

Você quer ser um cientista de dados?

Leandro Humberto,
Data Engineer
Data Scientist



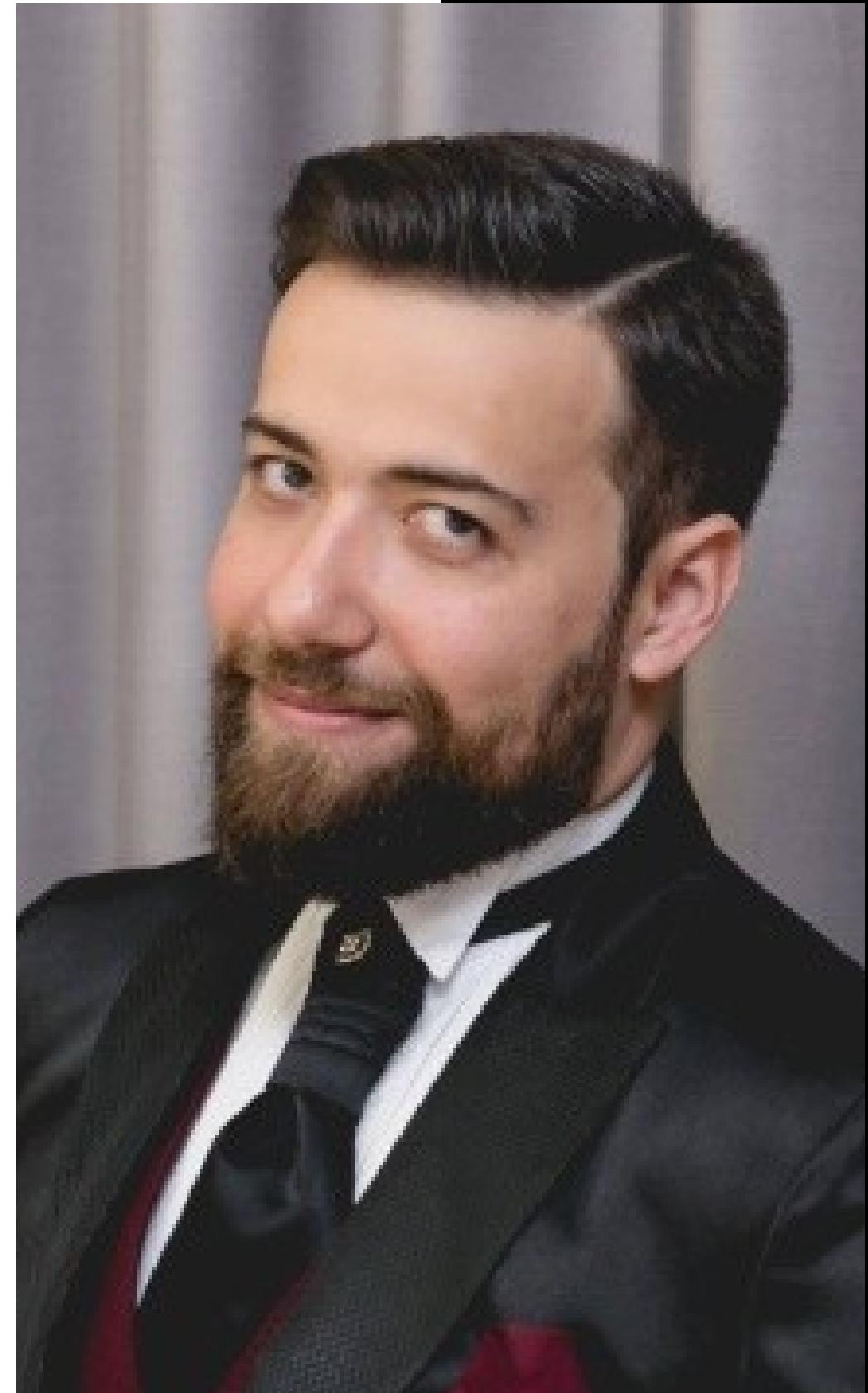
22º PYDATA
BRASILIA

Leandro Humberto



DATA @ BBTS

6 anos de experiência em projetos e tecnologias baseadas em dados



@PROFISSIONALDEDADOS



/IN/BEDATADRIVEN/

Objetivos da apresentação

- Expectativa da ciência de dados
- Problemas que ninguém te conta
- Carreiras na área de dados



Como me tornei cientista de dados



2017

PROCURA

Analista de BI queria inventar uma modinha

Procurei na internet como implementar a solução que queria

PROFISSÃO

Descobri que tinha um cargo somente para fazer este tipo de soluções

Tenha um Salário de R\$ 12.000 à
R\$ 25.000 reais atuando como
um Cientista de Dados.

Executivos **estão à caça** de profissionais
especialistas em **análise de dados**

Falta de talentos com capacidade analítica é uma das principais dificuldades encontradas por empregadores, num mercado que exige competências específicas. A potência desse conhecimento é cada vez mais **imprescindível**

≡ EXAME

Previdência FGTS STF Revista Newsletter

CARREIRA – VOCÊ S/A

Cientista de dados: a profissão do futuro continua em alta

Profissionais que atuam com métodos estatísticos e computacionais são cada vez mais disputados pelas empresas

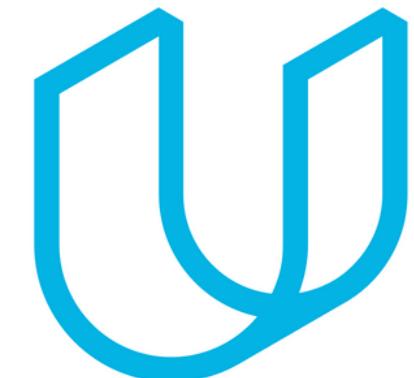
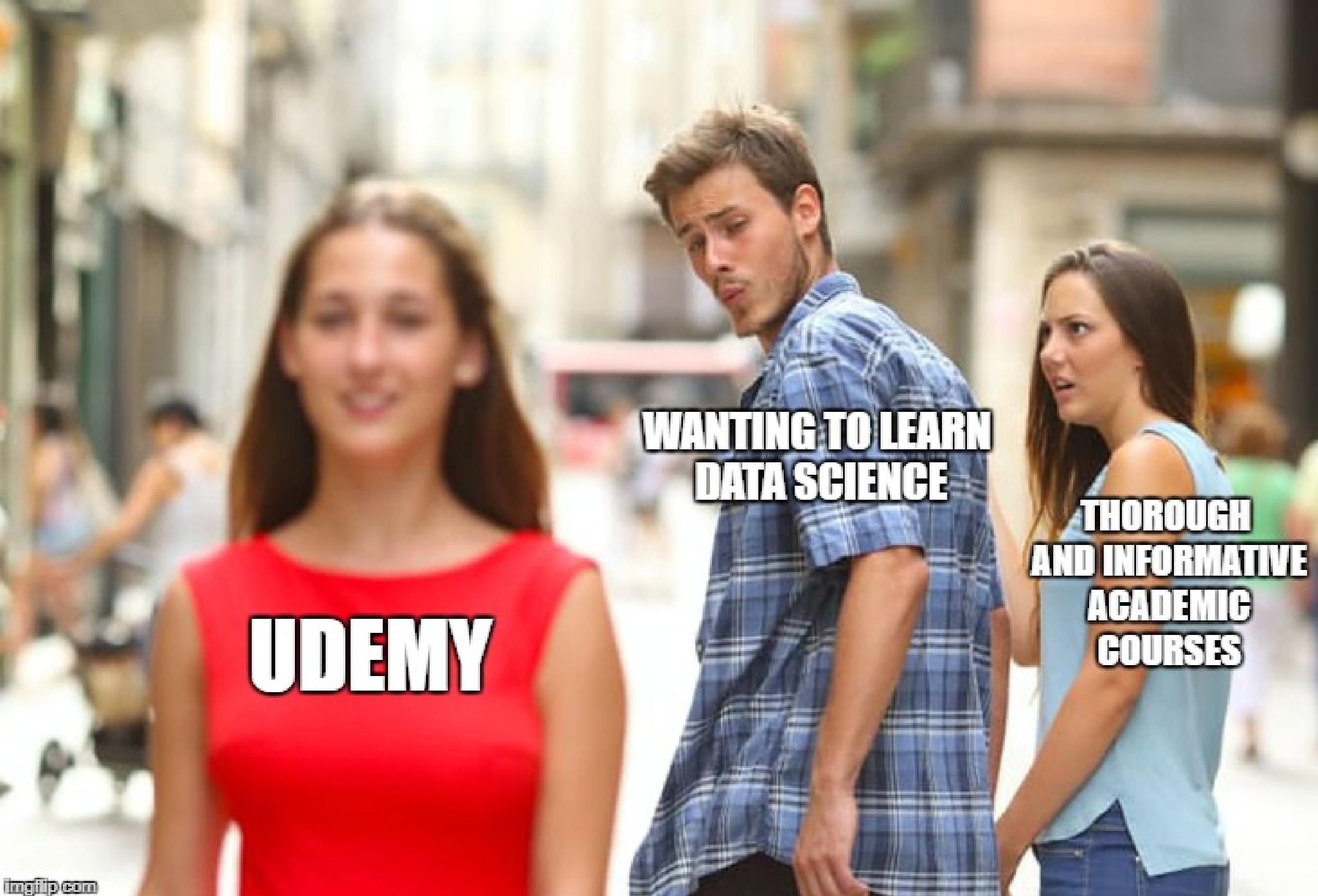
Home > Carreira

