GENE7033 – Tópicos Especiais em Genética I:

Visualização de dados para publicações científicas

Profª Drª Chirlei Glienke

Drª Desirrê Petters-Vandresen

Visualizando associação entre variáveis

Drª Desirrê Petters-Vandresen

01/12/2022

Finalidade

- Visualização de como duas ou mais variáveis quantitativas estão relacionadas entre si

- Exemplos
  - Associação entre quantidade de eletrodomésticos e consumo de energia em uma residência
  - Dimensões corporais de um animal e necessidades energéticas diárias

- Escolha de gráfico de acordo com a quantidade de variáveis:
  - Duas variáveis: gráfico de dispersão
  - Mais de duas variáveis: gráfico de bolhas, matriz de dispersão, correlogramas
  - Múltiplas variáveis: redução de dimensionalidade (ex: análise de componentes principais)

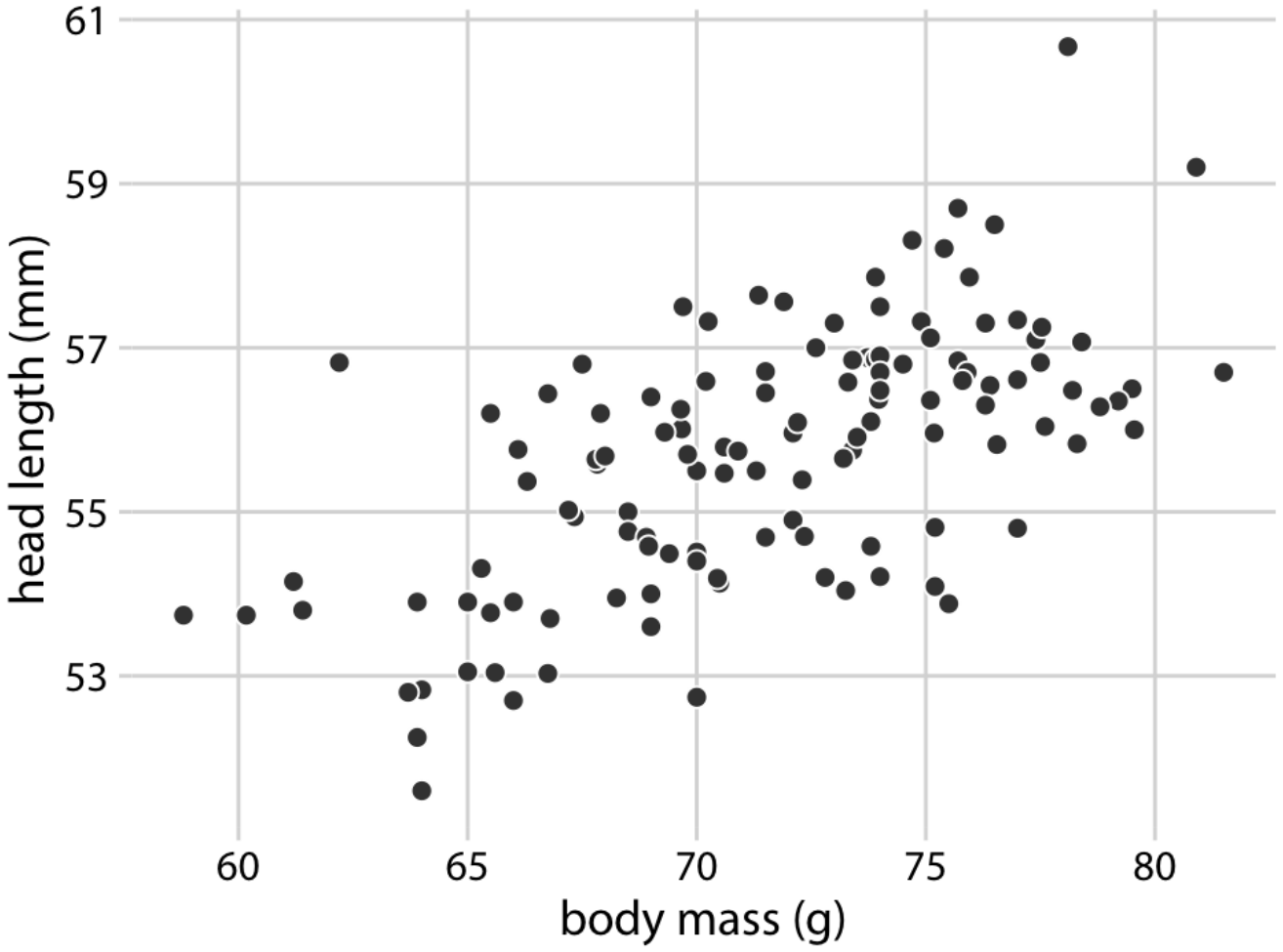
Gráfico de dispersão

- Conjunto de dados com medidas de 123 indivíduos de gaio-azul
  - Comprimento da cabeça (da ponta do bico até a parte de trás da cabeça)
  - Comprimento do crânio (comprimento da cabeça menos o comprimento do bico)
  - Massa corporal
- Relações esperadas:
  - Pássaros com bicos mais compridos teriam crânios maiores
  - Pássaros com maior massa corporal teriam crânios e bicos maiores

Fonte

Gráfico de dispersão

- Plotar o comprimento da cabeça (eixo y) vs. massa corporal (eixo x)
  - Sempre descrevemos “variável do eixo y plotada contra/vs. variável do eixo x”
- Cada indivíduo representado por um ponto
- Apesar de existir dispersão, em geral, pássaros de maior massa corporal tem cabeças maiores

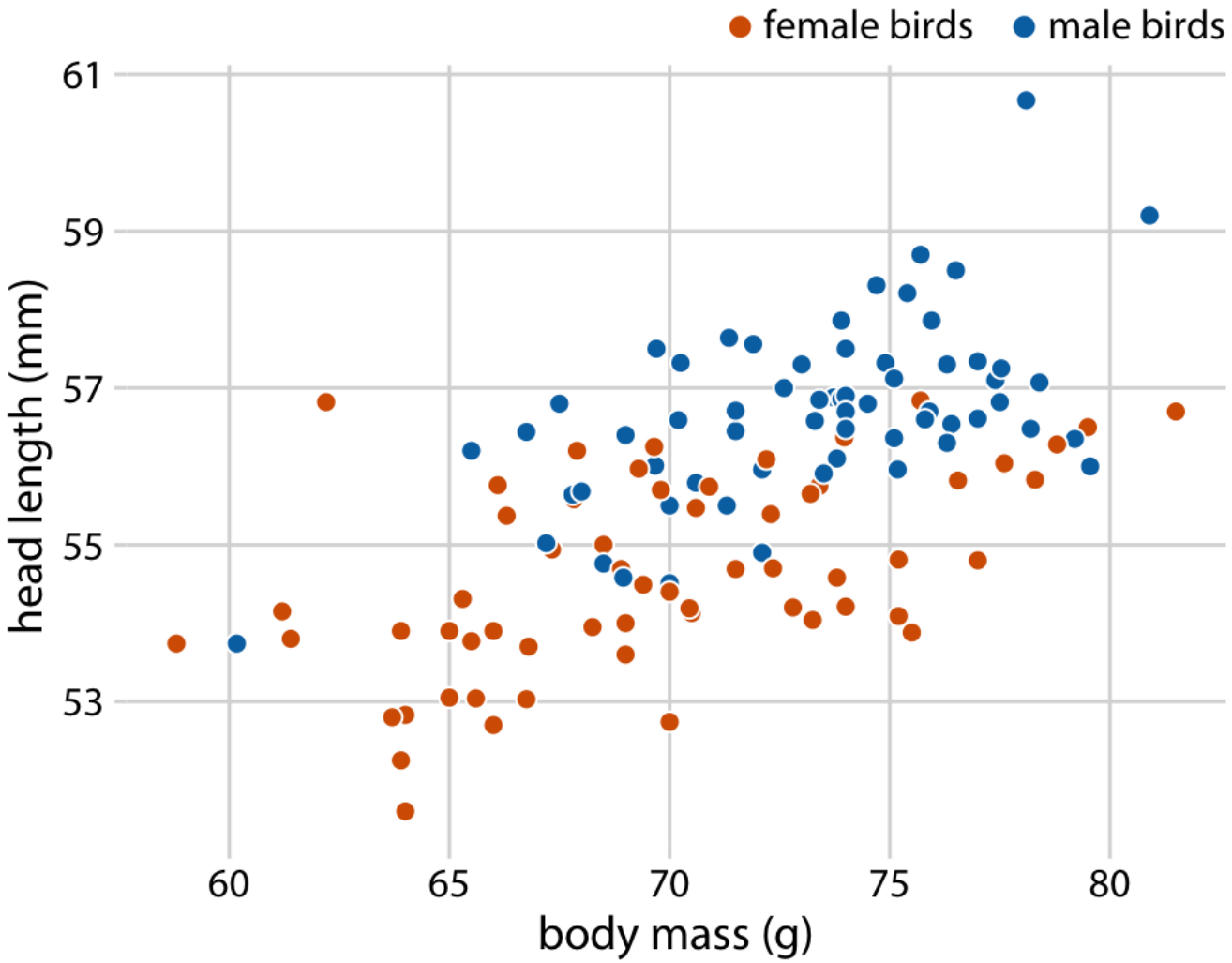


Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O’Reilly Media, 2019.

Gráfico de dispersão

- Possibilidade de adicionar mais um nível de codificação com o uso de cores

- Parte da diferenciação observada se deve ao sexo dos pássaros:
  - Para a mesma massa corporal, fêmeas tem cabeças menores que os machos
  - Fêmeas são mais leves que os machos de forma geral



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Gráfico de bolhas (bubble chart)

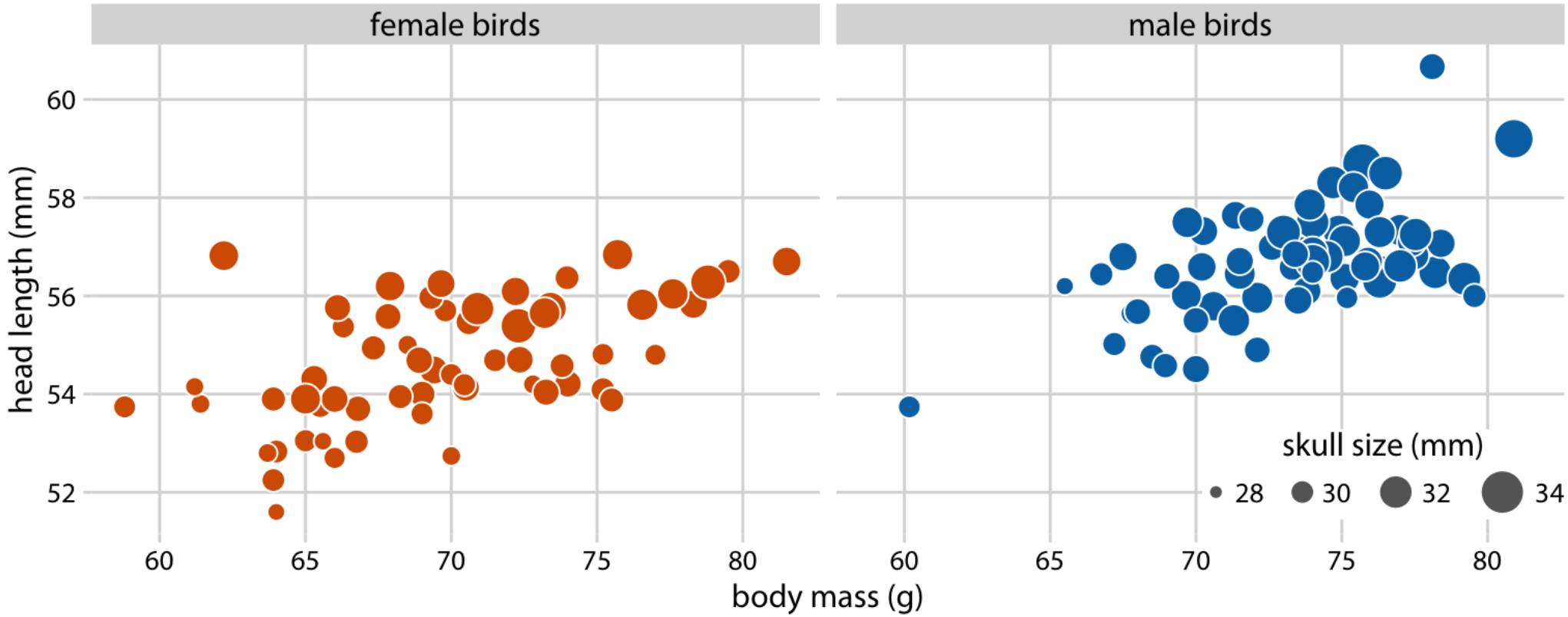
- Comprimento da cabeça definido como distância entre a ponta do bico até a parte de trás da cabeça

- Comprimento grande pode indicar:
  - Bicos longos
  - Crânios longos
  - Bicos e crânios longos

- Adicionar mais uma variável no gráfico por meio da variação no tamanho dos pontos

Gráfico de bolhas (bubble chart)

- Tamanho do crânio representado pela variação no tamanho dos pontos: em geral, comprimento da cabeça está relacionado ao tamanho do crânio, mas há exceções



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Gráfico de bolhas (bubble chart)

- Desvantagem:
  - Variáveis quantitativas com dois tipos diferentes de escala (posição e tamanho):
    - Dificuldade de percepção da existência e relevância da associação entre as variáveis apresentadas
  - Dificuldade de percepção dos valores codificados pelo tamanho do pontos em relação aos valores codificados por posição
    - Dificuldade de interpretar áreas circulares

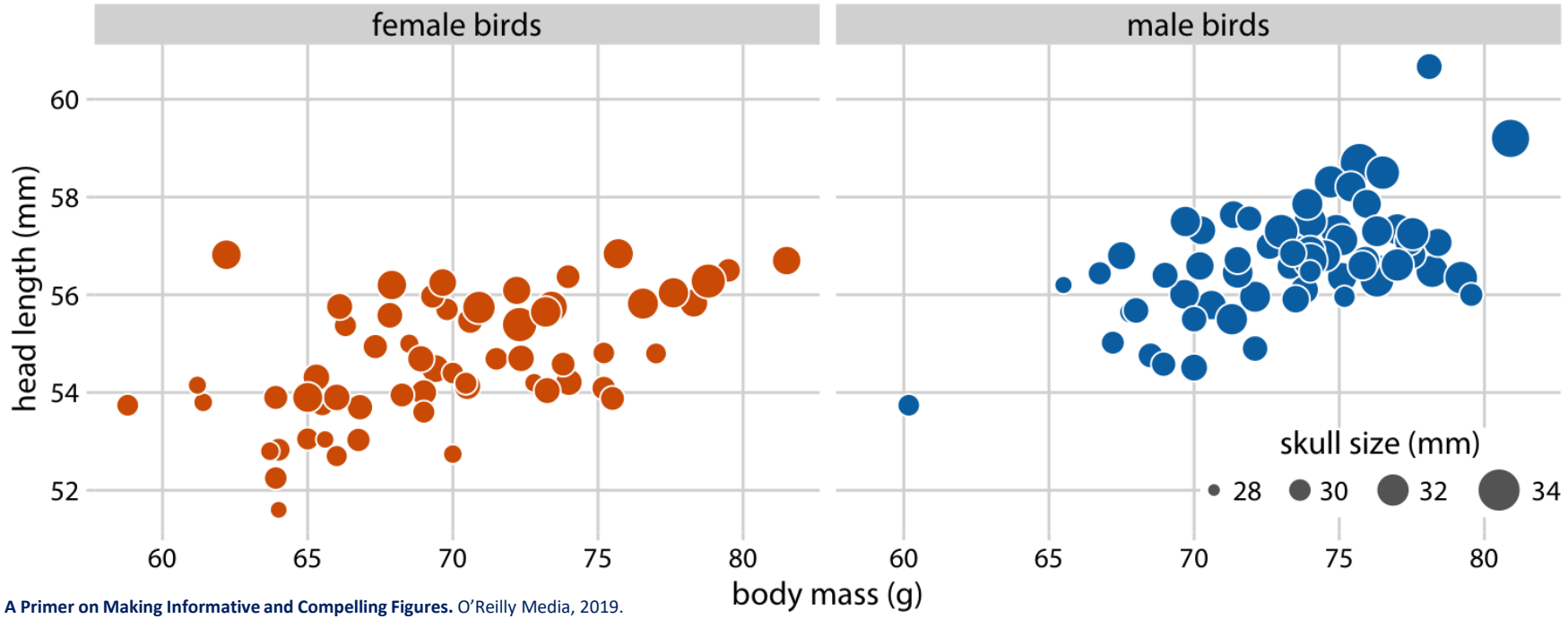
Gráfico de bolhas (bubble chart)

- Desvantagem:
  - Restrição de tamanho:
    - Mesmo as bolhas que representam valores maiores precisam ser pequenas em relação ao tamanho total da figura
    - Diminuição do nível de diferença de tamanho para comparação entre bolhas grandes e pequenas
    - Diferenças pequenas nos valores correspondem a diferenças quase imperceptíveis no tamanho das bolhas, praticamente impossíveis de visualizar

Gráfico de bolhas (bubble chart)

- Padrão de tamanho utilizado amplifica a diferença de tamanho entre os crânios pequenos (~28mm) e os maiores (~34mm)

- Ainda é difícil perceber a relação entre tamanho do crânio e massa corporal/tamanho da cabeça



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Matriz de dispersão

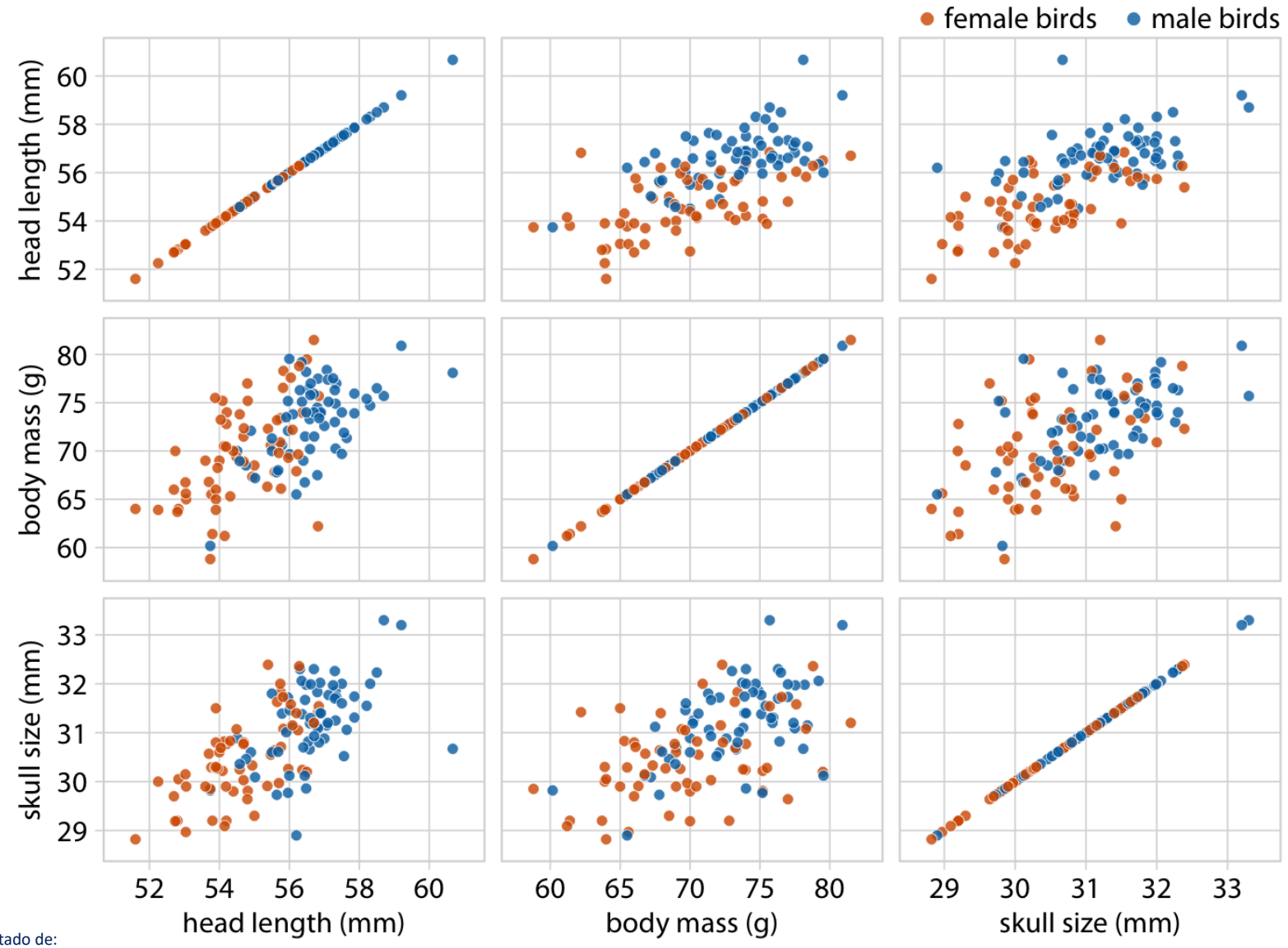
- Alternativa ao gráfico de bolhas, apresentando gráficos de dispersão individuais para cada combinação de variáveis (“all-against-all”)

