

Introdução à Ciência de Dados





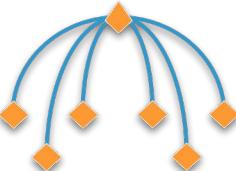
Jailson Pereira

Machine Learning
Researcher no
LSI/UEA.

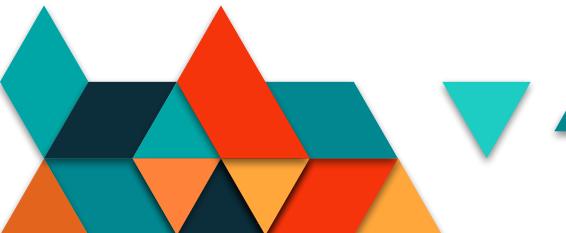


Juliany Raiol

Software Developer
no INDT.



@pydatamanaus





SUMÁRIO

- 
- 01 BIG DATA
 - 02 O QUE É CIÊNCIA DE DADOS?
 - 03 FLUXO DE PROCESSO
 - 04 CASES DE SUCESSO



Dados nos dias atuais

- A quantidade de dados vem crescendo mais rápido do que nunca
- 1,7 MB por segundo por cada pessoas



Dados nos dias atuais



Redes Sociais



Governo



Cidades Inteligentes

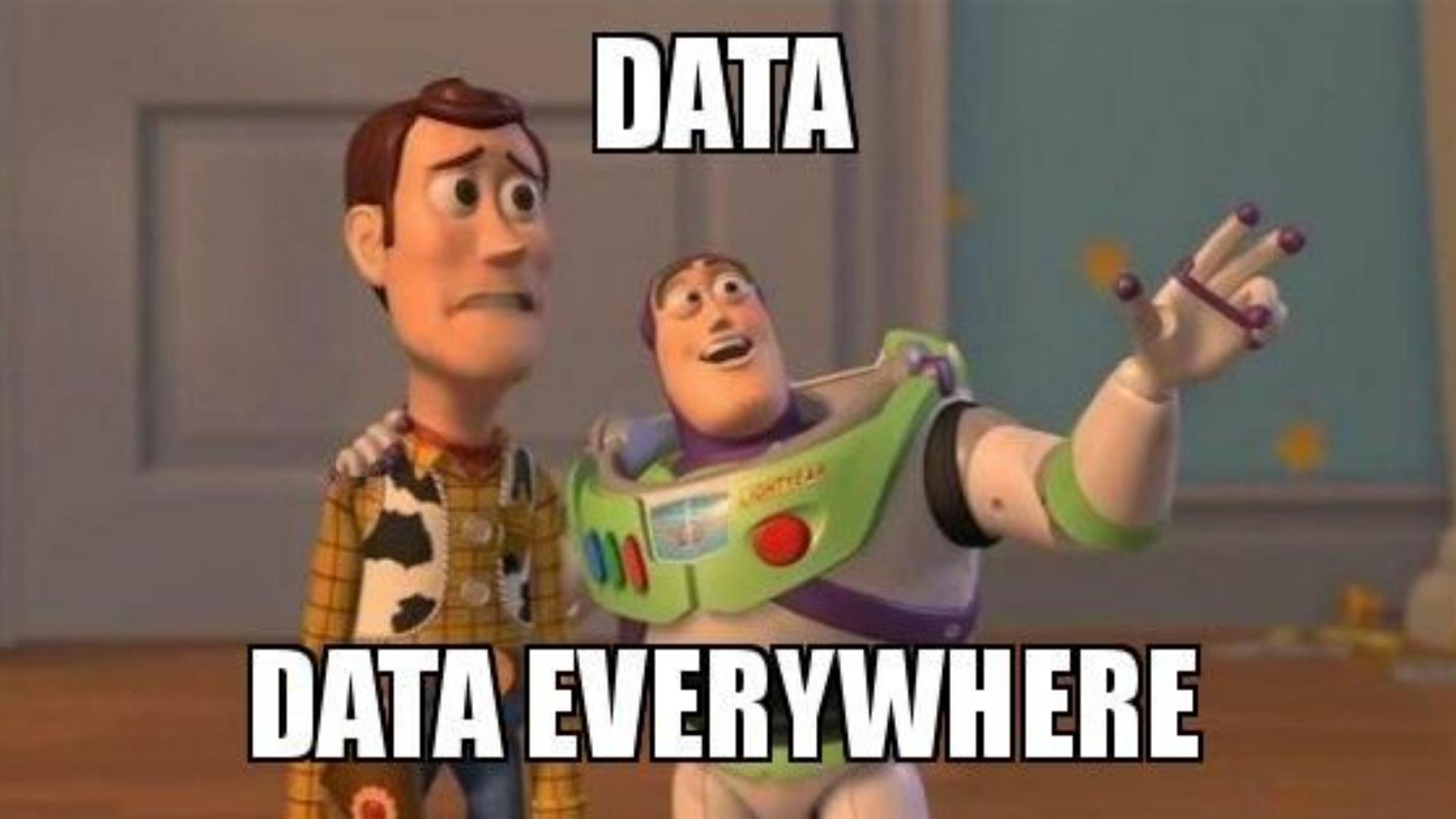


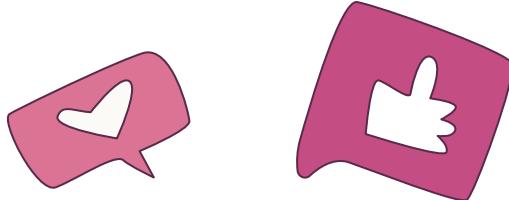
Dispositivos
Pessoais



DATA

DATA EVERYWHERE

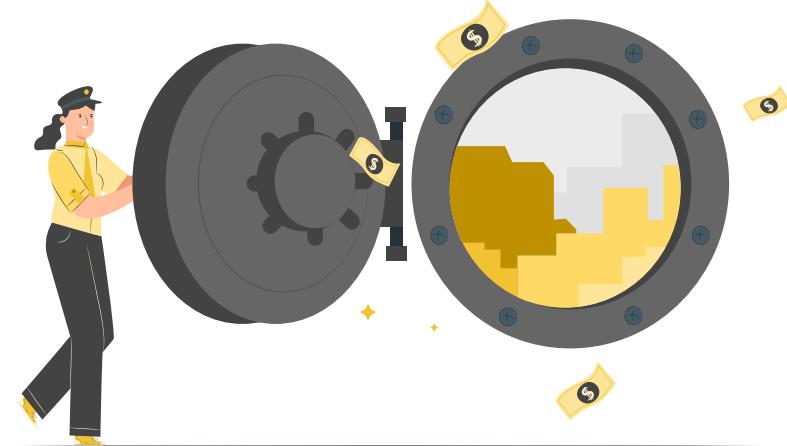




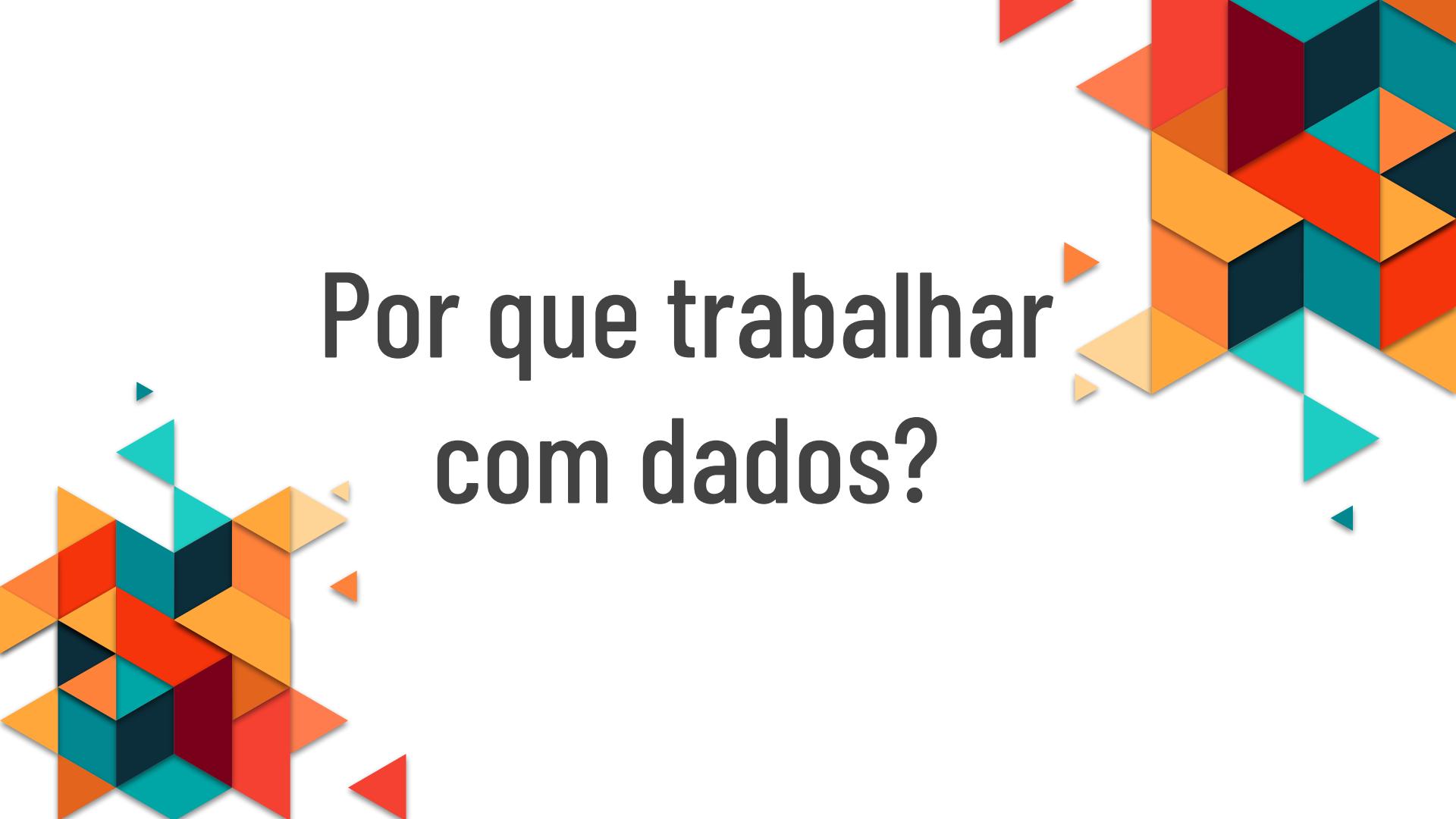
**65 bilhões de mensagens
são enviadas no WhatsApp**



