

# Almanaque de gráficos para visualização de dados

## Taxonomia, funções & formas

Prof. Walmes Zeviani

walmes@ufpr.br

Laboratório de Estatística e Geoinformação

Departamento de Estatística

Universidade Federal do Paraná

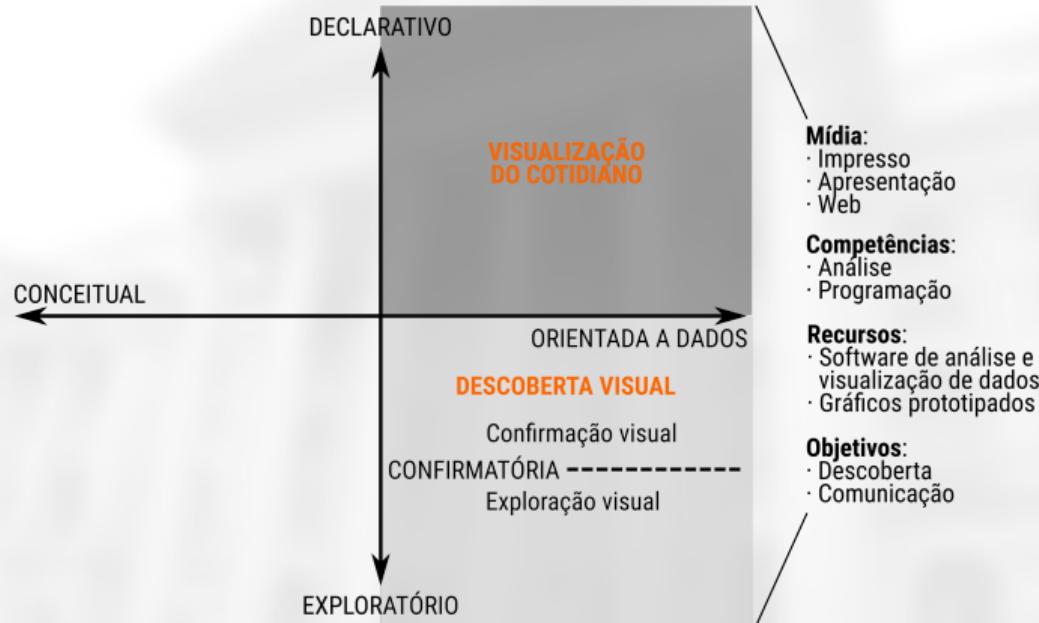
26 de maio de 2020

# Motivação

# Por que conhecer os tipos de gráficos?

- ▶ Conhecer os tipos de gráficos mais usados.
- ▶ Compreender os aspectos físicos, funcionais e de forma dos gráficos.
- ▶ Desenvolver senso crítico para escolha de representações.
- ▶ Nutrir a criatividade para desenvolver novas visualizações.

# Tipos de visualização



**Figura 1.** Aspectos da visualização orientada a dados. Fonte: o autor.

# A escolha do gráfico

A escolha de um gráfico depende de uma série de elementos. Os principais são:

- ▶ Escala das variáveis.
- ▶ Propósito da visualização → função do gráfico.
- ▶ Questões visuais/estéticas → forma do gráfico.

# Os 4 tipos de escala



# Taxonomia

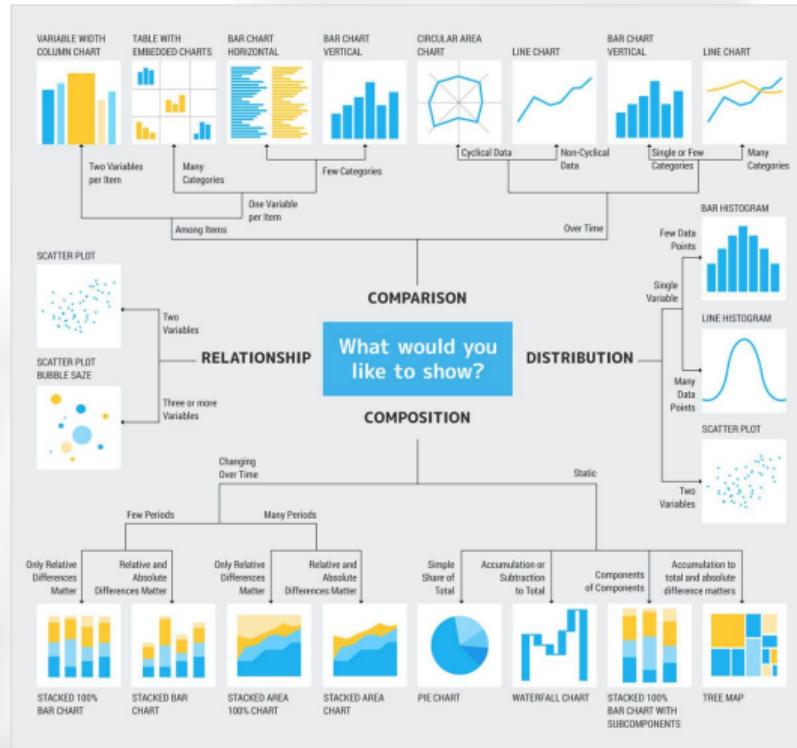


Figura 2. Taxonomia de gráficos baseada na função por Andrew Abela. Visite: <<https://extremepresentation.com/>>.

# Taxonomia

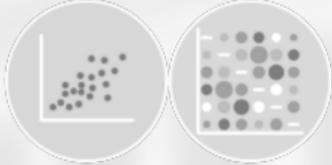
- ▶ Agrupamento dos gráficos baseado nas características comuns.
- ▶ **Função**
  - ▶ Classificação principal: função do gráfico.
  - ▶ Relacionado às quantidades a serem representadas.
  - ▶ Quais quantidades/propriedades serão exibidas → **o que?**
- ▶ **Forma**
  - ▶ Classificação secundária: forma do gráfico.
  - ▶ Relacionado às decisões estéticas e de mídia.
  - ▶ Como as quantidades/propriedades serão exibidas → **como?**

# Função de um gráfico

Distribuição



Relação



Ordenação



Evolução



Mapa



Organização



Fluxo



# Função de um gráfico

**Distribuição** Caracterizam a **distribuição ou divisão** de variáveis qualitativas ou quantitativas.

**Relação** Exibem a forma da **relação ou contraste** entre variáveis.

**Ordenação** Priorizam exibir a **classificação** de categorias/elementos baseada nas variáveis.

**Evolução** Exploram o comportamento de dados **cronológicos**.

**Mapa** Representam dados com propriedades **geográficas**.

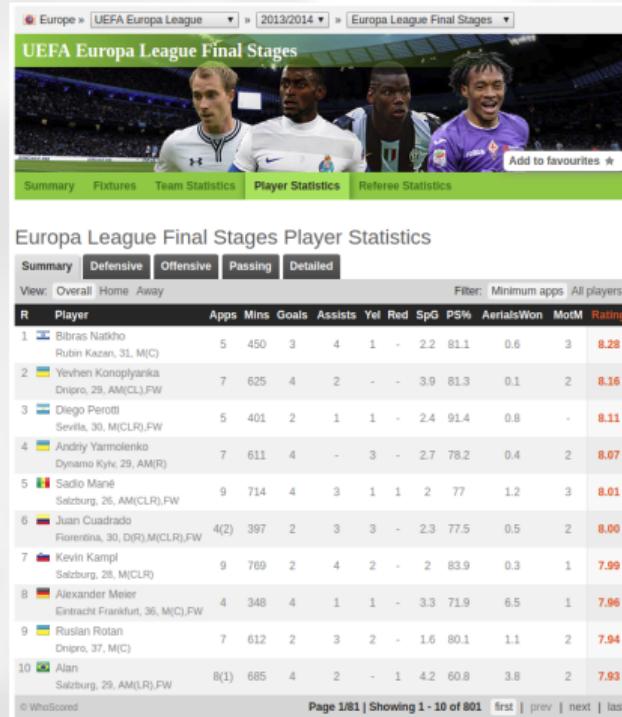
**Organização** Retratam a **estrutura** de organização de coleções de categorias/elementos.

**Fluxo** Descrevem o **movimento ou conexão** entre categorias/elementos.

# Distribuição

# Dataset

- Dados biométricos e esportivos de jogadores de futebol da liga européia (*Europe UEFA Europa League 2013/2014*).
- Extraídos do [www.whoscored.com](http://www.whoscored.com).



The screenshot shows a web browser displaying the 'Europa League Final Stages' section of the WhoScored website. The main title is 'UEFA Europa League Final Stages' with a background image of four football players. Below the title, there are tabs for 'Summary', 'Fixtures', 'Team Statistics', 'Player Statistics' (which is selected), and 'Referee Statistics'. The page displays a table of player statistics for the Europa League Final Stages. The table has columns for Rank (R), Player, Apps, Mins, Goals, Assists, Yel, Red, SpG, PS%, AerialsWon, MotM, and Rating. The data includes 10 players from various clubs like Brøndby, Rubin Kazan, Dinamo Zagreb, Sevilla, Dynamo Kyiv, Salzburg, Fiorentina, Salzburg, Eintracht Frankfurt, and Alania. The 'Rating' column shows values such as 8.28, 8.16, 8.11, 8.07, 8.01, 8.00, 7.99, 7.96, 7.94, and 7.93.

R	Player	Apps	Mins	Goals	Assists	Yel	Red	SpG	PS%	AerialsWon	MotM	Rating
1	Brøndby Natkho Rubin Kazan, 31, M(C)	5	450	3	4	1	-	2.2	81.1	0.6	3	8.28
2	Yevhen Konoplyanka Dinamo, 29, AM(CL),FW	7	625	4	2	-	-	3.9	81.3	0.1	2	8.16
3	Diego Perotti Sevilla, 30, M(CL,R),FW	5	401	2	1	1	-	2.4	91.4	0.8	-	8.11
4	Andriy Yarmolenko Dynamo Kyiv, 29, AM(R)	7	611	4	-	3	-	2.7	78.2	0.4	2	8.07
5	Sadio Mané Salzburg, 26, AM(CL),FW	9	714	4	3	1	1	2	77	1.2	3	8.01
6	Juan Cuadrado Fiorentina, 30, Di(R),M(CL),FW	4(2)	397	2	3	3	-	2.3	77.5	0.5	2	8.00
7	Kevin Kampl Salzburg, 28, M(CL)	9	769	2	4	2	-	2	83.9	0.3	1	7.99
8	Alexander Meier Eintracht Frankfurt, 36, M(C),FW	4	348	4	1	1	-	3.3	71.9	6.5	1	7.96
9	Ruslan Rotan Dinamo, 37, M(C)	7	612	2	3	2	-	1.6	80.1	1.1	2	7.94
10	Alan Salzburg, 29, AM(LR),FW	8(1)	685	4	2	-	1	4.2	60.8	3.8	2	7.93

# Dataset

```
## Rows: 1,326
## Columns: 21
## $ country <chr> "Austria", "Austria", "Austria", "Aust...
## $ team      <chr> "Salzburg", "Salzburg", "Salzburg", "S...
## $ name      <chr> "Sadio Mané", "Kevin Kampl", "Alan", "...
## $ pos       <chr> "M(L)", "M(R)", "FW", "D(C)", "M", "M...
## $ age       <dbl> 21, 23, 24, 22, 23, 28, 28, 24, 21, 28...
## $ cm        <dbl> 175, 180, 182, 182, 180, 172, 180, 186...
## $ kg        <dbl> 69, 63, 73, 77, 71, 69, 71, 74, 78, 69...
## $ tit       <int> 9, 9, 8, 9, 1, 7, 7, 8, 8, 8, 8, 3, 3, ...
## $ res       <dbl> 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 2, 0, ...
## $ goal      <dbl> 4, 2, 4, 1, 0, 0, 8, 0, 0, 0, 0, 0, 0, ...
## $ ass       <dbl> 3, 4, 2, NA, NA, 1, 1, NA, NA, NA, 1, ...
## $ yel       <dbl> 1, 2, 0, 4, 0, 3, 1, 1, 3, 0, 3, 1, 1, ...
## $ red       <dbl> 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...
## $ spg       <dbl> 2.0, 2.0, 4.2, 0.9, 0.5, 1.6, 4.1, 0.8...
## $ ps        <dbl> 77.0, 83.9, 60.8, 72.3, 86.3, 79.4, 72...
## $ aw        <dbl> 1.2, 0.3, 3.8, 3.2, 3.0, 0.5, 0.3, 4.0...
## $ mom      <dbl> 3, 1, 2, 1, NA, NA, NA, NA, NA, NA, NA...
## $ rt        <dbl> 7.98, 7.93, 7.91, 7.67, 7.59, 7.55, 7. ...
## $ bmi      <dbl> 22.53061, 19.44444, 22.03840, 23.24598...
## $ games    <dbl> 9, 9, 9, 9, 2, 8, 7, 9, 9, 8, 8, 5, 3, ...
## $ position <fct> M, M, FW, D, M, M, FW, M, D, D, D, ...
```

# Diagrama de ramos e folhas

```
## Número de jogadores: 81
##
## The decimal point is 1 digit(s) to the right of the |
##
## 16 | 5588
## 17 | 000133344
## 17 | 5555556678888889
## 18 | 0000000111122333333333334
## 18 | 555566777777888899
## 19 | 000111114
```

**Figura 3.** Diagrama de ramos e folhas para a altura (cm) dos jogadores ingleses.

# Diagrama de ramos e folhas

- ▶ **Nome:** diagrama de ramos e folhas.
- ▶ **Função:** representar a distribuição dos valores.
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** pontos empilhados.
- ▶ **Requisitos definíveis:** intervalos de classe.
- ▶ **Variáveis calculadas:** frequência absoluta.

# Diagrama de pontos (stacked dot plot)

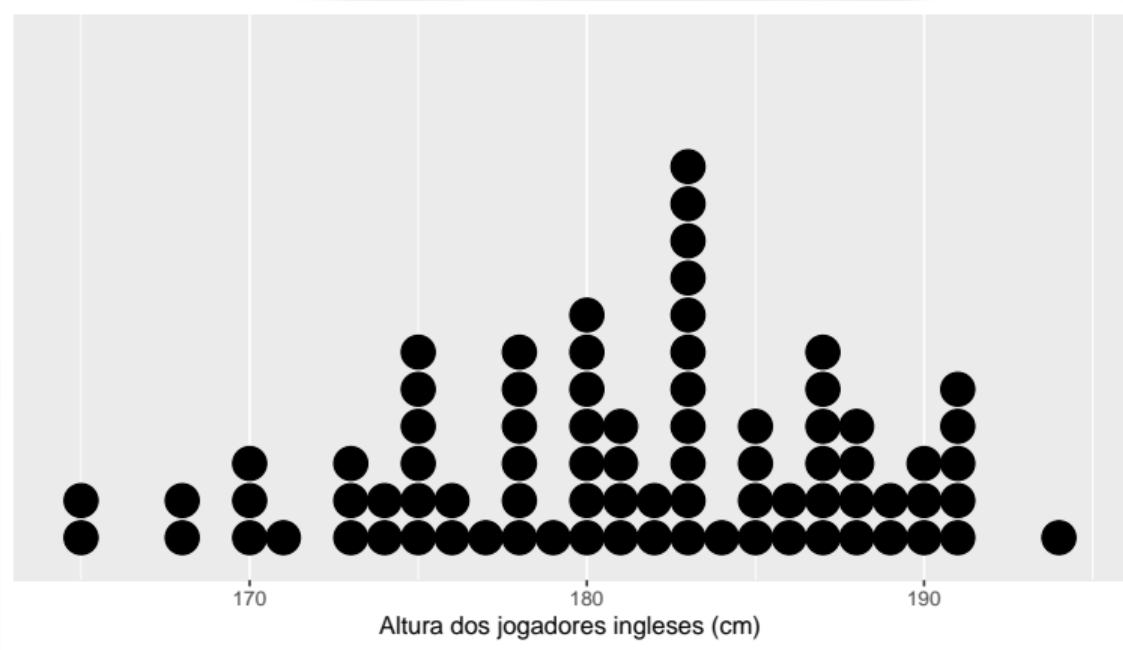


Figura 4. Distribuição da altura dos jogadores ingleses.

# Diagrama de pontos (stacked dot plot)

- ▶ **Nome:** *stacked dot plot*.
- ▶ **Função:** representar a distribuição dos valores.
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** pontos empilhados.
- ▶ **Requisitos definíveis:** intervalos de classe.
- ▶ **Variáveis calculadas:** frequência absoluta.

# Histograma

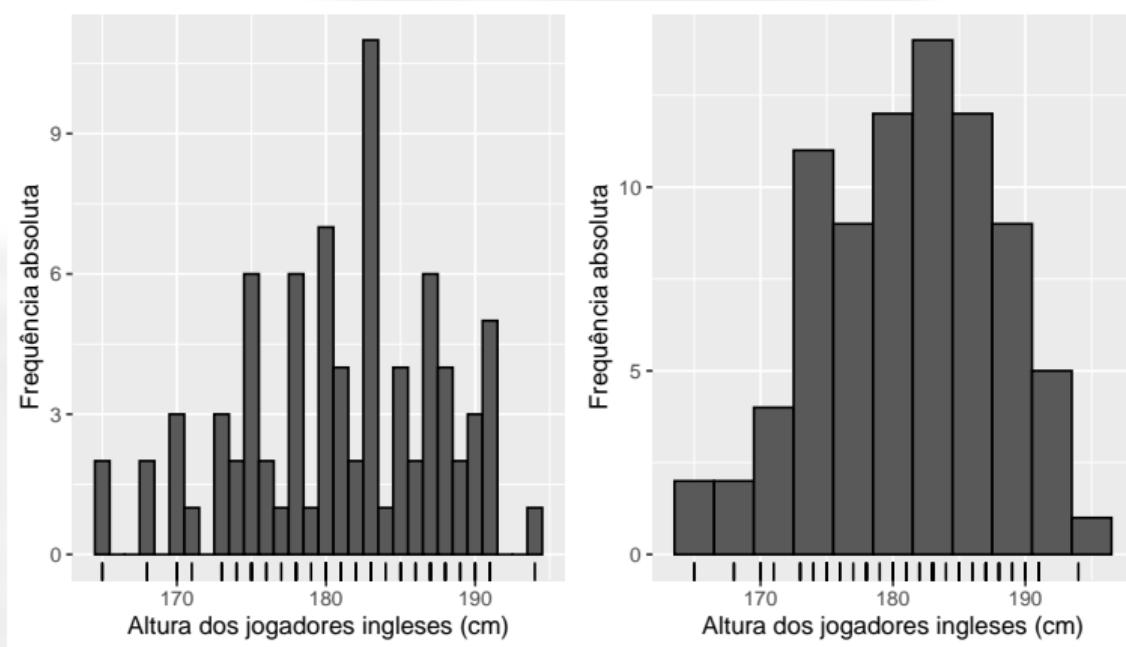
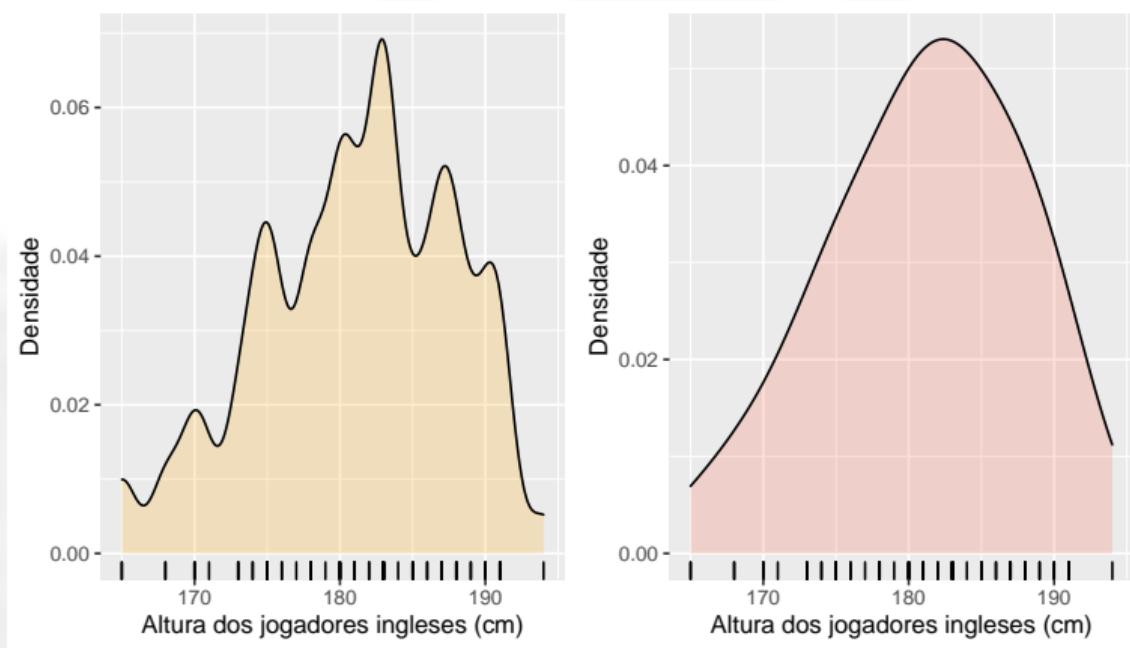


Figura 5. Distribuição da altura dos jogadores ingleses.

# Histograma

- ▶ **Nome:** histograma.
- ▶ **Função:** representar a distribuição dos valores.
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** barras.
- ▶ **Requisitos definíveis:** intervalos de classe.
- ▶ **Variáveis calculadas:** frequência absoluta, relativa ou densidade.

# Densidade



**Figura 6.** Distribuição da altura dos jogadores ingleses.

# Densidade

- ▶ **Nome:** distribuição empírica de densidade.
- ▶ **Função:** representar a distribuição dos valores.
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** linhas.
- ▶ **Requisitos definíveis:** função kernel e largura de banda.
- ▶ **Variáveis calculadas:** densidade empírica.
- ▶ **Aplicação Shiny:** <<http://shiny.leg.ufpr.br/walmes/density/>>.