- Como a percepção de diferenças entre posições funciona melhor que a diferença de tamanhos, correlações são percebidas mais facilmente

- Desvantagem: para uma grande quantidade de variáveis, podem ocupar muito espaço
  - Selecionar as variáveis mais relevantes para apresentar na figura principal, e apresentar a figura completa com todas as variáveis em material suplementar

Matriz de dispersão

- Relação entre o tamanho do crânio e a massa corporal é similar entre machos e fêmeas
- Relação entre tamanho da cabeça e massa corporal apresenta uma separação por sexo
  - Machos tendem a ter bicos mais longos que as fêmeas



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Correlogramas

- Matrizes de dispersão se tornam inviáveis para muito mais de três/quatro variáveis quantitativas

- Alternativa: visualizar o nível de associação entre pares de variáveis nos gráficos, ao invés dos dados brutos

- Cálculo de coeficiente de correlação: em geral, quando não especificado no gráfico/legenda, tende a ser o coeficiente de correlação de Pearson

Correlogramas

- Coeficiente de correlação de Pearson:
  - Grau de correlação linear entre duas variáveis quantitativas, variando entre -1 e 1
  - 0 indica ausência de associação linear, -1 e 1 indicam associação linear perfeita
  - Valores positivos: correlação positiva, valores maiores numa das variáveis coincidem com valores maiores na outra
  - Valores negativos: correlação negativa (anticorrelação), valores maiores numa das variáveis coincidem com valores menores na outra

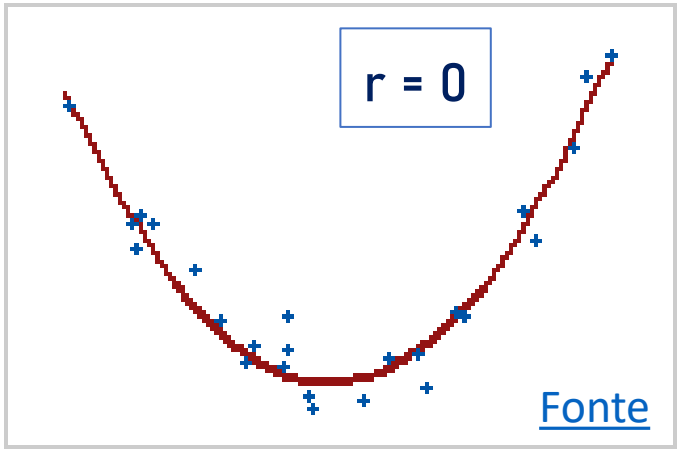
Correlogramas

- Coeficiente de correlação de Spearman:
  - Método não paramétrico, não pressupõe relação linear entre as variáveis, e não requer variáveis quantitativas
  - Utilização de postos/rankings das variáveis e avaliação da correlação entre os rankings
  - Também varia entre -1 e 1
  - 0 indica ausência de associação entre os rankings, -1 e 1 indicam associação linear perfeita
  - Valores positivos: correlação positiva, rankings maiores numa das variáveis coincidem com rankings maiores na outra
  - Valores negativos: correlação negativa (anticorrelação), rankings maiores numa das variáveis coincidem com rankings menores na outra

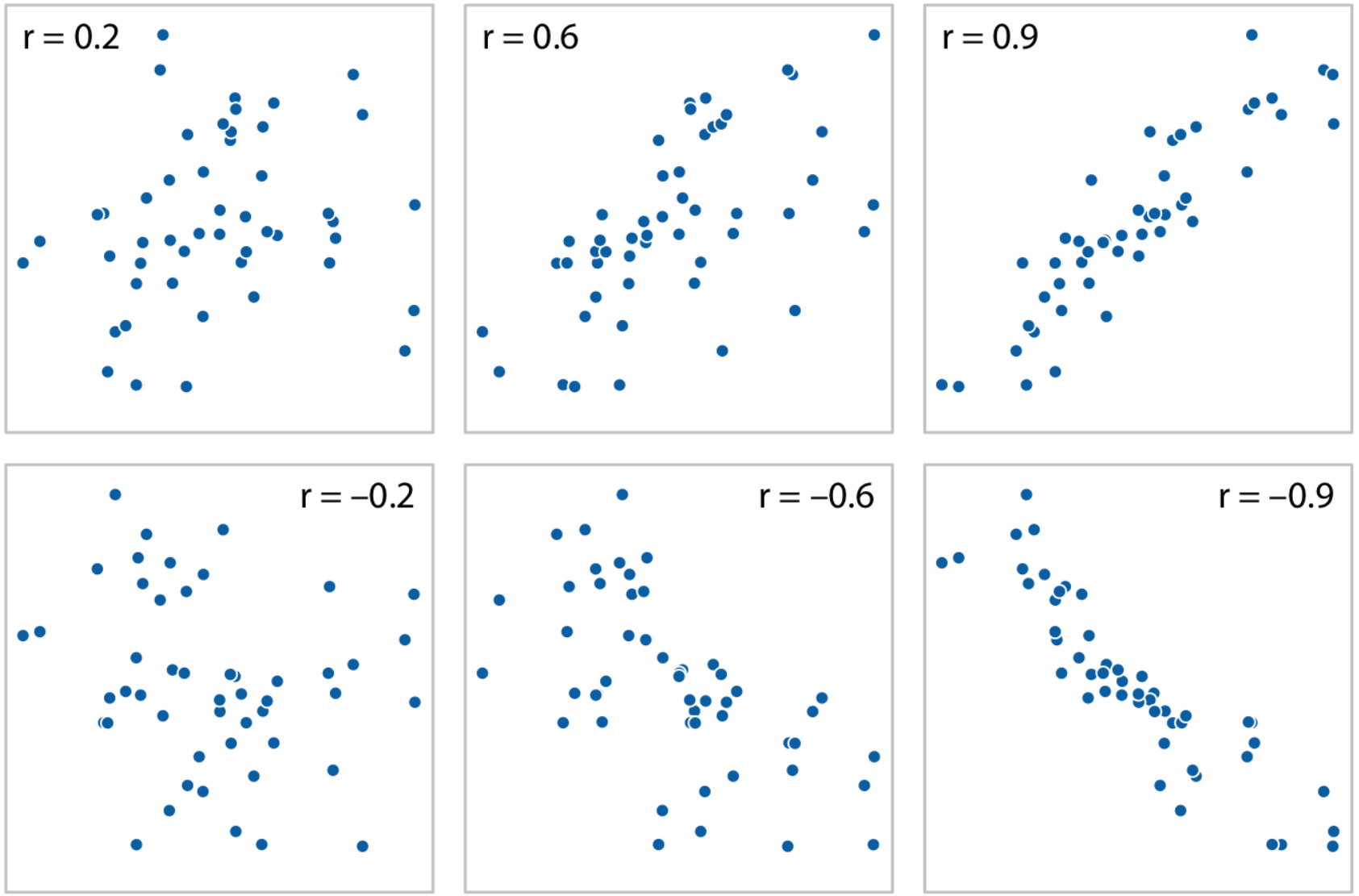
Correlogramas

- Exemplos de diferentes níveis de correlação para relações lineares

- Coeficiente 0 indica ausência de correlação linear, mas ainda pode existir relação entre as variáveis

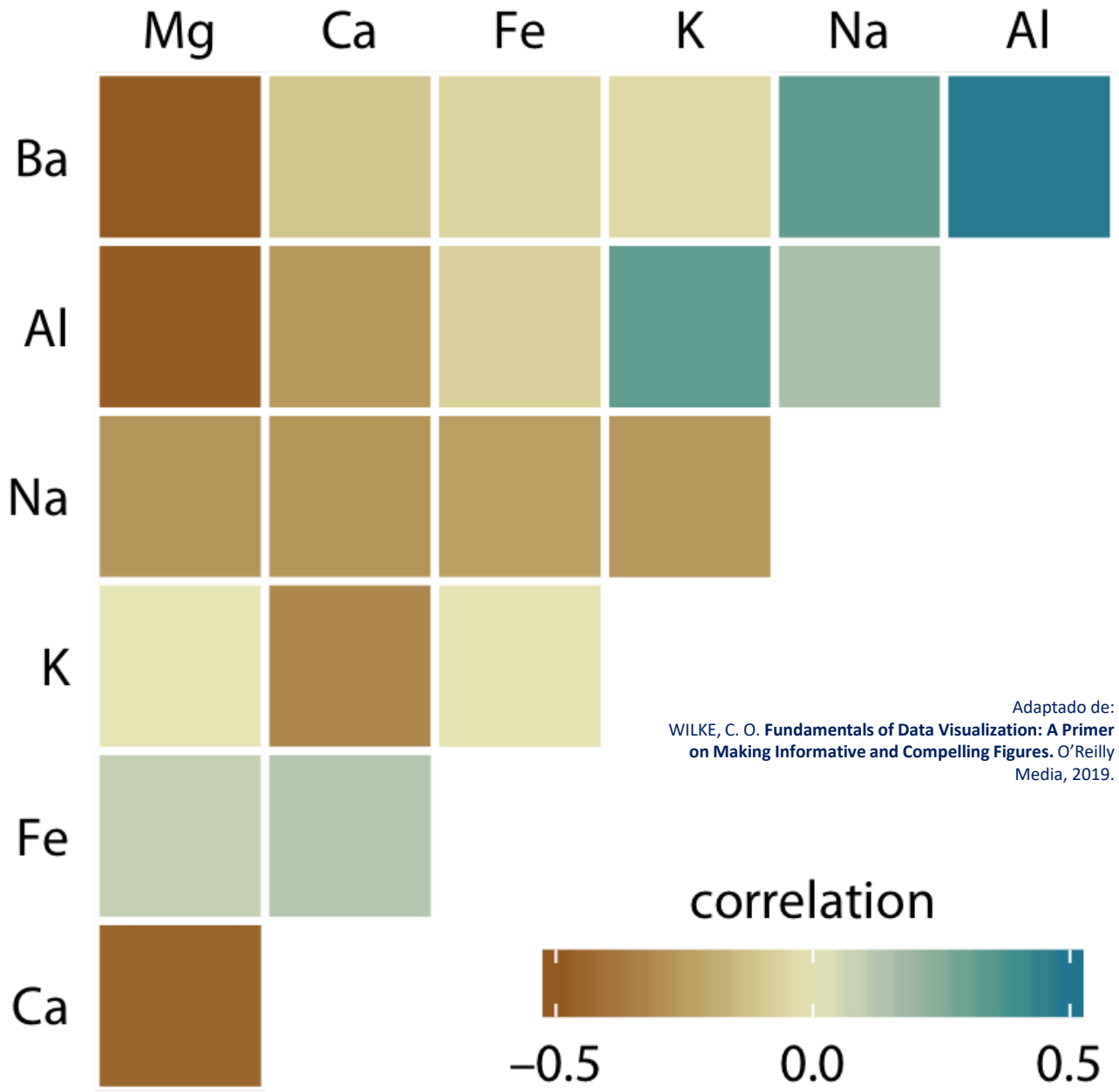


Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.



Correlogramas

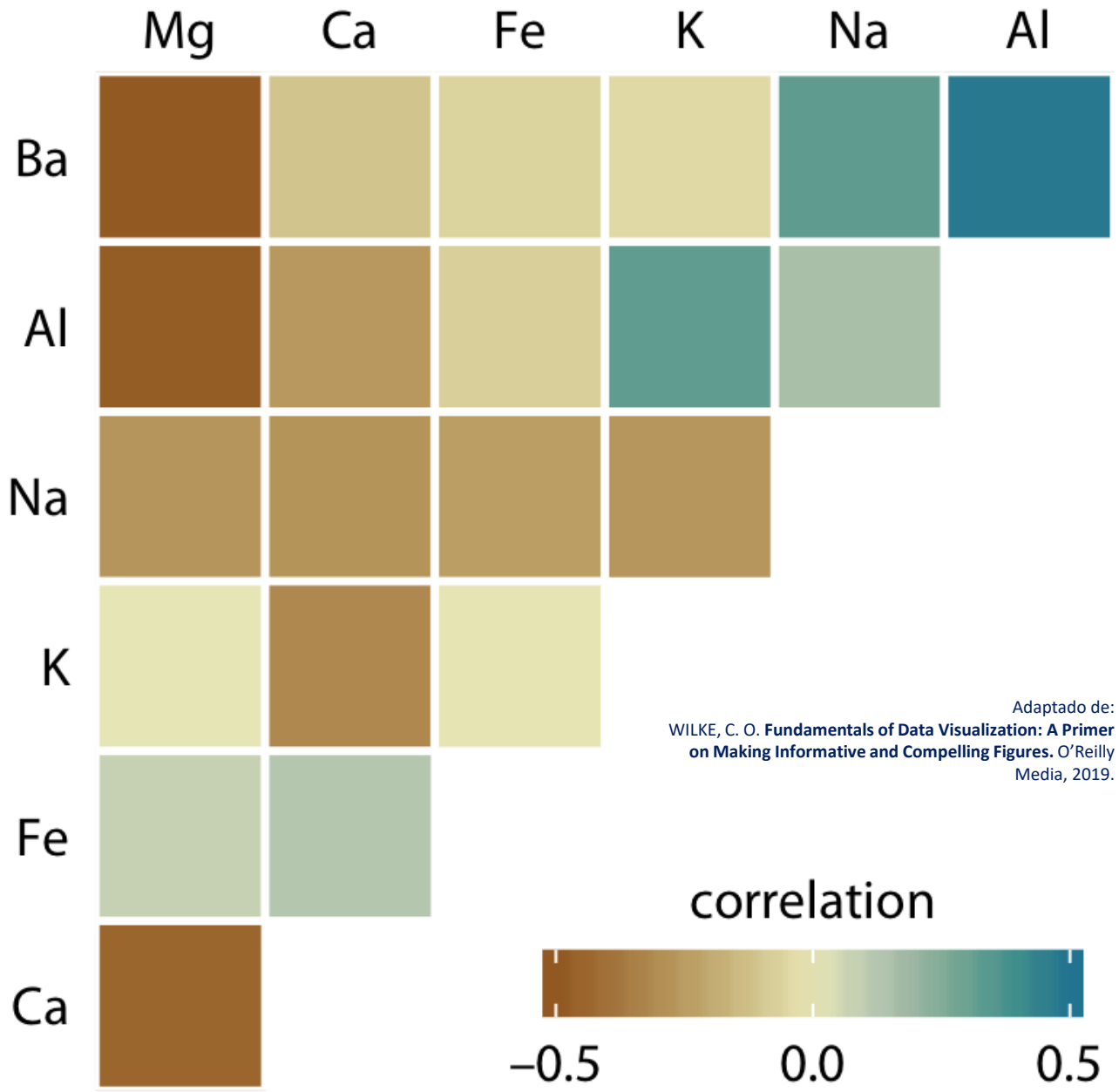
- Conjunto de dados de 200 fragmentos de vidro obtidos a partir de amostras forenses
  - Medidas de composição (porcentagem de diferentes óxidos minerais na composição)
  - Sete óxidos diferentes: total de 21 comparações par a par
  - Matriz de comparação com codificação por cor para representar o coeficiente de correlação



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Correlogramas

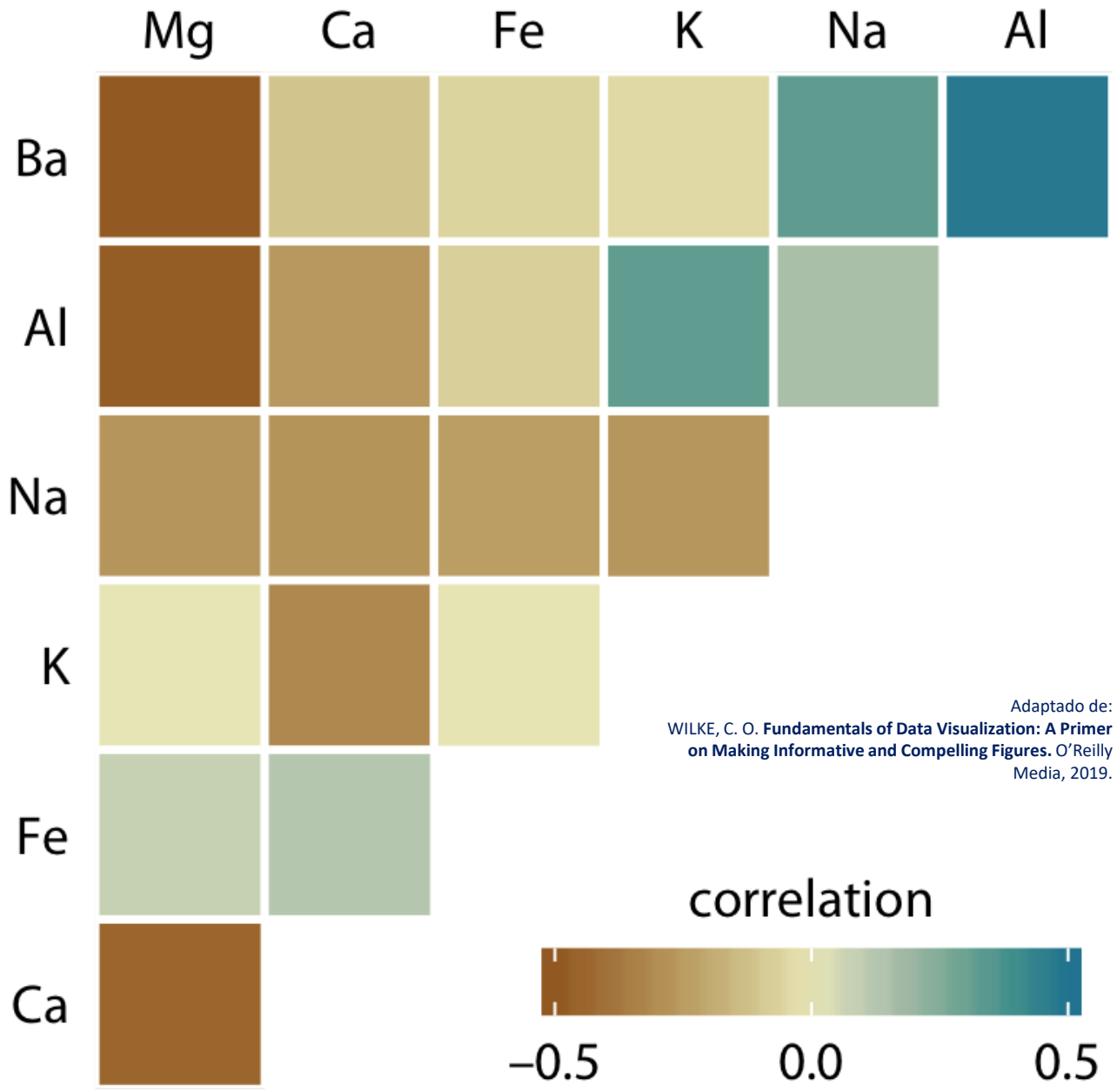
- Conjunto de dados de 200 fragmentos de vidro obtidos a partir de amostras forenses
  - Óxidos de magnésio negativamente correlacionados com quase todos os outros óxidos
  - Correlação positiva entre óxidos de alumínio e bário



