

contributing to the tidyverse

Mara Averick
tidyverse developer advocate
@dataandme

<http://bit.ly/tverse-contrib>

rstudio::conf
from RStudio

CONTRIBUTING TO THE

tidyverse

What is the tidyverse?

The tidyverse is an opinionated collection of R packages designed for data science.

All packages share an underlying design philosophy, grammar, and data structures.

Key FOSS actors

- maintainers
- contributors
- consumers/users
- sustainers

GitHub HQ (SF) | June 19, 2017

Sustain

A ONE DAY CONVERSATION
for Open Source Software sustainers

THE REPORT

the thoughtful user



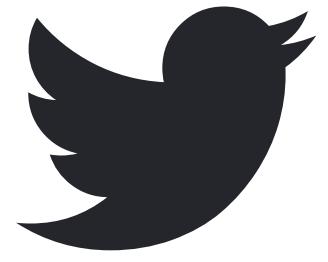
asking questions

The most useless problem statement that one can face is "it doesn't work", yet we seem to get it far too often.

- Thiago Maciera



where to ask



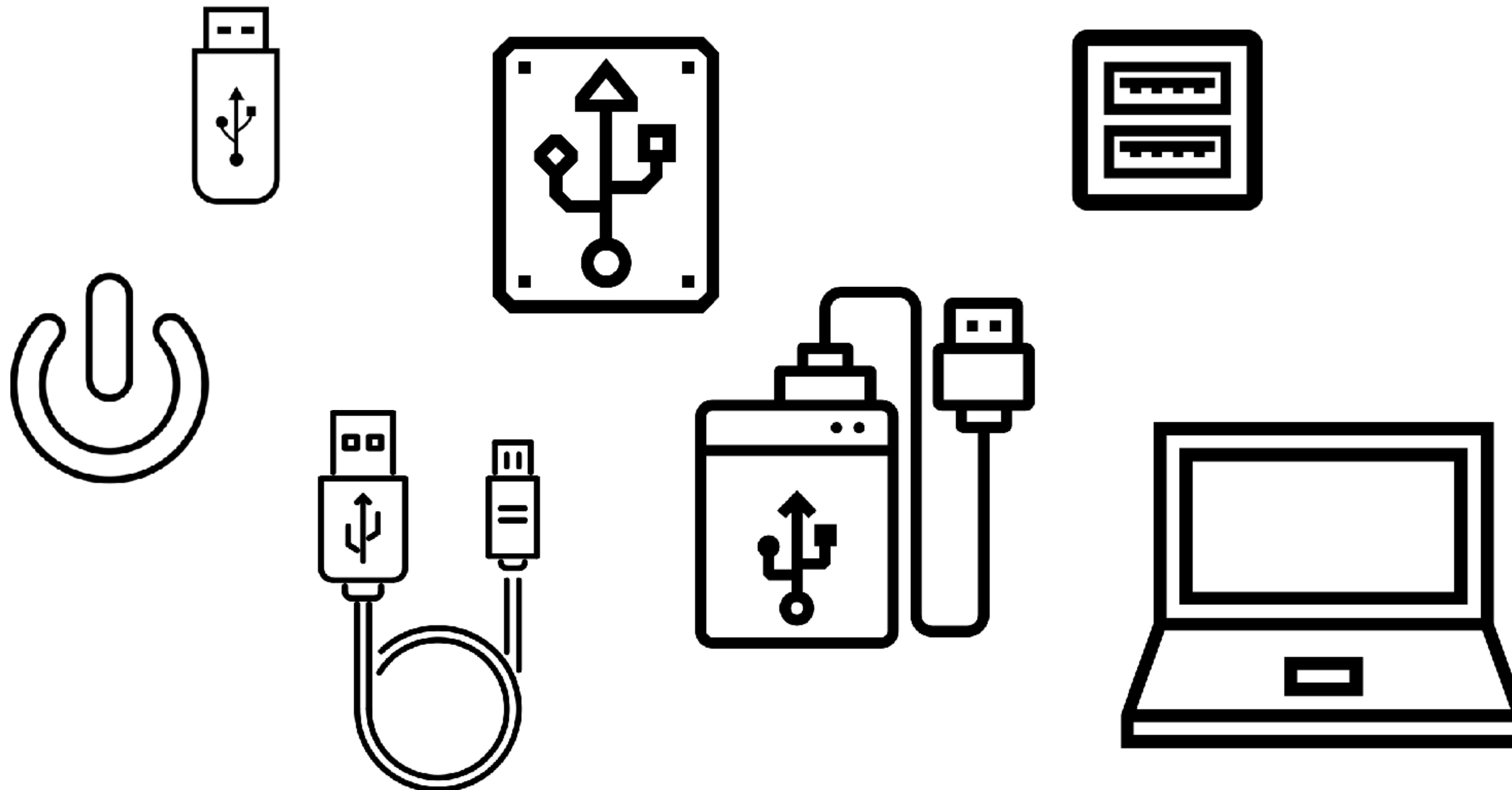
Twitter



RStudio Community



StackOverflow



The newcomer's paradox...



When you ask for help, some friendly soul will no doubt tell you that "*it's easy, just do foo, bar and baz.*" Except for you, it is not easy, there may be no documentation for foo, bar is not doing what it is supposed to be doing and what is this baz thing anyway with its eight disambiguation entries on Wikipedia?

— Leslie Hawthorne

"You'll Eventually Know Everything They've Forgotten." In *Open Advice: FOSS: What We Wish We Had Known When We Started*, edited by Lydia Pintscher, 29–32.

meta-help

helping others to help themselves...

10



meta-help



Why does head() show 6 rows by default?



hadley RStudio Employee

Nov '17

To answer a question like this, I first start looking through the S books I have on hand (e.g. [The New S Language](#) 11). I don't see `head()` mentioned in the index, so that suggests it's a function introduced by R.

Since it's an R function, I can next search [@winston](#)'s GitHub mirror of the R sources:

<https://github.com/wch/r-source> 16, finding the source at <https://github.com/wch/r-source/blob/af7f52f70101960861e5d995d3a4bec010bc89e6/src/library/utils/R/head.R> 20

This includes a comment which suggests we should ask Patrick Burns:

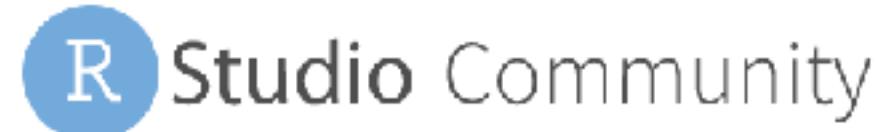
```
### placed in the public domain 2002
### Patrick Burns patrick@burns-stat.com
###
### Adapted for negative arguments by Vincent Goulet
### <vincent.goulet@act.ulaval.ca>, 2006
```

But it's worth checking just to make sure it's always used 6. I click on history, and then find the first version:

<https://github.com/wch/r-source/commit/37271cdbcd7e5d82c79bdb536ef305d93b644ad#diff-941bf47bf09f67538338535bd512d521> 28 - so it has been six from the very beginning.

So next step, I'll email Patrick and see if he recollects...

meta-help



Why does head() show 6 rows by default?



hadley RStudio Employee

Nov '17

To answer a question like this, I first start looking through the S books I have on hand (e.g. [The New S Language](#) 11). I don't see `head()` mentioned in the index, so that suggests it's a function introduced by R.

Look in S books

Since it's an R function, I can next search [@winston](#)'s GitHub mirror of the R sources: <https://github.com/wch/r-source> 16, finding the source at <https://github.com/wch/r-source/blob/af7f52f70101960861e5d995d3a4bec010bc89e6/src/library/utils/R/head.R> 20

Go to R-source mirror

This includes a comment which suggests we should ask Patrick Burns:

```
### placed in the public domain 2002
### Patrick Burns patrick@burns-stat.com
###
### Adapted for negative arguments by Vincent Goulet
### <vincent.goulet@act.ulaval.ca>, 2006
```

Find relevant bit

But it's worth checking just to make sure it's always used 6. I click on history, and then find the first version: <https://github.com/wch/r-source/commit/37271cdbcd7e5d82c79bdb536ef305d93b644ad#diff-941bf47bf09f67538338535bd512d521> 28 - so it has been six from the very beginning.

See the history

So next step, I'll email Patrick and see if he recollects...

Send an email

why does n = 6?



hadley RStudio Employee

Nov '17

From Pat (via email):

I came upon ‘head’ and ‘tail’ at one of my clients. That implementation had $n = 5$. I didn’t think there would ever be an issue regarding ownership of the code, but I changed to 6 just to help if there were a conflict.