# Distribuição acumulada

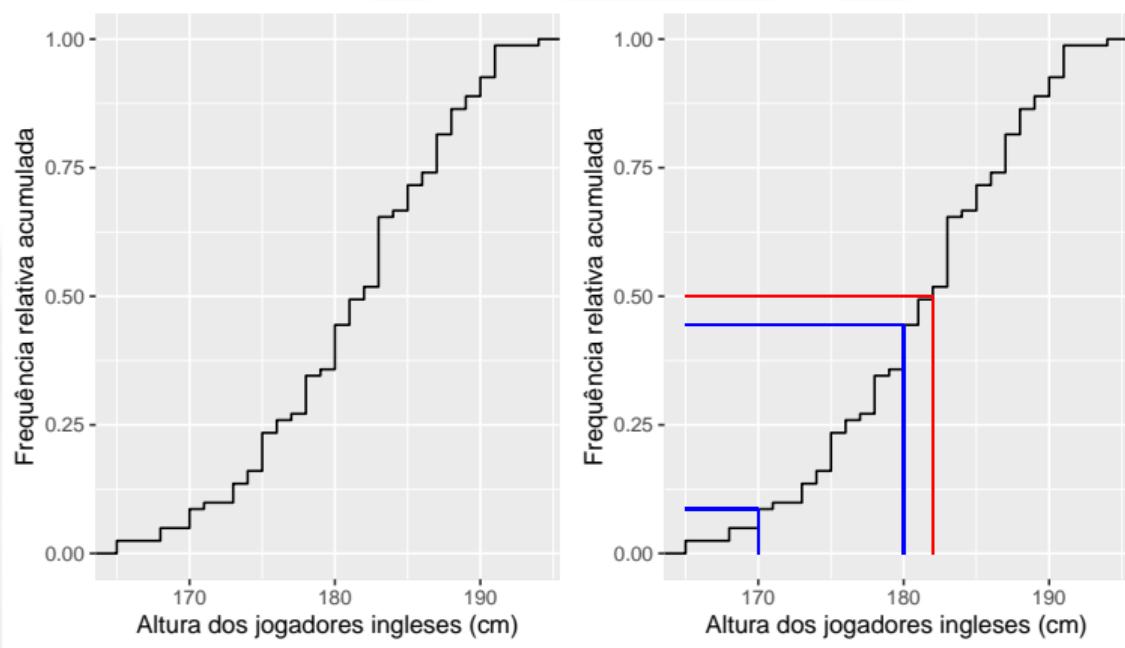


Figura 7. Distribuição da altura dos jogadores ingleses.

# Distribuição acumulada

- ▶ **Nome:** *empirical cumulative distribution function* (ECDF).
- ▶ **Função:** representar a distribuição acumulada dos valores.
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** linhas.
- ▶ **Variáveis calculadas:** frequência relativa acumulada.

# Caixas e bigodes

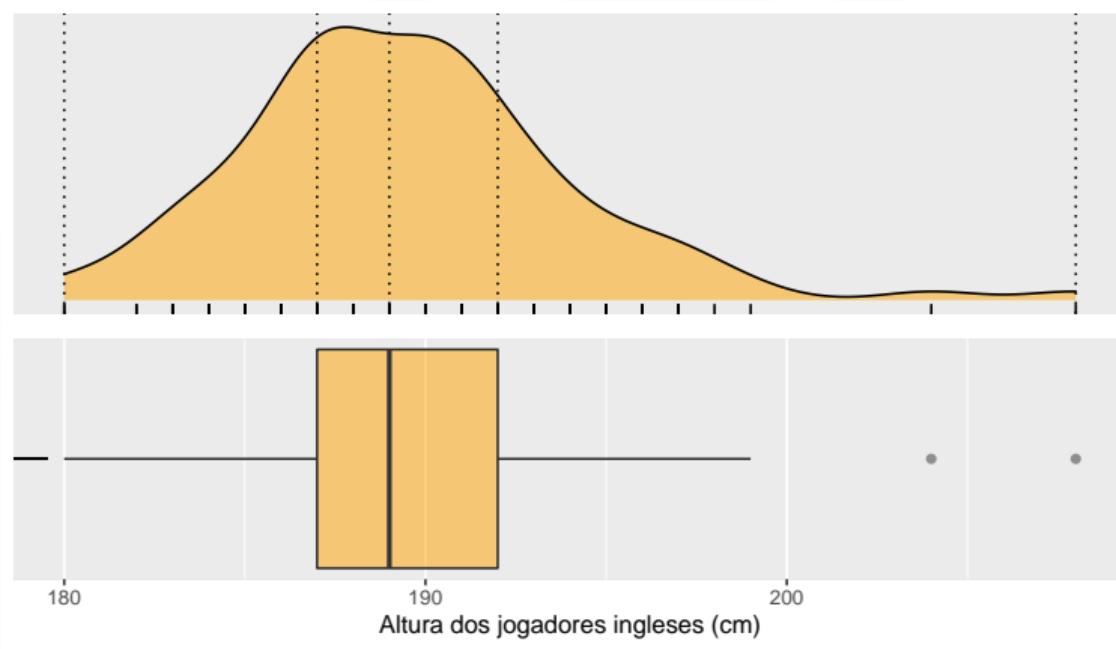


Figura 8. Distribuição da altura dos goleiros.

# Caixas e bigodes

- ▶ **Nome:** diagrama de caixas (e bigodes) (box (and whiskers)).
- ▶ **Função:** representar a distribuição dos valores demarcando as separatrizes.
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** caixas, segmentos e pontos.
- ▶ **Variáveis calculadas:** quartis, extremos, amplitude interquatílica e IC para mediana\*.

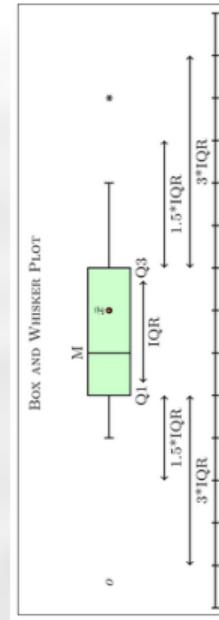


Figura 9. Componentes do gráfico de caixas e bigodes.

# Violin plot

- **Nome:** *violin plot*.
- **Inspiração:** densidade empírica.

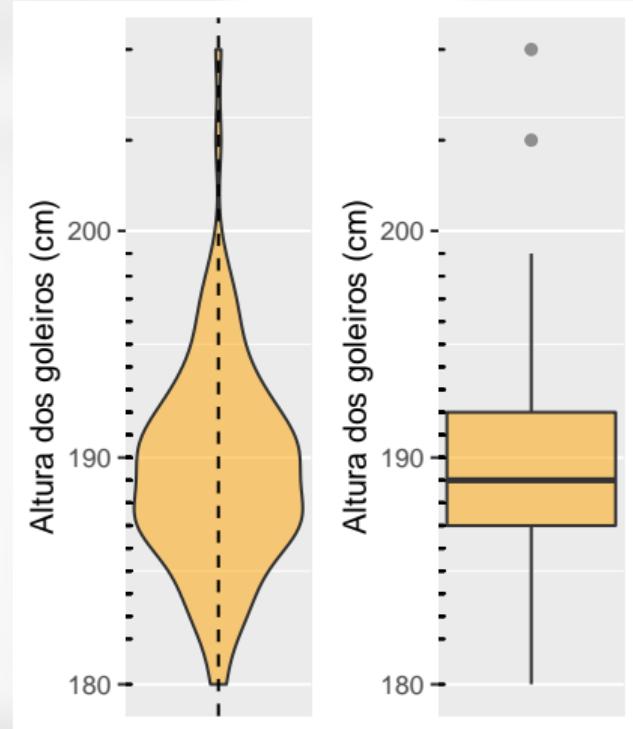


Figura 10. Distribuição da altura dos goleiros.

# Joyplot

- **Nome:** Joyplot ou *ridgeline plots*.
- **Inspiração:** densidade empírica.

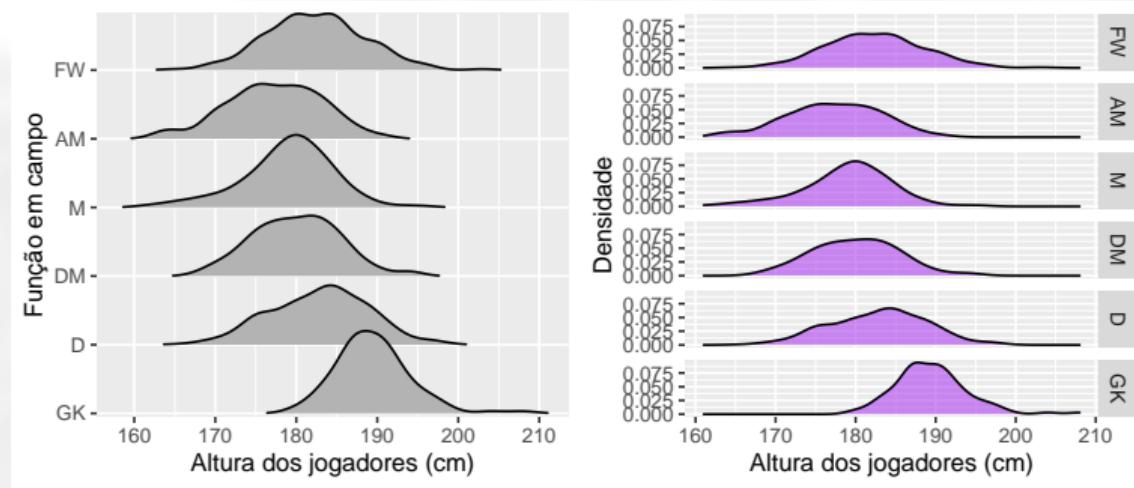


Figura 11. Distribuição da altura dos jogadores por função em campo.

# Comparação

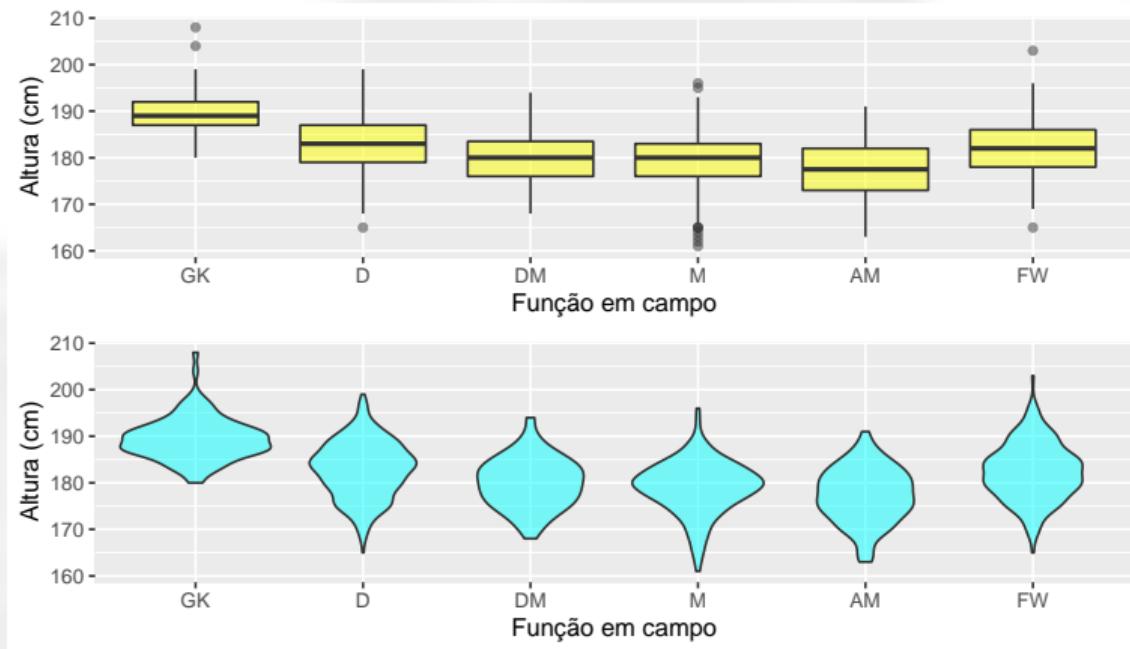


Figura 12. Distribuição da altura dos jogadores por função em campo.

# Recomendações

Qual usar para:

1. Reconhecer forma da distribuição?
2. Examinar medidas de locação e dispersão?
3. Determinar proporções e separatrizes?

# Gráfico de barras

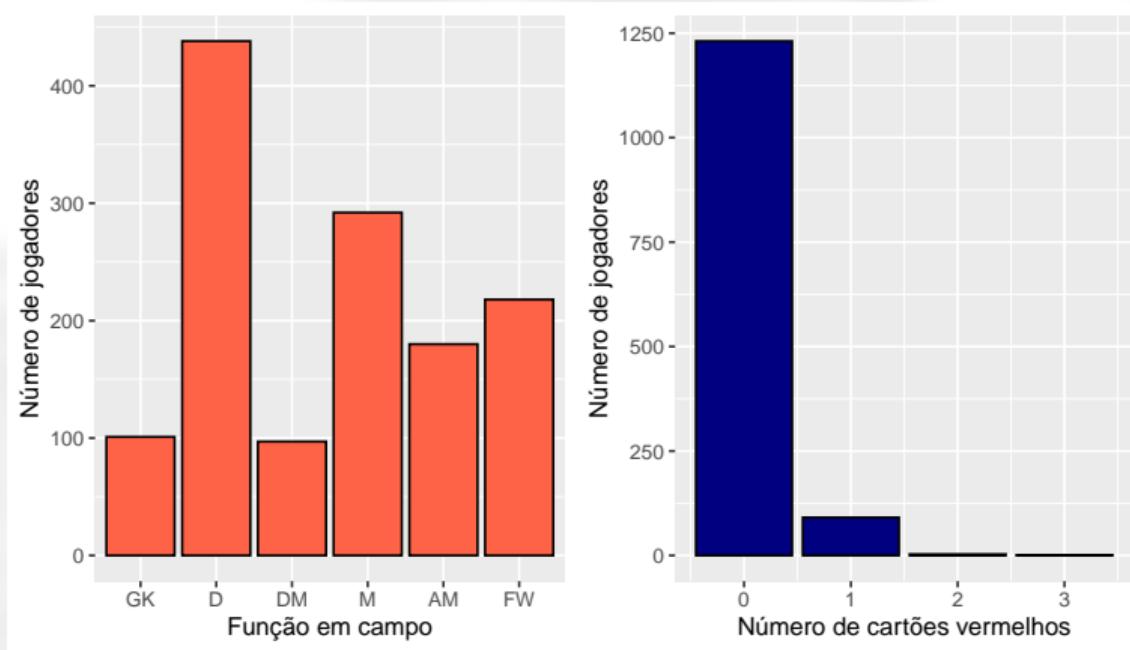


Figura 13. Distribuição dos jogadores quanto a função em campo e número de cartões vermelhos.

# Gráfico de barras

- ▶ **Nome:** Gráfico de barras (*bar chart*).
- ▶ **Função:** exibir a frequência das categorias.
- ▶ **Tipo de escala:** qualitativa.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** barras.
- ▶ **Variáveis calculadas:** frequência absoluta ou relativa.
- ▶ **Variações:**
  - ▶ Barras agrupadas (*grouped bar chart*).
  - ▶ Barras empilhadas (*stacked bar chart*).
  - ▶ Barras normalizadas (*normalized stacked bar chart*).

# Gráfico de setores

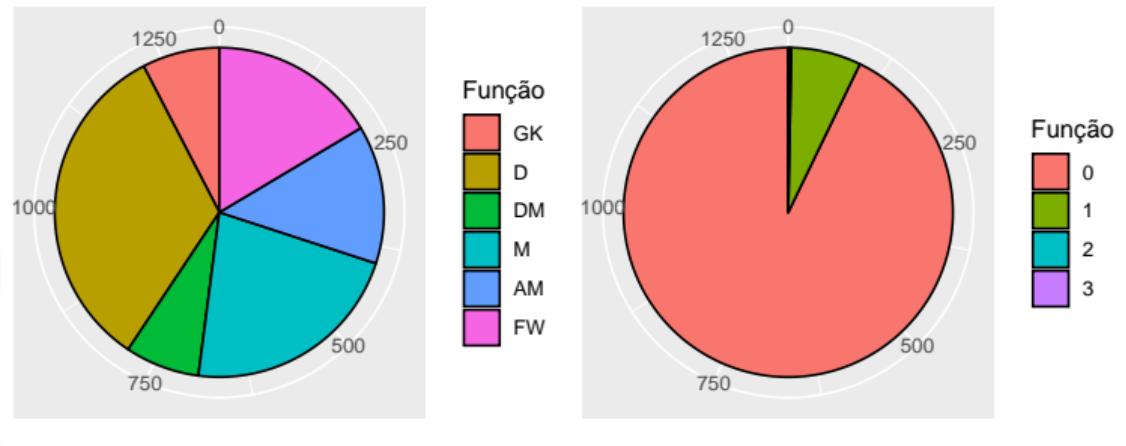


Figura 14. Distribuição dos jogadores quanto a função em campo e número de cartões vermelhos.

# Gráfico de setores

- ▶ **Nome:** Gráfico de setores (*pie chart*).
- ▶ **Função:** exibir a frequência relativa de categorias.
- ▶ **Tipo de escala:** qualitativa.
- ▶ **Variáveis visuais:** ângulo.
- ▶ **Geometria:** setores circulares.
- ▶ **Variáveis calculadas:** frequência relativa.
- ▶ **Variações:** *donut chart* ou *doughnut chart*.

# Gráfico de rosca

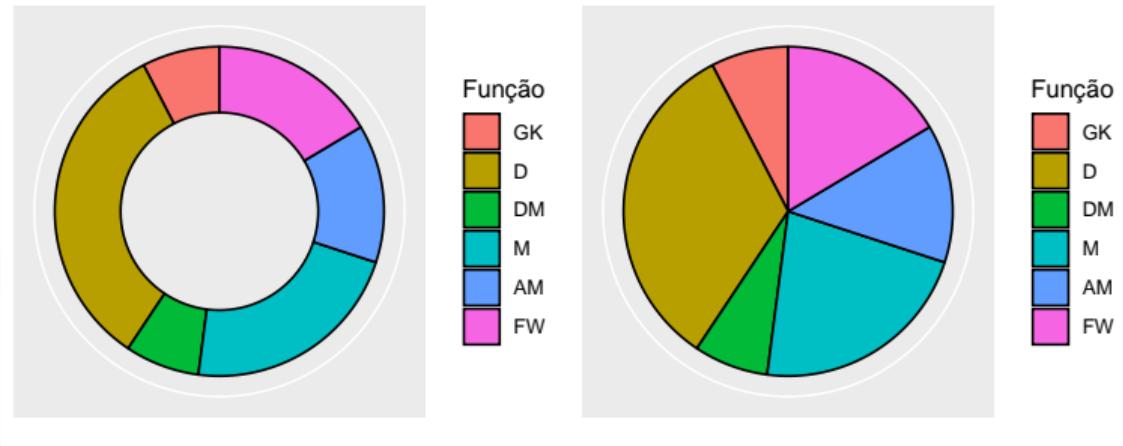
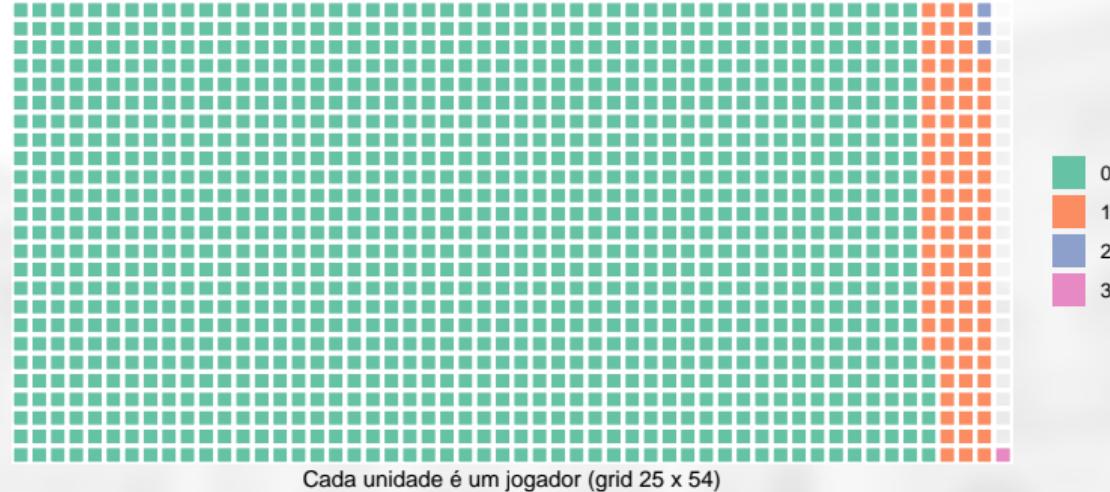


Figura 15. Distribuição dos jogadores quanto a função em campo.

# Gráfico de rosca

- ▶ **Nome:** Gráfico de rosca (*donut chart*).
- ▶ **Inspiração:** gráfico de setores.
- ▶ **Geometria:** arcos.

# Waffle chart



**Figura 16.** Distribuição dos jogadores quanto ao número de cartões vermelhos.

# Waffle chart

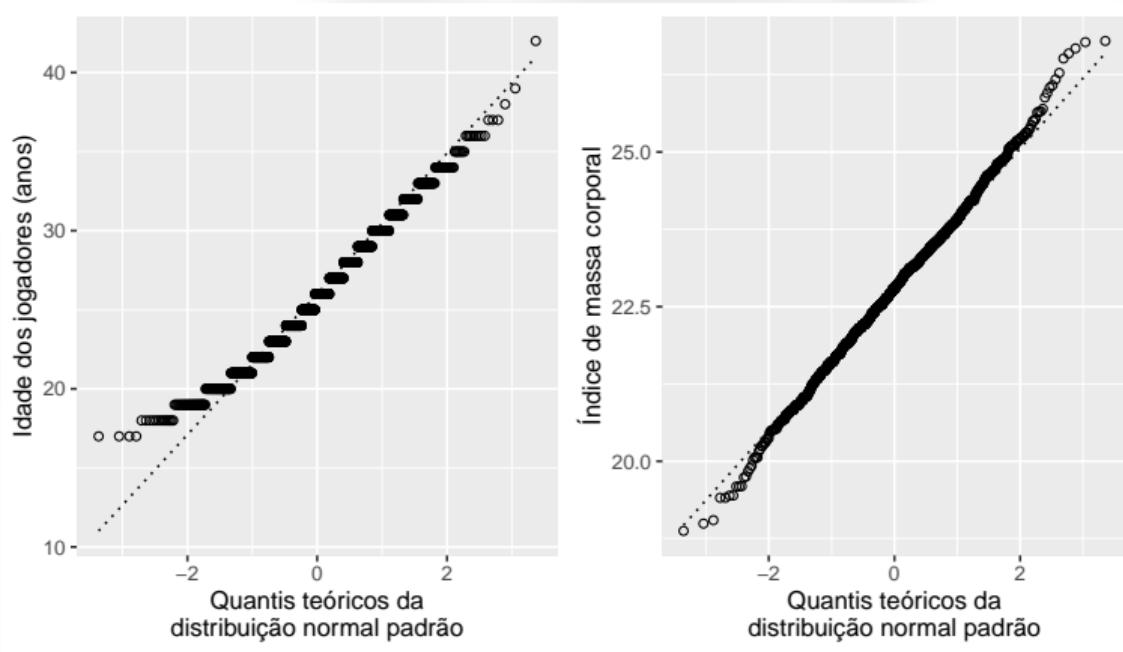
- ▶ **Nome:** *Waffle chart*.
- ▶ **Função:** exibir a frequência de categorias.
- ▶ **Tipo de escala:** qualitativa.
- ▶ **Variáveis visuais:** área e cor.
- ▶ **Geometria:** pixels.
- ▶ **Variáveis calculadas:** frequência absoluta ou relativa.
- ▶ **Variações:** *area size chart*.