Profissão de cientista de dados é a mais procurada do século 21

DATA SCIENTIST

A NEW BREED OF DATA JEDI

I know Machine Learning



UDACITY



Data Science
Academy



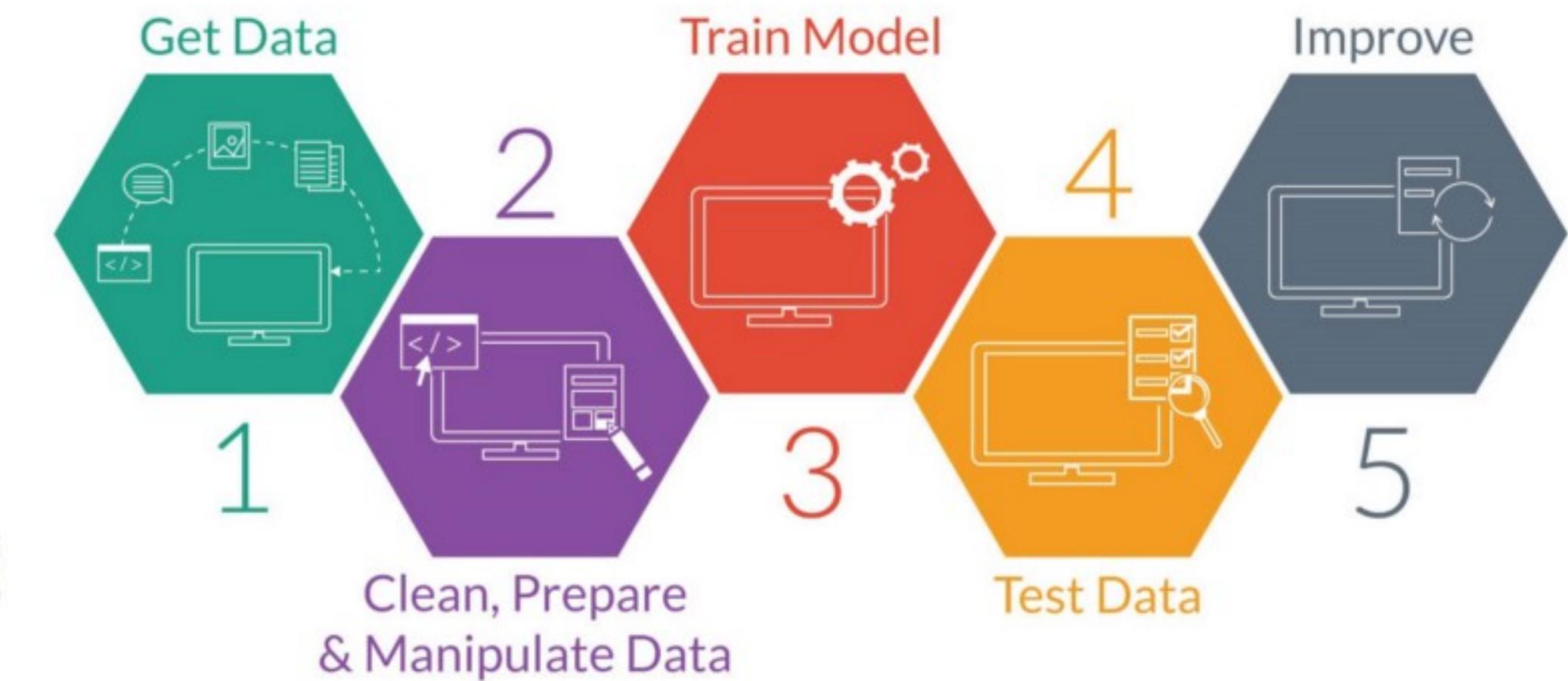
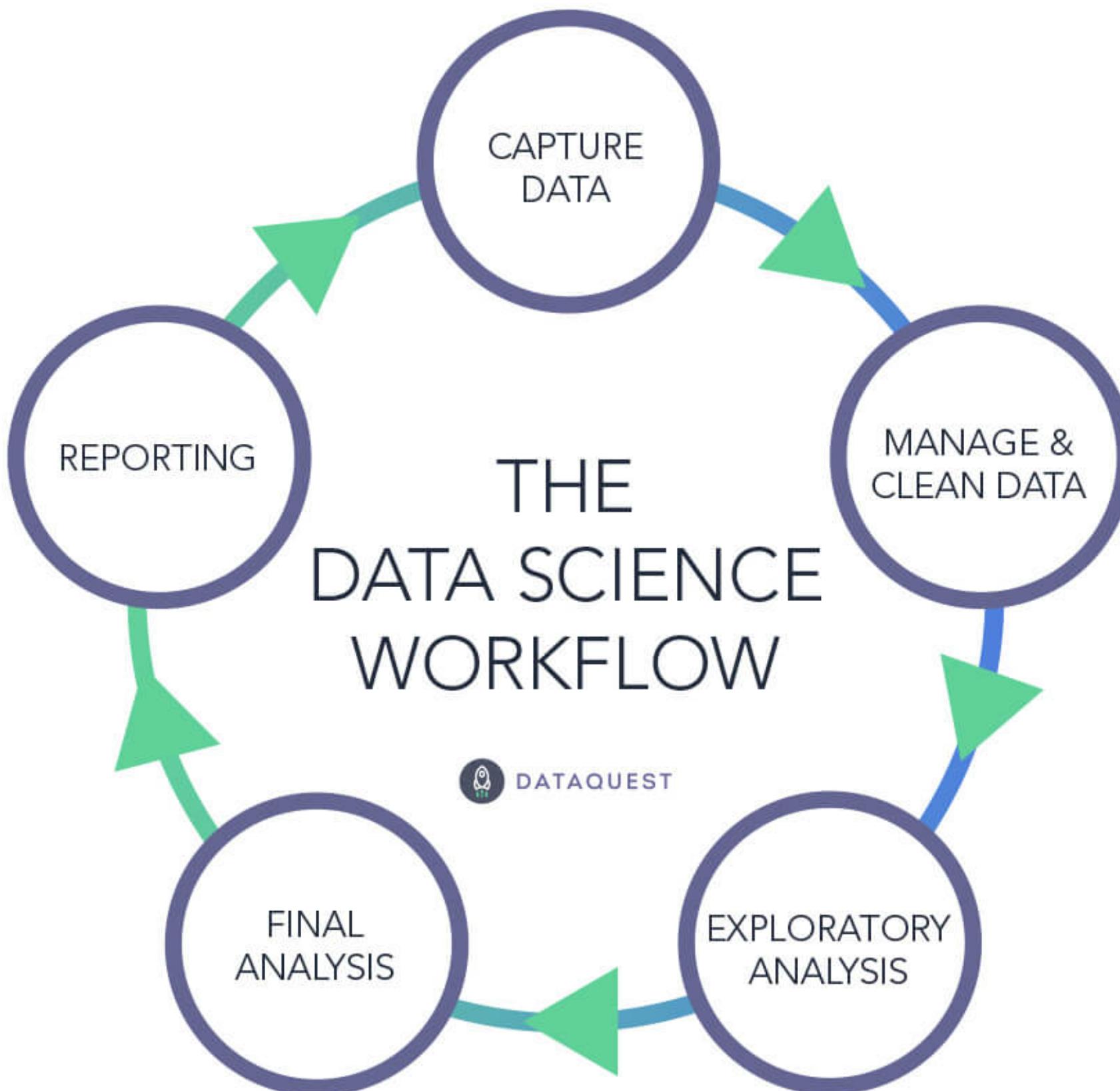
DataCamp

