

Introdução ao Pandas



CRONOGRAMA DA AULA

1

O que é Pandas?

2

Origem e
Importância

3

Primeiros Passos
com Pandas

4

Series: Estrutura
Unidimensional

5

Operações com
Series

6

DataFrames:
Estrutura
Bidimensional

7

Trabalhando com
DataFrames

8

Pandas na Análise
Crítica e Estratégica

CRONOGRAMA DA AULA

Biblioteca de código aberto



Desenvolvida para a linguagem Python com foco em análise de dados.

Estruturas de dados flexíveis



Oferece ferramentas para trabalhar com dados estruturados ou tabulares.

Análise eficiente



Ideal para manipular, processar e limpar grandes conjuntos de dados.

ORIGEM E IMPORTÂNCIA

Nome e Origem

O nome "pandas" deriva de "panel data", um termo da econometria.

Foi inicialmente construído para resolver problemas financeiros e de análise de negócios.

Comparação com NumPy

NumPy é adequado para dados numéricos homogêneos em arrays.

Pandas foi projetado para trabalhar com dados tabulares e heterogêneos.

Ecossistema Python

Frequentemente usado com NumPy, SciPy, matplotlib e scikit-learn.

Fundamental para preparação de dados, que ocupa 80% do tempo de análise.

PRIMEIROS PASSOS COM PANDAS

Instalação

Instale o Pandas usando pip ou através de distribuições como Anaconda.

Importação

Importe a biblioteca usando a convenção: `import pandas as pd`

Primeiras Operações

Experimente com exemplos simples para entender o fluxo de trabalho.



```
install pannow  
+ > pip install pandas  
  
pip install pandas (instah-PJ)
```



```
Install For panthow  
  
import pandas as pd
```


SERIES: ESTRUTURA UNIDIMENSIONAL

Definição

- Estrutura de dados unidimensional que armazena dados de qualquer tipo.
- Semelhante a um array com rótulos de índice.

Criação

- Criada passando uma lista ou array para o construtor Series.
- Exemplo: `pd.Series([1, 2, 3, 4])`

Indexação

- Permite seleção por posição ou pelos rótulos do índice.
- Facilita a manipulação eficiente de subconjuntos de dados.

Pandas Series

Index	Values
A	A
B	C,
B	30
.	4D

Index	Values
10	10
20	30
30	40
40	40

OPERAÇÕES COM SERIES



Operações Matemáticas

Series suportam adição, subtração, multiplicação e divisão.



Operações Lógicas

Permitem filtrar dados com condições booleanas.



Alinhamento Automático

Operações entre Series alinham dados com base nos índices.



Estilo Vetorizado

Operações sem laços for, herdadas do NumPy, tornam o código eficiente.

pandas Series

```
data = A
data = B

{
    data = 5
    data = B C D
    data + 2
}

*example: (lambda x*x*2)
data.apply(lambda x x*2)
datas.coply squared)
} !
```

DATAFRAMES: ESTRUTURA BIDIMENSIONAL

Estrutura Tabular

Semelhante a uma planilha ou tabela de banco de dados.

