

Módulo R Fundamentals

Fundamentos de R studio y sus paquetes

Sílvia Ariza Sentís

- 15 min. Big Data
- 15 min. Fuentes de datos
- 15 min. Data Scientist
- 15 min. Ciencia de datos
- 30 min. ¿Qué es R?
- 30 min. Entorno R
- 30 min. Caso práctico
- 5 min. Conclusión



Calentamiento



Alfredo Lozano

Estudié matemáticas aplicadas.

Como científico de datos he trabajado para startups data-driven como Kubo financiero y SinDelantal.

He trabajado en investigación para ONG's como TuElección y el Datalab ITAM.

He publicado artículos de divulgación en temas como procesos de decisión en sistemas multiagentes, procesamiento de lenguaje natural, teoría de categorías y sistemas dinámicos.

¿Qué es el Big Data?

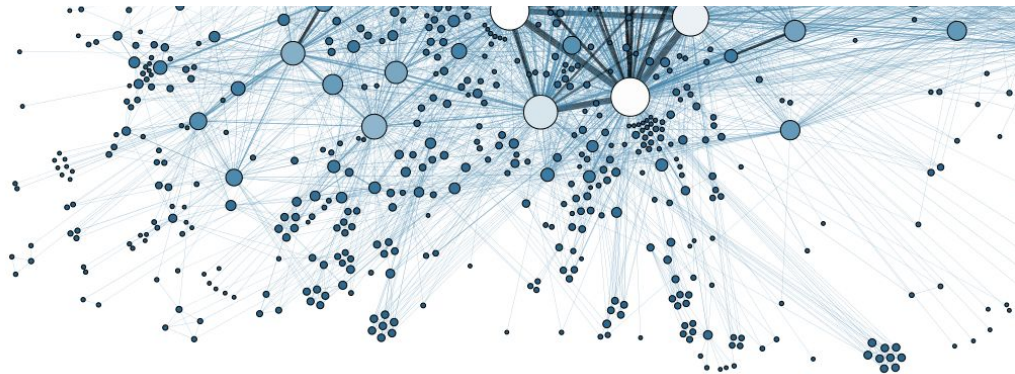
Gran volumen de datos, pero no sólo es eso

Todas las empresas almacenan datos diariamente; de **clientes**, **productos** y **operaciones**

Se pueden encontrar **patrones** de conducta, **correlaciones** entre variables y crear un puente de **comunicación** entre los clientes y la compañía

Disminuir **costos**, aumentar su **eficiencia** y comprender las **necesidades** de su público

Solución matemática - estadística a un problema de negocio



¿Cuáles son sus pilares?

Ejes Big Data

Volumen

Volumen que requiere almacenar datos pasados, presentes y futuros sin problemas de seguridad a los usuarios

Velocidad

Velocidad necesaria para
manipular bases de datos
multivariantes y extensas

Variedad

Diversidad del origen y tipología de datos; desde correos electrónicos, videos, audios, bases de datos estructuradas y no estructuradas



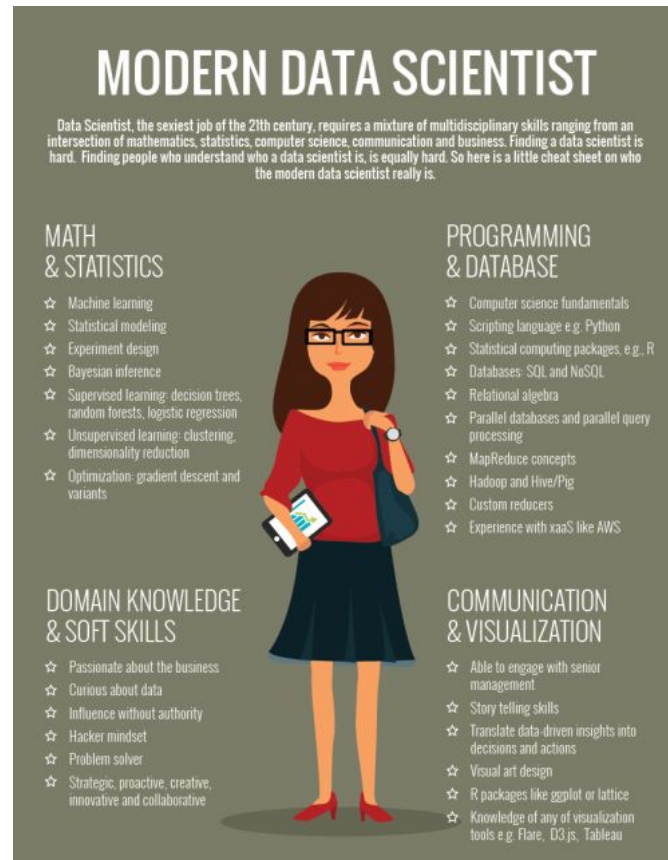
-

Tenemos mucha información. ¿Y ahora qué hacemos con ella?