É dado pra dar e vender!



Por que trabalhar com dados?



“Os dados podem ser o microscópio ou o telescópio de nossa era, possibilitando percepções importantes e até revolucionárias”

Seth Stephens-Davidowitz



Por que trabalhar com dados?



X



Intuição / Experiência

Dados



Por que trabalhar com dados?



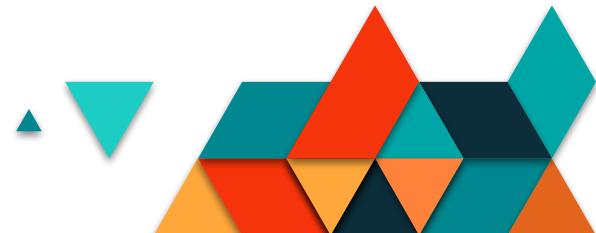
Ideias ou
Questionamentos



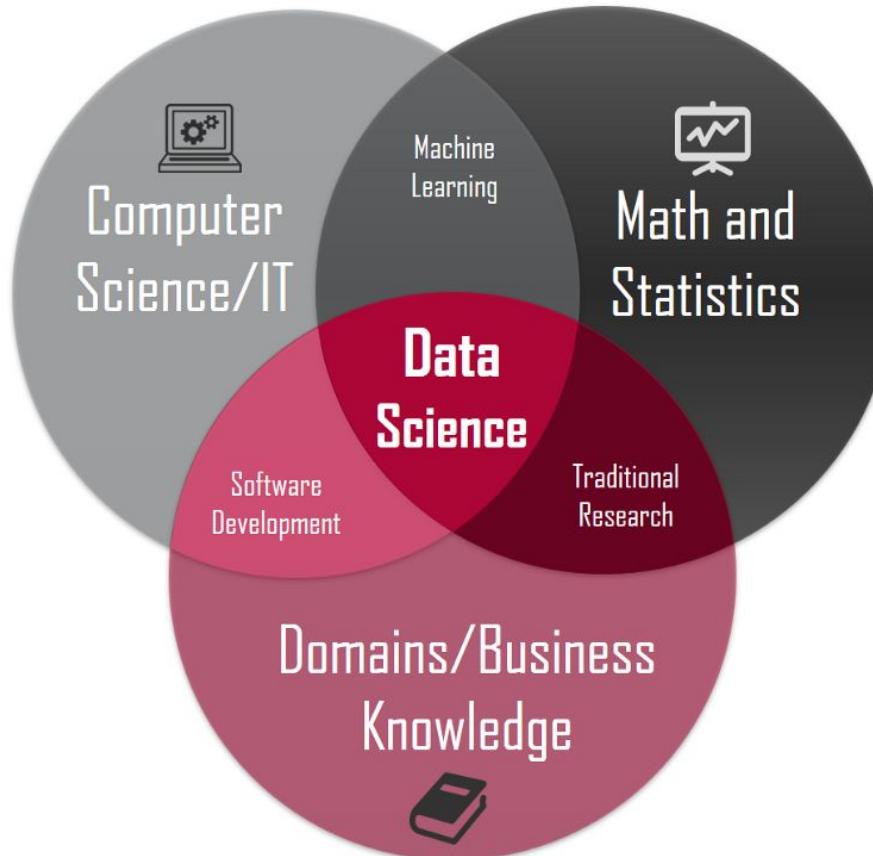
Dados



Descobertas

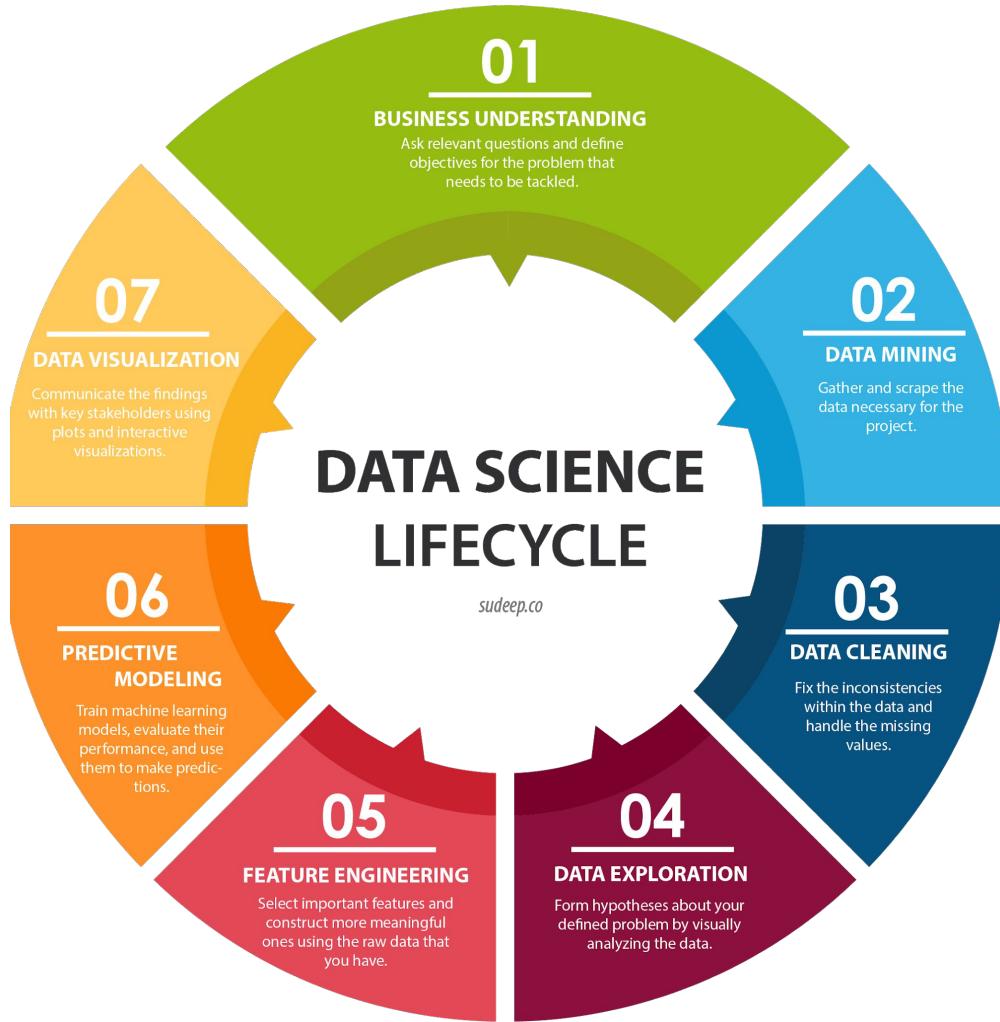


Ciência de Dados



Mas como funciona na prática?





