useR!2017

Warwick R User Group

Ella Kaye & David Selby

20th July 2017

PLACE DE
BELGIQUE
BELGIË
PLEIN
EXPO
DINO WORLD
BRUSSELS
useR
2017

Selecta Cars De Smet

Alice Data
@alice_data

Follow

There is NO tea to be found at
@useR_Brussels #sadbrit

3:14 PM - 4 Jul 2017

4 Likes

Welcoming and inclusive

- newbies session
- 25 diversity scholarships
- transport for people with impairments
- breast-feeding area
- quiet zone
- reasonable fees for students
- childcare
- Heather Turner / Forwards / R Foundation

Welcoming and inclusive

- newbies session
- 25 diversity scholarships
- transport for people with impairments
- breast-feeding area
- quiet zone
- reasonable fees for students
- childcare
- Heather Turner / Forwards / R Foundation

People come

- 1161 attendees
- 54 countries
- 941 litres of beer

The R Ecosystem

R Foundation

R Core Group

#rstats

CRAN

R-bloggers

Bioconductor

R User Groups

GitHub

R-Ladies

stackoverflow

RDocumentation

consortium

Microsoft

RStudio

ROpenSci

software carpentry

rOpenGov

RLadies.org | @RLadiesGlobal | @RLadiesLondon

R-Ladies <-

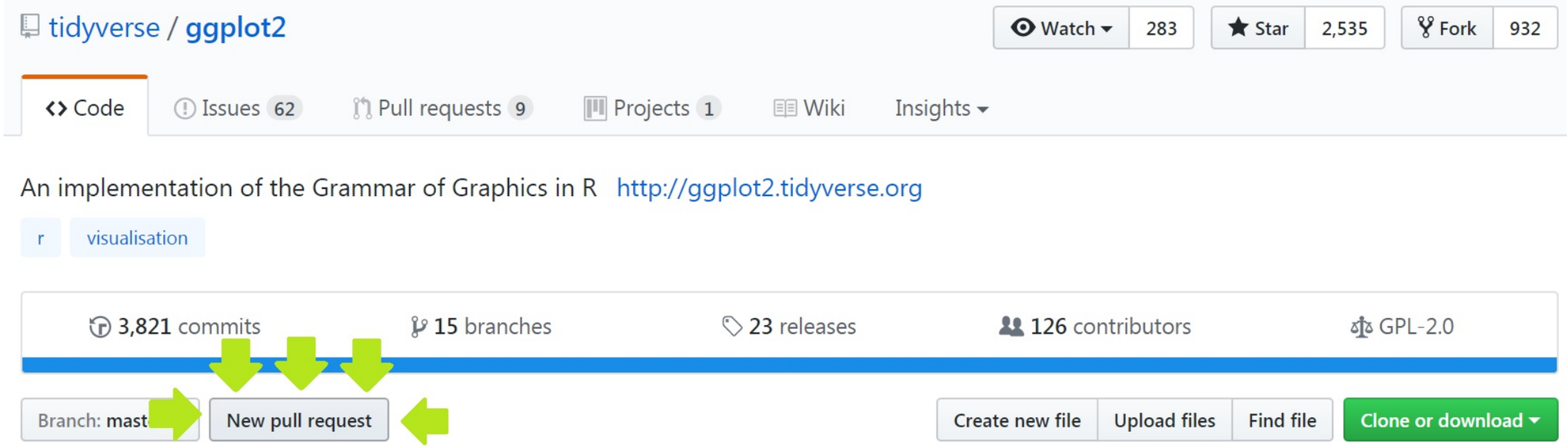
Worldwide organization that promotes gender diversity in the R community via meetups and mentorship in a friendly and safe environment

20+ countries | 45+ cities | 6000+ members

Collaborative coding*

> Step-by-step on contributing to package development, especially on GitHub

> If you're here, you're ready to contribute to R!



* *Charlotte Wickham, Oregon State University & DataCamp* (slides)

20 years of CRAN*

1976 <-

“At the beginning was S”

1992 <-

R development begins

1997 <-

R Core team and CRAN founded

2000 (February 29) <-

R-1.0.0 released (compatible with S3)

2013 <-

R-3.0.0 released

* *Uwe Ligges, R Core* (video)

20 years of CRAN*

11,100 packages depend on other packages that depend on other packages that depend on other packages...