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Correlogramas

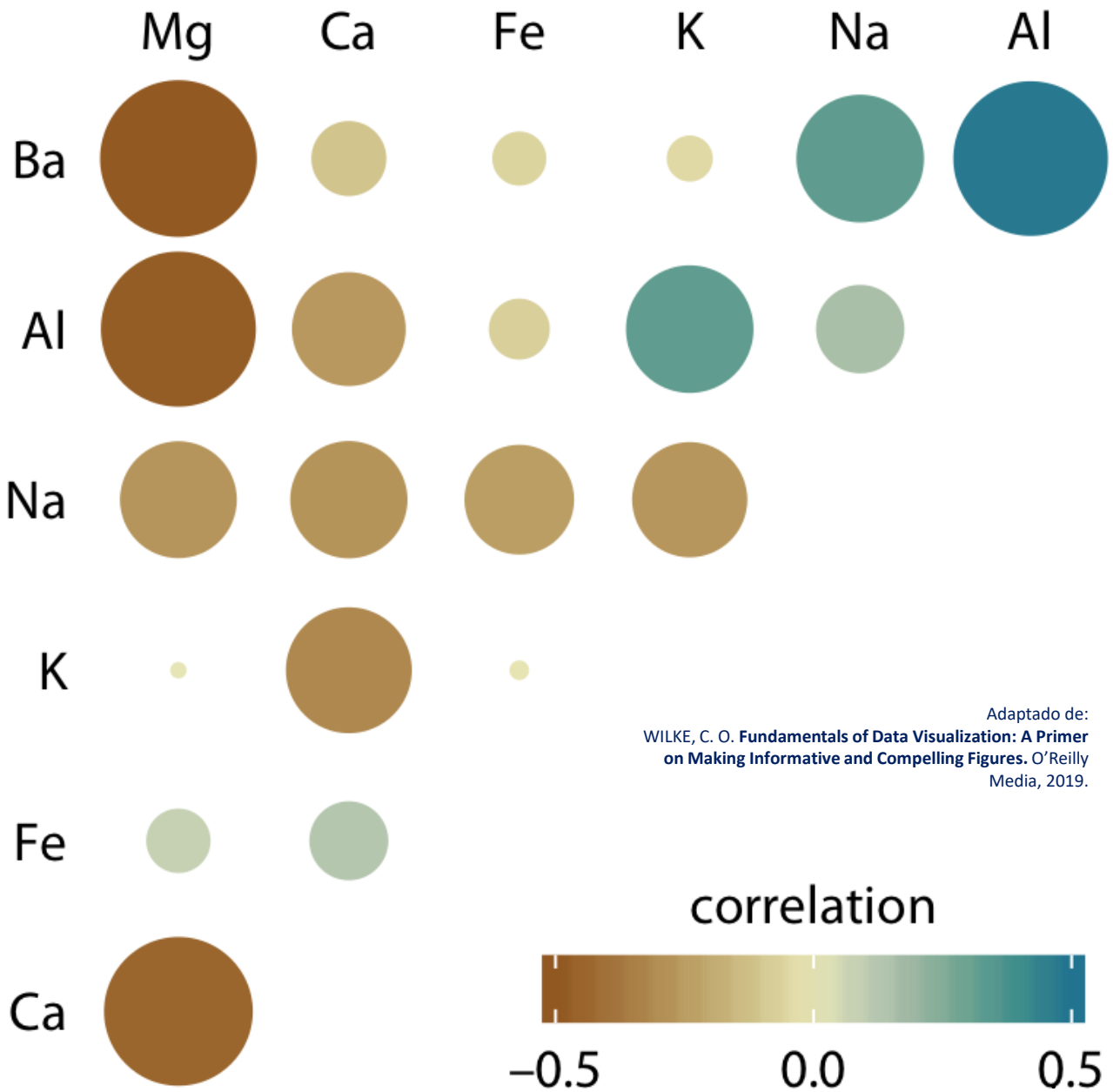
- Desvantagem: ausência de correlação ainda tem peso visual no gráfico:
  - Óxidos de magnésio e potássio não estão correlacionados, mas a percepção não é óbvia
  - Óxidos de magnésio e ferro são ligeiramente mais correlacionados que os de magnésio e potássio



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Correlogramas

- Alternativa:
  - Além da codificação por cor, ajustar o tamanho dos círculos de acordo com o valor absoluto do coeficiente de correlação
  - Diminuição do peso visual das comparações em que há ausência de correlação



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Correlogramas

- Desvantagem geral dos correlogramas: altamente abstratos
- Bons para demonstrar padrões gerais, mas dificuldade de representar valores precisos e exatos
- Sempre preferível representar os dados brutos ao invés de quantidades e valores derivados a partir dos dados brutos

Redução de dimensionalidade

- Meio termo entre apresentar padrões gerais e ao mesmo tempo apresentar os dados brutos

- Premissa geral: a maior parte dos conjuntos de dados multidimensionais consiste em múltiplas variáveis correlacionadas que possuem sobreposição de informação

- Se há sobreposição, é possível reduzir a quantidade de dimensões do conjunto de dados para poucas dimensões essenciais ao entendimento, sem perda de informação

Redução de dimensionalidade

- Por exemplo, conjunto de dados com medidas de características físicas de indivíduos:
  - Altura, largura, comprimento de braços e pernas, circunferência de cintura, quadril, tórax

- Todas as variáveis estão relacionadas ao tamanho geral do indivíduo
  - Em geral, pessoas maiores serão mais altas, mais largas, com braços e pernas mais compridos e maiores circunferências de cintura, quadril, tórax

- Variáveis também influenciadas pelo sexo
  - Em geral, mulheres tem a ter circunferências de quadril maiores que de homens

Redução de dimensionalidade - PCA

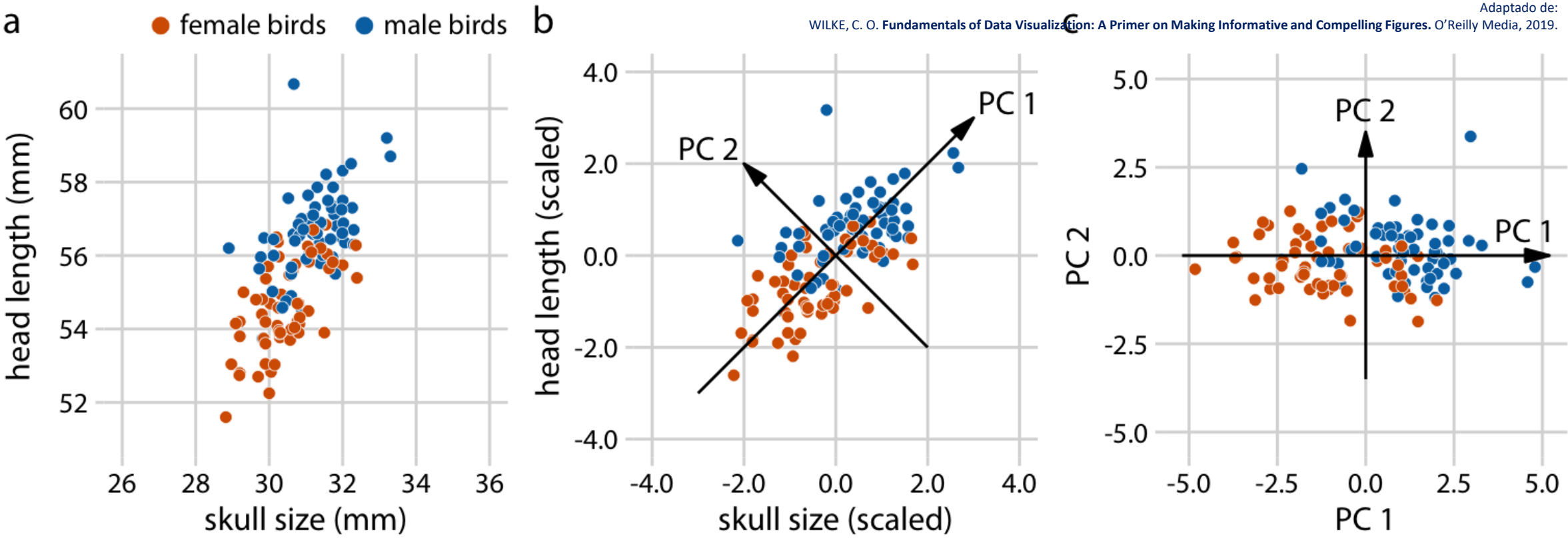
- Várias possibilidades de redução de dimensionalidade, análise de componentes principais (*principal component analysis*, PCA) é a mais difundida

- Criação de um novo conjunto de variáveis (componentes principais, PCs) pela combinação linear das variáveis originais no conjunto de dados, escalonadas para média 0 e variância 1

- Escolha dos PCs para que são não correlacionados, e ordenados de modo que o primeiro componente explique a maior parte da variação existente nos dados, e assim sucessivamente

- Em geral, as principais características do conjunto de dados tendem a ser demonstradas pelo uso dos primeiros dois ou três PCs

Redução de dimensionalidade - PCA

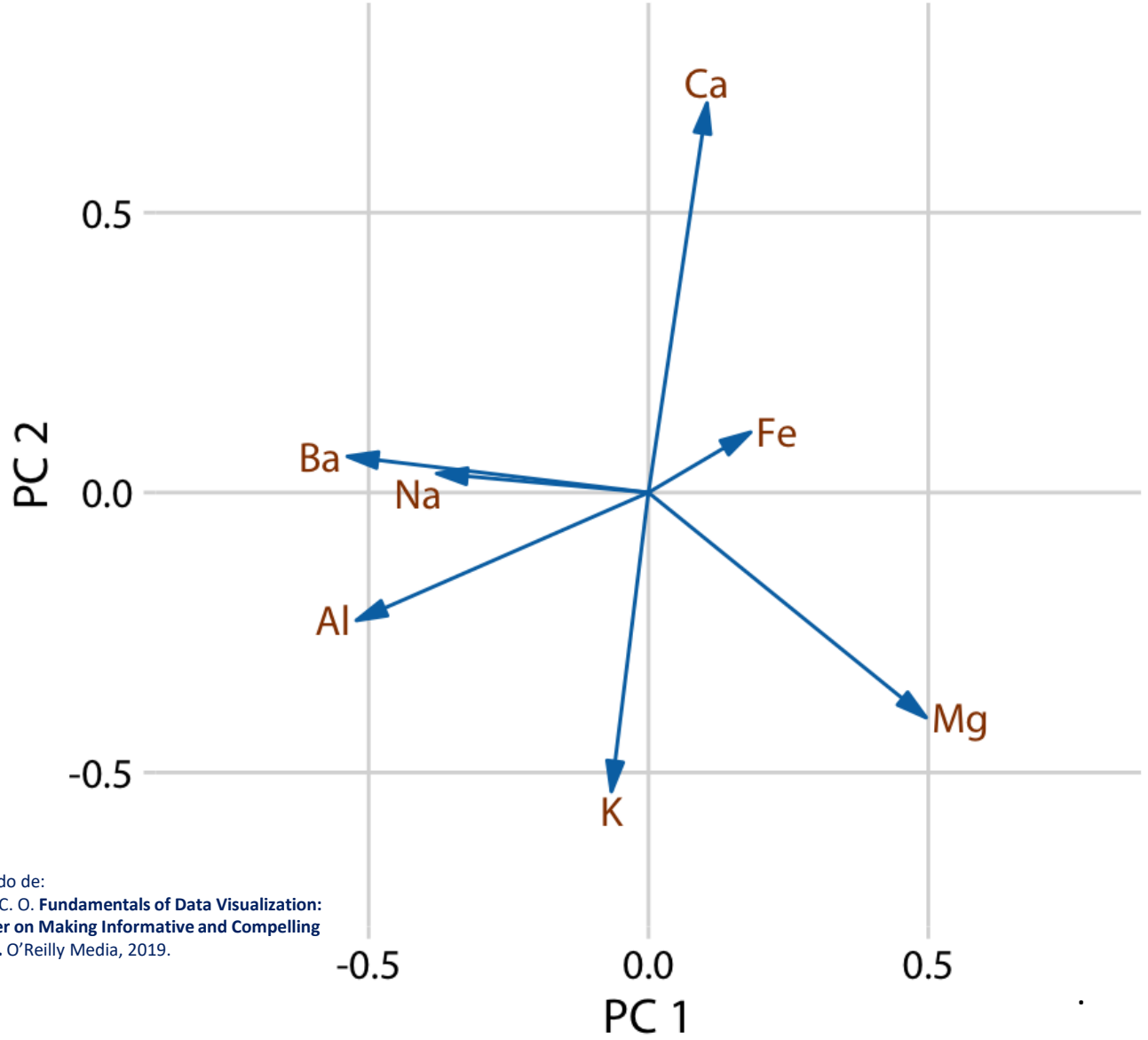


Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

- Dados escalonados para média 0 e variância 1, definição dos PCs ao longo das direções de maior variação dos dados, e projeção dos dados nas novas coordenadas

Redução de dimensionalidade - PCA

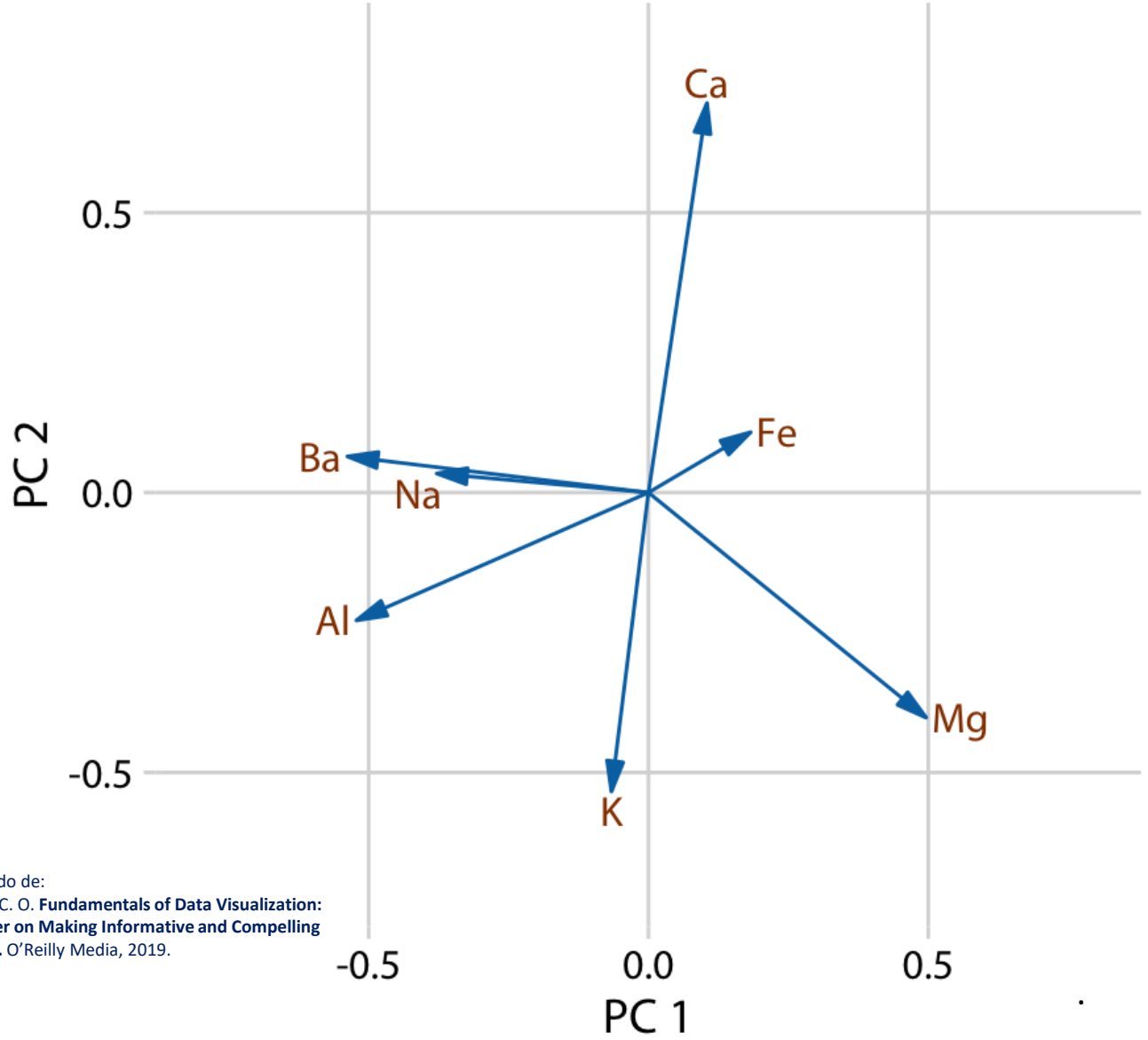
- Principais informações de interesse:
  - Composição dos PCs
  - Localização individual de cada observação no espaço dos PCs
- Se os PCs são combinações lineares das variáveis originais, cada variável pode ser representada como flechas, indicando sua contribuição para cada PC



