



Estadística y gráficos con *R*

Colaboración entre PDI (1)

* PDI = Personal Docente e Investigador

Dr. Xavier de Pedro
Dep. de Física Aplicada y Óptica
y
Grupo por el Conocimiento Libre en la UB (gclub)
Universidad de Barcelona

xavier.depedro@ub.edu - xavi@member.fsf.org

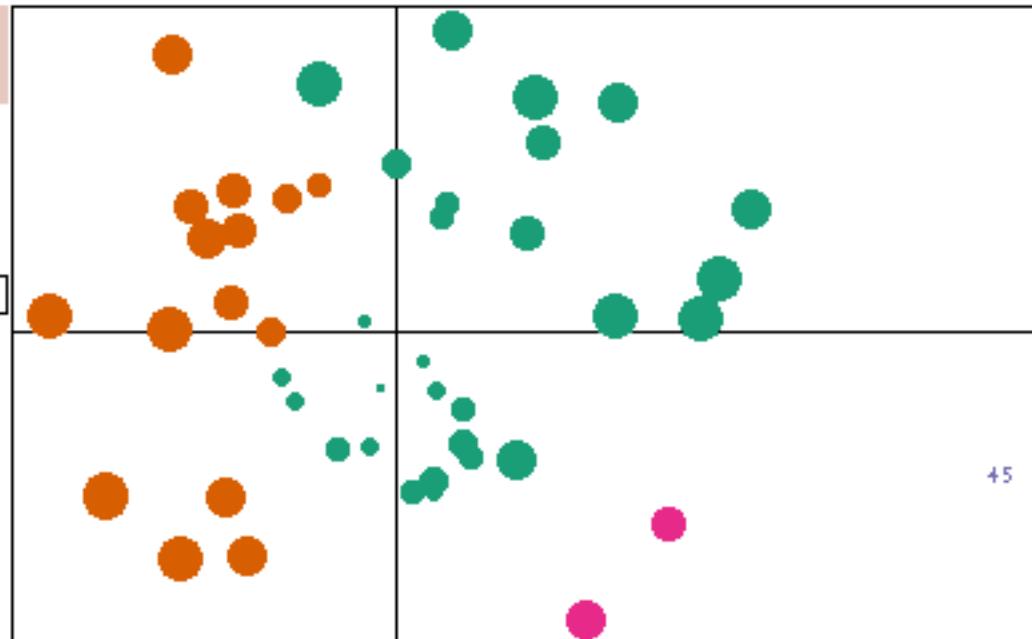
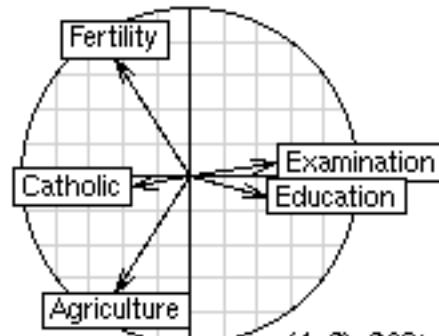


<http://www.ub.edu/gclub>

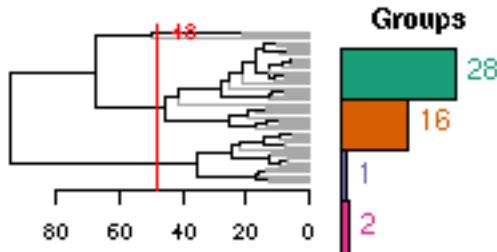


PCA 5 vars

```
princomp(x = data, cor = cor)
```

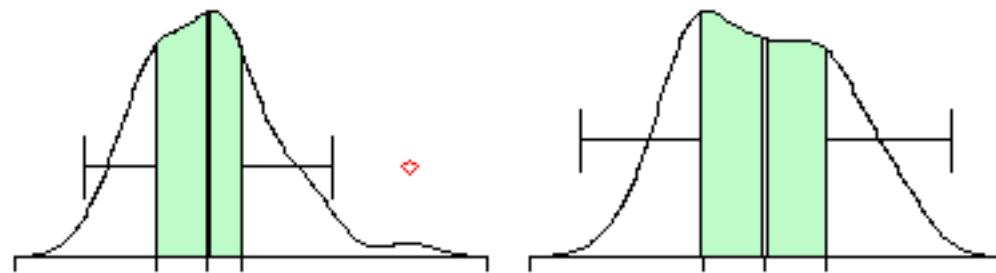


Clustering 4 groups



Factor 1 [41%]