Packages with >1000 reverse dependencies:

```
> knitr
> testthat
> Rcpp
> MASS   #  >75% of all packages
> Matrix   #  >75% of all packages
> survival   #  >75% of all packages
```

* *Uwe Ligges, R Core* (video)

Dirk Eddelbuettel
@eddelbuettel

Follow

Prof Ripley, r-devel:

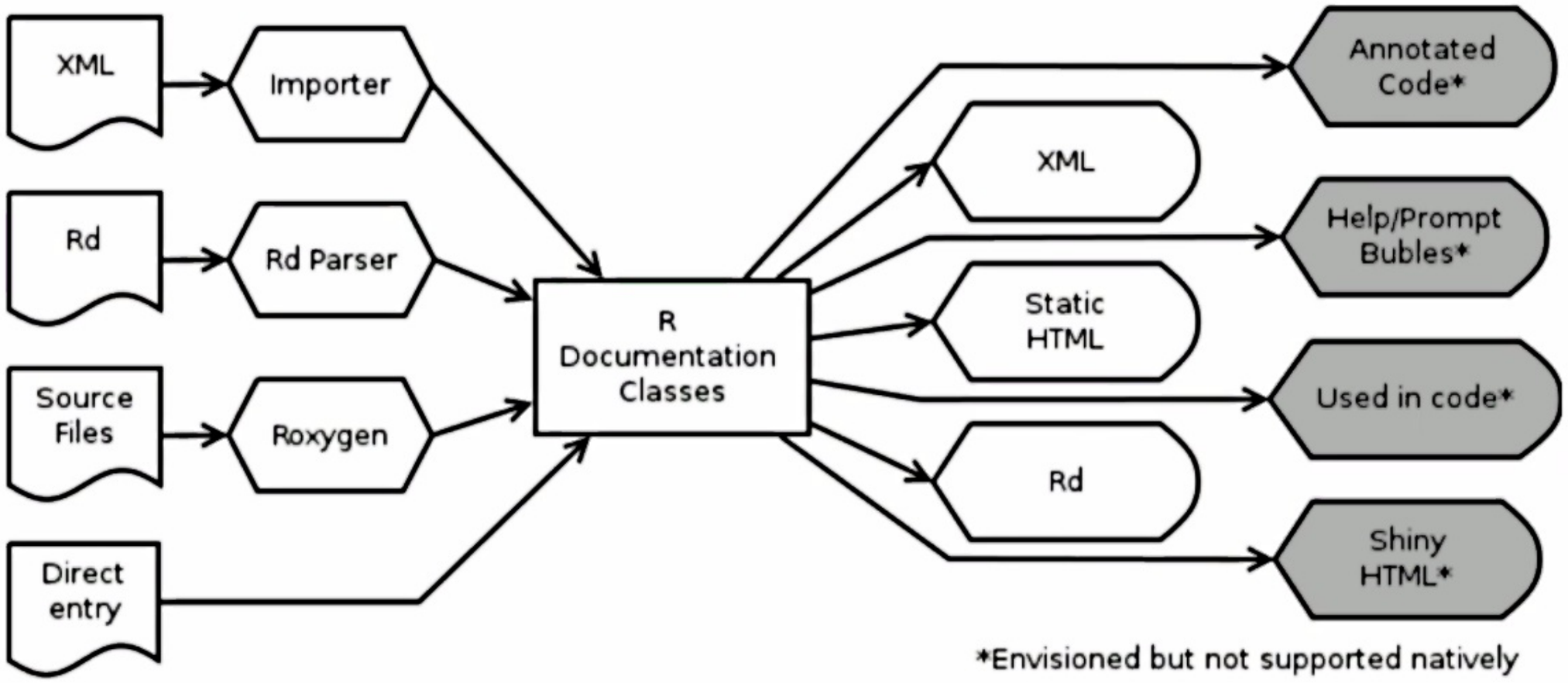
The Sparc hardware used for CRAN died during an unplanned power outage, and it seems unlikely it can be resurrected.

