# Microsoft

# R at (almost) 21
A look back on R's history and impact

David Smith
@revodavid

Photo credit: twitter.com/HannahOish

# About me

Cloud Advocate, Microsoft
Director, R Consortium
Co-author, *Introduction to R*
Editor, blog.revolutionanalytics.com

davidsmi@microsoft.com
@revodavid



"Bit" art by @ashleymcnamara

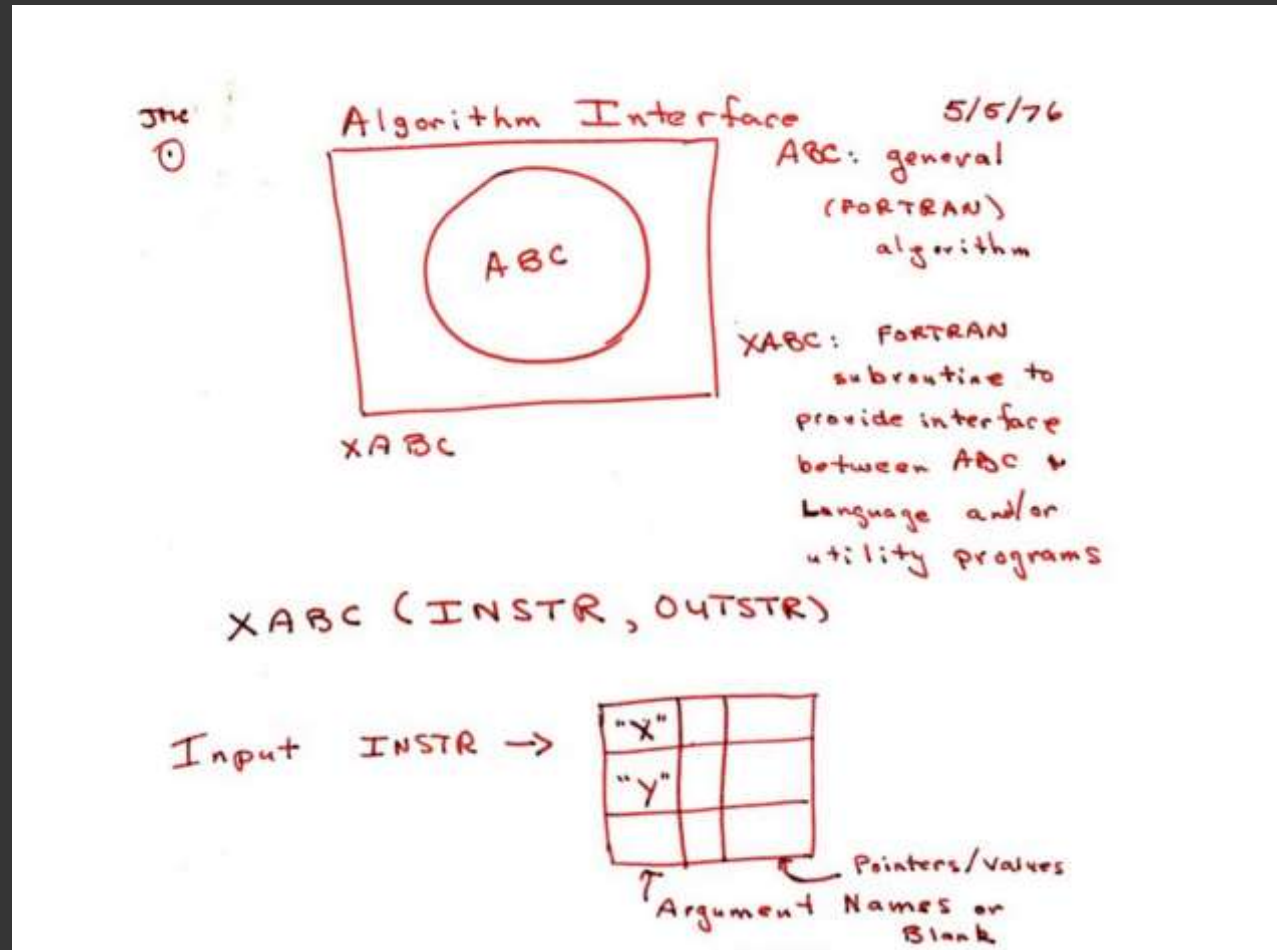# 1976: First, there was S



Photo credit: John Chambers

# 1992: R is born



Photo credit: Robert Gentleman

# 1993-08-04: The First Announcement

```
Date: Wed, 4 Aug 93 14:01:31 NZS
From: Ross Ihaka
To: S-news@utstat.toronto.edu
Subject: Re: Is  S  available for a Macintosh personal computer?

Joseph B Kruskal writes:

If anyone knows of an S available for a Macintosh computer, I
would be pleased to hear about it.

About a year ago Robert Gentleman and I considered the problem of
obtaining decent statistical software for our undergraduate
Macintosh lab.  After considering the options, we decided that
the most satisfactory alternative was to write our own.  We
started by writing a small lisp interpreter.  Next we expanded
its data structures with atomic vector types and altered its
evaluation semantics to include lazy evaluation of closure
arguments and argument binding by tag as well as order.  Finally
we added some syntactic sugar to make it look somewhat like S.
We call the result "R".
```

```
R Installation on the Macintosh (Pre-alpha Version)
----------------------------------------------------

Warning:   This is VERY preliminary.

Pre-requisites:

        A Macintosh computer (floating point is preferable)
        At least 2.5 megabytes of memory free (after the OS loads)
```

