



SME0803 Visualização e Exploração de Dados

Representação tabular e gráfica de dados qualitativos

Prof. Cibeles Russo

cibele@icmc.usp.br

Baseado em

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

Bussab, W.O.; Morettin, P.A. Estatística Básica. 6ª ed., São Paulo: Saraiva, 2009.

Representação tabular

Alguns elementos textuais para representações ilustrativas de informações, regras ou resultados de análises:

- Quadros
- Tabelas
- Figuras

Quadro

Quadros são usados para **informações qualitativas**, predominantemente preenchidos com palavras.

Algumas regras de apresentação:

- No topo do quadro apresentar a numeração e título;
- Na parte inferior citar a fonte (mesmo que seja o próprio autor do trabalho);
- Centralizar o quadro, título, legenda e fonte;

Quadro

Numeração e título do quadro

Quadro 12.1: Resumo do teste $H_0: \mu = 155$, $H_1: \mu = 145$, com $RC =]-\infty, 150]$.

Origem Real dos Parafusos	Decisão	
	A	B
A	conteúdo Sem erro	Erro tipo II $\beta = 1,88\%$
B	Erro tipo I $\alpha = 10,56\%$	Sem erro

Fonte: Bussab, W.O.; Morettin, P.A. "Estatística Básica". 6ª ed., São Paulo: Saraiva, 2009.

Fonte

fechado nas laterais

fechado nas laterais

Tabela

Tabelas são usadas para a apresentação de dados com destaque para os elementos numéricos.

- No topo da tabela inserir a numeração e título;
- Na parte inferior citar a fonte (mesmo que seja o próprio autor do trabalho);
- Em geral aberta nas laterais;
- Pode apresentar informações sumarizadas na última linha e/ou na última coluna.

Tabela

Numeração e título da tabela

Tabela 2.6: Distribuição de freqüências da variável S, salário dos empregados da seção de orçamentos da Companhia MB.

	Classes de salários	Ponto médio s_i	Frequência n_i	Porcentagem $100 f_i$
aberta nas laterais →	4,00 ┤ 8,00	6,00	10	27,78
	8,00 ┤ 12,00	10,00	12	33,33
	12,00 ┤ 16,00	14,00	8	22,22
	16,00 ┤ 20,00	18,00	5	13,89
	20,00 ┤ 24,00	22,00	1	2,78
informações sumarizadas	Total	—	36	100,00

Fonte: Tabela 2.4.

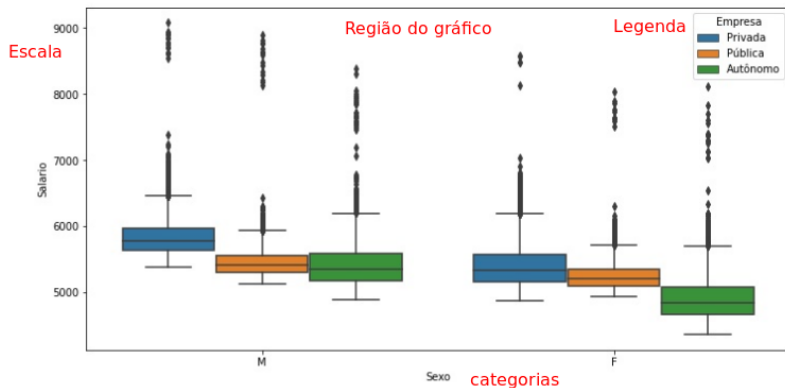
Fonte: Bussab, W.O.; Morettin, P.A. "Estatística Básica". 6ª ed., São Paulo: Saraiva, 2009.

Fonte

Representação gráfica

Numeração e título

Figura 1: Gráficos de caixas de salários por sexo e tipo de empresa.



Fonte: Elaborado pela autora.

Fonte

Representação tabular e gráfica de variáveis qualitativas

- Tabelas de frequências.
- Gráficos de barras
- Gráficos de Pareto.
- Gráficos de setores.

Tabelas de frequências

Sejam n elementos observados de uma população (uma amostra observada de tamanho n).

Tabelas de frequências

Sejam n elementos observados de uma população (uma amostra observada de tamanho n).

