



PANDAS: EXPLORANDO O PODER DOS DADOS"

RAFAEL FREITAS

01

INTRODUÇÃO AO PANDAS

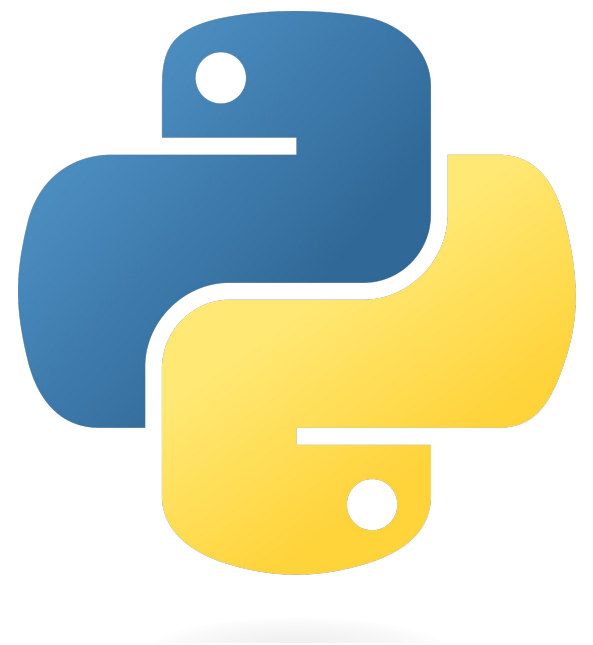
Pandas é uma biblioteca Python amplamente utilizada para manipulação e análise de dados.

Introdução ao Pandas

A biblioteca fornece estruturas de dados poderosas, como DataFrame e Series, que facilitam a manipulação de conjuntos de dados complexos de forma eficiente.



pandas



02

CARREGANDO E EXAMINANDO DADOS

Carregando e Examinando Dados

Ao trabalhar com dados, é crucial carregá-los em um formato adequado. O Pandas oferece o método `read_csv()` para carregar dados de um arquivo CSV em um `DataFrame`, permitindo uma fácil visualização e manipulação.

```
import pandas as pd
```

```
# Carregando dados de um arquivo CSV  
dados = pd.read_csv('dados.csv')
```

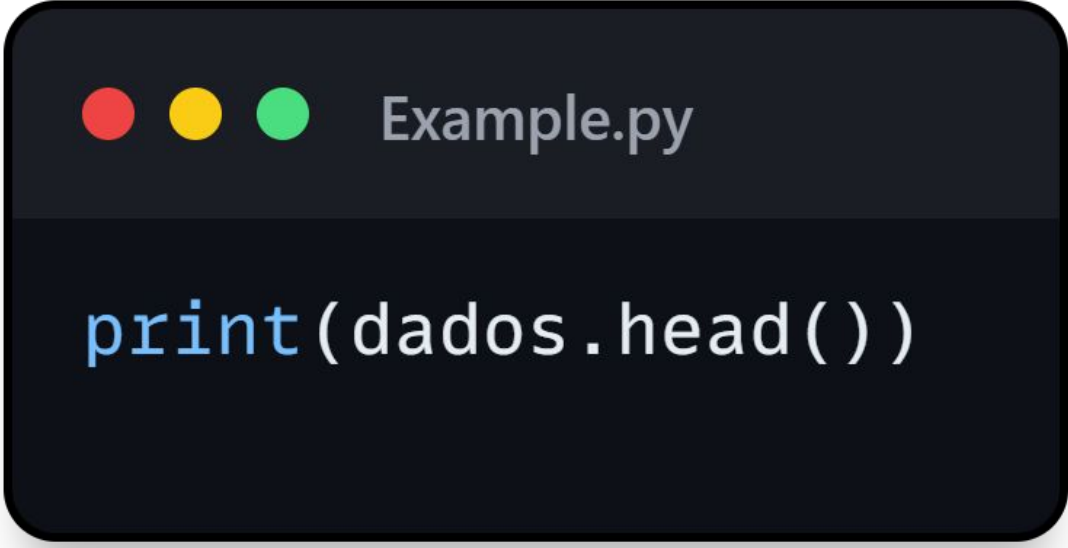
```
# Carregando dados de um arquivo Excel  
dados = pd.read_excel('dados.xlsx')
```

