



# Estadística Visualização de Dados

Quadro 1 – Modelo de quadro

<b>ÁREAS</b>	<b>UNESP</b>	<b>UNICAMP</b>	<b>USP</b>	<b>TOTAL</b>
<b>Interdisciplinar</b>	2	2	2	6
<b>Biológicas e da Saúde</b>	2	2	2	6
<b>Exatas e Tecnológicas</b>	2	2	2	6
<b>Humanas e Artes</b>	2	2	2	6
<b>TOTAL</b>	8	8	8	24

Fonte: Modelo de quadro.

A tabela deve ser completa o suficiente para entendimento, o leitor não deve recorrer ao texto para interpretar a tabela;

Tabela 1 – Modelo de tabela

<b>ÁREAS</b>	<b>UNESP</b>	<b>UNICAMP</b>	<b>USP</b>	<b>TOTAL</b>
Interdisciplinar	2	2	2	6
Biológicas e da Saúde	2	2	2	6
Exatas e Tecnológicas	2	2	2	6
Humanas e Artes	2	2	2	6
<b>TOTAL</b>	<b>8</b>	<b>8</b>	<b>8</b>	<b>24</b>

Fonte: Modelo de fonte.

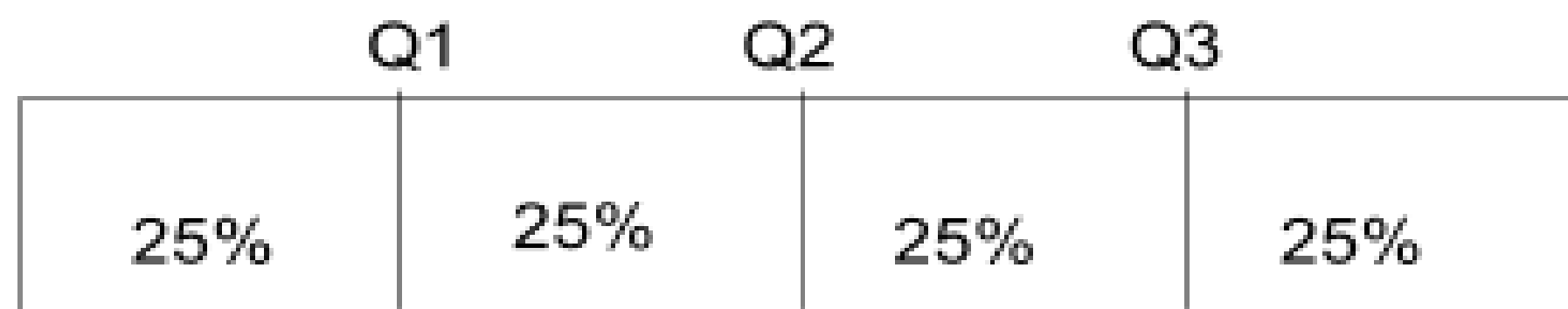
**Título e fonte deve deixar claro a resposta de duas questões: dizer onde e quando.**

# Quartis

Se dividirmos a série ordenada em quatro partes, cada uma ficara com seus 25% de seus elementos. Os elementos que separam estes grupos são chamados de **quartis**.

Assim, o primeiro quartil, que indicaremos por Q1, separa a sequência ordenada deixando 25% de seus valores à esquerda e 75% de seus valores a direita. Q3 é o contrário.

O segundo quartil, que indicaremos por Q2, separa a sequência ordenada deixando 50% de seus valores à esquerda e 50% de seus valores à direita. É a **mediana**



# Quintis

Se dividirmos a série ordenada em cinco partes, cada uma ficara com seus 20% de seus elementos. Os elementos que separam estes grupos são chamados de **quintis**.

Assim, o primeiro quintil, que indicaremos por K1, separa a sequência ordenada deixando 20% de seus valores à esquerda e 80% de seus valores a direita.

De modo análogo são definidos os outros quintis.

	K1	K2	K3	K4	
	20%	20%	20%	20%	20%

# Decis

Se dividirmos a série ordenada em dez partes, cada uma ficara com seus 10% de seus elementos. Os elementos que separam estes grupos são chamados de **decis**.

Assim, o primeiro quintil, que indicaremos por D1, separa a sequência ordenada deixando 10% de seus valores à esquerda e 90% de seus valores a direita.

De modo análogo são definidos os outros decis.

	D1	D2	.....	D9
10%	10%	.....10%.....		10%

# Percentis

Se dividirmos a série ordenada em cem partes, cada uma ficara com seus 1% de seus elementos. Os elementos que separam estes grupos são chamados de **percentis ou centis**.

Assim, o primeiro quintil, que indicaremos por P1, separa a sequência ordenada deixando 1% de seus valores à esquerda e 99% de seus valores a direita.

De modo análogo são definidos os outros percentis.

P1	P2 A P100
1%	99%

# Percentis

Note que:

## QUARTIS- PERCENTIS

$$Q_1 = P_{25}$$

$$Q_2 = P_{50}$$

$$Q_3 = P_{75}$$

## QUINTIS-PERCENTIS

$$K_1 = P_{20}$$

$$K_2 = P_{40}$$

$$K_3 = P_{60}$$

$$K_4 = P_{80}$$

## DECIS - PERCENTIS

$$D_1 = P_{10}$$

$$D_2 = P_{20}$$

$$D_3 = P_{30}$$

$$D_4 = P_{40}$$

$$D_5 = P_{50}$$

$$D_6 = P_{60}$$

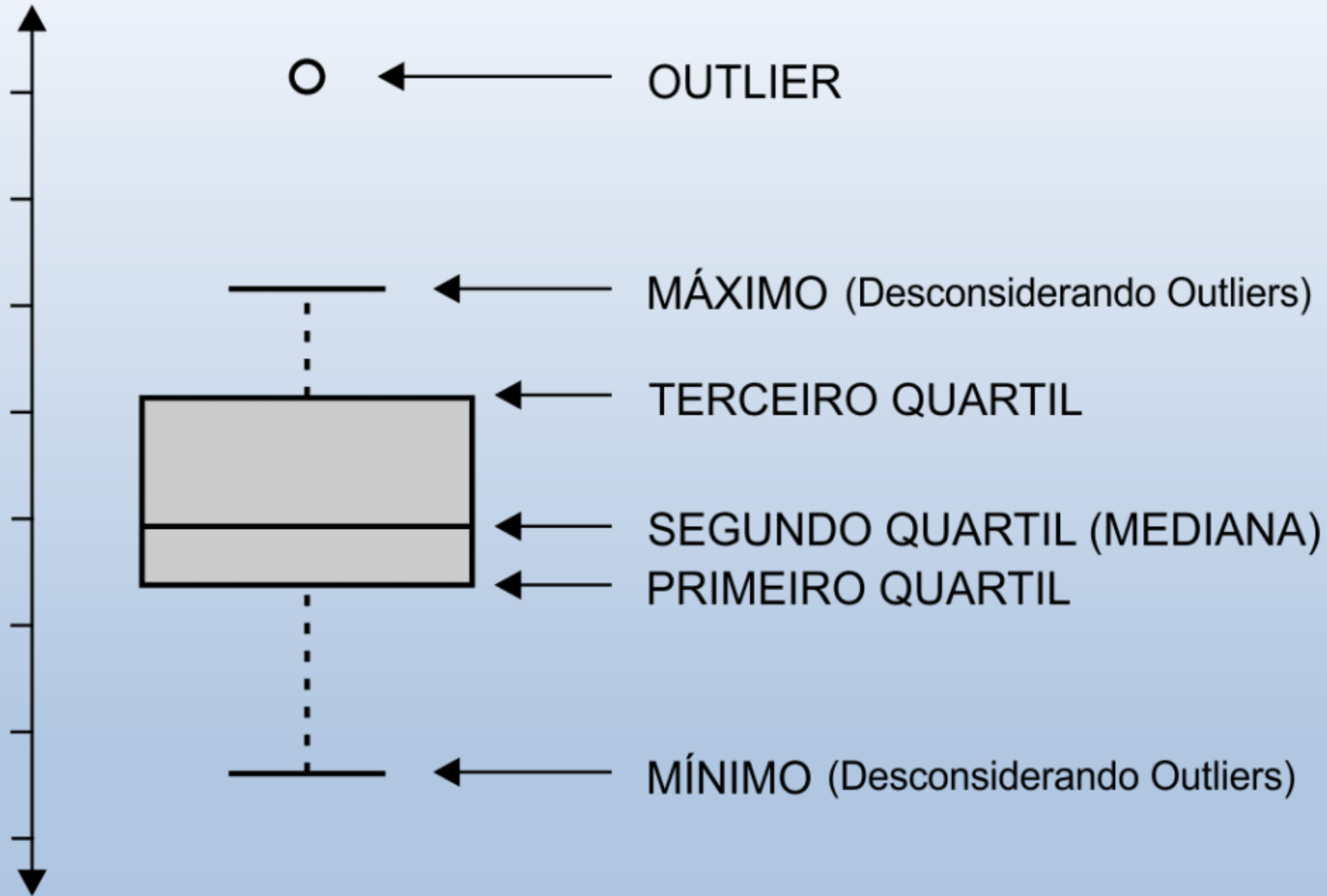
$$D_7 = P_{70}$$

$$D_8 = P_{80}$$

$$D_9 = P_{90}$$



# BOX PLOT



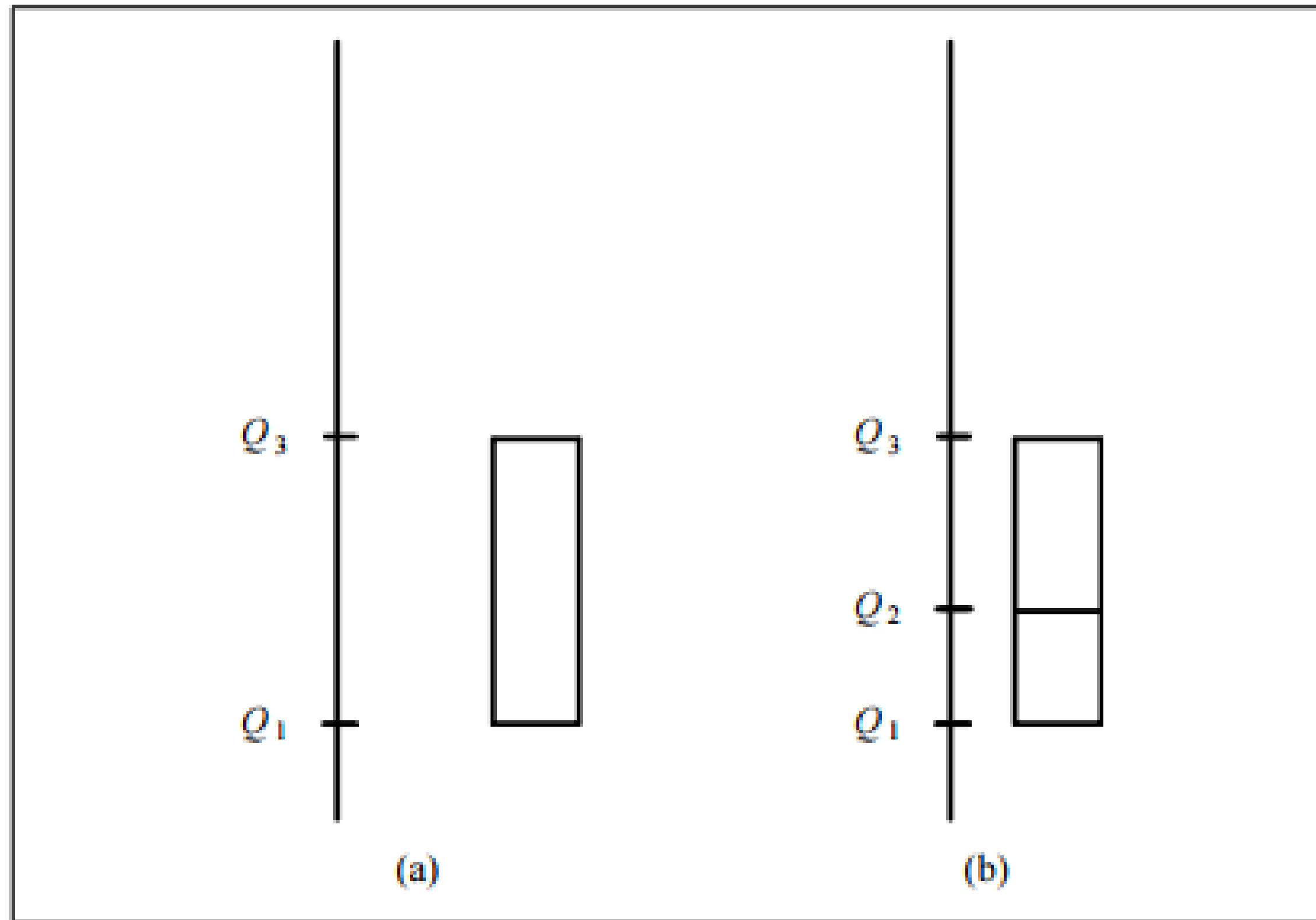
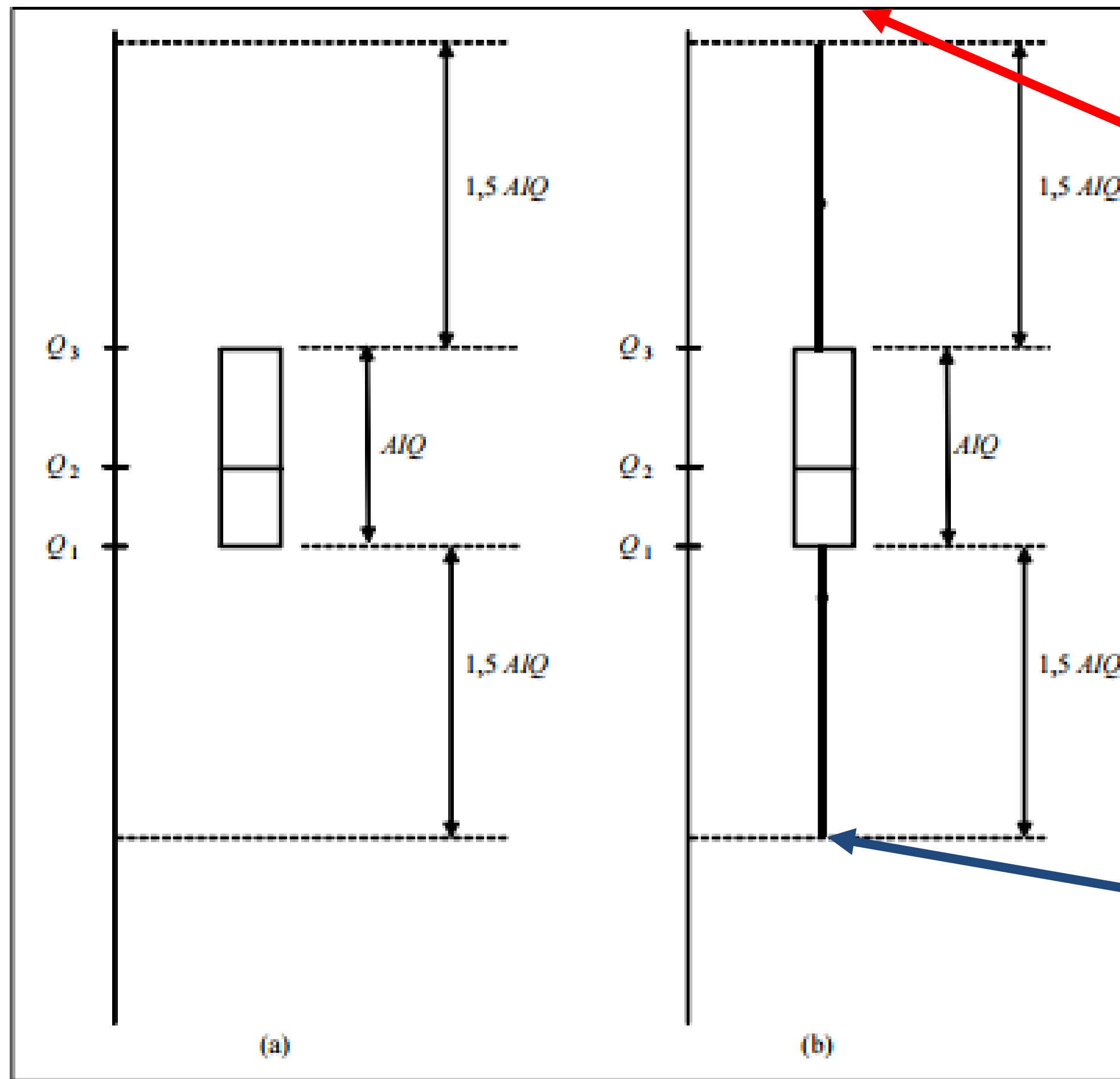


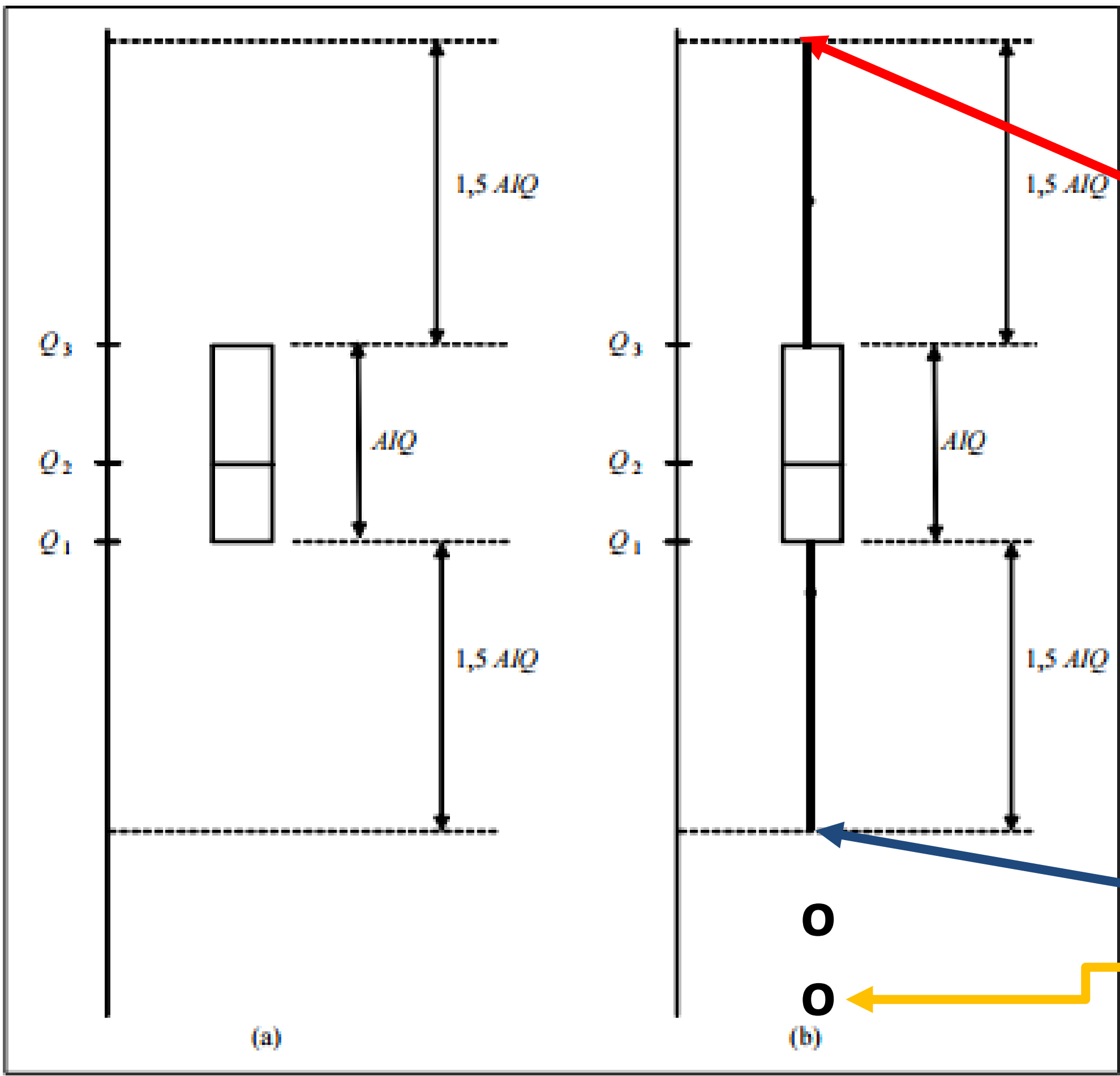
Figura 2.7 – Construção do boxplot - Parte 1



