

# Educação Profissional Paulista

Técnico em  
**Ciência de  
Dados**

# Bibliotecas: Pandas, NumPy, SciPy, Matplotlib e Seaborn

## Pandas: combinando *DataFrames*

Aula 4

Código da aula: [DADOS]ANO1C2B4S26A4

**Bibliotecas: Pandas,  
NumPy, SciPy,  
Matplotlib e Seaborn**

## Mapa da Unidade 5 Componente 3

semana

**25**

Pandas:  
transformando  
DataFrame

**Você está aqui!**

Pandas: combinando  
DataFrames

semana

**26**

semana

**23**

Pandas: acesso  
e seleção

semana

**24**

Pandas: ler e  
escrever

**Bibliotecas: Pandas,  
NumPy, SciPy,  
Matplotlib e Seaborn**

## **Mapa da Unidade 5 Componente 3**

# **Você está aqui!**

**Pandas: combinando  
DataFrames**

### **Aula 4**

Código da aula: [DADOS]ANO1C2B4S26A4

**26**



## Objetivos da Aula

- Praticar conceitos de manipulação e combinação de DataFrames usando a biblioteca Pandas do Python.



## Recursos Didáticos

- Recurso audiovisual para exibição de vídeos e imagens;
- Acesso ao laboratório de informática e/ou internet;
- Software Anaconda/Jupyter Notebook instalado ou similar.



## Duração da Aula

50 minutos.



## Competências Técnicas

- Ser proficiente em linguagens de programação para manipular e analisar grandes conjuntos de dados;
- Usar técnicas para explorar e analisar dados, aplicar modelos estatísticos, identificar padrões, realizar inferências e tomar decisões baseadas em evidências.



## Competências Socioemocionais

- Colaborar efetivamente com outros profissionais, como cientistas de dados e engenheiros de dados;
- Trabalhar em equipes multifuncionais, colaborando com colegas, gestores e clientes.





Vamos  
fazer um  
**quiz**

Qual método do Pandas é usado para combinar  
*DataFrames* com base em colunas comuns?

*join()*

*merge()*

*concat()*

*combine()*



Vamos  
fazer um  
**quiz**

Qual tipo de junção (*join*) inclui apenas os valores comuns nas colunas de ambos os *DataFrames*?

*Inner join*

*Outer join*

*Left join*

*Right join*



Vamos  
fazer um  
**quiz**

Como filtrar as linhas de um *DataFrame* onde o valor da coluna “Preço” é menor que 100?

`df[df["Preço"] > 100]`

`df[df["Preço"] < 100]`

`df[df["Preço"] == 100]`

`df[df["Preço"] >= 100]`





Colocando  
em **prática**

# Combinando a tabela de vendas de sorvete

Você é o gerente de uma sorveteria chamada "Scoops & Smiles" e está responsável por registrar as vendas diárias de sorvetes.

O objetivo é acompanhar as vendas ao longo do mês e analisar os sabores mais populares, bem como a quantidade vendida.

Confira a seguir as orientações para a atividade.



**Durante a aula**



**Em grupos de quatro alunos**



**Enviem o arquivo em .ipynb no AVA**

Situação fictícia elaborada especialmente para o curso.



Elaborado especialmente para o curso com apoio da ferramenta Microsoft Copilot.



Colocando  
em **prática**

# Combinando a tabela de vendas de sorvete

Porém, seu estagiário separou as vendas em duas tabelas e você precisa juntar as tabelas de vendas.

Com os DataFrames ao lado, crie uma única tabela com todas as vendas.

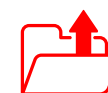
Situação fictícia elaborada especialmente para o curso.