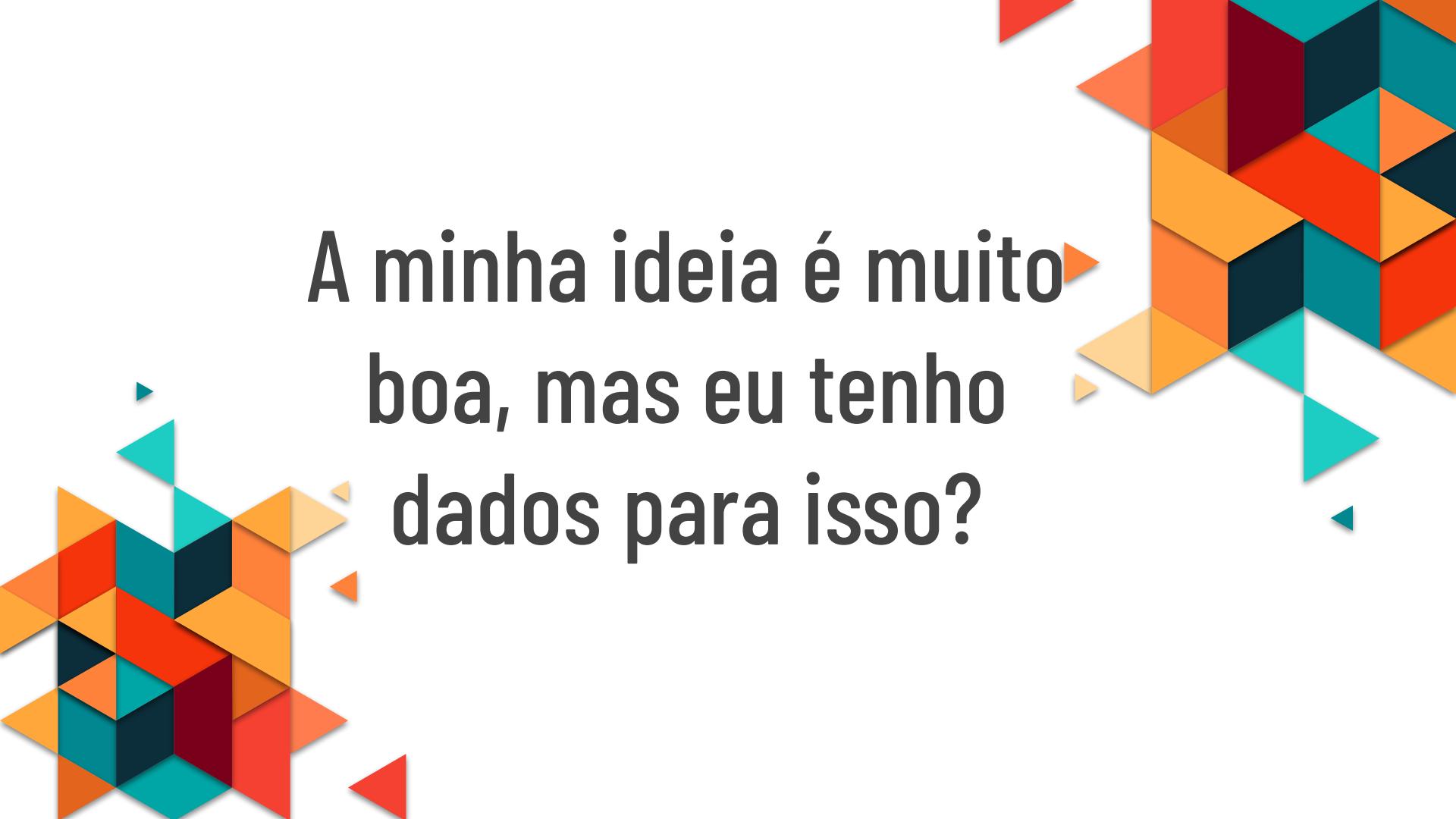
Trabalhando com dados

Mais importante do que extrair informações dos dados é fazer as perguntas corretas.

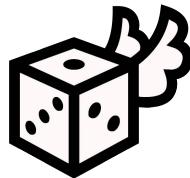
Você não precisa de muitos dados, precisa dos dados certos.



A minha ideia é muito
boa, mas eu tenho
dados para isso?



Como ter acesso aos dados?



dados.gov.br
®



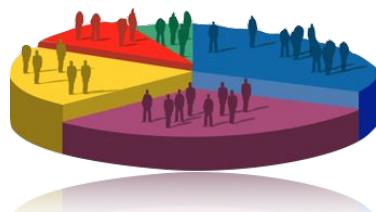
kaggle™



Algumas habilidades necessárias



Comunicação



Estatística



Programação



Entender sobre o
negócio ou o
problema que se
quer resolver





Ferramentas para trabalhar com dados



Por onde começar?

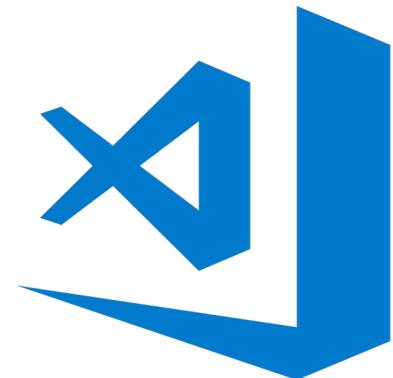
O primeiro passo é aprender **Python <3** (pode ser **R** tbm rs)



Ambientes de desenvolvimento



colab



Pandas



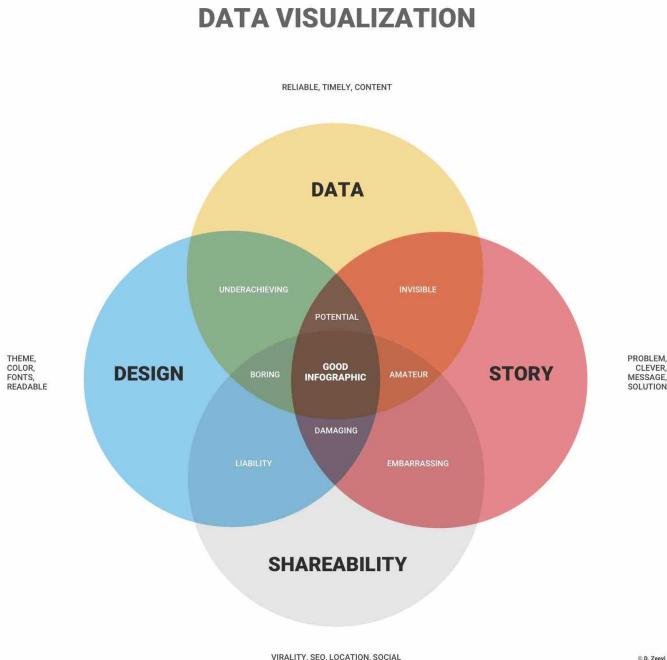
É possível trabalhar com



Entre outros...



