

Métodos Matemáticos para Gestão da Informação

Aula 03

Taxas de variação e
função lineares I



Dalton Martins
dmartins@gmail.com

Bacharelado em Gestão da Informação
Faculdade de Informação e Comunicação
Universidade Federal de Goiás

Introdução

- As taxas de variação nos permitem responder perguntas fundamentais para análise de dados do tipo:
 - Como a população de um país varia ao longo do tempo?
 - Como a altura das crianças se altera à medida que elas crescem?
- Taxas médias de variação fornecem uma ferramenta para medir como a mudança de uma variável afeta uma segunda variável.
- Quando as taxas médias de variação são constantes, a relação é linear.

Objetivos deste tópico

- Calcular e interpretar taxas médias de variação
- Entender como representações de dados podem ser tendenciosas
- Reconher que uma taxa constante de variação indica uma relação linear
- Representar funções lineares com equações, tabelas, gráficos ou palavras
- Deduzir, à mão, um modelo linear para um conjunto de dados
- Usar retas de regressão para resumir tendências lineares a partir de gráficos de dispersão.

Taxas médias de variação

- No capítulo anterior, vimos como mudanças em uma variável poderiam afetar mudanças em uma segunda variável.
- Veremos agora como medir essas mudanças

Exemplo

- Podemos pensar na população dos EUA como uma função do tempo:
 - O tempo é a variável independente e o tamanho da população é a variável dependente.

População dos Estados Unidos:
1790–2010

Ano	População em Milhões
1790	3,9
1800	5,3
1810	7,2
1820	9,6
1830	12,9
1840	17,1
1850	23,2
1860	31,4
1870	39,8
1880	50,2
1890	63,0
1900	76,2
1910	92,2
1920	106,0
1930	123,2
1940	132,2
1950	151,3
1960	179,3
1970	203,3
1980	226,5
1990	248,7
2000	281,4
2010	309,2 (est.)

Tabela 2.1

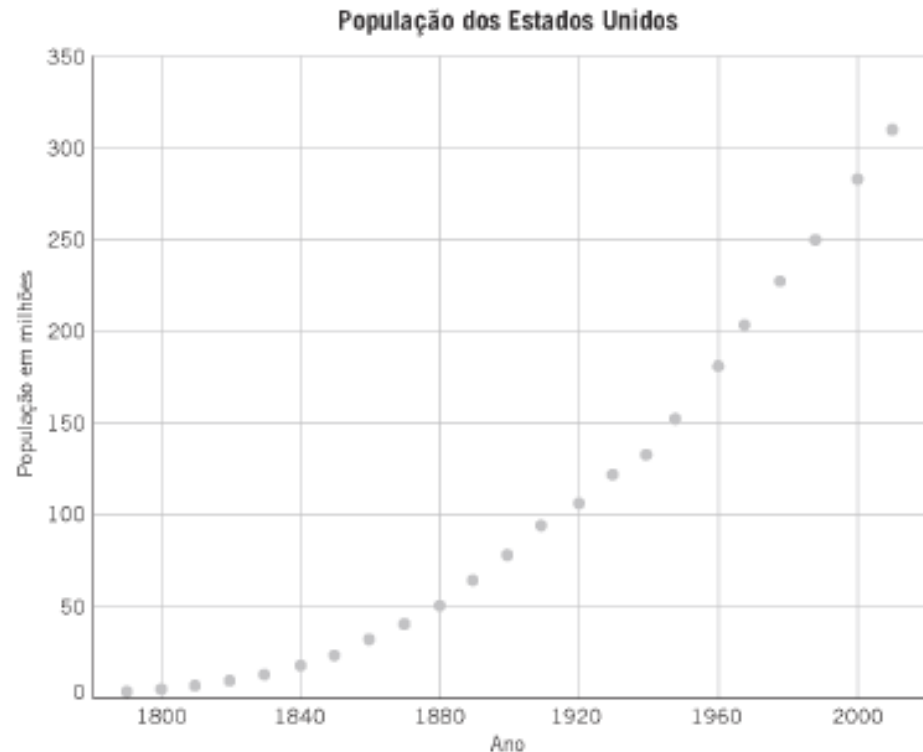


Figura 2.1

Fonte: U.S. Bureau of the Census, *Statistical Abstract of the United States*: 2010.

Exemplo

- O gráfico vem mostrando que a população vem crescendo ao longo dos dois últimos séculos e, ao que parece, a uma taxa que cresce cada vez mais.
- Como pode ser descrita quantitativamente a variação da população ao longo do tempo?
 - Uma maneira é tomar dois pontos no gráfico dos dados e calcular o quanto a população variou durante o período de tempo entre estes dois pontos.