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

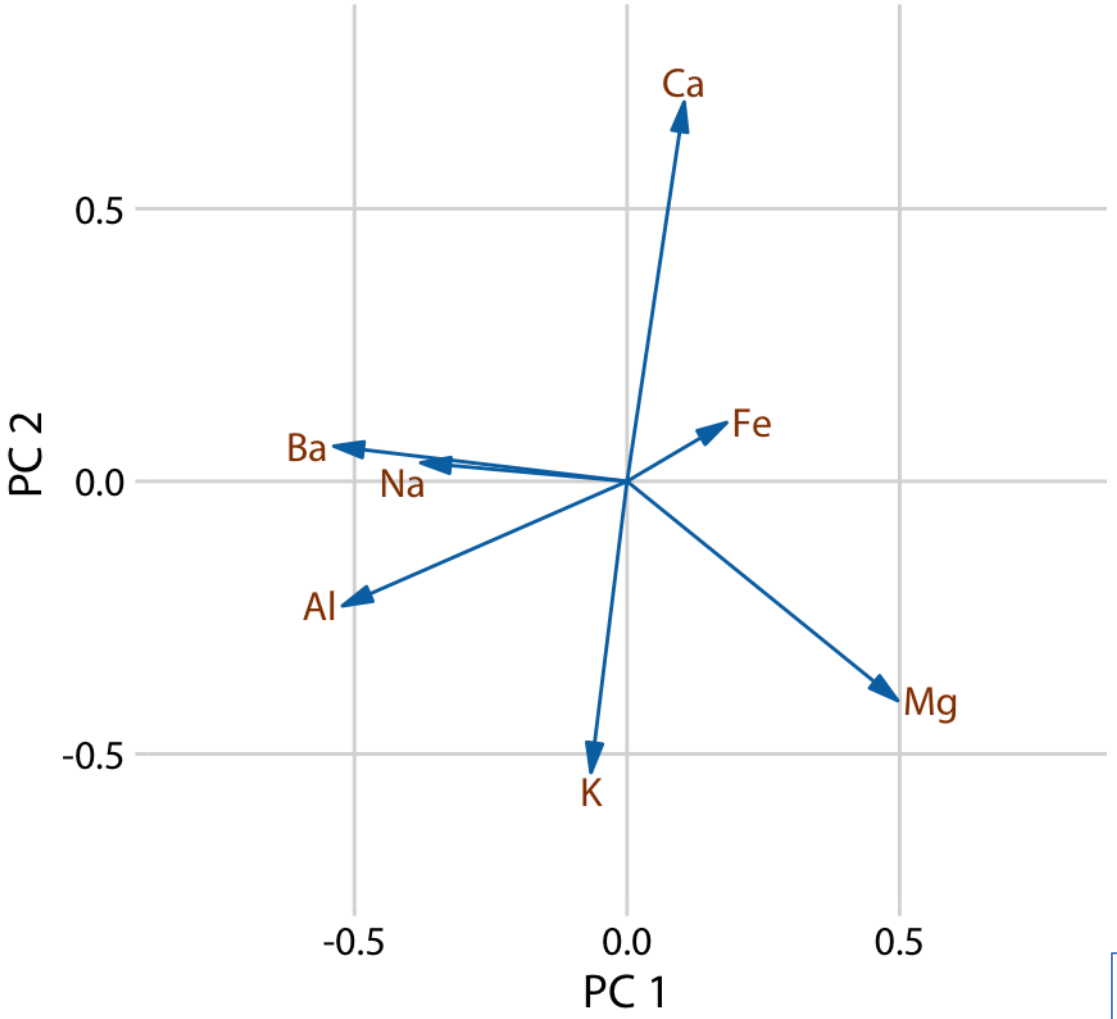
Redução de dimensionalidade - PCA

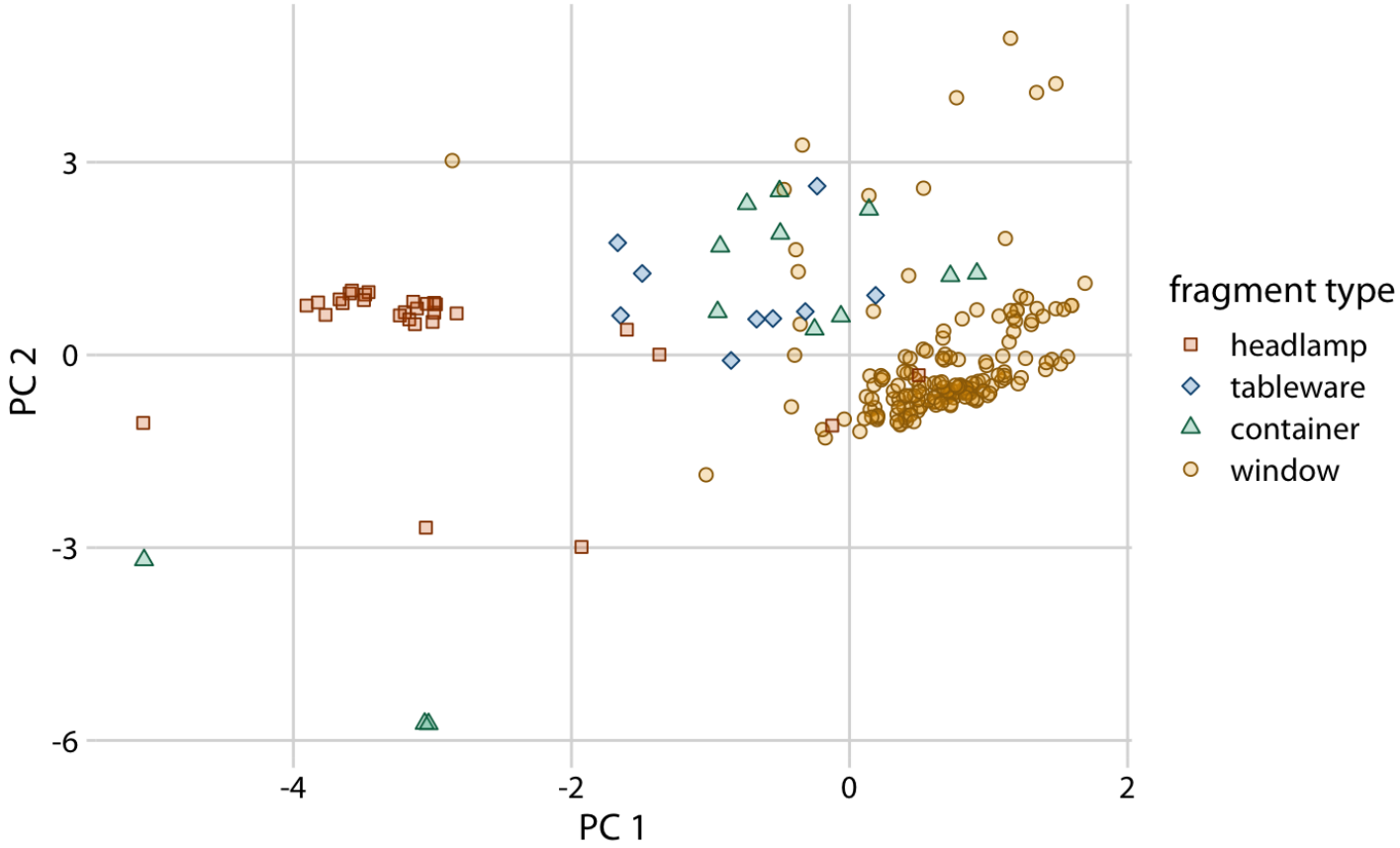
- PC1: contribuição principal por bário e sódio
- PC2: contribuição principal por cálcio e potássio
- Restante dos óxidos contribuindo tanto para PC1 e PC2, de formas distintas
- Comprimento das flechas varia pela existência de mais de dois PCs (não representados no gráfico)



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Redução de dimensionalidade - PCA



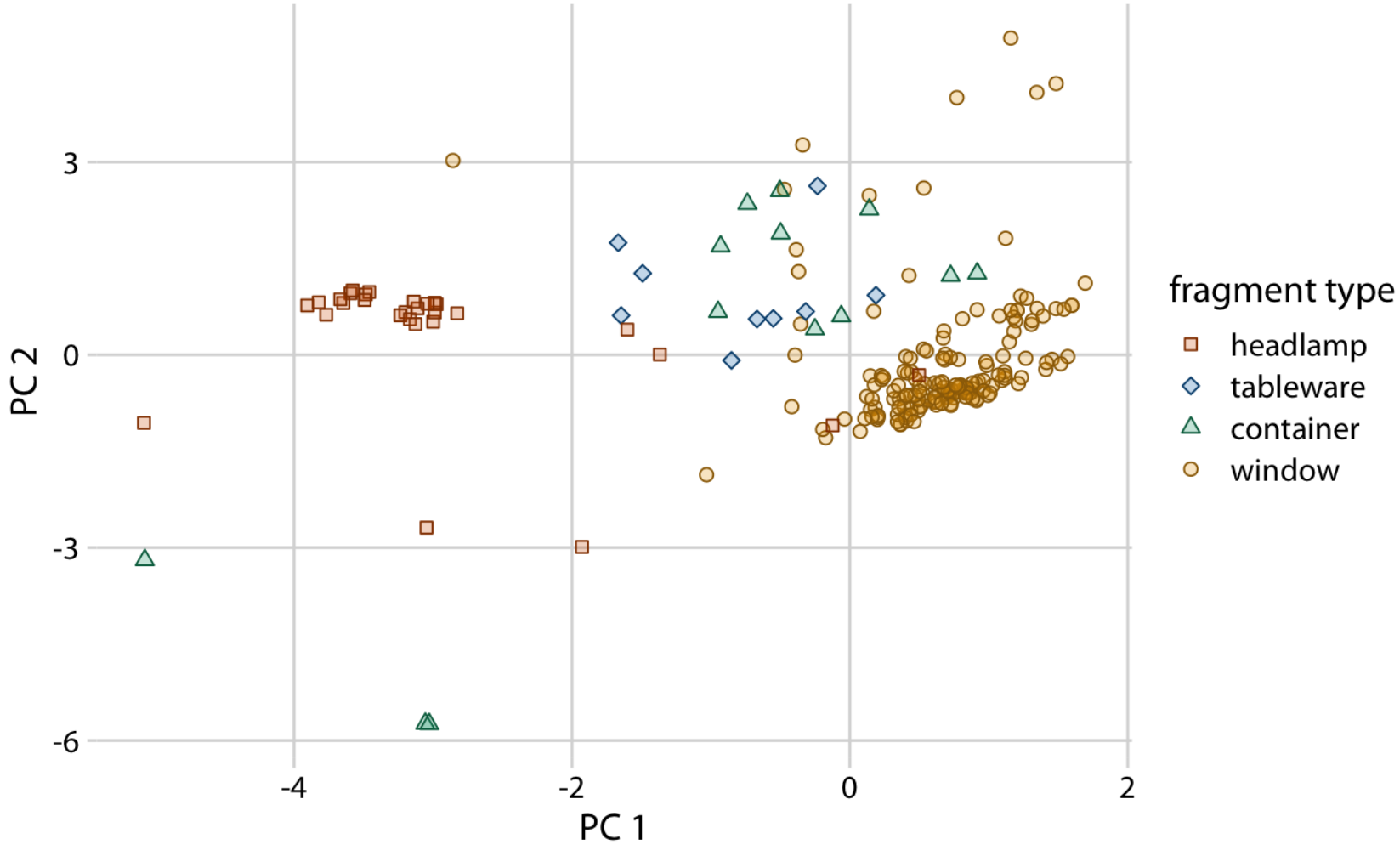


Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Fragmentos provenientes de janelas apresentam mais magnésio e menos bário, sódio e alumínio que a média

Redução de dimensionalidade - PCA

- Projeção dos dados originais no espaço dos PCs
- Agrupamento de tipos distintos de fragmentos de vidro (codificação por cor e forma)
- Diferenciação clara entre vidro de janelas e vidro de farol de carros
- Sobreposição entre utensílios de mesa e containers, mas distintos de janelas e faróis



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Dados pareados

- Conjuntos de dados em que duas ou mais observações de uma mesma variável foram feitas em condições ligeiramente diferentes

- Exemplos:
  - Comprimento do braço direito e do braço esquerdo de um indivíduo
  - Peso de um indivíduo em diferentes momentos de um ano
  - Altura de dois gêmeos idênticos

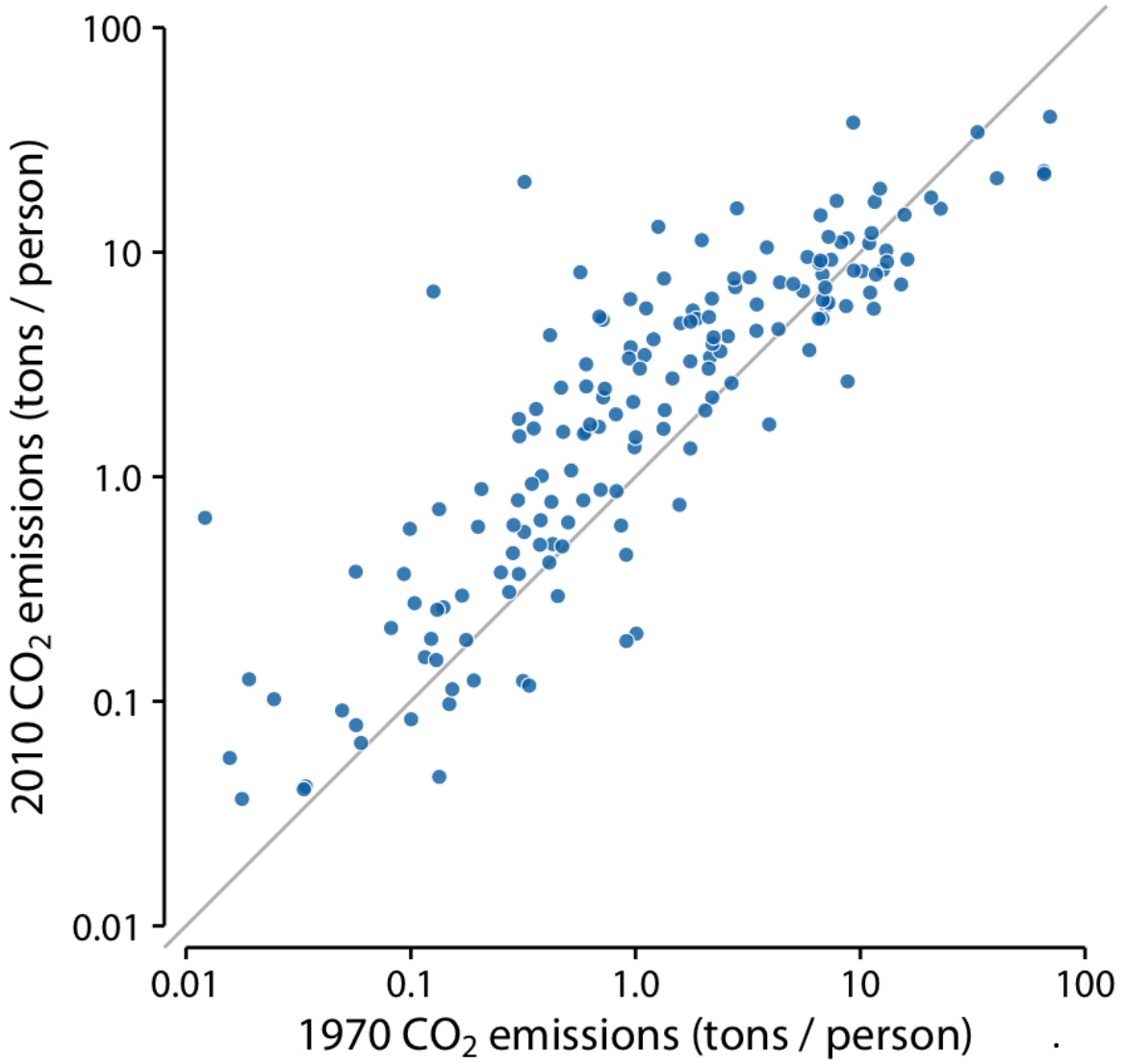
- Em dados pareados, a premissa é de que medições pertencentes a um par serão mais parecidas entre si do que com medidas pertencentes a outros pares

- Estratégias de visualização precisam reforçar esta premissa, e também eventuais diferenças que existam dentro de pares

Dados pareados – Gráfico de dispersão

- Plotar os dados junto de uma linha diagonal em que x = y
- Se a diferença entre as medidas dos pares forem mero ruído aleatório, os pontos estarão dispostos de maneira simétrica ao longo da linha
- Distorções significativas dentro dos pares serão visíveis pelo deslocamento dos pontos em relação à diagonal

Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.



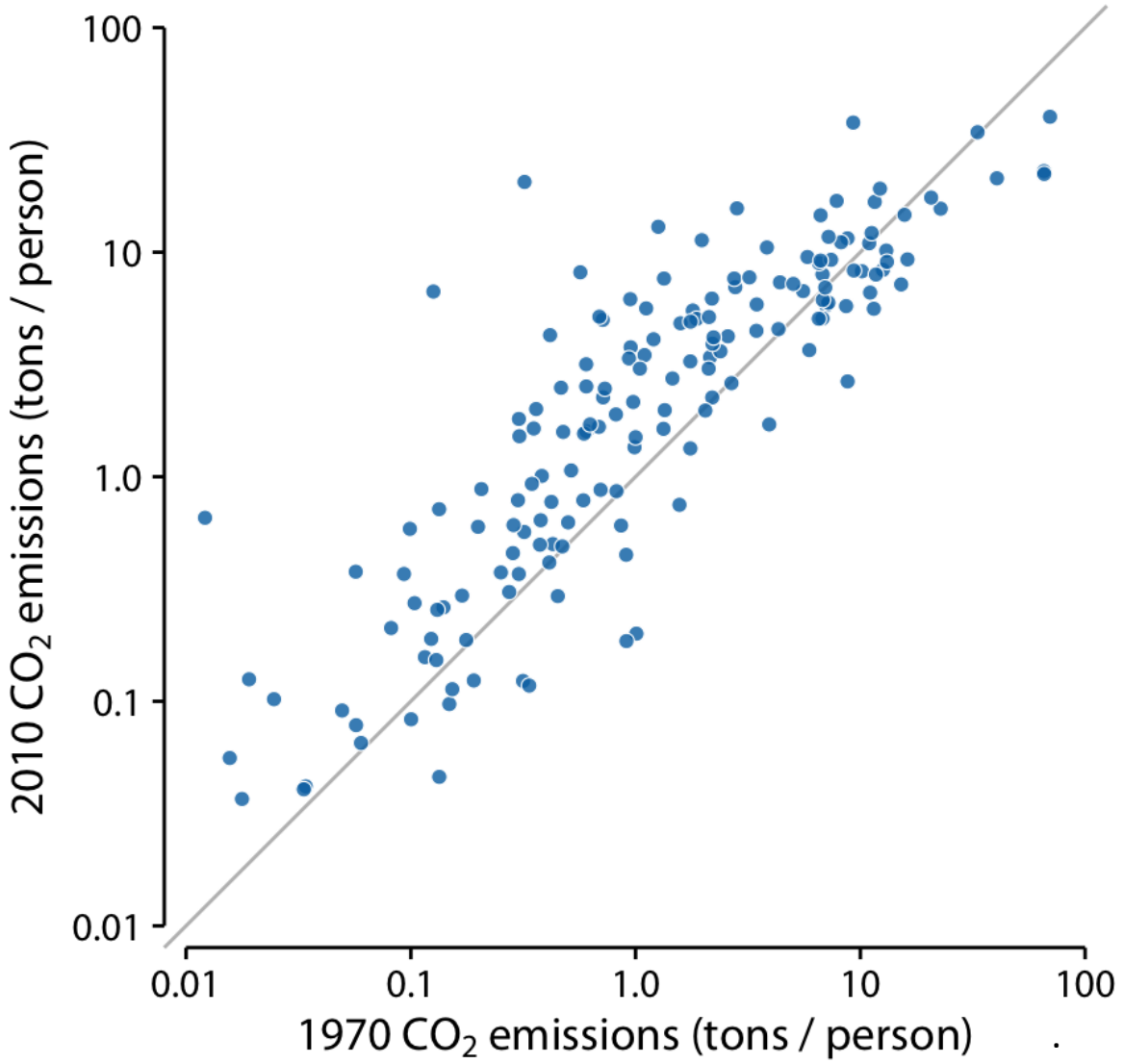
Dados pareados – Gráfico de dispersão

- Emissões de dióxido de carbono por pessoa, medidas em 166 países em 1970 e 2010

