

1

Tipos de variáveis

CONSIDERE A SEGUINTE SITUAÇÃO:
CARACTERÍSTICAS

Ind	EST CIVIL	ESCOLA	FILHOS	W	REGIÃO
1	S/C/DIV	anos	(0,)	$w < w < \bar{w}$	Qual
2	<u>Qualitativa</u>	<u>Quant</u>	<u>Quant</u>	<u>Quant</u>	<u>Qual</u>
3					
4					
5					
6					
7					

a construção

$w < w < \bar{w}$

Qualitativas:	<u>NOMINAL</u> Est. civil Região	<u>ORDINAL</u> grupos de a escola F < M < S
	<u>DISCRETA</u> Filhos / Anos de escola	<u>CONTÍNUA</u> W

glossário

dummy - (dicotômica/binária) $\begin{matrix} < \\ < \end{matrix} \begin{matrix} 1 \\ 0 \end{matrix}$
 $S = 1 (0)$

Proxy - variável de alta correlação

Intensidade - $X_1 \rightarrow$ tamanho $Y =$ Imp. pago

controle - Independent \leftrightarrow Dependente

aleatório - (N sem controle) $\Pi = f(e, \Sigma, \theta)$

Independente -

Distribuição - (μ, σ^2)

2 Distribuição de Frequência

Sempre estaremos interessados em
 → comportamentos da variável $\left\{ \begin{array}{l} \# \text{ Média} \checkmark \\ \# \text{ Variância} \checkmark \end{array} \right.$

Podemos organizar um conj. de dados
 p/ verificarmos sua distribuição.
 Tab. 1 - Amostra de um departamento

ENSINO	FREQ (n _i)	PROP (f _i)	% (100 · f _i)
FUNDAMENTAL	12	0,333	33,33%
MÉDIO	18	0,5	50%
SUPERIOR	6	0,1667	16,67%
TOTAL	36	1,0	100%

Como chegamos nisso?

$$f_{\text{freq}} = \sum_{i=1}^n n_i$$

↳ freq. absoluta

$$\frac{\sum_{i=1}^n n_i}{n} = f_i$$

Proporção (freq. relativa)

$$f_i = \frac{\sum_{i=1}^n n_i}{n}$$

n → n.º total de elementos

Tab. 2 - DADOS DA EMPRESA

ENSINO	FREQ	PROPOÇÃO	%
FUNDAMENTAL	650	0,325	32,5%
MÉDIO	1020	0,51	51,0%
SUPERIOR	330	0,165	16,5%
TOTAL	2000	1,00	100%

A amostra pode representar algo
 da população?
 ↳ funcionários da empresa //

3) Como tratar as variáveis contínuas?

↳ lembre-se do q. se trata //

⇒ Agrupar por faixas

↳ Preciso me preocupar em montar os grupos corretamente.

Exemplo:

classe salarial	Freq.	%
4,0 a 6,0	4	0,2
6,0 a 8,0	6	0,3
8,0 a 10,0	8	0,4
A cima de 10	2	0,1

