

ciclo1

April 18, 2024

1 Imports

```
/home/arthur/anaconda3/envs/busca-apartamentos-env/lib/python3.11/site-  
packages/tqdm/auto.py:21: TqdmWarning: IProgress not found. Please update  
jupyter and ipywidgets. See  
https://ipywidgets.readthedocs.io/en/stable/user\_install.html  
from .autonotebook import tqdm as notebook_tqdm
```

1.1 Funções Auxiliares

2 Carregar Dados

3 Análise Descritiva dos Dados

- Quantidade de Dados
- Quantidade de Dados Faltantes
- Tipo de Variáveis
- Estatística Descritiva

3.1 Dimensão dos Dados

Total de Linhas 878

Total de Colunas 22

3.2 Tratando NAs

site	0
titulo	0
link	0
data_coleta	0
endereco	0
bairro	1
cidade	0
aluguel	0
condominio	0
seguro_incendio	0
iptu	0
area	0

quartos	0
suites	0
banheiros	0
vagas_garagem	0
mobiliado	0
piscina	0
academia	0
sacada	0
churrasqueira	0
salao_de_festas	0
dtype:	int64

3.2.1 Removendo NAs

site	0
titulo	0
link	0
data_coleta	0
endereco	0
bairro	0
cidade	0
aluguel	0
condominio	0
seguro_incendio	0
iptu	0
area	0
quartos	0
suites	0
banheiros	0
vagas_garagem	0
mobiliado	0
piscina	0
academia	0
sacada	0
churrasqueira	0
salao_de_festas	0
dtype:	int64

3.3 Tipos de Variáveis

site	object
titulo	object
link	object
data_coleta	object
endereco	object
bairro	object

cidade	object
aluguel	float64
condominio	float64
seguro_incendio	float64
iptu	float64
area	float64
quartos	int64
suites	int64
banheiros	int64
vagas_garagem	int64
mobiliado	object
piscina	object
academia	object
sacada	object
churrasqueira	object
salao_de_festas	object
dtype:	object

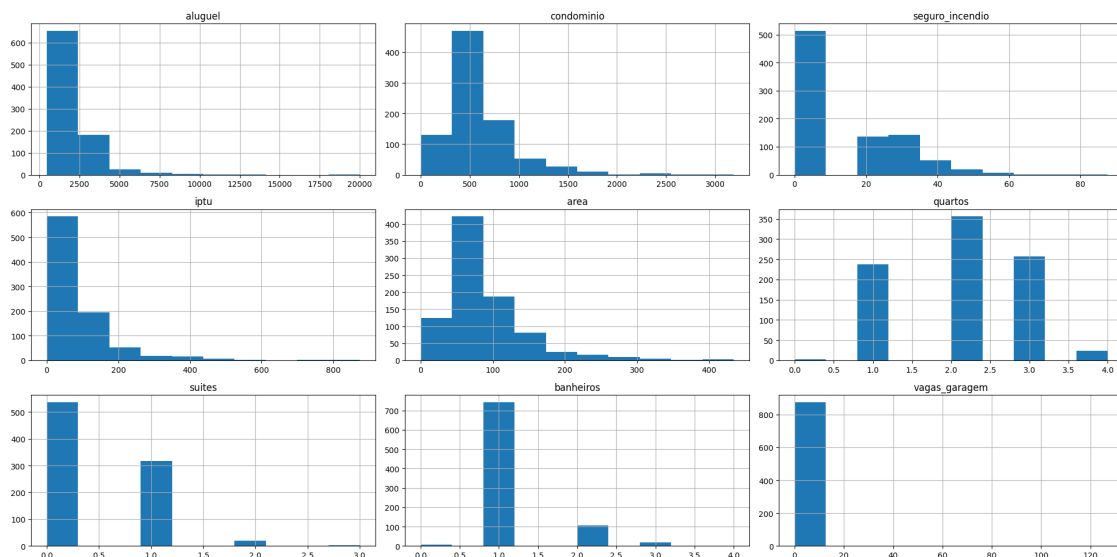
site	object
titulo	object
link	object
data_coleta	datetime64[ns]
endereco	object
bairro	object
cidade	object
aluguel	float64
condominio	float64
seguro_incendio	float64
iptu	float64
area	float64
quartos	int64
suites	int64
banheiros	int64
vagas_garagem	int64
mobiliado	object
piscina	object
academia	object
sacada	object
churrasqueira	object
salao_de_festas	object
dtype:	object

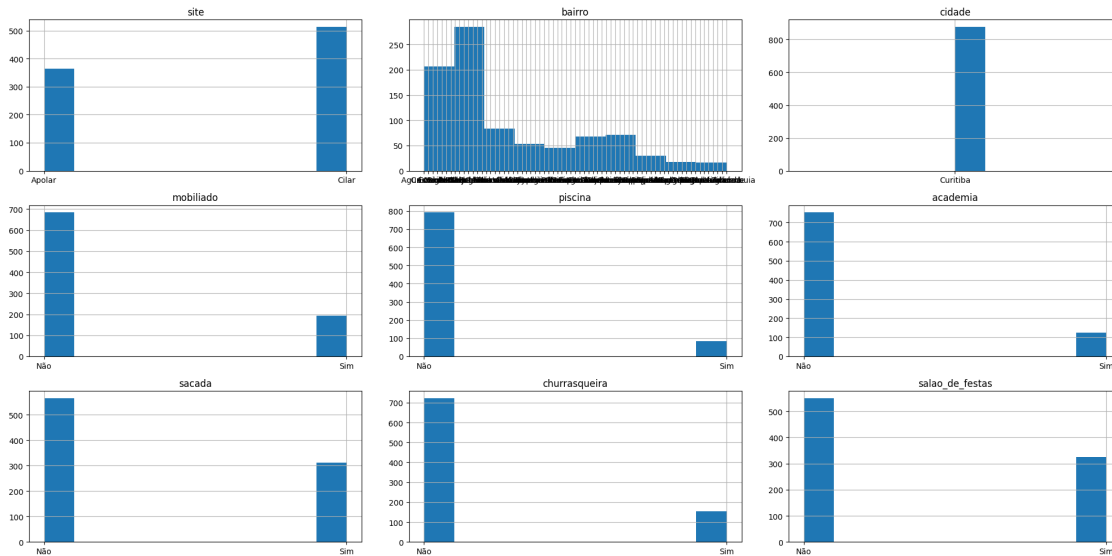
3.4 Distribuição dos Dados

	aluguel	condominio	seguro_incendio	iptu	area	quartos	suites	\
0	20000.0	0.00	62.56	0.00	203.0	0	0	
1	12000.0	840.00	71.93	0.00	250.0	3	1	
2	10500.0	1549.67	87.58	21.37	330.0	4	1	
3	7000.0	1690.58	53.19	341.74	163.0	2	1	
4	5000.0	635.61	39.14	230.98	107.0	2	1	
..	
873	850.0	320.00	0.00	42.25	52.0	2	0	
874	850.0	465.00	0.00	52.99	52.0	2	0	
875	800.0	382.00	0.00	51.19	48.0	1	0	
876	750.0	526.74	0.00	20.62	29.0	1	0	
877	450.0	450.00	0.00	89.97	51.0	1	0	

	banheiros	vagas_garagem
0	0	0
1	1	3
2	1	2
3	2	2
4	4	2
..
873	1	1
874	1	0
875	1	0
876	1	0
877	1	0

[877 rows x 9 columns]





Notamos aqui que possuímos valores extremos de 'aluguel' e 'vagas_garagem' que indicam a presença de outliers

Vamos investigar esses valores:

3.5 Presença de Outliers

	attributes	min	max	range	mean	median	\
0	aluguel	450.0	20000.00	19550.00	2144.564424	1800.00	
1	condominio	0.0	3189.70	3189.70	575.287320	500.00	
2	seguro_incendio	0.0	87.58	87.58	13.333136	0.00	
3	iptu	0.0	874.00	874.00	86.694629	59.61	
4	area	0.0	434.00	434.00	87.636260	72.00	
5	quartos	0.0	4.00	4.00	2.070696	2.00	
6	suites	0.0	3.00	3.00	0.416192	0.00	
7	banheiros	0.0	4.00	4.00	1.157355	1.00	
8	vagas_garagem	0.0	127.00	127.00	1.020525	1.00	
	std	skew	kurtosis				
0	1447.917203	4.798805	38.898546				
1	377.009722	2.031490	8.114300				
2	16.716970	0.770463	-0.436915				
3	101.498083	3.106794	14.624240				
4	59.153965	2.007615	6.006188				
5	0.825490	0.123748	-0.810291				
6	0.553923	1.020143	0.720944				
7	0.448340	2.400052	6.982625				
8	4.302638	28.670098	839.820872				

Além dos valores extremos em 'aluguel' e 'vagas_garagem', percebemos tbm imóveis com área, quartos e banheiros iguais a zero