Factor 3 [19%]





R: ¿Qué es?

- Programa para Estadística y gráficos científicos
- Alternativa libre al “S” (y Matlab, SPSS...)
- Importa fácilmente de Excel, .csv, ...
- Gran número de extensiones
- Funciona por comandos, y algunas Interfícies Gráficas de Usuario (“GUI's”)

<http://cran.r-project.org>





R: ¿Qué es? (ii)

- Ejemplo: uso por consola

```
xavi@ward: ~/MeusDocuments/Models/ommel/ □ □ X
Fitxer Edita Visualitza Terminal Pestanyes Ajuda
Calculating mass balances...
Saving results into output files (*.out)...
xavi@ward:~/MeusDocuments/Models/ommel/bin$ R

R version 2.5.1 (2007-06-27)
Copyright (C) 2007 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

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.

Natural language support but running in an English locale

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.

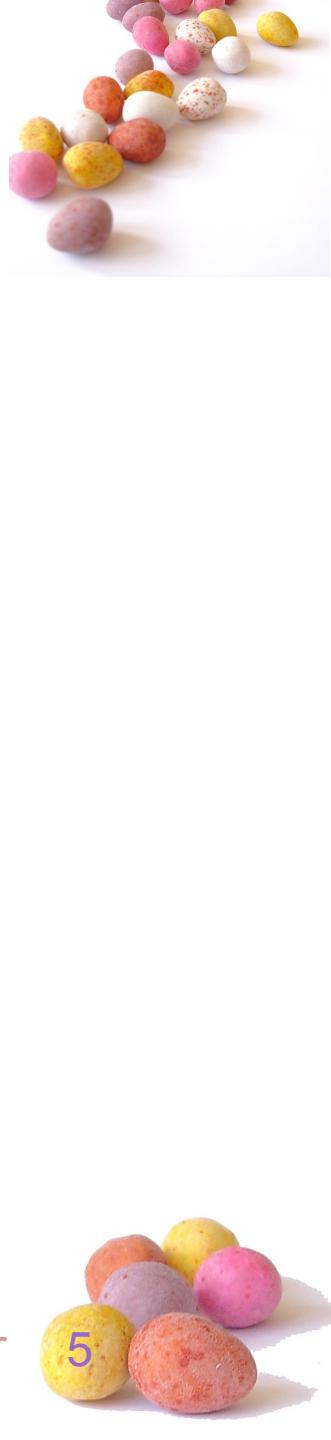
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> setwd("../")
> source("gommel.r")■
```



R: ¿Qué es? (iii)

- Ejemplo: script .r



A screenshot of a KDE desktop environment. In the bottom right corner, there is a decorative icon of several colorful Easter eggs (pink, yellow, orange) on a white surface. The main focus is a terminal window titled "gommel_all.r [modificat] - Kate". The window has a menu bar with Spanish labels: Fitxer, Edita, Visualitza, Vés, Document, Punts, Sessions, Finestra, Eines, and a separator. Below the menu is a toolbar with icons for Nou (New), Obre (Open), Enrere (Back), Endavant (Forward), Desa (Close), and Desa com a (Close All). On the left side of the terminal window, there is a sidebar titled "Navegador del sistema de fitxers" (File system navigator) which shows a tree view of directory structures. The terminal itself contains the following R script code:

```
## gommel_all.r
# Copyleft 2007 Xavier de Pedro (GNU/GPL'd)
# (1) To specify common settings for all graphs
# (2) To plot output from ommel model 2D, v1.7 onwards

# Clear all objects from R workspace memory
rm(list=ls())

# (1) Settings for all graphs
# -----
# Name and version number of the program to generate the files with data
ommel_version="OMMEL v1.9675"

# Version number of the set of gommel R scripts
ommel_r_scripts_set_version="0.12"

# Specify where to send the graphs to (0=screen, 1=.png files, 2=both)
ommel_send_graphs_to=2
|
# Subdirectory from "ommel" base directory where the desired output files
# will be generated
ommel_out_dir="graphs"

# Set the working directory to the ommel directory
setwd("ommel")
```