**Máximo desconsiderando o outlier**

**Mínimo desconsiderando o outlier**

Figura 2.8 - Construção do boxplot - Parte 2



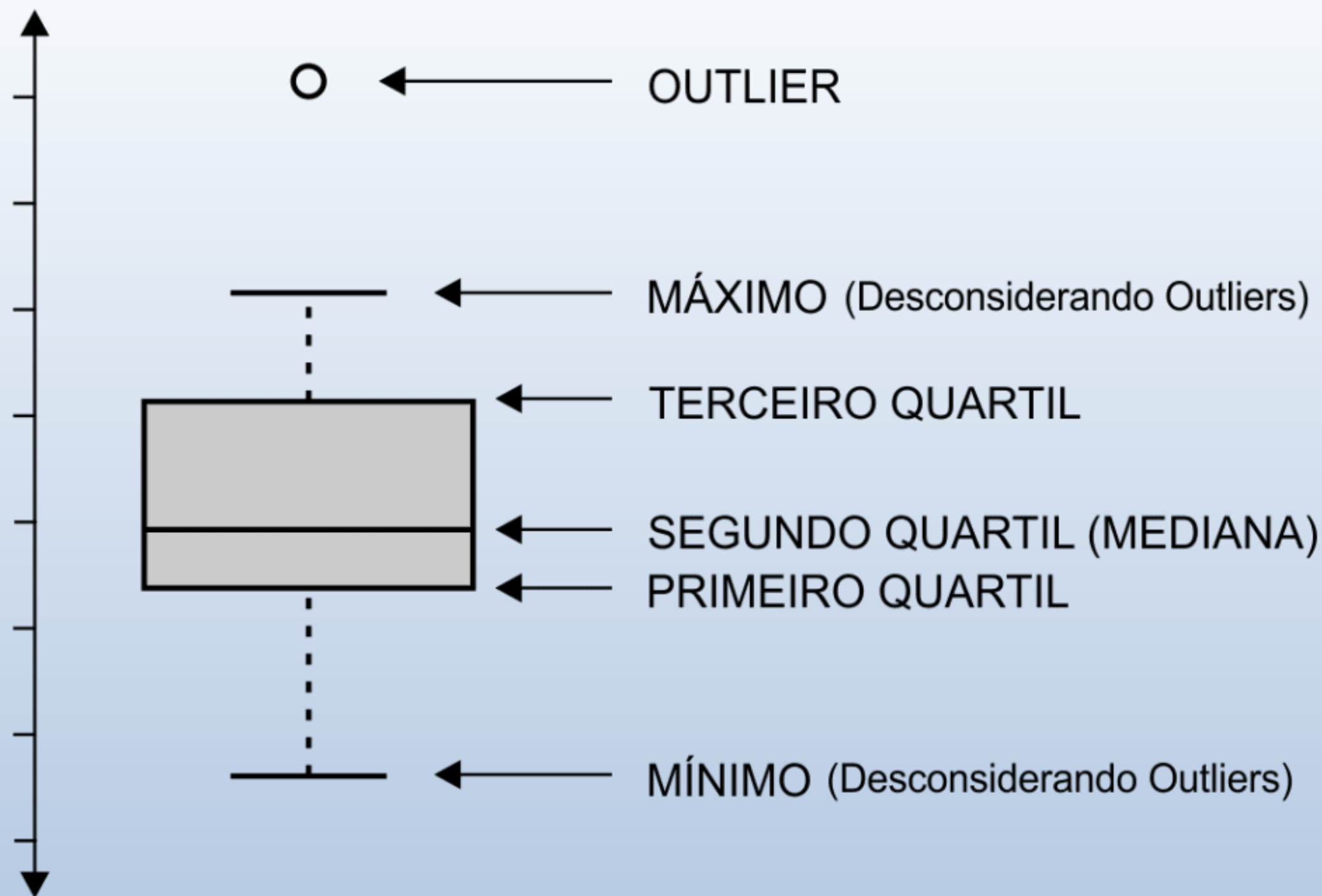
**Máximo desconsiderando os outliers**

**Mínimo desconsiderando os outliers**

**os outliers**

Figura 2.8 - Construção do boxplot - Parte 2

# BOX PLOT



Intervalo Interquartilico  
 $IQ=Q3-Q1$

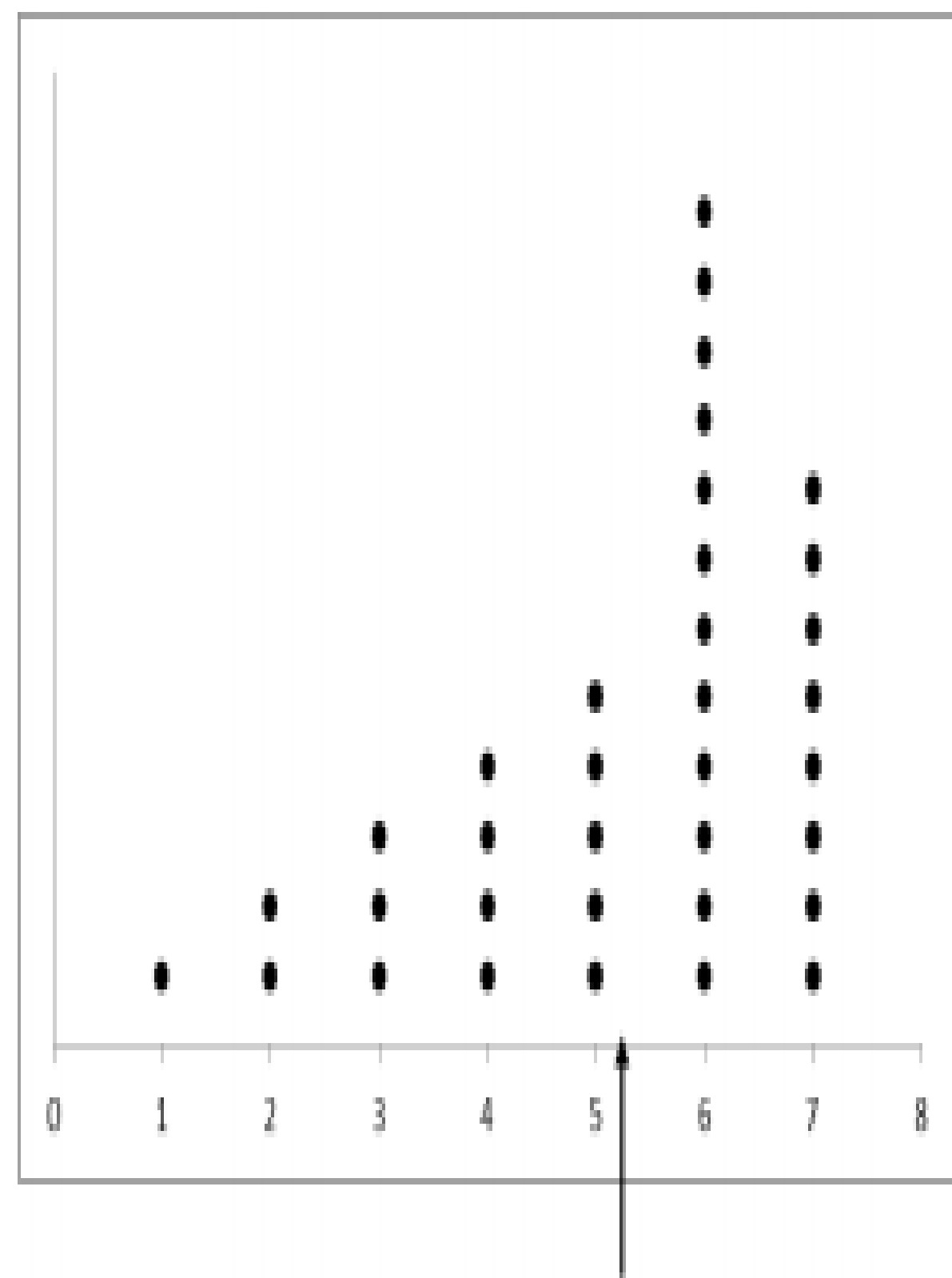
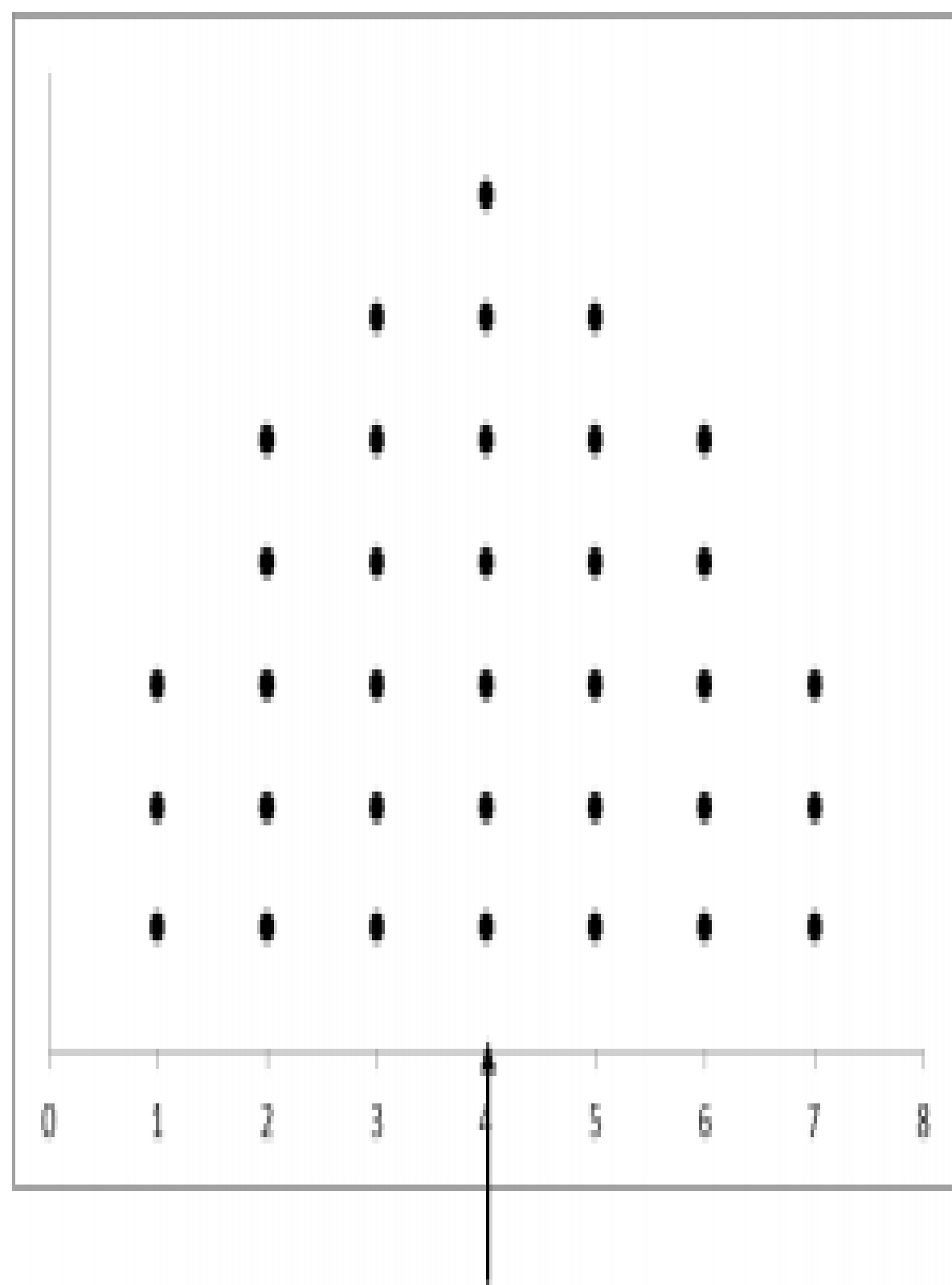
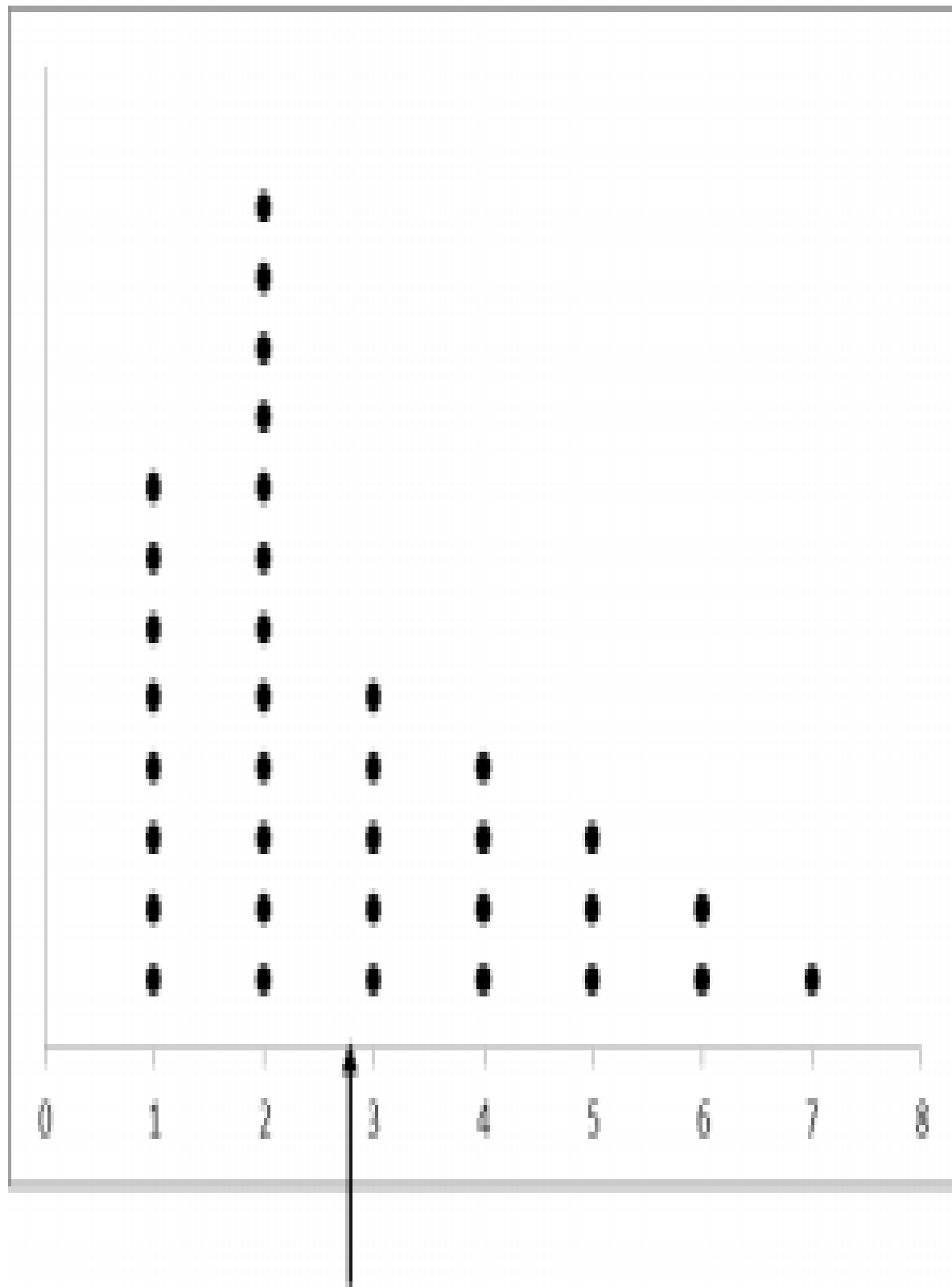
Já os outliers indicam possíveis valores discrepantes. No boxplot, os outliers estão abaixo ou acima do limite de detecção de outliers (Máximo e Mínimo)