Exemplo

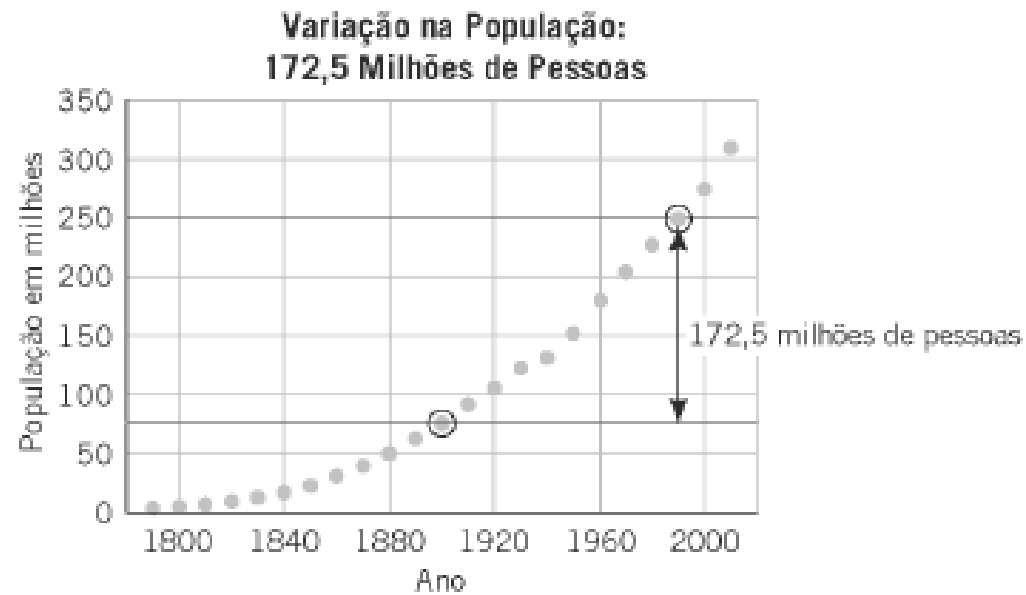


Figura 2.2

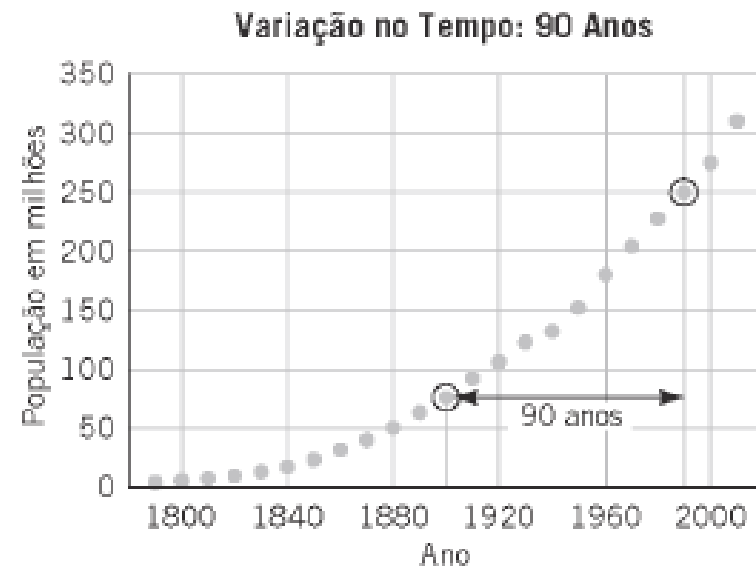


Figura 2.3

Exemplo

- Variação na população: 172,5 milhões
- Variação no tempo: 90 anos
- Para determinar a taxa média de variação na população por ano, de 1900 a 1990, divida a variação na população pela variação em anos:
 - Quanto dá para essa relação?

$$\text{taxa média de variação} = \frac{\text{variação na população}}{\text{variação em anos}}$$

Exemplo

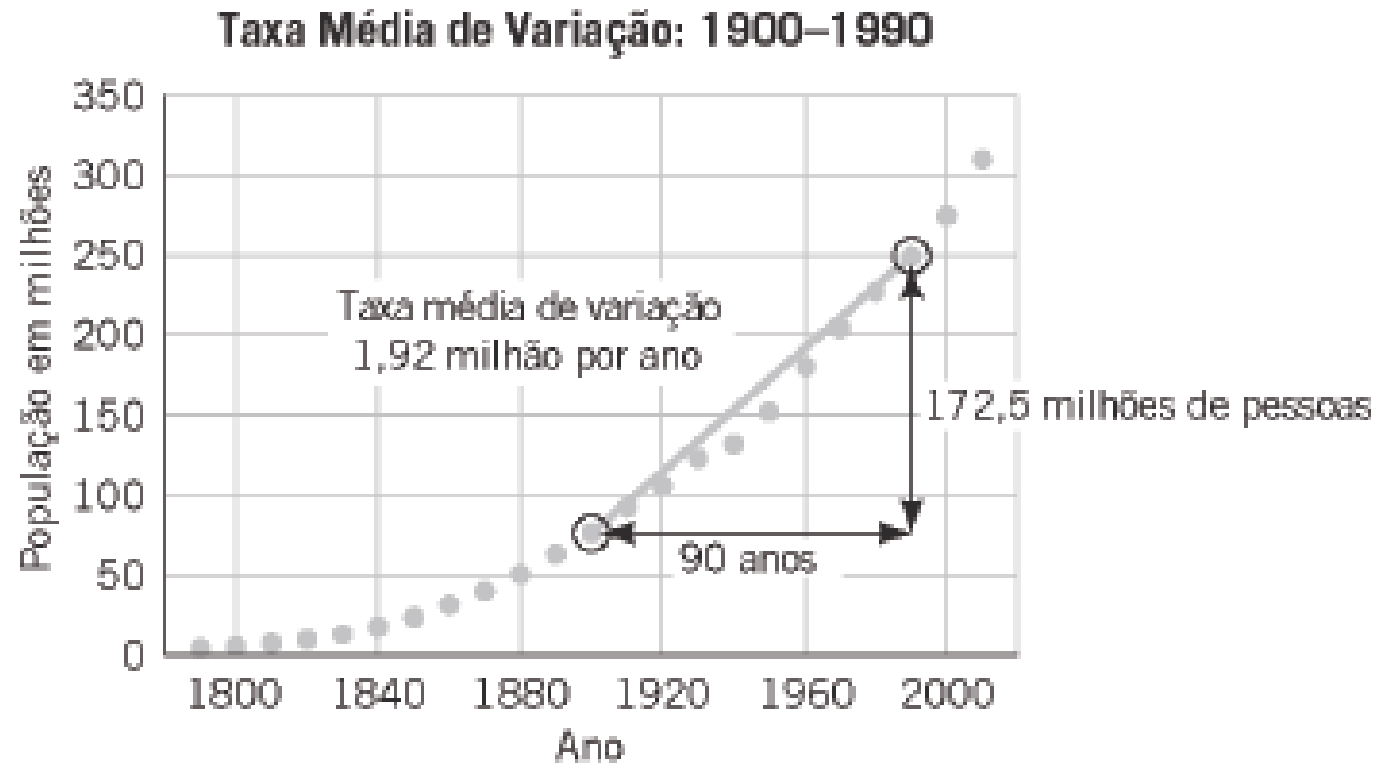


Figura 2.4

Definindo taxa média de variação

- A noção de taxa média de variação pode ser usada para descrever a variação de qualquer variável com relação a alguma outra.