The terminal status bar at the bottom shows "Línia: 19 Col.: 1 INS LÍNIA gommel_all.r" and a "Terminal" icon.

de <http://ommel.sf.net> con las equivalencias de scripts .m y .r

Cálculos estadísticos y programación



- Documentos base:
 - Introducción a R (2000): A. y S. González
<http://cran.r-project.org/doc/contrib/R-intro-1.1.0-espanol.1.pdf>
 - R para principiantes (2002). E. Paradis
http://cran.r-project.org/doc/contrib/rdebuts_es.pdf
- Conversión de lenguajes
 - Matlab/Octave | Python | R
<http://mathesaurus.sourceforge.net/matlab-python-xref.pdf>

3 Vectors

Desc.
Row vector, $1 \times n$ -matrix
Column vector, $m \times 1$ -matrix

MATLAB/Octave
a=[2 3 4 5];
adash=[2 3 4 5]';

Python
a=array([2,3,4,5])
array([2,3,4,5])[:,NewAxis]
array([2,3,4,5]).reshape(-1,1)
r_[1:10,'c']

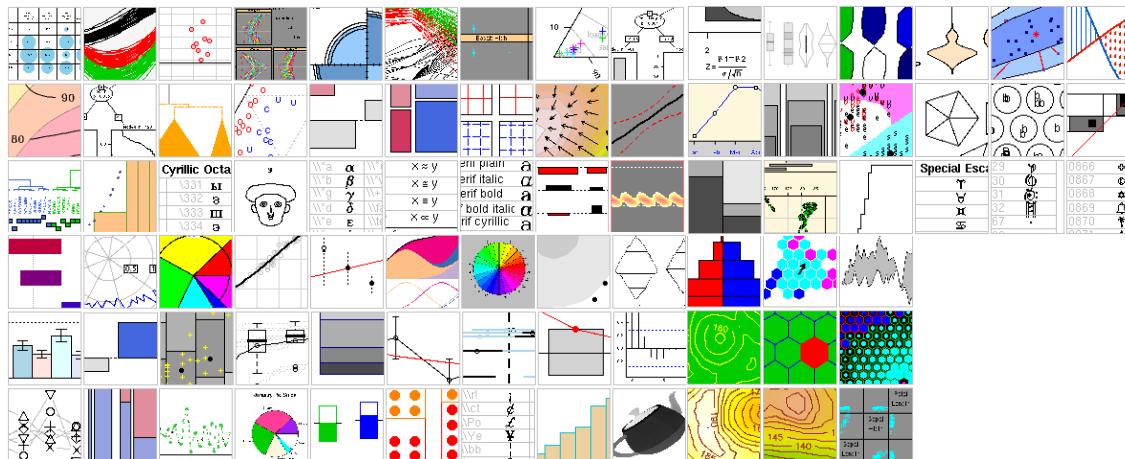
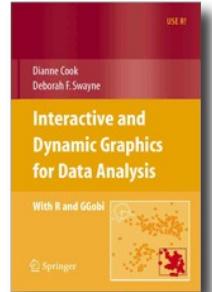
R
a <- c(2,3,4,5)
adash <- t(c(2,3,4,5))