# 1995-06: R released under GPL



**R : Past and Future History**

*A Draft of a Paper for Interface '98*

Ross Ihaka
Statistics Department
The University of Auckland
Auckland, New Zealand

# 1997: CRAN founded



CRAN on Nov 9, 2000

# 1997: CRAN founded



CRAN on Mar 26, 2020

# 2000-02-29: R 1.0.0 released
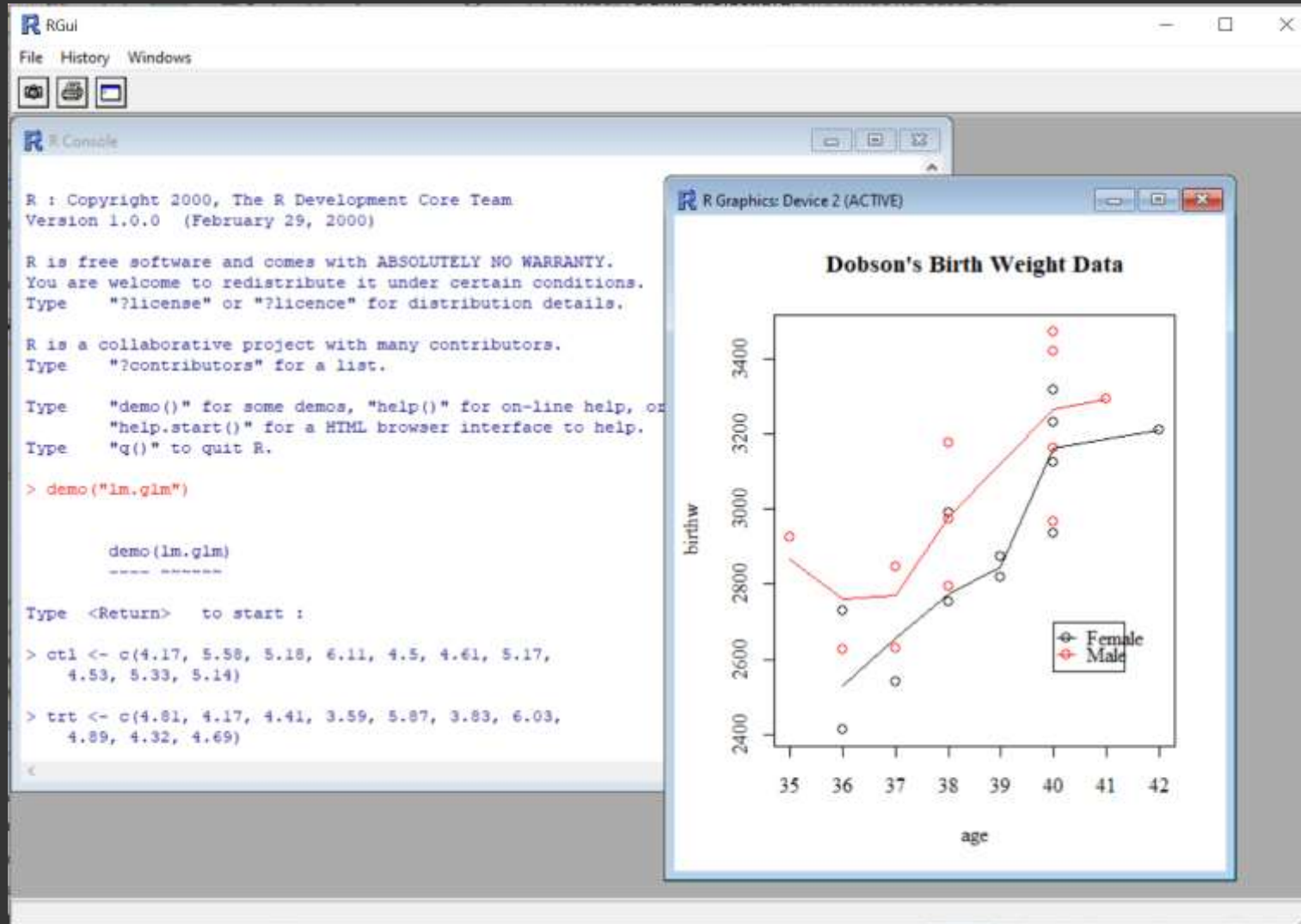


2.8Mb source code
- cf 32Mb today

8.6Mb Windows installer
- cf 84Mb today
- Still runs on Windows 10!

Features a Y2K compliance statement

# 2000-04-03: Incorporation of the R Foundation



Founded in Vienna, Austria

Presidents (2002): Robert Gentleman, Ross Ihaka

Holds R copyright

Funds servers, R Journal, and useR! conference

Now accepts online donations

# 2007: Revolution Computing founded

# 2009-01-06: R hits the Front Page*



The New York Times

**Business Computing**

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION

## Data Analysts Captivated by R's Power

Left; Stuart Isett for The New York Times; right, Kieran Scott for The New York Times

R first appeared in 1996, when the statistics professors Robert Gentleman, left, and Ross Ihaka released the code as a free software package.

By ASHLEE VANCE
Published: January 6, 2009

SCOPING:
function (x)
function (y)
x + y

*of the technology section of the NYT

# 2011-02-28: RStudio initial public release

## [R] A new IDE for R: RStudio

**johannes rara** johannesraja at gmail.com
*Mon Feb 28 19:29:57 CET 2011*

- Previous message: [R] plot data frame
- Next message: [R] power of the comparison of two ROC curves
- **Messages sorted by:** [ date ] [ thread ] [ subject ] [ author ]

Hi,

This is not an official announcement, I just discovered that a new IDE
for R named RStudio was released today. Looks very very promising to
me!

http://www.rstudio.org

-J

- Previous message: [R] plot data frame
- Next message: [R] power of the comparison of two ROC curves
- **Messages sorted by:** [ date ] [ thread ] [ subject ] [ author ]

More information about the R-help mailing list

---

| R | Home | Screenshots | Download | Docs | Support | Development | Blog |

### Introducing RStudio

RStudio™ is a new integrated development environment (IDE) for R. RStudio combines an intuitive
user interface with powerful coding tools to help you get the most out of R.