¿Qué podemos hacer con los datos?

¿Cómo optimizamos los resultados de nuestra empresa?

Departamento de inteligencia del negocio y Big Data



BI

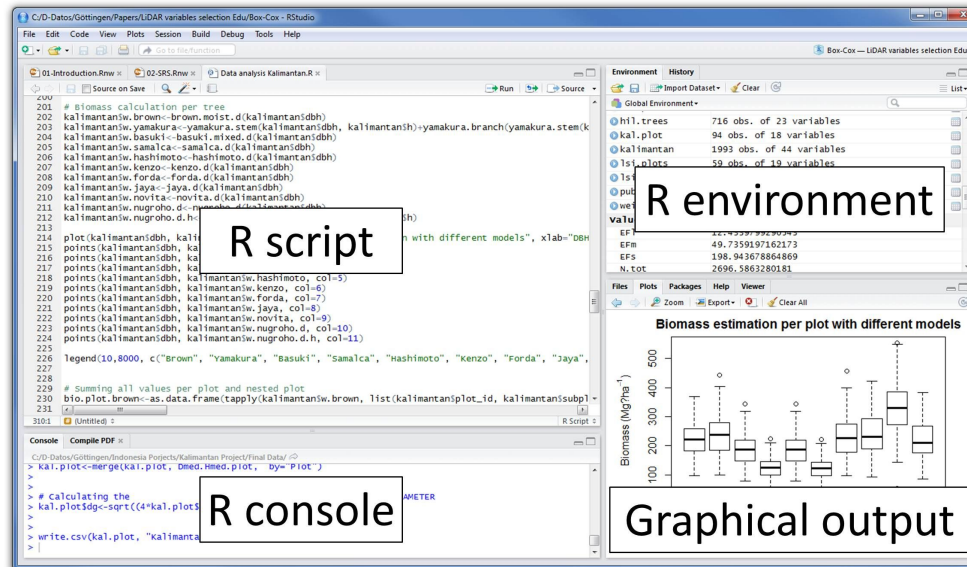
Extracción de la
información
Dashboards
Automatización
de procesos
SQL

DS

Machine Learning
Modelos
estadísticos
R / Python

- Universidad de Auckland, 1993
- Idioma: Inglés
- Software libre más utilizado en las universidades del mundo
- Lenguaje de programación
- Análisis estadístico
- Software libre
- ≈ 8.000 packages para tareas computacionales
- Stack Overflow

- <https://www.rstudio.com/products/rstudio/download/>
- Script: editor de texto
- Consola
- Entorno: espacio para los objetos creados
- Resultados gráficos: visualizador de objetos gráficos





Actividad

Ejemplo 01

Calculadora:

Vamos a realizar una operación, usaremos R como calculadora. ¿Cuántas horas tiene un año de 365 días?

¿CUÁNTAS HORAS TIENE 1 AÑO?

