For the Longest Time: Continuity and Change in One Teaching-Related Subreddit

K. Bret Staudt Willet & Jeffrey P. Carpenter
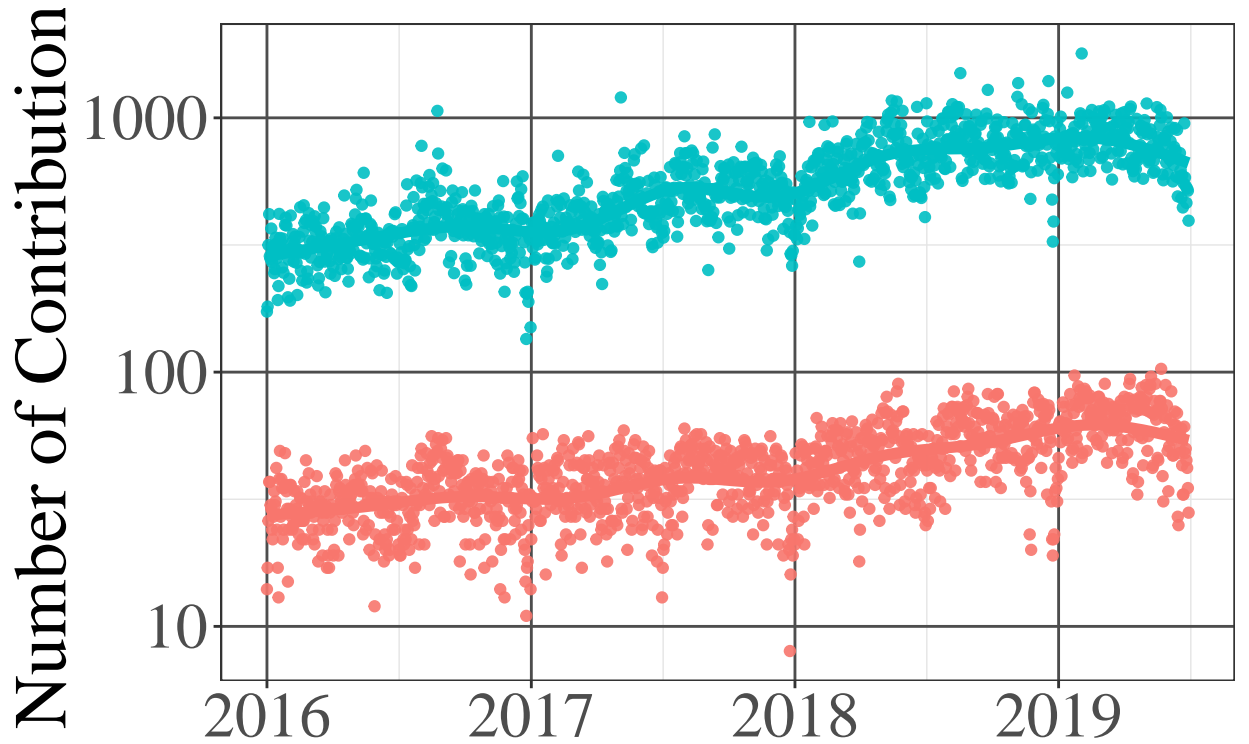
12/12/2019

Contents

Get set up

This section loads the data and packages and starts to process the data, but doesn't calculate any statistics or create any results.

1. Load packages

2. Set up settings for sharing

3. Set timeframe

4. Load subreddit posts

5. Load subreddit responses

6. Create a merged dataframe of both posts and responses.

Analysis and Results

Figure 1. Contributions over time

```
## [1] "In total, we collected 696660 contributions from 55146 contributors"
```

```
## [1] "dated between 2016-01-01 and 2019-06-30 (41.98 months),"
```

```
## [1] "a total of 54933 posts and 641727 responses to those posts,"
```

```
## [1] "to one teaching-related subreddit: r/Teachers."
```

```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

Type of Contribution: ━ post ━ respo

```
## [1] "The slope of the `post` linear regression is 0.03 (p=0),"