Suponha que existem k diferentes valores x_1, x_2, \dots, x_k , com $1 \leq k \leq n$. Cada um deles ocorre f_j vezes na amostra, $j = 1, \dots, k$.

Tabelas de frequências

Sejam n elementos observados de uma população (uma amostra observada de tamanho n).

Suponha que existem k diferentes valores x_1, x_2, \dots, x_k , com $1 \leq k \leq n$. Cada um deles ocorre f_j vezes na amostra, $j = 1, \dots, k$.

Exemplo

Dados observados (ordenados): 2, 2, 3, 3, 3, 3, 4, 5, 7, 8.

Valores distintos: 2, 3, 4, 5, 7, 8 com frequências 2, 4, 1, 1, 1, 1, respectivamente.

Tabelas de frequências

Definições

- ➊ **Frequência absoluta:** f_j : é o número de vezes que a variável X assume um determinado valor x_j na amostra.

Tabelas de frequências

Definições

- 1 **Frequência absoluta:** f_j : é o número de vezes que a variável X assume um determinado valor x_j na amostra.
- 2 **Frequência relativa:** f_j^* : é a frequência percentual do valor x_j na amostra, ou seja, com $f_j^* = \frac{f_j}{n}, j = 1, \dots, k$.

Tabelas de frequências

Definições

- 1 **Frequência absoluta:** f_j : é o número de vezes que a variável X assume um determinado valor x_j na amostra.
- 2 **Frequência relativa:** f_j^* : é a frequência percentual do valor x_j na amostra, ou seja, com $f_j^* = \frac{f_j}{n}, j = 1, \dots, k$.

Propriedades

- 1
$$\sum_{j=1}^k f_j = n$$

Tabelas de frequências

Definições

- 1 **Frequência absoluta:** f_j : é o número de vezes que a variável X assume um determinado valor x_j na amostra.
- 2 **Frequência relativa:** f_j^* : é a frequência percentual do valor x_j na amostra, ou seja, com $f_j^* = \frac{f_j}{n}, j = 1, \dots, k$.

Propriedades

- 1
$$\sum_{j=1}^k f_j = n$$
- 2
$$\sum_{j=1}^k f_j^* = 1.$$

Tabelas de frequências

Definições

- ❶ **Frequência acumulada:** F_j : é o número de vezes que a variável X um valor menor ou igual a x_j na amostra.

$$F_j = f_1 + f_2 + \dots + f_j$$

Tabelas de frequências

Definições

- ❶ **Frequência acumulada:** F_j : é o número de vezes que a variável X um valor menor ou igual a x_j na amostra.

$$F_j = f_1 + f_2 + \dots + f_j$$

- ❷ **Frequência relativa acumulada :** F_j^* : é a frequência percentual acumulada do valor x_j na amostra.

$$F_j^* = f_1^* + f_2^* + \dots + f_j^*$$

Tabelas de frequências

Definições

- ❶ **Frequência acumulada:** F_j : é o número de vezes que a variável X um valor menor ou igual a x_j na amostra.

$$F_j = f_1 + f_2 + \dots + f_j$$

- ❷ **Frequência relativa acumulada :** F_j^* : é a frequência percentual acumulada do valor x_j na amostra.

$$F_j^* = f_1^* + f_2^* + \dots + f_j^*$$

Propriedades

- ❶ $F_k = n$

Tabelas de frequências

Definições

- ❶ **Frequência acumulada:** F_j : é o número de vezes que a variável X um valor menor ou igual a x_j na amostra.

$$F_j = f_1 + f_2 + \dots + f_j$$

- ❷ **Frequência relativa acumulada :** F_j^* : é a frequência percentual acumulada do valor x_j na amostra.

$$F_j^* = f_1^* + f_2^* + \dots + f_j^*$$

Propriedades

❶ $F_k = n$

❷ $F_k^* = 1$

Tabelas de frequências

Tabela 1: Tabela de frequências de estado civil na Companhia MB
(Bussab & Morettin, 2009)

Estado civil	Frequência absoluta	Frequência relativa (%)	Frequência relativa acumulada (%)
Casado	20	55,56	55,56
Solteiro	16	44,44	100
Total	36	100	100

Fonte: Elaborada pela autora.