Vamos investigar esses valores:

	site	titulo	\
0	Apolar	Apartamento Residencial para Locação no Água V...	
1	Apolar	Apartamento Residencial para Locação no Cristo...	
364	Cilar	Apartamento no 15º andar com 3 quartos	
416	Cilar	Apartamento no 16º andar com 3 quartos	

	link	data_coleta	\
0	https://www.apolar.com.br/alugar/curitiba/agua...	2024-03-26	
1	https://www.apolar.com.br/alugar/curitiba/cris...	2024-03-26	
364	https://cilar.com.br/alugar/apartamento-no-15-...	2024-03-26	
416	https://cilar.com.br/alugar/apartamento-no-16-...	2024-03-26	

	endereco	bairro	cidade	\
0	Avenida Silva Jardim, 2733, Água Verde - Curitiba	Agua verde	Curitiba	
1	Rua Urbano Lopes, 238, Cristo Rei - Curitiba	Cristo rei	Curitiba	
364	Rua Professor Pedro Viriato Parigot Souza, 306...	Ecoville	Curitiba	
416	Rua Deputado Heitor Alencar Furtado, 3180 - Ec...	Ecoville	Curitiba	

	aluguel	condominio	seguro_incendio	iptu	area	quartos	suites	\
0	20000.0	0.0	62.56	0.00	203.0	0	0	
1	12000.0	840.0	71.93	0.00	250.0	3	1	
364	13000.0	3189.7	0.00	874.00	421.0	3	3	
416	12800.0	1700.0	0.00	605.54	294.0	3	3	

	banheiros	vagas_garagem	mobiliado	piscina	academia	sacada	churrasqueira	\
0	0	0	Não	Não	Não	Não	Não	
1	1	3	Sim	Não	Sim	Sim	Sim	
364	1	4	Não	Sim	Sim	Sim	Não	
416	1	3	Sim	Sim	Não	Não	Não	

	salao_de_festas
0	Não
1	Sim
364	Sim
416	Não

	site	titulo	\
580	Cilar	Apartamento no 2º andar com 2 quartos	

	link	data_coleta	\
580	https://cilar.com.br/alugar/apartamento-no-2-a...	2024-03-26	

	endereco	bairro	cidade	\
--	----------	--------	--------	---

580	Professor Pedro Viriato Parigot de Souza, 4554...	Ecoville	Curitiba
	aluguel	condominio	seguro_incendio
580	2400.0	550.0	0.0
	iptu	area	quartos
	0.0	90.0	2
	suites	\	
	2	2	
	banheiros	vagas_garagem	mobiliado
580	0	127	Não
			Não
			Não
			Não
			Sim
	salao_de_festas		
580	Sim		

	titulo	\
177	Apartamento Residencial para Locação no Água V...	
459	Apartamento no 13º andar com 3 quartos	
480	Apartamento no 4º andar com 4 quartos	
487	Apartamento no 5º andar com 3 quartos	
493	Apartamento no 4º andar com 3 quartos	

	endereco	bairro	\
177	Rua Palmeiras, 331, Água Verde - Curitiba	Agua verde	
459	Rua Dona Alice Tibiriçá, 629 - Bigorriho - Cu...	Bigorriho	
480	Rua Lamenha Lins, 585 - Centro - Curitiba	Centro	
487	Rua Luiz Ronaldo Canalli, 3175 - Campo Comprid...	Campo comprido	
493	Rua Luiz Ronaldo Canalli, 3175 - Campo Comprid...	Campo comprido	

	aluguel	condominio	seguro_incendio	iptu	area	quartos	banheiros	\
177	1580.0	410.0	26.26	98.55	0.0	2	1	
459	3900.0	1233.9	0.00	170.95	0.0	3	0	
480	3490.0	1301.0	0.00	0.00	0.0	4	3	
487	3200.0	400.0	0.00	0.00	0.0	3	1	
493	3100.0	400.0	0.00	0.00	0.0	3	1	

	suites	vagas_garagem
177	0	1
459	1	1
480	1	1
487	1	1
493	1	1

	site	titulo	\
0	Apolar	Apartamento Residencial para Locação no Água V...	
494	Cilar	Apartamento no Centro	
614	Cilar	Apartamento no BATEL	

	link_data_coleta	\
0	https://www.apolar.com.br/alugar/curitiba/agua...	2024-03-26
494	https://cilar.com.br/alugar/apartamento-no-cen...	2024-03-26

614 https://cilar.com.br/alugar/apartamento-no-bat... 2024-03-26

	endereco	bairro	cidade	\
0	Avenida Silva Jardim, 2733, Água Verde - Curitiba	Agua verde	Curitiba	
494	Avenida Visconde de Guarapuava, 3806 - Centro ...	Centro	Curitiba	
614	R VICENTE MACHADO, 1059 - BATEL - Curitiba	Batel	Curitiba	

	aluguel	condominio	seguro_incendio	iptu	area	quartos	suites	\
0	20000.0	0.0	62.56	0.00	203.0	0	0	
494	3100.0	637.0	0.00	0.00	0.0	0	0	
614	2250.0	0.0	0.00	190.36	152.0	0	0	

	banheiros	vagas_garagem	mobiliado	piscina	academia	sacada	churrasqueira	\
0	0	0	Não	Não	Não	Não	Não	
494	1	1	Sim	Não	Não	Sim	Não	
614	1	0	Não	Não	Não	Sim	Não	

	salao_de_festas
0	Não
494	Não
614	Não

	site	titulo	\
0	Apolar	Apartamento Residencial para Locação no Água V...	
374	Cilar	Apartamento com 1 quarto	
458	Cilar	Apartamento no 6º andar com 2 quartos	
459	Cilar	Apartamento no 13º andar com 3 quartos	
580	Cilar	Apartamento no 2º andar com 2 quartos	
649	Cilar	Apartamento com 2 quartos	
716	Cilar	Apartamento com 1 quarto	
808	Cilar	Apartamento no 1º andar com 3 quartos	

	link	data_coleta	\
0	https://www.apolar.com.br/alugar/curitiba/agua...	2024-03-26	
374	https://cilar.com.br/alugar/apartamento-com-1-...	2024-03-26	
458	https://cilar.com.br/alugar/apartamento-no-6-a...	2024-03-26	
459	https://cilar.com.br/alugar/apartamento-no-13-...	2024-03-26	
580	https://cilar.com.br/alugar/apartamento-no-2-a...	2024-03-26	
649	https://cilar.com.br/alugar/apartamento-com-2-...	2024-03-26	
716	https://cilar.com.br/alugar/apartamento-com-1-...	2024-03-26	
808	https://cilar.com.br/alugar/apartamento-no-1-a...	2024-03-26	

	endereco	bairro	\
0	Avenida Silva Jardim, 2733, Água Verde - Curitiba	Agua verde	
374	Rua Desembargador Motta, 2350 - Centro - Curitiba	Centro	
458	Rua Professor Arthur Loyola, 387 - Cabral - Cu...	Cabral	
459	Rua Dona Alice Tibiriçá, 629 - Bigorrrilho - Cu...	Bigorrrilho	

580	Professor Pedro Viriato Parigot de Souza, 4554...	Ecoville
649	Nilo Peçanha, 115 - São Francisco - Curitiba	Sao francisco
716	Rua Prefeito Ângelo Ferrário Lopes, 1737 - Hug...	Hugo lange
808	Rua José Fernandes Maldonado, 900 - Bairro Alt...	Bairro alto

	cidade	aluguel	condominio	seguro_incendio	iptu	area	quartos	\
0	Curitiba	20000.0	0.00	62.56	0.00	203.0	0	
374	Curitiba	2500.0	751.63	0.00	118.00	99.0	1	
458	Curitiba	3950.0	450.00	0.00	360.00	103.0	2	
459	Curitiba	3900.0	1233.90	0.00	170.95	0.0	3	
580	Curitiba	2400.0	550.00	0.00	0.00	90.0	2	
649	Curitiba	1980.0	250.00	0.00	0.00	70.0	2	
716	Curitiba	1700.0	130.00	0.00	60.00	77.0	1	
808	Curitiba	1300.0	526.00	0.00	53.72	0.0	3	

	suites	banheiros	vagas_garagem	mobiliado	piscina	academia	sacada	\
0	0	0	0	Não	Não	Não	Não	
374	1	0	1	Não	Não	Não	Sim	
458	2	0	1	Não	Não	Não	Não	
459	1	0	1	Não	Não	Não	Sim	
580	2	0	127	Não	Não	Não	Não	
649	1	0	1	Não	Não	Não	Não	
716	1	0	1	Não	Não	Não	Não	
808	1	0	1	Não	Não	Não	Sim	

	churrasqueira	salao_de_festas
0	Não	Não
374	Não	Não
458	Não	Não
459	Não	Não
580	Sim	Sim
649	Sim	Sim
716	Não	Não
808	Não	Não