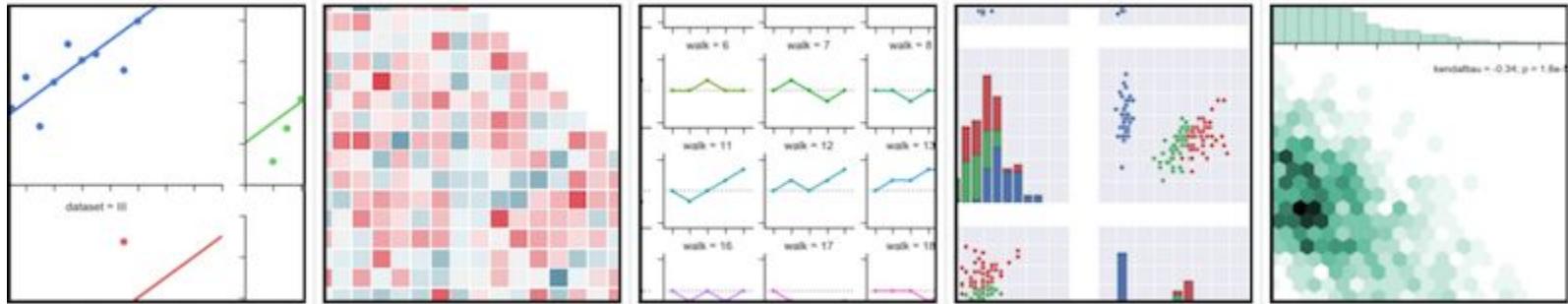
Visualização de Dados



“Ideias complexas comunicadas com clareza, precisão e eficiência”

Edward Tufte

Seaborn



- Biblioteca de visualização de dados
- Gráficos estatísticos e informativos
- <https://seaborn.pydata.org/>

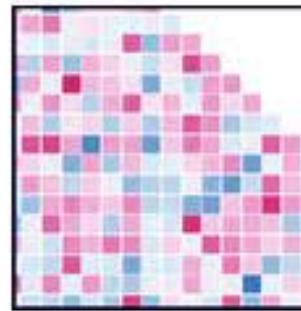




Bokeh



matplotlib



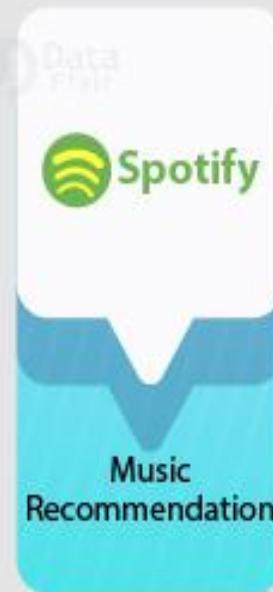
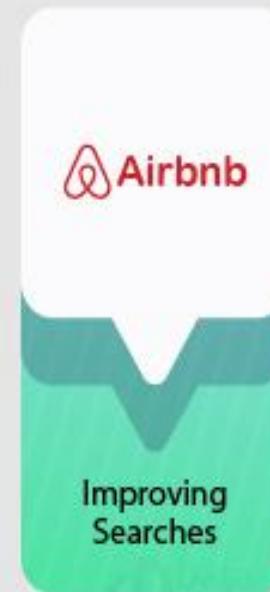
Seaborn



