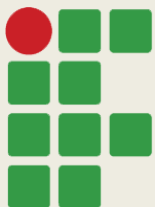
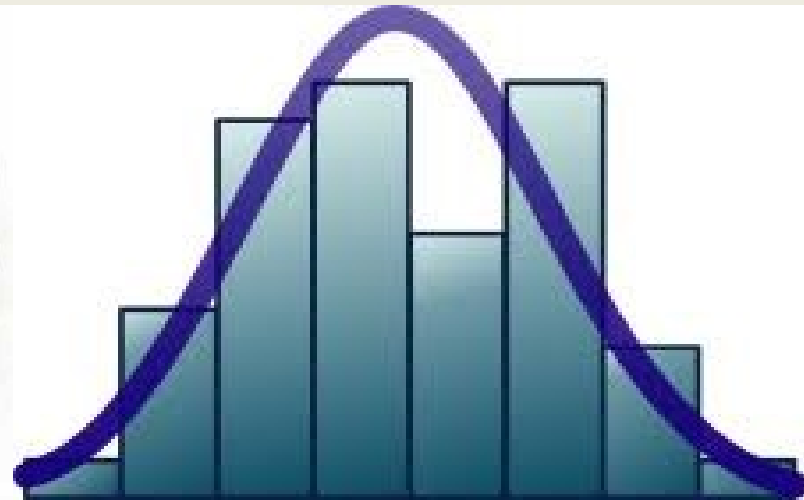


Probabilidade e Estatística



INSTITUTO FEDERAL
Catarinense
Campus Blumenau

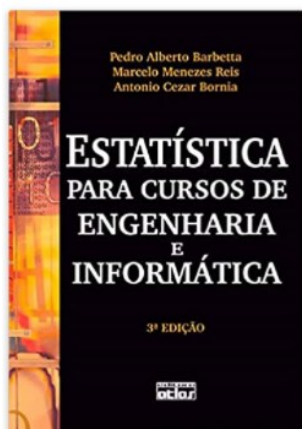
Professor Jeovani Schmitt



Probabilidade e Estatística

Aula 1

Conceitos básicos



- ✓ Perguntas iniciais e aplicações da Estatística nas diversas áreas
- ✓ Definição de Estatística
- ✓ A Estatística na Engenharia
- ✓ A Estatística e a Informática
- ✓ Modelos
- ✓ Subdivisões da Estatística
- ✓ Conceitos básicos
- ✓ Tipos de variáveis
- ✓ R e orientações para instalação

Capítulo 1

Perguntas iniciais

O que é Estatística?

Onde se utiliza?

Há necessidade de aprender Estatística?

Softwares que fazem análise Estatística?

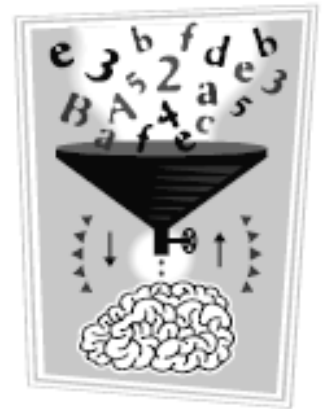
O papel da Estatística



Perguntas



Estudos



Dados

* A palavra Estatística vem do latim **status**, que significa Estado.

Síndrome de Burnout: o que é, sintomas, teste on-line e tratamento

A síndrome de Burnout, ou **síndrome do desgaste profissional**, é uma situação caracterizada por exaustão física, emocional ou mental que surge geralmente devido ao acúmulo de estresse no trabalho ou relacionado aos estudos, e que ocorre com mais frequência em profissionais que têm que lidar com pressão e responsabilidade constante, como professores ou profissionais de saúde por exemplo.



www.Chafic.com.br

Suporte e formação continuada para educadores

Brasília-DF | chafic@chafic.com.br

QUESTIONÁRIO JBEILI PARA IDENTIFICAÇÃO PRELIMINAR DA BURNOUT

Elaborado e adaptado por Chafic Jbeili, inspirado no Maslach Burnout Inventory – MBI

Obs.: este instrumento é de uso informativo apenas e não deve substituir o diagnóstico realizado por Médico ou Psicoterapeuta.

MARQUE "X" na coluna correspondente:

1- Nunca | 2- Anualmente | 3- Mensalmente | 4- Semanalmente | 5- Diariamente

Psicologia – Síndrome de *Burnout*

MARQUE "X" na coluna correspondente:

1- Nunca | 2- Anualmente | 3- Mensalmente | 4- Semanalmente | 5- Diariamente

Nº	Características psicofísicas em relação ao trabalho	1	2	3	4	5
1	Sinto-me esgotado(a) emocionalmente em relação ao meu trabalho					
2	Sinto-me excessivamente exausto ao final da minha jornada de trabalho					
3	Levanto-me cansado(a) e sem disposição para realizar o meu trabalho					
4	Envolver-me com facilidade nos problemas dos outros					
5	Trato algumas pessoas como se fossem da minha família					
6	Tenho que desprender grande esforço para realizar minhas tarefas laborais					
7	Acredito que eu poderia fazer mais pelas pessoas assistidas por mim					
8	Sinto que meu salário é desproporcional às funções que executo					
9	Sinto que sou uma referência para as pessoas que lido diariamente					
10	Sinto-me com pouca vitalidade, desanimado(a)					
11	Não me sinto realizado(a) com o meu trabalho					
12	Não sinto mais tanto amor pelo meu trabalho como antes					
13	Não acredito mais naquilo que realizo profissionalmente					
14	Sinto-me sem forças para conseguir algum resultado significativo					
15	Sinto que estou no emprego apenas por causa do salário					
16	Tenho me sentido mais estressado(a) com as pessoas que atendo					
17	Sinto-me responsável pelos problemas das pessoas que atendo					
18	Sinto que as pessoas me culpam pelos seus problemas					
19	Penso que não importa o que eu faça, nada vai mudar no meu trabalho					
20	Sinto que não acredito mais na profissão que exerço					
Totais (multiplique o numero de X pelo valor da coluna)						
Score						

Psicologia – Síndrome de *Burnout*

Resultados:

De 0 a 20 pontos: Nenhum indício da Burnout.

