

Eixo: GESTÃO AMBIENTAL

CURSO

Análise e visualização de dados ambientais

2024



GOVERNO DO ESTADO
DO ESPÍRITO SANTO

Secretaria de Gestão
e Recursos Humanos

Secretaria do Meio Ambiente
e Recursos Hídricos

CONTRATO DIDÁTICO



EXPECTATIVAS E CONTRIBUIÇÕES

ROTEIRO

- **Objetivos do módulo**
- **O que é estatística**
- **Estatística descritiva**
 - **Medidas de centralidade**
 - **Medidas de dispersão**
 - **Correlação**
- **Ciclo PPDAC**
- **Estatística inferencial**
- **Modelagem**

OBJETIVOS DO MÓDULO

- Criar uma base sobre estatística
- Desenvolver o pensamento estatístico
- Entender como aplicar a estatística no dia-a-dia

ESTATÍSTICA SE APRENDE PRATICANDO!!!!

O QUE É ESTATÍSTICA??

A estatística é uma ciência exata e incerta, devido à aleatoriedade da natureza.

A ESTATÍSTICA É UMA DELÍCIA!!!

INTRODUÇÃO À ESTATÍSTICA

O QUE É ESTATÍSTICA??

A estatística é dividida em quatro partes:

- Estatística descritiva
- Estatística probabilística
- Estatística inferencial
- Modelagem

ESTATÍSTICA DESCRITIVA

- É A BASE DE TODA ANÁLISE ESTATÍSTICA!!!
- Como o nome diz, a estatística descritiva descreve os dados e seus padrões através de medidas de tendência central ou dispersão, por exemplo.

ESTATÍSTICA DESCRITIVA



ESTATÍSTICA DESCRITIVA

Quando se analisa dados, um conceito importante aparece, o de variáveis.

VARIÁVEIS – Uma característica que está sendo observada ou medida

- Turbidez da água
- Concentração de CO na atmosfera
- Poeira sedimentável coletada em cada estação

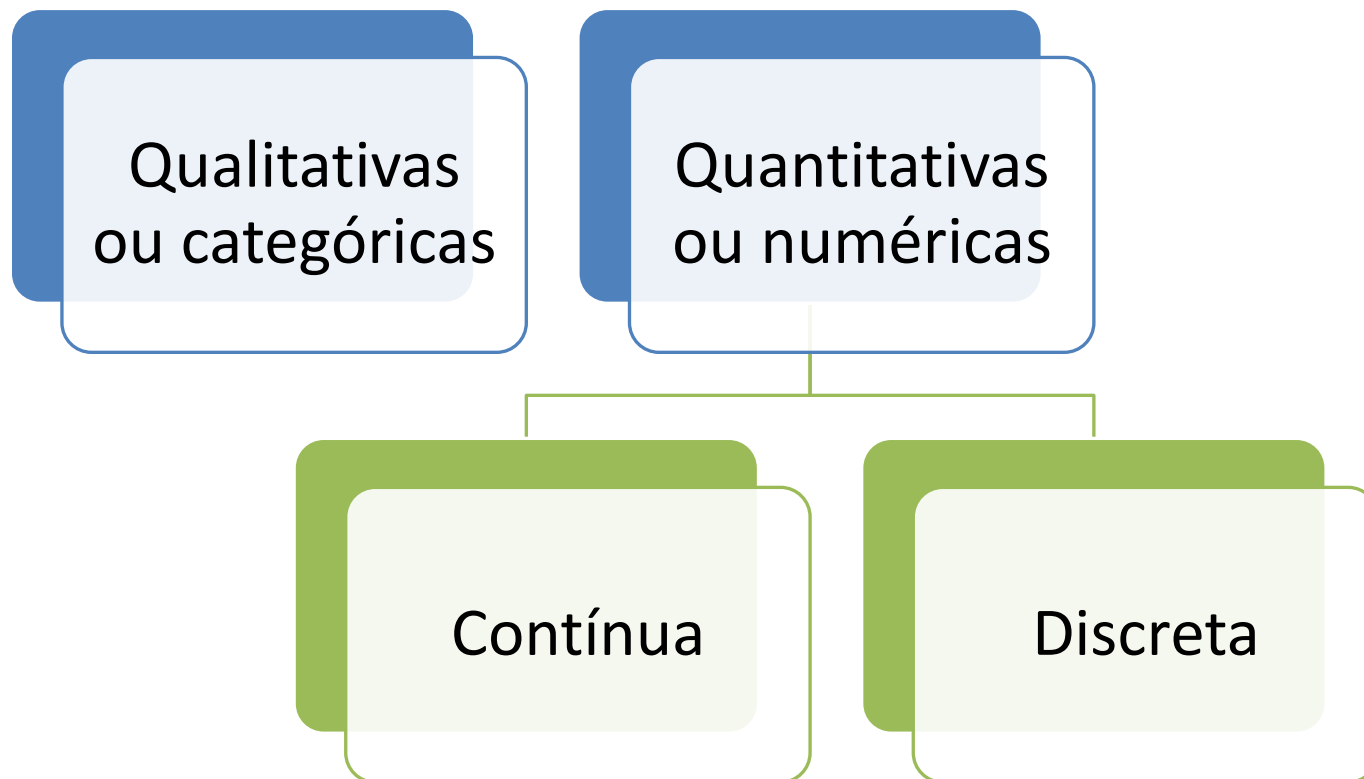
ESTATÍSTICA DESCRITIVA

VARIÁVEIS – Tipos:

- Qualitativas ou categóricas
 - Tipo sanguíneo (A, B, O, AB)
 - Ordinal (percepção sobre obras públicas)
- Quantitativas ou numéricas
 - Assume valores numéricos
 - Precipitação (30mm, 45mm...)

INTRODUÇÃO À ESTATÍSTICA

Apoio



ESTATÍSTICA DESCRITIVA

VARIÁVEIS NUMÉRICAS – Tipos:

– Discreta

- Dias de chuva
- Levantamento de fauna

– Contínua

- DAP (diâmetro na altura do peito)
- Área desmatada

ESTATÍSTICA DESCRITIVA

VARIÁVEIS – Escala:

Determina como a variável será medida e sua
representação

ESTATÍSTICA DESCRITIVA

VARIÁVEIS – Escala:

- Qualitativas ou categóricas
 - Nominal (espécies arbóreas)
 - Ordinal (percepção sobre obras públicas)
- Quantitativas ou numéricas
 - Intervalo (árvores por hectare, monitoramento de fauna)
 - Razão (Temperatura do ar, pH da água)

VARIÁVEIS – Escala:

-

ESTATÍSTICA DESCRITIVA

VARIÁVEIS – Escala:

– Qualitativa ordinal:

- Categorizam dados por ranqueamento (sem distâncias marcadas)
 - Índice de aprovação do governador pela população
 - Índice de satisfação de clientes

ESTATÍSTICA DESCRITIVA

VARIÁVEIS – Escala:

- Quantitativa intervalar:
 - Dados sem zero absoluto
 - Temperatura do ar (em °C)
 - Hora do dia
- *Zero não implica em ausência

ESTATÍSTICA DESCRITIVA

VARIÁVEIS – Escala:

- Quantitativa de razão:
 - Zero significa ausência e há múltiplos
 - Temperatura (em K)
 - Concentração de poluentes atmosféricos
 - PIB
 - * Possibilidade de cálculos aritméticos

ESTATÍSTICA DESCRITIVA

Exercício

1. Número de árvores em um parque
2. Índice de qualidade do ar – IQAr
3. Índice de desenvolvimento sustentável das cidades
4. Classe de uso de água
5. DAP de árvores
6. Módulos do curso

ESTATÍSTICA DESCRITIVA

Exercício

Quais variáveis podem ser identificadas na imagem?

Classificação ↑	Cidade	Estado	Pontuação
1	 Alfredo Marcondes	SP	66,76 ●
2	 Uru	SP	65,55 ●
3	 São Jorge do Ivaí	PR	64,71 ●
4	 Cruzália	SP	64,70 ●
5	 Jumirim	SP	64,64 ●

ESTATÍSTICA DESCRITIVA

QUALITATIVA ORDINAL

Exercício

Quais variáveis podem ser identificadas na imagem?

Classificação ↑	Cidade	Estado	Pontuação
1	 Alfredo Marcondes	SP	66,76 ●
2	 Uru	SP	65,55 ●
3	 São Jorge do Ivaí	PR	64,71 ●
4	 Cruzália	SP	64,70 ●
5	 Jumarim	SP	64,64 ●

ESTATÍSTICA DESCRITIVA

Exercício

**QUALITATIVA
NOMINAL**

Quais variáveis qualitativas foram identificadas na imagem?

Classificação ↑	Cidade	Estado	Pontuação
1	 Alfredo Marcondes	SP	66,76 ●
2	 Uru	SP	65,55 ●
3	 São Jorge do Ivaí	PR	64,71 ●
4	 Cruzália	SP	64,70 ●
5	 Jumirim	SP	64,64 ●

ESTATÍSTICA DESCRITIVA

Exercício

QUALITATIVA NOMINAL

Quais variáveis podem ser consideradas qualitativas na imagem?

