



By @kakashi_copiador

APRESENTAÇÃO DO MATERIAL

Queridos alunos!!

Sabemos que os **resumos** das disciplinas **são fundamentais para fixação de conteúdos** e, também, para **realização de revisões**. Um resumo bem feito garante que os principais pontos de cada matéria sejam revisados de forma rápida, **aumentando a produtividade dos estudos e a eficiência das revisões**.

Além disso, sabemos que, principalmente para os grandes concursos, o número de matérias cobradas no edital é muito grande. Dessa forma, além de revisar os pontos marcados em seus materiais, um bom resumo pode encurtar o tempo de revisão, garantindo, assim, que todo o material possa ser revisado em um período de tempo mais curto.

Com isso em mente, apresentamos a vocês o **Resumo de Estatística - Apresentação de Dados**. Trata-se de um material pensado para lhe ajudar em todo esse processo, visando, inclusive, uma economia de tempo de confecção de materiais, tempo que é o bem mais precioso de um concurseiro, não é mesmo?

Esperamos poder ajudá-los!

Conte sempre com o Estratégia em sua caminhada!

Estratégia Concursos



Esse é um material resumido. Em momento algum ele substitui o estudo do material completo. Trata-se de um complemento aos estudos e um facilitador de revisões!

RESUMO DE ESTATÍSTICA

Introdução à Estatística

A Estatística pode ser dividida em três grandes ramos: **Estatística Descritiva** (ou dedutiva), **Estatística Probabilística** e **Estatística Inferencial** (ou indutiva).

ESTATÍSTICA DESCRITIVA

É responsável pela coleta, organização, descrição e resumo dos dados observados.

ESTATÍSTICA PROBABILÍSTICA

É responsável por estabelecer o modelo matemático adotado para explicar fenômenos aleatórios.

ESTATÍSTICA INFERENCIAL

É responsável pela análise e interpretação dos dados.

Conceitos Iniciais

POPULAÇÃO: **CONJUNTO** de **TODOS** os elementos a serem estudados, que apresentam uma ou mais características em comum.

CENSO: **ESTUDO** dos dados relativos a **TODOS** os **ELEMENTOS** de uma população.

AMOSTRA: **SUBCONJUNTO** extraído **DA POPULAÇÃO** para análise, devendo ser representativo daquele grupo.

AMOSTRAGEM: **PROCESSO** que consiste na seleção criteriosa dos elementos a serem submetidos à investigação.

PARÂMETROS: Descrições numéricas de características da **POPULAÇÃO**, que normalmente precisam ser estimadas.

ESTATÍSTICAS: Medidas numéricas extraídas de **AMOSTRAS** representativas extraídas da população.

Método Experimental X Método Estatístico

MÉTODO EXPERIMENTAL	MÉTODO ESTATÍSTICO
As CAUSAS são mantidas CONSTANTES , COM EXCEÇÃO DE UMA , que é VARIADA para que seus efeitos sejam descobertos.	Admite e REGISTRA TODAS AS POSSÍVEIS VARIAÇÕES DAS CAUSAS PRESENTES , procurando determinar a influência de cada fator no resultado.

Dados Estatísticos

Com relação ao número de observações coletadas, os dados são classificados em **univariados, bivariados e multivariados**:

DADOS UNIVARIADOS	DADOS BIVARIADOS	DADOS MULTIVALORADOS
É quando uma única observação de cada indivíduo é registrada.	É quando duas observações de cada indivíduo são registradas.	É quando mais de duas observações acerca de cada indivíduo são registradas.

Quanto à forma de apresentação, os dados podem ser classificados em dados brutos ou rol.