kaggle



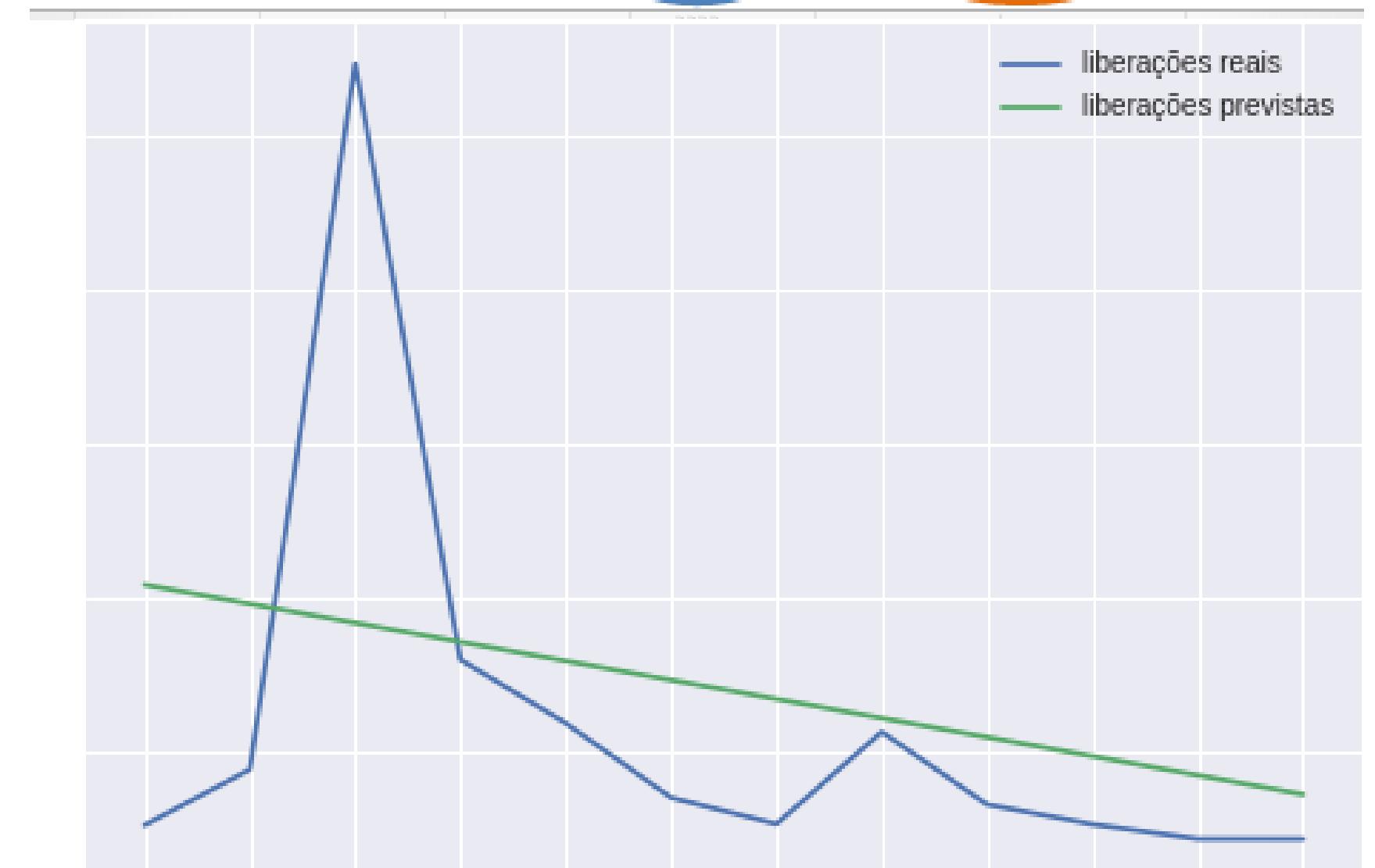
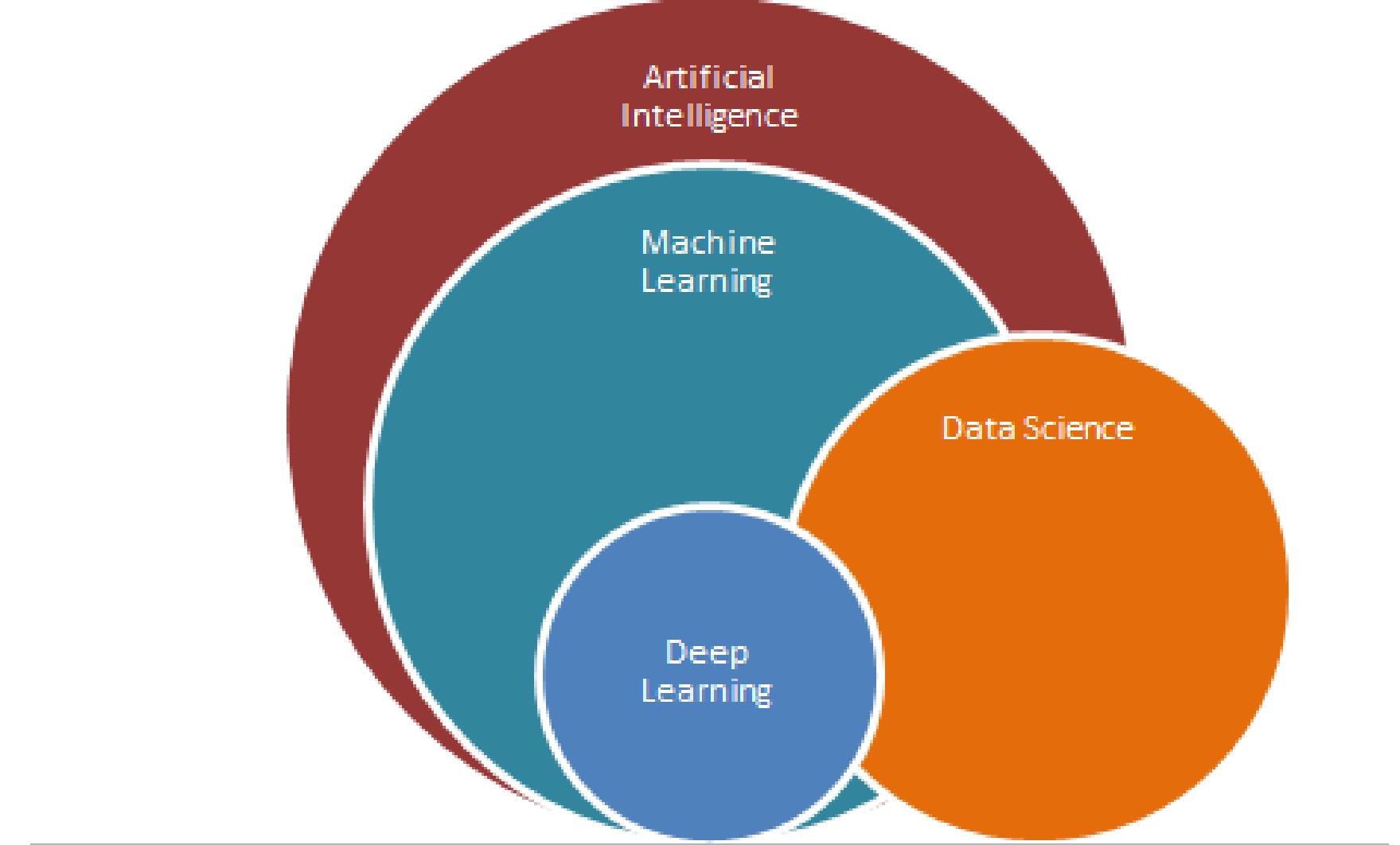
DATAQUEST

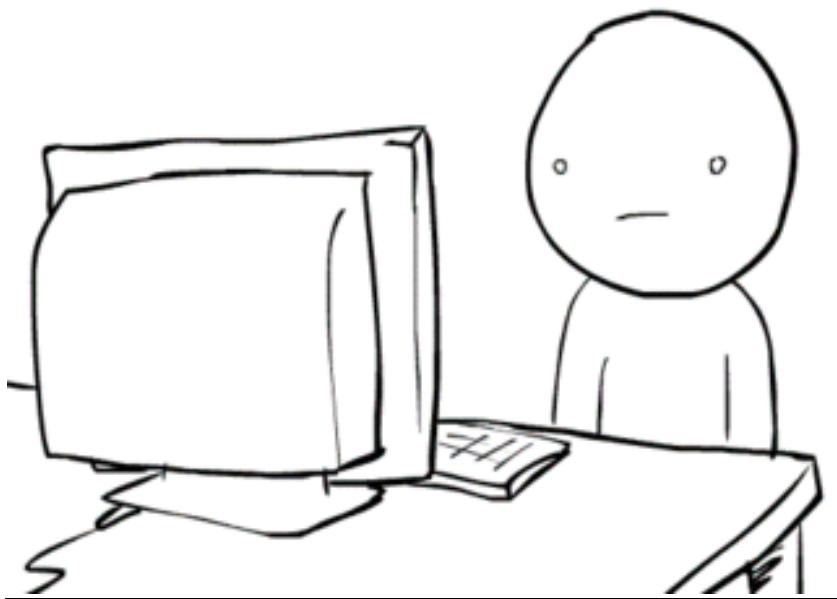
Ciência de dados (expectativa)



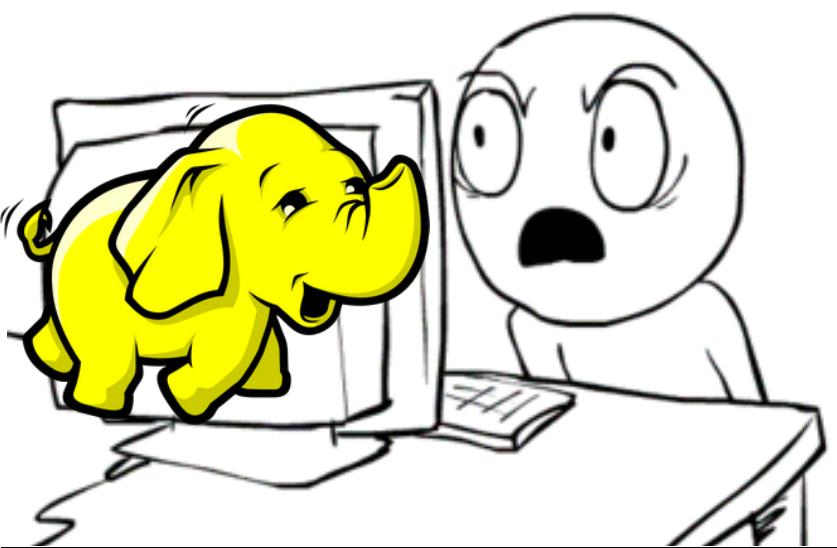
Machine Learning

não é bala de prata





**WHERE THE
F*CK IS THE
DATA?**



**HOW THE
F*CK I USE
THIS DATA?**



Os dados te odeiam

Os dados nunca, NUNCA estarão na forma correta para o seu problema e sempre trarão problemas adicionais para a solução



ENTENDA O PROBLEMA

O problema talvez já está sendo resolvido de outra forma, cabe a você mostrar a melhoria