Diamond Pricing

**Productive**

RStudio brings together everything you need to be
productive with R in a single, customizable
environment. Its intuitive interface and powerful
coding tools help you get work done faster.

**Runs Everywhere**

RStudio is available for all major platforms
including Windows, Mac OS X, and Linux. It can
even run alongside R on a server, enabling
multiple users to access the RStudio IDE using a
web browser.

**Free & Open**

Like R, RStudio is available under an open source
license that guarantees the freedom to share and
change the software, and to make sure it remains
free software for all its users.

Download RStudio Beta

### News

**RStudio Beta Available (2/28/2011)**

We are pleased to announce availability of the beta version of RStudio! RStudio is a new open-source IDE for R that runs
either on your desktop or on a server (where it is accessed using a browser). We invite everyone to check out the
screenshots for more details; download the product to try it out, and follow its ongoing development on github.

# 2015-04-06: Microsoft acquires Revolution Analytics



**Microsoft** | Official Microsoft Blog

Microsoft to acquire Revolution Analytics to help customers find big data value with advanced statistical analysis

Jan 23, 2015 | Joseph Sirosh - Corporate Vice President, Data Group, Microsoft
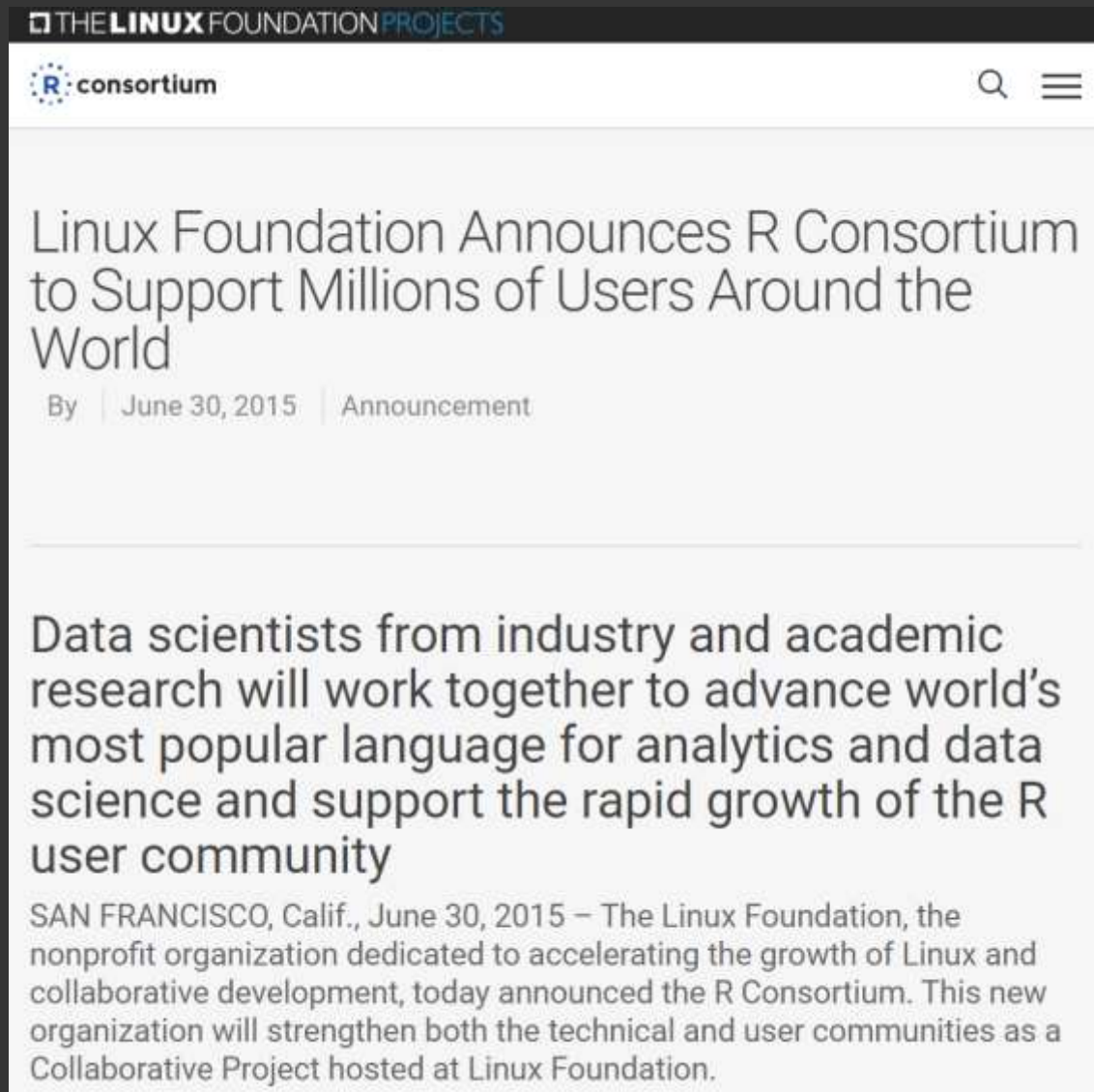
**Update: April 6, 2015:** *Microsoft has closed the acquisition of Revolution Analytics. For more details, please read the blog post by Joseph Sirosh here.*

I'm very pleased to announce that Microsoft has reached an agreement to acquire Revolution Analytics. Revolution Analytics is the leading commercial provider of software and services for R, the world's most widely used programming language for statistical computing and predictive analytics. We are making this acquisition to help more companies use the power of R and data science to unlock big data insights with advanced analytics.