Classificação ↑	Cidade	Estado	Pontuação
1	 Alfredo Marcondes	SP	66,76 ●
2	 Uru	SP	65,55 ●
3	 São Jorge do Ivaí	PR	64,71 ●
4	 Cruzália	SP	64,70 ●
5	 Jumirim	SP	64,64 ●

ESTATÍSTICA DESCRITIVA

Exercício

QUANTITATIVA DE RAZÃO

Quais variáveis podem ser identificadas?

Classificação ↑	Cidade	Estado	Pontuação
1	 Alfredo Marcondes	SP	66,76 ●
2	 Uru	SP	65,55 ●
3	 São Jorge do Ivaí	PR	64,71 ●
4	 Cruzália	SP	64,70 ●
5	 Jumirim	SP	64,64 ●

INTRODUÇÃO À ESTATÍSTICA

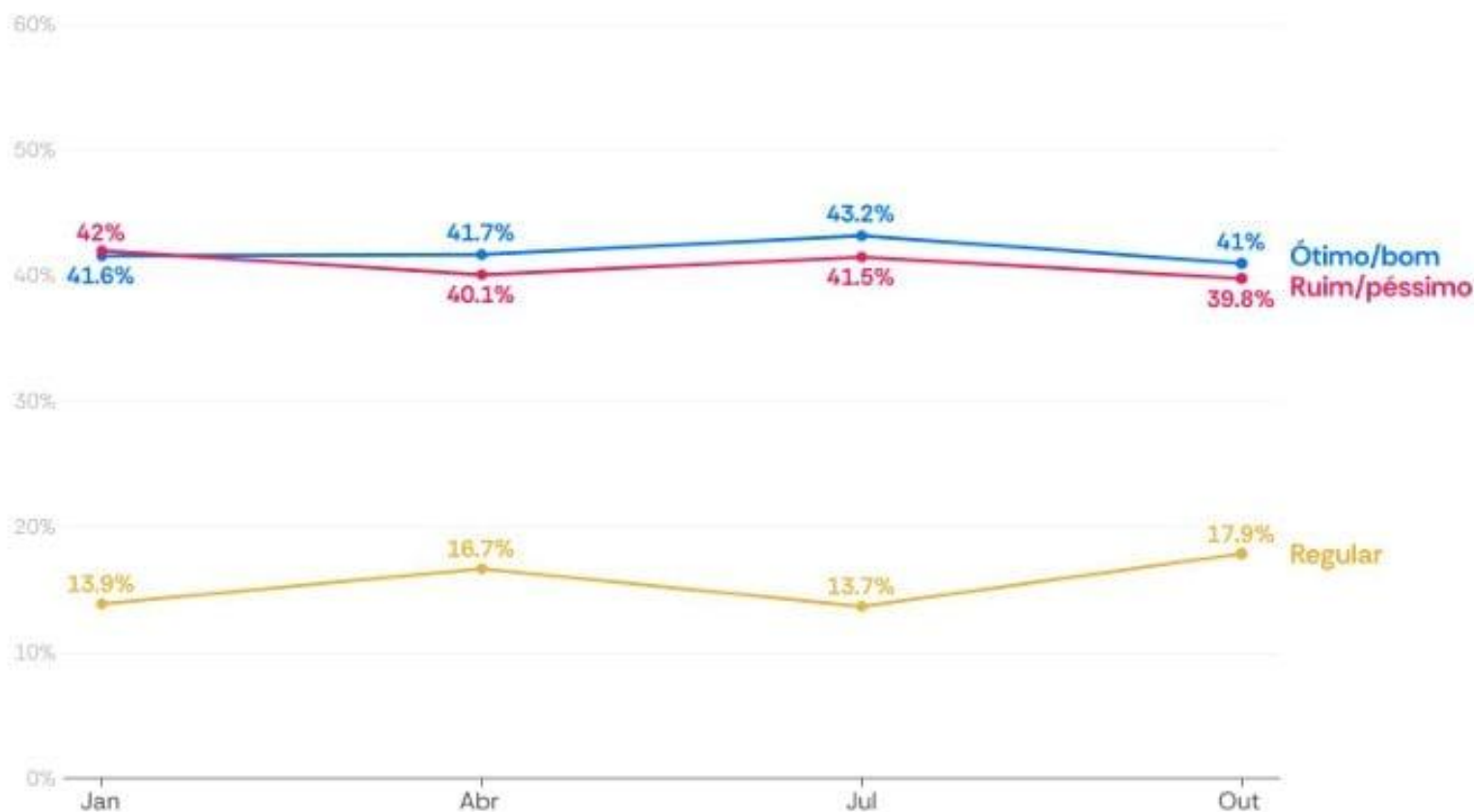
ESTATÍSTICA DESCRITIVA

Estudo de caso: Escala Likert

Criada em 1932 pelo psicólogo Rensis Likert para medir atitudes comportamentais

- Excelente – 5
- Bom – 4
- Médio – 3
- Ruim – 2
- Péssimo - 1

Como você avalia o governo do presidente Lula?



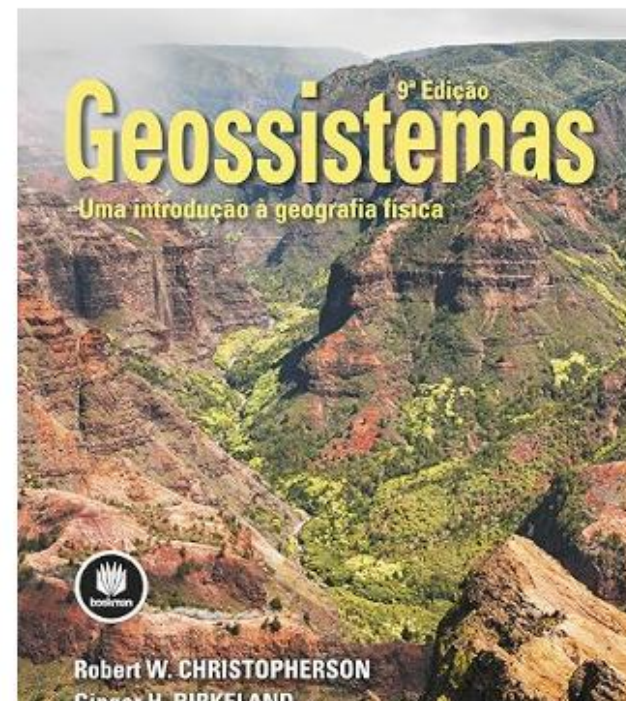
INTRODUÇÃO À ESTATÍSTICA

NET PROMOTER SCORE

$$\text{NPS} = \% \text{ PROMOTORES} - \% \text{ DETRATORES}$$



Livros › Ciências › Ciência da Terra



Geossistemas: Uma Introdução à Geografia Física

Capa comum – 2 agosto

Edição Português

por Robert Christopherson (Autor), &

4,8 ★★★★★
clientes

56 avaliações de

Ver todos os f
edições

Em até 6x R\$ 54,80 sem juros Ver parcelas disponíveis ▼

Obra reconhecida por sua meticulosa atenção aos detalhes, precisão e rica integração da ciência das mudanças climáticas, apresenta os sistemas terrestres como eles ocorrem - atmosfera, hidrosfera, litosfera e biosfera - e suas interações. Ricamente

ESTATÍSTICA DESCRITIVA

Análise de frequências:

- A frequência de uma variável mede o volume da ocorrência dessa variável
 - Frequência absoluta – número de ocorrências
 - Frequência relativa – proporção do volume

ESTATÍSTICA DESCRITIVA

Análise de frequências:

O estudo da distribuição das frequências é fundamental para a organização dos dados, além da observação do comportamento das variáveis em questão

INTRODUÇÃO À ESTATÍSTICA

ESTATÍSTICA DESCRITIVA

JW		JX		JY	JZ		KA	KB	
EMQAr - RGV9 (Cidade Continental)									
Qualidade do Ar									
Ar Ambiente									
Nitrogênio		Monóxido de Nitrogênio			Óxidos de Nitrogênio			Partículas Respiráveis	
Horária/ Alt.: 7.0 Média:1 h/ Freq.: horária/ Alt.: 7.0		Horária/ Alt.: 7.0 Média:1 h/ Freq.: horária/ Alt.: 7.0			Horária/ Alt.: 7.0 Média:1 h/ Freq.: horária/ Alt.: 7.0			Horária/ Alt.: 7.0 Média:1 h/ Freq.: horária/ Alt.: 7.0	
a		Rotina			Rotina			Rotina	
[	Flag	Valor [µg/m³]		Flag	Valor [µg/m³]		Flag	Valor [µg/m³]	
4,85			2,75			7,6			
4,92			2,94			7,86			
3,18			2,69			5,87			
3,68			2,79			6,48			
3,44			2,72			6,16			
4,25			2,84			7,09			
3,26			3,04			6,3			
5,09			3,96			9,05			
5,65			4,16			9,8			
5,18			4,13			9,31			

ESTATÍSTICA DESCRITIVA

Estudo de caso: RAMQAr RGV

Estação	N parâmetros
Laranjeiras	8
Ponta Formosa	2
Carapina	2
Jardim Camburi	6
Enseada do Suá	12
Vitória-Centro	7
Ibes	9
Vila Velha-Centro	3
Vila Capixaba	8
Cidade Continental	6

