

## 2 ESTATÍSTICA DESCRITIVA

### 2.3 Descrição gráfica

Silvio Sandoval Zocchi

11 de agosto de 2008

## 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
- Outros

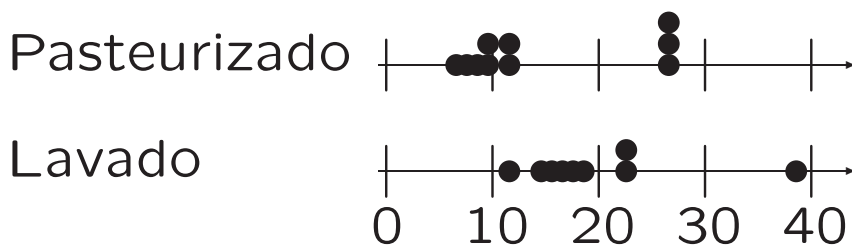
## Gráfico de pontos

Exercício 1. Cogumelo *Pleurotus sajor-caju*



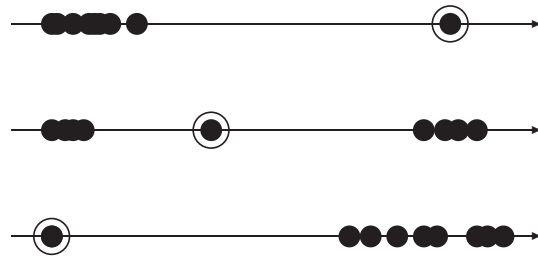
Tabela 6. Eficiência biológica.

Pasteurizado	11%	26%	26%	6%	26%
	9%	8%	11%	9%	7%
Lavado	14%	15%	18%	16%	11%
	22%	38%	17%	22%	



- Para conjuntos pequenos ( $n \leq 15$ ) de dados numéricos
- Mostram bem a **dispersão** (variabilidade) dos dados e sugerem valores ao redor dos quais os dados se agrupam (medidas de **posição**)
- Evidenciam a presença eventual de **dados atípicos** (dados distintamente diferentes dos grupos que incluem maioria dos dados)

## Conjuntos de dados com um dado atípico

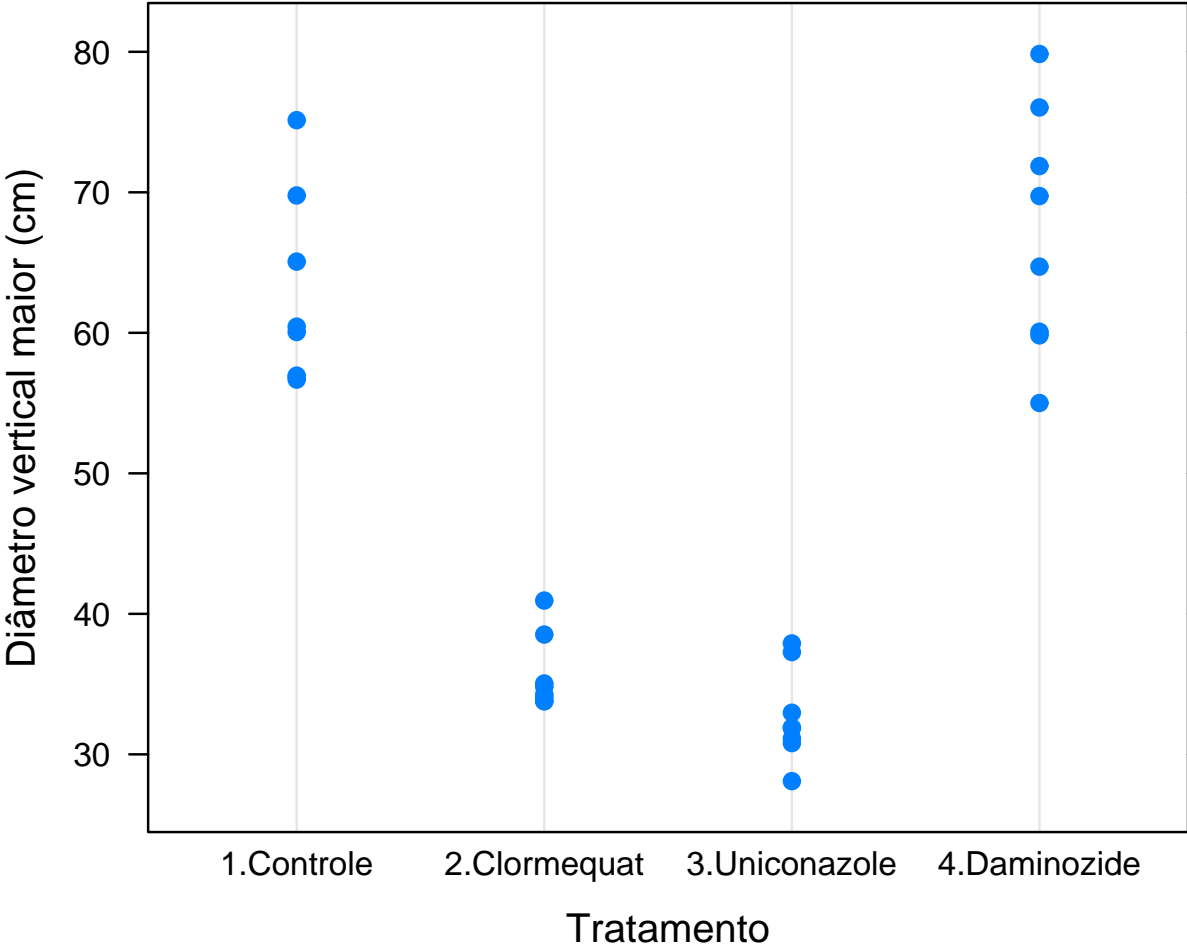


A presença de dados atípicos nos leva a refletir sobre a **qualidade** do dados

Exercício 2. Retardantes de crescimento em *Hibiscus sp* envasados



# Resposta



## Diagrama de ramo e folhas (Tukey, 1977)

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

Idéia básica: dividir cada dado em duas partes: a primeira (o ramo) e a segunda (a folha)

Exemplo: Comprimentos de asas de insetos, em mm (Silveira Neto *et al*, 1976)

