

# Relatório de Análise VII

## Criando Agrupamentos

In [1]:

```
import pandas as pd
```

In [2]:

```
dados = pd.read_csv('dados/aluguel_residencial.csv', sep=';')
```

In [3]:

```
dados.head(10)
```

Out[3]:

	Tipo	Bairro	Quartos	Vagas	Suites	Area	Valor	Condominio	IPTU	Valor_I
0	Quitinete	Copacabana	1	0	0	40	1700.0	500.0	60.0	42
1	Casa	Jardim Botânico	2	0	1	100	7000.0	0.0	0.0	70
2	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0	53
3	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	0.0	16
4	Apartamento	Cachambi	2	0	0	50	1300.0	301.0	17.0	26
5	Casa de Condomínio	Barra da Tijuca	5	4	5	750	22000.0	0.0	0.0	29
6	Casa de Condomínio	Ramos	2	2	0	65	1000.0	0.0	0.0	15
7	Apartamento	Grajaú	2	1	0	70	1500.0	642.0	74.0	21
8	Apartamento	Lins de Vasconcelos	3	1	1	90	1500.0	455.0	14.0	16
9	Apartamento	Copacabana	1	0	1	40	2000.0	561.0	50.0	50

In [4]:

```
dados['Valor'].mean()
```

Out[4]:

```
5046.172821405663
```

In [5]:

```
dados['Bairro'].drop_duplicates()
```

Out[5]:

```
0          Copacabana
1    Jardim Botânico
2          Centro
3    Higienópolis
4          Cachambi
...
13855    Bancários
15277    Cavalcanti
15280    Rio da Prata
15412    Cidade Jardim
16428    Coelho Neto
Name: Bairro, Length: 152, dtype: object
```

<https://pandas.pydata.org/pandas-docs/stable/api.html#api-dataframe-stats> (<https://pandas.pydata.org/pandas-docs/stable/api.html#api-dataframe-stats>).

In [6]:

```
bairros = ['Barra da Tijuca', 'Copacabana', 'Ipanema', 'Leblon', 'Botafogo', 'Flamengo', 'T
```

In [7]:

```
# selecao = dados['Bairro'].isin(bairros)
# dados = dados[selecao]
# dados.head(10)
```

In [8]:

```
dados['Bairro'].drop_duplicates()
```

Out[8]:

```
0          Copacabana
1    Jardim Botânico
2          Centro
3    Higienópolis
4          Cachambi
...
13855    Bancários
15277    Cavalcanti
15280    Rio da Prata
15412    Cidade Jardim
16428    Coelho Neto
Name: Bairro, Length: 152, dtype: object
```

In [9]:

```
grupo_bairros = dados.groupby('Bairro')  
grupo_bairros
```

Out[9]:

<pandas.core.groupby.generic.DataFrameGroupBy object at 0x0000021E70219648>

In [10]:

```
type(grupo_bairros)
```

Out[10]:

pandas.core.groupby.generic.DataFrameGroupBy

In [11]:

```
grupo_bairros.groups
```

Out[11]:

```
20, 4528, 4534, 4555, 4574, 4578, 4613, 4715, 4736, 4950, ...], 'Cascadur  
a': [558, 879, 1625, 1925, 4452, 5816, 6283, 6614, 8551, 9319, 10272, 1084  
4, 11782, 11842, 11998, 12250, 12723, 14902, 15569, 16030, 16035, 16075, 1  
6737, 17067, 17580, 19602, 19606, 20057, 20095, 20268, 21751], 'Catete':  
[94, 98, 120, 193, 300, 339, 682, 1174, 1351, 1402, 1692, 1798, 1824, 190  
3, 2011, 2125, 2286, 2397, 2423, 2930, 2931, 2987, 3244, 3529, 3673, 3708,  
3758, 3814, 3920, 4120, 4215, 4219, 4291, 4323, 4407, 4770, 4817, 5477, 56  
12, 5617, 5635, 5707, 5710, 5727, 5889, 5948, 6160, 6250, 6599, 6888, 695  
7, 7018, 7020, 7050, 7073, 7162, 7345, 7386, 7858, 8435, 8652, 8776, 8876,  
8942, 9089, 9284, 9327, 9349, 9451, 9566, 9833, 9854, 9978, 10134, 10172,  
10241, 10249, 10276, 10366, 10557, 10832, 10992, 11166, 11501, 11550, 1157  
3, 11615, 11730, 11740, 11789, 11907, 12013, 12206, 12279, 12289, 12348, 1  
2603, 12654, 13104, 13358, ...], 'Catumbi': [1343, 3152, 4506, 8394, 8769,  
16685, 17119, 21002], 'Cavalcanti': [15277, 18152], 'Centro': [2, 38, 65,  
71, 146, 181, 191, 312, 427, 585, 665, 826, 870, 939, 944, 954, 1016, 111  
1, 1115, 1136, 1321, 1385, 1483, 1496, 1613, 1633, 1691, 1697, 1733, 1737,  
1845, 1891, 1901, 2006, 2085, 2098, 2156, 2167, 2191, 2236, 2239, 2258, 23
```