ESTATÍSTICA DESCRITIVA

Estudo de caso: RAMQAr RGV

Estação	Frequência
Laranjeiras	8
Ponta Formosa	2
Carapina	2
Jardim Camburi	6
Enseada do Suá	12
Vitória-Centro	7
Ibes	9
Vila Velha-Centro	3
Vila Capixaba	8
Cidade Continental	6

Média = 6,3

V_{máx} = 12

V_{mín} = 2

Total = 10

Amplitude = 10

Moda = 2, 6, 8

ESTATÍSTICA DESCRITIVA

Estudo de caso: RAMQAr RGV

Estação	Frequência	Porcentagem
Laranjeiras	8	0,13
Ponta Formosa	2	0,03
Carapina	2	0,03
Jardim Camburi	6	0,10
Enseada do Suá	12	0,19
Vitória-Centro	7	0,11
Ibes	9	0,14
Vila Velha- Centro	3	0,05
Vila Capixaba	8	0,13
Cidade Continental	6	0,10

ESTATÍSTICA DESCRITIVA

Estudo de caso: PS da Grande Vitória

Estação	Ultrapassagens	Classe
Enseada do Suá	49	e
Laranjeiras	21	e
Vila Capixaba	14	c
Ibes	7	b
Ilha do Boi - Hotel Senac	5	b
Vitória Centro	4	a
Jardim Camburi	3	a
Ilha do Boi - Clube Ítalo	3	a
Cidade Continental	1	a
Carapina	0	a
Praia do canto	0	a
Carapebus	0	a

ESTATÍSTICA DESCRITIVA

Estudo de caso: PS da Grande Vitória

Classe	Frequência	Frequência relativa
a	7	58%
b	2	17%
c	1	8%
d	0	0%
e	2	17%

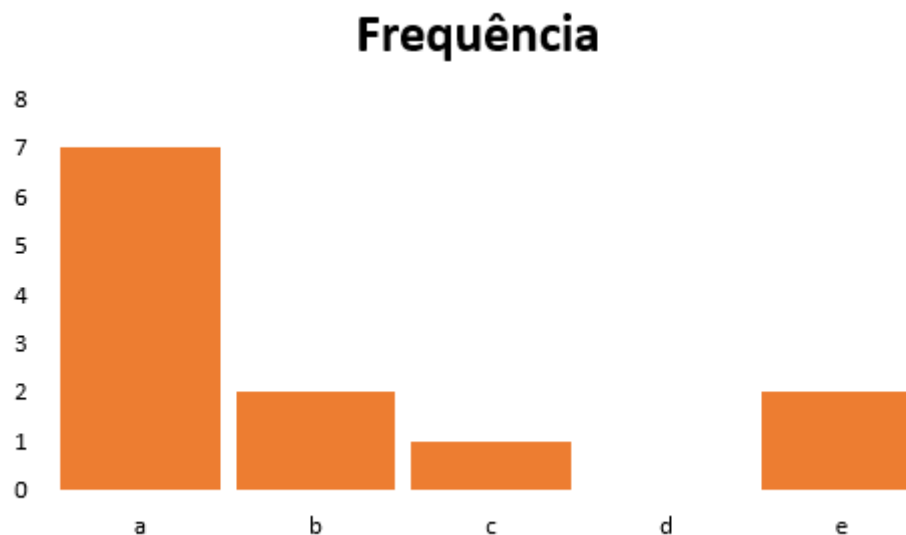
ESTATÍSTICA DESCRITIVA

Estudo de caso: PS da Grande Vitória

Padrão	Frequência	Proporção
PI-12g	3	0,02
PI-10g	8	0,05
PI-8g	24	0,14
PI-6g	51	0,29
PF	87	0,50

ESTATÍSTICA DESCRITIVA

Estudo de caso: PS da Grande Vitória



ESTATÍSTICA DESCRITIVA

Estudo de caso: IDSC Brasil



ESTATÍSTICA DESCRITIVA

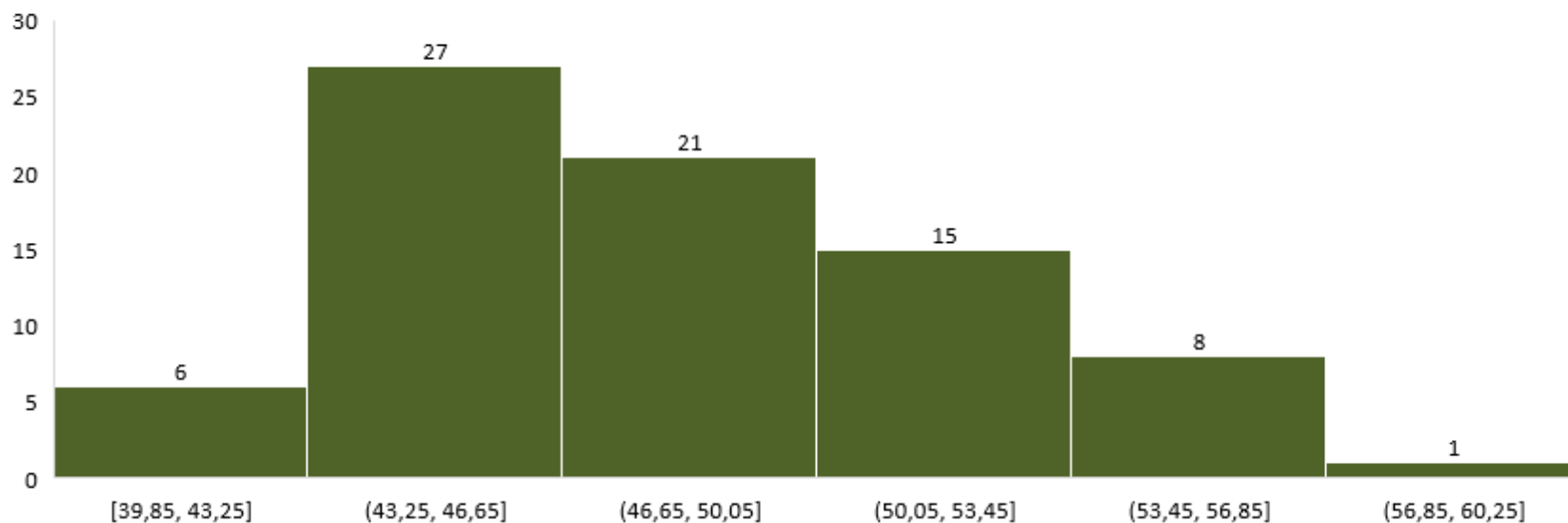
Estudo de caso: IDSC Brasil

	A	B	C	D	E	
1	COD_MUN ▼	MUNICIPIO ▼	SIGLA_UF ▼	IDSC-BR 2024 ▼	Classif. ▼	Valor
3099	3203163	Laranja da Terra	ES	58,94011807	162	
3100	3203601	Mucurici	ES	56,72033885	386	
3101	3202256	Governador Lindenberg	ES	56,23281913	442	
3102	3205069	Venda Nova do Imigrante	ES	56,02808587	475	
3103	3204609	Santa Teresa	ES	55,81277863	505	
3104	3204252	Ponto Belo	ES	55,43023901	573	
3105	3201803	Divino de São Lourenço	ES	55,4097513	577	
3106	3203163	Laranja da Terra	ES	58,94011807	162	

ESTATÍSTICA DESCRITIVA

Estudo de caso: IDSC Brasil

IDSC Espírito Santo (Excel)



ESTATÍSTICA DESCRITIVA

Estudo de caso: IDSC Brasil

Definindo as categorias

- Regra de Sturges

$$k = 1 + 3,3 * \log(n)$$

$$A = V_{\text{máx}} - V_{\text{mín}}$$

$$h = A/k$$

Onde:

k = número de classes

n = tamanho da amostra

A = Amplitude

h = intervalo

V_{máx} = Valor máximo

V_{mín} = Valor mínimo

ESTATÍSTICA DESCRITIVA

Estudo de caso: IDSC Brasil

Definindo as categorias
- Regra de Sturges

$$k = 1 + 3,3 * \log(n)$$

$$A = V_{\text{máx}} - V_{\text{mín}}$$

$$h = A/k$$

$$V_{\text{máx}} = 58,94$$

$$V_{\text{mín}} = 39,85$$

$$A = 58,94 - 39,85 = 19,09$$

$$k = 1 + 3,3 * \log(78)$$

$$k = 7,24 = 7 \text{ classes}$$

$$h = 19,09/7 = 2,64$$

ESTATÍSTICA DESCRITIVA

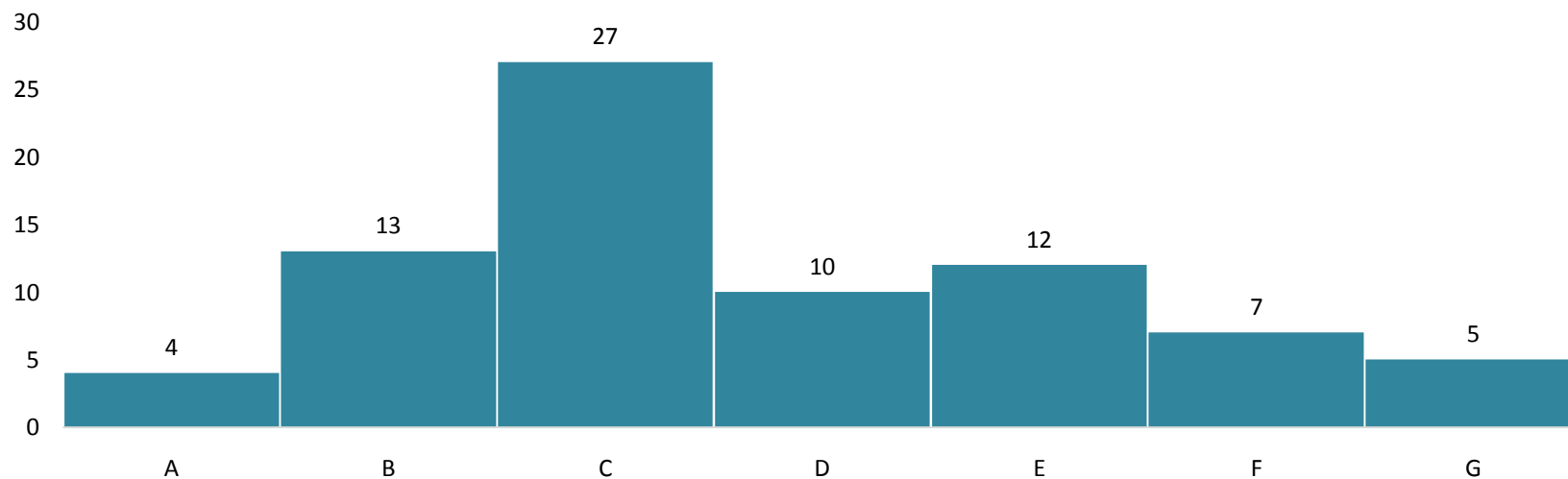
Estudo de caso: IDSC Brasil

Classe	Freq. Abs.	Freq. Rel.	F. Abs. Ac.	F. Rel. Ac.
A	4	0,05	4	0,05
B	13	0,17	17	0,22
C	27	0,35	44	0,56
D	10	0,13	54	0,69
E	12	0,15	66	0,85
F	7	0,09	73	0,94
G	5	0,06	78	1,00

ESTATÍSTICA DESCRITIVA

Estudo de caso: IDSC Brasil

IDSC Espírito Santo (Sturges)



ESTATÍSTICA DESCRITIVA



ESTATÍSTICA DESCRITIVA

Medidas de centralidade

- Representam um resumo das informações em um único dado
- Também são muito usados para comparar situações diferentes

ESTATÍSTICA DESCRITIVA

Medidas de centralidade

Média

“valor que demonstra a concentração dos dados de uma distribuição, como o ponto de equilíbrio das frequências em um histograma.” (Wikipedia)

$$\bar{x} = \frac{\sum x_i}{n}$$

ESTATÍSTICA DESCRITIVA

Medidas de centralidade

Mediana

“indica qual é o valor que está exatamente no meio de um conjunto de dados, quando eles estão ordenados”

(IBGE)

$$\tilde{X} \begin{cases} \left(\frac{X_n + 1}{2} \right), & \text{se } n \text{ é ímpar} \\ \left(\frac{X_n}{2} \right), & \text{se } n \text{ é par} \end{cases}$$

ESTATÍSTICA DESCRITIVA

Medidas de centralidade

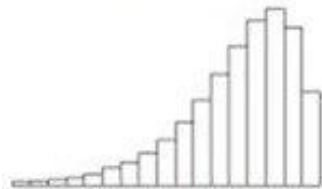
Moda

Moda é o valor mais frequente em uma distribuição de dados, podendo esta ser uniforme, unimodal, bimodal ou multimodal.

ESTATÍSTICA DESCRITIVA

Moda

Skewed Left



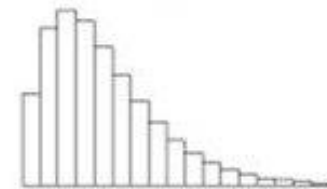
One Mode

Symmetric



Bell-Shaped

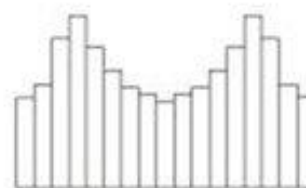
Skewed Right



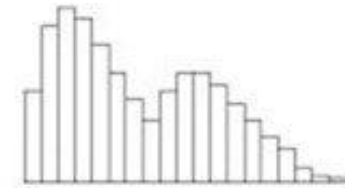
One Mode



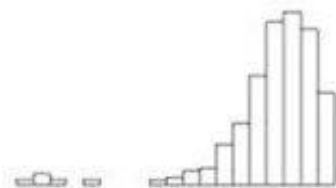
Two Modes



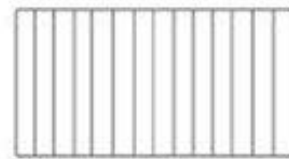
Bimodal



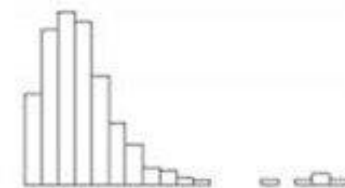
Bimodal



Left Tail Extremes



Uniform (no mode)



Right Tail Extremes

ESTATÍSTICA DESCRITIVA

Medidas de centralidade

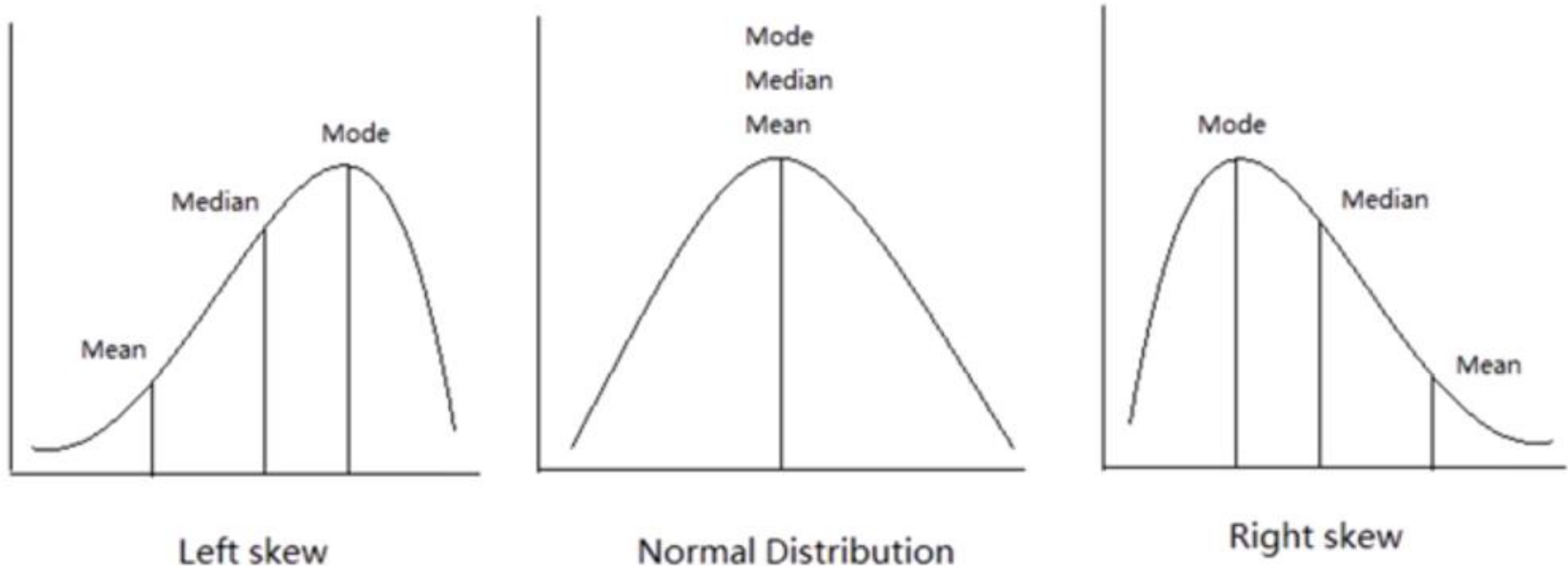
Média x Mediana x Moda

**VALORES PARA APRESENTAR E VALORES
PARA ANALISAR CONJUNTOS DE
DADOS!!!**

ESTATÍSTICA DESCRITIVA

Medidas de centralidade

Média x Mediana x Moda



ESTATÍSTICA DESCRITIVA



INTRODUÇÃO À ESTATÍSTICA

ESTATÍSTICA DESCRITIVA

Medidas de dispersão

Amplitude

$$At = x_{máx} - x_{mín}$$

Valor simples e simplista, pois apenas representa a diferença entre o maior e o menor valor.