$$\text{taxa média de variação de } y \text{ em relação a } x = \frac{\text{variação em } y}{\text{variação em } x}$$

Exercício

- Entre 1850 e 1950, a idade mediana dos EUA cresceu de 18,9 anos para 30,2 anos, mas até 1970 ela havia caído para 28,0.
 - Calcule a taxa média de variação na idade mediana entre 1850 e 1950
 - Compare a sua resposta no item anterior com a taxa média de variação entre 1950 e 1970 em termos numéricos e contexto da sociedade
 - Você esperaria que idade mediana crescesse ou diminuísse após 1970?

Limites da taxa média de variação

- A taxa média de variação é uma medida de tendência central, portanto, ela tem os limites desse tipo de medida;
- Calculamos, por exemplo, que a taxa média de variação da população dos EUA entre 1900 e 1990 foi de 1,92 milhão de pessoas/ano
 - É altamente improvável que a cada ano a população tenha crescido exatamente 1,92 milhão de pessoas.
- A taxa média de variação depende do ponto inicial e do ponto final. Se nem todos os pontos de dados estiverem em uma linha reta, a taxa média de variação não será a mesma se tomarmos intervalos de tempo diferentes.

Exercício

- Calcule e faça o gráfico da taxa média de variação da população dos EUA nos seguintes períodos de tempo:
 - 1840 a 1940
 - 1880 a 1980.

Exercício

Intervalo de Tempo	Variação no Tempo	Variação na População	Taxa Média de Variação
1840–1940	100 anos	$132,2 - 17,1 = 115,1$ milhões	$\frac{115,1 \text{ milhões}}{100 \text{ anos}} \approx 1,15 \text{ milhão/ano}$
1880–1980	100 anos	$226,5 - 50,2 = 176,3$ milhões	$\frac{176,3 \text{ milhões}}{100 \text{ anos}} \approx 1,76 \text{ milhão/ano}$

Tabela 2.2

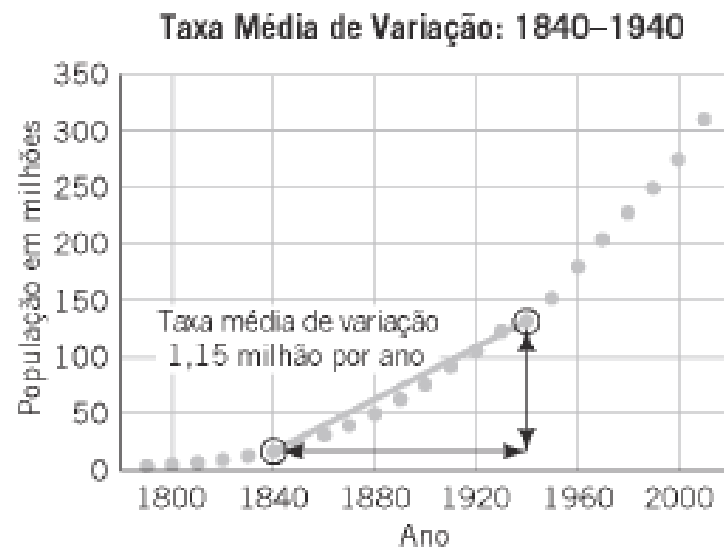


Figura 2.5

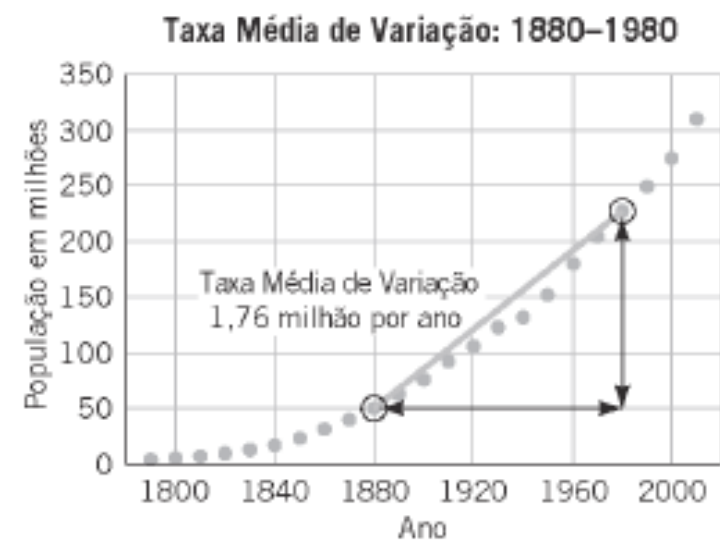


Figura 2.6

Lista de exercícios 1

Variação da taxa média de variação

- Podemos obter uma noção ainda melhor de padrões na população dos EUA se observarmos como a taxa média de variação varia ao longo do tempo;
- Uma maneira de se fazer isto é escolher um intervalo fixo para o tempo e depois calcular a taxa média de variação para cada período de tempo consecutivo.

Exemplo

- Como temos os dados da população dos EUA em intervalos de 10 anos, podemos calcular a taxa média de variação para cada década consecutiva;
- A terceira coluna da próxima tabela mostra esses resultados;
- Cada entrada representa o crescimento da população médio por ano (a taxa de variação anual média) durante a década anterior.

