

Iniciação à Estatística

Prof. Fernando de Souza Bastos
fernando.bastos@ufv.br

Departamento de Estatística
Universidade Federal de Viçosa
Campus UFV - Viçosa



Sumário

1 Preliminares

2 Introdução

Em alguma fase de seu trabalho, o pesquisador depara-se com o problema de analisar e entender um conjunto de dados relevante ao seu particular objeto de estudos. Ele necessitará trabalhar os dados para transformá-los em informações, para compará-los com outros resultados, ou ainda para julgar sua adequação a alguma teoria.

Não Tem Como Escapar dos Dados!

Figura: Era do Big Data



Fonte: Revista Exame 2017 Oliveira (2017)

- Google: são mais de 5,5 bilhões de buscas por dia (ARDORSEO, 2021).

- Google: são mais de 5,5 bilhões de buscas por dia (ARDORSEO, 2021).
- Youtube (2021): mais de 2 bilhões de usuários e são assistidas mais de 1 bilhão de horas de vídeos por dia.

- Google: são mais de 5,5 bilhões de buscas por dia (ARDORSEO, 2021).
- Youtube (2021): mais de 2 bilhões de usuários e são assistidas mais de 1 bilhão de horas de vídeos por dia.
- Facebook: Há picos de 2,8 bilhões de usuários ativos diariamente (INTERNET LIVE STATS, 2021a).

- Google: são mais de 5,5 bilhões de buscas por dia (ARDORSEO, 2021).
- Youtube (2021): mais de 2 bilhões de usuários e são assistidas mais de 1 bilhão de horas de vídeos por dia.
- Facebook: Há picos de 2,8 bilhões de usuários ativos diariamente (INTERNET LIVE STATS, 2021a).
- Instagram: Mais de 1 bilhão de usuários, mais de 500 milhões de pessoas usam o *stories* todos os dias (NEWBERRY, 2021).

- Google: são mais de 5,5 bilhões de buscas por dia (ARDORSEO, 2021).
- Youtube (2021): mais de 2 bilhões de usuários e são assistidas mais de 1 bilhão de horas de vídeos por dia.
- Facebook: Há picos de 2,8 bilhões de usuários ativos diariamente (INTERNET LIVE STATS, 2021a).
- Instagram: Mais de 1 bilhão de usuários, mais de 500 milhões de pessoas usam o *stories* todos os dias (NEWBERRY, 2021).
- WhatsApp: mais de 100 bilhões de mensagens por dia e mais de 2 bilhões de usuários (DEAN, 2021).

- Google: são mais de 5,5 bilhões de buscas por dia (ARDORSEO, 2021).
- Youtube (2021): mais de 2 bilhões de usuários e são assistidas mais de 1 bilhão de horas de vídeos por dia.
- Facebook: Há picos de 2,8 bilhões de usuários ativos diariamente (INTERNET LIVE STATS, 2021a).
- Instagram: Mais de 1 bilhão de usuários, mais de 500 milhões de pessoas usam o *stories* todos os dias (NEWBERRY, 2021).
- WhatsApp: mais de 100 bilhões de mensagens por dia e mais de 2 bilhões de usuários (DEAN, 2021).
- Twitter: 500 milhões de tweets por dia e 340 milhões de usuários ativos (ASLAM, 2021).

- Internet: mais de 1,8 bilhões de sites (**INTERNET LIVE STATS, 2021b**) e 4,6 bilhões de pessoas conectadas (**STATISTA, 2021**).

- Internet: mais de 1,8 bilhões de sites (**INTERNET LIVE STATS, 2021b**) e 4,6 bilhões de pessoas conectadas (**STATISTA, 2021**).
- Telefones celulares: mais de 5,7 bilhões (**DATAREPORTAL, 2021**).

- Internet: mais de 1,8 bilhões de sites ([INTERNET LIVE STATS, 2021b](#)) e 4,6 bilhões de pessoas conectadas ([STATISTA, 2021](#)).
- Telefones celulares: mais de 5,7 bilhões ([DATAREPORTAL, 2021](#)).
- Dispositivos conectados: serão 50 bilhões em 2030 ([STATISTA, 2021](#)).

- Internet: mais de 1,8 bilhões de sites ([INTERNET LIVE STATS, 2021b](#)) e 4,6 bilhões de pessoas conectadas ([STATISTA, 2021](#)).
- Telefones celulares: mais de 5,7 bilhões ([DATAREPORTAL, 2021](#)).
- Dispositivos conectados: serão 50 bilhões em 2030 ([STATISTA, 2021](#)).
- De acordo com [Vuleta \(2021\)](#), em 2025, a quantidade de dados gerados a cada dia deve chegar a 463 exabytes (10^{18} bytes) globalmente.

