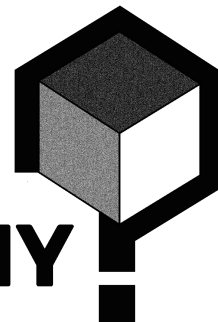


# Data Science & Points

Um projeto de Data Science do início ao fim

instituto  
AARON  
SWARTZ

**TÉO  
ME WHY!**



03/06/2024

# Trajetória

2016



2017



2018



2019



2020



2021



2022



2023



2024



# Combinados

- Sempre das **9AM às 11:20 horas** estaremos online
- Pausa às **10:00hrs** até as 10:15
- Notificações serão **desligadas**, mas sua contribuição é sempre bem-vinda
- Quando houver ADs, a explicação é **pausada**
- Evite **spamar** a mesma coisa no chat, moderadores estarão de olho
- Comentários discriminatórios e babacas serão **banidos** sem aviso prévio
- **!ppt** para acessar todo material o curso





# Avisos

## Link do repositório:

[github.com/TeoMeWhy/ds-points](https://github.com/TeoMeWhy/ds-points)

## Link dos dados:

[drive.google.com/drive/folders/1JLzofrtaVQdo0PdUysNWjNsBdAaI21EJ?usp=sharing](https://drive.google.com/drive/folders/1JLzofrtaVQdo0PdUysNWjNsBdAaI21EJ?usp=sharing)

## No último dia (07/06) de curso teremos sorteio de:

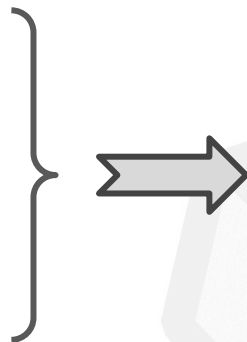
- Livros
- Cursos
- Adesivos





# Avisos

- Nosso foco é o básico bem feito;
- Estamos melhorando ano após ano!
- Tem algo que não abordei?
- Tem algo que você explicaria diferente?
- Tem algo mais complexo que deixei passar?



Faça um  
vídeo/material





# Nos Apoie

## Nossos cursos são 100% gratuitos

e você pode nos apoiar em diferentes plataformas:

- Apoia.se: [apoia.se/teomewhy](https://apoia.se/teomewhy)
- LivePix: [livepix.gg/teomewhy](https://livepix.gg/teomewhy)
- Assinante na Twitch: [twitch.tv/teomewhy](https://twitch.tv/teomewhy)
- Membros no YouTube: [youtube.com/@teomewhy](https://youtube.com/@teomewhy)
- GitHub Sponsors: [github.com/sponsors/TeoMeWhy](https://github.com/sponsors/TeoMeWhy)





# Ementa (1/2)

## **Qual o nosso desafio?**

- Definição do problema de negócio
- CRISP-DM

## **Conhecendo o banco de dados**

- Quais são as tabelas
- O que elas representam
- Quais variáveis podem nos ajudar neste desafio?

## **Feature Store**

- Definição das feature store por contexto
- Criação das Safras

## **Definição da Variável Resposta**

- Geração da ABT (Analytics Base Table)







# Ementa (2/2)

## **Treinar algoritmos**

- SEMMA
- Tuning de modelos

## **Deploy**

- Usando modelo em nosso Bot

## **Tópicos adicionais**

- MLFlow
- Orquestração com Mage







“Big  
stories need  
**human**  
stakes”

Aaron Swartz



@teomewhy



@teomewhy



/teocalvo



/teomewhy



instituto  
AARON  
SWARTZ