$$\text{Máximo} = Q3 + 1.5 * (Q3 - Q1)$$

$$\text{Mínimo} = Q1 - 1.5 * (Q3 - Q1)$$

# BOX PLOT

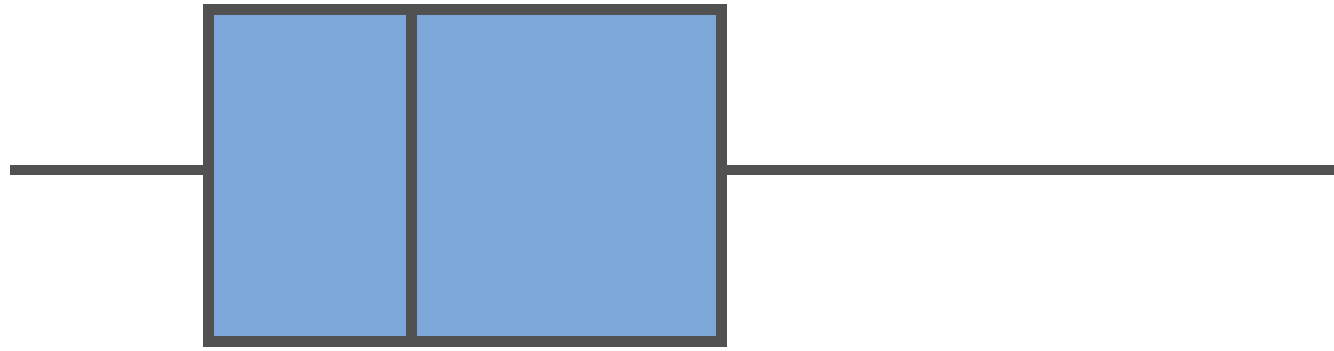




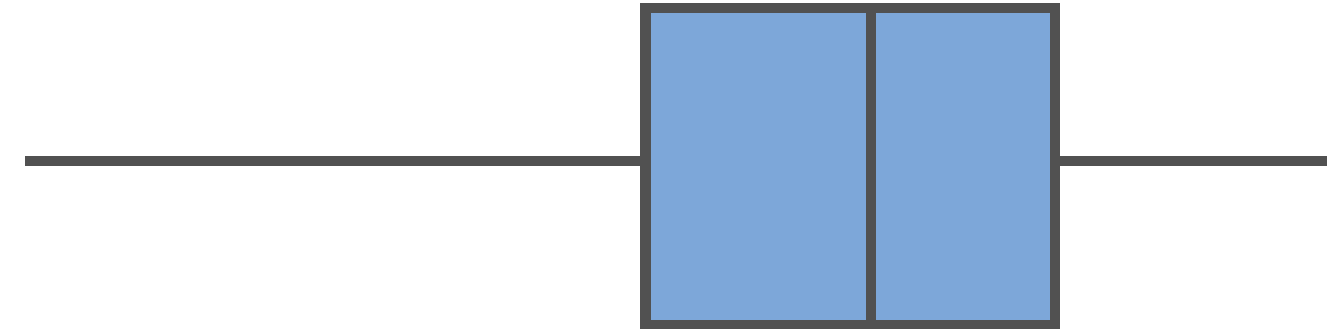
**Setas representam a média**

Figura 2.5 – Distribuições com diferentes tipos de assimetria

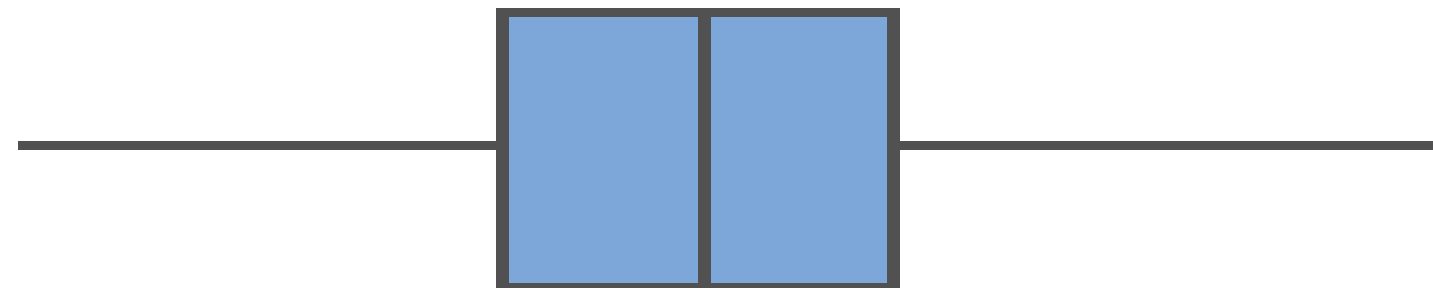
## Boxplot Assimétrico Positivo



## Boxplot Assimétrico Negativo

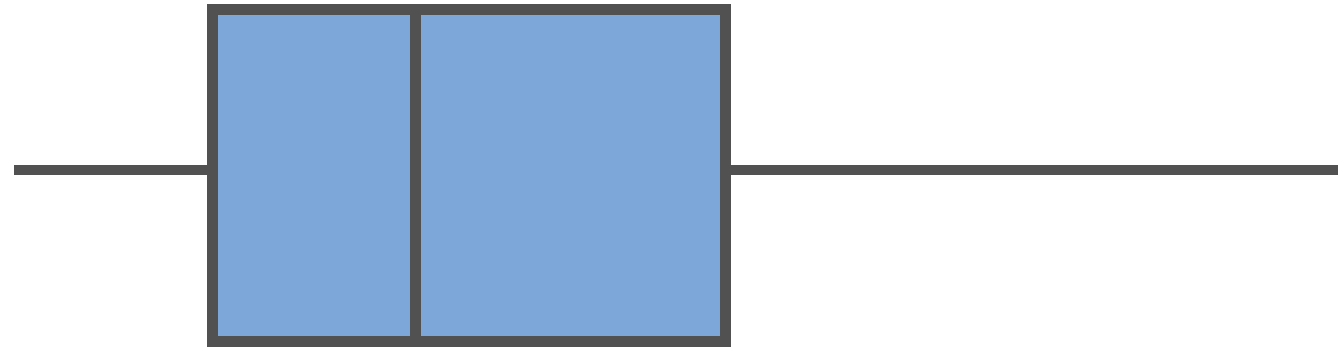


## Boxplot Simétrico





# Boxplot Assimétrico Positivo



# Boxplot Assimétrico Negativo

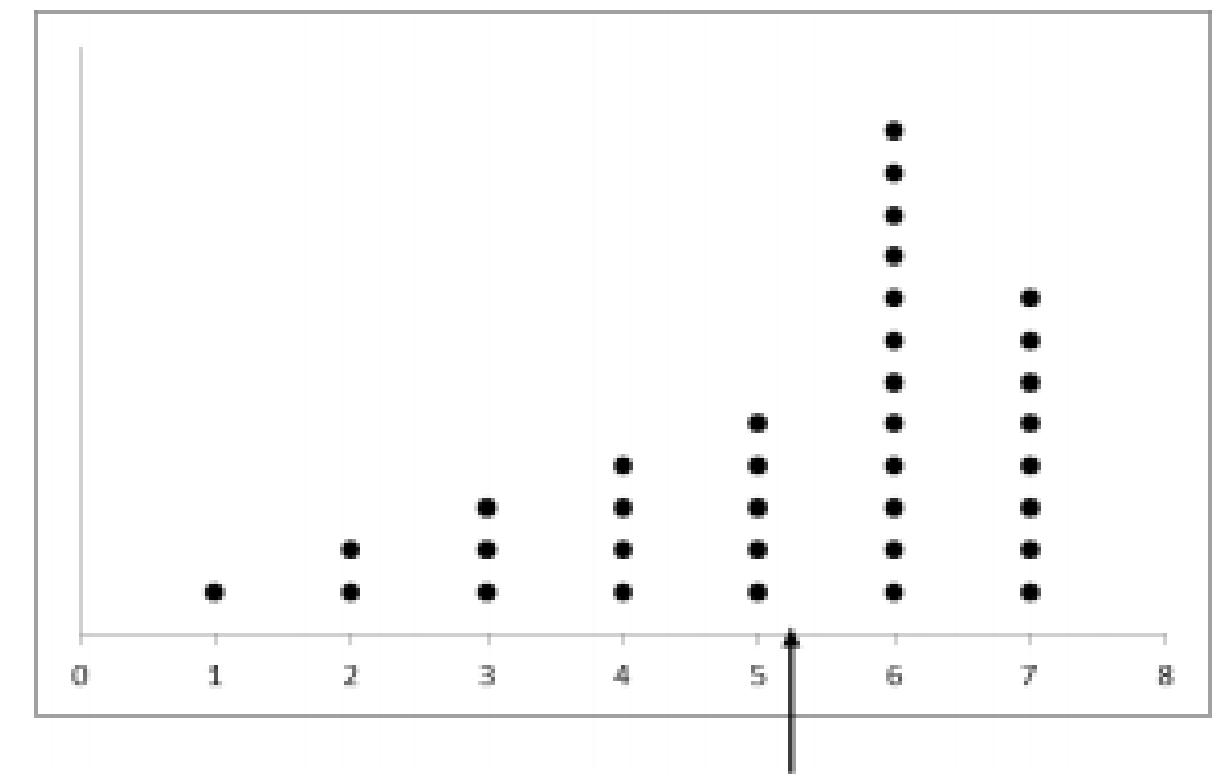
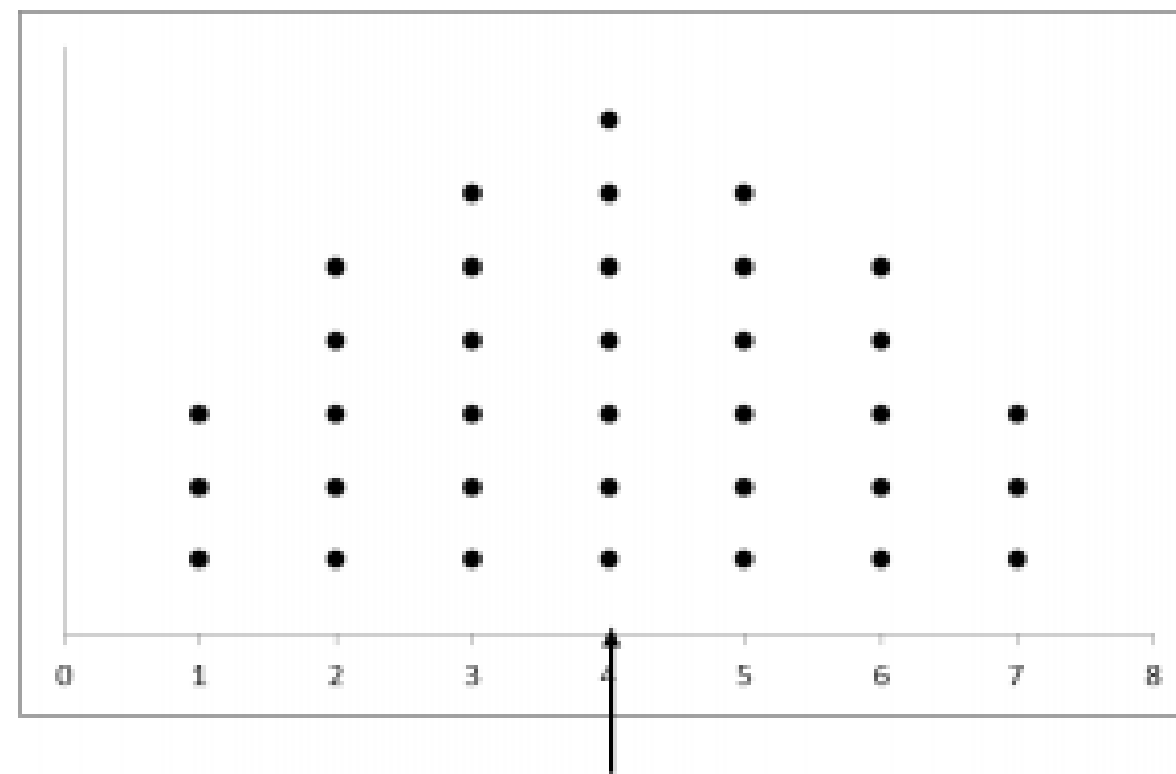
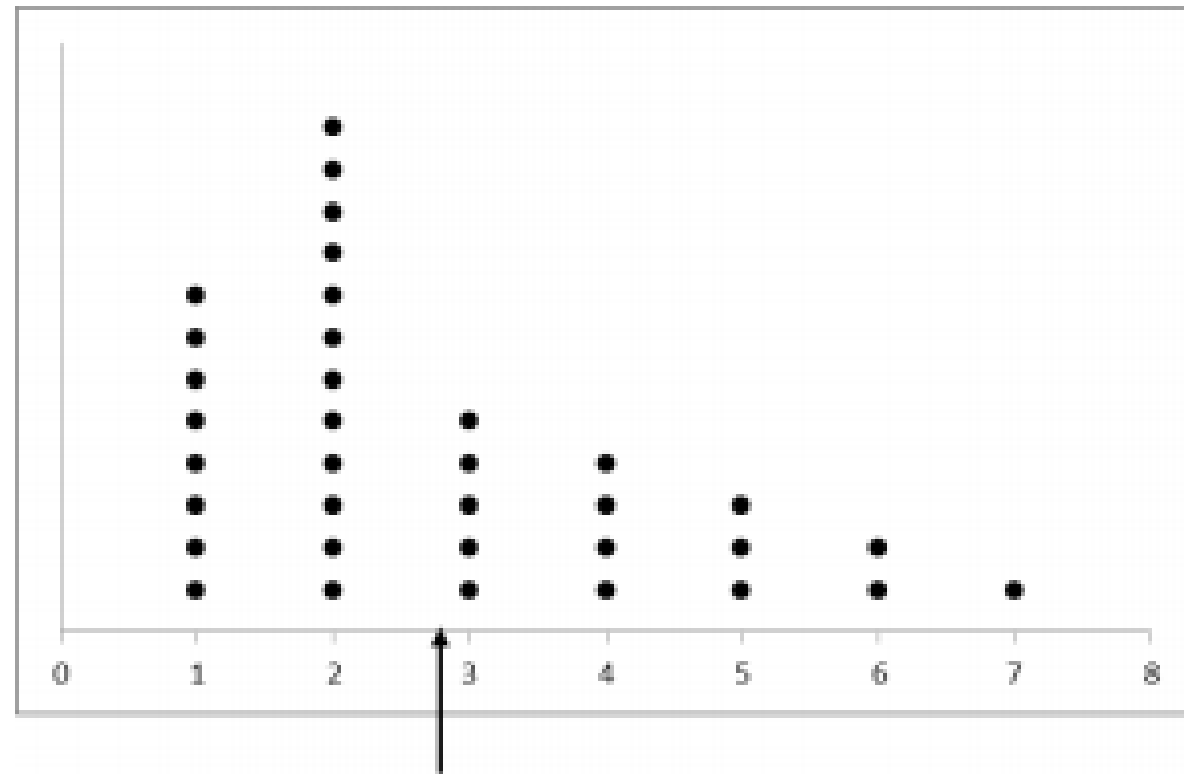
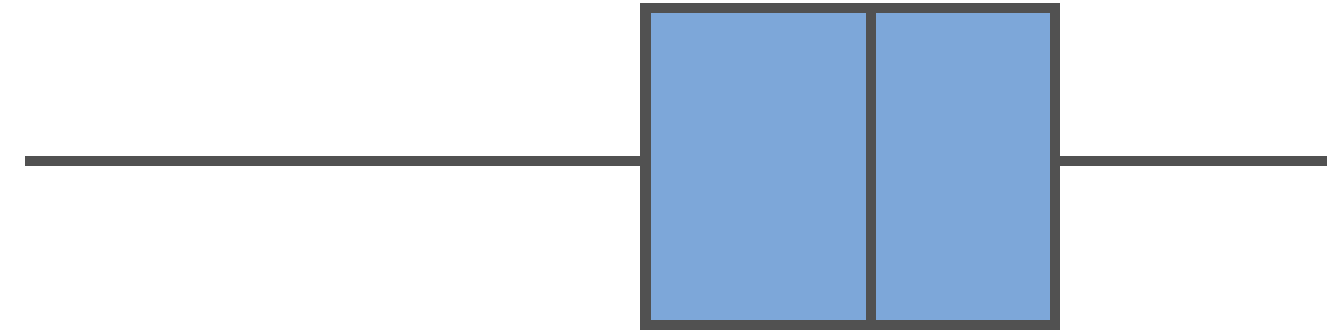
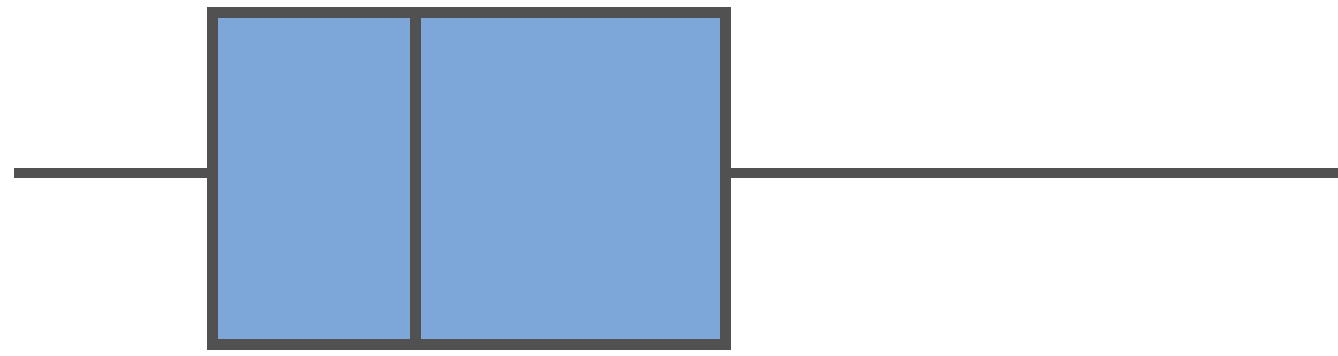
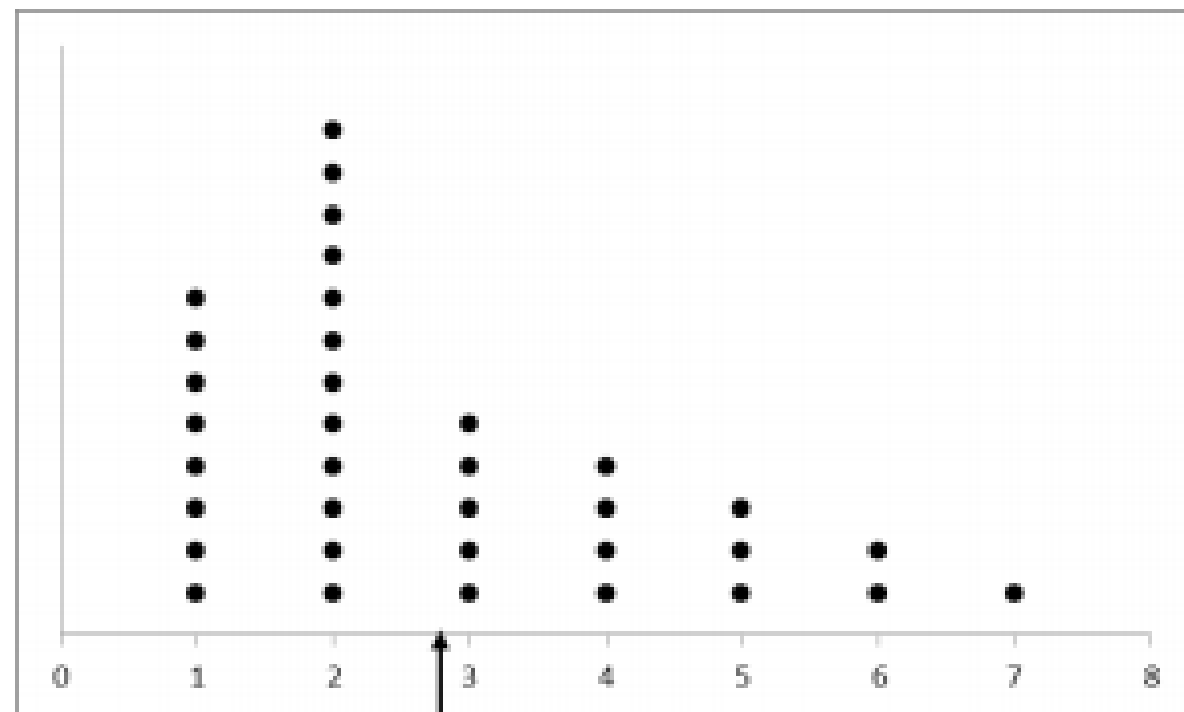
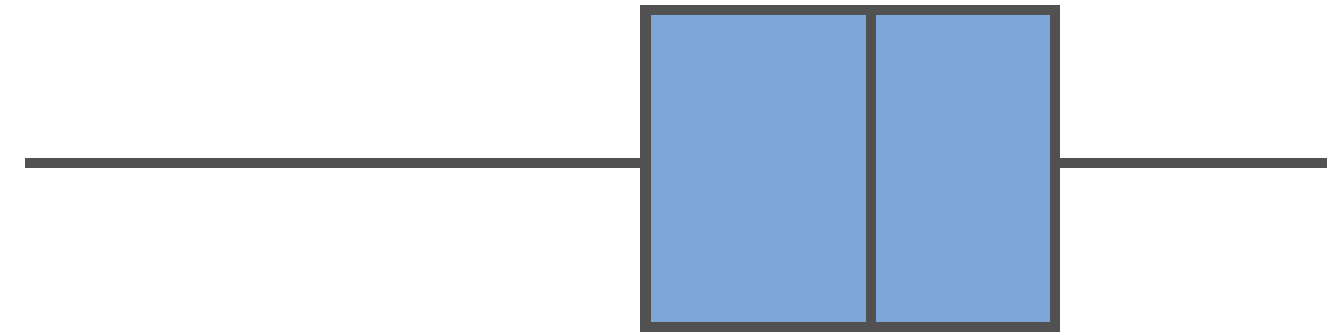