2		2356
3		055889
4		111112334469
5		0038
6		0000059
7		127
8		35
9		
10		
11		3
12		
13		8

Legenda: 13|8=13,8mm

Exemplo: Dados de eficiência biológica (Lavado)

1		458617
2		22
3		8

Ordenação

1		145678
2		22
3		8

Legenda:  $1|5=15\%$

Dividindo os ramos

1 *		14	}	Dados de 10 a 14
1 o		5678	}	Dados de 15 a 19
2 *		22		
2 o				
3 *				
3 o		8		

Legenda:  $1o|5=15\%$



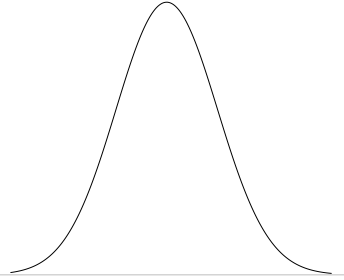
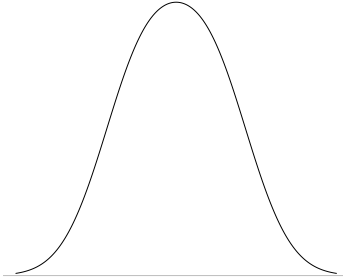
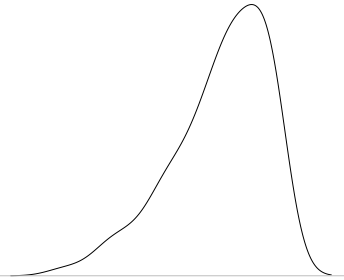
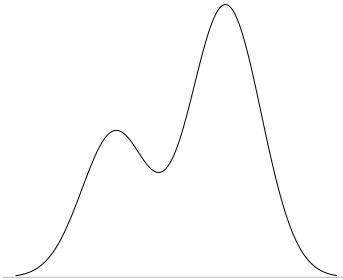
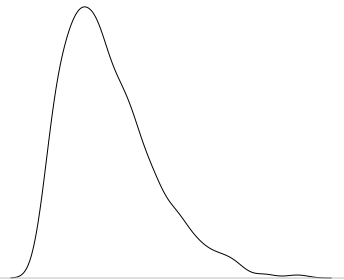
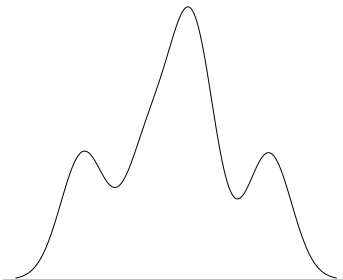
Exercício 3. Resistência de uvas Niágara, em N (Tabela 5)

Produtor 7		Produtor 58	
4		4	4
5	02489	5	4
6	134479	6	128
7	0233556666	7	4478
8	0223479	8	22445566667
9	249	9	234
10	044	10	37
11	0	11	034467
12		12	288
13		13	
14		14	
15		15	
16		16	
17		17	
18	0	18	

Legenda: 5|4=0,54 N

Resumidamente, os diagramas de ramos e folhas:

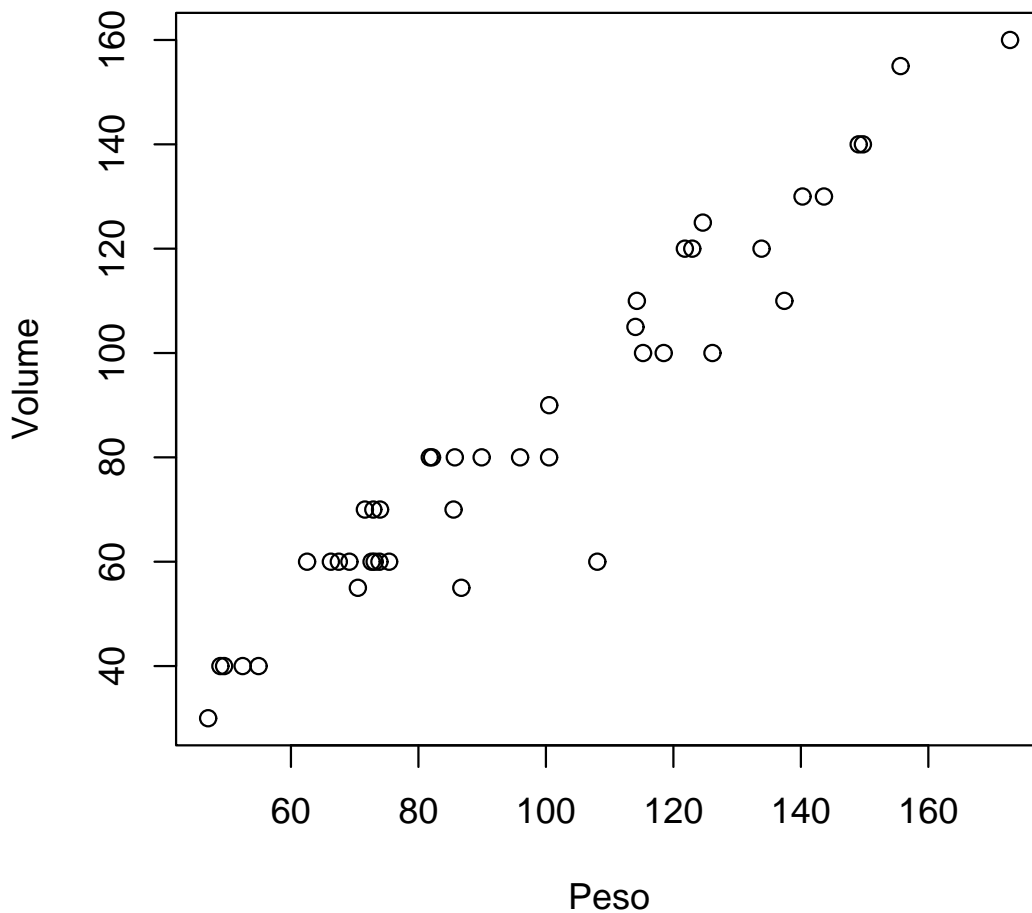
- São recomendados para conjuntos de dados numéricos de quaisquer tamanhos
- Mostram bem a **dispersão** (variabilidade) dos dados e sugerem valores ao redor dos quais os dados se agrupam (medidas de **posição**)
- Mostram bem o formato da **distribuição** dos dados
- Evidenciam a presença de **dados atípicos**

