

# O ambiente de trabalho R

Rui Paulo, ISEG e CEMAPRE

Colóquio do Departamento de Matemática do ISEG

8 de Março de 2013

# Sumário

1. O que é o R?
2. Software livre
3. Vantagens
4. Packages
5. Documentação
6. Produção de gráficos
7. Curta visita guiada

# O que é o R?

WWW.R-PROJECT.ORG — *The R Project for Statistical Computing*

- Ambiente de trabalho onde cálculos e gráficos estatísticos podem ser efectuados;
- armazenamento e manipulação de dados;
- manipular *arrays*, e em particular matrizes;
- grande variedade de métodos estatísticos, e.g. regressão linear, modelos lineares generalizados, séries temporais, classificação, testes, etc;
- produção de gráficos de qualidade profissional, quer no ecrã quer em ficheiros, tendo o utilizador controle sobre todos os parâmetros;
- linguagem de programação simples e intuitiva, apresentando todas as usuais estruturas: ciclos, condicionais, input/output, etc.

# Software livre

O R é um exemplo de *free software* — software livre

- Significa bastante mais do que ser distribuído gratuitamente:  
*free as in “free speech” and not as in “free beer”*.
- Direito a aceder ao código-fonte do software, a alterá-lo e a redistribuí-lo — GNU General Public License, (GPL).
- – Free Software Foundation [WWW.FSF.ORG](http://WWW.FSF.ORG)
  - Open Source Initiative [WWW.OPENSOURCE.ORG](http://WWW.OPENSOURCE.ORG)
  - GNU [WWW.GNU.ORG](http://WWW.GNU.ORG)
- Existem implementações de R para Linux, Windows e (Mac) OS X disponíveis gratuitamente a partir de [CRAN.RSTUDIO.COM](http://CRAN.RSTUDIO.COM).

# Vantagens

- Ao mudar de empresa, de escola, mudará o software disponível.
- Toda uma comunidade de utilizadores a testar, a desenvolver aplicações, e a corrigir erros.
- É possível utilizar somente software livre:
  - Linux como sistema operativo, eg, Ubuntu ou Fedora
  - OpenOffice no lugar de Office
  - $\text{\LaTeX}$  em vez de Word
  - Chrome ou Firefox no lugar de Internet Explorer
  - Thunderbird no lugar de Outlook
  - Octave em vez de Matlab; Sage em vez de Mathematica
  - etc...

## Vantagens

- R pode ser visto como uma implementação livre do ambiente S desenvolvido no Bell Labs por John Chambers et al. O R foi inicialmente escrito por Robert Gentleman e Ross Ihaka.
- O R é escrito em R, C/C++ e FORTRAN. Como temos acesso ao código fonte, é fácil
  - saber exactamente o que as rotinas calculam e que algoritmos usam
  - criar as nossas próprias rotinas possivelmente alterando rotinas existentes
  - chamar a partir do R funções escritas em C/C++ ou FORTRAN
  - chamar a partir do R rotinas de bibliotecas numéricas escritas em C/FORTRAN — BLAS, LAPACK, GNU scientific library, etc

# Packages

- O ambiente base do R é extensível pela adição de packages que se instalam com facilidade a partir do R.
- The Comprehensive R Archive Network — [CRAN.RSTUDIO.COM](http://CRAN.RSTUDIO.COM) — é um repositório de packages disponíveis (4370 actualmente).
- Facilidade de rápida disseminação de ideias.
- R afirma-se como a *lingua franca* da Estatística.

# Documentação

- Manual “An Introduction to R” disponível em [WWW.R-PROJECT.ORG](http://WWW.R-PROJECT.ORG).
- Ajuda acessível a partir do ambiente R.
- Centenas de apresentações/tutoriais/manuais disponíveis online além de fóruns de ajuda.
- Editora Springer tem uma colecção dedicada ao R: *use R!*, que conta actualmente com 42 publicações.
- Largas dezenas de livros publicados que vão desde textos introdutórios de Estatística a monografias em áreas de aplicação muito específicas.
- *UseR!* Conferência anual que se realiza desde 2004.