DADOS BRUTOS

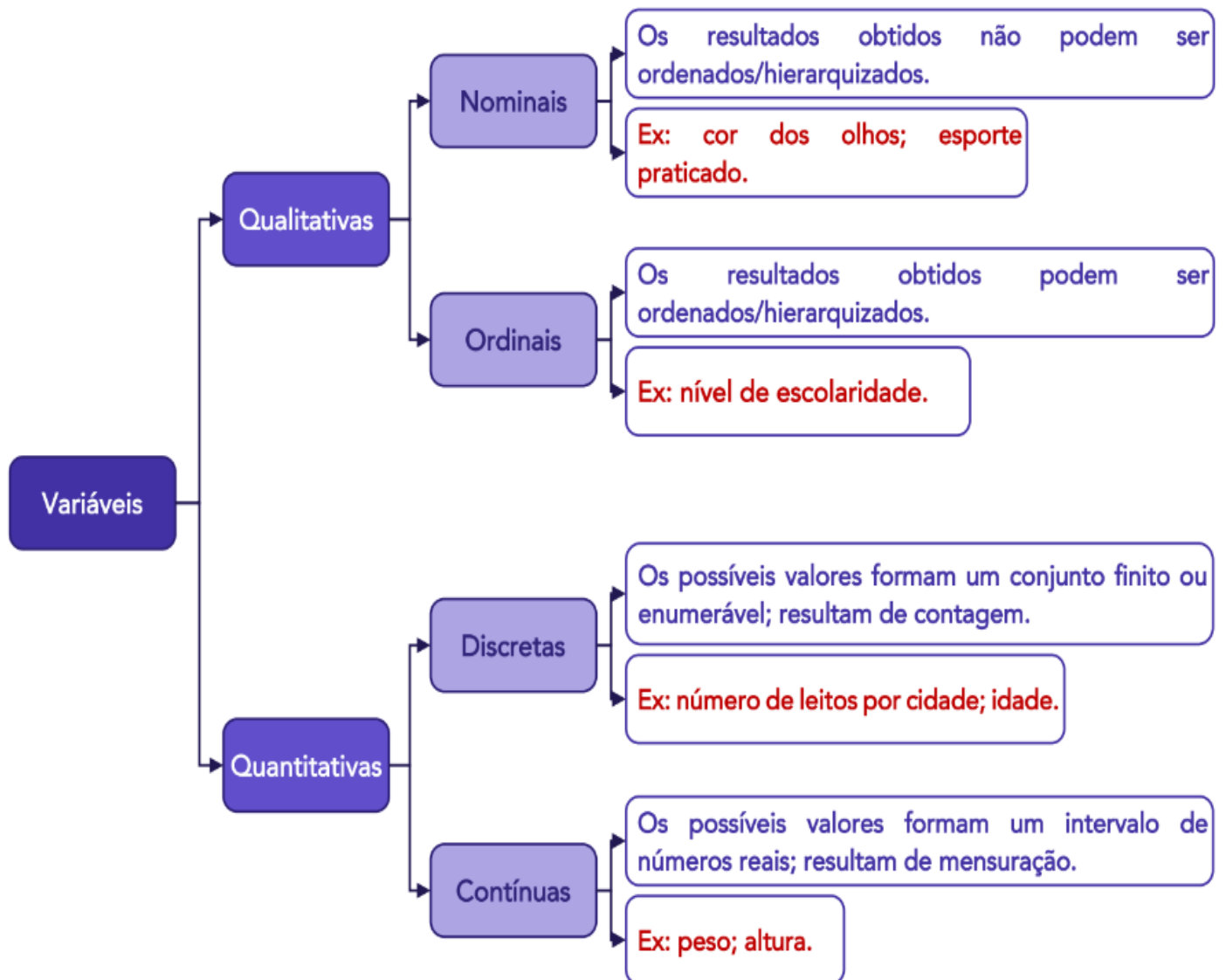
São aqueles que não foram numericamente organizados em ordem crescente ou decrescente, ou seja, estão na forma como foram coletados.

ROL

É a organização dos dados brutos em ordem de grandeza crescente ou decrescente.