Classificação de uma distribuição quanto à simetria	ao nº de modas
 <p data-bbox="428 856 685 907">Simétrica</p>	 <p data-bbox="1026 856 1279 907">Unimodal</p>
 <p data-bbox="246 1243 863 1293">Assimétrica à esquerda</p>	 <p data-bbox="1042 1243 1263 1293">Bimodal</p>
 <p data-bbox="279 1621 831 1671">Assimétrica à direita</p>	 <p data-bbox="1000 1621 1312 1671">Multimodal</p>

## Gráfico de dispersão (XY)

Sejam  $X$  e  $Y$  duas variáveis quantitativas com  $n$  observações cada. Gráfico de dispersão é o conjunto de pontos de coordenadas  $(x_i, y_i)$ ,  $(i = 1, \dots, n)$ .

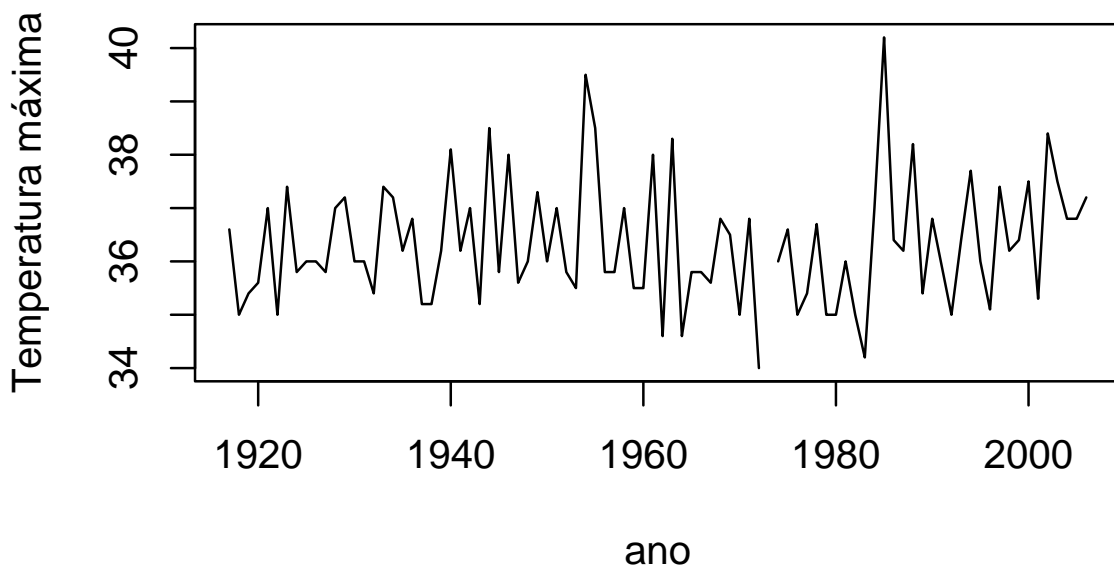
Ex: Gráfico de dispersão Peso  $\times$  Volume



Observação: Quando a variável  $Y$  é observada ao longo do tempo, ela é chamada **série temporal**.

Nesse caso, considera-se  $X$  como a variável **data** e costuma-se unir os pontos com linhas retas.

Exemplo: Série de temperaturas máximas anuais, em  $^{\circ}\text{C}$ , em Piracicaba.



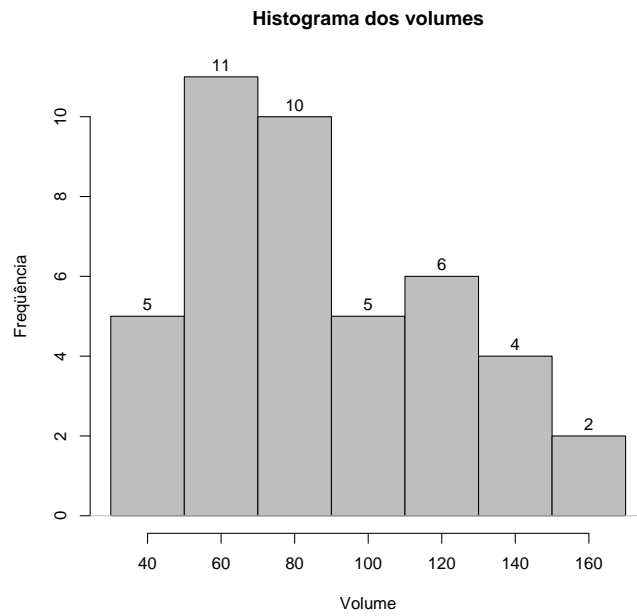
## Distribuição de freqüências

Tabela 3. Distribuição de freqüências para a variável volume de cachos de uva fornecidos pelo produtor n<sup>o</sup> 59, na safrinha de 2002.

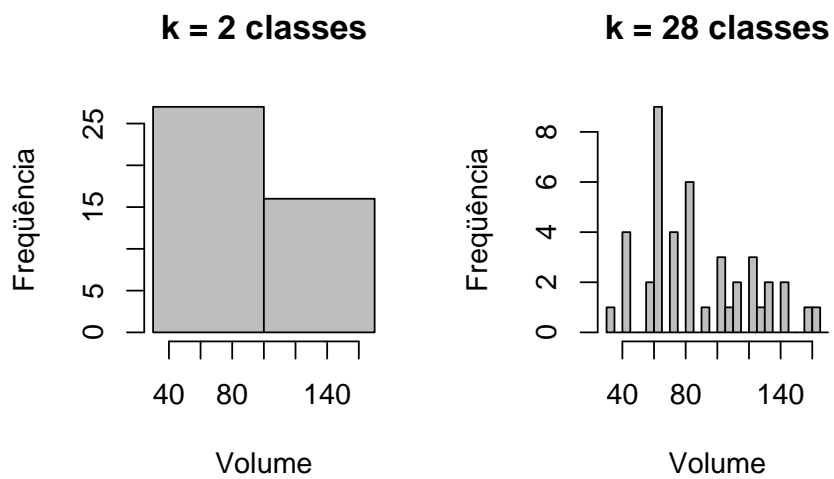
Volume	Ponto médio ( $m_i$ )	Freqüência	
		Absoluta ( $f_i$ )	Relativa ( $f'_i$ )
[ 30 ; 50[	40	5	0,116
[ 50 ; 70[	60	11	0,256
[ 70 ; 90[	80	10	0,233
[ 90 ; 110[	100	5	0,116
[110 ; 130[	120	6	0,140
[130 ; 150[	140	4	0,093
[150 ; 170[	160	2	0,047
Total		43	1,000

Como representá-la graficamente?