*Muito influenciada por valores extremos

ESTATÍSTICA DESCRITIVA

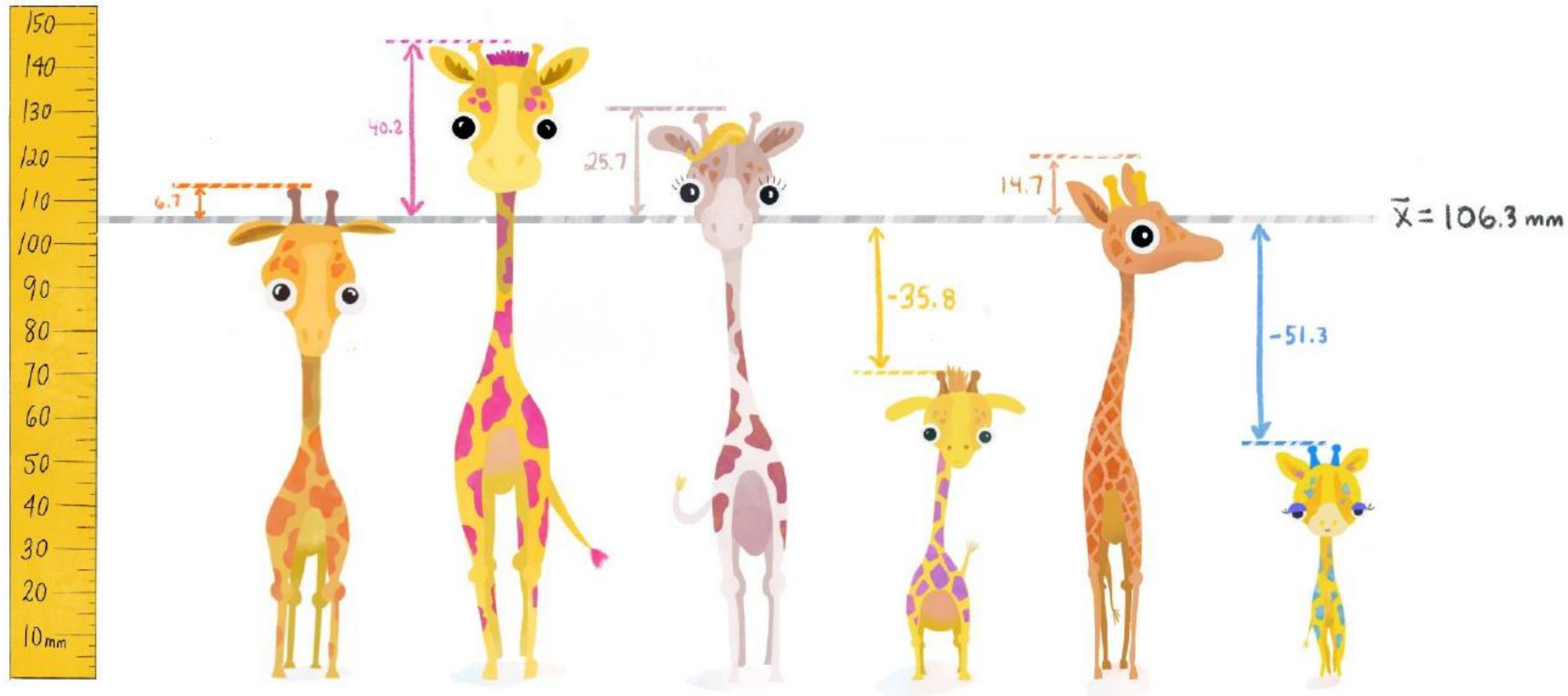
Medidas de dispersão

Variância

$$s^2 = \frac{\sum (X - \bar{X})^2}{N - 1}$$

“medida da dispersão estatística, indicando "o quão longe“, em geral, os seus valores se encontram do valor esperado” (Wikipédia)

ESTATÍSTICA DESCRITIVA



INTRODUÇÃO À ESTATÍSTICA

ESTATÍSTICA DESCRITIVA

Medidas de dispersão

Desvio padrão

“medida de dispersão em torno da média populacional de uma variável aleatória” (Wikipédia)

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

$$S = \sqrt{S^2}$$

ESTATÍSTICA DESCRITIVA

Medidas de dispersão

Desvio padrão

Pode ser utilizado para comparar a dispersão entre conjuntos de dados, pois o seu valor é atrelado à média.

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

$$S = \sqrt{S^2}$$

ESTATÍSTICA DESCRITIVA

Medidas de dispersão

Coeficiente de variação

$$CV = \frac{S}{\bar{x}} * 100$$

É a razão entre o desvio padrão (S)
e a média ($\bar{x}$).

Traz a variação dos dados em
relação ao seu valor médio.

ESTATÍSTICA DESCRITIVA

Medidas de dispersão

Coeficiente de variação

$$CV = \frac{s}{\bar{x}} * 100$$

Medida muito utilizada para
realizar comparações entre
conjuntos de dados com unidades
de medidas diferentes!

ESTATÍSTICA DESCRITIVA

Medidas de dispersão

Coeficiente de variação

$$CV = \frac{s}{\bar{x}} * 100$$

ESTATÍSTICA DESCRITIVA

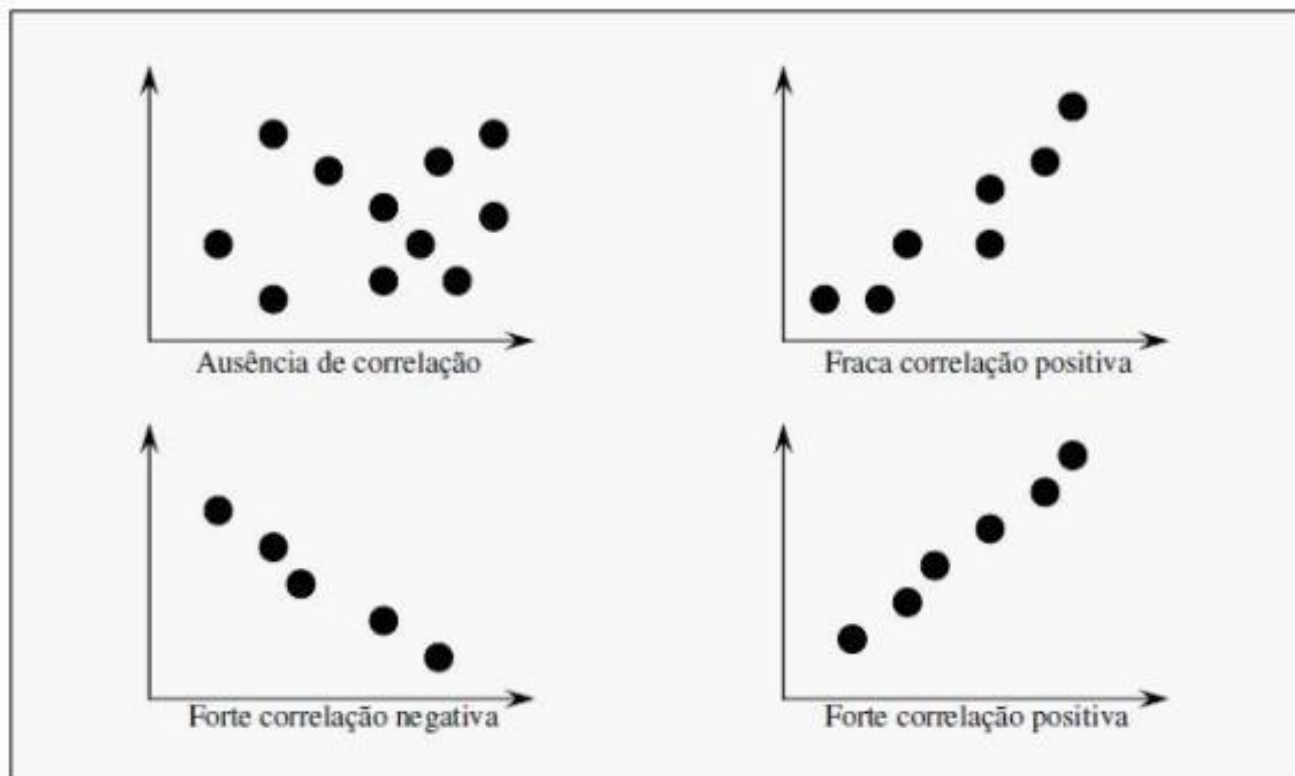
Correlação

Descreve o comportamento conjunto de duas variáveis distintas, podendo ser:

- Positiva: possuem a mesma inclinação
- Negativa: possuem inclinação inversa

ESTATÍSTICA DESCRITIVA

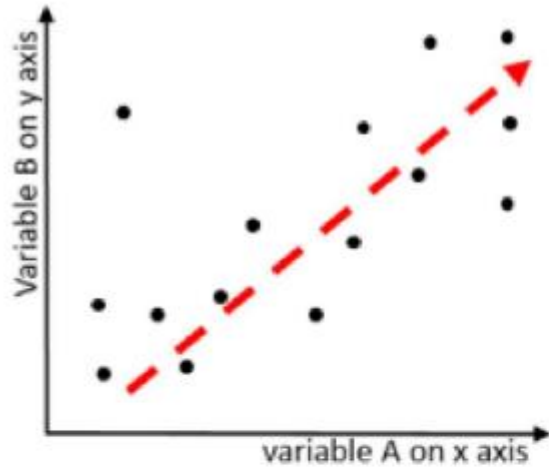
Correlação



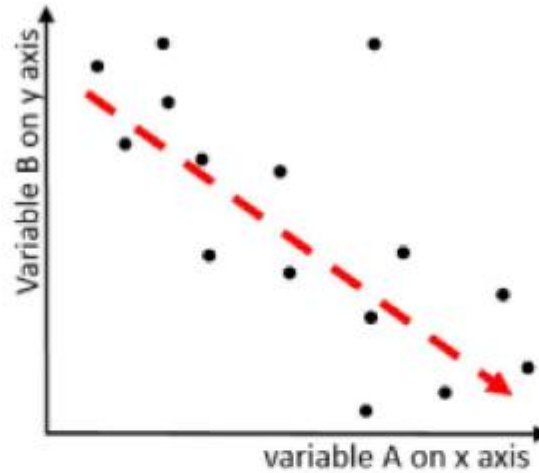
Diagramas de Dispersão.