Transformando dados em insights



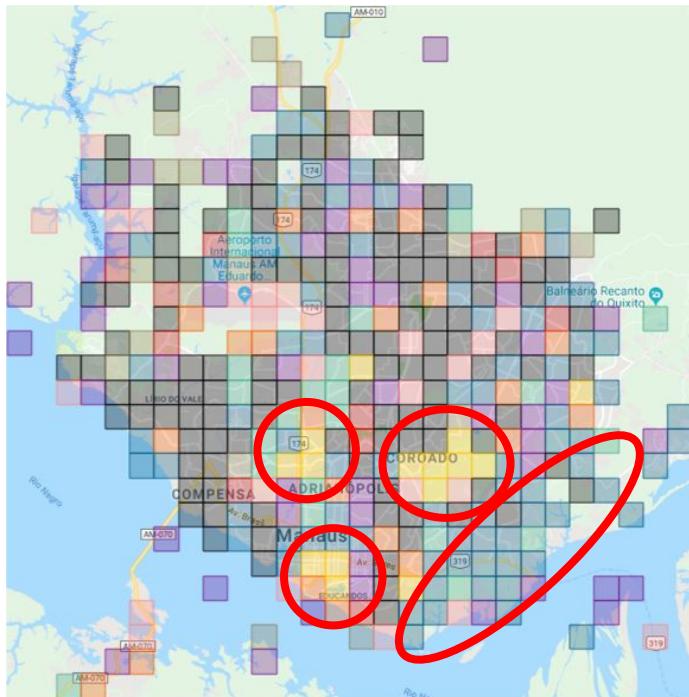
Data Science Use Cases



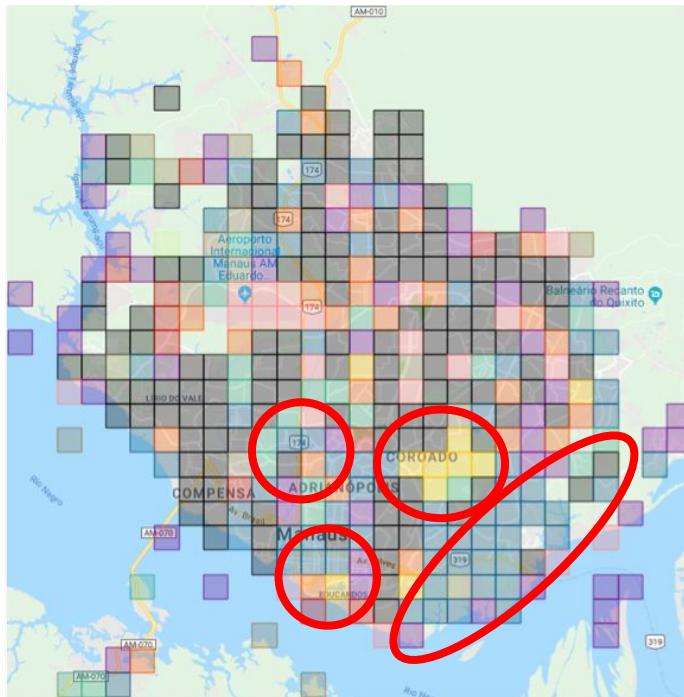
Ciência de Dados Espacial - Exemplo (Manaus)

Categoria do Foursquare	Exemplos de locais
<i>Arts and Entertainment</i>	Cinemas, museus e casinos.
<i>College and University</i>	Escolas, laboratórios universitários e centros de estudo.
<i>Food</i>	Restaurantes, cafeterias e padarias.
<i>Nightlife Spot</i>	Bares, clubes de <i>rock</i> e boates.
<i>Outdoors and Recreation</i>	Parques, academias e praias.
