

Análise exploratória de dados

Parte 3

Prof.: Eduardo Vargas Ferreira



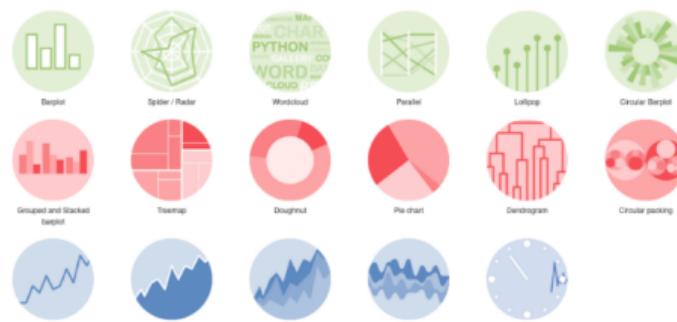
Gráficos

Introdução

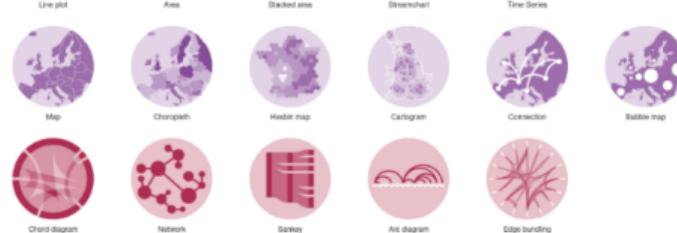
Distribuição: caracterizam a distribuição variáveis qualitativas ou quantitativas.



Relação: exibem a forma da relação ou contraste entre variáveis.



Ordenação: mostram a ordenação de categorias/elementos.



Evolução: exploram a evolução cronológica de variáveis.

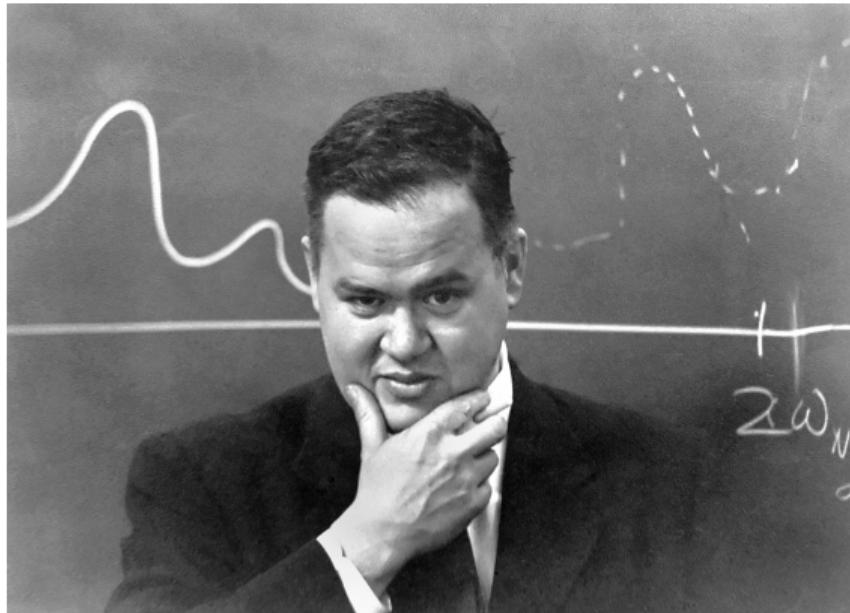
Mapas: representam dados com propriedades geográficas.

Fluxo: descrevem o movimento ou conexão entre categorias/elementos.

Imagens extraídas de www.r-graph-gallery.com

Os gráficos são essenciais

“O maior valor de um gráfico é quando ele nos força a perceber o que nós jamais esperávamos ver.” **(John Tukey · Estatístico)**



