

**Educação
Profissional
Paulista**

Técnico em
**Ciência de
Dados**

Bibliotecas: Pandas, NumPy, SciPy, Matplotlib e Seaborn

Pandas: acesso e seleção

Aula 3

Código da aula: [DADOS]ANO1C2B3S23A3

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

Mapa da Unidade 5 Componente 3

semana

22

NumPy:
manipulação
de array

semana

24

Pandas: ler e
escrever

semana

23

Você está aqui!
Pandas: acesso
e seleção

semana

25

Pandas: transformando
Data Frame

semana

26

Pandas: combinando
Data Frame

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

Mapa da Unidade 5 Componente 3

Você está aqui!

Pandas: acesso e seleção

Aula 3

Código da aula: [DADOS]ANO1C2B3S23A3

23



Objetivos da aula

- Aprender sobre como importar e criar um Data Frame.



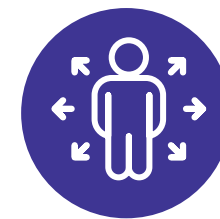
Recursos didáticos

- Recurso audiovisual para exibição de vídeos e imagens;
- Acesso ao laboratório de informática e/ou à 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 com base 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.

Construindo
o **conceito**

Pandas – Criando

Importe a biblioteca:

```
# Importando a biblioteca
```

```
import pandas as pd
```

Elaborado especialmente para o curso com a ferramenta Jupyter Notebook.

Construindo o **conceito**

Pandas – Criando

Em seguida, abra ou crie o Data Frame:

```
import pandas as pd

# Criando um DataFrame chamado primeiro_df com 2 colunas 'a' e 'b' e 2 valores 1 e 2
primeiro_df = pd.DataFrame({
    'a': [1, 2],
    'b': [1, 2]
})

# Exibindo o DataFrame
primeiro_df
```

	a	b
0	1	1
1	2	2

Elaborado especialmente para o curso com a ferramenta Jupyter Notebook.

Construindo o conceito

Pandas – Criando

Veja outro exemplo de criação de Data Frame:

```
import pandas as pd

# Criando um dicionário com dados básicos
dados_basicos = {
    "ID": [1, 2],
    "Nome": ["Alice", "Bob"],
    "Idade": [28, 32],
    "Cidade": ["São Paulo", "Rio de Janeiro"]
}

# Criando o DataFrame usando os dados básicos
df_basicos = pd.DataFrame(dados_basicos)

# Exibindo o DataFrame
df_basicos
```

	ID	Nome	Idade	Cidade
0	1	Alice	28	São Paulo
1	2	Bob	32	Rio de Janeiro

Elaborado especialmente para o curso com a ferramenta Jupyter Notebook.

Construindo
o conceito

Pandas – Criando

Veja outro exemplo de criação de Data Frame:

```
import pandas as pd

# Criando um dicionário com dados de saúde
dados = {
    "ID": [1, 2, 3, 4, 5],
    "Nome do Paciente": ["João Silva", "Maria Souza", "Ana Costa", "Lucas Martins", "Beatriz Santos"],
    "Idade": [30, 25, 40, 22, 35],
    "Sexo": ["M", "F", "F", "M", "F"],
    "Diagnóstico": ["Gripe", "Alergia", "Hipertensão", "Fratura no braço", "Ansiedade"],
    "Data da Consulta": ["2024-04-01", "2024-04-02", "2024-04-03", "2024-04-04", "2024-04-05"],
    "Tratamento Recomendado": ["Repouso e hidratação", "Antialérgicos", "Dieta e exercícios", "Imobilização", "Terapia"],
    "Custo do Tratamento (R$)": [150, 200, 300, 450, 500]
}

# Criando o DataFrame usando os dados
df_saude = pd.DataFrame(dados)

# Exibindo o DataFrame
df_saude
```

	ID	Nome do Paciente	Idade	Sexo	Diagnóstico	Data da Consulta	Tratamento Recomendado	Custo do Tratamento (R\$)
0	1	João Silva	30	M	Gripe	2024-04-01	Repouso e hidratação	150
1	2	Maria Souza	25	F	Alergia	2024-04-02	Antialérgicos	200
2	3	Ana Costa	40	F	Hipertensão	2024-04-03	Dieta e exercícios	300
3	4	Lucas Martins	22	M	Fratura no braço	2024-04-04	Imobilização	450
4	5	Beatriz Santos	35	F	Ansiedade	2024-04-05	Terapia	500

Elaborado especialmente para o curso com a ferramenta Jupyter Notebook.

