

Universidade Federal de São Paulo

Tabulação de dados (ou organização dos dados), Medidas de frequência e Gráficos para os diferentes tipos de variáveis



Professor

Julio Cezar



AULA DE HOJE —

- Tabulação de dados ou organização dos dados;
- Medidas de Frequência;
- Gráficos para as diferentes tipos de variáveis;



DADOS

Madeira	Massa_inicial_gramas	Massa_final_gramas	Temperatura_celsius	Tipo_de_arvore	Repeticao	Perda_de_massa_grama:
1	9,389	2,945	300	Morototoni	1	6,444
2	8,22	2,443	300	Morototoni	2	5,777
3	7,472	2,261	300	Morototoni	3	5,211
4	7,869	2,245	300	Morototoni	4	5,624
5	7,717	2,312	300	Morototoni	5	5,405
6	7,818	2,485	300	Morototoni	6	5,333
7	7,97	2,536	300	Morototoni	7	5,434
8	15,45	5,444	300	Tauari	1	10,006
9	14,794	5,085	300	Tauari	2	9,709
10	16,315	5,544	300	Tauari	3	10,771
11	13,978	4,786	300	Tauari	4	9,192
12	14,401	5,409	300	Tauari	5	8,992
13	16,913	6,046	300	Tauari	6	10,867
14	13,798	4,704	300	Tauari	7	9,094
15	12,353	5,278	300	Cedrilho	1	7,075
16	11,735	4,367	300	Cedrilho	2	7,368
17	12,04	5,118	300	Cedrilho	3	6,922
18	13,315	6,062	300	Cedrilho	4	7,253
19	13,079	5,579	300	Cedrilho	5	7,5
20	13,542	6,369	300	Cedrilho	6	7,173
21	13,193	5,909	300	Cedrilho	7	7,284
22	11,451	4,345	300	Pinus	1	7,106
23	11,79	4,846	300	Pinus	2	6,944
24	9,688	3,923	300	Pinus	3	5,765
25	11,043	4,46	300	Pinus	4	6,583
26	9,619	3,752	300	Pinus	5	5,867
					-	

• Quais informações posso tirar dessas

Tabelas?

• De que forma apresentar as conclusões?



No	ANO	DIA	MES	R.GLOBAL	INSO-	PRECIPI-	UMIDADE	VENTO	VENTO	TEMPER.	TEMPER.	TEMPER.	EVAPO-
				2	LACAO	TACAO	RELATIVA	MAXIMO	MEDIO	MAXIMA	MINIMA	MEDIA	RACAO
				cal/cm.d	h/d	mm	%	m/s	km/h	grau C	grau C	grau C	mm
1	2018	1	JAN	392	3,5	2,8	86	9,3	5,4	29,8	19,3		
2	2018	2	JAN	494	6,2	0,3	84	13,2	8,8	29,4	21,0		
3	2018	3	JAN	466	5,4		83	10,0	8,1	29,2	20,9		
4	2018	4	JAN	745	12,0			9,6			19,3		
5	2018	5	JAN	691	11,4			9,3	6,7	31,8	17,8		
6	2018	6	JAN	379	3,2			8,9	5,2	30,0	18,7		
7	2018	7	JAN	293	0,9			12,4	11,1	23,8	18,0		5,05
8	2018	8	JAN	253	0,0			4,9	6,4	24,4	18,5		
9	2018	9	JAN	321	1,6	0,0	85	9,5	6,9	26,7	18,9		5,50
10	2018	10	JAN	556	7,8	0,0	75	9,6	5,9	31,5	18,4		6,02
11	2018	11	JAN	445	4,9	28,4	86	14,4	5,3	31,2	18,8		6,04
12	2018	12	JAN	316	1,5	16,5	88	12,5	5,0	26,9	18,6		5,49
13	2018	13	JAN	416	4,1	7,4	84	11,5	5,5	29,6	18,3		5,78
14	2018	14	JAN	416	4,1	24,4	84	11,2	5,6	29,1	18,7		5,77
15	2018	15	JAN	524	7,0	34,3	82	8,9	6,6	31,3	19,7		6,16
16	2018	16	JAN	481	5,8	1,0	87	8,8	5,4	29,7	19,8		5,98
17	2018	17	JAN	694	11,5	13,0	77	12,0	4,9	32,6	20,4	26,5	6,40
18	2018	18	JAN	482	5,9	3,0	75	7,6	7,0	33,1	20,2	26,7	6,43
19	2018	19	JAN	401	3,7	2,3	80	13,7	5,6	31,4	19,8	25,6	6,18
20	2018	20	JAN	554	7,8	1,8	80	11,2	7,3	32,1	20,1	26,1	6,30
21	2018	21	JAN	646	10,2	0,0	77	15,6	6,2	33,9	20,1	27,0	6,52
22	2018	22	JAN	496	5,7	0,2	77	12,8	5,2	33,4	21,0	27,2	8,00
23	2018	23	JAN	492	9,4	39,4	78	10,1	7,4	33,8	20,0	26,9	6,49

IMPORTÂNCIA DA ESTATÍSTICA NO PROJETO DE PESQUISA :

1) Planejando a coleta de dados

2) Organizando os dados

3) Análises e Interpretação durante

4) Comparações, discussões e conclusões



População:

conjunto de todos os

elementos ou

resultados sob

investigação.

Amostra:

qualquer subconjunto

da população.

TIPOS DE VARIÁVEIS

O tipo define se a variável representa quantidades, sendo então denominado quantitativa ou numérica, ou qualidades, quando é chamado de qualitativa, simbólica ou categórica, pois os valores podem ser associados a categorias.

Exemplos de conjuntos de valores qualitativos são (pequeno, médio, grande) e

(matemática, física, química). Apesar de alguns atributos qualitativos poderem ter seus valores ordenados, operações aritméticas não podem ser aplicadas aos seus valores. As variáveis são numéricas, como no conjunto de valores (23, 45, 12). Os valores de uma variável quantitativa podem tanto ser ordenados quanto utilizados em operações aritméticas. Variáveis quantitativas podem ser ainda contínuas ou discretas.

TIPOS DE VARIÁVEIS -

Variáveis qualitativas

São definidas por várias categorias, ou seja, representam uma classificação dos indivíduos.

- não são numéricas;
- são classificadas como nominais e ordinais.



TIPOS DE VARIÁVEIS

Variáveis qualitativas

Variáveis qualitativas nominais

Não existe ordenação dentre as categorias.

Exemplos: Gênero (M/F), cor dos olhos(castanho, azul, verde,...), fumante/não fumante, doente/sadio.

Variáveis qualitativas ordinais

Existe uma ordenação entre as categorias

Exemplos estágio da doença (inicial, intermediário, terminal), mês de observação (janeiro, fevereiro,..., dezembro).



TIPOS DE VARIÁVEIS -

Variáveis quantitativas

São as características que podem ser medidas em uma escala quantitativa.

- são numéricas;
- são classificadas em discreta e contínua.