In [12]:

```
for bairros, data in grupo_bairros:
    print('{} -> {}'.format(bairros, data['Valor'].mean()))
    #print(type(dados))
```

```
Abolição -> 1195.3333333333333
Alto da Boa Vista -> 3966.6666666666665
Anchieta -> 875.0
Andaraí -> 1464.7113402061855
Anil -> 2048.8732394366198
Arpoador -> 12923.916666666666
Bancários -> 1825.0
Bangu -> 1016.0
Barra da Tijuca -> 7069.552938130986
Barra de Guaratiba -> 5550.0
Benfica -> 996.0
Bento Ribeiro -> 1030.8695652173913
Bonsucesso -> 1225.9322033898304
Botafogo -> 8791.828178694159
Braz de Pina -> 1115.0
Cachambi -> 1157.1742424242425
Cachamorra -> 3000.0
Cacuia -> 916.6666666666666
Caiu -> 850.0
```

In [13]:

```
grupo_bairros[['Valor', 'Condominio']].mean().round(2)
```

Out[13]:

	Valor	Condominio
Bairro		
<b>Abolição</b>	1195.33	191.40
<b>Alto da Boa Vista</b>	3966.67	885.00
<b>Anchieta</b>	875.00	19.75
<b>Andaraí</b>	1464.71	497.71
<b>Anil</b>	2048.87	455.85
...	...	...
<b>Vila Valqueire</b>	1769.58	276.88
<b>Vila da Penha</b>	1260.58	232.68
<b>Vista Alegre</b>	1114.38	210.06
<b>Zumbi</b>	2150.00	1050.00
<b>Água Santa</b>	861.11	269.44

152 rows × 2 columns

## Estatísticas Descritivas

In [14]:



```
grupo_bairros['Valor'].describe().round(2)
```

Out[14]:

	count	mean	std	min	25%	50%	75%	max
Bairro								
Abolição	15.0	1195.33	425.32	800.0	900.0	1180.0	1200.0	2300.0
Alto da Boa Vista	6.0	3966.67	5513.59	600.0	1150.0	1725.0	3275.0	15000.0
Anchieta	4.0	875.00	132.29	700.0	812.5	900.0	962.5	1000.0
Andaraí	97.0	1464.71	408.11	700.0	1200.0	1400.0	1650.0	2950.0
Anil	71.0	2048.87	1523.97	300.0	1100.0	1500.0	2675.0	7000.0
...	...	...	...	...	...	...	...	...
Vila Valqueire	48.0	1769.58	2203.01	800.0	1000.0	1250.0	1700.0	16000.0
Vila da Penha	104.0	1260.58	644.66	450.0	900.0	1100.0	1302.5	4500.0
Vista Alegre	16.0	1114.38	199.60	700.0	1012.5	1200.0	1225.0	1400.0
Zumbi	2.0	2150.00	777.82	1600.0	1875.0	2150.0	2425.0	2700.0
Água Santa	9.0	861.11	174.60	600.0	800.0	850.0	950.0	1200.0

152 rows × 8 columns

In [15]:

```
gurpo_bairros['Valor'].aggregate(['min', 'max', 'sum'])
```

Out[15]:

	min	max	sum
Bairro			
Abolição	800.0	2300.0	17930.0
Alto da Boa Vista	600.0	15000.0	23800.0
Anchieta	700.0	1000.0	3500.0
Andaraí	700.0	2950.0	142077.0
Anil	300.0	7000.0	145470.0
...	...	...	...
Vila Valqueire	800.0	16000.0	84940.0
Vila da Penha	450.0	4500.0	131100.0
Vista Alegre	700.0	1400.0	17830.0
Zumbi	1600.0	2700.0	4300.0
Água Santa	600.0	1200.0	7750.0

152 rows × 3 columns

In [16]:

```
gurpo_bairros['Valor'].aggregate(['min', 'max', 'sum']).rename(columns = {'min':'Mínimo', 'ma
```

Out[16]:

	Mínimo	Máximo	Soma
Bairro			
Abolição	800.0	2300.0	17930.0
Alto da Boa Vista	600.0	15000.0	23800.0
Anchieta	700.0	1000.0	3500.0
Andaraí	700.0	2950.0	142077.0
Anil	300.0	7000.0	145470.0
...	...	...	...
Vila Valqueire	800.0	16000.0	84940.0
Vila da Penha	450.0	4500.0	131100.0
Vista Alegre	700.0	1400.0	17830.0
Zumbi	1600.0	2700.0	4300.0
Água Santa	600.0	1200.0	7750.0

152 rows × 3 columns

In [17]:

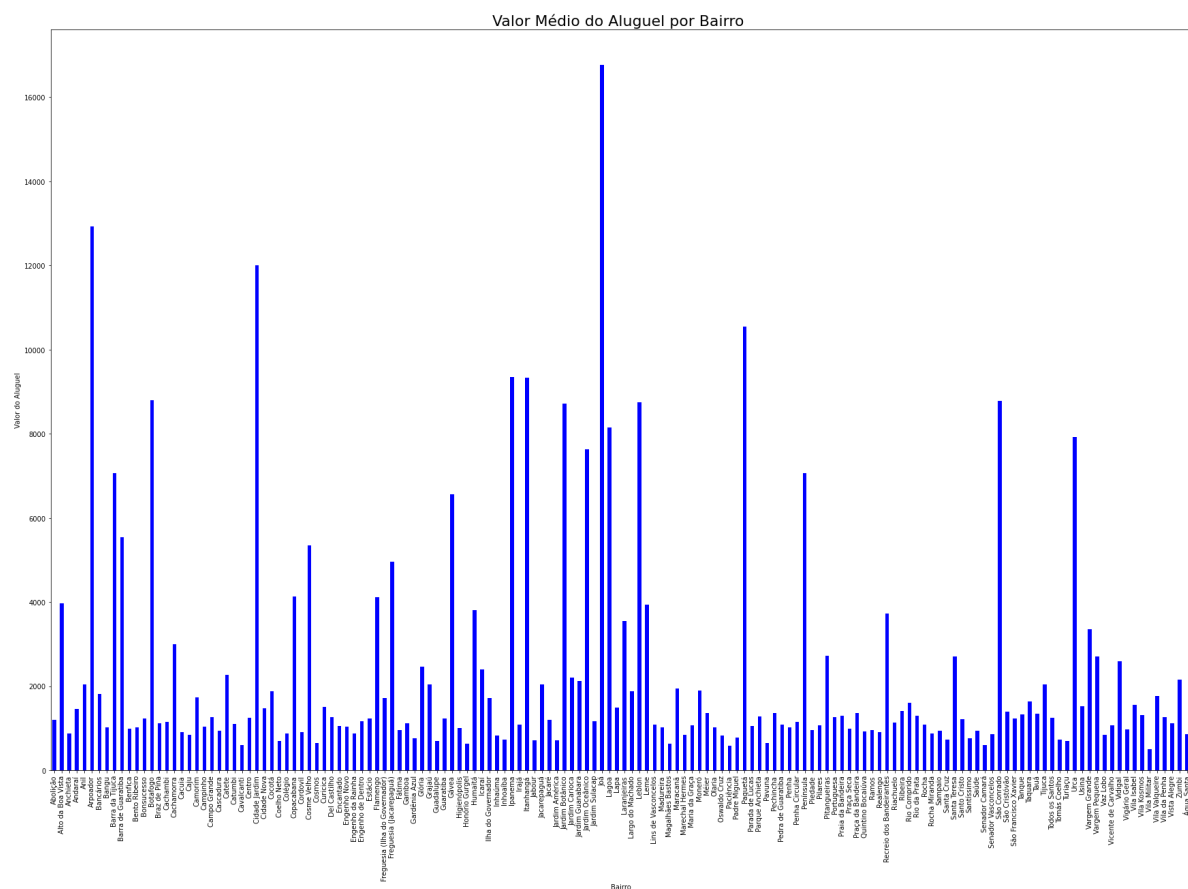
```
%matplotlib inline
import matplotlib.pyplot as plt
plt.rc('figure', figsize = (30, 20))
```

In [21]:

```
fig = gurpo_bairros['Valor'].mean().plot.bar(color = 'blue')
fig.set_ylabel('Valor do Aluguel')
fig.set_title('Valor Médio do Aluguel por Bairro ', {'fontsize':22})
```

Out[21]:

```
Text(0.5, 1.0, 'Valor Médio do Aluguel por Bairro ')
```





```
Text(0.5, 1.0, 'Valor Médio do Aluguel por Bairro ')
```

