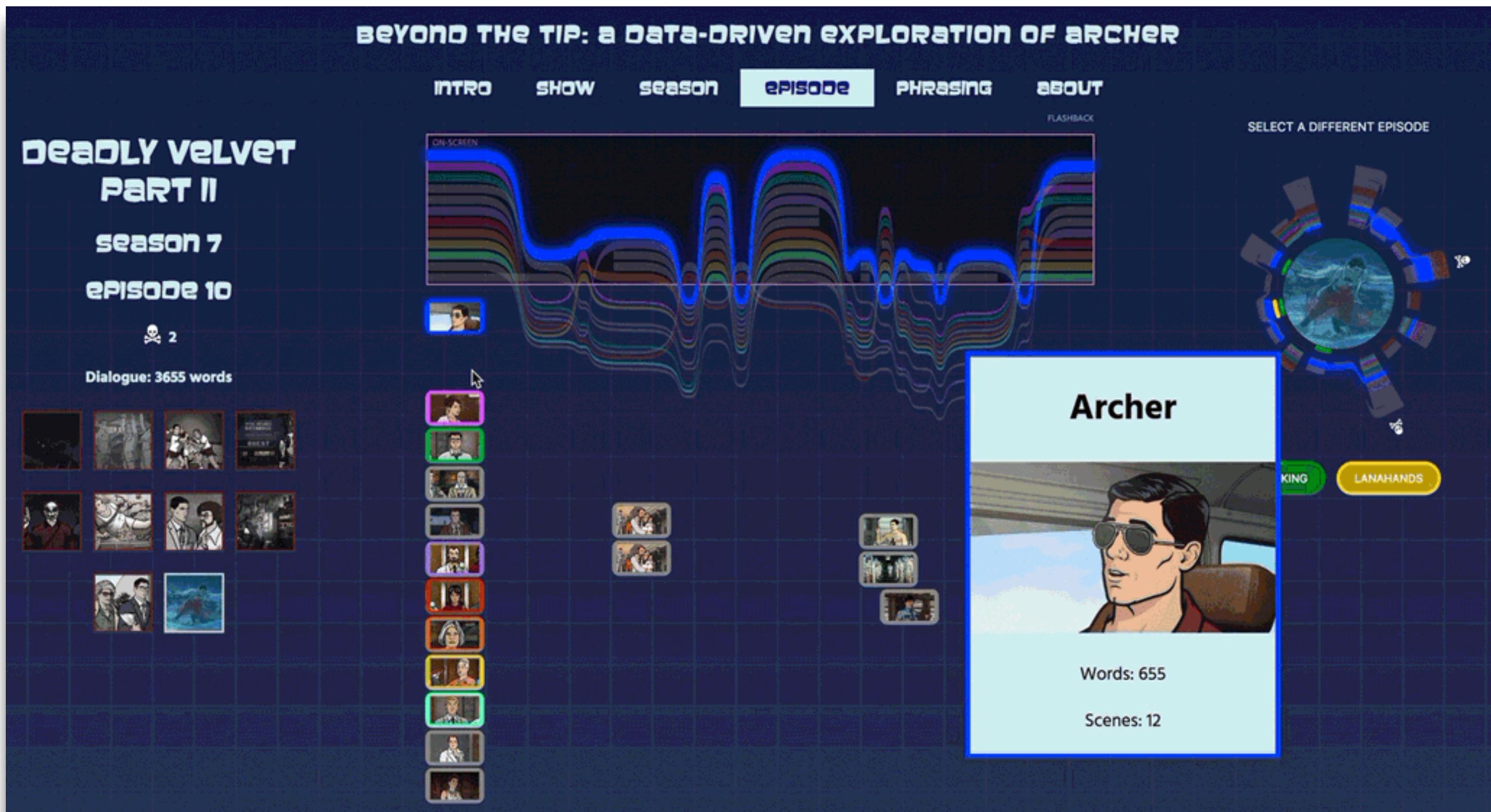
Automatically add to this .R file to the 'man' and 'NAMESPACE' files. It is very easy. You can read more about these, but in terms of what you need to take you to these tools anything goes.



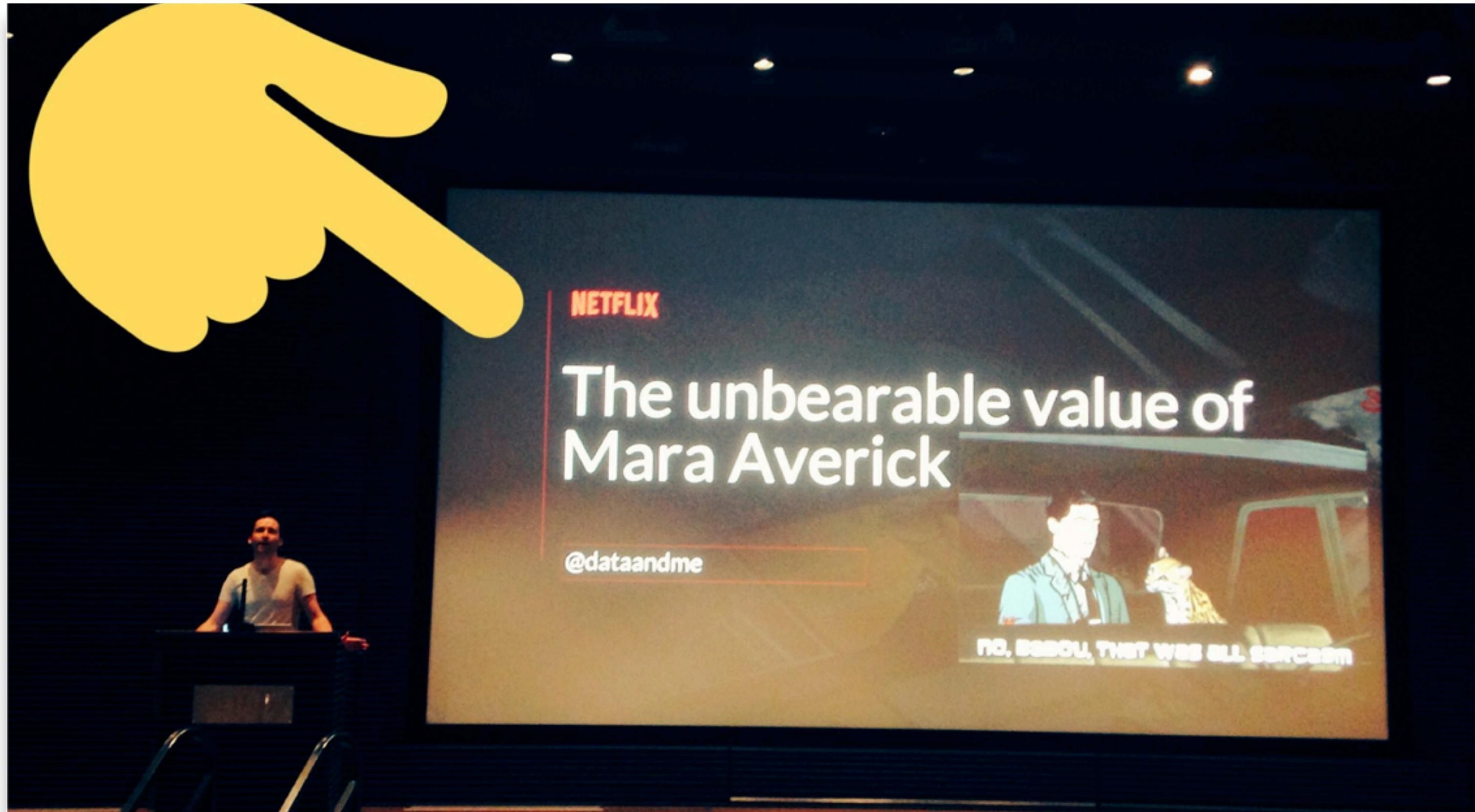
other people's content...



