



SME0803 Visualização e Exploração de Dados

Associação entre variáveis

Prof. Cibele Russo

cibele@icmc.usp.br

Baseado em

Murteira, B. J. F., Análise Exploratória de Dados. McGraw-Hill, Lisboa, 1993.

Notas de aula de Análise Exploratória de dados de Mário de Castro, ICMC-USP, 2010.

Associação entre variáveis

Objetivo: descrever simultaneamente a variabilidade de **duas (ou mais) variáveis**, de forma que cada conjunto de variáveis é observada para uma mesma unidade observacional (por exemplo pessoa, animal, planta, etc).

Associação entre variáveis

Sejam $(x_i, y_i), i = 1, \dots, n$ pares de observações de duas variáveis x e y , que podem ser qualitativas ou quantitativas.

Associação entre variáveis

Sejam $(x_i, y_i), i = 1, \dots, n$ pares de observações de duas variáveis x e y , que podem ser qualitativas ou quantitativas.

A técnica mais adequada para avaliar a associação entre x e y por meio dos valores observados (x_i, y_i) depende do tipo de variáveis envolvidas.

Associação entre variáveis

Sejam $(x_i, y_i), i = 1, \dots, n$ pares de observações de duas variáveis x e y , que podem ser qualitativas ou quantitativas.

A técnica mais adequada para avaliar a associação entre x e y por meio dos valores observados (x_i, y_i) depende do tipo de variáveis envolvidas.

Para estudar a **associação** entre x e y precisamos de uma **medida de associação**, que deve avaliar se associação é forte ou fraca, positiva ou negativa. Outra possibilidade é a **representação gráfica** da associação entre x e y .

Associação entre variáveis

Sentido da associação: Positiva ou negativa

Intensidade da associação: Forte, moderada ou fraca.

Tipos de relação entre variáveis

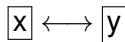
(a) **Causal unilateral:** y depende de x **ou** x depende de y .

$$\boxed{x} \rightarrow \boxed{y}$$

Exemplo: Preço da venda de um produto (y) depende do local de venda (x).

Tipos de relação entre variáveis

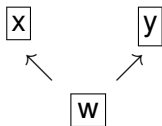
(b) **Causal bilateral:** y depende de x e x depende de y .



Exemplo: Peso (x) e a circunferência abdominal (y) de uma pessoa.

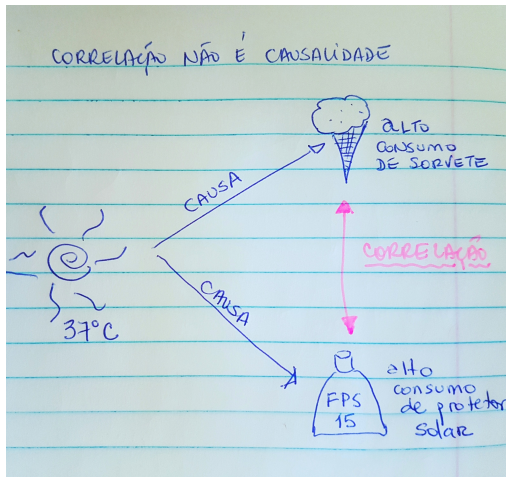
Tipos de relação entre variáveis

(c) Dependência indireta:



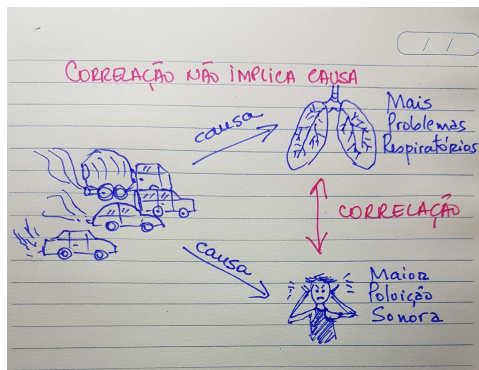
Exemplo: Venda de sorvete na praia (x), casos de afogamento (y) e temperatura (w).

Correlação x Causalidade



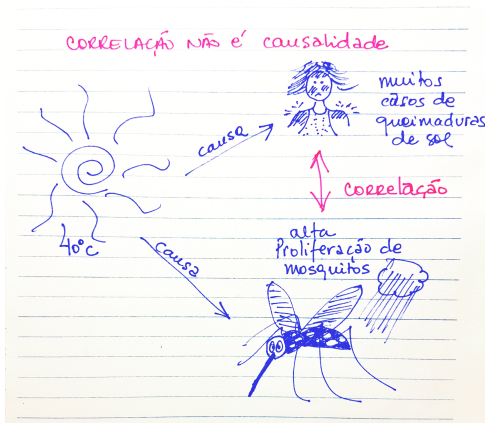
Por Doris Fontes, presidente do CONRE-3

Correlação x Causalidade



Por Doris Fontes, presidente do CONRE-3

Correlação x Causalidade



Por Doris Fontes, presidente do CONRE-3

Correlação x Causalidade

Correlação não implica em causalidade!

Exemplos de correlações espúrias:

<https://www.tylervigen.com/spurious-correlations>

Associação entre duas variáveis

- 1 Variável qualitativa x variável qualitativa: tabelas de contingência, teste qui-quadrado, gráfico de mosaico.

Associação entre duas variáveis

- 1 Variável qualitativa x variável qualitativa: tabelas de contingência, teste qui-quadrado, gráfico de mosaico.
- 2 Variável quantitativa x variável quantitativa: gráfico de dispersão, covariância, correlação.