Variáveis Estatísticas

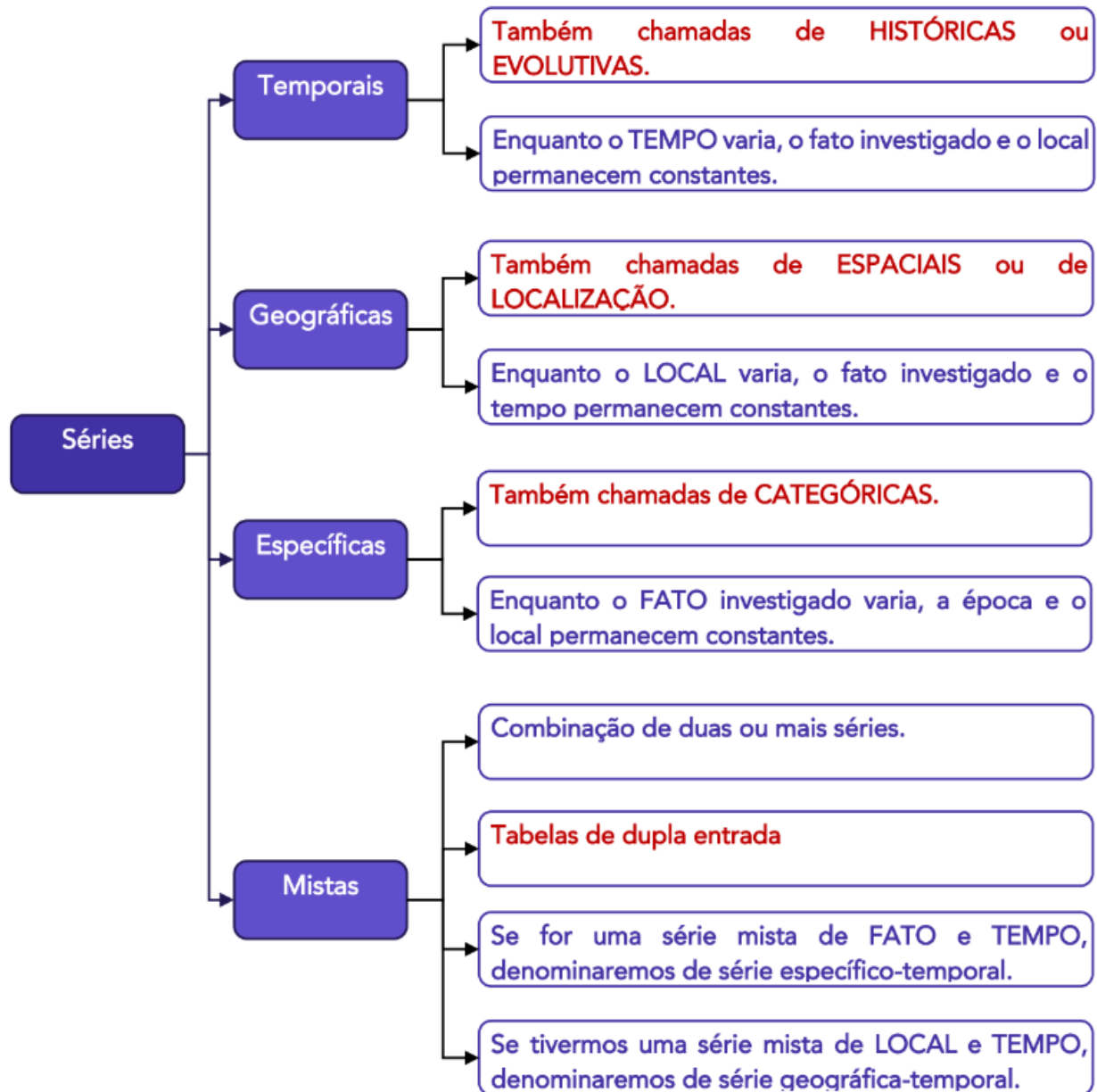
As variáveis estatísticas podem ser classificadas, inicialmente, em duas categorias: **qualitativas e quantitativas**.



Séries Estatísticas



As séries estatísticas podem ser classificadas em: temporais, geográficas, específicas ou mistas:



Distribuição de Frequências

As distribuições de frequências podem ser classificadas em dois tipos: distribuição de frequências **pontual (ou discreta)** e distribuição de frequências **intervalar (ou contínua)**.

