

ACADEMIA R

Manual Básico de Estatística Médica com a Linguagem R



Manual Básico de Estatística Médica com a Linguagem R

Introdução à análise de dados com a linguagem R e o RStudio

Henrique Alvarenga da Silva

Esse livro está à venda em <http://leanpub.com/estatisticabasicanorstudio>

Essa versão foi publicada em 2018-07-22



Esse é um livro [Leanpub](#). A Leanpub dá poderes aos autores e editores a partir do processo de Publicação Lean. [Publicação Lean](#) é a ação de publicar um ebook em desenvolvimento com ferramentas leves e muitas iterações para conseguir feedbacks dos leitores, pivotar até que você tenha o livro ideal e então conseguir tração.

© 2018 Henrique Alvarenga da Silva

Conteúdo

Dedicatória	1
Capa	2
Prefácio	3
Introdução	4
R - uma linguagem para análise de dados e gráficos	5
A importância da Estatística na Medicina	6
Breve História do R	7
Timeline	7
Comprehensive R Archive Network, CRAN.	7
Popularidade do R	7
RStudio	7
Referências.	7
Instalação do R e do RStudio	8
Instalando o R	8
Instalando o RStudio	8
Video da instalação no Youtube	8
RStudio - primeiros passos	9
Os Painéis do RStudio	9
R Script	9
R Notebook	10
Variáveis no R	11
Objetos no R	11
Variáveis no R	11
Vetores no R	13
Tipos de Dados	15
Variáveis categóricas (qualitativas)	15

CONTEÚDO

Variáveis quantitativas (numéricas)	16
Variáveis Lógicas	16
NA Values	17
Funções no R	18
Pacotes do R	20
Data Frames	21
O Operador \$	21
Lendo Dados	22
RStudio Projects	23
Working Directory	23
Projects do RStudio	23
Estatística Descritiva no R	24
Introdução	24
Parte 1 - Carregamento do dataset mpg	24
Parte 2 - Tabelas de frequências	24
Parte 3 - Medidas de Tendência Central	24
Parte 4 - Medidas de Dispersão	24
Manipulando dados no R	26
O conjunto de pacotes tidyverse	26
O pacote dplyr	26
A função 'filter()'	26
A função select()	26
Visualizando dados - Gráficos básicos do R	27
Introdução	27
Parte 1 - Visualizando dados categóricos.	27
Gráficos de Pizza - picharts	27
Parte 2 - Visualizando dados numéricos	28
Histogramas	28
Gráficos de densidade	28
Boxplots	28
Parte 3 - Visualizando relações entre variáveis	28
Correlação entre uma variável numérica e uma variável categórica	28
Correlação entre duas variáveis numéricas	28
Correlação	30
Introdução.	30
Análise Univariada e Bivariada	30

CONTEÚDO

Um scatter plot histórico	30
Covariância	30
Coefficiente de Correlação	30
Calculando o coeficiente de correlação no R	31
Modelos Científicos	32
Introdução	32
O que é um modelo científico?	32
Modelos de relação linear no R	32
O Coeficiente de Determinação - magnitude do efeito	32
As limitações e utilidades dos modelos	32
Distribuições de probabilidade	33
Introdução	33
A Distribuição uniforme e um jogo de dados	33
A distribuição normal	33
A curva normal padrão	33
A normalização e os z-scores.	33
Calculando as probabilidades - parte 1: pnorm(x)	34
Os parâmetros da curva normal - média e desvio padrão	34
Calculando as probabilidades - parte 2: pnorm(x, mean= , sd=)	34
Calculando probabilidades parte 3 - a função qnorm()	34
Simulando uma distribuição normal	34
Verificando a normalidade.	35
População e amostra	36
Distribuição amostral	36
Distribuição nula	37
Teste de Hipóteses	38
Introdução	38
A hipótese alternativa	38
A hipótese nula	38
Distribuição nula	38
Nível de significância do teste	38
Erros tipo I e tipo II	38
Teste unicaudal ou bicaudal	39
p-values	39
p-values	40
Introdução	40
Origens e confusões	40
Interpretação gráfica do valor de p	40

Dedicatória

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Capa

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Prefácio

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Introdução

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

R - uma linguagem para análise de dados e gráficos

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

A importância da Estatística na Medicina

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Breve História do R

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Timeline

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Comprehensive R Archive Network, CRAN.

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Popularidade do R

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

RStudio

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Referências.

