R & Python:
A Data Science Love Story

James Blair
February 2020

RStudio

Overview

RStudio Overview

Why R & Python in RStudio?

In this talk, you will learn:

- How Data Scientists can use R & Python together in a single Data Science project
- How IT/DevOps can support those Data Scientists in a single development infrastructure
- How bilingual Data Science teams can efficiently collaborate and share their work with business stakeholders

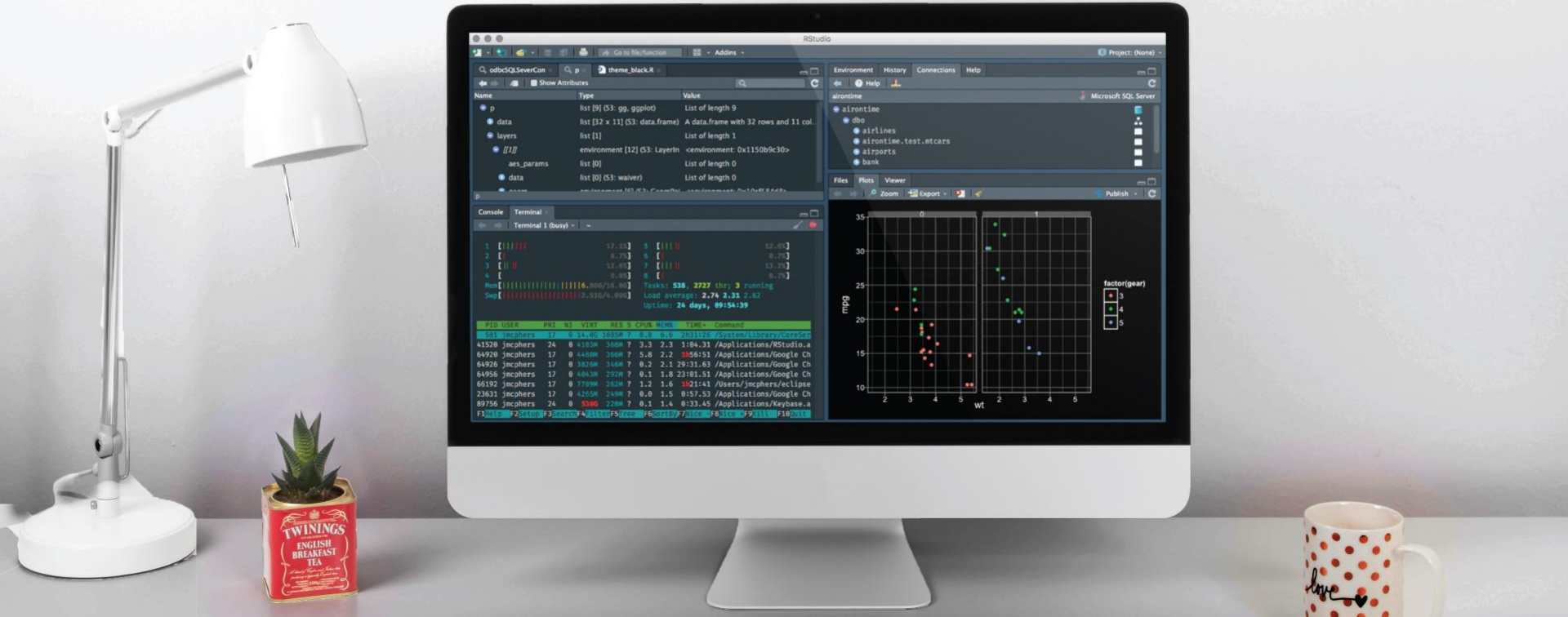
Ongoing Community Investment

Get more info

RStudio®

RStudio Overview

February 2020



About RStudio

2009
Founded by
J.J. Allaire

2011
RStudio IDE

2012
Shiny

2014
RStudio Server Pro
Shiny Server Pro

2016
Spark
Tidyverse

2017
RStudio Connect

2018
Databases
Ursa Labs
TensorFlow
Package Manager

2019
Python & Jupyter
Notebook support

2020
RStudio, PBC
rstudio::conf 2300 attendees

RStudio

RStudio®

RStudio


googledrive
readxl
rvest
readr
tidyr
haven
googlesheets
lubridate
tidymodels
stringr
sparklyr
modelr
hms
ggplot2
dplyr
recipes
R2D3
blogdown
plumber
rmarkdown
Shiny
xaringan
bookdown
knitr