outcomes...

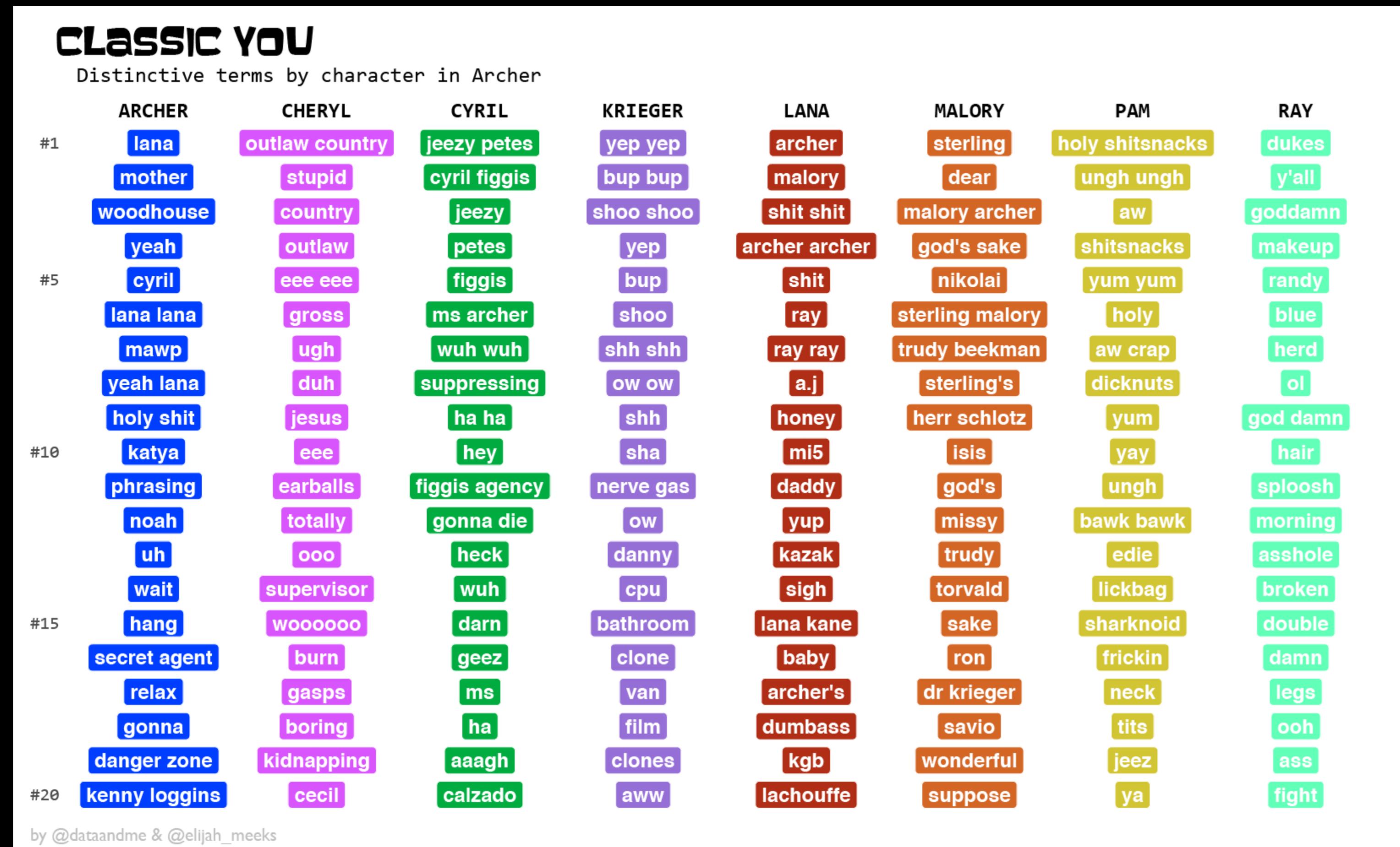


outcomes...