- A maior parte dos pontos está próxima à diagonal, sugerindo que para vários países o consumo foi similar após 40 anos

- Pontos deslocados para a região superior à linha diagonal: aumento na emissão de CO_2 durante o período avaliado

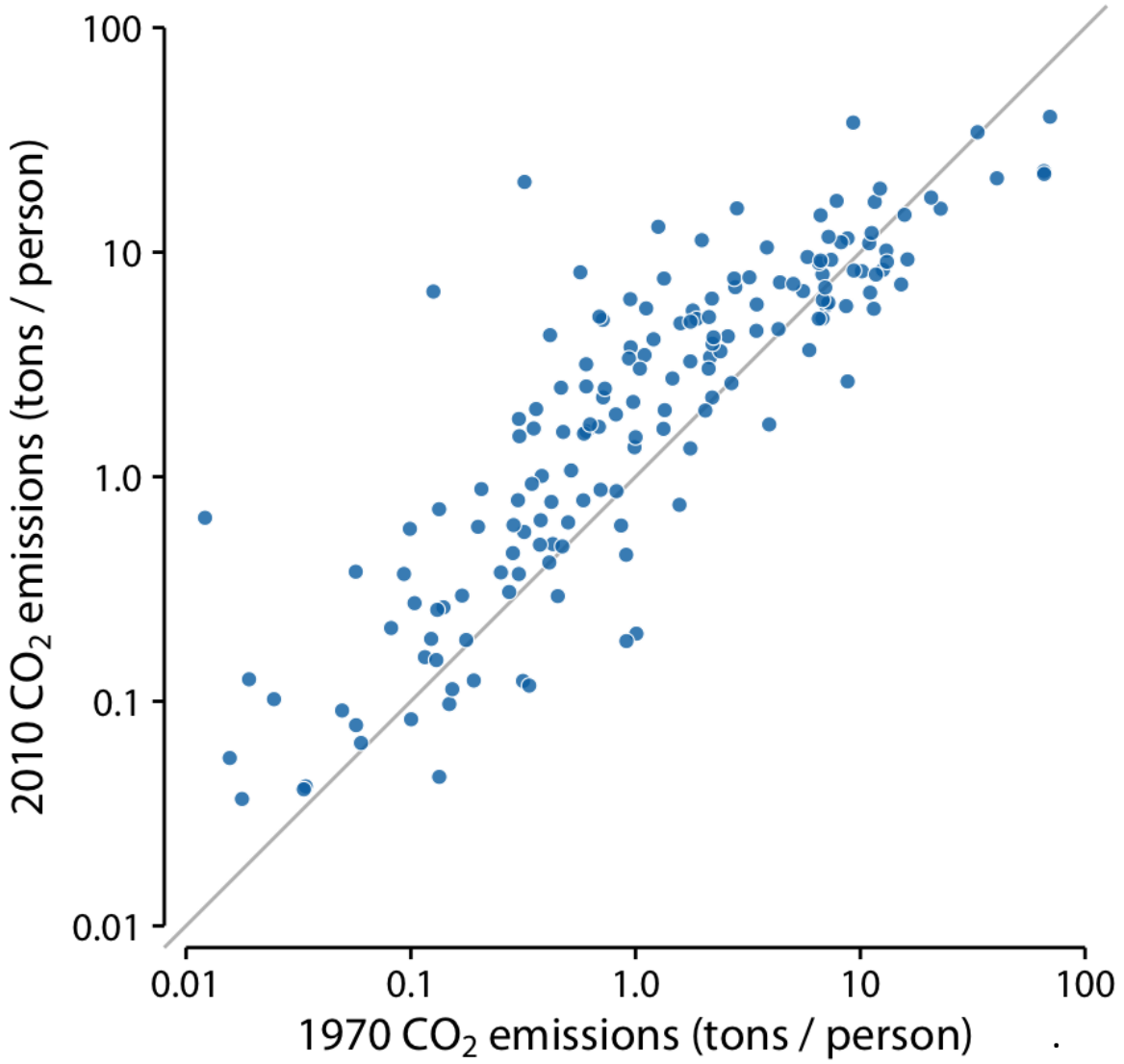
Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.



Dados pareados – Gráfico de dispersão

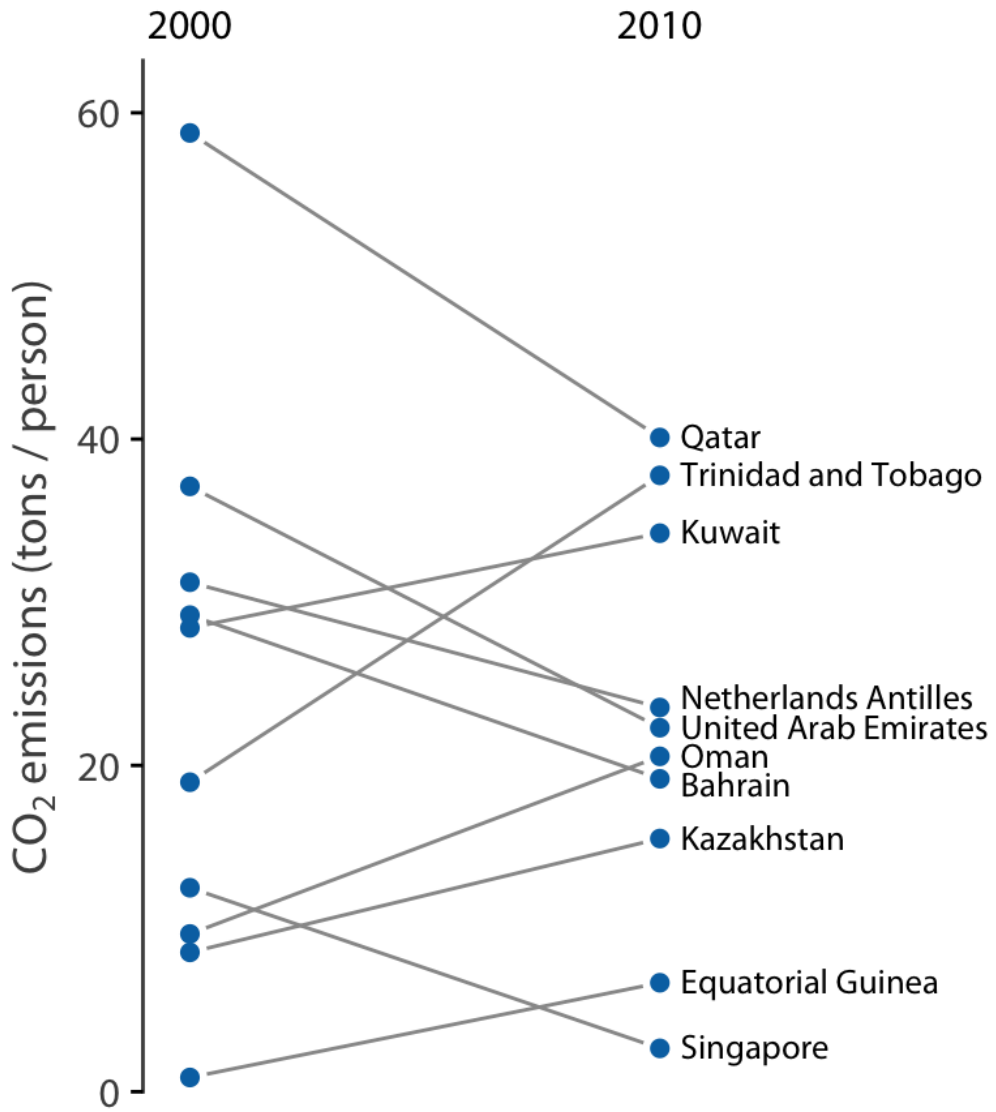
- Boa estratégia para conjuntos de dados com muitas observações
- Ressaltar a tendência e aspecto geral
- Ressaltar distorção em relação à hipótese nula (x = y) quando ocorre em boa parte do conjunto de dados

Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.



Dados pareados – Gráfico de inclinação

- Ideal para conjuntos de dados menores, com foco na identificação de cada um dos pares
- Medidas dos pares plotadas como pontos em duas colunas, pareados por linhas conectando os pontos
- Inclinação das linhas indica a magnitude e direção da mudança (quando ocorre)

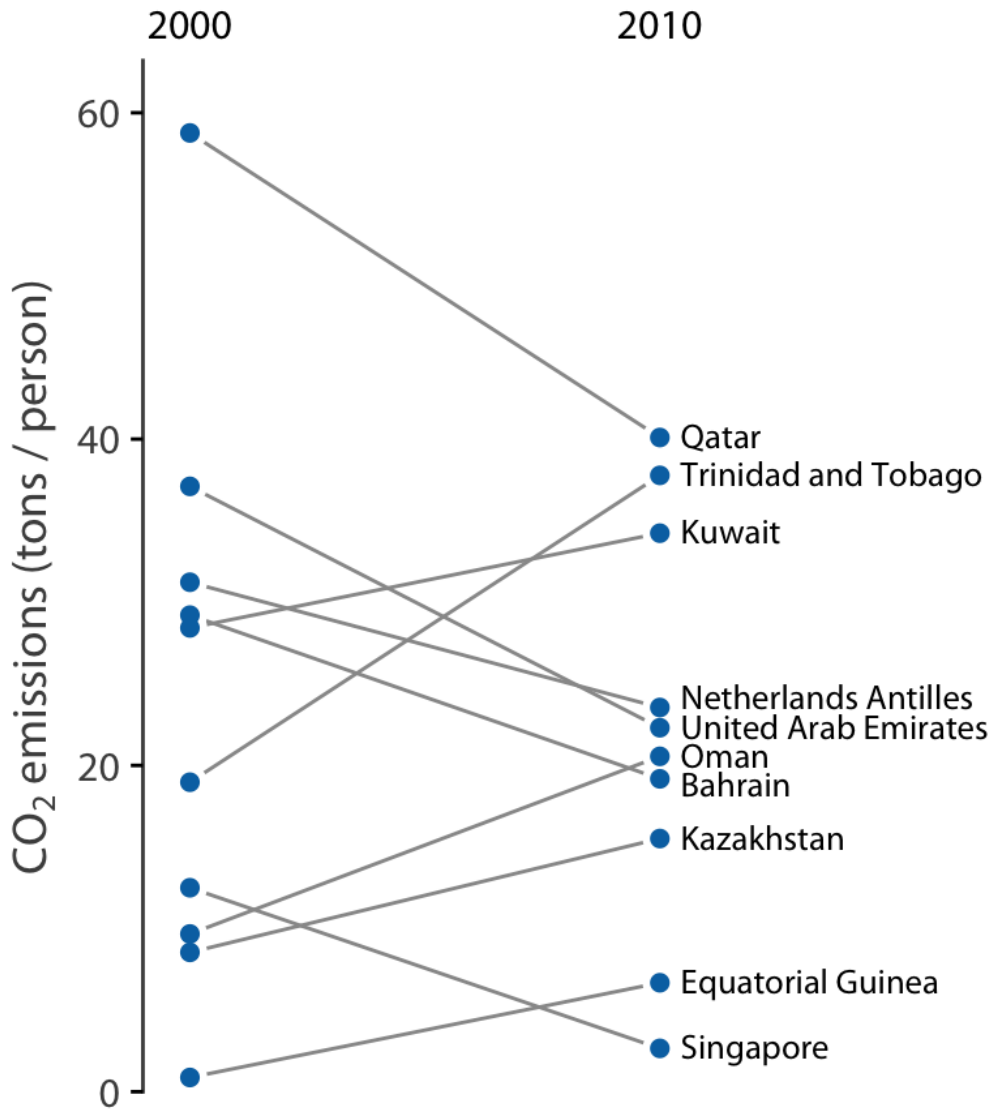


Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Dados pareados – Gráfico de inclinação

- Dez países com a maior diferença de emissão de dióxido de carbono entre 2000 e 2010

- Fácil percepção de quais países aumentaram ou diminuíram as emissões ao longo do período analisado

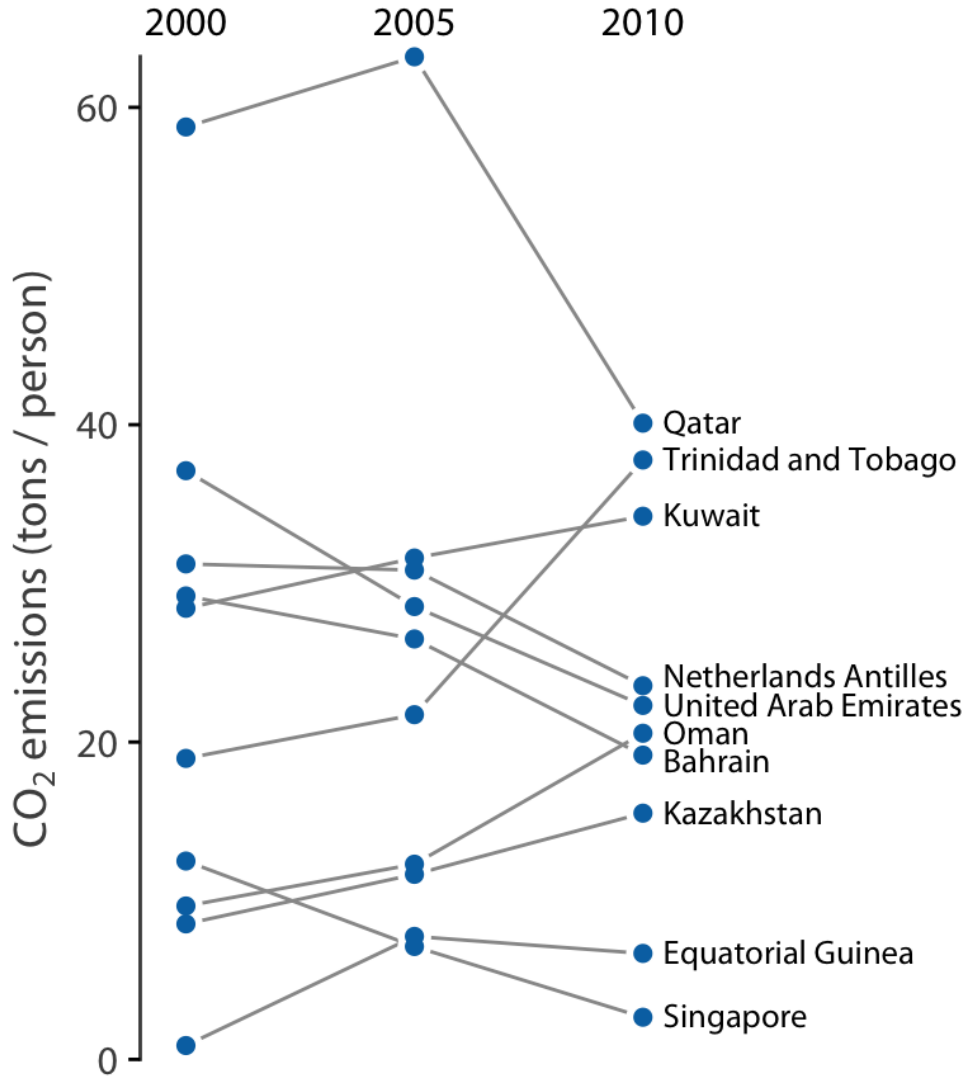


Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Dados pareados – Gráfico de inclinação

- Vantagem adicional: possibilidade de comparar mais de duas observações, não são restritos somente à pares como os gráficos de dispersão

- Adição de 2005 no gráfico, melhora na percepção do padrão de aumento ou diminuição da emissão
  - Catar aumentou as emissões antes de diminuir (gráfico anterior mostrava somente a diminuição)



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Visualizando séries temporais

Drª Desirrê Petters-Vandresen

01/12/2022

Finalidade

- Visualização de associação: visualização de como duas ou mais variáveis quantitativas estão relacionadas entre si

- Se uma das variáveis é um período de tempo, há adição de um novo nível de estrutura aos dados
  - As observações possuem uma ordem cronológica
  - Pontos podem ser ordenados ao longo do tempo, estabelecendo predecessores e sucessores para cada observação

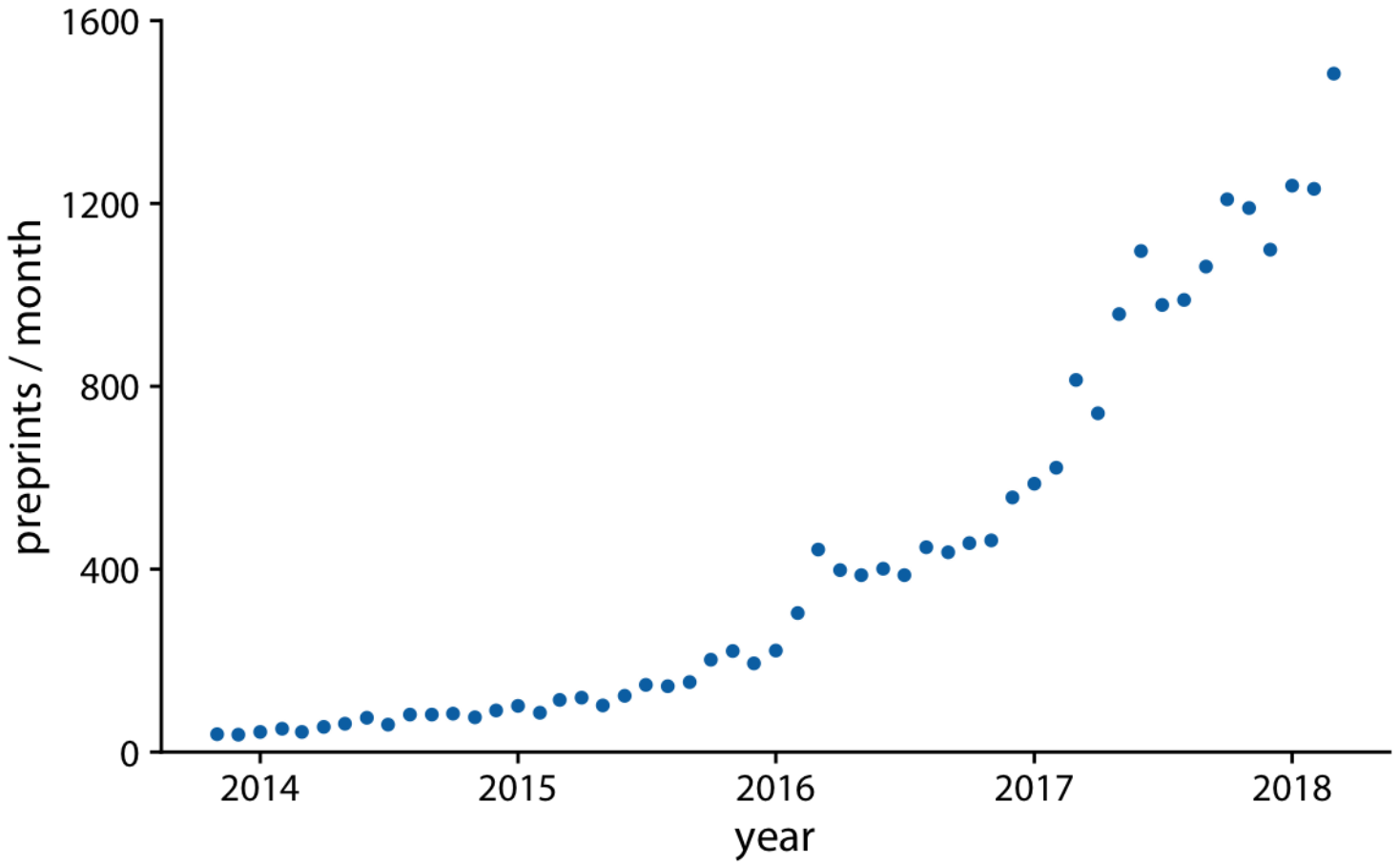
- Outros exemplos de nível de estrutura adicional aos dados:
  - Experimentos de dose-resposta (pontos ordenados de acordo com a concentração da dose)