As their volumes of data continually grow, organizations of all kinds around the world – financial, manufacturing, health care, retail, research – need powerful analytical models to make data-driven decisions. This requires high performance computation that is "close" to the data, and scales with the business' needs over time. At the same time, companies need to reduce the data science and analytics skills gap inside their organizations, so more employees can use and benefit from R. This acquisition is part of our effort to address these customer needs.
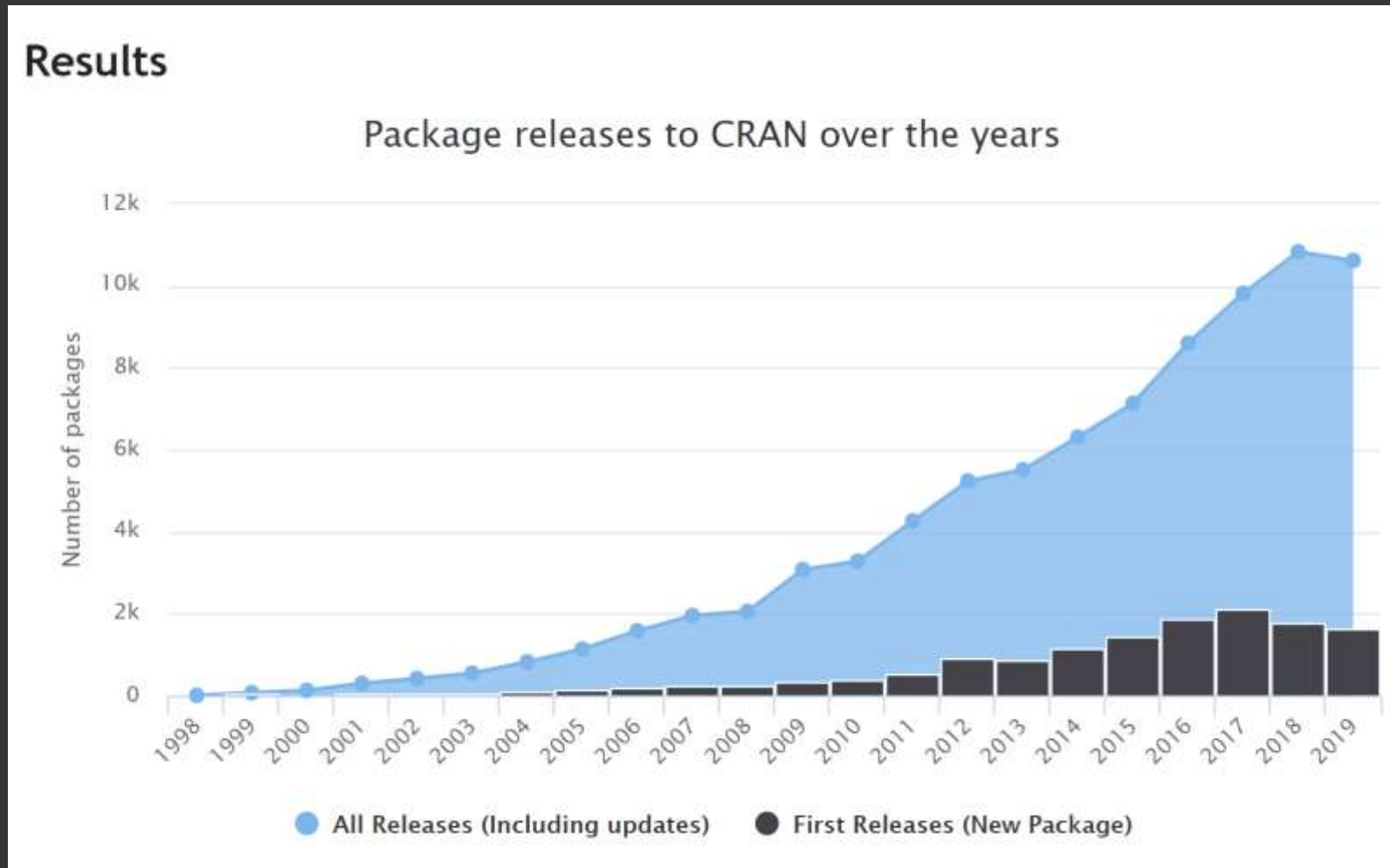
# 2015-06-30: R Consortium founded



**THE LINUX FOUNDATION** PROJECTS

R: consortium

Linux Foundation Announces R Consortium to Support Millions of Users Around the World

By | June 30, 2015 | Announcement

Data scientists from industry and academic research will work together to advance world's most popular language for analytics and data science and support the rapid growth of the R user community

SAN FRANCISCO, Calif., June 30, 2015 – The Linux Foundation, the nonprofit organization dedicated to accelerating the growth of Linux and collaborative development, today announced the R Consortium. This new organization will strengthen both the technical and user communities as a Collaborative Project hosted at Linux Foundation.

