

# Enseñando R en el Grado en Estadística

**Teresa González Arteaga**

Universidad de Valladolid

VII Jornadas de Usuarios de R

Salamanca, 5 de noviembre 2015

# Índice

---

1. Sobre el Grado en Estadística
2. Computación estadística: una asignatura particular
3. Tópicos de R y recursos
4. Comentarios finales

# Titulaciones de Grado en Estadística (2015 - 2016)

Universidades:

UAB

UB-UPC

UVIC

Carlos III

UCM

UMH

US

UGR

UJAEN

UEX

USAL

UVA



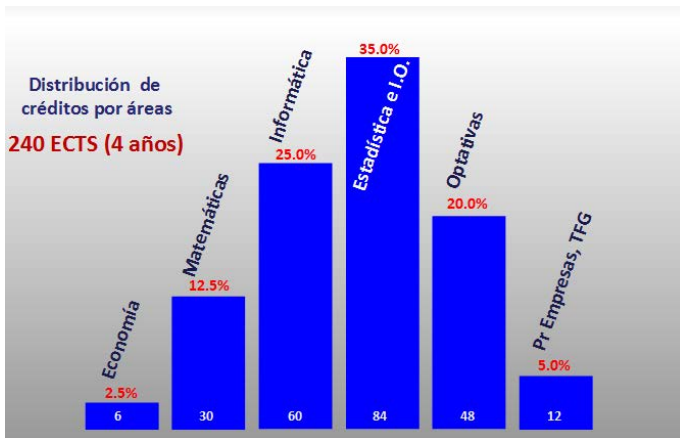


Tenemos enlaces de divulgación  
estadística desde: [www.eio.uva.es](http://www.eio.uva.es)

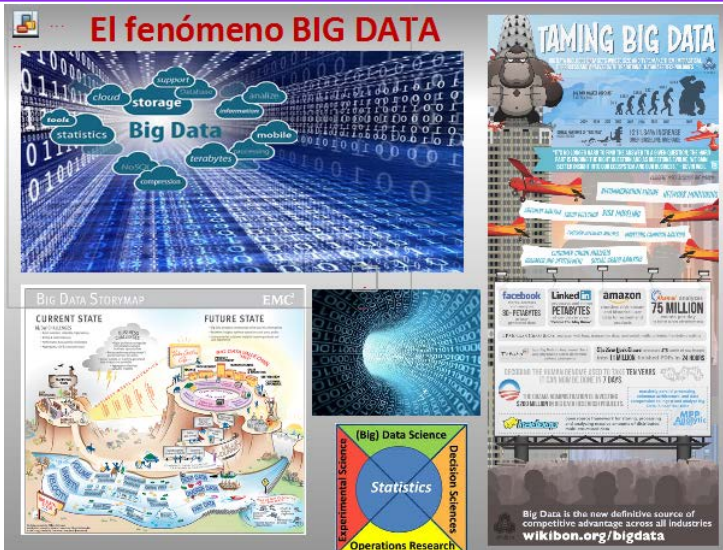


¡Síguenos en las redes sociales!

# Grado en Estadística en la UVa



# Nuevo reto ...





Grado en ESTADÍSTICA

4 años

Grado en INGENIERÍA INFORMÁTICA

4 años

Doble Titulación:

**ESTADÍSTICA +  
INFORMÁTICA**

**5 años**

# INdat

>estadística  
+informática

**Doble Grado: 5 años**

**Podrás trabajar en áreas como**

comercio

energía

farmacéuticas

industria

bioinformática

epidemiología

salud pública

logística

sistemas de información

asesoría TIC

analista financiero

meteorología

estudios de mercado

análisis de riesgos

comunicaciones

desarrollo de software

docencia

demografía

estadística pública

investigación científica

geología

ecología

seguridad de sistemas

bases de datos

medio ambiente

consultoría

genómica

bigdata

...

