In [1]:

```
!pip install pandas
```

```
Collecting pandas
  Downloading pandas-1.3.2-cp39-cp39-win_amd64.whl (10.2 MB)
Requirement already satisfied: python-dateutil>=2.7.3 in d:\curso-data-scien
ce\venv\lib\site-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2017.3 in d:\curso-data-science\venv\li
b\site-packages (from pandas) (2021.1)
Requirement already satisfied: numpy>=1.17.3 in d:\curso-data-science\venv\l
ib\site-packages (from pandas) (1.21.2)
Requirement already satisfied: six>=1.5 in d:\curso-data-science\venv\lib\si
te-packages (from python-dateutil>=2.7.3->pandas) (1.16.0)
Installing collected packages: pandas
Successfully installed pandas-1.3.2

WARNING: You are using pip version 21.2.3; however, version 21.2.4 is availa
ble.
You should consider upgrading via the 'D:\curso-data-science\venv\Scripts\py
thon.exe -m pip install --upgrade pip' command.
```

Criação de um DataFrame

In [2]:

```
import numpy as np
import pandas as pd
```

In [9]:

```
#df = pd.DataFrame([['a', 'b', 'c'], [1,2,3]])
#df = pd.DataFrame({'a': [1,2,3], 'b':[5,6,7]})
#df = pd.DataFrame(np.random.randn(3,2))
df = pd.DataFrame(np.random.randint(1,10,6).reshape(2,3),
                  columns=['São Paulo', 'Flamengo', 'Inter'],
                  index=['Vitorias', 'Derrotas'])
df
```

Out[9]:

	São Paulo	Flamengo	Inter
Vitorias	4	5	3
Derrotas	6	8	4

Seleção de Dados

In [10]:

```
df['São Paulo']
```

Out[10]:

```
Vitorias    4
Derrotas    6
Name: São Paulo, dtype: int32
```

In [11]:

```
type(df['São Paulo'])
```

Out[11]:

```
pandas.core.series.Series
```

In [23]:

```
df['São Paulo']['Vitorias']
```

Out[23]:

```
4
```

In [24]:

```
df
```

Out[24]:

	São Paulo	Flamengo	Inter
Vitorias	4	5	3
Derrotas	6	8	4

In [26]:

```
df.loc['Vitorias']
```

Out[26]:

```
São Paulo    4
Flamengo     5
Inter        3
Name: Vitorias, dtype: int32
```

In [30]:

```
df.loc[['Vitorias'], ['Flamengo']]
```

Out[30]:

	Flamengo
Vitorias	5

In [31]:

```
df.iloc[1]
```

Out[31]:

```
São Paulo    6
Flamengo     8
Inter        4
Name: Derrotas, dtype: int32
```

In [32]:

```
df.iloc[0,1:3]
```

Out[32]:

```
Flamengo    5
Inter       3
Name: Vitorias, dtype: int32
```

Adição e Remoção de linhas e colunas no DataFrame

In [33]:

```
df
```

Out[33]:

	São Paulo	Flamengo	Inter
Vitorias	4	5	3
Derrotas	6	8	4

In [34]:

```
df['Vasco'] = [3,17]
```

In [35]:

```
df
```

Out[35]:

	São Paulo	Flamengo	Inter	Vasco
Vitorias	4	5	3	3
Derrotas	6	8	4	17

In [36]:

```
df.loc['Empates']=[0,0,0,0]
df
```

Out[36]:

	São Paulo	Flamengo	Inter	Vasco
Vitorias	4	5	3	3
Derrotas	6	8	4	17
Empates	0	0	0	0

In [40]:

```
df.drop('Vasco', axis=1)
```

Out[40]:

	São Paulo	Flamengo	Inter
Vitorias	4	5	3
Derrotas	6	8	4
Empates	0	0	0

In [41]:

```
df.drop('Vitorias', axis=0)
```

Out[41]:

	São Paulo	Flamengo	Inter	Vasco
Derrotas	6	8	4	17
Empates	0	0	0	0

In [43]:

```
df.drop('Inter', axis=1, inplace=True)
```

In [44]:

```
df
```

Out[44]:

	São Paulo	Flamengo	Vasco
Vitorias	4	5	3
Derrotas	6	8	17
Empates	0	0	0

In [45]:

```
df2 = pd.DataFrame({'a': [1,2,np.nan], 'b':[5,6,7], 'c':[2,4, np.nan]})
```

In [46]:

```
df2
```

Out[46]:

	a	b	c
0	1.0	5	2.0
1	2.0	6	4.0
2	NaN	7	NaN

In [47]:

```
df2.dropna()
```

Out[47]:

	a	b	c
0	1.0	5	2.0
1	2.0	6	4.0

In [48]:

```
df2
```

Out[48]:

	a	b	c
0	1.0	5	2.0
1	2.0	6	4.0
2	NaN	7	NaN

In [50]:

```
df2.fillna('SEM')
```

Out[50]:

	a	b	c
0	1.0	5	2.0
1	2.0	6	4.0
2	SEM	7	SEM

In [51]:

```
df2.fillna(value=df2['b'].mean())
```

Out[51]:

	a	b	c
0	1.0	5	2.0
1	2.0	6	4.0
2	6.0	7	6.0

In [ ]:

```

```