Séries temporais individuais

- Gráfico de dispersão
  - Eixo x: marcos temporais
  - Eixo y: variável independente

- Crescimento na quantidade de preprints submetidos ao servidor bioRxiv desde sua fundação em 2013

- Pontos representando a quantidade de submissões por mês

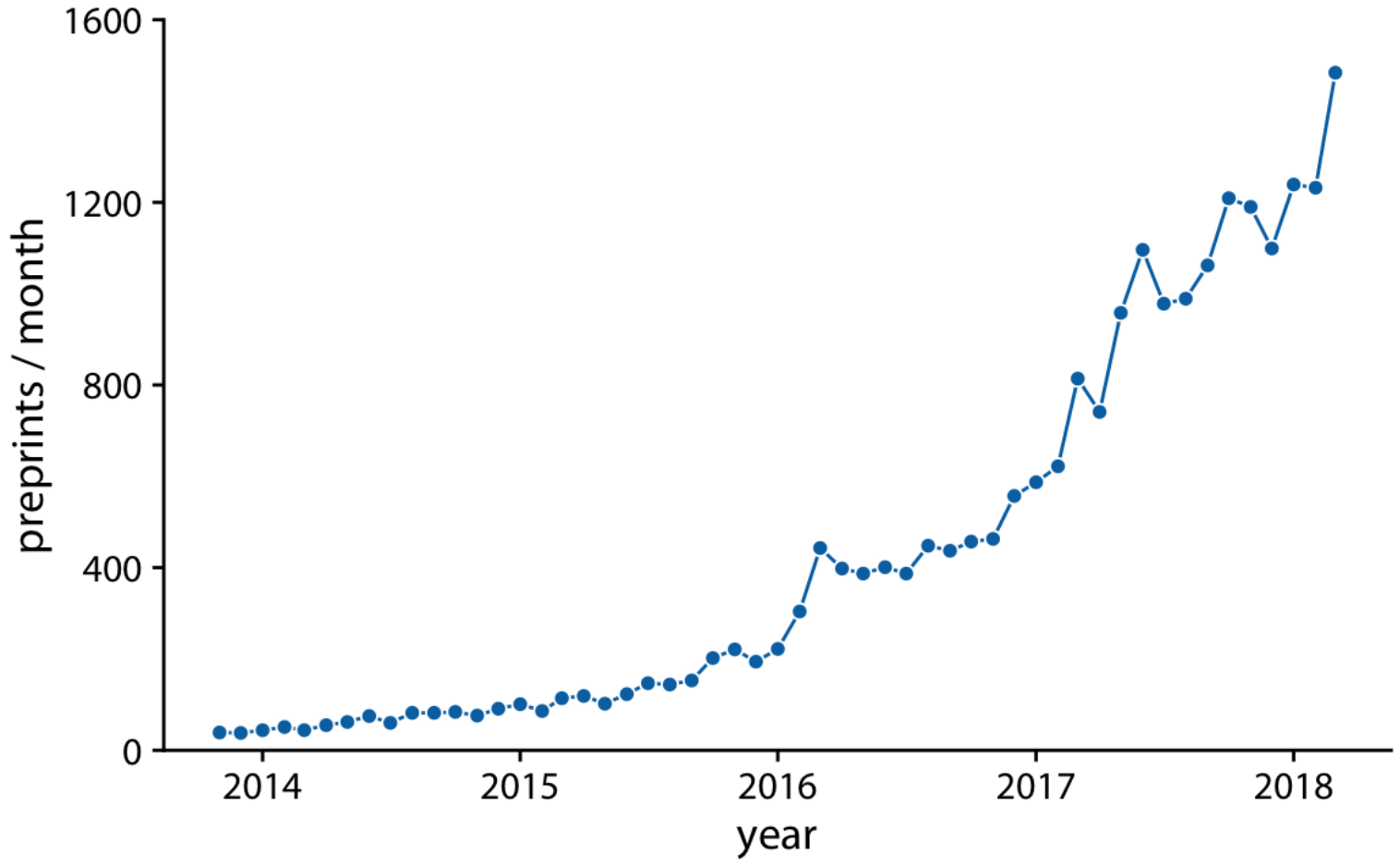


Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais individuais

- Diferença principal para um gráfico de dispersão convencional:
  - Pontos regularmente espaçados ao longo do eixo x
  - Ordem definida para os pontos (ordem cronológica)

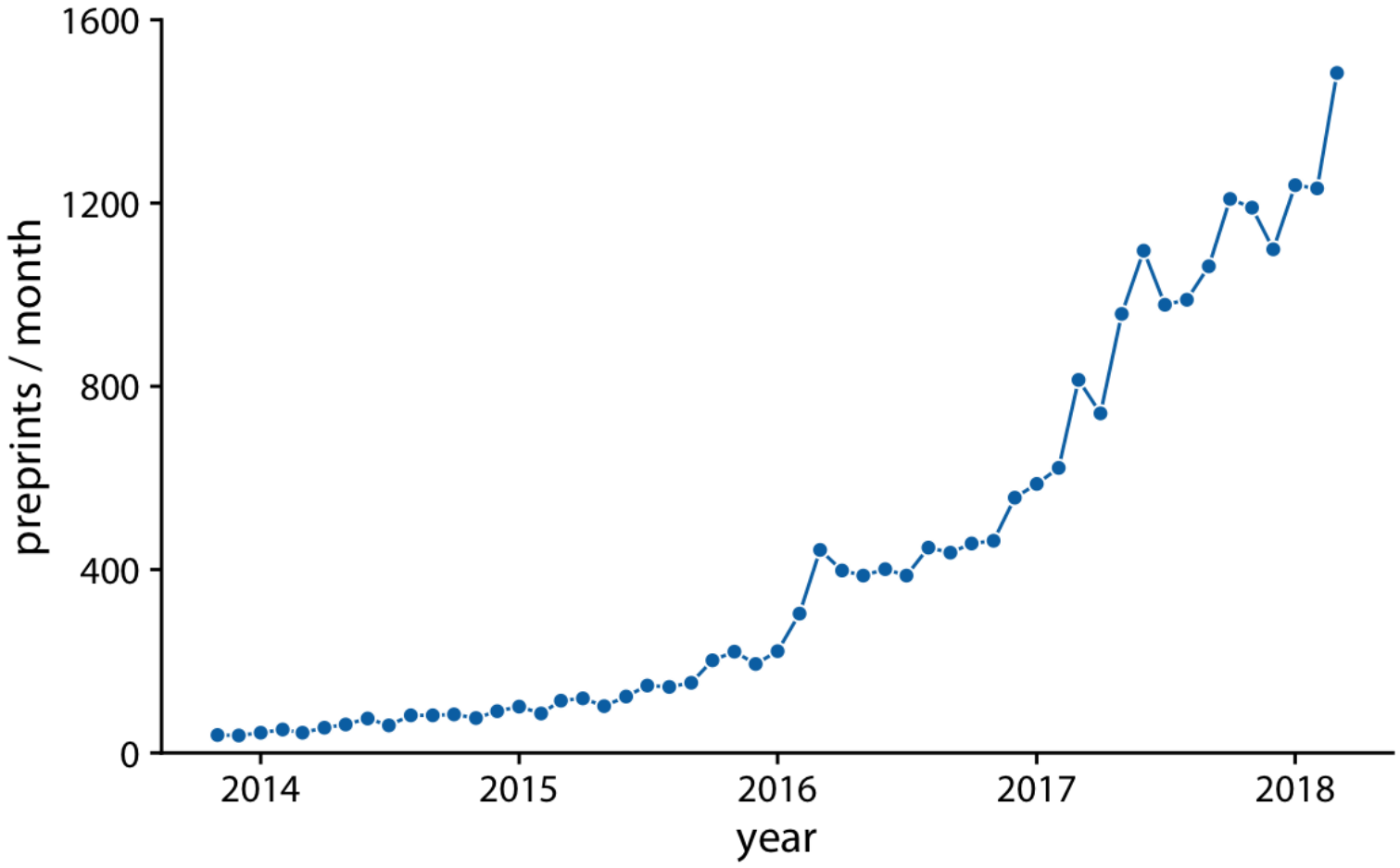
- Adição de uma linha conectando pontos adjacentes para ênfase na ordenação dos pontos: gráfico de linhas



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais individuais

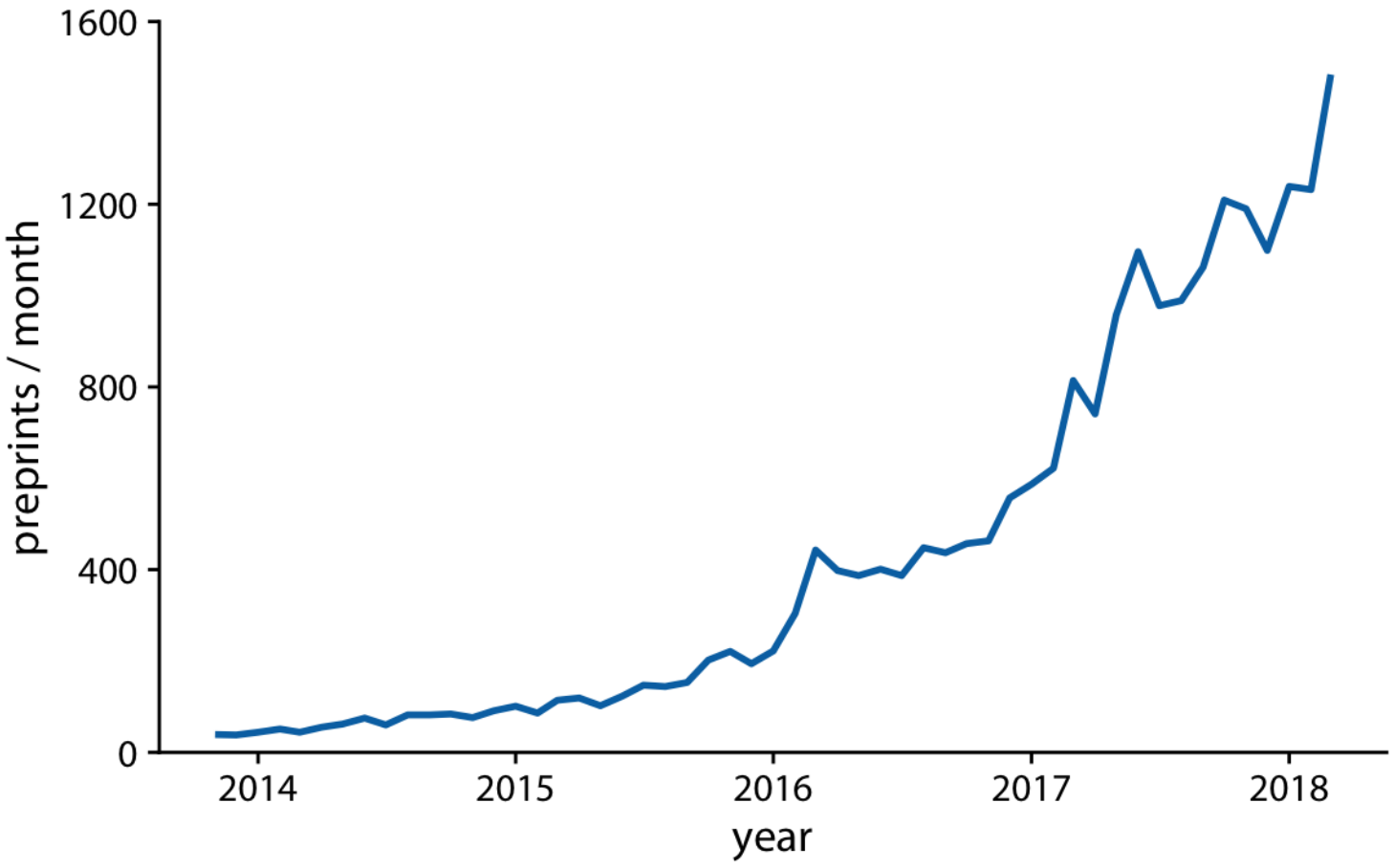
- Principais objeções ao uso de linhas:
  - Não representam dados reais, são somente um apoio visual
  - Se existissem observações para as posições ocupadas pelas linhas, pode ser que as observações não correspondessem às linhas



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais individuais

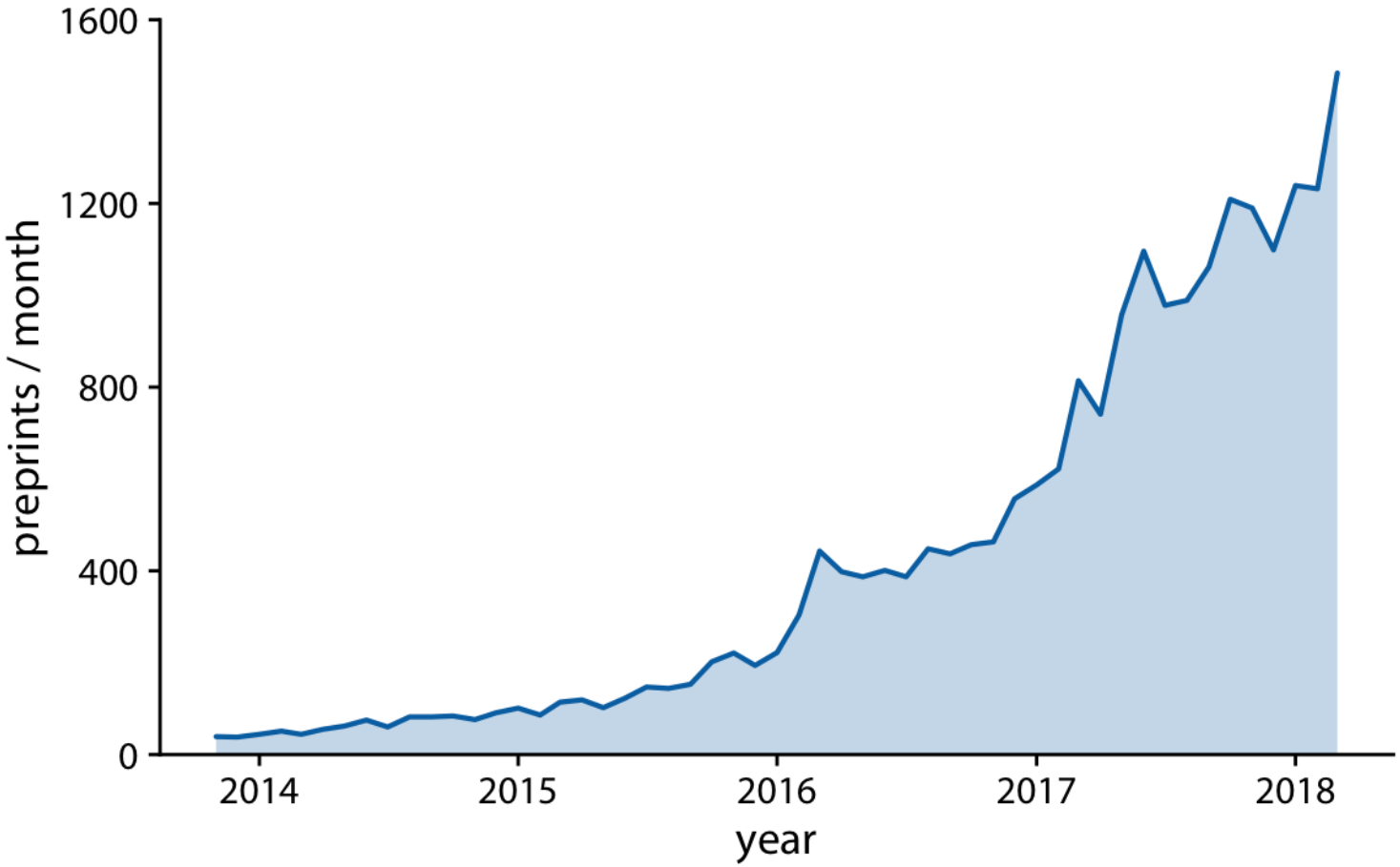
- Idealmente, mencionar na legenda da figura que as linhas não correspondem à dados, e são somente apoio visual
- Em muitos casos, os pontos são omitidos do gráfico
- Maior ênfase na tendência geral e não nas observações individuais
- Quanto mais densa a série temporal, mais benefícios em omitir os pontos



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais individuais

- Adição de cor na área sob a curva
- Maior ênfase na tendência do conjunto de dados, ao separar a área abaixo da curva da área acima da curva
- Somente utilizar esta estratégia se o eixo y começar em zero, para que a altura da área colorida seja proporcional à cada observação na série temporal

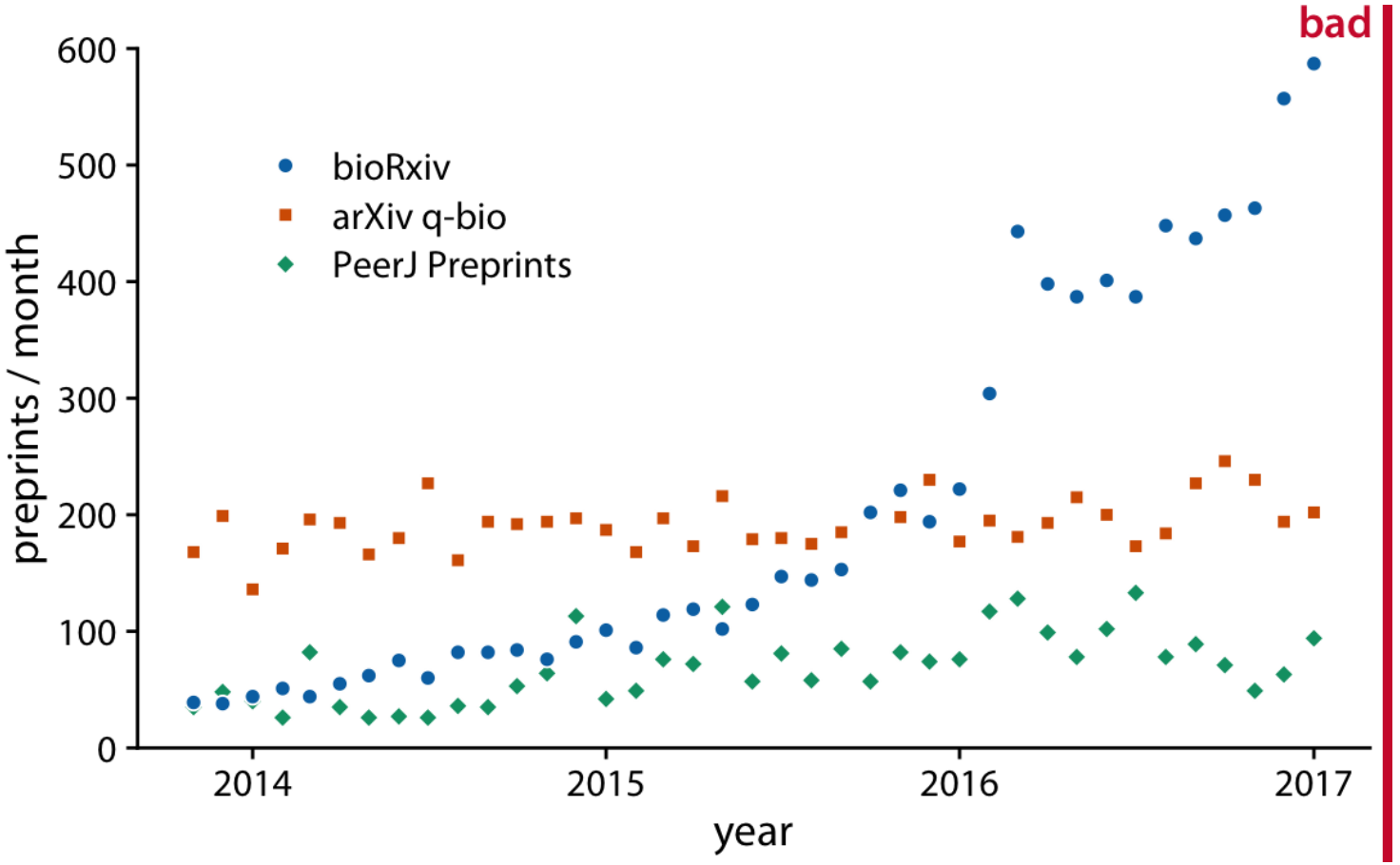


Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Múltiplas séries temporais e curvas dose-resposta

