



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



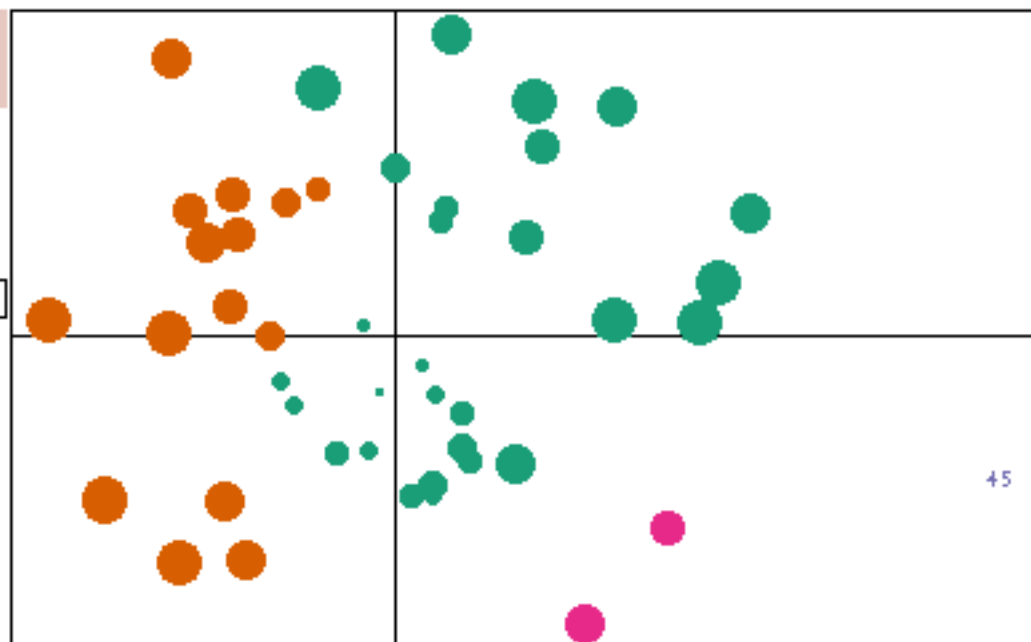
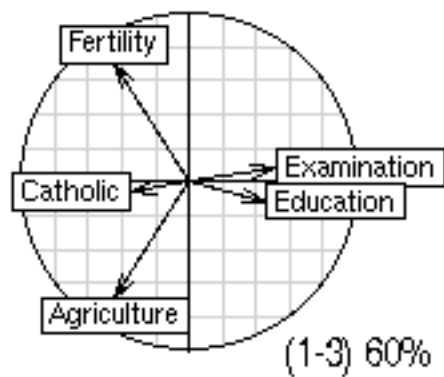
<http://www.ub.edu/gclub>



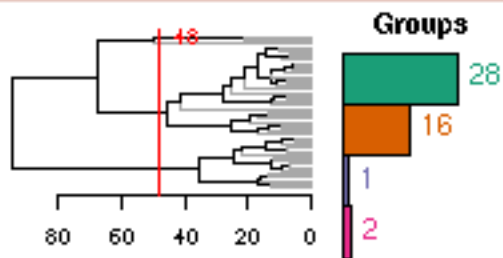


PCA 5 vars

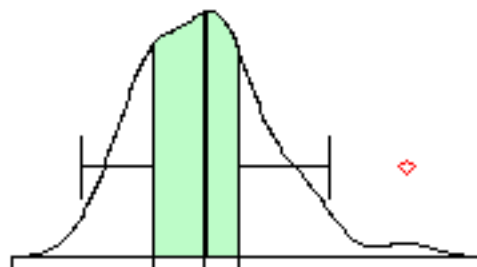
`princomp(x = data, cor = cor)`



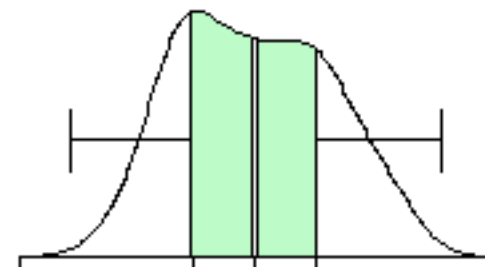
Clustering 4 groups



Factor 1 [41%]



Factor 3 [19%]



45

2



R: ¿Qué es?



