



# Linguagem de Programação

Professor Anderson I. S. Abreu

# Introdução a biblioteca Pandas



Tópicos da nossa aula:

1. Introdução a biblioteca Pandas;
2. Séries;
3. Leitura de dados estruturados com a biblioteca Pandas

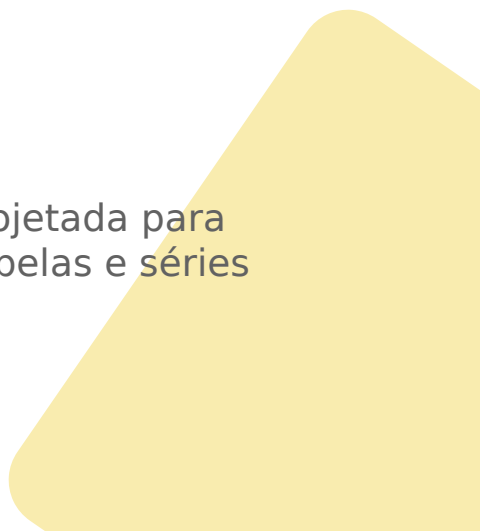


1

# Introdução a biblioteca Pandas

---

Uma biblioteca Python de alto desempenho e código aberto projetada para simplificar a manipulação e análise de dados organizados em tabelas e séries temporais.



# O famoso Pandas I

**DataFrames e Series:** estruturas de dados flexíveis, semelhante a tabela(DF) e lista.

**Manipulação de Dados:** filtragem, seleção, ordenação, agrupamento e agregação.

**Leitura e Escrita de Dados:** CSV, Excel, SQL, HDF5 e muitos outros, tornando-o uma ferramenta versátil para lidar com dados de diferentes fontes.



# O famoso Pandas II

**Tratamento de Dados Ausentes:** simplifica o tratamento de dados faltantes.

**Visualização de Dados:** integrado com bibliotecas de visualização, como Matplotlib e Seaborn.

**Integração com NumPy:** construído sobre a biblioteca NumPy, combinação de cálculos e manipulação.

**Comunidade Ativa:** tem uma comunidade de usuários e desenvolvedores ativa.



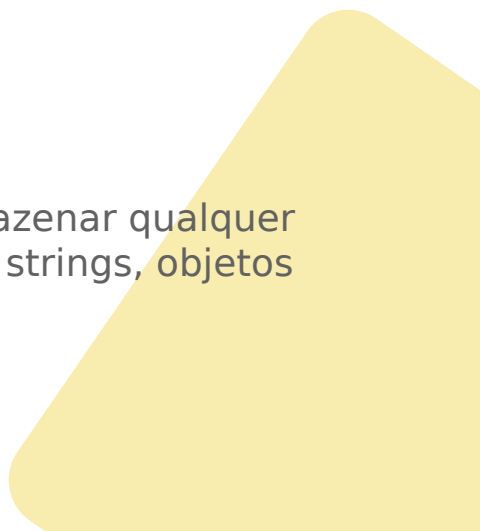


2

# Séries

---

Refere-se a uma estrutura unidimensional de dados que pode armazenar qualquer tipo de dado, como números inteiros, números de ponto flutuante, strings, objetos Python.

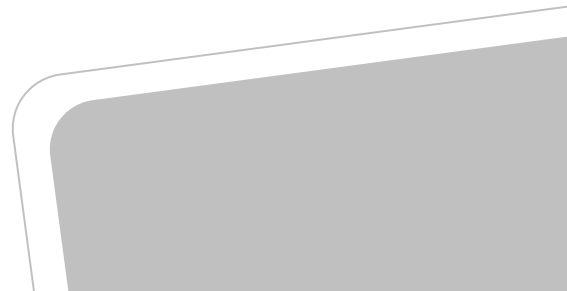


# Criando uma Série

O principal parâmetro é "data", que pode conter um único valor, uma lista de valores ou um dicionário.

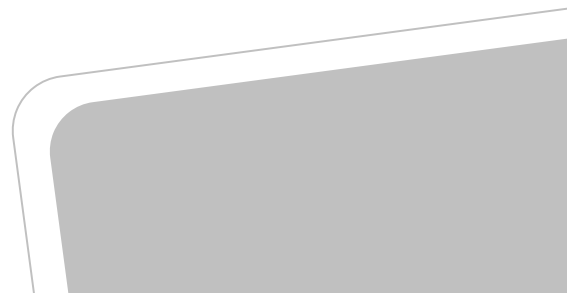
Os outros parâmetros, como "index", "dtype", "name" e outros, têm valores padrão predefinidos, tornando sua especificação opcional.

Vamos criar uma Série a partir de uma lista.



# Criando uma Série a partir de um dicionário

Vamos criar uma Série a partir de uma dicionário.







3

# Leitura de dados estruturados com a biblioteca Pandas

---

A biblioteca oferece vários métodos de leitura de dados, identificados pelo padrão "read", como `pandas.read_XXXXX`



# Ler é importante

Um recurso poderoso no pandas é a capacidade de ler dados estruturados e armazená-los em um DataFrame.

`pandas.read_XXXXX` ou `pd.read_xxxx`

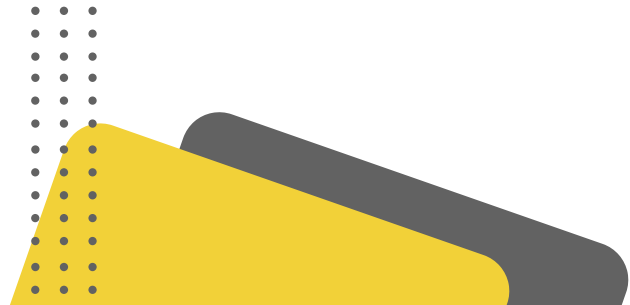
Cada um desses métodos é projetado para ler diferentes tipos de fontes de dados.



# Vamos de exemplo

Para exemplificar, vamos explorar o método `pandas.read_html()`, que é utilizado para extrair tabelas de uma página da web.

Na URL <https://www.fdic.gov/bank/individual/failed/banklist.html>, encontra-se uma tabela com bancos norte-americanos que faliram desde 1º de outubro de 2000, cada linha representa um banco.





# Aplicando a aula!

---

Suponha que estamos gerenciando o cadastro de uma loja, e essa loja, precisa de uma orientação em qual público deve investir, sendo assim, querem saber a idade média dos clientes.

# Obrigado!