Associação entre duas variáveis

- 1 Variável qualitativa x variável qualitativa: tabelas de contingência, teste qui-quadrado, gráfico de mosaico.
- 2 Variável quantitativa x variável quantitativa: gráfico de dispersão, covariância, correlação.
- 3 Variável quantitativa x variável qualitativa: boxplots, teste para comparação de médias.

Exemplo: variável qualitativa x variável qualitativa

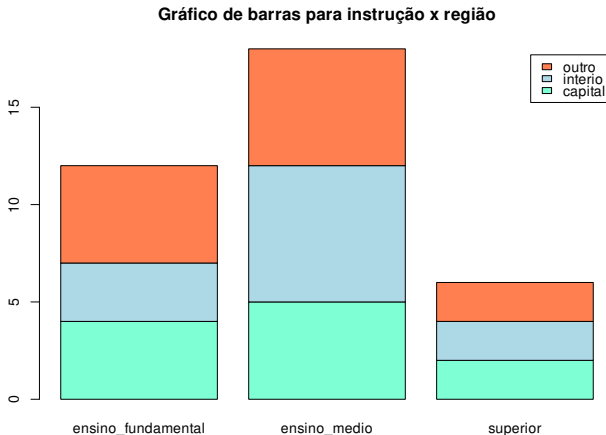


Figura: Gráfico de barras empilhadas de grau de instrução e região para os dados da Companhia MB. Fonte: Elaborado pela autora.

Exemplo: variável qualitativa x variável qualitativa

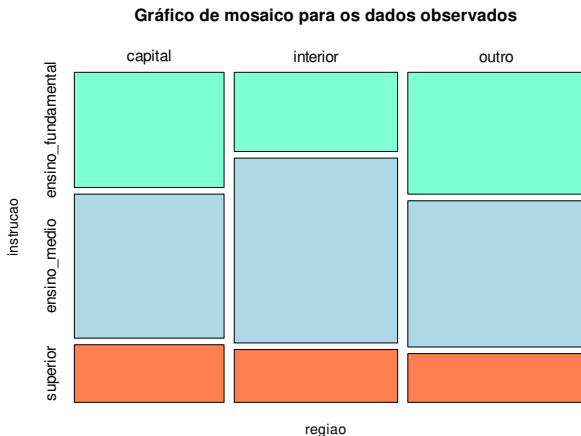


Figura: Gráfico de mosaico para grau de instrução e região para os dados da Companhia MB. Fonte: Elaborado pela autora.

Exemplo: variável quantitativa x variável quantitativa

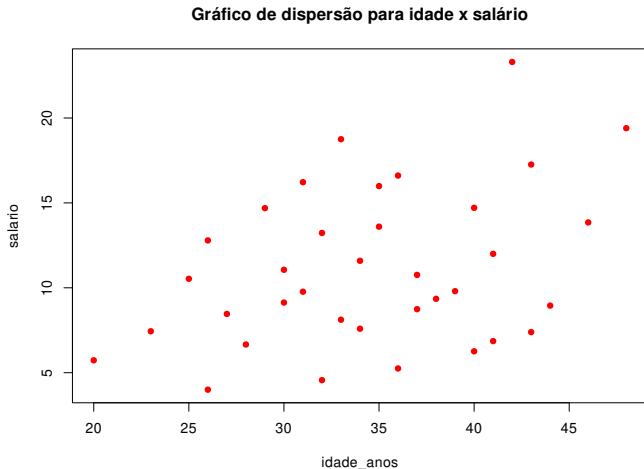


Figura: Dispersão de salário e idade para os dados da Companhia MB.

Fonte: Elaborado pela autora.

Exemplo: variável quantitativa x variável quantitativa

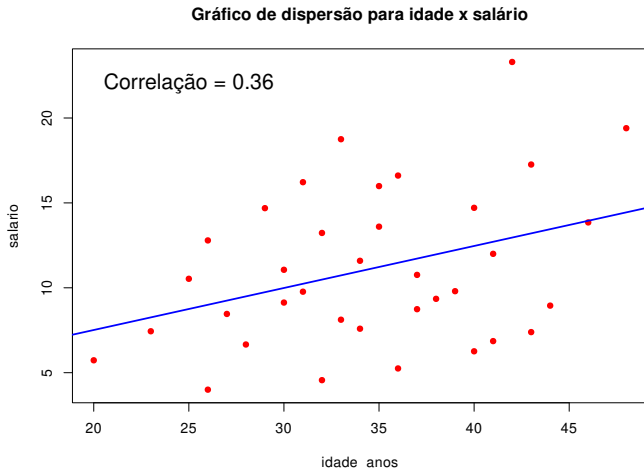


Figura: Dispersão e correlação entre salário e idade para os dados da Companhia MB. Fonte: Elaborado pela autora.

Exemplo: variável quantitativa x variável qualitativa

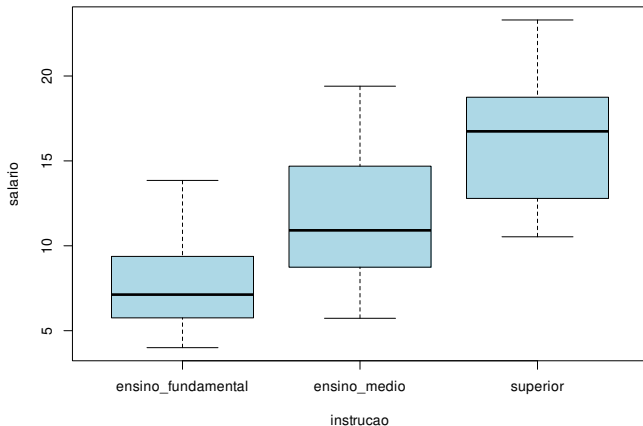


Figura: Boxplots para salário x grau de instrução para os dados da Companhia MB. Fonte: Elaborado pela autora.