INTRODUÇÃO À ESTATÍSTICA

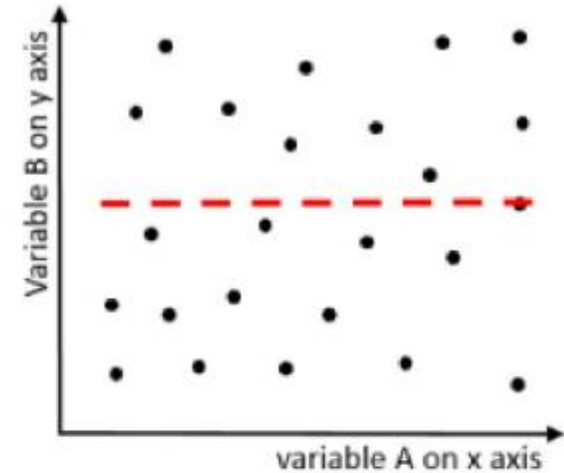
Positive correlation



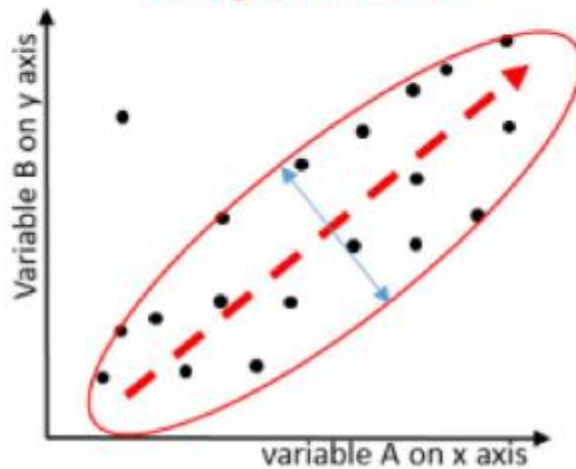
Negative correlation



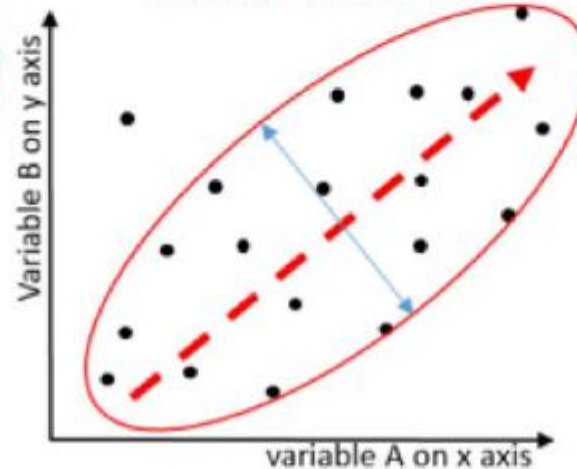
No correlation



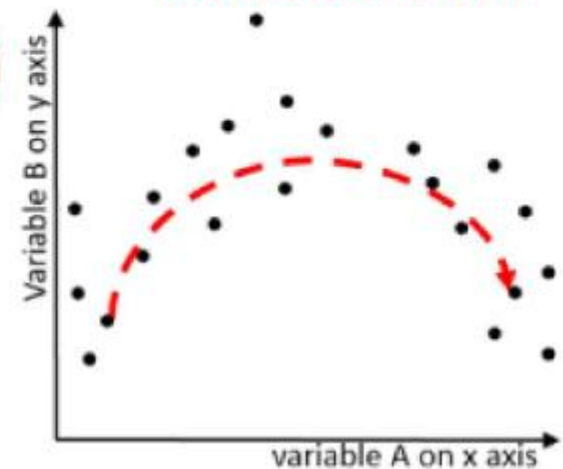
Stronger correlation



Weaker correlation



Non linear correlation



ESTATÍSTICA DESCRITIVA

Correlação: estudo de caso - Target



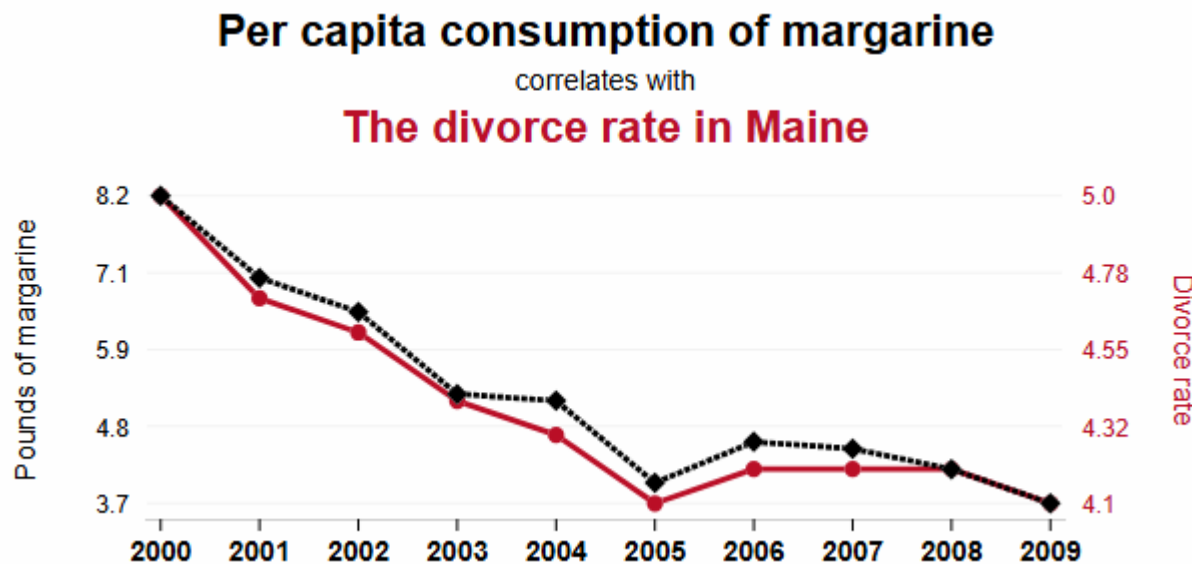
ESTATÍSTICA DESCRITIVA

Correlação: Correlações espúrias

**UMA DAS COISAS MAIS PERIGOSAS
NO MUNDO DOS DADOS!!!!**

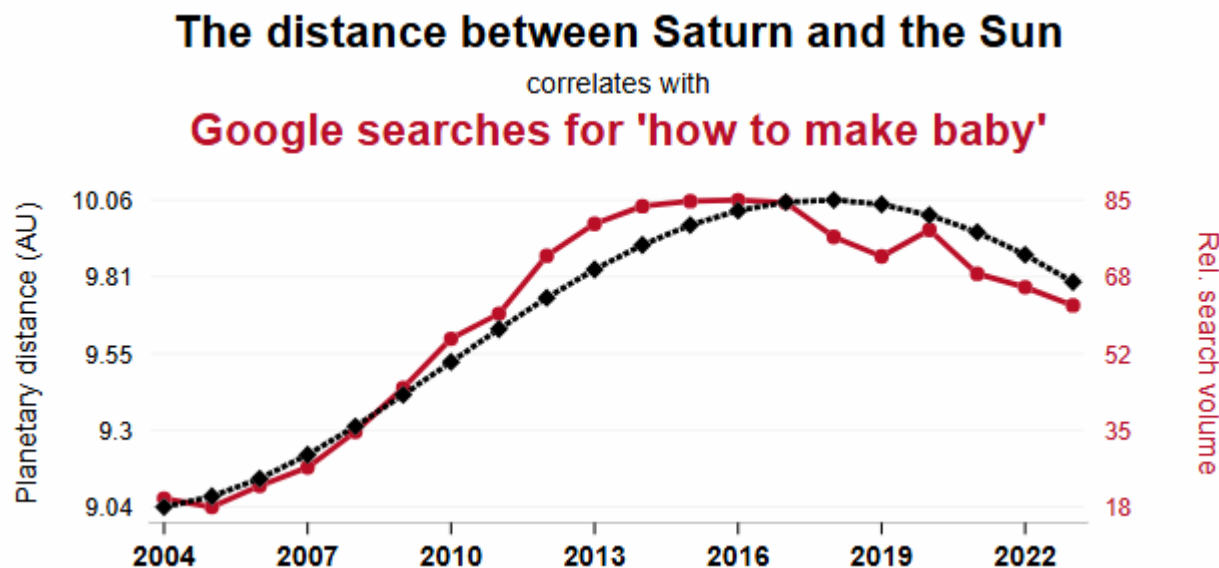
ESTATÍSTICA DESCRITIVA

Correlação: Correlações espúrias



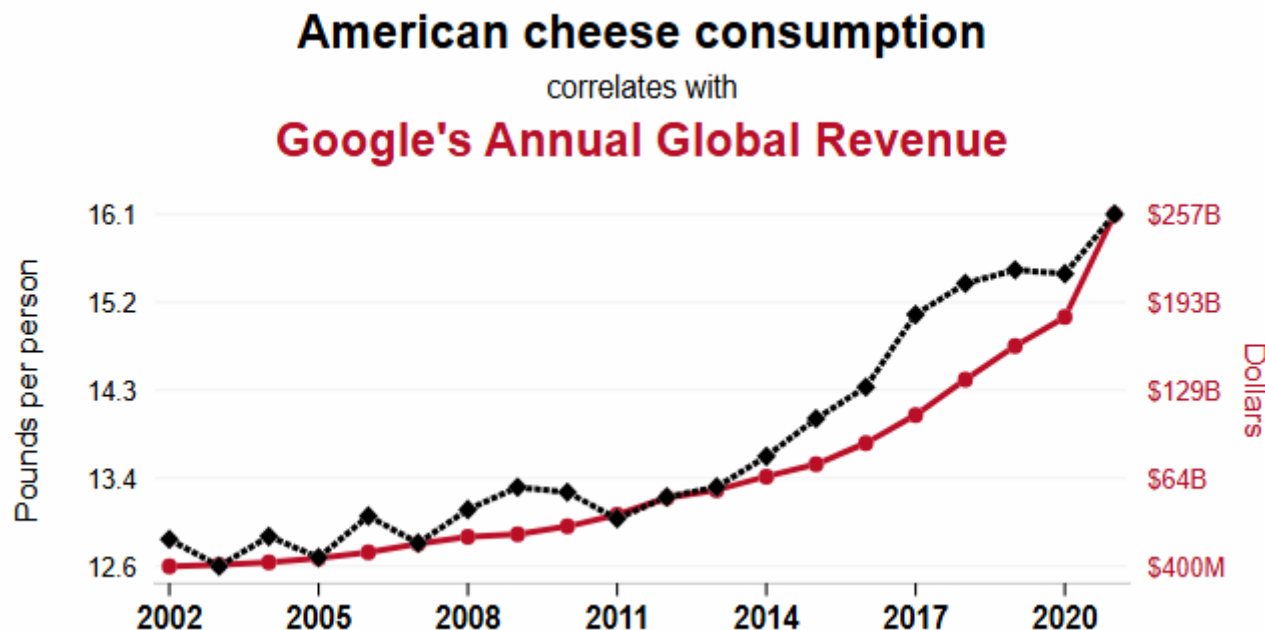
ESTATÍSTICA DESCRITIVA

Correlação: Correlações espúrias



ESTATÍSTICA DESCRITIVA

Correlação: Correlações espúrias



TYLERVIGEN.COM