03

VISUALIZANDO OS DADOS

Primeiras Colunas

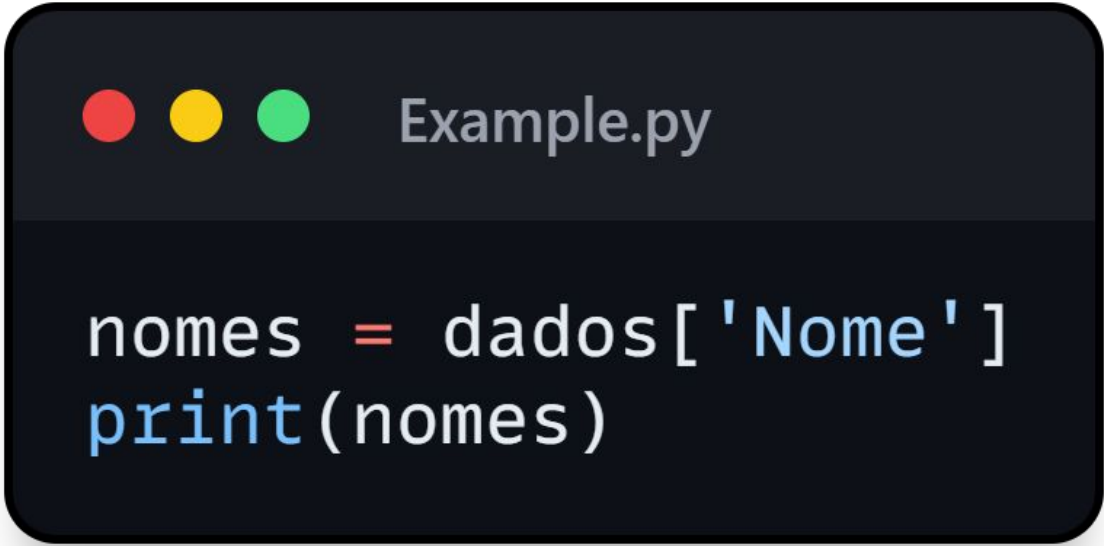
Para entender melhor seus dados, use `head()` para visualizar as primeiras linhas do DataFrame. Isso ajuda a ter uma visão geral dos dados e identificar possíveis problemas de formatação.



```
print(dados.head())
```

Selecionando Colunas

Com o Pandas, você pode acessar colunas específicas de um DataFrame usando o nome da coluna entre colchetes. Isso é útil para focar em dados específicos durante a análise.



```
nomes = dados['Nome']  
print(nomes)
```


Filtrando Dados

Filtrar dados com base em condições é uma tarefa comum na análise de dados. O Pandas permite filtrar dados de acordo com critérios específicos, como mostrando apenas registros de mulheres.

Example.py

```
mulheres = dados[dados['Gênero'] == 'Feminino']  
print(mulheres)
```

04

ADICIONANDO COLUNAS OU DADOS

Adicionando Colunas

Você pode criar novas colunas em um DataFrame com base em cálculos ou manipulações de dados existentes. Por exemplo, adicionar uma coluna de "Idade" com base na diferença do ano atual com o ano de nascimento.

```
Example.py

from datetime import datetime

# Obtendo o ano atual
ano_atual = datetime.now().year

# Calculando a idade e adicionando como uma nova coluna
dados['Idade'] = ano_atual - dados['Ano_Nascimento']
```

Adicionando Dados

Você pode adicionar novos dados em um DataFrame com base nas colunas existentes. Por exemplo, criar um dict com os valores de Nome e Idade. As colunas não mencionadas receberão o valor de NaN (*Not a Number*).

```
Example.py

# Criando um dict com os valores
nova_linha = {'Nome': 'Ana', 'Idade': 28}

# Adicionando uma nova linha ao DataFrame
dados = dados.append(nova_linha, ignore_index=True)
```

05

VISUALIZAÇÃO GRÁFICA

Visualização Gráfica

Além da manipulação dos dados, o Pandas se integra bem com bibliotecas de visualização, como o Matplotlib. Isso permite a criação de gráficos informativos diretamente a partir dos dados do DataFrame.



Example.py

```
import matplotlib.pyplot as plt  
  
dados['Idade'].plot(kind='hist')  
plt.show()
```


06

EXPORTANDO DADOS

Exportando Dados

Após realizar as manipulações necessárias, você pode exportar seus dados para um novo arquivo CSV ou outro formato desejado usando o Pandas.



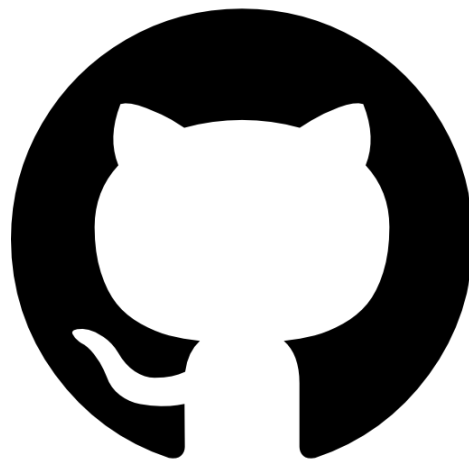
Example.py

```
dados.to_csv('dados_processados.csv', index=False)
```

AGRADECIMENTOS

Obrigado por ler até aqui! 😊

Esse ebook foi gerado por IA e diagramado por humano... No caso, eu.



<https://github.com/Anotherrafael>