# Recomendações

Qual usar para:

1. Reconhecer a distribuição dos valores?
2. Comparar as proporções?
3. Enfatizar categorias de baixa frequência?

# Gráfico quantil-quantil



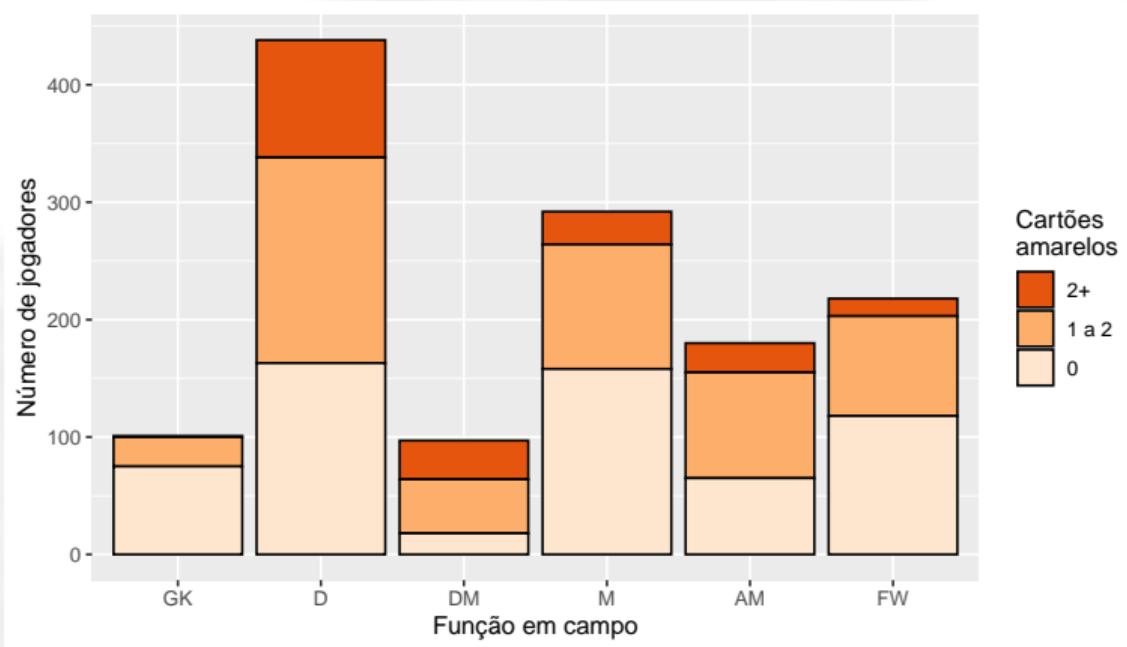
**Figura 17.** Distribuição da idade e índice de massa corporal dos jogadores.

# Gráfico quantil-quantil

- ▶ **Nome:** gráfico quantil-quantil.
- ▶ **Função:** examinar a adequação de uma distribuição de probabilidades.
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** pontos e linhas.
- ▶ **Variáveis calculadas:** quantis teóricos da distribuição sob hipótese.

# Relação

# Gráficos de barras agrupadas e empilhadas



**Figura 18.** Quantidade de cartões amarelos e a relação com a função em campo.

# Gráficos de barras agrupadas e empilhadas

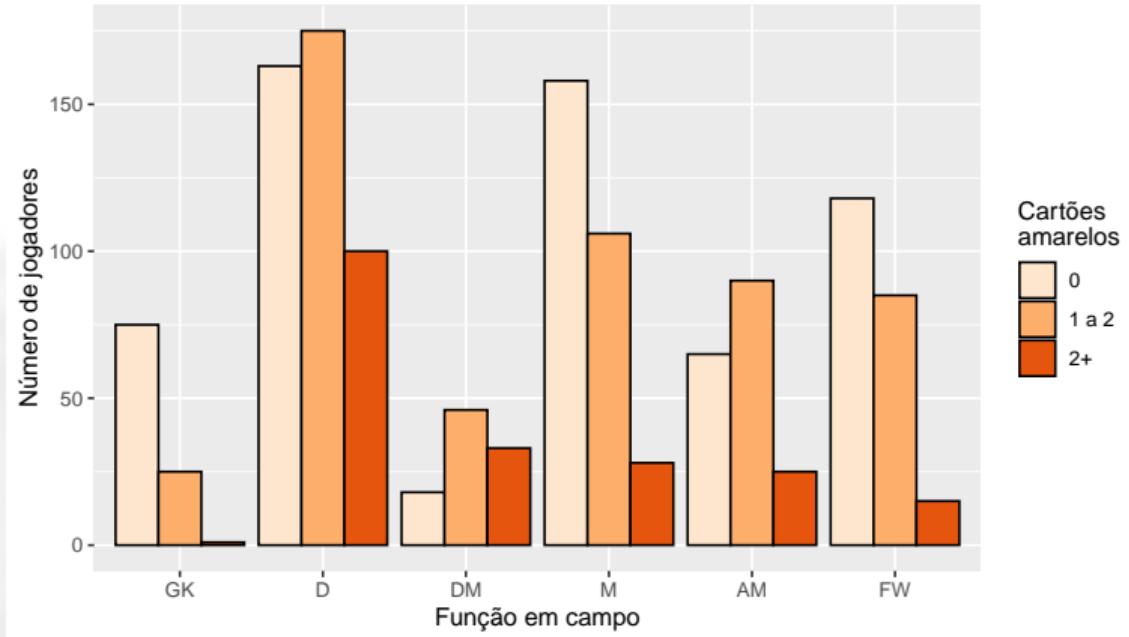


Figura 19. Quantidade de cartões amarelos e a relação com a função em campo.

# Gráficos de barras agrupadas e empilhadas

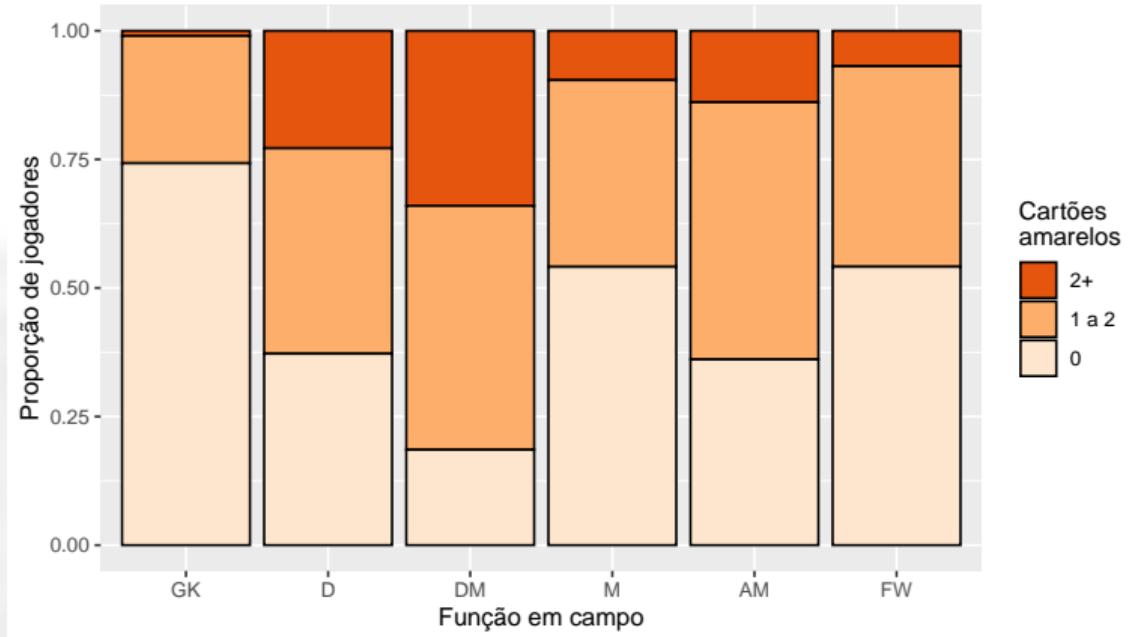
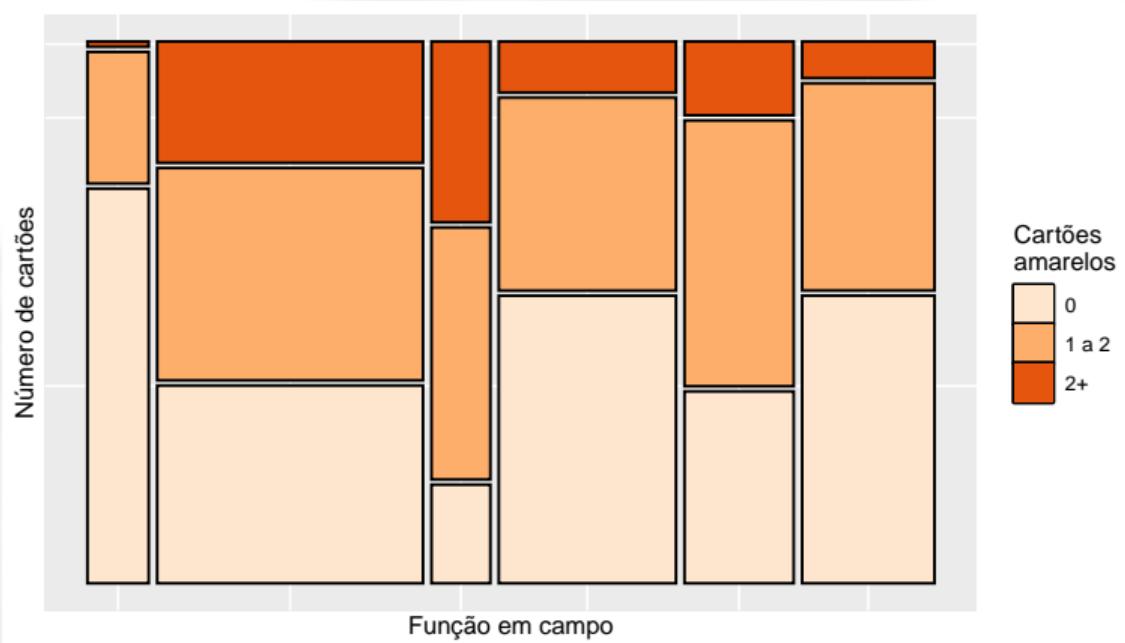


Figura 20. Quantidade de cartões amarelos e a relação com a função em campo.

# Gráficos de barras agrupadas e empilhadas

- ▶ **Nome:** Gráfico de barras (*bar plot*).
- ▶ **Função:** permitir a comparação de quantidades entre categorias.
- ▶ **Tipo de escala:** qualitativa × qualitativas.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** barras.

# Gráfico de mosaico



**Figura 21.** Quantidade de cartões amarelos e a relação com a função em campo.

# Gráfico de mosaico

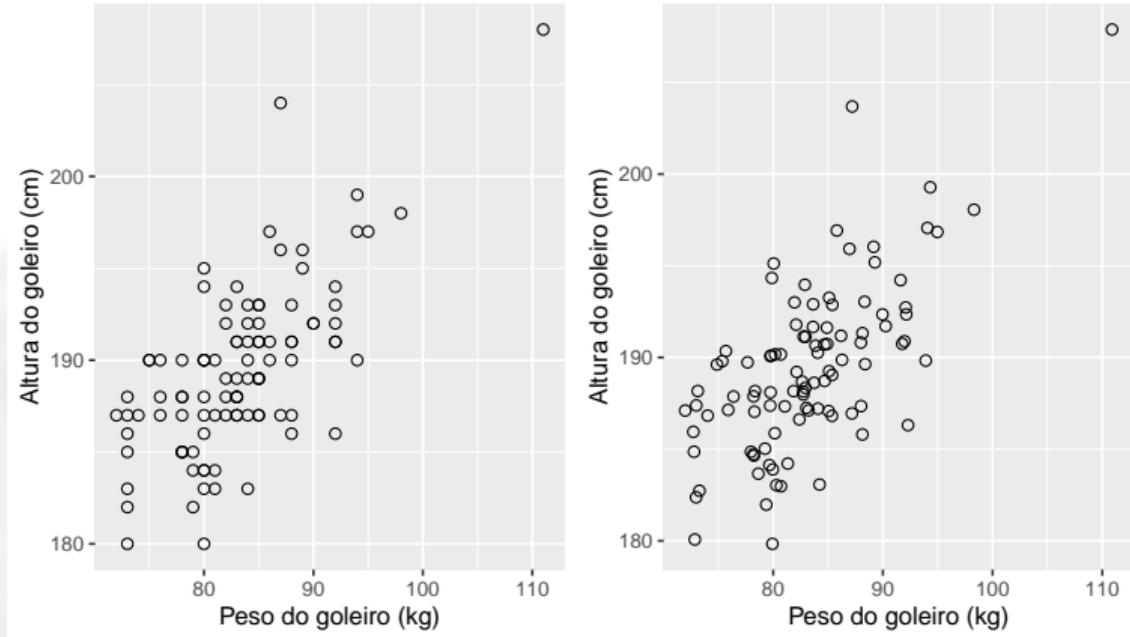
- ▶ **Nome:** Gráfico de mosaico (*mosaic plot* ou *Marimekko chart*).
- ▶ **Função:** exibir a frequência de combinações de categorias.
- ▶ **Tipo de escala:** qualitativa.
- ▶ **Variáveis visuais:** posição  $x \times y$  e área.
- ▶ **Geometria:** retângulos.
- ▶ **Variáveis calculadas:** frequência relativa.

# Recomendações

Qual usar para:

1. Comparar valores totais entre categorias?
2. Comparar valores relativos entre categorias?
3. Examinar a relação entre variáveis categóricas?

# Diagrama de dispersão



**Figura 22.** Relação altura e peso dos goleiros. O gráfico da direita foi adicionada perturbação para evitar sobreposição dos pontos.

# Diagrama de dispersão

- ▶ **Nome:** diagrama de dispersão (*scatter plot*).
- ▶ **Função:** retratar a relação entre variáveis.
- ▶ **Tipo de escala:** quantitativa × quantitativa.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** pontos.
- ▶ **Variações:** diagrama de dispersão 3D, diagrama ternário.

# Diagrama de dispersão

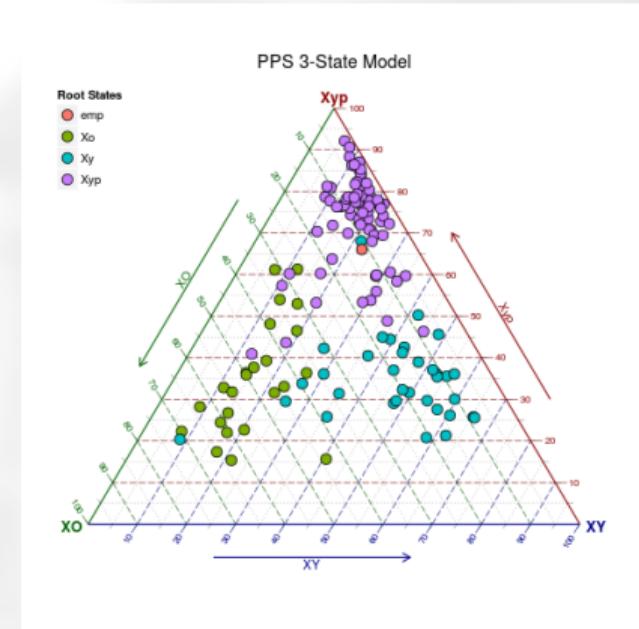
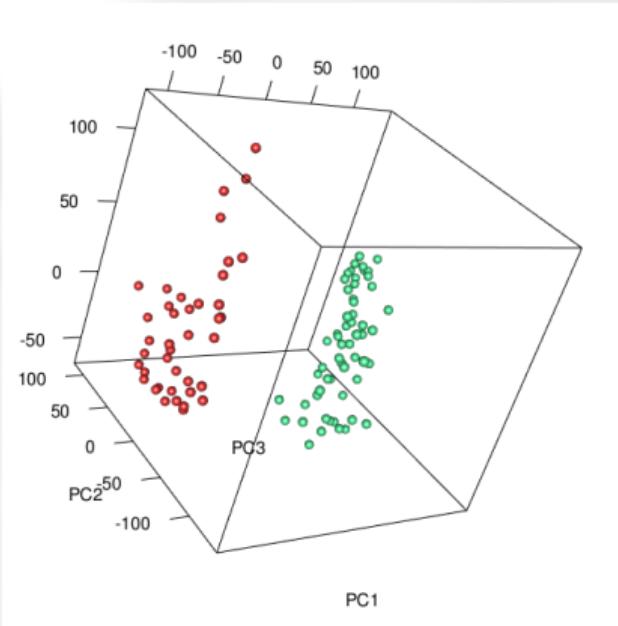


Figura 23. Diagrama de dispersão ternário. Fonte: <<http://www.ggtern.com/2013/12/16/pps-3-state-model/>>.

# Diagrama de dispersão



**Figura 24.** Diagrama de dispersão 3D. Fonte:  
<https://stackoverflow.com/questions/30904198/how-to-export-an-interactive-rgl-3d-plot-to-share-or-publish>.

# Bubble chart

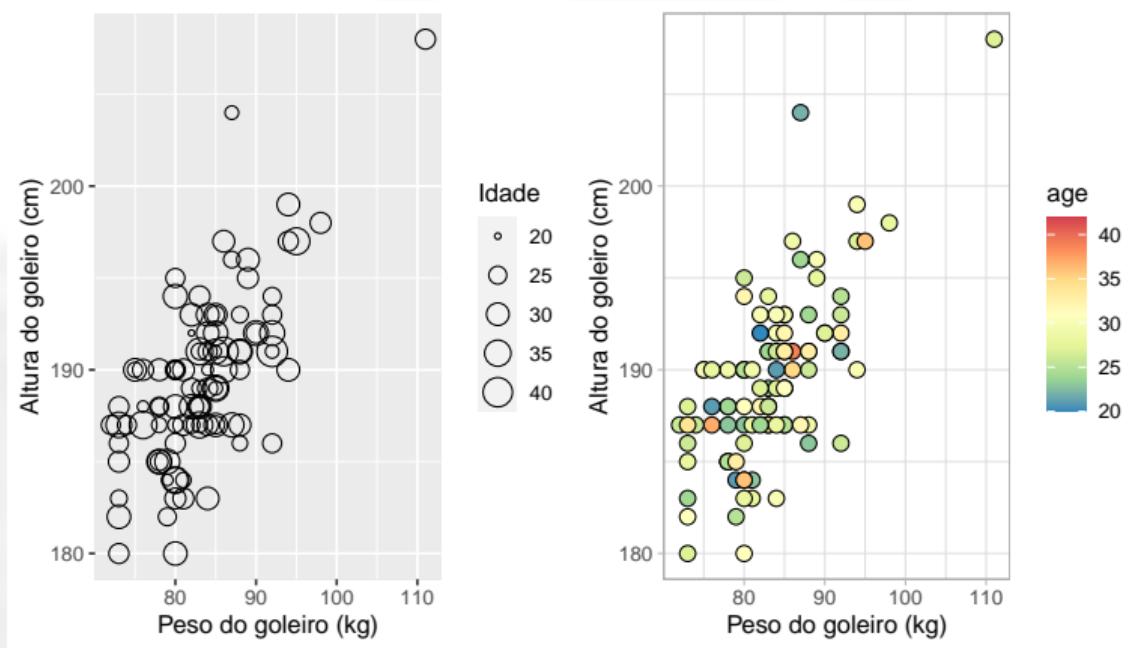


Figura 25. Relação altura, peso e idade dos goleiros.

# Bubble chart

- ▶ **Nome:** diagrama de dispersão com bolhas (*bubble plot*).
- ▶ **Função:** retratar a relação entre variáveis quantitativas.
- ▶ **Tipo de escala:** todas quantitativas.
- ▶ **Variáveis visuais:** posição  $x \times y$  e tamanho.
- ▶ **Geometria:** pontos.

# Matriz de diagramas de dispersão

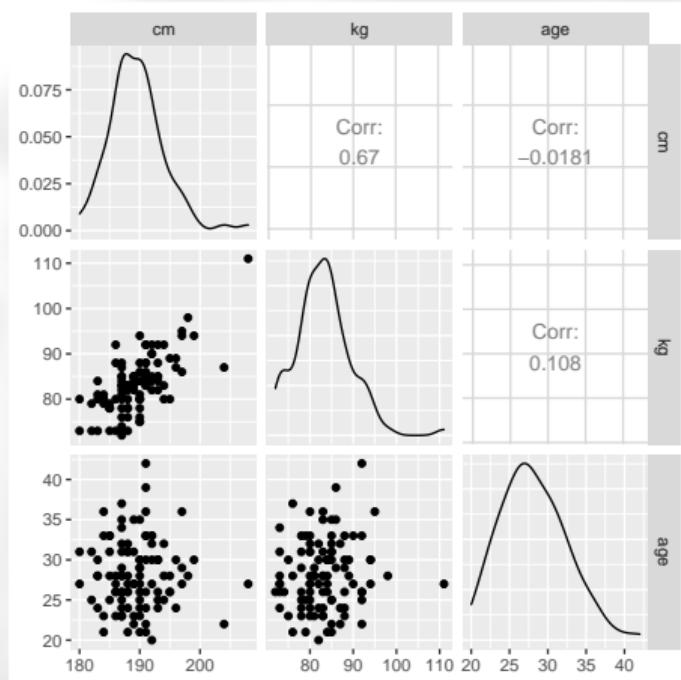


Figura 26. Matriz de diagramas de dispersão.

# Matriz de diagramas de dispersão

- ▶ **Nome:** matriz de diagramas de dispersão (*scatter plot matrix*).
- ▶ **Função:** exibir a relação entre variáveis e distribuições marginais.
- ▶ **Tipo de escala:** preferencialmente quantitativas.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** pontos, linhas e/ou barras.
- ▶ **Inspiração:** diagrama de dispersão e gráficos para distribuição.

# Correlograma

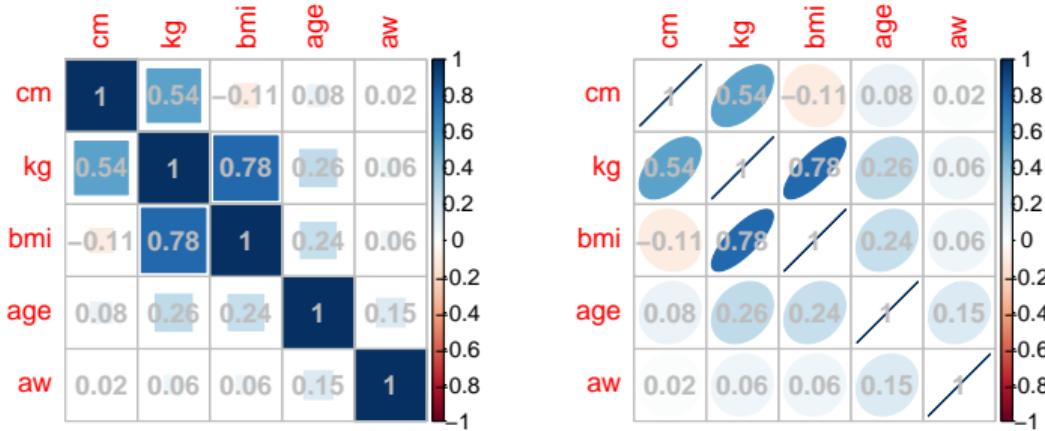


Figura 27. Correlograma das variáveis para os goleiros.

