

A escolha do *R* como software estatístico

Brasília
Novembro/2010

Perguntas freqüentes nas listas de e-mail

- **Qual software estatístico devo aprender?**
- **Quais softwares estão sendo exigidos nos processos seletivos?**



Perguntas freqüentes nas listas de e-mail

O aprendizado de um software está vinculado a:

- Necessidade por recursos computacionais para realizar análises estatísticas
 - Calculadora científica
 - *SPSS*
 - *SAS*
 - *R*
- Disponibilidade dos softwares (licença)
- Incentivo de terceiros e treinamento

Início da Graduação em Estatística

Estatística Exploratória – 1º Trabalho

- Diversas medidas descritivas
- Diversos Histograma

Usei as ferramentas do ensino médio

- Lápis
- Calculadora



Início da Graduação em Estatística

Estatística Exploratória – 1º Trabalho

- Uns 20 minutos para cada histograma
- Muito tempo revisando os cálculos medidas
- Baixa qualidade de apresentação

**PRECISO DO COMPUTADOR NAS
PRÓXIMAS VEZES!!**

Início da Graduação em Estatística

Uso com computador:



Início da Graduação em Estatística

Início da utilização do *Excel*

- Boa interface com o usuário
- Boas funções gráficas
- Boas funções de tabulação de dados
- Fácil utilização das funções em português

Empresa Junior de Estatística

Análise Descritiva de muitas variáveis

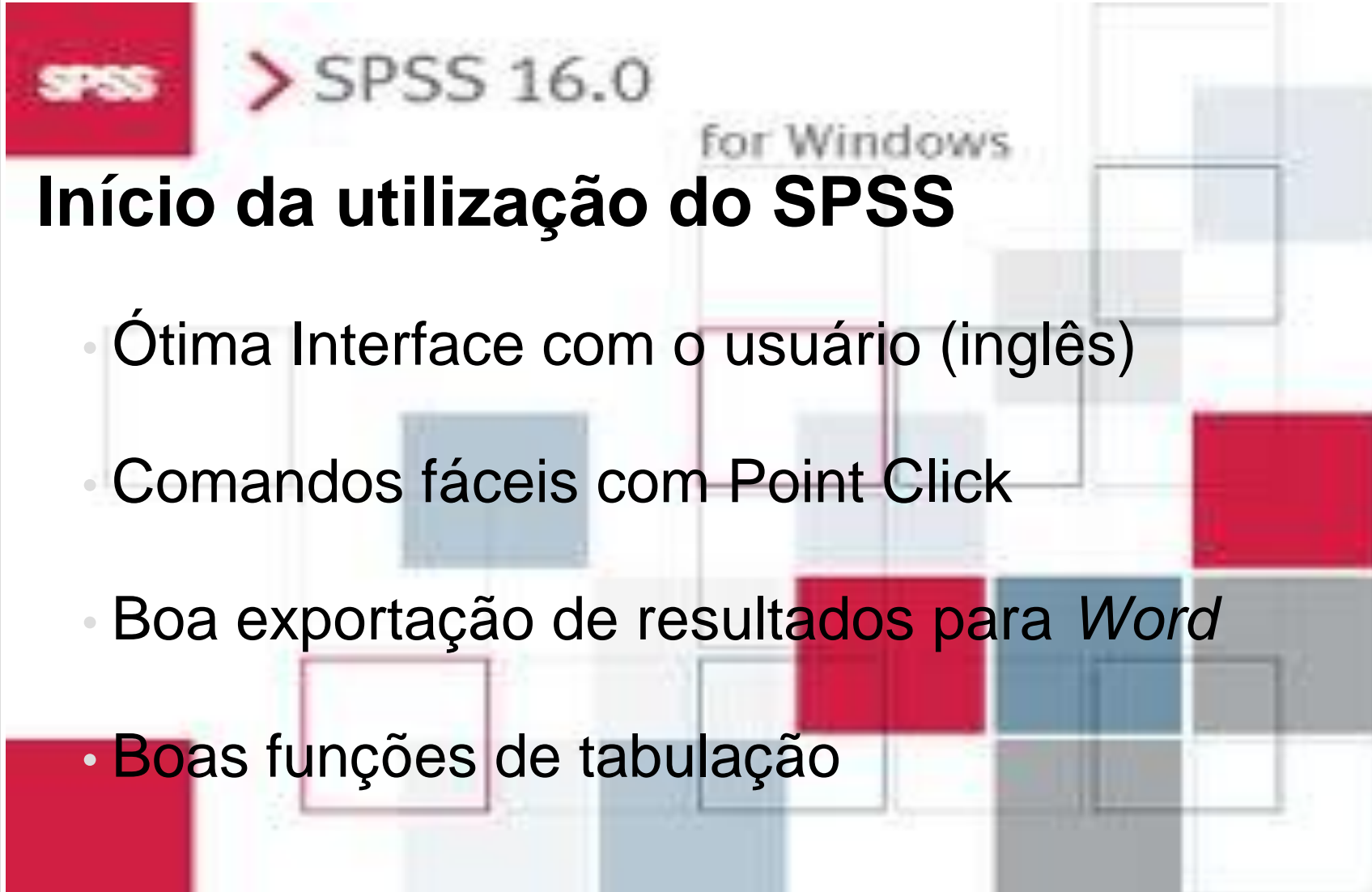
- Muito tempo para produzir relatório usando apenas os *Excel* e o *Word*
- Dificuldade para criação de macros no *Excel*

