## 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]: **IPTU** Tipo Valor Condominio Bairro Quartos Vagas Suites Area 0 1700.0 60.0 Quitinete 40 500.0 Copacabana 0 1 Casa Jardim Botânico 0 1 100 7000.0 NaN NaN 2 Conjunto Comercial/Sala 5200.0 4020.0 1111.0 Barra da Tijuca 0 150 3 Apartamento 0 0 15 0.008 390.0 20.0 4 0 0 48 800.0 230.0 NaN Apartamento Higienópolis 5 Apartamento Vista Alegre 1 0 70 1200.0 NaN NaN 0 1300.0 6 Apartamento Cachambi 0 50 301.0 17.0 7 Casa de Condomínio Barra da Tijuca 5 750 22000.0 NaN NaN 8 Casa de Condomínio Ramos 2 65 1000.0 NaN 0 NaN 9 Conjunto Comercial/Sala Centro 3 695 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 True 2 False 3 True 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]: Bairro Quartos Vagas Suites Area **IPTU** Tipo Valor Condominio 0 Quitinete Copacabana 40 1700.0 500.0 60.0 1 7000.0 Casa Jardim Botânico 0 1 100 NaN NaN 5200.0 2 Conjunto Comercial/Sala Barra da Tijuca 0 150 4020.0 1111.0 0 3 0.008 390.0 20.0 Apartamento Centro 1 0 15 Apartamento Higienópolis 0 48 800.0 230.0 NaN 5 1 1200.0 NaN Apartamento Vista Alegre 0 70 NaN 0 1300.0 6 Cachambi 0 50 301.0 17.0 Apartamento 4 22000.0 7 Casa de Condomínio Barra da Tijuca 5 750 NaN NaN 1000.0 Casa de Condomínio Ramos 65 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]: Valor Condominio IPTU Tipo Bairro Quartos Vagas Suites Area 0 Quitinete Copacabana 0 40 1700.0 500.0 60.0 1 Casa Jardim Botânico 0 1 100 7000.0 NaN NaN 3 Apartamento Centro 0 15 800.0 390.0 20.0 Higienópolis 4 Apartamento 0 0 48 800.0 230.0 NaN 70 1200.0 NaN NaN **5** Apartamento Vista Alegre 32953 Apartamento Méier 70 900.0 490.0 48.0 0 27 800.0 350.0 25.0 32955 Quitinete Centro Jacarepaguá 78 1800.0 800.0 40.0 32956 Apartamento 48 1400.0 32957 Apartamento São Francisco Xavier 1 509.0 37.0 2 0 32958 Apartamento Leblon 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 500.0 60.0 Quitinete Copacabana 40 1700.0 Jardim Botânico 1 100 7000.0 1 Casa 0 NaN NaN 2 Apartamento 800.0 390.0 20.0 Centro 15 3 Apartamento Higienópolis 0 800.0 48 230.0 NaN Vista Alegre 70 1200.0 4 Apartamento NaN NaN Méier 22575 Apartamento 900.0 490.0 48.0 0 70 350.0 25.0 22576 Quitinete Centro 27 800.0 **22577** Apartamento Jacarepaguá 78 1800.0 800.0 40.0 **22578** Apartamento São Francisco Xavier 1 48 1400.0 509.0 37.0 22579 Apartamento Leblon 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]: Bairro Quartos Vagas Suites Area Valor Condominio IPTU Unnamed: 0 Tipo 0 Quitinete Copacabana 40 1700.0 500.0 60.0 Casa Jardim Botânico 0 1 100 7000.0 1 1 NaN NaN 2 Apartamento Centro 15 800.0 390.0 20.0 0 230.0 4 Apartamento Vista Alegre 70 1200.0 NaN NaN ... 22575 22575 Apartamento Méier 70 900.0 490.0 48.0 22576 22576 0 27 800.0 Quitinete Centro 0 350.0 25.0 22577 22577 Apartamento Jacarepaguá 78 1800.0 0.008 40.0 22578 22578 Apartamento São Francisco Xavier 0 48 1400.0 509.0 37.0 1 22579 22579 Apartamento Leblon 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]: Bairro Quartos Vagas Suites Area Valor Condominio IPTU Tipo 500.0 60.0 Quitinete Copacabana 40 1700.0 1 Jardim Botânico 1 100 7000.0 Casa NaN NaN 2 Apartamento Centro 15 800.0 390.0 20.0 Higienópolis 0 48 800.0 230.0 NaN 3 Apartamento 0

22575 Apartamento Méier 0 70 900.0 490.0 48.0 22576 Centro 0 27 800.0 350.0 25.0 Quitinete 78 1800.0 22577 Apartamento Jacarepaguá 800.0 40.0 22578 Apartamento São Francisco Xavier 48 1400.0 509.0 37.0 1 22579 Apartamento Leblon 70 3000.0 760.0 NaN 22580 rows × 9 columns

Vista Alegre

4 Apartamento

70 1200.0

NaN NaN