

R aplicado a suinocultura

Marcio Valk

2020-07-09

Contents

Objetivo	1
Sobre a Academia	1
Sobre o autor	1
1 Tutorial básico R e RStudio	3
1.1 Apresentação da linguagem R	3
1.2 Instalando o R	3
1.3 Instalando o RStudio	3
1.4 Diretório de trabalho	5
1.5 RStudio Cloud	5
1.6 Instalação de pacotes	7
1.7 Funcionalidades básicas	7
1.8 Variáveis no R	7
1.9 Vetores	7
1.10 Matrizes	7
1.11 Uso do for (loop)	7
1.12 Funções	7
2 Importando dados	9
2.1 Dados de diferentes formatos	9
2.2 Análise exploratória	9
2.3 Análise exploratória	9
2.4 Análise estatística de dados - Modelagem	10

3	Visualização de dados com o R	11
3.1	ggplot, anatomia e a estrutura data.frame	11
3.2	Funções geom e aes	11
3.3	Segmentação e junção de gráficos do ggplot	11
3.4	Diferentes tipos de gráficos	11
4	Análise de delineamento de experimentos	13
4.1	Delineamento inteiramente casualizados	13
4.2	Delineamento em blocos ao acaso	13
4.3	Fatorial	13
4.4	Parcela subdividida	13
5	Relatórios e artigos científicos com o rmarkdown	15
5.1	Relatório dinâmico com o Rmarkdown	15
5.2	Relatórios em word, pdf ou html	15
5.3	Apresentações em Power Point	15

Objetivo

O R tem se caracterizado por ser uma ferramenta completa para quem trabalha com pesquisa, seja aplicada ou teórica. As diversas áreas do conhecimento acabam convergindo para o R, por se tratar de uma linguagem moderna, dinâmica, colaborativa e integradora.

A pesquisa e o desenvolvimento tecnológico na área da suinocultura geram um volume considerável de informações que necessitam ser analisadas. O R é uma ferramenta que permite, ler, integrar e tratar grandes bancos de dados, é bem desenvolvido na área de visualização de dados e possui uma diversidade de métodos estatísticos implementados. Além disso, ferramentas para gerar relatórios, apresentações e até mesmo a criação de aplicativos que possibilitam a interação do usuário final, tornam o R uma ferramenta completa.

Sobre a Academia

A Academia Suína foi fundada em 2018 para revolucionar a disseminação da educação na Suinocultura Brasileira. Contamos os maiores professores da Suinocultura Internacional, representando mais de 28 universidades, e com um foco prático e de alto impacto ao nível de granja. (Academia-Suína, 2018)

Sobre o autor

Meu nome é Marcio, sou professor do Departamento de Estatística da UFRGS e sou um grande fã do R. Encarei esse desafio pois gosto de pensar que outras áreas, além da Estatística, possam usufruir das fantásticas ferramentas que essa linguagem oferece. Além disso, contribuindo para o desenvolvimento tecnológico da área da suinocultura, estaremos contribuindo para o desenvolvimento do país.

Chapter 1

Tutorial básico R e RStudio

1.1 Apresentação da linguagem R

R é uma linguagem de programação caracterizada como Software Livre sob os termos da *General Public License (GNU)* da *Free Software Foundation* no formato *open source*. É voltada a manipulação, análise e visualização de dados e tem como característica o aspecto colaborativo, sendo que as ferramentas desenvolvidas são compartilhadas online pelos desenvolvedores, podendo ter acesso a elas qualquer pessoa, sem restrições. Uma breve história do R pode ser encontrada no wikipedia.

1.2 Instalando o R

Para instalar no computador, O R deve ser baixado do CRAN.

Se o sistema operacional for Linux, uma versão *base* do R já vem instalada. No caso de outros sistemas operacionais, como o Windows, é necessário instalar o R base.

1.3 Instalando o RStudio

Como quase toda linguagem *Open Source* a utilização se dá por meio de linhas de comando. Para tornar a linguagem mais amigável aos usuários, várias IDEs (*integrated development environment*) são utilizadas. No caso do R, a mais desenvolvida e utilizada é o RStudio.

Uma versão *Free* do RStudio para o seu desktop pode ser baixado de <https://rstudio.com/products/rstudio/download/>.

