Wrap Up - R Resources

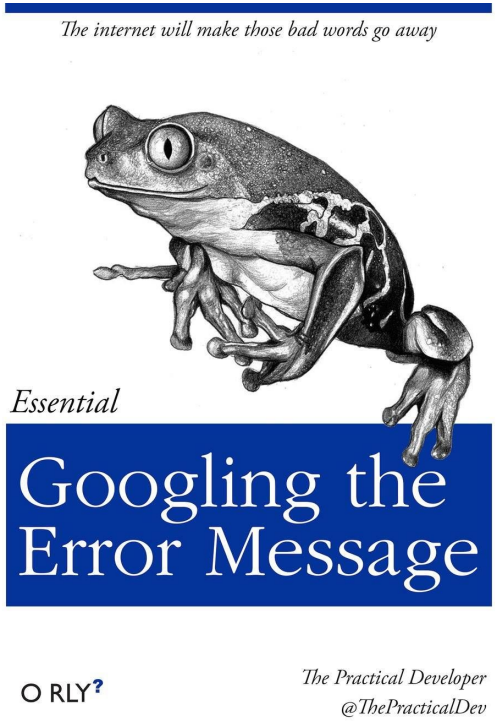
Guy J. Abel

Online Help

- If you get stuck first investigate the help file
  - The help file (e.g. `?log`)
- Most of the topics covered in this course are described basic level in *R for Data Science* by Garrett Grolemund and Hadley Wickham.
  - Available for free online: `http://r4ds.had.co.nz/`
- If this fails use the internet. . .

Online Help

- R is very popular. A lot of people have learnt how to use R, so there is lots of help available.
  - More often than not someone else has had the same problem before and sought help on the internet.
- Google (search) error and warning messages if you do not understand them.

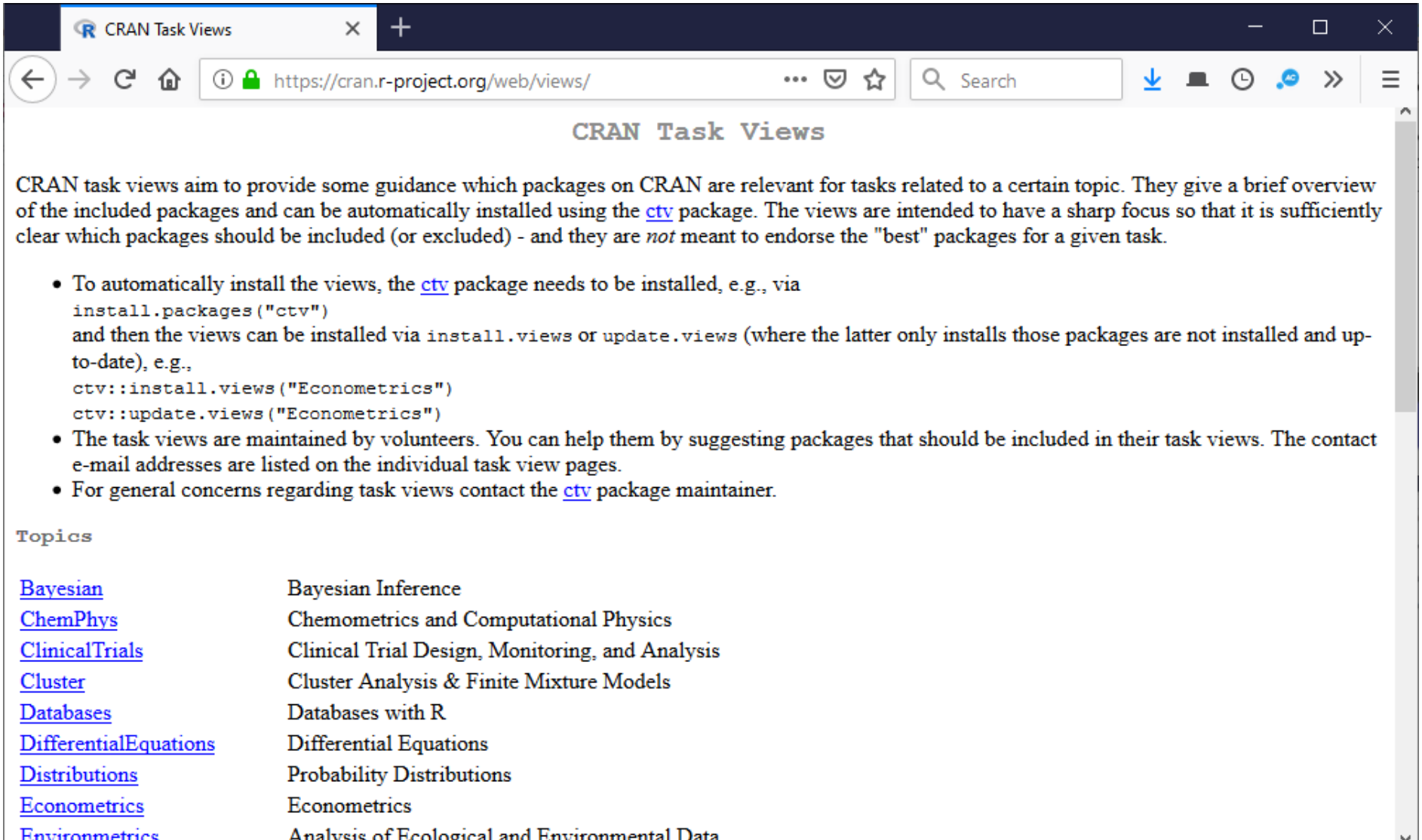


Online Help

- R-Help mailing list `https://stat.ethz.ch/mailman/listinfo/r-help`
  - Long history of discussions on R.
  - Archive at `https://stat.ethz.ch/pipermail/r-help/`
  - Reputation for aggressive responders.