# Correlograma

- ▶ **Nome:** Correlograma (*correlogram*).
- ▶ **Função:** mostrar a estrutura de correlação entre várias variáveis contínuas (admite qualitativas mas requer especificar apropriadamente uma medida de correlação).
- ▶ **Tipo de escala:** preferencialmente para quantitativas.
- ▶ **Variáveis visuais:** área, inclinação ou outra retinal quantitativa.
- ▶ **Geometria:** elipses, círculos ou retângulos.

# Mapa de calor

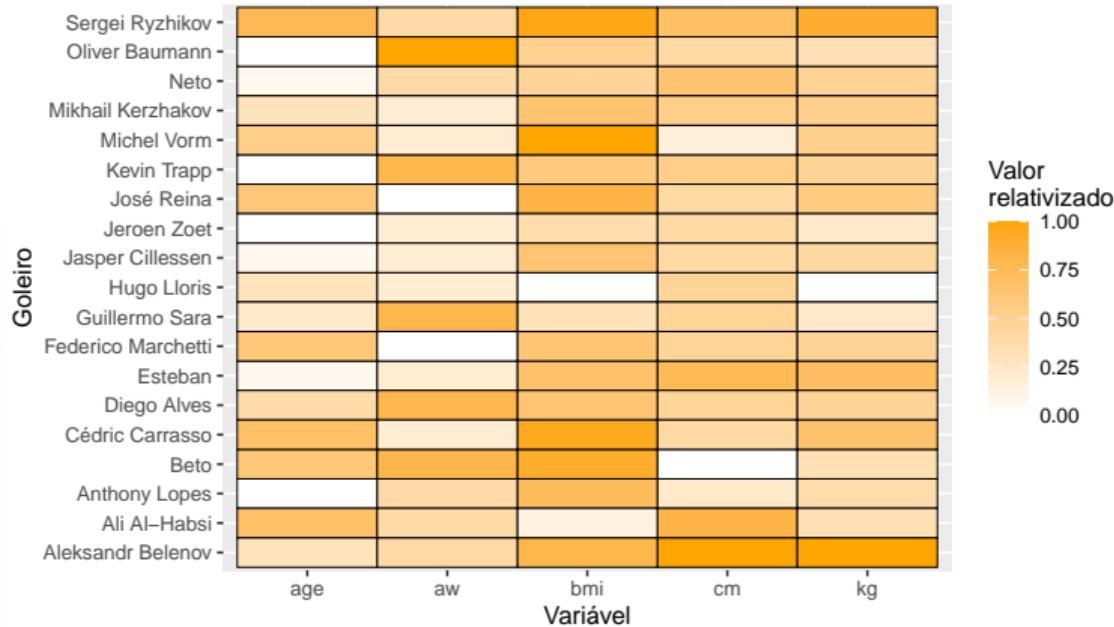


Figura 28. Mapa de calor das variáveis para os goleiros.

# Mapa de calor

- ▶ **Nome:** Mapa de calor (*heatmap*).
- ▶ **Função:** comparar variável quantitativa entre categorias.
- ▶ **Tipo de escala:** quantitativas × categorias.
- ▶ **Variáveis visuais:** posição  $x \times y$ , cor ou saturação.
- ▶ **Geometria:** pixels coloridos.

# Parallel plot

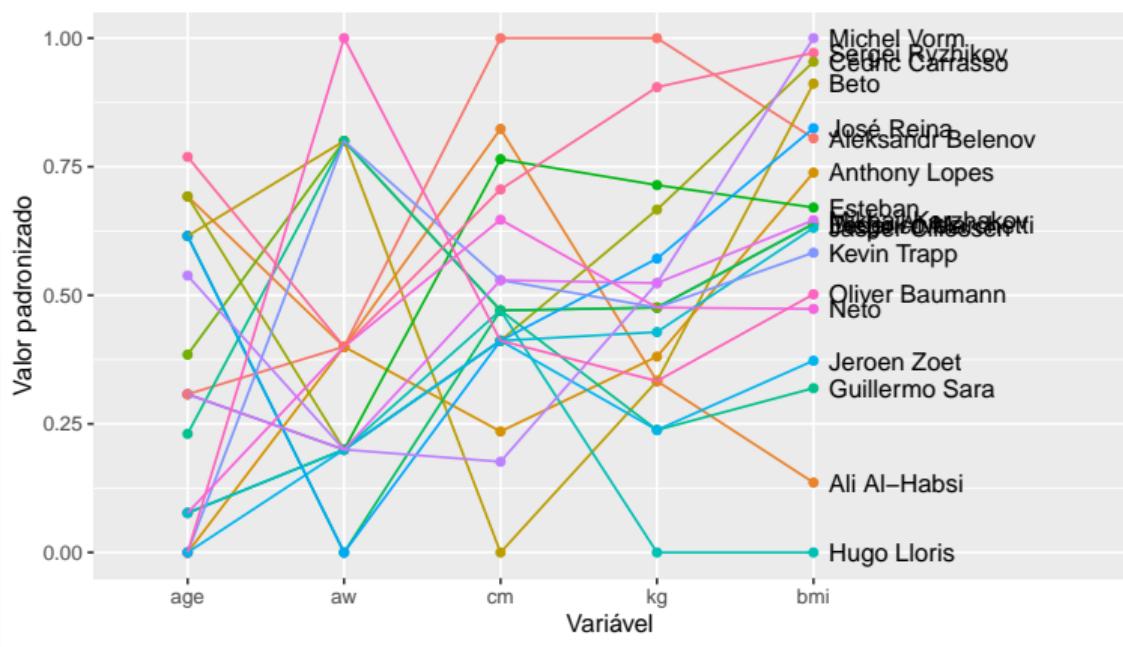


Figura 29. Gráfico de eixos paralelos das variáveis para os goleiros.

# Parallel plot

- ▶ **Nome:** *Parallel plot*.
- ▶ **Função:** comparar variável quantitativa entre categorias e examinar a relação entre variáveis (perfil).
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** linhas.
- ▶ **Inspiração:** *slopegraph*.
- ▶ **Variações:** radar/spider plot.

# Radial/Spider plot

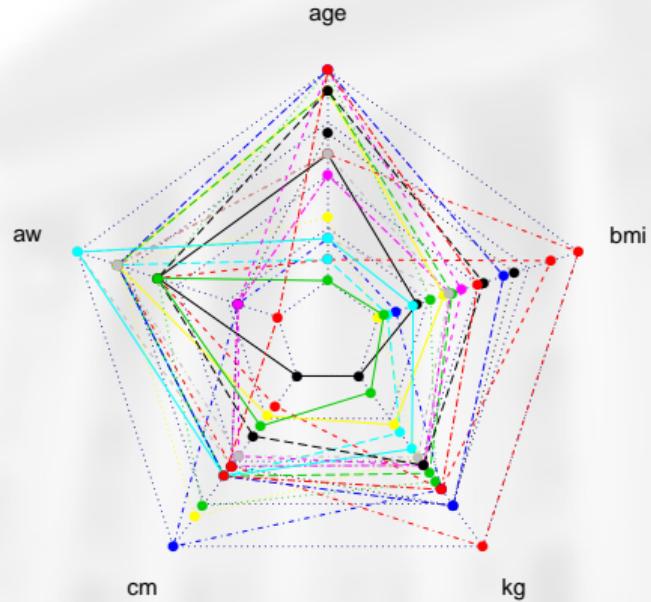


Figura 30. Gráfico de radar das variáveis para os goleiros.

# Radial/Spider plot

- ▶ **Nome:** *Radar plot* ou *Spider plot*.
- ▶ **Função:** comparar variável quantitativa entre categorias.
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** posição  $r \times \theta$ .
- ▶ **Geometria:** linhas.
- ▶ **Inscrição:** *parallel plot*.

# Cleveland dot plot

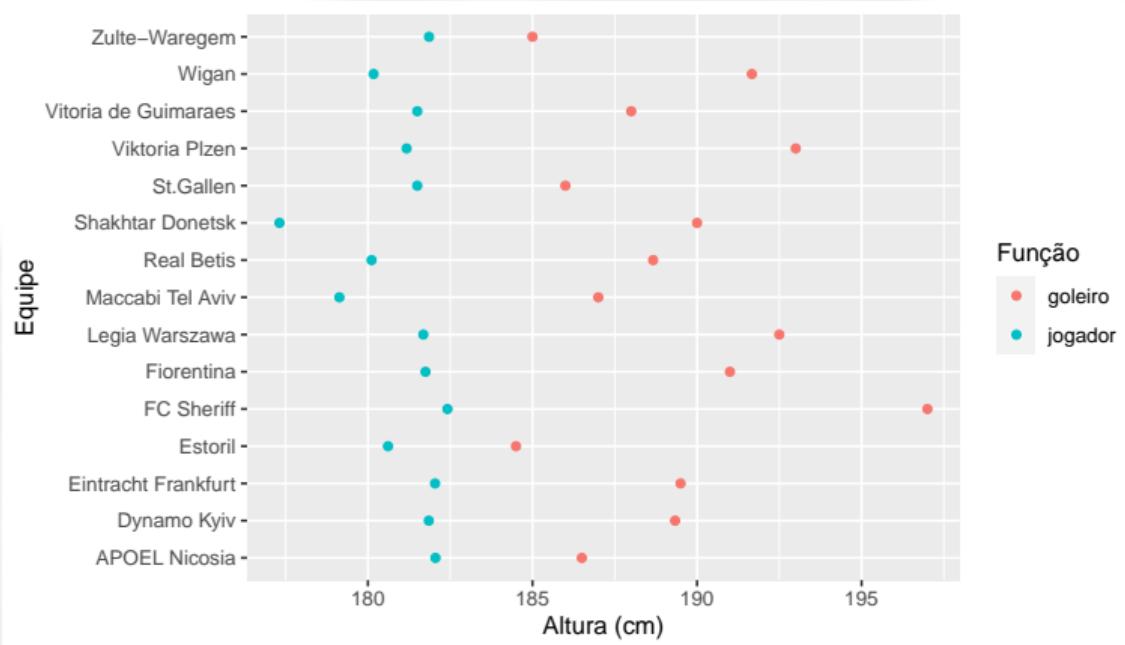


Figura 31. Altura média por equipe para goleiros e jogadores da linha.

# Cleveland dot plot

- ▶ **Nome:** *Cleveland dot plot*.
- ▶ **Função:** comparar variável contínua entre categorias.
- ▶ **Tipo de escala:** quantitativa × qualitativas.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** pontos e segmentos.

# Boxplot

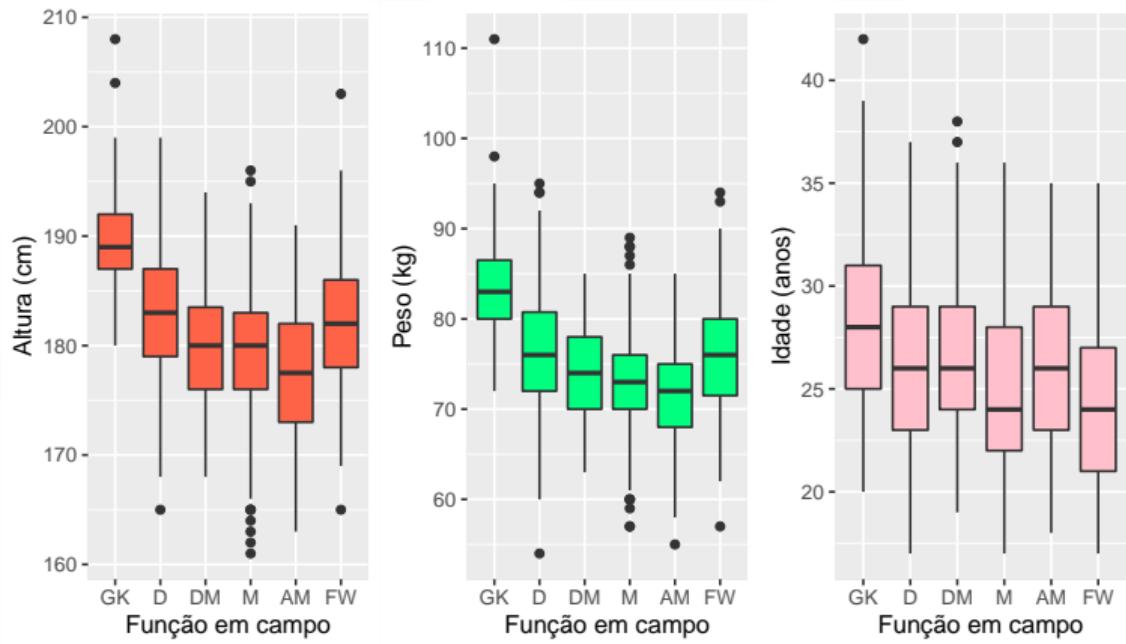


Figura 32. Diagramas de caixa para altura, peso e idade dos jogadores por função.

# Boxplot

- ▶ **Nome:** diagrama de caixas (*boxplot*).
- ▶ **Função:** comparar aspectos da distribuição de variável quantitativa entre categorias.
- ▶ **Tipo de escala:** quantitativa × qualitativa.
- ▶ **Inspiração:** boxplot para única grupo.
- ▶ **Variação:** *minimal box plot*.

# Densidade 2D

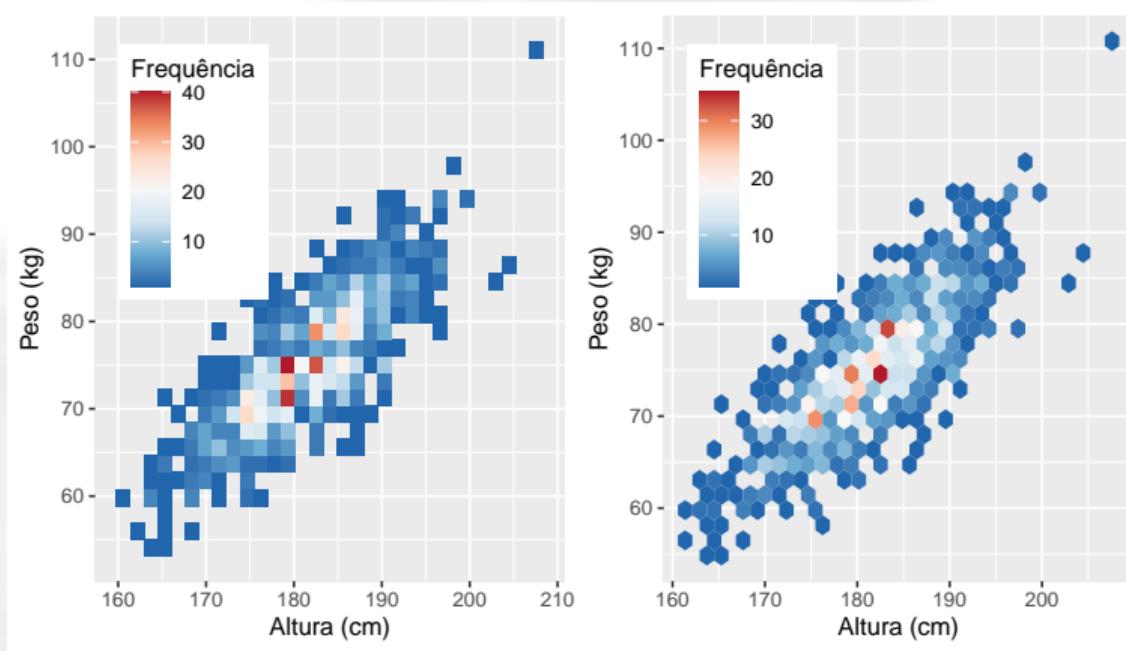


Figura 33. Densidade de jogadores conforme a classificação por altura e peso.

# Densidade 2D

- ▶ **Nome:** Densidade 2D.
- ▶ **Função:** relação entre variáveis e distribuição de valores agrupados em classe.
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** posição  $x \times y$ , cor ou saturação.
- ▶ **Geometria:** pixels coloridos ou linhas de contorno.
- ▶ **Requisitos definíveis:** aspectos da determinação das classes (dimensões, forma e posição dos pixels).
- ▶ **Variáveis calculadas:** frequência absoluta ou relativa.

# Faces de Chernoff

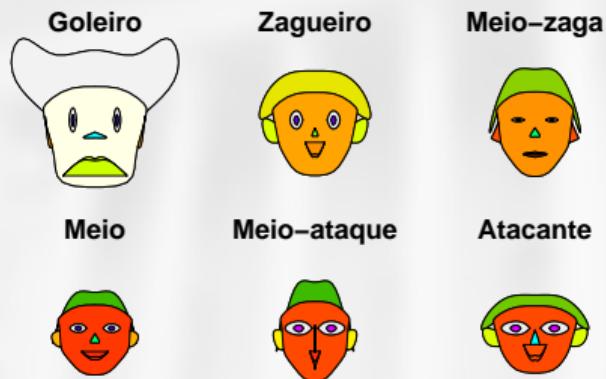


Figura 34. Faces de Chernoff para medidas descritivas dos jogadores conforme função em campo.

# Faces de Chernoff

- ▶ **Nome:** Faces de Chernoff (*Chernoff faces*).
- ▶ **Função:** comparação entre un conjunto de variáveis quantitativas entre categorias/elementos.
- ▶ **Tipo de escala:** quantitativas × categorias.
- ▶ **Variáveis visuais:** posição e comprimentos.
- ▶ **Geometria:** elementos do rosto humano.

# Recomendações

Qual usar para:

1. Examinar a forma da relação entre variáveis contínuas?
2. Comparar valores de variável contínua entre categorias?
3. Comparar a distribuição entre categorias?
4. Exibir a estrutura de correlação entre variáveis?

# Ordenação

# Barras ordenadas

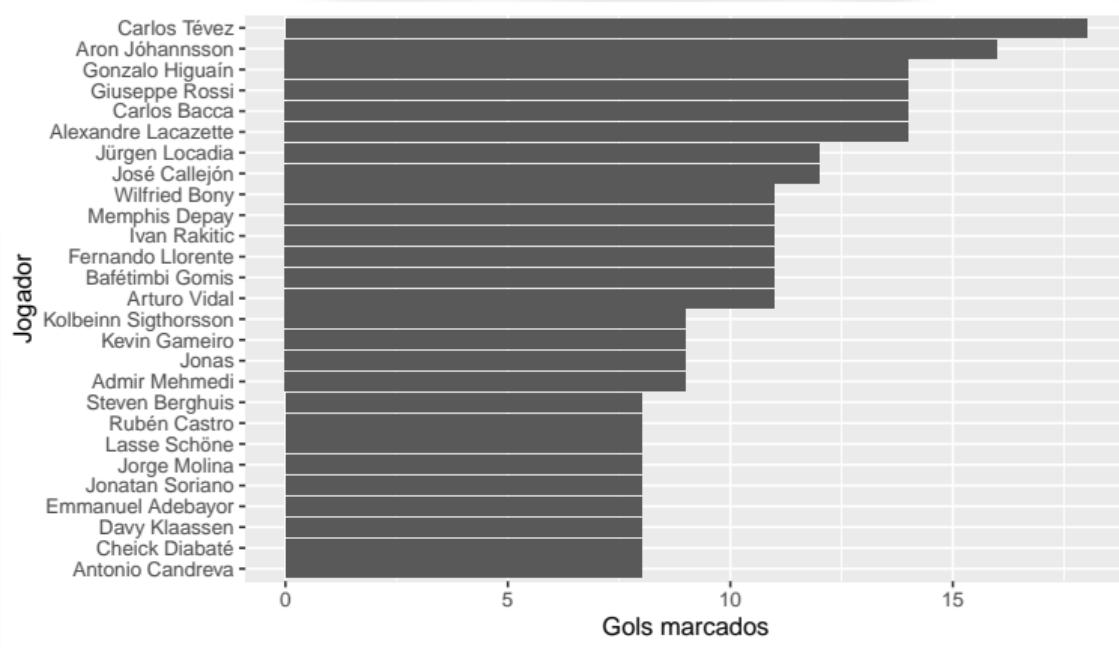


Figura 35. Jogadores classificados pelo número de gols feito.

# Barras ordenadas

- ▶ **Nome:** barras ordenadas.
- ▶ **Função:** ordenar as categorias por uma variável quantitativa.
- ▶ **Tipo de escala:** quantitativa × categorias.
- ▶ **Variáveis visuais:** posição e comprimento.
- ▶ **Geometria:** barras.
- ▶ **Variação:** gráfico de Pareto.

# Lollipop

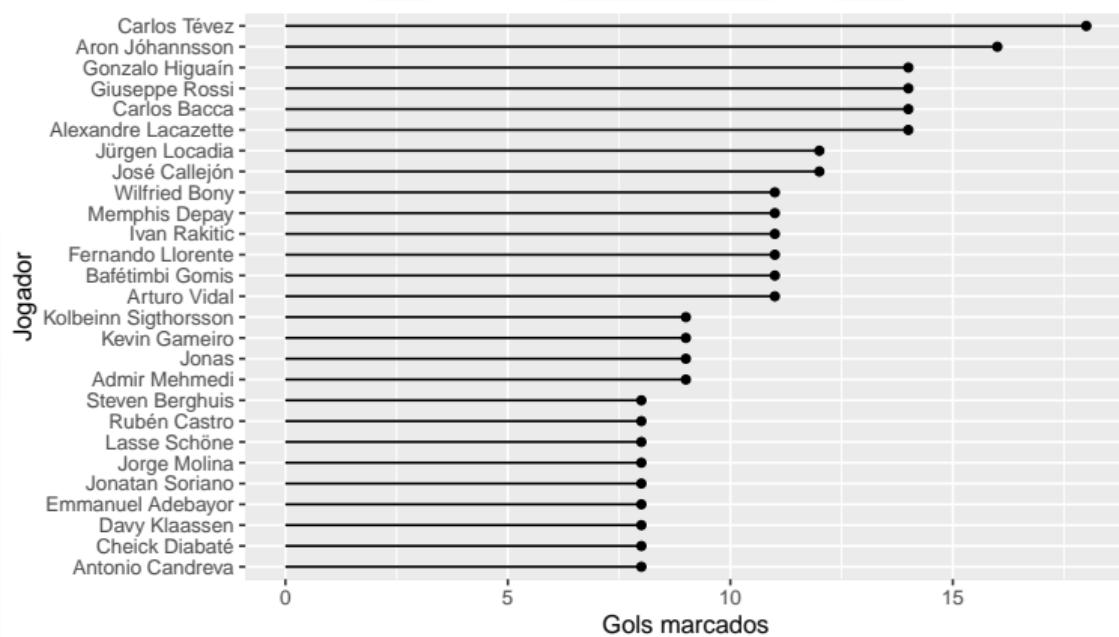


Figura 36. Jogadores classificados pelo número de gols feito.