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Instalação do R e do RStudio

Instalando o R

O R pode ser baixado para instalação no site do CRAN:

<https://cran.r-project.org/index.html>

Instalando o RStudio

O RStudio pode ser instalado a partir do site do RStudio:

<https://www.rstudio.com>

Video da instalação no Youtube

Há um pequeno video de instalação do R e do RStudio no Mac no link a seguir:

<https://youtu.be/zhhxPbGhI8Q>

RStudio - primeiros passos

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Os Painéis do RStudio

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

O Painei Console

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

O Painei Editor

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

O Painei Ambiente

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

O Painei Files

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

R Script

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

R Notebook

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Video Aula

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Variáveis no R

Objetos no R

“Everything that exists is an object” “Everything that happens is a function”

John Chambers, Creator of the S programming language

O R é uma linguagem de programação de alto nível e como tal, usa do conceito de objetos. As linguagens de programação desse tipo são chamadas de linguagens orientadas a objetos. Essa abstração simplifica muito a programação e torna muito mais fácil resolver problemas complexos. Os objetos mais comuns do R são chamados de variáveis.

Variáveis no R

Uma variável é um objeto que armazena dados, tais como valores numéricos, datas, caracteres, palavras, valores lógicos (veremos isso adiante) etc.

Operador de Atribuição

Para armazenar um dado numa variável usamos um operador especial, denominado de **operador de atribuição**:

```
1 <-
```

O operador de atribuição serve para atribuímos um dado a um objeto e tem a forma de uma seta para esquerda, formada pelo sinal de menor imediatamente seguida do sinal de menos <-.

Isso forma uma seta, indicando que o resultado da operação será colocado no objeto à esquerda da seta.

```
1 > x <- 2
2 > y <- 3
3 > z <- "Maria"
```

Observe que ao criar uma variável para armazenar uma palavra ou um caractere, é necessário colocar essa palavra entre aspas. O R entende que está entre aspas é uma palavra ou um texto e não um número.

Operadores Aritméticos

Os operadores mais comuns do R são os aritméticos:

operador	ação
+	somar
-	subtrair
/	dividir
*	multiplicar
^	elevar à uma potência

Veja nos exemplos abaixo algumas operações matemáticas simples:

```
1 > x <- 2
2 > y <- 3
3 > z <- (x + y)^2
```

Precedência de Operações

Assim como na matemática, também no R existem regras de precedência de operações com um detalhe: os parênteses sempre tem preferência, ou precedência, sendo usados para colocarmos as operações na ordem desejada. ou seja, as operações entre parênteses tem prioridade sobre outras operações. Veja o exemplo abaixo

```
1 < 4 * 3 + 2
2 [1] 14
3
4 > (4 * 3) + 2
5 [1] 14
6
7 > 4 * (3 + 2)
8 [1] 20
```

Comentários no R

Um outro símbolo importante no R é o hashtag #. Esse símbolo indica que o texto a seguir é um comentário, ou seja, é um texto para ser lido por humanos e que o computador simplesmente ignora.

```
1 # atribuindo o valor de 30 (anos) a variável idade
2 > idade <- 30
3
4 # Calculando a idade em meses e armazenando
5 # esse novo dado na variável idade.meses
6 > idade.meses <- 30 * 12
```

Estilos de nomeação de variáveis e objetos

Como você deve ter notado, variáveis devem ter nomes fáceis de serem compreendidos, nomes que mostrem o que significam. Se uma variável serve para armazenar a glicemia é mais adequado que essa variável seja denominada `glicemia` do que apenas `x`. Por outro lado, devemos usar nomes sucintos e evitar nomes grandes, `glicemia` é mais apropriado do que `niveis.de.glicose.dos.pacientes`, que é demasiadamente extenso.

Nomes compostos

Entretanto, quando houver necessidade de usar nomes compostos, o modo adequado é usar um ponto separando as palavras, tais como: `glico.fem` ou `idade.media`.

Evite usar o underline em variáveis, a separação de palavras pelo underline é geralmente usado para nomear arquivos tais como: `research_results_fase_1.csv`.

Evite usar maiúsculas para separar as palavras de uma variável, pois esse estilo geralmente é usado para nomear funções, tal como em `solveEquation`. Veja que as palavras são separadas pelo uso de uma maiúscula no início das palavras, exceto a primeira. Esse modo criar nomes é geralmente usado para nomearmos funções, portanto, vamos evitar fazer isso ao criarmos nomes de variáveis.

Regras para criação de nomes de objetos

Além dessa dica, existem regras formais para criar nomes de variáveis:

1. O nome de uma variável deve SEMPRE começar com uma letra
2. O nome de uma variável NÃO pode começar com números ou caracteres especiais
3. O nome de uma variável NÃO pode conter espaços
4. O nome de uma variável NÃO pode conter caracteres com acentos gramaticais.