>informática  
+estadística

**...PRESENTE... y FUTURO!**

[www.inf.uva.es](http://www.inf.uva.es)

UvA

[www.eio.uva.es](http://www.eio.uva.es)  
[www.inf.uva.es](http://www.inf.uva.es)

UvA

Necesitamos: médicos, economistas, ingenieros, físicos, telecos, sociólogos, biotecnólogos, meteorólogos..., también filósofos, historiadores, geógrafos... Cualquier profesión es interesante y necesaria, y todas tienen en común el manejo y la producción de DATOS, de los que es preciso extraer conocimiento.

La ESTADÍSTICA junto a la INFORMÁTICA, son las herramientas básicas para el **Análisis de Datos**: La Informática aporta las tecnologías de las bases de datos y la computación, para aplicar de forma eficiente los métodos científicos de análisis que aporta la Estadística.

El doble título **INdat**, con un enfoque profesional, te proporcionará las herramientas y la formación necesarias para tu futuro desarrollo personal y profesional en el campo de la "Ingeniería de Datos".

El futuro de los profesionales de la estadística, y de **INdat** en particular, es muy prometedor. No existe desempleo entre los profesionales del sector y la demanda es creciente en todo el mundo, debido al auge del fenómeno **BigData** y sus aplicaciones.

Con **INdat** tendrás ambos títulos oficiales:

- Grado en Estadística y Grado en Ingeniería Informática
- Consta de 336 créditos, organizados en 5 años
- Número de plazas limitado
- Facultad de Ciencias y ETSI Informática (Valladolid)

[www.inf.uva.es](http://www.inf.uva.es)  
[www.eio.uva.es](http://www.eio.uva.es)

UvA

[www.inf.uva.es](http://www.inf.uva.es)  
[www.eio.uva.es](http://www.eio.uva.es)

UvA

[www.eio.uva.es](http://www.eio.uva.es)  
[www.inf.uva.es](http://www.inf.uva.es)

UvA

>informática  
+estadística

# INdat



# Asignaturas de Estadística en el Grado en la UVA

---


- Estadística descriptiva
- Modelos probabilísticos
- Probabilidad
- **Computación estadística**
- Inferencia estadística I, II
- Muestreo estadístico I, II
- Regresión y anova
- Análisis de datos
- Modelos lineales
- Análisis de datos categóricos
- Análisis de series temporales
- Análisis multivariante
- Modelos estadísticos avanzados
- Análisis de supervivencia
- Almacenes y minería de datos
- etc.