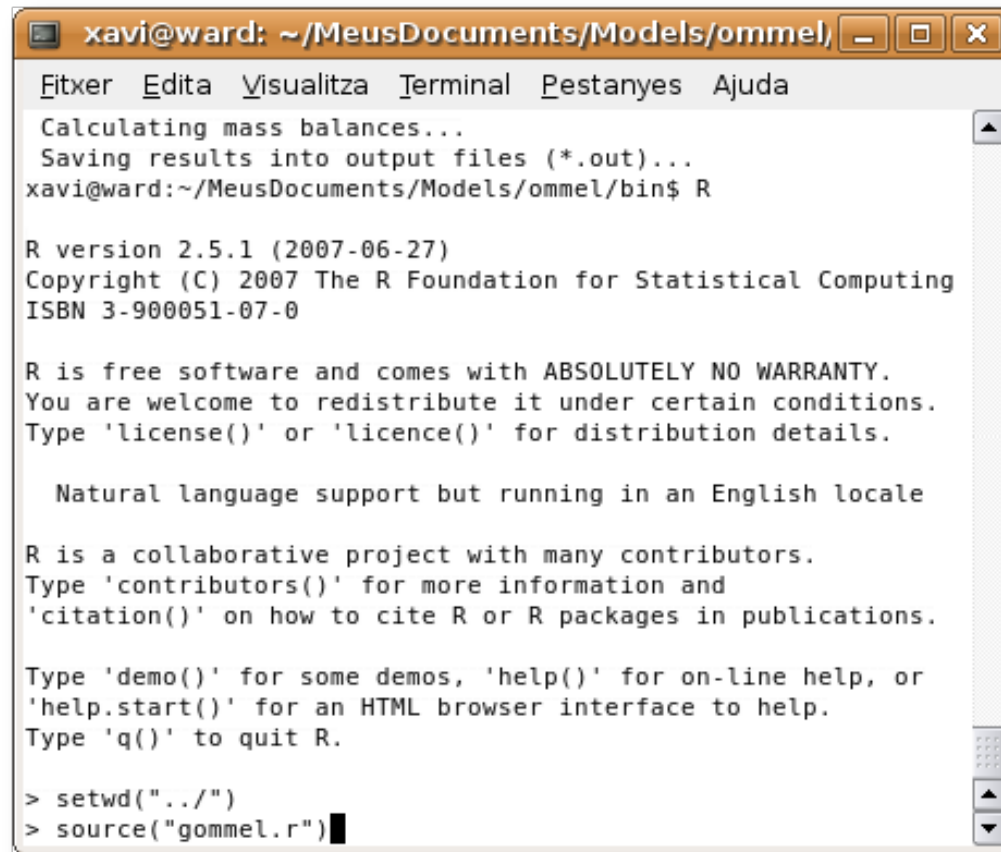
- Programa para Estadística y gráficos científicos
- Alternativa libre al “S” (y Matlab, SPSS...)
- Importa fácilmente de Excel, .csv, ...
- Gran número de extensiones
- Funciona por comandos, y algunas Interficies Gráficas de Usuario (“GUI's”)

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



R: ¿Qué es? (ii)