Case Sensitive

Um ponto importante: o R é **case sensitive**, ou seja, maiúsculas e minúsculas são considerados caracteres diferentes: portanto as variáveis `idade` e `Idade` são diferentes. A dica é evitar usar maiúsculas em nomes de variáveis, para não criar confusão.

Vetores no R

Vetores são objetos fundamentais de todas linguagens computacionais. Um vetor é um conjunto de elementos da mesma natureza. Por exemplo, um conjunto de números, um conjunto de palavras, etc.

A forma mais comum de criar um vetor no R é através do uso do comando `c` que significa **combine**, como mostrado a seguir.

```
1  # criando um vetor numérico com idades dos pacientes
2  > idades <- c(45, 32, 24, 23, 55, 56)
3
4  # criando um vetor com nomes dos pacientes:
5  > nomes <- c("Eduardo", "José", "Antônio", "Pedro", "Maria", "Gustavo")
```

Tipos de Dados

A ciência depende de dados, que são gerados a partir de alguma forma de coleta. Tudo que for quantificado será armazenado numa variável. Uma variável deve ser entendida como um objeto que contém os resultados dessa coleta de dados.

Mas dados podem ser coletados de diferentes modos: algumas dados são provenientes de algo que foi contado, outros dados provêm de algo que foi medido. Contar e medir fornecem diferentes tipos de dados. Podemos por exemplo, contar o nº de pessoas com AIDS, o nº de eleitores de um determinado político, o nº de óbitos, de nascimentos etc. Por outro lado, podemos medir o perímetro cefálico de crianças recém-nascidas, o nível pressórico ou de glicemia de um grupo de pacientes, etc.

Variáveis com dados provenientes de contagem são denominadas variáveis **categóricas**, também chamadas de variáveis **qualitativas**, pois podem expressar uma qualidade (qual o nº de asiáticos no Brasil)

Variáveis com dados provenientes de medidas são denominadas variáveis **numéricas**, também chamadas de variáveis quantitativas.

As variáveis categóricas (ou qualitativas) se dividem em **nominais** e **ordinais**.

As variáveis numéricas por sua vez são tradicionalmente divididas em variáveis numéricas **discretas** e numéricas **contínuas**.

A necessidade de classificarmos as variáveis em diferentes tipos é devido ao fato de que o tipo de variável determina os tipos de operações matemáticas que podem ser realizadas e, conseqüentemente, as medidas estatísticas e os testes estatísticos que podem ser realizados.

Variáveis categóricas (qualitativas)

Algumas variáveis são nomes que expressão uma qualidade (que podem ter ou não uma ordenação). Em estatística essas variáveis são chamadas categóricas e podem se dividir em **nominais** ou **ordinais**. Variáveis nominais expressam qualidades, mas sem uma ordenação. Variáveis ordinais também expressão qualidades, mas tem uma ordenação. As variáveis categóricas não podem ser usadas em operações aritméticas. Não podemos, por exemplo, calcular a média desse tipo de variável. O que podemos fazer com variáveis categóricas é construir tabelas de frequências de cada categoria.

No R as variáveis categóricas são chamados de **factor**.

Os dados das variáveis que não possuem nenhuma ordenação, são chamadas de variáveis categóricas **nominais**. Quando uma variável categórica tem uma ordenação ela recebe o nome de variável **ordinal**. O R não tem nomes distintos para esses dois tipos de variáveis categóricas, sendo ambas classificadas no R como sendo do tipo **factor**.

Entretanto, é possível indicar para o R que a ordenação de uma variável é importante. A função `factor()` permite atribuir uma ordem às variáveis nominais, tornando-as assim variáveis ordinais. Isso é feito configurando o parâmetro de ordem para `TRUE` e atribuindo um vetor com a hierarquia de nível desejada. Sem essa ordenação vamos ter problemas ao fazer gráficos que precisam de uma ordem, pois não controlamos a ordem que será colocada no gráfico.