# Lollipop

- ▶ **Nome:** Lollipop (pirulito) ou Cleveland dot plot.
- ▶ **Função:** ordenar as categorias por variável uma quantitativa.
- ▶ **Tipo de escala:** quantitativa × categorias.
- ▶ **Variáveis visuais:** posição e comprimento.
- ▶ **Geometria:** pontos e segmentos.

# Núvem de palavras



Figura 37. Países conforme o número de jogadores de cada nacionalidade e jogadores conforme o número de gols.

# Núvem de palavras

- ▶ **Nome:** *Wordcloud* ou núvem de palavras.
- ▶ **Função:** retratar a frequência de ocorrência de palavras.
- ▶ **Variáveis visuais:** posição e tamanho.
- ▶ **Geometria:** texto.

# Evolução

# Série ou pontos conectados

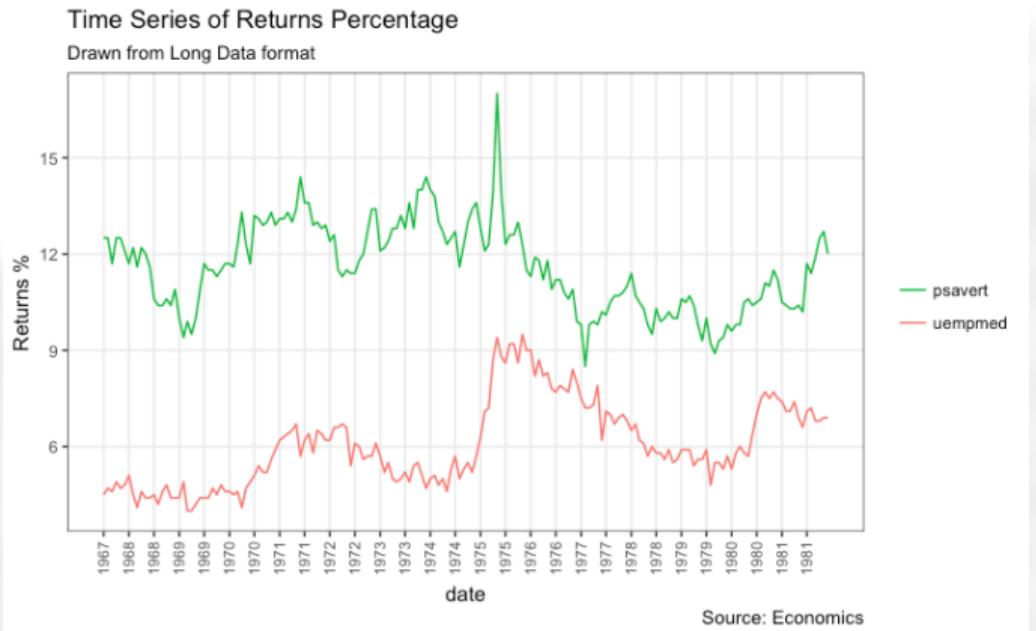
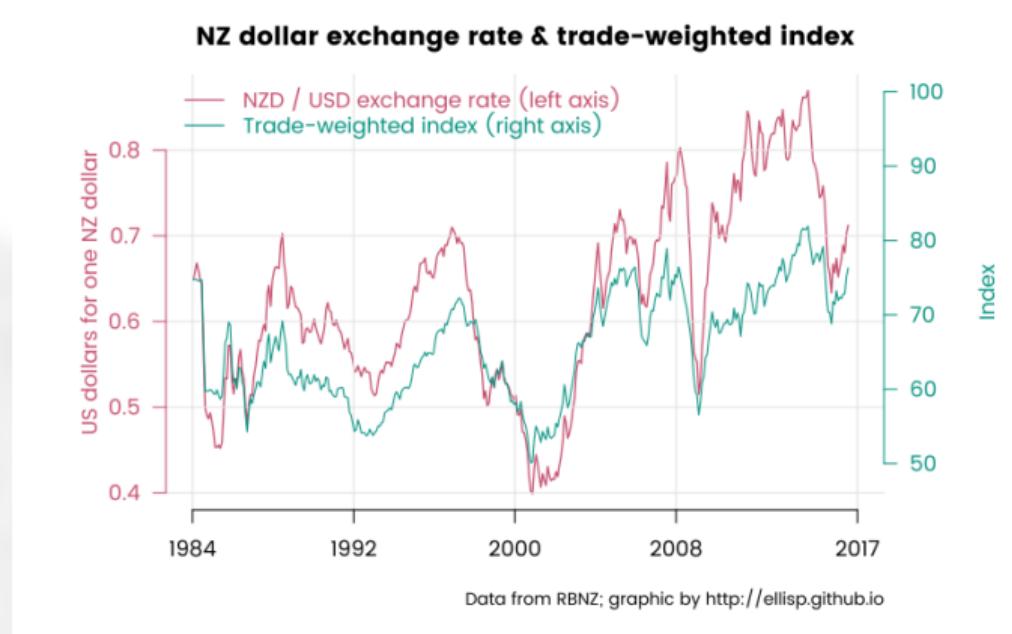


Figura 38. Gráfico de uma série temporal. Fonte:  
<<http://r-statistics.co/Top50-Ggplot2-Visualizations-MasterList-R-Code.html>>.

# Série ou pontos conectados

- ▶ **Nome:** Pontos conectados (*connected scatterplot*).
- ▶ **Função:** mostrar a trajetória de uma variável ao longo do tempo.
- ▶ **Tipo de escala:** quantitativa × cronológica.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** linhas ou pontos conectados.
- ▶ **Variação:**
  - ▶ Área preenchida entre a série e o eixo da origem.
  - ▶ Dois eixos verticais para séries de grandeza diferentes.

# Séries com dois eixos verticais

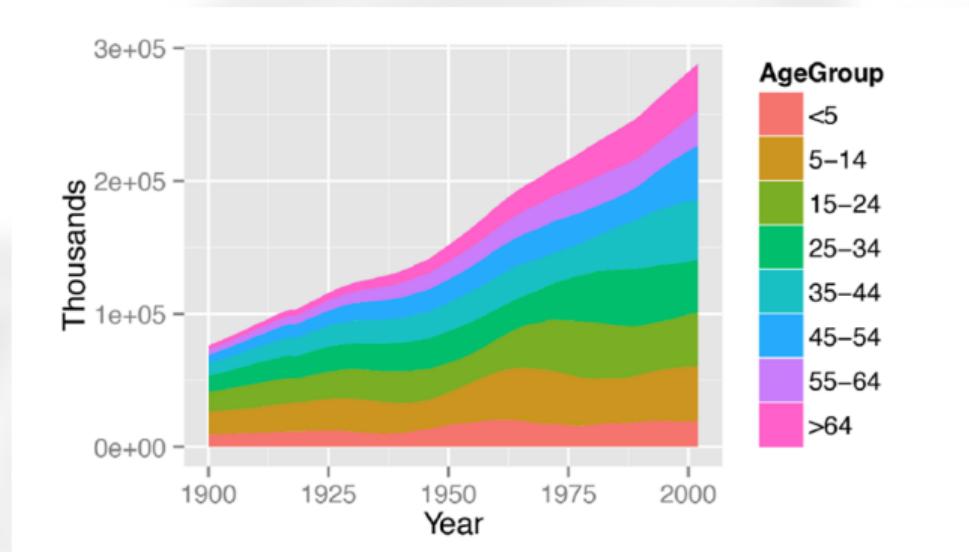


**Figura 39.** Séries com dois eixos. Fonte: <<http://freerangestats.info/blog/2016/08/18/dualaxes>>.

# Séries com dois eixos verticais

- ▶ **Nome:** Série com dois eixos verticais.
- ▶ **Geometria:** linhas ou pontos conectados.

# Áreas empilhadas

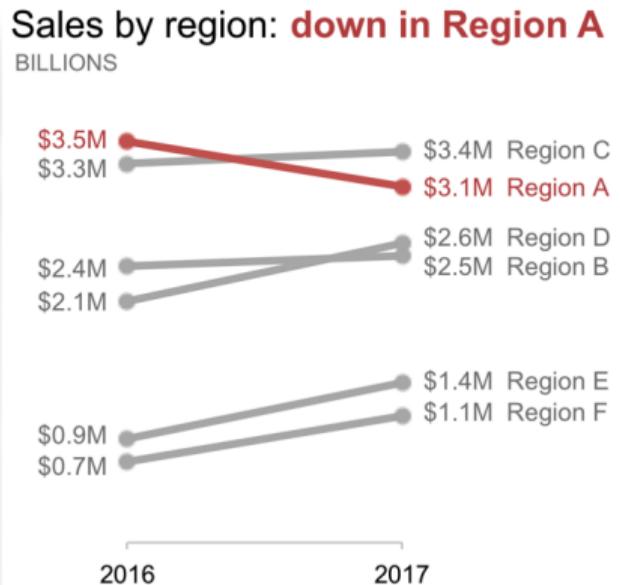


**Figura 40.** Gráfico de áreas empilhadas. Fonte:  
<https://stackoverflow.com/questions/35587571/make-a-stacked-area-graph-in-ggplot2>.

# Áreas empilhadas

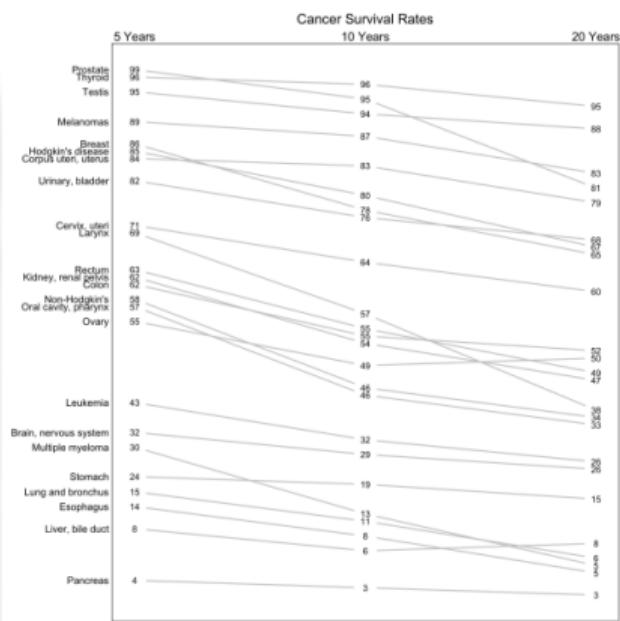
- ▶ **Nome:** Áreas empilhadas (*stacked area plot*).
- ▶ **Função:** mostrar a trajetória de várias séries.
- ▶ **Tipo de escala:** quantitativa × cronológica × categórica.
- ▶ **Variáveis visuais:** posição  $x \times y$  e preenchimento.
- ▶ **Geometria:** polígono preenchido.
- ▶ **Variáveis calculadas:** valores acumulados (relativos) nas categorias em um ponto da variável cronológica.

# Slopegraph



**Figura 41.** Um exemplo de *Slopegraph*. Fonte:  
<<http://www.storytellingwithdata.com/blog/2018/6/1/swdchallenge-slopegraph>>.

# Slopegraph



**Figura 42.** Slopegraph com mais condições. Fonte:  
<<https://shivathudi.com/2017/04/19/data-visualization-in-r-using-ggplot-and-shiny/>>.

# Slopegraph

- ▶ **Nome:** Pontos conectados (*connected scatterplot*).
- ▶ **Função:** mostrar a trajetória de uma variável ao longo do tempo.
- ▶ **Tipo de escala:** quantitativa × cronológica.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** pontos conectados e texto (opcional).

# Calendar plot

Calendar view of DJIA Price Deviation

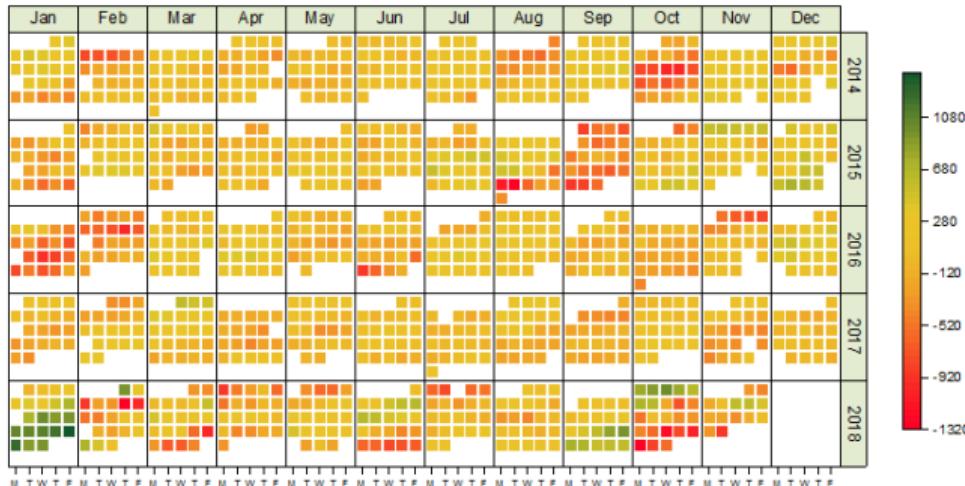


Figura 43. Exemplo de gráfico de calendário. Fonte: <[https://d2mvzyuse3lwjc.cloudfront.net/www/resources/graph\\_gallery/images\\_galleries\\_new/Calendar\\_Chart.png](https://d2mvzyuse3lwjc.cloudfront.net/www/resources/graph_gallery/images_galleries_new/Calendar_Chart.png)>.

# Calendar plot

- ▶ **Nome:** Diagrama de calendário (*calendar plot*).
- ▶ **Função:** retratar uma variável conforme disposição das datas no calendário.
- ▶ **Tipo de escala:** quantitativa × cronológica diária.
- ▶ **Variáveis visuais:** posição  $x \times y$  e preenchimento.
- ▶ **Geometria:** linhas, colunas e cédulas (grid).

# Waterfall chart

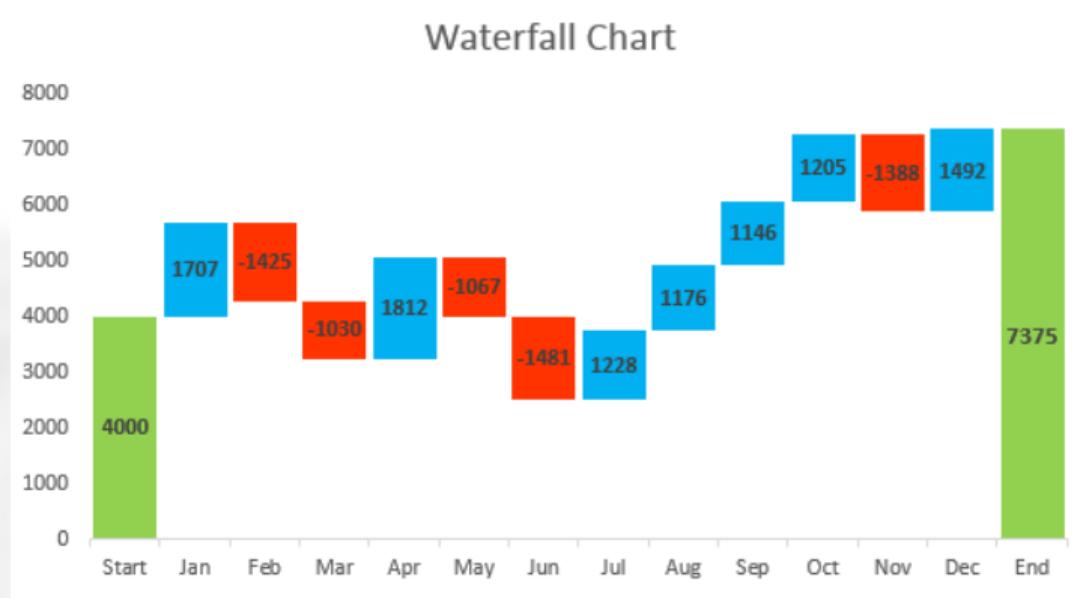
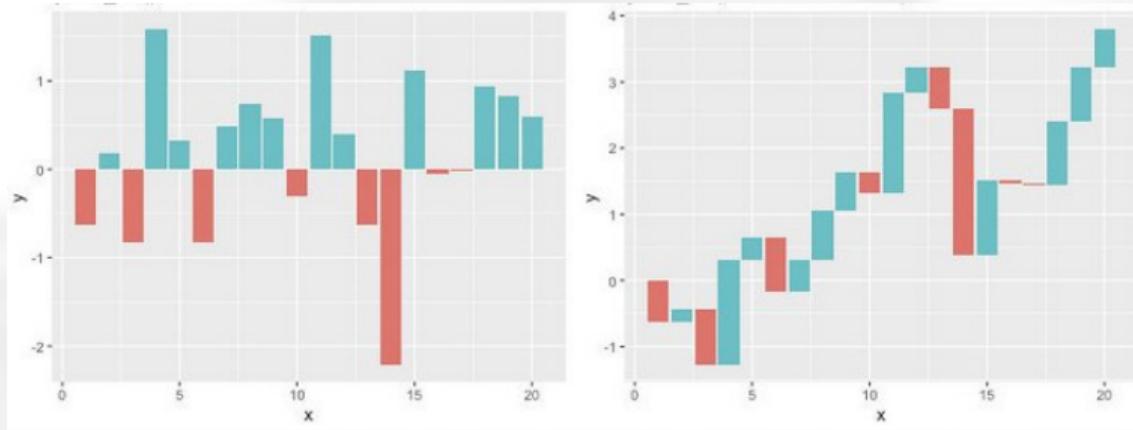


Figura 44. Exemplo de um gráfico de barras em cascata. Fonte:  
<<https://www.ablebits.com/office-addins-blog/2014/07/25/waterfall-chart-in-excel/>>.

# Waterfall chart



**Figura 45.** Diferença de exibição com barras ou barras em cascata. Fonte:  
<https://twitter.com/BrodieGaslam/status/1002932910696157186>.

# Waterfall chart

- ▶ **Nome:** *Waterfall chart* (gráfico de cascata).
- ▶ **Função:** exibe a decomposição/evolução de uma quantidade inicial em termos de somas e diferenças.
- ▶ **Tipo de escala:** quantitativa  $\times$  categorias.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** barras.

# Gantt chart

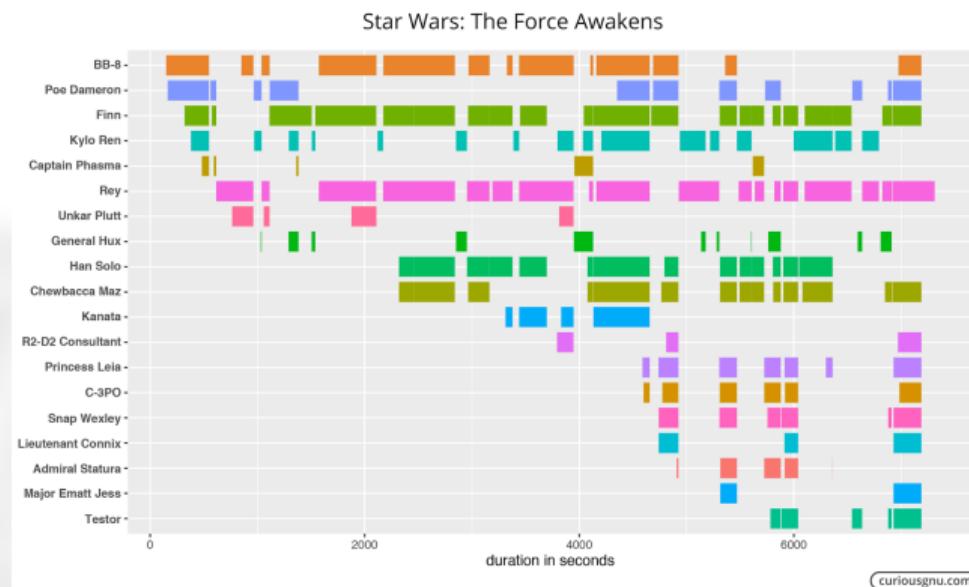
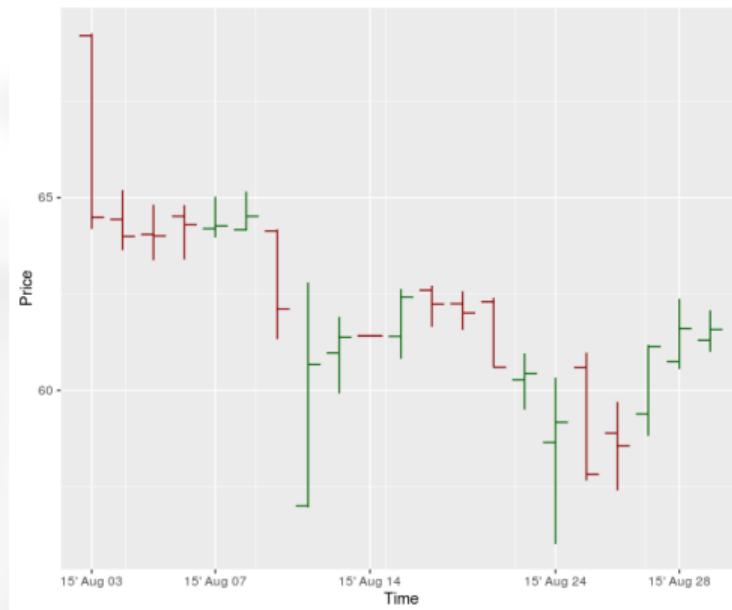


Figura 46. Exemplo de diagrama de Gantt. Fonte: <[http://curiousgnu1.rssing.com/chan-66482562/all\\_p1.html](http://curiousgnu1.rssing.com/chan-66482562/all_p1.html)>.

# Gantt chart

- ▶ **Nome:** Diagrama de Gantt (*Gantt chart*).
- ▶ **Função:** descrever uma variável de duração.
- ▶ **Tipo de escala:** quantitativa × categórica.
- ▶ **Variáveis visuais:** posição  $x \times y$ .
- ▶ **Geometria:** barras ou segmentos.
- ▶ **Inspiração:** *Cleveland dot plot*.

