Dwindling software freedom in the R ecosystem

rm

* What R is and why we like it

* Trend away from fReedom

* Wrong arguments about why this is fine

* What we can do

R (and S)

* Programming language

* Statistical computing, graphics

* Popular in genomics, statistics

* Ancestor of the "data frame"
  interface now available
  in many languages

HISTORY

S

* Used in Bell Labs from 1976
* Distributed with source code from 1981 (dunno license)
* Used in many academic settings, including by
  Douglas Martin at the University of Washington.

S-PLUS

* Extension of S, developed by Statistical Sciences, Inc, a company founded by Douglas Martin in the 1980s
* TIBCO now owns S-PLUS and publishes it under a proprietary license.

R

* Developed by Ross and Robert,
  starting in August 1993

* Became a GNU project in 1997

* The most popular implementation of S today

Comprehensive R Archive Network (CRAN)

* Main repository of R extensions

* Created in 1997 by Kurt Hornik and Friedrich Leisch

* Packages generally must choose licenses from a short list.
  Some non-free licenses are tolerated, e.g. Artistic License

Bioconductor

* Extensions for analysis of genomic data

* Started in 2001 at the Dana Farber Cancer Institute

* Same licensing expectations as CRAN

* Core Bioconductor packages are Artistic License 2.0.

Me

* I took a statistics class in 2006.

* Then I wanted to do statistics calculations with free software.

* The options I found were R and PSPP, both GNU projects.

* I chose R, probably because it seemed easier to install.

Tidyverse

* Set of freely licensed R packages
  with a common theory of data structure

* Originally the Hadleyverse, developed
  by Hadley Wickham starting in 2000s

* Much development is now done by Posit.

* Hadley Wickham now works at Posit.

RStudio (AGPL-licensed R IDE)

* RStudio Inc. released RStudio in 2011

* RStudio Inc. renames to RStudio PBC in 2020

* RStudio PBC renames to Posit PBC in 2022

Other Posit products

* Often proprietary

* More on this later

References

* https://en.wikipedia.org/wiki/S_(programming_language)

* https://en.wikipedia.org/wiki/S-PLUS

* https://en.wikipedia.org/wiki/R_(programming_language)

* https://homes.cs.washington.edu/~lazowska/impact/statsci.html

* https://posit.co/blog/rstudio-pbc/

* https://en.wikipedia.org/wiki/RStudio

* https://www.r-project.org/Licenses/

* https://svn.r-project.org/R/trunk/share/licenses/license.db

Why we like R and PSPP: Four freedoms

GNU project:
a fully free
operating system.

Distribution	Brief Description
	Dragora GNU/Linux-Libre, an independent GNU/Linux distribution based on concepts of simplicity.
	Dyne:bolic, a GNU/Linux distribution with special emphasis on audio and video editing. This is a "static" distro, normally run from a live CD. Since it will not receive security updates, it should be used offline.
guix	*Guix System*, an advanced GNU/Linux distro built on top of GNU Guix (pronounced "geeks"), a purely functional package manager for the GNU system.
	Hyperbola GNU/Linux-libre, a long-term support simplicity-focused distribution based on Arch GNU/Linux.
	Parabola GNU/Linux-libre, a distribution based on Arch that prioritizes simple package and system management.
PureOS	*PureOS*, a GNU/Linux distribution based on Debian with a focus on privacy, security, and convenience.
	Trisquel, a GNU/Linux distribution based on Ubuntu that's oriented toward small enterprises, domestic users and educational centers.
	Ututo S, a GNU/Linux 100% free distribution. It was the first fully free GNU/Linux system recognized by the GNU Project.

Trends towards unethical practices

Great: R itself remains free
(GNU General Public License version 2 or greater),
a member of the GNU project, and
well-supported on free operating systems.

Good: tidyverse packages remain free, typically under MIT license.

Good: RStudio itself remains free
(GNU Affero General Public License).

Good: Popular packages sometimes work only on non-free operating systems or use interact with proprietary software, but the R packages themselves generally are free.