A linguagem R trabalha com diferentes tipos de variáveis para armazenar as diferentes categorias de dados. As variáveis no R podem ser classificadas como numéricas (para armazenar dados numéricos), caracteres (para armazenar palavras), datas etc. Para verificar o tipo de dado de uma variável ou vetor, basta usar a função `class()`. Vejamos alguns exemplos.

```
1 > nomes <- c("Henris", "Leo", "Gustavo")
2 > class(nomes)
3 [1] "character"
```

Variáveis quantitativas (numéricas)

As variáveis quantitativas (numéricas) são resultado de alguma medida realizada. Podem ser números inteiros (discretas) ou reais (contínuas). A grande diferença dessas variáveis com as qualitativas é que com as variáveis numéricas podemos fazer todas operações matemáticas: somar, dividir, calcular a média, a variância, o desvio padrão etc.

```
1 > idade <- c(45,10,12)
2 > class(idade)
3 [1] "numeric"
```

Variáveis Lógicas

Existem variáveis chamadas de lógicas. Variáveis lógicas armazenam o resultado de uma operação lógica. Operações lógicas são aquelas realizadas através de operadores lógicos: igual, maior, maior ou igual, menor, menor ou igual, etc.

operador	significado
<	menor que ..
<=	menor ou igual a ...
>	maior que ...
>=	maior ou igual a ...
==	exatamente igual a ...
!	não / negação
!=	não igual ou diferente de ...
	OR
&	AND

Variáveis lógicas são aquelas que armazenam resultados de operações lógicas. No exemplo uma operação lógica é realizada e seu resultado é colocado numa variável.

```
1 > resultado.1 <- (4 < 5)
2 > resultado.1
3 [1] TRUE
4
5 > resultado.2 <- (10 < 5)
6 > resultado.2
7 [1] FALSE
```

Veja que o valor da variável resultado.1 é TRUE, pois quatro é realmente menor que cinco. Da mesma forma o valor da variável resultado.2 é FALSE, pois 10 não é menor que 5.

Usando a função `class()`, podemos verificar que as variáveis resultado.1 e resultado.2 são do tipo logical (lógico).

```
1 > class(resultado.1)
2 [1] "logical"
3
4 > class(resultado.2)
5 [1] "logical"
```

Um ponto importante a ser memorizado é que é usual em linguagens de programação que TRUE tenha o valor de 1 e FALSE tenha o valor de 0.

portanto:

```
1 > resultado.1 * 5
2 [1] 5
3
4 > resultado.2 * 5
5 [1] 0
```

NA Values

É muito frequente que faltem dados em pesquisas. Às vezes uma questão de um questionário deixou de ser respondida, às vezes um dado não foi encontrado etc. Esses dados são representados no R como NA, que significa NOT AVAILABLE. É importante reconhecer a existência desses dados faltantes pois a presença desses dados faltantes pode impedir que sejam executados cálculos matemáticos. Afinal de contas, o que poderia significar $3 \times \text{NA}$? Experimente fazer essa conta no R.

Funções no R

As funções são a base de uma linguagem de programação. Assim como uma função matemática, na computação uma função recebe um conjunto de dados, processa esses dados e retorna um resultado.

Uma função $y=2x$ recebe como input um valor e dá como resultado o dobro desse valor.

A função $y=\text{sen}(x)$ recebe como input um determinado valor numérico e dá como resultado o seno desse valor.

A linguagem R tem inúmeras funções estatísticas que executam os inúmeros cálculos matemáticos necessários para resolvermos problemas estatísticos, desde os mais básicos até os mais complexos, facilitando bastante o ensino e a aprendizagem da estatística.

Argumentos de uma função

Toda função no R tem a seguinte estrutura: `nome(argumentos)`. Ou seja, o nome da função, entre parênteses, os argumentos da função. Observe que não há espaços entre o nome da função e os parênteses com os argumentos.

Denominam-se argumentos de uma função os dados que são inseridos na função para o cálculo desejado. Vejamos como exemplo função para extrair a raiz quadrada de um número: `sqrt()`.

O nome da função é `sqrt`, é uma abreviação de **square root**, ou raiz quadrada. Entre parênteses colocamos os argumentos, nesse caso, o número que desejamos que seja extraída a raiz quadrada. Portanto, para obter a raiz quadrada de 225 basta usar a função `sqrt()` com o argumento 225:

```
1 > sqrt(225)
2 [1] 15
```

Uma função pode ter mais de um argumento, o que é o mais comum na verdade. E alguns argumentos servem para indicar como a função deve se comportar. Uma função fundamental do R é a que lê os arquivos com dados, para que possamos fazer nossos cálculos. Um argumento dessa função é o nome do arquivo, outros argumentos são: se esse arquivo contém cabeçalho, se os dados estão separados por vírgula ou ponto e vírgula, se o decimal é uma vírgula ou um ponto etc. Para usar essa função devemos informar todos esses argumentos.