De 21 a 40 pontos: Possibilidade de desenvolver Burnout, procure trabalhar as recomendações de prevenção da Síndrome.

De 41 a 60 pontos: Fase inicial da Burnout, procure ajuda profissional para debelar os sintomas e garantir, assim, a qualidade no seu desempenho profissional e a sua qualidade de vida.

De 61 a 80 pontos: A Burnout começa a se instalar. Procure ajuda profissional para prevenir o agravamento dos sintomas.

De 81 a 100 pontos: Você pode estar em uma fase considerável da Burnout, mas esse quadro é perfeitamente reversível. Procure o profissional competente de sua confiança e inicie o quanto antes o tratamento.

ATENÇÃO: este instrumento é de uso informativo apenas e não deve substituir o diagnóstico realizado por médico ou psicoterapeuta de sua preferência e confiança.

Onde mais podemos utilizar a Estatística?

Na comunicação de pesquisas e trabalhos acadêmicos, tais como:

- Artigos científicos;
- Trabalhos de conclusão de cursos;
- Dissertações e Teses

Definição de Estatística

A Estatística envolve técnicas para coletar, organizar, descrever, analisar e interpretar dados, ou provenientes de experimentos, ou vindos de estudos observacionais. (BARBETTA, p. 11)





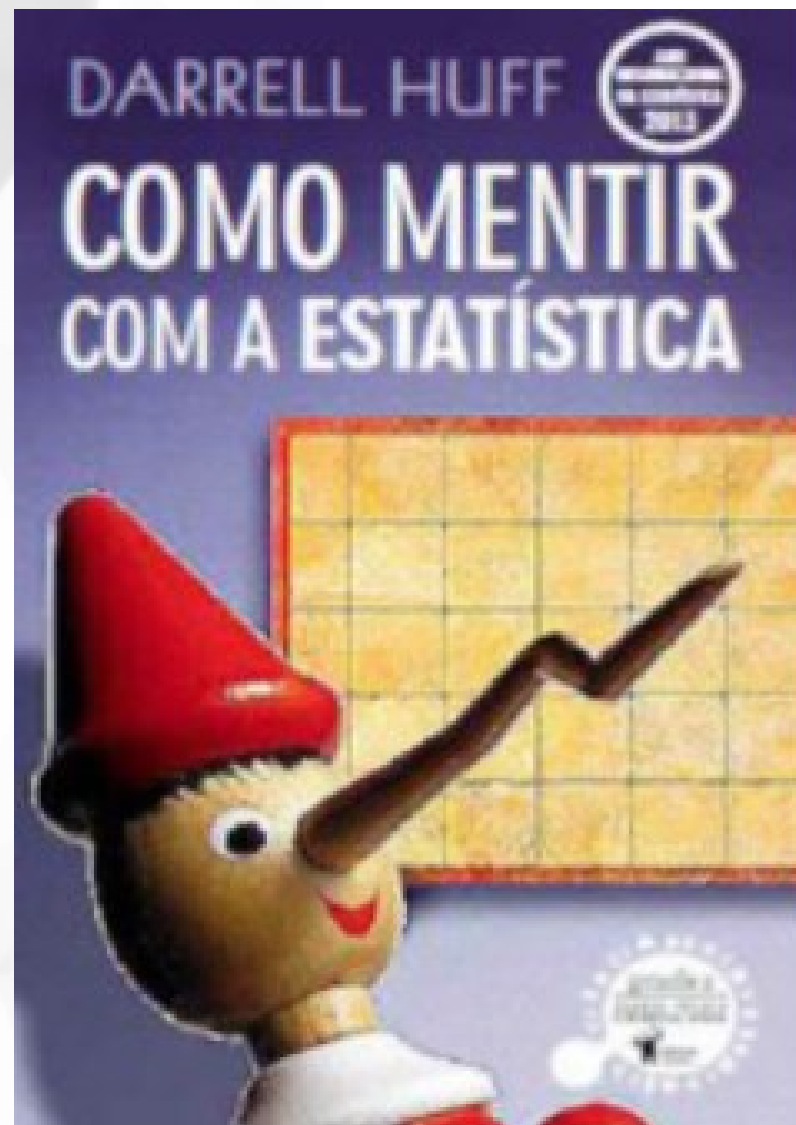
“Os números não mentem.”

**“É possível contestar as palavras mas
não é possível contestar os números.”**

COMO MENTIR COM ESTATÍSTICAS



Darrell Huff



Como contestar Estatísticas

- ✓ Quem é que diz isso?
- ✓ Como é que ele/ela sabe?
- ✓ O que é que está faltando?
- ✓ Alguém mudou de assunto?
- ✓ Isso faz sentido?

A Estatística na Engenharia

- Avaliação da qualidade de produtos e processos;
- Realização de experimentos planejados;

Exemplo: Projeto de lâmpadas LED

Prof. Thiago Farias dos Santos (IFC)



Objetivo: Analisar os níveis de distorções harmônicas frente a associação de uma quantidade considerável de lâmpadas baseadas na tecnologia LED. Serão ensaiadas 600 lâmpadas (200 de 9W, 200 de 12W e 200 de 15W).

