

# R Básico

## Aula 1

João Matheus  
Lineu Alberto

PET - Estatística  
Universidade Federal do Paraná

2019



**PET-Estatística**  
**UFPR**

# Sumário

- 1 Funções
- 2 Objetos e Classes
- 3 Leitura de Dados
- 4 Gráficos Simples
- 5 Links úteis



# Conceito de Funções

- Como já contextualizamos anteriormente: funções são códigos encapsulados para realização de uma tarefa.
- Utilizando uma função faz com que não seja necessário escrever e copiar o código todas as vezes que precisar executar aquela operação.
- As funções no R têm a seguinte estrutura: `nome_da_funcao(arg1, arg2, arg3)`
- Vejamos as funções básicas do R.



# Funções Básicas

Função	O que faz
<code>c()</code>	Cria um Vetor
<code>&lt;-</code>	Atribui um objeto para uma variável
<code>setwd()</code>	Muda o Diretório de Trabalho Atual
<code>getwd()</code>	Mostra o Diretório de Trabalho
<code>dir()</code>	Lista os Arquivos do Diretório de Trabalho Atual
<code>sessionInfo()</code>	Mostra algumas informações da sessão instalada
<code>install.packages()</code>	Instala um pacote
<code>library()</code>	Carrega um pacote
<code>require()</code>	Carrega um pacote
<code>help.start()</code>	Mostra um tutorial para novos usuários de R
<code>help()</code>	Mostra a documentação de um pacote ou função
<code>?</code>	Mostra a documentação de um pacote ou função



# Funções Básicas

Função	O que faz
help.search()	Faz uma pesquisa mais geral para o pacote ou função
??	Faz uma pesquisa mais geral para o pacote ou função
example()	Mostra exemplos de alguma função
print()	Imprime o resultado de uma variável
q()	Fecha a Sessão
rm(x,y)	Remove o objeto 'x' e 'y'
ls()	Exibe os objetos que foram armazenados
objects()	Exibe os objetos que foram armazenados
rm(list = ls())	Remove o objeto 'x' e 'y'
rm(list = objects())	Remove todos os objetos que foram armazenados
str()	Mostra a estrutura de um objeto



# Operadores e Funções Matemáticas

Função	O que faz
+	Adição
-	Subtração
*	Multiplicação
/	Divisão
^	Potência
**	Exponenciação
%%	Módulo
sqrt()	Raiz quadrada
factorial()	Fatorial
exp()	Exponencial
abs()	Valor absoluto

Função	O que faz
log()	Logaritmo
round()	Arredondamento
ceiling()	Arredondamento para cima
floor()	Arredondamento para baixo
length()	Tamanho de um vetor
mean()	Média
median()	Mediana
min()	Mínimo
max()	Máximo
var()	Variância
sd()	Desvio padrão



# Objetos, Classes e Estruturas

- Um objeto é uma variável capaz de armazenar qualquer valor ou estrutura de dados.
- A forma de um objeto define sua classe.
- No R temos as classes:
  - 1 numeric.
  - 2 integer.
  - 3 character.
  - 4 complex.
  - 5 logical.
- Com objetos podemos criar estruturas, no R existem as seguintes estruturas:
  - 1 Vetores.
  - 2 Matrizes.
  - 3 Listas.
  - 4 Data Frames.



# Vetores

- 1 Conjunto de valores.
- 2 Unidimensional.
- 3 Apenas uma linha ou apenas uma coluna.

Vector



# Matrizes

- 1 Conjunto de valores.
- 2 Bidimensional.
- 3 Valores na linha e na coluna.

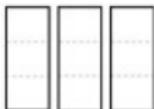
Matrix



# Data Frames

- 1 Um conjunto de vetores ou fatores, de mesmo comprimento, concatenados.
- 2 É o formato mais comum de armazenamento no R.

Data frame



# Listas

- 1 Pode combinar objetos de diferentes classes.
- 2 Trata-se de uma estrutura unidimensional.
- 3 Contudo, cada elemento de uma lista pode ter um diferente formato.
- 4 Por exemplo: o primeiro elemento pode ser um vetor, o segundo uma matriz, o terceiro um data frame, etc.

List



# Leitura de Dados

- Os formatos mais comuns de arquivos para leitura são:
  - Texto (txt).
  - Comma Separated Values (csv).
  - Planilhas Excel (xlsx).
- Todos estes formatos podem ser importados para o R.
- Outros formatos também são possíveis, mas abordaremos apenas estes por serem os mais comuns.
- Para que a importação seja feita de forma adequada há a necessidade de que os dados estejam tabulados e organizados de forma correta, isto é:

Observações nas LINHAS.

Variáveis nas COLUNAS.

Cada coluna representando uma única variável.