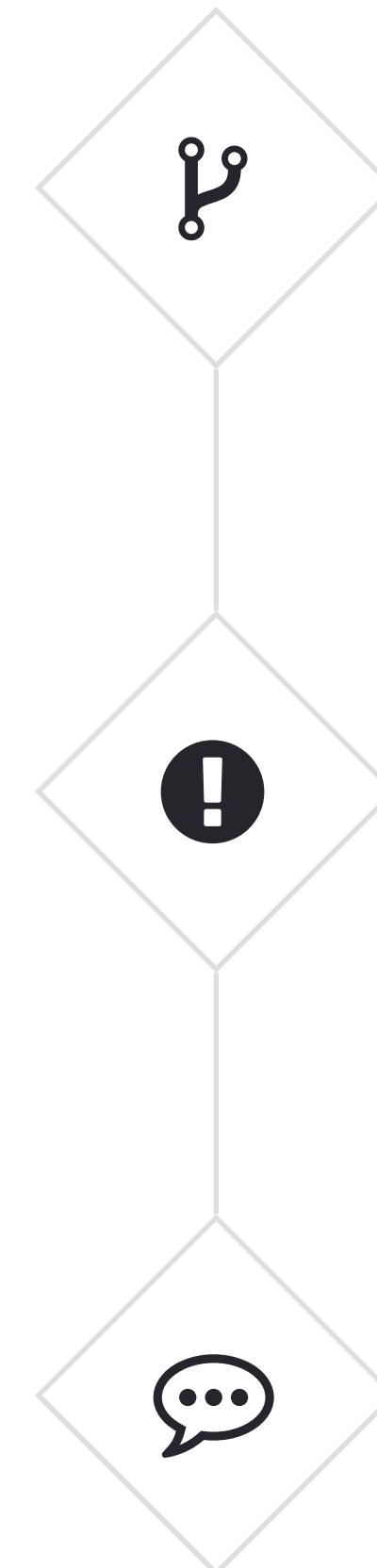


mara

Nov '17

n = 6 : R :: brown M&Ms : Van Halen

Ways to contribute



PULL REQUESTS

Contributing code/making fixes.

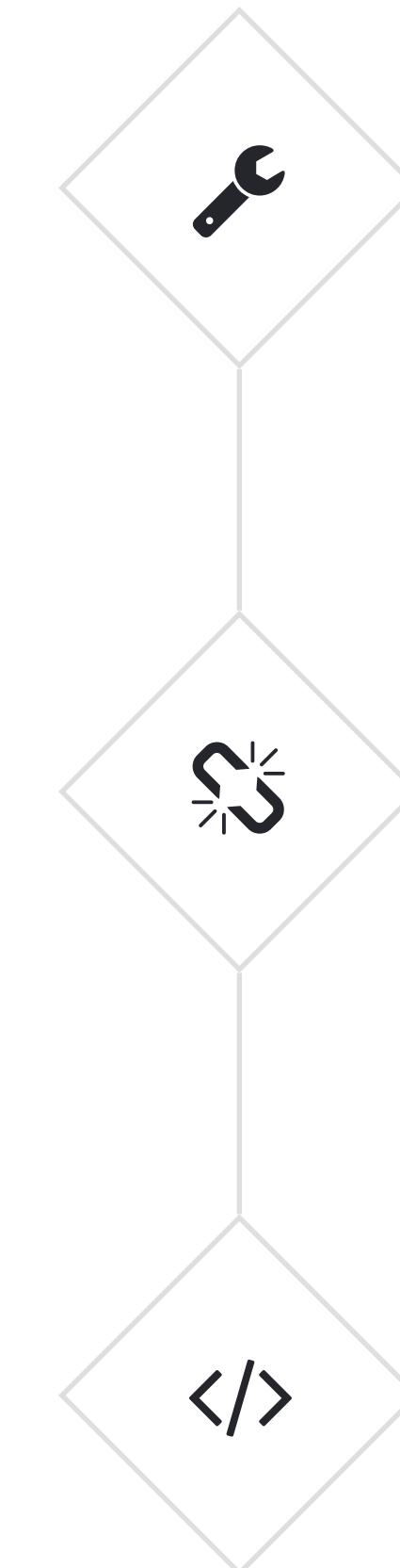
ISSUES

Identifying a problem, trying your best to isolate its source.

COMMENTS

Help maintainers answer questions, triage issues. Help newcomers learn how to ask better questions (e.g. the art of the reprex).

the anatomy of an issue

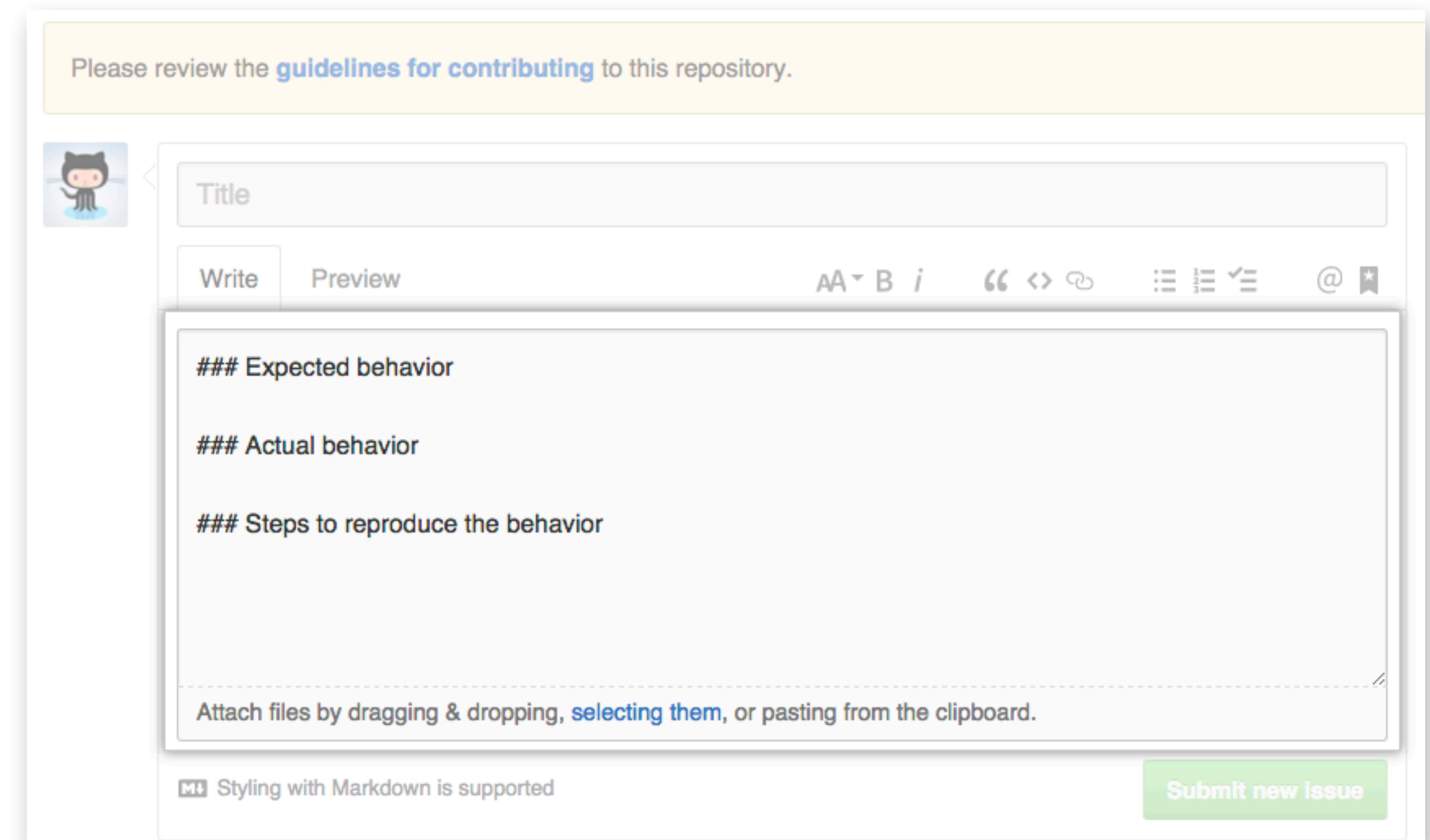


PROBLEM DESCRIPTION

EXPECTED BEHAVIOUR

MINIMAL REPRODUCIBLE EXAMPLE

the anatomy of an issue



"Creating an issue template for your repository" – GitHub Help <<https://help.github.com/articles/creating-an-issue-template-for-your-repository/>>

My first “contribution”

The screenshot shows a GitHub pull request page for the repository `hadley / r4ds`. The pull request is titled "Fixed spelling of purrr on line 121 (#208)". It was committed by `batpigandme` and `hadley` on July 29, 2016. The commit message is "Fixed spelling of purrr on line 121". The code change is shown in the file `model-basics.Rmd`, specifically line 121. A circular highlight surrounds the word "purrrr" in the commit message and its corresponding line in the code diff.