Exemplo

Taxas da Variação Anuais Médias na População dos EUA: 1790–2010

Ano	População (milhões)	Taxa de Variação Anual Média da Década Anterior (milhões/ano)	Exemplos dos Cálculos
1790	3,9	Dados não disponíveis	
1800	5,3	0,14	$0,14 = (5,3 - 3,9)/(1800 - 1790)$
1810	7,2	0,19	
1820	9,6	0,24	
1830	12,9	0,33	
1840	17,1	0,42	$0,42 = (17,1 - 12,9)/(1840 - 1830)$
1850	23,2	0,61	
1860	31,4	0,82	
1870	39,8	0,84	
1880	50,2	1,04	
1890	63,0	1,28	
1900	76,2	1,32	
1910	92,2	1,60	
1920	106,0	1,38	
1930	123,2	1,72	
1940	132,2	0,90	$0,90 = (132,2 - 123,2)/(1940 - 1930)$
1950	151,3	1,91	
1960	179,3	2,80	
1970	203,3	2,40	
1980	226,5	2,32	
1990	248,7	2,22	
2000	281,4	3,27	
2010	309,2 (est.)	2,78	

Tabela 2.3

Fonte: U.S. Bureau of the Census, *Statistical Abstract of the United States: 2010*.

O que está acontecendo com a taxa média de variação ao longo do tempo?

- Perceba que até 1910 a taxa média de variação crescia a cada ano;
- Não apenas a população estava crescendo a cada década até 1910, mas ela estava crescendo a um ritmo cada vez mais intenso:
 - Não importa só saber se cresce ou decresce, mas **COMO** cresce ou decresce o fenômeno que estamos estudando;
- Um aspecto que não era tão óbvio nos dados originais agora é evidente: nos intervalos de 1910 a 1920, de 1930 a 1940 e de 1960 a 1990 vemos uma população aumentando, mas uma taxa de crescimento diminuindo.

Exercício

- Variações na taxa de variação da população dos EUA
 - Faça o gráfico e interprete a taxa média de variação da população ao longo do tempo;
 - O que isto lhe diz a respeito da população dos EUA?

Exercício - resolvido

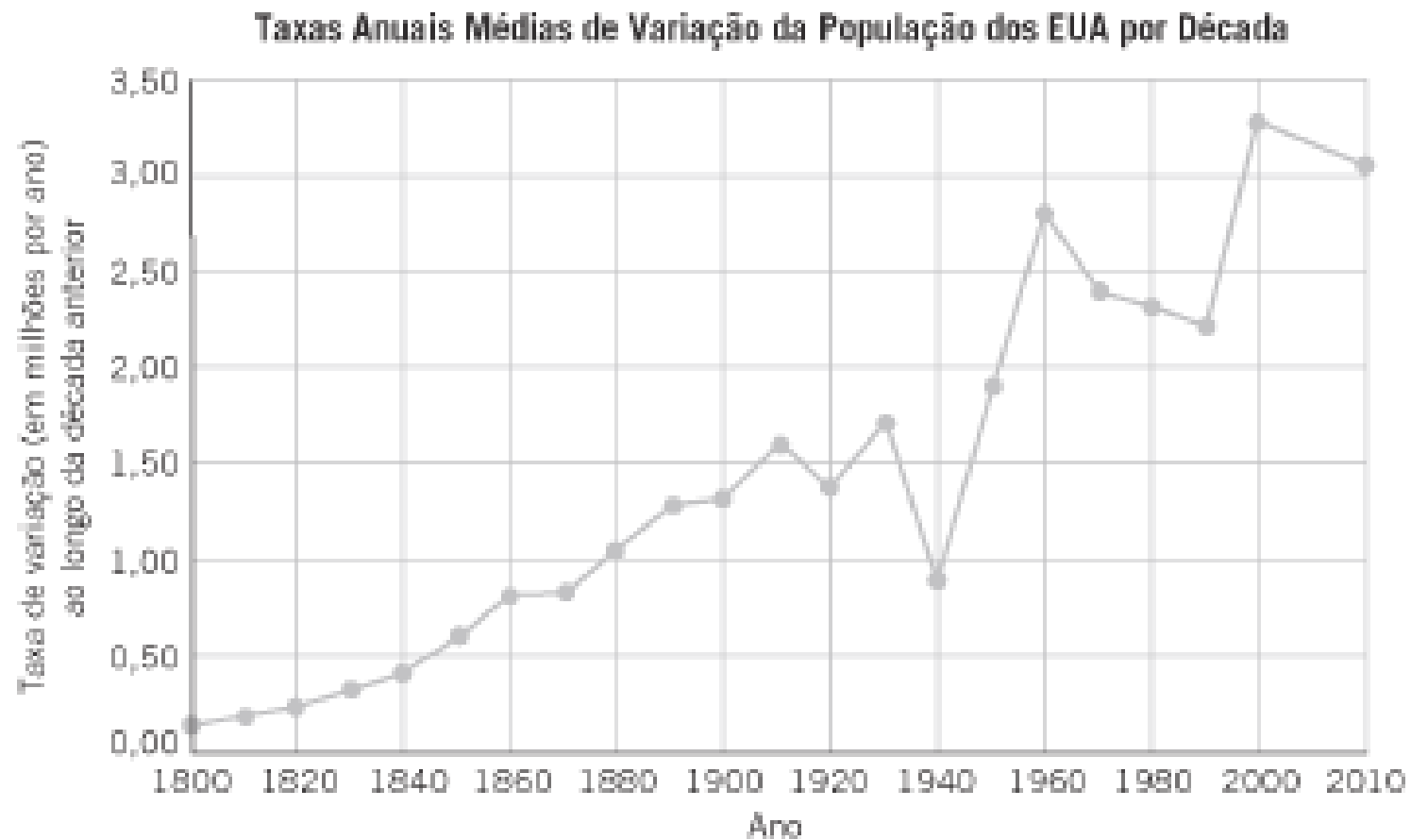


Figura 2.7

Lista de exercícios 2

A taxa média de variação é uma declividade

- Em um gráfico, a taxa média de variação é a declividade da reta que liga dois pontos;
- A declividade é um indicador de quão inclinada é a reta;
- Calculando declividades:
 - Se (x_1, y_1) e (x_2, y_2) são dois pontos, então a variação em y é igual a $y_2 - y_1$;
 - Em x , acontece da mesma forma e aí teremos o seguinte:

A taxa média de variação é uma declividade

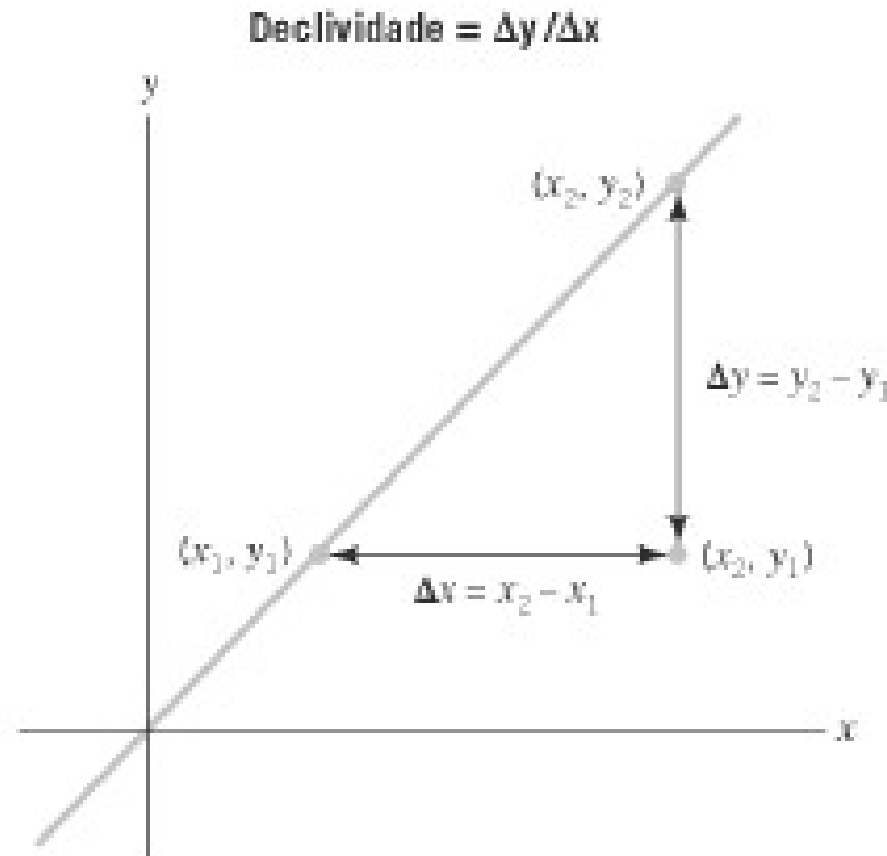


Figura 2.8

$$\text{taxa média de variação} = \frac{\text{variação em } y}{\text{variação em } x} = \frac{y_2 - y_1}{x_2 - x_1}$$

A taxa média de variação é uma declividade

- Dados dois pontos, (x_1, y_1) e (x_2, y_2) , não importa qual deles utilizamos como o primeiro ponto quando calculamos a declividade;
- Em outras palavras, podemos calcular colocando x_1 primeiro ou x_2 ;
- Os dois cálculos resultam no mesmo valor:
 - Os sinais se anulam e se compensam no resultado final.

Exemplo

- Marque no gráfico os dois pontos $(-2, -6)$ e $(7, 12)$ e calcule a declividade da reta que passa por eles.

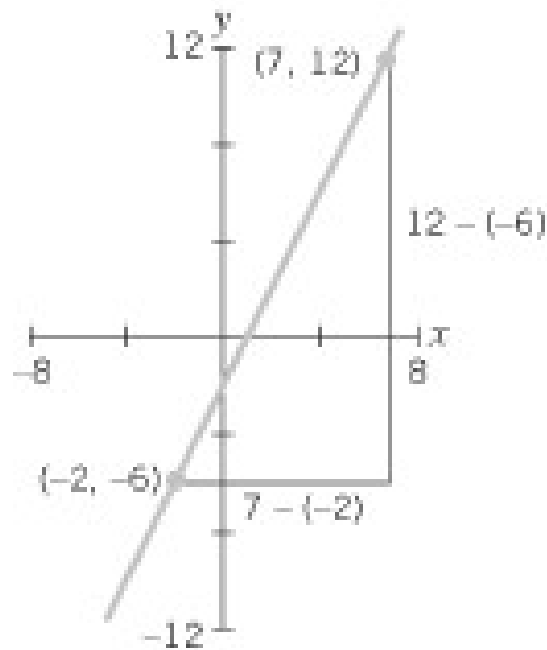


Figura 2.9

Exemplo

- O percentual da população dos EUA que vive em áreas rurais diminuiu de 84,7% em 1850 para 21,0% em 2000. Marque estes dados em um gráfico, depois calcule e interprete a taxa média de variação da população rural ao longo do tempo.

