





Programación en R para Ciencia de Datos Miguel Jorquera

DBDC-2021 Educación Profesional Escuela de Ingeniería

El uso de apuntes de clases estará reservado para finalidades académicas. La reproducción total o parcial de los mismos por cualquier medio, así como su difusión y distribución a terceras personas no está permitida, salvo con autorización del autor.



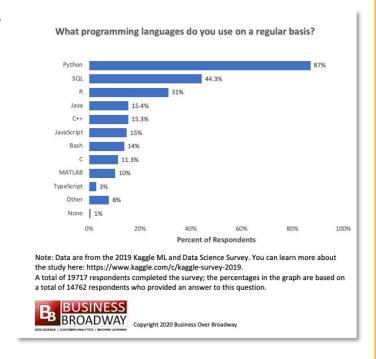


INTRODUCCIÓN

INTRODUCCIÓN

¿Qué es R?

- Software libre, basado en lenguaje S.
- Lenguaje funcional.
- Uso: Principalmente investigación, academia, análisis de datos.
- Ventajas (intuitivo, especializado, buena documentación, comunidad y CRAN)
- Desventajas (paradigma poco flexible, nativamente poco escalable*)







CRAN

Comunidad global



Academia

Metodologías actualizadas



Open Source

Colaboración masiva



Empresa

Auge y soporte



INTRODUCCIÓN



¿Para qué es?

- Software altamente especializado en el análisis de datos.
- Ideal para análisis descriptivo, modelamiento predictivo e implementación de modelos de machine learning.
- Por ejemplo...







INSTALACIÓN LOCAL

INSTALACIÓN



¿Qué es R?

- Instalación del R en local Ultima versión disponible desde CRAN (https://cran.r-project.org/).
- Interfaz de desarrollo local Rstudio (https://www.rstudio.com/)
- Interfaz cloud
 Google Colab (R) (colab.fan/r) (crear script desde cero con kernel R)
 Nuestro
 repositorio(https://github.com/majorquev/DBDC 202103 Programacion en R para ciencia de datos.git)



CRAN
Mirrors
What's new?
Task Views
Search

About R R Homepage The R Journa

Software
R Sources
R Binaries
Packages
Other

Documentation
Manuals
FAQs
Contributed

The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, Windows and Mac users most likely want one of these versions of R:

- Download R for Linux
- Download R for (Mac) OS X
 Download R for Windows
- Download R for Windows

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (2020-06-22, Taking Off Again) <u>R-4.0.2.tar.gz</u>, read <u>what's new</u> in the latest version.
- Sources of <u>R alpha and beta releases</u> (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are <u>available here</u>. Please read
 about <u>new features and bug fixes</u> before filing corresponding feature requests or bug reports.
- . Source code of older versions of R is available here
- · Contributed extension packages

Questions About

If you have questions about R like how to download and install the software, or what the license
terms are, please read our <u>answers to frequently asked questions</u> before you send an email.