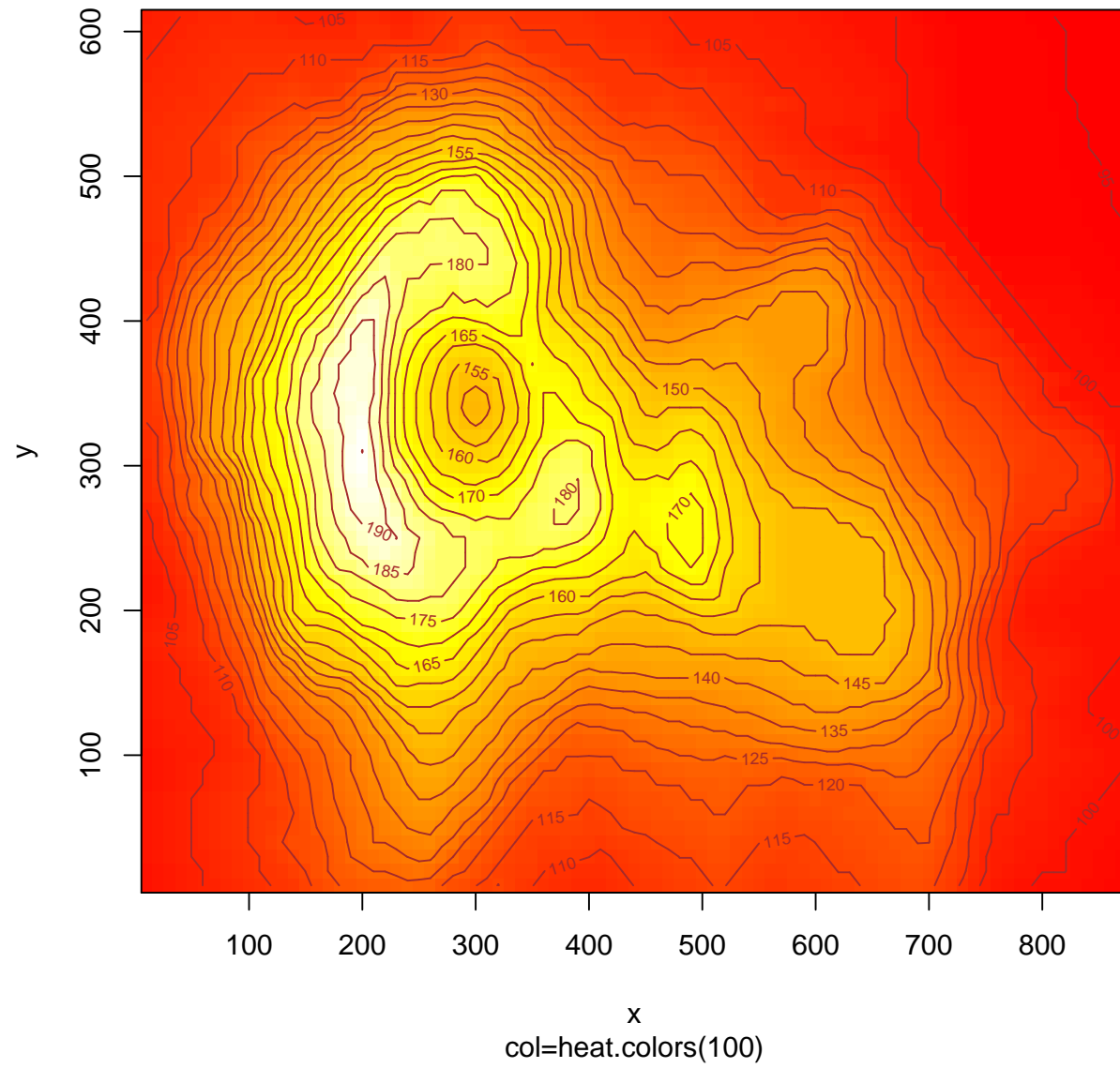
## Exemplos

- Soetaert, Cash and Mazzia. (2012) *Solving Differential Equations in R*. Springer.
- Pfaff. (2012) *Financial Risk Modelling and Portfolio Optimisation with R*. Wiley.
- Gilli, Maringer, and Schumann. (2011) *Numerical Methods and Optimization in Finance*. Academic Press.
- Kleiber and Zeileis. (2008) *Applied Econometrics with R*. Springer.
- Cowpertwait and Metcalfe. (2009) *Introductory Time Series with R*. Springer.
- Muenchen and Hilbe. (2010) *R for Stata Users*. Springer.
- Muenchen. (2009) *R for SAS and SPSS Users*. Springer.