A IDC (empresa líder em inteligência de mercado e consultoria nas indústrias de tecnologia da informação, telecomunicações e mercados de consumo em massa de tecnologia) estimou, em 2014, que do total de dados no mundo, 22% contêm informação útil. E apenas 5% foram analisados e utilizados de alguma forma (GREGO, 2014).

Introdução

Estatística é a ciência que nos ajuda a tomar decisões e tirar conclusões na presença de variabilidade, é uma maneira de raciocinar que pode ajudar você a tomar decisões mais bem fundamentadas e pode ajudá-lo a solucionar problemas que envolvem decisões que estão baseadas em dados que tenham sido coletados.

Introdução

A estatística é a arte de torturar os números até que eles confessem!
*Como mentir com estatísticas (HUFF, 2016) (reedição de 1954)

Na primeira parte deste curso estaremos interessados em:

- Reduzir ou organizar, limpar e resumir, analisar e interpretar dados!

Na primeira parte deste curso estaremos interessados em:

- Reduzir ou organizar, limpar e resumir, analisar e interpretar dados!

Essa será a parte inicial do curso denominada Análise exploratória de dados (AED). Queremos nessa fase obter dos dados a maior quantidade possível de informação, para indicar modelos plausíveis a serem utilizados na inferência estatística.

Introdução

Vamos aprender e calcular algumas medidas de posição e de variabilidade. E aprender técnicas gráficas para a análise exploratória de dados.

Vamos utilizar:

- Roteiro de Aulas (EST 105) (s.d.).
- (Sugestão de Livro) Estatística básica (2017) dos autores Wilton de Oliveira Bussab e Pedro Alberto Morettin;
- (Sugestão de Livro) Introdução à Estatística com R (2020) dos autores Eric Batista Ferreira e Marcelo Silva de Oliveira.

Referências I

ARDORSEO. 11 Search Statistics You Need to Know in 2021. 2021. Disponível em: <<https://ardorseo.com/blog/how-many-google-searches-per-day/>>.

ASLAM, Salman. Twitter by the Numbers: Stats, Demographics & Fun Facts. 2021. Disponível em: <<https://www.omnicoreagency.com/twitter-statistics/>>.

DATAREPORTAL. DIGITAL AROUND THE WORLD. 2021. Disponível em: <<https://datareportal.com/global-digital-overview>>.

DEAN, Brian. WhatsApp 2021 User Statistics: How Many People Use WhatsApp? 2021. Disponível em: <<https://backlinko.com/whatsapp-users>>.

Referências II

FERREIRA, Eric Batista; OLIVEIRA, Marcelo Silva de. Introdução à Estatística com R. [S.l.]: Editora Universidade Federal de Alfenas, 2020. Disponível em:

<https://www.unifal-mg.edu.br/bibliotecas/wp-content/uploads/sites/125/2021/12/32-EBR_Unifal.pdf>.

GREGO, Maurício. Conteúdo digital dobra a cada dois anos no mundo. 2014. Disponível em:

<<https://exame.abril.com.br/tecnologia/conteudo-digital-dobra-a-cada-dois-anos-no-mundo/>>.

HUFF, Darrell. Como mentir com estatística. [S.l.]: Editora Intrínseca, 2016.

INTERNET LIVE STATS. Real Time Statistics Project. 2021. Disponível em: <<https://www.internetlivestats.com/>>.

Referências III

INTERNET LIVE STATS. Total number of Websites. 2021.
Disponível em: <<https://www.internetlivestats.com/total-number-of-websites/>>.

MORETTIN, Pedro Alberto; BUSSAB, Wilton Oliveira. Estatística básica. [S.l.]: Saraiva Educação, 2017.

NEWBERRY, Christina. 44 Instagram Stats That Matter to Marketers in 2021. 2021. Disponível em:
<<https://blog.hootsuite.com/instagram-statistics/>>.

OLIVEIRA, M. Chamam de Big Data, mas deveria ser Best Data. 2017. Disponível em:
<<https://exame.abril.com.br/blog/relacionamento-antes-do-marketing/chamam-de-big-data-mas-deveria-ser-best-data/>>.

Referências IV

REGAZZI, Adair José; SILVA, Carlos Henrique Osório; SANTOS, Gerson Rodrigues dos; EMILIANO, Paulo César; BARBOSA, Eduardo Campana. **Roteiro de Aulas (EST 105)**. [S.l.: s.n.]. Disponível no PVANet - Moodle.

STATISTA. Global digital population as of January 2021 (in billions). 2021. Disponível em: <<https://www.statista.com/statistics/617136/digital-population-worldwide/>>.

VULETA, Branka. How Much Data Is on the Internet? 2021. Disponível em: <<https://seedscientific.com/how-much-data-is-created-every-day/>>.

YOUTUBE. 2021. Disponível em: <<https://www.youtube.com/intl/pt-BR/about/press/>>.