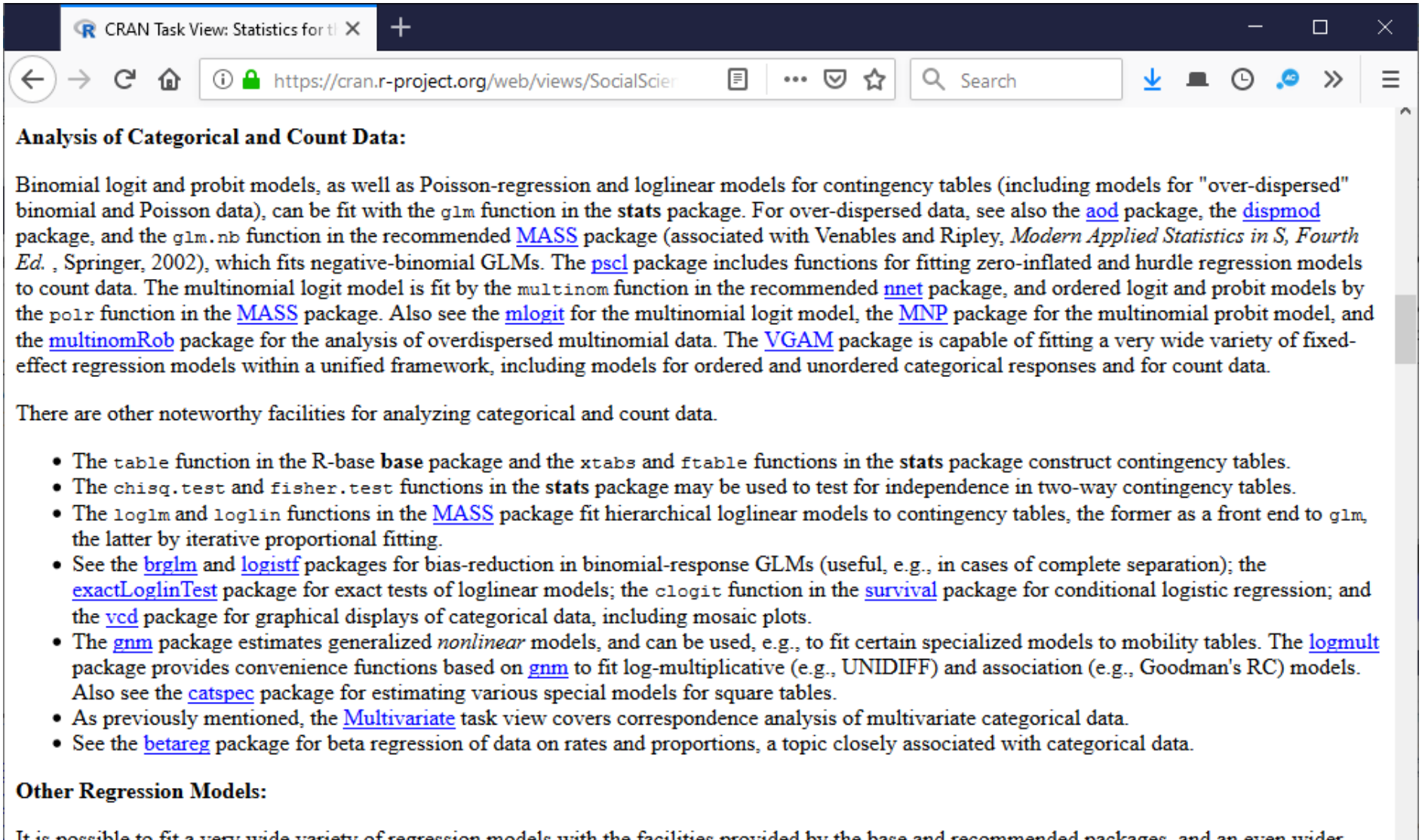
CRAN Task Views

- CRAN Task Views provide regularly updated overview of packages for particular methods.



CRAN Task Views

- CRAN Task Views provide regularly updated overview of packages for particular methods.



Analysis of Categorical and Count Data:

Binomial logit and probit models, as well as Poisson-regression and loglinear models for contingency tables (including models for "over-dispersed" binomial and Poisson data), can be fit with the `glm` function in the **stats** package. For over-dispersed data, see also the aod package, the dispmod package, and the `glm.nb` function in the recommended MASS package (associated with Venables and Ripley, *Modern Applied Statistics in S, Fourth Ed.* , Springer, 2002), which fits negative-binomial GLMs. The pscl package includes functions for fitting zero-inflated and hurdle regression models to count data. The multinomial logit model is fit by the `multinom` function in the recommended nnet package, and ordered logit and probit models by the `polr` function in the MASS package. Also see the mlogit for the multinomial logit model, the MNP package for the multinomial probit model, and the multinomRob package for the analysis of overdispersed multinomial data. The VGAM package is capable of fitting a very wide variety of fixed-effect regression models within a unified framework, including models for ordered and unordered categorical responses and for count data.

There are other noteworthy facilities for analyzing categorical and count data.

