

# Introdução ao Pandas



# Ementa

- Apresentação
- Biblioteca Pandas
- ETL x ELT
- Buscando novos Conhecimentos
- Codificar





# Lucas Ribeiro

Data Analyst

Grupo Zap

+5 anos em análise de dados

@lucasmbribeiro



---

# Introdução a Biblioteca Pandas





**Pandas é uma biblioteca de código aberto para análise de dados com base na linguagem Python.**

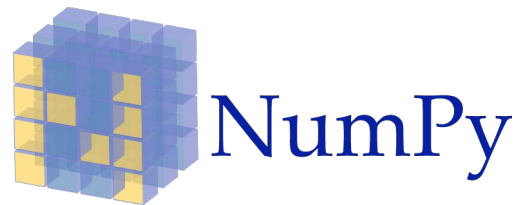
É a ferramenta perfeita para análises exploratórias, pois foi desenvolvida para **leitura, manipulação, agregação e visualização** de dados

# Por que pandas é tão popular?

## ***Vantagens de usar Pandas:***

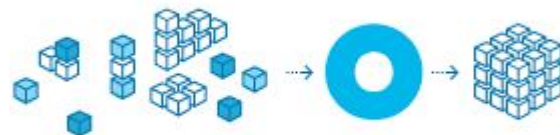
- *Biblioteca de fácil entendimento e uso.*
- *Torna fácil a manipulação de dados.*
- *A porta de entrada para o mundo de data science.*
- *Código semelhante ao SQL, isso trás velocidade para aprender.*
- *Uma das melhores bibliotecas para análise de dados <3 \*Opinião Particular rs...*

seaborn





# ETL x ELT



## Outdated ETL

O processo de ETL é um dos principais se não o mais importante passo no mundo de ciência de dados.

## Modern ELT

**E** Extração  
**T** Transformação  
**L** Carga

## Extração → Transformação → Carga

Esse procedimento é indicado para soluções de Business Intelligence. Onde os dados precisam ser tratados antes de serem carregados.

Soluções: Data Warehouse e seus derivados.

## Extração → Carga → Transformação

Usamos esse novo modelo quando temos um Data Lake. Na fonte do dado, simplesmente carregamos para lake e quando necessário transformamos.

Soluções: Soluções de dados de forma geral, também serve para BI.



# Buscando Novos Conhecimentos

→ ***Introdução a Pandas e Python***

[10 Minutos com Pandas](#)

[Avançando no Pandas](#)

[Configurando seu Ambiente Python](#)

→ ***Leituras de Data***

<https://medium.com/data-hackers>

<https://towardsdatascience.com/>

---

# Vamos Colocar a Mão na Massa

[Link do Repositório](#)

[Link do Arquivo](#)





# Muito Obrigado :)