Regressão Linear múltipla

Análise de Experimento

O *Excel* não atendia mais todas as minhas necessidades

Empresa Junior de Estatística



Início da utilização do SPSS

- Ótima Interface com o usuário (inglês)
- Comandos fáceis com Point Click
- Boa exportação de resultados para *Word*
- Boas funções de tabulação

6º Semestre da Graduação

Na UnB, três porquinhos:

- Técnicas de Amostragem
- Análise Multivariada 1
- Planejamento e Pesquisa 1

O point click do SPSS começou a parecer um labirinto de opções estatísticas

Incentivo dos professores ao uso do SAS

7º Semestre da Graduação

- Boa Interface com o usuário (inglês)
- Muitas opções nas PROC's (labirinto resolvido)
- PROC IML
- Boa opção de Point Click
- Boas funções de exportação de resultados
- Boa documentação (help)
- Grande demanda do mercado de trabalho

8º Semestre da Graduação

Apresentado ao *R* no curso de Estatística Não-Paramétrica

- Teste de Hipótese - bootstrap
- Apenas com scripts de programação
- Utilizaram poucas funções prontas

Resolvi que não utilizaria o *R*

- Programação difícil
- Pobre interface com o usuário
- Poucos colegas profissionais utilizavam

Início do Mestrado

Primeiras disciplinas

- Muita teoria e pouca computação
- No entanto, constatei que todos os professores usavam o SAS ou o *R*
- Continue usando o SAS

Estatística Matemática (Inferência)

- Trabalho sobre propriedades de Estimadores
- Simulação
- Não consegui adaptar funções no SAS

Mestrado

Início da utilização do *R*

- Resolvido o problema de simulação e propriedades de estimadores
- Passei a utilizar o SAS e o *R* conjuntamente

Dissertação

- Modelos geoestatísticos bivariados
- Estimação e predição complexa
- Nenhum software possui tudo implementado

TERMINEI O MESTRADO USANDO APENAS O R

1º Emprego como Estatístico

TST - Tribunal Superior do Trabalho

- Possuía licenças de *Minitab*
- Servidores treinados para utilizar o software
- Problemas com renovação de contrato para o Minitab
- Problema com licitação para outros softwares

Atualmente usamos o *R* para realizar:

- Validação de dados
- Análise de séries temporais
- Análise de *Clusters* e Análise Fatorial

Qual software devo apreender?



O uso dos softwares vincula-se com:

- Necessidade
- Incentivo a utilização (professores, colegas...)
- Disponibilidade de licença

Qual software devo apreender?



**APROVEITE AS OPORTUNIDADES E TENHA
APRENDER UM POUCO DE TODOS!!**

- Processos seletivos com diversas exigências de conhecimento
- Possibilidade de comparação e crítica dos resultados

Em qual software devo me especializar?



ESPECIALIZE-SE NO *R*

Por que?

Principais motivos para especializar-se no *R*

- Software gratuito
- Todos os códigos são abertos, reproduzíveis e adaptáveis
- Desenvolvido para Windows, diversas plataformas UNIX e MacOS
- Comunidade acadêmica disseminando conhecimento (Grupos muito ativos; no Brasil, um bom grupo é o R_STAT do yahoo)
- Criação da revista *The R Journal*
- Empresas privadas e públicas estão se atentando ao R (diminuição de despesas e flexibilidade de análises)
- Link com *C*, *C++* e *Fortran*
- Ótima documentação (alterações de versões e help's)

**“Não pergunte se o *R* é capaz de fazer algo!
Pergunte como ele faz!”** *Bill Venables*

Baixando o R

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



The R Project for Statistical Computing

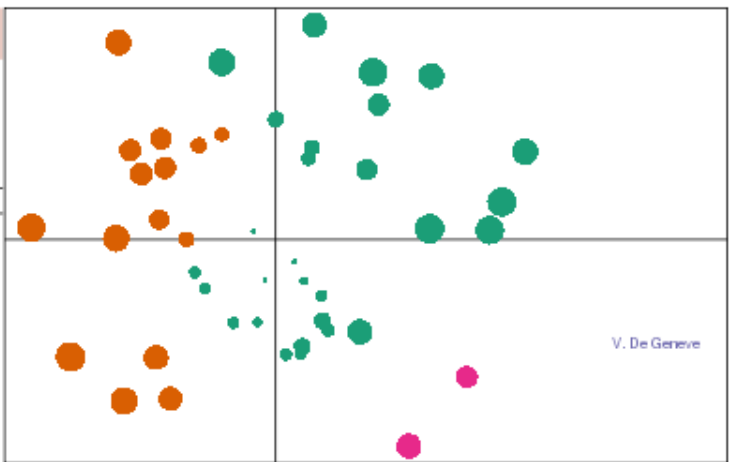
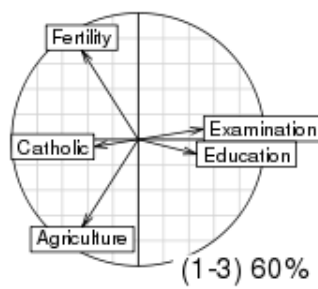
About R
[What is R?](#)
[Contributors](#)
[Screenshots](#)
[What's new?](#)