Observação: Será considerada a associação mista de tais potências e quantidades (exemplo: 100 de 9W + 100 de 15W).

A Estatística e a Informática

- Análise de desempenho de sistemas computacionais;
- Estudos de simulação;
- Banco de dados x Estatística x Inteligência artificial
- *Machine learning*
- *Data Science*

A Estatística e a Informática

INTRODUCTION TO DATA MINING

The evolution of technology helped internet expand lightning fast. Over time, internet access became accessible to more and more people. This led to the development of million websites and the use of databases for storing these data. The creation of the first commercial and social webpages created the need for storing and managing large amount of data.

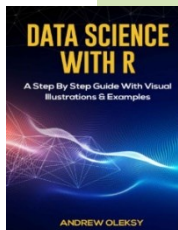
Today, the amount of available data is huge and is growing exponentially every day. The need for minimizing the costs for collecting and storing these data was one of the biggest reasons for the growth of this scientific field.

The huge amount of data stored in databases and data warehouses could not be utilized as is. In order to get useful conclusions, some necessary actions are required in order to structure the data. On this chapter we will view which are the fundamental stages in order to extract valuable and usable information from data.

.1 DATA SCIENCE

Data Science is a new term, which came to replace former terms like Knowledge Discovery in Database or Data Mining.

FONTE: ANDREW, OLEKSY. [Data Science with R](#), 2018.



Modelos: uma representação da realidade

- Determinístico: Exemplo: Lei de Ohm

Dados a tensão (x_1) e a resistência (x_2) de um circuito simples, podemos calcular o fluxo da corrente elétrica (y) por:

$$y = \frac{x_1}{x_2}$$

$$\textit{fluxo corrente elétrica} = \frac{\textit{tensão}}{\textit{resistência}}$$

Modelos: uma representação da realidade

- Probabilístico

$$p(k) = \frac{\lambda^k \cdot e^{-\lambda}}{k!} \quad (k = 0, 1, 2, \dots e \approx 2,72)$$

λ taxa média chegadas por segundo

k pacotes

$p(k)$ probabilidade de chegar exatamente k pacotes num dado segundo

Esse tipo de modelo pode auxiliar o projetista a planejar a capacidade de um sistema computacional

Modelos: uma representação da realidade

- Empíricos: construídos com base em observações reais sobre o problema em estudo.

Exemplo: Conhecer a relação entre a **resistência à compressão de um concreto** e seu **tempo de hidratação**

Modelos: uma representação da realidade

Exemplo: Conhecer a relação entre a **resistência à compressão de um concreto** e seu **tempo de hidratação**

Corpo de prova	Tempo de hidratação (dias)	Resistência (MPa)
1	10	11,3
2	11	12,1
3	12	16,4
4	13	16,3
5	14	20,2
6	15	20,5
7	16	25,0
8	17	26,4
9	18	26,2
10	19	28,4
11	20	30,2

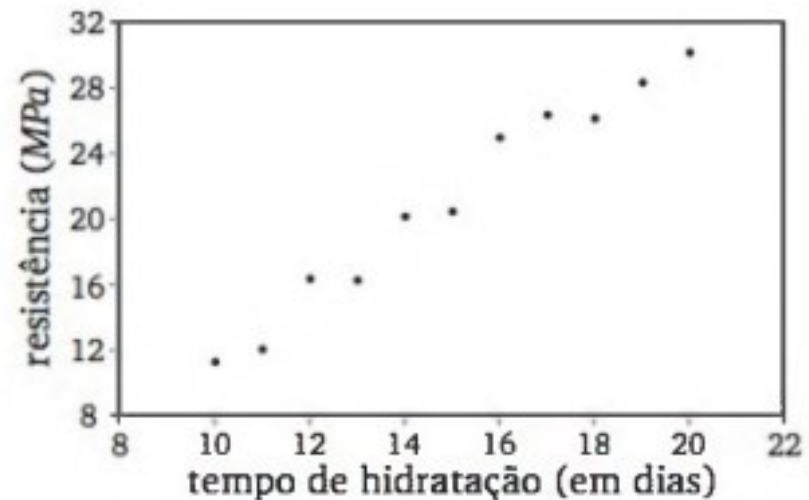


Figura 1.2 Resultados de um experimento sobre resistência à compressão de concreto, em função do tempo de hidratação – dados e gráfico.

Modelos: uma representação da realidade

Exemplo: Conhecer a relação entre a **resistência à compressão de um concreto** e seu **tempo de hidratação**

$$\hat{y} = 8,01 + 1,95x$$

onde x é o tempo de hidratação e \hat{y} é o valor de resistência predito pelo modelo.

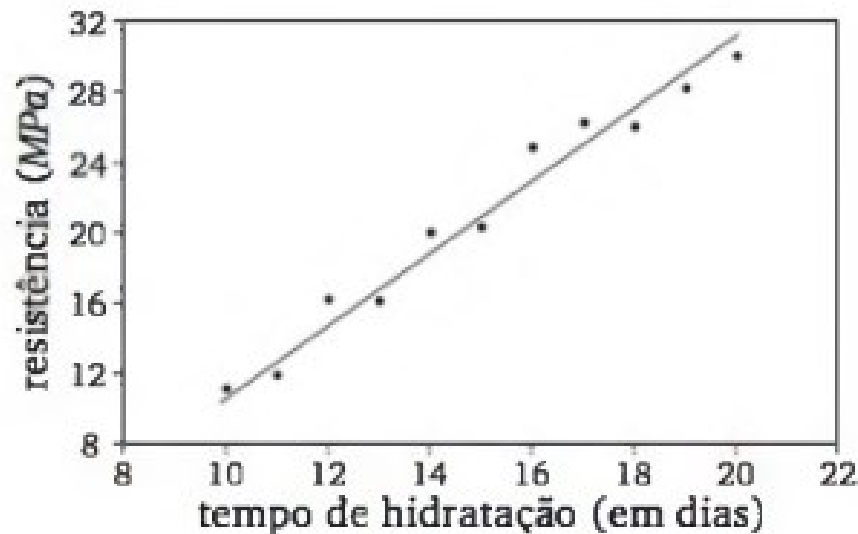


Figura 1.3 *Exemplo de um modelo empírico para explicar a resistência de um concreto, em função do tempo de hidratação.*