**Durante a aula**



**Em grupos de quatro alunos**



**Enviem o arquivo em .ipynb no AVA**

```
1 vendas = pd.DataFrame(  
2     {  
3         "Data": [  
4             "2024-05-01", "2024-05-02", "2024-05-03", "2024-05-04", "2024-05-05",  
5             "2024-05-06", "2024-05-07", "2024-05-08", "2024-05-09", "2024-05-10"  
6         ],  
7         "Quantidade Vendida": [48, 38, 24, 17, 30, 48, 28, 32, 20, 20],  
8         "Sabor": [  
9             "Baunilha", "Creme", "Creme", "Creme", "Baunilha",  
10            "Baunilha", "Baunilha", "Morango", "Baunilha", "Baunilha"  
11        ]  
12     }  
13 )
```

```
1 vendas_novas = pd.DataFrame(  
2     {  
3         "Data": [  
4             "2024-05-11", "2024-05-12", "2024-05-13", "2024-05-14", "2024-05-15",  
5             "2024-05-16", "2024-05-17", "2024-05-18", "2024-05-19", "2024-05-20"  
6         ],  
7         "Quantidade Vendida": [33, 45, 49, 33, 12, 31, 11, 33, 39, 47],  
8         "Sabor": [  
9             "Baunilha", "Morango", "Baunilha", "Creme", "Baunilha",  
10            "Creme", "Creme", "Creme", "Morango", "Morango"  
11        ]  
12     }  
13 )
```

Elaborado especialmente para o curso com a ferramenta Jupyter Notebook.



Colocando  
em **prática**

## Analizando as vendas da sorveteria

Com os dados de vendas completos, combine com a tabela de preços e filtre apenas as datas onde a receita da venda de sorvetes foi maior que 200.

Abaixo, é possível conferir os valores dos sabores dos sorvetes.

**Lembre-se de que Receita é a quantidade\*preço.**

```
1 precos = pd.DataFrame(  
2     {  
3         "Sabor": ["Baunilha", "Chocolate", "Morango", "Creme"],  
4         "Preço": [5.0, 6.0, 5.5, 4.5]  
5     }  
6 )
```