Provides funding for:

- R user groups
- R-Ladies
- R-hub
- useR! video streaming
- Community projects, events, and working groups

# 2019-12-31: CRAN Package Growth



**Results**

Package releases to CRAN over the years

Number of packages

All Releases (Including updates)  •  First Releases (New Package)

CRAN packages in 2019:
15368 available
10620 released / updated
1656 new

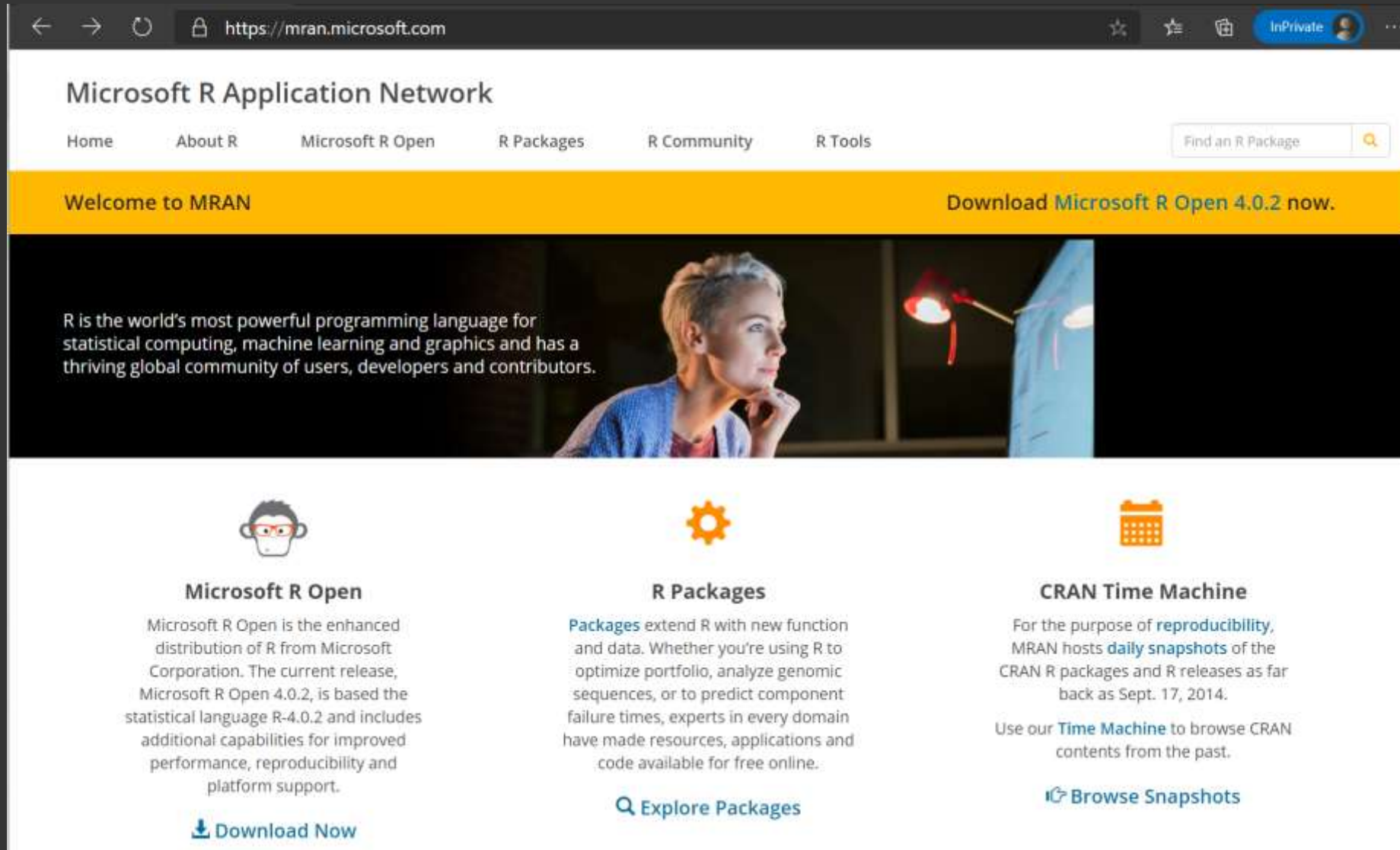Chart: Jozef Hajnala

# 2020-02-29 (R's 20th birthday): R 3.6.3 released

# 2020-04-17 R 4.0.0

- Reference counting (better memory usage)
- Changes to defaults (`stringsAsFactors`)
- New syntax `r("raw strings")`
- Unification of `matrix` and `array` objects
- Performance improvements
- New color palletes