Download,
Packages
[CRAN](#)

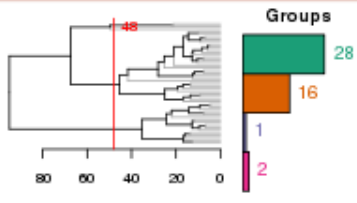
R Project
[Foundation](#)
[Members & Donors](#)
[Mailing Lists](#)
[Bug Tracking](#)
[Developer Page](#)
[Conferences](#)
[Search](#)

Documentation
[Manuals](#)
[FAQs](#)
[The R Journal](#)

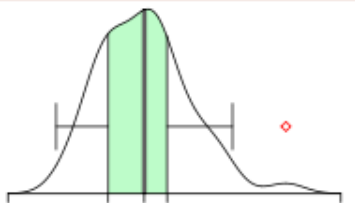
PCA 5 vars
`princomp(x = data, cor = cor)`



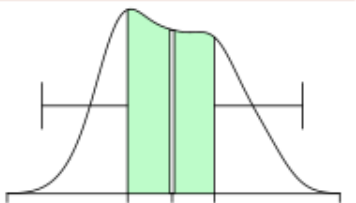
Clustering 4 groups



Factor 1 [41%]



Factor 3 [19%]



Getting Started:

- R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX

Baixando o R

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



CRAN Mirrors

The Comprehensive R Archive Network is available at the following URLs, please choose a location close to you. Some statistics on the status of the mirrors can be found [here](#).

- About R
- [What is R?](#)
- [Contributors](#)
- [Screenshots](#)
- [What's new?](#)
- Download, Packages
- [CRAN](#)
- R Project
- [Foundation](#)
- [Members & Donors](#)
- [Mailing Lists](#)
- [Bug Tracking](#)
- [Developer Page](#)
- [Conferences](#)
- [Search](#)
- Documentation
- [Manuals](#)
- [FAQs](#)
- [The R Journal](#)

Australia	http://cran.ms.unimelb.edu.au/	University of Melbourne
Austria	http://cran.at.r-project.org/	Wirtschaftsuniversitaet Wien
Belgium	http://www.freeststatistics.org/cran/	K.U.Leuven Association
Brazil	http://cran-r.c3sl.ufpr.br/	Universidade Federal do Parana
	http://cran.fiocruz.br/	Oswaldo Cruz Foundation, Rio de Janeiro
	http://www.vps.fmvz.usp.br/CRAN/	University of Sao Paulo, Sao Paulo
	http://brieger.esalq.usp.br/CRAN/	University of Sao Paulo, Piracicaba
Canada	http://cran.stat.sfu.ca/	Simon Fraser University, Burnaby
	http://mirror.its.dal.ca/cran/	Dalhousie University, Halifax
	http://probability.ca/cran/	University of Toronto
	http://cran.skazkaforyou.com/	iWeb, Montreal
	http://cran.parentingamerica.com/	iWeb, Montreal
Chile	http://dirichlet.mat.puc.cl/	Pontificia Universidad Catolica de Chile, Santiago
China		



Baixando o R

<http://cran-r.c3sl.ufpr.br>



The Comprehensive R Archive Network

Frequently used pages

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Linux](#)
- [MacOS X](#)
- [Windows](#)



Source Code for all Platforms

Windows and Mac users most likely want the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- **The latest release** (2010-10-15): [R-2.12.0.tar.gz](#) (read [what's new](#) in the latest version).
- Sources of [R alpha and beta releases](#) (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are [available here](#). Please read about [new features and bug fixes](#) before filing corresponding feature requests or bug reports.

CRAN

[Mirrors](#)

[What's new?](#)

[Task Views](#)

[Search](#)

About R

[R Homepage](#)

[The R Journal](#)

Software

[R Sources](#)

[R Binaries](#)

[Packages](#)

[Other](#)

Documentation

[Manuals](#)

[FAQs](#)

[Contributed](#)

Baixando o *R*

<http://cran-r.c3sl.ufpr.br>



R for Windows

This directory contains 32-bit binaries for a base distribution and packages to run on i386/x64 Windows. See [here](#) for a 64-bit Windows port.

Note: CRAN does not have Windows systems and cannot check these binaries for viruses. Use the normal precautions with downloaded executables.

Subdirectories:

[base](#) ← Binaries for base distribution (managed by Duncan Murdoch)
[contrib](#) Binaries of contributed packages (managed by Uwe Ligges)

Please do not submit binaries to CRAN. Package developers might want to contact Duncan Murdoch or Uwe Ligges directly in case of questions / suggestions related to Windows binaries.

You may also want to read the [R FAQ](#) and [R for Windows FAQ](#).

Last modified: April 4, 2004, by Friedrich Leisch

CRAN

[Mirrors](#)

[What's new?](#)

[Task Views](#)

[Search](#)

About R

[R Homepage](#)

[The R Journal](#)

Software

[R Sources](#)

[R Binaries](#)

[Packages](#)

[Other](#)

Documentation

[Manuals](#)

[FAQs](#)

[Contributed](#)

Baixando o R

<http://cran-r.c3sl.ufpr.br>

R-2.12.0 for Windows (32/64 bit)

[Download R 2.12.0 for Windows](#) (37 megabytes, 32/64 bit) 

[Installation and other instructions](#)

New features in this version: [Windows specific](#), [all platforms](#).

