

Introdução À Análise E Visualização De Dados Através Do Software R



Dhara Avelino
Mestranda em Genética UFRJ

Por que
estudar
análise de
dados?



Por que estudar análise de dados?

Análise crítica

Autonomia

Qualidade do
seu trabalho

Ampliar
possibilidades



Objetivos:



Etapas para uma boa análise de dados



Como produzir gráficos bonitos e informativos



Onde buscar ajuda

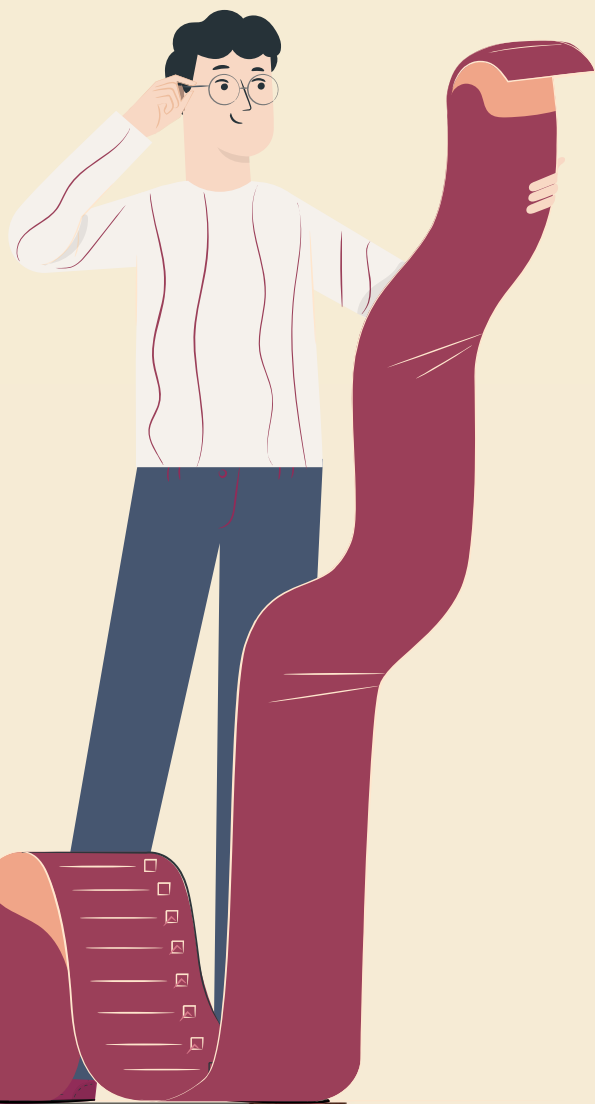


Por onde começar?

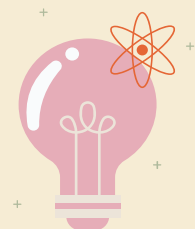


A screenshot of a web browser window. The browser's address bar shows a URL. The main content area displays a circular profile picture of a woman with long dark hair, smiling. To the right of the photo, the name 'Dra. Letícia Raposo' is displayed in a dark purple box, followed by the URL <https://leticia raposo.netlify.app/>. At the bottom of the browser window, the URL <https://www.datacamp.com/> is visible, along with the DataCamp logo, which consists of a green stylized 'd' followed by the word 'datacamp' in black lowercase letters.

Etapas da pesquisa:



Leia a literatura



Gere uma ideia



Gere uma hipótese



Faça o
delineamento do
estudo



Realize o estudo

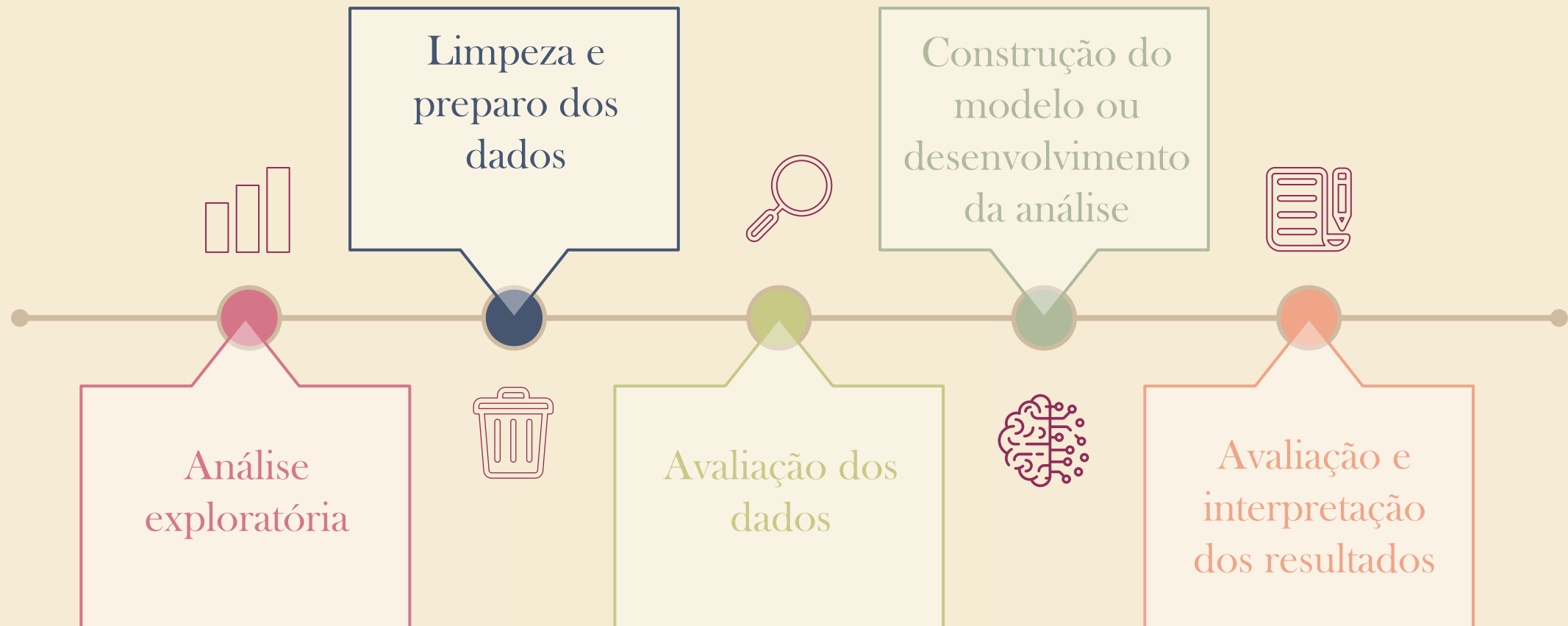


Analise os dados



Decida sobre a
veracidade da resposta

Etapas da análise de dados:



O que é uma variável?

Resumo de Bioestatística

O que é uma variável?



variável

adjetivo de dois gêneros

1. sujeito a variações ou mudanças; que pode variar; inconstante, instável.
"clima v."
2. que pode assumir qualquer um dos valores em um conjunto de valores.
"na equação, certos símbolos têm valor v."

Características dos elementos de uma amostra



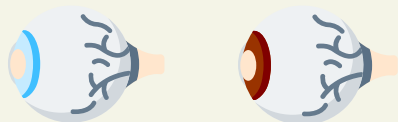


Variações

Qualitativa

Nominal

Classificação relacionada à característica
Não há relação de grandeza



Ordinal

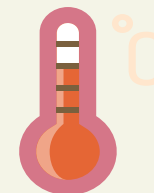
Há relação de grandeza entre as ordens porém sem intervalos iguais