$w < w < w$

20
 $6 \leq w_2 < 8,0$
 $8 \leq w_3 < 10$
 $w_4 \geq 10$

DADOS DA EMPRESA

Nº	Estado civil	Grau de instrução	Nº de filhos	Salário (x sal. mín.)	Idade		Região de procedência
					anos	meses	
1	solteiro	ensino fundamental	—	4,00	26	03	interior
2	casado	ensino fundamental	1	4,56	32	10	capital
3	casado	ensino fundamental	2	5,25	36	05	capital
4	solteiro	ensino médio	—	5,73	20	10	outra
5	solteiro	ensino fundamental	—	6,26	40	07	outra
6	casado	ensino fundamental	0	6,66	28	00	interior
7	solteiro	ensino fundamental	—	6,86	41	00	interior
8	solteiro	ensino fundamental	—	7,39	43	04	capital
9	casado	ensino médio	1	7,59	34	10	capital
10	solteiro	ensino médio	—	7,44	23	06	outra
11	casado	ensino médio	2	8,12	33	06	interior
12	solteiro	ensino fundamental	—	8,46	27	11	capital
13	solteiro	ensino médio	—	8,74	37	05	outra
14	casado	ensino fundamental	3	8,95	44	02	outra
15	casado	ensino médio	0	9,13	30	05	interior
16	solteiro	ensino médio	—	9,35	38	08	outra
17	casado	ensino médio	1	9,77	31	07	capital
18	casado	ensino fundamental	2	9,80	39	07	outra
19	solteiro	superior	—	10,53	25	08	interior
20	solteiro	ensino médio	—	10,76	37	04	interior

$\frac{4}{20} \quad \frac{6}{20} \quad \frac{8}{20} \quad \frac{2}{20}$

$0,2 \quad 0,3 \quad 0,4 \quad 0,1$
 $0,10 \quad 0,15 \quad 0,2 \quad 0,05$

4

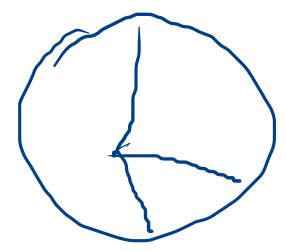
Tipos de Gráficos

CONSIDERAR SEMPRE O TIPO DE VARIÁVEL E A MENSAGEM

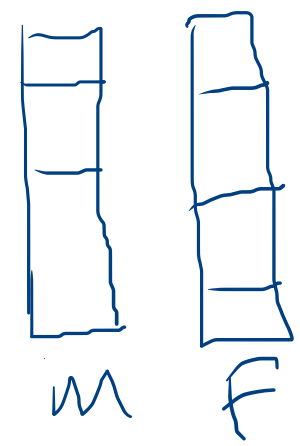
Gráficos p/ variáveis QUALITATIVAS:

A importância de considerar os eixos.