If you want to double-check that the package you have downloaded exactly matches the package distributed by R, you can compare the [md5sum](#) of the .exe to the [true fingerprint](#). You will need a version of md5sum for windows: both [graphical](#) and [command line versions](#) are available.

Frequently asked questions

- [How do I install R when using Windows Vista?](#)
- [How do I update packages in my previous version of R?](#)
- [Should I run 32-bit or 64-bit R?](#)

Please see the [R FAQ](#) for general information about R and the [R Windows FAQ](#) for Windows-specific information.

Other builds

- Patches to this release are incorporated in the [r-patched snapshot build](#).
- A build of the development version (which will eventually become the next major release of R) is available in the [r-devel snapshot build](#).
- [Previous releases](#)



CRAN

[Mirrors](#)

[What's new?](#)

[Task Views](#)

[Search](#)

About R

[R Homepage](#)

[The R Journal](#)

Software

[R Sources](#)

[R Binaries](#)

[Packages](#)

[Other](#)

Documentation

[Manuals](#)

[FAQs](#)

[Contributed](#)

Abrindo o R – Usando demonstrações prontas

```
R Console
Arquivo  Editar  Misc  Pacotes  Ajuda

R version 2.9.1 (2009-06-26)
Copyright (C) 2009 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R é um software livre e vem sem GARANTIA ALGUMA.
Você pode redistribuí-lo sob certas circunstâncias.
Digite 'license()' ou 'licence()' para detalhes de distribuição.

R é um projeto colaborativo com muitos contribuidores.
Digite 'contributors()' para obter mais informações e
'citation()' para saber como citar o R ou pacotes do R em publicações.

Digite 'demo()' para demonstrações, 'help.start()' para o sistema on-line de ajuda,
ou 'help.start()' para abrir o sistema de ajuda em HTML no seu navegador.
Digite 'q()' para sair do R.

Carregando pacotes exigidos: tcltk
Loading Tcl/Tk interface ... done
Carregando pacotes exigidos: Hmisc

Anexando pacote: 'Hmisc'

The following object(s) are masked from package:base :

  format.pval,
  round.POSIXt,
  trunc.POSIXt,
```

Exemplo de função interativa – Gincana do XXI ENEST

```
gincana <- function() {  
  cat("Digite o nome da universidade que irá ganhar a gincana (em letras minúsculas)\n")  
  universidade <- readLines(n = 1)  
  if (universidade == "unb")  
    cat("ACHO QUE VOCÊ ACERTOU!!!\n")  
  else cat("IHhh, ESCOLHA ERRADA...\n")  
  return(invisible())  
}  
gincana()
```

Exemplo gráfico – Amor pela cidade

```

cidade=function(nome="Brasília"){
  x=seq(-2,2,0.001)
  y1=sqrt(1-((abs(x)-1)^2))
  y2=-3*sqrt(1-(sqrt(abs(x))/sqrt(2)))
  plot(main=paste("XXI ENEST\n",nome," achei uma fórmula pra explicar o que sinto por
  você...",sep=""),
  ylab="",xlab="",c(x,x),c(y1,y2),lwd=3,pch=20,cex=6)
  grid()
  cores=paste(c("green","yellow"),sort(rep(1:4,2)),sep="")
  eq=c("( AM + BC ) * X = AM ( X + BOC ) - BCTE","AMX + BCX = AMX + AMBOC -
  BCTE","BCX = AMX - AMX + BC (AMO - TE)","BCX = BC ( AMO - TE )","X = ???")
  pos=c(-0.3,-0.6,-0.9,-1.2,-1.5)
  cont=0
  for (i in 1:30){
    points(c(x,x),c(y1,y2),col=cores[sample(1:8,1)],lwd=sample(1:4,1),pch=20,cex=round(r
    unif(1,1,4)))
    if (i%%5 == 0 & i <=25) {
      cont=cont+1
      text(0,pos[cont],eq[cont])
    }
  }
  text(1.2,-2.5,"AMO-TE",col="red",cex=3)}
cidade()
## Caso queira trocar o nome da cidade, basta executar cidade("Nome da cidade")

```

Entendendo o sistema de pacotes do R

- Ao instalar o R, diversos pacotes básicos são automaticamente baixados e instalados
- Ao iniciar a execução do R, *por default*, alguns pacotes ficam disponíveis para utilização
- Os demais pacotes pré-instalados podem ser carregados para o R com as funções:
 - `library(nome do pacote)`
 - `require(MASS)`, obs.: MASS é exemplo
- Outros pacotes não são instalados ao baixar o R, como por exemplo, o pacote Rcmdr, para instalá-lo utilize:
 - `install.packages("Rcmdr")`

Entendendo o sistema de pacotes do R

Todos os Pacotes

**Pacotes Instalados
no Computador**

**Pacotes instalados
automaticamente**

**Pacotes de
instalação opcional:
install.packages()**

**Pacotes
carregados para
memória do R
(disponíveis
para utilização)**

**Pacotes carregados
automaticamente**

**Pacotes de
carregamento opcional:
require() ou library()**

Preciso saber programar para iniciar o uso do R?