about · email me · subscribe

spurious correlations

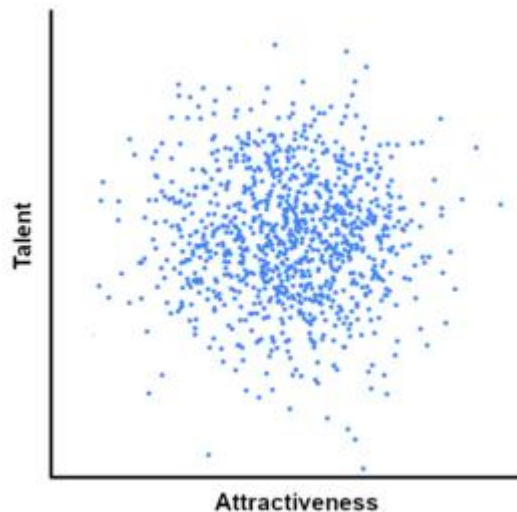
correlation is not causation

[random](#) · [discover](#) · [next page](#) →

don't miss [spurious scholar](#),
where each of these is an academic paper

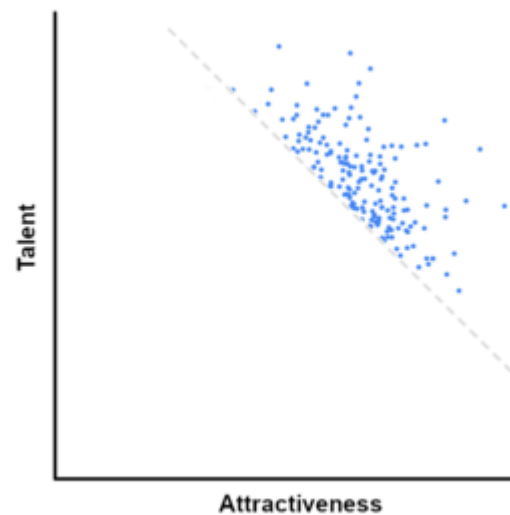
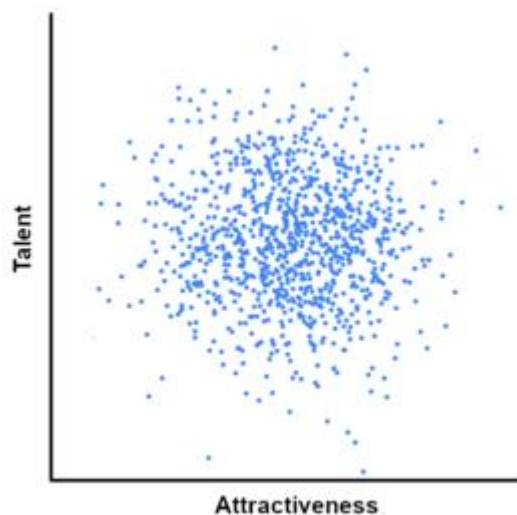
ESTATÍSTICA DESCRITIVA

Correlação: Paradoxo de Berkson



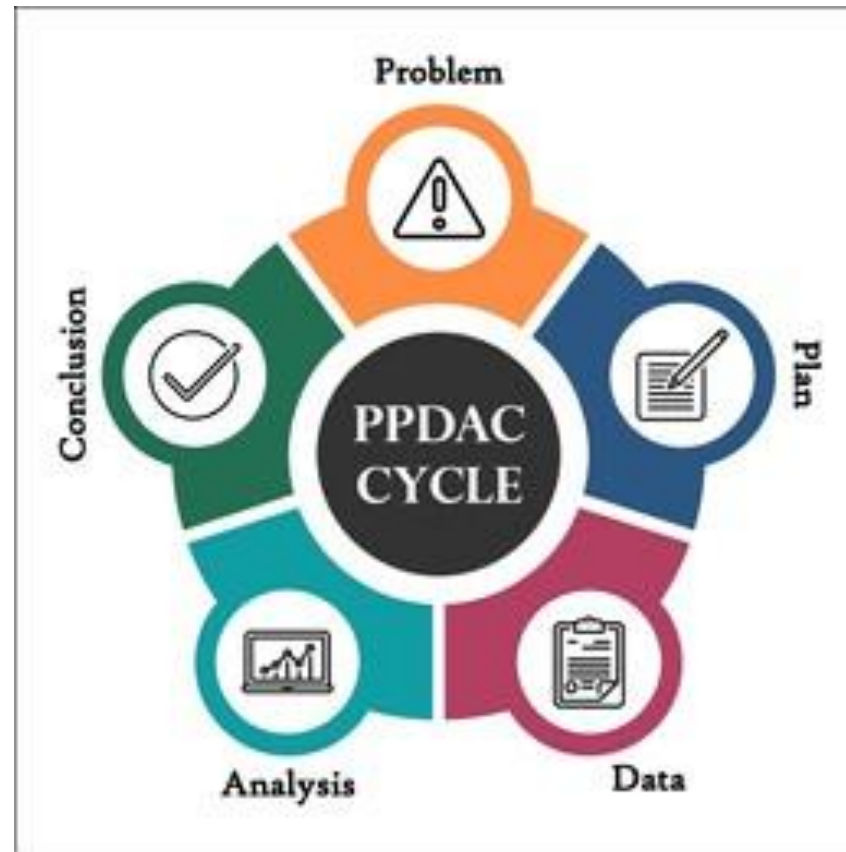
ESTATÍSTICA DESCRITIVA

Correlação: Paradoxo de Berkson



CORRELAÇÃO $\neq$ CAUSALIDADE

Ciclo PPDAC



shutterstock.com • 2247139183

Ciclo PPDAC

P - Problema

- Especificar
- Compreender
- Definir
- Geralmente é uma pergunta

Ciclo PPDAC

P- Plano

- Elaborar um plano para responder a pergunta
- Pode ser simples como contabilizar ocorrências de um fenômeno, ou complexo como uma coleta estratificada de dados