Figura 2.5 Distribuições com diferentes tipos de assimetria

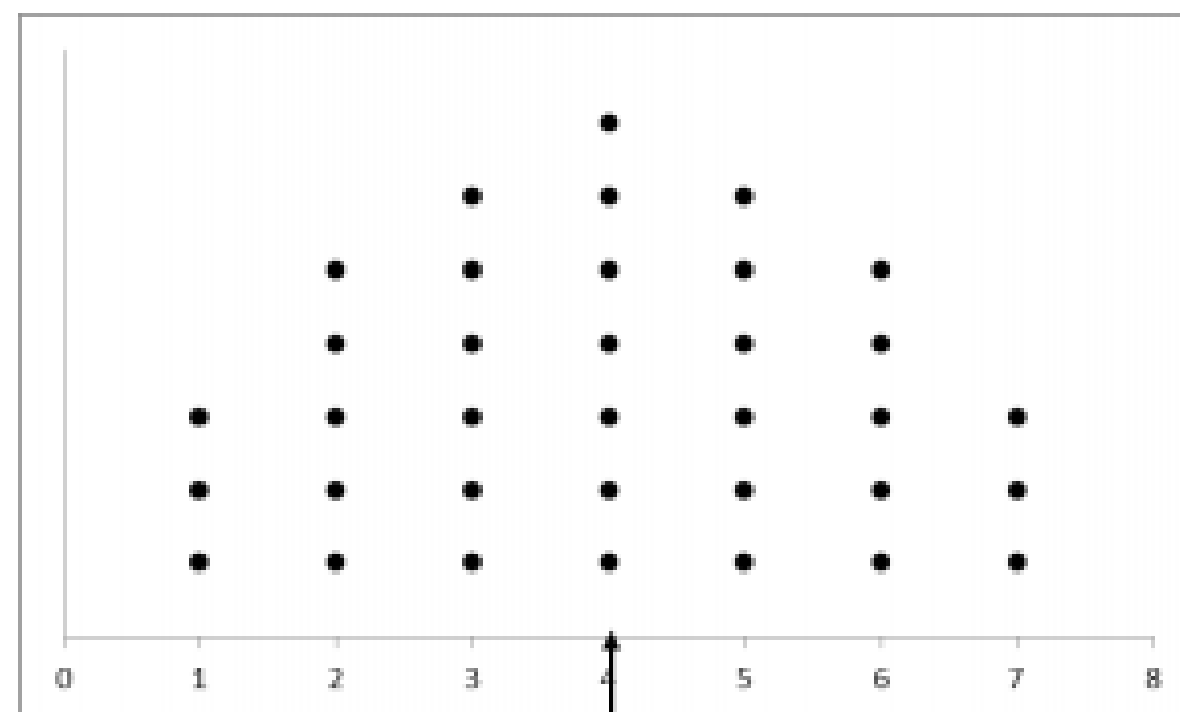
# Boxplot Assimétrico Positivo



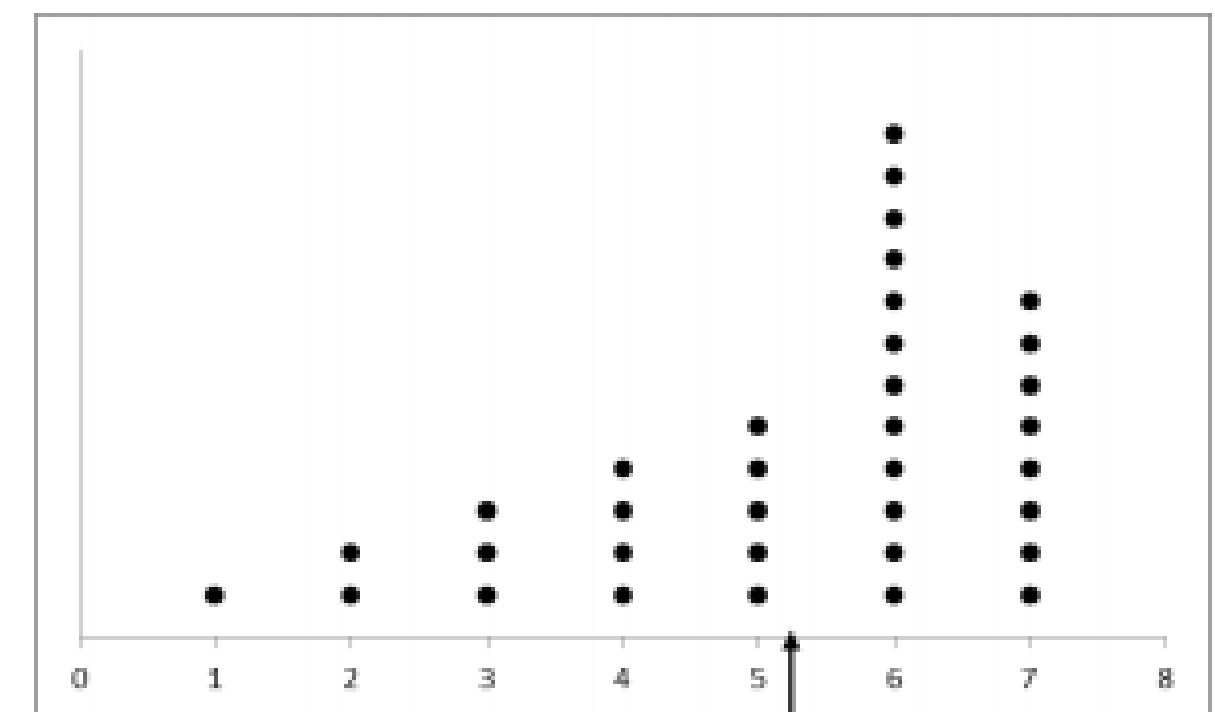
# Boxplot Assimétrico Negativo



Concentração valores baixos



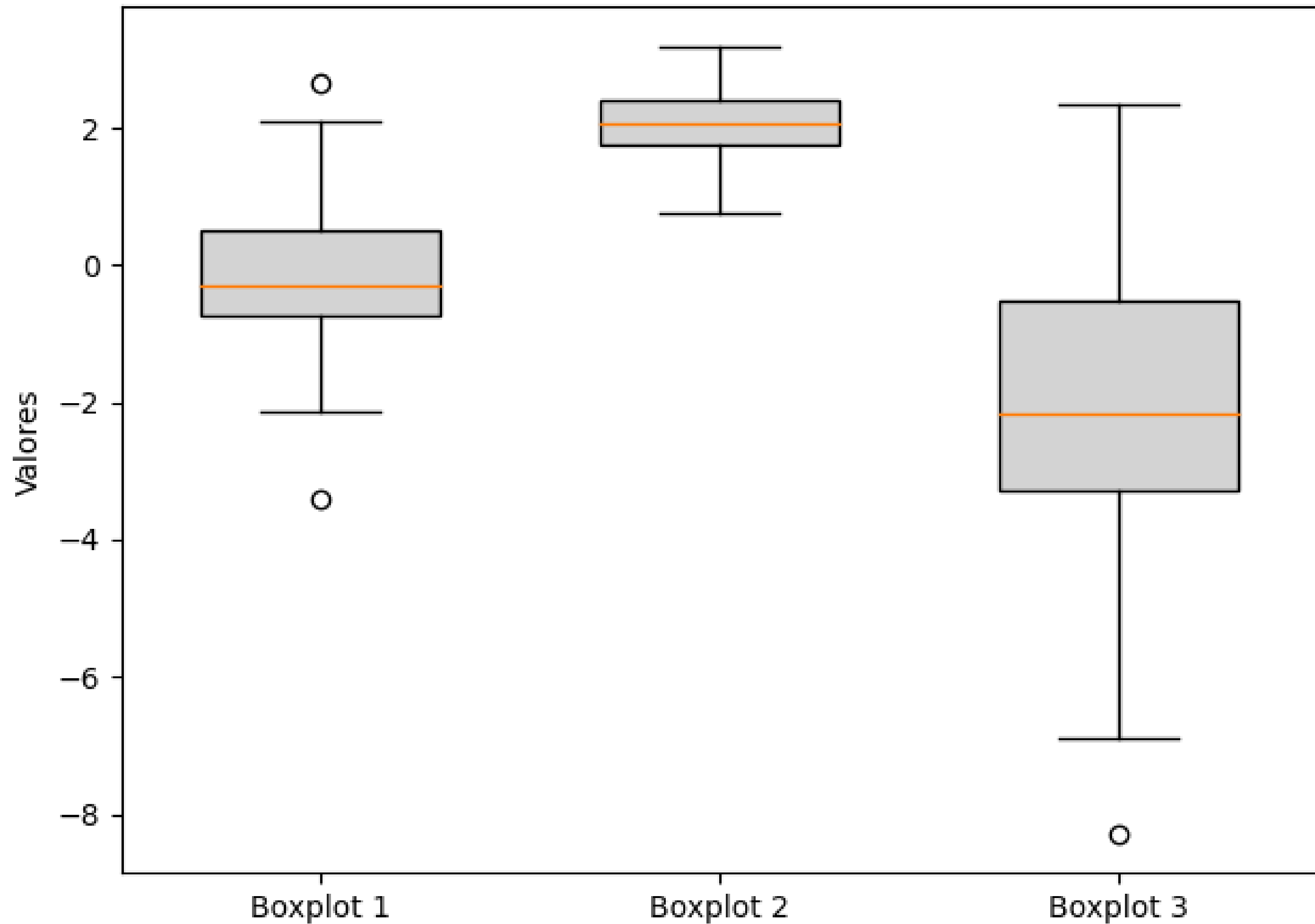
equilibrado



concentração valores altos

Figura 2.5 Distribuições com diferentes tipos de assimetria

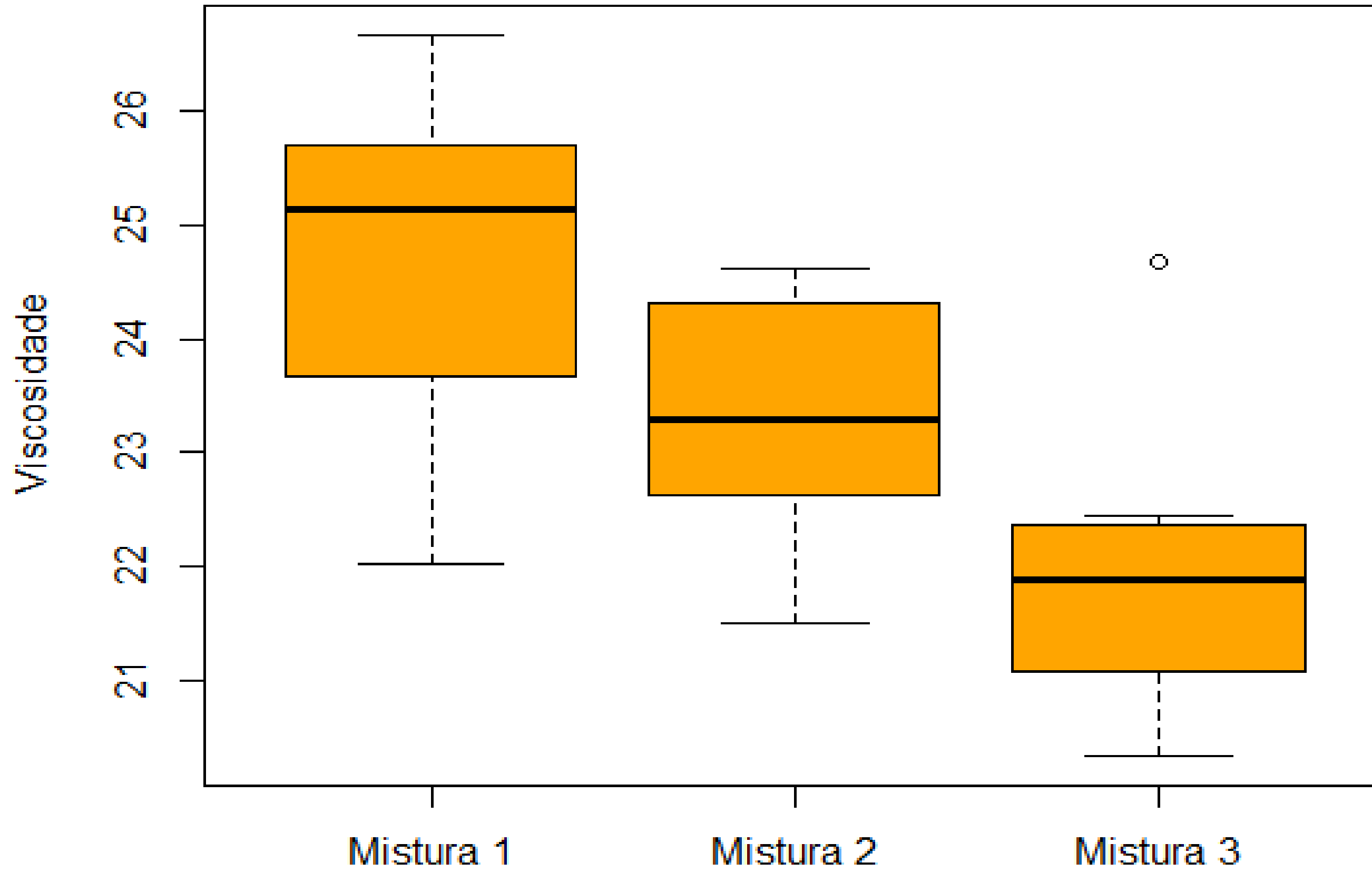
# Boxplots com Medianas e Quartis Diferentes



O conjunto de dados apresenta dados de viscosidade de três misturas diferentes

	Mistura 1	Mistura 2	Mistura 3
1	22.02	21.49	20.33
2	23.83	22.67	21.67
3	26.67	24.62	24.67
4	25.38	24.18	22.45
5	25.49	22.78	22.29
6	23.50	22.56	21.95
7	25.90	24.46	20.49
8	24.89	23.79	21.81

O Box Plot comparativo é apresentado a seguir:

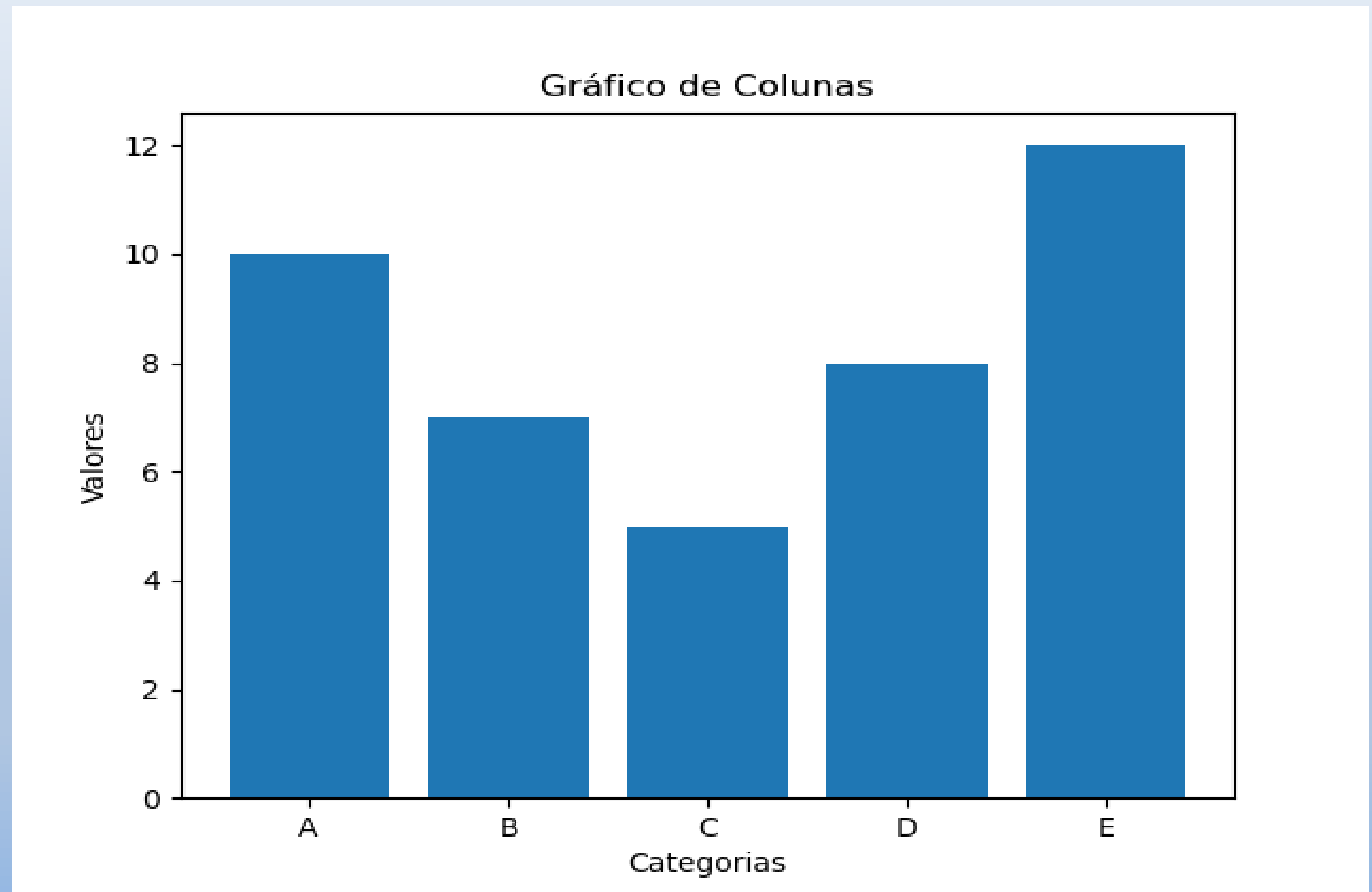


# Como Representar Dados Qualitativos?

- Gráficos de colunas
- Gráficos de barras
- Gráficos de pizza ou setorial

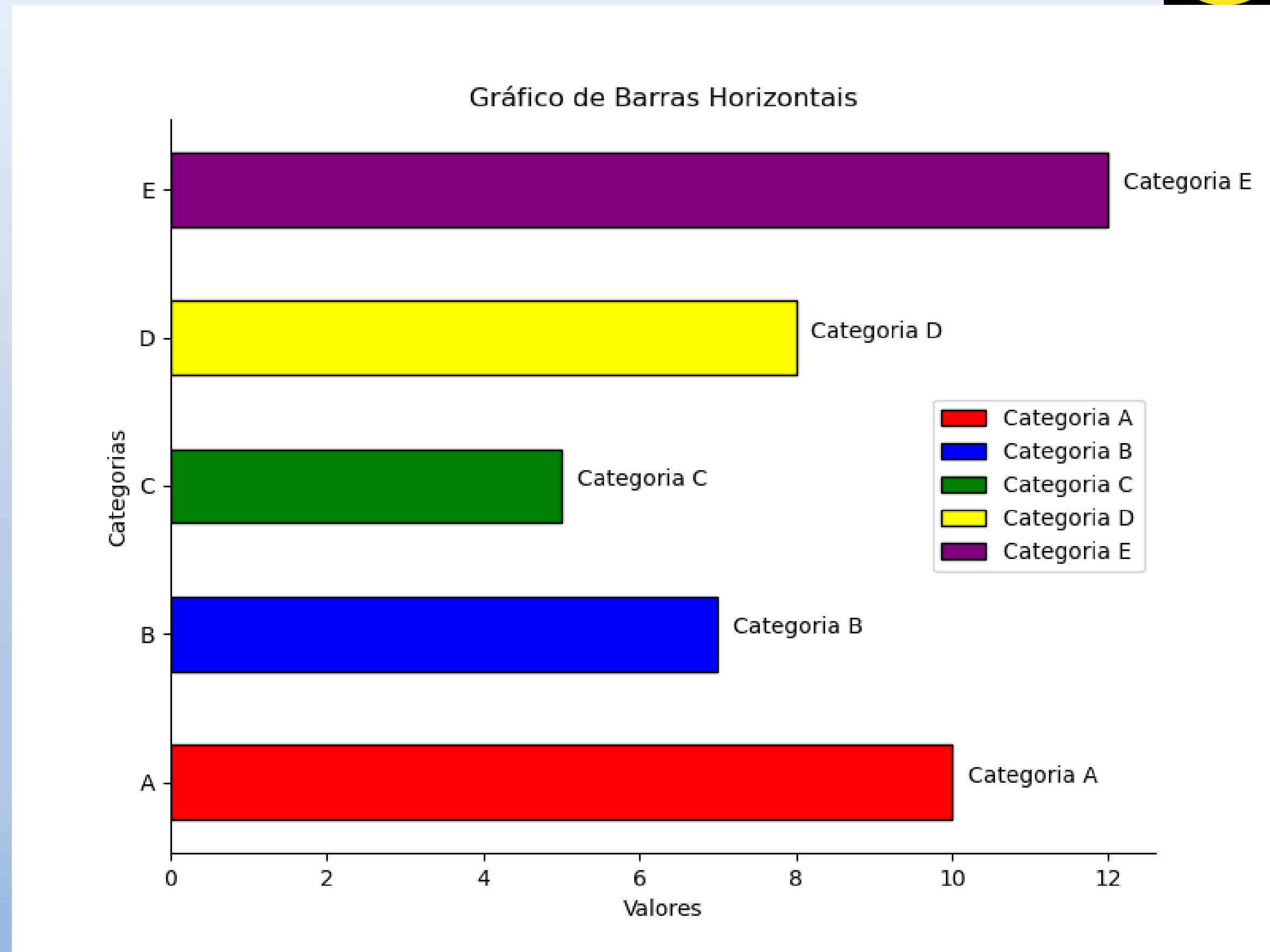
# Como Representar Dados Qualitativos?

- Gráficos de colunas



# Como Representar Dados Qualitativos?

- Gráficos de barras

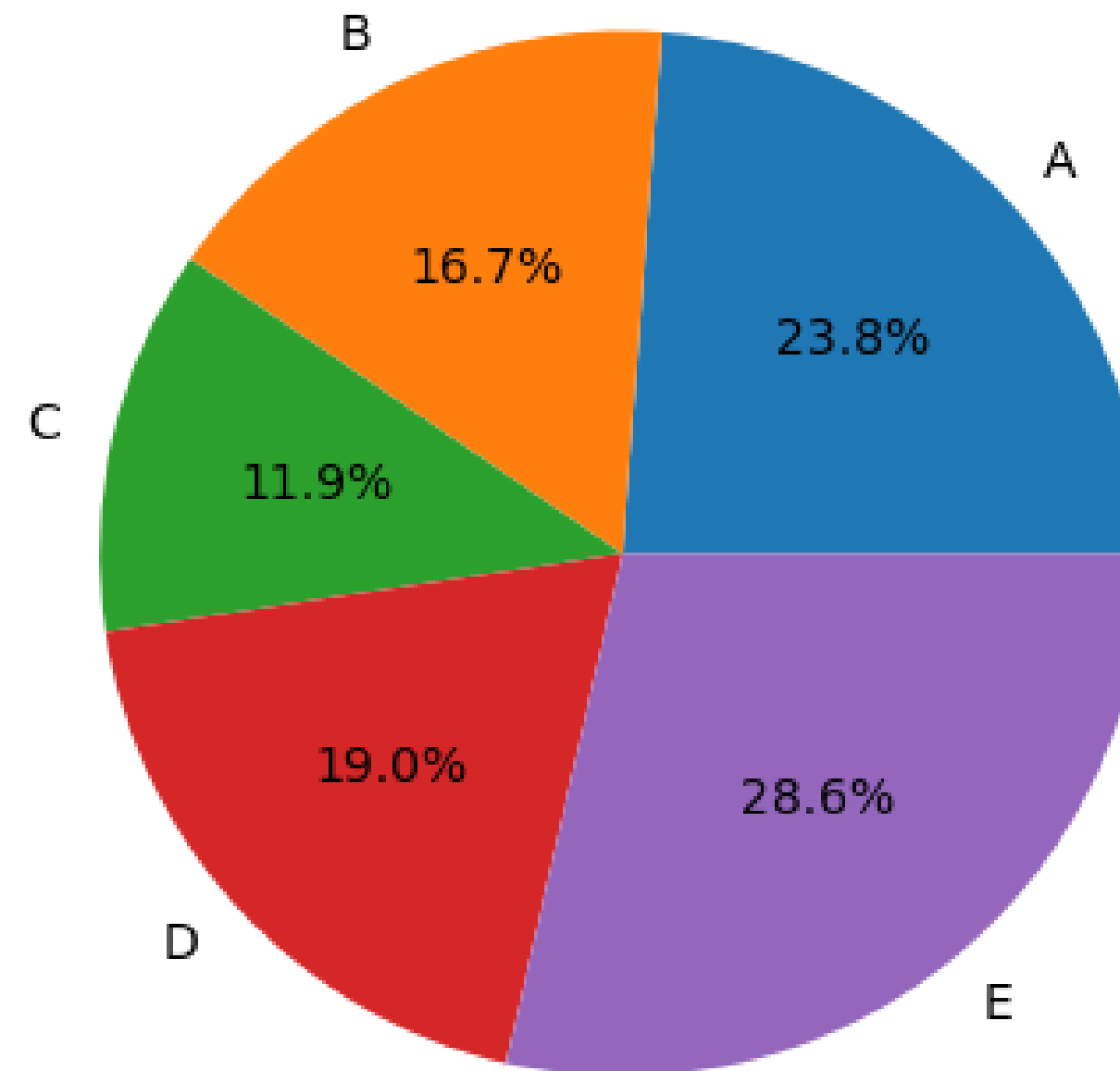




# Como Representar Dados Qualitativos?

- Gráficos de pizza  
ou setores

Gráfico de Setores

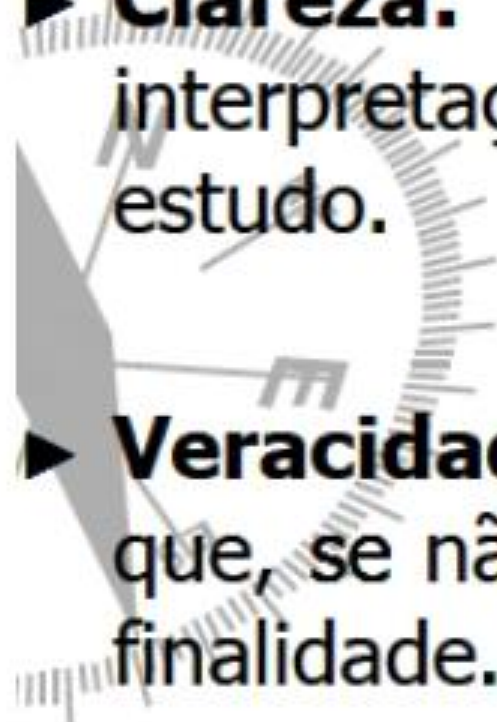


**Devemos sempre nos preocupar com a questão  
Da estética, da simplicidade e da clareza!!!!**

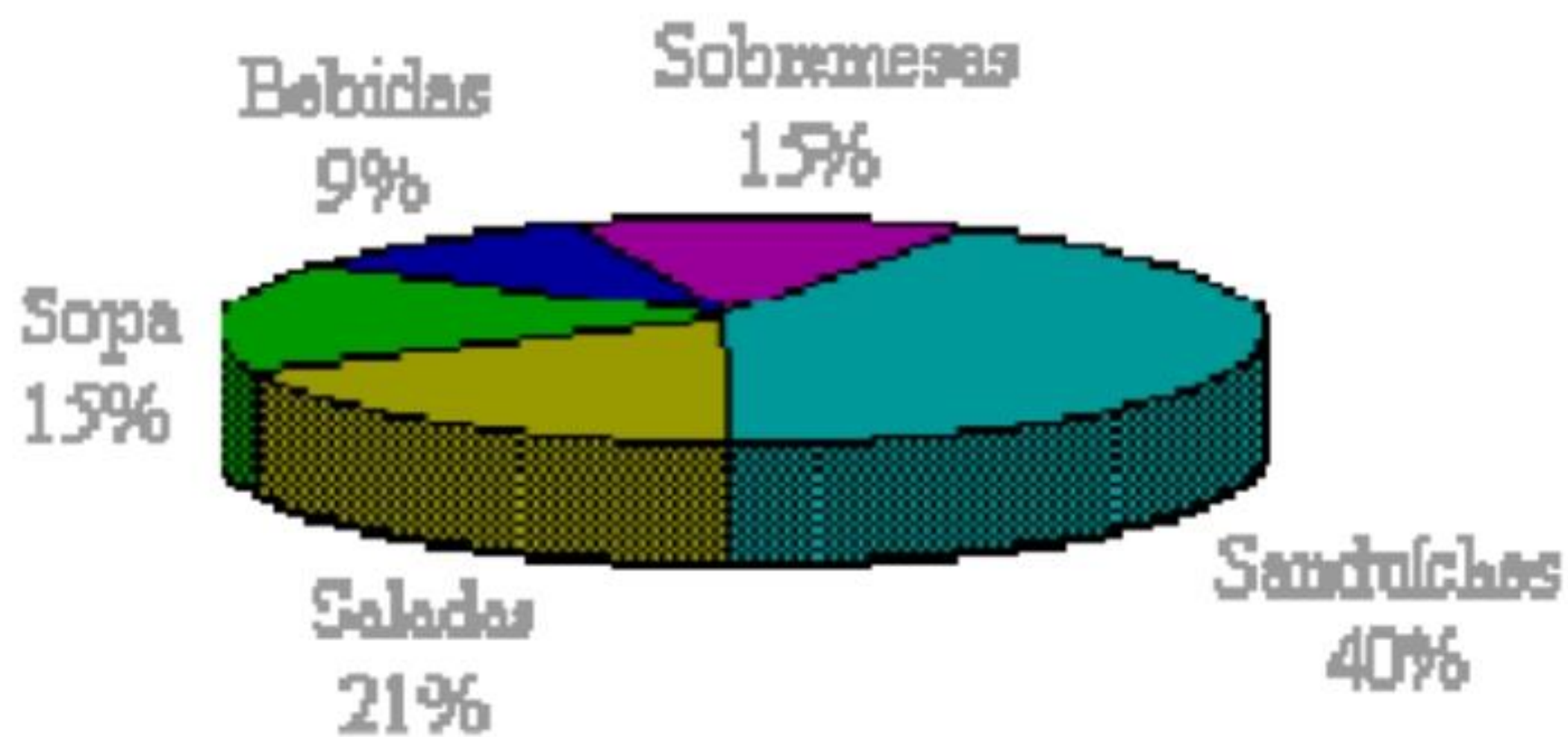
**Por que?**

A representação gráfica de um fenômeno deve obedecer aos seguintes requisitos primordiais:

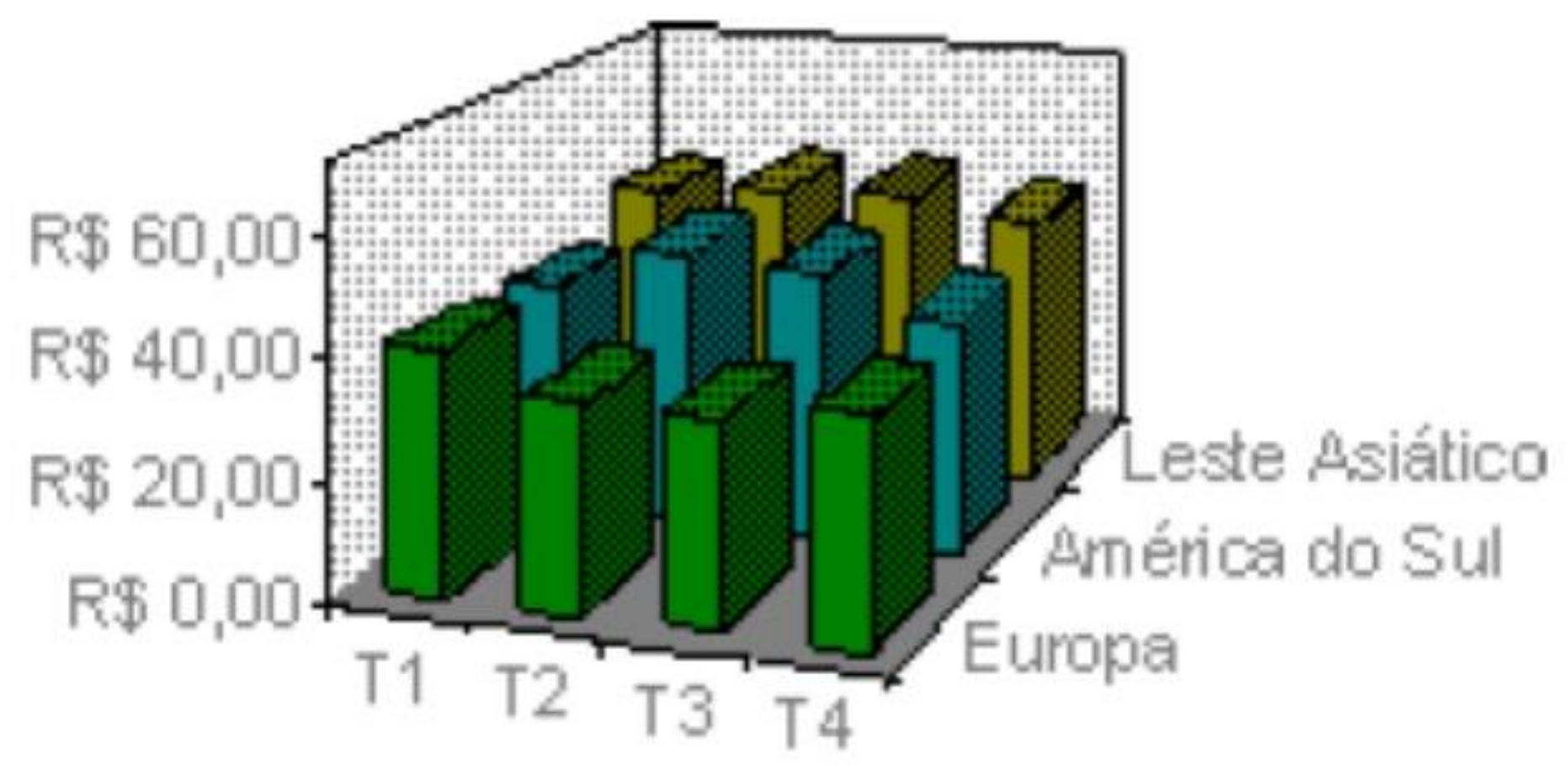
- ▶ **Simplicidade:** indispensável devido à necessidade de levar a uma rápida apreensão do sentido geral do fenômeno apresentado a fim de não nos perdermos na observação de minúcias de importância secundária.
- ▶ **Clareza:** o gráfico deve possibilitar uma correta interpretação dos valores representativos do fenômeno em estudo.
- ▶ **Veracidade:** indispensável qualquer comentário, posto que, se não representa uma realidade, o gráfico perde sua finalidade.