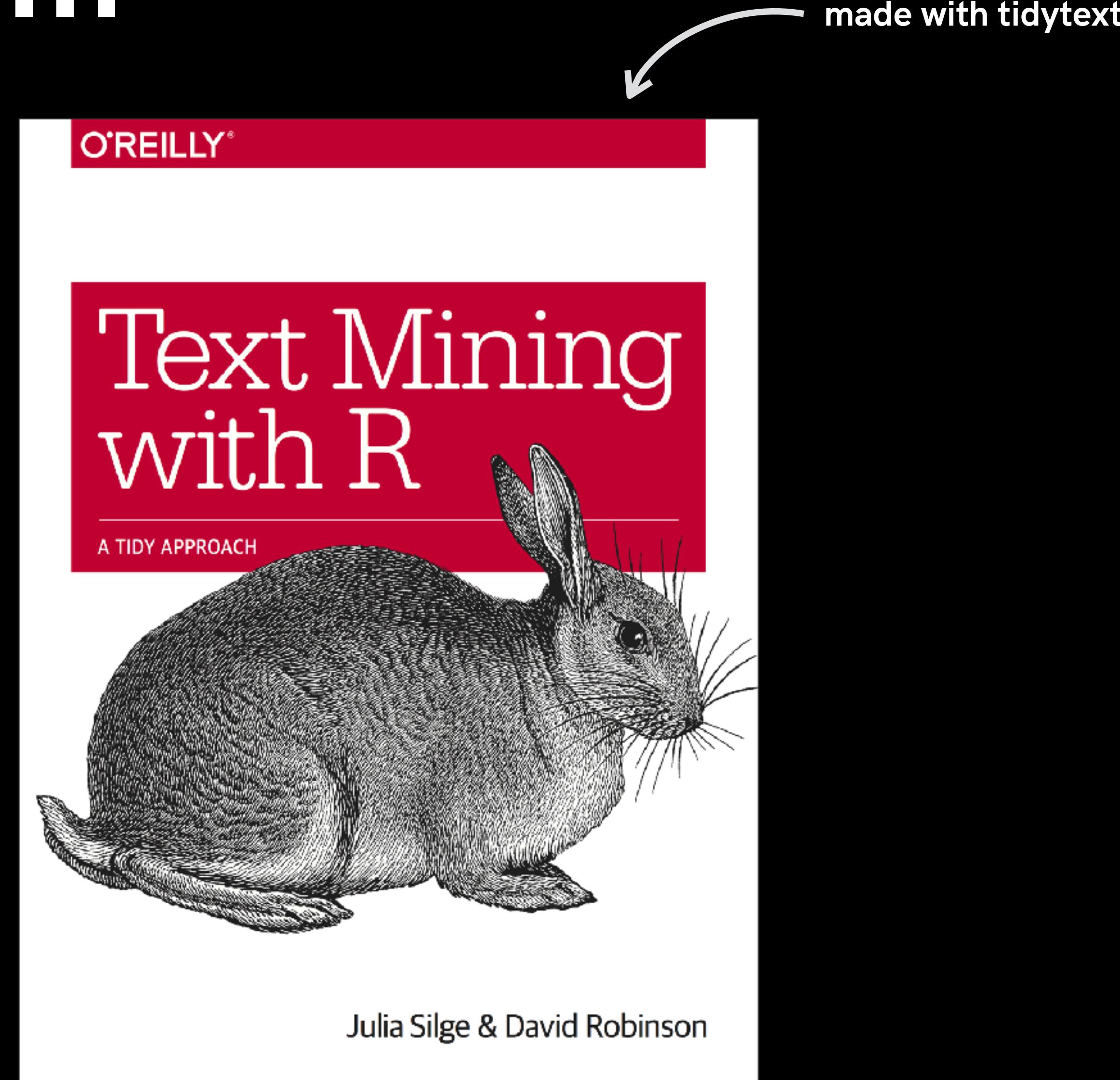


outcomes....

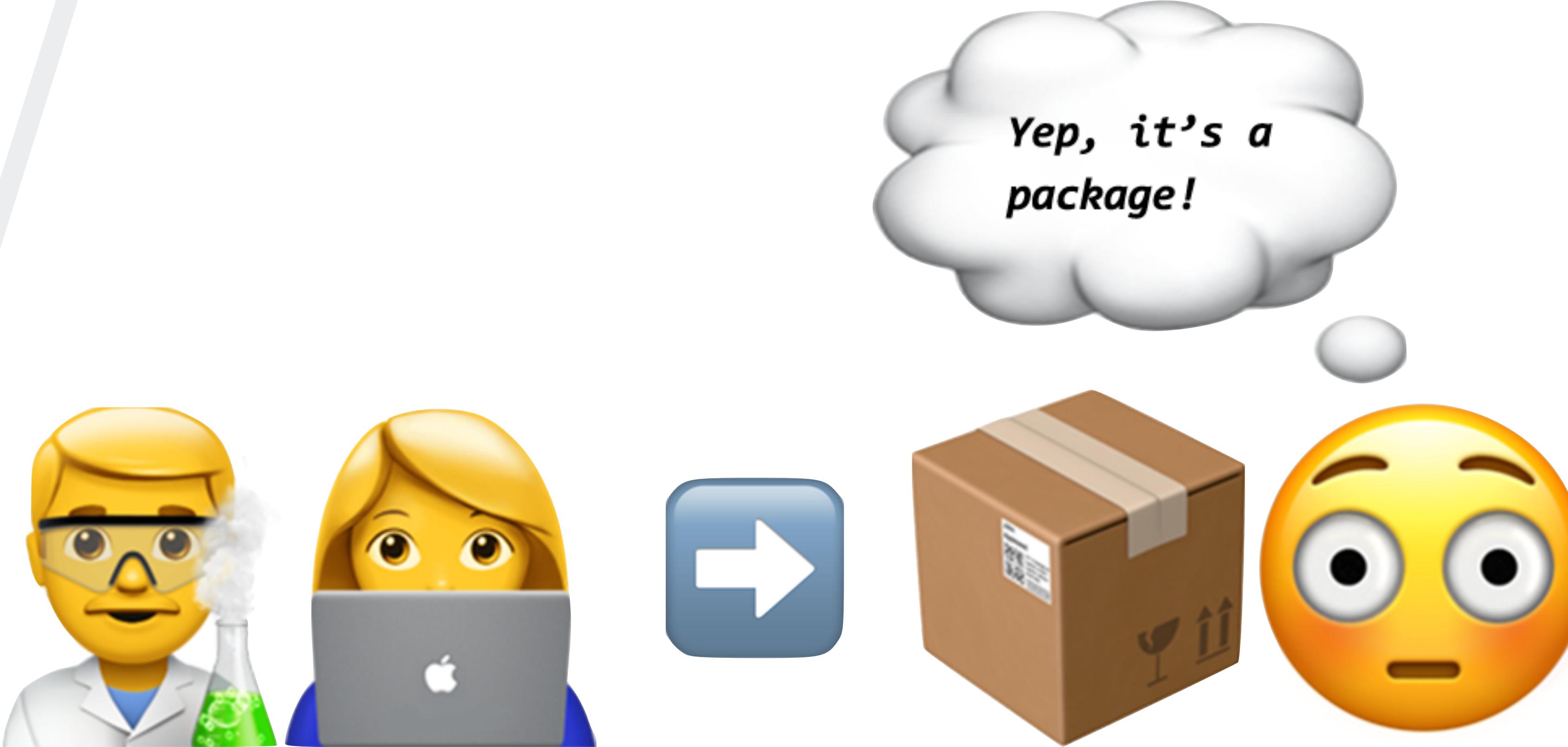
made with tidytext



outcomes...



outcomes...



outcomes...