<i>Professional and Other Places</i>	Fábricas, auditórios e centros médicos.
<i>Shop and Service</i>	Lojas, salões de beleza e supermercados.
<i>Travel and Transport</i>	Aeroportos, estradas e hotéis.
<i>Residence</i>	Propriedades privadas, prédios e reboques residenciais.

Ciência de Dados Espacial - Exemplo (Manaus)



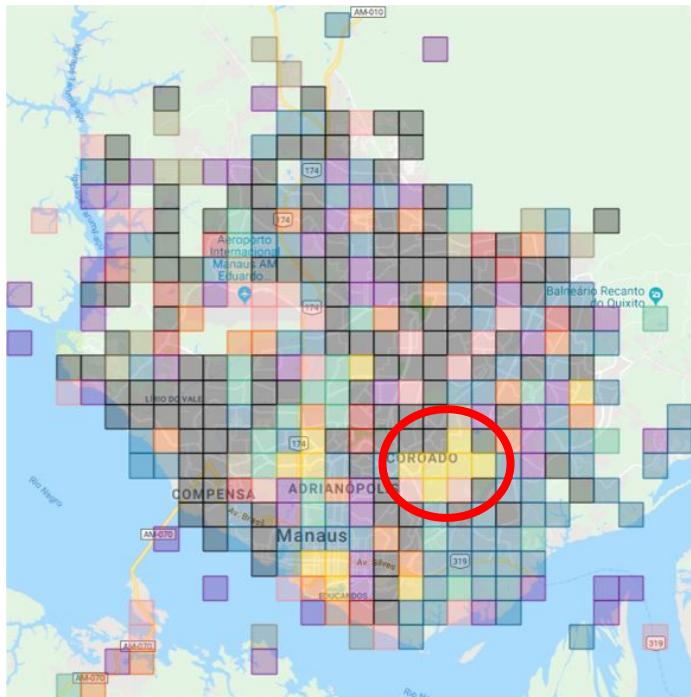
Período ativo



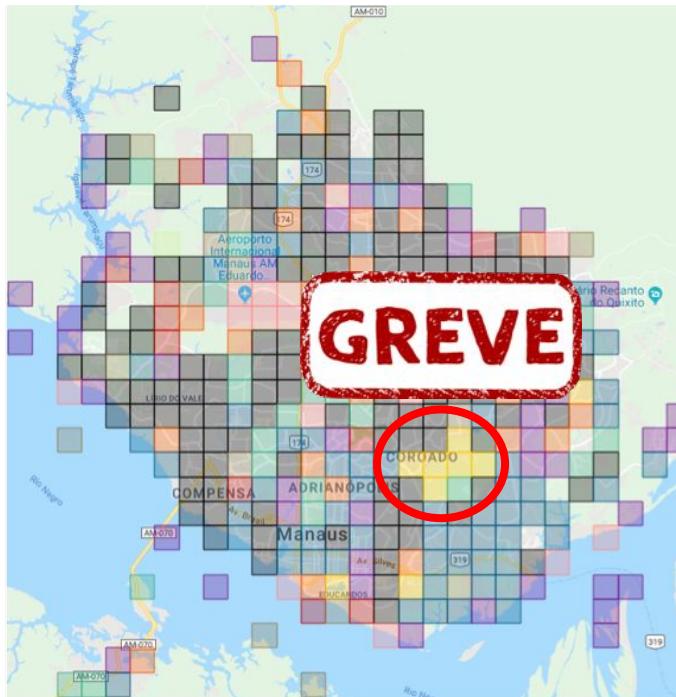
Período inativo



Ciência de Dados Espacial - Exemplo (Manaus)