Gráficos



- Libros:
 - *R Graphics* (2005) P. Murrel
<http://www.stat.auckland.ac.nz/~paul/RGraphics/rgraphics.html>
 - *The GGobi Book (2007): Interactive and Dynamic Graphics for Data Analysis: with examples using R and Ggobi.* Springer. D. Cook & D. F. Swayne. <http://www.ggobi.org/book/>
- Galeria colaborativa de gráficos



Gráficos (ii)

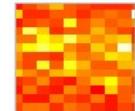
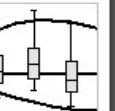
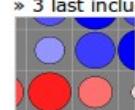
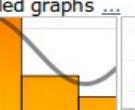
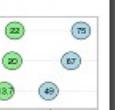
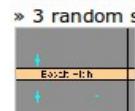
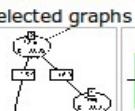
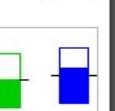
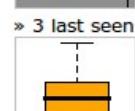
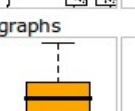
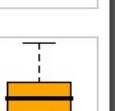
read my blog [RELATED SITES: R-project](#) | [CRAN](#) | [Bioconductor](#) | [R Movies Gallery](#) | [R Wiki](#) | [search :](#)

Graph Gallery : Enhance your data visualization with R

Home **Browse** **Related** **Source code** **Graphics List** **Thumbnails**

a agreement analysis and association back bar barplot boxplot boxplots chart classification cluster color colored colors conditional conditionnal conditioning contour coplot correlation cumulative curve curves **d** dem dendrogram density diagram distribution double ellipses escape estimator extended filled fonts for function geographic hershey hexbin highest histogram in kernel lattice map mathematical matrices matrix maunga model more mosaic **of** parallel perspective pie **plot** plots plotting quiver r regions regression rgb roc rose sample scatter scatterplot seasonal **sequences** simple som space special splom teapot ternary **the** tree use vector volcano whau with

Enter the gallery

- » 3 best ranked graphs ...
-   
Cadillac Fleetwood
- » 3 last included graphs ...
-   
- » 3 random selected graphs
-   
- » 3 last seen graphs
-   

Galeria colaborativa de gráficos: <http://addictedtor.free.fr/graphiques/>

GUI's (i): GNU/Linux



RKWard

<http://rkward.sf.net>

The screenshot shows the RKWard application window. The main area is a code editor titled "rkward_for_new_users" containing an R script named "gommel_all.r". The script includes comments specifying the OMMEL version (v1.9673), output settings (to screen or .png), and a comment code for .png files. It also sources three R scripts: "gdo2Db.r", "gdo2Dm.r", and "gdo2Dk.r". Below the code editor is a terminal window showing the command "getwd()" and its output, followed by "gommel.r" and "source("gommel.r")". On the left, there's a file browser titled "Espai de treball" showing various files and folders related to the "ommel" project. At the bottom, there are tabs for "Bitàcola de les ordres", "Treballs pendents", "Consola R", and "Busca ajuda". The status bar at the bottom right indicates "El motor R està inactiu".

```
1 # gommel_all.r
2 # Copyright 2007 Xavier de Pedro (GNU/GPL'd)
3 # (1) To specify common settings for all graphs
4 # (2) To plot output from ommel model 2D, v1.7 onwards
5
6 # (1) Settings for all graphs
7 # -----
8 # Name and version number of the program
9 ommel.version="OMMEL v1.9673".
10 # Specify whether output is driven to screen only (0) or to .png
11 # files (1)
12 ommel.graphs.on.files=0
13 # Subdirectory from "ommel" base directory where the desired output
14 # files generated by ommel are stored, to be read by r scripts.
15 # If ommel.graphs.on.files=1, the png are stored also in this
16 # directory
17 ommel.path.to.output.files=".binaries_windows/"
18 # Show line of text with comments on top of all graphs (1=yes/0=no)
19 ommel.show.comments=1
20 # Set a short comment code or word to be included in all .png file
21 # names and figures on screen
22 ommel.comment.on.graph.set="ommel_linux"
23
24 # R Scripts to plot OMMEL results
25 source("./R_script/gdo2Db.r")
26 source("./R_script/gdo2Dm.r")
27 source("./R_script/gdo2Dk.r")
```