Pizza



BARRAS



Variáveis Quantitativas

→ discritas

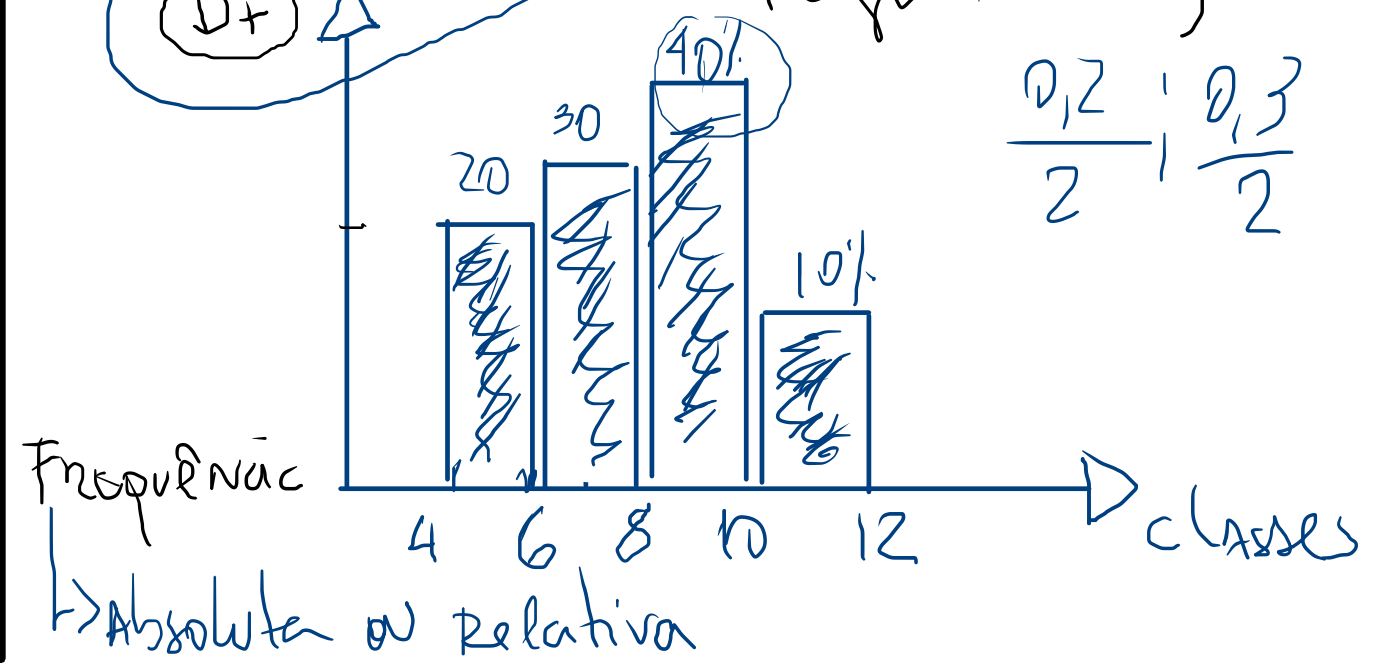
→ contínuas

$$DF = \frac{f_i}{\Delta \text{ amplitude}}$$

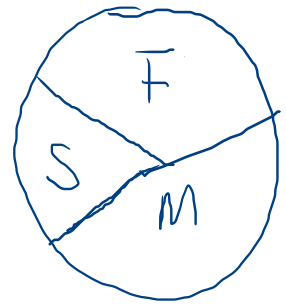
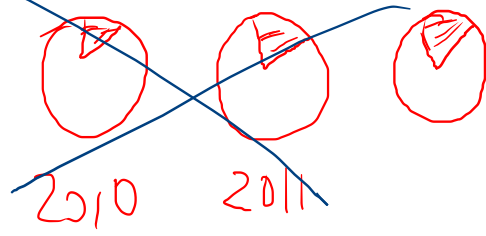
HISTOGRAMA:

Densidade de Frequência

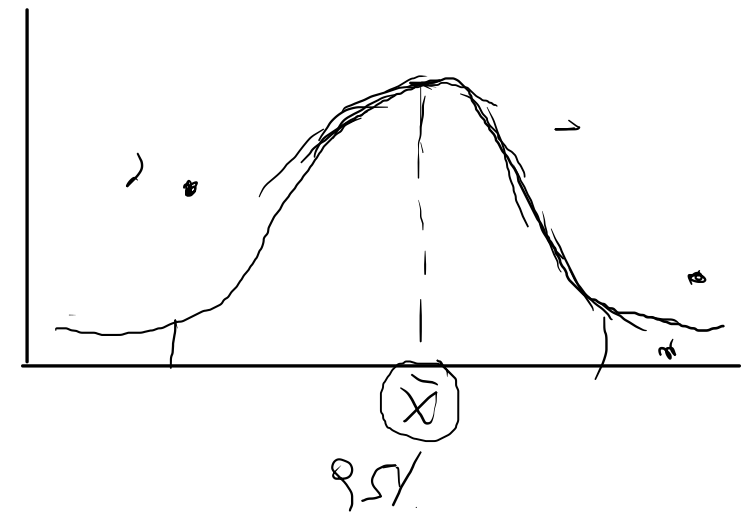
DF Δ



Rascunho



2020



Resíduo

Tributação

↳ clientes

↳ DADOS

Interesse (explicativa) Independente

Dependente (explicada)

*4 Ativ. Econômica (CNAE)

CNAE	Func	W
X4	X3	- X2

↳ variáveis

Tamanho

x_1

Tributo

y

$y = f(\dots)$

$y = \alpha + \beta_1 x_1 + \dots$

$\hat{\beta}_1$

ϵ

resíduo

1
2
3
4
...

Estado → e ou ϵ