6:19 PM - 11 Jul 2017

19 Retweets **33** Likes

New R documentation*

Documentation stored/manipulated as S4 objects



* *Andrew Redd, University of Utah* (video)

New R documentation*

```
hw <-
function( greet = "Hello" #< A greeting
        , who   = "World" #< who to greet.
        ){
    #! A more complicated hello world
    print(paste(greet, who))
    #< Called for the side effect of printing,
    #^ but returns the pasted arguments invisibly.
}
```

Implicit unit testing

```
if (FALSE) {#@testing
  ... }
```

* *Andrew Redd, University of Utah* (video)

Teaching data science to new useRs*

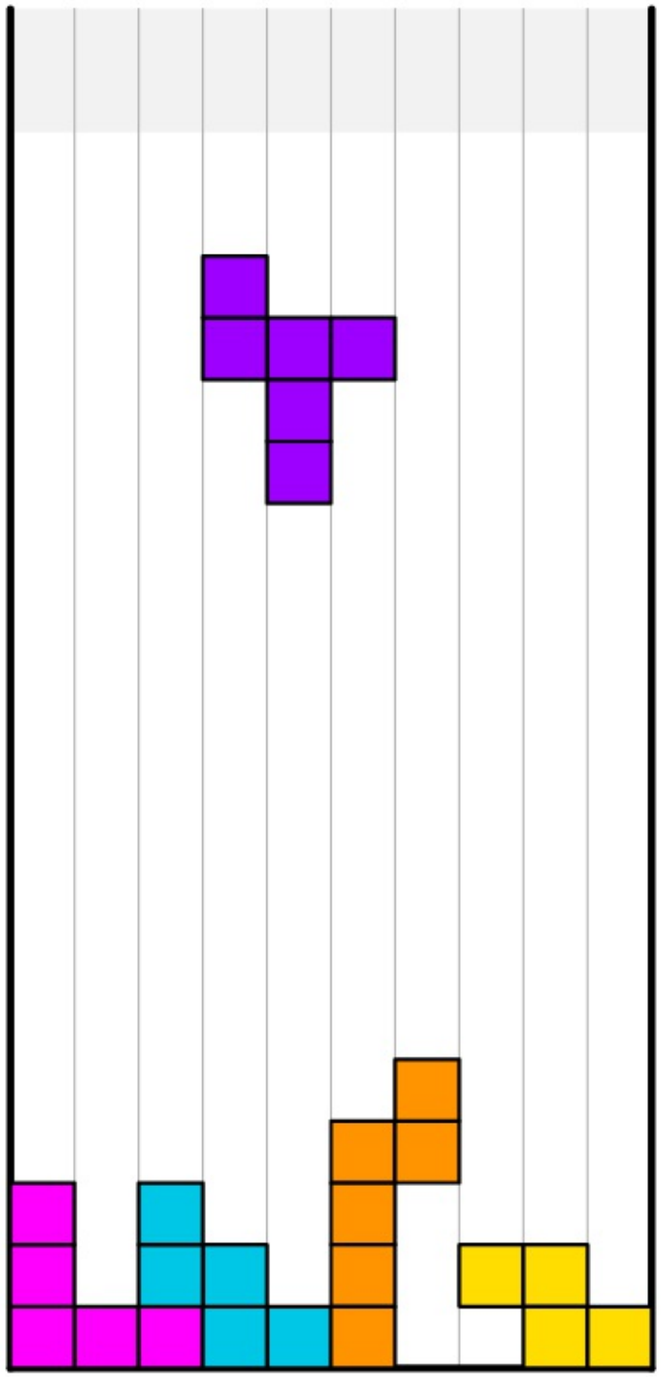
Teaching data science to first-year undergraduates at Duke

- "**tidyverse** first"
- RStudio instances
- Reproducibility with R Markdown
- Version control with GitHub
- Supplement with DataCamp and **learnr** package

* *Mine Çetinkaya-Rundel, Duke University & RStudio* (video)