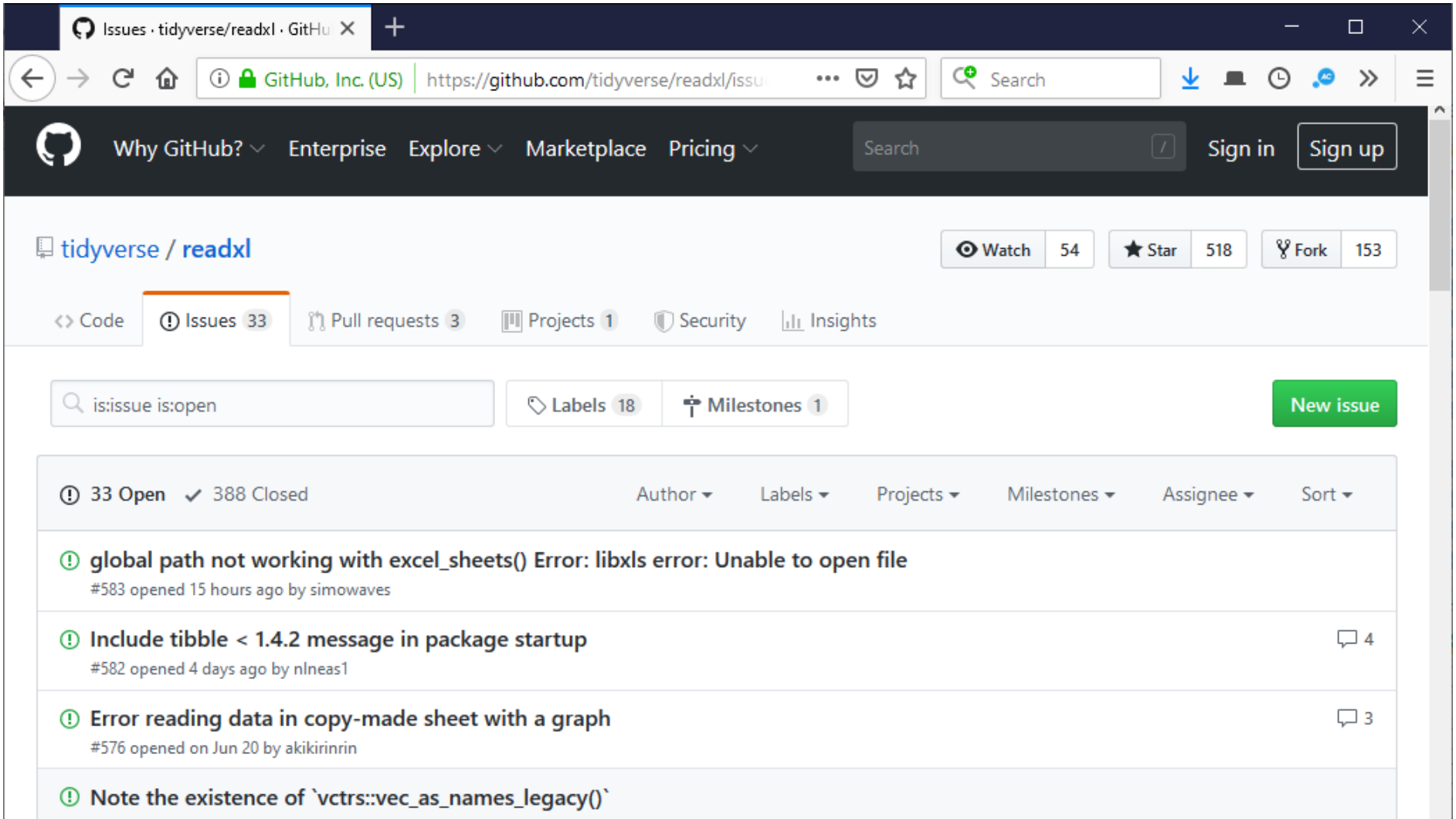
- The `table` function in the R-base **base** package and the `xtabs` and `ftable` functions in the **stats** package construct contingency tables.
- The `chisq.test` and `fisher.test` functions in the **stats** package may be used to test for independence in two-way contingency tables.
- The `loglm` and `loglin` functions in the MASS package fit hierarchical loglinear models to contingency tables, the former as a front end to `glm`, the latter by iterative proportional fitting.
- See the brglm and logistf packages for bias-reduction in binomial-response GLMs (useful, e.g., in cases of complete separation); the exactLoglinTest package for exact tests of loglinear models; the `clogit` function in the survival package for conditional logistic regression; and the vcd package for graphical displays of categorical data, including mosaic plots.
- The gnm package estimates generalized *nonlinear* models, and can be used, e.g., to fit certain specialized models to mobility tables. The logmult package provides convenience functions based on gnm to fit log-multiplicative (e.g., UNIDIFF) and association (e.g., Goodman's RC) models. Also see the catspec package for estimating various special models for square tables.
- As previously mentioned, the Multivariate task view covers correspondence analysis of multivariate categorical data.
- See the betareg package for beta regression of data on rates and proportions, a topic closely associated with categorical data.

Other Regression Models:

It is possible to fit a very wide variety of regression models with the facilities provided by the base and recommended packages, and an even wider