Tabelas de frequências

Tabela 1: Tabela de frequências de estado civil na Companhia MB
(Bussab & Morettin, 2009)

Estado civil	Frequência absoluta	Frequência relativa (%)	Frequência relativa acumulada (%)
Casado	20	55,56	55,56
Solteiro	16	44,44	100
Total	36	100	100

Fonte: Elaborada pela autora.

As classes podem ser ordenadas de forma decrescente pela frequência absoluta. para **variáveis qualitativas nominais**. Para **variáveis qualitativas ordinais**, costuma-se ordenar pela ordem das categorias.

Representação tabular e gráfica de variáveis qualitativas

- **Tabelas de frequências.**

Representação tabular e gráfica de variáveis qualitativas

- **Tabelas de frequências.**
- **Gráficos de barras:** Retângulos verticais (ou horizontais) espaçados com alturas (ou bases) iguais às frequências dos valores da variável.

Representação tabular e gráfica de variáveis qualitativas

- **Tabelas de frequências.**
- **Gráficos de barras:** Retângulos verticais (ou horizontais) espaçados com alturas (ou bases) iguais às frequências dos valores da variável.
- **Gráficos de Pareto:** Gráfico de barras com os valores da variável em ordem decrescente de frequências e com as frequências relativas acumuladas no segundo eixo vertical.

Representação tabular e gráfica de variáveis qualitativas

- **Tabelas de frequências.**
- **Gráficos de barras:** Retângulos verticais (ou horizontais) espaçados com alturas (ou bases) iguais às frequências dos valores da variável.
- **Gráficos de Pareto:** Gráfico de barras com os valores da variável em ordem decrescente de frequências e com as frequências relativas acumuladas no segundo eixo vertical.
- **Gráficos de setores:** Gráfico circular utilizado para destacar a composição das partes de um todo.

Gráfico de barras (R)