ALINHE AS EXPECTATIVAS DO NEGÓCIO

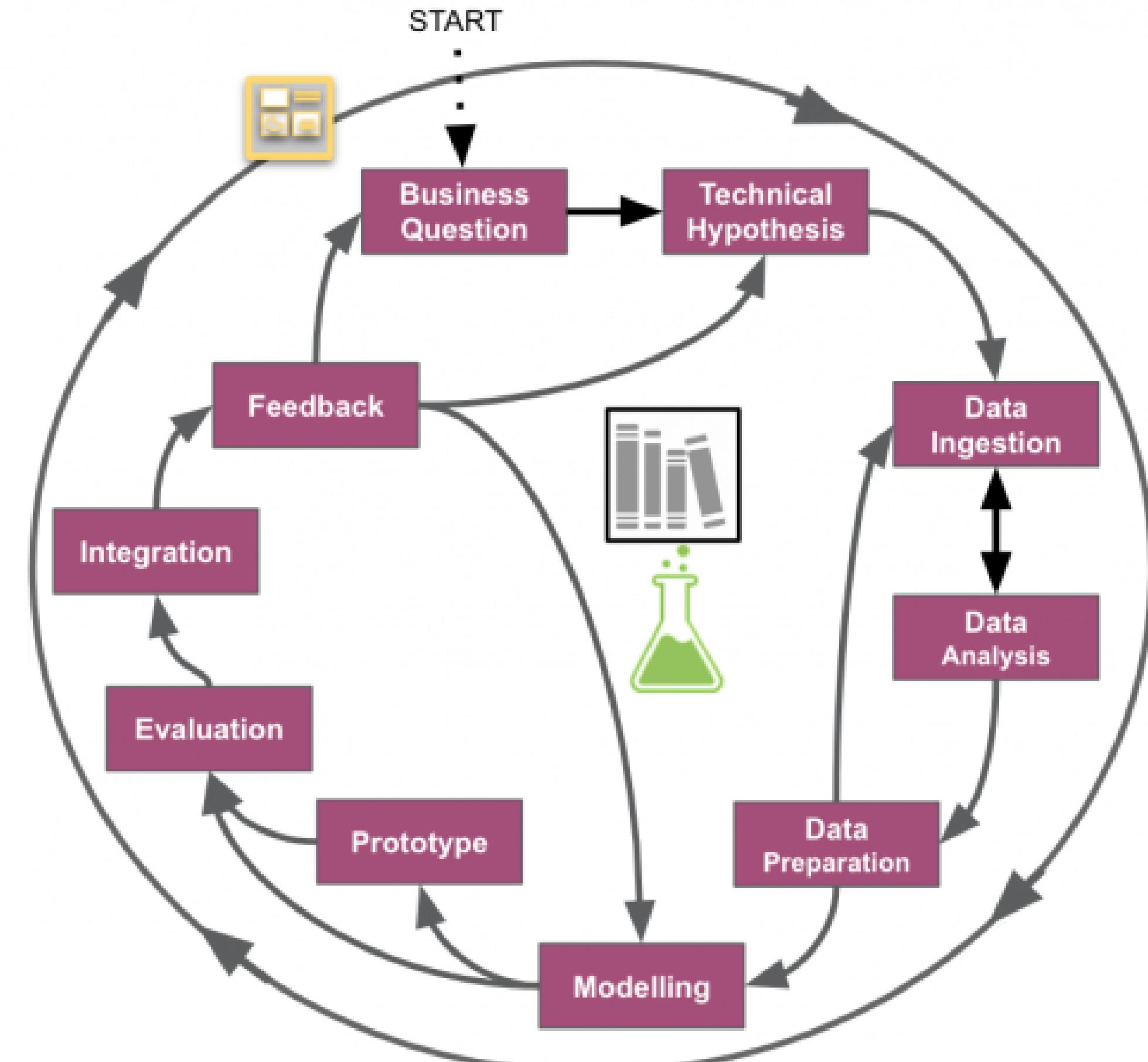
Data Science não é mágica, é baseada em experimentos

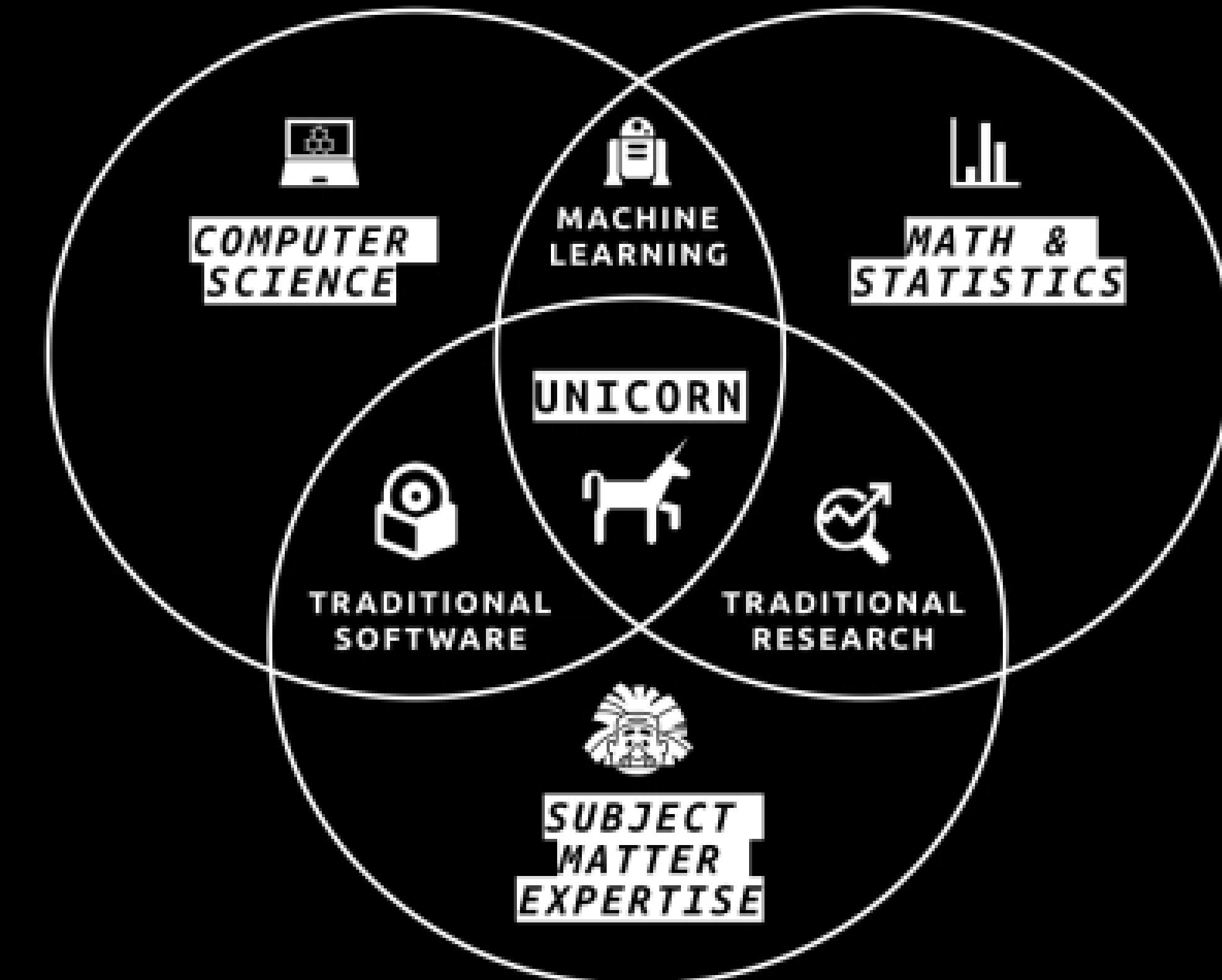
ENTENDA COMO GERAR AÇÕES BASEADAS EM DADOS

Devemos entregar soluções completas

Tudo é uma questão de comunicação com o negócio

Ciência de dados (ciclo completo)





Time de
dados de
acordo com
a hype



Analytics is a team sport

Analytic teams will always have a variety of roles, skills, favorite tools, and titles.