## Vendas do almoço

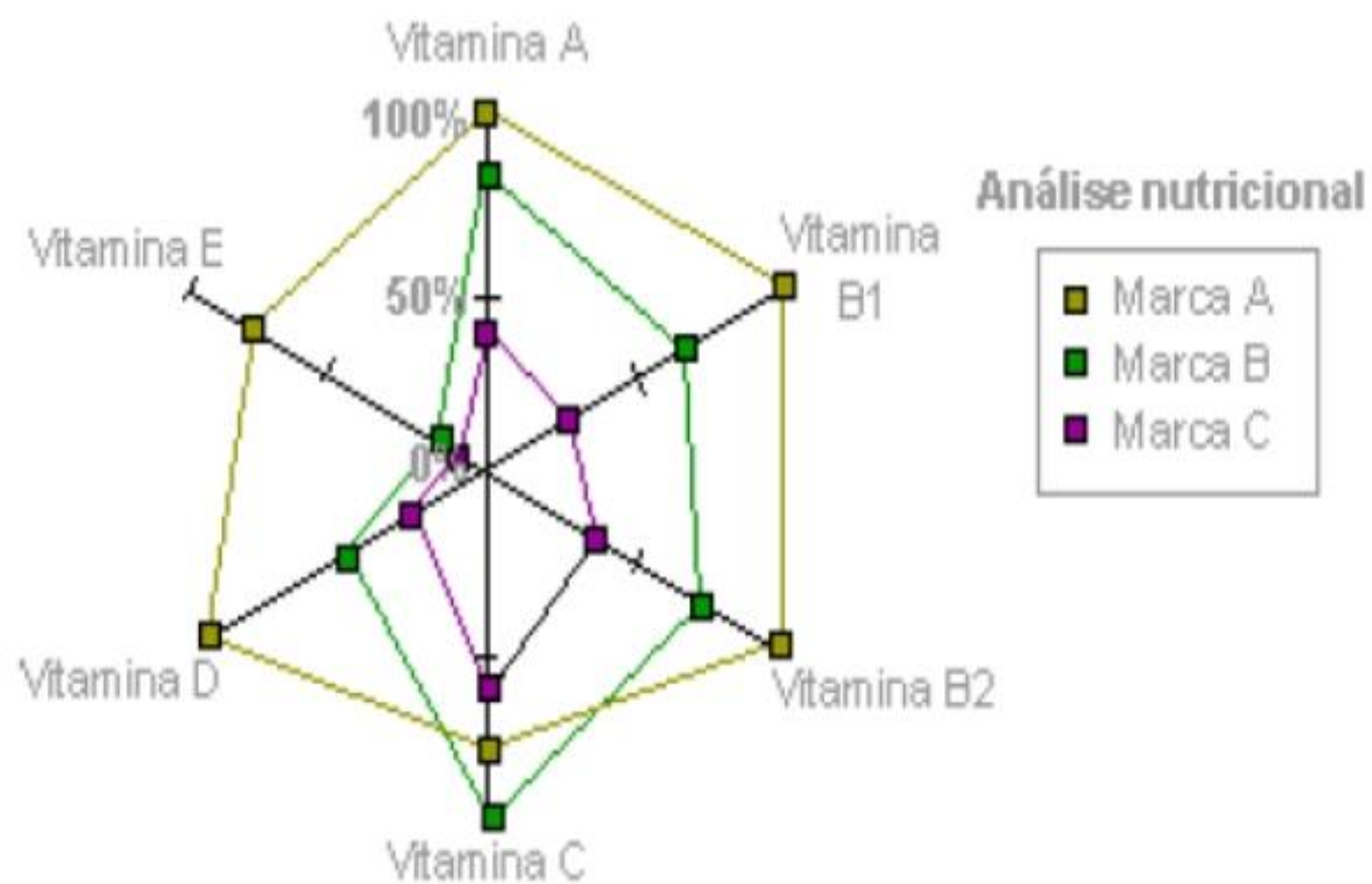






## Radar

Um gráfico de radar compara os valores agregados de várias séries de dados.



# Como Representar Dados Quantitativos?

- Gráficos de linha
- Gráficos de dispersão

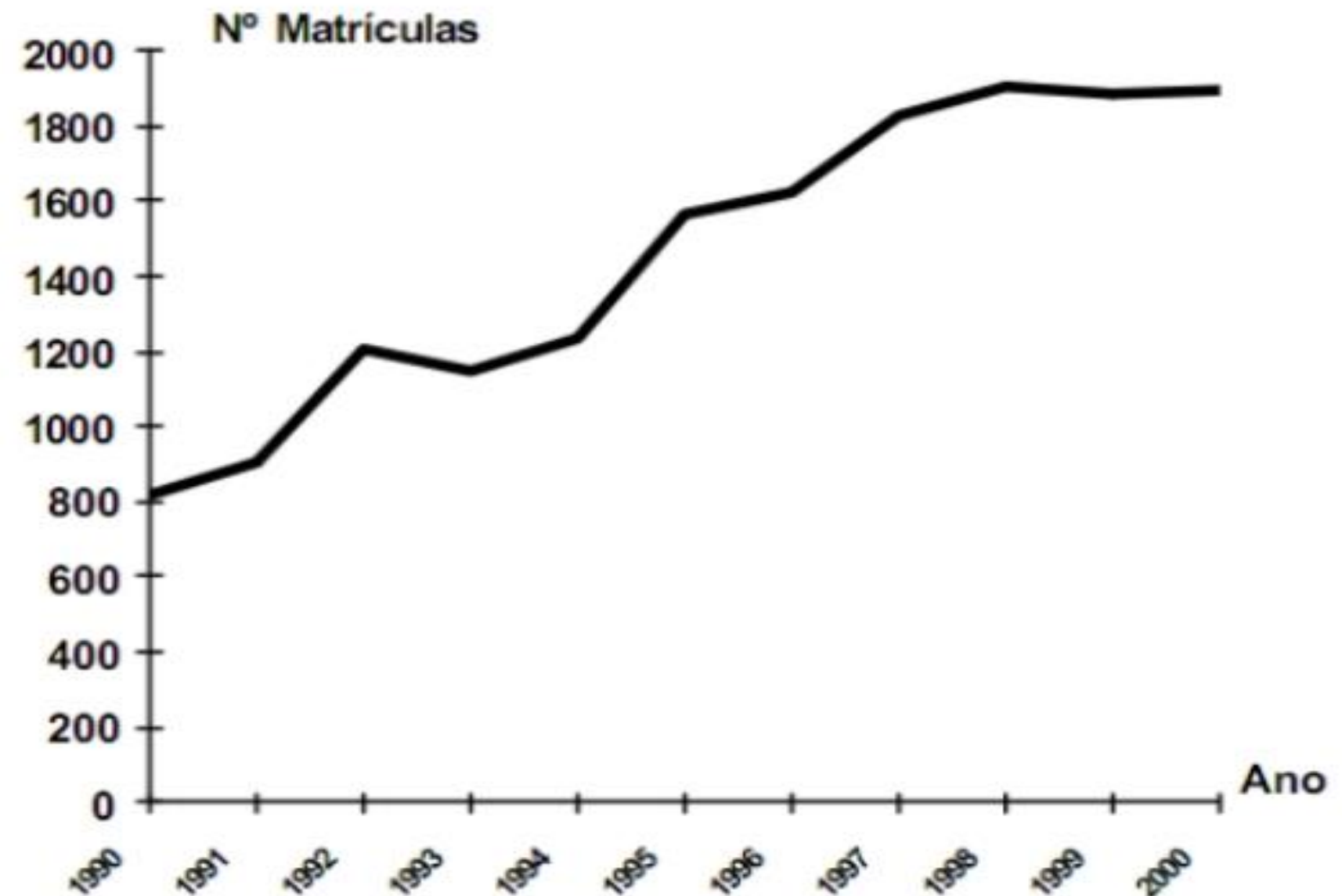
# Como Representar Dados Quantitativos?

## GRÁFICOS EM LINHA OU EM CURVA

MATRÍCULAS INICIAIS DO COLÉGIO NAVAL

RIO DE JANEIRO - BRASIL

PERÍODO	Nº Matrículas	ESC. 1 : 200
1990	816	4,1 cm
1991	904	4,5 cm
1992	1203	6,0 cm
1993	1147	5,7 cm
1994	1239	6,2 cm
1995	1565	7,8 cm
1996	1620	8,1 cm
1997	1833	9,2 cm
1998	1910	9,6 cm
1999	1890	9,4 cm
2000	1903	9,5 cm

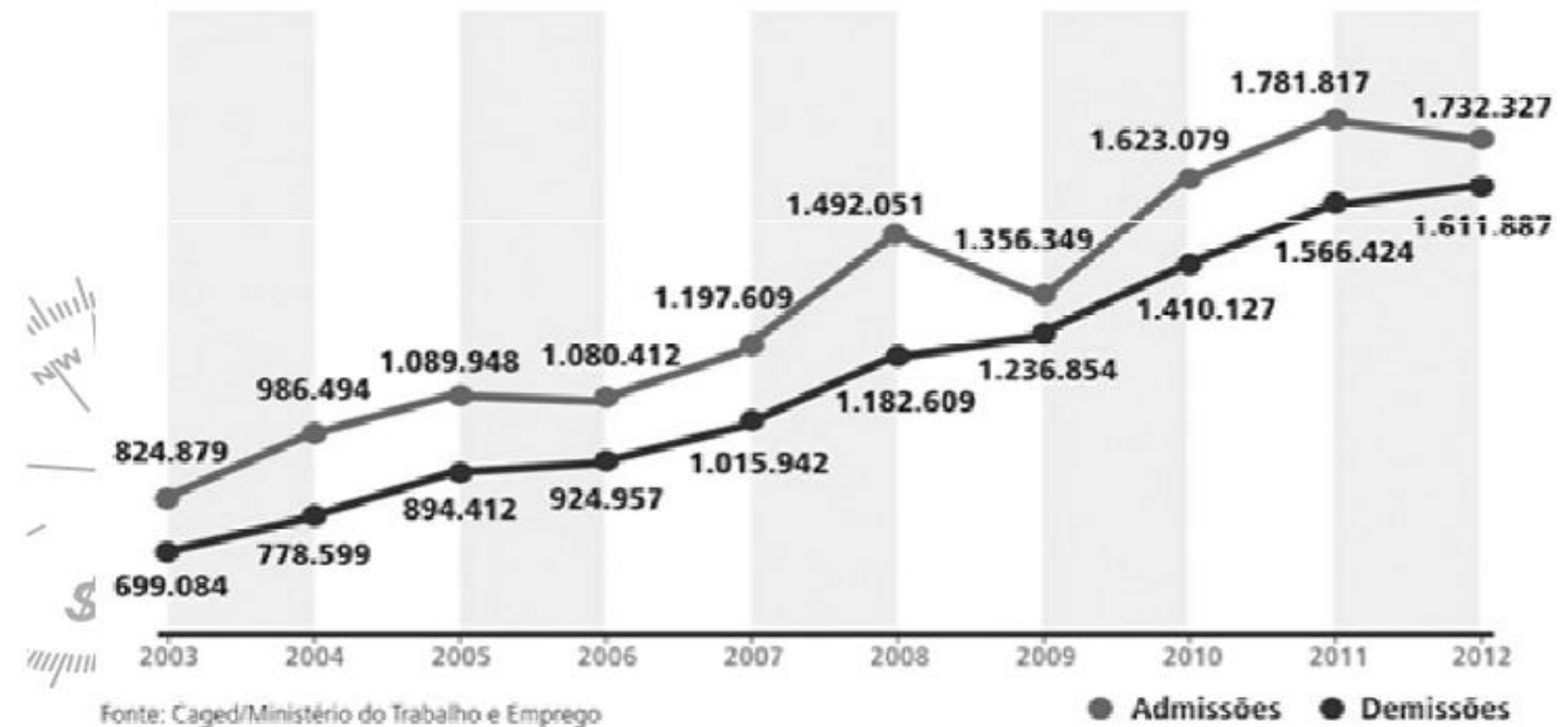




# Como Representar Dados Quantitativos?

- Gráficos de linha

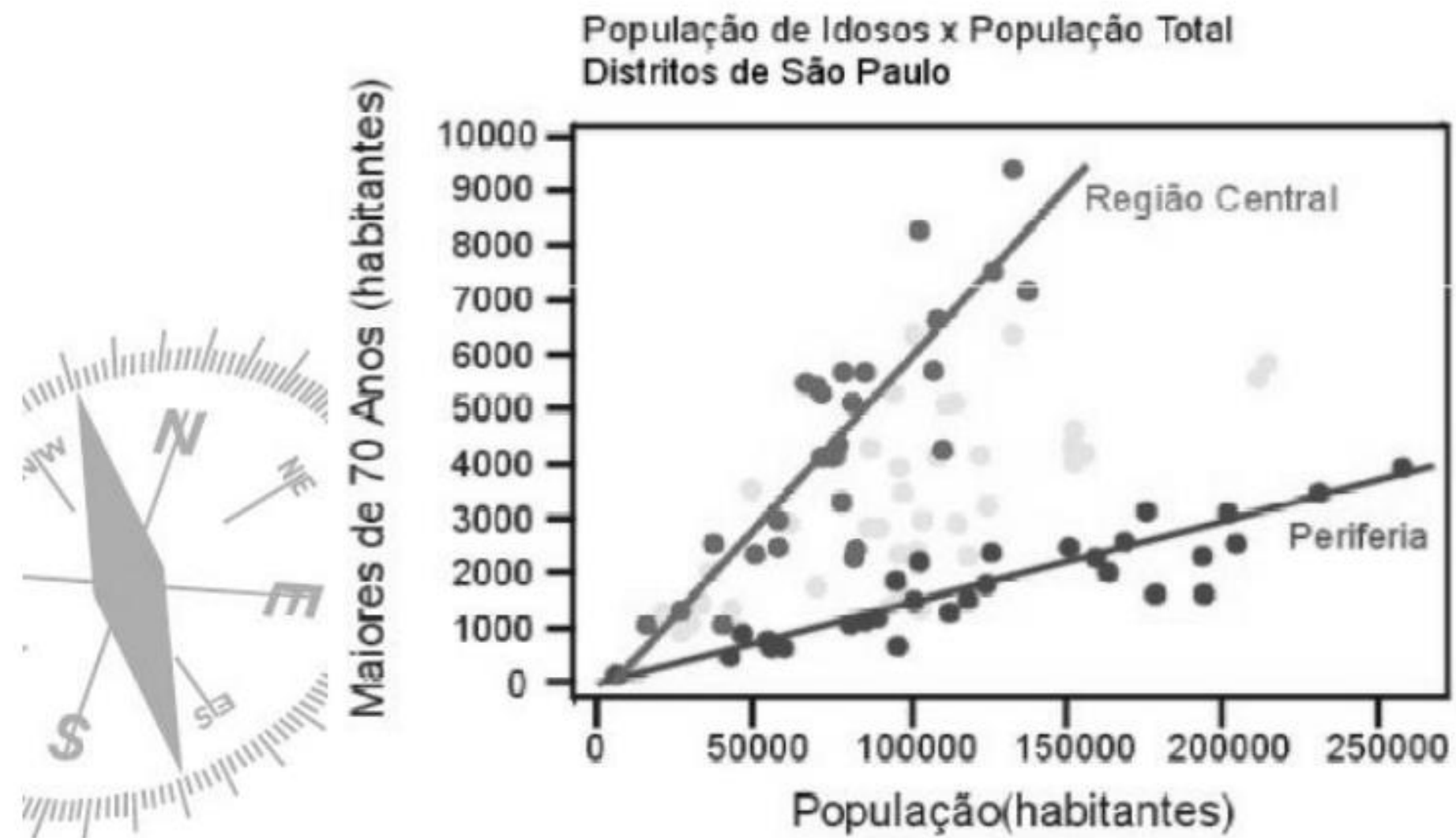
4) **Gráfico de Linha:** Utilizado usualmente para verificar o comportamento de uma determinada variável ao longo do tempo (série temporal). O eixo x representa anos, meses, semestres, entre outros. As linhas são mais eficientes neste tipo de gráfico porque permitem a detecção de flutuações ou mudanças intensas nas séries e também possibilitam a representação de várias séries no mesmo gráfico. Para construir um gráfico em linhas, basta marcar os pontos correspondentes às grandezas e uní-los através de segmentos de reta.



# Como Representar Dados Quantitativos?

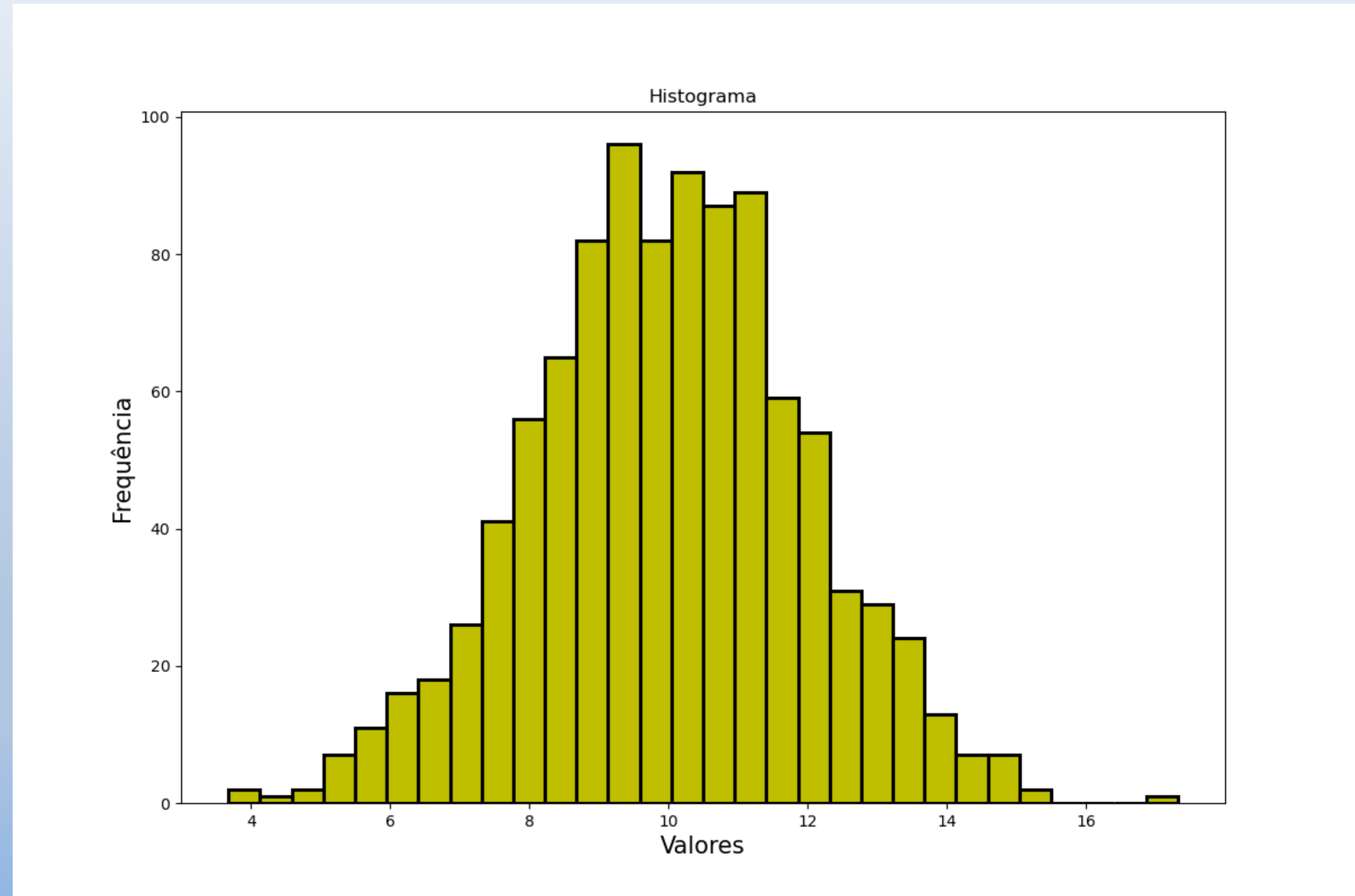
- Gráficos de dispersão

7) Diagrama de Dispersão: Utilizado para verificar possíveis relações entre variáveis quantitativas.



