

I Semana Tecnológica do IFMT

Minicurso de Introdução ao Software R

Jhessica Letícia Kirch – UFMT

jhessicakirch@gmail.com



Lucas do Rio Verde – MT

2016

Sumário

Como baixar e instalar o Software R	1
Como instalar pacotes no Software R.....	5
Comandos importantes para o uso do R	5
Criando objetos	6
Funções matemáticas básicas	6
Funções matemáticas.....	6
Vetor.....	7
Matriz	7
Importando um banco de dados.....	7
Estatística descritiva	8
Anexando dados.....	8
Tabela de frequência.....	8
Funções estatísticas básicas	8
Resumo numérico	8
Alterando o tipo de variável.....	9
Criando um novo banco de dados a partir do anterior.....	9
Filtro	9
Juntando dois bancos de dados em um só	9
Análise visual de dados - Gráficos.....	9
Regressão linear simples	10
Análise de Variância	10
Análise de Experimentos.....	11
Delineamento Inteiramente Casualizado.....	11
Delineamento em Blocos Casualizado	11
Delineamento em Quadrado Latino.....	11

O R é um software gratuito de linguagem de programação amplamente utilizado em análises estatísticas. Todos os códigos do R são abertos, reproduzíveis e adaptáveis.

Foi criado originalmente por Ross Ihaka e por Robert Gentleman no departamento de Estatística da universidade de Auckland, Nova Zelândia, e foi desenvolvido em um esforço colaborativo de pessoas em vários locais do mundo.

O R disponibiliza uma ampla variedade de técnicas estatísticas e gráficas, incluindo modelação linear e não linear, testes estatísticos clássicos, análise de séries temporais, classificação, agrupamento e outras.

Como baixar e instalar o Software R

- Acesse a página oficial do projeto R: <http://www.r-project.org/>
- No canto superior esquerdo, acessa a seção de *download CRAN*



[Home]

Download

CRAN

- Escolha um servidor, de preferência em seu país e mais próximo da sua cidade

Brazil

<http://nbcgib.uesc.br/mirrors/cran/>

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

<https://cran.fiocruz.br/>

<http://cran.fiocruz.br/>

<https://vps.fmvz.usp.br/CRAN/>

<http://vps.fmvz.usp.br/CRAN/>

<http://brieger.esalq.usp.br/CRAN/>

- Na página seguinte, escolha seu sistema operacional

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:

- [Download R for Linux](#)
- [Download R for \(Mac\) OS X](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

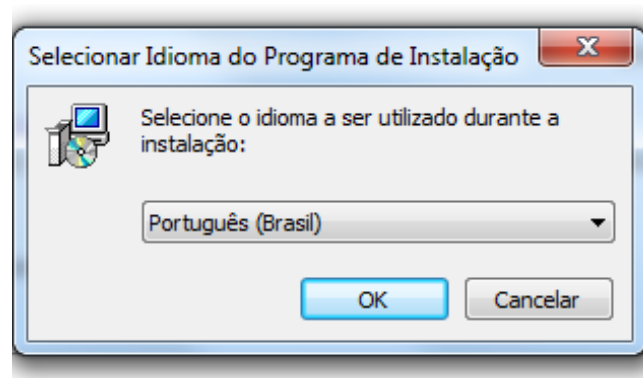
- Em seguida, clique em “*install R for the first time*”

Binaries for base distribution (managed by Duncan Murdoch). This is what you want to [install R for the first time](#).
Binaries of contributed CRAN packages (for R \geq 2.11.x; managed by Uwe Ligges). There is also information on [third](#)