> getwd()
[1] "/media/USBDISK/ommel"
> gommel.r
> source("gommel.r")
> |

Bitàcola de les ordres Treballs pendents Consola R Busca ajuda

Llistat. /media/USBDISK/ommel El motor R està inactiu



GUI's (ii): Windows

- (A) SciViews

<http://www.sciviews.org>

- (B) RKWard

(Junio'09)

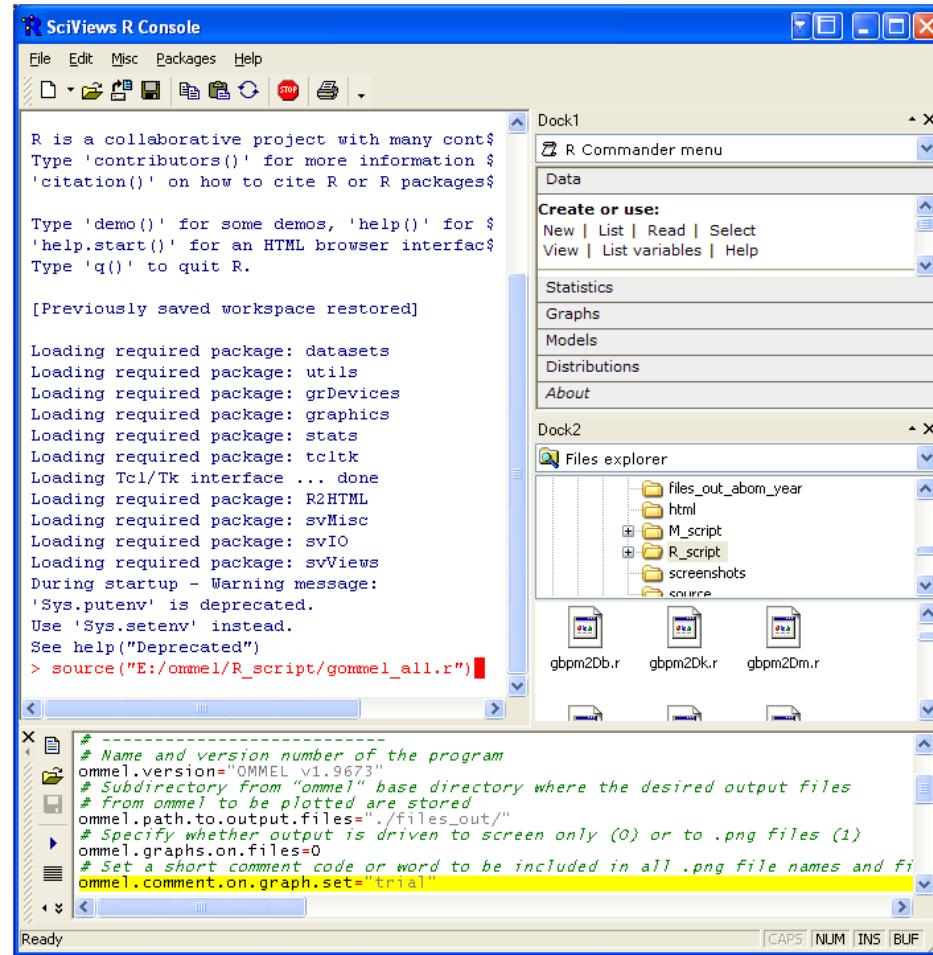
(B1) sobre WinKDE (.exe)