DISTRIBUIÇÃO DE FREQUÊNCIAS PONTUAL

São apresentados todos os dados coletados juntamente com suas respectivas frequências, não havendo perda de valores.

DISTRIBUIÇÃO DE FREQUÊNCIAS INTERVALAR

É agrupamento os valores por intervalos de classe.

Elementos de uma Distribuição de Frequências

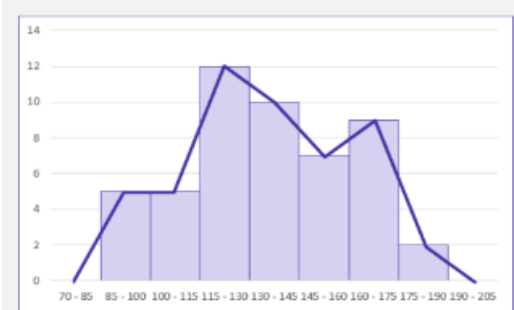
Item	Definição	Símbolos e Fórmulas
Número de Classes	As classes são os intervalos nos quais o fenômeno é subdividido.	$k = 1 + 3,3 \times \log n$ $k = \sqrt{n}$
Limites de Classe	Correspondem aos valores extremos.	l_{inf} e l_{sup}
Amplitude de um Intervalo de Classe	Distância entre os limites inferiores (ou superiores) de classes consecutivas.	$h = l_{sup} - l_{inf}$
Amplitude total	Diferença entre o limite superior da última classe (limite superior máximo) e o limite inferior da primeira classe (limite inferior mínimo).	$AT = l_{máx} - l_{mín}$ $AT = h \times k$
Ponto Médio	Média aritmética simples dos valores extremos de uma classe.	$PM = \frac{(l_{inf} + l_{sup})}{2}$ $PM = l_{inf} + \frac{h}{2}$ $PM = l_{sup} - \frac{h}{2}$
Frequência Absoluta Simples	Número de observações correspondentes a uma determinada classe ou a um determinado valor.	f_i

Frequência Absoluta Acumulada	Total das frequências de todos os valores inferiores ao limite superior do intervalo de uma dada classe	$f_{ac_i} = f_1 + f_2 + f_3 + \dots + f_i$
Frequência Relativa Simples	Proporção de dados existentes em uma determinada classe.	$F_i = \frac{f_i}{\sum f_i} = \frac{f_i}{n}$
Frequência Relativa Acumulada	Proporção de valores inferiores ao limite superior do intervalo de uma dada classe.	$F_{ac_i} = F_1 + F_2 + F_3 + \dots + F_i$
Densidade de Frequência	Quociente entre a frequência da classe (absoluta ou relativa) e sua amplitude	$d = \frac{f}{h}$

