

Relatório de Análise VI

Criando Novas Variáveis

In [105... `import pandas as pd`

In [106... `dados = pd.read_csv('../dados/aluguel_residencial_notnull.csv', sep=';')
dados.head(10)`

Out[106...

	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

Criando uma Variável (coluna)

Somando valores de outras colunas

In [107... `dados['Valor Bruto'] = dados['Valor'] + dados['Condominio'] + dados['IPTU']
dados.head(10)`

Out[107...

	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

	Tipo	Bairro	Quartos	Vagas	Suites	Area	Valor	Condominio	IPTU	Valor Bruto
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

Usando Divisão

```
In [108... dados['Valor m2'] = dados['Valor'] / dados['Area']
dados.head(10).sort_index(axis=1, ascending=False)
# para melhor visualização o cabeçalho da coluna foi organizado de trás para
# para não cortar a coluna quando exportar o arquivo para pdf
# utilizou-se então a função sort_index, com os parâmetros:
# axis=1 => para organizar as colunas
# ascending=False => para organizar de maneira decrescente
```

Out[108...

	Valor m2	Valor Bruto	Valor	Vagas	Tipo	Suites	Quartos	IPTU	Condominio	Ba
0	42.500000	2260.0	1700.0	0	Quitinete	0	1	60.0	500.0	Copacab
1	70.000000	7000.0	7000.0	0	Casa	1	2	0.0	0.0	Jai Botã
2	53.333333	1210.0	800.0	0	Apartamento	0	1	20.0	390.0	Ce
3	16.666667	1030.0	800.0	0	Apartamento	0	1	0.0	230.0	Higienóp
4	26.000000	1618.0	1300.0	0	Apartamento	0	2	17.0	301.0	Cacha
5	29.333333	22000.0	22000.0	4	Casa de Condomínio	5	5	0.0	0.0	Barri Ti
6	15.384615	1000.0	1000.0	2	Casa de Condomínio	0	2	0.0	0.0	Ra
7	21.428571	2216.0	1500.0	1	Apartamento	0	2	74.0	642.0	Gr
8	16.666667	1969.0	1500.0	1	Apartamento	1	3	14.0	455.0	Lins Vasconc
9	50.000000	2611.0	2000.0	0	Apartamento	1	1	50.0	561.0	Copacab

Arredondando as casas decimais

```
In [109... dados['Valor m2'] = dados['Valor m2'].round(2)
dados.head(10).sort_index(axis=1, ascending=False)
# também é possível arredondar diretamente, colocando tudo entre parênteses
# dados['Valor m2'] = (dados['Valor'] / dados['Area']).round(2)
```

Out[109...

	Valor m2	Valor Bruto	Valor	Vagas	Tipo	Suites	Quartos	IPTU	Condominio	Bairro
0	42.50	2260.0	1700.0	0	Quitinete	0	1	60.0	500.0	Copacabana
1	70.00	7000.0	7000.0	0	Casa	1	2	0.0	0.0	Jardim Botânico
2	53.33	1210.0	800.0	0	Apartamento	0	1	20.0	390.0	Centro
3	16.67	1030.0	800.0	0	Apartamento	0	1	0.0	230.0	Higienópolis
4	26.00	1618.0	1300.0	0	Apartamento	0	2	17.0	301.0	Cachambi

	Valor m2	Valor Bruto	Valor	Vagas	Tipo	Suites	Quartos	IPTU	Condominio	Bairro
5	29.33	22000.0	22000.0	4	Casa de Condomínio	5	5	0.0	0.0	Barra da Tijuca
6	15.38	1000.0	1000.0	2	Casa de Condomínio	0	2	0.0	0.0	Ramos
7	21.43	2216.0	1500.0	1	Apartamento	0	2	74.0	642.0	Grajaú
8	16.67	1969.0	1500.0	1	Apartamento	1	3	14.0	455.0	Lins de Vasconcelos
9	50.00	2611.0	2000.0	0	Apartamento	1	1	50.0	561.0	Copacabana

Continuando com a divisão

In [110]...

```
dados['Valor Bruto m2'] = (dados['Valor Bruto'] / dados['Area']).round(2)
dados.head(10).sort_index(axis=1, ascending=False)
```

Out[110]...

	Valor m2	Valor Bruto m2	Valor Bruto	Valor	Vagas	Tipo	Suites	Quartos	IPTU	Condominio	
0	42.50	56.50	2260.0	1700.0	0	Quitinete	0	1	60.0	500.0	Copa
1	70.00	70.00	7000.0	7000.0	0	Casa	1	2	0.0	0.0	B
2	53.33	80.67	1210.0	800.0	0	Apartamento	0	1	20.0	390.0	
3	16.67	21.46	1030.0	800.0	0	Apartamento	0	1	0.0	230.0	Higie
4	26.00	32.36	1618.0	1300.0	0	Apartamento	0	2	17.0	301.0	Ca
5	29.33	29.33	22000.0	22000.0	4	Casa de Condomínio	5	5	0.0	0.0	B
6	15.38	15.38	1000.0	1000.0	2	Casa de Condomínio	0	2	0.0	0.0	
7	21.43	31.66	2216.0	1500.0	1	Apartamento	0	2	74.0	642.0	
8	16.67	21.88	1969.0	1500.0	1	Apartamento	1	3	14.0	455.0	Vasci
9	50.00	65.28	2611.0	2000.0	0	Apartamento	1	1	50.0	561.0	Copa

Filtrando Valores para Criar uma Nova Variável

Listando quais valores existem em Tipo

In [111]...

```
list(dados['Tipo'].unique())
```

Out[111]...

```
['Quitinete', 'Casa', 'Apartamento', 'Casa de Condomínio', 'Casa de Vila']
```

Selecionando a variável desejada

In [112]...

```
casa = ['Casa', 'Casa de Condomínio', 'Casa de Vila']
```

Criando a nova coluna com aplicação de uma função lambda para cada registro

In [113...

```
dados['Tipo Agregado'] = dados['Tipo'].apply(lambda x: 'Casa'
                                              if x in casa else 'Apartamento')
dados.head(10).sort_index(axis=1, ascending=False)
```

Out[113...