Packages

poissontris: Shiny game

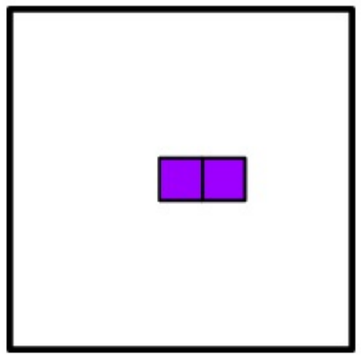


λ: 3.5

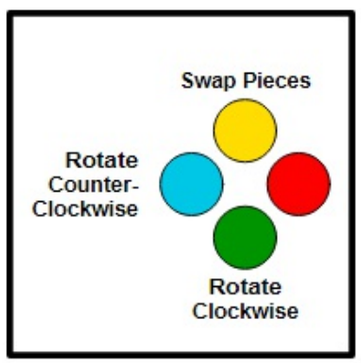
Score: 0

34

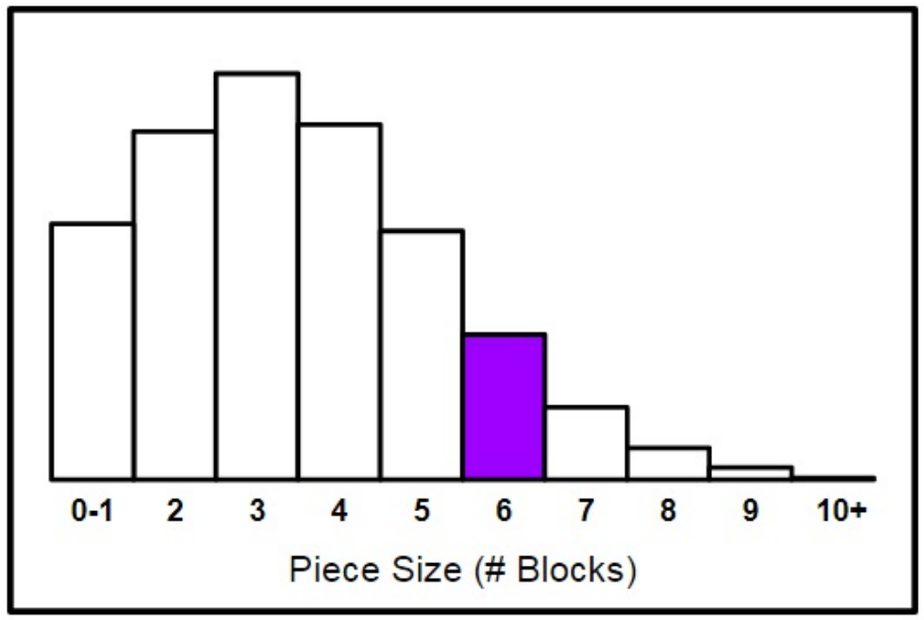
Next Piece:



Controls:



Poisson PDF (λ = 3.5)