Good: Popular packages sometimes work only in RStudio, but RStudio is free.

Not good: RStudio can become proprietary.

Despite the AGPLv3 license,

Posit has the right to make proprietary

RStudio forks, from the contributor

agreement that contributors must sign.

Not good: Contributors to many projects (e.g., RStudio and tidyverse) must contribute by an unethical repository service, GitHub.

Good: R uses an ethical approach, svn and patches by email.

Site	Grade	Date	Criteria Version
savannah.gnu.org	A	2015-10-01	1.1
sr.ht	B	2021-03-05	1.1
notabug.org	C	2021-03-23	1.1
gitlab.com	F	2021-05-05	1.1
github.com	F	2021-04-02	1.1
sourceforge.net	F	2015-10-07	1.1

How GitHub is unethical:

Service as software substitute (SaaS)

Case study: repexp and usethis

1. I used a free package by RStudio Inc without RStudio IDE.
2. It didn't work.
3. I traced the bug to usethis::proj_activate.
4. The usethis tests did not work on my free operating system.
5. I patched the function anyway.
6. While I could not run the tests (reprex), I wrote an untested test.

7. I sent my contribution by email
   to the listed maintainers.
8. They informed me that contributions are
   through GitHub pull requests and recommended
   I create an issue.

8. Valuing my freedom, I requested a less
   principled friend to post an issue
   with my patch.

9. It got no response until I happened to meet
   one of the usethis developers in person.

10. He made a better patch.
11. He explained that the particular issue demonstrated the importance of writing reprexes.

Summary of bug

* Package is free.

* RStudio is free.

* Problem: Package works only in RStudio

Disregard for freedom

* Cannot report issue

* Cannot write test: the testing framework is free but in practice needs a proprietary operating systems

* Cannot publish patch

My opinion on cause

* Mostly not malice

* People accept slavery out of convenience

Severity categories

* Great

* Good

* Not Good

Severity categories

* Great

* Good

* Not Good

* Bad

Bad: Clearly unethical practices

What's popular in R today?

Let us look at the Posit blog.

posit   PRODUCTS ⌄   SOLUTIONS ⌄   LEARN & SUPPORT ⌄   EXPLORE MORE ⌄   PRICING   🔍   **DOWNLOAD RSTUDIO**

RESOURCES

The Posit Blog

The Posit Blog is all things data science and the world that benefits from our products, community, and events. Check back regularly for new content, and share the things that really make an impact on you!

773+
POSTS

19+
CATEGORIES

310+
TAGS

Search 🔍   Filter by

COMPANY NEWS AND EVENTS   COMPANY NEWS AND EVENTS   PRODUCTS AND TECHNOLOGY

https://posit.co/blog/

posit   PRODUCTS ⌄   SOLUTIONS ⌄   LEARN & SUPPORT ⌄   EXPLORE MORE ⌄   PRICING

DOWNLOAD RSTUDIO

773+
POSTS

19+
CATEGORIES

310+
TAGS

Search

Subscribe to more inspiring open-source data science content.

We want to help you go further with open-source data science. By subscribing, you'll get alerted whenever we publish something new.

https://posit.co/blog/

30%

posit PRODUCTS SOLUTIONS LEARN & SUPPORT EXPLORE MORE PRICING DOWNLOAD RSTUDIO

RESOURCES

The Posit Blog

The Posit Blog is all things data science and the world that benefits
from our products, community, and events. Check back regularly for
new content, and share the things that really make an impact on you!