# Histograma



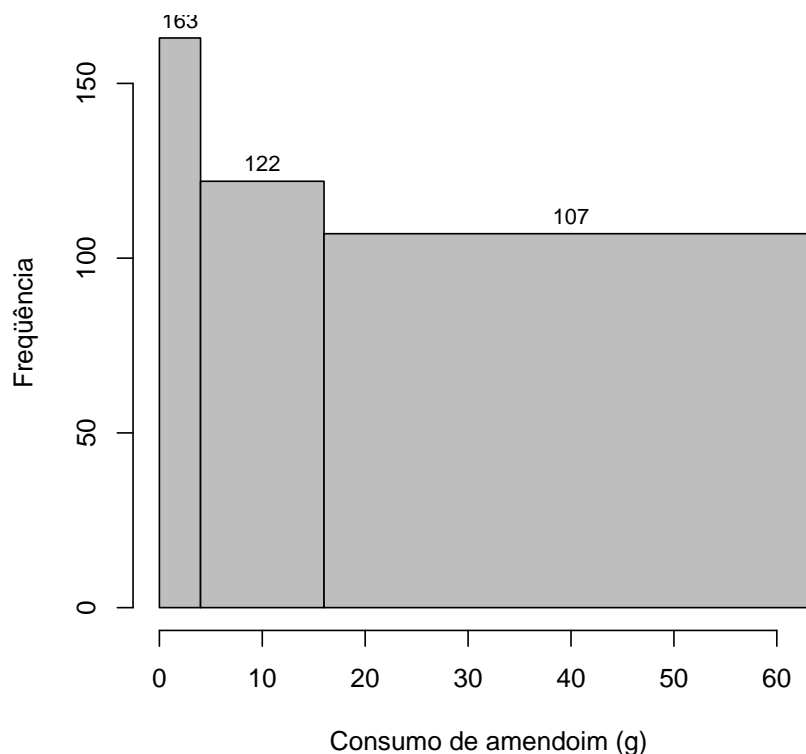
- Influência do número  $k$  de classes



- Classes com amplitudes diferentes

São recomendadas especialmente para variáveis com distribuições assimétricas.

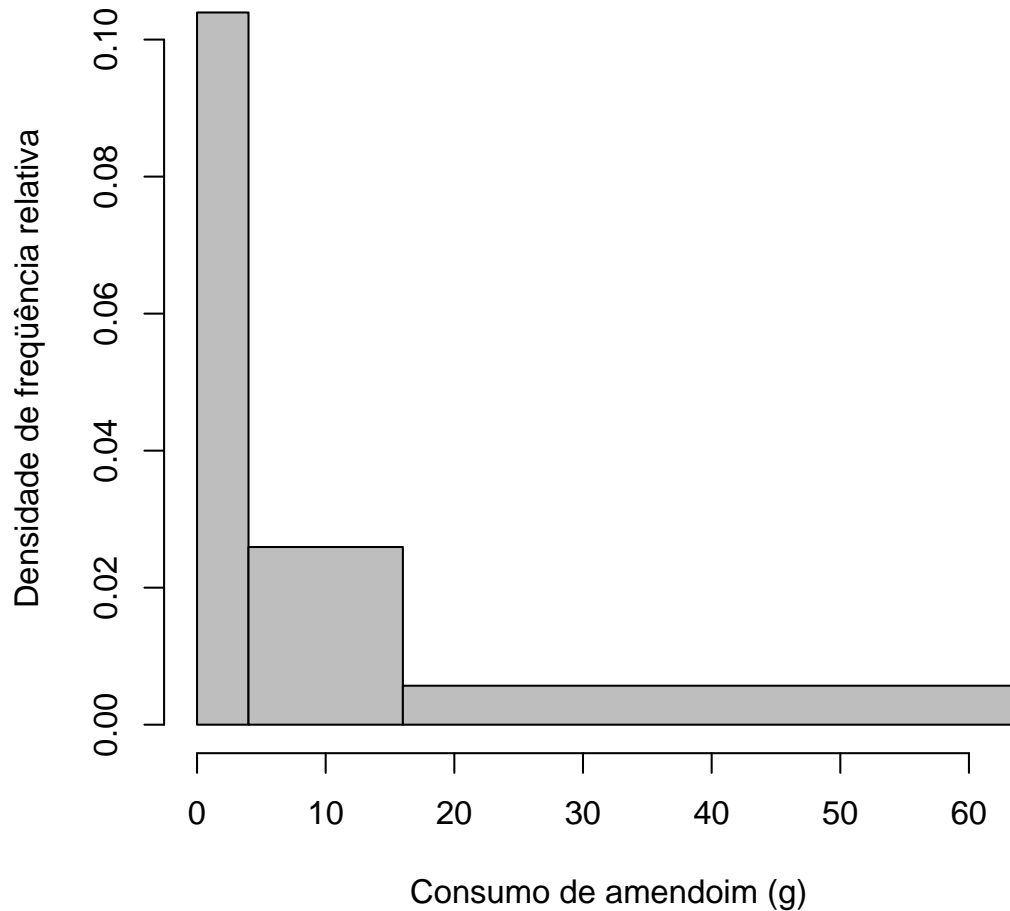
Exemplo. Consumo diário, em g, de amendoim em Piracicaba.



Cuidado!!! Histograma incorreto!!!



Atenção! Nesses casos, no entanto, devemos utilizar as **densidades de frequência** como altura dos retângulos.