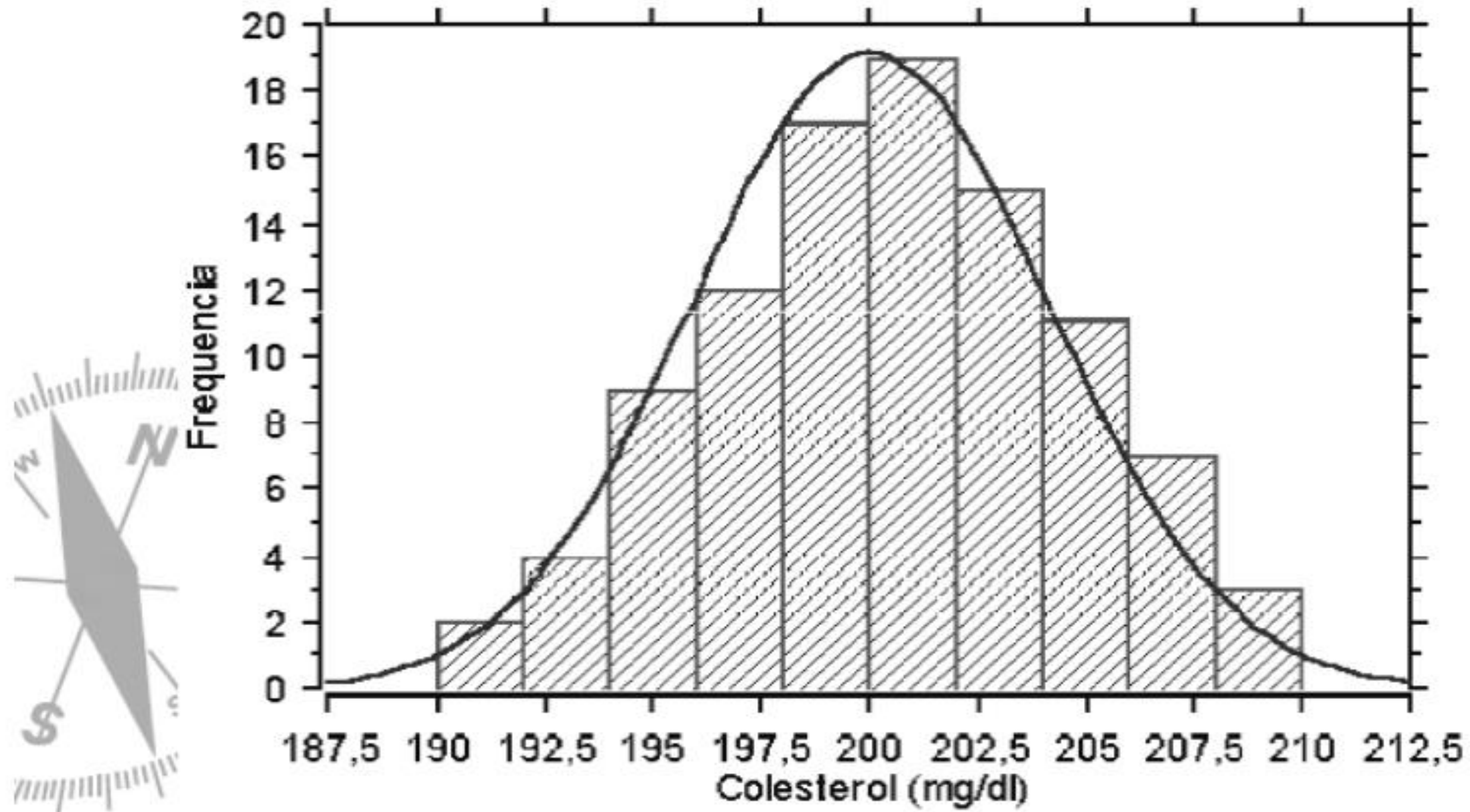
# Como Representar Dados Quantitativos?

- Histograma



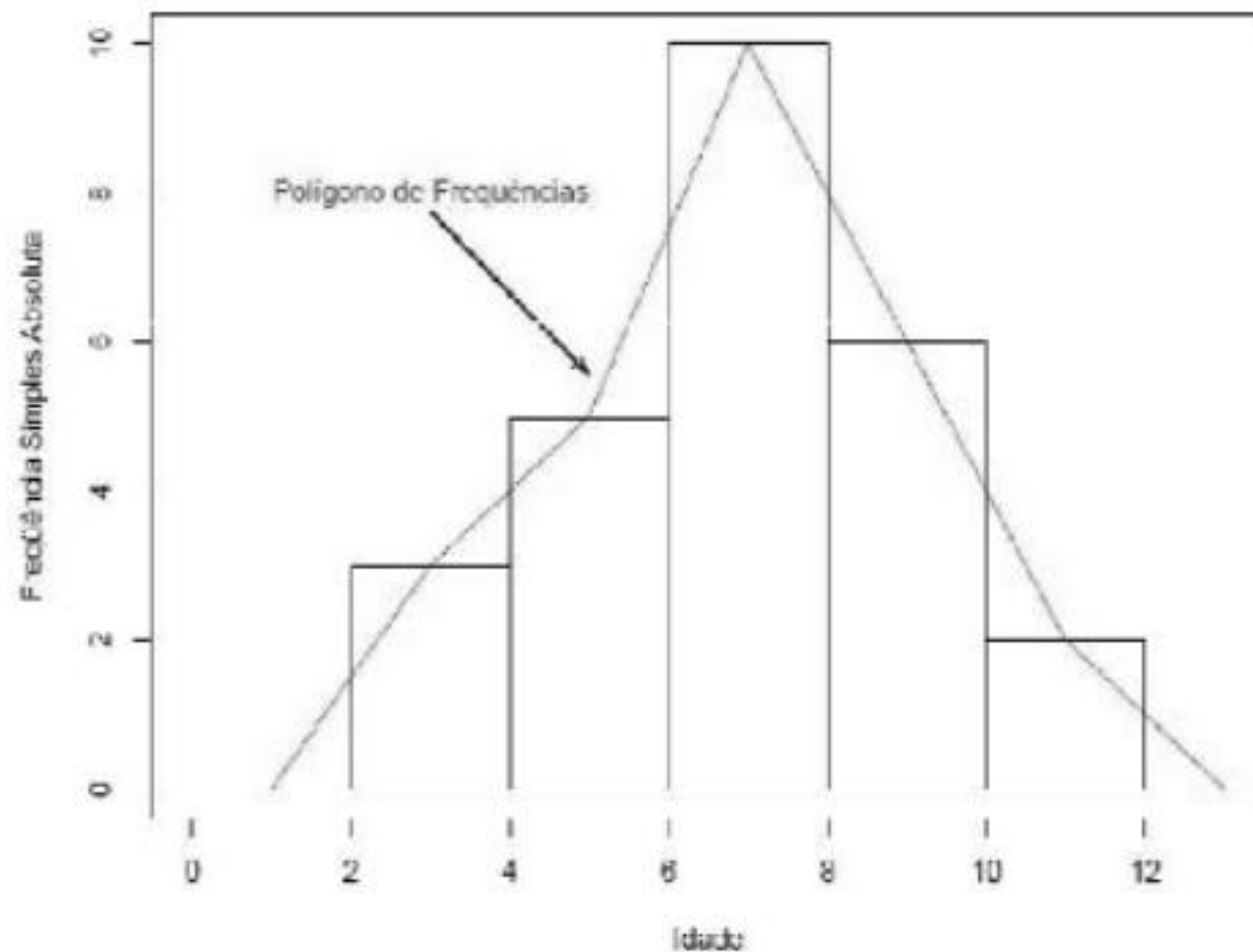


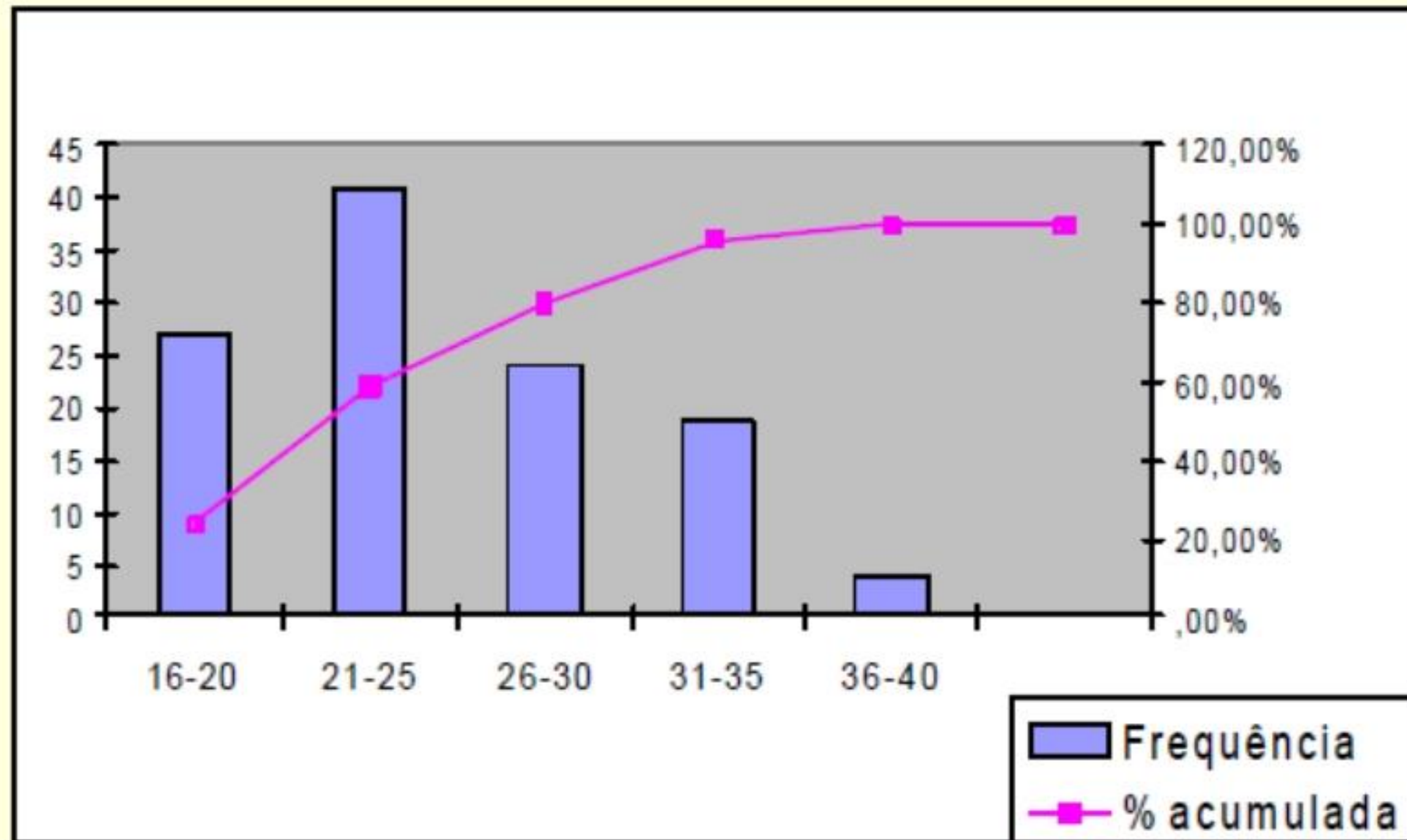
10) **Histograma:** É a representação gráfica de uma distribuição de frequência por meio de retângulos justapostos, cujas áreas são proporcionais às frequências das classes. Vale mencionar que, tanto as frequências absolutas simples quanto as relativas simples podem ser representadas através de histogramas.



Idade	$f_i$	$F_i^{ab}$	$F_i^{ac}$
2   4	3	3	26
4   6	5	8	23
6   8	10	18	18
8   10	6	24	8
10   12	2	26	2
$\Sigma$	26		

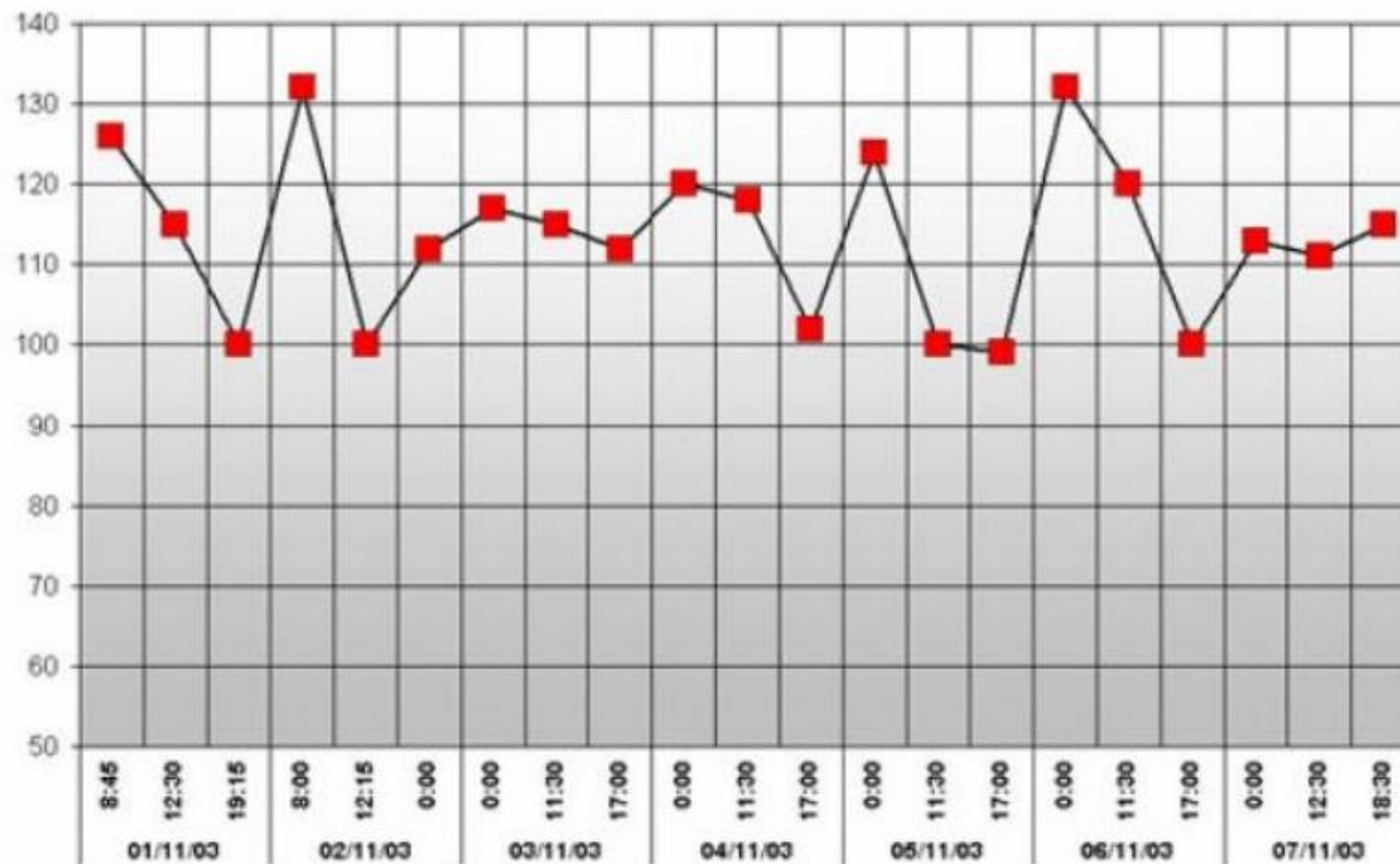
Histograma das Idades dos Indivíduos tratados



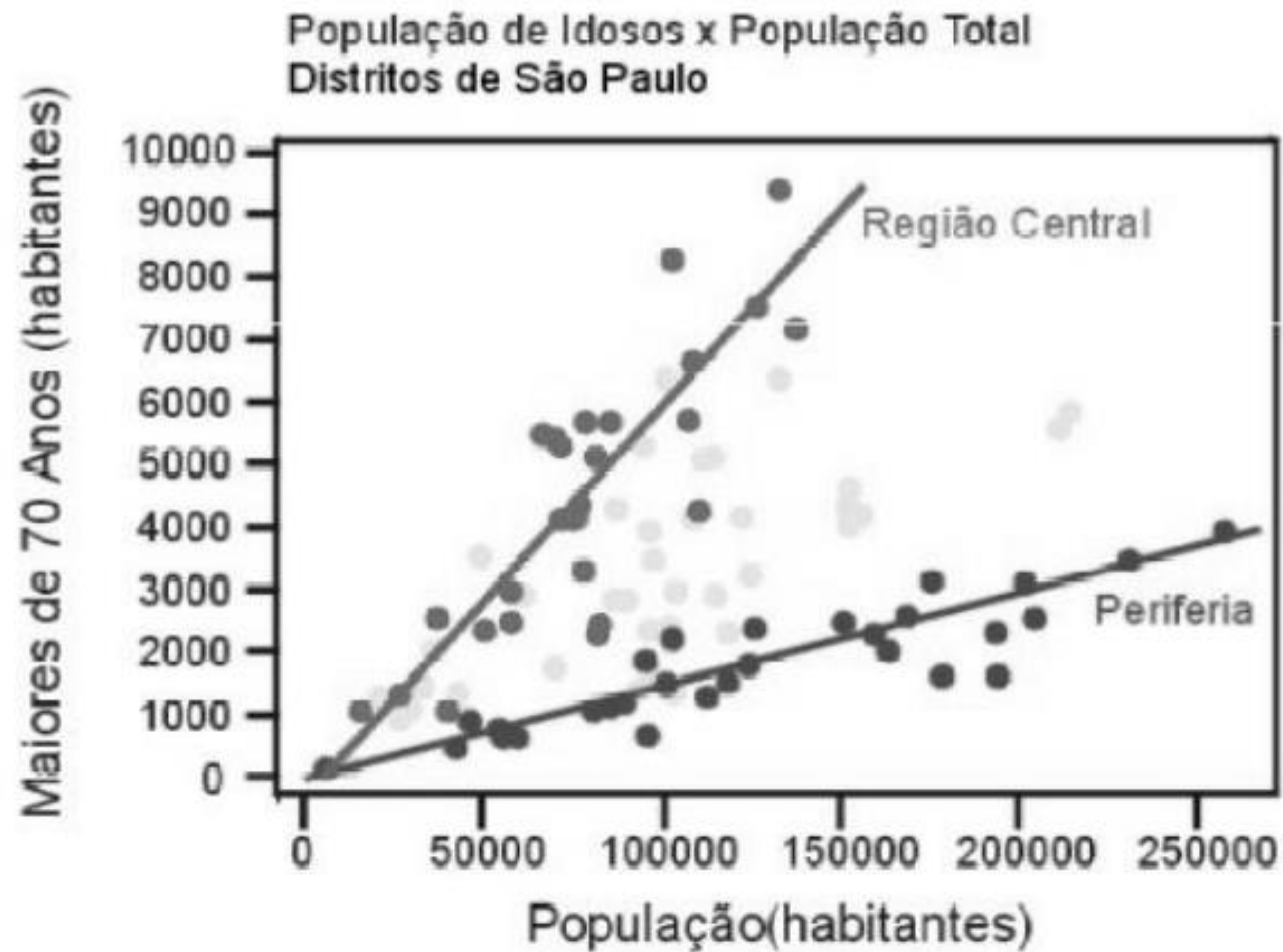




■ Nível de açúcar no sangue



7) Diagrama de Dispersão: Utilizado para verificar possíveis relações entre variáveis quantitativas.



