



Tratando dados em Python

Criando um DataFrame

Um DataFrame pode ser criado a partir de um conjunto de Series.

```
aluno1 = pd.Series({'Nome': 'Carlos', 'Geografia': 6, 'História': 7})
```

```
aluno2 = pd.Series({'Nome': 'José', 'Geografia': 8, 'História': 8})
```

```
aluno3 = pd.Series({'Nome': 'Rafaela', 'Geografia': 4, 'História': 5.3})
```

```
df = pd.DataFrame([aluno1, aluno2, aluno3])
```

Criando um DataFrame

Um DataFrame também pode ser criado de forma direta através da importação de uma planilha.

```
df = pd.read_csv('nota_dos_alunos.csv')
```

Selecionando colunas de um DataFrame

A seleção de elementos de um DataFrame pode ser feito indicando o nome da coluna e serão retornados todos os dados da coluna.

```
df['Nome']
```

```
df['Geografia']
```

Selecionando linhas de um DataFrame

Para selecionar apenas uma linha específica, deve ser utilizada a função *loc*. Sendo possível indicar somente a linha ou a linha e coluna.

```
df.loc[1]
```

```
df.loc[1, 'Nome']
```

Selecionando linhas de um DataFrame

Existe a possibilidade de selecionar um conjunto de linhas:

```
df.loc[1:2, 'Nome']
```

Alterando atributo de um DataFrame

Também é possível alterar o valor de um atributo indicando a linha e a coluna em que o valor será atribuído.

```
df.loc[1, 'Nome'] = "Carla"
```

Renomeando colunas DataFrame

Determinadas situações podem requerer a nomeação dos nomes das colunas do DataFrame. Para isso, pode ser feita uma atualização, devendo ser especificadas todas as colunas.

```
df.columns = "Nome", "Física", "Química"
```


Renomeando colunas DataFrame

A renomeação também pode ser feita utilizando a função *rename* e a especificação dos nomes das colunas.

```
df = df.rename(columns={'Nome.': 'Nome', 'Matemática.': 'Matemática', 'Português.':  
'Português'})
```

Filtrando dados do DataFrame

Os dados de um DataFrame podem ser filtrados diretamente, indicando uma condição específica para seleção das linhas.

```
df.loc[df['Geografia'] < 7]
```

Filtrando dados do DataFrame

É possível filtrar dados de duas condições diferentes para um mesmo DataFrame. No exemplo, são filtrados os alunos que foram aprovados nas disciplinas de Geografia e História.

```
df[(df['Geografia'] > 6) & (df['História'] > 6)]
```

Transposição de um DataFrame

A transposição é a transformação da estrutura da “planilha” do DataFrame, convertendo o que é coluna em linha e linha em colunas.

`df.T`

Removendo coluna de um DataFrame

A remoção de coluna de um DataFrame pode ser feita especificando o nome da coluna e utilização da função *drop*.

```
df.drop(columns=[ 'Nome' ] )
```

Removendo coluna de um DataFrame

Os dados podem ser ordenados por alguma coluna específica, através da função *sort_values*.

```
df.sort_values(by='Geografia')
```



Obrigada!

Ana Laurentino