## [1] "and the slope of the `response` linear regression is 0.49 (p=0)."
```

Table 1. Overall contributions in r/Teachers

Year	Contributors	Posters	Responders	Multiple Posts	Multiple Responses	Posters with Self Responses
2016	14736	6715	12613	81.53	63.58	56.31
2017	18022	7971	15704	84.19	63.42	59.19
2018	25997	10411	23100	82.25	61.75	58.34
2019	18368	7062	16241	78.58	61.92	55.49

Table 2. Descriptive Statistics of Posts per Poster and Responses per Responder

Year	Mean	SD	Median	Min	Max	Mean	SD	Median	Min	Max
2016	17.98	53.75	7	1	2169	9.44	39.67	2	1	2186
2017	20.08	64.71	7	1	3811	10.10	50.21	2	1	3480
2018	22.83	91.93	7	1	6581	10.19	48.50	2	1	3033
2019	18.49	62.33	6	1	2567	7.93	33.40	2	1	1906

Table 3. Content Interaction by Posts, Threads, and Responses

Year	Posts	Threads	Responses	Response Rate	Responses per Thread	Words per Post	Words per Response
2016	11413	9761	119091	85.53	12.20	159.44	63.34
2017	13564	12066	158591	88.96	13.14	166.85	62.41
2018	18473	16055	235279	86.91	14.65	156.38	56.56
2019	11483	9649	128766	84.03	13.35	161.45	54.61

Table 4. Descriptive Statistics of Voting Scores of Posts and Responses

Year	Mean	SD	Median	Min	Max	Mean	SD	Median	Min	Max
2016	9.72	56.97	2	0	5074	3.54	13.32	2	-235	2620
2017	15.22	131.14	3	0	10471	4.66	14.81	2	-238	1954
2018	21.85	137.32	2	0	8383	6.10	26.32	2	-530	4843
2019	29.86	134.61	2	0	3636	7.01	23.96	2	-156	1264

Table 5. Social Interaction: Network Statistics by Year

Year	Nodes	Edges	Transitivity	Reciprocity	Mean	SD	Median	Min	Max
2016	13934	117765	2.65	0.52	8.59	30.48	2	0	1368
2017	17332	156861	2.53	0.64	8.65	33.38	2	0	1819
2018	24878	231474	2.43	0.82	8.86	37.81	2	0	2247
2019	17496	126277	1.84	0.38	6.99	26.10	2	0	1352

Table 6. High-Mutuality Threads

```
## [1] "There are 91381 self-responses (14.08%) written by original posters."
```

We theorized *high-mutuality threads* as those posts with at least **10 responses** and where the original poster contributed not just a single response but also did not overly dominate the conversation (i.e., **20-50% self-responses**). Table 6 depicts a summary of these high-mutuality threads in terms of threads with self-responses.

Year	Threads n	Threads p	Mean n	SD n	Median n	min n	max n	Mean p	SD p	Median p	min
2016	1407	12	19.27434	13.11634	15	10	129	32.76006	8.582689	31.58	20
2017	1880	14	20.98617	19.53062	15	10	298	32.73147	8.638257	31.58	20
2018	2461	13	19.99147	16.89370	15	10	327	32.57887	8.588061	31.25	20
2019	1235	11	20.37733	21.23924	15	10	369	33.46908	8.759741	33.33	20

Sample posts for qualitative content analysis

Finally, we took a sample of 100 posts from the *high-mutuality threads* for further examination.

Version/dependencies

```
sessionInfo()
```

```
## R version 3.5.2 (2018-12-20)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS Mojave 10.14.6
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] ggraph_1.0.2    igraph_1.2.4.1  lubridate_1.7.4 anytime_0.3.3
##  [5] forcats_0.4.0   stringr_1.4.0   dplyr_0.8.0.1   purrr_0.3.2
##  [9] readr_1.3.1     tidyr_0.8.3     tibble_2.1.1    ggplot2_3.1.1
## [13] tidyverse_1.2.1 knitr_1.22
##
## loaded via a namespace (and not attached):
##  [1] ggrepel_0.8.0      Rcpp_1.0.1         lattice_0.20-38
##  [4] clisymbols_1.2.0   assertthat_0.2.1   rprojroot_1.3-2
##  [7] digest_0.6.18      ggforce_0.2.2      R6_2.4.0
## [10] cellranger_1.1.0   plyr_1.8.4         backports_1.1.4
## [13] RApiDatetime_0.0.4 evaluate_0.13      highr_0.8
## [16] httr_1.4.0         pillar_1.3.1       rlang_0.4.2
## [19] lazyeval_0.2.2     readxl_1.3.1       rstudioapi_0.10
## [22] Matrix_1.2-17      rmarkdown_1.12     splines_3.5.2
## [25] polyclip_1.10-0    munsell_0.5.0      broom_0.5.2
## [28] compiler_3.5.2     modelr_0.1.4       xfun_0.6
## [31] pkgconfig_2.0.2    mgcv_1.8-28        htmltools_0.4.0
## [34] tidyselect_0.2.5   gridExtra_2.3      viridisLite_0.3.0
## [37] crayon_1.3.4       withr_2.1.2        MASS_7.3-51.4
## [40] grid_3.5.2         nlme_3.1-139       jsonlite_1.6
## [43] gtable_0.3.0       magrittr_1.5       scales_1.0.0
## [46] cli_1.1.0          stringi_1.4.3      farver_1.1.0
## [49] viridis_0.5.1      fs_1.2.7           xml2_1.2.0
## [52] generics_0.0.2     tools_3.5.2        glue_1.3.1
## [55] tweenr_1.0.1       hms_0.4.2          yaml_2.2.0
## [58] colorspace_1.4-1   rvest_0.3.3        haven_2.1.0
## [61] usethis_1.5.0
```