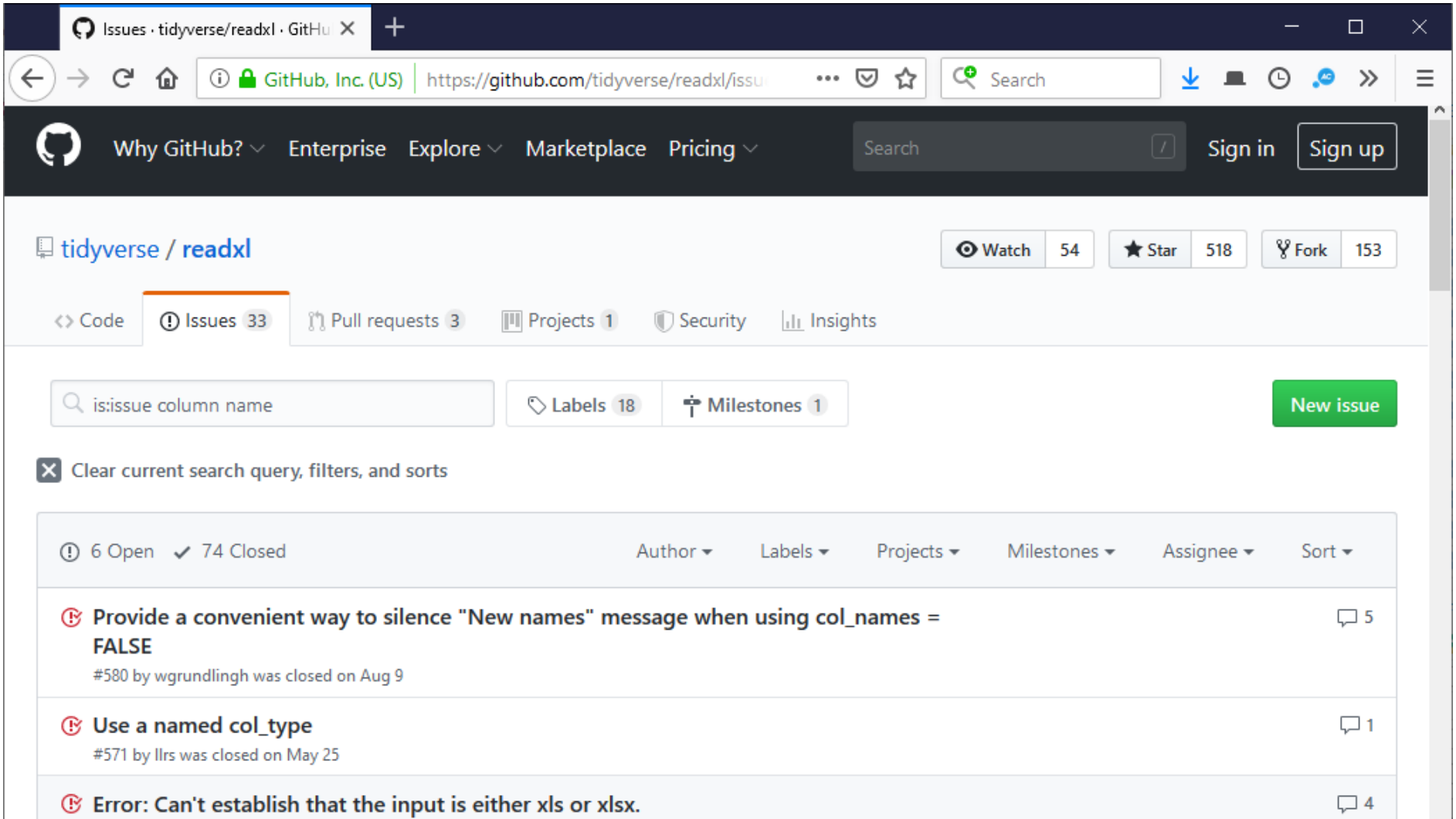
Github

- Github open and closed issues for a particular package.
  - Many package developers used Github.
  - Can access latest version of package, before go onto CRAN.



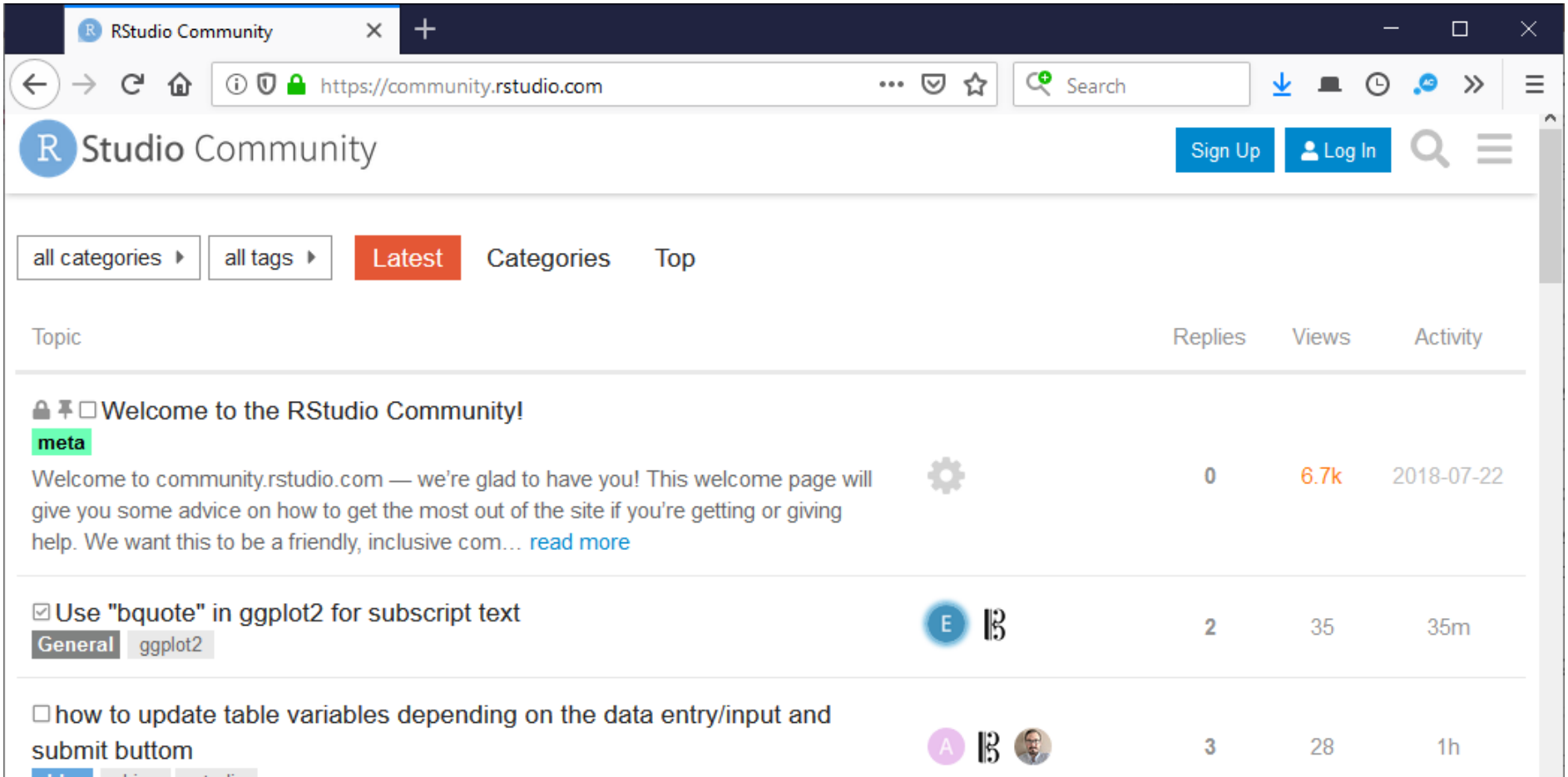
Github

- Github open and closed issues for a particular package.
  - Issues page often the best place to contact the author and get a response.
  - Search open and closed issues before posting a new issue.