- Ejemplo: uso por consola



```
xavi@ward: ~/MeusDocuments/Models/ommel/
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

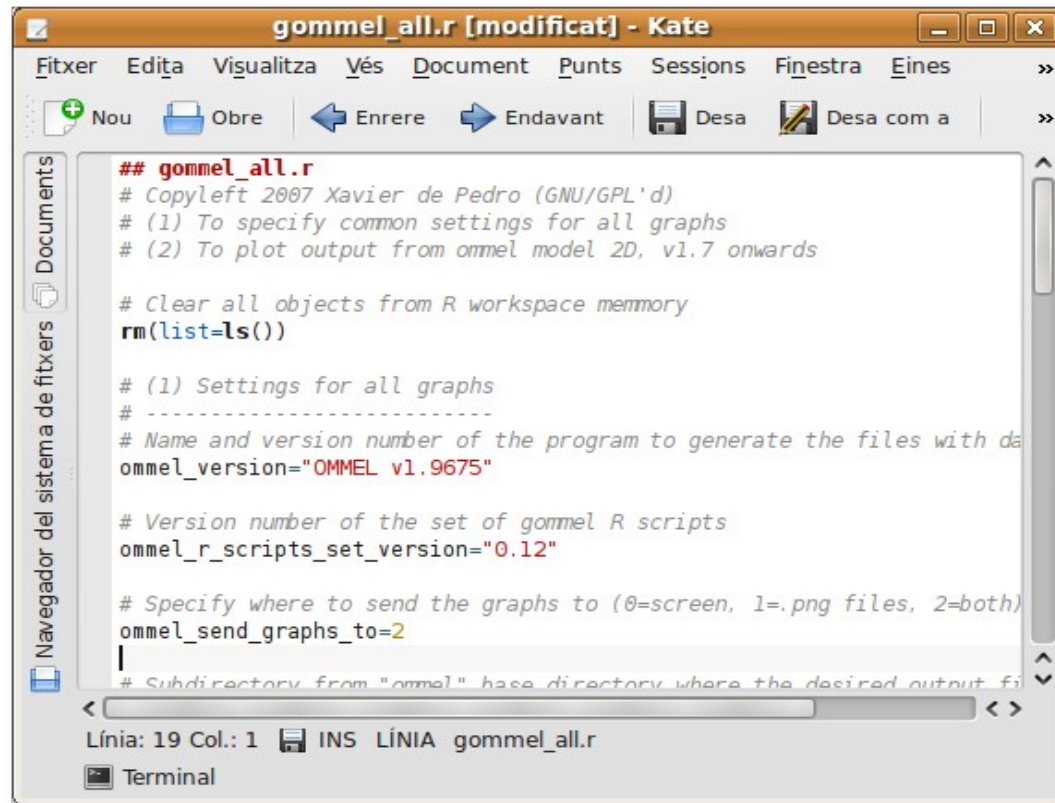
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

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