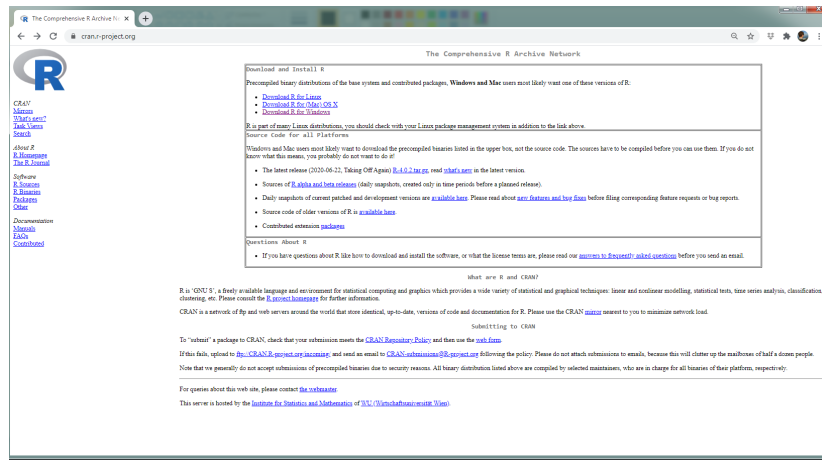


Figure 1.1: Comprehensive R archive network (CRAN)

Depois de instalar o R, o RStudio já estará integrado ao R e terá uma interface intuitiva e amigável ao usuário.

Ainda assim, é importante ressaltar que na linguagem R não encontraremos *botões* para realizar as análises.

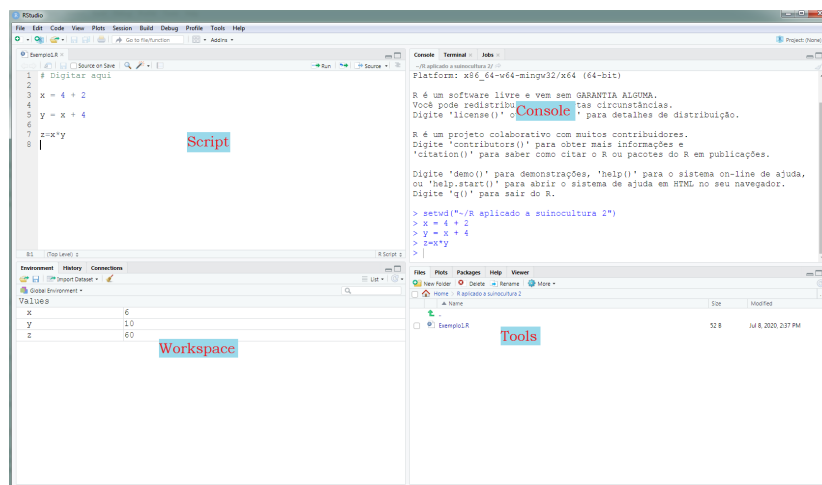


Figure 1.2: Interface do RStudio

1.4 Diretório de trabalho

Um ação importante que deve ser realizada pelo usuário é *setar o diretório* de trabalho. Para isso existem diferentes formas. Uma delas é usando a função `setwd()`.

```
setwd("~/R aplicado a suinocultura 2")
```

Outra opção é através do *Go to directory* que está disponível no Workspace do RStudio, conforme figura 1.3. Nessa opção o usuário escolhe o diretório de trabalho e depois usando a opção *set as working directory* esse diretório será “*setado*” como diretório de trabalho.

Todas os arquivos gerados, como figuras serão salvos nesse diretório. Para abrir um conjunto de dados, por exemplo, será muito mais simples se este também estiver salvo no mesmo diretório ou um sub-diretório. Além disso, toda vez que o RStudio for reinicializado, esse procedimento terá que ser feito.

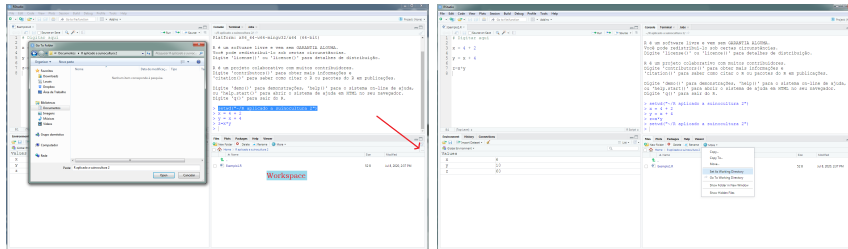


Figure 1.3: Diretório de trabalho

1.5 RStudio Cloud

Outra forma simples e prática para usar o R e o RStudio é usar o RStudio Cloud.

THE MISSION We created RStudio Cloud to make it easy for professionals, hobbyists, trainers, teachers and students to do, share, teach and learn data science. (RStudio-Cloud, 2020)

Na nuvem é possível usar o RStudio utilizando um login através da conta google, ou criar gratuitamente uma conta. Uma vez logado, escolhendo a opção *project* o usuário terá uma versão do RStudio perfeitamente funcional, que pode ser utilizada até no smartphone. Obviamente, é necessário ter conexão com a internet para que a ferramenta possa ser utilizada.