When assigning reviewers to a submission, we aim to pair experienced reviewers with new ones, or reviewers with expertise on a package's programming methods with those experienced in its field of application.

experts need n00bs



How rOpenSci uses Code Review to Promote Reproducible Science

ropensci.org highlighted with Highly



tradecraft

tweet anatomy



Your Name

@yourhandle

Cool tutorial using **package-name**:
"Title of the tutorial" by **@author_handle**
www.thelinktoit.com via **@sitename**
#relevanthash tag **#maybeanother**

tweet anatomy

Mara Averick
@dataandme

Follow

Another ⭐ chart resource: "ggplot2 extensions" (w/ more than shown) maintained by @Emaasit buff.ly/2eJI4HA
#rstats #ggplot2 #dataviz

ggplot2 extensions - gallery

The screenshot shows a grid of 12 charts from the ggplot2 extensions gallery. Each chart is accompanied by a brief description and links to the extension's GitHub page and documentation.

- ggQC: Use ggQC to plot single, faceted and multi-layered quality control charts.
- ggedit: ggedit is aimed at providing missing functionality to ggplot2 through the extension system introduced with ggplot2 v2.0.0.
- ggforce: ggforce is aimed at providing missing functionality to ggplot2 through the extension system introduced with ggplot2 v2.0.0.
- ggalt: A complement of 'ggname', 'cosine' and 'wedge' for 'ggplot2'.
- ggiraph: A collection to make ggplot2 graphics interactive.
- ggmuller: ggmmuller plots for visualizing evolutionary dynamics.
- ggstance: ggstance implements horizontal versions of common ggplot2 geom_s.
- ggrepel: A geom for repelling text labels away from each other.
- ggraph: ggraph is tailored for plotting graph-like data structures (graphs, networks, trees, hierarchies.)
- ggpmisc: Miscellaneous Extensions to ggplot2.

Maintained by Dale Emasit (@Emaasit). Design inspired by @hrbrmst.

tweet anatomy

Mara Averick
@dataandme

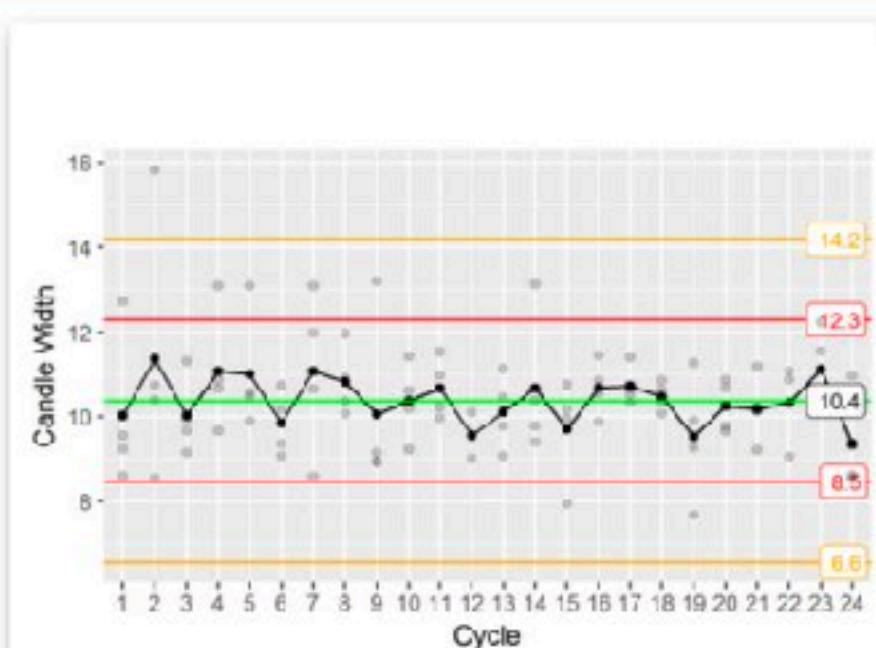
Follow

chart showing "ggplot2 extensions" for more than 1000 extensions by [@dataandme](#)

[ggplot2 extensions - gallery](#)

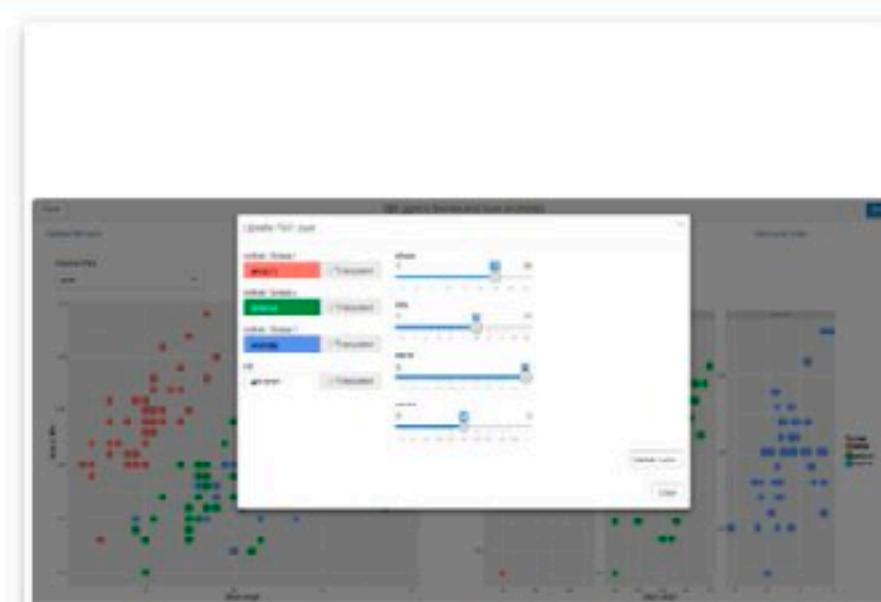
The screenshot shows a Twitter profile for Mara Averick (@dataandme). Below the profile picture and name, there is a blurred tweet. To the right of the tweet is a blue 'Follow' button and a dropdown arrow. Below the tweet is a link to the 'ggplot2 extensions - gallery'. The gallery page displays a grid of 12 ggplot2 extension packages, each with a thumbnail image, package name, star count, and a brief description. The packages are: ggQC, ggedit, ggforce, ggalt, ggiraph, ggmuller, ggstance, ggrepel, ggraph, and ggpmisc. Each package entry includes the author's name, tags, and js libraries.