Construindo o conceito

Pandas – Criando

Veja outro exemplo de criação de Data Frame:

```
import pandas as pd

dados_futebol = {
    "ID": [1, 2, 3, 4, 5],
    "Nome do Jogador": ["Cristiano Ronaldo", "Lionel Messi", "Neymar Jr", "Kylian Mbappé", "Robert Lewandowski"],
    "Idade": [36, 34, 29, 22, 33],
    "Clube Atual": ["Manchester United", "Paris Saint-Germain", "Paris Saint-Germain", "Paris Saint-Germain", "Bayern Munich"],
    "Gols na Temporada": [24, 30, 19, 27, 41]
}

# Criando o DataFrame usando os dados de futebol
df_futebol = pd.DataFrame(dados_futebol)

# Exibindo o DataFrame
df_futebol
```

	ID	Nome do Jogador	Idade	Clube Atual	Gols na Temporada
0	1	Cristiano Ronaldo	36	Manchester United	24
1	2	Lionel Messi	34	Paris Saint-Germain	30
2	3	Neymar Jr	29	Paris Saint-Germain	19
3	4	Kylian Mbappé	22	Paris Saint-Germain	27
4	5	Robert Lewandowski	33	Bayern Munich	41

Elaborado especialmente para o curso com a ferramenta Jupyter Notebook.

Construindo o conceito

Pandas – Criando

Para criar Data Frames, usamos a função `pd.DataFrame()`.

`pd.DataFrame()`: esta é a função mais direta para criar um Data Frame a partir de um dicionário de listas, uma lista de dicionários ou até mesmo um NumPy array.

```
d = {'col1': [1, 2], 'col2': [3, 4]}  
df = pd.DataFrame(data=d)  
df
```

	col1	col2
0	1	3
1	2	4

```
d = {'col1': [0, 1, 2, 3], 'col2': pd.Series([2, 3], index=[2, 3])}  
pd.DataFrame(data=d, index=[0, 1, 2, 3])
```

	col1	col2
0	0	NaN
1	1	NaN
2	2	2.0
3	3	3.0

Elaborado especialmente para o curso com a ferramenta Jupyter Notebook.



Colocando
em **prática**

Construindo um Data Frame

Com base no Data Frame fornecido abaixo, desenvolva o código no Jupyter Notebook, utilizando o Pandas, para gerar essa visão de tabela:

	ID	Título	Ano	Diretor	Gênero	Duração (min)
0	1	O Poderoso Chefão	1972	Francis Ford Coppola	Crime, Drama	175
1	2	A Lista de Schindler	1993	Steven Spielberg	História, Drama	195
2	3	Forrest Gump	1994	Robert Zemeckis	Romance, Drama	142
3	4	O Senhor dos Anéis	2001	Peter Jackson	Aventura, Fantasia	178
4	5	O Silêncio dos Inocentes	1991	Jonathan Demme	Crime, Drama	118



Próxima aula



Em grupo



Enviar no AVA

Elaborado especialmente para o curso com a ferramenta Jupyter Notebook.



© Getty Images

O que nós
**aprendemos
hoje?**

Então ficamos assim...

- 1** Aprendemos como criar um Data Frame;
- 2** Praticamos a criação de um Data Frame com a função `pd.DataFrame`.

Saiba mais

Quer dominar a análise de dados com Python?
Aprenda tudo sobre Pandas com este curso da Alura.

ALURA. Pandas: conhecendo a biblioteca.

Disponível em:

<https://cursos.alura.com.br/course/pandas-conhecendo-biblioteca>.

Acesso em: 28 jun. 2024.

Referências da aula

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

PANDAS. Ferramenta de análise e manipulação de dados, 2024. Disponível em: <https://pandas.pydata.org/>. Acesso em: 28 jun. 2024.

Identidade visual: imagens © Getty Images.

**Educação
Profissional
Paulista**

Técnico em
**Ciência de
Dados**