**Definição:** Histograma é uma seqüência de retângulos postos lado a lado onde cada retângulo tem, como base, a amplitude da classe e como altura:

- a freqüência absoluta ou relativa, no caso de classes com amplitudes iguais, ou
- a densidade de freqüência absoluta ou relativa, no caso de classes com amplitudes diferentes.

$$\text{Densidade de freqüência} = \frac{\text{freqüência}}{\text{amplitude}}$$

- Para conjuntos grandes de dados ( $n \geq 30$ )
- 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

## Polígono de freqüências

**Definição:** Polígono de freqüências é a poligonal formada unindo-se os pontos de coordenadas  $(m_i, f_i)$  ou  $(m_i, f'_i)$ ,  $(i = 0, 1, \dots, k, k+1)$ , sendo:

$k$ , o número de classes

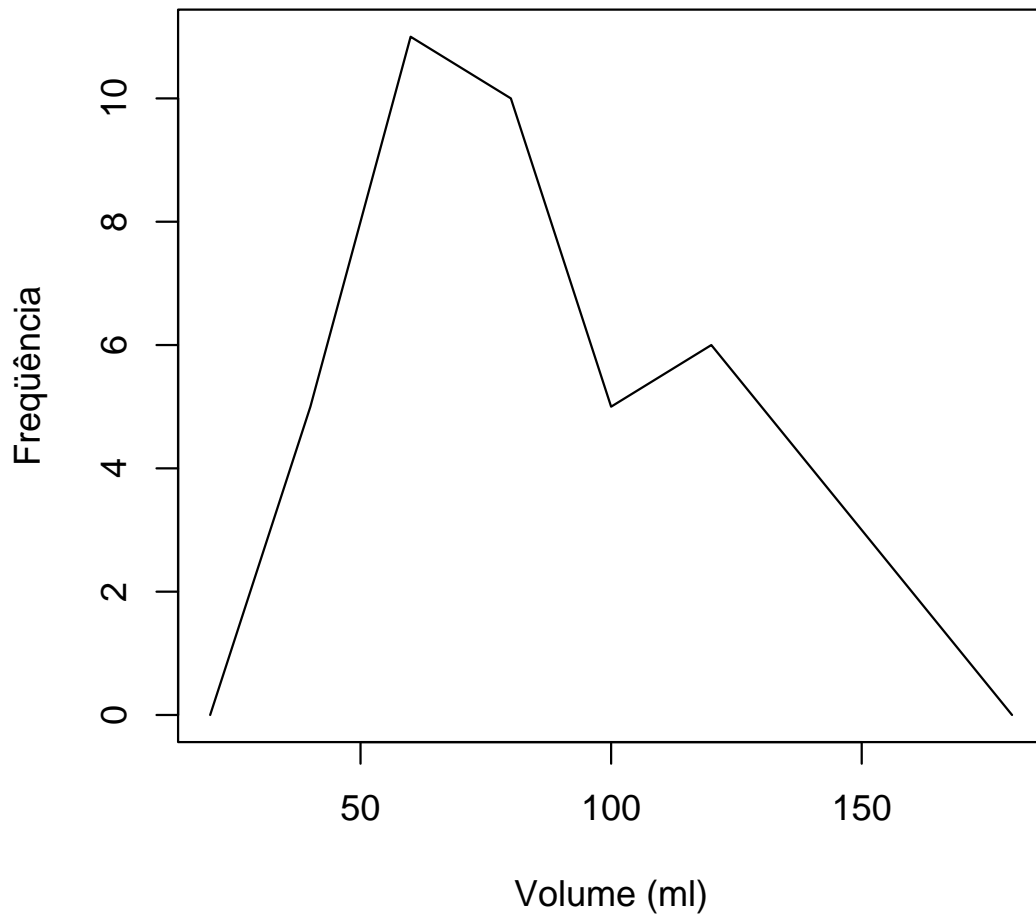
$m_i$ , o ponto médio da  $i$ -ésima classe

$f_i$ , a freq. absoluta associada à  $i$ -ésima classe

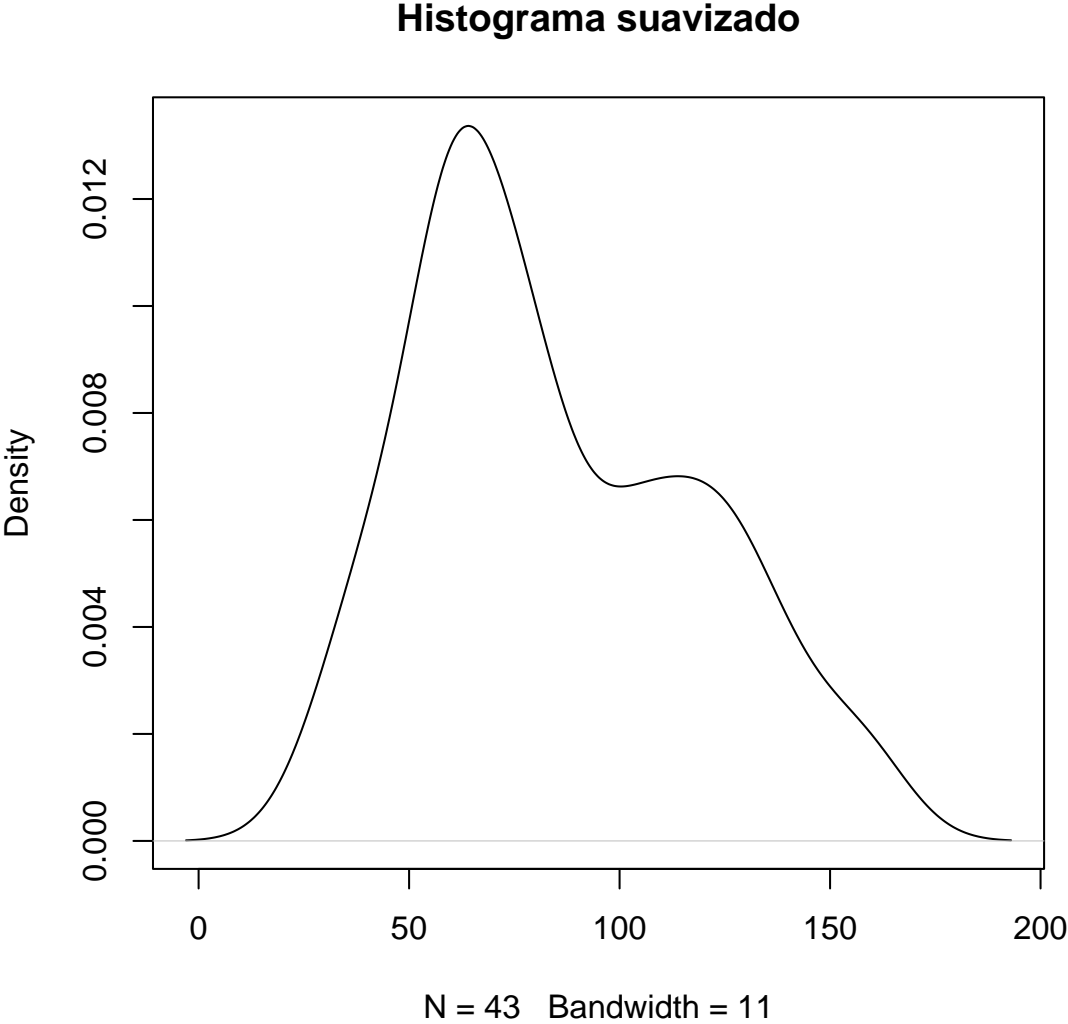
$f'_i$ , a freq. relativa associada à  $i$ -ésima classe