Package	Author	Tags	js Libraries
ggQC	Kenith Grey	QC, XbarR, XbarR, SixSigma, Visualization	
ggedit	janicd	visualization, interactive, shiny, general	
ggforce	thomasp85	visualization, interactive, shiny, general	
ggalt	hrbrmstr	visualization, general	
ggiraph	davidoghieh	visualization, general	
ggmuller	robjphillip	visualization, evolution, dynamics	
ggstance	loneli	visualization, general	
ggrepel	slokov	visualization, general	
ggraph	thomasp85	visualization, general	
ggpmisc		miscellaneous Extensions to 'ggplot2'	

**ggQC**

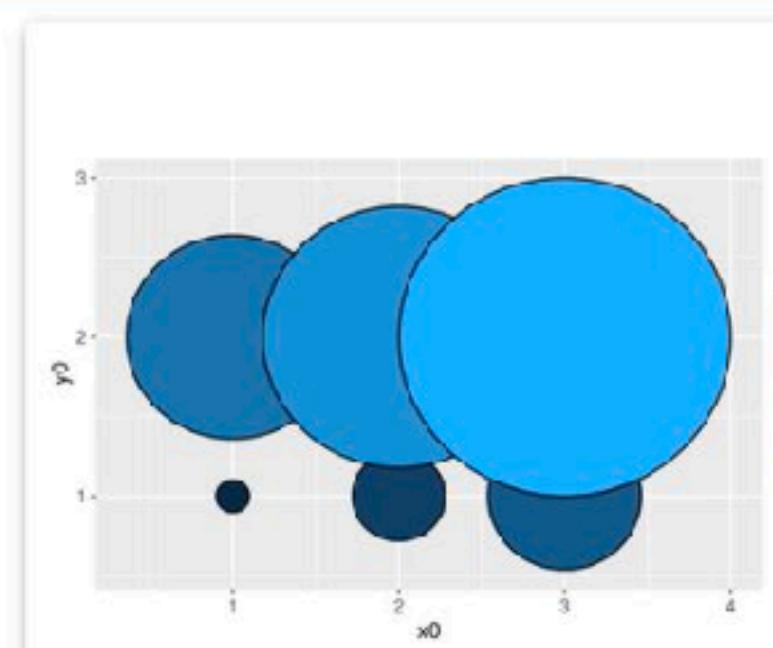
Use ggQC to plot single, faceted and multi-layered quality control charts .

- **author:** Kenneth Grey
- **tags:** QC, XmR, XbarR, SixSigma, Visualization
- **js libraries:**

**ggedit**

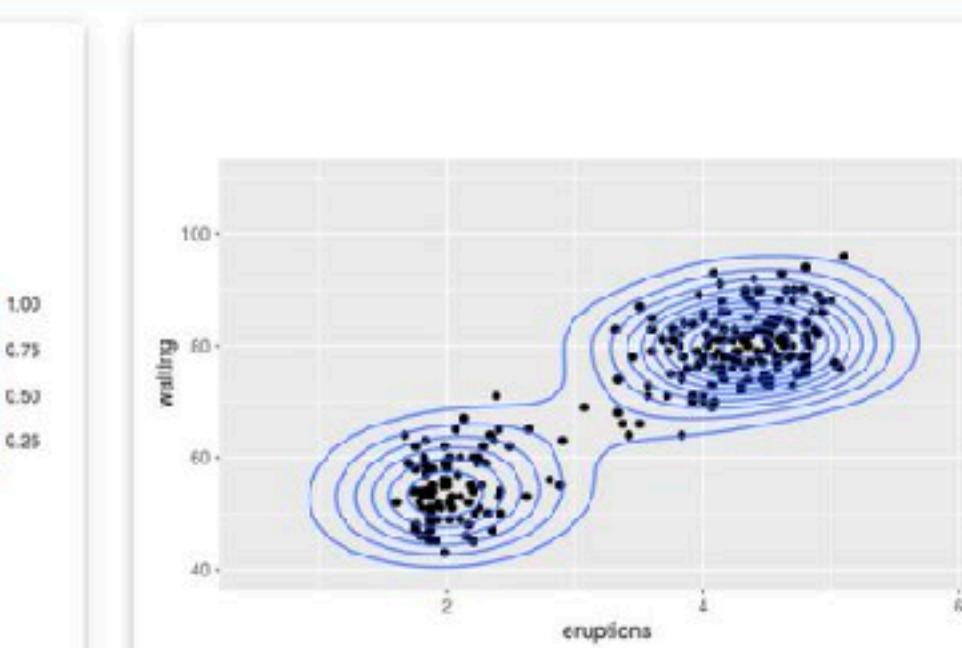
ggedit is aimed to interactively edit ggplot layers, scales and themes aesthetics

- **author:** yonied
- **tags:** visualization, interactive, shiny, general
- **js libraries:**

