



ESCUELA DE INGENIERÍA
FACULTAD DE INGENIERÍA

EDUCACIÓN
PROFESIONAL

Programación en R para ciencia de datos

Diplomado en Big Data y
ciencia de datos

Profesor:

Miguel Jorquera Viguera





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*

Características de R



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...



Características de R



CRAN

Comunidad global



Academia

Metodologías actualizadas



Open Source

Colaboración masiva



Empresa

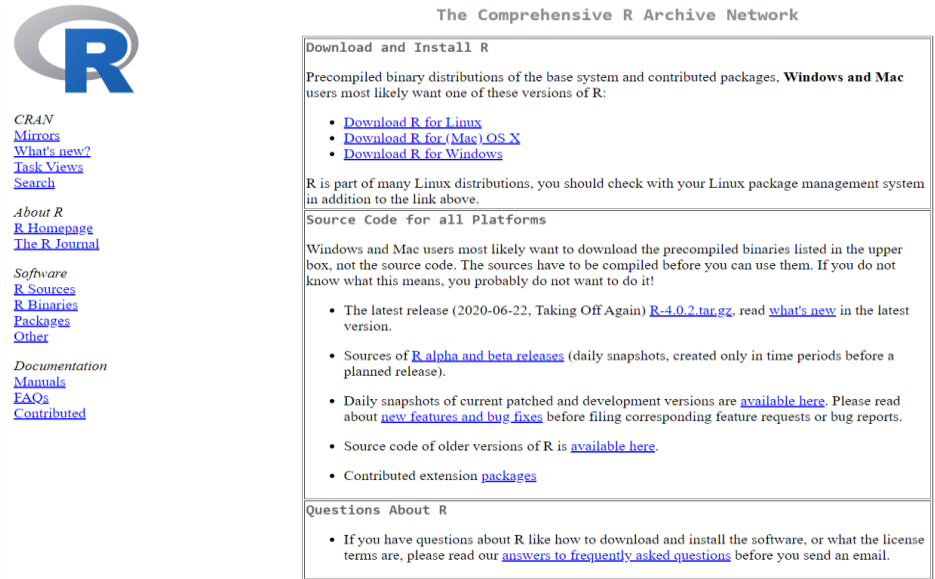
Auge y soporte



INSTALACIÓN LOCAL

Acerca del curso

- **Instalación del R en local**
Última versión disponible desde CRAN (<https://cran.r-project.org/>).
- **Interfaz de desarrollo local**
Rstudio (<https://posit.co/>)
- **Interfaz cloud**
Google Colab (R): colab.to/r (crear script desde cero con kernel R)
Nuestro repositorio:
https://github.com/majorquev/DBDC_202307_Programacion_en_R_para_ciencia_de_datos.git





ESCUELA DE INGENIERÍA
FACULTAD DE INGENIERÍA

EDUCACIÓN
PROFESIONAL



www.educacionprofesional.ing.uc.cl