TIPOS DE VARIÁVEIS

Variáveis quantitativas

Variáveis quantitativas discretas

São aquelas variáveis que podem assumir somente valores inteiros num conjunto de valores.

Exemplos: Número de alunos, número de funcionários, número de processos.

Variáveis quantitativas contínuas

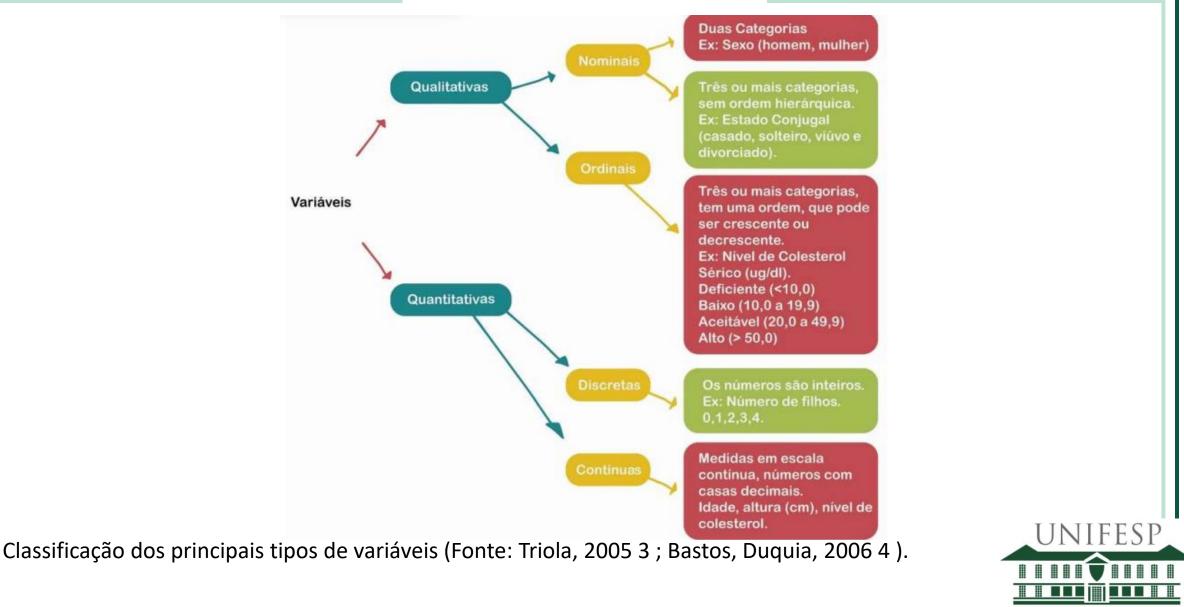
São aquelas variáveis que podem assumir um valor dentro de um intervalo de valores.

Exemplos: Pressão arterial, idade, salário, atraso de transmissão de bytes por uma rede de internet.

TIPOS DE VARIÁVEIS -



TIPOS DE VARIÁVEIS



TABULAÇÃO DE DADOS -

Coletar dados podem envolver diversas atividades tais como experimentos em laboratório, observações de campo, pesquisa de opinião, exame de registros históricos,...

A quantidade de dados em geral é grande e torna-se difícil, por simples inspeção obter informações sobre aspectos importantes dos dados.

É importante obter uma Estatística Descritiva dos dados: resumos em forma de tabelas, gráficos e indicadores numéricos de centro e variabilidade



TABULAÇÃO DE DADOS

Métodos de Estatística descritiva são aplicáveis a situações em que temos amostras ou um censo da população.

Para um censo (toda a população), a estatística descritiva destes dados podem ser apresentados em relatórios que podem ser apresentados ao público, podem ser partes de dados para ações governamentais, etc...

Para dados amostrais, a estatística descritiva serve como ponto de partida para estudos posteriores, tal que inferências sobre a população possam ser feitas.



TABULAÇÃO DE DADOS -

Principais aspectos ao descrever um conjunto de dados

Resumo e descrição do padrão geral dos dados por:

- Apresentação de tabelas e gráficos;
- Observação das características importantes nas formas dos gráficos, tais como simetria ou assimetria;
- Explorar nos gráficos a presença de observações não usuais, ou seja, observações que parecem estar distantes dos valores onde ocorrem a maior parte dos dados

Cálculo de medidas numéricas:

- Um valor indicando o centro dos dados (média, mediana, moda,...);
- Uma medida da variação nos dados.



TABELAS E GRÁFICOS

Alguns tipos de tabelas e gráficos

- Tabela de frequência;
- Gráfico de barras e gráfico de pizza;
- Diagrama de pareto;
- Diagrama de pontos;
- Diagrama de Ramos-e-Folhas;
- Distribuições de Frequência, Histogramas e Polígonos;
- Distribuições Acumuladas;
- Tabelas de Contingência;
- Gráfico de Dispersão e de Séries Temporais.

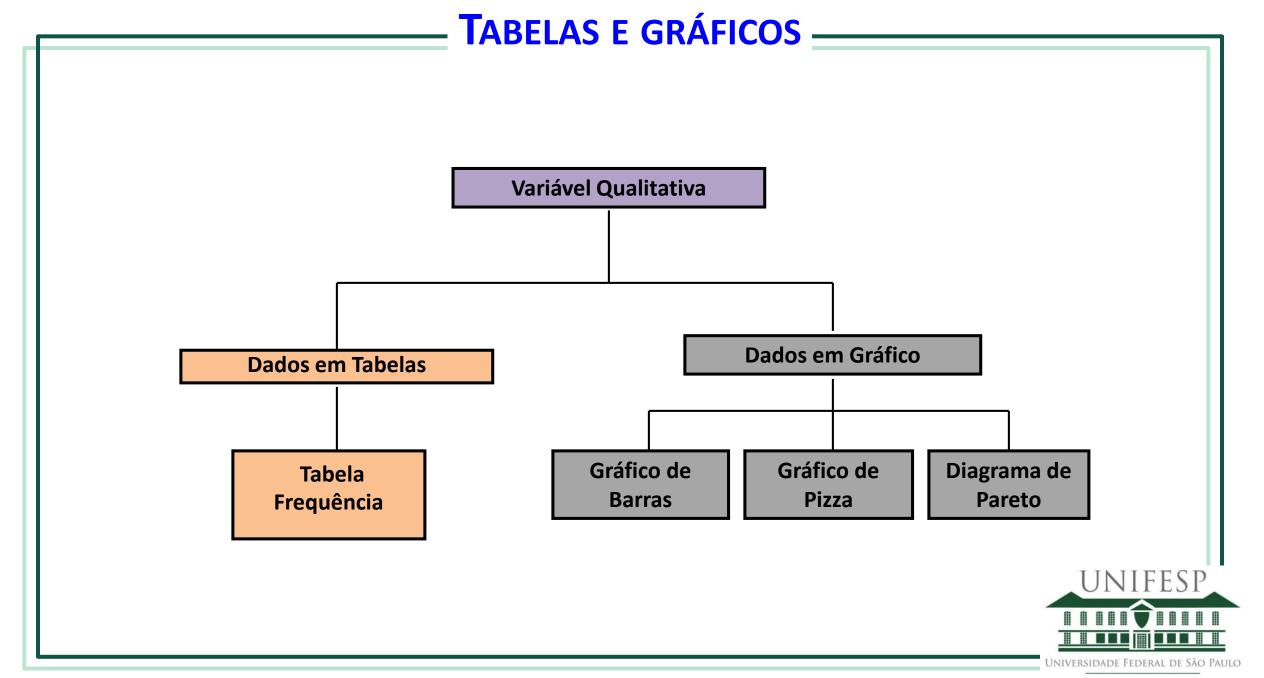


TABELAS E GRÁFICOS -

Uma tabela deve conter:

- Título
- Corpo
- Cabeçalho
- Coluna indicadora
- O **título** explica o conteúdo da tabela;
- O corpo da tabela e composto pelos dados, organizados em linhas e colunas que se cruzam;
- **Célula** e a interseção de uma linha com uma coluna;
- Linha e uma série de células organizadas horizontalmente;
- **Coluna** e uma serie de células organizadas verticalmente;
- O cabeçalho especifica o conteúdo das colunas;
- A coluna indicadora especifica o conteúdo das linhas.





DISTRIBUIÇÃO DE FREQUÊNCIA

> Resumo e visualização de um conjunto de dados

→ <u>comportamento da variável X</u>

- Uma distribuição de frequência pode ser apresentada em tabelas ou gráficos!
- Uma tabela de frequências possui informações sobre o número de pesquisados, a porcentagem, a proporção e a proporção acumulada de cada classe da variável analisada.
- Os métodos gráficos são utilizados para buscar padrões e relações; confirmar (ou não) certas expectativas que se tinha sobre os dados; descobrir novos fenômenos; confirmar (ou não) suposições feitas sobre os procedimentos estatísticos usados e un ligidades de modo mais rápido e fácil.

Distribuições de Frequência: Uma distribuição de frequência é uma lista ou uma tabela contendo a frequência com que cada dado está dentro de uma determinada categoria, ou de um determinado valor, ou de um intervalo de valores.

n: número total de observações

Frequência absoluta (f): é o número de vezes que um dado valor foi observado.

Frequência relativa (fr): é o quociente entre a frequência absoluta e o número

total de observações. Ou proporção.

$$\mathbf{fr} = \frac{\mathbf{f}}{\mathbf{n}}$$

Porcentagem (%) = fr x 100



Frequência acumulada (Fa): é a soma de todas as frequências absolutas até o valor analisado.

Frequência acumulada relativa (Far): é a soma de todas as frequências relativas até o valor analisado.

Far %: frequência acumulada relativa em porcentagem (Far x 100).





Universidade Federal de São Paulo

Dados não agrupados

Tabela 2.1 Informações sobre estado civil, grau de instrução, número de filhos, salário (expresso como fração do salário mínimo), idade (medida em anos e meses) e procedência de 36 empregados da seção de orçamentos da Companhia MB.

3.10	Parada aladi	Com de la servición	NIO 4- 611	C-16-1- (v16-)	Id	ade	Danilla da assaultanta
IN-	Estado civil	Grau de instrução	Nº de filhos	Salário (× sal. mín.)	anos	meses	Região de procedência
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
21	casado	ensino médio	1	11,06	30	09	outra
22	solteiro	ensino médio	_	11,59	34	02	capital
23	solteiro	ensino fundamental	_	12,00	41	00	outra
24	casado	superior	0	12,79	26	01	outra
25	casado	ensino médio	2	13,23	32	05	interior
26	casado	ensino médio	2	13,60	35	00	outra
27	solteiro	ensino fundamental	_	13,85	46	07	outra
28	casado	ensino médio	0	14,69	29	08	interior
29	casado	ensino médio	5	14,71	40	06	interior
30	casado	ensino médio	2	15,99	35	10	capital
31	solteiro	superior	_	16,22	31	05	outra
32	casado	ensino médio	1	16,61	36	04	interior
33	casado	superior	3	17,26	43	07	capital
34	solteiro	superior	_	18,75	33	07	capital
35	casado	ensino médio	2	19,40	48	11	capital
36	casado	superior	3	23,30	42	02	interior



Tabela de frequência contendo frequências acumuladas

Exemplo:

Tabela 1: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i	Frequência acumulada Fa _i	Frequência acumulada relativa Far _i	Porcentagem 100 x Far _i
Fundamental						
Médio						
Superior						
Total						

Tabela de frequência contendo frequências acumuladas

Exemplo:

Tabela 1: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

(proporção) 100 x fr _i Fa _i relativa Far _i	Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i	Frequência acumulada Fa _i		Porcentagem 100 x Far _i
---	-------------------	---------------------------------------	--	--------------------------------------	--	--	---------------------------------------

Note que, as frequências são associadas a i-ésima classe (classe i = 1, ..., K).

Tabela de frequência contendo frequências acumuladas

Exemplo:

Tabela 1: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

	Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i	Frequência acumulada Fa _i	Frequência acumulada relativa Far _i	Porcentagem 100 x Far _i
	Fundamental						
\dashv	Médio						
	Superior						
	Total						

Dados não agrupados

UNIFESP

Tabela de frequência contendo frequências acumuladas

Exemplo:

Tabela 1: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

	Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i	Frequência acumulada Fa _i	Frequência acumulada relativa Far _i	Porcentagem 100 x Far _i
	Fundamental	12					
4	Médio	18					
	Superior	6					
	Total	36					

Dados não agrupados

Tabela de frequência contendo frequências acumuladas

Exemplo:

Tabela 1: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

	Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i	Frequência acumulada Fa _i	Frequência acumulada relativa Far _i	Porcentagem 100 x Far _i
	Fundamental	12	0,3333				
4	Médio	18	0,5000				
	Superior	6	0,1667				
	Total	36	1,000				

Dados não agrupados

Tabela de frequência contendo frequências acumuladas

Exemplo:

Tabela 1: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

	Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i	Frequência acumulada Fa _i	Frequência acumulada relativa Far _i	Porcentagem 100 x Far _i
	Fundamental	12	0,3333	33,33			
1	Médio	18	0,5000	50,00			
	Superior	6	0,1667	16,67			
	Total	36	1,000	100,00			

Dados não agrupados

Fonte: BUSSAB e MORETTIN, 2017
UNIFESP

Tabela de frequência contendo frequências acumuladas

Exemplo:

Tabela 1: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

	Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i	Frequência acumulada Fa _i	Frequência acumulada relativa Far _i	Porcentagem 100 x Far _i
	Fundamental	12	0,3333	33,33	12		
4	Médio	18	0,5000	50,00	12+18 = 30		
	Superior	6	0,1667	16,67	30+6 = 36		
L	Total	36	1,000	100,00			

Dados não agrupados

Tabela de frequência contendo frequências acumuladas

Exemplo:

Tabela 1: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i	Frequência acumulada Fa _i	Frequência acumulada relativa Far _i	Porcentagem 100 x Far _i
Fundamental	12	0,3333	33,33	12	0,3333	
Médio	18	0,5000	50,00	12+18 = 30	0,3333+0,500 = 0,8333	
Superior	6	0,1667	16,67	30+6 = 36	0,8333+0,1667 = 1,000	
Total	36	1,000	100,00			

Dados não agrupados

Fonte: BUSSAB e MORETTIN, 2017,

Tabela de frequência contendo frequências acumuladas

Exemplo:

Tabela 1: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

	Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i	Frequência acumulada Fa _i	Frequência acumulada relativa Far _i	Porcentagem 100 x Far _i
	Fundamental	12	0,3333	33,33	12	0,3333	33,33
	Médio	18	0,5000	50,00	12+18 = 30	0,3333+0,500 = 0,8333	33,33+50,00 = 83,33
L	Superior	6	0,1667	16,67	30+6 = 36	0,8333+0,1667 = 1,000	83,33+ 16,67 = 100
	Total Dados nã	o agrupados	1,000	100,00			

Fonte: BUSSAB e MORETTIN, 2017 UNIFEST

Tabela de frequência contendo frequências acumuladas

Exemplo:

Tabela 1: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

	Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i	Frequência acumulada Fa _i	Frequência acumulada relativa Far _i	Porcentagem 100 x Far _i
	Fundamental	12	0,3333	33,33	12	0,3333	33,33
4	Médio	18	0,5000	50,00	30	0,8333	83,33
	Superior	6	0,1667	16,67	36	1,000	100
	Total	36	1,000	100,00			

Dados não agrupados

Quais informações podemos tirar destas novas medidas ? ?

Tabela de frequência contendo frequências acumuladas

Exemplo:

Tabela 1: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i	Frequência acumulada Fa _i	Frequência acumulada relativa Far _i	Porcentagem 100 x Far _i
Fundamental	12	0,3333	33,33	12	0,3333	33,33
Médio	18	0,5000	50,00	30	0,8333	83,33
Superior	6	0,1667	16,67	36	1,000	100
Total	36	1,000	100,00			

Dados não agrupados

A frequência acumulada 12+18 = 30 (ou 83,33%) representa o número (ou a %) de pessoas com escolaridade até o ensino médio.



Note que:

A soma de todas as frequências absolutas = n

A soma de todas as frequências relativas = 1

A soma de todas as frequências relativas em % = 100%



- Observando os resultados da segunda coluna, vê-se que dos 36 empregados da companhia, 12 têm ensino fundamental, 18 ensino médio e 6 possuem curso superior.
- A frequência relativa e a porcentagem são bastante úteis quando decide-se comparar o resultado de pesquisas distintas, por exemplo, quando deseja-se pesquisar o total de empregados da companhia dentro de cada categoria (ou classe).
- A **frequência acumulada** é utilizada para verificar onde encontra-se a maior parte da população pesquisada.





Universidade Federal de São Paulo

Dados agrupados

Organização dos dados em intervalos de classes

Muitas vezes, devido a grande dispersão dos dados, é interessante agrupar os dados em classes.

classe: intervalo no qual é agrupado um conjunto de valores dos dados.

Número de valores diferentes muito grande



Agrupamento dos dados em classes



Características

- as classes devem abranger todas as observações.
- o extremo superior de uma classe é o extremo inferior da classe subsequente.
- cada valor observado deve estar presente em apenas uma classe.



É possível escolher os limites de uma classe de várias maneiras.

Vamos escolher as classes de frequências com intervalos do tipo fechado a esquerda e aberto a direita:

 $[a,b) = \{ a \le x < b \}$, representado também como: $a \vdash b$

a: limite inferior da classe

b: limite superior da classe

intervalo de classe = b - a

Obs.: Também é possível obter o **ponto médio da classe (pmc):** é definido como o valor médio entre os limites superiores e inferiores de uma determinada classe.

UNIFES

$$pmc = (a + b)/2$$

Construindo uma distribuição de frequências

- (a) Obter o Rol (é um arranjo dos dados brutos em ordem crescente, ou seja, é a tabela obtida após a ordenação dos dados).
- (b) Determinar a quantidade de classes (K): a escolha do número de classes é arbitrária, no entanto, existem na literatura algumas alternativas como as duas dadas a seguir:
- $K = \sqrt{n}$ (regra da raíz quadrada).
- Fórmula de Sturges: $K = 1 + 3,32x\log_{10}(n) = 1 + \log_2(n)$ (**obs.:** deve-se aproximar para o interior mais próximo).

Construindo uma distribuição de frequências

(c) Calcular a amplitude da classe (amplitude do intervalo da classe), h:

- Limite Inferior: \mathbf{x}_{min}
- Limite Superior: x_{máx}
- Amplitude total: $\Delta = \mathbf{x}_{máx} \mathbf{x}_{mín}$

$$\mathbf{h} = \frac{\Delta}{K}$$

Assim, têm-se:

- 1^a classe: x_{mín} até x_{mín}+h
- $-2^{\underline{a}}$ classe: \mathbf{x}_{min} + \mathbf{h} até \mathbf{x}_{min} + $\mathbf{2.h}$
- k^a classe: $x_{min}+(k-1).h$ até $x_{min}+K.h$



Construindo uma distribuição de frequências:

(d) Determine a frequência de classe (número de observações em cada classe).

(e) Determine a frequência relativa de cada classe, que pode ser obtida por:

Frequência relativa de classe
$$=\frac{\text{Frequência de classe}}{n}$$



TABELAS DE FREQUÊNCIAS ____ Salário (× sal. mín.) 4,00 4,56 5,25 5,73 6,26 6,66 6,86 7,39 7,59 7,44 8,12 8,46 8,74 8,95 9,13 9,35 9,77 9,80 10,53 10,76 11,06 11,59 12,00 12,79 13,23 13,60 13,85 14,69 14,71 15,99 16,22 16,61 17,26 18,75

19,40

Universidade Federal de São Paulo

Salário (× sal. mín.) 4,56 5,25 5,73 6,26 6,66 6,86 7,39 7,59 7,44 8,12 8,46 8,74 8,95 9,13 9,35 9,77 9,80 10,53 10,76 11,06 11,59 12,00 12,79 13,23 13,60 13,85 14,69 14,71 15,99 16,22 16,61 17,26

18,75 19,40 Tabela 2: Frequências dos salários dos empregados de uma empresa.

Classe de salários	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Frequência acumulada relativa Far _i	Porcentagem 100 x fr _i



Salário (× sal. mín.) 4,56 5,25 5,73 6,26 6,66 6,86 7,39 7,59 7,44 8,12 8,46 8,74 8,95 9,13 9,35 9,77 9,80 10,53 10,76 11,06 11,59 12,00 12,79 13,23 13,60 13,85 14,69 14,71 15,99 16,22 16,61 17,26 18,75 19,40

Tabela 2: Frequências dos salários dos empregados de uma empresa.