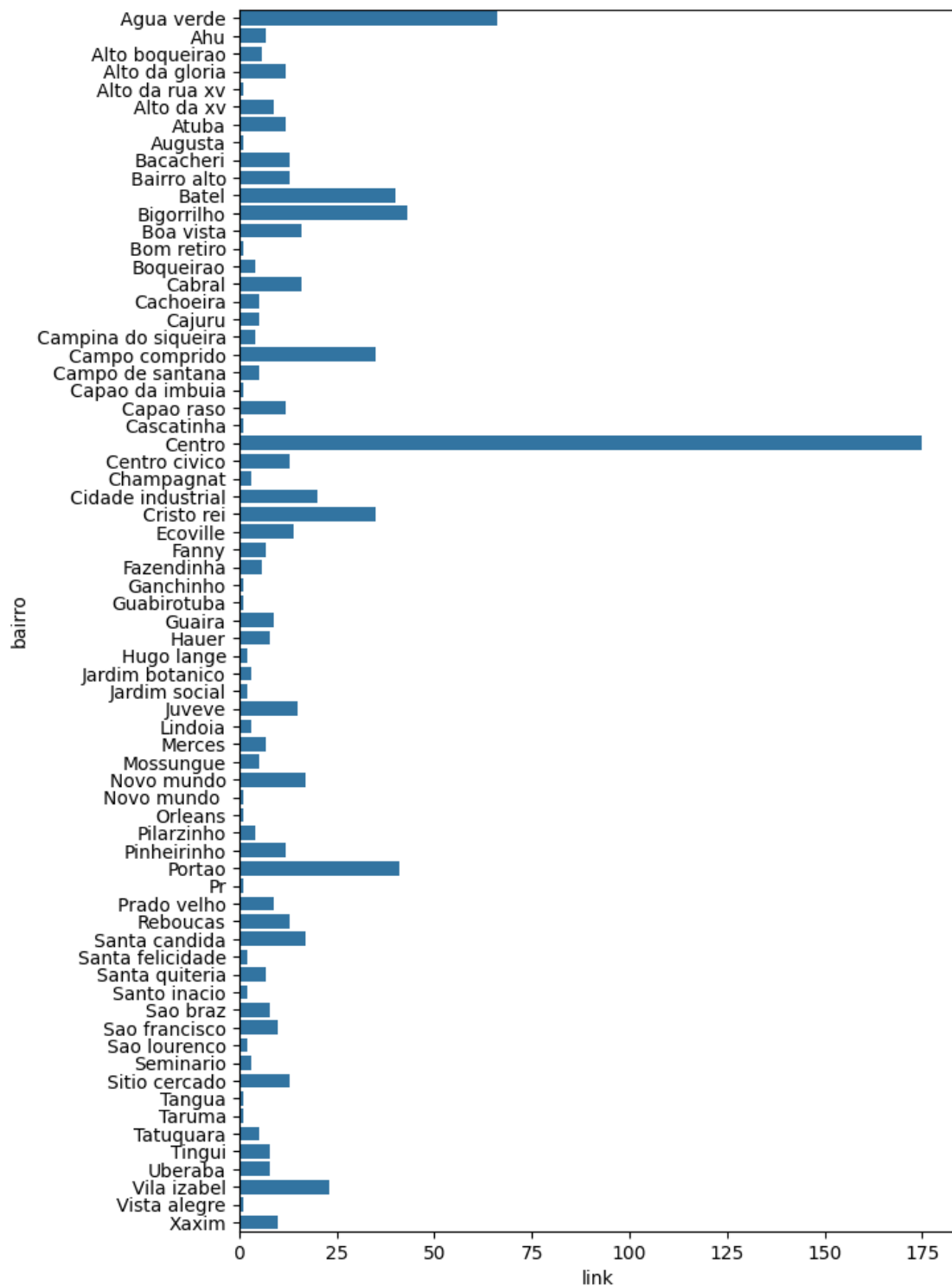
Sobre o valor dos alugueis, fica claro que tratam-se de imóveis caros mas não de erros, então eles serão mantidos no dataset

Já o número de 'vagas_garagem' trata-se de um valor absurdo o que indica muito provavelmente um erro, logo esse imóvel será removido

Quanto aos imóveis com área igual a 0, vemos que os demais valores do imóvel fazem sentido portanto, concluí-se que esse dado de área está errado e esses imóveis serão removidos do dataset

Os imóveis com quantidade de quartos e banheiros iguais a zero tratam-se de erros, portanto também serão removidos

```
<Axes: xlabel='link', ylabel='bairro'>
```



Analisando a variável ‘bairro’ notamos alguns nomes repetidos e bairros escritos nomes populares diferentes ou apelidos (assim como Ecoville e Mossungue), logo o nome desses bairros será padronizado

4 Filtrando Variáveis

4.1 Removendo Erros do dataset

5 Feature Engineering

5.0.1 Info dos bairros

Informações coletadas da wikipedia sobre os bairros de curitiba

5.0.2 Valor Total

5.0.3 Preço m²

5.0.4 Banheiros

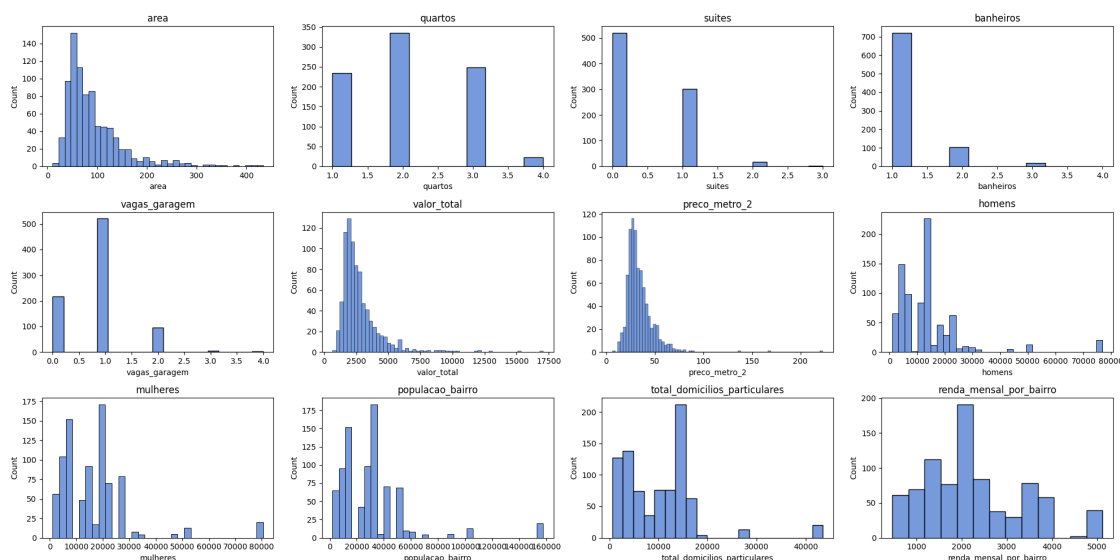
6 Análise Exploratória de Dados

- Neste tópico, vamos entender a distribuição e a relação entre as variáveis

6.1 Análise Univariada

- Entender a distribuição da variável em relação à ela mesma

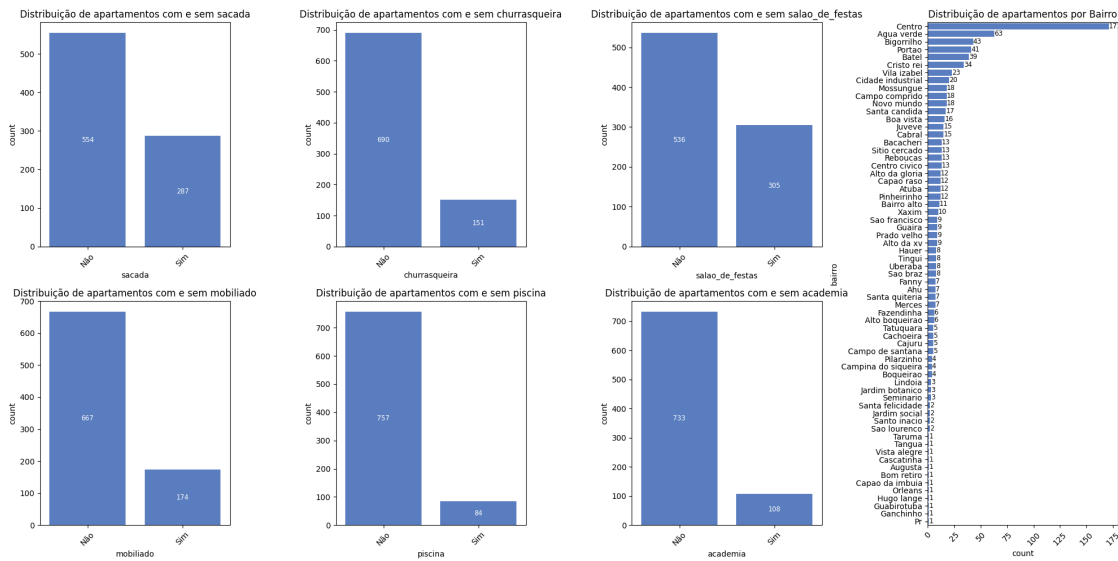
6.1.1 Variáveis Numéricas



Aqui, podemos perceber que as variáveis estão mais ajustadas, sem a presença de valores sem sentido ou outliers muito significativos.