Gráficos para variáveis qualitativas

Gráficos para variáveis qualitativas

Gráfico de Barras



Gráfico de Setores

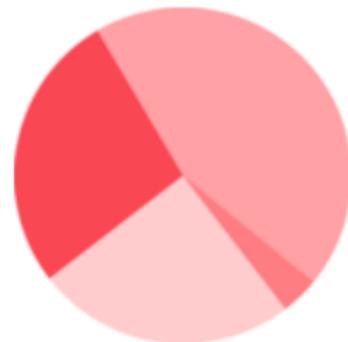
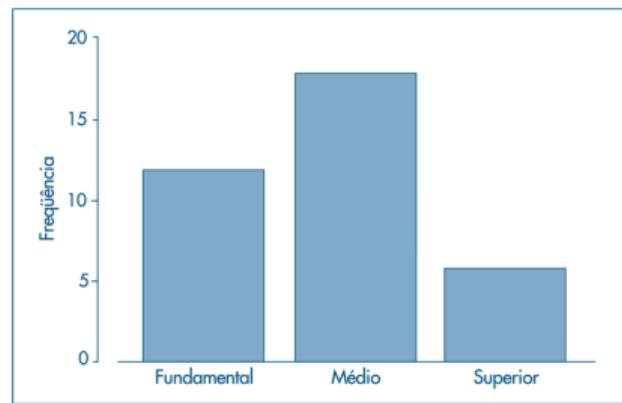


Gráfico de Barras

- É um gráfico formado por retângulos de larguras iguais, em que cada um deles representa a intensidade de um atributo;

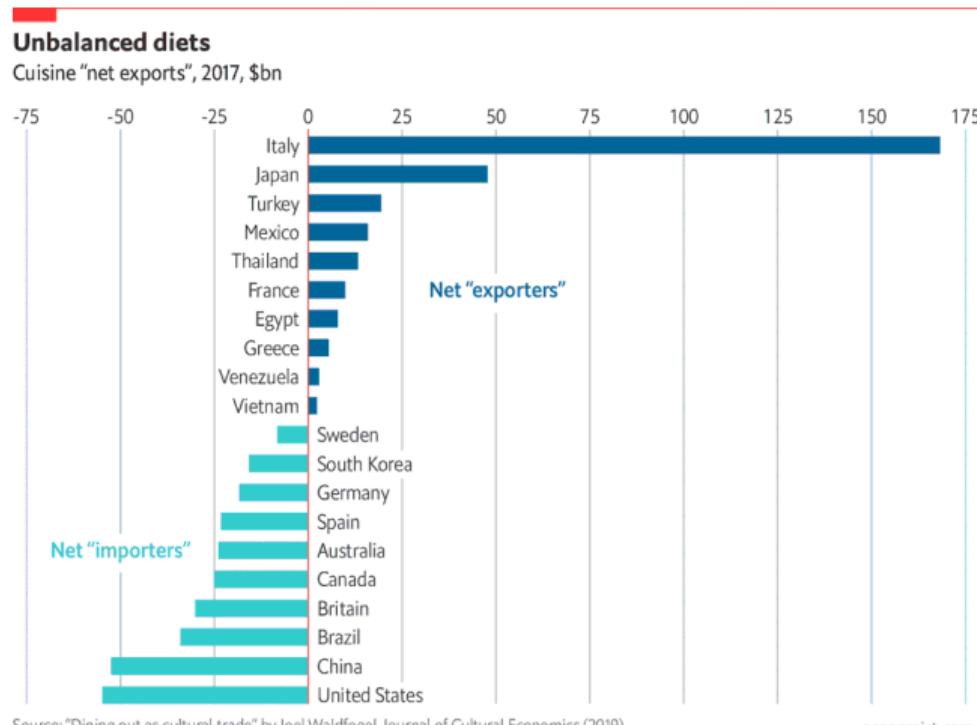
Grau de instrução	Frequência n_i	Proporção f_i	Porcentagem $100f_i$
Fundamental	12	0.33	33%
Médio	18	0.50	50%
Superior	6	0.17	17%
Total	36	1.00	100%