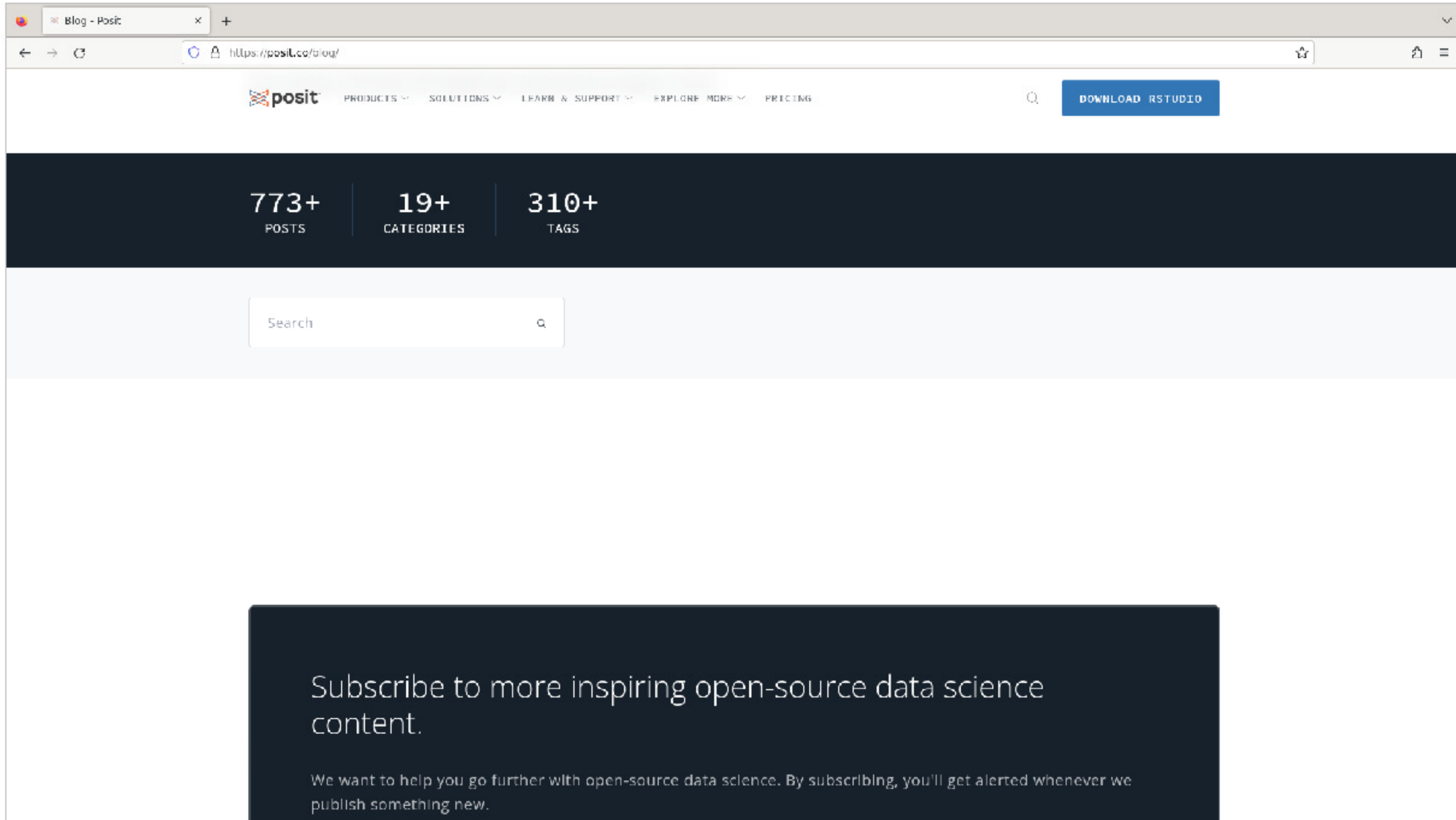**773+**
POSTS

19+
CATEGORIES

310+
TAGS

Search

Subscribe to more inspiring open-source data science
content.