No exemplo abaixo, mostramos a função `read.csv()` e como essa função pode ser usada para ler um arquivo chamado *resultados.csv*, que contém cabeçalhos, cujos dados estão separados por vírgulas e cujos valores decimais são separados por um ponto:

```
1 read.csv(file = "resultados.csv",  
2         header = TRUE,  
3         sep = ",",  
4         dec = ".")
```

na.rm = TRUE

Um argumento importante de muitas funções é o `na.rm = TRUE`. a expressão `na.rm` é uma abreviação de REMOVE NOT AVAILABLE DATA, ou seja, remova os dados faltantes. Ao indicarmos que esse argumento é verdadeiro (TRUE), o R irá desconsiderar dados em branco ou faltantes ao fazer os cálculos. Sem esse argumento, frequentemente os cálculos não são realizados.

Algumas funções matemáticas comuns:

função	ação
<code>sqrt()</code>	calcula a raiz quadrada
<code>log()</code>	calcula o logaritmo natural
<code>abs()</code>	retorna o valor absoluto

Algumas funções estatísticas comuns:

função	ação
<code>mean()</code>	calcula a média de um conjunto de valores
<code>median()</code>	calcula a mediana de um conjunto de valores
<code>sum()</code>	calcula a soma de um conjunto de valores
<code>min()</code>	retorna o valor mínimo de um conjunto de valores
<code>max()</code>	retorna o valor máximo de um conjunto de valores
<code>var()</code>	calcula a variância de um conjunto de valores
<code>sd()</code>	calcula o desvio padrão de um conjunto de valores

Pacotes do R

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Data Frames

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

O Operador \$

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Datasets do R

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Datasets externos

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Lendo Dados

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Argumentos default

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

RStudio Projects

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Working Directory

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Projects do RStudio

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Criando um Projeto

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Video de Criação de Projetos no RStudio

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Estatística Descritiva no R

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Introdução

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Parte 1 - Carregamento do dataset mpg

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Parte 2 - Tabelas de frequências

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Parte 3 - Medidas de Tendência Central

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Média e Mediana

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Parte 4 - Medidas de Dispersão

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Amplitude

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Quartis, Percentil, Qualtil e Amplitude Interquartil

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Variância e Desvio Padrão

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Variância

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Desvio Padrão

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Propriedades do Desvio Padrão

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Conclusões finais

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Referências

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Manipulando dados no R

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

O conjunto de pacotes `tidyverse`

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

O pacote `dplyr`

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

A função `'filter()`

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

A função `select()`

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Visualizando dados - Gráficos básicos do R

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Introdução

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Objetivos de aprendizagem

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Carregamento dos dados

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Parte 1 - Visualizando dados categóricos.

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Gráficos de Pizza - picharts

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Gráficos de Barra - barplot

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Parte 2 - Visualizando dados numéricos

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Histogramas

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Gráficos de densidade

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Boxplots

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Parte 3 - Visualizando relações entre variáveis

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Correlação entre uma variável numérica e uma variável categórica

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Correlação entre duas variáveis numéricas

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Referências

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Correlação

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Introdução.

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Carregamento dos pacotes necessários

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Análise Univariada e Bivariada

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Um scatter plot histórico

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Covariância

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Coeficiente de Correlação

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Calculando o coeficiente de correlação no R

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Apêndice Extra - Transformações logarítmicas

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Referências

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Modelos Científicos

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Introdução

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

O que é um modelo científico?

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Modelos de relação linear no R

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

O Coeficiente de Determinação - magnitude do efeito

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

As limitações e utilidades dos modelos

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Distribuições de probabilidade

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

carregando os pacotes necessários para o capítulo

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Introdução

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

A Distribuição uniforme e um jogo de dados

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

A distribuição normal

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

A curva normal padrão

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

A normalização e os z-scores.

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Equação de normalização

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Interpretando o z-score:

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Calculando as probabilidades - parte 1: pnorm(x)

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Os parâmetros da curva normal - média e desvio padrão

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Calculando as probabilidades - parte 2: pnorm(x, mean= , sd=)

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Calculando probabilidades parte 3 - a função qnorm()

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Simulando uma distribuição normal

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Verificando a normalidade.

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

População e amostra

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Distribuição amostral

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Exemplo 1:

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Referências

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Distribuição nula

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Teste de Hipóteses

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Introdução

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

A hipótese alternativa

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

A hipótese nula

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Distribuição nula

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Nível de significância do teste

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Erros tipo I e tipo II

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Teste unicaudal ou bicaudal

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

p-values

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Conclusão

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Referências

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

p-values

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Introdução

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Origens e confusões

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Interpretação gráfica do valor de p

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.

Referências

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/estatisticabasicanorstudio>.