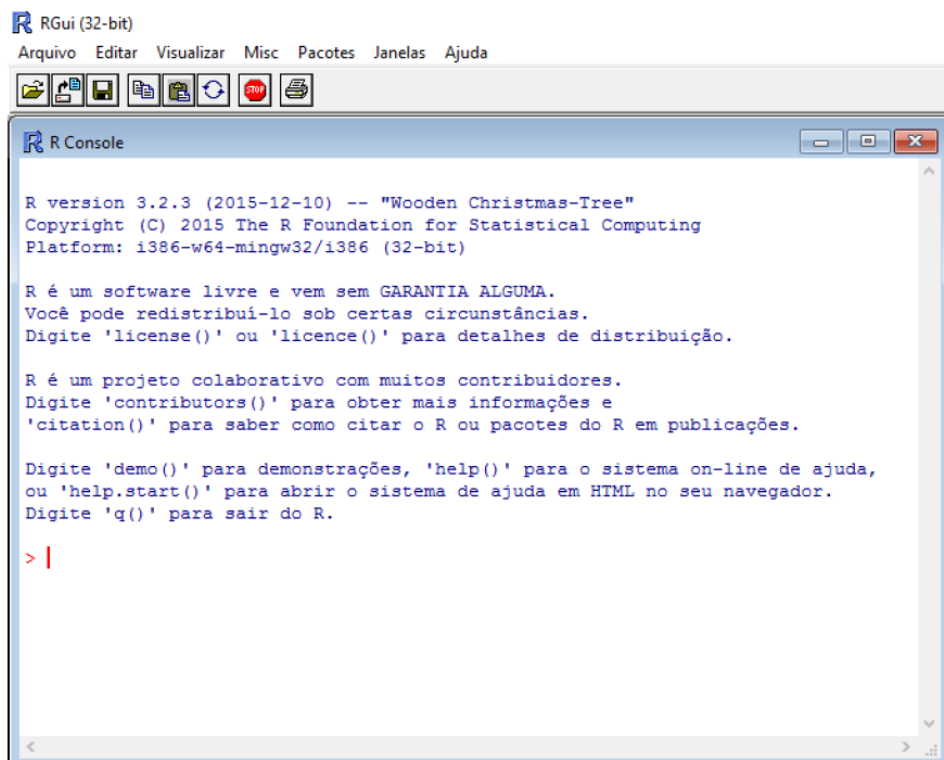
- Na página seguinte, clique em “*Download R...*”

[Download R 3.3.2 for Windows](#) (62 megabytes, 32/64 bit)

- Para a instalação do software, clique no arquivo executável (download)
- Selecione o idioma a ser utilizado durante a instalação. O idioma **Português (Brasil)** está disponível.

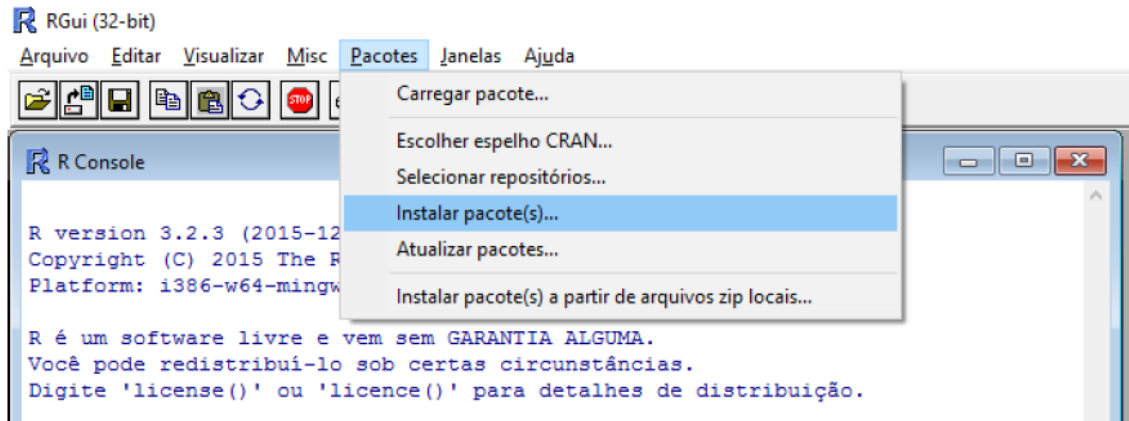


- Nas demais etapas, basta clicar repetidamente no botão **Avançar**. Ao final do processo, execute o R