Resposta: NÃO

- **O pacote *Rcmdr* é uma ótima solução**
 - Ótima interface gráfica
 - Menus de entrada em português
 - Janela de Output com scripts
- **Execute nas linhas de comando do R:**
 - `install.packages("Rcmdr")`
 - `require(Rcmdr)` ou `library(Rcmdr)`

Interface do pacote Rcmdr

The screenshot displays the R Commander application window. The title bar reads "R Commander". The menu bar includes "Arquivo", "Editar", "Dados", "Estatísticas", "Gráficos", "Modelos", "Distribuições", "Ferramentas", and "Ajuda". Below the menu, the "Conjunto de Dados:" field shows "<Não há conjunto de dados ativo>" with buttons for "Editar conjunto de dados" and "Ver conjunto de dados". The "Modelo:" field shows "<sem modelo ativo>".

The main interface is divided into two primary sections: "Janela do Script" (Script Window) and "Janela de Resultados" (Results Window). Both are currently empty. A "Submeter" button is located in the top right corner of the Results Window.

At the bottom, the "Mensagens" (Messages) pane displays the following text: "[1] NOTA: Versão do R Commander 1.5-4: Mon Nov 08 18:07:51 2010".

The Windows taskbar at the bottom shows the Start button and several open applications: "Entrada - bruno.fon...", "Microsoft PowerPoin...", "[R_STAT] Resumo 1...", "R Console", and "R Commander". The system tray on the right shows the date and time as "18:28".

Introdução aos scripts do R

- **Operações Aritméticas:**

```
> 5+4
```

```
[1] 9
```

```
> ((5+4)*25)/18
```

```
[1] 12.5
```

```
> 3^2
```

```
[1] 9
```

```
> 3**2
```

```
[1] 9
```

- **Funções de operações:**

- `sqrt()` raiz quadrada
- `abs()` valor absoluto
- `sin()` `cos()` `tan()` trigonométricas
- `asin()` `acos()` `atan()` trigonométricas inversas
- `inh()` `cosh()` `tanh()` hiperbólicas
- `asinh()` `acosh()` `atanh()` hiperbólicas inversas
- `exp()` `log()` exponencial e logaritmo natural
- `log10()` `log2()` logaritmo base-10 e 2
- `gamma()` função Gamma de Euler
- `factorial()` fatorial (n!)
- `choose()` número de combinações

Introdução aos scripts do R

Cada pacote possui diversas análises implementadas via funções

Funcionamento básico de uma função:

- Similar às funções do Excel
- Possui basicamente um nome e argumentos de entrada
- Os argumentos devem ser informados entre parênteses
- Os argumentos devem ser separados por vírgula
- Para executar o comando basta um clique no *Enter*

Exemplos:

```
> c(10,15,12,9,18) #criando um vetor sem armazenar em objeto
```

```
> dados <- c(10,15,12,9,18) #armazenando como objeto de nome dados
```

```
> quantile(x=dados,probs=c(0,0.25,0.5,0.75,1)) #calcula os quartis
```

```
> quartil_enest <- quantile(x=dados,probs=c(0,0.25,0.5,0.75,1))  
#armazenando como objeto de nome quartil_enest
```

Introdução aos scripts do *R*

• **Objetos:**

- Linguagem orientada à objetos
- Os objetos são armazenados na memória ativa do computador
- O símbolo “<-” atribui algo para um objeto

```
> x <- sqrt(100)
> x
[1] 10
```
- O símbolo “=” também é atribui algo para um objeto, mas cuidado com seu uso.

Introdução aos scripts do *R*

Tipos básicos de objetos:

- Escalar (numérico ou caractere)
- Vetor (numérico ou caractere)
- Matriz (numérico)
- Data Frame (matriz de banco de dados, numérico e/ou caractere)
- Lista (combina diversos tipos de objetos)

Introdução aos scripts do R

- **Procurando funções na documentação interna:**
(tomando como exemplo a palavra clusters)

```
> help.search("clusters");
```

ou

```
> ??clusters
```

- **É disponibilizada uma lista:**
 - Nome do pacote
 - Nome da função
 - Breve descrição da função

Introdução aos scripts do *R*

- **Procurando funções na web (search no site do R):**
(tomando como exemplo a palavra clusters)

```
>RSiteSearch("clusters");
```

- **Disponibiliza uma lista de dúvidas respondidas:**
 - Palavra Chave
 - Frases

Introdução aos scripts do *R*

- **Acessando o help de uma função:**
(tomando como exemplo a função `quantile`)