# Leitura de Dados

country	year	cases	population
Afghanistan	1999	745	12257071
Afghanistan	2000	2666	20995360
Brazil	1999	37737	172006362
Brazil	2000	80488	174004898
China	1999	212258	1272115272
China	2000	212766	128028583

variables

country	year	cases	population
Afghanistan	1999	745	12257071
Afghanistan	2000	2666	20995360
Brazil	1999	37737	172006362
Brazil	2000	80488	174004898
China	1999	212258	1272115272
China	2000	212766	128028583

observations

# Leitura de Dados

Existem diversos pacotes e funções para importação de dados, as 3 a seguir são apenas opções:

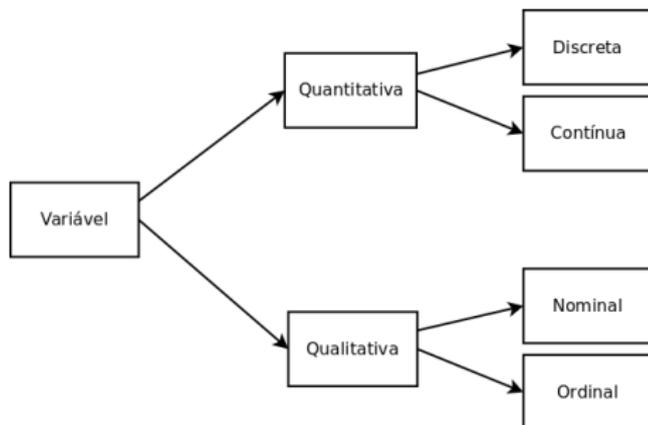
Leitura de arquivos `.txt`: `read.table()`.

Leitura de arquivos `.csv`: `read.csv()`.

Leitura de arquivos `.xlsx`: `readxl::read_excel()`.

## Gráficos Simples

- Visualização de dados é um dos mais interessantes, importantes e úteis tópicos em Estatística.
- Nesta aula mostraremos as mais básicas formas de visualização para os mais comuns tipos de variáveis em análises.
- Para escolher a visualização adequada é preciso ter em mente os tipos de variáveis:



# Gráficos Simples

Dentre as visualizações mais comuns, destacam-se:

**Gráfico de Barras:** Utilizado para representar frequências.

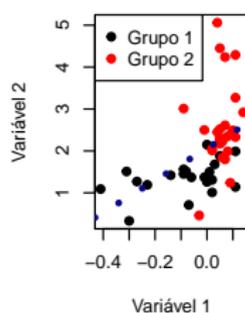
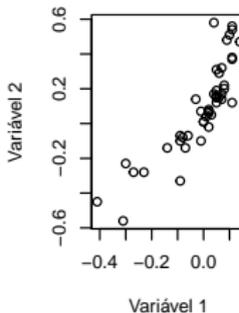
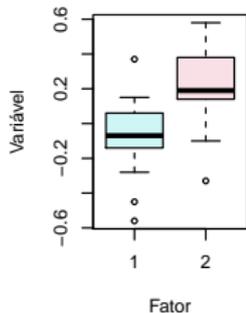
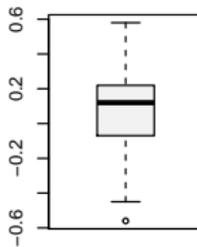
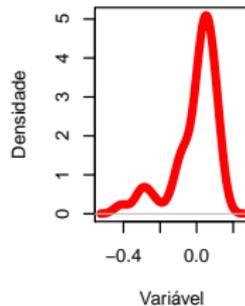
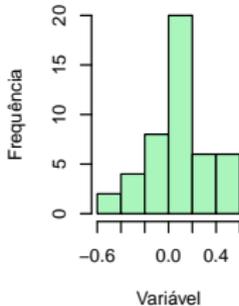
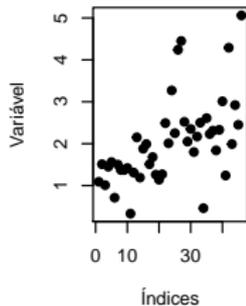
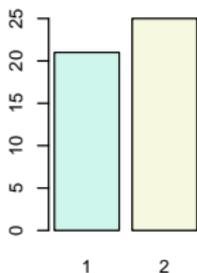
**Histograma:** Diagrama constituído por retângulos, em que a posição no eixo x representa o valor ou a amplitude de uma variável e a altura, o valor correspondente.

**Densidade:** Suavização da forma do histograma.

**Boxplot:** Gráfico de caixas; pode ser usado para uma variável numérica; ou para uma variável numérica em função de uma variável categórica, gerando um boxplot para cada nível do fator.

**Gráfico de Dispersão:** Gráfico de pontos com duas variáveis numéricas: uma no eixo x e outra no eixo y.

# Exemplos de Gráficos Básicos em R



## Comentários

- Experimente o comando `demo(graphics)` para ver mais gráficos com o R base.
- O R base é a forma mais simples e rápida de gerar visualizações, porém nem sempre é a mais elegante.
- Existem pacotes no R específicos para geração de gráficos.
- Utilizaremos, ainda neste curso, o `ggplot2`.

# Próxima Aula

## Parte 1

- 1 if, else, for, while
- 2 Família apply
- 3 Criação de funções

## Parte 2

- 1 Introdução ao Tidyverse
- 2 %>%
- 3 readr
- 4 tibble

## Links úteis

- Página do professor Fernando Mayer ([link](#)).
- Material sobre Funções Matemáticas e Estatísticas da EcoR ([link](#)).
- Material R Básico da Curso-R ([link](#)).



# PET-Estatística UFPR