Fixed spelling of purrr on line 121 (#208)

master (#208)

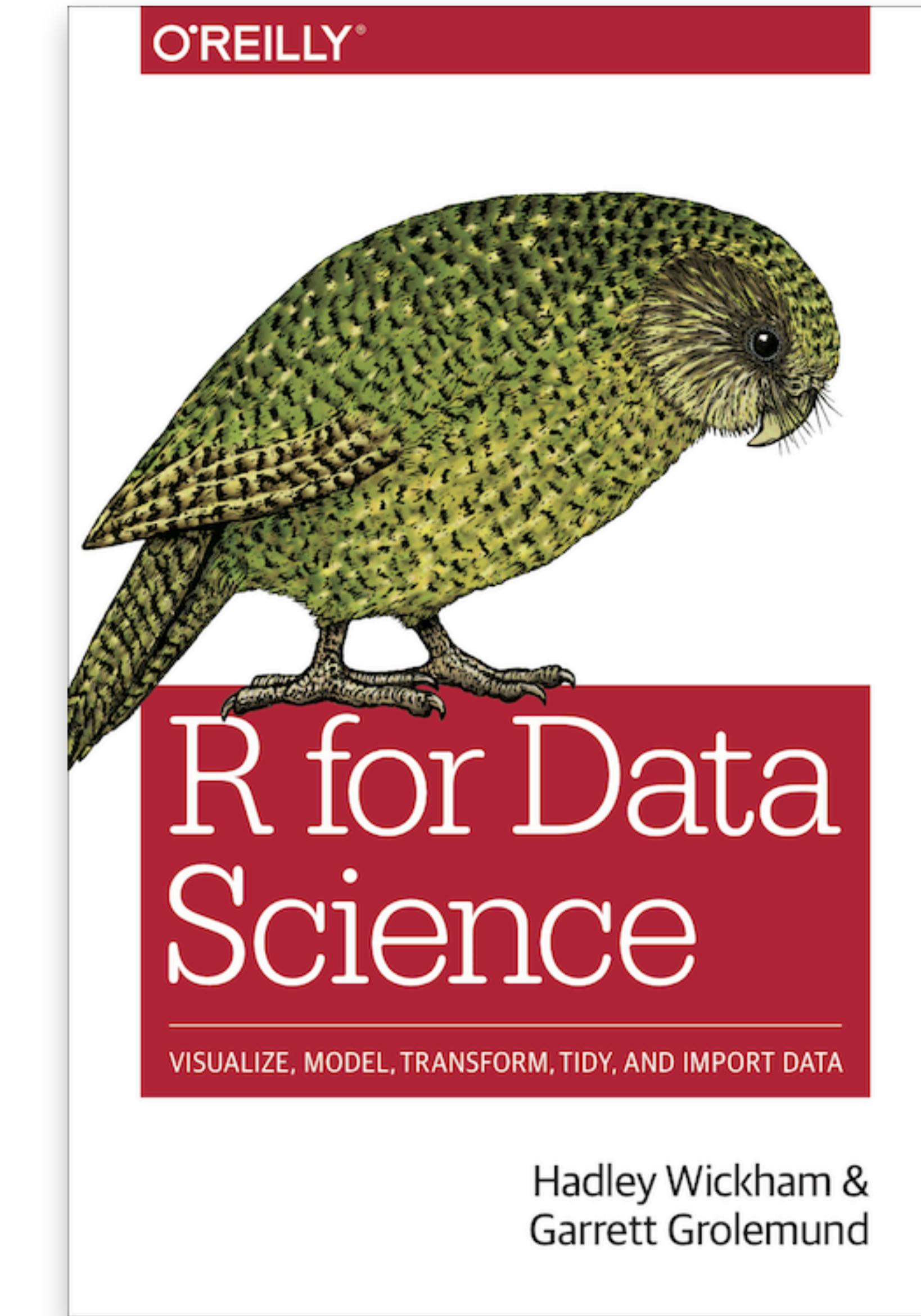
batpigandme committed with hadley on Jul 29, 2016

Showing 1 changed file with 1 addition and 1 deletion.