> `help(quantile)`

ou

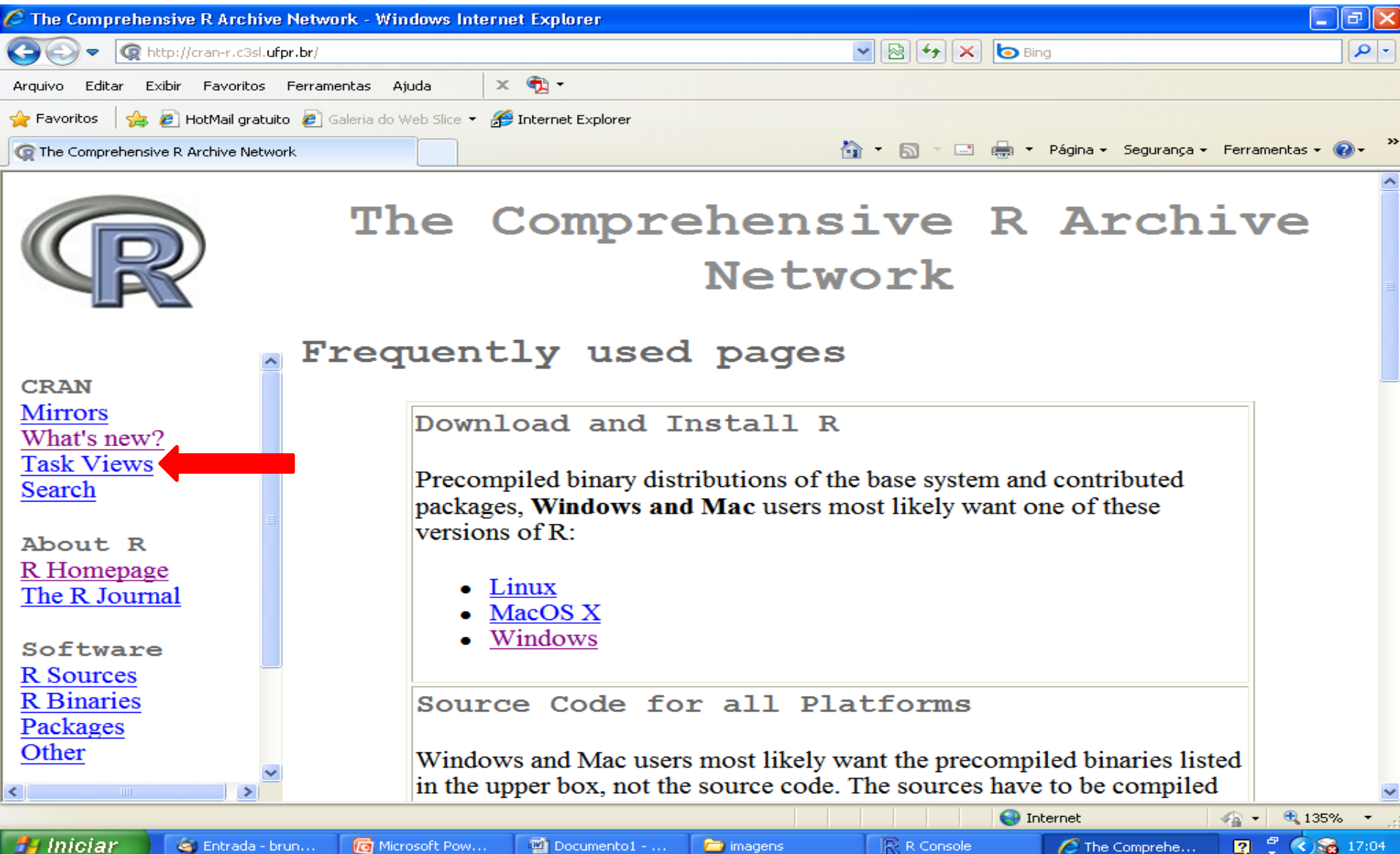
> `?quantile`

- **O arquivo da ajuda apresenta:**

- Uso dos argumentos de entrada para função
- Detalhes técnicos (tratamento de desempates por exemplo)
- Autor e referências bibliográficas
- Lista funções relacionadas
- Exemplos reproduzíveis (com um simples Ctrl C; Ctrl V)

Como encontrar os pacotes do R que preciso?

<http://cran-r.c3sl.ufpr.br>



The screenshot shows the CRAN website in Internet Explorer. The browser title is "The Comprehensive R Archive Network - Windows Internet Explorer". The address bar shows "http://cran-r.c3sl.ufpr.br/". The page content includes the CRAN logo, the title "The Comprehensive R Archive Network", and a section titled "Frequently used pages". On the left, there is a navigation menu with links: "CRAN", "Mirrors", "What's new?", "Task Views", "Search", "About R", "R Homepage", "The R Journal", "Software", "R Sources", "R Binaries", "Packages", and "Other". A red arrow points to the "Task Views" link. The main content area has two boxes: "Download and Install R" and "Source Code for all Platforms". The "Download and Install R" box contains the text "Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:" followed by a bulleted list: "Linux", "MacOS X", and "Windows". The "Source Code for all Platforms" box contains the text "Windows and Mac users most likely want the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled".

The Comprehensive R Archive Network

The Comprehensive R Archive Network

Frequently used pages

CRAN
[Mirrors](#)
[What's new?](#)
[Task Views](#)
[Search](#)

About R
[R Homepage](#)
[The R Journal](#)

Software
[R Sources](#)
[R Binaries](#)
[Packages](#)
[Other](#)

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Linux](#)
- [MacOS X](#)
- [Windows](#)

Source Code for all Platforms

Windows and Mac users most likely want the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled

Como encontrar os pacotes do R que preciso?

http://cran-r.c3sl.ufpr.br

The Comprehensive R Archive Network - Windows Internet Explorer

http://cran-r.c3sl.ufpr.br/

Arquivo Editar Exibir Favoritos Ferramentas Ajuda

Favoritos HotMail gratuito Galeria do Web Slice Internet Explorer

The Comprehensive R Archive Network

CRAN

[Mirrors](#)

[What's new?](#)