# Candlestick chart



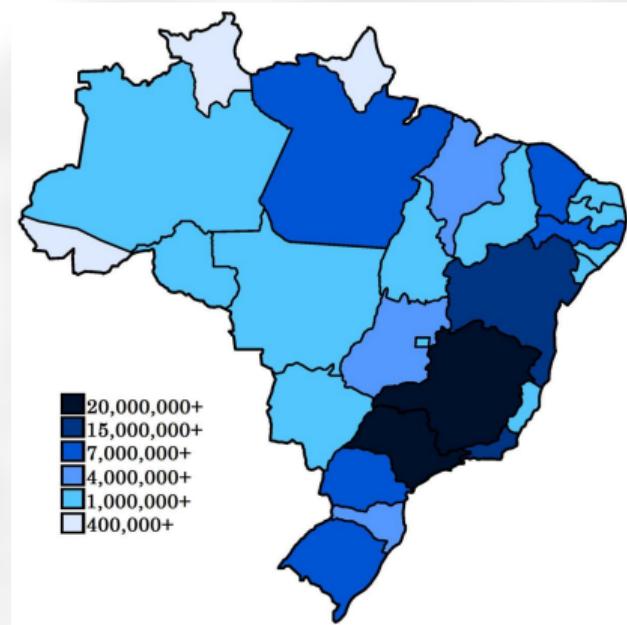
**Figura 47.** Exemplo de gráfico de candelabro. Fonte:  
<<https://stackoverflow.com/questions/28201587/ggplot-for-ohlc-chart-using-bars-not-candles>>.

# Candlestick chart

- ▶ **Nome:** *Candlestick chart* (gráfico de candelabro).
- ▶ **Função:** Exibir valores de abertura, fechamento, máximo e mínimo diariamente.
- ▶ **Tipo de escala:** quantitativa × diária.
- ▶ **Variáveis visuais:** posição  $x \times y$  e cor.
- ▶ **Geometria:** segmentos e/ou retângulos.

# Mapa

# Choropleth map



**Figura 48.** Mapa de densidade populacional por Estado. Fonte:  
<https://gl.maps-brazil.com/brasil-densidade-de-poboación-mapa>.

# Choropleth map

- ▶ **Nome:** *Choropleth map* (grego: “área/região” + “magnitude”) ou mapa temático.
- ▶ **Função:** retratar uma variável quantitativa em um espaço demarcado.
- ▶ **Tipo de escala:** quantitativa × geográfica.
- ▶ **Variáveis visuais:** preenchimento.
- ▶ **Geometria:** polígonos.
- ▶ **Aplicação Shiny:** <<http://shiny.leg.ufpr.br/walmes/mapaufpr/>>.

# Bubble map

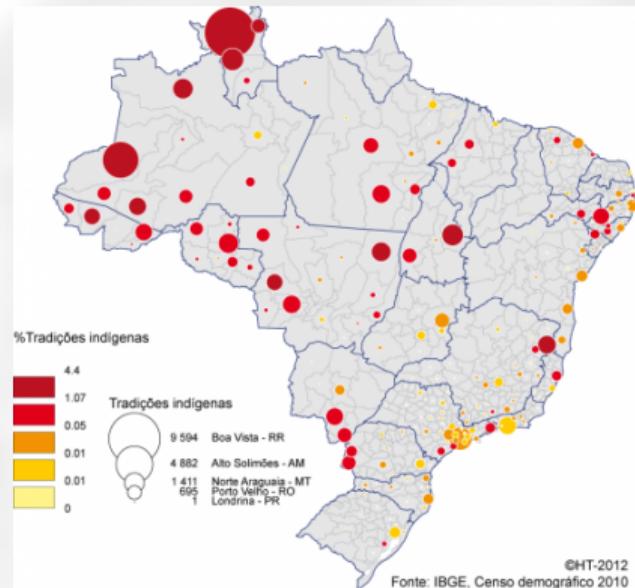


Figura 49. Tradição indígenas no Brasil. Fonte: <<https://journals.openedition.org/confins/7785>>.

# Bubble map

- ▶ **Nome:** *Bubble map* (mapa de bolhas).
- ▶ **Função:** exibir a ocorrência e magnitude de variável quantitativa em um território.
- ▶ **Tipo de escala:** quantitativa × geográfica.
- ▶ **Variáveis visuais:** latitude, longitude e tamanho.
- ▶ **Geometria:** pontos e polígonos.
- ▶ **Variações:** uso de símbolos/marcadores para representar variável qualitativa.
- ▶ **Aplicação Shiny:** <<http://shiny.leg.ufpr.br/walmes/geor/>>.

# Mapa topográfico ou de elevação

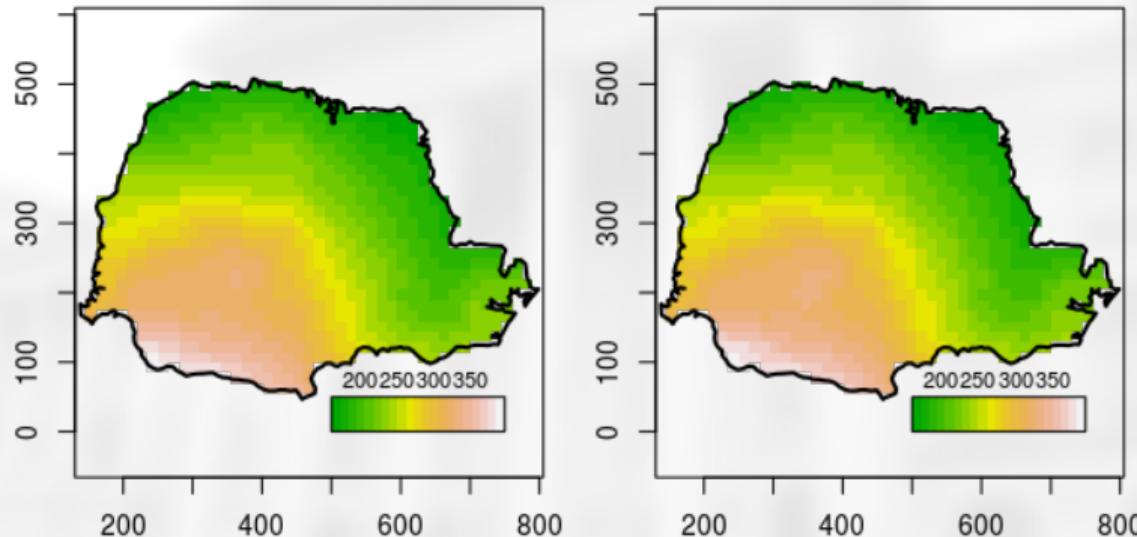


Figura 50. Mapa de densidade populacional por Estado. Fonte: <<http://www.leg.ufpr.br/geoR/PASI/>>.

# Mapa topográfico ou de elevação

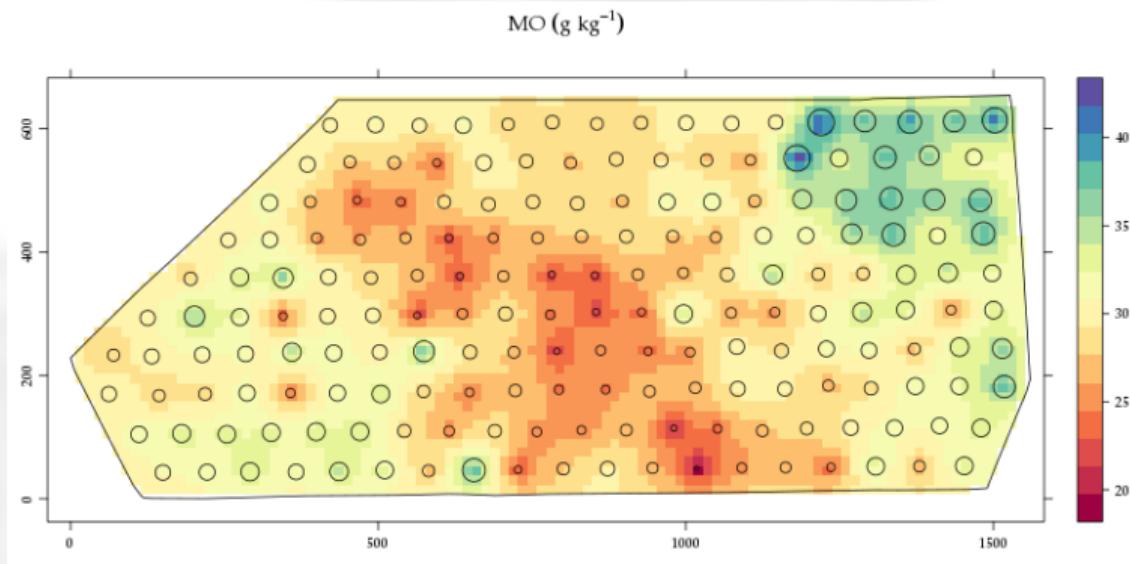
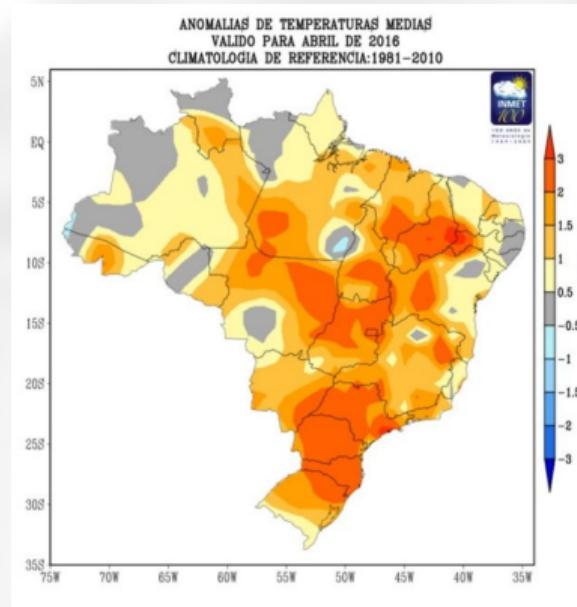


Figura 51. Valores para matéria orgânica do solo. Fonte: <<http://shiny.leg.ufpr.br/walmes/geor/>>.

# Mapa topográfico ou de elevação

- ▶ **Nome:** Mapa topográfico ou de elevação.
- ▶ **Função:** exibir uma variável contínua no espaço geográfico.
- ▶ **Tipo de escala:** quantitativa × geográfica.
- ▶ **Variáveis visuais:** latitude, longitude e cor.
- ▶ **Geometria:** pontos ou pixels.
- ▶ **Variações:** uso de isolinhas de contorno.

# Mapa de contornos de nível

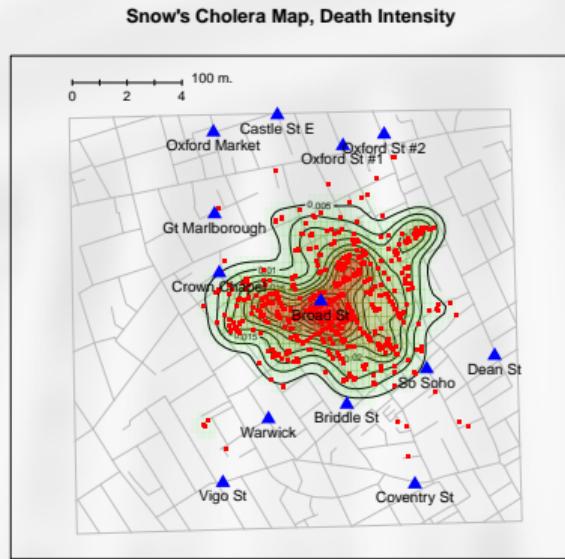


**Figura 52.** Anomalias de temperaturas médias no Brasil para Abril de 2016. Fonte: <[https://www.bbc.com/portuguese/internacional/2016/05/160516\\_nasa\\_aquecimento\\_np](https://www.bbc.com/portuguese/internacional/2016/05/160516_nasa_aquecimento_np)>.

# Mapa de contornos de nível

- ▶ **Nome:** Mapa de contornos de nível.
- ▶ **Função:** exibir a ocorrência de classes de uma variável contínua no espaço geográfico.
- ▶ **Tipo de escala:** catgórica ordenada × geográfica.
- ▶ **Variáveis visuais:** latitude, longitude e cor.
- ▶ **Geometria:** polígonos.
- ▶ **Requisitos definíveis:** intervalos de classe.

# Mapa de densidade



**Figura 53.** Mapa de densidade de ocorrência. Fonte: <<https://cran.r-project.org/package=HistData>>.

# Mapa de densidade

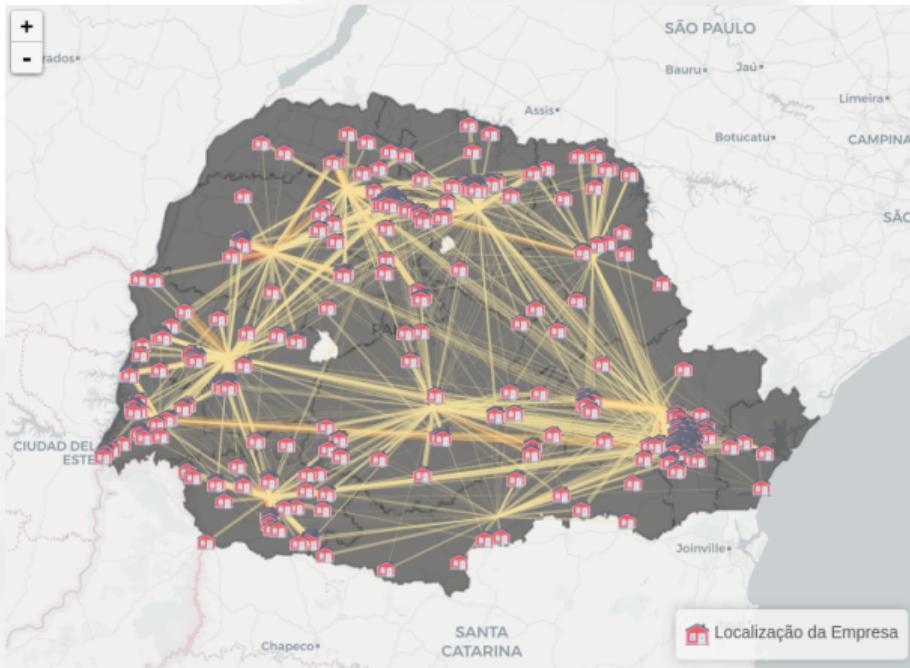
- ▶ **Nome:** Mapa de desidade.
- ▶ **Função:** exibe a intensidade da ocorrência de eventos no espaço.
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** latitude, longitude e cor.
- ▶ **Geometria:** pontos.
- ▶ **Requisitos definíveis:** função de densidade, largura de banda.
- ▶ **Variáveis calculadas:** densidade.
- ▶ **Inspiração:** gráfico de densidade empírica.
- ▶ **Variação:** intensidade por píxel/hexágono.

# Mapa de conexões



Figura 54. Mapa de conexões. Fonte: <[https://datavizcatalogue.com/methods/connection\\_map.html](https://datavizcatalogue.com/methods/connection_map.html)>.

# Mapa de conexões



**Figura 55.** Mapa de conexões entre empresas e prefeituras no Paraná. Fonte: Trabalho de Conclusão de Curso de Especialização em Data Science & Big Data, Paulo Jhonny Scheleider da Costa Rosa.

# Mapa de conexões

- ▶ **Nome:** Mapa de conexões.
- ▶ **Função:** exibe rotas e conexões entre elementos no mapa.
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** latitude, longitude e tamanho.
- ▶ **Geometria:** linhas/arcos.

# Cartograma

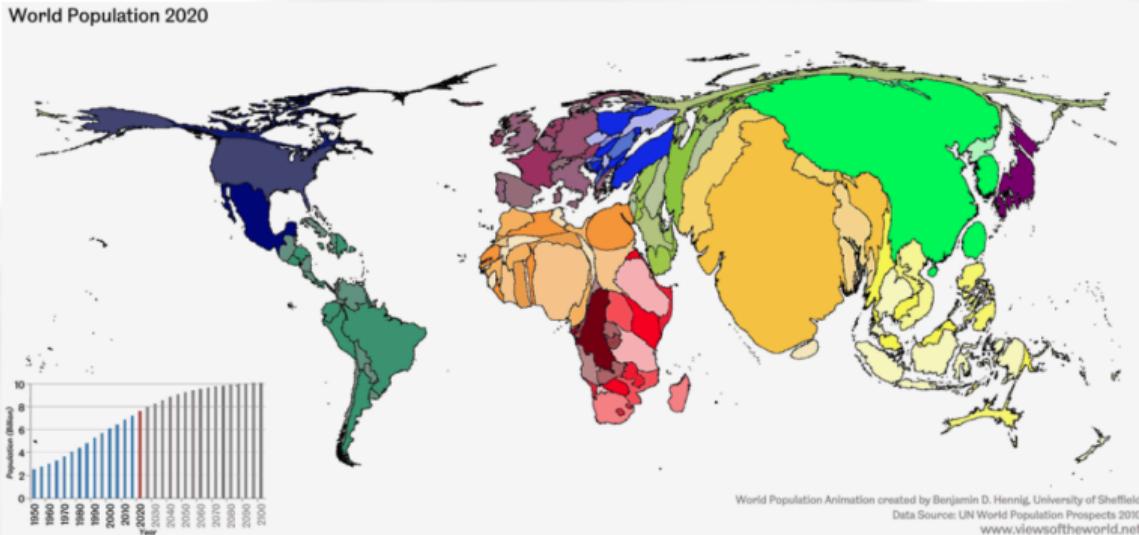


Figura 56. Cartograma da população mundial em 2020. Fonte:  
<<https://giphy.com/gifs/geography-jpJ6VkJ9k17cA>>.

# Cartograma

- ▶ **Nome:** Cartograma (*cartogram*).
- ▶ **Função:** retrata uma variável quantitativa pelo tamanho dos polígonos ou distância fazendo distorção dos mesmos.
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** latitude, longitude e tamanho.
- ▶ **Geometria:** polígonos.

# Organização

# Dendrograma

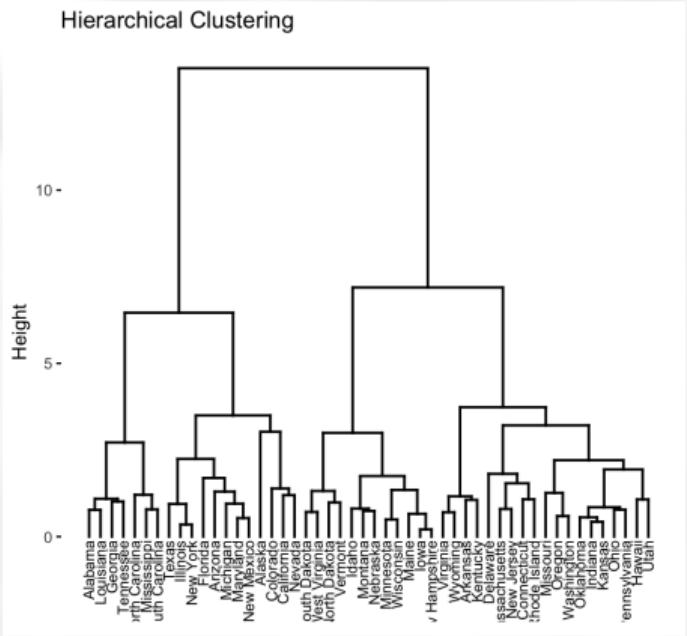
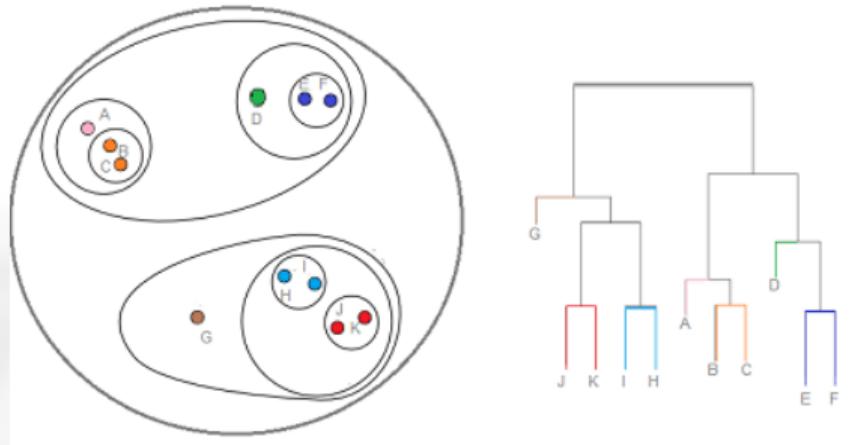


Figura 57. Exemplo de dendrograma. Fonte:  
<https://www.datanovia.com/en/courses/hierarchical-clustering-in-r-the-essentials/>.

# Dendrograma



**Figura 58.** Ideia para a construção do dendrograma. Fonte:  
<https://www.statisticshowto.datasciencecentral.com/hierarchical-clustering/>.

# Dendrograma

- ▶ **Nome:** Dendrorama (dendro = árvore).
- ▶ **Função:** retrata o grau similaridade ou estrutura hierárquica (perfilhamento) entre categorias.
- ▶ **Tipo de escala:** quantitativas × categorias.
- ▶ **Variáveis visuais:** comprimento.
- ▶ **Geometria:** linhas.
- ▶ **Requisitos definíveis:** medida de similaridade, método de agrupamento.
- ▶ **Variáveis calculadas:** similaridade ou distância entre (grupos de) categorias.

# Treemap

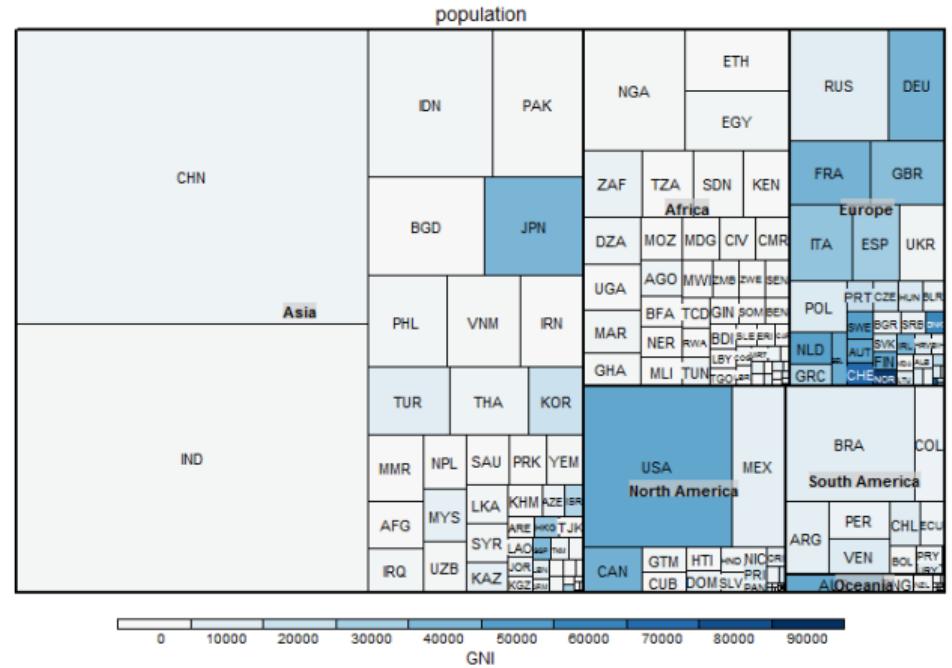


Figura 59. Exemplo de treemap. Fonte: <<https://github.com/mtennekes/treemap>>.