**Computación Estadística**

Statistical Computing

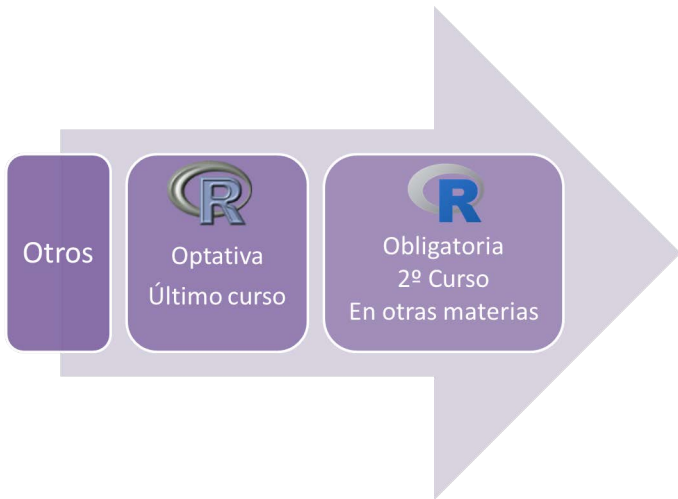
**Estadística computacional**

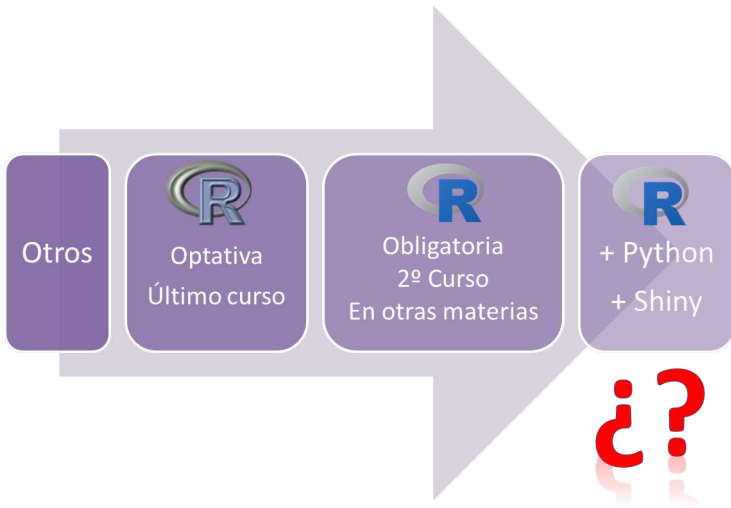
Computational Statistics

¿Por qué  ?

---







# Objetivos

---

- Trabajar sobre el razonamiento estadístico y computacional
- Fundamentos de la programación para el análisis de datos

- Lenguaje de comandos de R para análisis estadísticos
- Construir gráficos estadísticos elaborados
- Manejar recursos de ayuda disponible en R
- Análisis reproducible

- Explorar datos
- Manejar conceptos importantes de probabilidad e inferencia estadística mediante simulaciones



# Tópicos

---

- Entorno R
- Estructuras de datos y manipulación de objetos
- Gráficos
- Programación
- Entradas y salidas digitales
- Herramientas de presentación de informes



# Tópicos

## Estructuras de datos y manipulación de objetos

Vector, factor, matrix, array, data.frame, list, data(), attach, [, [[, \$, subset, split, cut, sort, which.min, which.max, which, dim(), cbind, rbind, c(), colnames, names, levels, labels, funciones matemáticas básicas, estadísticos básicos **apply, tapply, lapply, sapply**

## Creación de funciones propias

if else, ifelse, for, while, any, all, break, next, warning, stop, system.time

## Entradas y salidas digitales

source, scan, load, save, dump, write.table, read.table, pdf, png, SavePlot

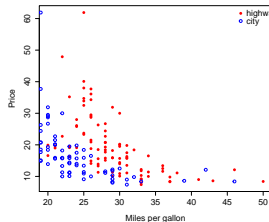
# Tópicos

## Gráficos tradicionales

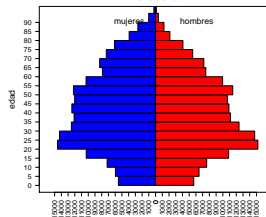
**Funciones gráficas:** plot, points, lines, text, mtext, legend, title, abline, segments, grid, axis, colors, palette, matplot, boxplot, stripchart, barplot, pie, mosaicplot, pairs, par, layout, split.screen

**Prámetros gráficos:** type, xlab, ylab, xlim, ylim, main, sub, add, col, lty, lwd, pch, cex

## Prices versus MPG



## Pirámide poblacional Valladolid



# Tópicos

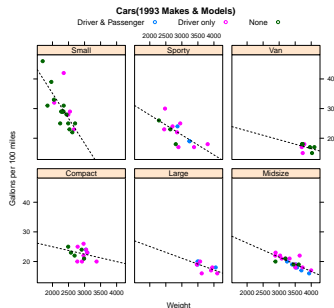
## Gáficos LATTICE

### Funciones alto nivel:

xyplot, histogram, density-plot, bwplot, barchart, dotplot, ...  
make.groups, equal.count, trellis.par.get, ...

### Funciones bajo nivel y parámetros:

type, autokey, layout, groups, strip, between, panel, par.settings, panel.xyplot, panel.abline, panel.lmline, panel.rug, panel.grid, ...



## Sobre metodología

---

- Exposiciones cortas para introducir los distintos conceptos
- Realizar ejercicios y prácticas
- Explorar y sacar partido de los recursos online

- Se trata de aprender haciendo
- Centrarnos en enseñar cómo *desarrollar nuevo código* en lugar de sólo usar plantillas o llamar funciones ya existentes

## Sobre metodología

---

- Exposiciones cortas para introducir los distintos conceptos
- Realizar ejercicios y prácticas
- Explorar y sacar partido de los recursos online

- Se trata de aprender haciendo
- Centrarnos en enseñar cómo *desarrollar nuevo código* en lugar de sólo usar plantillas o llamar funciones ya existentes

No la librería *Rcommander*, hasta el final.