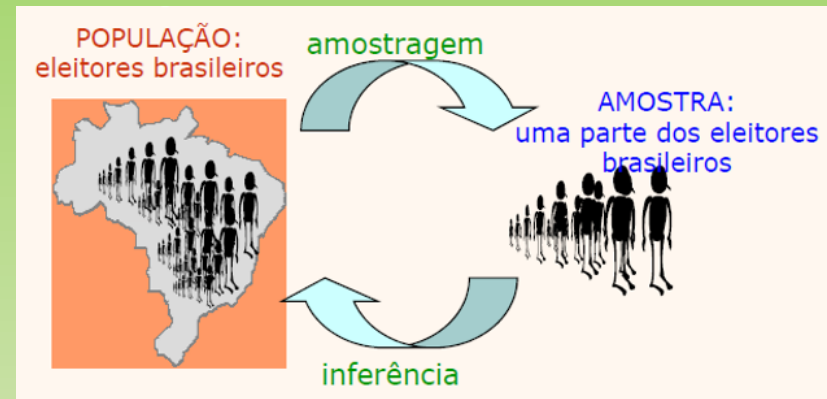
Subdivisões da Estatística

- **Estatística Descritiva**
(análise exploratória de dados)



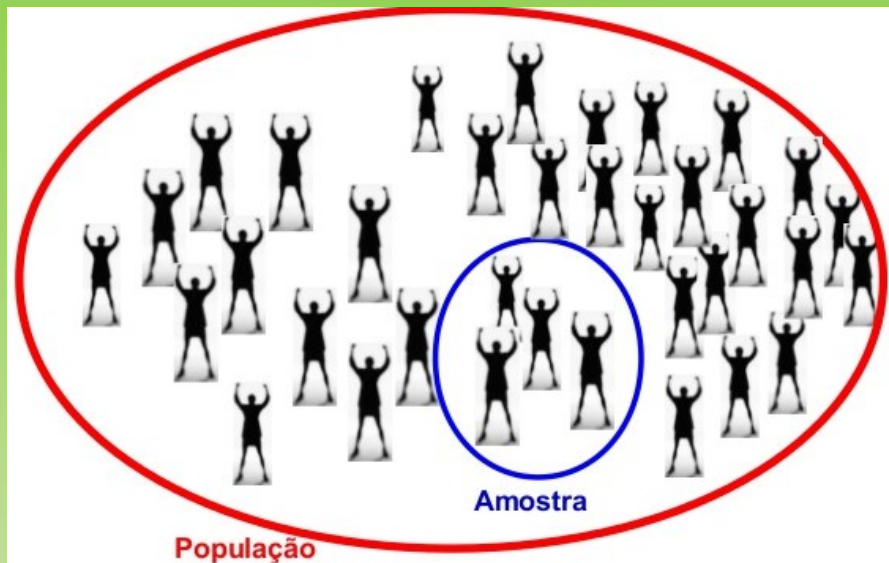
“É uma tentativa de captar a essência das informações contidas nos dados, através da descrição adequada em tabelas e, principalmente, em gráficos. (BARBETTA, p. 16)

- **Inferência Estatística**



É quando o pesquisador utiliza os resultados da amostra para estender para a população. (generalizar os resultados da parte para o todo)

Conceitos básicos



AMOSTRA: parte dos elementos de uma população.



AMOSTRAGEM: processo de seleção da amostra

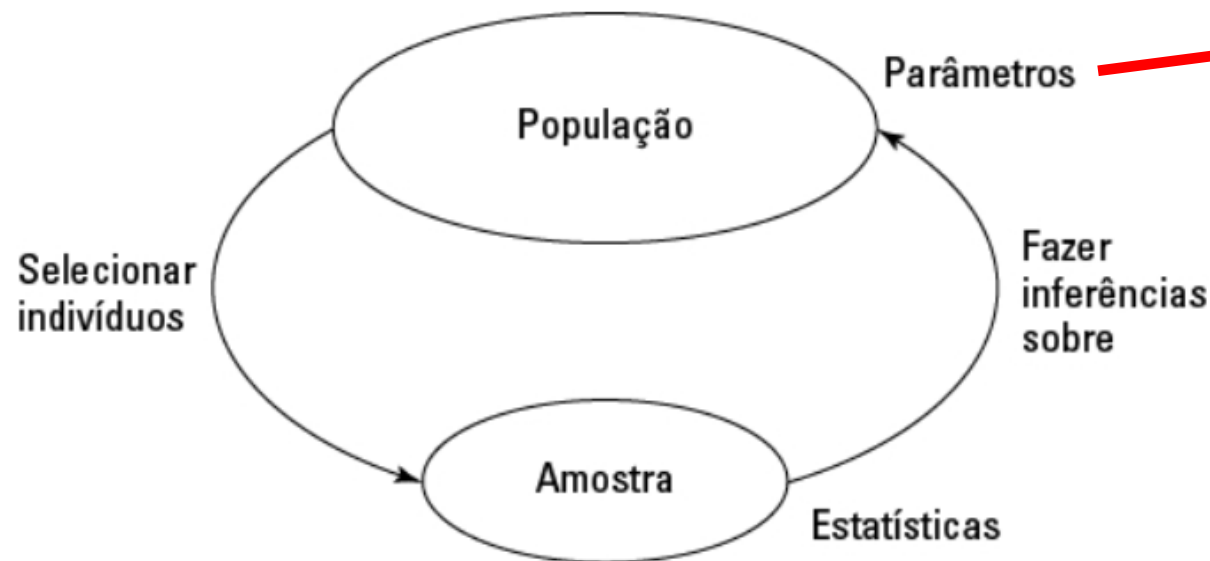
POPULAÇÃO: conjunto de elementos que formam o universo de nosso estudo que são passíveis de serem observados, sob as mesmas condições.