winkde.org + rkward.sf.net

ó

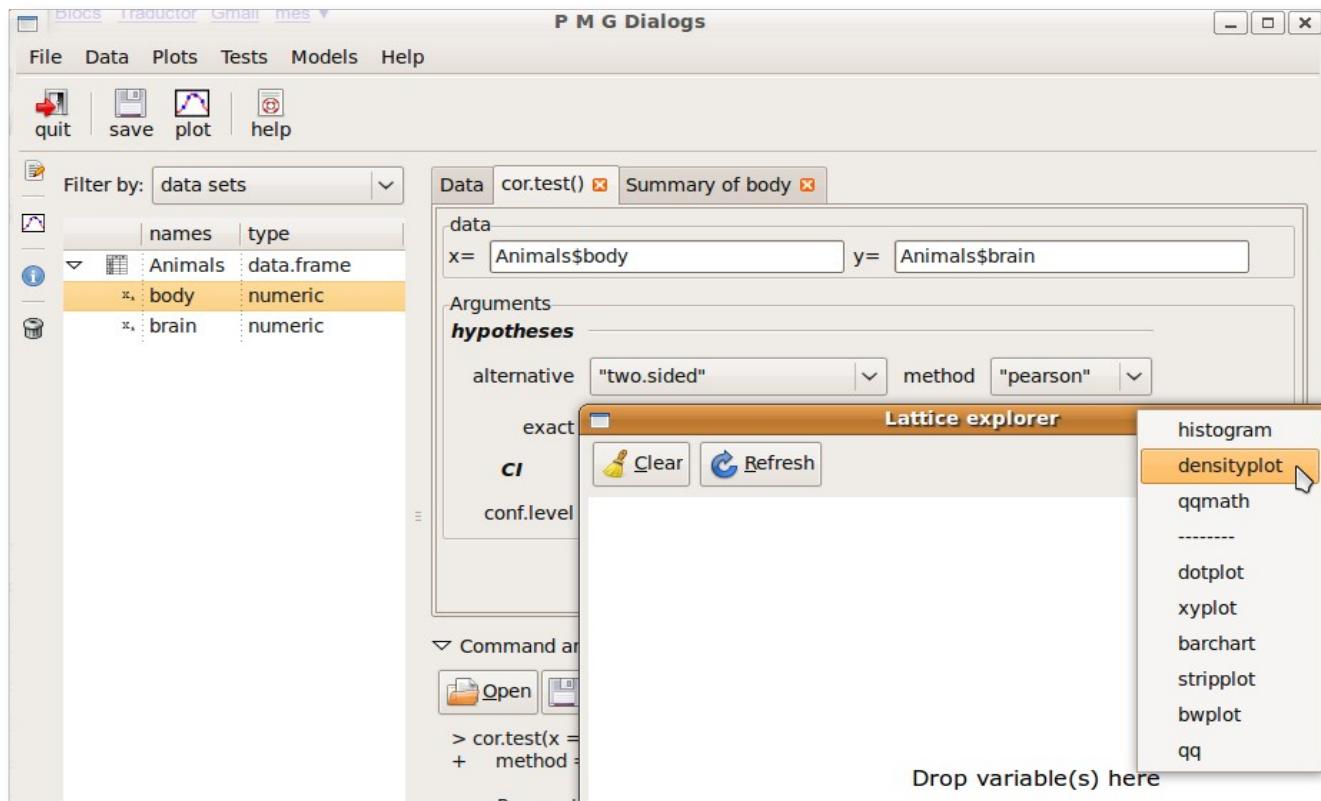
(B2) via “andLinux” (.deb)