```
## 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
```

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

Dese.
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])[::-1]`
`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

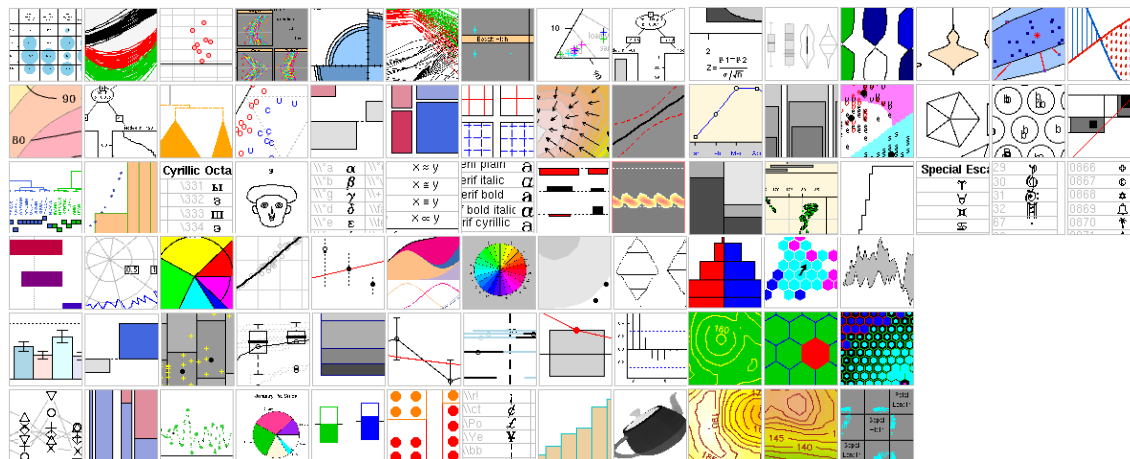
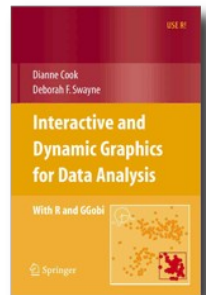
- Libros:

- *R Graphics* (2005) P. Murrell

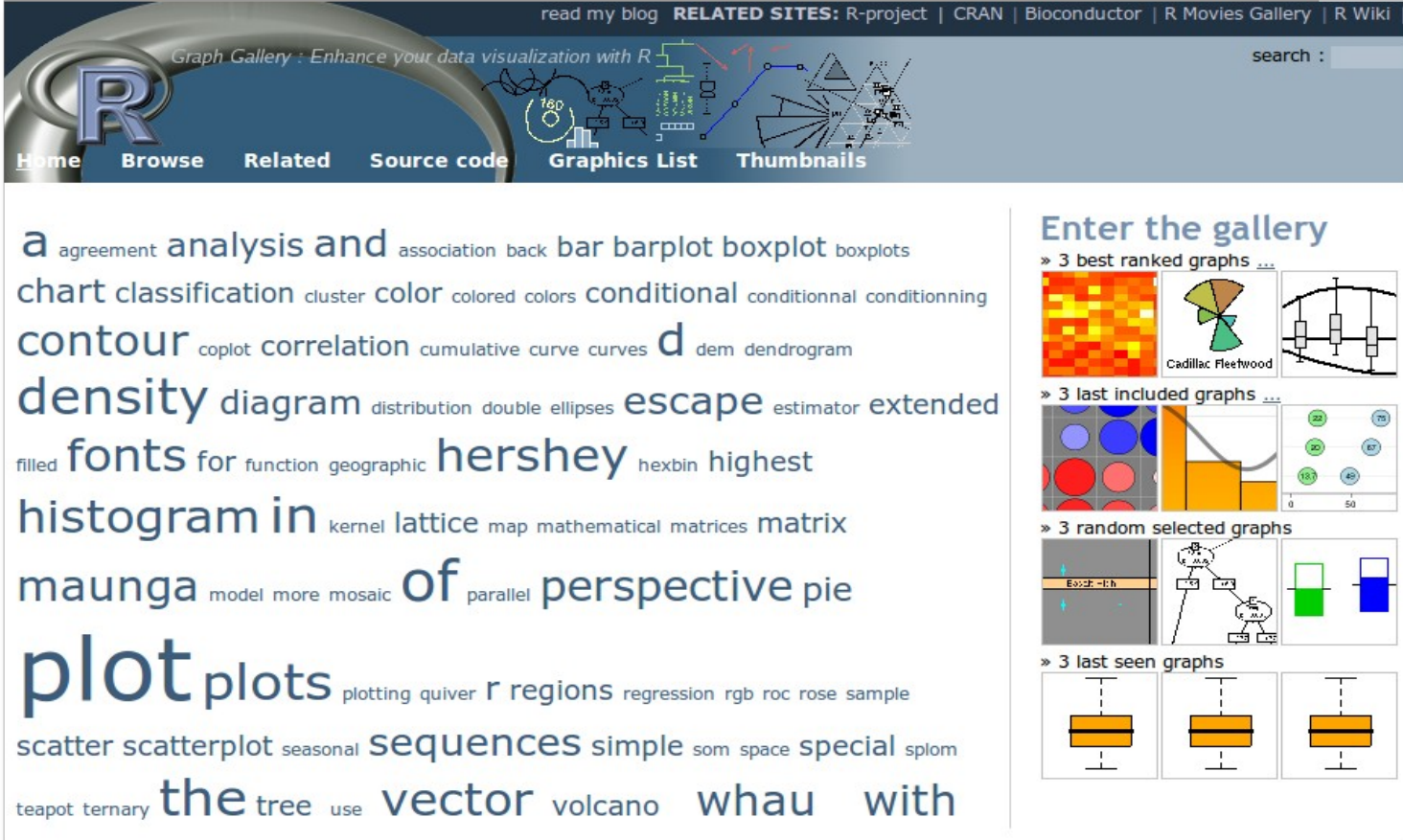
- <http://www.stat.auckland.ac.nz/~paul/RGraphics/rgraphics.html>

- *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)