Período ativo



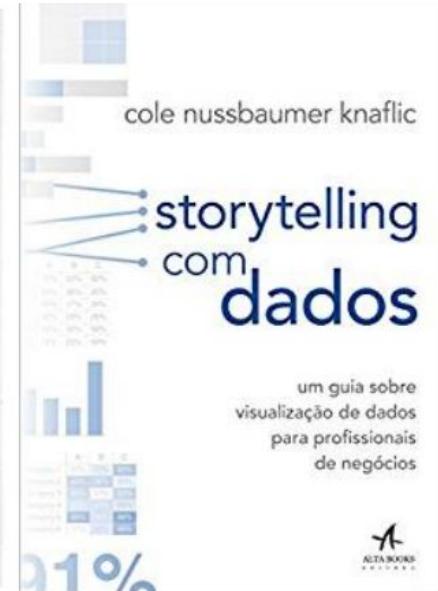
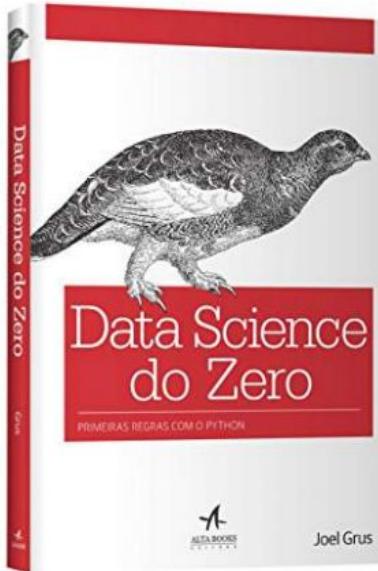
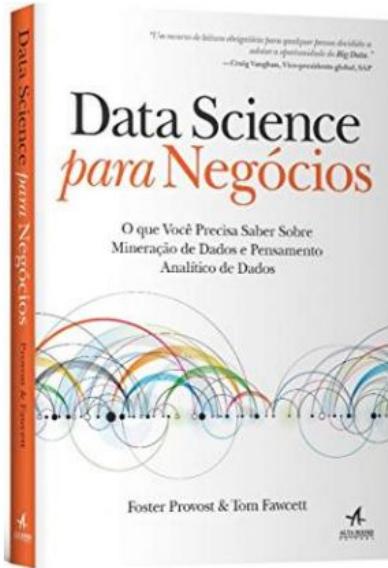
Período inativo



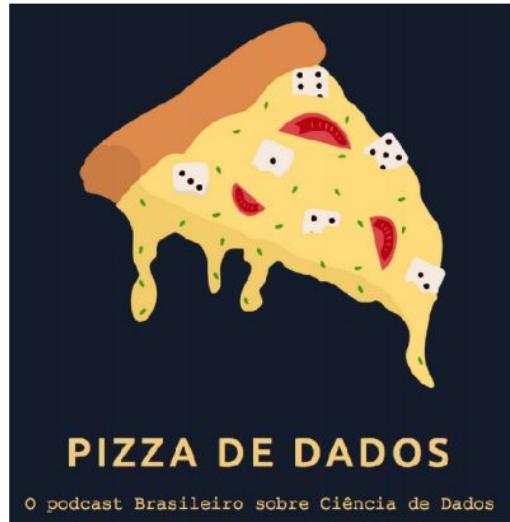
Quero continuar
aprendendo. E
agora?



Livros!



Podcasts!



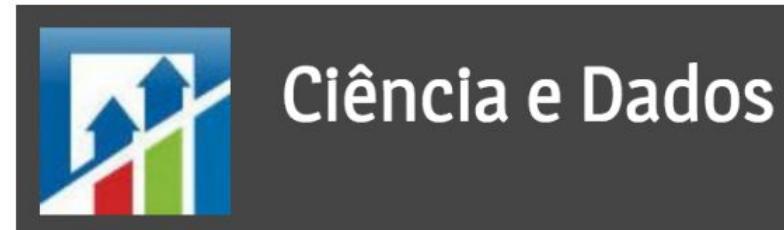
Cursos Online



Data Science
Academy



Blogs





PyData
MANAUS

bit.ly/manaus-digital

NUMFOCUS
OPEN CODE = BETTER SCIENCE

pydata.org



PyData

MANAUS



t.me/pydatamanaus



pydata.manaus@gmail.com



[@pydatamanaus](https://www.facebook.com/@pydatamanaus)



[@pydatamanaus](https://twitter.com/@pydatamanaus)



[@pydatamanaus](https://www.instagram.com/@pydatamanaus)