[Home]

## Download

CRAN

## R Project

About R  
Contributors  
What's New?  
Mailing Lists  
Bug Tracking  
Conferences  
Search

## R Foundation

Foundation  
Board  
Members  
Donors  
Donate

## Documentation

Manuals

# The R Project for Statistical Computing

## Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To [download R](#), please choose your preferred [CRAN mirror](#).

If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

## News

- [R version 3.2.2 \(Fire Safety\)](#)
- [The R Journal Volume 7/1](#)
- [R version 3.1.3 \(Smooth Sailing\)](#)
- [useR! 2015](#), will take place at the University of Vienna
- [useR! 2014](#), took place at the University of Vienna



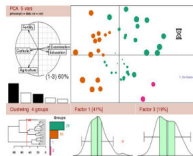
About R  
What's New?  
Contributors  
Conferences  
What's new?

Download, Package  
CRAN

R Project  
Foundation  
Windows & MacOS  
Mailing Lists  
Bug Tracking  
Conferences/Events  
Conferences  
Search

Documentation  
Manuals  
FAQs  
The R Journal  
Web  
Books  
Contributors  
Index

The R Project for Statistical Computing



### Getting Started:

- For a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To [download R](#), please choose your preferred [CRAN mirror](#).
- If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.



```
RGui (64-bit)
File Edit Packages Windows Help

R Console
R version 3.1.0 (2014-04-10) -- "Spring Dance"
Copyright (C) 2014 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]

> |

R F:\CE\CE_2014_2015\Tema 2 programa R\ordenes-diapositivas-8-lattice.R -
barchart(VADeaths, groups = FALSE)

barchart(VADeaths, groups = FALSE, layout = c(4, 1))

dotplot(VADeaths)
dotplot(VADeaths, groups = FALSE, layout = c(4, 1))
dotplot(VADeaths, groups = T, layout = c(4, 1))

dotplot(VADeaths, type = "c", auto.key = list(points = TRUE, lines
dotplot(VADeaths, type = "c", auto.key = list(points = TRUE, lines
dotplot(VADeaths, type = "c", auto.key = list(points = TRUE, lines

#####

data(Earthquake, package = "nlme")

help(Earthquake, package = "nlme")

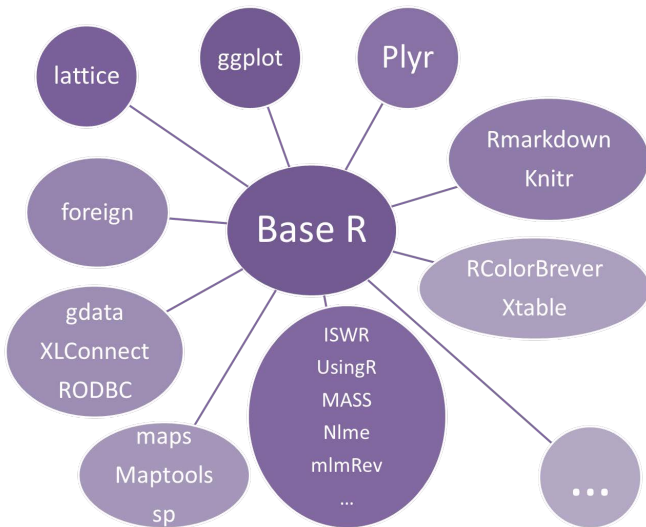
xyplot(accel ~ distance, data = Earthquake)

xyplot(accel ~ distance, data = Earthquake, scales = list(log = TR
xyplot(accel ~ distance, data = Earthquake, scales = list(log = TR
xyplot(accel ~ distance, groups=soil, data = Earthquake, scales = 1.

<
```

