

## CELEBRATION 2020 - R WIRD 20! Q WAR IN KOPENHAGEN DABEI.

Am 28. und 29. Februar fand in Kopenhagen die CelebRation 2020 statt, eine Konferenz zum 20. Jubiläum der Version 1.0 der Programmiersprache R. Sie fand in den Räumen der Universität Kopenhagen im Maersk-Turm statt. Etwa 250 R-Nutzer und Entwickler nahmen teil.



## Tag 1: Workshops

Thomas Lin Pedersen (Softwareentwickler bei RStudio) hielt einen Workshop zum Thema Datenvisualisierung mit ggplot2 (Slides). Dirk Eddelbuettel von der Universität Illinois referierte zu Rcpp, einem R Package, dass es erlaubt, C++ Code in R aufzurufen (Slides). Ich besuchte den ersten Workshop und konnte einiges über ggplot2 lernen. Mit Updates (bzw. Releases) der Packages ggforce, ggraph, gganimate und patchwork sind im vergangenen Jahr viele neue Funktionen dazu gekommen, die in unserer täglichen Arbeit bei Q von Nutzen sind. Das Schlüsselwort ist Storytelling: Mit neuen Annotationen und Animationen kann der Leser durch eine Datenanalyse geführt werden. Unter #tidytuesday auf Twitter lassen sich tausende von Beispielen finden.



Foto: Thomas Lin Pedersen zeigt, wie mit ggforce Communitys in Netzwerken markiert werden.

Tag 2: Vorträge und CelebRation

Die Eröffnung machte Peter Dalgaard, Mitglied des R Core Teams. Er zeigte ein echtes Artefakt: Die CD mit der ersten Version von R. Außerdem führte er live den Release der neuesten R Version 3.6.3 durch.



Nordic) wie man R in Docker-Container verpackt und Skripte auf Google Cloud (und anderen Cloud-Plattformen wie AWS und Azure) automatisiert.

Im Kontrast zum nostalgischen Rückblick auf die Entstehung von R zeigte Mark Edmonson (IIH



Entwickler von Data Science Packages sie nutzen können. Einen Eindruck der vielfältigen Anwendungen von R vermittelte Therese Graversen: Sie erzählte, wie sie mittels statistischer Analyse in R DNA-Spuren in Kriminalfällen untersucht.

Überraschend wurde auch ein Vortrag von Tidyverse-Erfinder Hadley Wickham eingeschoben. Er

zeigte eine Reihe von ungewöhnlichen Eigenschaften der R-Programmiersprache und wie

Eine kreative Anwendung von R stellte Thomas Lin Pedersen vor: generative Kunst mit ggplot2. Ein Unterthema, dass sich in vielen Vorträgen wiederfand, war die Spaltung der R Community in klassische Statistiker und Data Scientists. Typisch für Statistiker:

Regressionsmodelle und Hypothesentests

Visualisierung mit Base / Lattice (Beispiele)

Base-R Funktionen R lokal auf dem PC

Akademische Präsentationen mit Formeln & Referenzen

Slides mit Beamer (LaTeX)

Typisch für Data Scientists:

Visualisierung mit ggplot2 (Beispiele)

Machine Learning mit Fokus auf Predictions

Data Scientists allerdings noch nicht vertreten.

schrittweise von SAS auf R umsteigen.

R in der Cloud

Tidyverse

Bildgewaltige Visualisierungen Slides mit Xaringan oder Keynote

Gruppen in der R Community stärken soll. Sie ergänzt damit R-Ladies Global. Turner forderte in ihrer Rede außerdem mehr Transparenz zu den Bedingungen für eine Mitgliedschaft im R Core Team, welches seit 1997 fast unverändert ist und nur aus Männern besteht.

Mit dem breiten Package Ökosystem bedient R beide Interessengruppen. Im R Core Team sind die

Zum Abschluss stellte Heather Turner die Forwards Initiative vor, welche unterrepräsentierte

Rückblick Die Konferenz war eine tolle Gelegenheit, den bekannten Namen aus der Entwickler-Community Gesichter zuzuordnen. Den Code der anwesenden Teilnehmer habe ich schon hundertfach benutzt und jetzt konnte ich einige der Entwickler endlich persönlich kennenlernen. Neben erfahrenen

Entwicklern waren aber auch viele Neulinge in der R-Community dabei. So erfuhr ich

## R bei Q

Der 20. Geburtstag von R bietet auch eine Gelegenheit, über die Entwicklung der Programmiersprache und ihre Bedeutung für Q zu reflektieren. Bei Q ist R mittlerweile im täglichen Einsatz. Wir nutzen es in (fast) jedem Schritt der

beispielsweise, dass das statistische Amt in Dänemark und das Pharmaunternehmen Novo Nordisk

Explorative Datenanalyse Text Mining

Datenerhebung durch Web-Scraping oder Einlesen der Daten von Social Media Monitoring Tools

Datenvisualisierung

Social Media Marktforschung:

Im Oktober 2019 haben wir ein wöchentliches R-Meeting eingerichtet, in dem wir Wissen zu Best

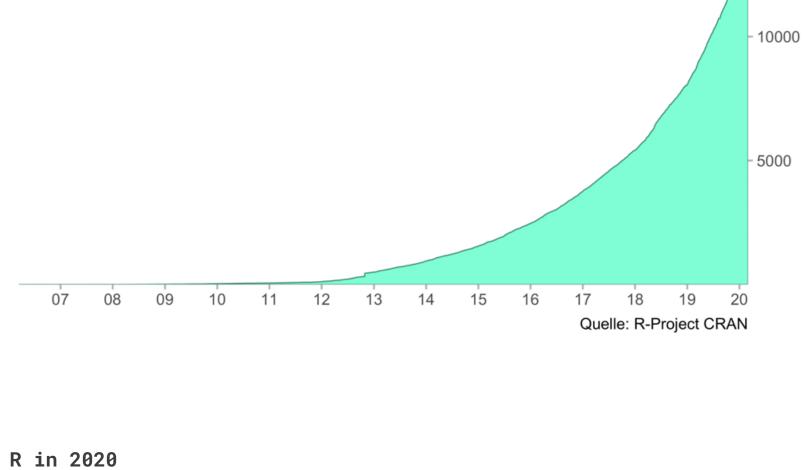
gespeichert und sind alle kostenfrei verfügbar.

Practices und den neusten Packages austauschen. Letztes Jahr erschienen im Durchschnitt 416 neue Packages pro Monat und viele tausend wurden aktualisiert – eine Open Source

Erfolgsstory. Packages werden zentral auf dem CRAN (Comprehensive R Archive Network)

Erstellen von Reports, interaktiven Webseiten und Apps mit R Markdown und Shiny

Packages auf CRAN, März 2006 - Februar 2020 Das Package-Ökosystem wächst exponentiell - 15000



Auch 2020 beginnt R stark - die Programmiersprache wird stetig weiterentwickelt und bleibt doch ihren Wurzeln treu. Während das R Core Team auf die Stabilität von Base-R bedacht ist, entwickelt sich die Sprache mit Packages immer weiter. Sie erweitern R in alle möglichen Richtungen: neue statistische Modelle, neue Visualisierungen, neue Anbindungen an APIs und Datenbanken, neue Reportingformate. Q setzt weiter auf Datenanalyse mit Open Source Software. Wir kombinieren die Datenanalyse-Funktionen von R mit Deep Learning Modulen in Python. Wir bleiben gespannt auf die Neuerungen

im Bereich Data Science, die 2020 bietet. Autor: Paul Simmering

Jobs

Kontakt

Datenschutz