Classe de salários	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Frequência acumulada relativa Far _i	Porcentagem 100 x fr _i
[4,00; 8,00)				
[8,00; 12,00)				
[12,00; 16,00)				
[16,00; 20,00)				
[20,00; 24,00)				
Total				





Tabela 2: Frequências dos salários dos empregados de uma empresa.

	Classe de salários	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Frequência acumulada relativa Far _i	Porcentagem 100 x fr _i
	[4,00; 8,00)				
	[8,00; 12,00)				
4	[12,00; 16,00)				
	[16,00; 20,00)				
	[20,00; 24,00)				
\	Total Dados agr				
	Jo agi	upados	E. A. DU	OOAD MODET	TIN 0040 III

Salário (× sal. mín.) 4,56 5,25 5,73 6,26 6,66 6,86 7,39 7,59 7,44 8,12 8,46 8,74 8,95 9,13 9,35 9,77 9,80 10,53 10,76 11,06 11,59 12,00 12,79 13,23 13,60 13,85 14,69 14,71 15,99 16,22 16,61 17,26 18,75 19,40

Tabela 2: Frequências dos salários dos empregados de uma empresa.

Classe de salários	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Frequência acumulada relativa Far _i	Porcentagem 100 x fr _i
[4,00; 8,00)	10			
[8,00; 12,00)	12			
[12,00; 16,00)	8			
[16,00; 20,00)	5			
[20,00; 24,00)	1			
Total	36			



Salário (× sal. mín.) 4,56 5,25 5,73 6,26 6,86 7,39 7,59 7,44 8,12 8,46 8,74 8,95 9,13 9,35 9,77 9,80 10,53 10,76 11,06 11,59 12,00 12,79 13,23 13,60 13,85 14,69 14,71 15,99 16,22 16,61 17,26 18,75 19,40

Tabela 2: Frequências dos salários dos empregados de uma empresa.

Classe de salários	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Frequência acumulada relativa Far _i	Porcentagem 100 x fr _i
[4,00; 8,00)	10	10/36 =0,278		
[8,00; 12,00)	12	12/36 =0,333		
[12,00; 16,00)	8	8/36 =0,222		
[16,00; 20,00)	5	5/36 =0,139		
[20,00; 24,00)	1	1/36 =0,029		
Total	36	1		



Salário (× sal. mín.) 4,56 5,25 5,73 6,26 6,66 6,86 7,39 7,59 7,44 8,12 8,46 8,74 8,95 9,13 9,35 9,77 9,80 10,53 10,76 11,06 11,59 12,00 12,79 13,23 13,60 13,85 14,69 14,71 15,99 16,22 16,61 17,26 18,75 19,40

Tabela 2: Frequências dos salários dos empregados de uma empresa.

Classe de salários	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Frequência acumulada relativa Far _i	Porcentagem 100 x fr _i
[4,00; 8,00)	10	10/36 =0,278	0,278	
[8,00; 12,00)	12	12/36 =0,333	0,611	
[12,00; 16,00)	8	8/36 =0,222	0,833	
[16,00; 20,00)	5	5/36 =0,139	0,972	
[20,00; 24,00)	1	1/36 =0,029	1,000	
Total	36	1		



Salário (× sal. mín.) 4,56 5,25 5,73 6,26 6,66 6,86 7,39 7,59 7,44 8,12 8,46 8,74 8,95 9,13 9,35 9,77 9,80 10,53 10,76 11,06 11,59 12,00 12,79 13,23 13,60 13,85 14,69 14,71 15,99 16,22 16,61 17,26 18,75 19,40

Tabela 2: Frequências dos salários dos empregados de uma empresa.

Classe de salários	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Frequência acumulada relativa	Porcentagem 100 x fr _i
[4,00; 8,00)	10	10/36 =0,278	Far _i 0,278	27,78%
[8,00; 12,00)	12	12/36 =0,333	0,611	33,33%
[12,00; 16,00)	8	8/36 =0,222	0,833	22,22%
[16,00; 20,00)	5	5/36 =0,139	0,972	13,89%
[20,00; 24,00)	1	1/36 =0,028	1,000	2,78%
Total	36	1		100%



Interpretando tabela de frequência dos salários dos empregados

Agrupando as classes por faixas, perde-se alguma informação. **Por exemplo**, não sabemos quais são os oito salários da classe de 12 a 16, a não ser que investigamos a tabela original.



Uma forma de interpretação é dizer que todos os oito salários são iguais ao ponto médio, 14.

Observação: Ao resumir os dados referentes a uma variável contínua, perde-se alguma informação.



Tabela de distribuição de frequências

Tabela 3: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i	
Fundamental	12	0,3333	33,33	1
Médio	18	0,5000	50,00	
Superior	6	0,1667	16,67	
Total	36	1,000	100,00	

Qual o tipo de gráfico podemos construir com base na Tabela 🕴 🕻

Tabela 3: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i
Fundamental	12	0,3333	33,33
Médio	18	0,5000	50,00
Superior	6	0,1667	16,67
Total	36	1,000	100,00

A altura da barra representa a frequência ou a porcentagem de cada categoria.

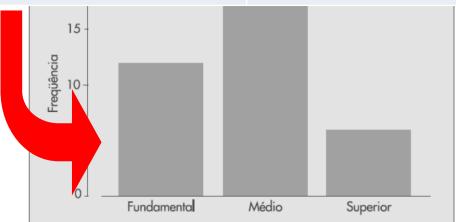


Tabela 3: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i
Fundamental	12	0,3333	33,33
Médio	18	0,5000	50,00
Superior	6	0,1667	16,67
Total	36	1,000	100,00

A "fatia" da pizza representa a frequência ou a porcentagem de cada categoria.

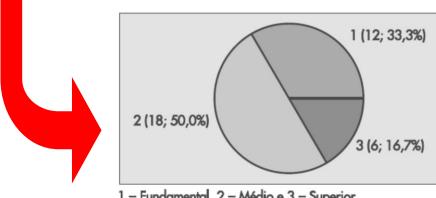




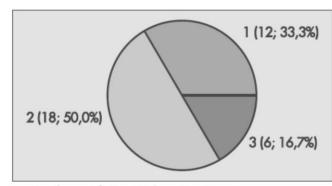
Tabela 3: Distribuição de frequências e porcentagens dos 36 empregados da companhia MB segundo o grau de instrução.

Grau de instrução	Frequência absoluta f _i	Frequência relativa (proporção) fr _i	Porcentagem 100 x fr _i
Fundamental	12	0,3333	33,33
Médio	18	0,5000	50,00
Superior	6	0,1667	16,67
Total	36	1,000	100,00

O ângulo de cada "fatia" é obtido, utilizando a regra de três

> 360 ----- n (ou 100%) x ----- f (ou fr%)





1 = Fundamental, 2 = Médio e 3 = Superior



Na construção de um gráfico circular deve-se ter em conta que:

- a amplitude de cada setor é proporcional à frequência que representa;
- a legenda pode ser dispensada, inscrevendo-se os valores da variável e as suas frequências nos respectivos setores circulares;
- podem-se usar cores diferentes para os diferentes setores;
- o gráfico deve ter um título adequado.