PRODUCTS SOLUTIONS LEARN & SUPPORT EXPLORE MORE PRICING

DOWNLOAD RSTUDIO
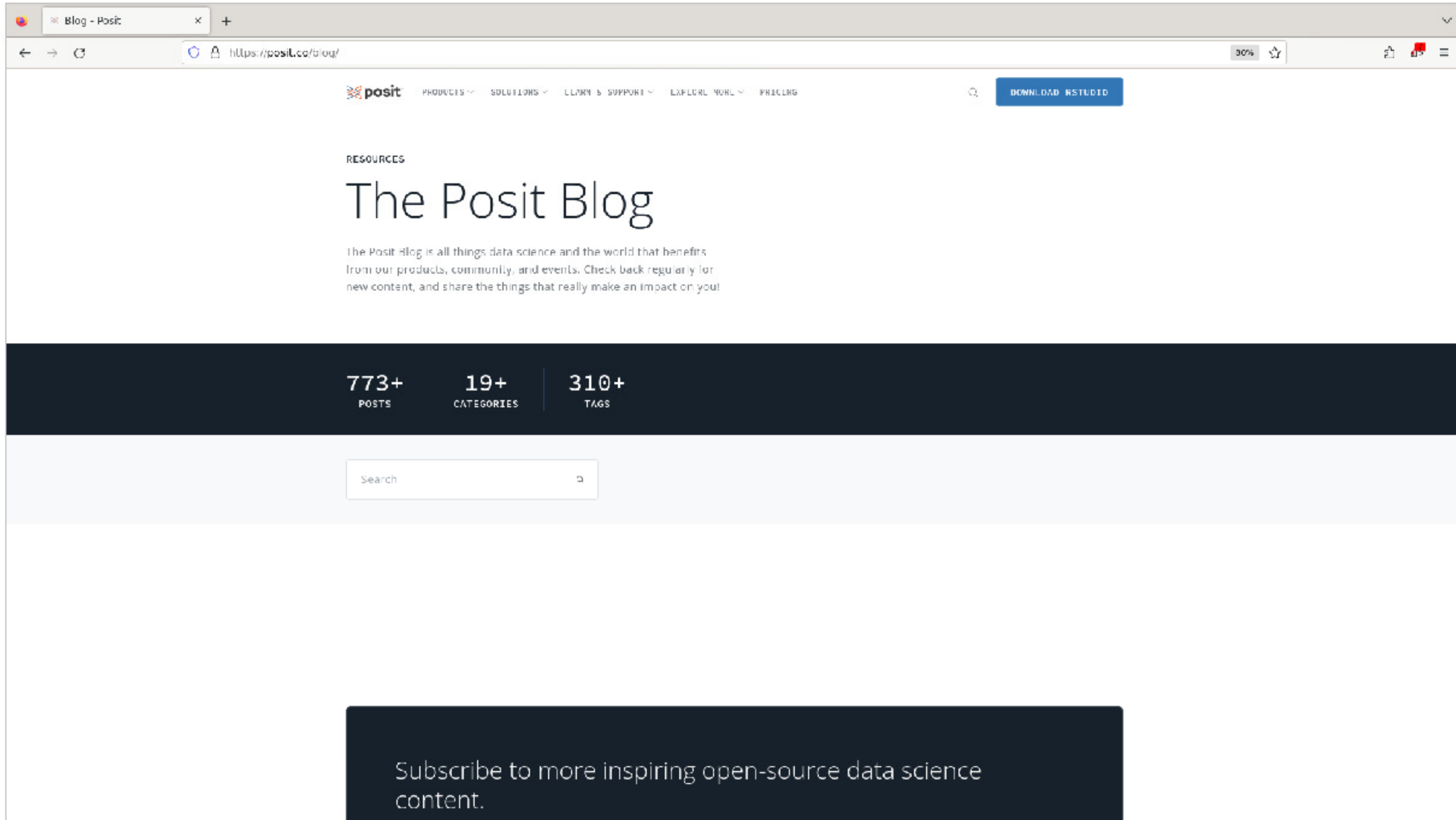
Subscribe to more inspiring open-source data science content.

We want to help you go further with open-source data science. By subscribing, you'll get alerted whenever we publish something new.