# Treemap

- ▶ **Nome:** Treemap (mapa da árvore).
- ▶ **Função:** retrata a organização hierárquica com o tamanho das divisões.
- ▶ **Tipo de escala:** quantitativa × categorias.
- ▶ **Variáveis visuais:** área.
- ▶ **Geometria:** retângulos.
- ▶ **Inspiração:** dendrograma e gráfico de mosaico.
- ▶ **Variação:** *circular packing*.
- ▶ <<http://newsmap.jp/>>.

# Circular packing

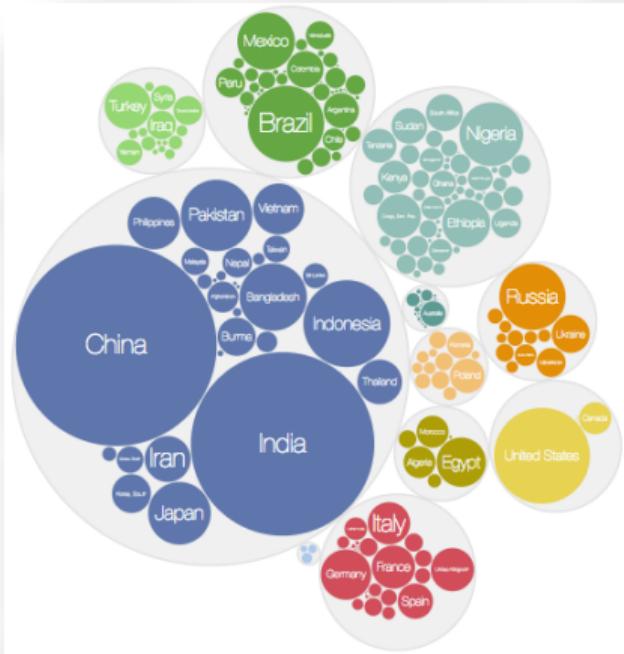
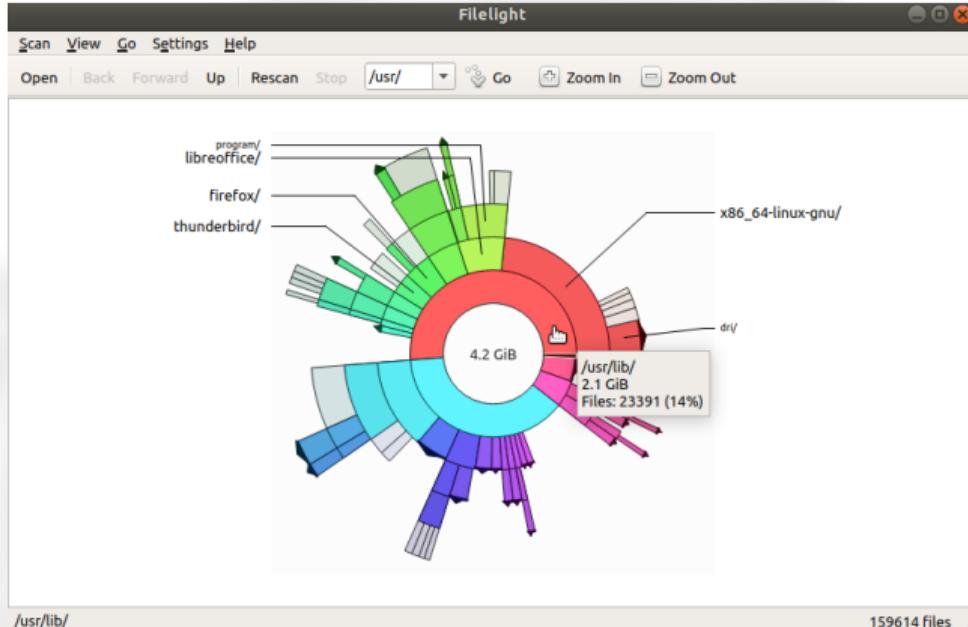


Figura 60. Exemplo de *circular packing*. Fonte: <<https://d3indepth.com/layouts/>>.

# Circular packing

- ▶ **Nome:** *Circular packing*.
- ▶ **Inspiração:** treemap e diagrama de Venn.
- ▶ **Geometria:** círculos.
- ▶ **Variação:** *circular packing*.

# Sunburst



**Figura 61.** Exemplo de diagrama *sunburst*. Fonte:  
<<https://www.debugpoint.com/2018/07/apps-visual-disk-usage-linux-ubuntu-mint/>>.

# Sunburst

- ▶ **Nome:** *Sunburst* (= raios de sol).
- ▶ **Função:** retrata a organização hierárquica com o tamanho das divisões.
- ▶ **Tipo de escala:** quantitativa × categorias.
- ▶ **Variáveis visuais:** ângulo.
- ▶ **Geometria:** arcos.

# Fluxo

# Network

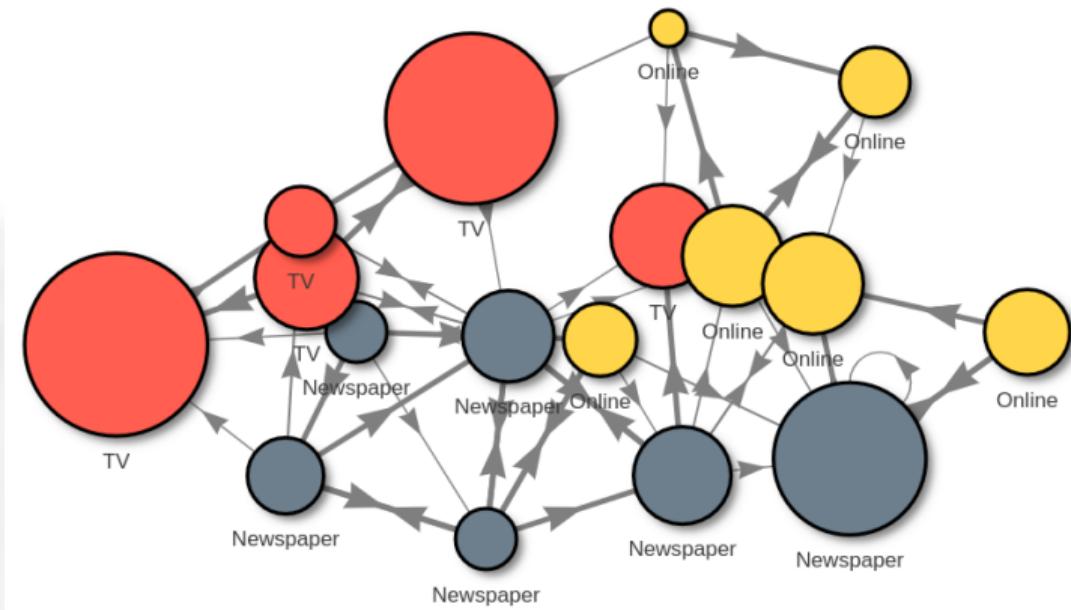
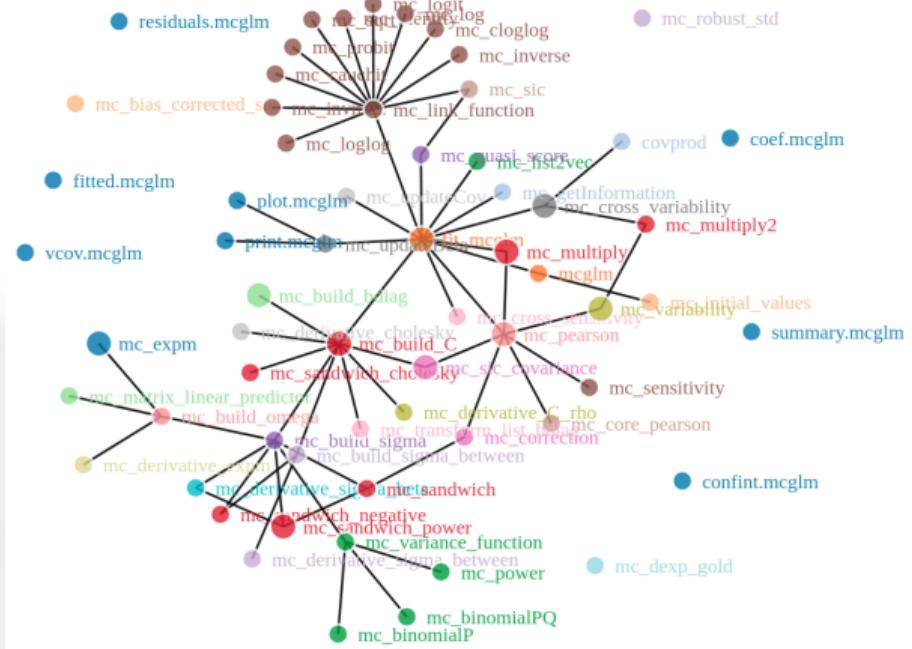


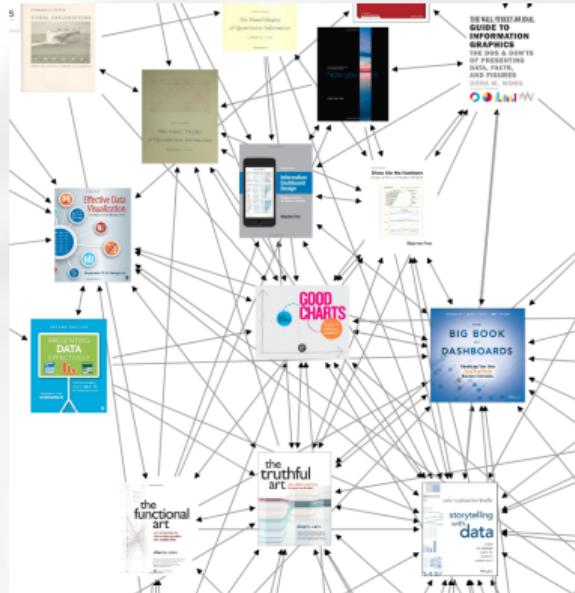
Figura 62. Exemplos de redes. Fonte: <<https://kateto.net/network-visualization>>.

# Network



**Figura 63.** Relação entre as funções do pacote `mcglm`. Fonte:  
[http://leg.ufpr.br/~leg/mcglm/vignettes/functions\\_network.html](http://leg.ufpr.br/~leg/mcglm/vignettes/functions_network.html).

# Network



**Figura 64.** Relação entre livros sobre visualização de dados comprados na Amazon. mcglm. Fonte: <<http://yasiv.com/#/Search?q=Data%20visualization>>.

# Network

- ▶ **Nome:** Network (rede) ou gráfo.
- ▶ **Função:** compreender o relacionamento entre um grupo de entidades.
- ▶ **Tipo de escala:** qualitativa e quantitativa.
- ▶ **Variáveis visuais:** tamanho e preenchimento.
- ▶ **Geometria:** pontos e segmentos.
- ▶ **Variações:** *radial network*.
- ▶ Rede de relacionamento de livros na Amazon: <<http://yasiv.com/>>.
- ▶ Redes com o R: <<https://kateto.net/network-visualization>>.

# Radial network

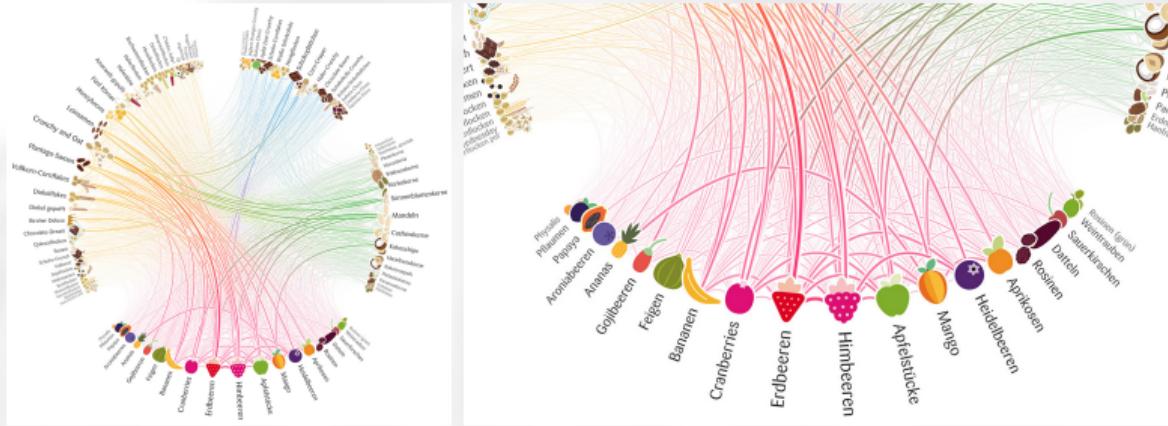


Figura 65. Exemplo de rede radial. Fonte: <<https://truth-and-beauty.net/projects/muesli-ingredient-network>>.

# Radial network

- ▶ **Nome:** *Radial network* ou *hierarchical edge bundling*.
- ▶ **Função:** visualizar a conexão entre todas entidades de uma coleção.
- ▶ **Variações:** diagrama de arcos.

# Diagrama de arcos

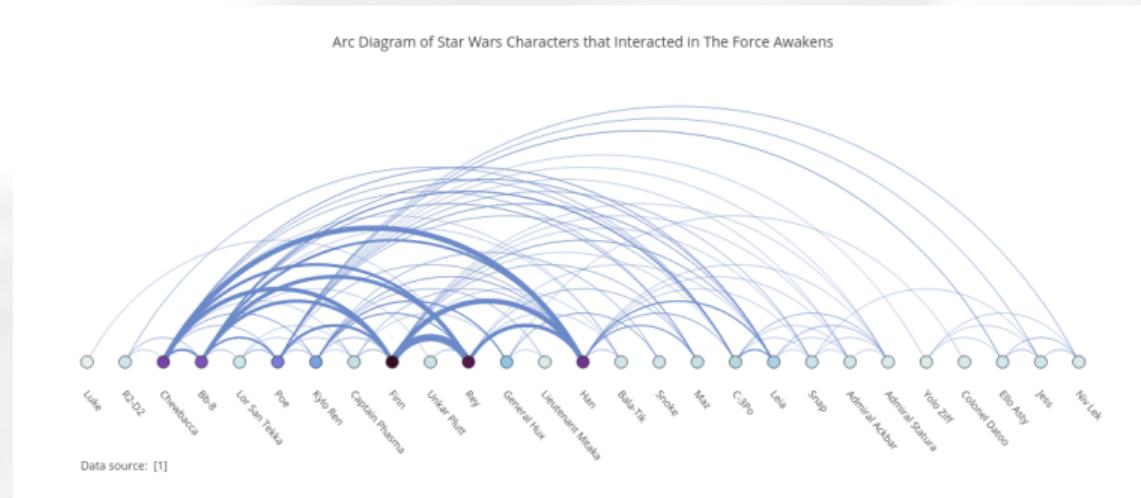


Figura 66. Exemplo de diagrama de arco. Fonte:

<<https://plot.ly/~empet/13574/arc-diagram-of-star-wars-characters-that-interacted-in-the-force-awakens>>.

# Diagrama de arcos

- ▶ **Nome:** Diagrama de arcos (*arc diagram*).
- ▶ **Função:** visualizar a conexão entre entidades de uma coleção.

# Sankey diagram ou alluvial chart

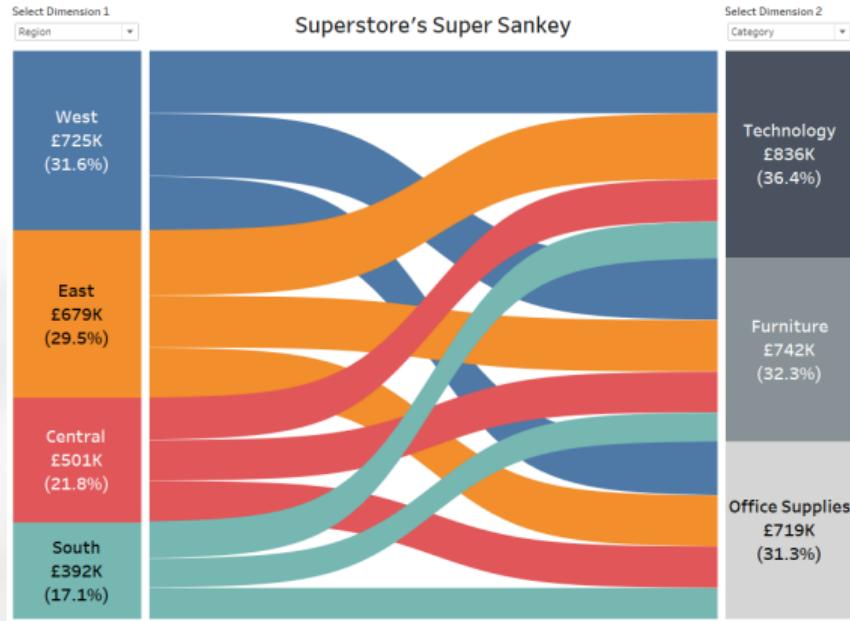


Figura 67. Exemplo de Sankey diagram. Fonte: <<https://www.theinformationlab.co.uk/2018/03/09/build-sankey-diagram-tableau-without-data-prep-beforehand/>>.

# Sankey diagram ou alluvial chart

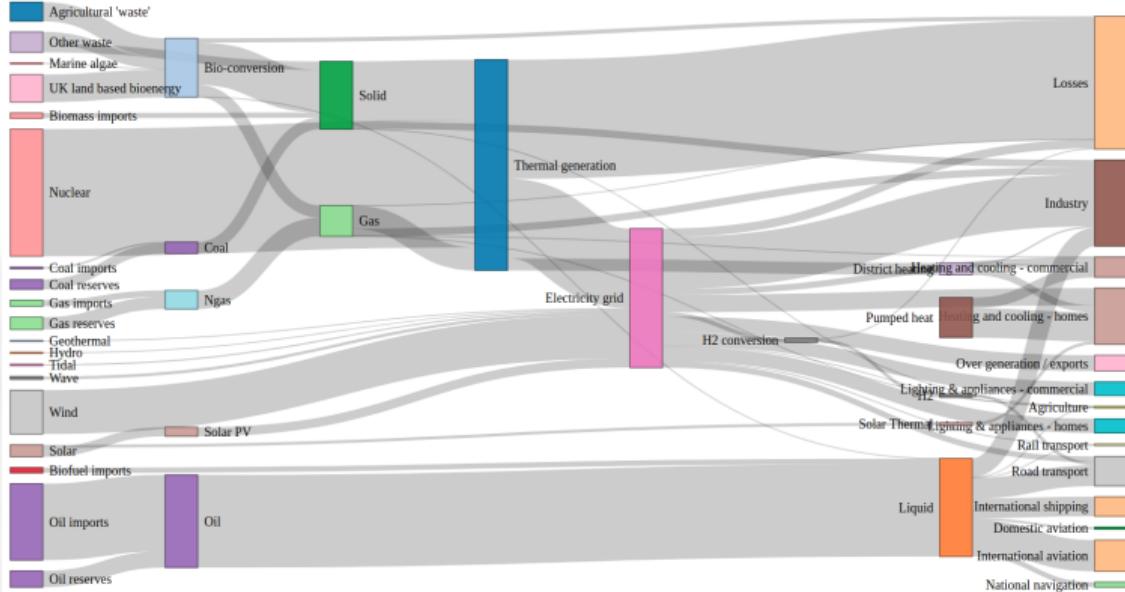


Figura 68. Exemplo de Sankey diagram. Fonte: <<https://www.r-graph-gallery.com/sankey-diagram/>>.

# Sankey diagram ou alluvial chart

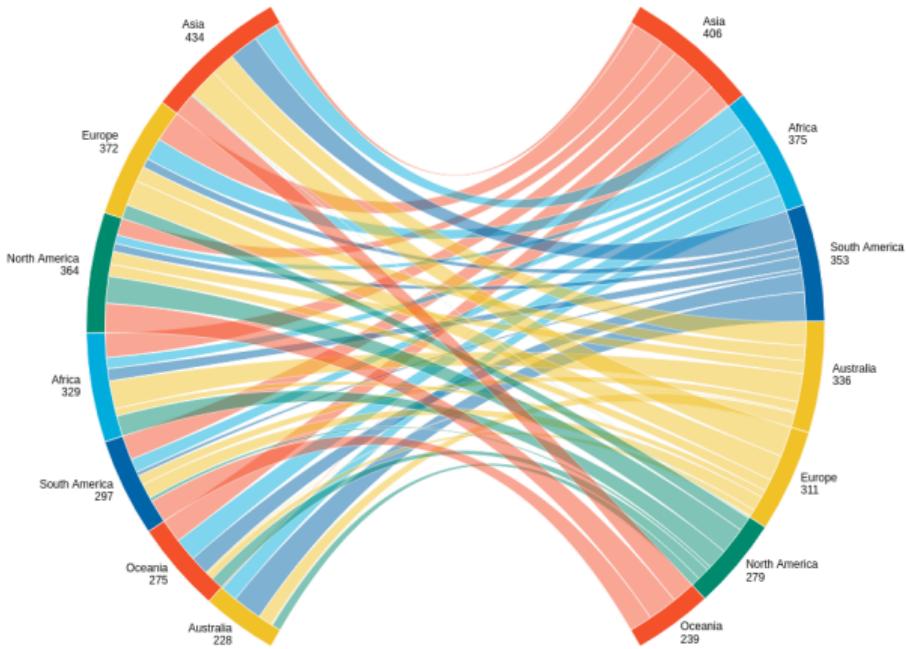


Figura 69. Exemplo de Sankey diagram. Fonte: <<https://plotdb.com/chart/1039>>.

# Sankey diagram ou alluvial chart

- ▶ **Nome:** *Sankey diagram* ou *alluvial chart*.
- ▶ **Função:** exibe o fluxo de quantidade/proporção de uma variável.
- ▶ **Tipo de escala:** quantitativa.
- ▶ **Variáveis visuais:** tamanho.
- ▶ **Geometria:** arcos.

# Chord diagram

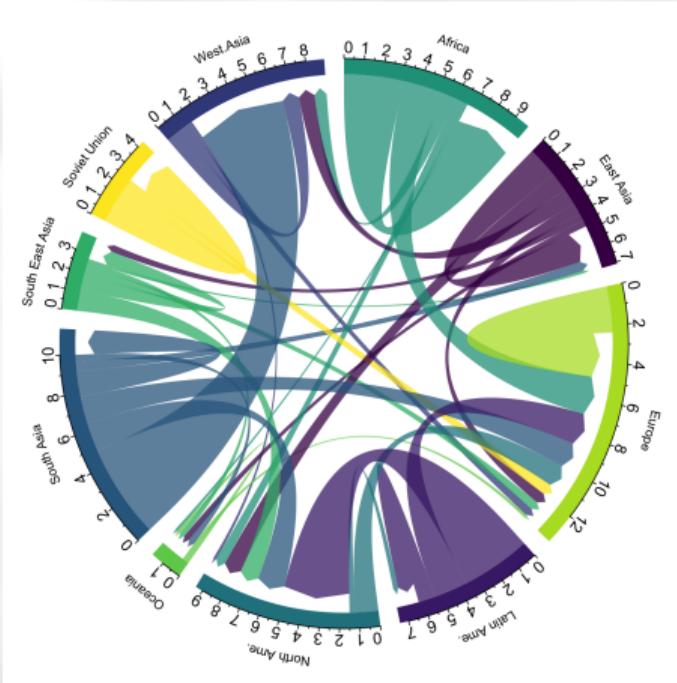


Figura 70. Exemplos de diagramas de corda. Fonte: <<https://www.data-to-viz.com/graph/chord.html>>.