# 2020: Microsoft supports R

# R in SQL Server

R code that is run from "inside" SQL Server is executed by calling a stored procedure. Thus, any application that can make a stored procedure call can initiate execution of R code. Thereafter SQL Server manages the execution of R code as summarized in the following diagram.
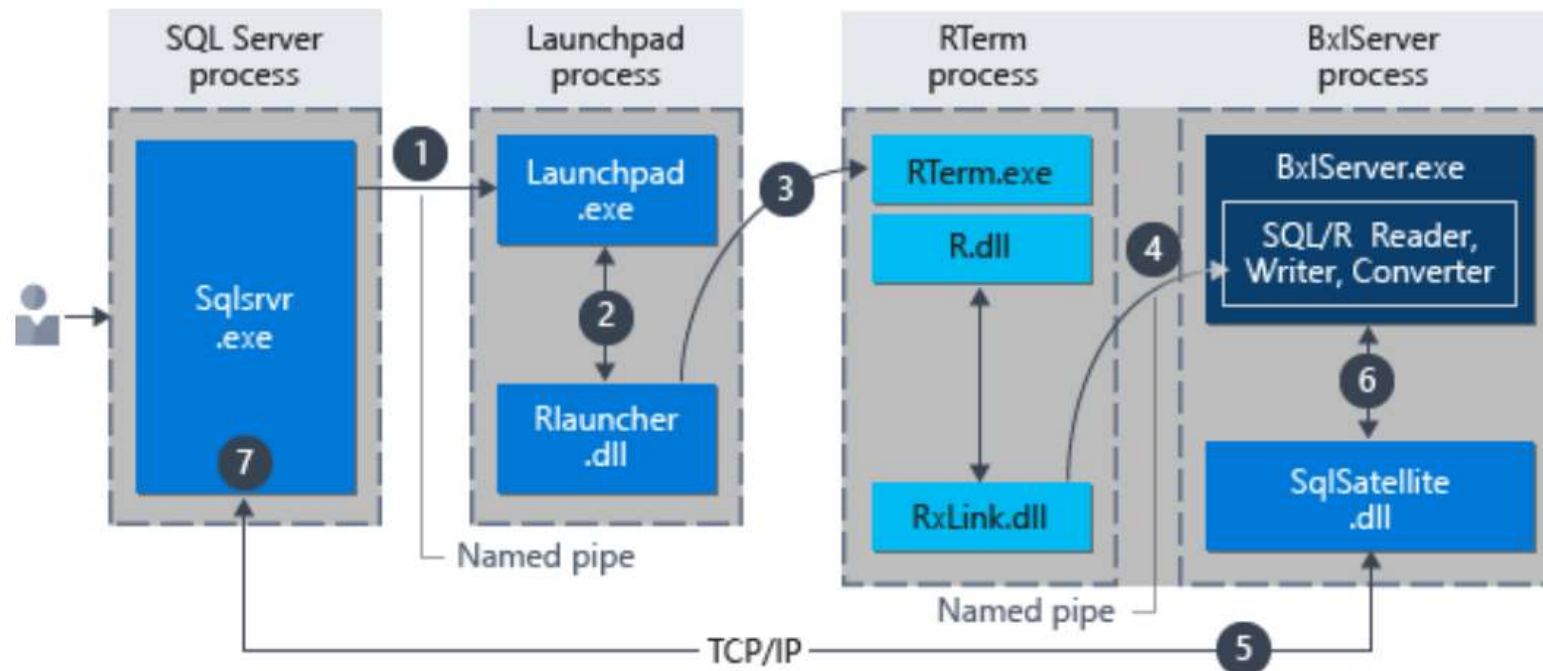
# R in Power BI
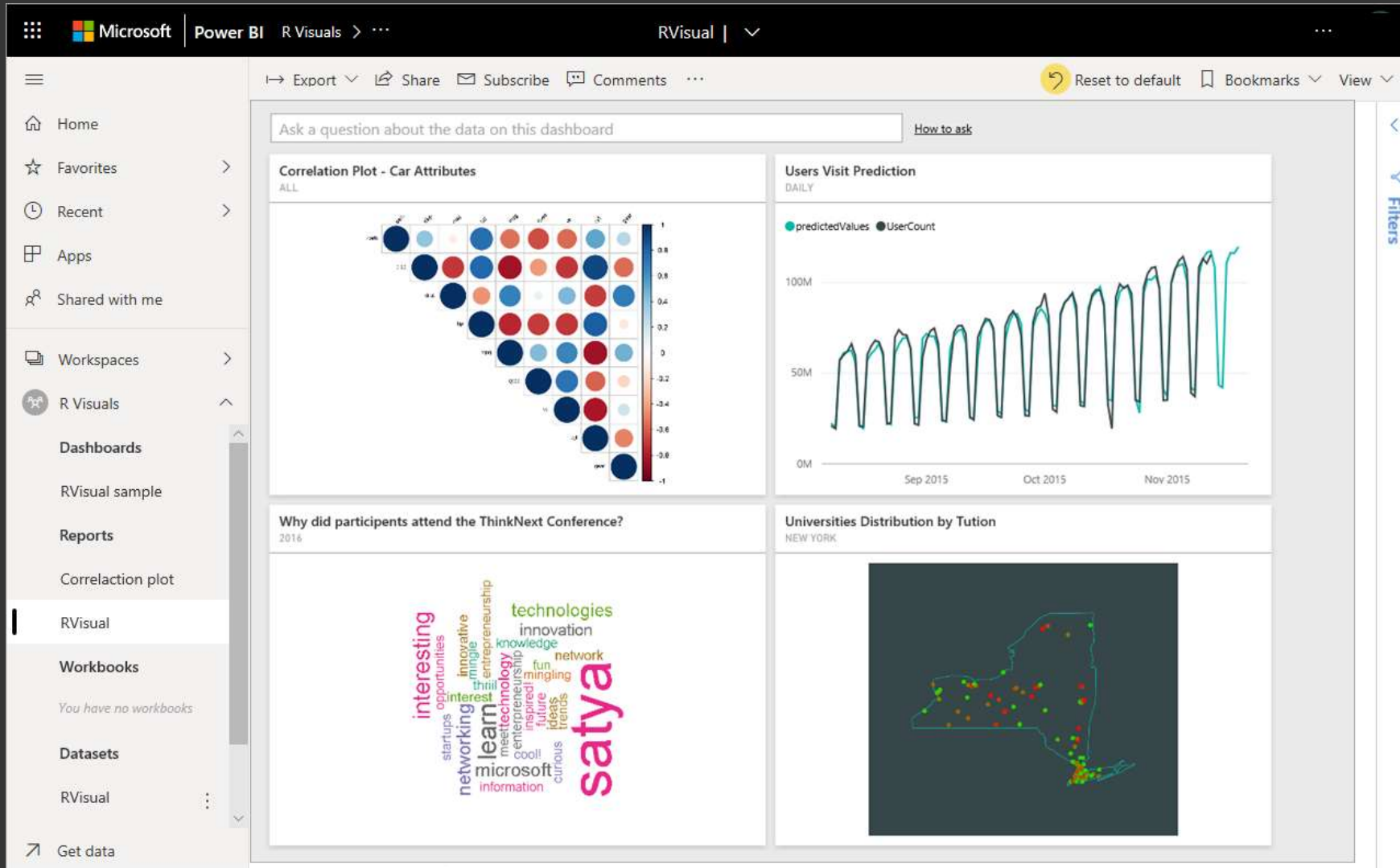
# R in Azure