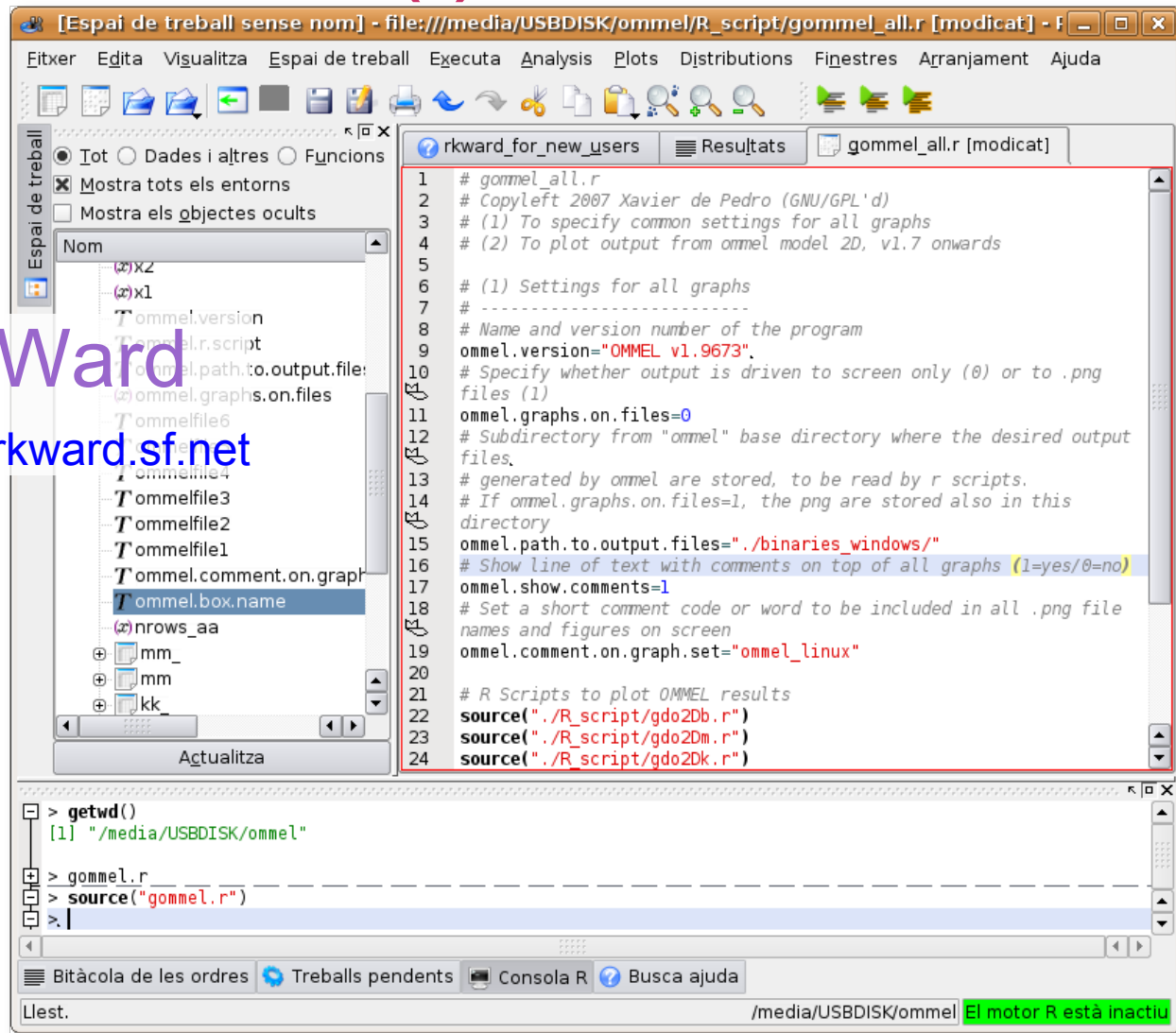
The screenshot shows the R Graph Gallery website. At the top, there's a navigation bar with links: Home, Browse, Related, Source code, Graphics List, and Thumbnails. A search bar is on the right. The main content area features a large word cloud of R plot types, including 'a', 'agreement', 'analysis', 'and', 'association', 'back', 'bar', 'barplot', 'boxplot', 'boxplots', 'chart', 'classification', 'cluster', 'color', 'colored', 'colors', 