

Title: Monthly Summary of Selected Trends, Activities and Insights for R - JULY

Abstract: R is a leading language in the data science domain. In the following article, a summary of selected trends, activities and insights around the R language in July, 2018 are presented.

Data for the trends and activities summarized here were obtained from popular websites used by the R community such as Google, GitHub, StackOverflow, Rstudio, METACRAN and R-Bloggers

1. StackOverflow

Number of StackOverflow Questions tagged R: 4,929 (1.09% decrease from June)

Number of Answers for R questions: 4,497 (1.09% decrease from June)

Number of Comments for R questions: 9,589 (1.11% decrease from June)

Page Views for R questions: 161,300 (1.10% decrease from June)

Source: <https://data.stackexchange.com>

The pie chart below in Figure 1 shows distribution of the questions, answers and comments for R at StackOverflow

Figure 1

2. R/Packages Downloads

The data for this section is obtained through the API of the METACRAN's (<https://r-pkg.org/services#cranlogs>) service that could be found here:

<https://github.com/metacran/cranlogs.app>

METACRAN obtains the download summaries from the RStudio (<https://www.rstudio.com/>) CRAN (<https://cran.r-project.org/>) mirror taken from <http://cran-logs.rstudio.com/> This is one very popular download mirror for the R language due to the popularity of the RStudio IDE for R.

Base R downloads: 56,238 (1.06% increase from June)

The chart in Figure 2 shows the distribution of base R downloads among the computer operating systems from which R was downloaded.

Figure 2

The chart in Figure 3 shows the number of downloads by R versions. From the chart, it is clear that there are users of R that still make use of older R versions e.g. version 2.x.x series.

Figure 3

R Packages Downloads: 42,285,633 (1.06% decrease from June)

Figure 4 is a chart that shows the daily download variation of R packages in July.

Figure 4

Ratio of Base R downloads to R Package downloads: 1 : 752

Thus, for each download of base R, there are over 752 extension packages downloaded. The use of R still depends largely on extension packages.

Figure 5 is a chart visualizing the ratio of Base R downloads to R packages downloads

Figure 5

Top Packages (July, 2018):

Sum of Downloads for Top 50 packages: 15,301,517 (1.11% decrease from June)

Download contribution of Top 50 packages amongst other 12,000+ CRAN packages: 36.2% (i.e, over 36% of total R package downloads came from Top 50 packages in July)

Figure 6

Below is a chart in Figure 7 showing the top 50 downloaded packages based on their download counts.

Figure 7

Sum of Downloads for Top 100 packages: 23,006,206

Download contribution of Top 100 amongst other 12,000+ CRAN packages: 54.4% (i.e, over 54% of R package downloads came from Top 100)

Figure 8

Top 10 Packages and their Primary Maintainers:

	Package	Authors
1	Rcpp	Dirk Eddelbuettel
2	stringi	Marek Gagolewski
3	ggplot2	Hadley Wickham
4	rlang	Lionel Henry
5	stringr	Hadley Wickham
6	glue	Jim Hester
7	dplyr	Hadley Wickham
8	pillar	Kirill Müller

9	digest	Dirk Eddelbuettel
10	tibble	Kirill Müller

Hadley Wickham, Dirk Eddelbuettel and Kirill Müller dominate the top 10 by both number of packages and downloads from the list above.

3. GitHub

The Top 10 R repositories that appeared on GitHub's trends in July are:

1. gganimate (<https://github.com/thomasp85/gganimate>) by Thomas Lin Pedersen
2. ggplot2 (<https://github.com/tidyverse/ggplot2>) by Hadley Wickham
3. shiny (<https://github.com/rstudio/shiny>) by Winston Chang
4. datapasta (<https://github.com/MilesMcBain/datapasta>) by Miles McBain
5. dplyr (<https://github.com/tidyverse/dplyr>) by Hadley Wickham
6. infer (<https://github.com/tidymodels/infer>) by Andrew Bray
7. drake (<https://github.com/ropensci/drake>) by William Michael Landau
8. polite (<https://github.com/dmi3kno/polite>) by Dmytro Perepolkin
9. rmarkdown (<https://github.com/rstudio/rmarkdown>) by JJ Allaire
10. dataviz (<https://github.com/clauswilke/dataviz>) by Claus Wilke

The chart of Figure 9 shows the top 10 projects that did trend in July based on data from <https://github.com/trending/r?since=monthly> on August 1, 2018,

Figure 9

4. R Meetings/Events

The Rstudio Community website provides a weekly list (<https://community.rstudio.com/c/irl>) of R user-group meetings and conference events curated from meetup.com and elsewhere. The data found on this website is the basis for the following analysis.