Quantitativa

Intervalar

Os intervalos são iguais mas não tem um ponto de origem igual a zero



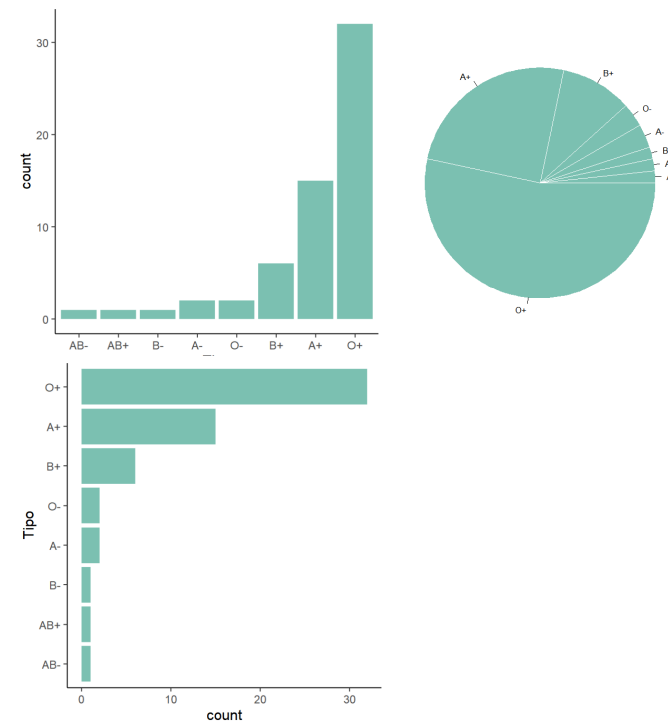
Razão

O valor zero representa a ausência de quantidade

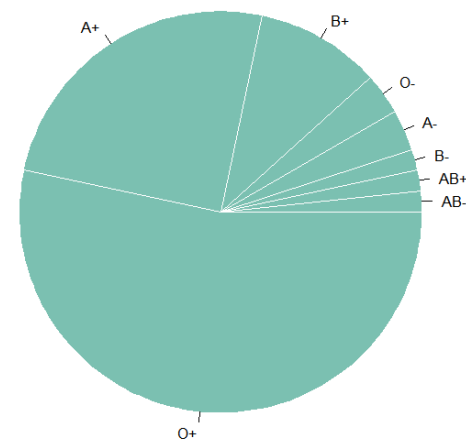
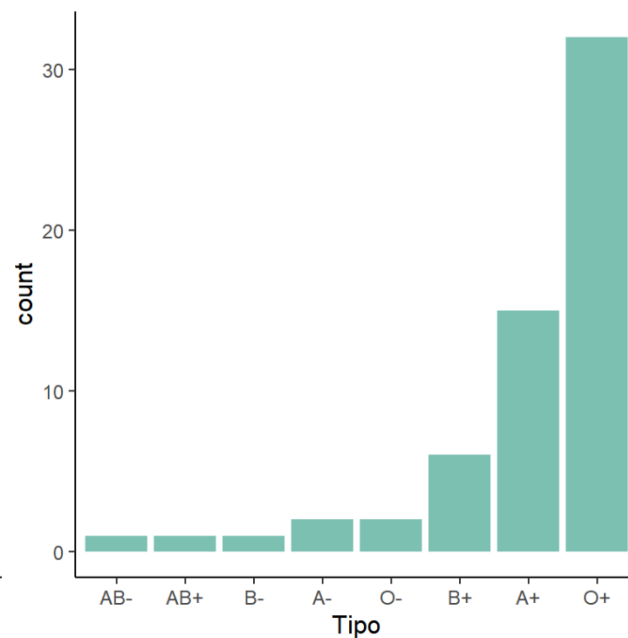
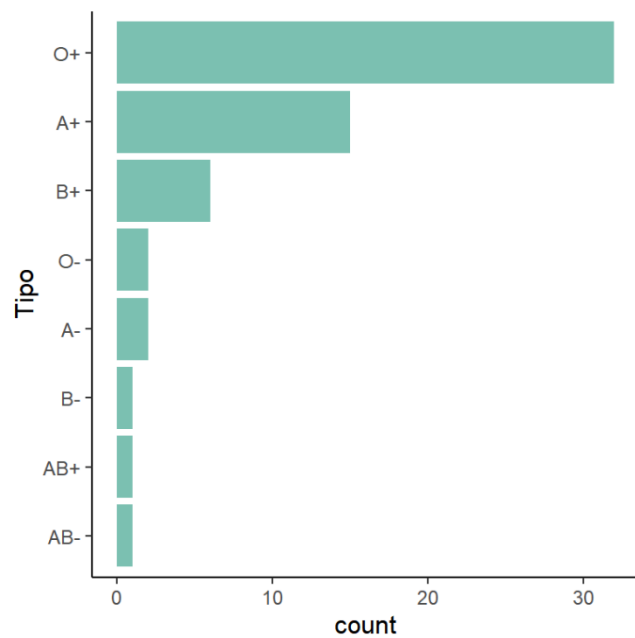