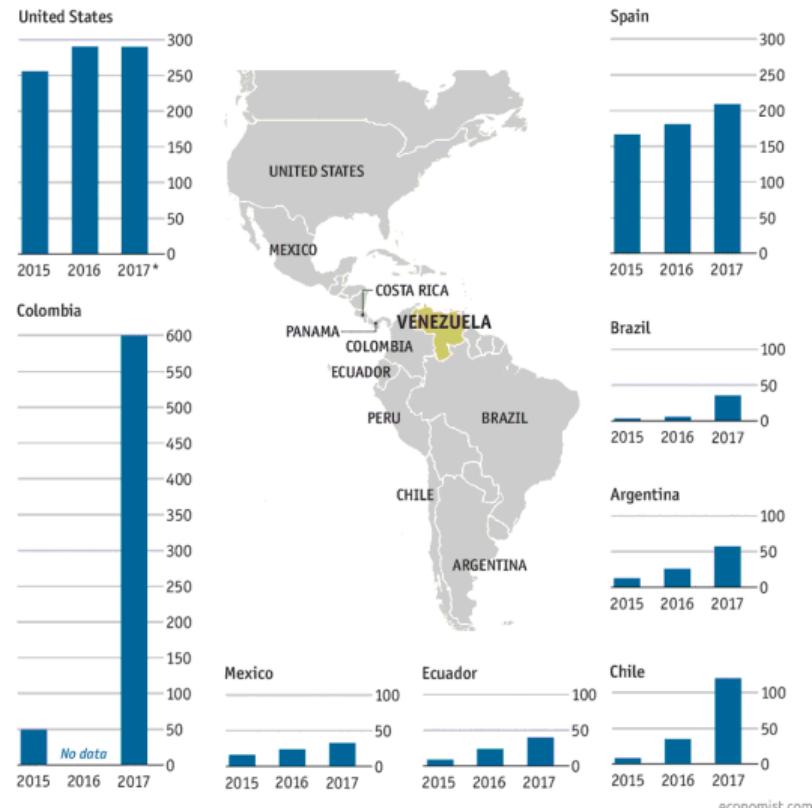
Fonte: Estatística Básica (Bussab e Morettin, 2017)

- As barras são dispostas paralelamente umas às outras, horizontal ou verticalmente.

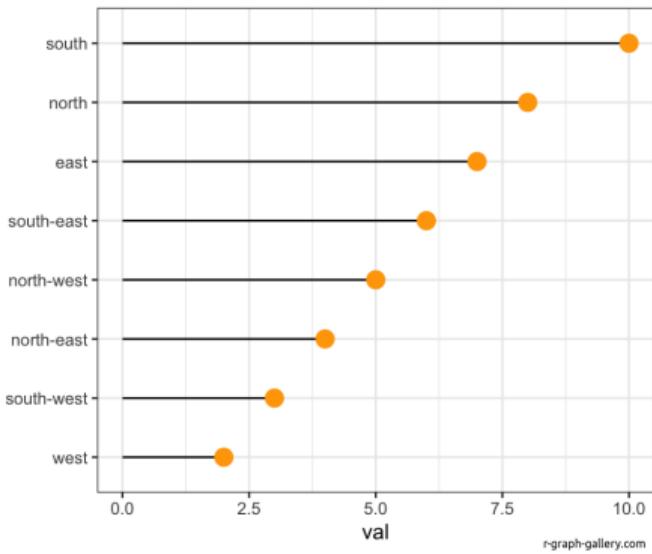
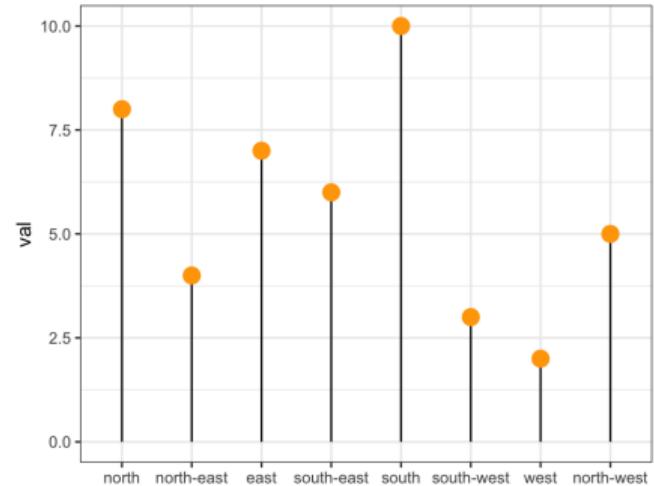
Exemplo: quais países têm maior influência no paladar do mundo?