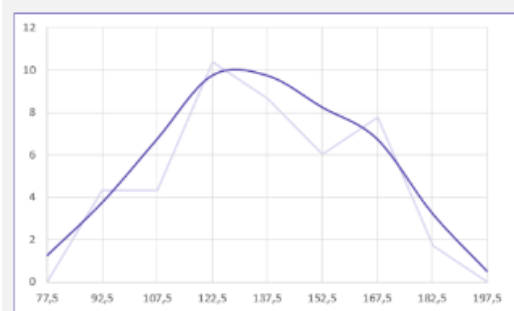


Representações Gráficas das Distribuições de Frequências

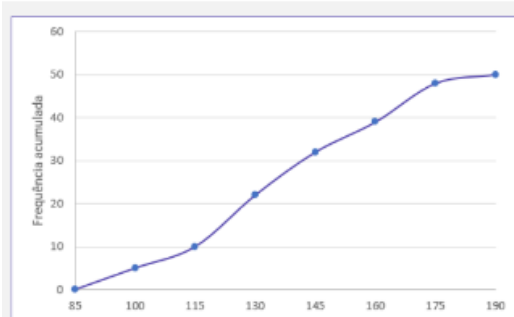
Gráfico	Definição																
<p>Gráfico de hastes ou bastões (Stem Plot) mostrando a distribuição de frequências para dados discretos. O eixo horizontal (X) representa os valores discretos (0 a 6), e o eixo vertical (Y) representa a frequência (0 a 30). As hastes são representadas por pontos no topo de linhas verticais (hastes) que se estendem do eixo X.</p> <table border="1"> <thead> <tr> <th>Valor (X)</th> <th>Frequência (Y)</th> </tr> </thead> <tbody> <tr><td>0</td><td>4</td></tr> <tr><td>1</td><td>10</td></tr> <tr><td>2</td><td>25</td></tr> <tr><td>3</td><td>15</td></tr> <tr><td>4</td><td>10</td></tr> <tr><td>5</td><td>5</td></tr> <tr><td>6</td><td>1</td></tr> </tbody> </table>	Valor (X)	Frequência (Y)	0	4	1	10	2	25	3	15	4	10	5	5	6	1	<p>O gráfico de hastes ou bastões é muito utilizado para representar dados não agrupados em classes, o que normalmente ocorre com dados discretos.</p>
Valor (X)	Frequência (Y)																
0	4																
1	10																
2	25																
3	15																
4	10																
5	5																
6	1																
<p>Histograma mostrando a distribuição de frequências para dados agrupados em classes. O eixo horizontal (X) representa as classes (85-100, 100-115, 115-130, 130-145, 145-160, 160-175, 175-190), e o eixo vertical (Y) representa a frequência (0 a 14). As barras são retângulos contíguos (justapostos).</p> <table border="1"> <thead> <tr> <th>Classe (X)</th> <th>Frequência (Y)</th> </tr> </thead> <tbody> <tr><td>85 - 100</td><td>5</td></tr> <tr><td>100 - 115</td><td>5</td></tr> <tr><td>115 - 130</td><td>12</td></tr> <tr><td>130 - 145</td><td>10</td></tr> <tr><td>145 - 160</td><td>7</td></tr> <tr><td>160 - 175</td><td>9</td></tr> <tr><td>175 - 190</td><td>2</td></tr> </tbody> </table>	Classe (X)	Frequência (Y)	85 - 100	5	100 - 115	5	115 - 130	12	130 - 145	10	145 - 160	7	160 - 175	9	175 - 190	2	<p>O histograma é um gráfico destinado a representar dados agrupados em classe, sendo composto por um conjunto de retângulos contíguos (justapostos).</p>
Classe (X)	Frequência (Y)																
85 - 100	5																
100 - 115	5																
115 - 130	12																
130 - 145	10																
145 - 160	7																
160 - 175	9																
175 - 190	2																
<p>Poligonal característica (Frequency Polygon) mostrando a distribuição de frequências para dados agrupados em classes. O eixo horizontal (X) representa as classes (85-100, 100-115, 115-130, 130-145, 145-160, 160-175, 175-190), e o eixo vertical (Y) representa a frequência (0 a 14). A linha poligonal conecta os pontos médios do topo de cada barra do histograma.</p> <table border="1"> <thead> <tr> <th>Classe (X)</th> <th>Frequência (Y)</th> </tr> </thead> <tbody> <tr><td>85 - 100</td><td>5</td></tr> <tr><td>100 - 115</td><td>5</td></tr> <tr><td>115 - 130</td><td>12</td></tr> <tr><td>130 - 145</td><td>10</td></tr> <tr><td>145 - 160</td><td>7</td></tr> <tr><td>160 - 175</td><td>9</td></tr> <tr><td>175 - 190</td><td>2</td></tr> </tbody> </table>	Classe (X)	Frequência (Y)	85 - 100	5	100 - 115	5	115 - 130	12	130 - 145	10	145 - 160	7	160 - 175	9	175 - 190	2	<p>A poligonal característica é construída utilizando apenas os contornos do histograma.</p>
Classe (X)	Frequência (Y)																
85 - 100	5																
100 - 115	5																
115 - 130	12																
130 - 145	10																
145 - 160	7																
160 - 175	9																
175 - 190	2																



O **polígono de frequências** é um gráfico em linha obtido por meio da ligação, por segmentos de reta, dos pontos médios das bases superiores dos retângulos de um histograma.