- Ao plotar múltiplos conjuntos, necessário maior cuidado para evitar dificuldade de interpretação

- Gráfico de dispersão não é uma boa estratégia, visto que os conjuntos de dados podem se sobrepor e misturar entre si

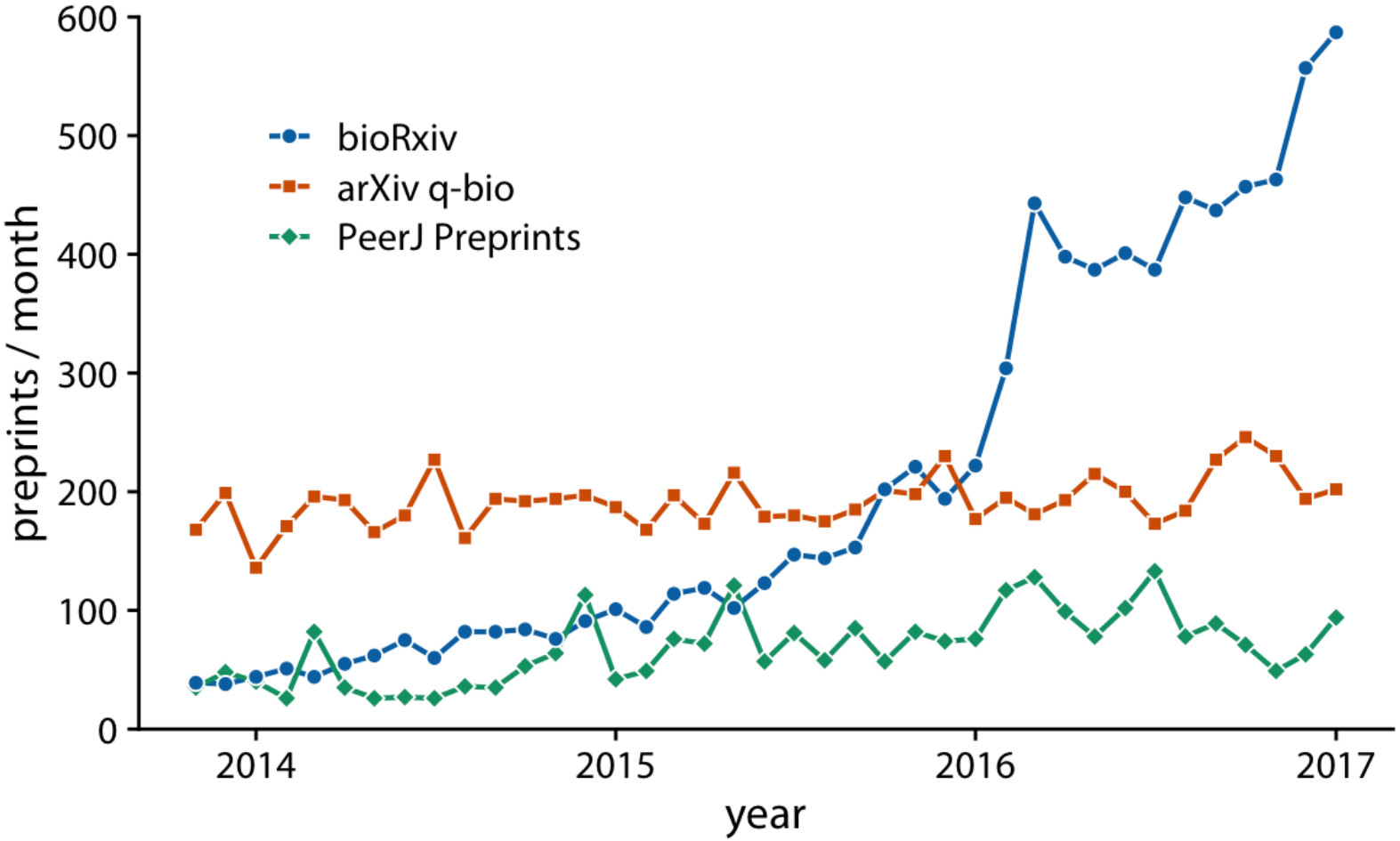


Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Múltiplas séries temporais e curvas dose-resposta

- O uso de diferentes cores e formas não é suficiente para facilitar a interpretação

- A adição de linhas conectando os pontos alivia um pouco a dificuldade de interpretação

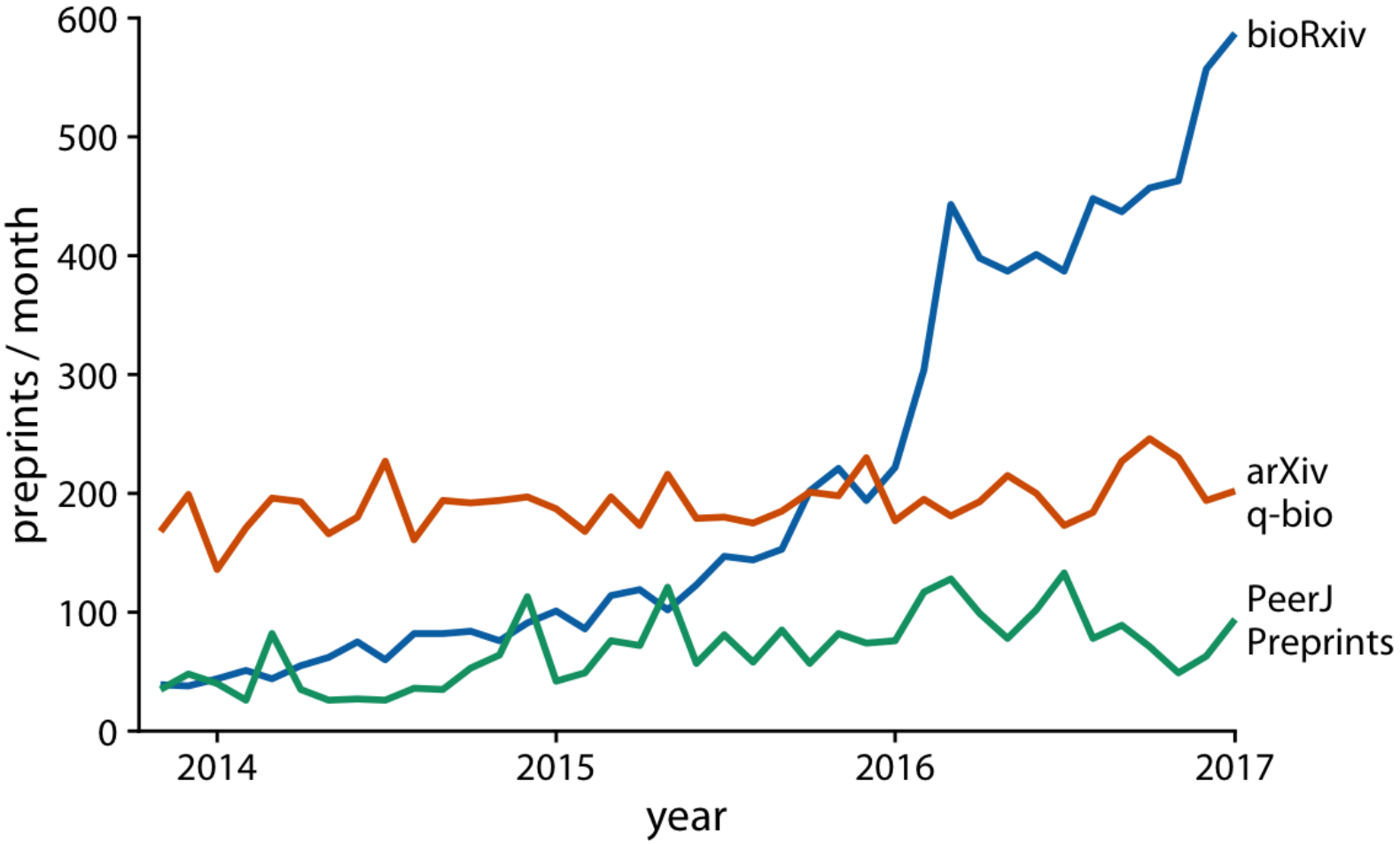


Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Múltiplas séries temporais e curvas dose-resposta

- Eliminação dos pontos individuais e legenda posicionada ao lado de cada série temporal facilita a leitura e interpretação

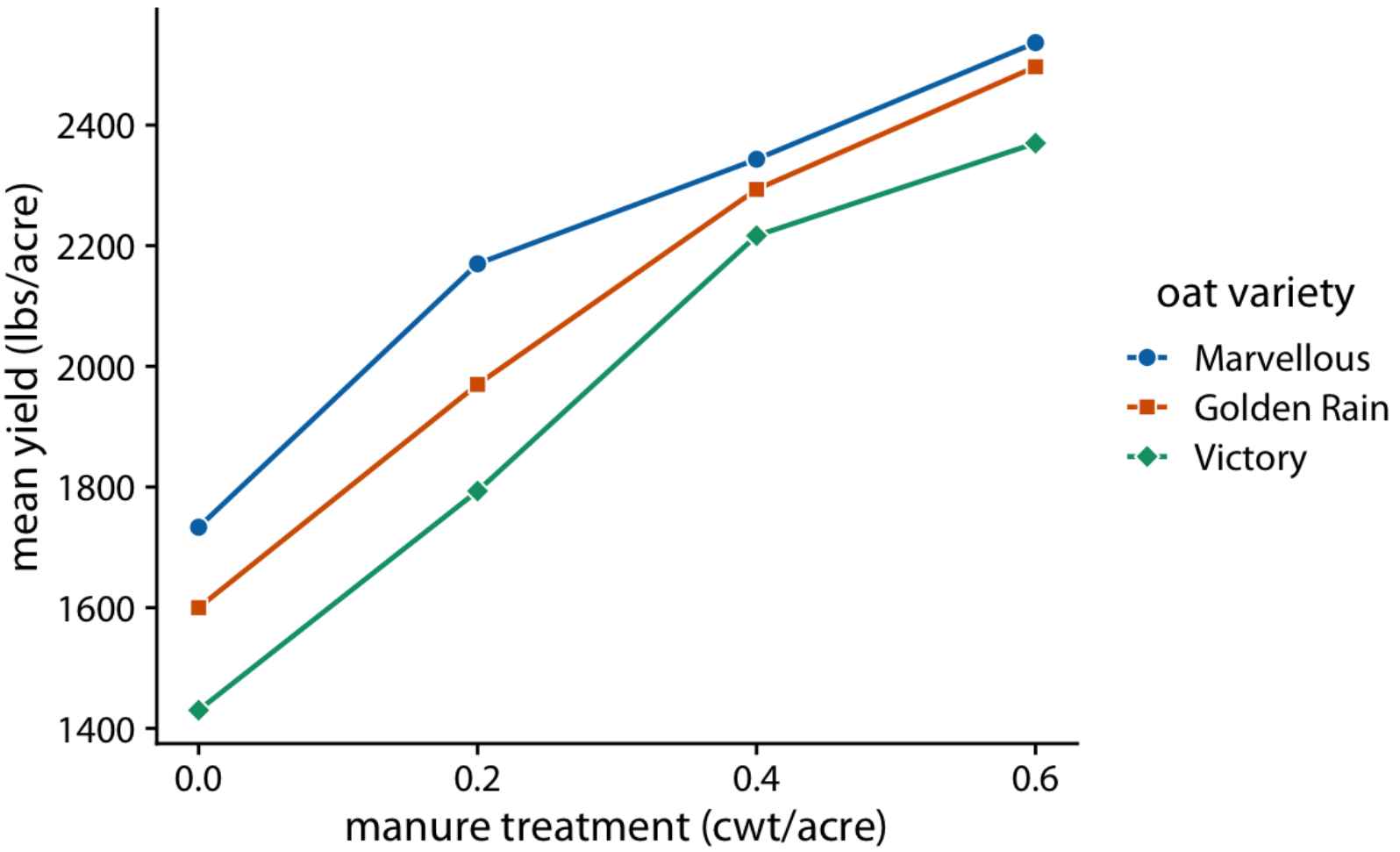
- Imagem menos poluída e esteticamente mais agradável



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Múltiplas séries temporais e curvas dose-resposta

- Estratégia funcional para visualização de curvas dose-resposta
- Ordenação de acordo com aumento de concentração do tratamento
- Ênfase na resposta geral: diferentes variedades de aveia respondem de forma similar ao aumento de fertilizante



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais de duas ou mais variáveis resposta

- Em muitos casos estamos interessados na variação de mais de uma variável ao longo do tempo

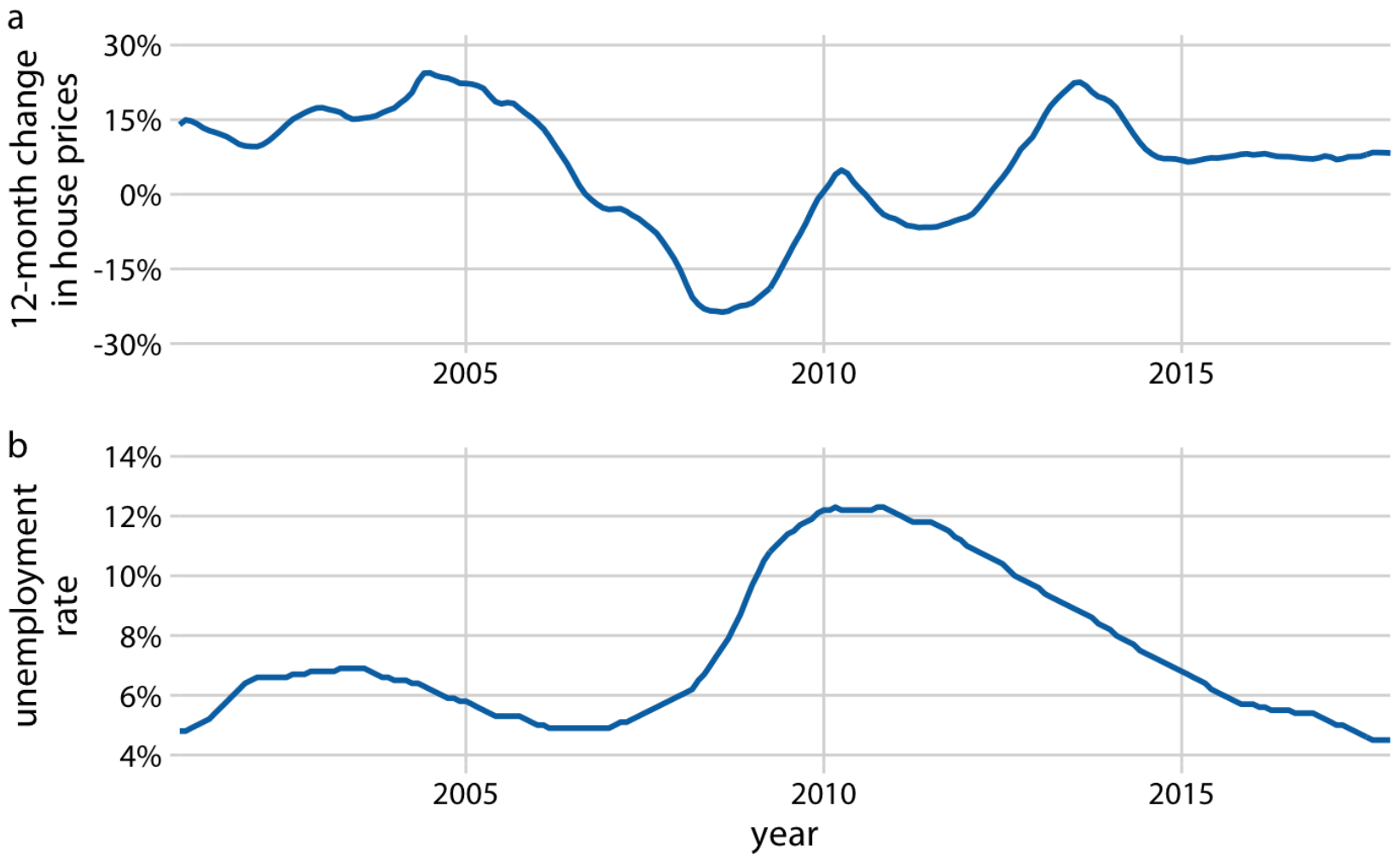
- Exemplo:
  - Variação de preço de residências e taxa de desemprego ao longo de doze meses

- Hipótese:
  - Preços sobem quando a taxa de desemprego é alta (e vice-versa)

Séries temporais de duas ou mais variáveis resposta

- Primeira possibilidade: visualização com dois gráficos de linha separados, um sobre o outro