Show PC Controls

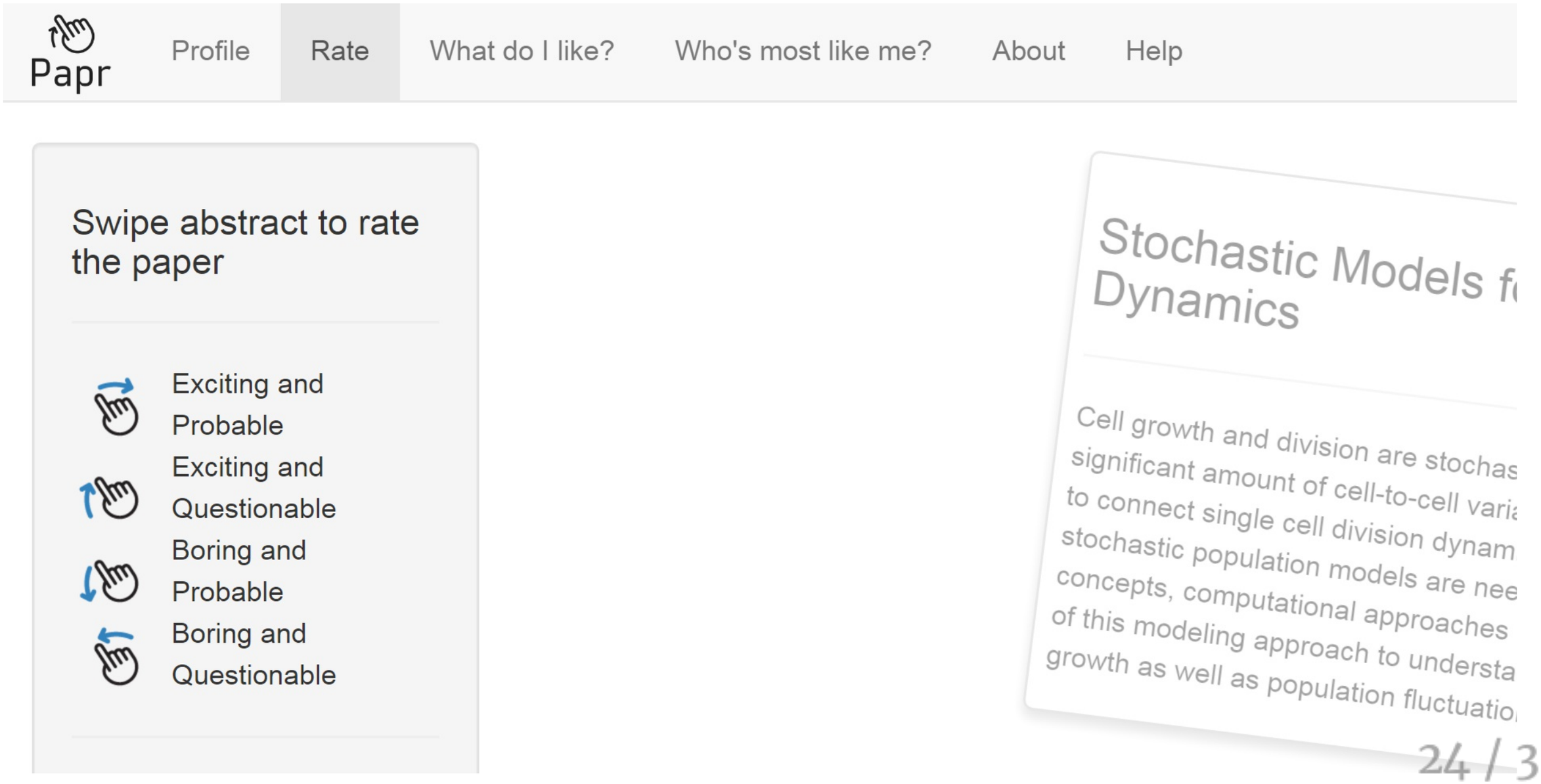
miner: R interface for Minecraft

miner: R interface for Minecraft

miner: R interface for Minecraft

papr: Tinder for pre-prints

- By Jeff Leek (simplystatistics.org)
- Supports bioRxiv (maybe arXiv in the future)