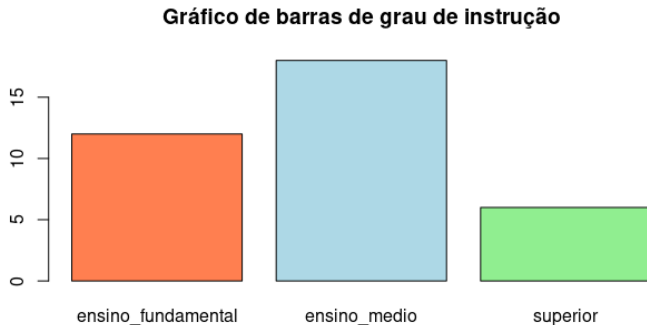


Gráfico de barras R: ggplot {ggplot2}

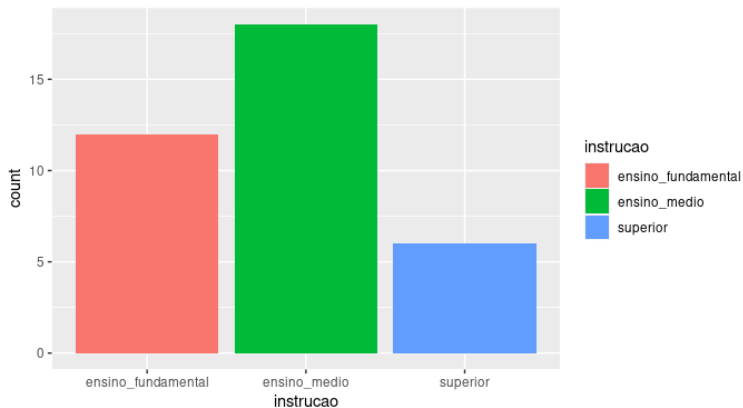


Gráfico de barras R: ggplot {ggplot2}

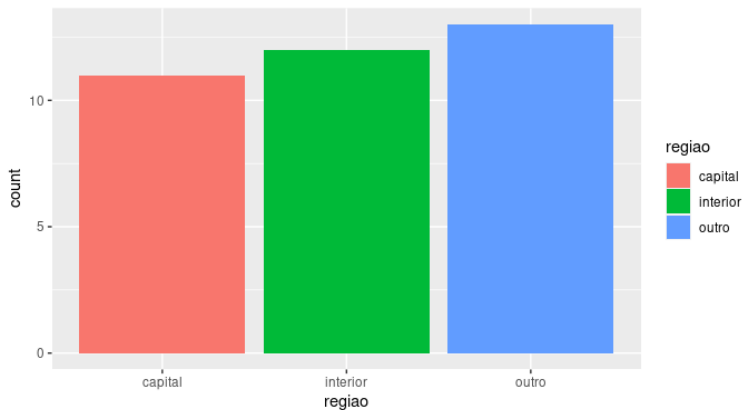


Gráfico de Pareto (R)

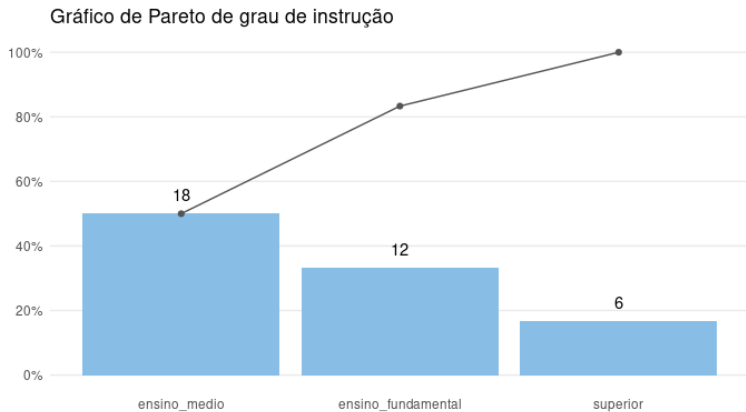


Gráfico de Pareto (Python)

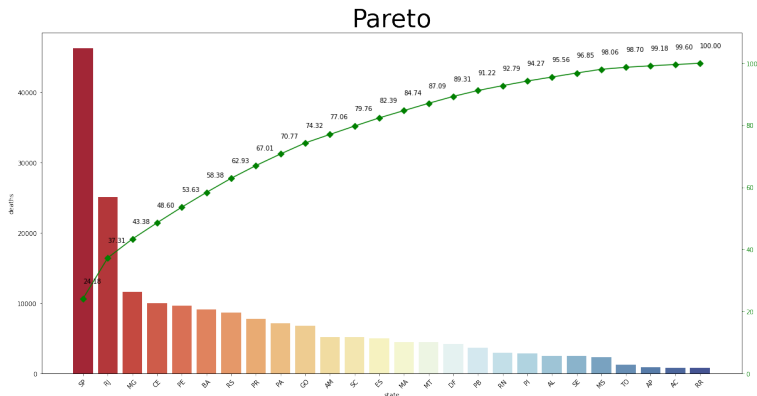


Figura: Gráfico de Pareto para mortes acumuladas por COVID-19 em 29/12/2020 em estados brasileiros.

Gráfico de setores (R)

Gráfico de setores para grau de instrução

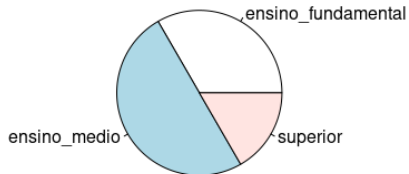


Gráfico de setores R: ggplot {ggplot2}

Gráfico de setores de grau de instrução

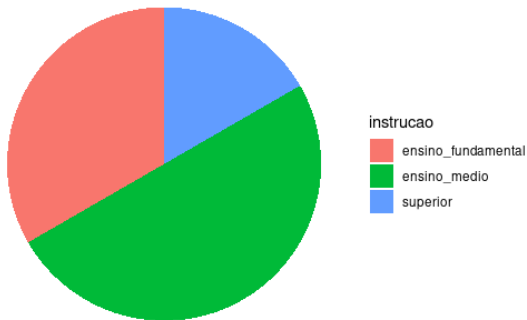


Gráfico de setores (R: ggplot {ggplot2})

Gráfico de setores de estado civil