CENSO: quando toda a população é pesquisada

A essência de uma análise estatística é tirar conclusões sobre uma população, ou universo, com base em um amostra de observações.

Conceitos básicos



*letras gregas do
alfabeto grego
 μ, σ, π, ρ*

FIGURA 1-1:

O relacionamento entre populações, amostras, parâmetros e estatísticas.

SCHMULLER, J., p. 16)

PARÂMETRO: é uma medida numérica que descreve uma característica de uma **população**. São valores fixos, geralmente desconhecidos e usualmente representados por letras gregas

EXEMPLOS

μ = média populacional,

σ = desvio padrão populacional,

π = proporção populacional

ρ = coeficiente de correlação populacional

Conceitos básicos

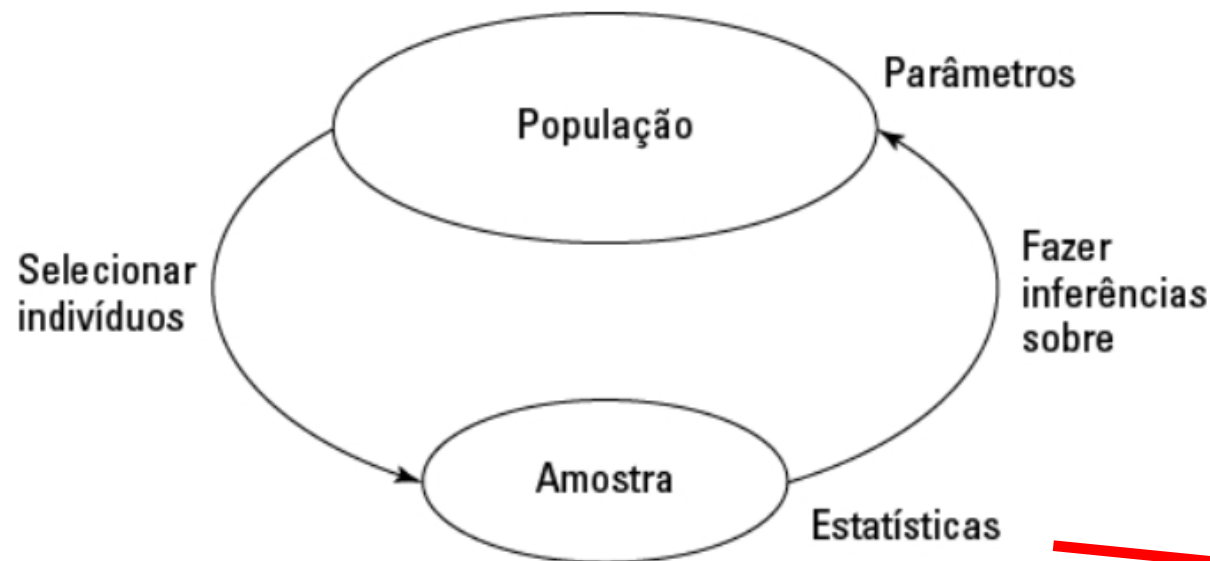


FIGURA 1-1:

O relacionamento entre populações, amostras, parâmetros e estatísticas.

*letras do
alfabeto latino
 \bar{x}, s, p, r*

SCHMULLER, J., p. 16)

ESTATÍSTICA: medida numérica que descreve uma característica de uma **amostra**.
Representada por letras do alfabeto latino

EXEMPLOS

\bar{x} = média amostral,

s = desvio padrão amostral,

p = proporção amostral

r = coeficiente de correlação amostral

Tipos de pesquisa

- **Levantamento (observacional ou *survey*)**

Um ou mais grupos são observados e suas características de interesse registradas (observadas ou medidas) para análise, mas sem manipulação.

- **Experimento**

Grupos de indivíduos (ou animais, ou objetos) são manipulados para se avaliar o efeito de diferentes tratamentos.

Há interferência do pesquisador na formação dos grupos de comparação. (uma manobra controlada – um procedimento, um tratamento, um medicamento).

Variáveis

Dependente: é aquilo que o pesquisador “mede”