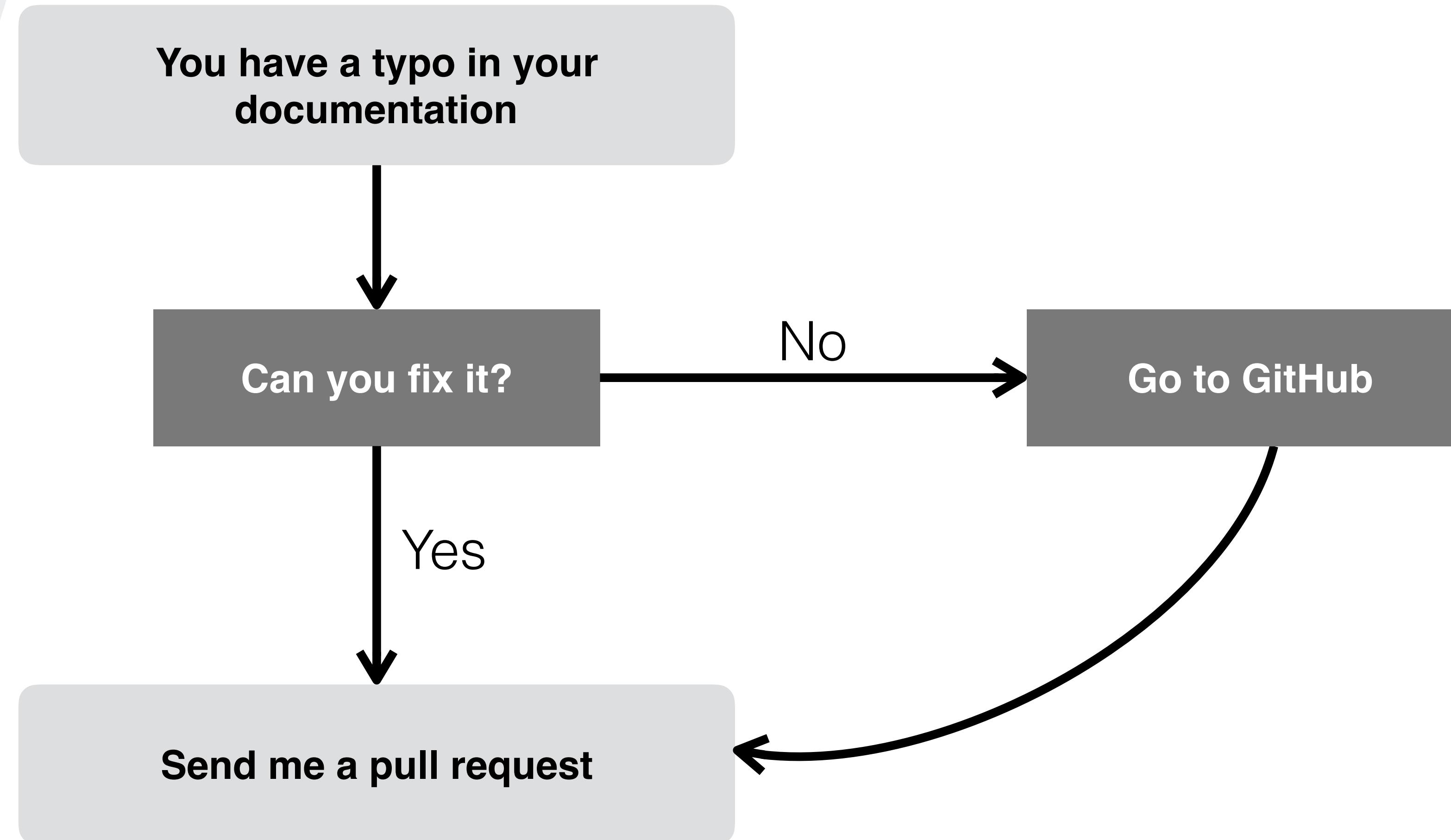
model-basics.Rmd

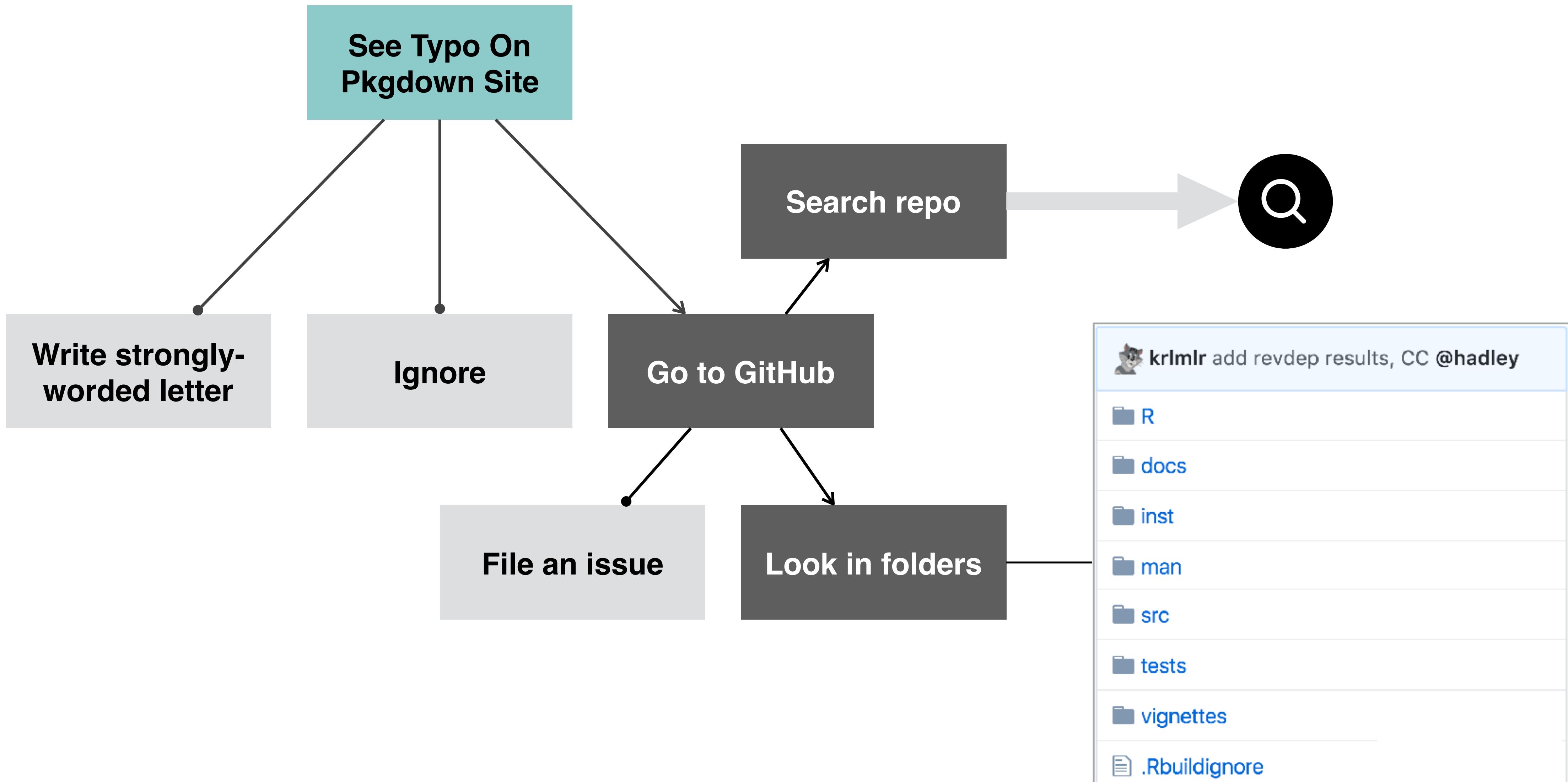
```
@@ -118,7 +118,7 @@ measure_distance <- function(mod, data) {  
 118 118  measure_distance('1.5), sim1)  
 119 119  ...  
 120 120  
 121 -Now we  
 121 -distan  
 121 +Now we  
 121 +distan  
 122 122  
 123 123  ``{r}  
 124 124  sim1_dist <- function(a1, a2) {  
 125 125  
 126 126  
 127 127  
 128 128  
 129 129  
 130 130  
 131 131  
 132 132  
 133 133  
 134 134  
 135 135  
 136 136  
 137 137  
 138 138  
 139 139  
 140 140  
 141 141  
 142 142  
 143 143  
 144 144  
 145 145  
 146 146  
 147 147  
 148 148  
 149 149  
 150 150  
 151 151  
 152 152  
 153 153  
 154 154  
 155 155  
 156 156  
 157 157  
 158 158  
 159 159  
 160 160  
 161 161  
 162 162  
 163 163  
 164 164  
 165 165  
 166 166  
 167 167  
 168 168  
 169 169  
 170 170  
 171 171  
 172 172  
 173 173  
 174 174  
 175 175  
 176 176  
 177 177  
 178 178  
 179 179  
 180 180  
 181 181  
 182 182  
 183 183  
 184 184  
 185 185  
 186 186  
 187 187  
 188 188  
 189 189  
 190 190  
 191 191  
 192 192  
 193 193  
 194 194  
 195 195  
 196 196  
 197 197  
 198 198  
 199 199  
 200 200  
 201 201  
 202 202  
 203 203  
 204 204  
 205 205  
 206 206  
 207 207  
 208 208  
 209 209  
 210 210  
 211 211  
 212 212  
 213 213  
 214 214  
 215 215  
 216 216  
 217 217  
 218 218  
 219 219  
 220 220  
 221 221  
 222 222  
 223 223  
 224 224  
 225 225  
 226 226  
 227 227  
 228 228  
 229 229  
 230 230  
 231 231  
 232 232  
 233 233  
 234 234  
 235 235  
 236 236  
 237 237  
 238 238  
 239 239  
 240 240  
 241 241  
 242 242  
 243 243  
 244 244  
 245 245  
 246 246  
 247 247  
 248 248  
 249 249  
 250 250  
 251 251  
 252 252  
 253 253  
 254 254  
 255 255  
 256 256  
 257 257  
 258 258  
 259 259  
 260 260  
 261 261  
 262 262  
 263 263  
 264 264  
 265 265  
 266 266  
 267 267  
 268 268  
 269 269  
 270 270  
 271 271  
 272 272  
 273 273  
 274 274  
 275 275  
 276 276  
 277 277  
 278 278  
 279 279  
 280 280  
 281 281  
 282 282  
 283 283  
 284 284  
 285 285  
 286 286  
 287 287  
 288 288  
 289 289  
 290 290  
 291 291  
 292 292  
 293 293  
 294 294  
 295 295  
 296 296  
 297 297  
 298 298  
 299 299  
 300 300  
 301 301  
 302 302  
 303 303  
 304 304  
 305 305  
 306 306  
 307 307  
 308 308  
 309 309  
 310 310  
 311 311  
 312 312  
 313 313  
 314 314  
 315 315  
 316 316  
 317 317  
 318 318  
 319 319  
 320 320  
 321 321  
 322 322  
 323 323  
 324 324  
 325 325  
 326 326  
 327 327  
 328 328  
 329 329  
 330 330  
 331 331  
 332 332  
 333 333  
 334 334  
 335 335  
 336 336  
 337 337  
 338 338  
 339 339  
 340 340  
 341 341  
 342 342  
 343 343  
 344 344  
 345 345  
 346 346  
 347 347  
 348 348  
 349 349  
 350 350  
 351 351  
 352 352  
 353 353  
 354 354  
 355 355  
 356 356  
 357 357  
 358 358  
 359 359  
 360 360  
 361 361  
 362 362  
 363 363  
 364 364  
 365 365  
 366 366  
 367 367  
 368 368  
 369 369  
 370 370  
 371 371  
 372 372  
 373 373  
 374 374  
 375 375  
 376 376  
 377 377  
 378 378  
 379 379  
 380 380  
 381 381  
 382 382  
 383 383  
 384 384  
 385 385  
 386 386  
 387 387  
 388 388  
 389 389  
 390 390  
 391 391  
 392 392  
 393 393  
 394 394  
 395 395  
 396 396  
 397 397  
 398 398  
 399 399  
 400 400  
 401 401  
 402 402  
 403 403  
 404 404  
 405 405  
 406 406  
 407 407  
 408 408  
 409 409  
 410 410  
 411 411  
 412 412  
 413 413  
 414 414  
 415 415  
 416 416  
 417 417  
 418 418  
 419 419  
 420 420  
 421 421  
 422 422  
 423 423  
 424 424  
 425 425  
 426 426  
 427 427  
 428 428  
 429 429  
 430 430  
 431 431  
 432 432  
 433 433  
 434 434  
 435 435  
 436 436  
 437 437  
 438 438  
 439 439  
 440 440  
 441 441  
 442 442  
 443 443  
 444 444  
 445 445  
 446 446  
 447 447  
 448 448  
 449 449  
 450 450  
 451 451  
 452 452  
 453 453  
 454 454  
 455 455  
 456 456  
 457 457  
 458 458  
 459 459  
 460 460  
 461 461  
 462 462  
 463 463  
 464 464  
 465 465  
 466 466  
 467 467  
 468 468  
 469 469  
 470 470  
