

Projeto Airbnb

Neste Notebook serão analisados dados do Airbnb.

Para mais informações sobre os dados, acesse o repositório no GitHub.

Bruna Mendes (brunnamendes@gmail.com)

Última alteração: 18/01/2021

Análise Exploratória

Primeiras linhas das bases utilizadas:

	Unnamed: 0	Unnamed: 0.1	airbnb_listing_id	booked_on	date	price_string	occupied
0	0	2148	40201349	blank	2020-11-15	250.0	0
1	1	2159	40201349	blank	2020-11-26	250.0	0
2	2	2160	40201349	blank	2020-11-27	250.0	0
3	3	2173	40201349	blank	2020-12-10	250.0	0
4	4	2226	40201349	blank	2021-02-01	250.0	0

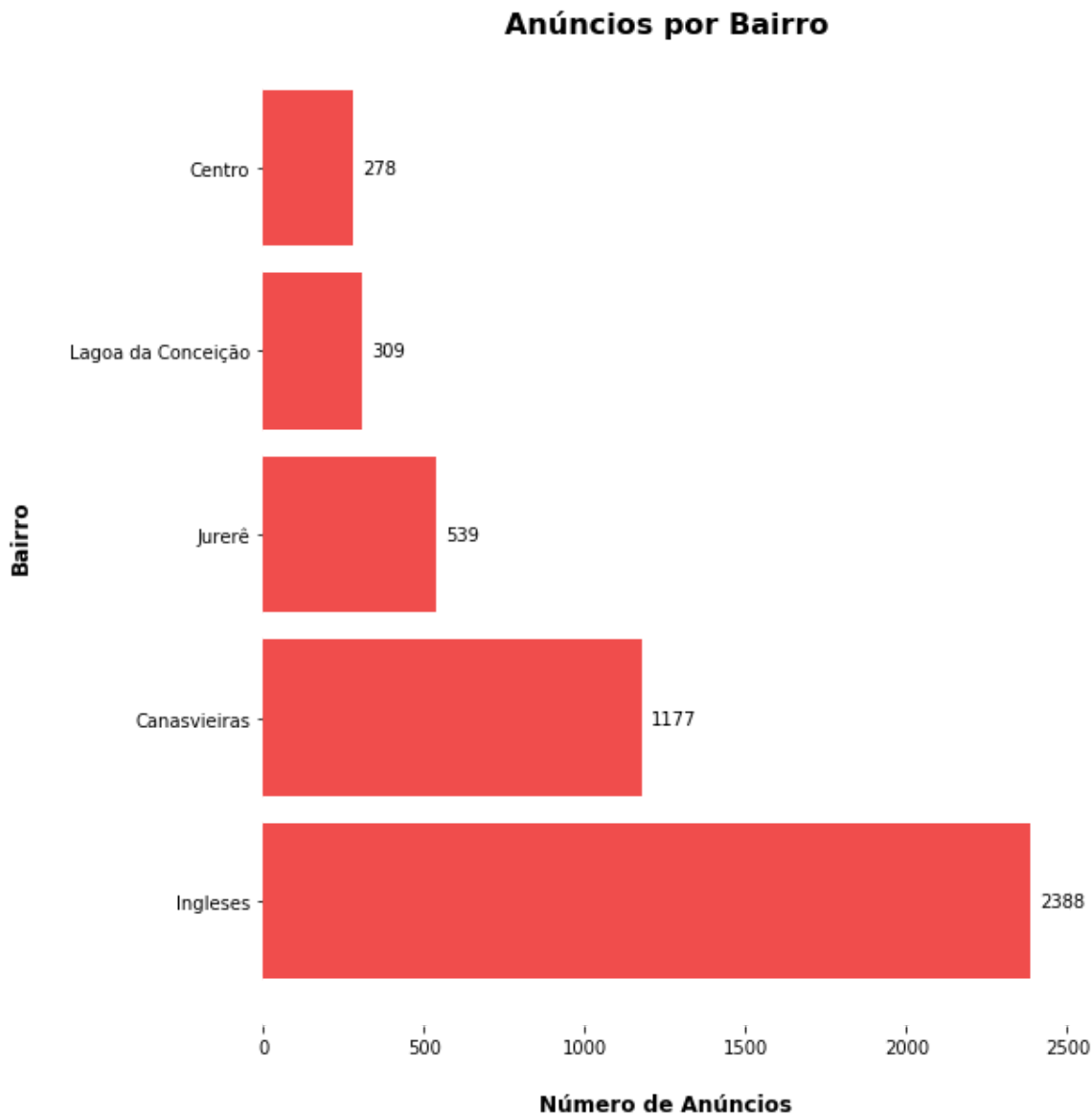
Unnamed: 0	airbnb_listing_id	suburb	ad_name	number_of_bedrooms	number_of_l
0	0	31389869	Jurerê Lindo Apartamento em Jurerê	2.0	
1	1	40010667	Canasvieiras Residencial Arruda, 1 quarto	1.0	
2	2	38905997	Ingleses Apartamento NOVO Completo - Moderno e Sofisticado	1.0	
3	3	22343656	Ingleses 06- Apartamento 02 habitaciones	2.0	
4	4	18328184	Canasvieiras Apto 2 quartos em Canasvieiras, Florianopolis!	2.0	