Somos capazes de perceber uma clara distribuição normal em algumas variáveis como 'area', 'valor_total' e 'preco_metro_2' enquanto outras variáveis não possuem uma distribuição muito clara

6.1.2 Variáveis Categóricas



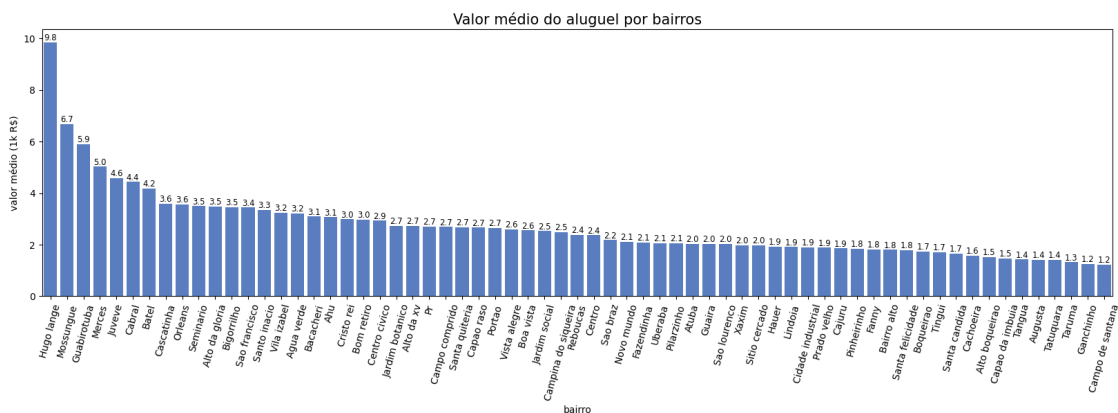
Aqui conseguimos notar o quão raro são alguns atributos como ‘piscina’ e ‘academia’ nos condomínios

Vemos também a predominância do centro da cidade como região com mais imóveis a sendo alugados

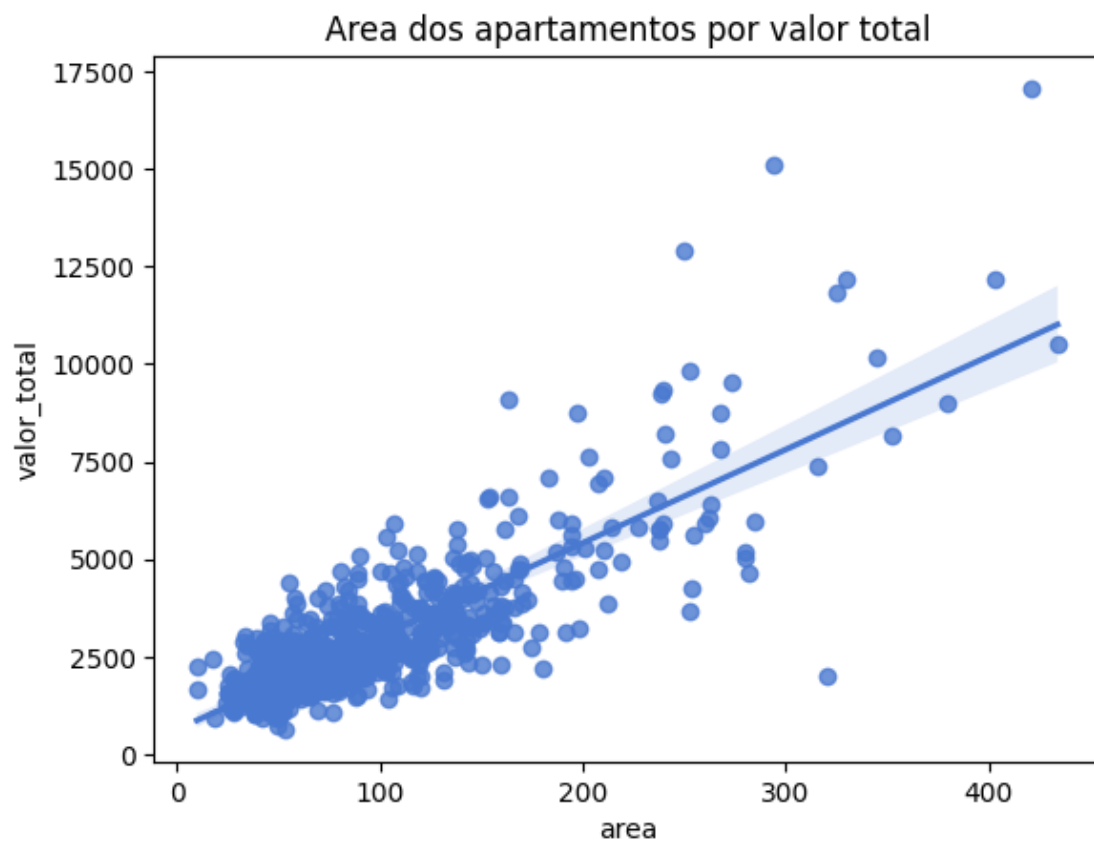
6.2 Análise Bivariada

- Nesta etapa, vamos analisar o comportamento das principais features em relação ao ‘valor_total’ do imóvel