SUBSCRIBE

posit

The open source data science company

PRODUCTS	SOLUTIONS	RESOURCES	SUPPORT	CONTACT US
OPEN SOURCE	AMAZON SAGEMAKER	BLOG	DOCUMENTATION	250 NORTHERN AVE SUITE 420 BOSTON, MA 02210
ENTERPRISE	MICROSOFT AZURE	CHEATSHEETS	POSIT COMMUNITY	
CLOUD	POSIT CLOUD	VIDEOS	ALL SUPPORT	844 448 1212
ALL PRODUCTS	ALL SOLUTIONS	CHAMPIONS	LEGAL TERMS	CHAT WITH US
		DATA SCIENCE HANGOUT	STATE OF CALIFORNIA POLICY	INFO@POSIT.CO

Terms & Conditions   Privacy Policy

LibreJS

Free JavaScript in your browser

This site https://posit.co/*

Whitelist | Blacklist | Forget | Reload

No **WHITELISTED** scripts on this page.

ACCEPTED scripts in https://posit.co/blog/:

Show | <SCRIPT> at line 4768 of https://posit.co/blog/

Script appears to be trivial.

Whitelist | Blacklist | Forget

Show | <SCRIPT> at line 4911 of https://posit.co/blog/

Script appears to be trivial.

Whitelist | Blacklist | Forget

Show | <SCRIPT> at line 4918 of https://posit.co/blog/

Script appears to be trivial but defines functions.

Whitelist | Blacklist | Forget

Show | <SCRIPT> at line 4923 of https://posit.co/blog/

Script appears to be trivial.

Whitelist | Blacklist | Forget

Show | <SCRIPT> at line 5065 of https://posit.co/blog/

Script appears to be trivial.

Whitelist | Blacklist | Forget

BLOCKED scripts in https://posit.co/blog/:

Show | <SCRIPT> at line 9 of https://posit.co/blog/

Script uses a reserved object (XMLHttpRequest)

Whitelist | Blacklist | Forget

Show | <SCRIPT> at line 35 of https://posit.co/blog/

NONTRIVIAL: nontrivial token: 'sessionStorage'

Whitelist | Blacklist | Forget

Blog - Posit ✕ | Display JS Monitoring Panel ✕ | +

Eliendajo (GNU LibreJS) moz-extension://93291d56-0406-4917-02a9-c6c7ed1ec161/html/display_panel/content/display-panel.html#fromTab=1

Script appears to be trivial.

Whitelist | Blacklist | Forget

Show | <SCRIPT> at line 5065 of https://posit.co/blog/:

Script appears to be trivial.

Whitelist | Blacklist | Forget

BLOCKED scripts in https://posit.co/blog/:

Show | <SCRIPT> at line 9 of https://posit.co/blog/:

Script uses a reserved object (XMLHttpRequest)

Whitelist | Blacklist | Forget

Show | <SCRIPT> at line 35 of https://posit.co/blog/:

NONTRIVIAL: nontrivial token: "sessionStorage"

Whitelist | Blacklist | Forget

Show | https://posit.co/wp-content/themes/Posit/dist/algolia.js:

External script with no known license

Whitelist | Blacklist | Forget

Show | https://posit.co/wp-content/plugins/ewww-image-optimizer/includes/lazysizes.min.js:

External script with no known license

Whitelist | Blacklist | Forget

Show | https://posit.co/wp-content/plugins/enlighter/cache/enlighterjs.min.js:

External script with no known license

Whitelist | Blacklist | Forget

Show | https://posit.co/wp-content/plugins/complianz-gdpr/cookiebanner/js/complianz.min.js:

External script with no known license

Whitelist | Blacklist | Forget

Show | https://posit.co/wp-content/themes/Posit/dist/app.js:

External script with no known license

Whitelist | Blacklist | Forget

No **BLACKLISTED** scripts on this page.

posit   PRODUCTS   SOLUTIONS   LEARN & SUPPORT   EXPLORE MORE   PRICING   **DOWNLOAD RSTUDIO**

NOW OPEN   AUGUST 12–14 ◎ SEATTLE

posit conf(2024)

snowflake   posit

POST | HADLEY WICKHAM

Get ready for posit::conf(2024)!

Registration for posit::conf(2024) in-person tickets is now open!