Foram utilizadas duas bases, a details, que está a nível do imóvel(anúncio) e traz características do imóvel como bairro, número de banheiros e quartos, nome do anúncio, avaliação em estrelas, superhost e número de avaliações. A segunda base, price, está a nível do imóvel e da data de ocupação e traz informações como data, data da reserva, preço da diária e ocupação.

1. Ordenando os bairros por ordem crescente do número de listings

Para esse exercício será utilizada a base details. Agruparemos a base por bairro e contaremos o número de anúncios

Um gráfico de barras será construído para visualizar o resultado

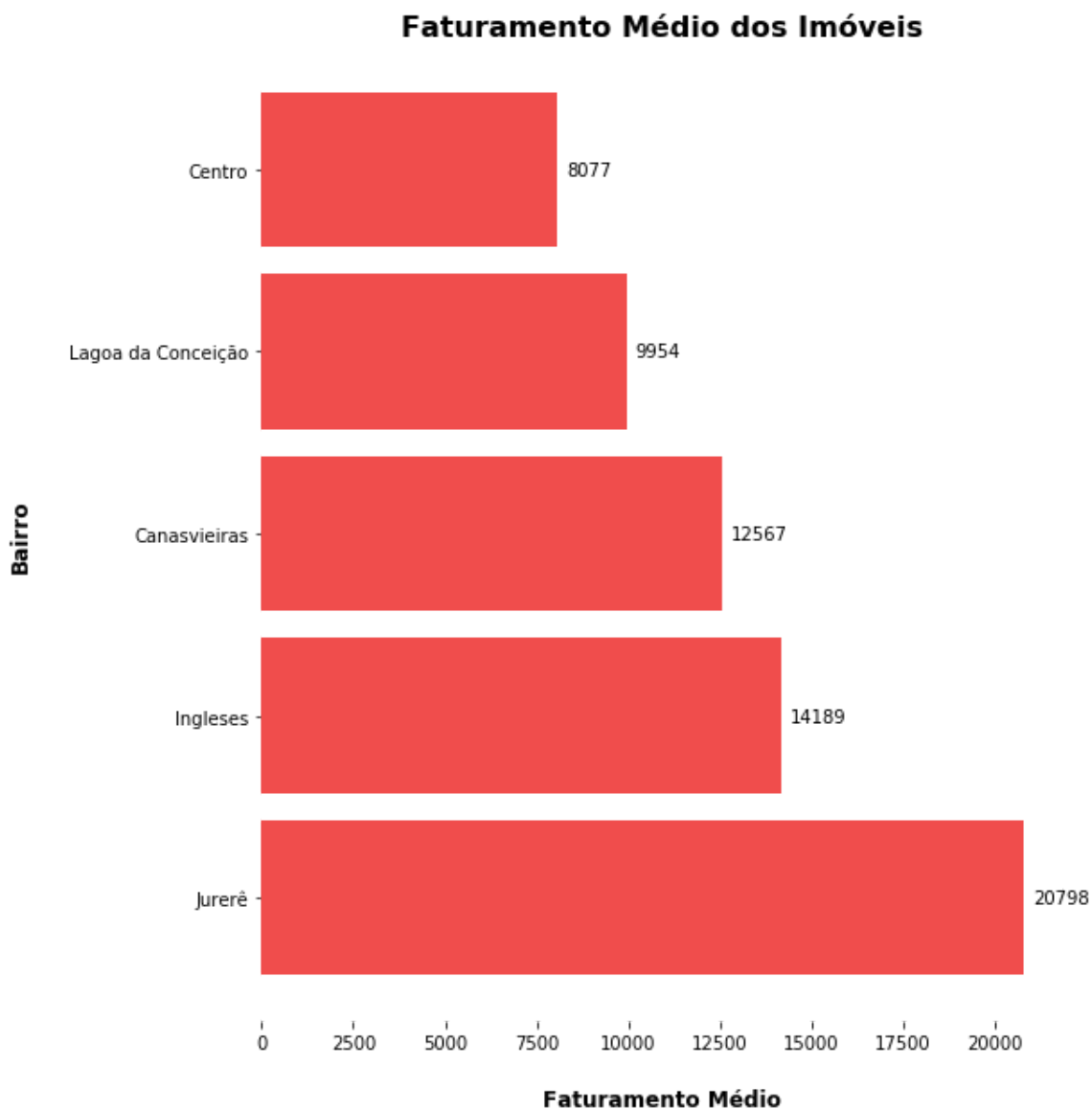


O bairro com maior número de anúncios é o bairro dos Ingleses, com um total de 2388. O Centro possui o menor número de anúncios entre os cinco bairros abrangidos pela base, 278.

2. Ordenando os bairros em ordem crescente de faturamento médio dos listings

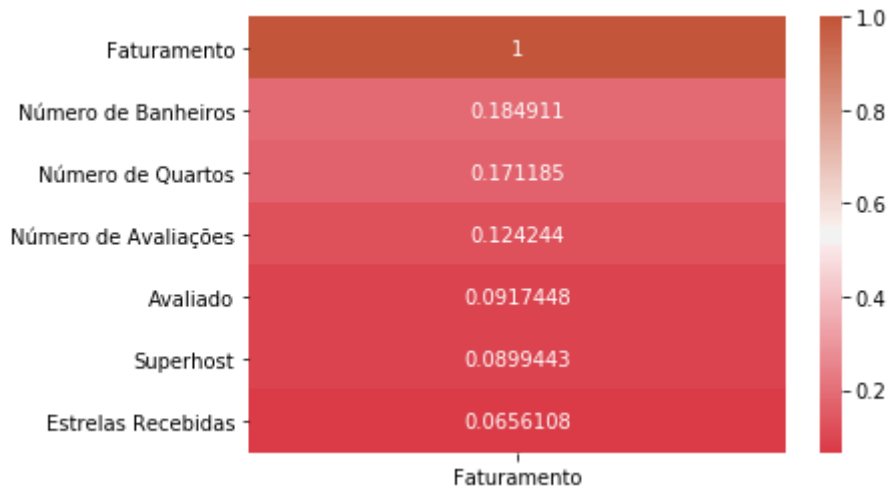
Para esse exercício serão utilizadas as duas bases. Primeiramente a price para criar a variável faturamento e depois uma união das duas bases para analisar o faturamento médio dos imóveis por bairro.

Constrói um gráfico de barras horizontais para visualizar o resultado.



O bairro Jurerê possui um faturamento médio por imóvel de 20.798,00 reais, sendo o maior entre os bairros. O faturamento médio do Centro, o menor entre os bairros analisados é de 8.077,00 reais.

3. Calculando correlação entre o faturamento e outras variáveis



Não há variáveis altamente correlacionadas com o faturamento. As variáveis com maior correlação com o faturamento são: número de banheiros e números de quartos, respectivamente. Isso pode ter relação com o fato de imóveis maiores terem diárias mais caras. A terceira característica com maior correlação com o faturamento é o número de avaliações, o que pode ser um sinal positivo para quem busca um imóvel para aluguel, elevando as chances de uma nova reserva.

4. Calculando antecedência média da reserva

Para calcular a antecedência média das reservas, a base price foi utilizada para calcular a diferença em dias da data da reserva até a primeira diária. As duas variáveis possuem formatos diferentes de data. Uma inclui horário e a outra não.

Como verificamos que não há variação de horário, padronizou-se para o formato dd/mm/aaaa.

A antecedência média das reservas é de 15.5 dias

Considerando apenas as reservas de fim de semana, a média cai para 8.55 dias de antecedência.