Não é aconselhável construir um gráfico circular:

- para variáveis que tenham mais de cinco ou seis modalidades;
- para situações em que os sectores resultam aproximadamente com a mesma amplitude;
- para setores com amplitudes muito pequenas.



Diagrama de Pareto

- Gráfico de Barras ordenadas (da mais alta para a mais baixa);
- Muito utilizado em controle de qualidade (busca por erros mais comuns);
 - Barras mais baixas erros de menor ocorrência;
- Ordem em que os erros devem ser sanados, reduzidas as perdas ou solucionadas as reclamações;
- Identificar causas de sucesso.



Diagrama de Pareto: Exemplo

As reclamações no PROCON, em determinado ano, foram enquadradas em algumas áreas.

Tabela 4: Reclamações no PROCON por área em um determinado ano.

Área	Frequência absoluta f _i	Frequência relativa (proporção) fr _i
Alimentos	15	0,01
Assuntos Financeiros	345	0,23
Habitação	120	0,08
Produtos	315	0,21
Saúde	105	0,07
Serviços	600	0,40
Total	1500	1,00

Diagrama de Pareto: Exemplo

As reclamações no PROCON, em determinado ano, foram enquadradas em algumas áreas. (**Obs.:** Ordenando a frequência absoluta).

Tabela 4: Reclamações no PROCON por área em um determinado ano.

Área	Frequência absoluta f _i	Frequência relativa (proporção) fr _i
Serviços	600	0,40
Assuntos Financeiros	345	0,23
Produtos	315	0,21
Habitação	120	0,08
Saúde	105	0,07
Alimentos	15	0.01
Total	1500	1,00

Diagrama de Pareto: Exemplo

As reclamações no PROCON, em determinado ano, foram enquadradas em algumas áreas.

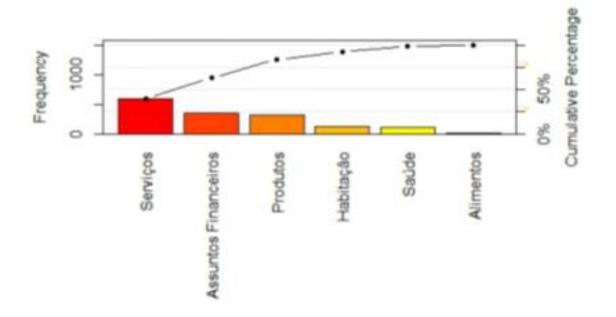


Figura: Reclamações no PROCON por área em determinado ano.



A Disposição Ordenada

Sequência de dados em uma ordem de classificação:

- Em ordem crescente, do menor valor para o maior valor;
- Fornece alguns sinais de variação dentro do intervalo;
- Pode ajudar a identificar pontos extremos (distantes da maioria);
- Se o conjunto de dados é muito grande, a disposição ordenada é pouco utilizada.



Exemplo de disposição ordenada

Dados em sua forma bruta (como coletados):

24, 26, 24, 21, 27, 27, 30, 24, 27, 32, 38, 27

Dados em disposição ordenada, do menor para o maior:

21, 24, 24, 24, 26, 27, 27, 27, 27, 30, 32, 38



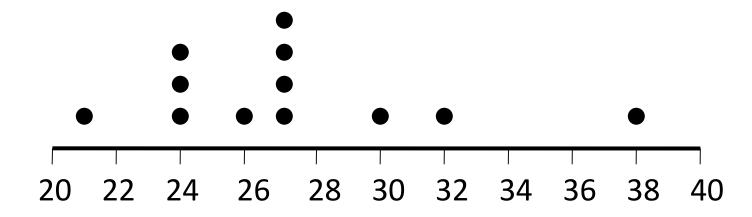
Exemplo de disposição ordenada

Dados em sua **forma bruta** (como coletados):

24, 26, 24, 21, 27, 27, 30, 24, 27, 32, 38, 27

Dados em disposição ordenada, do menor para o maior:

21, 24, 24, 24, 26, 27, 27, 27, 27, 30, 32, 38





Exemplo de disposição ordenada

Dados em sua **forma bruta** (como coletados):

24, 26, 24, 21, 27, 27, 30, 24, 27, 32, 38, 27

Dados em disposição ordenada, do menor para o maior:

21, 24, 24, 24, 26, 27, 27, 27, 27, 30, 32, 38

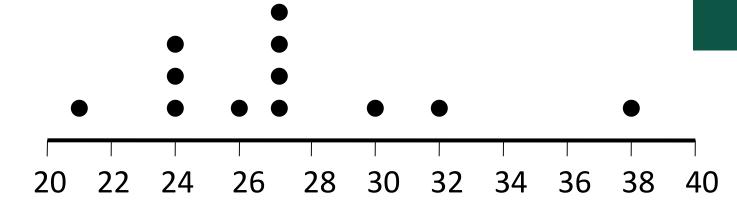


Diagrama de pontos: uma reta contendo a escala de valores dos dados com pontos representando os valores obtidos. O número de pontos sobre um dado valor representa a frequência deste.



Exemplo de disposição ordenada

Dados em sua **forma bruta** (como coletados):

24, 26, 24, 21, 27, 27, 30, 24, 27, 32, 38, 27

Dados em disposição ordenada, do menor para o maior:

21, 24, 24, 24, 26, 27, 27, 27, 27, 30, 32, 38

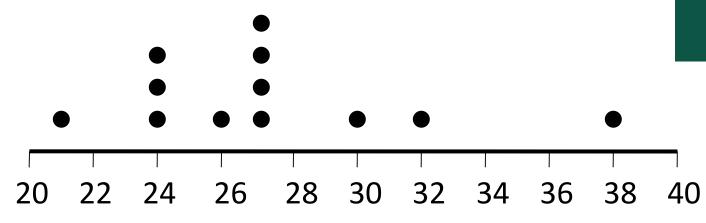


Diagrama de pontos: uma reta contendo a escala de valores dos dados com pontos representando os valores obtidos. O número de pontos sobre um dado valor representa a frequência deste.

Útil para um conjunto pequeno de dados !!!



Ramo-e-Folhas

É maneira simples de verificar os detalhes da distribuição de um conjunto pequenode dados.

Vantagem: Não se perdem as informações sobre os dados originais.

É obtido: dividindo-se cada valor dos dados em duas partes, o ramo e a folha.

Por exemplo, se os todos dados contém dois dígitos, tal como 36, podemos escolher a dezenas como ramo e a unidade como folha:



Ramo-e-Folhas

Exemplo: 21, 24, 24, 24, 26, 27, 27, 27, 27, 30, 32, 38

```
ramo folha
2 144467777
3 028
```



Ramo-e-Folhas

Exemplo: 21, 24, 24, 24, 26, 27, 27, 27, 27, 30, 32, 38

```
ramo folha
2 144467777
3 028
```

Os ramos devem ser escolhidos de modo que o gráfico ramos e folhas resultante seja informativo sobre a distribuição dos dados.