Exemplo: êxodo dos venezuelanos



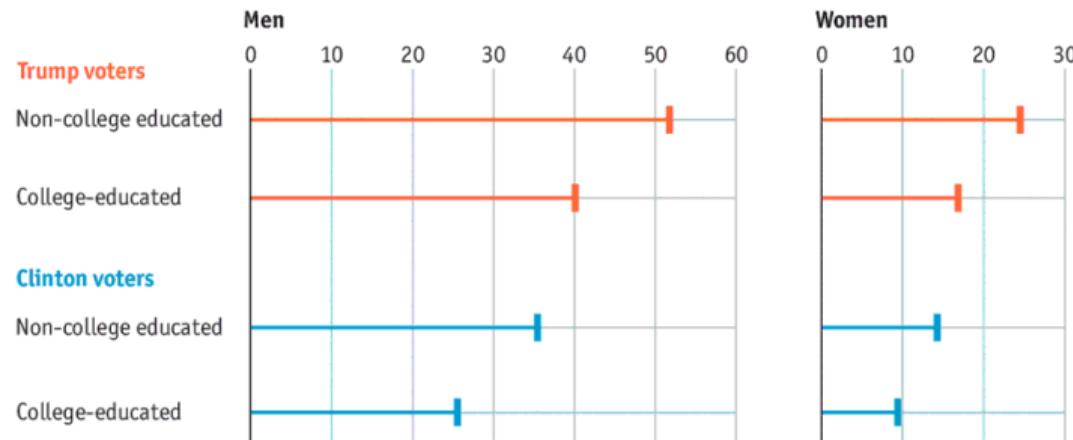
Lollipop plot



Exemplo: intenção de voto

Slurred speech

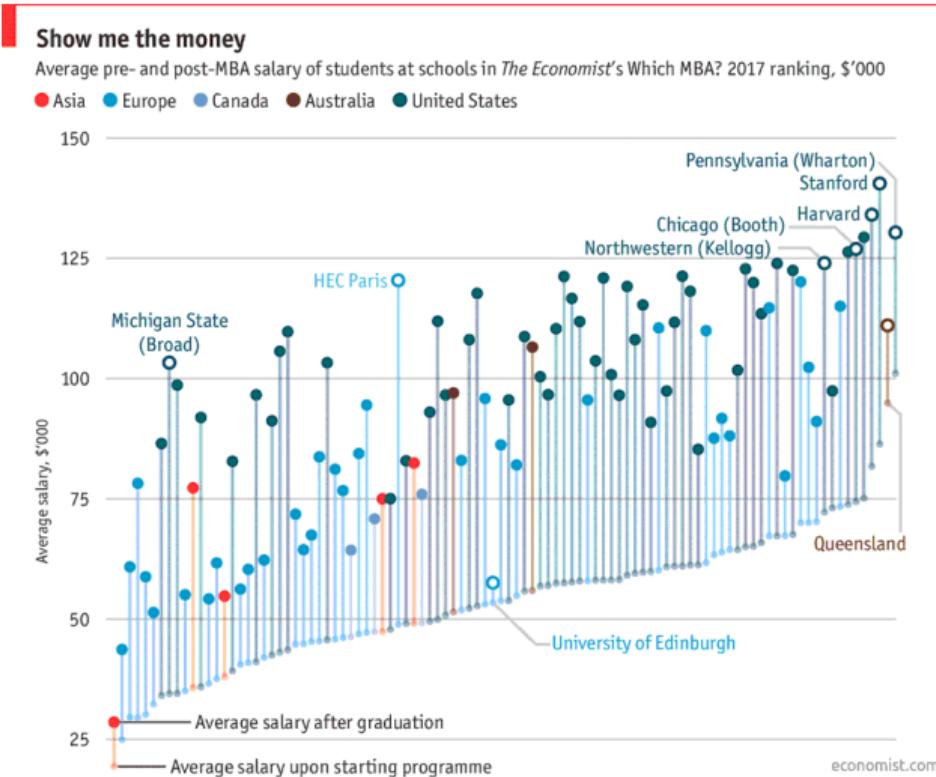
United States, probability of saying the N-word, %
August 19th–21st 2018