<http://andlinux.org>



GUI's (iii): Multiplataforma

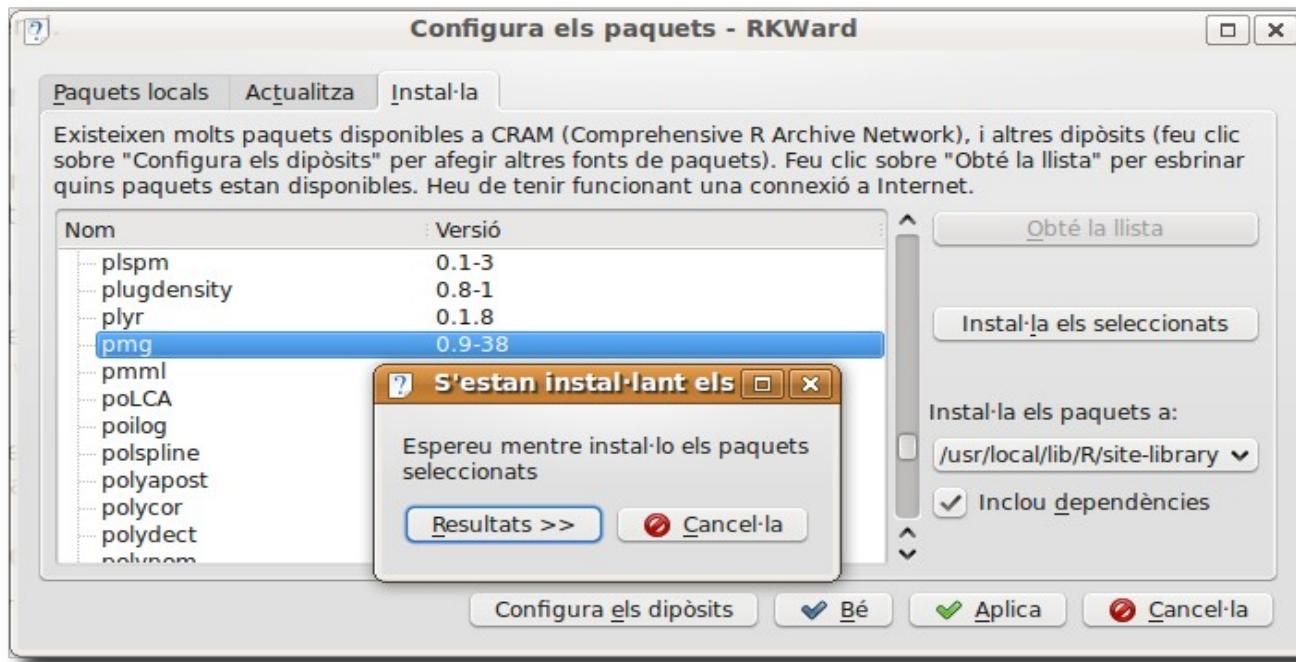
- Java (Jaguar) y R GUI (Rcmdr, PMG, ...)
 - PMG = Poor Man GUI (RGTK2 (gWidgets))



Instalación de paquetes



- Listado enorme, y creciendo año tras año



Iniciativas en REDIRIS



- **IRIS-LIBRE: Proyecto “eReros”**

“Promocionar el uso de R en las Universidades españolas y colaboración entre usuarios”

<https://forja.rediris.es/pipermail/ereros-general/>

- Web preparada y congreso en preparación
- Propuestas:
 - Cluster alta disponibilidad para cálculos mayores
 - Formación de R via CENATIC-CRUE en unis
 - Desarrollo 2009: alternativa a otras aplicaciones comerciales (stata, statgraphics, sas, ...)



Alternativas a R



Libres vs. *versión propietaria*

- **GNU Octave**
 - <http://www.gnu.org/software/octave/>
- **GNU PSPP**
 - <http://gnu.org/software/pspp/>
- **Scilab**
 - <http://scilab.org>
- **Sage**
 - <http://sagemath.org>
- *Matlab,*
Mathematica
- *SPSS*
- *Matlab,*
Mathematica
- *Magma, Maple,*
Mathematica and
Matlab



Gracias



- Presentación en:
<http://www.ub.edu/gclub/jornadespl2009>
- Referencia:
 - *Colaboración entre PDI (1): Estadística y Gráficos con R.* 2009. Xavier de Pedro. VIII Jornades de Programari Lliure. 1 - 4 Juliol. Free Software, Open Source i Coneixement Lliure. <http://jornadespl.org>