RStudio Community Discussion Board

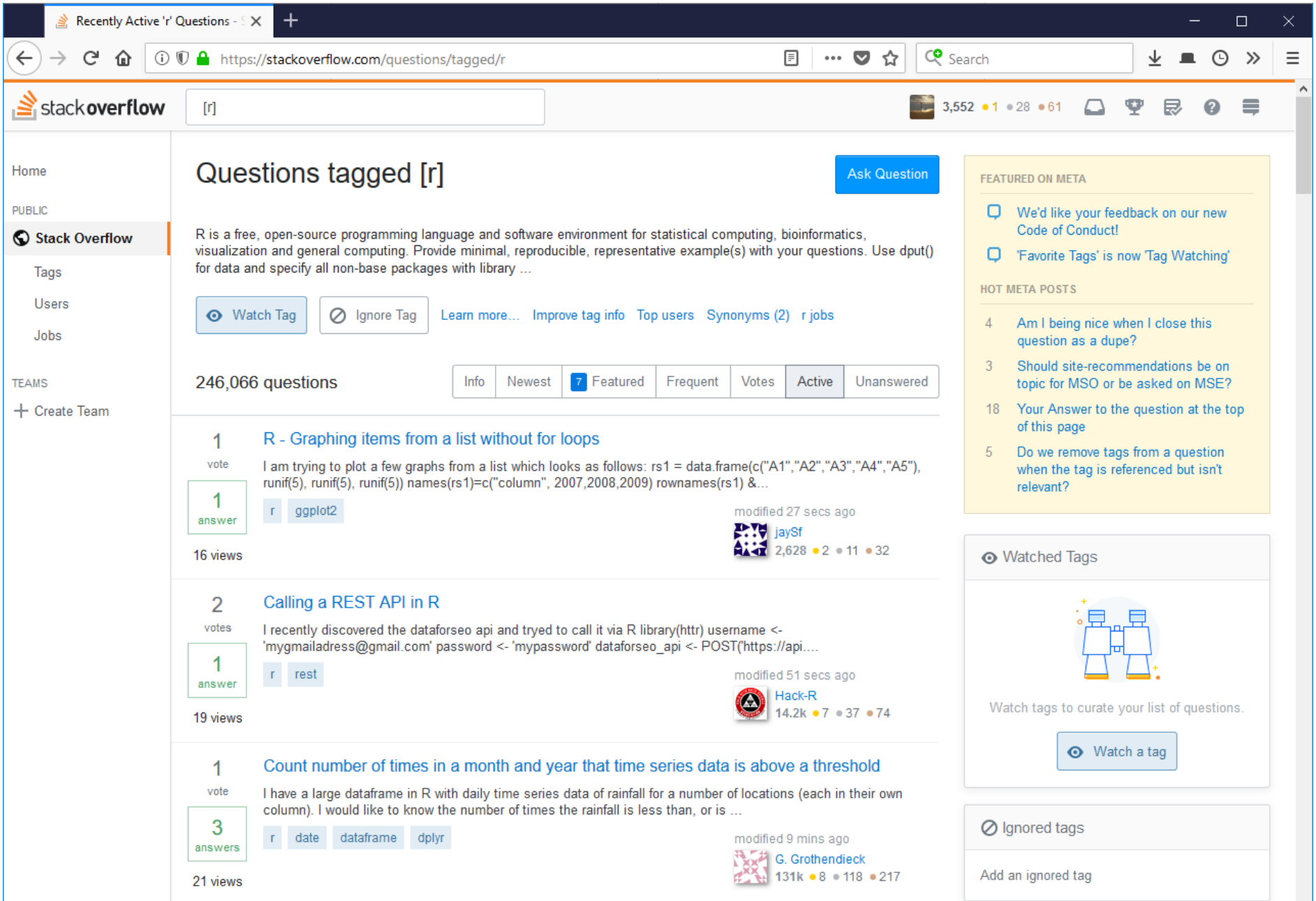
- RStudio Community discussion board at `https://community.rstudio.com/`
  - Well maintained question and answer board for asking questions on tidyverse and beyond
  - More likely to get replies from people who work at RStudio