Source: *The Economist*/YouGov

economist.com

Exemplo: MBA's que dão o maior impulso aos salários



Devemos atrair a atenção da audiência para o que queremos



Devemos atrair a atenção da audiência para o que queremos



Exemplo: eleições na Venezuela



Devemos atrair a atenção da audiência para o que queremos



Devemos atrair a atenção da audiência para o que queremos

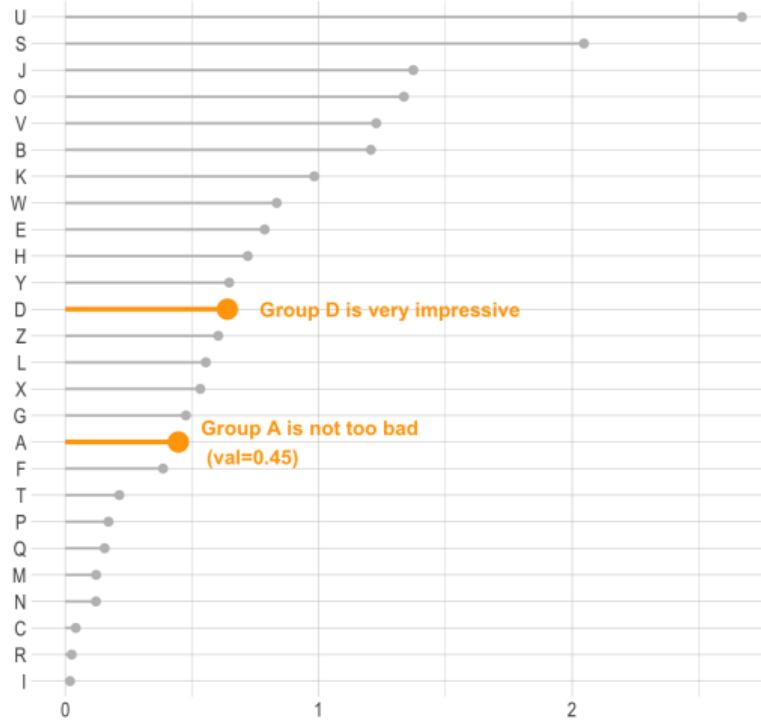


Devemos atrair a atenção da audiência para o que queremos



Lollipop plot

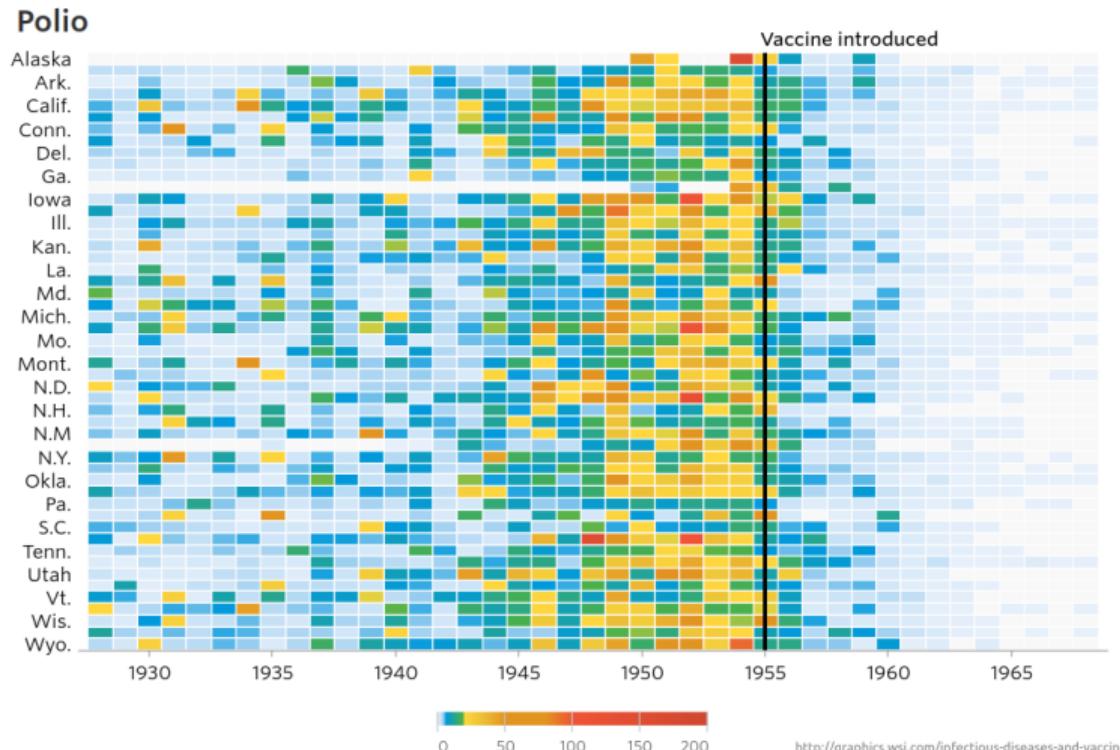
How did groups A and D perform?