Independente: é aquilo que o pesquisador “manipula” ou observa

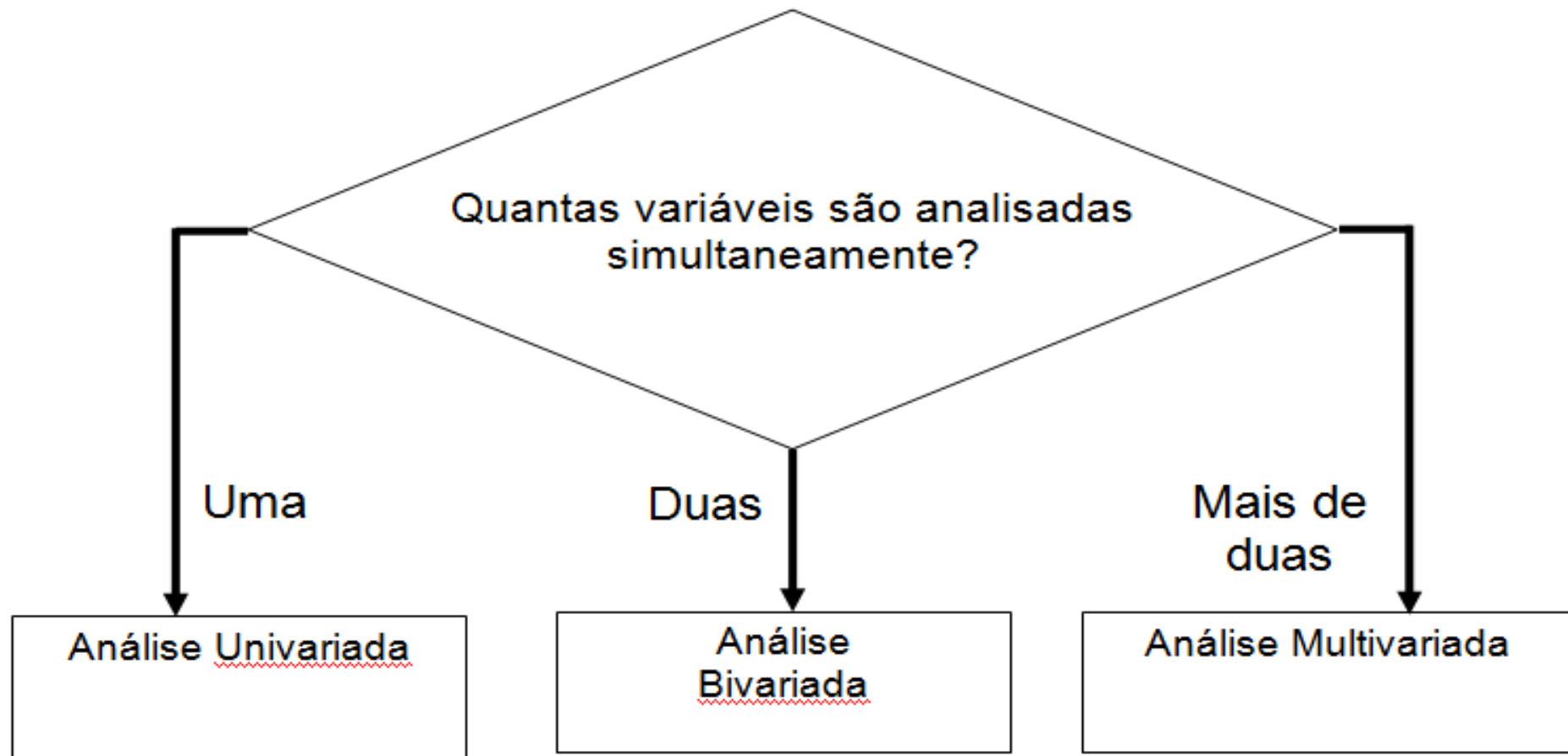


Variáveis

- Classificação por nível de mensuração: quantidade de informação que a variável apresenta.

Variáveis	Tipos	Exemplos
Qualitativas ou Categóricas	Nominal	cor dos olhos, sexo, estado civil
	Ordinal	nível de escolaridade, estágio da doença
Quantitativas	Discretas	número de filhos por casal, quantidade de leitos
	Contínuas	medida de altura e peso, taxa de glicose

Número de variáveis x Análise



Questionário da turma

Responder ao questionário para ilustrar uma coleta de dados primários.

Questionário

1 – Seu curso?

() BCC

() BEE

2 – Peso: _____ kg

3 – Altura: _____ metros

4 – Idade: _____ anos

5 – Quantos irmãos você tem? _____

6 – Cidade onde você mora?

7 – Você trabalha?

() Sim

() Não

8 – Você tem notebook?

() Sim

() Não

9 – Você gosta de Estatística?

() Sim

() Não

() Não sei

10 – Assinale seu nível de satisfação com o curso

() Muito insatisfeito

() Insatisfeito

() Neutro

() Satisfeito

() Muito Satisfeito

Exercício 1

Classificar as variáveis do questionário da turma

(1) Qualitativa nominal

(2) Qualitativa ordinal

(3) Quantitativa discreta

(4) Quantitativa contínua

☐ curso

☐ peso

☐ altura

☐ idade

☐ número de irmãos

☐ cidade mora

☐ trabalha

☐ possui notebook

☐ gosta de Estatística

☐ satisfação geral com o curso

Como analisar os dados de sua pesquisa?