Ciclo PPDAC

D - Dados

- Coletar dados
- Gerenciar
- Limpar
- Organizar

INTRODUÇÃO À ESTATÍSTICA

Ciclo PPDAC

A - Análise

- Seleção dos dados
- Construção de tabelas e gráficos
- Busca por padrões
- Geração de hipóteses (o que pode gerar novas perguntas)

INTRODUÇÃO À ESTATÍSTICA

Ciclo PPDAC

C - Conclusão

- Construção de narrativa
- Interpretação
- Comunicação
- Novas ideias

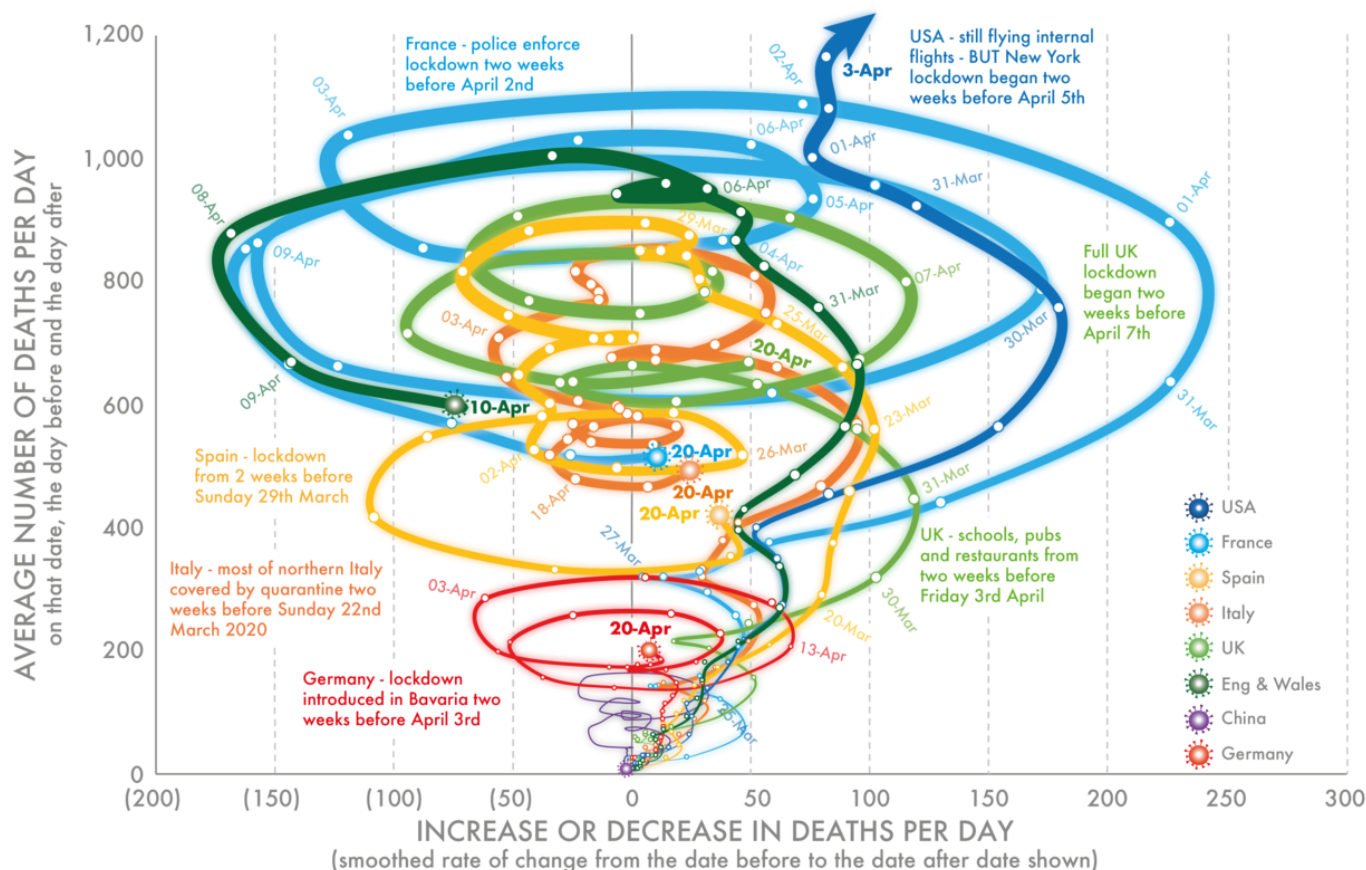
* Ciclo extraído do livro A Arte da Estatística – David Spiegelhalter

Dicas finais

- Porcentagens – enquadramento positivo e negativo
 - 95% de dar certo ou 5% de dar errado?
 - Variação de 250mg/L de DQO é muito ou é pouco?

Dicas finais

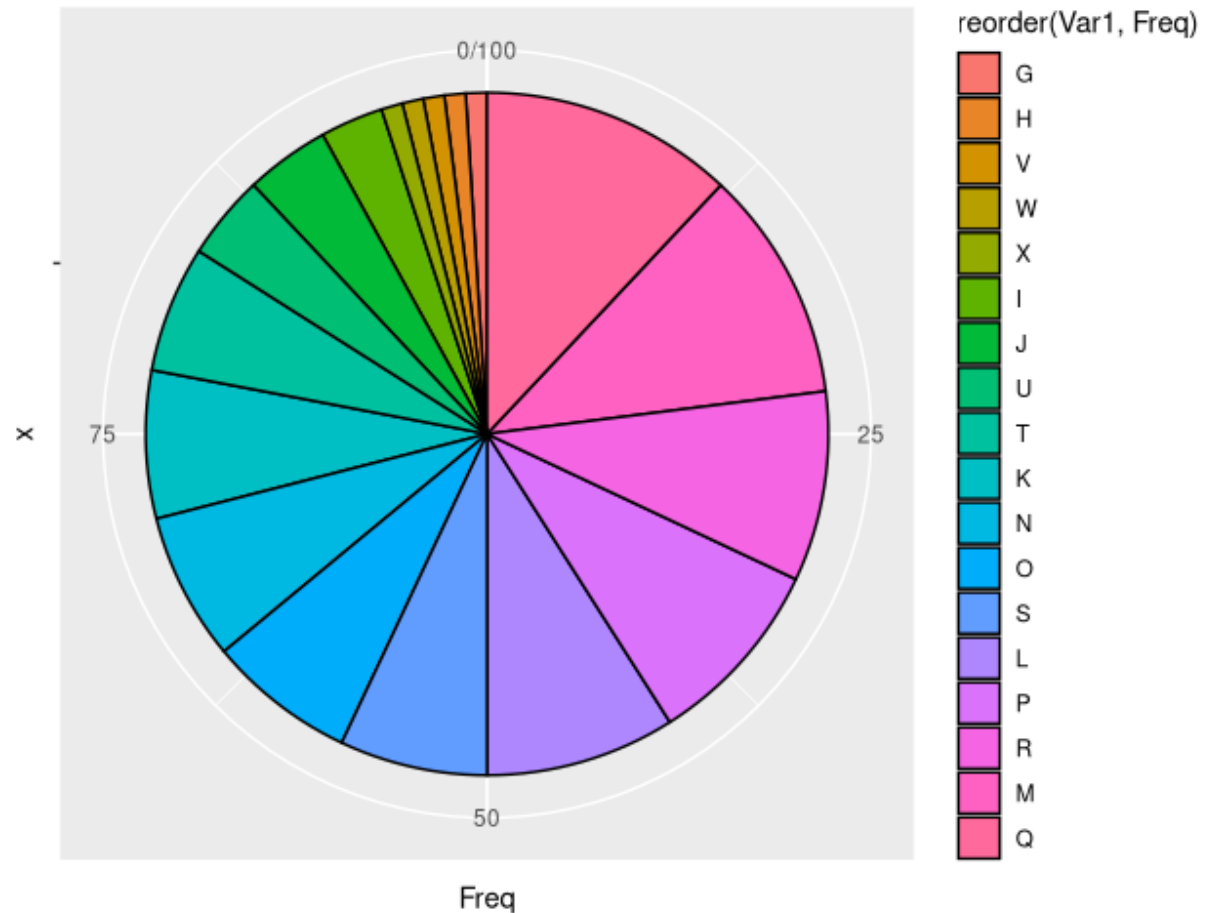
- Gráficos...



DannyDorling.org. Illustration by Kirsten McClure @orpheuscat

Dicas finais

- Gráficos...



Dicas finais

- Gráficos...



**ESTATÍSTICA É UMA ARTE
ANALISAR DADOS É POESIA
LER O MUNDO COM BASE EM DADOS E ANÁLISES
ESTATÍSTICAS, É ELEVAÇÃO ESPIRITUAL.**

SOLUÇÕES EDUCACIONAIS



Presenciais



A Distância



Customizadas



Lato e Stricto
Sensu

 **FaceEsesp**
esesp.es.gov.br