	Valor m2	Valor Bruto m2	Valor Bruto	Valor	Vagas	Tipo Agregado	Tipo	Suites	Quartos	IPTU	Cond
0	42.50	56.50	2260.0	1700.0	0	Apartamento	Quitinete	0	1	60.0	
1	70.00	70.00	7000.0	7000.0	0	Casa	Casa	1	2	0.0	
2	53.33	80.67	1210.0	800.0	0	Apartamento	Apartamento	0	1	20.0	
3	16.67	21.46	1030.0	800.0	0	Apartamento	Apartamento	0	1	0.0	
4	26.00	32.36	1618.0	1300.0	0	Apartamento	Apartamento	0	2	17.0	
5	29.33	29.33	22000.0	22000.0	4	Casa	Casa de Condomínio	5	5	0.0	
6	15.38	15.38	1000.0	1000.0	2	Casa	Casa de Condomínio	0	2	0.0	
7	21.43	31.66	2216.0	1500.0	1	Apartamento	Apartamento	0	2	74.0	
8	16.67	21.88	1969.0	1500.0	1	Apartamento	Apartamento	1	3	14.0	
9	50.00	65.28	2611.0	2000.0	0	Apartamento	Apartamento	1	1	50.0	

In [114...

```
# Verificando se Casa de Condomínio também recebeu o tratamento certo
dados[dados['Tipo'] == 'Casa de Condomínio'].sort_index(axis=1,
                                                         ascending=False).head
```

Out[114...

	Valor m2	Valor Bruto m2	Valor Bruto	Valor	Vagas	Tipo Agregado	Tipo	Suites	Quartos	IPTU	Co
5	29.33	29.33	22000.0	22000.0	4	Casa	Casa de Condomínio	5	5	0.0	
6	15.38	15.38	1000.0	1000.0	2	Casa	Casa de Condomínio	0	2	0.0	
12	17.39	17.39	2000.0	2000.0	1	Casa	Casa de Condomínio	1	3	0.0	
16	16.09	21.88	10195.0	7500.0	3	Casa	Casa de Condomínio	2	4	0.0	
42	15.62	16.62	26600.0	25000.0	10	Casa	Casa de Condomínio	5	5	0.0	
58	17.86	26.43	3700.0	2500.0	3	Casa	Casa de Condomínio	1	5	0.0	
166	50.42	58.84	35010.0	30000.0	2	Casa	Casa de Condomínio	4	4	2510.0	
168	24.34	39.36	10508.0	6500.0	4	Casa	Casa de Condomínio	4	4	3308.0	
183	22.00	27.00	6750.0	5500.0	2	Casa	Casa de Condomínio	1	3	500.0	

	Valor m2	Valor Bruto m2	Valor Bruto	Valor	Vagas	Tipo Agregado	Tipo	Suites	Quartos	IPTU	Co
207	30.00	56.41	56410.0	30000.0	4	Casa	Casa de Condomínio	4	4	23715.0	

Aplicando a função lambda com várias condicionais

In [115...

```
# cuidado com os parênteses
# colocar as outras condicionais dentro de parênteses do primeiro else
# e cada if também entre parênteses
dados['Tamanho'] = dados['Quartos'].apply(lambda x: 'Grande' if x > 3
                                           else ('Médio' if (x >= 2 and x <= 3)
                                           else 'Pequeno'))
dados.head(10).sort_index(axis=1, ascending=False)
```

Out[115...

	Valor m2	Valor Bruto m2	Valor Bruto	Valor	Vagas	Tipo Agregado	Tipo	Tamanho	Suites	Quartos	II
0	42.50	56.50	2260.0	1700.0	0	Apartamento	Quitinete	Pequeno	0	1	
1	70.00	70.00	7000.0	7000.0	0	Casa	Casa	Médio	1	2	
2	53.33	80.67	1210.0	800.0	0	Apartamento	Apartamento	Pequeno	0	1	
3	16.67	21.46	1030.0	800.0	0	Apartamento	Apartamento	Pequeno	0	1	
4	26.00	32.36	1618.0	1300.0	0	Apartamento	Apartamento	Médio	0	2	
5	29.33	29.33	22000.0	22000.0	4	Casa	Casa de Condomínio	Grande	5	5	
6	15.38	15.38	1000.0	1000.0	2	Casa	Casa de Condomínio	Médio	0	2	
7	21.43	31.66	2216.0	1500.0	1	Apartamento	Apartamento	Médio	0	2	
8	16.67	21.88	1969.0	1500.0	1	Apartamento	Apartamento	Médio	1	3	
9	50.00	65.28	2611.0	2000.0	0	Apartamento	Apartamento	Pequeno	1	1	