

Estatística Descritiva

Apresentação da Unidade Didática

Prof. Walmes M. Zeviani

Departamento de Estatística
Universidade Federal do Paraná



O papel da Estatística Descritiva

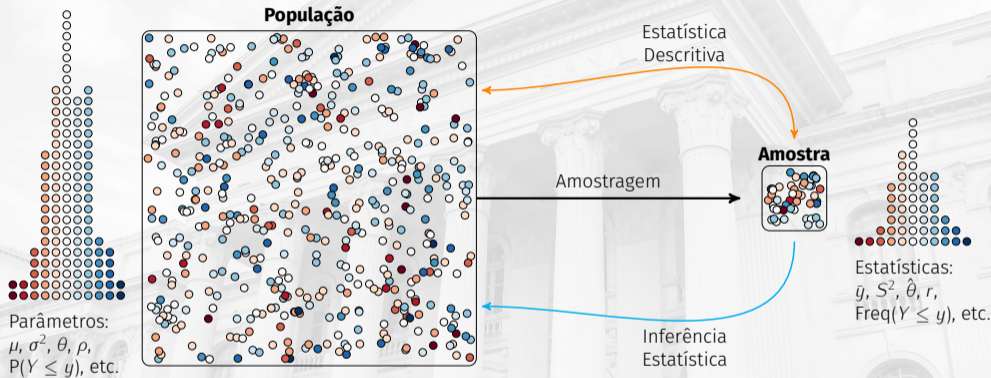


Figura 1. Representação esquemática do funcionamento da estatística descritiva e estatística inferencial.

Dashboards · Painéis de controle com Estatística Descritiva

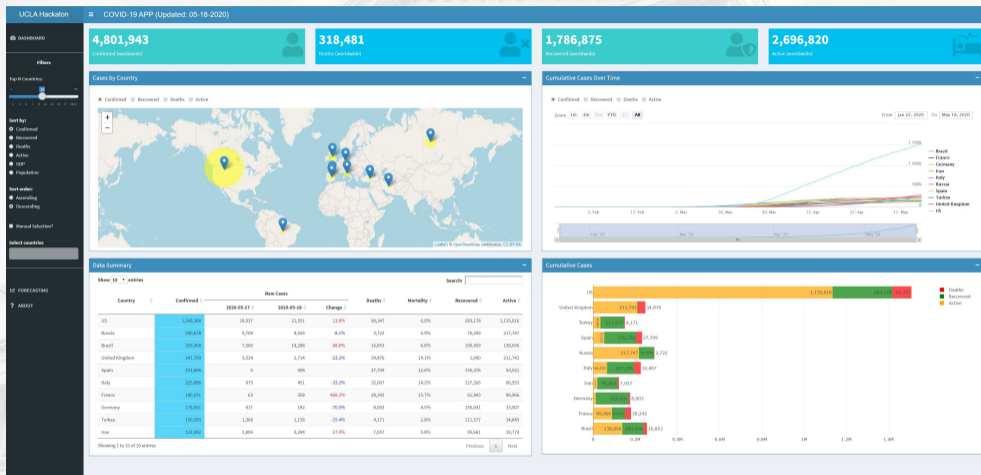


Figura 2. Dashboard de Covid-19. Fonte: <https://user-images.githubusercontent.com/44356224/82365851-21ed5e00-99c6-11ea-87b3-d190c2e0dba7.JPG>.

Dashboards · dados em tempo real



Figura 3. Exemplo de visualização de dados. Fonte: <https://www.bizsoft.co.za/wp-content/uploads/2018/08/mobile-data-analytics.png>.

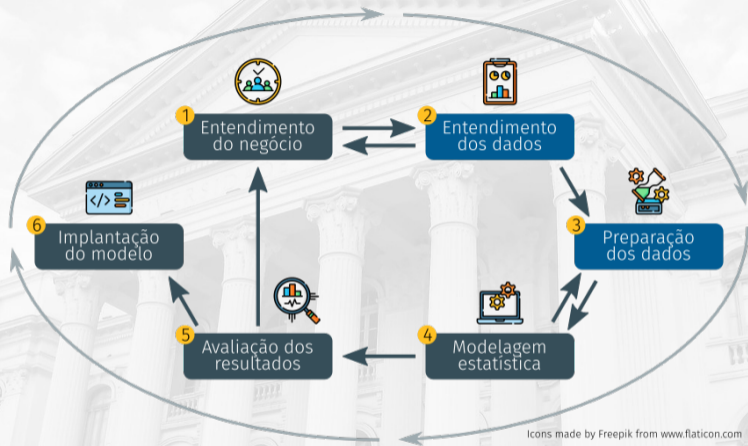
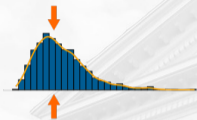


Figura 4. O CRISP-DM (Cross Industry Standard Process for Data Mining) estabelece etapas para a análise de dados. As caixas em azul são as etapas que envolvem análise exploratória de dados.

Medidas de posição



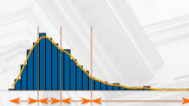
Expressam:

- Posição central ocupada pela variável.
- Valores que dividem a amostra em partes iguais.
- Valores típicos onde há maior densidade.

São elas:

- Média aritmética.
- Mediana.
- Moda.
- Média geométrica.
- Média harmônica.
- Média aparada.

Medidas de posição relativa



Expressam:

- Pontos no domínio da variável que definem porções com frequências conhecidas.

São elas:

- Quartis.
- Decis.
- Percentis.
- Máximo e mínimo.

Figura 5. Medidas de posição usadas em análise descritiva de dados.

Medidas de dispersão



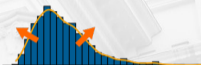
Expressam:

- O domínio observado da variável.
- O grau de dispersão ao redor do centro.
- O distanciamento médio dos valores.

São elas:

- Amplitude total.
- Variância.
- Desvio-padrão.
- Desvio absoluto médio e mediano.
- Coeficiente de variação.

Medidas de forma



Expressam:

- Aspectos da forma.
- Assimetria.
- Curtose.

São elas:

- Coeficiente de assimetria.
- Coeficiente de curtose.

Figura 6. Medidas de dispersão e forma usadas em análise descritiva de dados.

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



Figura 8. Foto de fauxels no Pexels. Fonte: <https://rb.gy/lksjxt>.

Distribuição: caracterizam a distribuição variáveis qualitativas ou quantitativas.

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

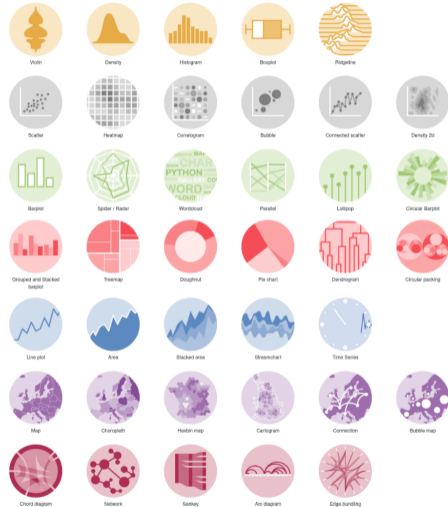
Ordenação: mostram a ordenação de categorias/elementos.

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

- ▶ Motivação e utilidade da Estatística Descritiva.
- ▶ Tipos de variáveis e formas de análise.
- ▶ Distribuição de frequências.
- ▶ Medidas descritivas numéricas.

Os números têm uma importante história para contar. Eles dependem de você dar-lhes uma clara e convincente voz.

– Stephen Few