Exemplo: gasta-se consertando estradas antigas, mas não constrói novas



Exemplo: incidência de doença entre estados durante os anos



Exemplo: nuvem de palavras dos comentários de um filme



Gráficos para variáveis qualitativas

Gráfico de Barras



Gráfico de Setores

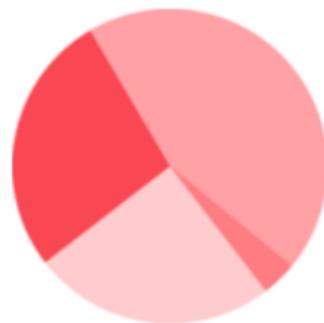
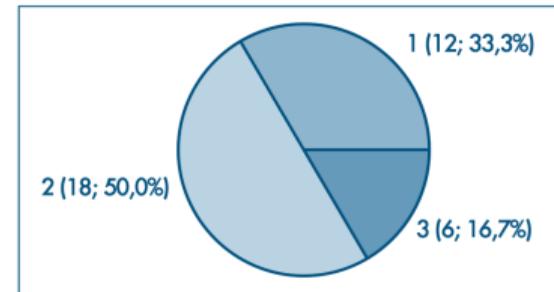


Gráfico de Setores

- O gráfico de setores (ou gráfico de pizza) destina-se a representar a composição, usualmente em porcentagem, de partes de um todo.

Grau de instrução	Frequência n_i	Proporção f_i	Porcentagem $100f_i$
Fundamental	12	0.33	33%
Médio	18	0.50	50%
Superior	6	0.17	17%
Total	36	1.00	100%

Figura: Gráfico de setores para a variável: grau de instrução.



1 = Fundamental, 2 = Médio e 3 = Superior

Fonte: Estatística Básica (Bussab e Morettin, 2017)

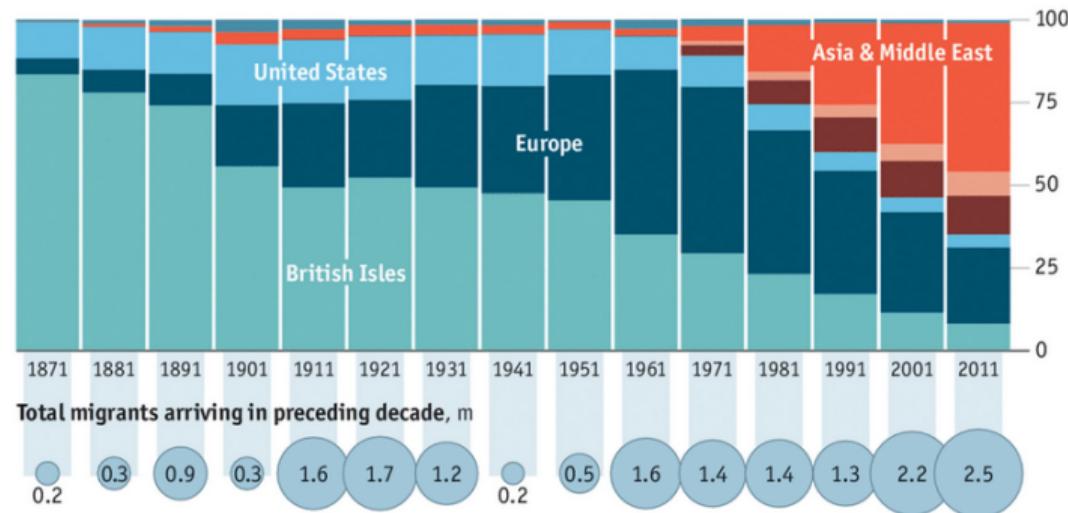
- Consiste num círculo de raio arbitrário, representando o todo, dividido em setores, que correspondem às partes de maneira proporcional.

Exemplo: de onde vêm os imigrantes do Canadá?