aka.ms/R-in-Azure

Azure DataBricks

Azure HDInsight

Azure Data Science VM

Azure ML Service

# Azure ML Service: MLOPS and GitHub Actions
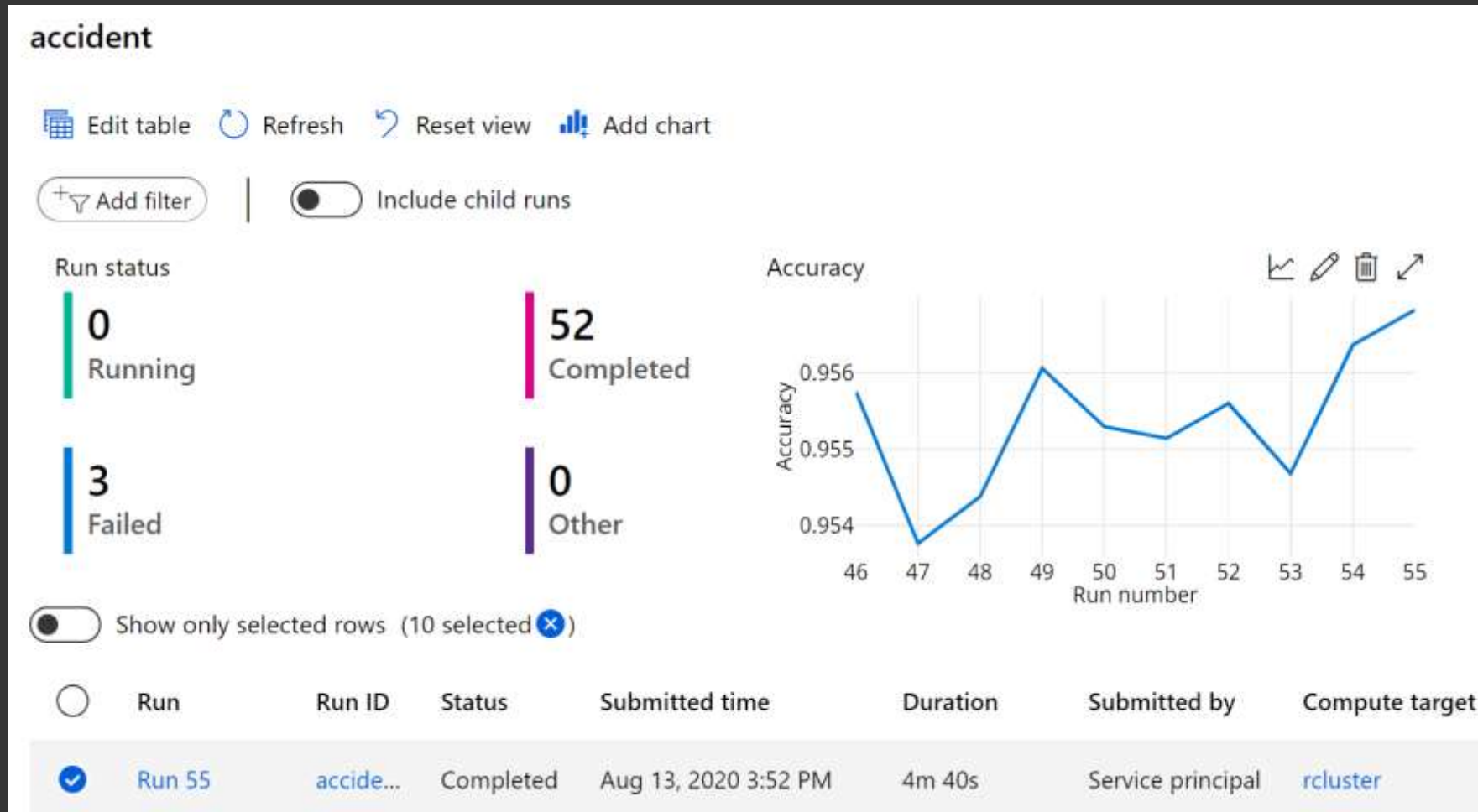
# AzureR: Manage Azure services from R

**AzureRMR**
- Lightweight, extensible R6-based interface to Azure Resource Manager
- Manage subscriptions and resource groups
- Create, update and delete resources and templates
- Work with role-based access control (RBAC)

**AzureKeyVault**
- Resource Manager and client interface to Azure Key Vault
- Secure facility for passwords, cryptographic keys, certificates, storage account logins
  - Encrypt and decrypt with keys
  - Sign and verify certificates
  - Rotate storage account access keys

**AzureStor**
- Resource manager and client interface to Azure blob storage, file storage and Data Lake storage gen2
- List, upload and download files
- Authenticate with access key, SAS or AAD token
- Fast parallel transfers for multiple files

**AzureContainers**
- Interface to Azure Container Registry, Azure Container Instances and Azure Kubernetes Service
- Push and pull images to and from ACR
- Spin up containers with ACI
- Create AKS clusters and deploy and manage services

[github.com/Azure/AzureR](github.com/Azure/AzureR)