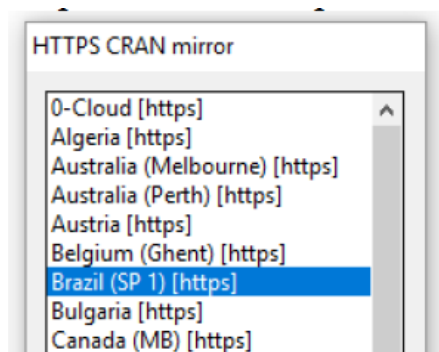


Como instalar pacotes no Software R

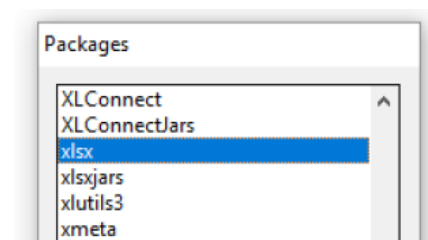
- Acesse o menu **Pacotes > Instalar pacote(s)...**



- Selecione o repositório disponível em seu país e clique em **OK**



- Selecione um pacote. Nesta apostila será utilizado o pacote **xlsx**. Procure, selecione e, em seguida, clique em **OK**



Comandos importantes para o uso do R

- Para utilizar um pacote já instalado é necessário utilizar a função:

```
require(nome_do_pacote)  
Exemplo: require(xlsx)
```

- Para ver quais pacotes já estão instalados:

```
library()
```

- Para saber quais pacotes já estão ativos:

```
search()
```

- Para obter ajuda sobre alguma função do R, utilize qualquer umas das funções abaixo:

```
help(nome_da_função)
?nome_da_função
Exemplo: help(plot)
```

- Para limpar a tela: `Ctrl+l`
- Para fechar o R: `q()`
- Para citar o R em artigos acadêmicos, utilizar a função: `citation()`
- Para saber quais cores o R conhece: `colors()`

<http://www.stat.columbia.edu/~tzheng/files/Rcolor.pdf>

Criando objetos

- Para criar um objeto:

```
x<-
```

- Para listar objetos:

```
ls()
```

- Para remover um objeto:

```
rm(nome do objeto)
rm(list=ls(all=TRUE))
```

Funções matemáticas básicas

FUNÇÕES MATEMÁTICAS

```
sqrt(número) #raiz quadrada
abs(número) #módulo
exp(número) #exponencial
log(número) #logatirmo
log(número,base) #logatirmo
factorial(número) #fatorial
```

VETOR

```
u<-c(número,número)
length(nome do vetor) #tamanho do vetor
seq(numero inicial, numero final, by=número da razão)
#gerar vetor de sequencia
rep(numero a ser repetido, quantas vezes repeti-lo)
#gerar vetor de repetição

sample(x, tamanho, replace) #gerar valores aleatórios
#Exemplos

Moeda<-c("CARA","COROA")
sample(moeda,10)

dados<-sample(1:10,5) # tira 5 amostras com valores
entre 1 e 10

sort(dados) #colocar os dados em ordem crescente
sort(dados, decreasing=TRUE) #colocar os dados em
ordem decrescente
order(dados) #retoma a posição original de cada valor
rank(dados) #atribui postos aos valores do objeto
```

MATRIZ

```
#matrix(c(digitar todos os numeros na ordem da
coluna), número de linhas,numero de colunas)
A<-matrix(c(2,1,3,4),2,2) #exemplo
A
t(A) #matriz transposta
t(A)%*%A #produto de matrizes
#sinal "%*%" designa produto de matrizes
sum(diag(A)) #traço da matriz
det(A) #determinante da matriz
solve(A) #matriz inversa
```

Importando um banco de dados

- Primeiramente é necessário escolher o local em que o arquivo está localizado por meio da função:

```
setwd("endereço_do_arquivo")
```

Observação: é necessário substituir as barras “\” por “/” ou “//”

```
Exemplo: setwd("C:/Users/CAEST/Minicurso R")
```

- Para verificar a pasta selecionada:

```
getwd()
```

- Para importar um arquivo *.txt*

```
#Colunas separadas por espaço
dados <- read.table("nome_do_arquivo.txt",header=T)

#Colunas separadas por ponto e vírgula
dados <- read.table("nome_do_arquivo.txt ", header= T,
sep=";")

#Colunas separadas por tabulação
dados <- read.table("nome_do_arquivo.txt ", header=T,
sep="\t")
```