Alternativa ao histograma com interpretação e uso semelhantes

Ex: Polígono de frequências para a variável volume



# Histograma suavizado



Como podemos representar a tabela de frequências acumuladas?

## Ogiva de Galton Crescente

**Definição:** Ogiva de Galton crescente é a poligonal formada unindo-se os pontos de coordenadas  $(LS_i, a_i)$  ou  $(LS_i, a'_i)$ ,  $(i = 1, \dots, k)$ , sendo:

$k$ , o número de classes

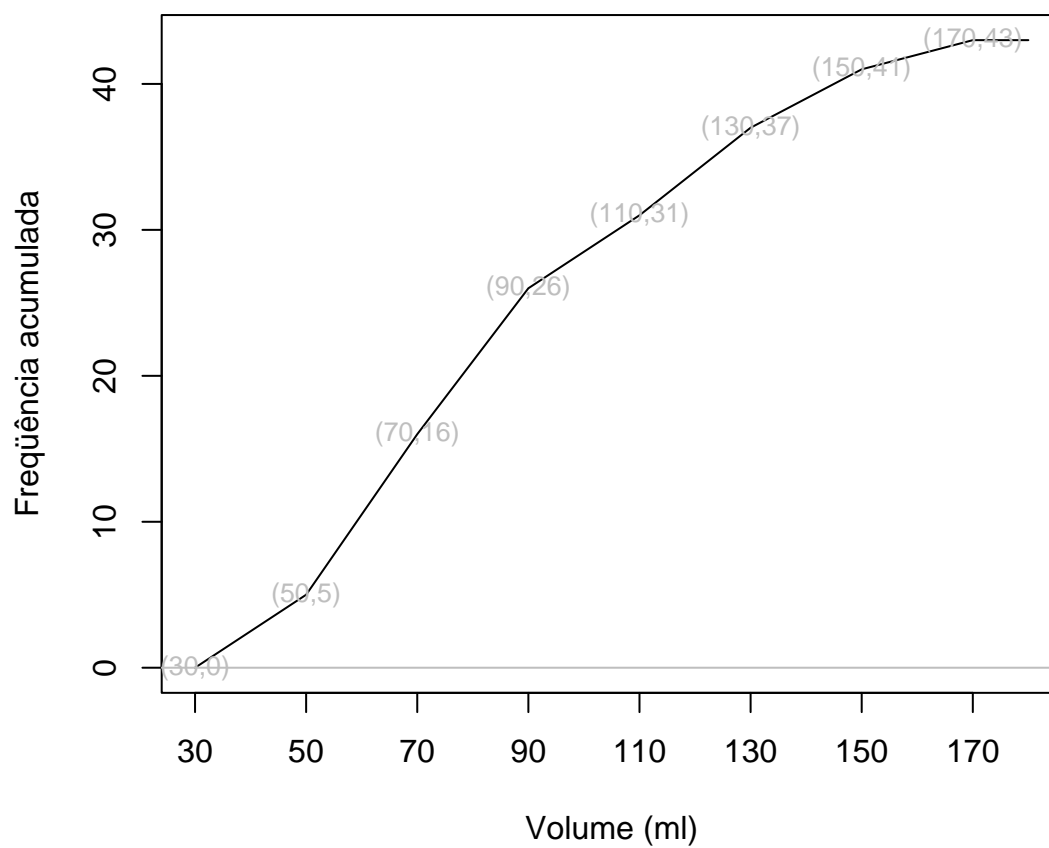
$LS_i$ , o limite superior da  $i$ -ésima classe

$a_i$  e  $a'_i$  as freqüências acumuladas absoluta e relativa associadas à  $i$ -ésima classe

### Exemplo

Volume (ml)	Freqüência acumulada	
	Absoluta ( $a_i$ )	Relativa ( $a'_i$ )
[ 30 ; 50[	5	0,116
[ 50 ; 70[	16	0,372
[ 70 ; 90[	26	0,605
[ 90 ; 110[	31	0,721
[110 ; 130[	37	0,860
[130 ; 150[	41	0,953
[150 ; 170[	43	1,000

Exemplo: Ogiva de Galton crescente para a variável volume



## Representação gráfica de tabelas de frequências para variáveis qualitativas ou quantitativas discretas

Tabela 10. Números de nematóides de solo, segundo o gênero e a época do ano.

Gênero	Época do ano		Total
	Antes do plantio	Depois da colheita	
Helicotylenchus	13972	29997	43969
Pratylenchus	1312	1924	3236
Mesocriconema	2044	1164	3208
Trichodorus	1088	970	2058
Scutellonema	720	980	1700
Heterodera	16	332	348
Meloidogyne	48	64	112
Aphelenchoides	108	0	108
Rotylenchulus	0	56	56
Xiphinema	8	0	8
Total	19316	35487	54803



Tabela. Distribuição de frequências da intensidade do odor de pedaços de abacaxi, emitidas por 40 provadores, segundo a variedade.