 471 471  
 472 472  
 473 473  
 474 474  
 475 475  
 476 476  
 477 477  
 478 478  
 479 479  
 480 480  
 481 481  
 482 482  
 483 483  
 484 484  
 485 485  
 486 486  
 487 487  
 488 488  
 489 489  
 490 490  
 491 491  
 492 492  
 493 493  
 494 494  
 495 495  
 496 496  
 497 497  
 498 498  
 499 499  
 500 500  
 501 501  
 502 502  
 503 503  
 504 504  
 505 505  
 506 506  
 507 507  
 508 508  
 509 509  
 510 510  
 511 511  
 512 512  
 513 513  
 514 514  
 515 515  
 516 516  
 517 517  
 518 518  
 519 519  
 520 520  
 521 521  
 522 522  
 523 523  
 524 524  
 525 525  
 526 526  
 527 527  
 528 528  
 529 529  
 530 530  
 531 531  
 532 532  
 533 533  
 534 534  
 535 535  
 536 536  
 537 537  
 538 538  
 539 539  
 540 540  
 541 541  
 542 542  
 543 543  
 544 544  
 545 545  
 546 546  
 547 547  
 548 548  
 549 549  
 550 550  
 551 551  
 552 552  
 553 553  
 554 554  
 555 555  
 556 556  
 557 557  
 558 558  
 559 559  
 560 560  
 561 561  
 562 562  
 563 563  
 564 564  
 565 565  
 566 566  
 567 567  
 568 568  
 569 569  
 570 570  
 571 571  
 572 572  
 573 573  
 574 574  
 575 575  
 576 576  
 577 577  
 578 578  
 579 579  
 580 580  
 581 581  
 582 582  
 583 583  
 584 584  
 585 585  
 586 586  
 587 587  
 588 588  
 589 589  
 590 590  
 591 591  
 592 592  
 593 593  
 594 594  
 595 595  
 596 596  
 597 597  
 598 598  
 599 599  
 600 600  
 601 601  
 602 602  
 603 603  
 604 604  
 605 605  
 606 606  
 607 607  
 608 608  
 609 609  
 610 610  
 611 611  
 612 612  
 613 613  
 614 614  
 615 615  
 616 616  
 617 617  
 618 618  
 619 619  
 620 620  
 621 621  
 622 622  
 623 623  
 624 624  
 625 625  
 626 626  
 627 627  
 628 628  
 629 629  
 630 630  
 631 631  
 632 632  
 633 633  
 634 634  
 635 635  
 636 636  
 637 637  
 638 638  
 639 639  
 640 640  
 641 641  
 642 642  
 643 643  
 644 644  
 645 645  
 646 646  
 647 647  
 648 648  
 649 649  
 650 650  
 651 651  
 652 652  
 653 653  
 654 654  
 655 655  
 656 656  
 657 657  
 658 658  
 659 659  
 660 660  
 661 661  
 662 662  
 663 663  
 664 664  
 665 665  
 666 666  
 667 667  
 668 668  
 669 669  
 670 670  
 671 671  
 672 672  
 673 673  
 674 674  
 675 675  
 676 676  
 677 677  
 678 678  
 679 679  
 680 680  
 681 681  
 682 682  
 683 683  
 684 684  
 685 685  
 686 686  
 687 687  
 688 688  
 689 689  
 690 690  
 691 691  
 692 692  
 693 693  
 694 694  
 695 695  
 696 696  
 697 697  
 698 698  
 699 699  
 700 700  
 701 701  
 702 702  
 703 703  
 704 704  
 705 705  
 706 706  
 707 707  
 708 708  
 709 709  
 710 710  
 711 711  
 712 712  
 713 713  
 714 714  
 715 715  
 716 716  
 717 717  
 718 718  
 719 719  
 720 720  
 721 721  
 722 722  
 723 723  
 724 724  
 725 725  
 726 726  
 727 727  
 728 728  
 729 729  
 730 730  
 731 731  
 732 732  
 733 733  
 734 734  
 735 735  
 736 736  
 737 737  
 738 738  
 739 739  
 740 740  
 741 741  
 742 742  
 743 743  
 744 744  
 745 745  
 746 746  
 747 747  
 748 748  
 749 749  
 750 750  
 751 751  
 752 752  
 753 753  
 754 754  
 755 755  
 756 756  
 757 757  
 758 758  
 759 759  
 760 760  
 761 761  
 762 762  
 763 763  
 764 764  
 765 765  
 766 766  
 767 767  
 768 768  
 769 769  
 770 770  
 771 771  
 772 772  
 773 773  
 774 774  
 775 775  
 776 776  
 777 777  
 778 778  
 779 779  
 780 780  
 781 781  
 782 782  
 783 783  
 784 784  
 785 785  
 786 786  
 787 787  
 788 788  
 789 789  
 790 790  
 791 791  
 792 792  
 793 793  
 794 794  
 795 795  
 796 796  
 797 797  
 798 798  
 799 799  
 800 800  
 801 801  
 802 802  
 803 803  
 804 804  
 805 805  
 806 806  
 807 807  
 808 808  
 809 809  
 810 810  
 811 811  
 812 812  
 813 813  
 814 814  
 815 815  
 816 816  
 817 817  
 818 818  
 819 819  
 820 820  
 821 821  
 822 822  
 823 823  
 824 824  
 825 825  
 826 826  
 827 827  
 828 828  
 829 829  
 830 830  
 831 831  
 832 832  
 833 833  
 834 834  
 835 835  
 836 836  
 837 837  
 838 838  
 839 839  
 840 840  
 841 841  
 842 842  
 843 843  
 844 844  
 845 845  
 846 846  
 847 847  
 848 848  
 849 849  
 85
```