**ggforce**

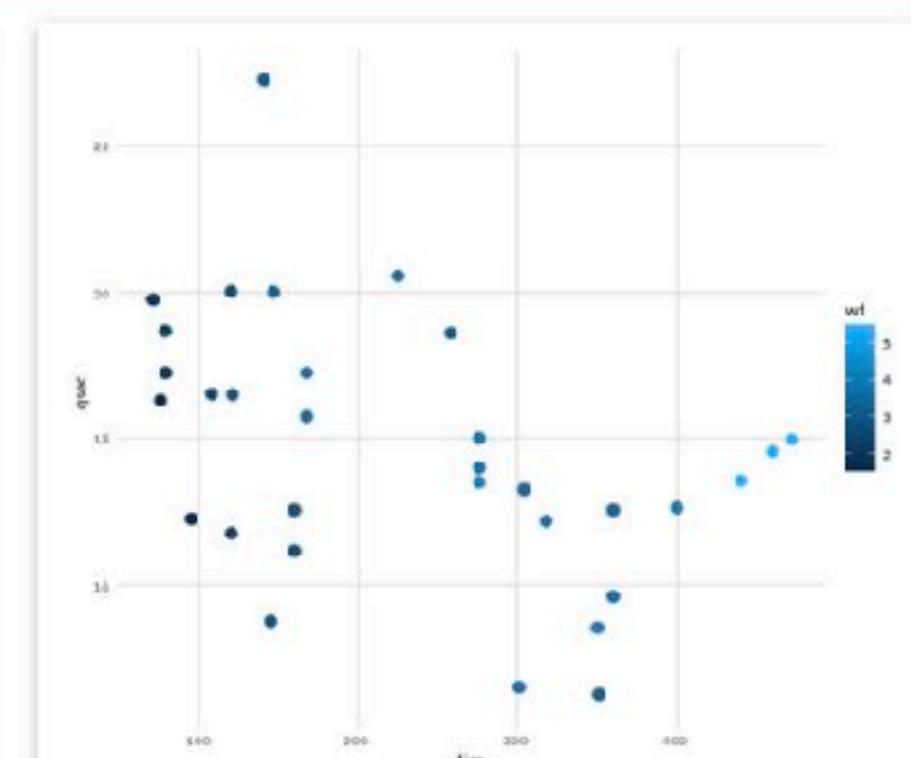
ggforce is aimed at providing missing functionality to ggplot2 through the extension system introduced with ggplot2 v2.0.0.

- **author:** thomasp85
- **tags:** visualization, general
- **js libraries:**

**ggalt**

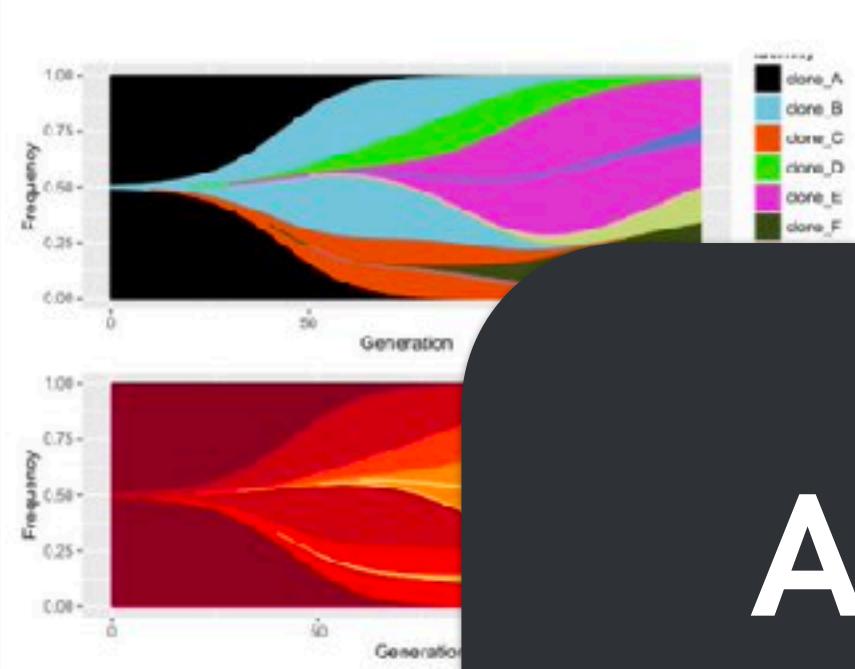
A compendium of 'geoms', 'coords' and 'stats' for 'ggplot2'.