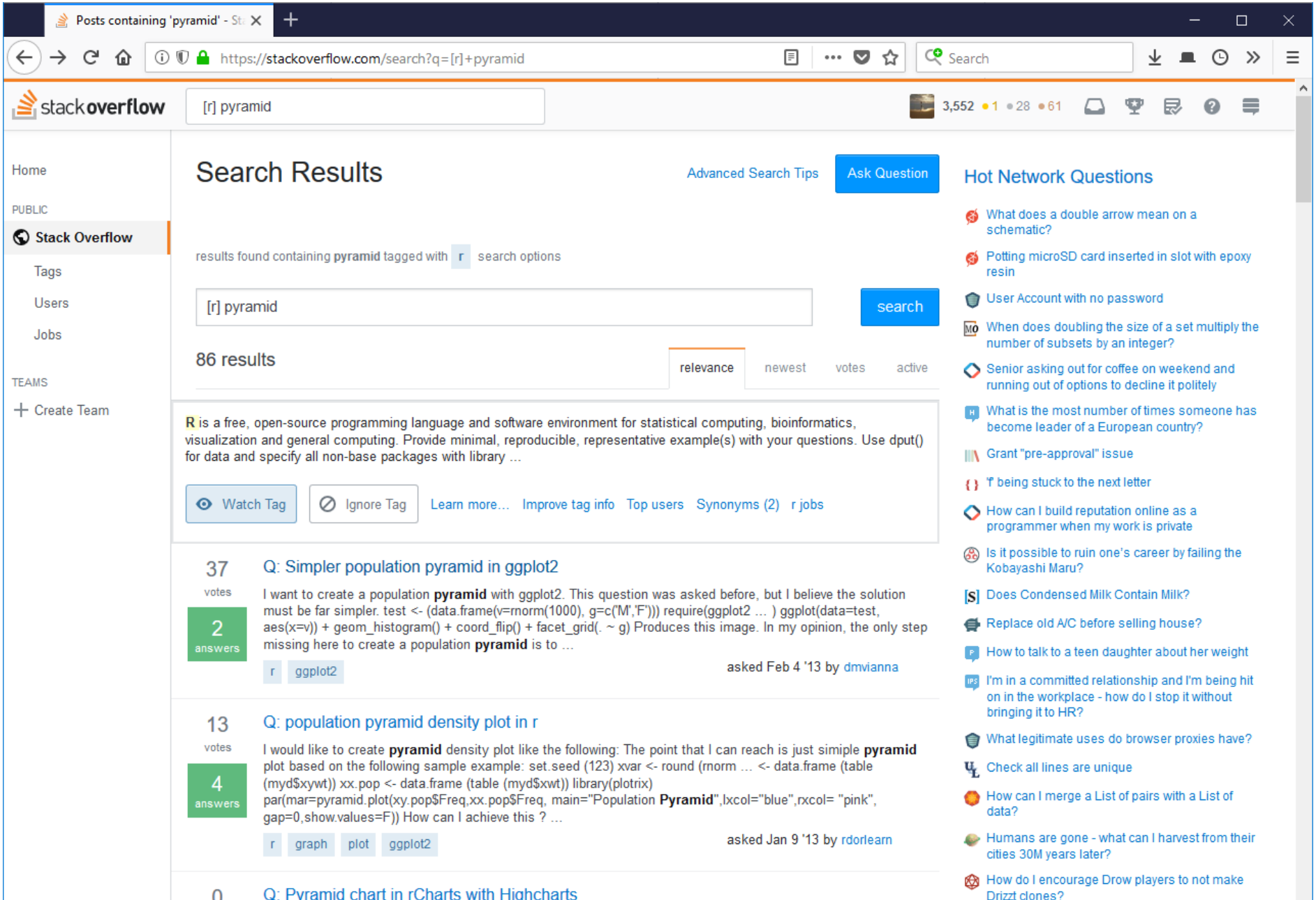
Stackoverflow R tag

- Stackoverflow uses tag to (cross) categorize questions.
- A little more wild than RStudio discussion board but pretty reliable.
- If you ask a good question users respond very quickly.
  - How do I ask a good question
  - How to make a great R reproducible example
- All questions on R are tagged `[r]`
  - `http://stackoverflow.com/questions/tagged/r`

Stackoverflow R tag



Search the stackoverflow R tag



Typical Question

r - Simpler population pyramid

https://stackoverflow.com/questions/14680075/simpler-population-pyramid-in-ggplot2

Search...

stackoverflow

Search...

3,552 ●1 ●28 ●61

Home

PUBLIC

Stack Overflow

Tags

Users

Jobs

TEAMS

Create Team

Simpler population pyramid in ggplot2

Ask Question

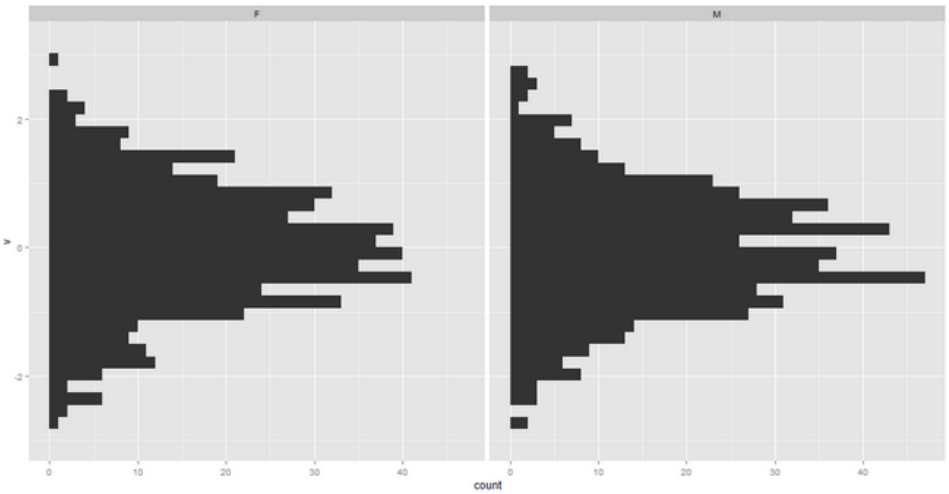
37

I want to create a population pyramid with ggplot2. This question was asked before, but I believe the solution must be far simpler.

```
test <- (data.frame(v=rnorm(1000), g=c('M','F')))
require(ggplot2)
ggplot(data=test, aes(x=v)) +
    geom_histogram() +
    coord_flip() +
    facet_grid(. ~ g)
```

29

Produces this image. In my opinion, the only step missing here to create a population pyramid is to invert the x axis of the first facet, so that is goes from 50 to 0, while keeping the second untouched. Can anyone help?



r ggplot2

asked 5 years, 5 months ago

viewed 12,342 times

active 4 months ago

FEATURED ON META

We'd like your feedback on our new Code of Conduct!

'Favorite Tags' is now 'Tag Watching'