**THE DATAOPS MANIFESTO -
PRINCIPLE N° 4**

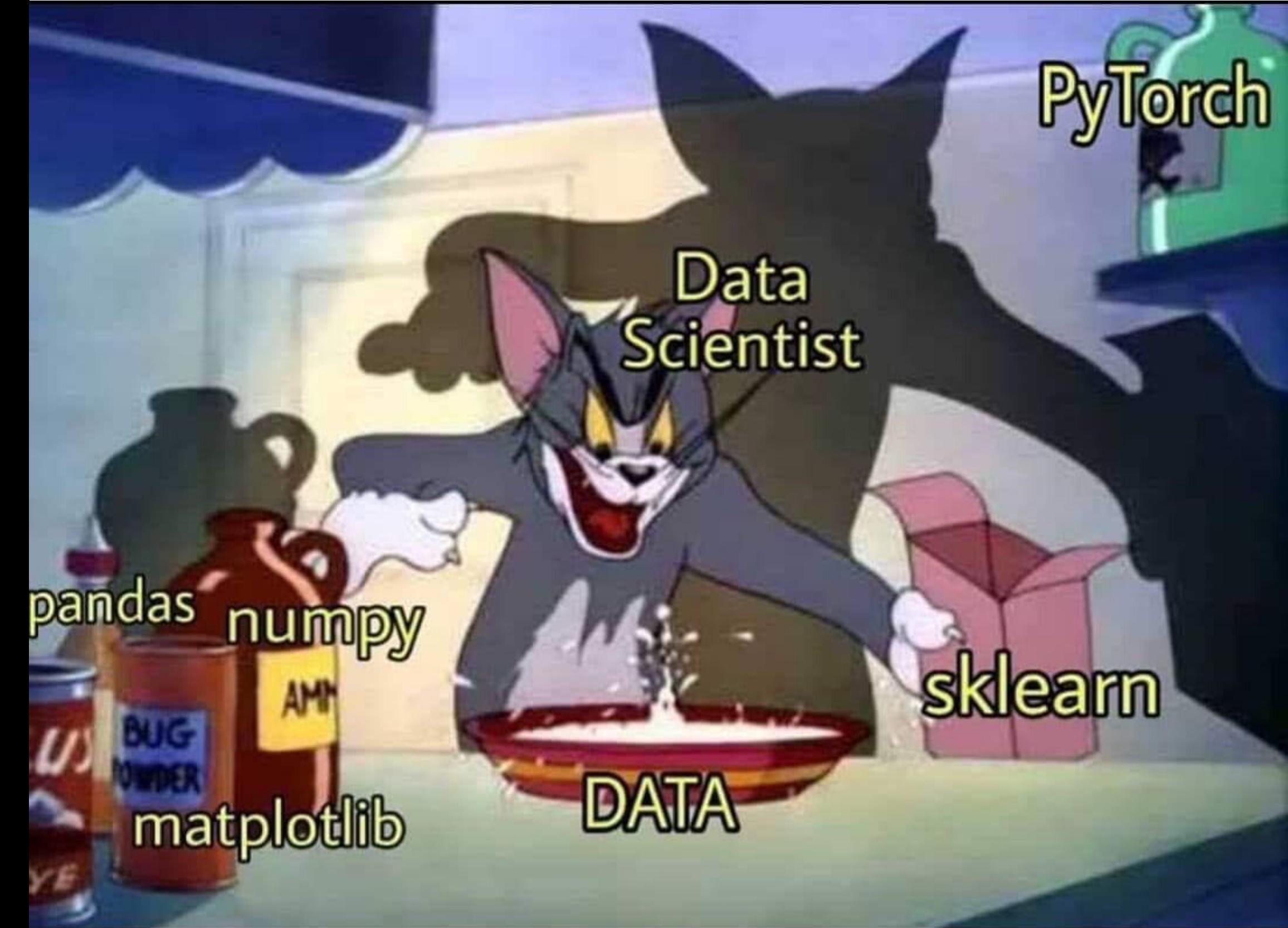


O time de dados funcional



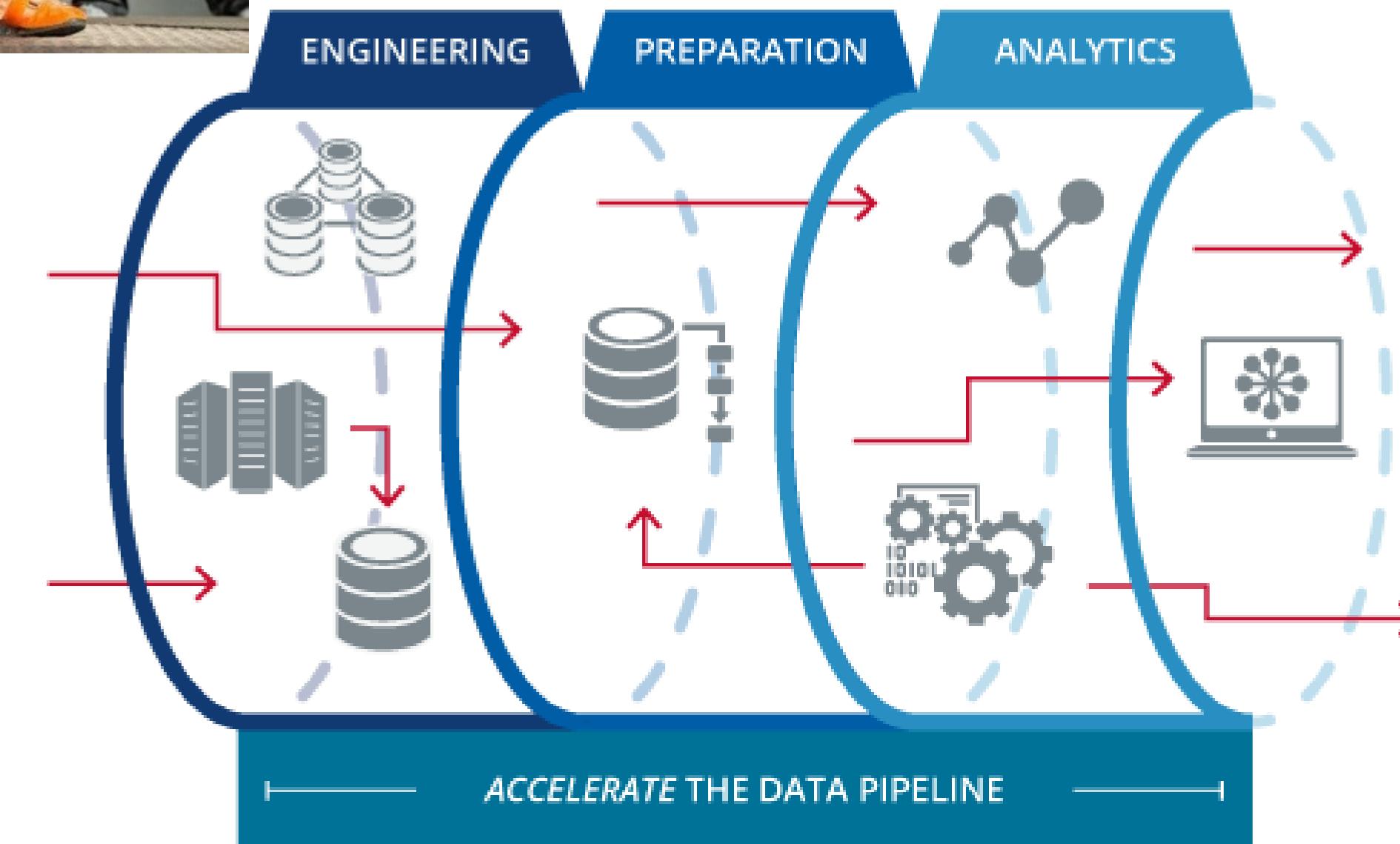
Cientista de dados

- FAZ EXPERIMENTOS COM OS DADOS
- POSSUI O DOM DE EXPLICAR CONCEITOS, MESMO QUE MATEMÁTICOS
- POSSUI DOMÍNIO DO MÉTODO CIENTÍFICO PARA VALIDAÇÃO DE HIPÓTESES
- PERSEVERANTE, CURIOSO, RESILIENTE

