

Relatórios de Análise VI

Criando Novas Variáveis

```
In [70]: # importando Pandas
import pandas as pd
```

```
In [71]: # importando a base de dados
dados = pd.read_csv('data/aluguel_residencial.csv', sep = ';')
```

```
In [72]: # exibindo os primeiros dez itens
dados.head(10)
```

Out[72]:

	Tipo	Bairro	Quartos	Vagas	Suites	Area	Valor	Condominio	IPTU
0	Quitinete	Copacabana	1	0	0	40	1700.0	500.0	60.0
1	Casa	Jardim Botânico	2	0	1	100	7000.0	0.0	0.0
2	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0
3	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	0.0
4	Apartamento	Cachambi	2	0	0	50	1300.0	301.0	17.0
5	Casa de Condomínio	Barra da Tijuca	5	4	5	750	22000.0	0.0	0.0
6	Casa de Condomínio	Ramos	2	2	0	65	1000.0	0.0	0.0
7	Apartamento	Grajaú	2	1	0	70	1500.0	642.0	74.0
8	Apartamento	Lins de Vasconcelos	3	1	1	90	1500.0	455.0	14.0
9	Apartamento	Copacabana	1	0	1	40	2000.0	561.0	50.0

```
In [73]: # criando uma nova variável chamada Valor Bruto
# consistirá na união dos valores das variáveis Valor, Condominio e IPTU
dados['Valor Bruto'] = dados['Valor'] + dados['Condominio'] + dados['IPTU']
```

```
In [74]: # exibindo os primeiros dez itens
dados.head(10)
```

Out[74]:

	Tipo	Bairro	Quartos	Vagas	Suites	Area	Valor	Condominio	IPTU	Valor Bruto
0	Quitinete	Copacabana	1	0	0	40	1700.0	500.0	60.0	2260.0
1	Casa	Jardim Botânico	2	0	1	100	7000.0	0.0	0.0	7000.0
2	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0	1210.0
3	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	0.0	1030.0
4	Apartamento	Cachambi	2	0	0	50	1300.0	301.0	17.0	1618.0
5	Casa de Condomínio	Barra da Tijuca	5	4	5	750	22000.0	0.0	0.0	22000.0
6	Casa de Condomínio	Ramos	2	2	0	65	1000.0	0.0	0.0	1000.0
7	Apartamento	Grajaú	2	1	0	70	1500.0	642.0	74.0	2216.0
8	Apartamento	Lins de Vasconcelos	3	1	1	90	1500.0	455.0	14.0	1969.0
9	Apartamento	Copacabana	1	0	1	40	2000.0	561.0	50.0	2611.0

```
In [75]: # criando variável que armazenará o valor por metro quadrado
dados['Valor m2'] = dados['Valor'] / dados['Area']
```

```
In [76]: # exibindo os primeiros dez itens
dados.head(10)
```

Out[76]:

	Tipo	Bairro	Quartos	Vagas	Suites	Area	Valor	Condominio	IPTU	Valor Bruto	Valor m2
0	Quitinete	Copacabana	1	0	0	40	1700.0	500.0	60.0	2260.0	42.500000
1	Casa	Jardim Botânico	2	0	1	100	7000.0	0.0	0.0	7000.0	70.000000
2	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0	1210.0	53.333333
3	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	0.0	1030.0	16.666667
4	Apartamento	Cachambi	2	0	0	50	1300.0	301.0	17.0	1618.0	26.000000
5	Casa de Condomínio	Barra da Tijuca	5	4	5	750	22000.0	0.0	0.0	22000.0	29.333333
6	Casa de Condomínio	Ramos	2	2	0	65	1000.0	0.0	0.0	1000.0	15.384615
7	Apartamento	Grajaú	2	1	0	70	1500.0	642.0	74.0	2216.0	21.428571
8	Apartamento	Lins de Vasconcelos	3	1	1	90	1500.0	455.0	14.0	1969.0	16.666667
9	Apartamento	Copacabana	1	0	1	40	2000.0	561.0	50.0	2611.0	50.000000

```
In [77]: # formatando à quantidade de casas decimais da variável 'Valor m2'
dados['Valor m2'] = dados['Valor m2'].round(2)
```

```
In [78]: # exibindo os primeiros dez itens
dados.head(10)
```

Out[78]:

	Tipo	Bairro	Quartos	Vagas	Suites	Area	Valor	Condominio	IPTU	Valor Bruto	Valor m2
0	Quitinete	Copacabana	1	0	0	40	1700.0	500.0	60.0	2260.0	42.50
1	Casa	Jardim Botânico	2	0	1	100	7000.0	0.0	0.0	7000.0	70.00
2	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0	1210.0	53.33
3	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	0.0	1030.0	16.67
4	Apartamento	Cachambi	2	0	0	50	1300.0	301.0	17.0	1618.0	26.00
5	Casa de Condomínio	Barra da Tijuca	5	4	5	750	22000.0	0.0	0.0	22000.0	29.33
6	Casa de Condomínio	Ramos	2	2	0	65	1000.0	0.0	0.0	1000.0	15.38
7	Apartamento	Grajaú	2	1	0	70	1500.0	642.0	74.0	2216.0	21.43
8	Apartamento	Lins de Vasconcelos	3	1	1	90	1500.0	455.0	14.0	1969.0	16.67
9	Apartamento	Copacabana	1	0	1	40	2000.0	561.0	50.0	2611.0	50.00

```
In [79]: # criando variável para o valor bruto do metro quadrado
dados['Valor Bruto m2'] = (dados['Valor Bruto'] / dados['Area']).round(2))
```

```
In [80]: # exibindo os primeiros dez itens
dados.head(10)
```

Out[80]:

	Tipo	Bairro	Quartos	Vagas	Suites	Area	Valor	Condominio	IPTU	Valor Bruto	Valor m2	Valor Bruto m2
0	Quitinete	Copacabana	1	0	0	40	1700.0	500.0	60.0	2260.0	42.50	56.500000
1	Casa	Jardim Botânico	2	0	1	100	7000.0	0.0	0.0	7000.0	70.00	70.000000
2	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0	1210.0	53.33	80.666667
3	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	0.0	1030.0	16.67	21.458333
4	Apartamento	Cachambi	2	0	0	50	1300.0	301.0	17.0	1618.0	26.00	32.360000
5	Casa de Condomínio	Barra da Tijuca	5	4	5	750	22000.0	0.0	0.0	22000.0	29.33	29.333333
6	Casa de Condomínio	Ramos	2	2	0	65	1000.0	0.0	0.0	1000.0	15.38	15.384615
7	Apartamento	Grajaú	2	1	0	70	1500.0	642.0	74.0	2216.0	21.43	31.657143
8	Apartamento	Lins de Vasconcelos	3	1	1	90	1500.0	455.0	14.0	1969.0	16.67	21.877778
9	Apartamento	Copacabana	1	0	1	40	2000.0	561.0	50.0	2611.0	50.00	65.275000

```
In [81]: # criando variável que abrigue os tipos de imóvel casa e apartamento
casa = ['Casa', 'Casa de Condomínio', 'Casa de Vila']
```

```
In [82]: # recolhendo e identificando os marcadores abrigados na variável 'casa'
# método apply(), permite que apliquemos uma função à cada registro do DataFrame
# função lambda para definir no tipo agregado se é casa ou apartamento
dados['Tipo Agregado'] = dados['Tipo'].apply(lambda x: 'Casa' if x in casa else 'Apartament
o')
```

```
In [83]: # exibindo os primeiros dez itens
dados.head(10)
```

Out[83]:

	Tipo	Bairro	Quartos	Vagas	Suites	Area	Valor	Condominio	IPTU	Valor Bruto	Valor m2	Valor Bruto m2	Tipo Agregad
0	Quitinete	Copacabana	1	0	0	40	1700.0	500.0	60.0	2260.0	42.50	56.500000	Apartament
1	Casa	Jardim Botânico	2	0	1	100	7000.0	0.0	0.0	7000.0	70.00	70.000000	Cas
2	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0	1210.0	53.33	80.666667	Apartament
3	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	0.0	1030.0	16.67	21.458333	Apartament
4	Apartamento	Cachambi	2	0	0	50	1300.0	301.0	17.0	1618.0	26.00	32.360000	Apartament
5	Casa de Condomínio	Barra da Tijuca	5	4	5	750	22000.0	0.0	0.0	22000.0	29.33	29.333333	Cas
6	Casa de Condomínio	Ramos	2	2	0	65	1000.0	0.0	0.0	1000.0	15.38	15.384615	Cas
7	Apartamento	Grajaú	2	1	0	70	1500.0	642.0	74.0	2216.0	21.43	31.657143	Apartament
8	Apartamento	Lins de Vasconcelos	3	1	1	90	1500.0	455.0	14.0	1969.0	16.67	21.877778	Apartament
9	Apartamento	Copacabana	1	0	1	40	2000.0	561.0	50.0	2611.0	50.00	65.275000	Apartament

