



Manipulação de datasets

Primeiro projeto prático

Sobre a base de dados



Palestras do TED Talks

Este conjunto de dados foi criado para **estudantes iniciantes** de Análise de Dados

O dataset apresenta 6 informações a respeito de cada palestra

- = **title**- Título da palestra
- = **author** - Autor da Palestra
- = **date** - Data em que a palestra ocorreu
- = **views**- Número de visualizações do Talk
- = **likes**- Número de curtidas do Talk
- = **link** - Link da palestra no ted.com

Queremos responder as seguintes perguntas:



- = Baseado na quantidade de likes, qual o top 10 palestras do TED Talks?
- = Quais os 5 autores com mais vídeos no TED Talks?
- = Quantas palestras ocorreram em cada ano?
- = Quantas palestras ocorreram em cada mês do ano de 2020?
- = Quantas palestras relacionadas à tag "education" existem no dataset?



Regras do jogo:

A cada dia uma dica é liberada

Pode pedir ajuda pros coleguinhas do grupo

Se você terminou rapidinho, não poste o código no grupo





Primeira dica:

Essas serão suas amigas:

https://pandas.pydata.org/docs/user_guide/index.html#user-guide

<https://numpy.org/doc/stable/user/index.html#user>