



(1) Instalación de R y RStudio. Importación y visualización básica de datos

Rubén Heradio
Universidad Nacional de Educación a Distancia

**Plasencia 2025 – Curso de Análisis y Visualización de Datos: Estadística
Práctica con R e Inteligencia Artificial**

¿Por qué merece la pena estudiar estadística?

Enfocándonos en nuestro trabajo como investigadores, la estadística es fundamental:

- Para el método científico: entender el estado del arte, si nuestro método funciona o mejora a la competencia, etc.
- Para decidir ¿sobre qué investigar? ¿donde publicar? ¿con quien conviene trabajar?
- Para escribir artículos:
 - Artículo convencional: **Abstract**, **intro**, related work, our approach, **experimental validation**, **conclusions**.
 - Meta-artículos: **literature review**, **meta-analysis**, **bibliometric study**, etc.

¿Por qué deberíamos darle una oportunidad a R?

- R es una herramienta de calidad: produce gráficos muy profesionales, es capaz de procesar eficientemente grandes volúmenes de datos, etc.
- R es Open Source:
 - Gratuito
 - Una comunidad de usuarios y desarrolladores muy dinámica: a 1 de marzo de 2016, 8.006 paquetes disponibles en <https://cran.r-project.org/web/packages/>
A 5 de febrero de 2017, ya hay 10.043 paquetes!
 - Facilita la reproducibilidad experimental
- R tiene una proyección muy buena. Según un estudio de IEEE Spectrum fue el 6º lenguaje de programación más relevante de 2015, por encima de PHP, Javascript, etc.
<http://spectrum.ieee.org/static/interactive-the-top-programming-languages-2015>
En 2016 ha ascendido una posición:
<http://spectrum.ieee.org/computing/software/the-2016-top-programming-languages>
- Si desarrollamos en R, el resultado es publicable: *Journal of Statistical Software*.
Q1 en STATISTICS & PROBABILITY / Q1 en COMPUTER SCIENCE,
INTERDISCIPLINARY APPLICATIONS

Contenido del curso de iniciación (i)

- Instalación de R
- Manejo básico
 - Tipo de dato básico: el dataframe

	group letter (A-C)	age	admission time	discharge
1	A	27	11:00 AM	11:00 AM
2	A	29	11:00 AM	11:00 AM
3	A	31	11:00 AM	11:00 AM
4	A	33	11:00 AM	11:00 AM

- Tipo de dato atómico: el vector
- Bucles y sentencias If
- Estadística descriptiva
 - Introducción a la librería estándar de R
 - Paquetes para “summaries”

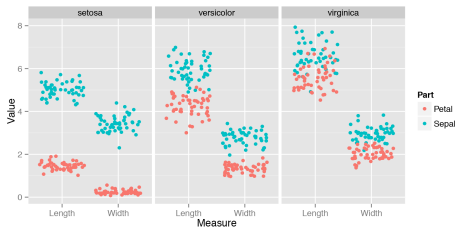
Contenido del curso de iniciación (ii)

- Inferencia estadística
 - Generación aleatoria de muestras
 - Simulaciones
 - Teorema Central del Límite
 - t-test
 - tamaño del efecto, potencia estadística, intervalo de confianza, etc.

Contenido del curso de avanzado

- Inferencia estadística: ANOVA
 - Justificación
 - ANOVA simple
 - ANOVA factorial
- Estadística descriptiva: ggplot2

```
> ggplot(iris.tidy, aes(x = Measure, y = Value, col = Part)) +  
  geom_jitter() +  
  facet_grid(. ~ Species)
```



Instalación de R y RStudio

R

<https://cran.r-project.org/mirrors.html>

Instalación de R y RStudio

R

<https://cran.r-project.org/mirrors.html>

RStudio

<https://www.rstudio.com/products/rstudio/>

Instalación de R y RStudio

R

<https://cran.r-project.org/mirrors.html>

RStudio

<https://www.rstudio.com/products/rstudio/>

Comprobando que todo es correcto

```
demo(graphics)
```

```
demo(persp)
```

Comentarios generales

- R es **interpretado**
- R sigue el paradigma de **programación funcional**
- **RStudio** facilita editar y ejecutar programas, así como visualizar los resultados