Variável Qualitativa

Tipo ABO	F_i	$F_{ri}(\%)$	F_{ac}	$Fr_{ac}(\%)$
A ⁺	15	25	15	25
A ⁻	2	3,33	17	28,33
B ⁺	6	10	23	38,33
B ⁻	1	1,67	24	40
AB ⁺	1	1,67	25	41,67
AB ⁻	1	1,67	26	43,33
O ⁺	32	53,33	58	96,67
O ⁻	2	3,33	60	100
Total	60	100		

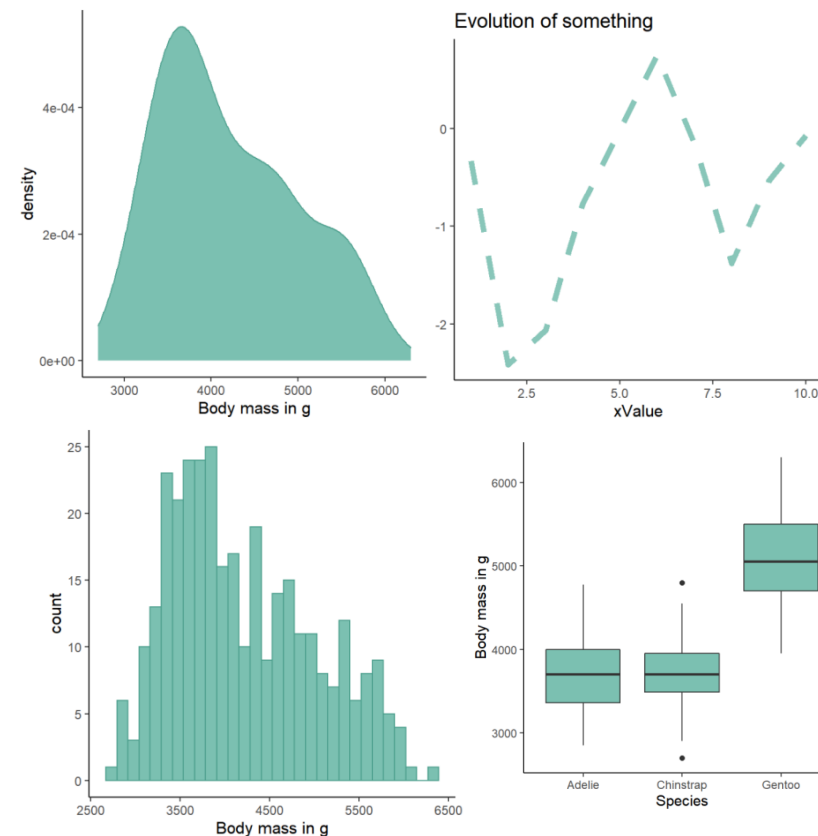


Variável Qualitativa

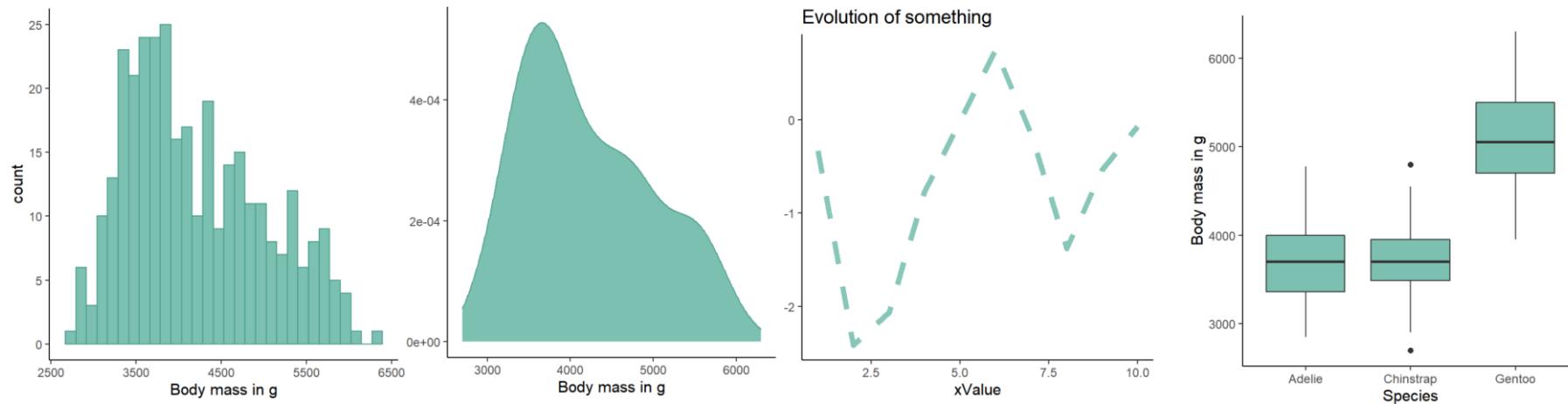


Variável Quantitativa

N° de filhos	F_i	$F_{ri}(\%)$	F_{ac}	$Fr_{ac}(\%)$
0	5	25	5	25
1-	7	35	12	60
2	6	30	18	90
3	2	10	20	100
Total	20	100		



Variável Quantitativa



Medidas resumo:

Tendência central



Média

Indicada quando não há valores extremos nos dados.



Mediana

Ótima quando há valores extremos nos dados.



Moda