[Task Views](#)

[Search](#)

About R

[R Homepage](#)

[The R Journal](#)

Software

[R Sources](#)

[R Binaries](#)

[Packages](#)

[Other](#)

[Bayesian](#)

[ChemPhys](#)

[ClinicalTrials](#)

[Cluster](#)

[Distributions](#)

[Econometrics](#)

[Environmetrics](#)

[ExperimentalDesign](#)

[Finance](#)

[Genetics](#)

[Graphics](#)

[gR](#)

[HighPerformanceComputing](#)

[MachineLearning](#)

[MedicalImaging](#)

CRAN Task Views

Bayesian Inference

Chemometrics and Computational Physics

Clinical Trial Design, Monitoring, and Analysis

Cluster Analysis & Finite Mixture Models

Probability Distributions

Computational Econometrics

Analysis of Ecological and Environmental Data

Design of Experiments (DoE) & Analysis of Experimental Data

Empirical Finance

Statistical Genetics

Graphic Displays & Dynamic Graphics & Graphic Devices & Visualization

gRaphical Models in R

High-Performance and Parallel Computing with R

Machine Learning & Statistical Learning

Medical Image Analysis

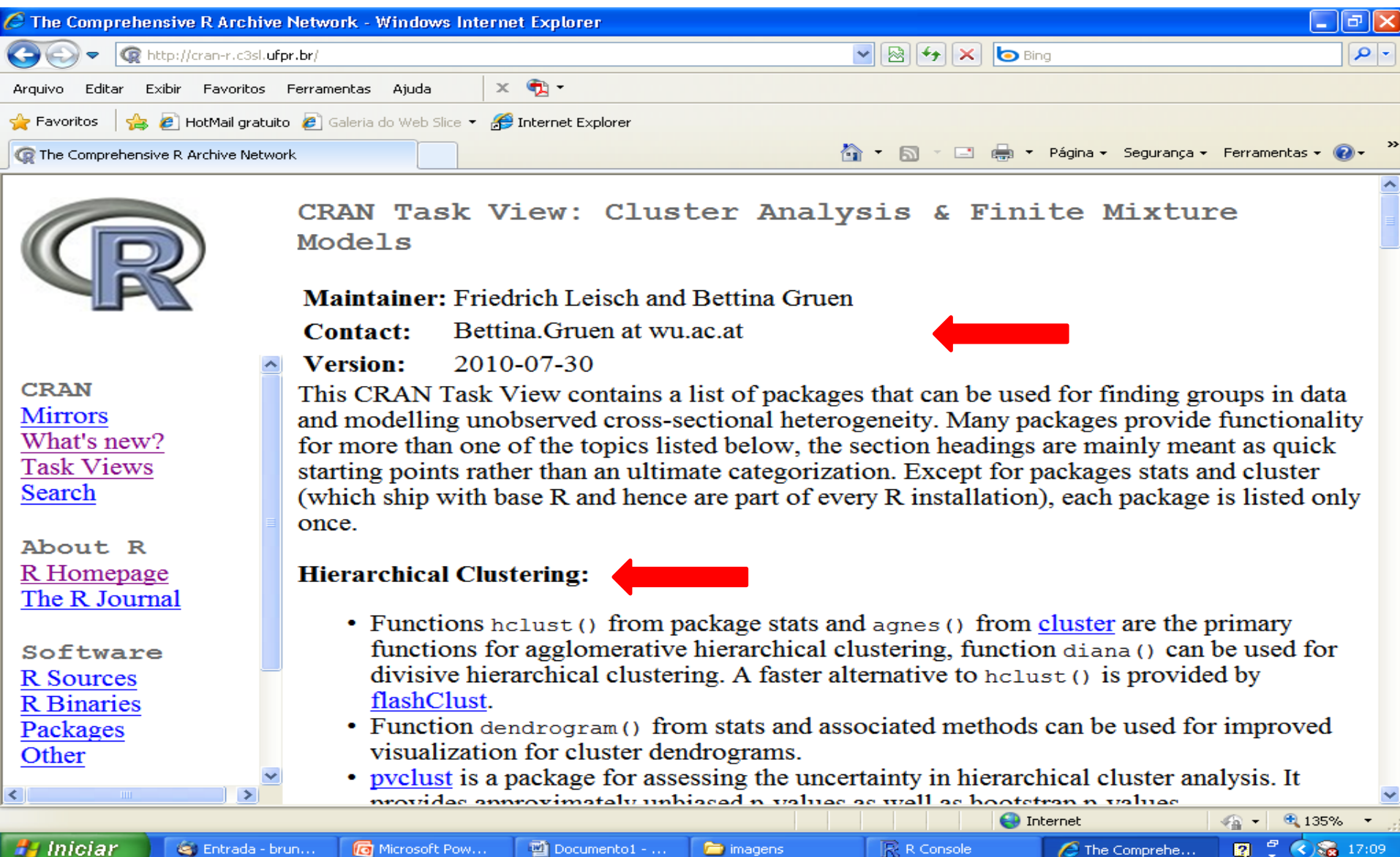
Concluído

Internet 135%

Iniciar Entrada - brun... Microsoft Pow... Documento1 - ... imagens R Console The Comprehe... 17:07

Como encontrar os pacotes do R que preciso?

http://cran-r.c3sl.ufpr.br



The screenshot shows a Windows Internet Explorer browser window with the address bar displaying `http://cran-r.c3sl.ufpr.br/`. The page content is titled "CRAN Task View: Cluster Analysis & Finite Mixture Models".