Text mining, the tidy way*

[1] Each variable is a column

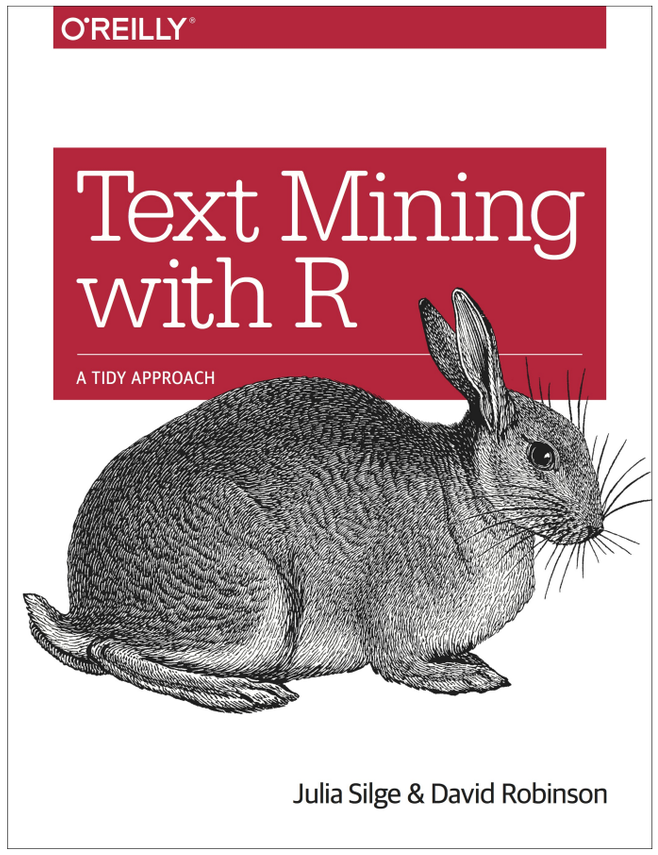
[2] Each observation is a row

[3] Each type of observation is a table

–>> one *token* per row

Tidy text packages:

- **tidytext**
- **manifestoR**
- **cleanNLP**
- **quanteda**



* *Julia Silge, Stack Overflow* (video)

Programming with the tidyverse*

```
band_members
```

```
## # A tibble: 3 x 2
##    name    band
##   <chr>   <chr>
## 1  Mick  Stones
## 2  John Beatles
## 3  Paul Beatles
```

```
x <- 'John'
filter(band_members, name == x)
```

```
## # A tibble: 1 x 2
##    name    band
##   <chr>   <chr>
## 1  John Beatles
```

Programming with the tidyverse*

```
band_members
```

```
## # A tibble: 3 x 2
##    name    band
##   <chr>   <chr>
## 1  Mick  Stones
## 2  John Beatles
## 3  Paul Beatles
```

```
x <- 'John'
filter(band_members, name == x)
```

```
## # A tibble: 1 x 2
##    name    band
##   <chr>   <chr>
## 1  John Beatles
```

```
y <- 'name'
filter(band_members, y == x)
# Nothing found!
```

```
z <- rlang::sym('name')
filter(band_members, (!! z) == x)
```

```
## # A tibble: 1 x 2
##    name    band
##   <chr>   <chr>
## 1  John Beatles
```

* *Lionel Henry, RStudio* (slides)

Using RStan in packages*

- **rstantools**
  - provides `rstan_package_skeleton`
- **rstanarm**: RStan for **a**pplied **r**egression **m**odels
  - `stan_lm`, `stan_aov` for Anova models
  - `stan_glm` for generalised linear models
  - `stan_lmer`, `stan_glmer` for mixed-effects models

* *Ben Goodrich, Columbia University* (video)