My first “contribution”

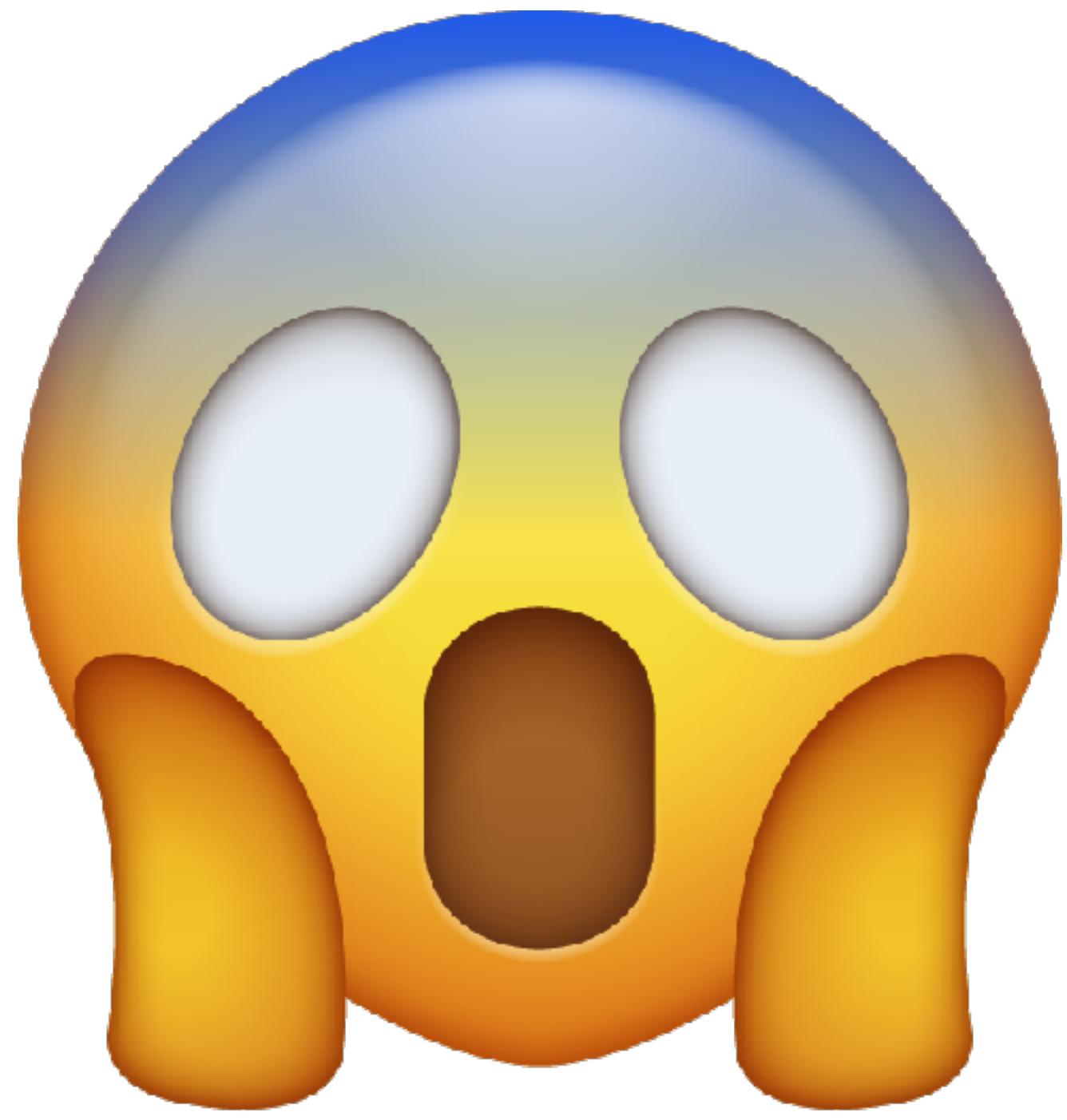


typos





Source code?!



rstudio::conf

18

A screenshot of a web browser window displaying R source code on GitHub. The URL in the address bar is <https://github.com/tidyverse/dplyr/blob/9f1ccccc220c905955a1f333c25401d0r/man/man/summarise.R>. The code is for the `summarise` function in the dplyr package. A specific line of code, which highlights the ability of data frames to create variables in summaries, is highlighted in yellow:

```
111 #' `summarise()` is typically used on grouped data created by [group_by()].
112 #' The output will have one row for each group.
113 #
114 #' @section Useful functions:
115 #
116 #' * Center: [mean()], [median()]
117 #' * Spread: [sd()], [IQR()], [mad()]
118 #' * Range: [min()], [max()], [quantile()]
119 #' * Position: [first()], [last()], [nth()],
120 #' * Count: [n()], [n_distinct()]
121 #' * Logical: [any()], [all()]
122 #
123 #' @section Backend variations:
124 #
125 #' Data frames are the only backend that supports creating a variable and
126 #' using it in the same summary. See examples for more details.
127 #
128 #' @export
129 #' @inheritParams filter
130 #' @inheritSection filter Tidy data
131 #' @param ... Name-value pairs of summary functions. The name will be the
132 #' name of the variable in the result. The value should be an expression
133 #' that returns a single value like `min(x)`, `n()`, or `sum(is.na(y))`.
134 #
135 #' These arguments are automatically [quoted][rlang::quo] and
136 #' [evaluated][rlang::eval_tidy] in the context of the data
137 #' frame. They support [unquoting][rlang::quasiquotation] and
138 #' splicing. See `vignette("programming")` for an introduction to
139 #' these concepts.
```

Source code...

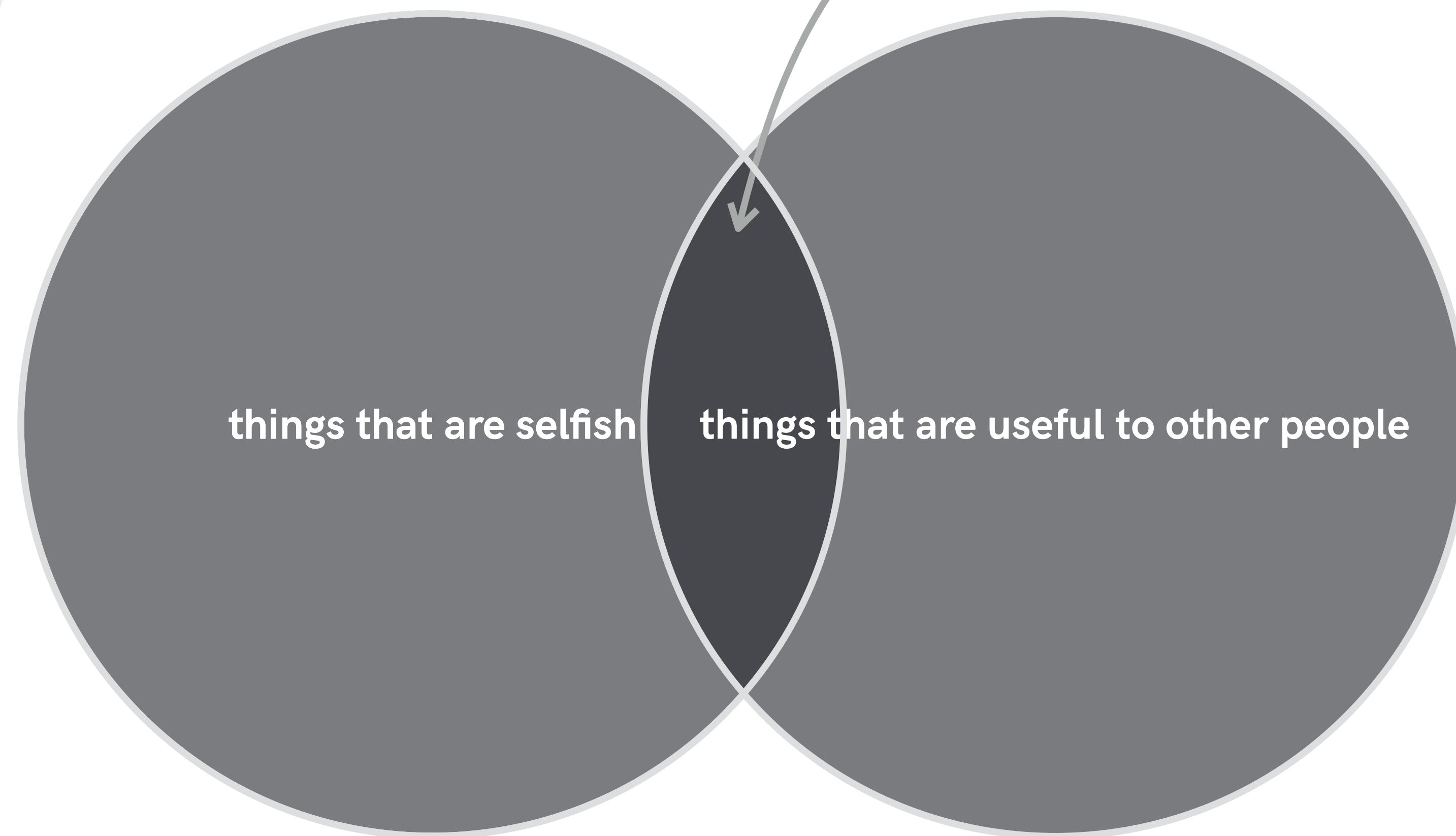


Contribute documentation

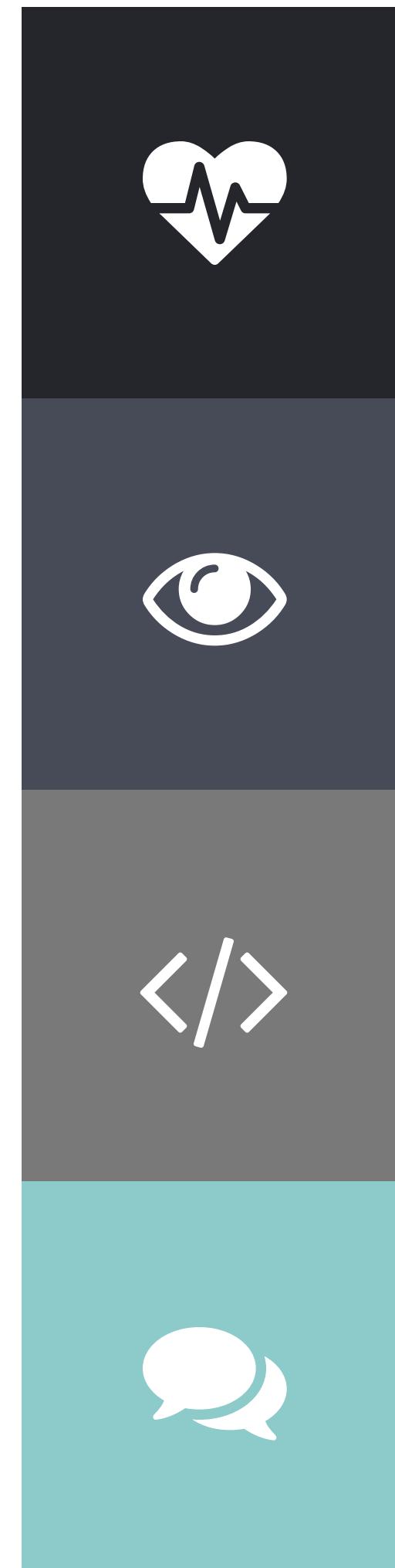
“Innocence lost is not easily regained. The designer simply cannot predict the problems people will have, the misinterpretations that will arise, and the errors that will get made.”