HOT META POSTS

4 Am I being nice when I close this question as a dupe?

3 Should site-recommendations be on topic for MSO or be asked on MSE?

18 Your Answer to the question at the top of this page

5 Do we remove tags from a question when the tag is referenced but isn't relevant?

Linked

0 Unusual pyramid plot in R or matplotlib

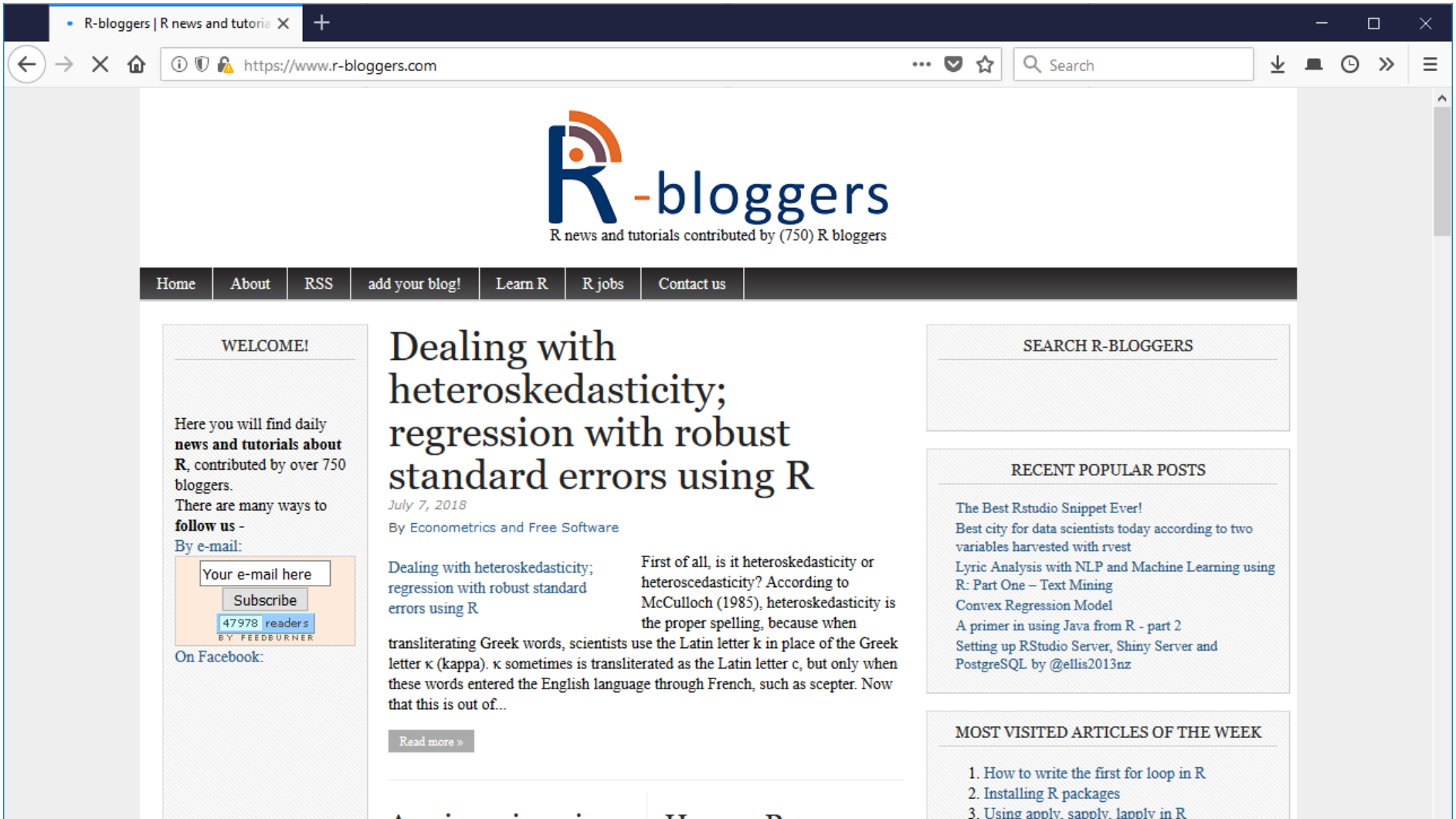
21 drawing pyramid plot using R and ggplot2