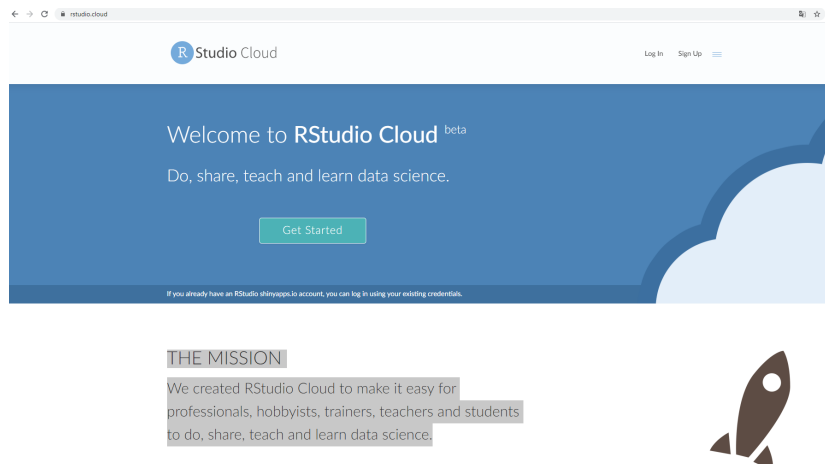


Figure 1.4: Rstudio Cloud

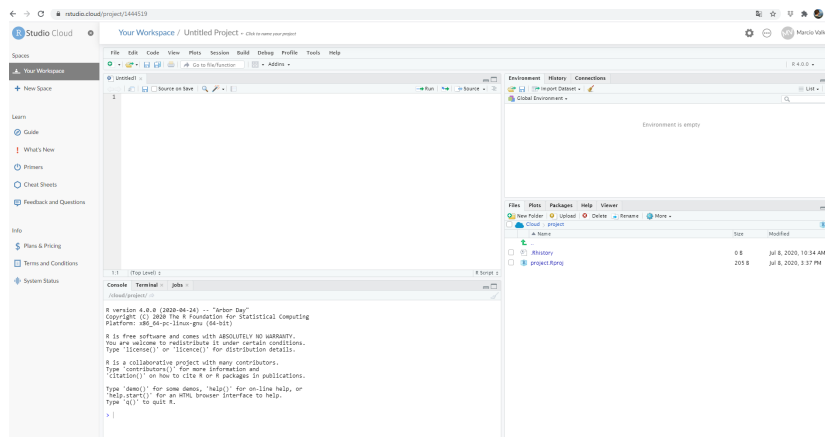


Figure 1.5: Rstudio Cloud

1.6 Instalação de pacotes

Na versão *base* do R, uma série de ferramentas, funções e métodos estatísticos são disponibilizados. Além disso, alguns pacotes também compõe a versão *base* do R. Depois de instalado, o usuário pode verificar quais pacotes estão instalados acessando o ícone *Packages* n

1.7 Funcionalidades básicas

1.7.1 Operações

1.7.2 Operadores lógicos

1.8 Variáveis no R

1.8.1 Operações

1.9 Vetores

1.9.1 Operações com vetores

1.10 Matrizes

1.10.1 Operações com Matrizes

1.11 Uso do for (loop)

1.12 Funções

Chapter 2

Importando dados

2.1 Dados de diferentes formatos

2.1.1 Dados no formato .txt

2.1.2 Dados no formato .csv

2.1.3 Dados no formato .xls e .xlsx

2.1.4 Dados do SAS, SPSS e STATA

2.1.5 Dados de um repositório web

2.2 Análise exploratória

2.2.1 Dados faltantes

2.3 Análise exploratória

2.3.1 Medidas de tendência central

Média, mediana, variância

2.3.2 Plot simple

2.3.3 Box plot simples

2.3.4 Box plot multiplas variáveis

2.3.5 Scatter plot

2.3.6 QQplot

2.4 Análise estatística de dados - Modelagem

2.4.1 Regressão linear simples

2.4.2 Regressão Múltipla

2.4.3 Gráficos de resíduos

2.4.3.1 studentized residuals

2.4.3.2 outliers

2.4.3.3 normalidade

2.4.4 Anova one way

Chapter 3

Visualização de dados com o R

3.1 ggplot, anatomia e a estrutura data.frame

3.2 Funções geom e aes

3.3 Segmentação e junção de gráficos do ggplot

3.4 Diferentes tipos de gráficos

3.4.1 Gráfico de relacionamento

3.4.2 Gráfico de barras

3.4.3 Gráfico de séries temporais

3.4.4 Mapas de calor

3.4.5 box plot

Chapter 4

Análise de delineamento de experimentos

4.1 Delineamento inteiramente casualizados

4.2 Delineamento em blocos ao acaso

4.3 Fatorial

4.4 Parcela subdividida

Chapter 5

Relatórios e artigos científicos com o rmarkdown

5.1 Relatório dinâmico com o Rmarkdown

5.2 Relatórios em word, pdf ou html

5.3 Apresentações em Power Point

Bibliography

Academia-Suína (2018). Quem somos | academia suína. <https://academiasuina.eadbox.com/quem-somos>. (Accessed on 07/09/2020).

RStudio-Cloud (2020). Rstudio cloud. <https://rstudio.cloud/>. (Accessed on 07/09/2020).