



Aula 25: Estruturas de Dados III: Dataframes

| | |
|----------------|------------------------------|
| ☰ Ciclo | Ciclo 03: Estrutura de Dados |
| # Aula | 25 |
| 🕒 Created | @November 11, 2022 6:00 PM |
| ☑ Reviewed | ☑ |
| 📎 Material PDF | |
| ☑ Finished | ☑ |

Objetivo da Aula:

1. Apresentar a estrutura de dados chamada dataframes
2. Na próxima aula

Conteúdo:

▼ 1. Criando Dataframes

▼ Exemplos:

```
usuarios = {  
    "nome" : ["Meigarom", "Claudio", "Fernando", "Diego", "André"],  
    "idade" : [34, 30, 36, 25, 36],  
    "cidade" : ["Indaiatuba", "Arcos", "Ouro Fino", "Itapira", "Curitiba"],  
    "estado": ["SP", "RJ", "SP", "RJ", "RJ"]  
}  
  
import pandas as pd
```

```
df = pd.DataFrame( usuarios )
```

▼ 2. Lista de Comandos:

O acesso aos valores do dicionário acontece somente através do nome da chave.

| Ação | Comandos em Python |
|---|--|
| Criar um dataframe vazio | <code>df = pd.DataFrame()</code> |
| Seleciona a primeira linha e primeira coluna | <code>df.loc[<linha>, <coluna>]</code> |
| Selecionar primeira linha e um intervalo de coluna | <code>df.loc[0, <inicio:fim>]</code> |
| Selecionar um intervalo de linha e uma coluna | <code>df.loc[<inicio:fim>, 0]</code> |
| Selecionar um intervalo de linha e um intervalo de coluna | <code>df.loc[<inicio:fim>, <inicio:fim>]</code> |
| Selecionar um intervalo de linha e uma coluna pelo nome | <code>df.loc[<inicio:fim>, "idade"]</code> |
| Selecionar um intervalo de linha e várias colunas pelo nome | <code>df.loc[<inicio:fim>, ["idade", "nome", "cidade"]]</code> |
| | |

▼ Exemplos:

```
# Biblioteca que contém o comando para criar dataframes
import pandas as pd

# Dicionario com as informacoes do dataframe
usuarios = {
    "nome" : ["Meigarom", "Claudio", "Fernando", "Diego", "André"],
    "idade" : [34, 30, 36, 25, 36],
    "cidade" : ["Indaiatuba", "Arcos", "Ouro Fino", "Itapira", "Curitiba"],
    "estado": ["SP", "RJ", "SP", "RJ", "RJ"]
}

# Criar um dataframe vazio
df_vazio = pd.DataFrame()

# criacao do dataframe a partir dos dados do dicionario
df = pd.DataFrame( usuarios )
```

```

# Seleciona a primeira linha e primeira coluna
df_selecionado = df.iloc[0, 0]
print( df_selecionado )

# Selecionar primeira linha e um intervalo de coluna
df_selecionado = df.iloc[0, 5:10]
print( df_selecionado )

# Selecionar um intervalo de linha e uma coluna
df_selecionado = df.iloc[0:25, 3]
print( df_selecionado )

# Selecionar um intervalo de linha e um intervalo de coluna
df_selecionado = df.iloc[0:25, 3:7]
print( df_selecionado )

# Selecionar um intervalo de linha e uma coluna pelo nome
df_selecionado = df.iloc[0:25, "idade"]
print( df_selecionado )

# Selecionar um intervalo de linha e várias colunas pelo nome
colunas = ["idade", "estado", "cidade"]
df_selecionado = df.iloc[0:25, colunas ]
print( df_selecionado )

```

▼ 3. Lendo arquivos externos:

Use esses comandos para ler arquivos externos, armazenados no seu computador, para dentro do Google Collabs.

```

# bibliotecas necessárias
import pandas as pd
import io
from google import colab as cl

# importando o arquivo para o computador
file_upload = cl.files.upload()

# lendo o arquivo importado
df = pd.read_csv( io.BytesIO( file_upload( <NOME DO ARQUIVO> ) ) )

```

▼ Exemplos:

```

# bibliotecas necessárias
import pandas as pd
import io
from google import colab as cl

# importando o arquivo para o computador

```

```
file_upload = cl.files.upload()

# lendo o arquivo importado
df = pd.read_csv( io.BytesIO( file_upload( 'train.csv' ) ) )
```

▼ Na próxima aula

Aula 26: Exercícios de fim de ciclo 03