Intensidade do odor	Variedade de abacaxi	
	Smooth Cayenne	Pérola
muito fraco	2	5
fraco	9	14
médio	15	12
forte	10	8
muito forte	4	1
Total	40	40

Tabela. Frequências observadas do número de sementes caídas por golpe, dados em quatro semeadoras manuais reguladas para caírem duas sementes por golpe.

N <sup>o</sup> de sementes por golpe	Semeadora			
	A	B	C	D
0	0	2	14	7
1	19	26	21	19
2	103	70	82	82
3	18	48	27	42
4	3	4	5	0
5	7	0	1	0
Total	150	150	150	150

## Gráfico de setores

Para variáveis qualitativas nominais

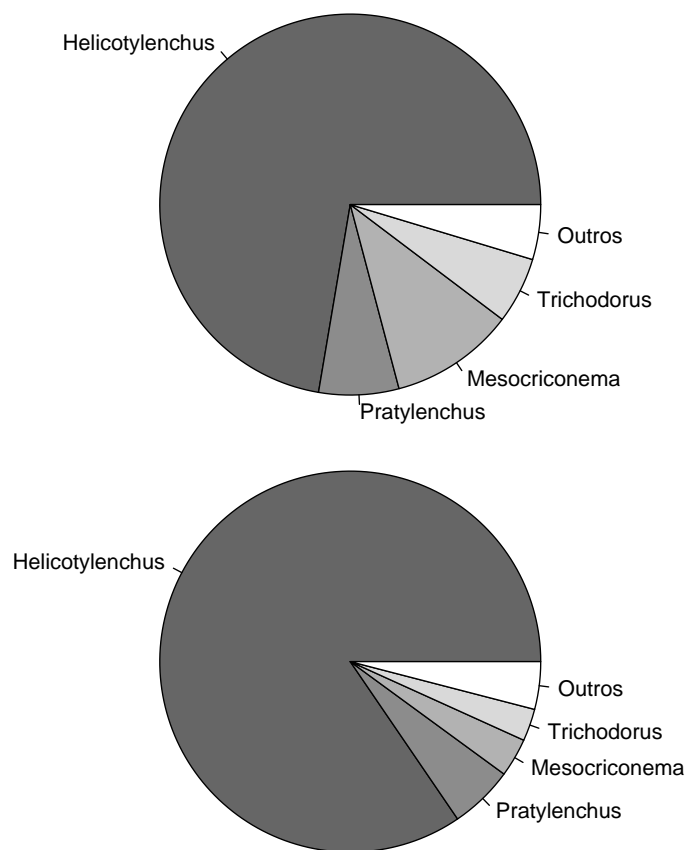
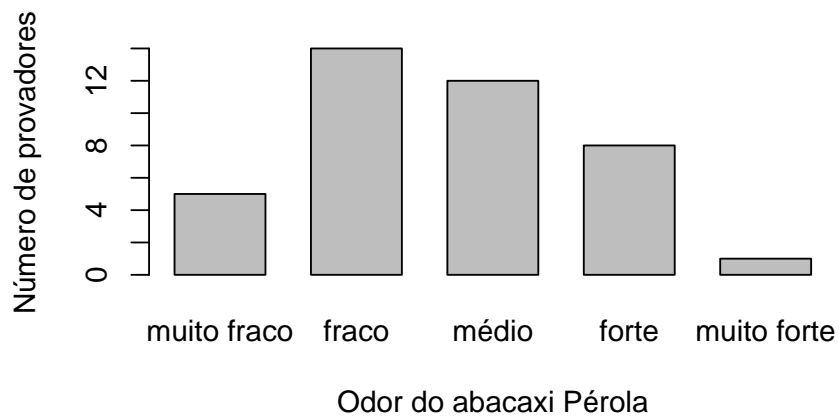
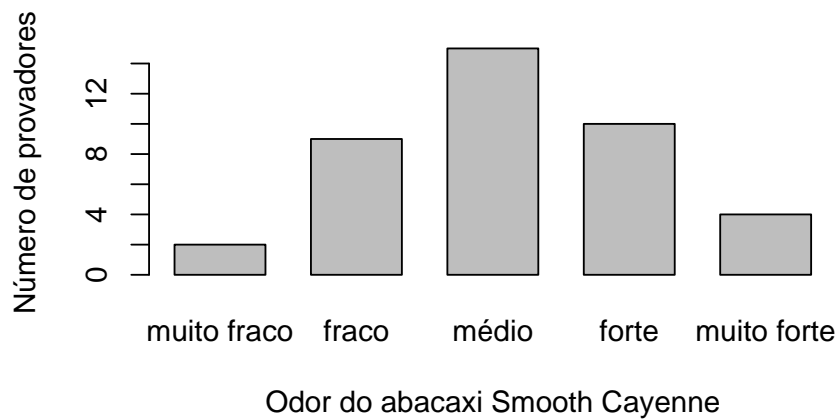


Figura. Gêneros de nematóides encontrados no solo antes do plantio da soja e depois da colheita.