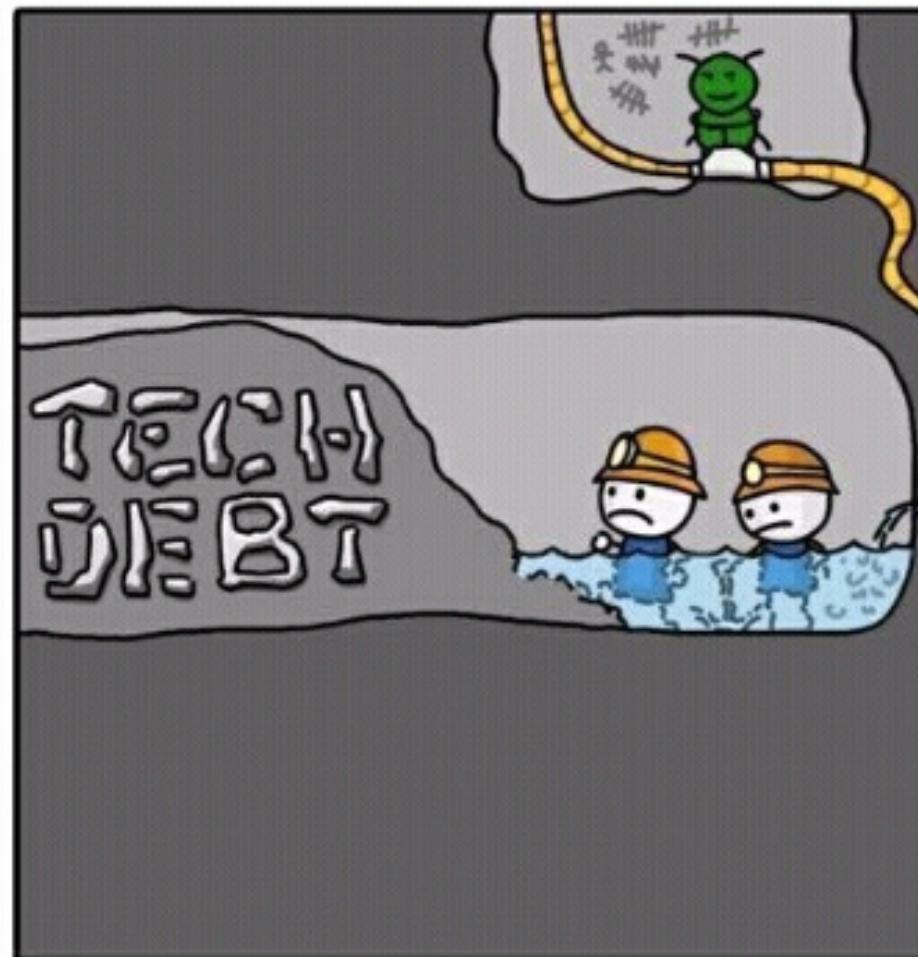
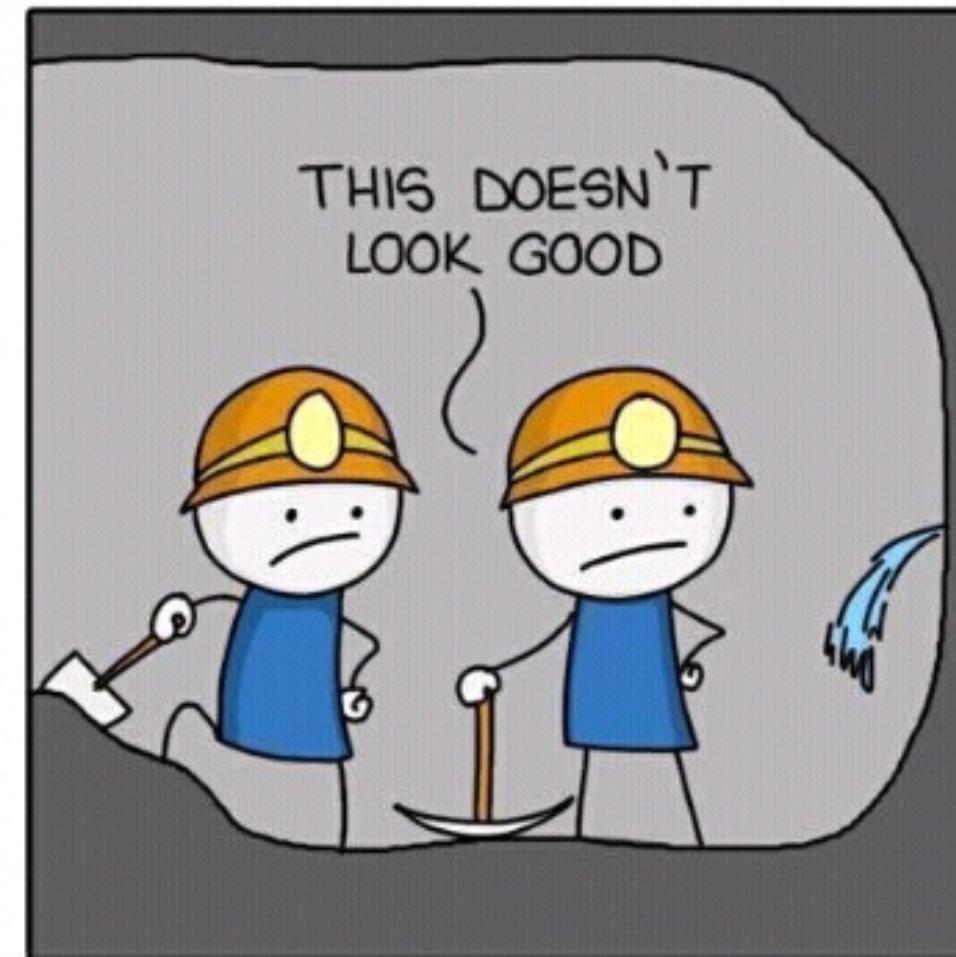
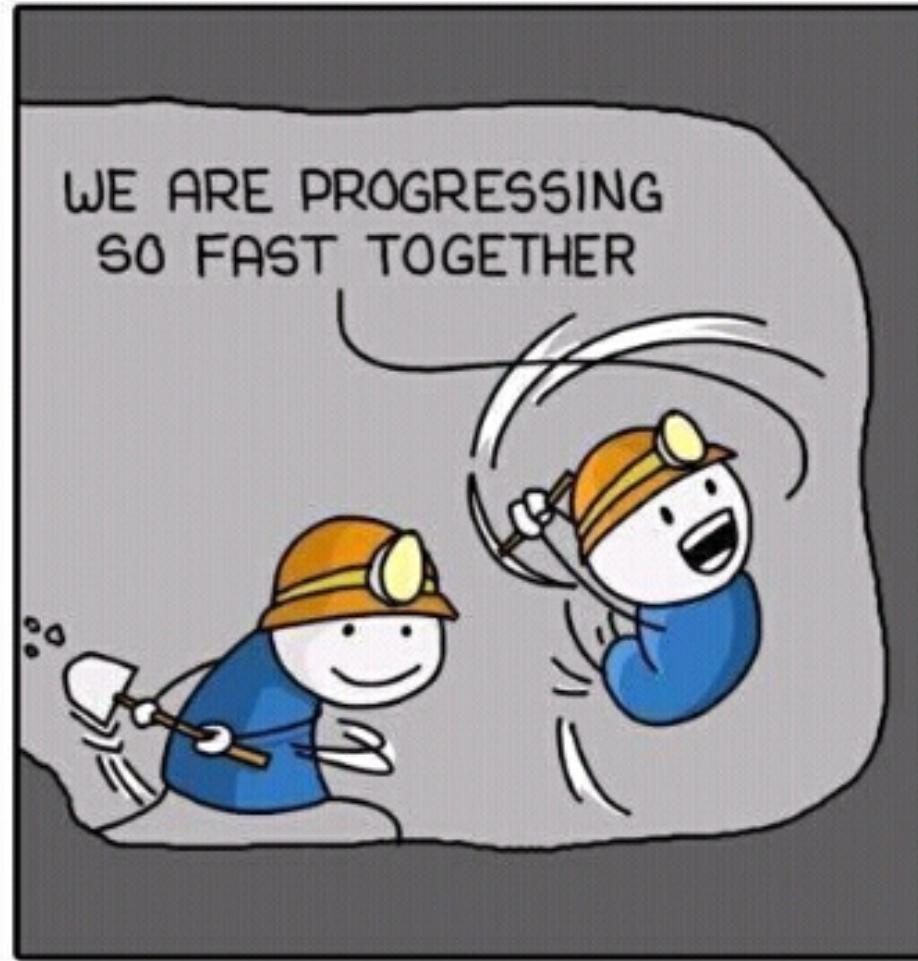
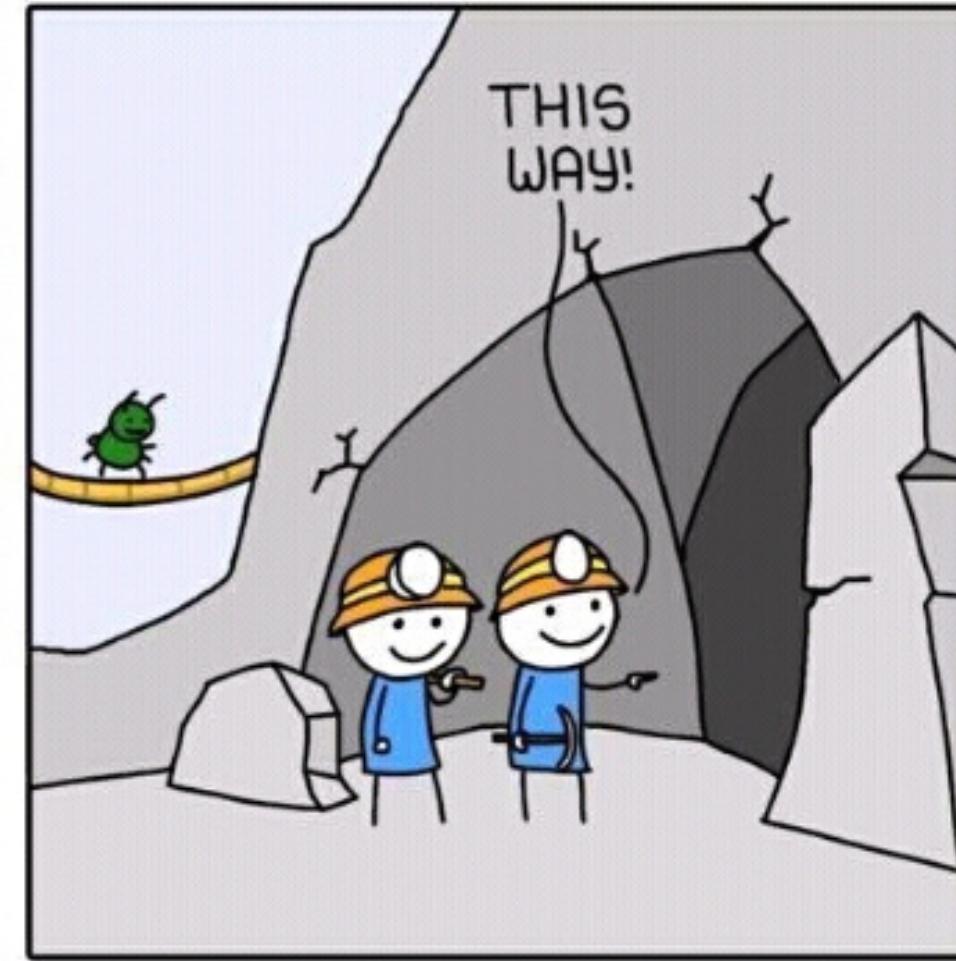