Ramo-e-Folhas

Exemplo: 21, 24, 24, 24, 26, 27, 27, 27, 27, 30, 32, 38

```
ramo folha
2 144467777
3 028
```

Os ramos devem ser escolhidos de modo que o gráfico ramos e folhas resultante seja informativo sobre a distribuição dos dados.

Exemplo: o mesmo exemplo com mais detalhes.

vamos dividir cada ramo em dois ramos: um contendo folhas de digitos ≤5 e outro com digitos >5.

Ramo-e-Folhas

Exemplo: 21, 24, 24, 24, 26, 27, 27, 27, 27, 30, 32, 38

ramo folha 2 144467777 3 028

Os ramos devem ser escolhidos de modo que o gráfico ramos e folhas resultante seja informativo sobre a distribuição dos dados.

Exemplo: o mesmo com mais detalhes. vamos dividir cada ramo em dois ramos: um contendo folhas de digitos ≤5 e outro com digitos >5.



Histograma:

O histograma é um gráfico de barras contíguas, com as bases proporcionais aos intervalos das classes e a área de cada retângulo proporcional à respectiva frequência. Pode-se usar tanto a frequência absoluta, f_i , como a relativa, f_i .

Indiquemos a amplitude do *i*-ésimo intervalo por h_i . Para que a área do retângulo respectivo seja proporcional a fr_i , a sua altura deve ser proporcional a fr_i/h_i (ou a f_i/h_i), que é chamada densidade de frequência da i-ésima classe.



Histograma:

Algumas informações:

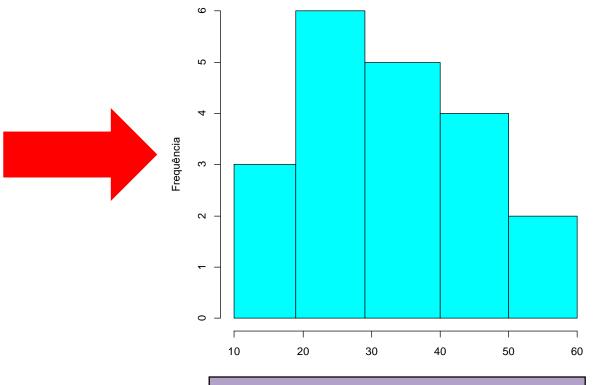
- Os pontos médios das classes são representados no eixo horizontal;
- No eixo vertical, além da densidade, também pode-se representar a frequência absoluta (f), a frequência relativa (fr) ou a porcentagem (fr%), então as barras são utilizadas para representar o número de observações (ou a frequência relativa ou a porcentagem) dentro de cada classe.



Histograma:

Classes	Ponto Médio	f _i
10 ├ 20	15	3
20 ⊢30	25	6
₃₀ ⊢ ₄₀	35	5
₄₀ ⊢ ₅₀	45	4
50 ├ 60	55	2

Histograma: Diâmetro das árvores de eucalipto.



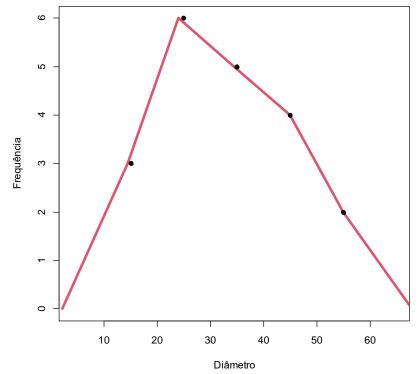
(não há espaço entre as barras)



Polígono de Frequência:

Classes	Ponto Médio	f _i
10 20	15	3
20 ⊢30	25	6
30 ⊢40	35	5
₄₀ ⊢ ₅₀	45	4
50 ├ 60	55	2

Polígono de Frequência: histograma onde as barras verticais são substituidas por linhas unindo a frequência dos pontos médios das classes.



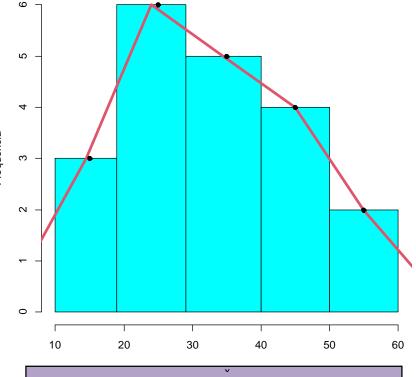


Histograma com ajuste do polígono:

Classes **Ponto** Médio 10 |--20 15 3 20 ⊢30 25 6 30 **⊢**40 35 5 40 **—**50 45 4 50 **⊢**60 55 2

Histograma com ajuste do polígono:

Diâmetro das árvores de eucalípto.



(não há espaço entre as barras)



Histograma:

Construção do histograma com as densidades de frequências e não com as frequências.

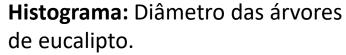
Quando os intervalos das classes (amplitudes) forem todos iguais a h, a densidade de frequência da i-ésima classe passa a ser $\mathrm{fr_i}$ /h (ou $\mathrm{f_i}$ /h), ou seja,

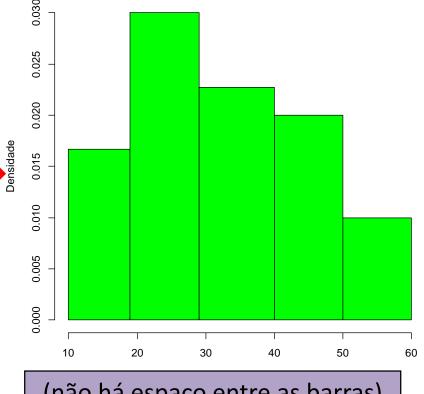
$$densidade_i = \frac{f_{\underline{r_i}}}{h}$$



Histograma:

Classes	Ponto Médio	fr	fr/h
10 20	15	0,15	0,015
20 ⊢30	25	0,30	0,030
₃₀ ⊢ ₄₀	35	0,25	0,025
₄₀ ⊢ ₅₀	45	0,20	0,020
50 ⊢ 60	55	0,10	0,010

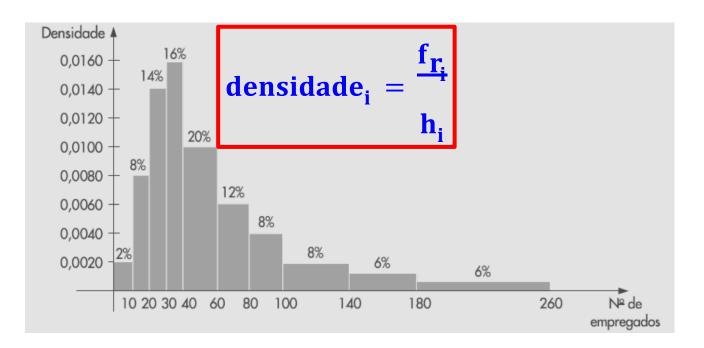




(não há espaço entre as barras)

Histograma:

No entanto ocorre muitas vezes que as amplitudes das classes serem desiguais, então, deve-se usar:





Histograma:

Observações:

- Para conjuntos grandes de dados (n ≥ 30).
- Mostram bem a dispersão (variabilidade) dos dados e sugerem valores ao redor dos quais os dados se agrupam (medidas de posição).
- Mostram bem o formato da distribuição dos dados .
- Não evidenciam bem a presença de dados atípicos.



