

Tipos de variáveis e formas de análise

Prof. Walmes M. Zeviani

Departamento de Estatística
Universidade Federal do Paraná





Tipos de variáveis

Tipos de variáveis



Figura 1. Principais tipos de variáveis. Fonte: Os autores.

Outros tipos de variáveis

O processamento e análise desses dados é particular de alguma forma.

- ▶ Dados circulares:
 - ▶ Direção do vento.
 - ▶ Hora de acesso ao Moodle.
- ▶ Dados composicionais:
 - ▶ Argila + Silte + Areia = 1.
 - ▶ Natação + Ciclismo + Corrida = tempo total do Triathlon.
- ▶ Dados de classificação:
 - ▶ 2º colocado, 71º colocado, etc.
 - ▶ Meu filme número 1, 2 e 3.
- ▶ Dados censurados:
 - ▶ Conteúdo de alumínio abaixo do limite de detecção: $< 0.5 \text{ cmol}_c/\text{dm}^3$.
 - ▶ Corrente elétrica acima da escala do voltímetro: $> 3 \text{ A}$.
 - ▶ A eclosão do inseto aconteceu entre 6h e 18h.
- ▶ Dados difusos (*fuzzy data*):
 - ▶ “frio”, “bem frio”, “uma friaca”, etc.
 - ▶ “está doendo”, “doendo muito”, “dor insuportável”, etc.
- ▶ Não estruturado:
 - ▶ Texto, imagem, áudio, vídeo, etc.



Formas de análise descritiva

Objetivos com a descrição dos dados

1. Investigar o **comportamento** de uma variável.
2. Examinar a **relação** entre variáveis.
3. Enfatizar a **ordenação**/classificação de elementos/categorias.
4. Compreender a estrutura **organização** dos elementos/categorias.
5. Explorar a **evolução** cronológica de uma variável.
6. Revelar padrões **espaciais** nos dados.
7. Descrever a **conexão** entre elementos/categorias.



Figura 3. Foto de fauxels no Pexels. Fonte: <https://rb.gy/lksjxt>.

As 7 funções gráficas

Distribuição: caracterizam a distribuição variáveis qualitativas ou quantitativas.

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

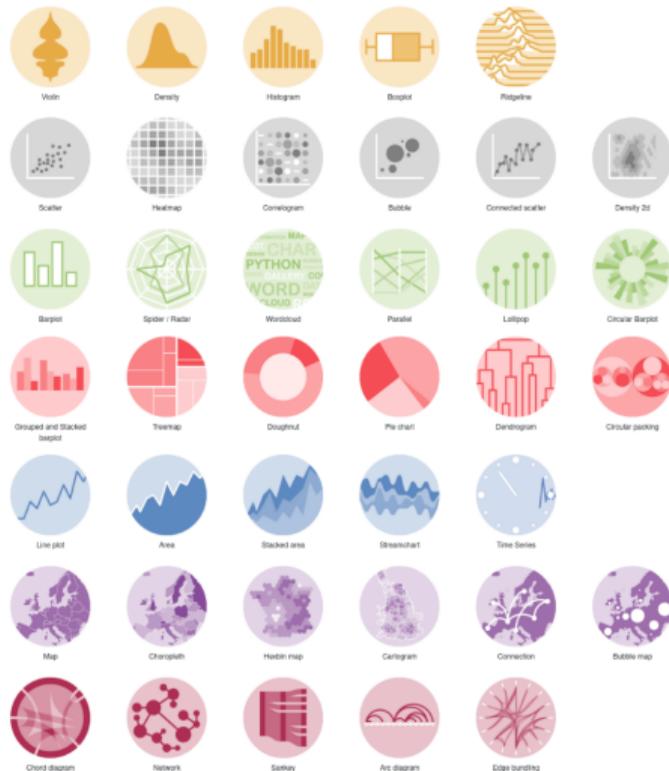
Ordenação: mostram a ordenação de categorias/elementos.

Organização: retratam a estrutura de organização de categorias/elementos.

Evolução: exploram a evolução cronológica de variáveis.

Mapas: representam dados com propriedades geográficas.

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



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

As medidas descritivas

- ▶ Variáveis qualitativas.
 - ▶ Tabelas de frequência (uma variável).
 - ▶ Tabelas de frequências cruzadas (mais de uma variável).
- ▶ Variáveis quantitativas.
 - ▶ Medidas de posição ou tendência central.
 - ▶ Medidas de dispersão ou variabilidade.
 - ▶ Separatrizes.
 - ▶ Medidas de associação.
 - ▶ Coeficientes/índices específicos.

```

Skin summary statistics
n obs: 336776
n variables: 19

--- Variable type:character ---
variable missing complete      n min max empty n_unique
carrier      0 336776 336776    2 2  0    16
dest         0 336776 336776    3 3  0   105
flight       0 336776 336776    1 4  0  3844
origin       0 336776 336776    3 3  0    3
tailnum     2512 334264 336776    5 6  0  4043

--- Variable type:integer ---
variable missing complete      n mean sd  p0  p25  p50  p75  p100  hist
arr_time      8713 328063 336776 1502.05 533.26  1 1104 1535 1940 2400
day           0 336776 336776   15.71  8.77  1  8  16  23  31
dep_time     8255 328521 336776 1349.11 488.28  1 907 1401 1744 2400
month         0 336776 336776    6.55  3.41  1  4  7  10  12
sched_arr_time 0 336776 336776 1536.38 497.46  1 1124 1556 1945 2359
sched_dep_time 0 336776 336776 1344.25 467.34 106 906 1359 1729 2359
year          0 336776 336776 2013 0 2013 2013 2013 2013

--- Variable type:numeric ---
variable missing complete      n mean sd  p0  p25  p50  p75  p100  hist
air_time     9430 327346 336776 150.69 93.69 20 82 129 192 695
arr_delay   9430 327346 336776    6.9 44.63 -86 -17 -5 14 1272
dep_delay   8255 328521 336776   12.64 40.21 -43 -5 -2 11 1301
distance    0 336776 336776 1039.91 733.23 17 502 872 1389 4983
hour        0 336776 336776   13.18  4.66  1  9  13  17  23
minute      0 336776 336776   26.23 19.3  0  8  29  44  59

--- Variable type:POSIXct ---
variable missing complete      n min max median n_unique
time_hour   0 336776 336776 2013-01-01 2013-12-31 2013-07-03 6936
  
```

Figura 4. Exemplos de medidas descritivas: Retirado de <https://www.pluralsight.com/guides/exploring-data-set-r>.

Considerações finais

Revisão

- ▶ Aspectos da análise descritiva de dados.
 - ▶ Tipos de variáveis.
 - ▶ Formas de descrição dos dados.
 - ▶ Objetivos da descrição de dados.
 - ▶ As principais medidas descritivas.
 - ▶ As principais funções gráficas.



Figura 5. Foto de fauxels no Pexels. Fonte: <https://rb.gy/lksjxt>.