Maintainer: Friedrich Leisch and Bettina Gruen
Contact: Bettina.Gruen at wu.ac.at
Version: 2010-07-30

This CRAN Task View contains a list of packages that can be used for finding groups in data and modelling unobserved cross-sectional heterogeneity. Many packages provide functionality for more than one of the topics listed below, the section headings are mainly meant as quick starting points rather than an ultimate categorization. Except for packages stats and cluster (which ship with base R and hence are part of every R installation), each package is listed only once.

Hierarchical Clustering:

- Functions `hclust()` from package stats and `agnes()` from [cluster](#) are the primary functions for agglomerative hierarchical clustering, function `diana()` can be used for divisive hierarchical clustering. A faster alternative to `hclust()` is provided by [flashClust](#).
- Function `dendrogram()` from stats and associated methods can be used for improved visualization for cluster dendrograms.
- [pvclust](#) is a package for assessing the uncertainty in hierarchical cluster analysis. It provides approximately unbiased p values as well as bootstrap p values.

Navigation links on the left include: CRAN Mirrors, What's new?, Task Views, Search, About R, R Homepage, The R Journal, Software, R Sources, R Binaries, Packages, Other.

Introdução aos scripts do R

Algumas formas para entrada de dados:

- **Vetor diretamente na linha de comando:**

> bd <- c(1,2,3,4,5) ou >bd <- 1:5 ou bd <- seq(1,5,1)

- **Importando arquivo de texto (".txt", ".csv") :**

>bd <- read.table(file="C:/Users/t/Desktop/enest/bd.txt",header=T)

- Use o help(read.table) para detalhes, maior atenção aos argumentos "header" e "dec"

- **Usando o teclado:**

>bd <- scan()

- Irão aparecer linhas para digitação dado por dado

- **Usando planilha do R:**

>bd <- edit(data.frame())

- Para alterar os dados use a função fix() (no exemplo, fix(bd))

Introdução aos scripts do R

Algumas formas para entrada de dados:

- **Importando de outros softwares:**
 - `read.dbf()` para arquivos DBASE
 - `read.epiinfo()` para arquivos .REC do Epi-Info
 - `read.mtp()` para arquivos "Minitab Portable Worksheet"
 - `read.S()` para arquivos do S-PLUS
 - `read.spss()` para dados do SPSS
 - `read.dta()` para dados do STATA
 - `read.octave()` para dados do OCTAVE (similar ao MATLAB)
 - `read.export()` ou `sas.get()` para dados SAS, requer instalação do sistema SAS.

Introdução aos scripts do R

Funções interessantes de ligação com outros aplicativos:

- **Cálculo de derivadas e integrais**

- >deriv() # função dos pacotes básicos

- >integrate() # função dos pacotes básicos

- **Criando relatórios combinando R e Latex**

- >Sweave() # função dos pacotes básicos

- **Enviando e-mail pelo R:**

- >sendmail() # função do pacote sendmailR

- **Utilizando o R para manipular o Twitter:**

- Pacote twitterR

- Diversas funcionalidades

- Enviar texto, encontrar usuários, administrar contatos...

Introdução aos scripts do R

Observações importantes:

- O linguagem do R é case sensitive (exemplo: Dados \neq dados)
- O símbolo “>” é o início de uma linha de comando, para executar o script da linha basta pressionar a tecla *enter*
- O símbolo “<-” atribui algo a um objeto, um pouco diferente do “=”
 - > teste <- quantile(x=dados) #é o mesmo que
 - > teste = quantile(x=dados) #porém, diferente de
 - > teste <- quantile(x<-dados) #além de criar o objeto teste, cria o objeto x
- O símbolo “==” é utilizado em funções lógicas (if, else...)
- símbolo “#” introduz comentário, que não é executado

Introdução aos scripts do R

Observações importantes:

- O aspas diferenciam objetos de simples textos (caracteres)
 - > mu1 <- mean(x=c(1,2,3,4)) #mu1 recebe a média do vetor
 - > mu2 <- mean(x=c(5,6,7,8)) #mu2 recebe a média do vetor
 - > mu <- c(mu1,mu2) #mu recebe um vetor 2x1 com as duas médias
 - > mu <- c(mu1,"mu2") #mu recebe um objeto 2x1 com uma média e um texto
- O "." separa a parte inteira das casas decimais dos números
- Utilize help(options) para ver as configurações do R
- Utilize a options() para alterar as configurações de interesse

Conclusão

Necessidade por softwares

- Todos que trabalham com estatística possuem ou possuem

Disponibilidade de licença

- O *R* é livre e de código aberto, então nunca irá se preocupar

Incentivo para utilização do *R*

- Espero ter motivado os futuros estatísticos para utilizar o *R*

Se preocupe mais com seu conhecimento teórico do que com seu conhecimento sobre softwares estatísticos



Contato:

- bruno.fonseca@tst.jus.br

Slides disponíveis em:

- www.leg.ufpr.br/doku.php/pessoais:bruno



Universidade de Brasília

XXI ENEST
Bem-vindos a Brasília
OBRIGADO