The long hello

Canada, foreign-born population, by place of birth, %

Latin America & Caribbean
Africa
Oceania



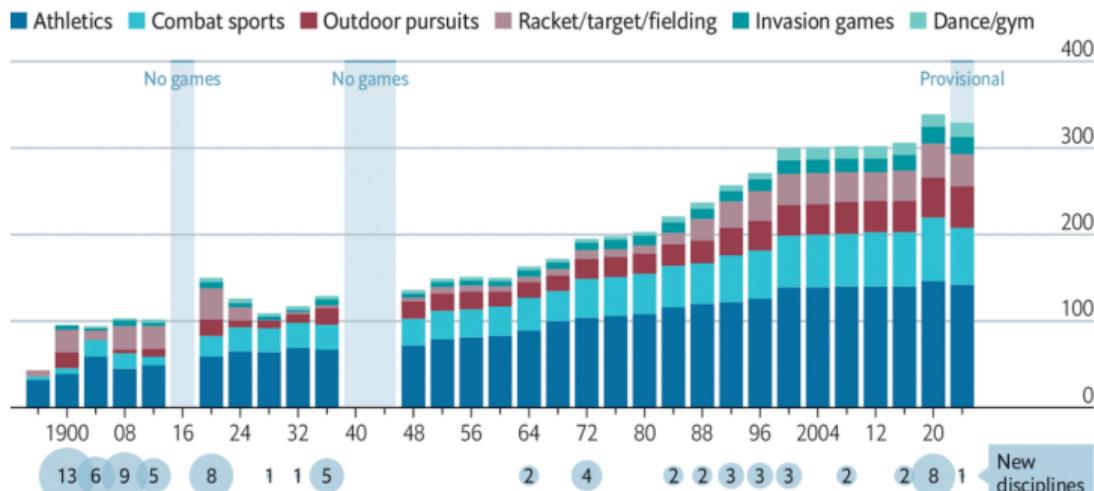
Source: Statistics Canada

Economist.com

Exemplo: olimpíadas maiores e mais diversificadas

Citius, altius, fortius

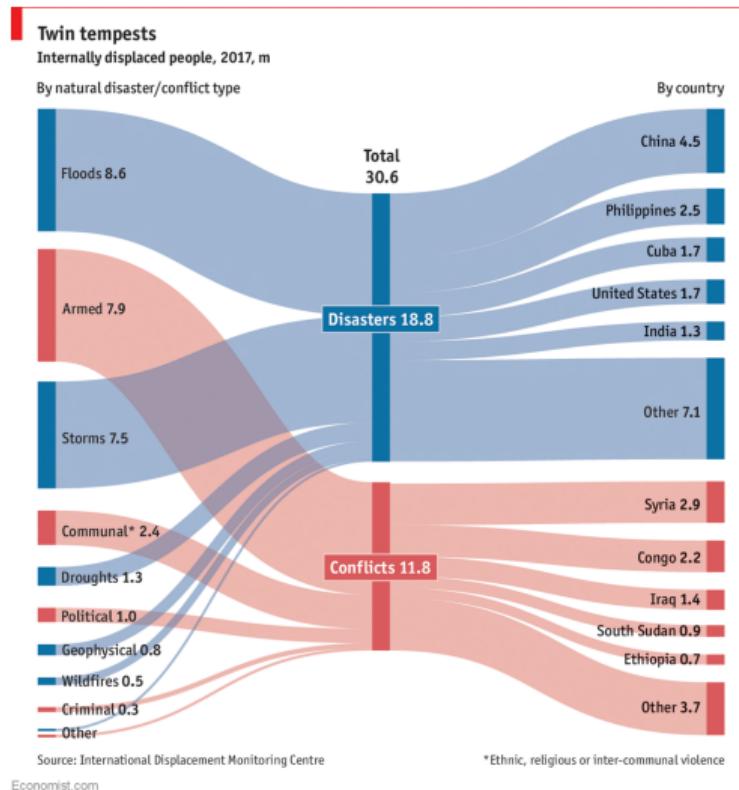
Number of events at summer Olympic games



Source: International Olympic Committee

economist.com

Exemplo: fluxo de imigração entre os continentes



Referências

- Bussab, WO; Morettin, PA. Estatística Básica. São Paulo: Editora Saraiva, 2006 (5^a Edição).
- Magalhães, MN; Lima, ACP. Noções de Probabilidade e Estatística. São Paulo: EDUSP, 2008.