- **author:** hrbrmstr
- **tags:** visualization, general
- **js libraries:**

**ggiraph**

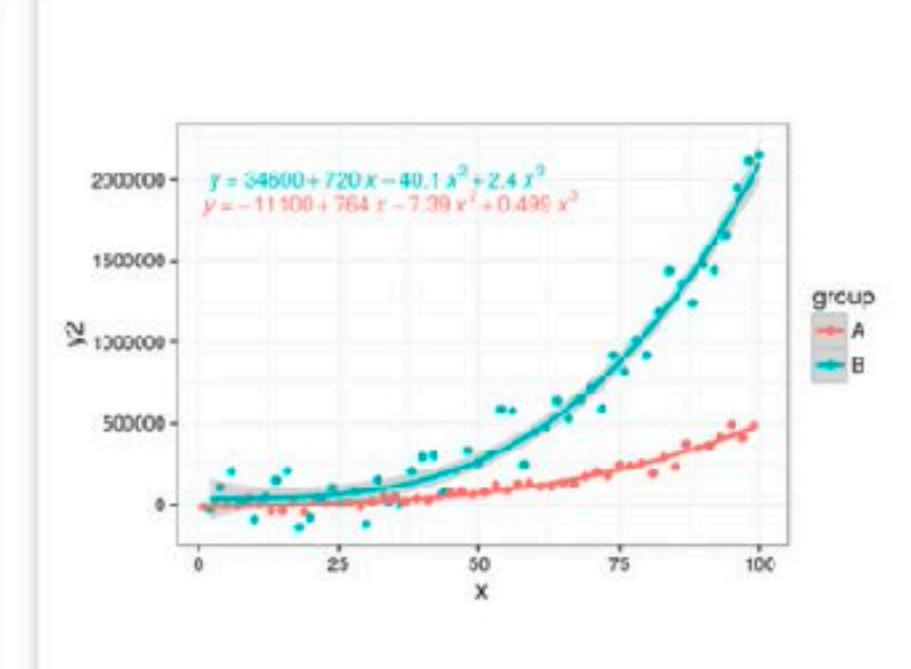
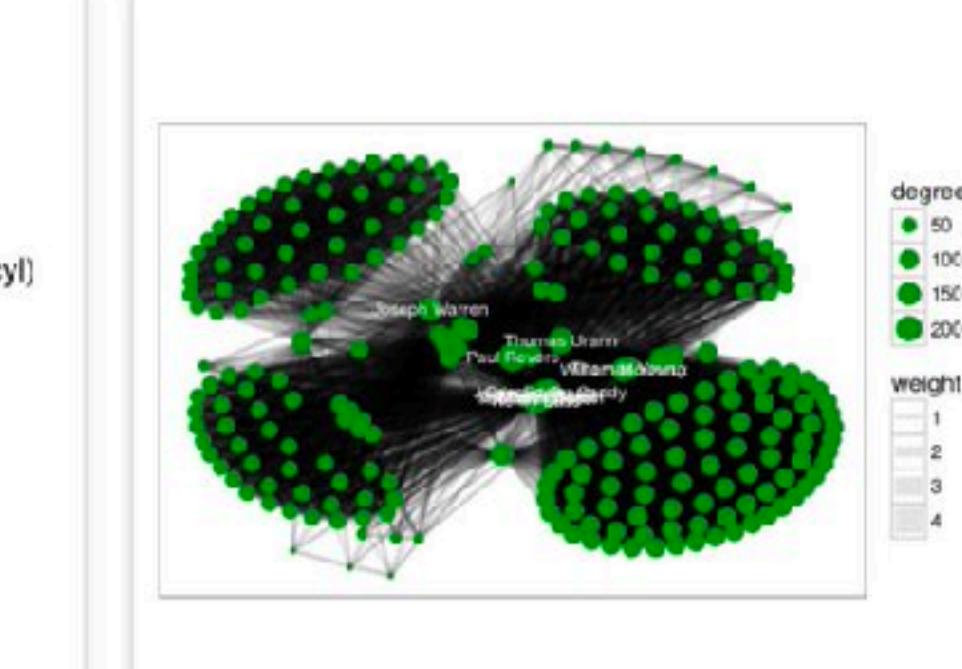
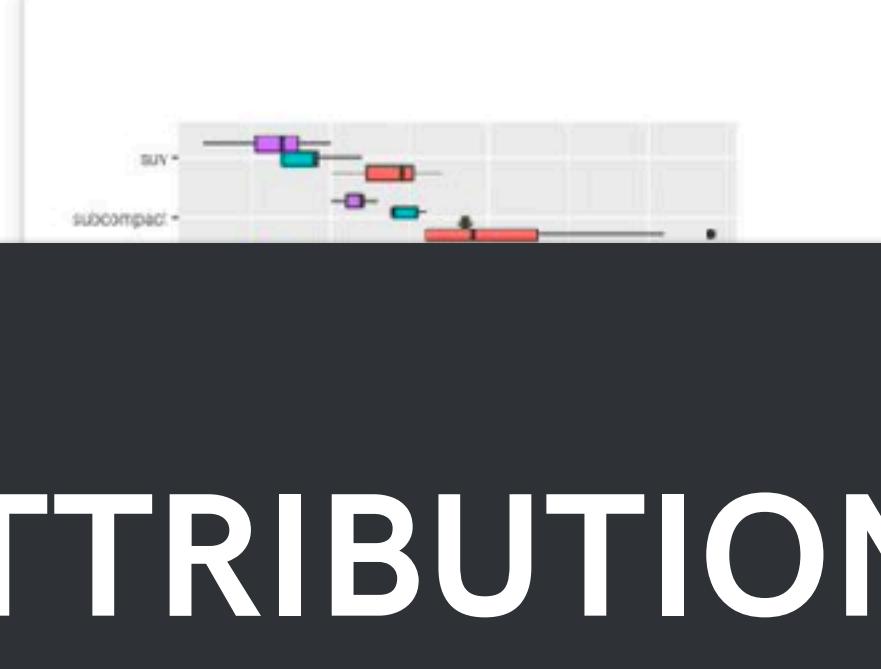
htmlwidget to make 'ggplot' graphics interactive.

- **author:** davidgohel
- **tags:** visualization, general
- **js libraries:**



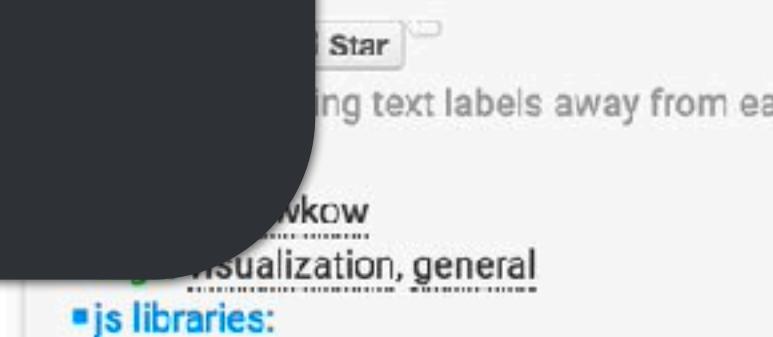
ATTRIBUTION

(e.g. @emaasit)

**ggmuller**

Creates Muller plots for visualizing dynamics.

- **author:** robjohnnoble
- **tags:** visualization, evolution, dynamics
- **js libraries:**

**ggraph**

ggraph is tailored at plotting graph-like data structures (graphs, networks, trees, hierarchies...).

- **author:** thomasp85
- **tags:** visualization, general
- **js libraries:**

ggpmisc

Miscellaneous Extensions to 'ggplot2'.

- **author:**
- **tags:** visualization, general
- **js libraries:**

the tweet != the content

Mara Averick (@dataandme) posted:

This should be mandatory reading before group projs ❤️: "Naming Things" by @JennyBryan buff.ly/2wa5QXz #rstats #SoDS17 #commonDecency

Jenny Bryan (@JennyBryan) responded:

Naming things
are by Jenny Bryan for
Introducible Science Workshop

emoji Mara's fault, though

↑
the real reason
we can't have
nice things

principles for (file) na
achine readable 🤖📖
man readable 🧑📖
ays well with default or
1 2 3
achine readable 🤖📖
y to search for files later
y to narrow file lists based on names
y to extract info from file names, e.g. by split
to regular expressions and globbing? be kind
self and avoid
spaces in file names
punctuation
accented characters
different files named "loc" and "Loc"