There were 76 events in about 23 countries (1.2% decrease in events from June)
39 events out of 76 (51% of events) were held in the United States of America

	Country	Number of events
1	Australia	2
2	Brazil	1
3	Canada	1
4	côte d'Ivoire	1
5	Ecuador	1
6	Ethiopia	1
7	France	1
8	Georgia	1
9	Germany	6
10	Hong Kong	1
11	India	1
12	Italy	1
13	Morocco	1

14	Netherlands	1
15	New Zealand	2
16	Nigeria	3
17	Philippines	1
18	Poland	1
19	South Africa	1
20	Switzerland	4
21	Taiwan	2
22	United Kingdom	3
23	United States	39

A world map showing the distribution of R events across 23 countries can be found in Figure 10a

Figure 10a

Figure 10b shows the event distribution across the 23 countries in a pie (area) chart.

Figure 10b

A weekly summary of events compared with countries for the month is shown in the bar chart of Figure 11.

Figure 11

R-Ladies

R-Ladies have a strong grip of the share of events in the R community. 34 events out of 76 (45% of events) were R-Ladies events

useR! 2018

This July, the useR! conference – largest annual conference for the R community – was held in Brisbane, Australia. Full list of YouTube Videos for presentations during this conference has been generously provided by the R Consortium here:

https://www.youtube.com/channel/UC_R5smHVXRYGhZYDJsnXTwg/videos

The most viewed useR! 2018 talks on this channel by August 7, 2018 are:

1. Teaching R to New Users: From tapply to Tidyverse. 2.4k views (<https://www.youtube.com/watch?v=5033jBHFfiHE>)
2. Code Feels and Smells. 1.9k views (<https://www.youtube.com/watch?v=7oyiPBjLAWY>)
3. What is R? 1.5k views (<https://www.youtube.com/watch?v=XcBLEVknqvY>)
4. The Grammar of Animation. 1.4k views (<https://www.youtube.com/watch?v=21ZWDrTukEs>)
5. Recipes of Data Processing .1.1k views (<https://www.youtube.com/watch?v=JapQdj1Vfc>)

5. R Bloggers

Rbloggers.com is the most popular news aggregation website for blog posts related to the R language. There were about 258 blog posts at Rbloggers.com in July. An average of about 8 posts / day.

(1.11% increase from June)

6. Google Trends

The chart below show Google Trends for the R language in July with search trends dipping only during the weekends.

Figure 12

Based on Interest by region, the top 5 countries in July with highest search activity on Google are:

1. Philippines
2. China
3. South Korea
4. Georgia
5. St. Helena

Source: <https://trends.google.com/trends/explore?date=2018-07-01%202018-07-31&q=%2Fm%2F0212jm>

Embed Code to place the chart in an HTML page:

```
<script type="text/javascript"
src="https://ssl.gstatic.com/trends_nrtr/1513_RC01/embed_loader.js"></script>
<script type="text/javascript">
  trends.embed.renderExploreWidget("TIMESERIES", {"comparisonItem":
[{"keyword":"/m/0212jm","geo":"","time":"2018-07-01 2018-07-
31"}], "category":0, "property":""}, {"exploreQuery":"date=2018-07-01%202018-07-31&q=
%2Fm%2F0212jm","guestPath":"https://trends.google.com:443/trends/embed/"}]);
</script>
```

7. Language Ranking:

TIOBE Index (<https://www.tiobe.com/tiobe-index/>): 14th in July, 2018

Redmonk (<https://redmonk.com/sogrady/2018/03/07/language-rankings-1-18/>): 12th in January, 2018

IEEE (<https://spectrum.ieee.org/static/interactive-the-top-programming-languages-2018>): 6th in July, 2018

PyPL (<http://pypl.github.io/PYPL.html>): 7th in July, 2018

8. R Consortium

The impact of R Consortium's user-base expansion program through the R User Group Support (RUGS) program and R-Ladies project can be seen from the number of events in July.

45% of R user-group events in July were R-Ladies events. Nigeria ranked 4th based on number

of user-group events in July, with only the Kano R User Group, an R Consortium funded group meeting three times in July.

9. Google Summer of Code

The second phase of evaluations for the Google Summer of Code program for participating students and mentors took place in July. There are 27 funded projects in this program this year. See full list of R-Google Summer of Code projects 2008 – date here: <http://r-central.com/gsoc>.

Biography

R Central is a developing aggregation website for news, trends and activities around the R language. It provides a dashboard which is suitable as a browser homepage that lists posts from popular sources like CRAN, R-devel, Twitter, Rbloggers, Crantastic, RWeekly, Microsoft, RConsortium, StackOverflow and several others.

Reference Links

Web: <https://r-central.com>

Twitter: @rcentrall