RStudio


googledrive
readxl
rvest
readr
tidyr
haven
googlesheets


lubridate
tidymodels
stringr
sparklyr
modelr
hms
ggplot2
dplyr
recipes


R2D3
blogdown
plumber
rmarkdown
Shiny
xaringan
bookdown
knitr


RStudio Server Pro


googledrive
readxl
rvest
readr
tidyr
haven
googlesheets


lubridate
tidymodels
stringr
sparklyr
modelr
hms
ggplot2
dplyr
recipes


R2D3
blogdown
plumber
rmarkdown
Shiny
xaringan
bookdown
knitr


RStudio Connect


RStudio Server Pro

RStudio Package Manager

RStudio Connect

RStudio Server Pro

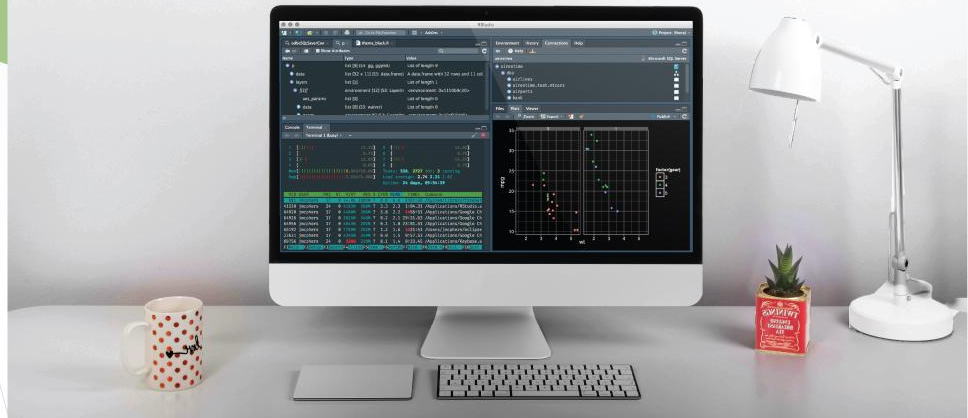
OPEN SOURCE TOOLS

Accessible to everyone

RStudio Package Manager

RStudio Connect

RStudio Server Pro




OPEN SOURCE TOOLS
Accessible to everyone
ENTERPRISE PRODUCTS
Invest in
Studio Package Manager
Studio Server Pro
Studio Connect
readr
tidyr
ggplot2
dplyr
recipes
rmarkdown
Shiny
plumber
bookdown
knitr
xaringan
haven
googlesheets
rvest
readxl
googledrive
sparklyr
modelr
R2D3

Strategy

50%* of company is Engineering

52%* of engineering is open source

* As of February 5, 2020

Why R & Python in RStudio?

- Help organizations realize the full value of all their Data Science investments
- Long term focus on integrating with the tools and environments that Data Scientists need. E.g.,
  - Shiny
  - Spark
  - SQL/Databases
- Now, we've expanded our Python support to help bilingual Data Science teams
  - Tensorflow integration
  - reticulate package
  - Python, Jupyter Notebooks and JupyterLab support

RStudio®

Common Challenges for Bilingual Data Science Teams

Data Scientists
Want to use the best tool for the job, so constantly need to switch contexts among multiple environments.

DevOps/IT
Spending time and resources to maintain, manage and scale separate environments for R and Python in a cost-effective way.

Data Science Leaders
Wrestle with how to share results consistently and deliver value to the larger organization, while providing tools for collaboration between R and Python users on their team.

Business Stakeholders
Not interested in underlying details, but simply want to leverage insights from Data Science to make better decisions.

RStudio

Maintaining separate environments for each data science language is a pain - especially across desktops! Packages fail, jobs conflict, and your team scrambles to learn unsupported platforms.

DevOps / IT Admin

RStudio

DevOps / IT Admin

Maintaining separate environments for each data science language is a pain - especially across desktops! Packages fail, jobs conflict, and your team scrambles to learn unsupported platforms.

Use RStudio Team to:

- Provide a single infrastructure for R & Python
- Enable users favorite tools: RStudio & Jupyter
- Configure, Integrate, Scale, and Secure **once**

RStudio


R Studio®
&
jupyter

Data Scientist

You want the right tool for the job, but mixing R and Python requires manual translation, duplicate code, and tedious conversion.

RStudio

Data Scientist

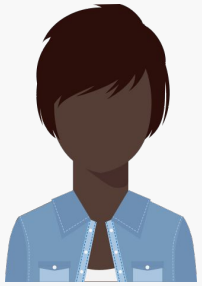
You want the right tool for the job, but mixing R and Python requires manual translation, duplicate code, and tedious conversion.

Use reticulate in RStudio to:

- Combine R & Python in R Markdown
- Source and Use Python Functions in R
- Edit and Run Python in RStudio

RStudio

RStudio



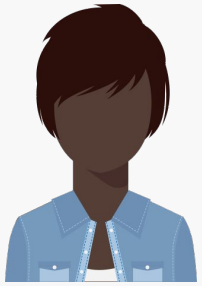
Data Science Leader

Business Stakeholder

Multilingual teams struggle to collaborate, forced to do repetitive analyses, reinvent wheels and translate work.