Engenheiro de dados

- CONSTRÓI AS "REFINARIAS DE DADOS"
- PROVÊ INFRA EM AMBIENTES ON PREMISES
- RESOLVE O PROBLEMA DO BIG DATA
- CRIA ARQUITETURAS ESCALÁVEIS DE DADOS E MACHINE LEARNING
- CRIA AS ESTRUTURAS DE DADOS ADEQUADAS ÀS NECESSIDADES
- CONSTRUTOR, DESAFIOS SÓLIDOS, PRECAVIDO



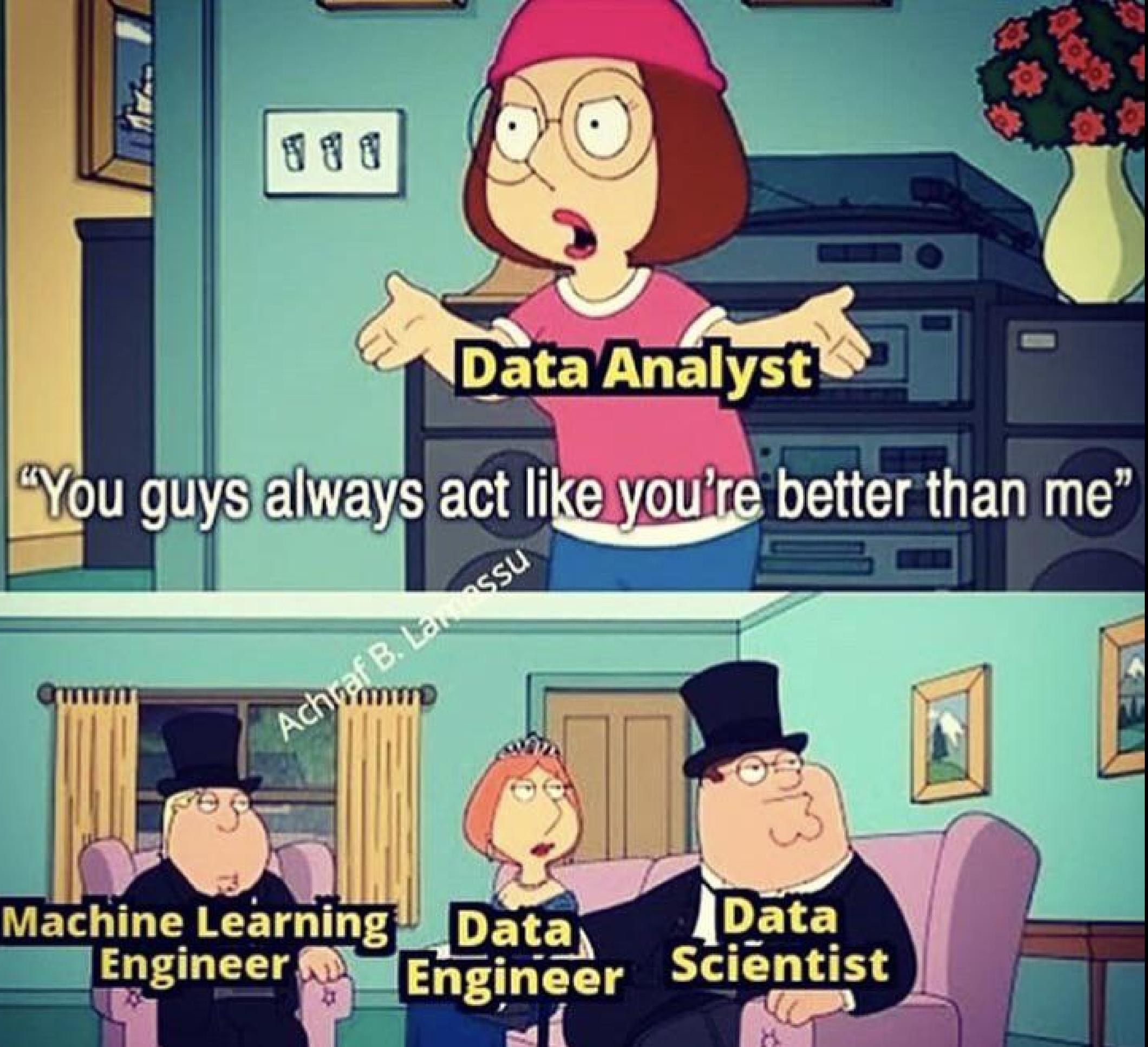
TECH DEBT



Arquiteto de dados

- Pensa nos dados de modo corporativo
- Pensa e desenvolve práticas e padrões que permitem o crescimento saudável da área de dados
- Lida com as normas da empresa em relação aos dados
- Mapeia os dados aos processos corporativos
- Orquestrador, diplomata, organizador

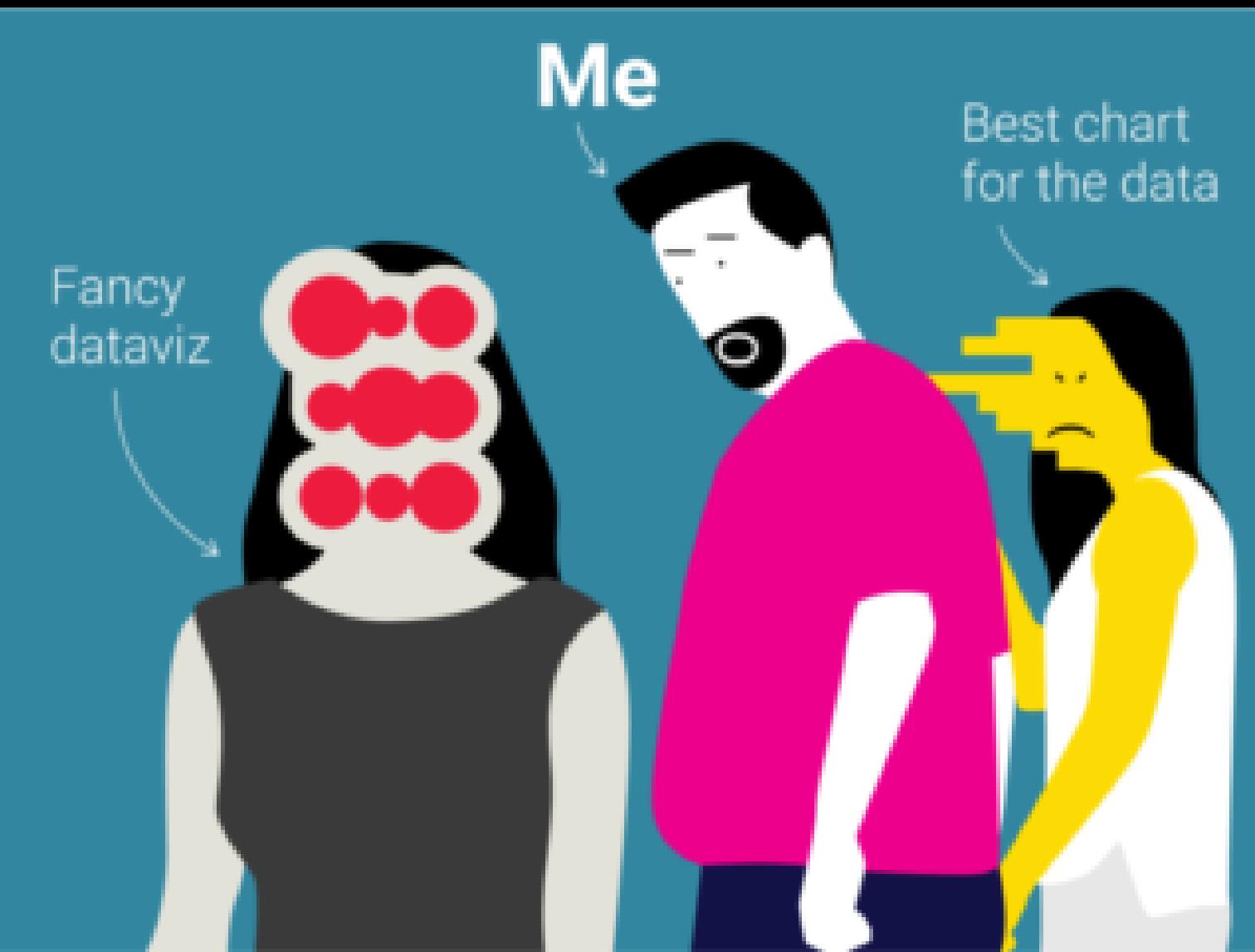
Analista de dados



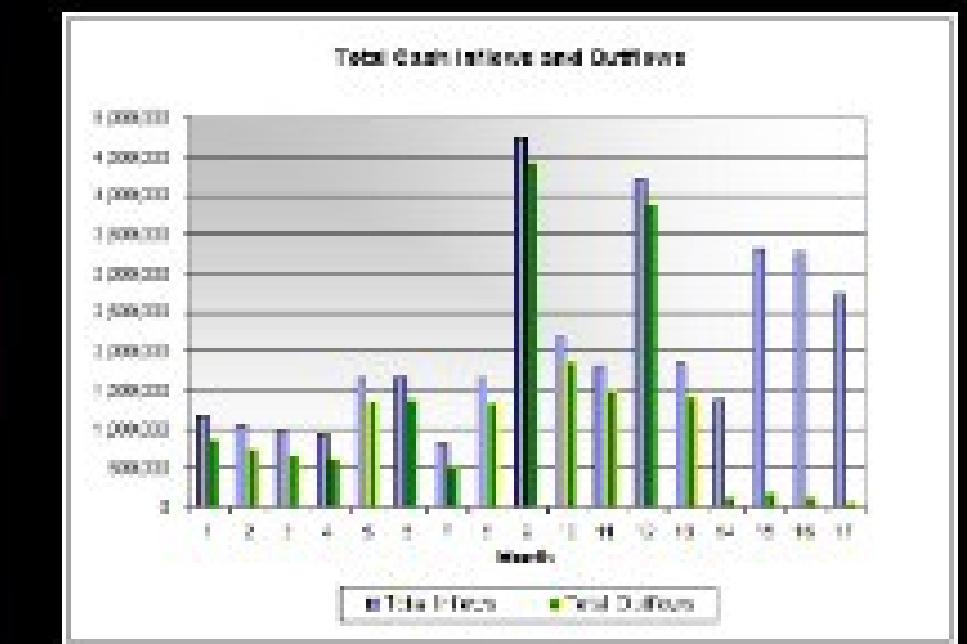
- Trabalha geralmente junto à área de negócio
- Possui o maior conhecimento do negócio
- Possui a capacidade de se tornar o dono da iniciativa de dados em sua área
- Conduz projetos de dados como Product Owner
- Intra-empreendedor, criativo