Using RStan in packages*

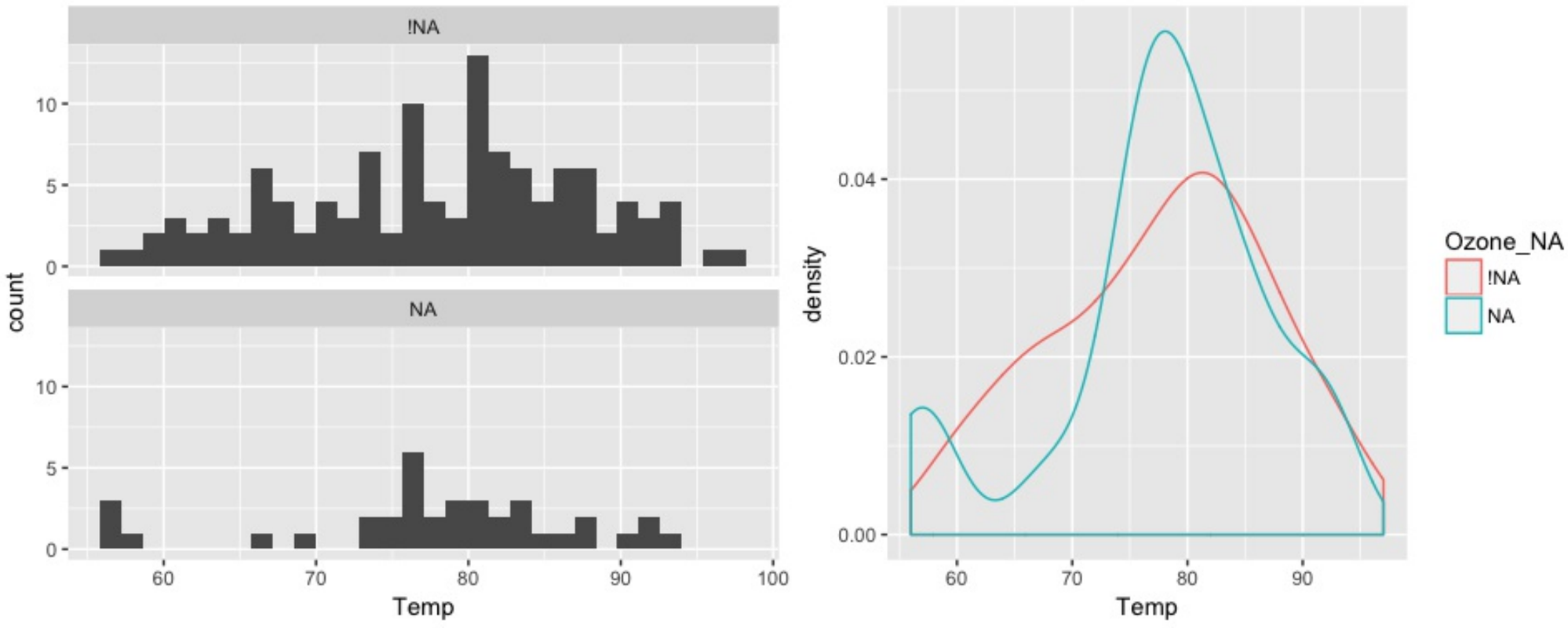
- **rstantools**
  - provides `rstan_package_skeleton`
- **rstanarm**: RStan for **a**pplied **r**egression **m**odels
  - `stan_lm`, `stan_aov` for Anova models
  - `stan_glm` for generalised linear models
  - `stan_lmer`, `stan_glmer` for mixed-effects models

* *Ben Goodrich, Columbia University* (video)

```
glm(y ~ roach1 + treatment + senior, offset = log(exposure2),
    data = roaches, family = poisson)

stan_glm(y ~ roach1 + treatment + senior, offset = log(exposure2),
         data = roaches, family = poisson,
         prior = normal(0,2.5), prior_intercept = normal(0,5),
         chains = CHAINS, cores = CORES, seed = SEED)
```

narnia: consistent exploration of missing data*



* *Nicholas Tierney, Monash University* (web site)

Other cool packages/functions for missing data

```
> visdat   # visualise missing data
```

```
> simputation   # make imputation simple
```

```
> mice::ampute   # simulate missing values
```

Top tips

Speed up your code

Add `ByteCompile: yes` to your package `DESCRIPTION`

Interactive comments

```
ggplot(iris, aes(Petal.Width, Petal.Length)) +
  geom_point() +
  labs(
      x = "Petal width"
    , y = "Petal length"
#   , title = "Fisher's iris data set"
    )
```

Resources

- These slides warwick.ac.uk/wrug
- Videos of *useR!2017* talks* bit.ly/2sUxtPJ
- R-Consortium.org
- RLadies.org
- satRdays.org

* Dootika's is up! Ella's talk isn't there (yet)

FutuRe

R User Groups

- Bay Area User Group
  - format: mix of 3—4 talks of different lengths
    - 12-min lightning talks
    - 20/40-min presentations
  - no permanent venue
- satRdays
- R Consortium
- Meetup.com
- eRum / useR!2018
- Online / recorded meetings

Warwick R User Group

- Now on Meetup.com/Warwick-useRs
  - R Consortium funding
  - Archived material at warwick.ac.uk/wrug
- Evening/weekend events
- Meeting venues
- Online / recorded meetings

Upcoming R events

August <-

Next WRUG meeting?

11–15 September <-

WDSI/IAS Vacation School

Principles and Practice of Data Analysis for Reproducible Research in R

12–14 September <-

EARL conference (London)

Enterprise Applications of the R Language

Thank you!