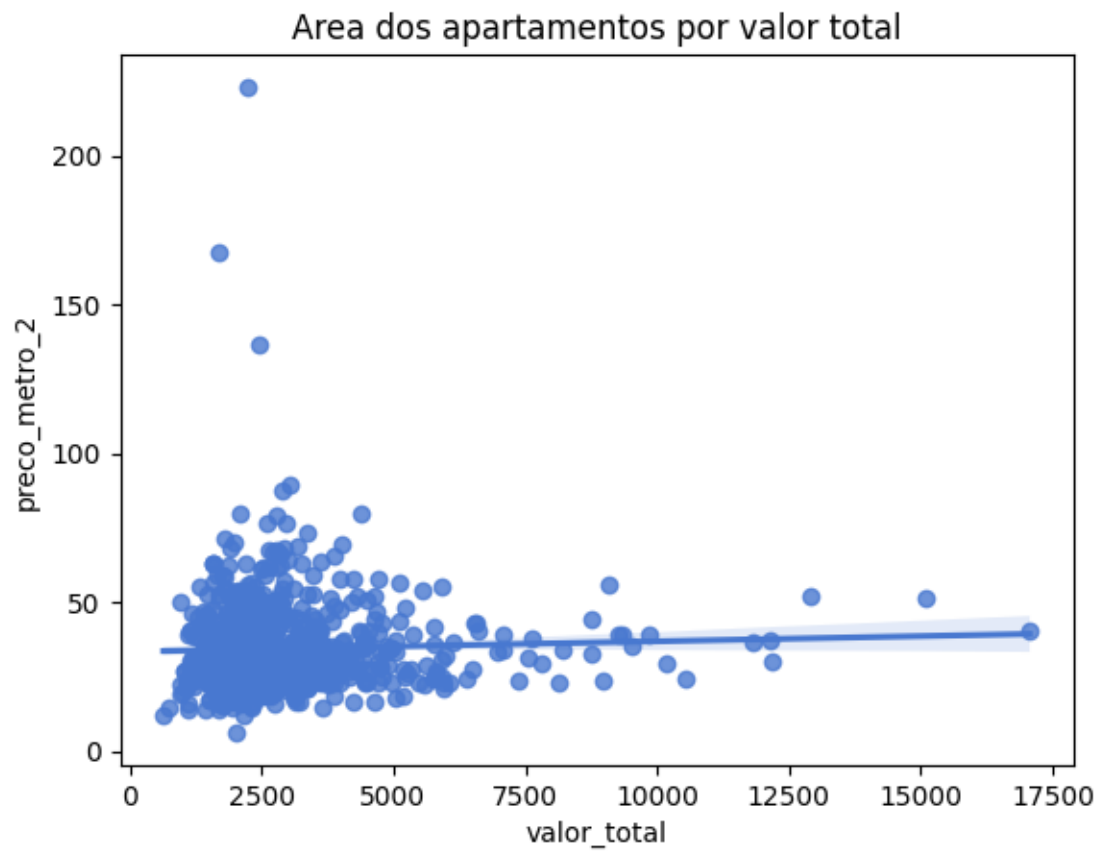
6.2.1 Bairro



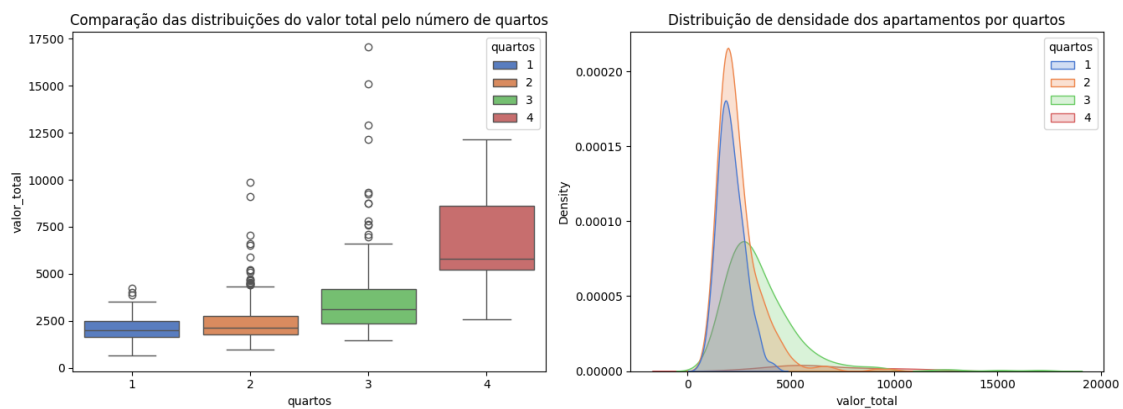
6.2.2 Area



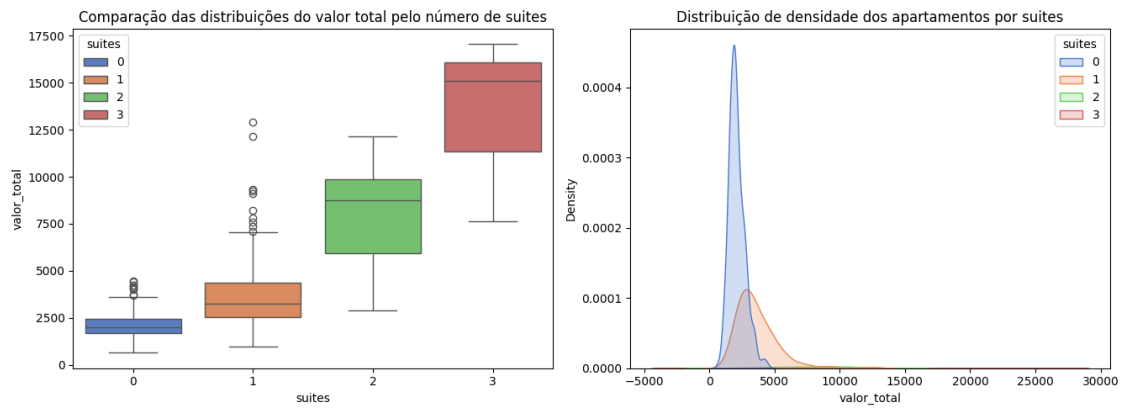
6.2.3 Preço m²



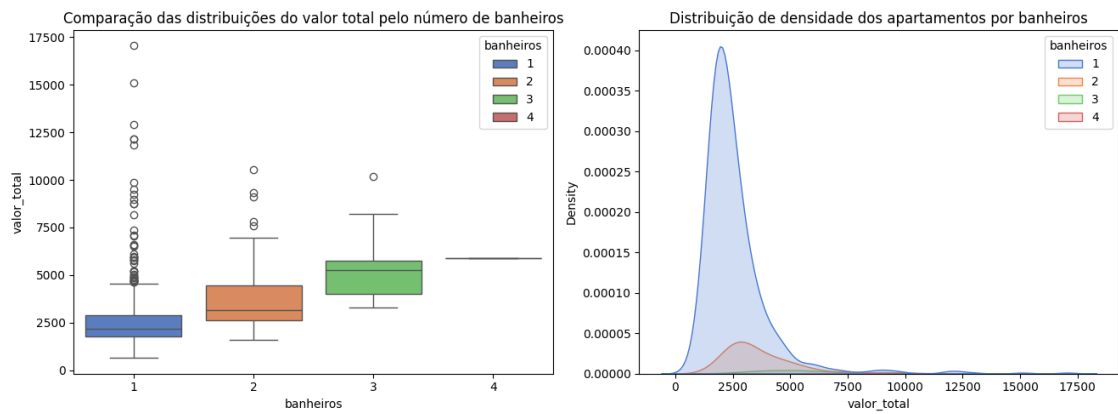
6.2.4 Quartos



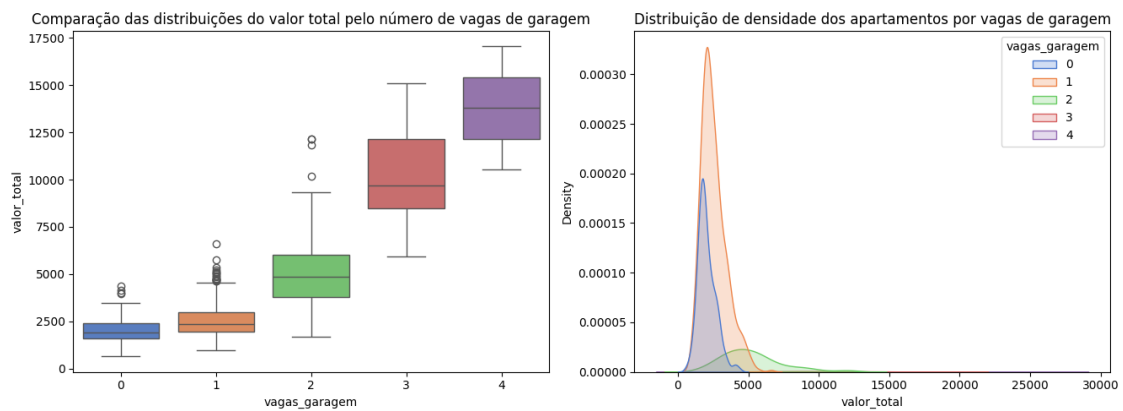
6.2.5 Suites



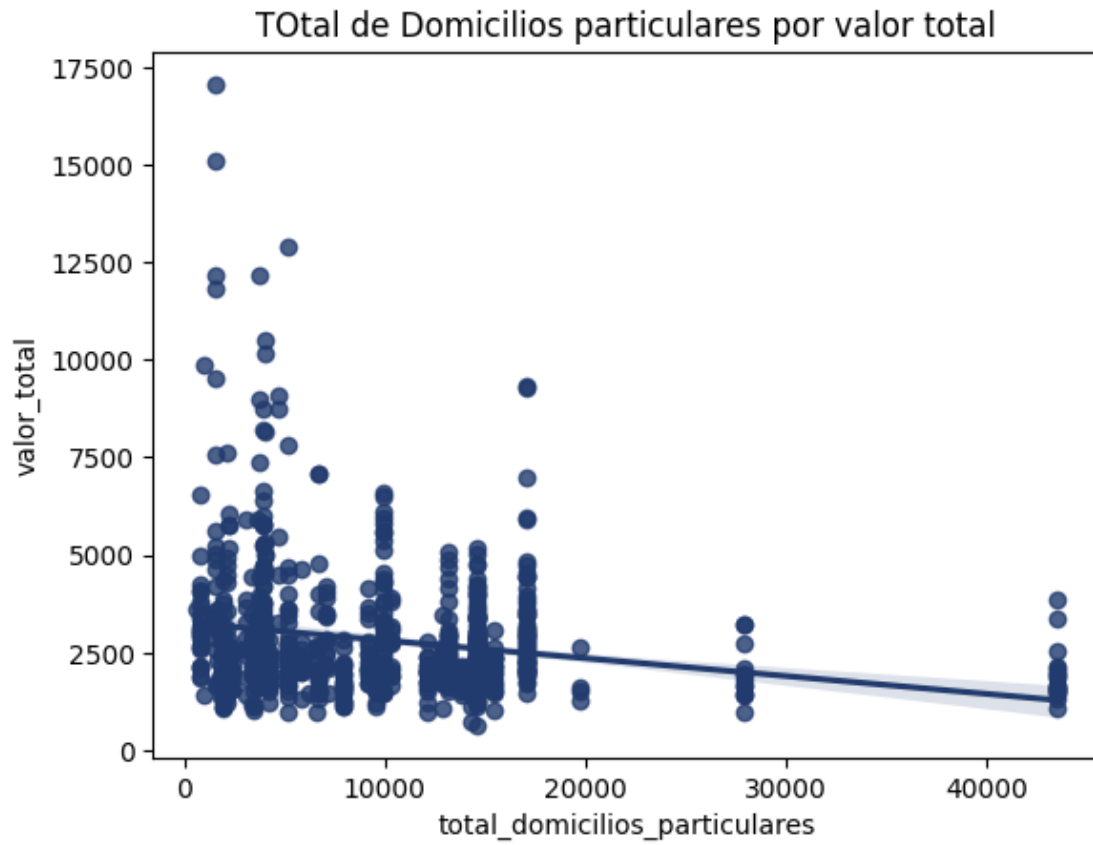