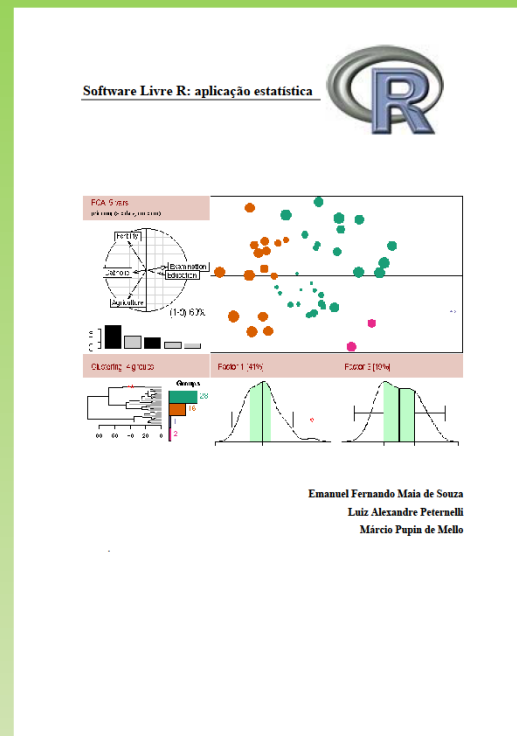
R

R

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

RStudio

<http://rstudio.org>



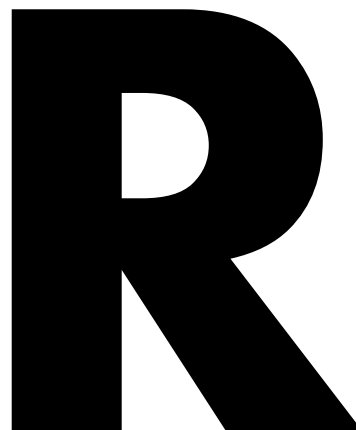
RStudio



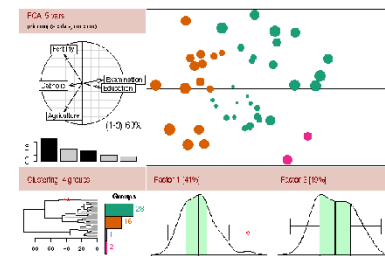
O que é o R?

- ✓ Uma linguagem?
- ✓ Um programa?
- ✓ Um Software?
- ✓ Uma Plataforma?
- ✓ Um Ambiente?





Software Livre R: aplicação estatística



Emanuel Fernando Maia de Souza
Luiz Alexandre Peterelli
Márcio Pupin de Mello

- O R é uma linguagem orientada a objetos criada em 1996 por Ross Ihaka e Robert Gentleman que aliada a um ambiente integrado permite a manipulação de dados, realização de cálculos e geração de gráficos.

R pode ser obtido gratuitamente em <http://cran.r-project.org>, onde é apresentado em versões de acordo com o sistema operacional UNIX, Windows ou Macintosh

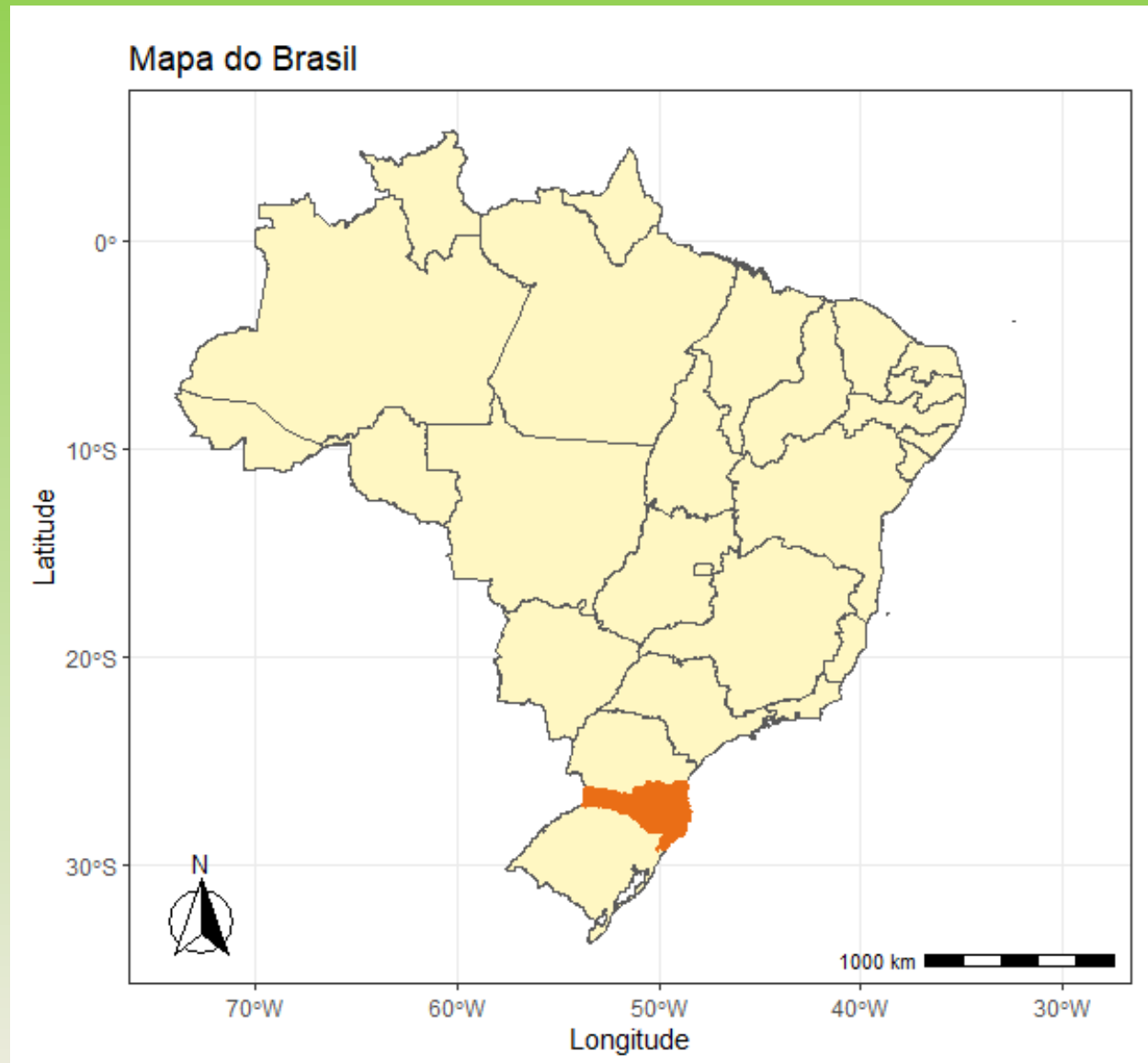
CRAN = The Comprehensive R Archive Network