'conditional', 'conditionnal', 'conditionning', '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', and 'with'. On the right side, there's a section titled 'Enter the gallery' with four sub-sections: '3 best ranked graphs', '3 last included graphs', '3 random selected graphs', and '3 last seen graphs'. Each sub-section displays a grid of various R plots, including heatmaps, bar charts, boxplots, and network diagrams.

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

GUI's (i): GNU/Linux



- RKWard
<http://rkward.sf.net>



GUI's (ii): Windows

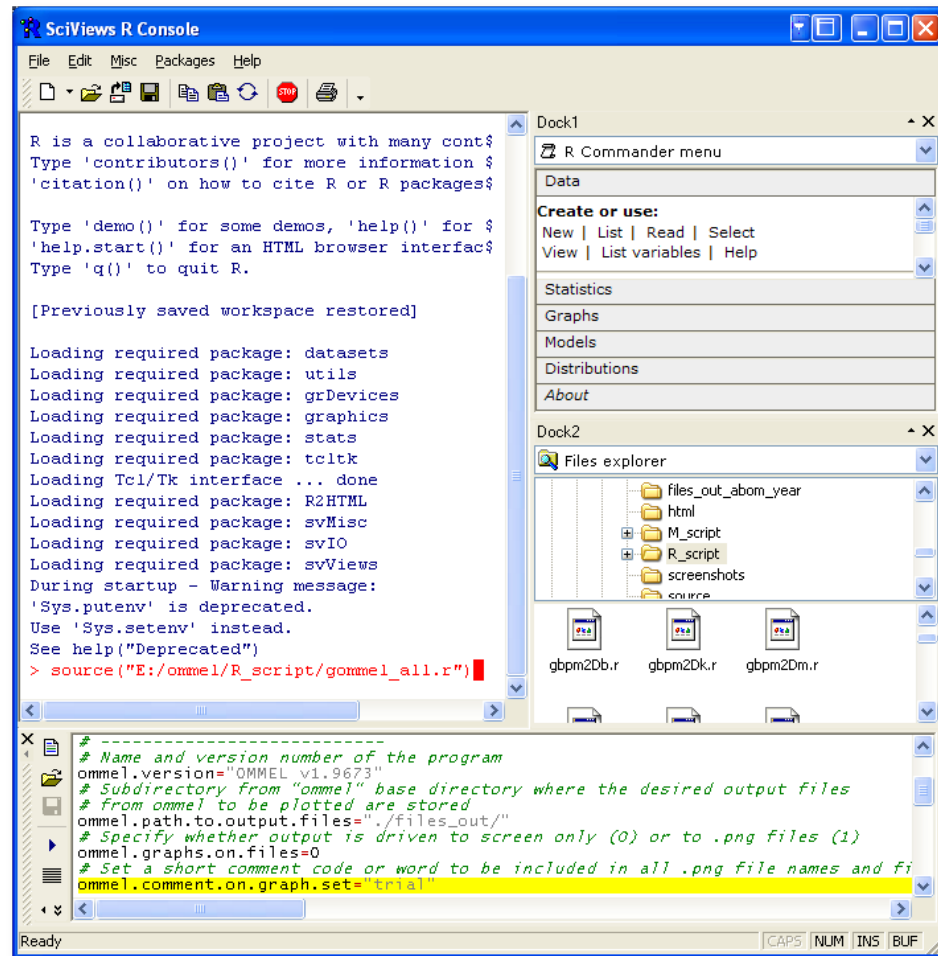
- (A) SciViews
<http://www.sciviews.org>

- (B) RKWard
(Junio'09)

(B1) sobre WinKDE (.exe)
winkde.org + rkward.sf.net

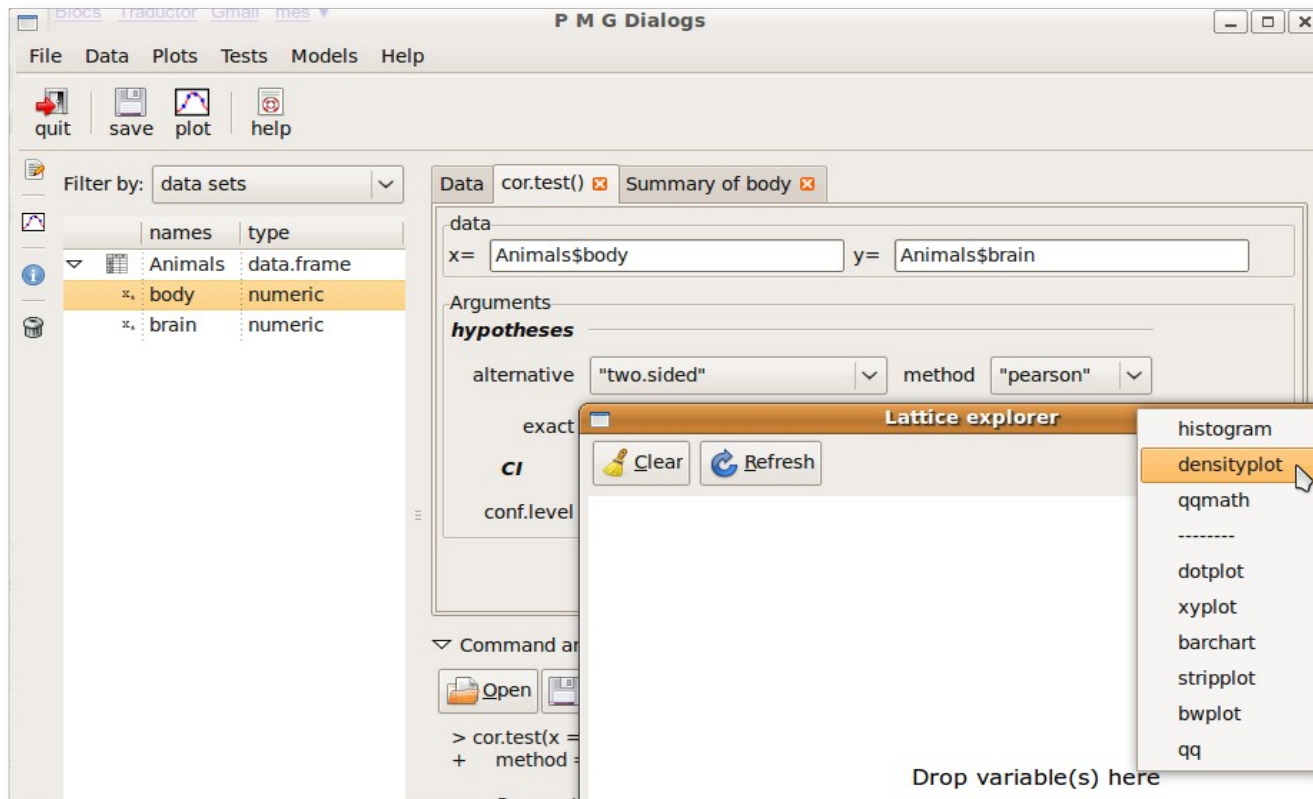
ó

(B2) via “andLinux” (.deb)
<http://andlinux.org>



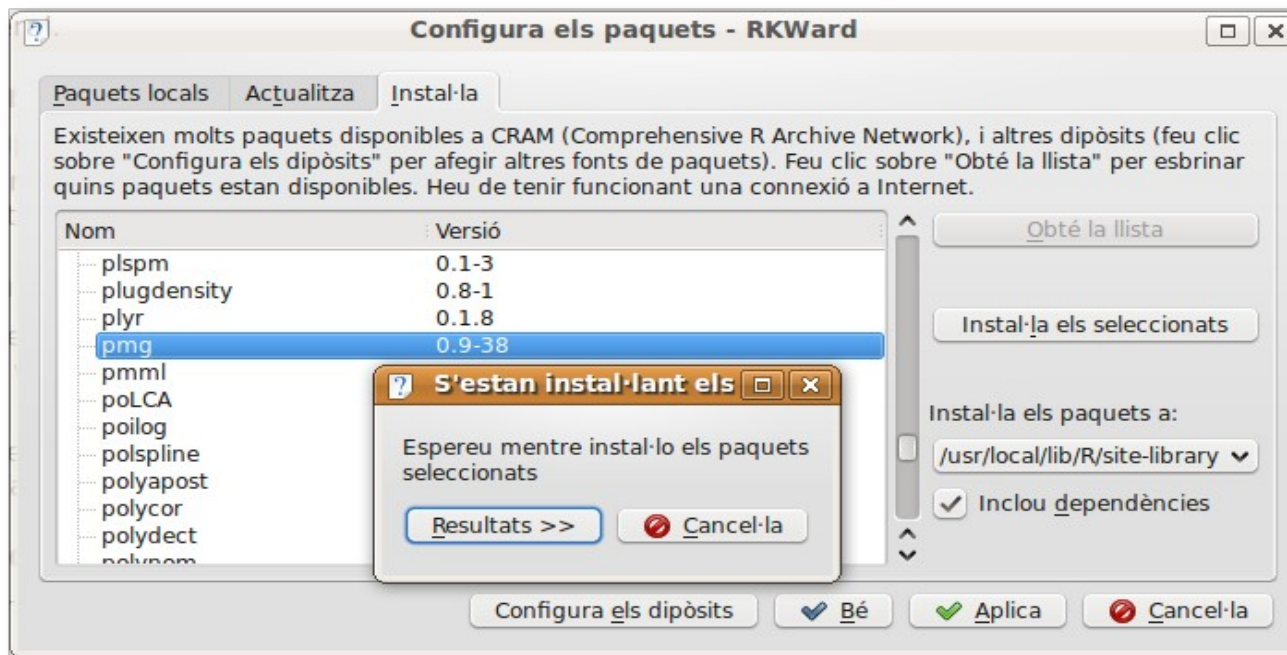
GUI's (iii): Multiplatforma

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



Instal·lació de paquets

- 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>