6.2.6 Banheiros



6.2.7 Vagas garagem

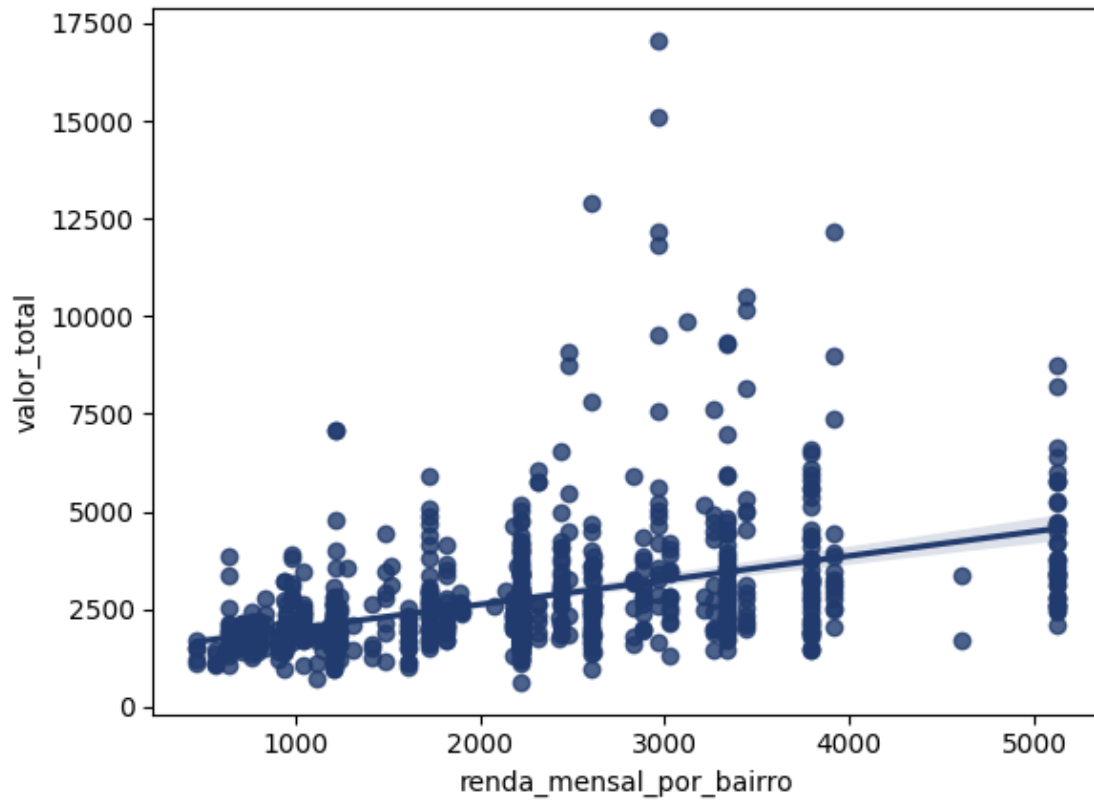


6.2.8 Total de Domicilios Particulares

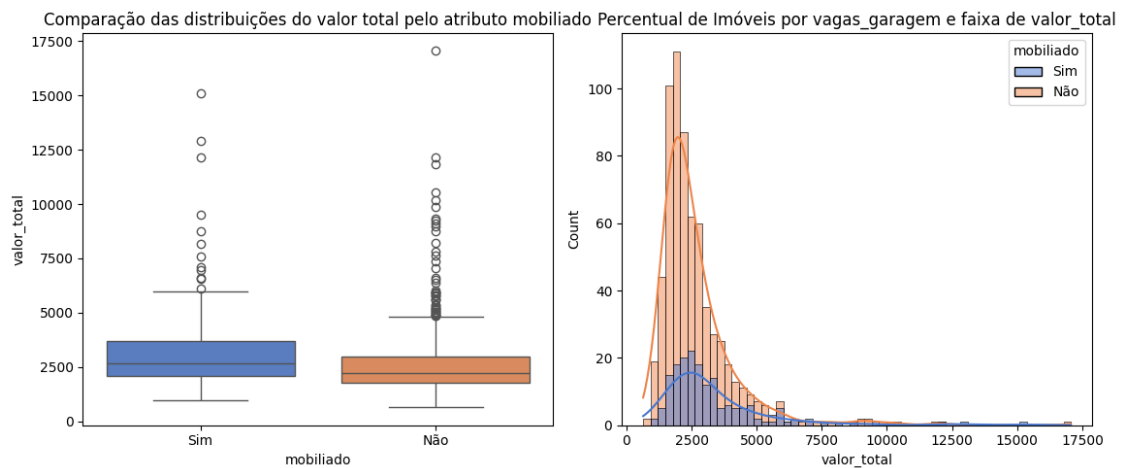


6.2.9 Renda Média Mensal por bairro

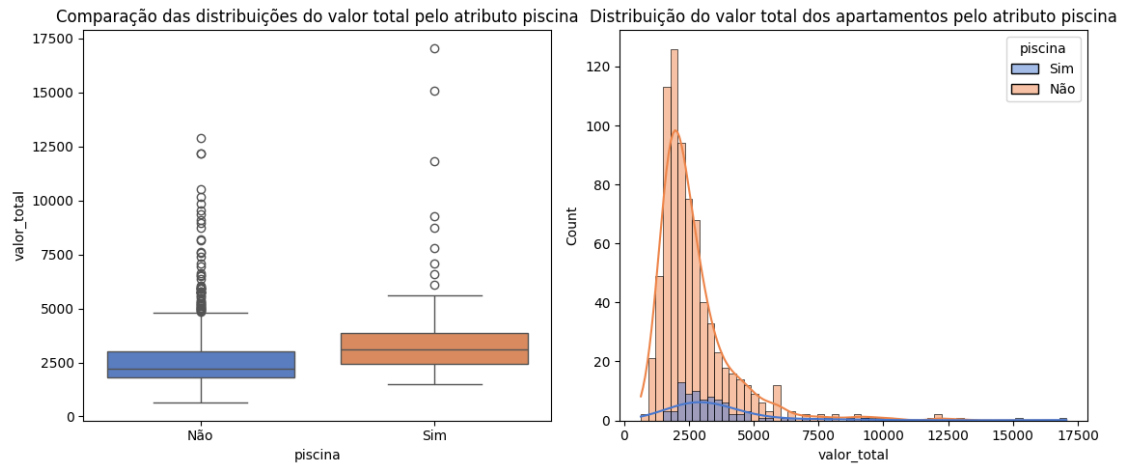
<Axes: xlabel='renda_mensal_por_bairro', ylabel='valor_total'>