Vantagens / motivações para uso do R

- ✓ A existência de uma comunidade grande e bastante entusiasmada, que compartilha conhecimento todo o tempo;
- ✓ o R é gratuito, *open source* (código aberto), de modo que você não precisa comprar licenças de software para instalá-lo;
- ✓ Tem inúmeras bibliotecas (pacotes) em estatística, *machine learning*, visualização, importação e tratamento de dados;
- ✓ Ferramentas poderosas para comunicação dos resultados da sua pesquisa, seja em forma de um website ou em pdf.;

Exemplo de gráfico no R destacando SC

<https://www.instagram.com/estatisticalivre/>



Exemplo de gráfico no R destacando SC

```
# MAPA COM OS DADOS DO IBGE
# https://www.instagram.com/p/CkijmdqNHSZ/

rm(list=ls(all=TRUE))
getwd()

setwd("D:/Users/Dell/Documents/Jeovani/AULAS/PROBABILIDADE E ESTATISTICA 2023")

library(ggplot2)
library(geobr)

# install.packages("geobr")
# install.packages("ggspatial")
# install.packages("showtext")
# install.packages("plotly")

library(geobr)
library(ggspatial)
library(showtext)
library(plotly)

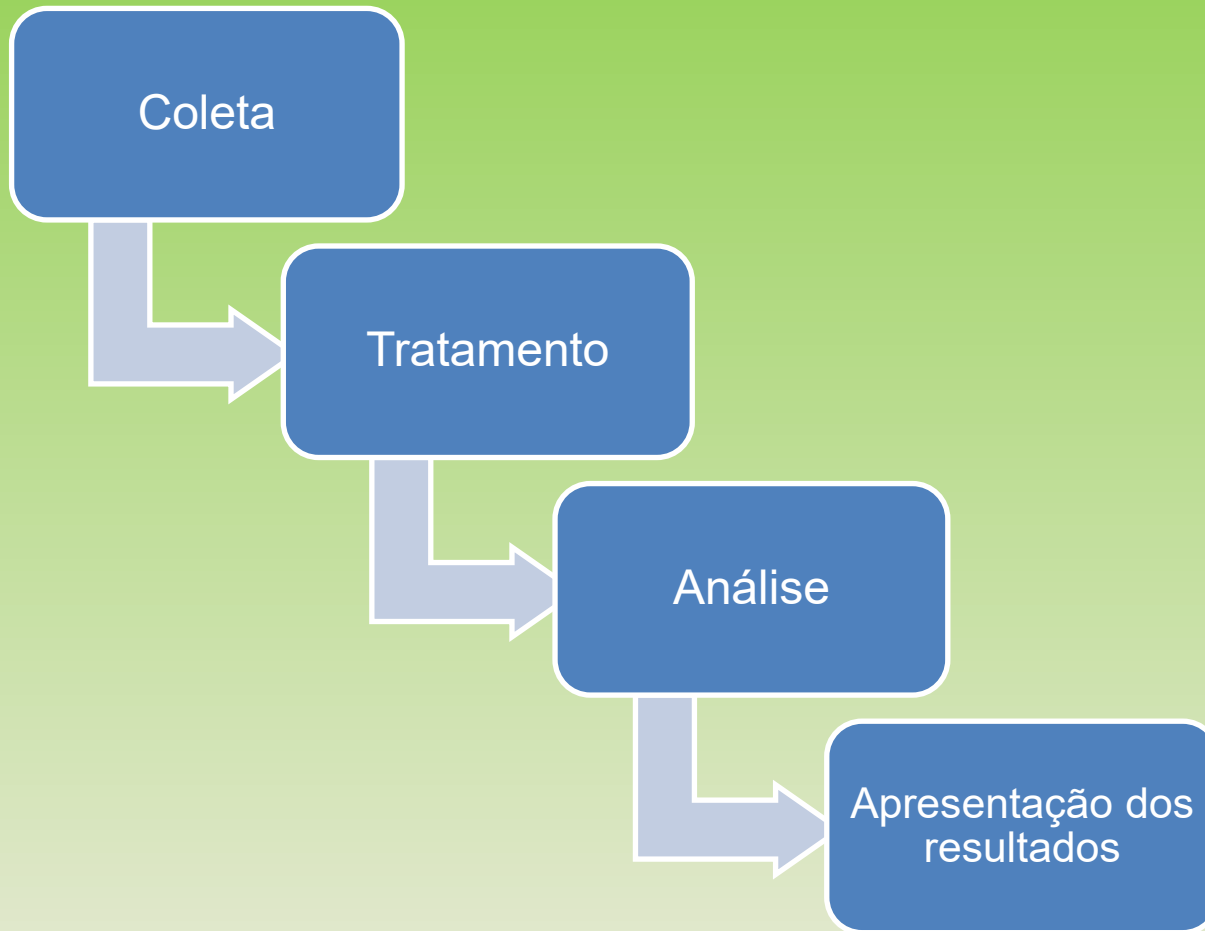
# Acessando os dados do IBGE

sc <- read_state(code_state = "SC", year = 2010)
br <- read_state(code_state = "all", year = 2010)

# Plotando o MAPA

ggplot(br) +
  aes(group = code_state) +
  geom_sf(size = 0.1, fill = "#FFF7c2") +
  geom_sf(aes(group=code_state), data=sc, fill="#EA6E16", col="#EA6E16") +
  labs(x = "Longitude", y = "Latitude", title = "Mapa do Brasil") +
  annotation_north_arrow(style = north_arrow_fancy_orienteering) +
  annotation_scale(location="br", height = unit(0.15, "cm")) +
  theme_bw()
```

O R permite integrar as etapas de *data analysis*:



Instalar o R e RStudio