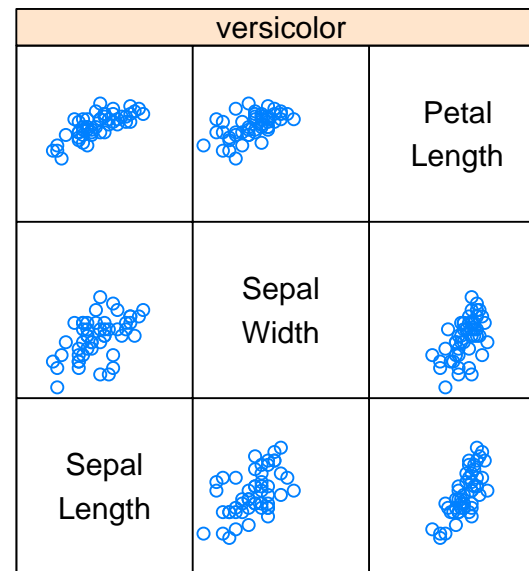
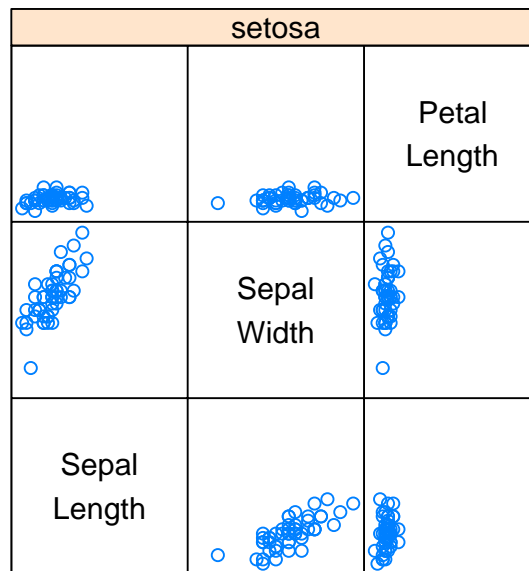
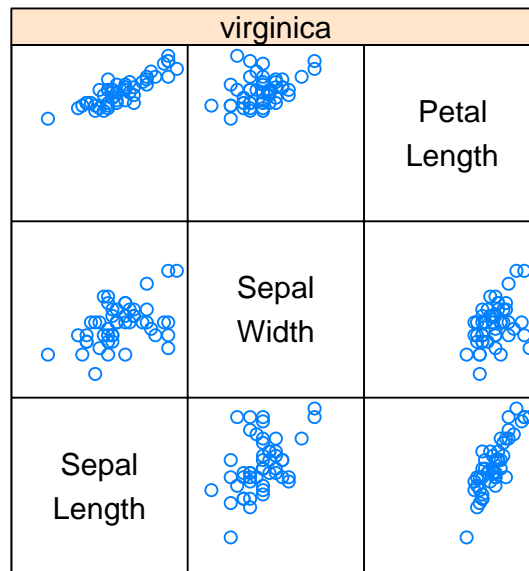
# Produção de Gráficos

- O R permite a produção de gráficos de altíssima qualidade em vários formatos, muitas vezes recorrendo a *packages*.
- Os valores padrão (*default*) dos parâmetros gráficos e de design foram estabelecidos por profissionais, mas todos podem ser alterados pelo utilizador.
- Não se trata de um sistema WYSIWG — *what you see is what you get*.
- Galeria de exemplos, incluindo código-fonte, em [GALLERY.R-ENTHUSIASTS.COM](http://GALLERY.R-ENTHUSIASTS.COM).

## Maunga Whau Volcano



# Three Varieties of Iris



Scatter Plot Matrix

**Curta visita guiada...**