— Donald Norman, *The Design of Everyday Things*

wait, but why?



Hints for happy code contributions in the tidyverse



GET THE PULSE OF A PROJECT

WATCH THE REPO

READ THE CODE

DISCUSS YOUR IDEAS

CONTRIBUTING.md



tidyverse & package conventions

KEY PACKAGES FOR CONTRIBUTING



ROXYGEN2

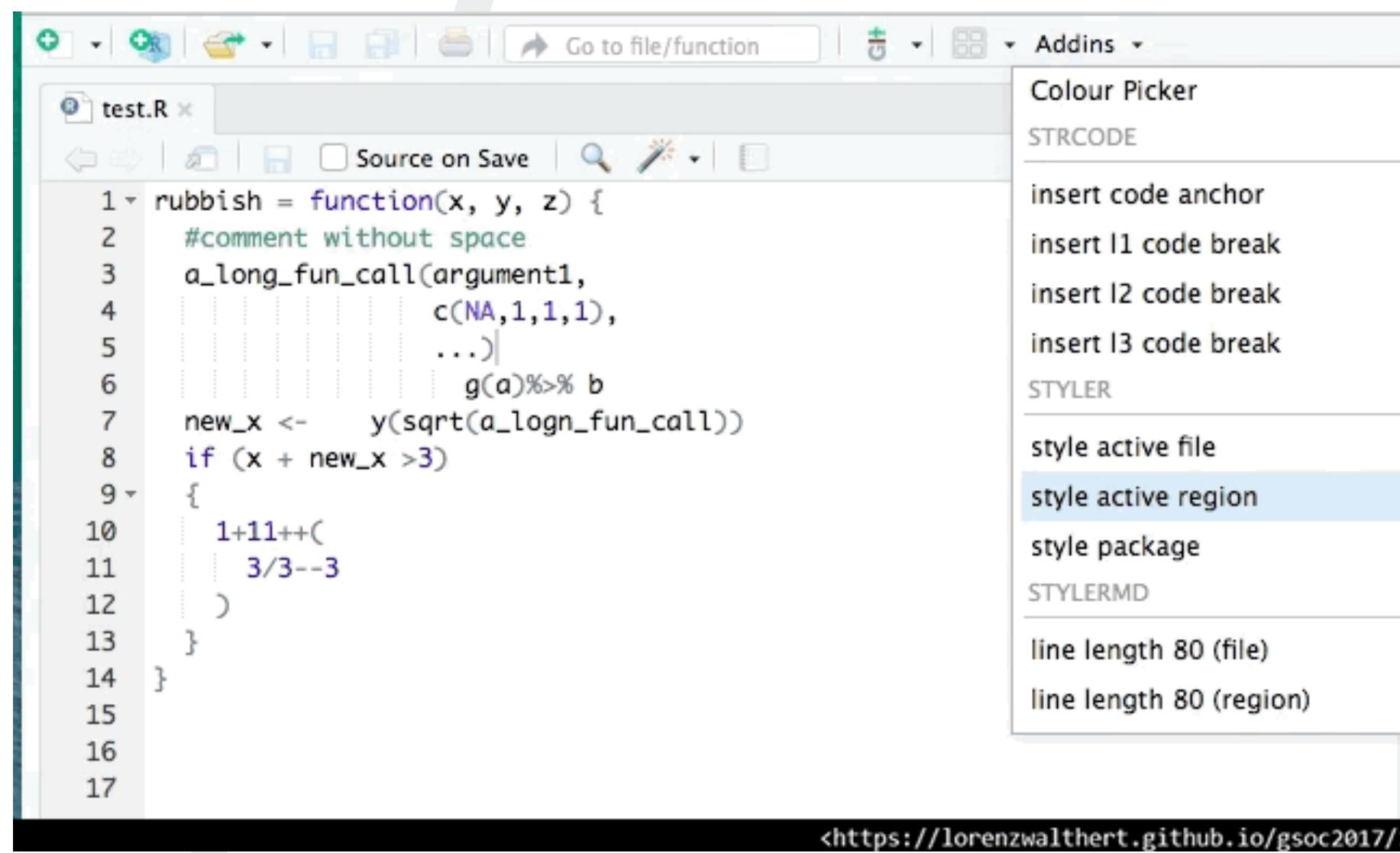


DEVTOOLS



TESTTHAT

stylish...



styler by Kirill Müller & Lorenz Walthert <http://styler.r-lib.org>

rstudio::conf

The image shows two views of the tidyverse style guide. On the left, a screenshot of a web browser displays the title "The tidyverse style guide" and author "Hadley Wickham". The main content is a "Welcome" section with text about good coding style and consistency. On the right, a screenshot of RStudio shows the "Addins" menu with the "STYLER" option selected. Below the menu, a search bar contains the word "style". The "Style selection" option in the STYLER submenu is highlighted with a blue selection bar.

The tidyverse style guide
Hadley Wickham

Welcome

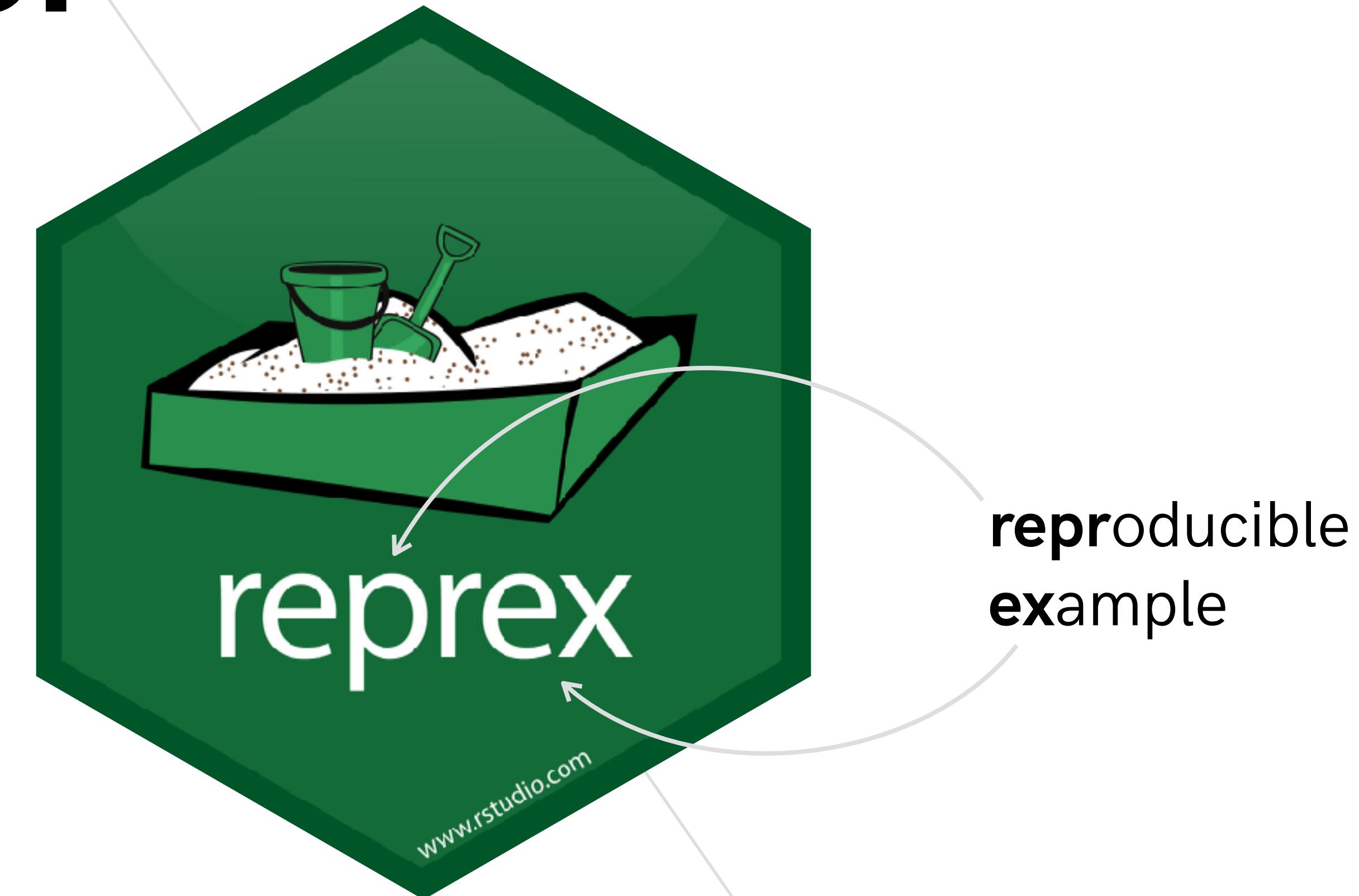
Good coding style is like correct punctuation: you can manage without it, but it sure makes things easier to read. This site describes the style used throughout the [tidyverse](#). It was originally derived from [Google's R style guide](#), but has evolved and expanded considerably over the years.