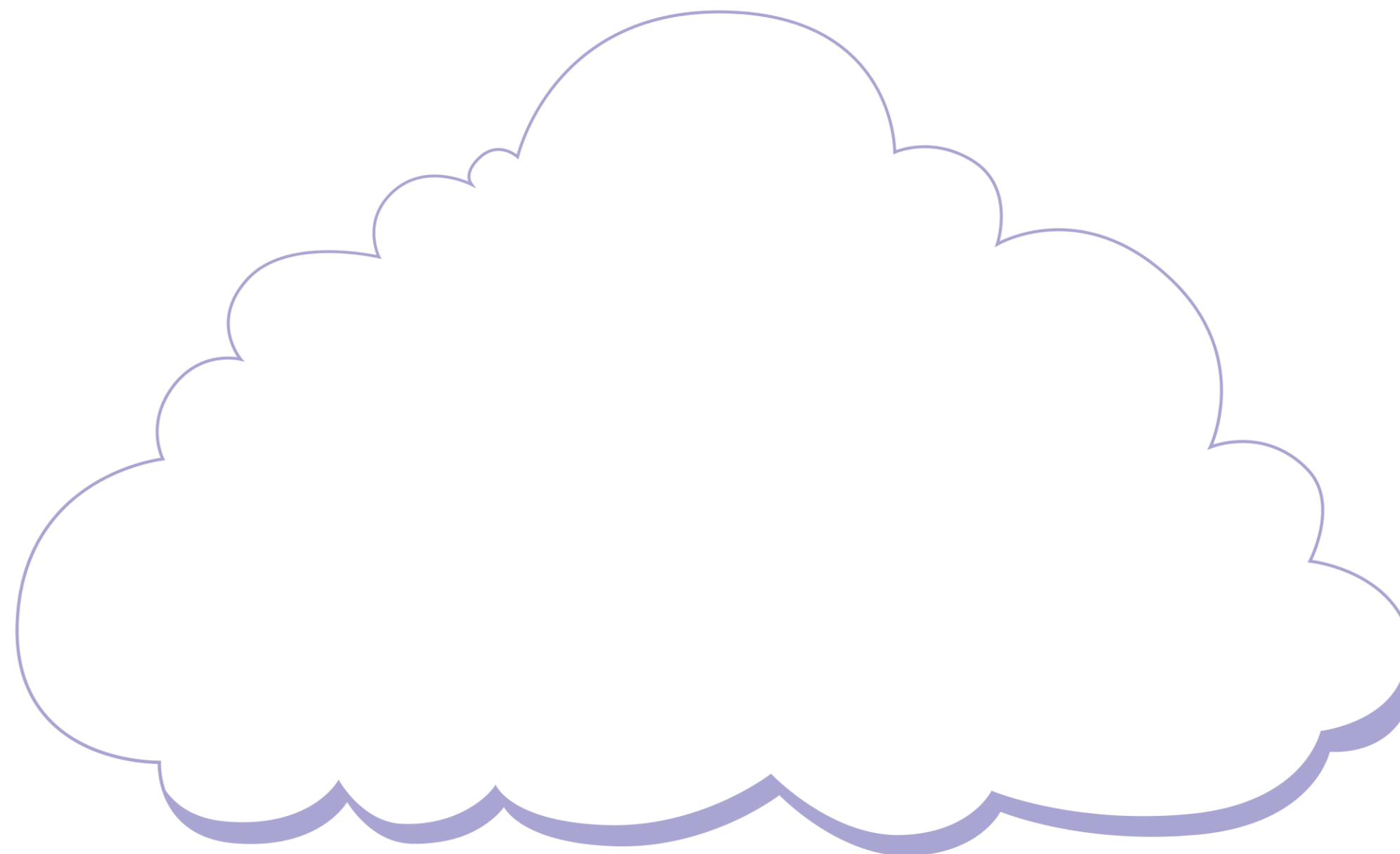
**Durante a aula**



**Em grupos de quatro alunos**

Elaborado especialmente para o curso com a ferramenta Jupyter Notebook.

# Nuvem de palavras



© Getty Images

O que nós  
**aprendemos  
hoje?**





© Getty Images

O que nós  
**aprendemos  
hoje?**

## Então ficamos assim...

- 1** Praticamos combinação entre DataFrames com os métodos *concat* e *merge*.
- 2** Filtramos DataFrames usando a biblioteca Pandas do Python.
- 3** Vimos na prática a importância de Pandas para os dados.

Ser  
sempre +

# Situação

## Muitas tabelas, muitos arquivos CSV

A sorveteria “Cone Cone” não só serve os sorvetes mais deliciosos da cidade, como também é conhecida por sua precisão matemática.

Todos os dias, ao se fecharem as portas, a equipe habilidosa registra todas as vendas, despesas e lucros em um arquivo CSV chamado “Data\_Sorveteria\_Cone\_Cone.csv”.

Situação fictícia elaborada especialmente para o curso.



Ser  
sempre +

## Ação

Na soverteria “Cone Cone”, você foi contratado para consolidar as vendas dos últimos três anos, ou seja, abrir e juntar todos os mais de mil arquivos em apenas um dia.



Elaborado especialmente para o curso com apoio da ferramenta Microsoft Copilot.

► Diante dessa situação, qual seria sua atitude?

Situação fictícia elaborada especialmente para o curso.



# Saiba mais

Quer saber como manipular arquivos e diretórios para ler vários arquivos .csv e juntar em um só?

Descubra em:

FELIPE, A. *Trabalhando com arquivos e diretórios no Python*. Alura, 14 jun. 2017. Disponível em:  
<https://www.alura.com.br/artigos/trabalhando-com-arquivos-e-diretorios-no-python>.  
Acesso em: 19 jul. 2024.



# Referências da aula

MCKINNEY, W. *Python para análise de dados: tratamento de dados com Pandas, NumPy & Jupyter*. São Paulo: Novatec, 2023.

PANDAS. *Pandas documentation*, 10 abr. 2024. Disponível em: <https://pandas.pydata.org/docs/>. Acesso em: 19 jul. 2024.

Identidade visual: imagens © Getty Images.

# Educação Profissional Paulista

Técnico em  
**Ciência de  
Dados**