"human readable" 🧑📖
1_nothingworks.r
2_gro-deal-filberting.r
3_dos-with-dos-and-r
4_explore-dna-sequences.r
5_lets-make-a-new-name-for-our-dna.r
6_perfect_load-source.r
7_perfect2_load-nep-data.r
8_perfect_load-cookies-tables.r
9_perfect4_reformat-and-tidy.r

the tweet != the content

naming things

prepared by Jenny Bryan for

Reproducible Science Workshop

emoji Mars's fault, though

↑
the real reason
we can't have
nice things



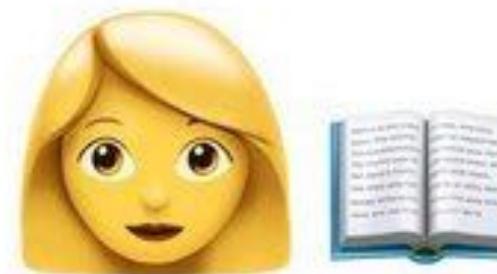
the tweet != the content

three principles for (file) names

machine readable



human readable



plays well with default ordering



tweet Rs

- Read (or run)
- Relatable
- Retrievable
- Relevant
- Real ←

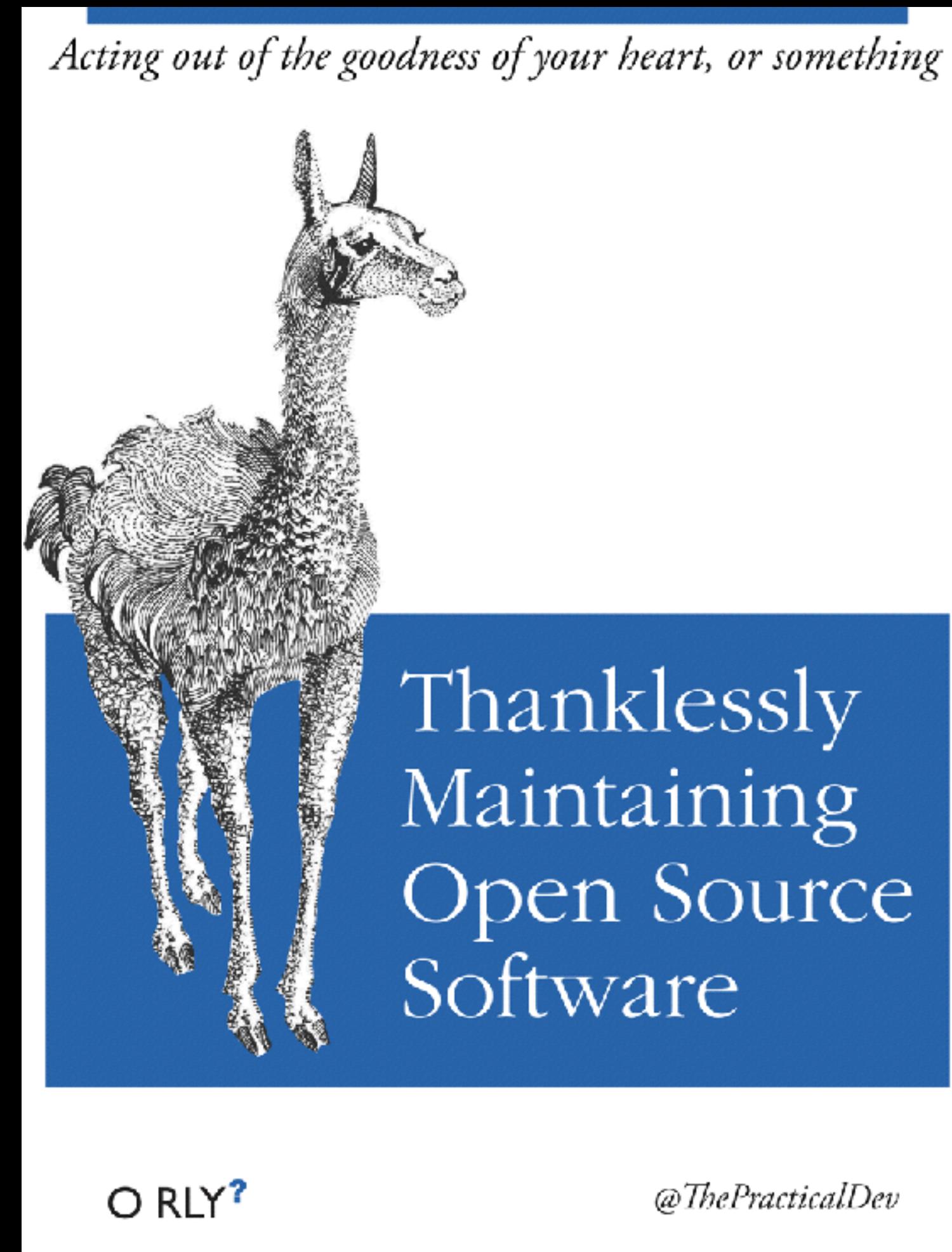
you do you 

* completely arbitrary rules I made up

classic misdirection?



classic misdirection?



Contributing to FOSS

WHAT HOLDS PEOPLE BACK?

- “*I can't write code.*”



Contributing to FOSS

WHAT HOLDS PEOPLE BACK?

- “*I can't write code.*”
- “*I'm not really good at this.*”



Contributing to FOSS

WHAT HOLDS PEOPLE BACK?

- “*I can't write code.*”
- “*I'm not really good at this.*”
- “*I'd just be a burden.*”



Contributing to FOSS

WHAT HOLDS PEOPLE BACK?

- “*I can't write code.*”
- “*I'm not really good at this.*”
- “*I'd just be a burden.*”
- “*They already have enough people smarter than me.*”



you never know...

 **Hadley Wickham**  @hadleywickham · 18 Nov 2017
If you're using databases with shiny, make sure to check out the pool 🎯:
blog.rstudio.com/2017/11/17/poo... Will make your life much easier! #rstats

7 120 395

 **Mara Averick** @dataandme · 18 Nov 2017
Never forget to check the pool...



2 19

you never know...



Matthew Dickinson

@midickinson

Follow



Replies to [@dataandme](#) [@hadleywickham](#)

If it weren't for the gif I would have totally just skipped over knowing about this package that looks to be a huge benefit to me.
Thanks!

6:13 PM - 18 Nov 2017

Thank You

<http://bit.ly/phrasing-ds>



Thank You

<http://bit.ly/phrasing-ds>