Mais útil quando há dados categóricos.

SEPARATIZES

Quartis

Dados \div 4 = 25 %

Centis

Dados \div 10 = 10 %

Percentis

Dados \div 100 = 1 %

DISPERSÃO

Amplitude

Diferença entre o maior e o menor valor do conjunto de observações.

Variância

Avalia o quanto os dados estão dispersos em relação à média aritmética.

Desvio padrão

Avalia a homogeneidade, quanto menor o desvio-padrão maior a Homogeneidade.

Medidas resumo:



Tendência central



Média

Indicada quando não há valores extremos nos dados.



Mediana

Ótima quando há valores extremos nos dados.

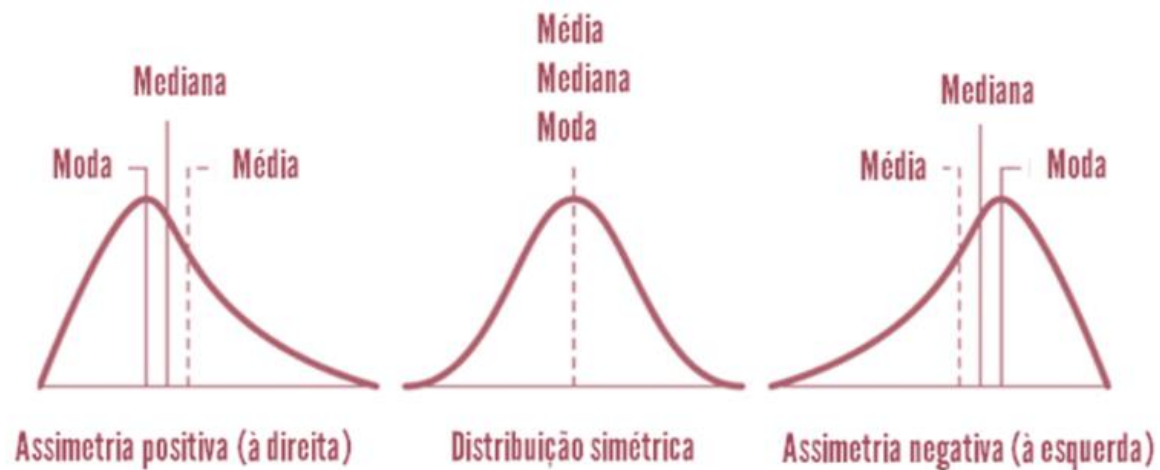


Moda

Mais útil quando há dados categóricos.

Medidas resumo:

Forma



Assimetria

Curtose