Excluindo Variáveis

```
In [84]: # criando um DataFrame auxiliar para explorar as possibilidades disponiveis de excluir variá
veis
dados_aux = pd.DataFrame(dados[['Tipo Agregado', 'Valor m2', 'Valor Bruto', 'Valor Bruto m2'
]])
```

```
In [85]: # exibindo os primeiros dez itens
dados_aux.head(10)
```

Out[85]:

	Tipo Agregado	Valor m2	Valor Bruto	Valor Bruto m2
0	Apartamento	42.50	2260.0	56.500000
1	Casa	70.00	7000.0	70.000000
2	Apartamento	53.33	1210.0	80.666667
3	Apartamento	16.67	1030.0	21.458333
4	Apartamento	26.00	1618.0	32.360000
5	Casa	29.33	22000.0	29.333333
6	Casa	15.38	1000.0	15.384615
7	Apartamento	21.43	2216.0	31.657143
8	Apartamento	16.67	1969.0	21.877778
9	Apartamento	50.00	2611.0	65.275000

```
In [86]: # excluindo a variável 'Valor Bruto' do banco de dados
del dados_aux['Valor Bruto']
```

```
In [87]: # exibindo os primeiros dez itens
# 'Valor Bruto' não estará mais disponível
dados_aux.head(10)
```

Out[87]:

	Tipo Agregado	Valor m2	Valor Bruto m2
0	Apartamento	42.50	56.500000
1	Casa	70.00	70.000000
2	Apartamento	53.33	80.666667
3	Apartamento	16.67	21.458333
4	Apartamento	26.00	32.360000
5	Casa	29.33	29.333333
6	Casa	15.38	15.384615
7	Apartamento	21.43	31.657143
8	Apartamento	16.67	21.877778
9	Apartamento	50.00	65.275000

```
In [88]: # outra maneira de excluir variáveis com método pop()
dados_aux.pop('Valor Bruto m2')
```

Out[88]:

0	56.500000
1	70.000000
2	80.666667
3	21.458333
4	32.360000
...	
21821	20.542857
21822	43.518519
21823	33.846154
21824	40.541667
21825	53.714286
Name: Valor Bruto m2, Length: 21826, dtype: float64	

```
In [89]: # exibindo os primeiros dez itens
# 'Valor Bruto m2' não estará mais disponível
dados_aux.head(10)
```

Out[89]:

	Tipo Agregado	Valor m2
0	Apartamento	42.50
1	Casa	70.00
2	Apartamento	53.33
3	Apartamento	16.67
4	Apartamento	26.00
5	Casa	29.33
6	Casa	15.38
7	Apartamento	21.43
8	Apartamento	16.67
9	Apartamento	50.00

```
In [90]: # utilizando nosso DataFrame original para excluirmos múltiplas variáveis
# utilizando o método drop() para essa exclusão
# para utilizar o método drop() é necessário especificar o eixo em que se dará a exclusão
# definimos axis = 1, pois o eixo 0 corresponde às linhas e o 1 colunas
dados.drop(['Valor Bruto', 'Valor Bruto m2'], axis = 1, inplace = True)
```

```
In [91]: # exibindo os primeiros dez itens
dados.head(10)
```

Out[91]:

	Tipo	Bairro	Quartos	Vagas	Suites	Area	Valor	Condominio	IPTU	Valor m2	Tipo Agregado
0	Quitinete	Copacabana	1	0	0	40	1700.0	500.0	60.0	42.50	Apartamento
1	Casa	Jardim Botânico	2	0	1	100	7000.0	0.0	0.0	70.00	Casa
2	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0	53.33	Apartamento
3	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	0.0	16.67	Apartamento
4	Apartamento	Cachambi	2	0	0	50	1300.0	301.0	17.0	26.00	Apartamento