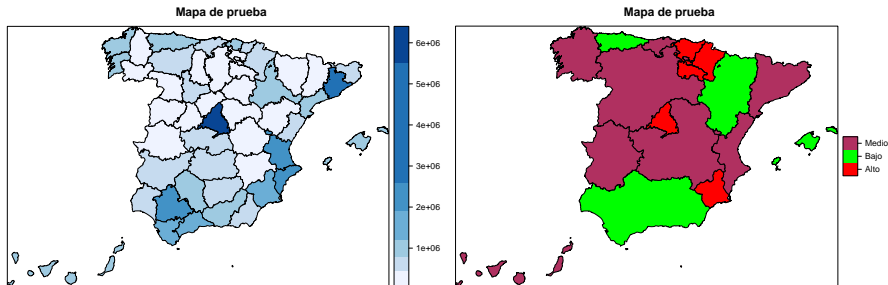




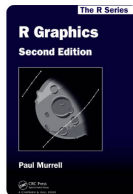
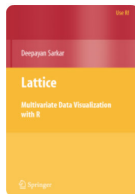
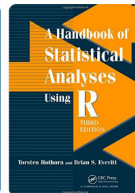
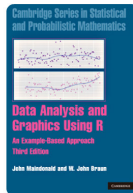
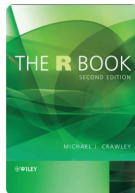
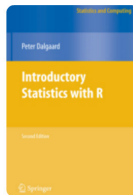
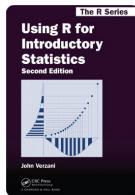


# Ejemplo

¿Es fácil realizar mapas (cartogramas) de este tipo con R? ¿Qué librerías son necesarias?



## Algunos libros útiles



## Ayuda de los paquetes (enlaces, viñetas)

An implementation of the grammar of graphics in R. It combines the advantages of both base and lattice graphics: conditioning and shared axes are handled automatically, and you can still build up a plot step by step from multiple data sources. It also implements a sophisticated multidimensional conditioning system and a consistent interface to map data to aesthetic attributes. See <http://ggplot2.org> for more information, documentation and examples.

Preference manual: [rustc2.pdf](#)  
 Vignettes: [Contribution to rustc2 development](#)  
 Package source: [rustc2-1.0.1.tar.gz](#)  
 Windows binaries: r-devel: [rustc2-1.0.1.exe](#) r-release: [rustc2-1.0.1.exe](#) r-oldrel: [rustc2-1.0.1.exe](#)  
 OS X 64-bit Leopard binaries: r-release: [rustc2-1.0.1.pkg](#) r-oldrel: [rustc2-1.0.1.pkg](#)  
 OS X 32-bit binaries: r-release: [rustc2-1.0.1.pkg](#)  
 Old versions: [rustc2\\_archive](#)

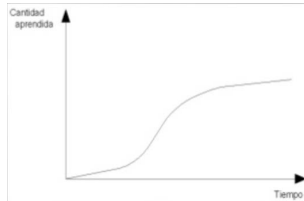
[illegible]

## R graph gallery

The blog is a collection of script examples with example data and output plots. R produce excellent quality graphs for data analysis, science and business presentation, publications and other purposes. Self-help codes and examples are provided. Enjoy nice graphs !!

[illegible]

## ¿Qué he aprendido fruto de mi experiencia personal?



- Curva de aprendizaje relativamente empinada. Debemos advertirlo e intentar evitar la frustración.
- **Aprender haciendo**



## Comentarios finales

---

- El objetivo principal: promover competencias fundamentales para los profesionales de la Estadística:
- Al principio el aprendizaje de R es muy “duro”. Luego se pueden abordar contenidos más avanzados más fácilmente.
- Enfatizar características muy relevantes del R: potencia, flexibilidad y extensibilidad.

# Cuestiones abiertas

---

- ¿Qué herramientas enseñamos para realizar gráficos?
  - Gráficos tradicionales, latttice
  - ggplot2, gmap, sp
  - ggvis, googleVis, otras librerías de gráficos interactivos
  - ...
- ¿Otras librerías instrumentales?
  - odfWeave
  - ...
- ¿Enseñamos a crear librerías?
- ¿Librerías para construir GUIs sencillos?





¡Gracias!