**AzureGraph**
- R6-based interface to Microsoft Graph
- Emphasis on registered apps and service principals, to support other packages in family
- Can be extended to work with other services in Graph: SharePoint, OneDrive, Outlook, device management, etc

**AzureVM**
- Flexible, configurable interface to virtual machines and virtual machine scale sets
- Customise your deployment by virtual network, security rules, IP address, load balancer, and more
- Prebuilt configurations for popular Windows and Linux images

**AzureKusto**
- Resource Manager and client interface to Azure Data Explorer, aka Kusto
- Query data using dplyr and DBI interfaces
  - Built in the manner of dbplyr (delayed execution)
- On control plane side, create and manage database principals

**AzureAuth**
- OAuth authentication for Azure Active Directory
- Supports multiple authentication flows: authorization_code, device_code, resource_owner, client_credentials
- Supports AAD v1.0 and v2.0
- Authenticate with password or private key (certificate)

# 2020: R in Industry
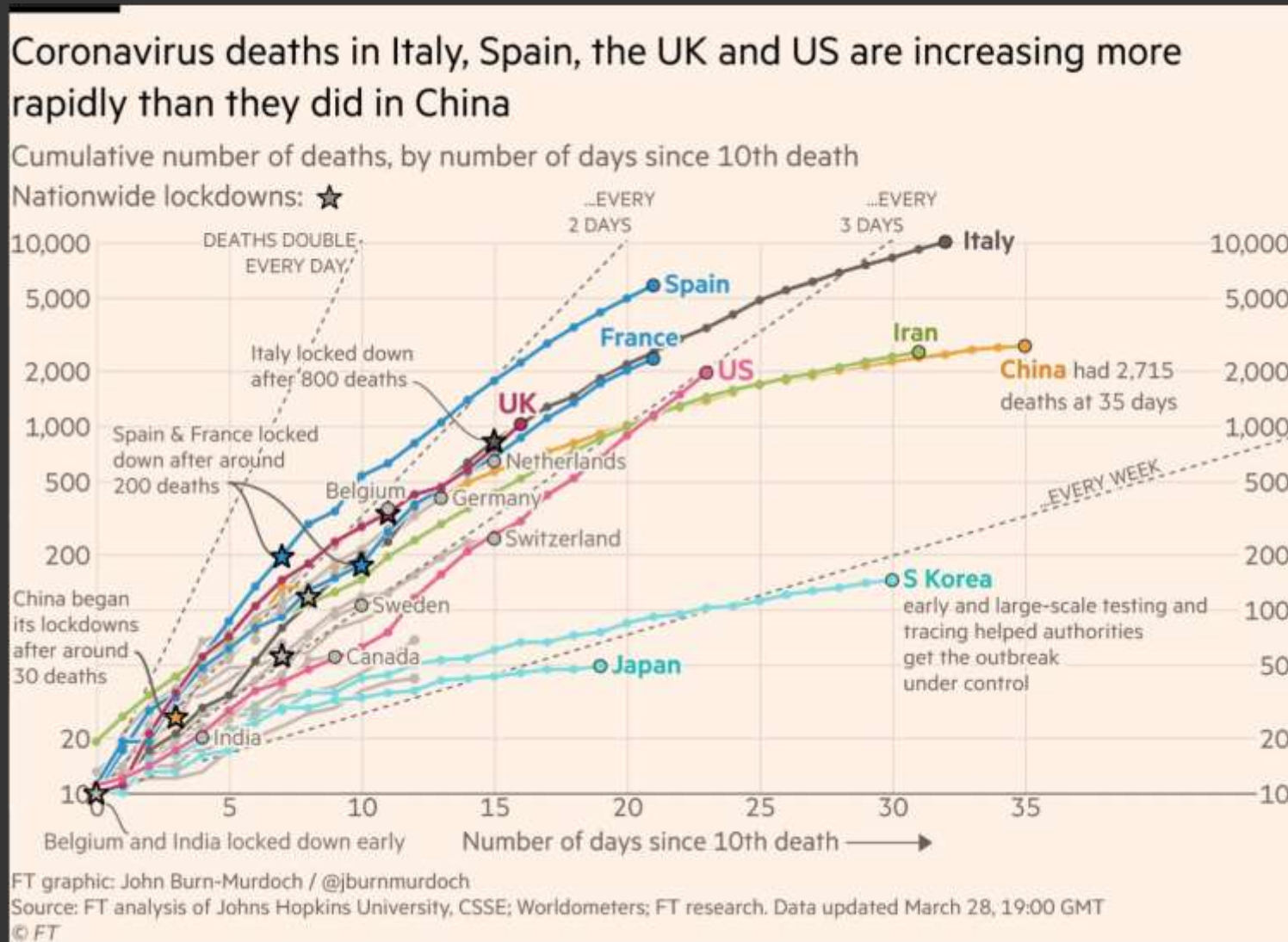
# ... and the next 20?

# Thanks to R Core Team members, past and present

Douglas Bates

John Chambers

Peter Dalgaard

Seth Falcon

Robert Gentleman

Kurt Hornik

Stefano Iacus

Ross Ihaka

Tomas Kalibera

Michael Lawrence

Friedrich Leisch

Uwe Ligges

Thomas Lumley

Martin Maechler

Guido Masarotto

Martin Morgan

Duncan Murdoch

Paul Murrell

Martyn Plummer

Brian Ripley

Deepayan Sarkar

Heiner Schwarte

Duncan Temple Lang

Luke Tierney

Simon Urbanek

# Thank you!

# Slides, links and credits:
# github.com/revodavid/20-years-of-R

David Smith
davidsmi@microsoft.com
@revodavid