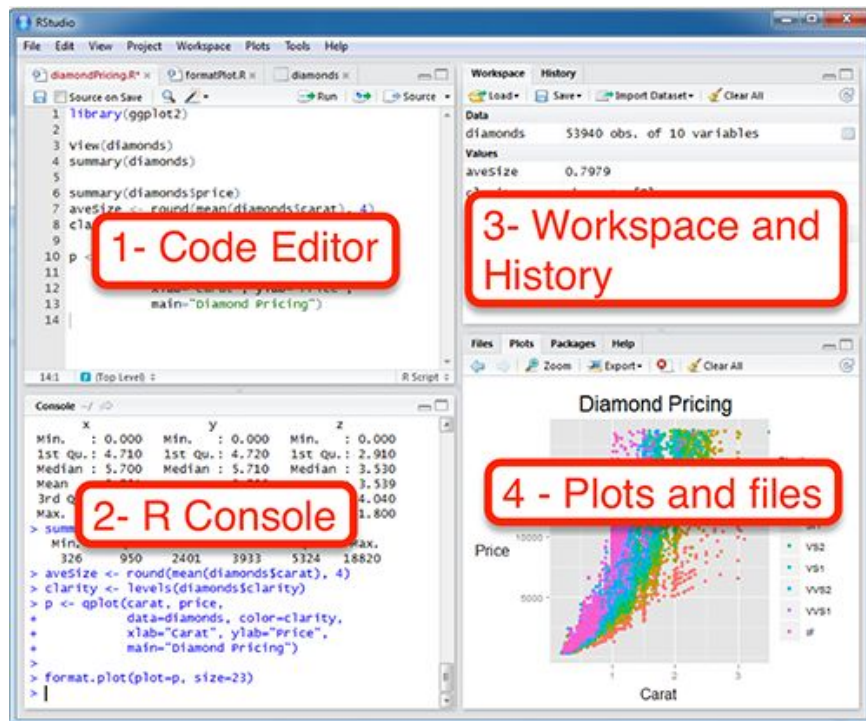




Actividad

Reto 01

Apoyate de R Studio para ejecutar una operación con la calculadora, la que tu quieras.



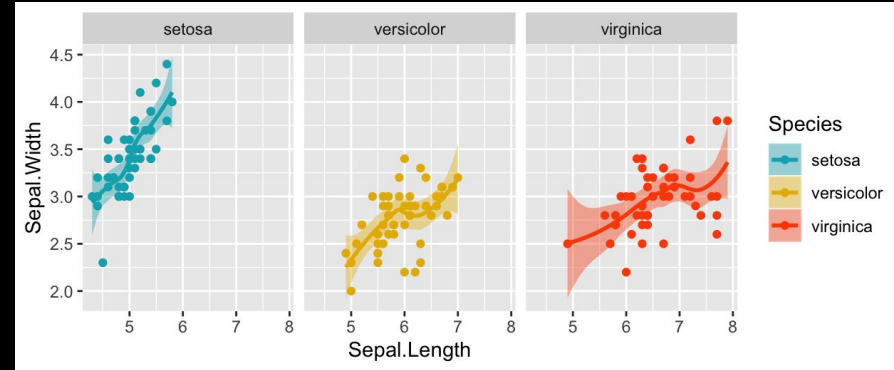


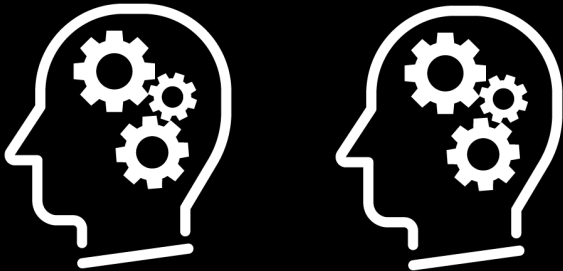
Actividad

Ejemplo 02

Paquetes:

Instalaremos un paquete de nuestra elección, en este caso, ggplot2, que me permitirá hacer ejercicios de visualización.

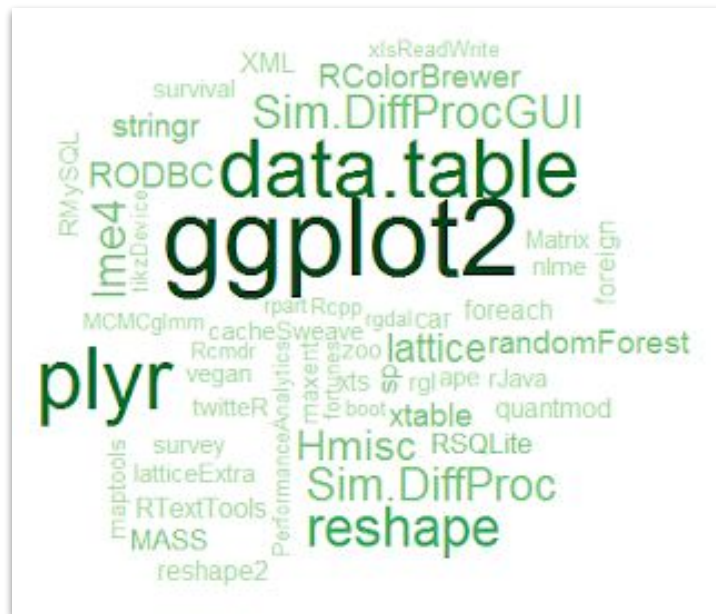




Actividad

Reto 02

De todos los paquetes existentes, instala uno, que ya quedará guardado en la memoria de tu R Studio para siempre!





Preguntas

BEEDU