# Chord diagram

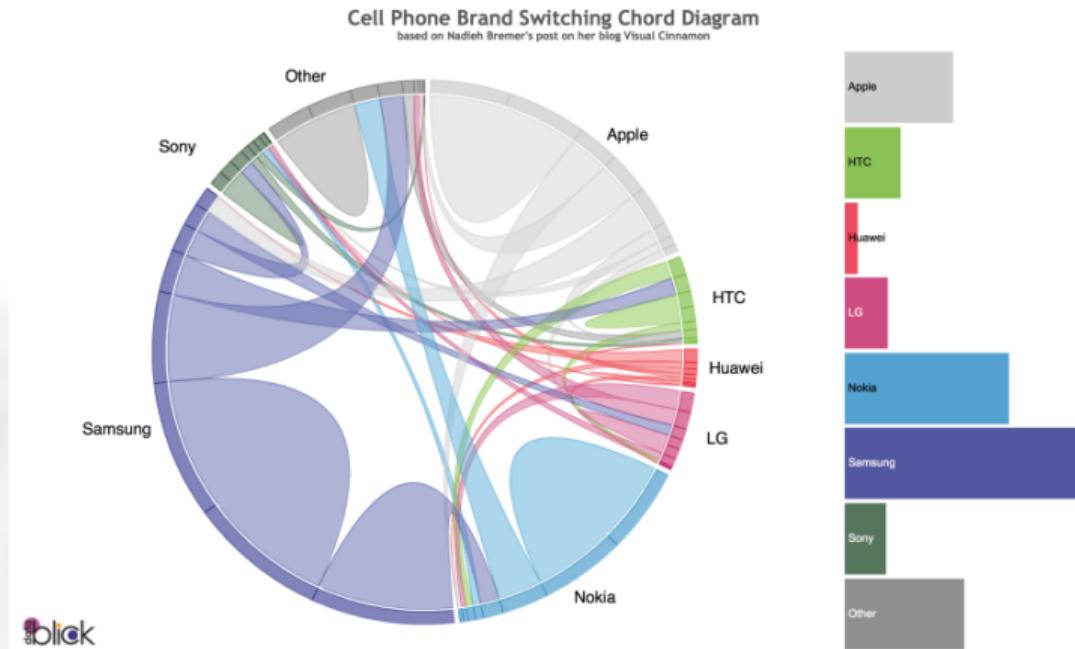
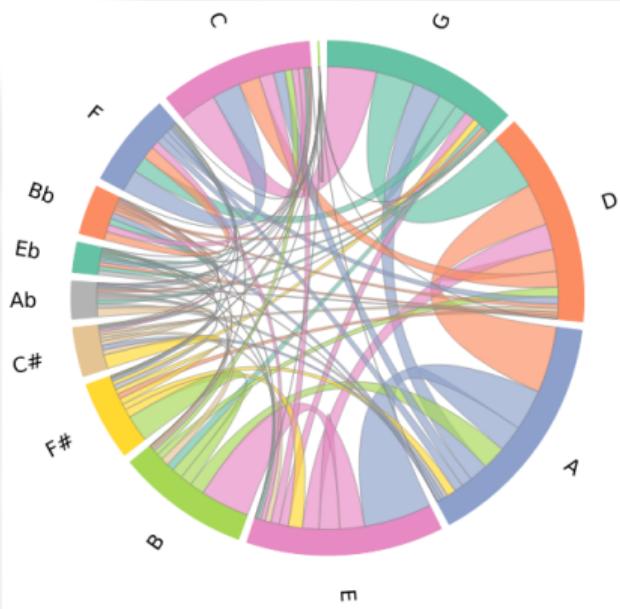


Figura 71. Exemplos de diagramas de corda. Fonte:  
<https://learningtableaublog.wordpress.com/2016/08/04/chord-diagrams/>.

# Chord diagram



**Figura 72.** Disposição dos acordes em sequência na música brasileira. Trabalho de conclusão de curso de Bruna Wunderwald e Rogério Filho (2017).

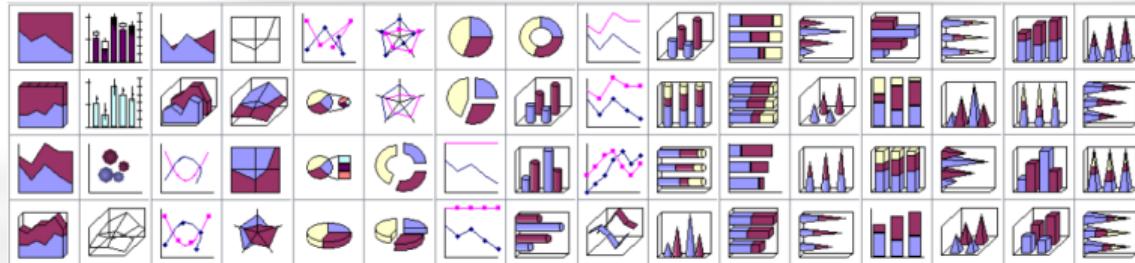
# Chord diagram

- ▶ **Nome:** Diagrama de cordas (*chord diagram*)
- ▶ **Função:** exibe o fluxo de quantidade/proporção de uma variável.
- ▶ **Tipo de escala:** quantitativa × categórica.
- ▶ **Variáveis visuais:** tamanho.
- ▶ **Geometria:** arcos.
- ▶ **Inspiração:** Sankey diagram e radial network.

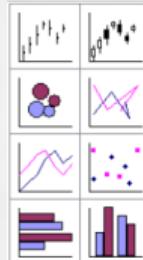
## O que você deve evitar

# O problema do viés de disponibilidade?

Lixo



Utéis



Adaptado: <https://excelcharts.com/excel-chart-gallery-a-difficult-equilibrium/>

# O problema do viés de disponibilidade?

**Figura 73.** Apenas 11% dos gráficos do MS Excel são úteis conforme Jorge Camoes. Adaptado: <<https://excelcharts.com/excel-chart-gallery-a-difficult-equilibrium/>>.

# O problema do viés de disponibilidade?

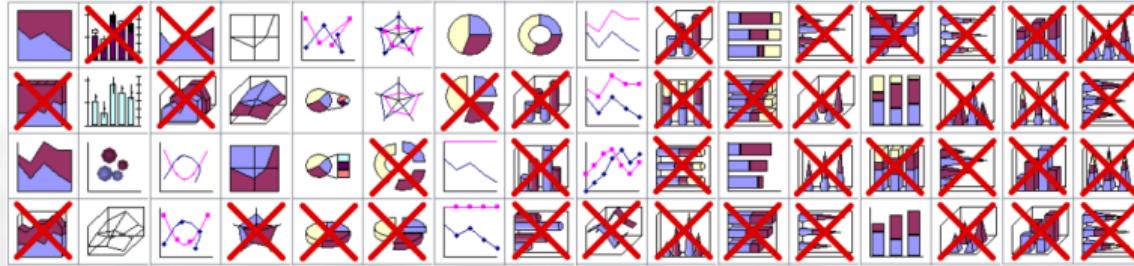


Figura 74. Gráficos de devem ser banidos.

# Gráficos ruins

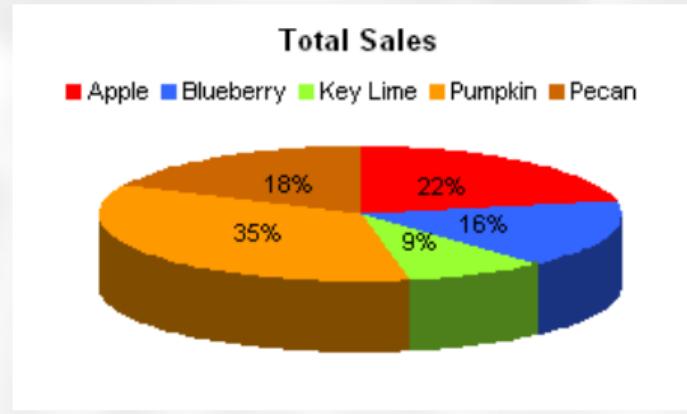


Figura 75. Gráfico de setores com perspectiva tridimensional gera distorção dos ângulos.

# Gráficos ruins

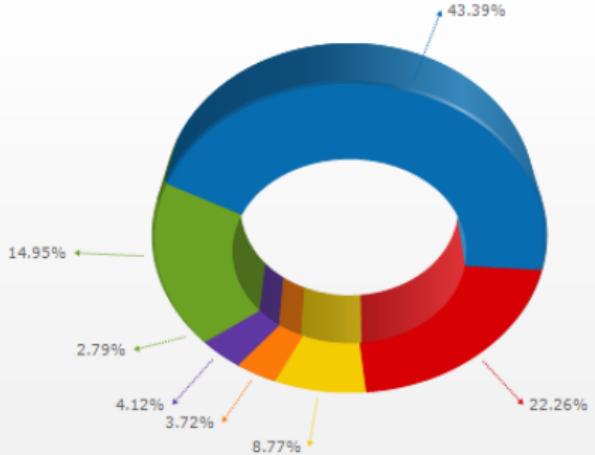
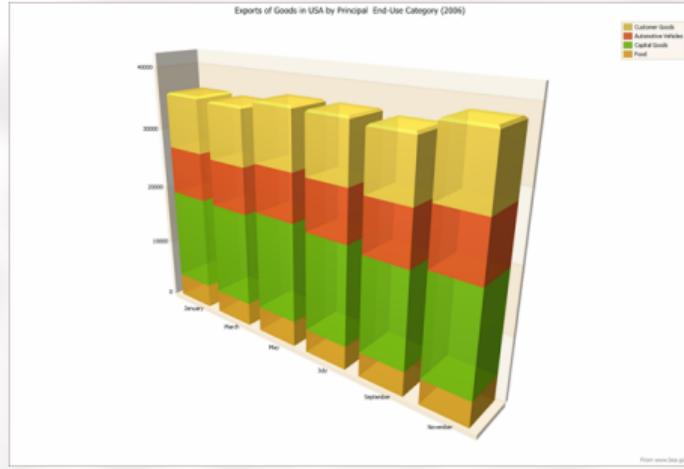


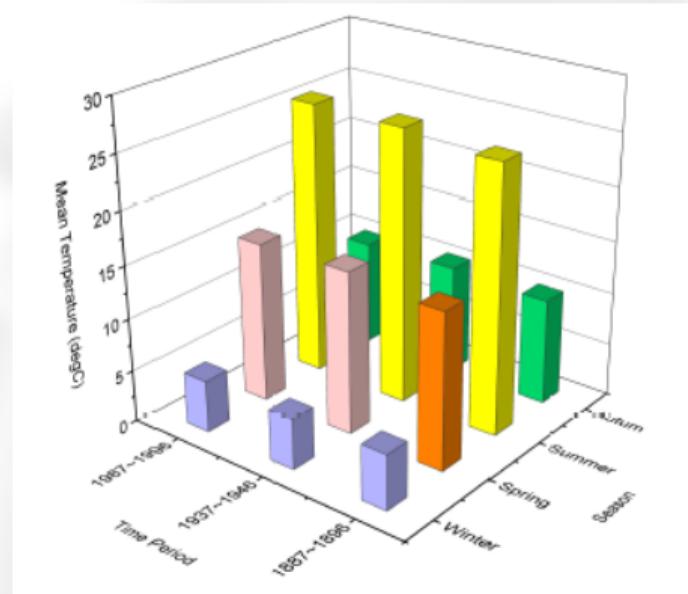
Figura 76. Comparação de comprimentos angulares é menos precisa, principalmente em uma perspectiva 3D.

# Gráficos ruins



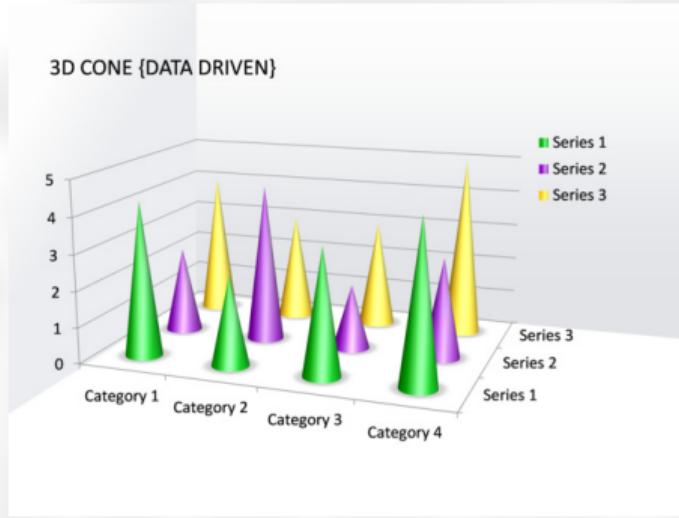
**Figura 77.** Perspectiva 3D e pontos de fuga deformam o sistema de coordenadas e dificultam perceber comprimentos.

# Gráficos ruins



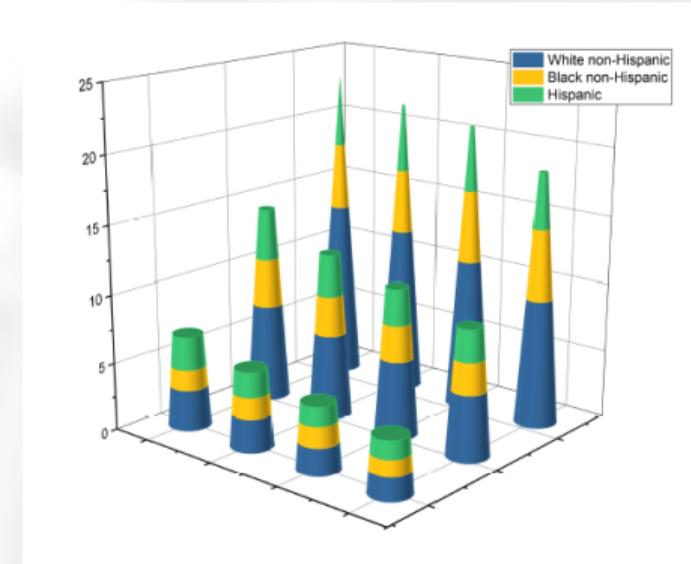
**Figura 78.** Mais do mesmo. Softwares comerciais de planilha eletrônica oferecem vários recursos gráficos tridimensionais com baixa efetividade. Isso popularizou ou mal uso de gráficos (*availability bias*).

# Gráficos ruins



**Figura 79.** Gráficos que violam os princípios de expressividade e efetividade.

# Gráficos ruins



**Figura 80.** Perspectiva e ornamentação que não agregam funcionalidade.  
<https://www.originlab.com/www/products/graphgallery.aspx>.

# Gráficos ruins

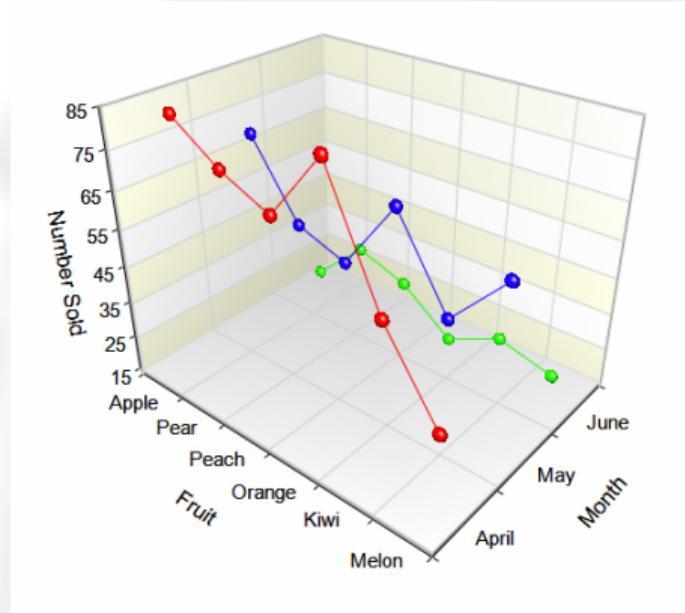
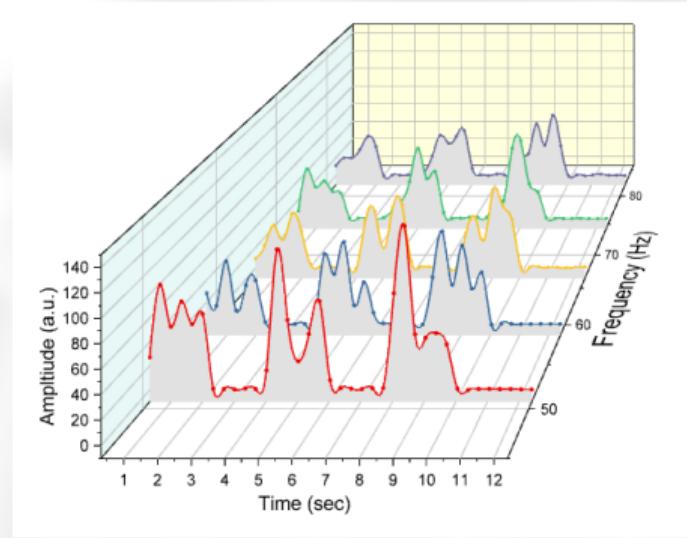


Figura 81. A falta de sistema de coordenadas para compreensão da informação o torna simplesmente uma alegoria.

# Gráficos ruins



**Figura 82.** A perspectiva reduz a acurácia da percepção.  
<https://www.originlab.com/www/products/graphgallery.aspx>.

# Gráficos ruins

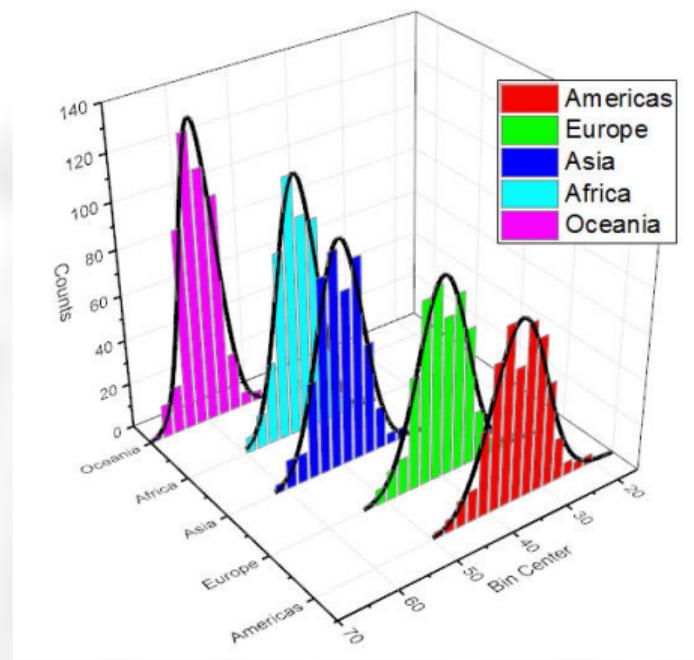


Figura 83. A perspectiva reduz a acurácia da percepção.  
<https://www.originlab.com/www/products/graphgallery.aspx>.

# Gráficos ruins

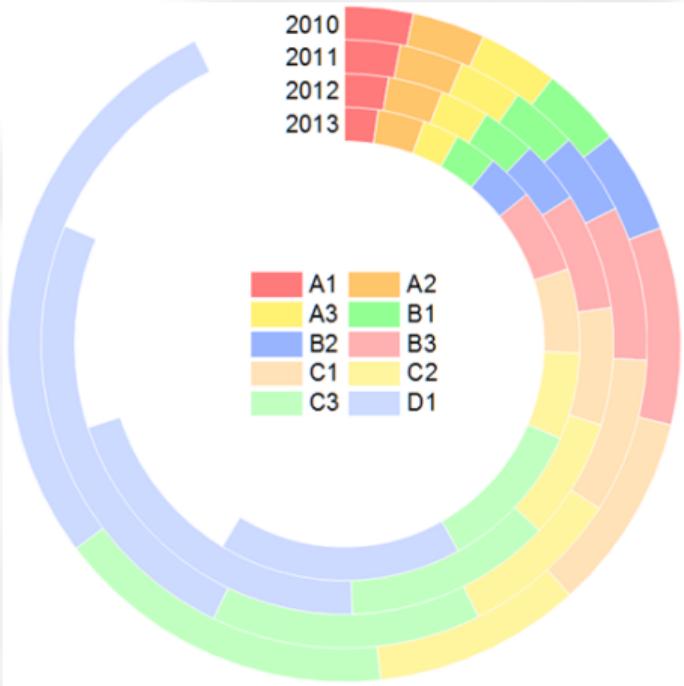


Figura 84. Visualizações polares têm menor acurácia para comparação dos comprimentos.  
<https://www.originlab.com/www/products/graphgallery.aspx>.

# Reflexão

*"O olho humano é um dispositivo maravilhoso. Com um pouco de esforço, pode deixar de ver até mesmo a injustiça mais flagrante."*

– Richard K. Morgan · Altered Carbon

*"Você obtém simplicidade na elaboração de gráficos e tabelas eficazes, lembrando três princípios fundamentais: restringir, reduzir, enfatizar."*

– Garr Reynolds

# Galerias online

1. <<https://www.data-to-viz.com/>>.
2. <<https://datavizcatalogue.com/>>.
3. <<https://www.r-graph-gallery.com/>>.
4. <<https://matplotlib.org/gallery/index.html>>.
5. <<https://public.tableau.com/en-us/s/gallery>>.
6. <<https://github.com/d3/d3/wiki/Gallery>>,  
<<http://biovisualize.github.io/d3visualization/>>.
7. <<http://jkunst.com/highcharter/>>.
8. <<https://developers.google.com/chart/interactive/docs/gallery>>.
9. <<https://dash.plot.ly/gallery>>.
10. <<https://www.zingchart.com/gallery/>>.
11. <<https://news-explorer.mybluemix.net/?query=data%20science>>.
12. <<https://truth-and-beauty.net/>>.

# Próximo encontro

- ▶ Recursos para elaboração de gráficos no R.
- ▶ Análise exploratória com *data storytelling*.
- ▶ Introdução ao pacote ggplot2.

**Pela oportunidade e atenção, agradeço.**

Prof. Walmes Zeviani  
[<walmes@ufpr.br>](mailto:walmes@ufpr.br)