## Gráfico em colunas

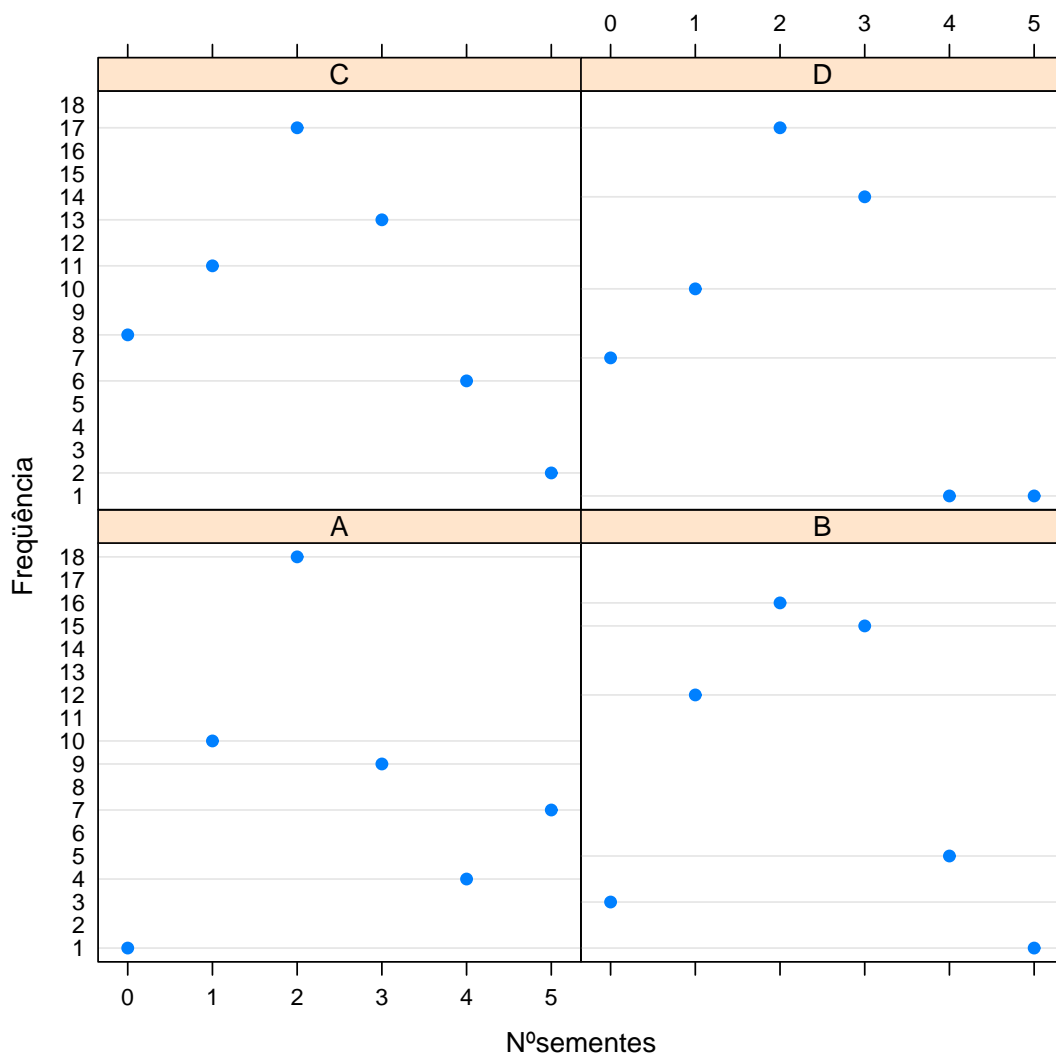
Para variáveis qualitativas ordinais ou quantitativas discretas.



## Gráfico de dispersão

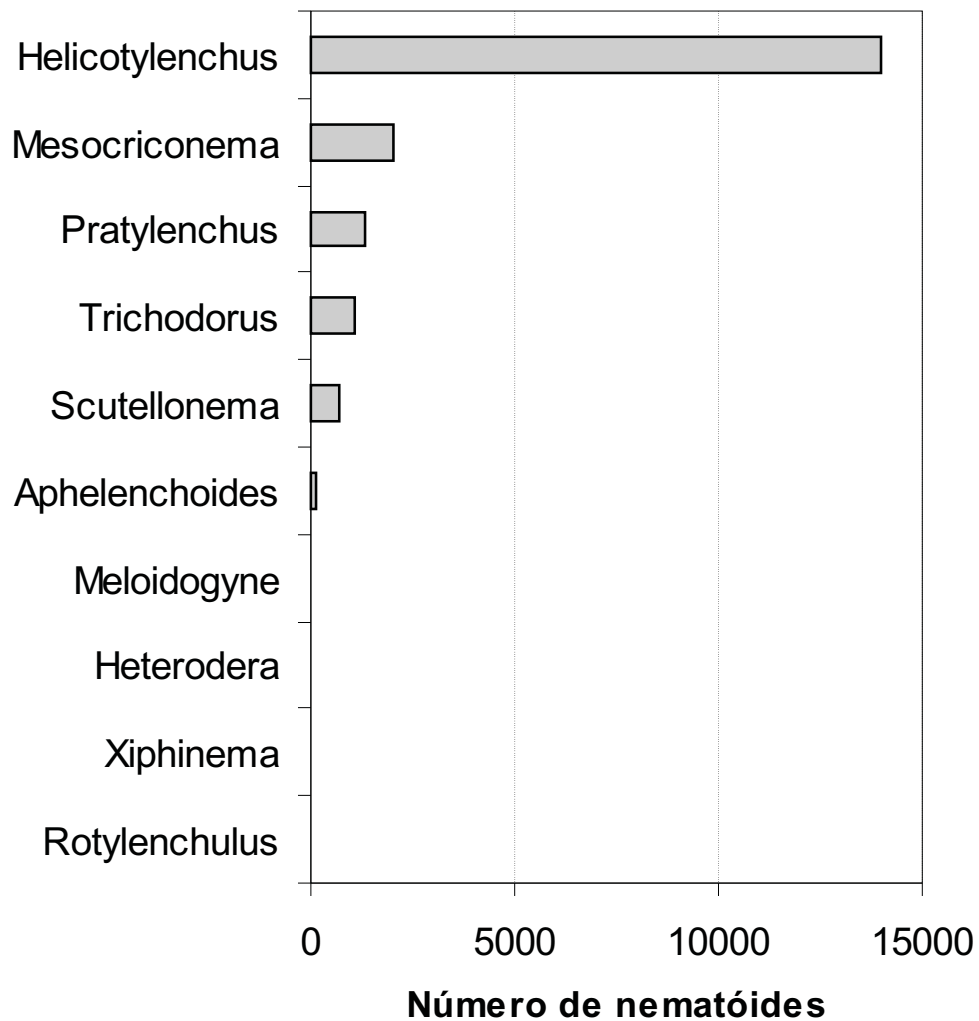
Conjunto de pontos de coordenadas  $(x_i, f_i)$ ,  
 $(i = 1, \dots, k)$

Para variáveis quantitativas discretas.



## Gráfico em barras horizontais

Para variáveis qualitativas com um número grande de categorias.



Outros gráficos