POST | JAMES BLAIR

Exploring the Future: Harnessing Data with Snowflake and Posit

Snowflake and Posit have joined forces, integrating Posit's user-friendly data science tools with Snowflake's scalable and secure platform. This blog post explores future enhancements for joint users.

COMPANY NEWS & EVENTS

CALL FOR TALKS

PRODUCTS AND TECHNOLOGY

https://posit.co/blog/harnessing-data-with-snowflake-and-posit/

https://posit.co/blog/

posit conf (2024)

SEATTLE

POST   POSIT TEAM

posit::conf() call for talks deadline extended!

We've extended the deadline for posit::conf(2024) call for talks to February 19!

POST   ALEX CHISHOLM

GitHub Copilot on Posit Cloud

Speed up your coding projects on Posit Cloud with GitHub Copilot, an AI coding assistant.

PRODUCTS AND TECHNOLOGY

PRODUCT RELEASE

Posit Package Manager 2023.12

Reporting and blocking security vulnerabilities and more

posit Package Manager

OPEN SOURCE   PRODUCTS AND TECHNOLOGY

ALL THE NEW FEATURES IN

gt 0.10.0

POST   JOE ROBERTS

POST   RICH IANNONE

posit   PRODUCTS   SOLUTIONS   LEARN & SUPPORT   EXPLORE MORE   PRICING   DOWNLOAD RSTUDIO

Vulnerabilities and More

The December release of Posit Package Manager adds native
integration with the Open Source Vulnerabilities (OSV) database, plus
Git Builders for Python packages and more.

PRODUCTS AND TECHNOLOGY

DATA SCIENCE LEADERSHIP

POST   CURTIS KEPHART

Quarto 1.4 has been released

Quarto 1.4 brings new formats for dashboards and Typst, a new
manuscript project type, a cross-reference overhaul, Shiny for Python
support, and a ton of other updates.

POST   RACHAEL DEMPSEY

Tips for communicating ROI of data
science projects

Successfully translating data science projects into Return on
Investment (ROI) is the key to getting buy-in from stakeholders.

We use cookies to bring you the most relevant experience by remembering your preferences between your visits to our website. By clicking "Accept All," you consent to the use of ALL cookies. However, you may click on "Cookie Settings" to select the types of cookies you choose
to use or avoid.

ACCEPT ALL   COOKIE SETTINGS

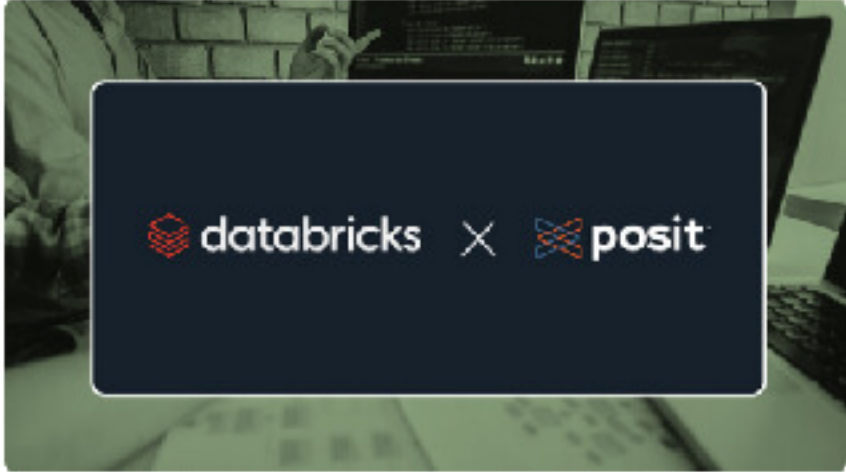https://posit.co/blog/tips-for-communicating-roi-of-data-science-projects/

POST JAMES BLAIR

Posit x Databricks: December '23
Webinar Recap

In December, we teamed up with Databricks for a webinar. Here's a
quick summary of what we covered and the Q&A from the session.