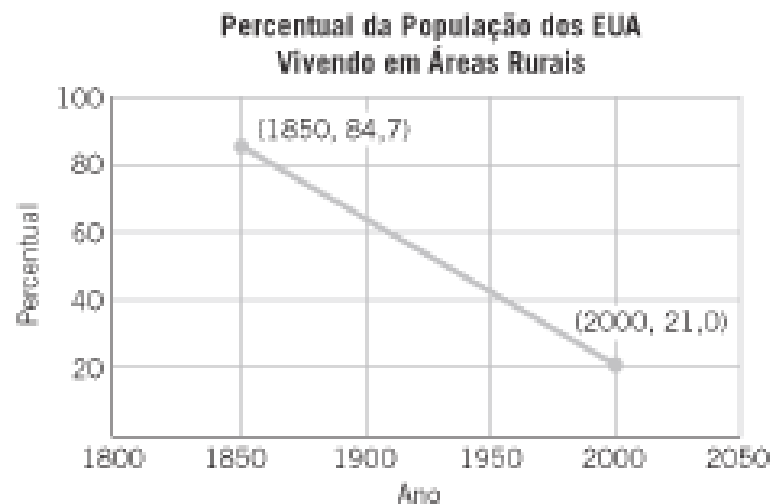


Figura 2.10

Fonte: U.S. Bureau of the Census, www.census.gov

Sinais da declividade

- Se a declividade da reta que liga quaisquer dois pontos de dados em um gráfico for positiva, a saída aumenta quando a entrada aumenta;
- Se a declividade for negativa, a saída diminui quando a entrada aumenta;
- Se a declividade for igual a nula, a reta é horizontal e a saída é constante para todas as entradas.

Exercício

- Use a tabela abaixo para traçar o gráfico das perturbações da ordem pública em cidades dos EUA (em intervalos de 3 meses) entre 1967 e 1977. Ligue os pontos e indique no gráfico quando a taxa de variação entre os pontos adjacentes for positiva, negativa e nula.
- Descreva o padrão que você identifica no gráfico.

Perturbações da Ordem Pública em Cidades dos EUA

Ano	Período	Número de Distúrbios	Ano	Período	Número de Distúrbios
1968	Jan.–Mar.	6	1971	Jan.–Mar.	12
	Abr.–Jun.	46		Abr.–Jun.	21
	Jul.–Set.	25		Jul.–Set.	5
	Out.–Dez.	3		Out.–Dez.	1
1969	Jan.–Mar.	5	1972	Jan.–Mar.	3
	Abr.–Jun.	27		Abr.–Jun.	8
	Jul.–Set.	19		Jul.–Set.	5
	Out.–Dez.	6		Out.–Dez.	5
1970	Jan.–Mar.	26			
	Abr.–Jun.	24			
	Jul.–Set.	20			
	Out.–Dez.	6			

Tabela 2.4

Fonte: D.S. Moore and G.P. McCabe, *Introduction to the Practice of Statistics*.
Copyright © 1989 de W.H. Freeman and Company. Usado com permissão.

Exercício - resultado

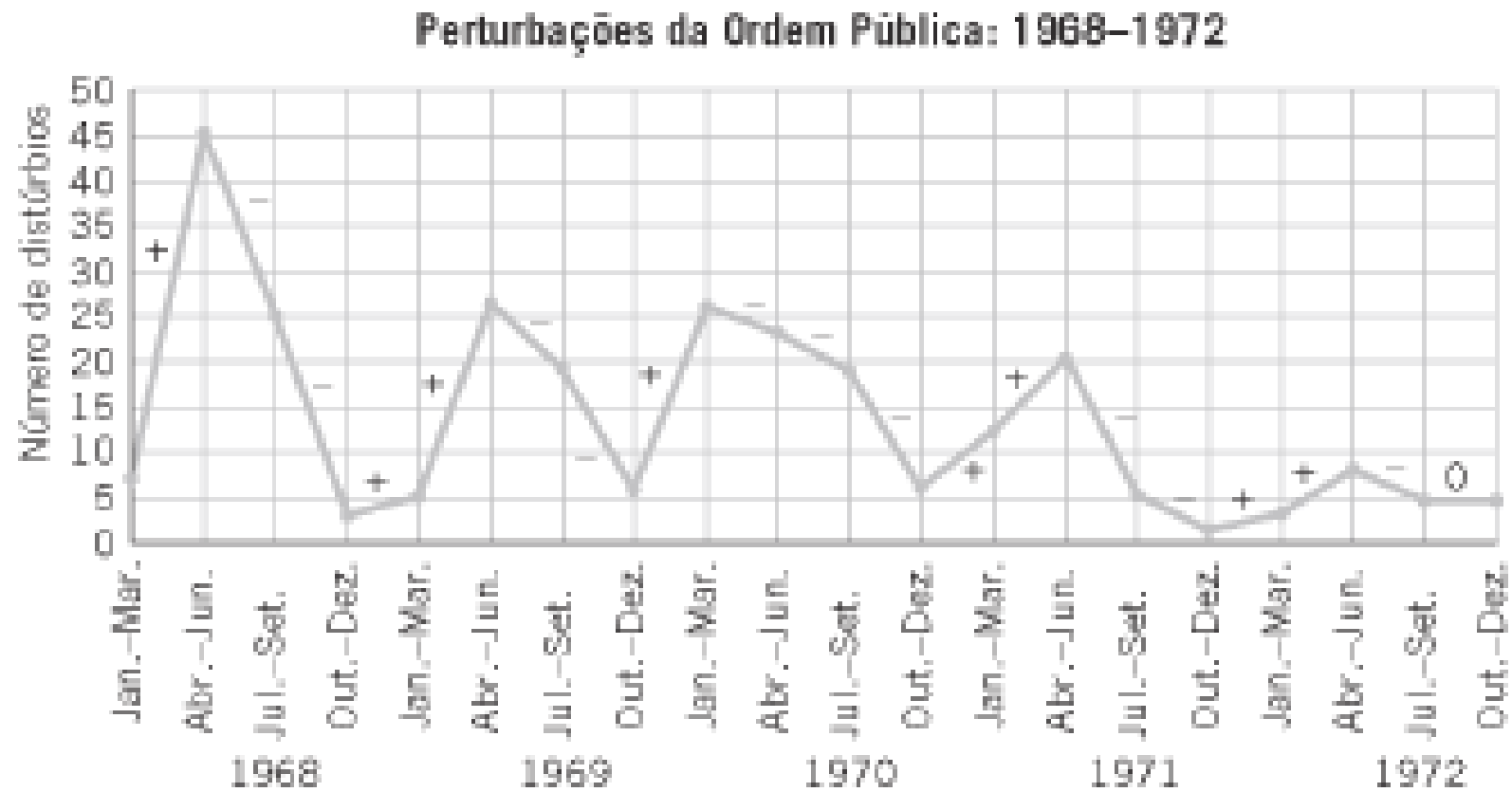


Figura 2.11

Fonte: D.S. Moore and G.P. McCabe, *Introduction to the Practice of Statistics*.
Copyright © 1989 by W.H. Freeman and Company. Usado com permissão.