Ogiva de Galton Crescente:

Definição: Ogiva de Galton crescente é a poligonal formada unindo-se os pontos de coordenadas (LS_i ; Fa_i) ou (LS_i , Far_i), (classe i = 1, ..., K), sendo:

- K: o número de classes;
- LS_i: o limite superior da *i*-ésima classe;
- Fa_i: frequência acumulada absoluta associadas à *i*-ésima classe;
- Far_i: frequência acumulada relativa associadas à *i*-ésima classe.



Ogiva de Galton Crescente:

Exemplo:

Tabela 5: Volume (ml).

Volume (ml)	Frequência absoluta f _i	Frequência relativa fr _i	Frequência acumulada Fa _i	Frequência acumulada relativa Far _i
[30; 50)	5	0,116	5	0,116
[50; 70)	11	0,256	16	0,372
[70; 90)	10	0,233	26	0,605
[90; 110)	5	0,116	31	0,721
[110; 130)	6	0,139	37	0,860
[130; 150)	4	0,093	41	0,953
[150; 170)	2	0,047	43	1,000
Total	43	1,000		

Fonte: Zocchi, 2008



Ogiva de Galton Crescente:

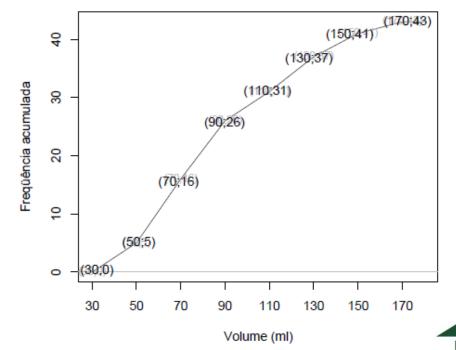
Exemplo: (Obs.: Note que, está sendo considerado (LS_i; Fa_i)).

Tabela 5: Volume (ml).

Fonte: Zocchi, 2008

Volume (ml)	Frequência acumulada Fa _i
[30; 50)	5
[50; 70)	16
[70; 90)	26
[90; 110)	31
[110; 130)	37
[130; 150)	41
[150; 170)	43
Total	

Exemplo: Ogiva de Galton crescente para a variável volume



Fonte: Zocchi, 2008

•

Universidade Federal de São Pau

Resumo

Baseados nos dados brutos:

- Gráfico de pontos .
- Diagrama de ramo e folhas.
- Gráfico de dispersão (XY).

Baseados em tabelas de frequências:

- Histograma e polígono de frequências.
- Ogiva de Galton crescente.
- Gráfico em barras horizontais/em colunas.
- Gráfico de setores.



O Mau Uso dos Gráficos e Questões Éticas

Orientações:

- Não distorcer os dados;
- Evitar adornos desnecessários;
- Utilizar uma escala para cada eixo em gráficos bi-dimensionais;
- A escala do eixo vertical deve começar em zero;
- Rotular todos os eixos;
- Colocar título no gráfico;
- Utilizar o modelo de gráfico mais simples possível.



CUIDADO: Figuras podem enganar





Para interpretar corretamente um gráfico, devemos analisar a informação numérica dada num gráfico, que pode ser distorcida pelo desenho (No exemplo a direita, pela escala vertical).



Informação pode ser distorcida ao usar-se desenhos para informar: Desenhos com área e volume podem passar falsa impressão para dados de natureza unidimensional

Exemplo:

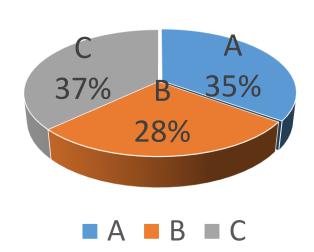
Preço em 1986

preço em 2003

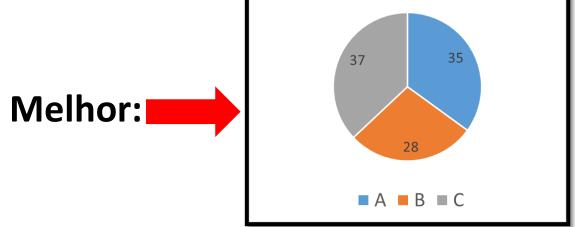


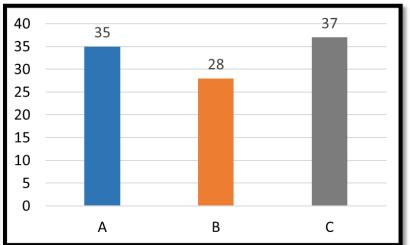


Note que o preço dobrou, mas a área da figura ficou 4 vezes maior, o que dá a impressão de um aumento muito maior



A figura distorce os valores: o ítem B tem menor porcentagem



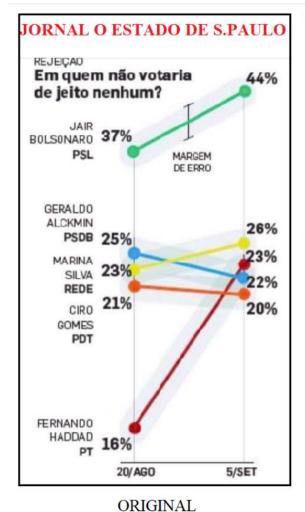


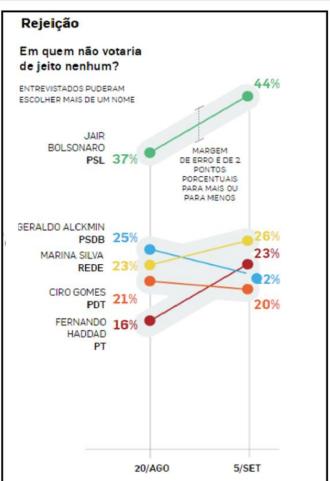




A apresentadora mostra que o número de casos confirmados é cerca de "um palmo" menor que o projetado pela Secretaria Estadual da Saúde da Bahia. Ufa, que alívio!







CORRIGIDO



Universidade Federal de São Paulo

PRÓXIMAS AULAS _____

- Medidas de posição;
- Medidas de variabilidade;
- Gráfico de Box-plot.



– REFERÊNCIAS –

MONTGOMERY, D. C.; RUNGER, G. C. Estatística aplicada e probabilidade para engenheiros. 6 ed. Rio de Janeiro: LTC, 2018. 628p.

MORETTIN, P. A.; BUSSAB, W. O. Estatística básica. 9 ed. São

Paulo: Saraiva, 2017. 554p.

ZOCCHY, S.S. Estatística descritiva, 2008.



CLASS FINISHED