Colunas Heterogêneas

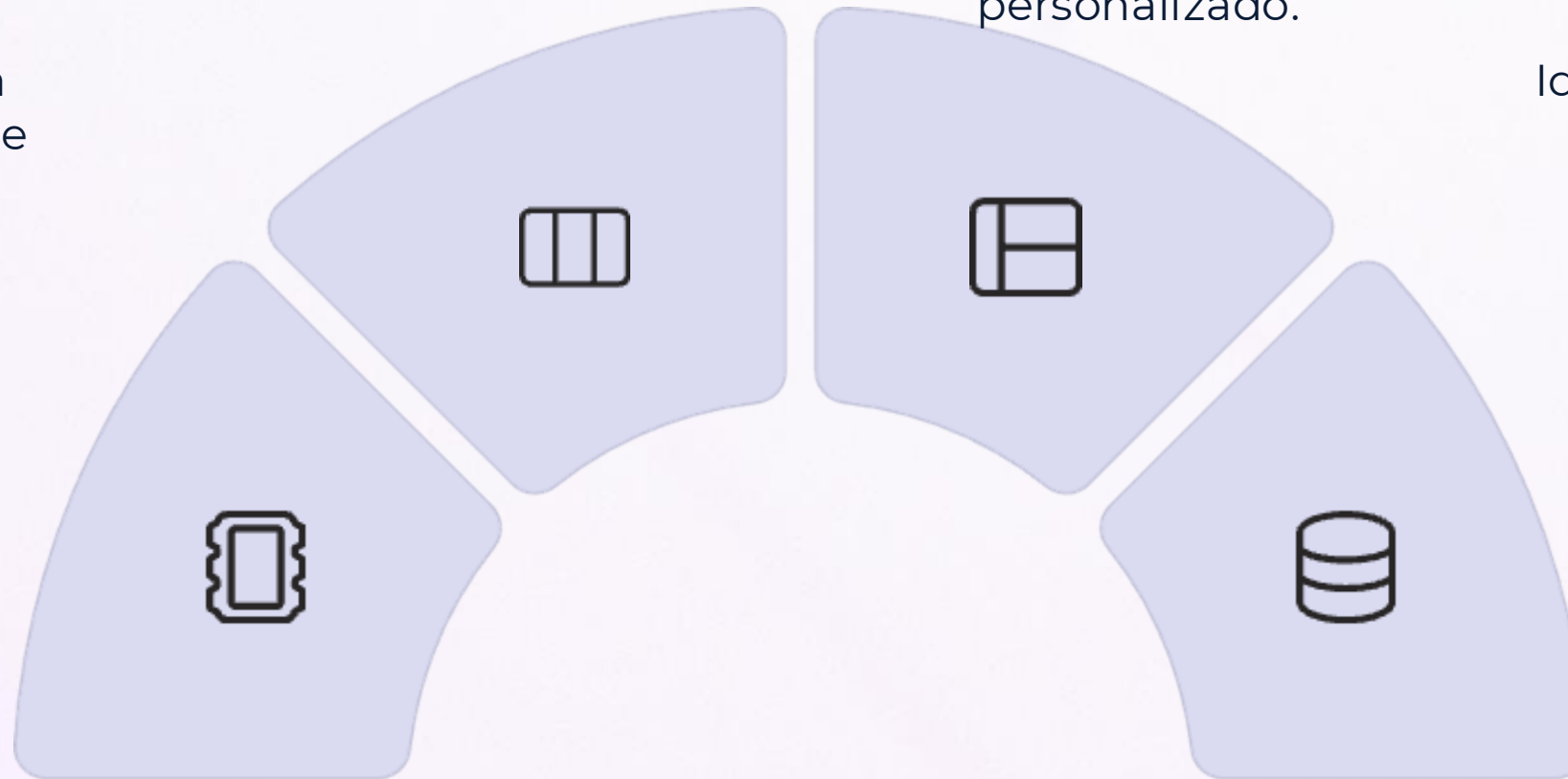
Cada coluna pode ter um tipo de dado diferente.

Linhas Indexadas

Linhas são identificadas por um índice personalizado.

Dados Tabulares

Ideal para trabalhar com conjuntos de dados estruturados.



TRABALHANDO COM DATAFRAMES

Carregando Dados

Use `pd.read_csv()` para importar dados de arquivos CSV.



Visualizando Dados

Use `.head()` e `.tail()` para inspecionar as primeiras e últimas linhas.



Informações Básicas

Use `.info()` para ver tipos de dados e `.shape` para dimensões.



Manipulação

Selecione, filtre e transforme dados com operações intuitivas.

PANDAS NA ANÁLISE CRÍTICA E ESTRATÉGICA





Vamos
avaliar o
encontro?