Lista de exercícios 3

Enviesando os dados

- Sempre que alguém resume um conjunto de dados, **estão sendo feitas escolhas**;
- Uma escolha pode não ser mais “correta” do que outra, mas estas escolhas podem transmitir, acidental ou propositalmente, impressões muito diferentes.

Enviesando a declividade: escolhendo diferentes pontos iniciais e finais

- Dentro de um mesmo conjunto de dados, uma escolha dos pontos extremos pode mostrar um quadro otimista, enquanto outra escolha pode retratar um resultado mais pessimista.
- Os dados da tabela abaixo e o gráfico mostram o número de pessoas abaixo da linha da pobreza nos EUA de 1960 a 2008. Como poderíamos usar essas informações para defender a ideia de que o nível de pobreza diminuiu? Que o nível de pobreza aumentou?

Pessoas na Pobreza nos Estados Unidos

Ano	Número de Pessoas na Pobreza (em milhares)
1960	39.851
1970	25.420
1980	29.272
1990	33.585
2000	31.581
2008	38.110

Tabela 2.5

Fonte: U.S. Bureau of the Census, www.census.gov.



Figura 2.12 Número de pessoas na pobreza nos Estados Unidos entre 1960 e 2008.

Visão otimista e pessimista

- Para criar uma visão positiva de que o número de pobres diminuiu, poderíamos escolher como pontos extremos do intervalo (1960,39.851) e (2008,38.110):
 - A taxa média de variação seria: -36 mil pessoas/ano
- Para criar uma visão deprimente de que os números de pobreza aumentaram, poderíamos escolher (1970,25.420) e (2008,38.110) como pontos extremos:
 - A taxa média de variação seria: 334 mil pessoas/ano.

As duas taxas médias de variação estão corretas, mas elas dão impressões bem diferentes da variação do número de pessoas vivendo na pobreza nos EUA.

Enviesando os dados com palavras e gráficos

- Se embalsarmos os dados em vocabulário sugestivo e dermos formas a gráficos que apoiem um ponto de vista particular, podemos influenciar a interpretação das informações;
- Vejamos 3 exemplos. Cada um deles contém exatamente a mesma realidade subjacente:
 - O mesmo cálculo da taxa média de variação e um gráfico com a marcação dos mesmos dois pontos de dados (2000,281,4) e (2010,309,2), representando a população dos EUA (em milhões) em 2000 e em 2010.

Visão 1: a população dos EUA cresceu apenas 2,78 milhões/ano entre 2000 e 2010

- Esticando a escala do eixo horizontal em relação à do eixo vertical faz com que a declividade parece mais achatada, minimizando desta maneira a impressão de variação.

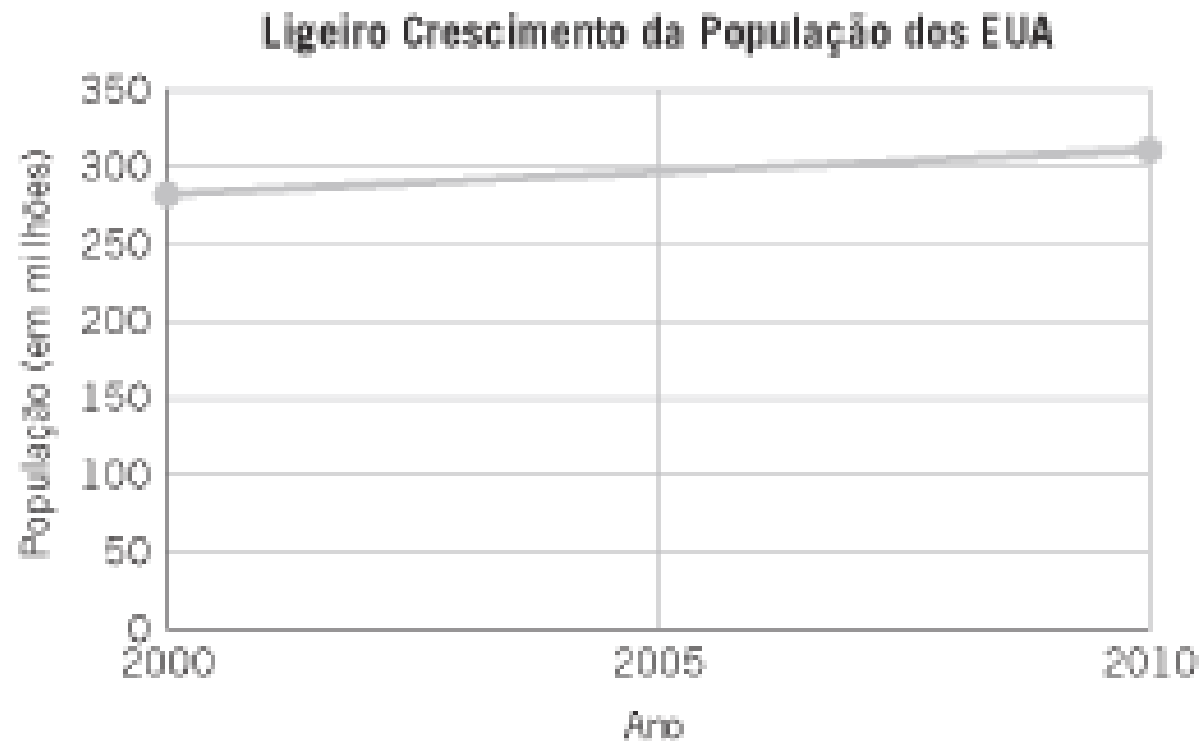


Figura 2.13

Visão 2: a população dos EUA teve um crescimento explosivo superior a 2,78 milhões/ano entre 2000 e 2010

- Cortando o eixo vertical (que agora começa em 280 em vez de 0) e esticando a escala do eixo vertical em relação ao eixo horizontal faz com que a declividade da reta pareça maior e reforça a impressão de variação dramática.

