

Relatório de Análise III

Imovéis Residenciais

In [2]:

importando Pandas
import pandas as pd

In [3]:

importando a base de dados
dados = pd.read_csv('data/aluguel.csv', sep = ';')

In [5]:

exibindo os primeiros dez itens
dados.head(10)

Out[5]:

	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	NaN	NaN
2	Conjunto Comercial/Sala	Barra da Tijuca	0	4	0	150	5200.0	4020.0	1111.0
3	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0
4	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	NaN
5	Apartamento	Vista Alegre	3	1	0	70	1200.0	NaN	NaN
6	Apartamento	Cachambi	2	0	0	50	1300.0	301.0	17.0
7	Casa de Condomínio	Barra da Tijuca	5	4	5	750	22000.0	NaN	NaN
8	Casa de Condomínio	Ramos	2	2	0	65	1000.0	NaN	NaN
9	Conjunto Comercial/Sala	Centro	0	3	0	695	35000.0	19193.0	3030.0

In [6]:

criando uma lista com todos os tipos de imóveis, sem repetição
list(dados['Tipo'].drop_duplicates())

Out[6]:

['Quitinete',
'Casa',
'Conjunto Comercial/Sala',
'Apartamento',
'Casa de Condomínio',
'Prédio Inteiro',
'Flat',
'Loja/Salão',
'Galpão/Depósito/Armazém',
'Casa Comercial',
'Casa de Vila',
'Terreno Padrão',
'Box/Garagem',
'Loft',
'Loja Shopping/ Ct Comercial',
'Chácara',
'Loteamento/Condomínio',
'Sítio',
'Pousada/Chalé',
'Studio',
'Hotel',
'Indústria']

In [9]:

selecionando apenas os imóveis que classificamos como residenciais
residencial = ['Quitinete',
'Casa',
'Apartamento',
'Casa de Condomínio',
'Casa de Vila']

In [10]:

saída
residencial

Out[10]:

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

In [14]:

verifica dentro do DataFrame se os elementos da nossa lista existem de fato
selecao = dados['Tipo'].isin(residencial)
selecao

Out[14]:

0 True
1 True
2 False
3 True
4 True
...
32955 True
32956 True
32957 True
32958 True
32959 False
Name: Tipo, Length: 32960, dtype: bool

In [11]:

exibindo os primeiros dez itens para comparar com a última saída
dados.head(10)

Out[11]:

	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	NaN	NaN
2	Conjunto Comercial/Sala	Barra da Tijuca	0	4	0	150	5200.0	4020.0	1111.0
3	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0
4	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	NaN
5	Apartamento	Vista Alegre	3	1	0	70	1200.0	NaN	NaN
6	Apartamento	Cachambi	2	0	0	50	1300.0	301.0	17.0
7	Casa de Condomínio	Barra da Tijuca	5	4	5	750	22000.0	NaN	NaN
8	Casa de Condomínio	Ramos	2	2	0	65	1000.0	NaN	NaN
9	Conjunto Comercial/Sala	Centro	0	3	0	695	35000.0	19193.0	3030.0

In [16]:

atribui a nossa lista apenas registros com marcação = True
dados_residencial = dados[selecao]

In [17]:

saída
dados_residencial

Out[17]:

	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	NaN	NaN
3	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0
4	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	NaN
5	Apartamento	Vista Alegre	3	1	0	70	1200.0	NaN	NaN
...
32953	Apartamento	Méier	2	0	0	70	900.0	490.0	48.0
32955	Quitinete	Centro	0	0	0	27	800.0	350.0	25.0
32956	Apartamento	Jacarepaguá	3	1	2	78	1800.0	800.0	40.0
32957	Apartamento	São Francisco Xavier	2	1	0	48	1400.0	509.0	37.0
32958	Apartamento	Leblon	2	0	0	70	3000.0	760.0	NaN

22580 rows × 9 columns

In [18]:

criando uma lista com apenas os imóveis residenciais, sem repetição
list(dados_residencial['Tipo'].drop_duplicates())

Out[18]:

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

In [19]:

conferindo o tamanho do DataFrame com os dados residenciais
dados_residencial.shape[0]

Out[19]:

22580

In [20]:

conferindo o tamanho do DataFrame com todos os dados de imóveis
dados.shape[0]

Out[20]:

32960

In [21]:

reconstrução do índice
dados_residencial.index = range(dados_residencial.shape[0])

In [22]:

saída
dados_residencial

Out[22]:

	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	NaN	NaN
2	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0
3	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	NaN
4	Apartamento	Vista Alegre	3	1	0	70	1200.0	NaN	NaN
...
22575	Apartamento	Méier	2	0	0	70	900.0	490.0	48.0
22576	Quitinete	Centro	0	0	0	27	800.0	350.0	25.0
22577	Apartamento	Jacarepaguá	3	1	2	78	1800.0	800.0	40.0
22578	Apartamento	São Francisco Xavier	2	1	0	48	1400.0	509.0	37.0
22579	Apartamento	Leblon	2	0	0	70	3000.0	760.0	NaN

22580 rows × 9 columns

Exportando a Base de Dados

In [26]:

método para exportar com formato '.csv'
dados_residencial.to_csv('data/aluguel_residencial.csv', sep = ';')

In [27]:

importando novamente nossa base de dados
dados_residencial_2 = pd.read_csv('data/aluguel_residencial.csv', sep = ';')

In [28]:

saída dos dados com o índice incorreto
dados_residencial_2

Out[28]:

	Unnamed: 0	Tipo	Bairro	Quartos	Vagas	Suites	Area	Valor	Condominio	IPTU
0	0	Quitinete	Copacabana	1	0	0	40	1700.0	500.0	60.0
1	1	Casa	Jardim Botânica	2	0	1	100	7000.0	NaN	NaN
2	2	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0
3	3	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	NaN
4	4	Apartamento	Vista Alegre	3	1	0	70	1200.0	NaN	NaN
...
22575	22575	Apartamento	Méier	2	0	0	70	900.0	490.0	48.0
22576	22576	Quitinete	Centro	0	0	0	27	800.0	350.0	25.0
22577	22577	Apartamento	Jacarepaguá	3	1	2	78	1800.0	800.0	40.0
22578	22578	Apartamento	São Francisco Xavier	2	1	0	48	1400.0	509.0	37.0
22579	22579	Apartamento	Leblon	2	0	0	70	3000.0	760.0	NaN

22580 rows × 10 columns

In [29]:

método para exportar com o índice correto
dados_residencial.to_csv('data/aluguel_residencial.csv', sep = ';', index = False)

In [30]:

importando novamente para conferir
dados_residencial_2 = pd.read_csv('data/aluguel_residencial.csv', sep = ';')

In [31]:

saída
dados_residencial_2

Out[31]:

	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	NaN	NaN
2	Apartamento	Centro	1	0	0	15	800.0	390.0	20.0
3	Apartamento	Higienópolis	1	0	0	48	800.0	230.0	NaN
4	Apartamento	Vista Alegre	3	1	0	70	1200.0	NaN	NaN
...
22575	Apartamento	Méier	2	0	0	70	900.0	490.0	48.0
22576	Quitinete	Centro	0	0	0	27	800.0	350.0	25.0
22577	Apartamento	Jacarepaguá	3	1	2	78	1800.0	800.0	40.0
22578	Apartamento	São Francisco Xavier	2	1	0	48	1400.0	509.0	37.0
22579	Apartamento	Leblon	2	0	0	70	3000.0	760.0	NaN

22580 rows × 9 columns