All style guides are fundamentally opinionated. Some decisions genuinely do make code easier to use (especially matching indenting to programming structure), but many decisions are arbitrary. The most important thing about a style guide is that it provides consistency, making code easier to write because you need to make fewer decisions.

Two R packages support this style guide:

- [styler](#) allows you to interactively restyle selected text, files, or entire projects. It includes an RStudio add-in, the easiest way to re-style existing code.
- [lintr](#) performs automated checks to confirm that your conform to the style guide.

the magic of reprex



Keys to reprex-cellence

- ✓ Code that **actually runs**
- ✓ Code that **doesn't have to be run**
- ✓ Code that **can be easily run**

Nailing those reprexes?

- Help others ask questions.
- Answer questions.
- Make *more* thoughtful users

contributing is
a relationship,
not a
transaction

29



baby steps build up trust

...and make life easier



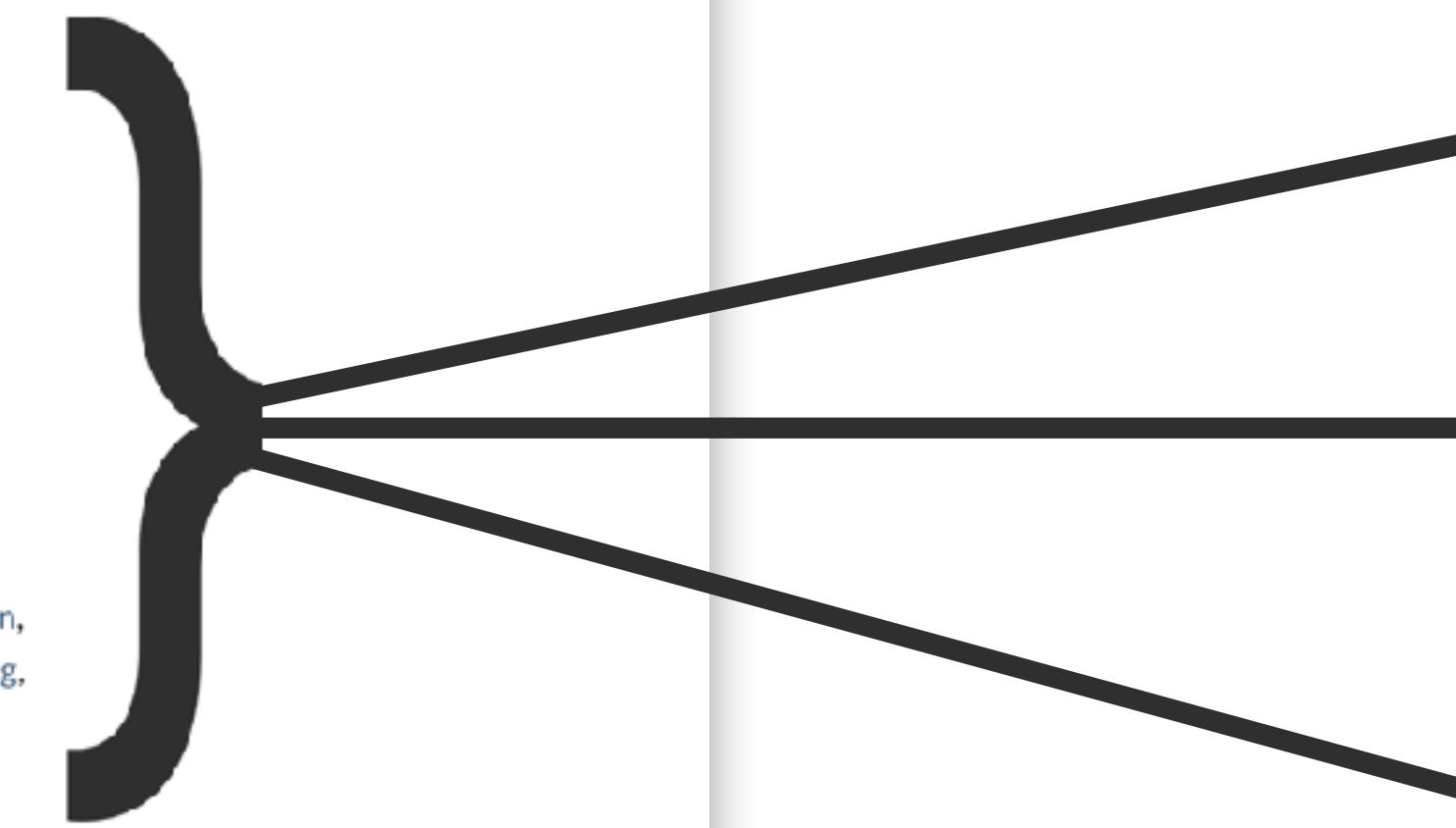
every contribution counts...

The screenshot shows the Tidyverse website's 'Acknowledgments' page. At the top, there's a navigation bar with links for 'Tidyverse', 'Packages', 'Articles' (which is highlighted), 'Learn', 'Help', and 'Contribute'. Below the navigation, the title 'Acknowledgments' is displayed. A large block of text lists numerous contributors, many of whom are linked to their GitHub profiles. At the bottom of the page, it says 'The tidyverse is proudly supported by R Studio' and features icons for GitHub and Twitter.

Acknowledgments

We received issues, pull requests, and comments from 109 people since tibble 1.2.0. Thanks to everyone: @alexandersson, @adnbps, @AkhilNairAmey, @alexhallam, @alibat, @amjuzi, @AndreMikulec, @andrewjpfeiffer, @ashiklom, @atribe, @bapfeld, @barnettjacob, @behrman, @BillDunlap, @BruceZhaoR, @cassiusoat, @cboettig, @cderv, @ckluss, @ClaytonJY, @colearendt, @csgillespie, @dalejbarr, @dan87134, @DavisVaughan, @ddiez, @dhicks, @dldpd, @drewgendaleau, @drolejoel, @echasnovski, @edzer, @ElsLommelen, @etiennebr, @FabianRoger, @garrettgman, @gavinsimpson, @geotheory, @ginolhac, @hadley, @happyshows, @heavywatal, @helix123, @holstius, @huftis, @ianmcook, @manuelcostigan, @janschulz, @javierluraschi, @jennybc, @jimhester, @joelgombin, @jonathan.g, @kendonB, @kevinushey, @khughitt, @kismsu, @krlmlr, @kwstat, @LaDilettante, @lcolladotor, @lionel-, @lpmarco, @m-sostero, @MarcusWalz, @matteodefelice, @mattfidler, @mgirlich, @michaellevy, @MikeBadescu, @mkearney, @mmuurr, @Monduiz, @mubeenarasack, @mundl, @nbenn, @ncarchedi, @NikNakk, @noamross, @ntguardian, @p0bs, @patperry, @pgensler, @phalexo, @pssguy, @r2evans, @rentrop, @richierocks, @Rongpeng, @s-fleck, @sainathadapa, @sebschub, @sibojan, @slonik-az, @sskim47, @t-kalinowski, @thercast, @thornend, @tjmahr, @trinker, @vnijjs, @vspinu, @vvrably, @wibom, @wpetry, @yedle, @yihui, @yutannihilation, and @Zedseyou.

The tidyverse is proudly supported by R Studio



PULL REQUESTS

ISSUES

COMMENTS

Thank You

<http://bit.ly/tverse-contrib>



resources

- Happy Git and GitHub for the useR by Jenny Bryan.
- Contribute to the tidyverse (incl. ways to contribute that *don't* involve writing code).
- Contributing Code to the Tidyverse by Jim Hester.
- R packages by Hadley Wickham.
 - Git and GitHub
 - Automated checking
 - Object documentation
 - Testing
- dplyr's NEWS.md is a good source of examples for both content and styling.
- Closing issues using keywords on GitHub.
- Autolinked references and URLs on GitHub.
- GitHub Guides: Forking Projects.