Business leaders struggle with the slow iteration time, worry about the value of their data science investment, and fall back on intuition - leading to suboptimal decisions.

RStudio



Data Science Leader

Business Stakeholder

Multilingual teams struggle to collaborate, forced to do repetitive analyses, reinvent wheels and translate work.

Business leaders struggle with the slow iteration time, worry about the value of their data science investment, and fall back on intuition - leading to suboptimal decisions.

Use RStudio Team to:

- Have multi-lingual team members collaborate and share work
- Deploy up-to-date interactive analyses, dashboards and custom reports to a central location for easy discover and controlled access
- Send custom emails with inline data science reports
- Make better decisions based on the latest information

RStudio

Connect data scientists with decision makers

http://rstd.io/gartner-movies

RStudio

Solving the Major Challenges for R & Python Teams

Data Scientists

- Combine R and Python in a single project
- Launch Jupyter Notebooks or JupyterLab from a single infrastructure

DevOps/IT

- Provide a single infrastructure for R & Python
- Enable users favorite tools: RStudio & Jupyter
- Configure, Integrate, Scale, and Secure ONE TIME

Data Science Leaders

- Enable collaboration across your team, maximize productivity
- Deliver regularly updated results, custom reports, and self-serve apps in a single location or via email

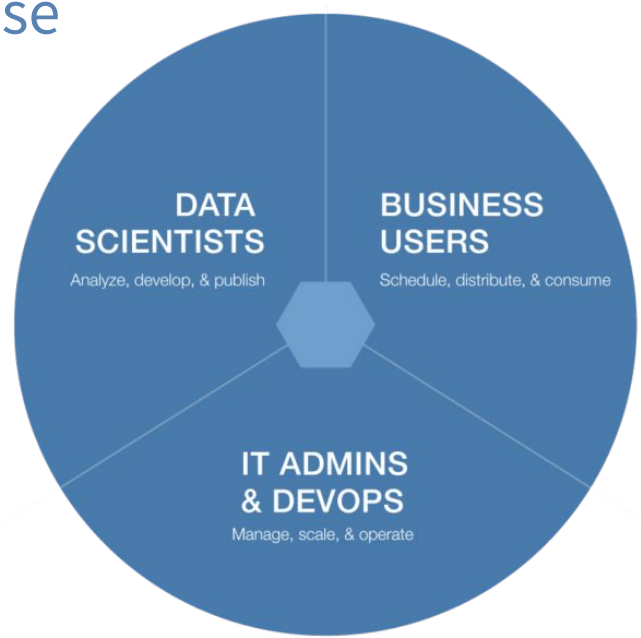
Business Stakeholders

- Access up-to-date interactive analyses, dashboards and emails
- Get answers when you need them to make better decisions

RStudio

RStudio Team Enterprise

Our recommended professional data science solutions for every team



RStudio Server Pro

ANALYZE DATA

Data Scientists use RStudio Server Pro to analyze data and create data products using R and Python

RStudio Connect

PUBLISH RESULTS

Business Users and Collaborators use data products on RStudio Connect that are published by Data Scientists

RStudio Package Manager

MANAGE PACKAGES

IT Administrators use RStudio Package Manager to control and manage R packages that Data Scientists need to create and share data products.

Click for Team Enterprise Demo Video


Walmart

Carnival

janssen
PHARMACEUTICAL COMPANIES
OF Johnson & Johnson

STITCH FIX

Farmacias del
Ahorro

accenture
High performance. Delivered.

AstraZeneca

NBC

GE

THE FAMILY
CONNECTION FOUNDATION

HONDA
The Power of Dreams

METRUM
RESEARCH GROUP

waze

EXPRESS SCRIPTS®

NASA

aridhia

ebay

DTCC

Liberty
Mutual

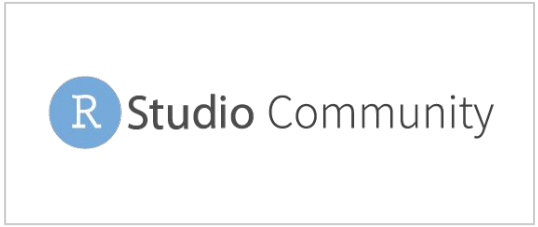
Santander

AVON

SAMSUNG

AMGEN

Ongoing Investment in R & Python Communities



For More Information

- Overview of R & Python: A Love Story: rstudio.com/python
- Contact us to learn more, get answers
- Webinar recording
- Talk Materials

- reticulate package
  - Package website
  - Examples and Webinar
- Documentation
  - RStudio Server Pro and Jupyter
  - Sharing Jupyter Notebooks in RStudio Connect
- RStudio Community

RStudio

Questions?

Or to set up a conversation:
rstudio.drift.click/R_and_Python

RStudio