- Para importar um arquivo *.csv*

```
dados <- read.csv("nome_do_arquivo.csv")
```

- Para importar um arquivo *.xlsx*

```
require(xlsx)
dados <- read.xlsx("nome_do_arquivo.xlsx ", planilha)
```

Estatística Descritiva

#ANEXANDO DADOS

```
attach(dados)
```

TABELA DE FREQUÊNCIA

```
table(variável)
#tabela de frequências cruzadas
table(variável1,variável2)
```

FUNÇÕES ESTATÍSTICAS BÁSICAS

```
mean(variável) #média
median(variável) #mediana
quantile(variável,0.5) #quartil
quantile (variável, 0.25) #quartil
max(variável) #máximo
min(variável) #mínimo
var(variável) #variância
sd(variável) #desvio padrão
cov(variável1, variável2) #covariância
cor(variável1, variável2) #correlação
```

RESUMO NUMÉRICO

```
summary(variável)
```


ALTERAR TIPO DE VARIÁVEL

```
str(dados)
#obs:criar outra a partir dela
#obs2:chamando uma variavel exemplo dados$ID

#mudando para variavel qualitativa
#character
dados$ID1<- as.character(dados1$ID)

#mudando para categórico
#factor
dados1$SMOKE2<- as.factor(dados1$SMOKE)

#mudando para numerico
#numeric
dados1$ID2<- as.numeric(dados1$ID1)
```

CRIANDO UM NOVO BANCO DE DADOS A PARTIR DE UM ANTERIOR

```
base1=dados[c(1,2,3)] #número da coluna das variáveis
que entrarão no novo banco de dados
base2=dados[c("variável1"," variável2"," variável3")]
#nome da coluna das variáveis que entrarão no novo banco de
dados
str(base1)
str(base2)
```

FILTRO

```
#somente variavel1=2
dados.S=dados[variavel1=="2",]
dados.S
str(dados.S)

#somente variavel1=2 e variavel2<30
dados.T=dados[variavel1=="2"&variavel2<45,]
dados.T

# e é "&" e ou é "|"
```

JUNTAR DOIS BANCOS DE DADOS EM UM SÓ

```
#merge
merge(obj,obj,by.x=c(chave),by.y=c(chave),all.y)
```

Análise visual de dados - Gráficos

```
#vários gráficos na mesma janela
par(mfrow=c(2,2))

#barras
barplot(variável, main = "Gráfico de barras")
```

```

#setores
  pie(table(variável))
  pie(table(variável), labels = c("TIPO 0","TIPO 1"),
    main = "Setores")

#gráfico de dispersão
  plot(variável, main = "Gráfico simples", ylab =
    "Valor", xlab = "Observação", col = "blue")

#boxplot simples
  boxplot(variável)

#boxplot por grupo
  boxplot(variável1~variável2)

#histograma
  hist(variável, main = "Histograma")
  text(coordenadaX,coordenadaY,"texto")
  text(locator(1),"Texto") #clique com o mouse na
  posição desejada do gráfico

#alterar o número de classes
  hist(variável, main = "Histograma", breaks = 15, ylab
    = "Frequência", xlab = "", col = "gray")

#probabilidades
  hist(variável, main = "Histograma", prob = T, breaks =
    15, ylab = "Prob", xlab = "", col = "gray")
  lines(density(variável))

```

Regressão Linear Simples

```

# REGRESSÃO LINEAR
  regressao<-lm(variável1~variável2)
  regressao
  plot(variável1~variável2)
  abline(regressao)

```

Análise de Variância

```

#ANOVA
  anova<-aov(Resp~Variavel)
  summary(anova)

```

Análise de Experimentos

#DIC

```
dic(Trat,Resp, mcomp="teste")  
?dic
```

#DBC

```
dbc(Trat,Bloc,Valor,mcomp="teste")  
?dbc
```

#DQL

```
dql(Trat,Linha,Coluna,Valor, mcomp="teste")  
?dql
```