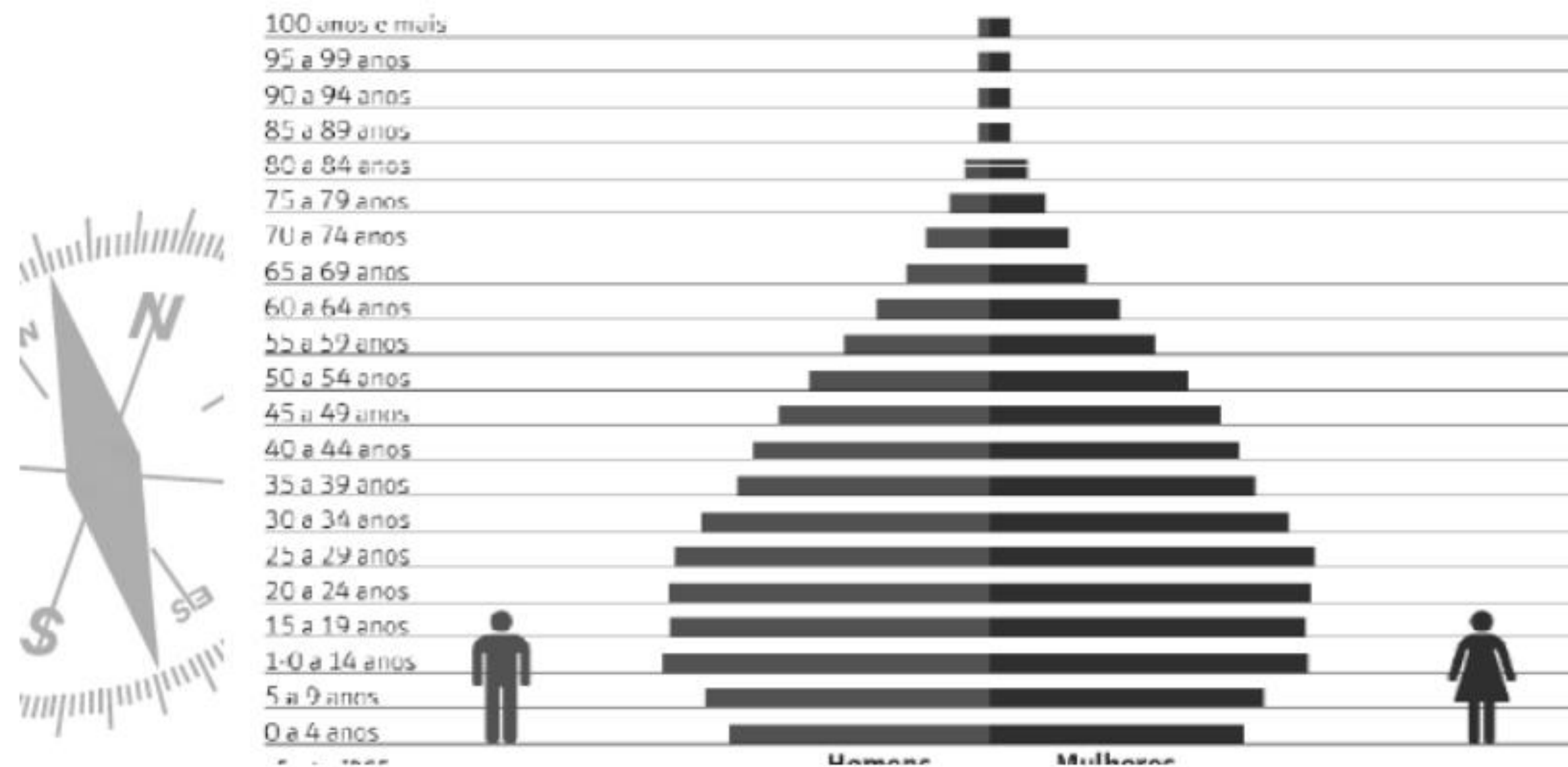


8) Pirâmide Etária: As pirâmides etárias permitem visualizar a forma como a população é distribuída conforme a variável faixa etária.

### ENVELHECIMENTO DA POPULAÇÃO BRASILEIRA

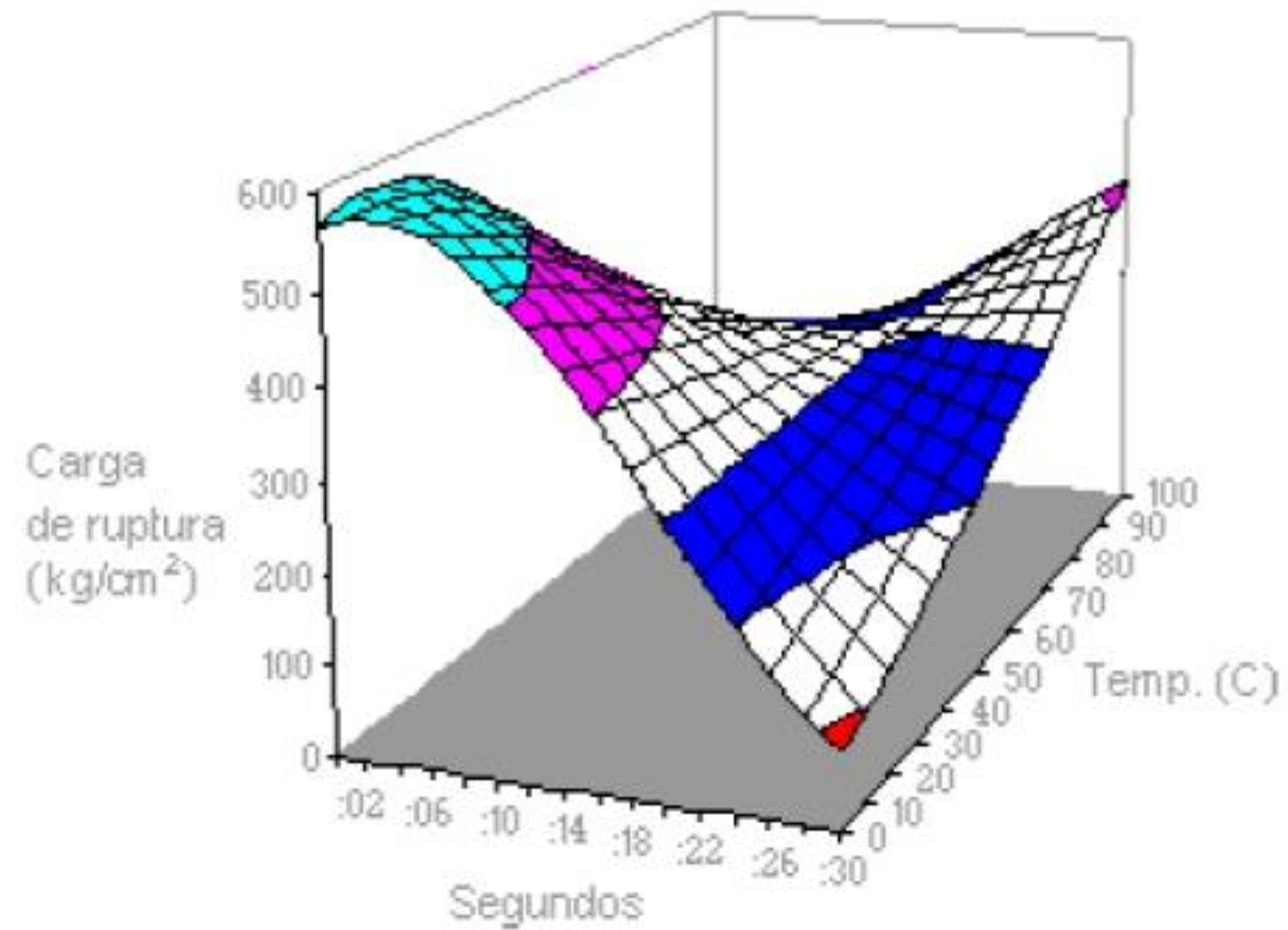
Em no máximo 40 anos, a pirâmide etária brasileira será semelhante à da França hoje

#### PIRÂMIDE ETÁRIA BRASIL 2010



## Superfície

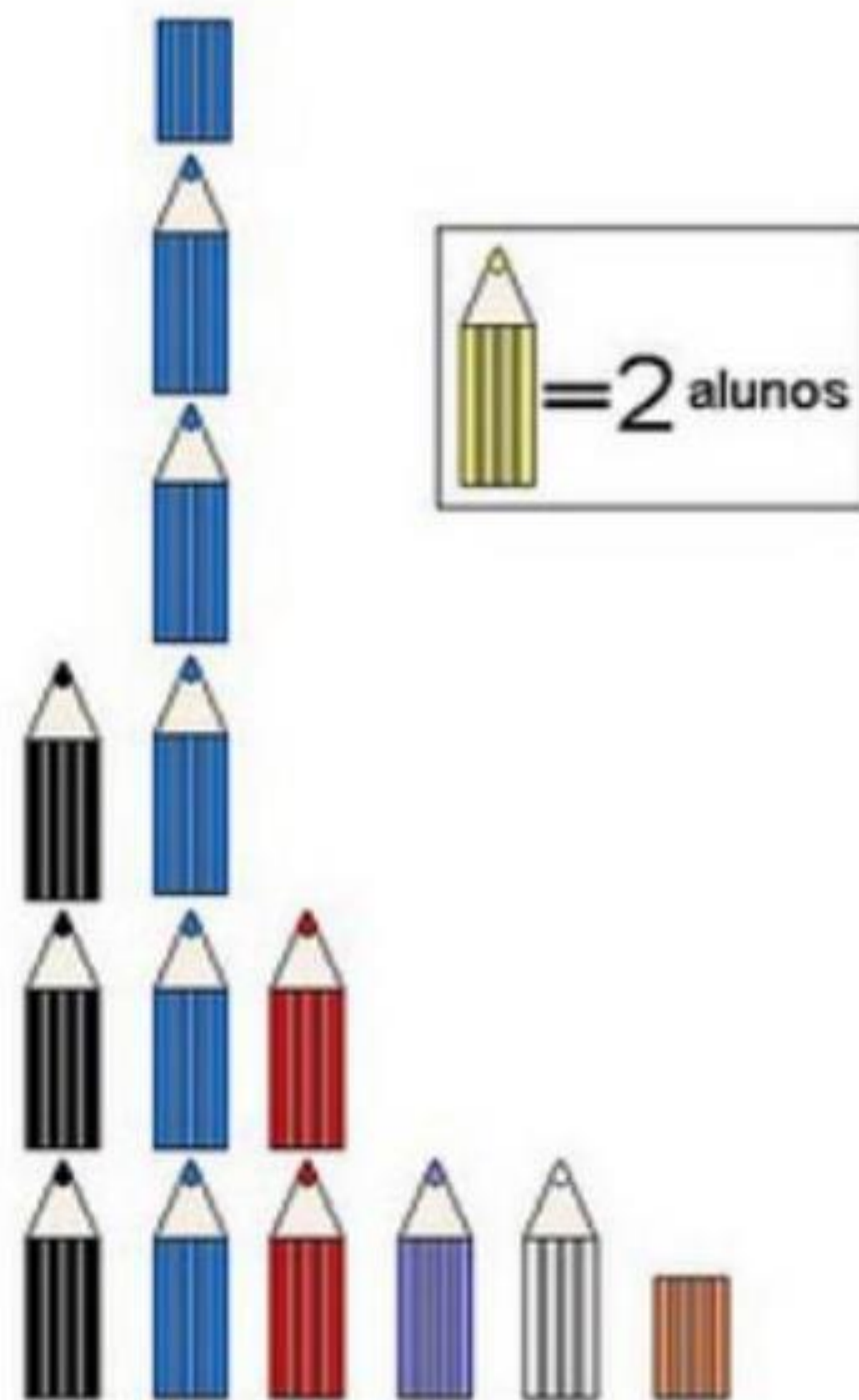
Um gráfico de superfície é útil quando você deseja localizar combinações vantajosas entre dois conjuntos de dados. Como em um mapa topográfico, as cores e os padrões indicam áreas que estão no mesmo intervalo de valores.



# PICTOGRAMA

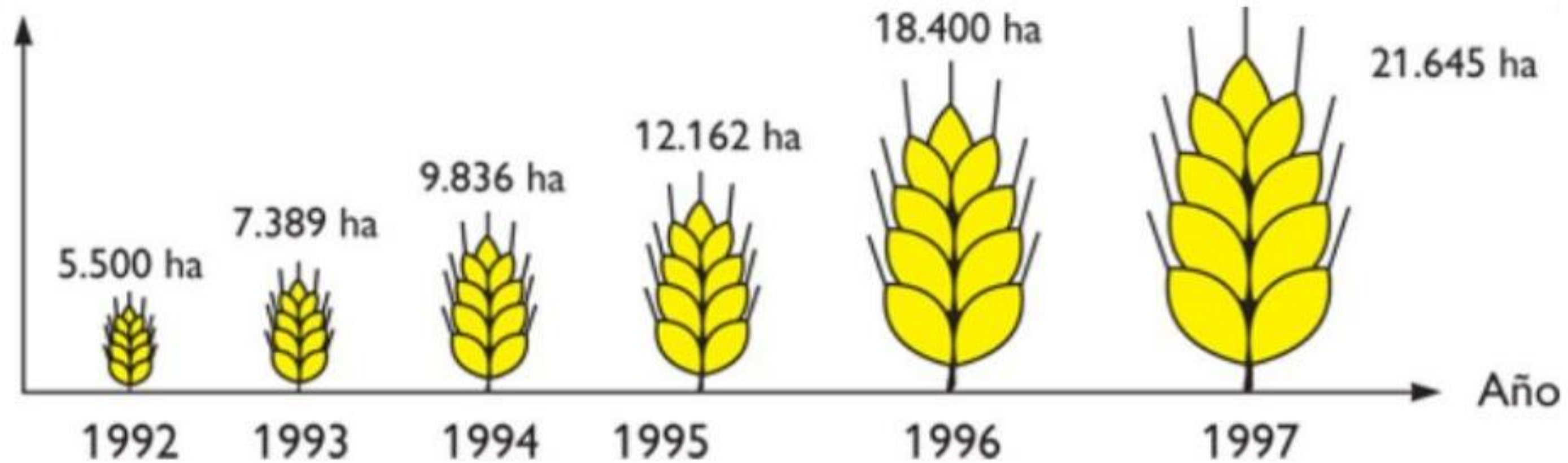
A cor que os alunos preferem em um determinado Jardim de Infância.

Quantos preferem a cor azul?





# PICTOGRAMAS

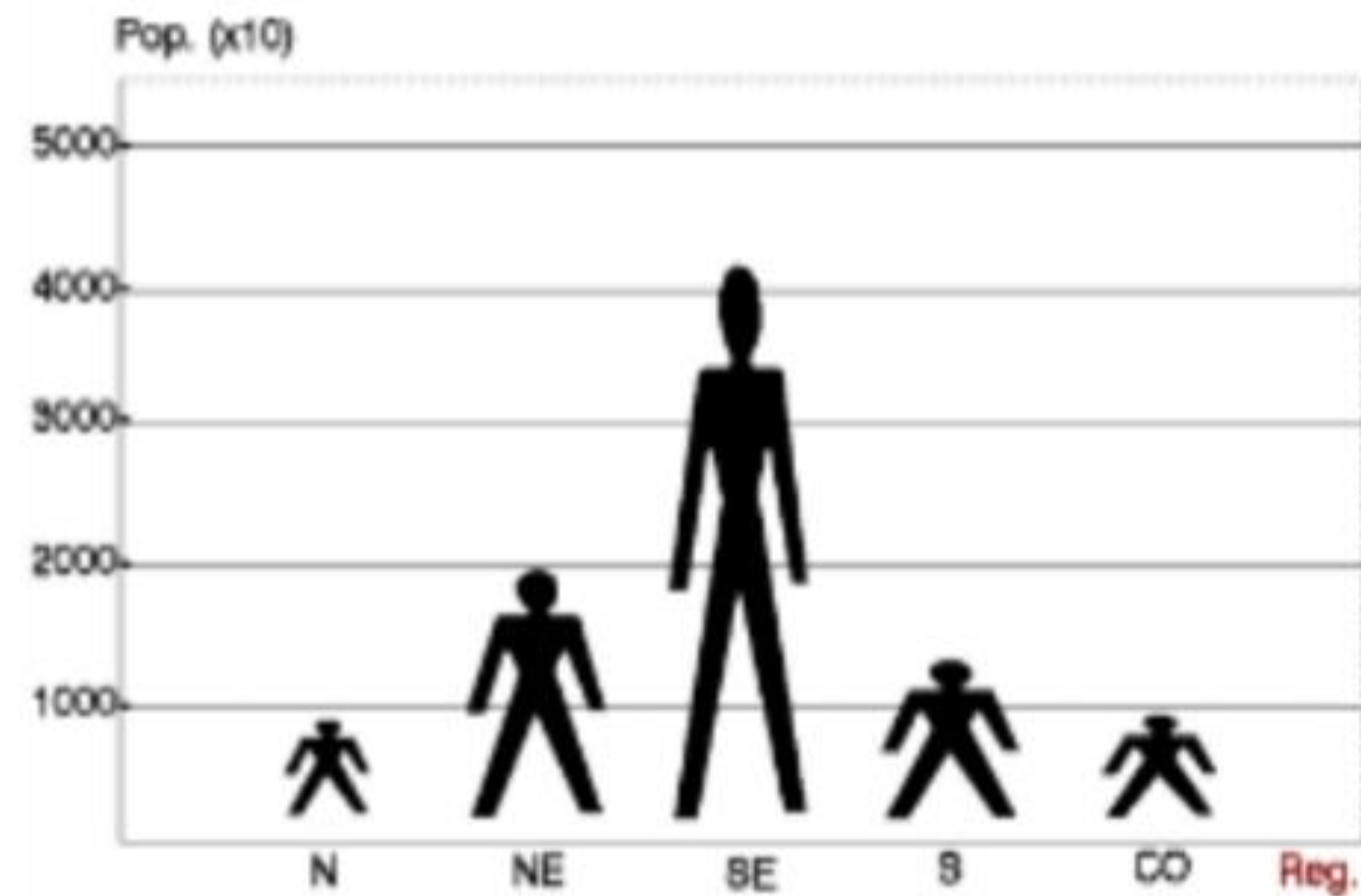


Qual foi o aumento de produção entre os anos de 1996 e 1997?

# PICTOGRAMAS

Processo gráfico no qual constam figuras.

População Urbana no Brasil em 1980 (x10)



Fonte: Anuário Estatístico (1984)

# PICTOGRAMAS

 = 32 mil hectares de floresta ardida

Ano 2003



Ano 2004



Ano 2005



Ano 2006



Ano 2007