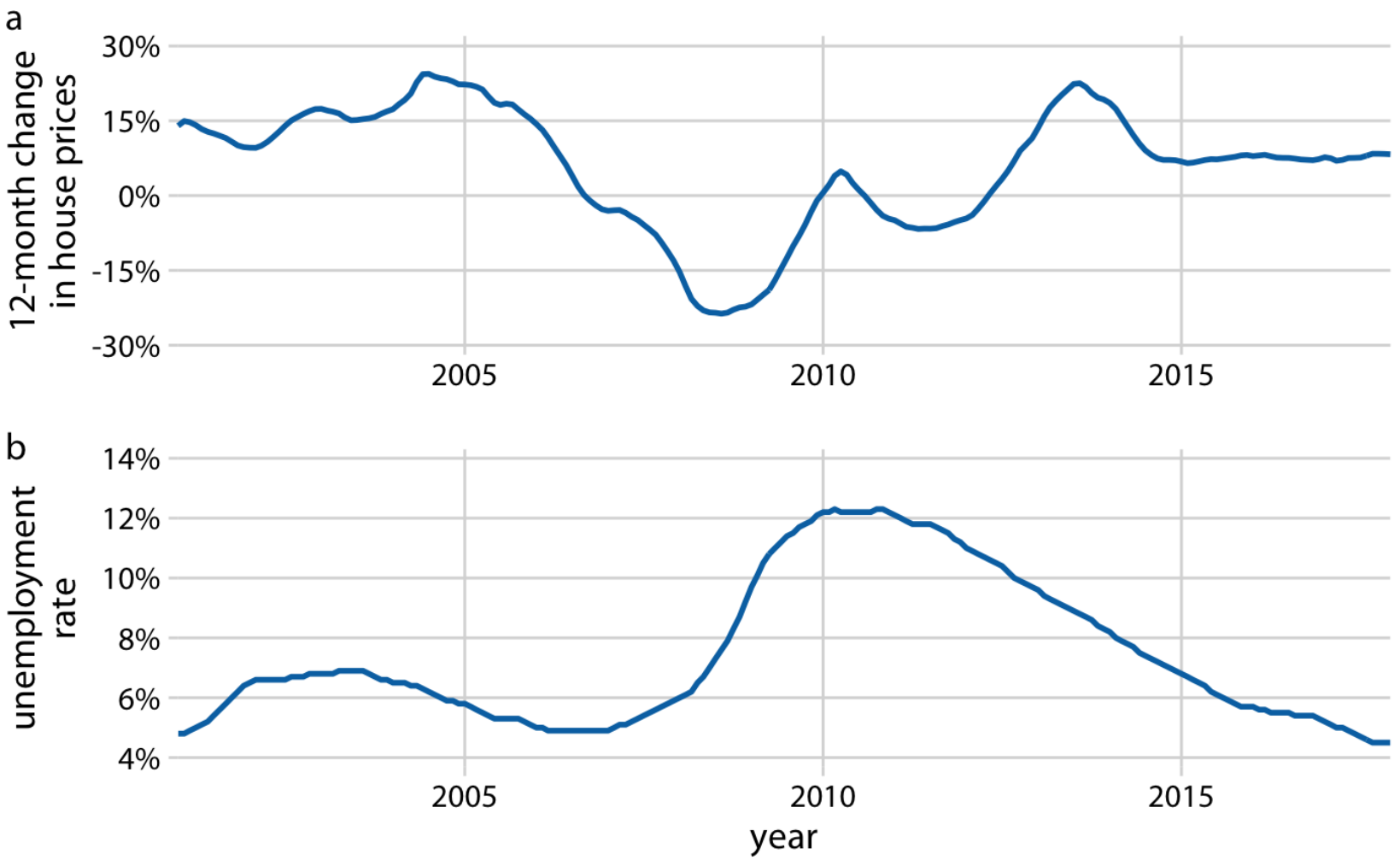
- Fácil interpretação, mas comparações podem ser trabalhosas



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais de duas ou mais variáveis resposta

- Por exemplo, para identificar momentos em que ambas as variáveis se moveram na mesma direção ou direções opostas:
  - Necessário comparar as inclinações nos dois gráficos separadamente



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais de duas ou mais variáveis resposta

- Alternativa: connected scatter plot (gráfico de dispersão conectado)

- Plotar as variáveis uma contra a outra, gerando um caminho visual do primeiro ponto da ordem cronológica até o último

Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais de duas ou mais variáveis resposta

- Linhas indo da região inferior esquerda para a região superior direita indicam crescimento das duas variáveis

- Linhas na direção oposta indicam crescimento anticorrelacionado (aumento de uma variável e diminuição da outra)

Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais de duas ou mais variáveis resposta

- Se a relação entre as variáveis for cíclica, visualizam-se círculos e espirais no gráfico

- Relação cíclica entre 2001 e 2005, e relação cíclica entre 2005 e 2017

Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais de duas ou mais variáveis resposta

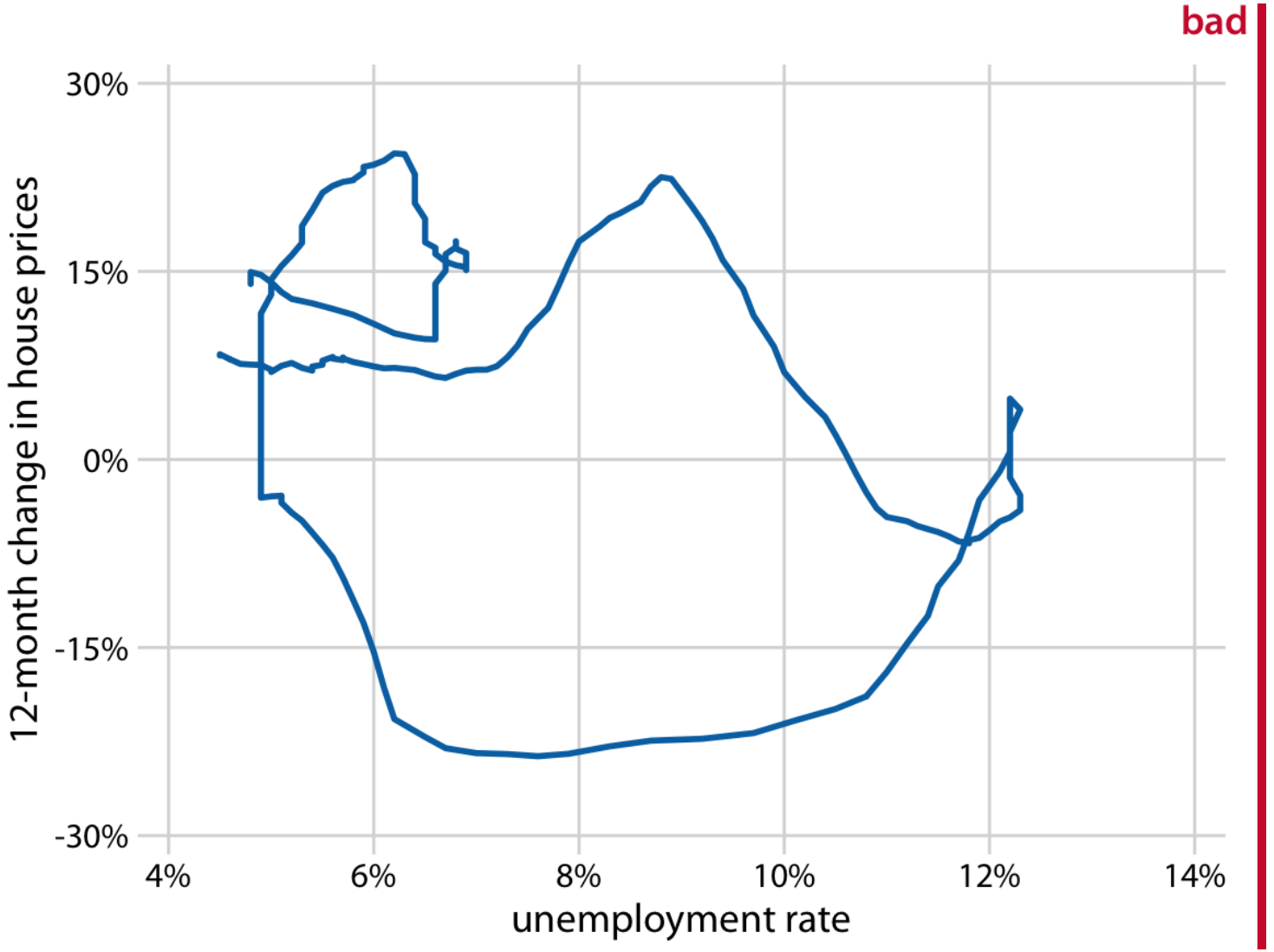
- Alguns leitores podem ter dificuldades de interpretação e inverter o sentido e a ordem em gráficos de dispersão conectados

- Dificuldade no reconhecimento e interpretação de correlações

Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais de duas ou mais variáveis resposta

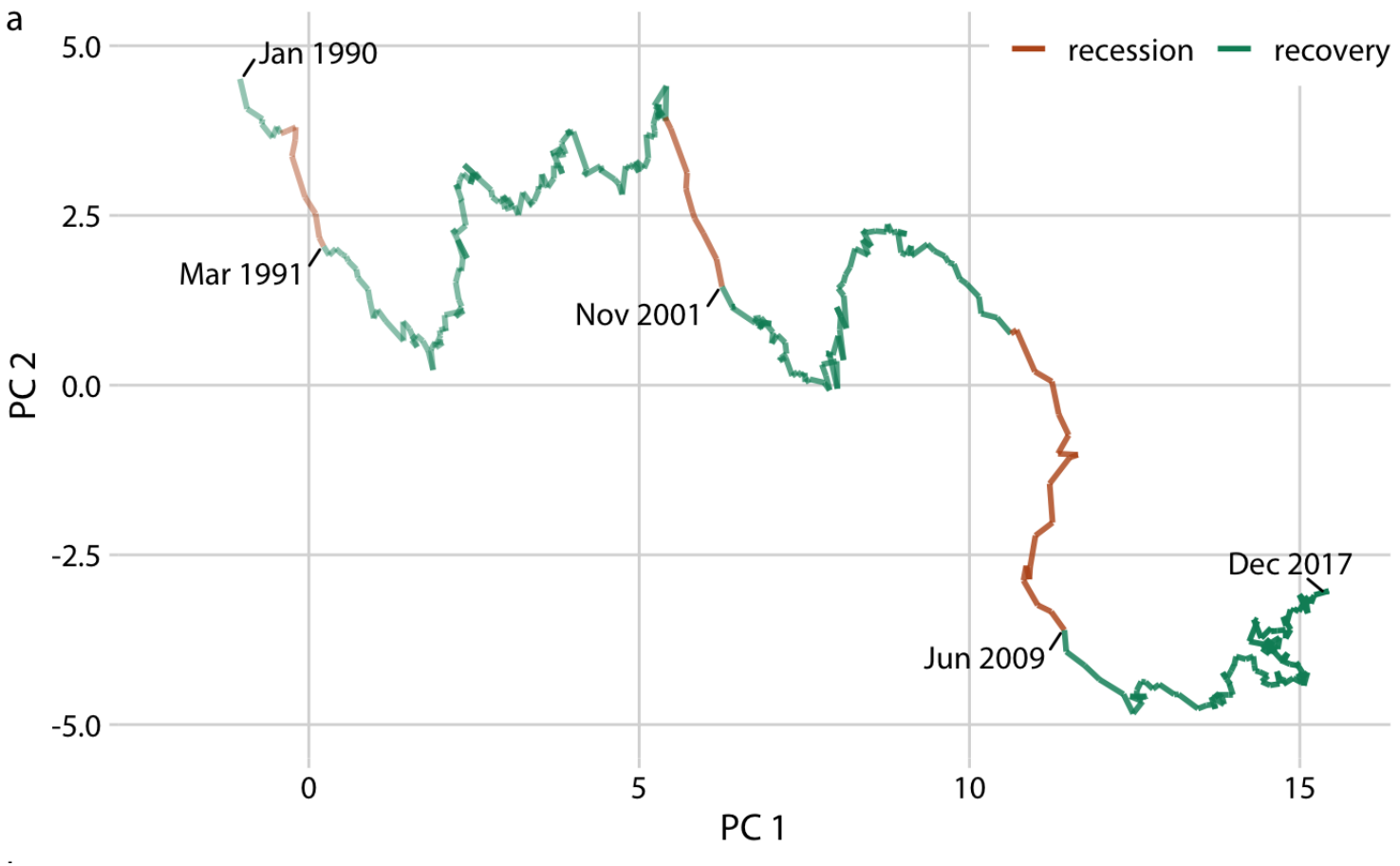
- Essencial indicar a direção e escala temporal dos dados
- Sem estas informações, o gráfico não passa de um rabisco
- Possibilidade de utilizar gradiente de cor ou setas



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais de duas ou mais variáveis resposta

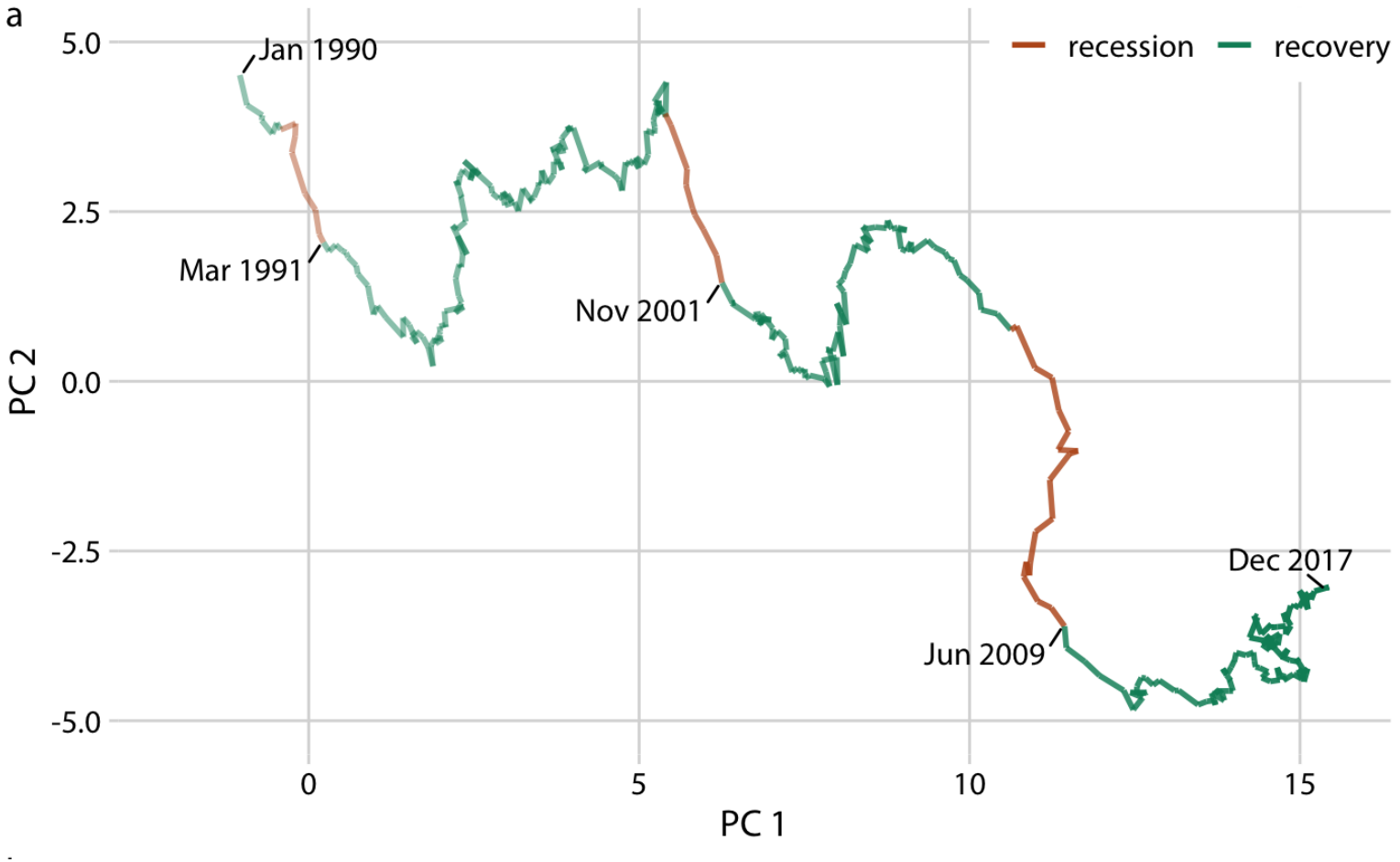
- Possibilidade de apresentar conjuntos multidimensionais após aplicar redução de dimensionalidade
- 100 indicadores macroeconômicos combinados
- Aspecto similar à um gráfico de linhas, com o tempo variando da esquerda para direita
- Aspecto comum em PCAs: PC1 normalidade mede o tamanho geral do sistema (nesse caso, tamanho da economia)



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais de duas ou mais variáveis resposta

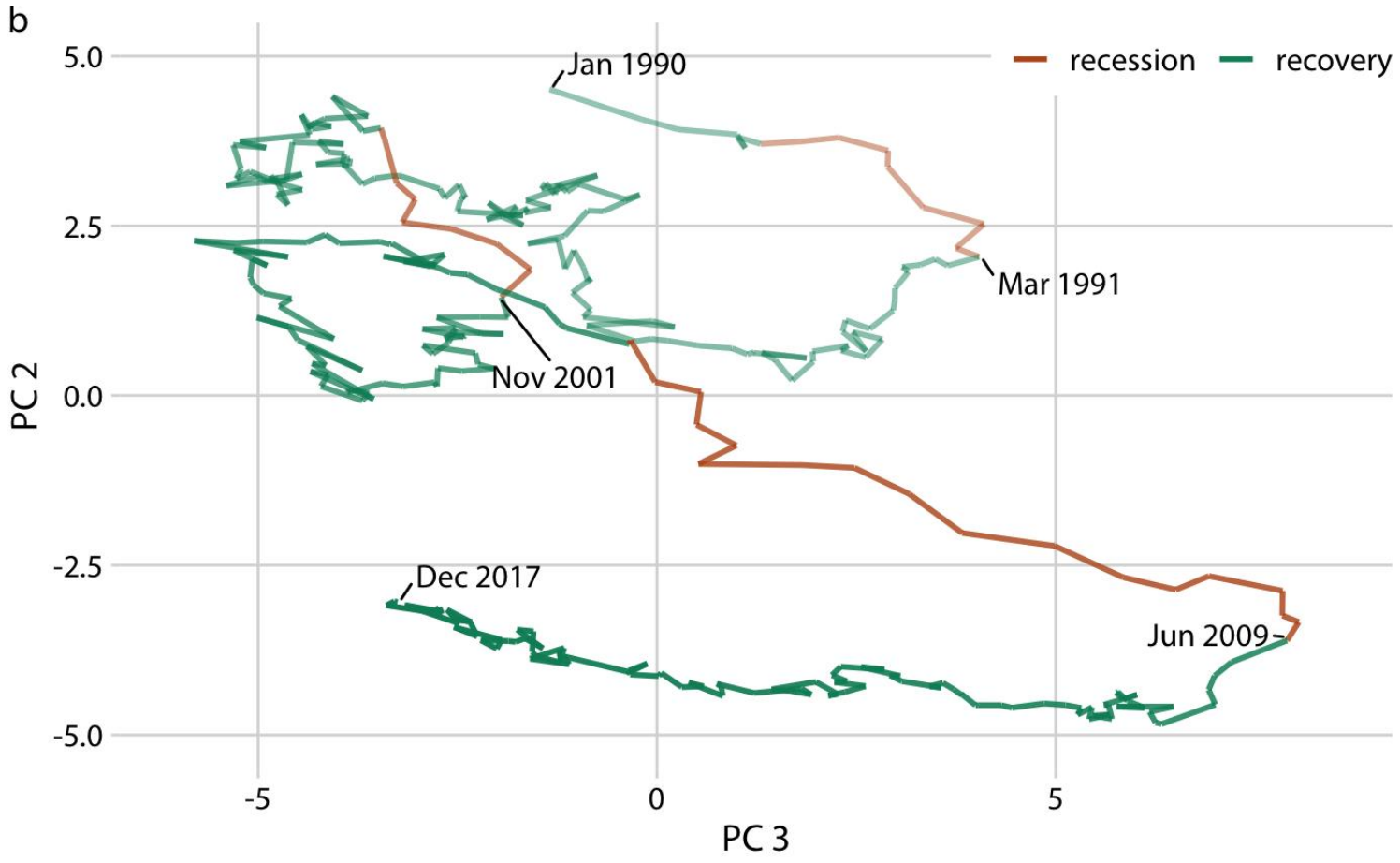
- Codificação por cor:
  - Recessões associadas com diminuição em PC2
  - Períodos de recuperação sem associação clara com PC1 ou PC3



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.

Séries temporais de duas ou mais variáveis resposta

- Codificação por cor:
  - Ao plotar PC2 vs. PC3, aspecto de espiral no sentido horário
  - Ênfase na natureza cíclica da economia: recessões seguidas de períodos de recuperação, oscilações ao longo de PC3



Adaptado de:
WILKE, C. O. **Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures.** O'Reilly Media, 2019.