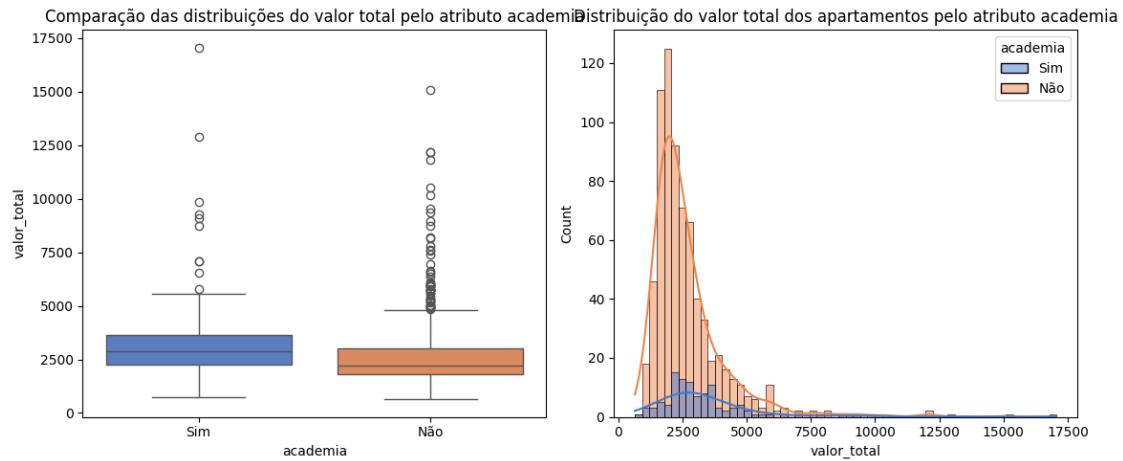
6.2.10 Mobiliado



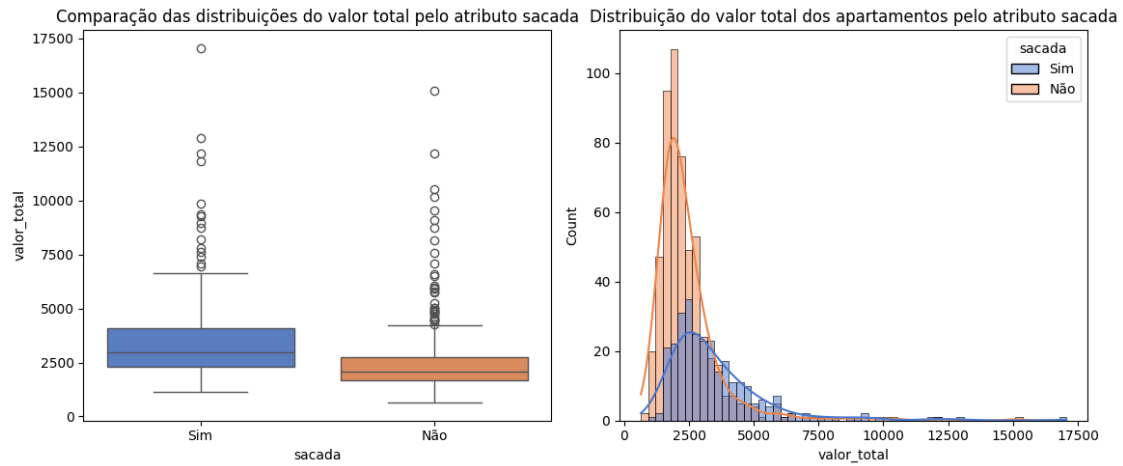
6.2.11 Piscina



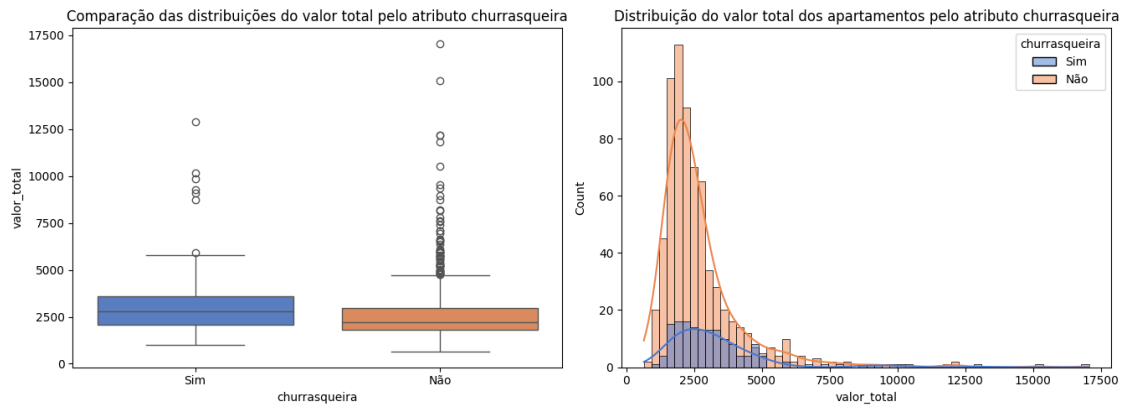
6.2.12 Academia



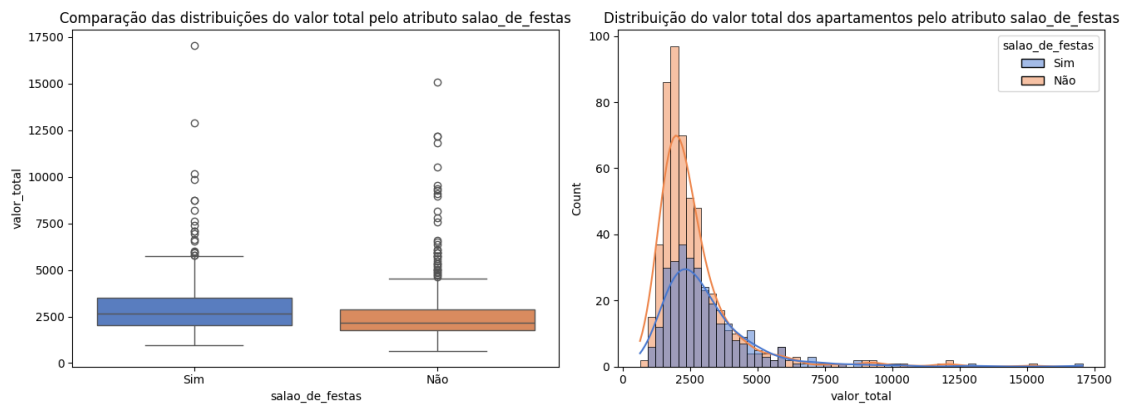
6.2.13 Sacada



6.2.14 Churrasqueira

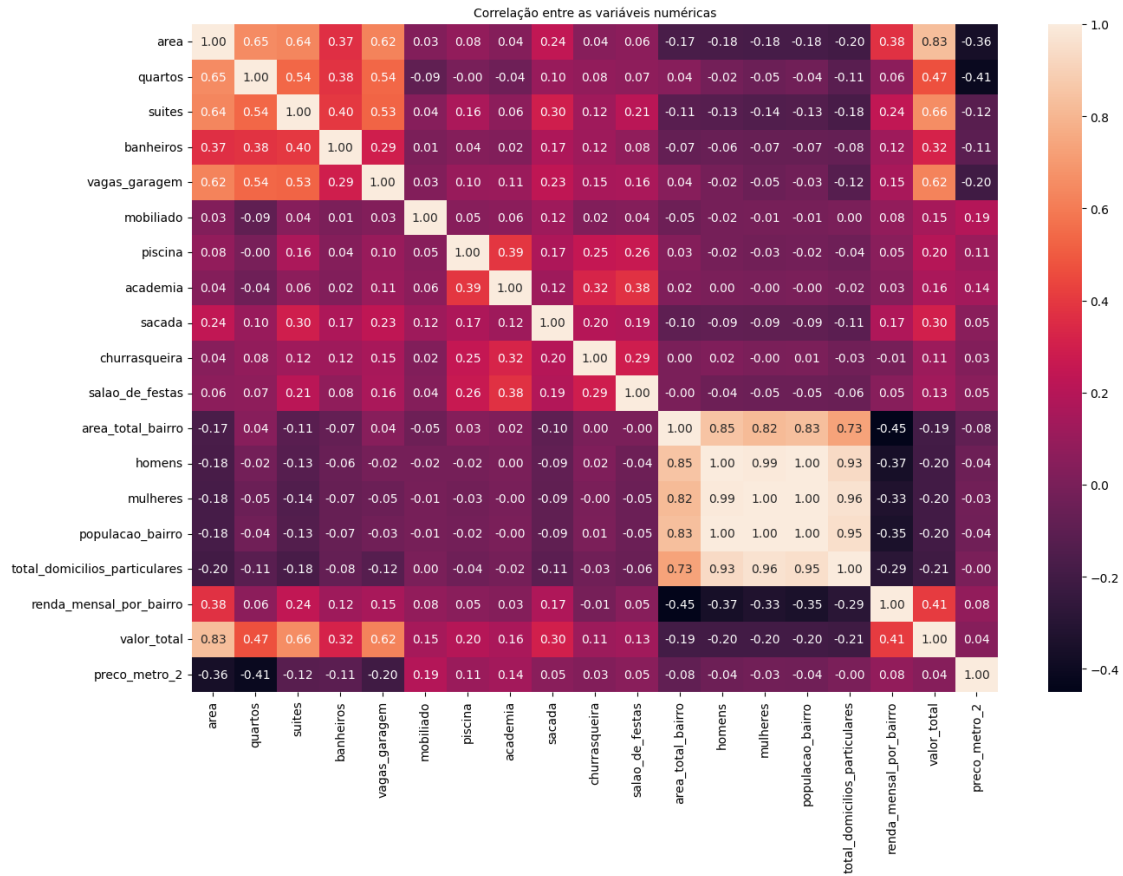


6.2.15 Salão de festas



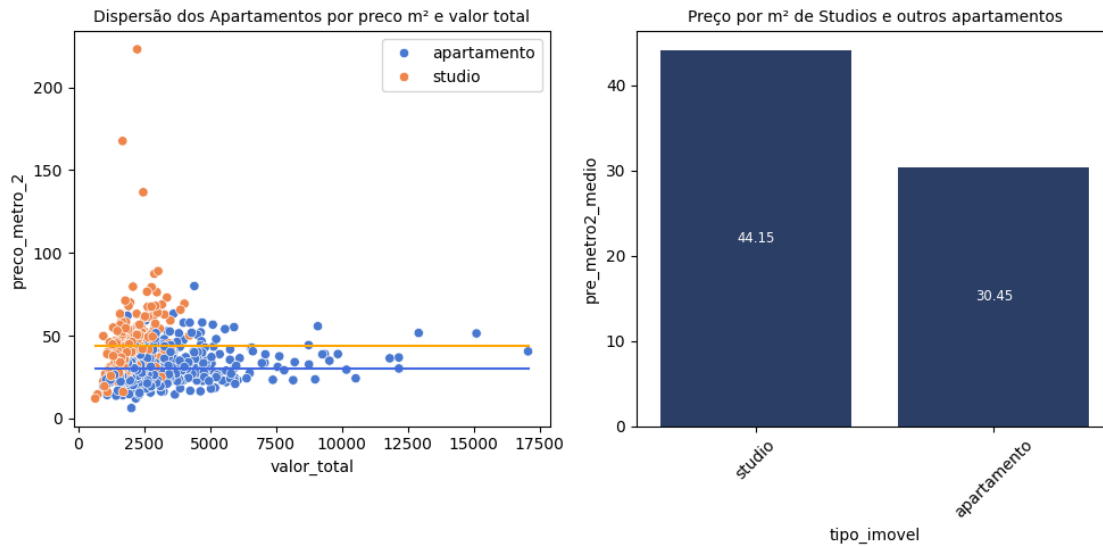
6.3 Análise Multivariada

- Nesta etapa o objetivo é entender a relação entre as diversas variáveis simultaneamente



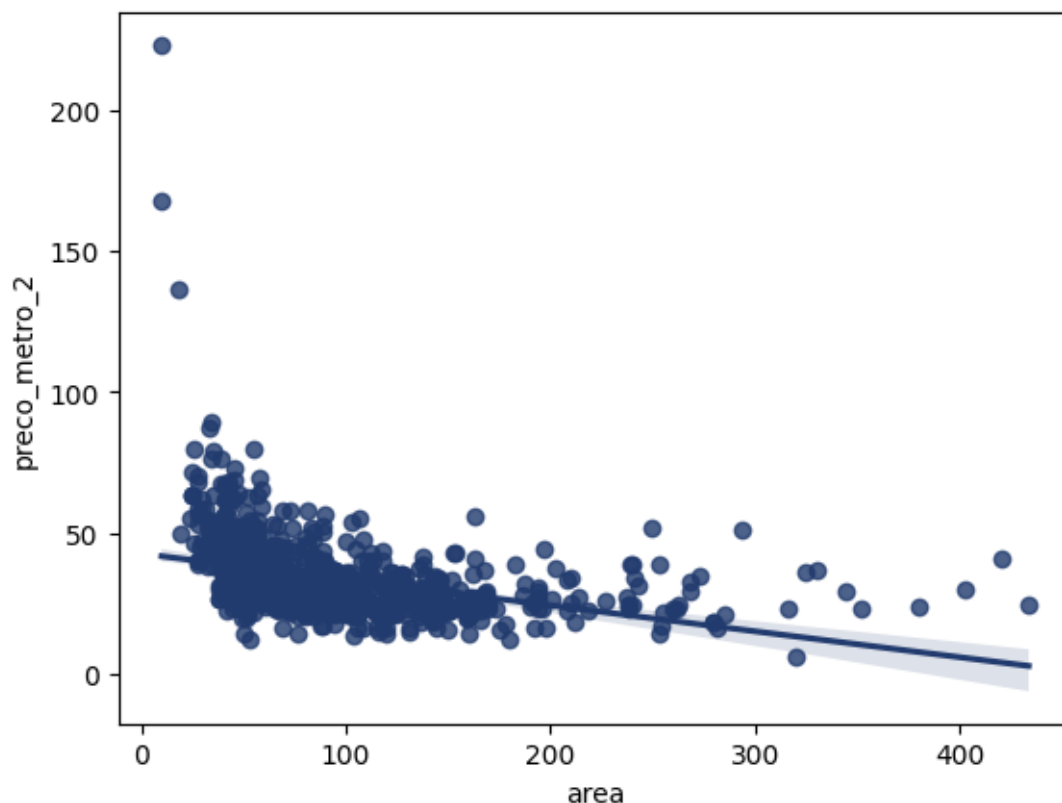
6.4 Teste de Hipóteses

6.4.1 1) O preço do metro quadrado de studios é maior que o preço do metro quadrado de outros apartamentos?