Visualização de dados

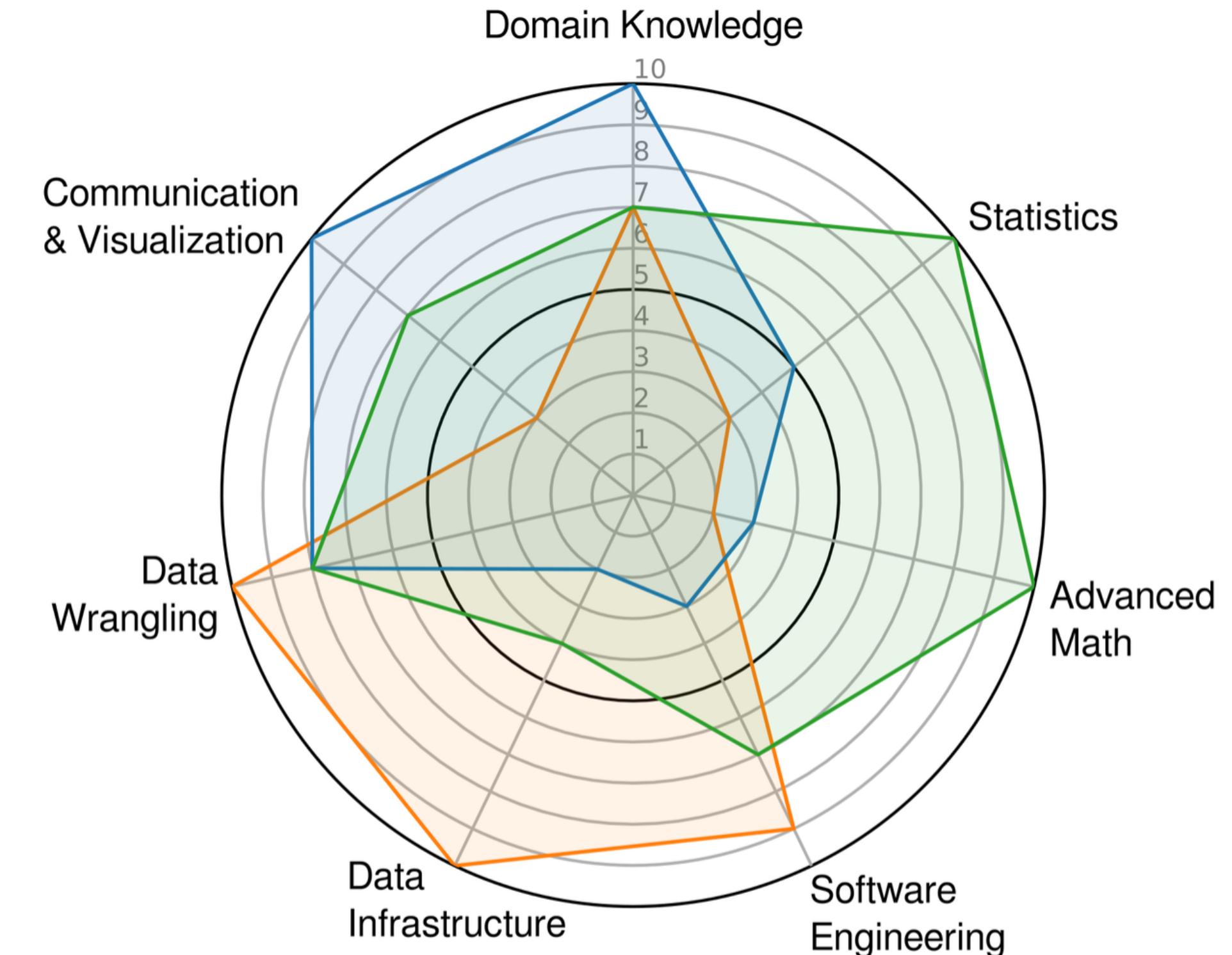
- Transforma os dados em visualizações que sintetizam terabytes de dados
- Criação de ferramentas que permitem ao usuário manipular e descobrir a informação
- Criativo, visual, abstratista, artístico



ID	NAME	CLASS	MARK	SEX
1	John Doh	Four	75	Female
2	Max Ruin	Three	85	male
3	Amelia	Three	95	male
4	Krish Shan	Four	80	Female
5	John Mike	Four	60	Female
6	Alex John	Four	85	male
7	My John Rob	Five	78	male
8	Aarid	Five	85	male
9	Tee Day	Six	78	male
10	Big John	Four	55	Female



Resultado de um time funcional



■ Data Analyst
 $\mu=5.86, \sigma=3.44$

■ Data Engineer
 $\mu=6.29, \sigma=3.55$

■ Data Scientist
 $\mu=7.57, \sigma=2.07$

Ainda na dúvida?

Participe da comunidade!



DATA HACKERS

Brasília DataGroup

Brasília, Brasil

953 membros · Grupo público

Organizado por Raul Diego de Queiroz Oliveira e mais 2 pessoas



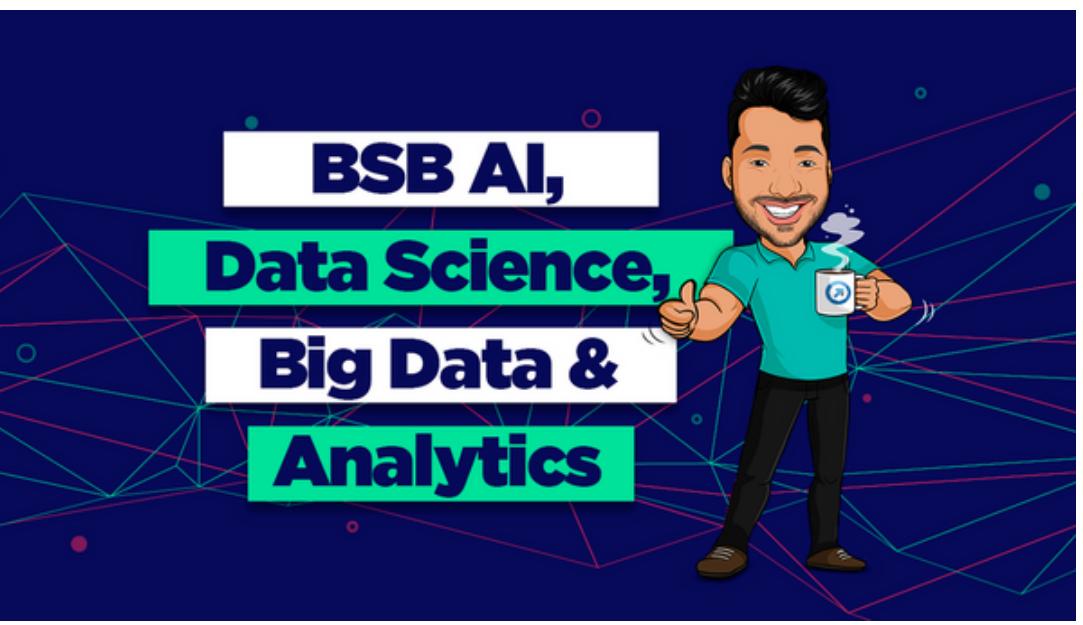
@datatips



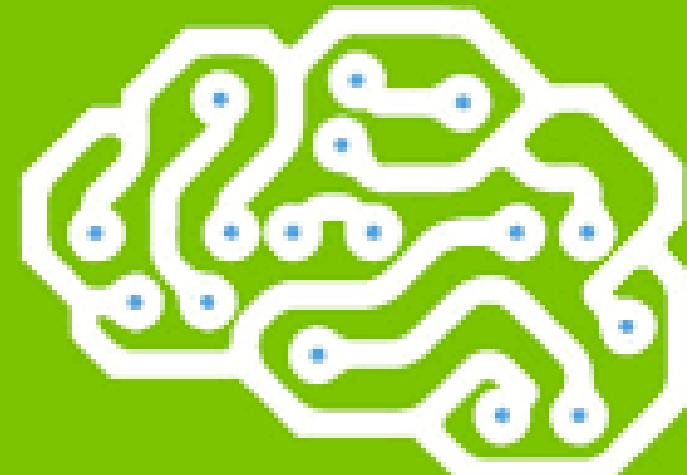
@engenhariadedados



@entendendobi



Machine
Learning
Brasília



Obrigado!

Dúvidas?

LEANDRO HUMBERTO
+55 (62) 9 8193 6680



USE AI
useai.com.br