< 1 2 3 4 5 6 7 >

We use cookies to bring you the most relevant experience by remembering your preferences between your visits to our website. By clicking "Accept All," you consent to the use of ALL cookies. However, you may click on "Cookie Settings" to select the types of cookies you choose to use or avoid.

ACCEPT ALL   COOKIE SETTINGS

posit   PRODUCTS   SOLUTIONS   LEARN & SUPPORT   EXPLORE MORE   PRICING

DOWNLOAD RSTUDIO

Subscribe to more inspiring open-source data science content.

We want to help you go further with open source data science. By subscribing, you'll get alerted whenever we publish something new.

SUBSCRIBE

---

posit

The open source data
science company

PRODUCTS	SOLUTIONS	RESOURCES	SUPPORT	CONTACT US
OPEN SOURCE	AMAZON SAGEMAKER	BLOG	DOCUMENTATION	250 NORTHERN AVE
ENTERPRISE	MICROSOFT AZURE	CHEATSHEETS	POSIT COMMUNITY	SUITE 420
CLOUD	POSIT CLOUD	VIDEOS	ALL SUPPORT	BOSTON, MA 02210
				844 448 1212

data

PRODUCTS

OPEN SOURCE

ENTERPRISE

CLOUD

ALL PRODUCTS

SOLUTIONS

AMAZON SAGEMAKER

MICROSOFT AZURE

POSIT CLOUD

ALL SOLUTIONS

Some normal dangers

of proprietary software

* Freedom

* Licensing compliance burden

  from different proprietary licenses

* Waste of time to adapt to proprietary software

* High likelihood of malware in proprietary software

Some dangers specific to R

Posit Cloud in schools?

* Tell students they may not
   learn how things work.

* Enslave students to a SaaSS.

See "The University of Costumed Heroes", by FSF.

Other peoples computers ("clouds")

* R is sometimes used with large
  datasets on several big computers.

* People choose to run other people's
  computers with proprietary APIs.

* Protect freedom (and even save money)
  by running on your own computers
  with standard APIs.

WRONG ARGUMENTS

Better than S-PLUS, SAS, &c?

* Freedom to choose among slave masters is not freedom.

Posit is a good company.

* Freedom to choose among slave masters is not freedom.

* Posit is already releasing proprietary software

Businesses need proprietary licensing

* Do they really?

* If they really do, change society

    to require ethical licensing.

Posit is a B Corporation.

* B Corporation certification does not require that Posit protect our freedom.

* B Corporation certification is a private initiative, so it has little legal weight anyway.

Posit is a Public Benefit Corporation (PBC).

* PBC is a legal entity type in USA:

  A corporation with bylaws requiring that

  the company consider public benefit

* But Public Benefit Corporation does not

  protect our software freedom.

https://en.wikipedia.org/wiki/Benefit_corporation

Kickstarter is also PBC and B Corporation. They

* fired workers for organizing,

* used other normal union-busting tactics,

* failed to bust the union, and

* make their customers run proprietary software.

WHAT CAN WE DO

Don't run the proprietary software.

Use something else.

* Use base R.

* Avoid using SaaSSs through R,
  even if the R package is free.

* Complain to R package maintainers
  if they require you to use GitHub

You really don't need the new packages.

* R community is running in circles
  just like everyone else

* The trendy packages are often overly
  complicated versions of base features

If you must run proprietary software, isolate it.

* privilege separation

* firewall

* virtual machine

* separate computer

* &c.

Make free equivalents of proprietary software.

* R is inspired by S and S-PLUS.

* PSPP is inspired by SPSS.

Complain if people promote proprietary software.

* Complain even if they also promote free software.

* Use the arguments about Posit I mentioned before.

We shall overcome!

Keep R fRee!

APPENDIX

Open source: Free software with the freedom removed

Problems

* attraction is practical, not rights

* people are scared of freedom

* expropriation of volunteer labor

* considering open source as means

   proprietary software is legitimate

Why license matters

* Proprietary: Unethical

* Permissive: Ethical but accepts
  that slavery will continue

* Copyleft: Ethical and fights
  for freedom