13 population pyramid density plot in r

15 Two horizontal bar charts with shared axis in ggplot2 (similar to population pyramid)

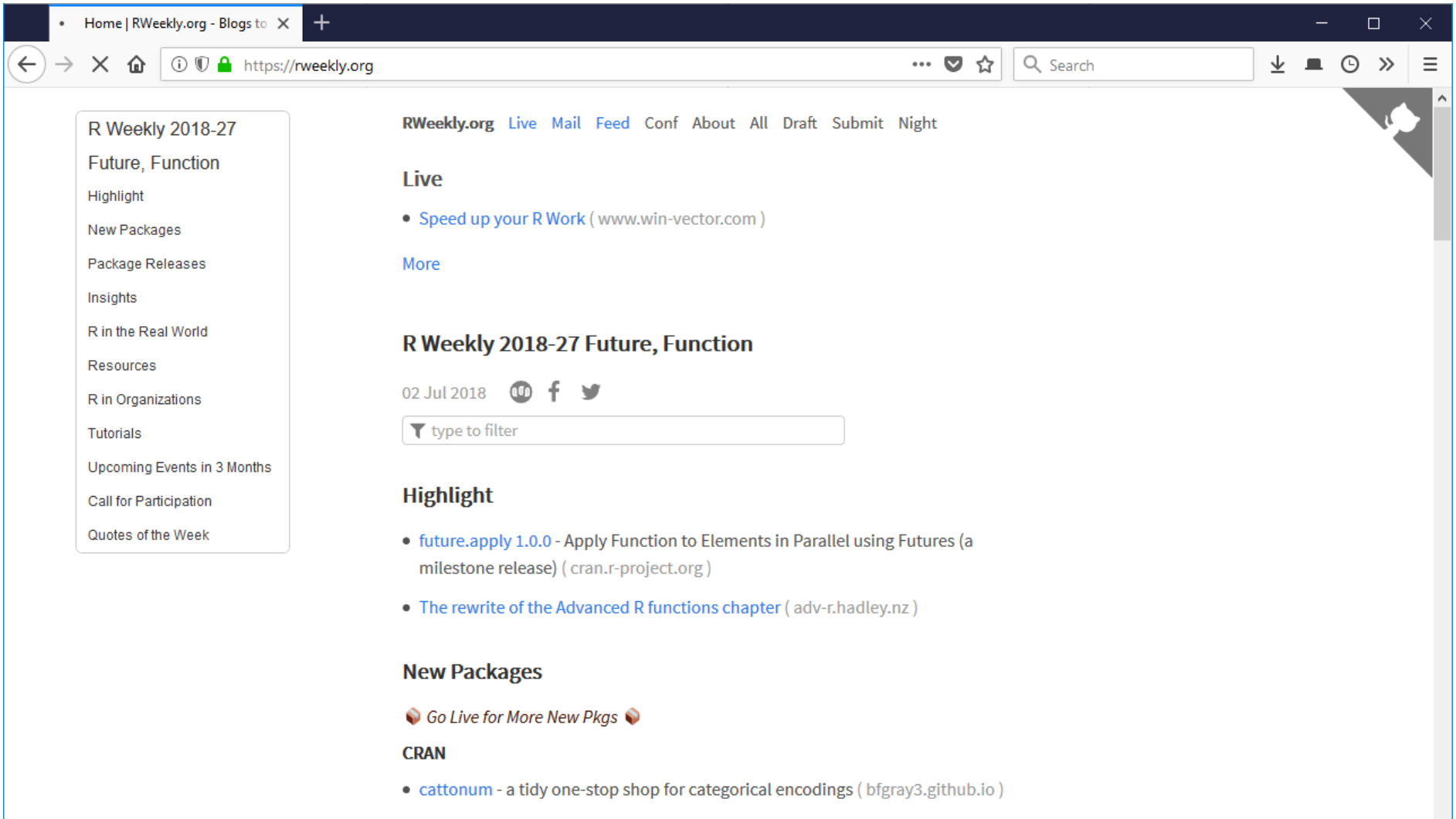
R-bloggers

- R-Bloggers collects blog post of R users
  - `https://www.r-bloggers.com/`
  - Email list that can send you updates.



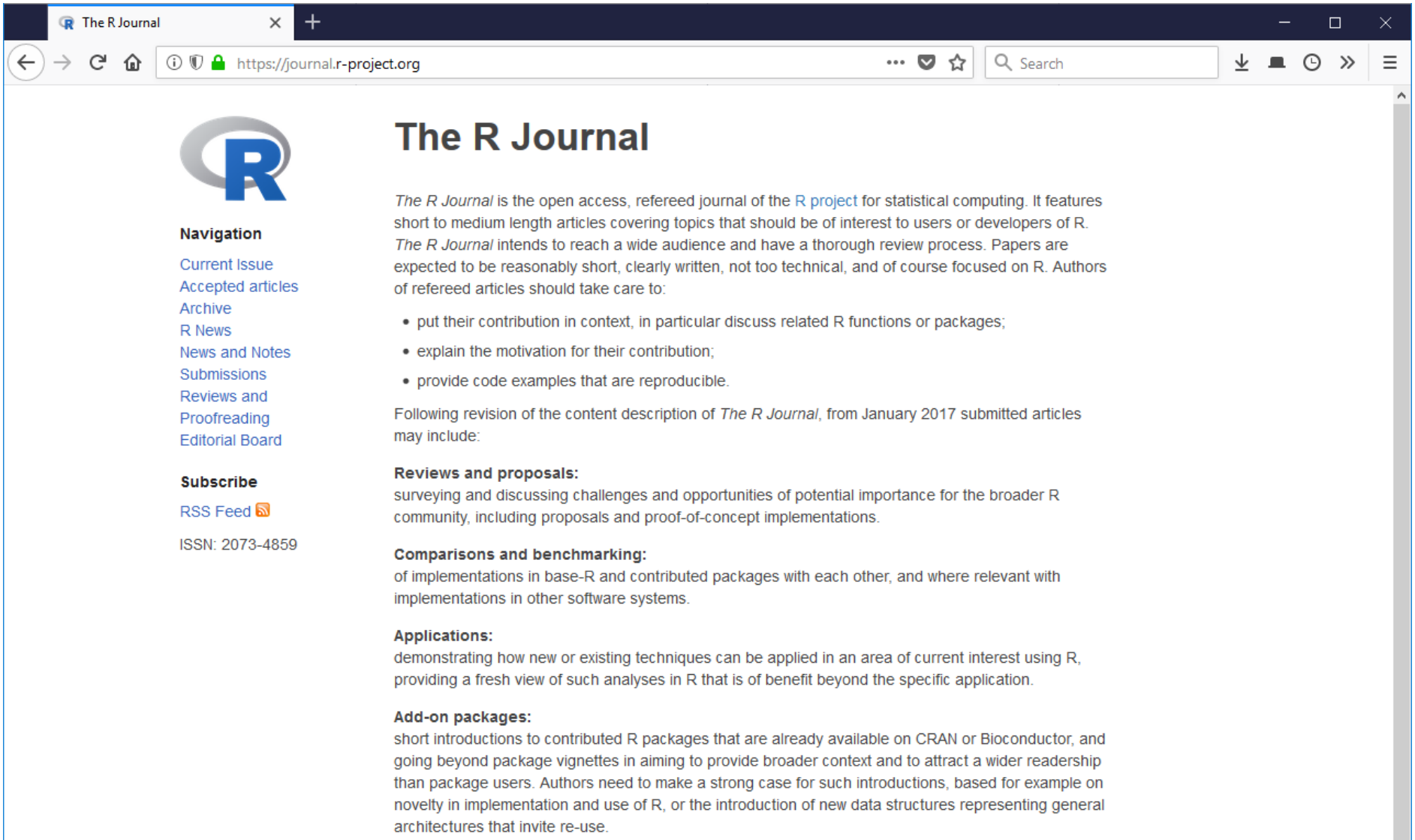
R Weekly

- R Weekly rounds up submitted posts and interesting links
  - `https://rweekly.org/`
  - Email list that can send you updates.



The R Journal

- The R Journal publishes articles on R packages and code snippets.
  - `https://journal.r-project.org/`



Journal of Statistical Software

- Journal of Statistical Software often includes articles on R packages
  - `https://www.jstatsoft.org/`

Final tips

- Copy and pasting code is done.
  - I do it all the time
  - It is not a bad way to get initial results.
- BUT, try and form a habits to
  - Read the help files for the function you are using.
  - Type code out yourself.
- You will gain a better understanding of what you are programming.
  - Allow you to tweak for your context
  - Will avoid using unsuitable settings in functions that might give a wrong answer (when copy and pasted)

Contact Details

- Contact Details
  - guy.abel@shu.edu.cn
  - guyabelguyabel
- Thank You!