A **curva de frequências** é obtida a partir do polígono de frequências



O **gráfico de ogiva** é empregado na representação de distribuições de frequências acumuladas, sejam elas crescentes ou decrescentes

**Este material contempla assuntos dispostos no livro digital (pdf) da aula 00.

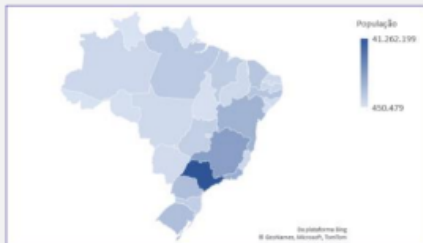


Principais formas de representação de dados estatísticos:

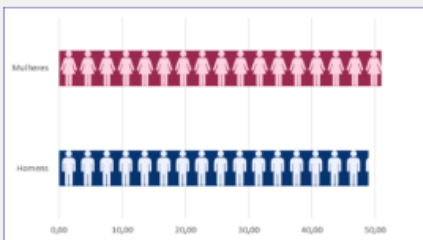
Gráfico	Definição
	<p>Os gráficos em linha normalmente são usados para representar a variação dos valores de uma variável ao longo do tempo. Esse tipo de gráfico permite-nos comparar duas variáveis: uma é traçada no eixo x (horizontal) e a outra no eixo y (vertical).</p>
	<p>Os gráficos em barra normalmente são usados para representar distribuições de dados categóricos ou qualitativos. Uma série estatística é representada por um conjunto de retângulos dispostos horizontalmente, cada um indicando uma categoria particular.</p>
	<p>Os gráficos em coluna também são usados para distribuições de dados categóricos ou qualitativos. A diferença básica é que, agora, uma série estatística é representada por um conjunto de retângulos dispostos verticalmente, cada um indicando uma categoria particular.</p>
	<p>O gráfico em setores é usado para representar a frequência relativa (porcentagem) de uma variável categórica, sendo formado por um círculo dividido em setores circulares, cada um representando uma categoria.</p>



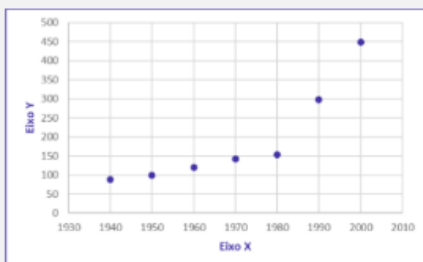
O **gráfico polar** consiste em uma sequência de eixos igualmente espaçados (ângulos iguais), cada um representando uma das variáveis. Uma linha é desenhada ligando os valores de cada eixo.



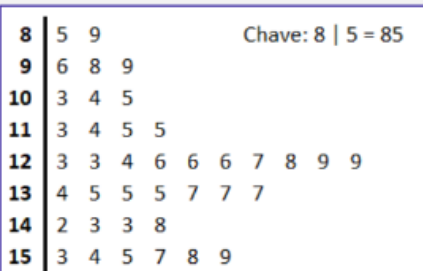
O **cartograma** é empregado com a finalidade de apresentar dados estatísticos diretamente relacionados com áreas geográficas.



O **pictograma** substitui valores por ícones, tornando os dados mais atraentes e facilitando o entendimento acerca de um determinado fenômeno.



O **gráfico de dispersão** é uma representação de pares ordenados em um plano cartesiano, composto por um eixo vertical (ordenada) e um eixo horizontal (abscissa). Os dados são representados como uma coleção de pontos.



O **diagrama de ramos e folhas** fornece uma maneira rápida de representar graficamente a distribuição dos dados. Nele, cada número é separado em duas partes. Em geral, de um lado ficam as unidades do número e do outro lado fica o restante desse número.