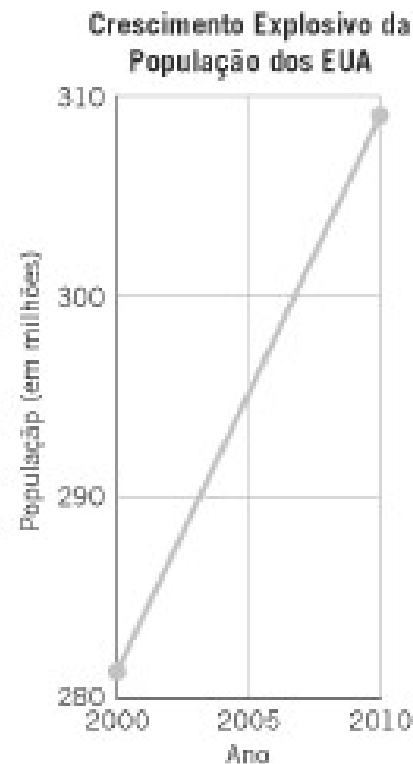


Figura 2.14

Visão 3: a população dos EUA cresceu a uma taxa razoável de 2,78 milhões/ano entre 2000 e 2010

- Visualmente a inclinação da reta abaixo parece estar grosso modo a meio caminho entre as dos dois gráficos anteriores. De fato, a declividade é a mesma para todos os 3 gráficos.

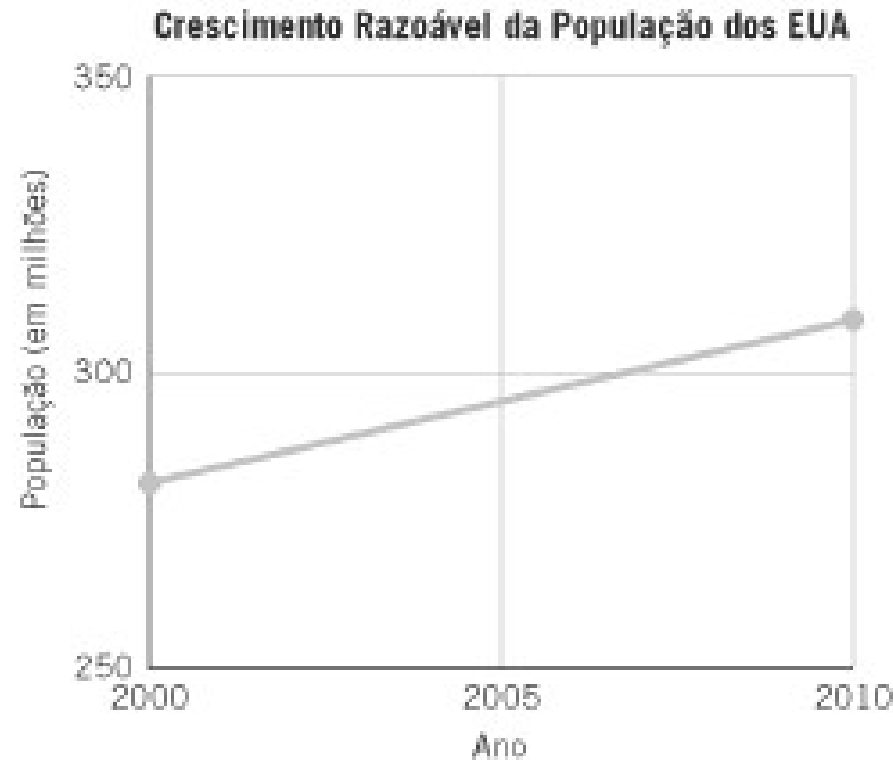


Figura 2.15

Como você poderia decidir sobre uma interpretação “justa” dos dados?

- Você poderia tentar contextualizar os dados perguntando:
 - Como o crescimento da população entre 2000 e 2010 se compara com o de outras décadas?
 - Como ela se compara com o crescimento em outros países na mesma época?
 - Esta taxa de crescimento foi facilmente acomodada, ou ela pressionou os recursos nacionais e sobrecarregou a infraestrutura?

Como você poderia decidir sobre uma interpretação “justa” dos dados?

- **FUNDAMENTAL:** a interpretação sempre depende do quadro de contexto que se deseja utilizar
 - Esse quadro é muitas vezes oculto (não dito) nos projetos e trabalhos que nos envolvemos;
 - É algo extremamente importante para o Gestor da Informação dar visibilidade a isso. Veremos mais sobre isso no semestre que vem, na disciplina de Gestão da Informação e Conhecimento.

Estatística e distorções

- Uma afirmação estatística nunca é livre de distorções;
- Para cada estatística que é citada, outras foram deixadas de fora. Isto não significa que você deva ignorar todas estatísticas.
- No entanto, você estará mais bem servido por uma abordagem criteriosa ao interpretar as estatísticas que você encontra diariamente:
 - Ao criar o hábito de fazer perguntas e depois chegar às suas próprias conclusões, você desenvolverá um bom senso sobre o entendimento dos dados.

Lista de exercícios 4