6.4.2 2) Imóveis com área menor tendem à ter maior valor por m²

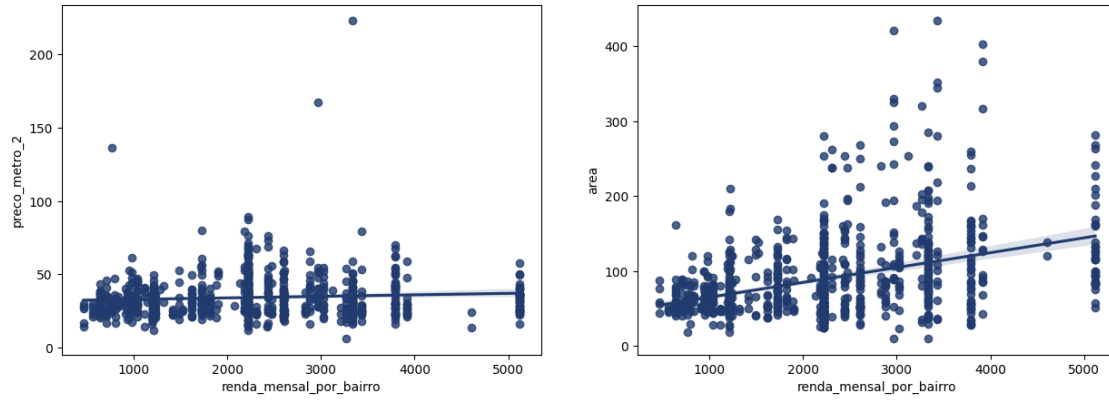
<Axes: xlabel='area', ylabel='preco_metro_2'>



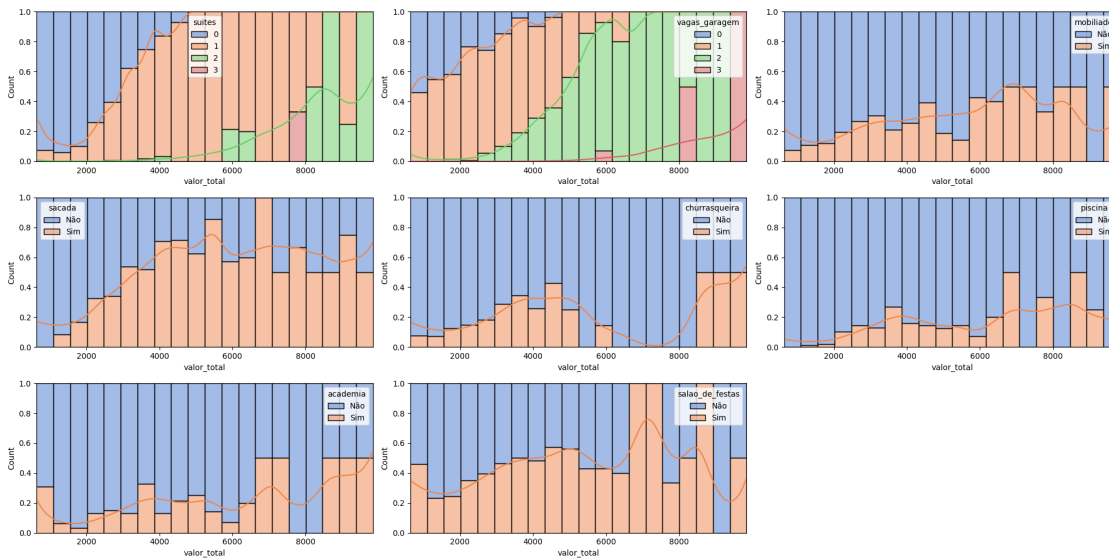
6.4.3 3) O preço por m² dos imóveis em bairros com renda média elevada é maior do que dos outros bairros

Vemos que o preço do m² é praticamente constante quando comparamos com a renda média por bairro, logo, bairros com renda média maior não necessariamente possuem imóveis com preço por m² mais caro

Em compensação, notamos que bairros com maior renda média possuem apartamentos com maior área, o que eleva o preço médio dos alugueis Além disso, vemos também que apartamentos em bairros mais ricos possuem mais atributos, como piscina, sacada e etc, que por sua vez também elevam o valor dos alugueis

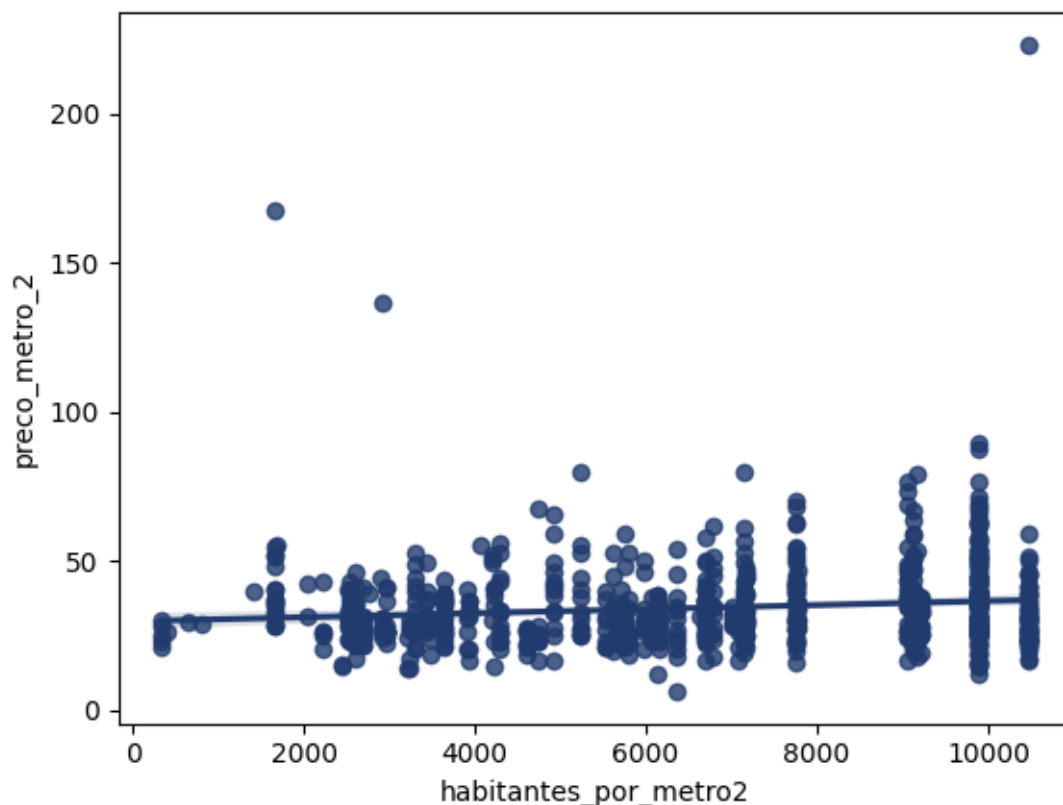


6.4.4 4) Imóveis mais caros possuem com mais frequência atributos como mobilia, sacada, churrasqueira e etc



6.4.5 5) Bairros com maior concentração de pessoas possuem um preço por m² maiores que outros bairros

<Axes: xlabel='habitantes_por_metro2', ylabel='preco_metro_2'>



7 Conclusão

Após explorar os dados e testar algumas hipóteses, podemos chegar à algumas conclusões sobre o comportamento dos preços do alugueis na cidade de Curitiba.

- 1) Imóveis menores, em média, são mais caros, ou seja, possuem um preço por m^2 maior do que os outros imóveis
 - Observamos que imóveis menores tem o preço por m^2 maior do que os outros imóveis, e isso pode ser devido à algumas hipóteses que podemos testar futuramente como o preço de venda do m^2 desses imóveis também é maior em relação aos outros ou então devido à maior procura por imóveis do tipo studio seu preço se eleve, mas o fato é que o preço de seu m^2 é maior quando comparado à imóveis maiores
- 2) O bairro do imóvel não encarece o imóvel em si. A área e a presença de mais atributos são mais relevantes para encarecer os alugueis
 - É notório que o valor médio dos alugueis de determinados bairros (com renda média maior) é maior do que o de outros bairros, mas fica claro que esse aumento no preço não se deve ao bairro em si e sim às características do imóvel. Imóvel em bairros com renda média maior, apresentam também maior área, e mais atributos como suite, vagas de garagem e condomínios mais sofisticados e isso sim parece ser mais determinante para o aumento do preço do que a própria localidade em si.