

João Pessoa - Análise dos dados do AirBnb



(Disponível em: https://www.pbtur.pb.gov.br/wp-content/uploads/2021/10/DJI_0022.jpg)

Introdução

João Pessoa é uma grande cidade no exterior do nordeste brasileiro, sendo considerada a terceira cidade mais antiga do Brasil. É uma região com uma população de 825.796 habitantes. A capital paraibana é a 8ª cidade mais populosa da **Região Nordeste**.

É conhecida como “Porta do Sol”, devido ao fato de, no município, estar localizada a **Ponta do Seixas**, que é o ponto mais oriental das **Américas**, o que faz a cidade ser conhecida como o lugar “onde o sol nasce primeiro no continente americano”.

Análise dos dados

AirBnb é um serviço ao qual pessoas do mundo inteiro podem oferecer suas casas para usuários que buscam acomodações mais em conta em qualquer lugar do mundo. Atualmente já é considerada como **a maior empresa hoteleira, mesmo não possuindo nenhum hotel**.

Esse serviço oferece uma plataforma inovadora para tornar a hospedagem alternativa. Fundada há quase 15 anos, a empresa responsável **já hospedou mais de 300 milhões de pessoas em todo o mundo**.

Conhecendo os dados

Uma das primeiras coisas que devemos fazer é conhecer a nossa base de dados. Ao verificar o site, decidimos utilizar algumas informações, as quais achamos mais pertinentes. Abaixo estão destacadas as principais variáveis para a análise:

Localização - informa o local em que está localizada a acomodação.

Tipo de acomodação - informa qual o tipo de acomodação.

Quantidade e tipo de camas - informa qual o tipo e quantas camas possuem na acomodação.

Preço por noite - informa qual o preço por noite na acomodação.

Avaliação das acomodações - informa qual a avaliação dada para a acomodação.

Quantidade de avaliações - informa quantas avaliações foram feitas para a acomodação.

Ao examinar o dicionário de dados constatamos que nossa base possui um total de 169 entradas com 6 variáveis.

Por conseguinte, foi verificado se a base possuía valores ausentes:

```
[ ] df.isnull().sum().sort_values(ascending=False) / df.shape[0] * 100
```

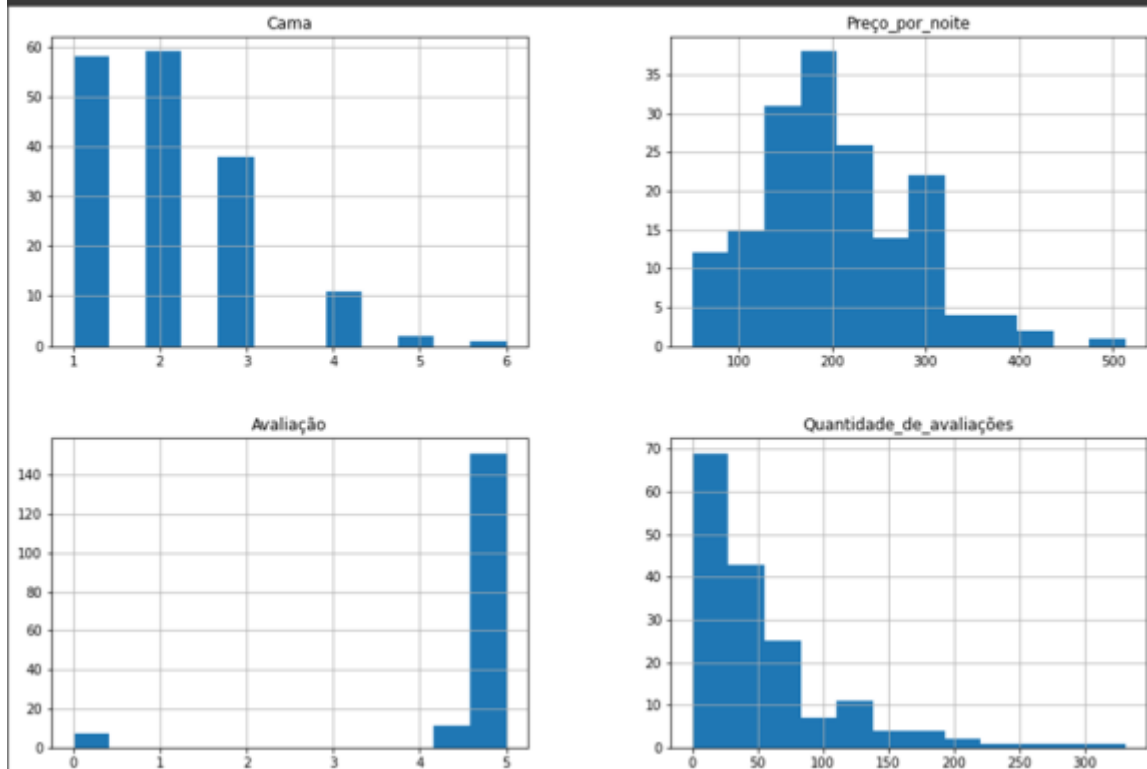
Quantidade_de_avaliações	4.142012
Local	0.000000
Tipo	0.000000
Cama	0.000000
Preço_por_noite	0.000000
Avaliação	0.000000
dtype: float64	

Como podemos verificar na imagem acima, a variável Quantidade_de_avaliações possui cerca de 4% de dados ausentes, então, faz-se necessária a remoção desses dados.

Verificando a distribuição dos dados

Para verificarmos a distribuição de dados podemos utilizar um histograma. As barras mais altas mostram que mais dados estão nesse intervalo.

```
# plotar o histograma das variáveis numéricas
df.hist(bins=12, figsize=(15,10));
```



Analisando a imagem acima podemos observar que as variáveis preço_por_noite e avaliação possuem uma variação significativa em seus dados. Tal característica pode indicar a presença de outliers.

Os [outliers](#) são dados que se diferenciam radicalmente de todos os outros, ou seja, é um valor anormal e que pode (e provavelmente irá) causar irregularidades nos resultados obtidos por meio de algoritmos e sistemas de análise.

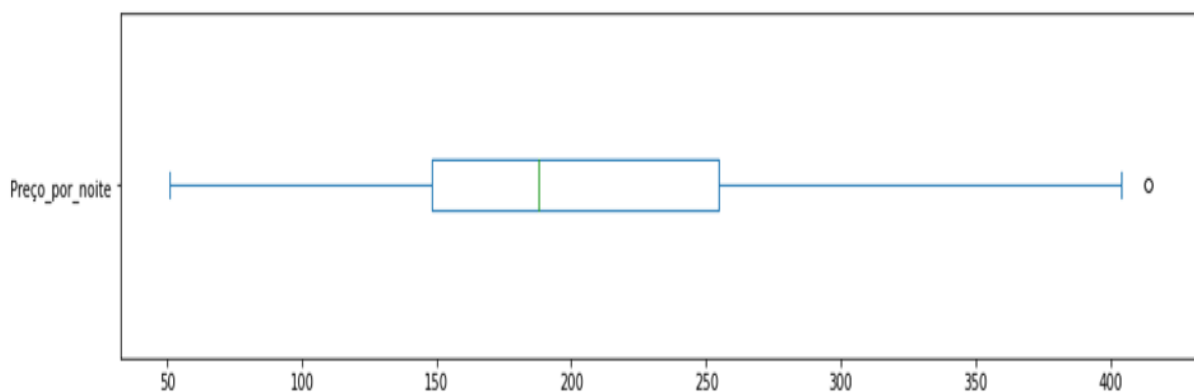
Para auxiliar nossa análise vamos verificar os dados estatísticos da base:

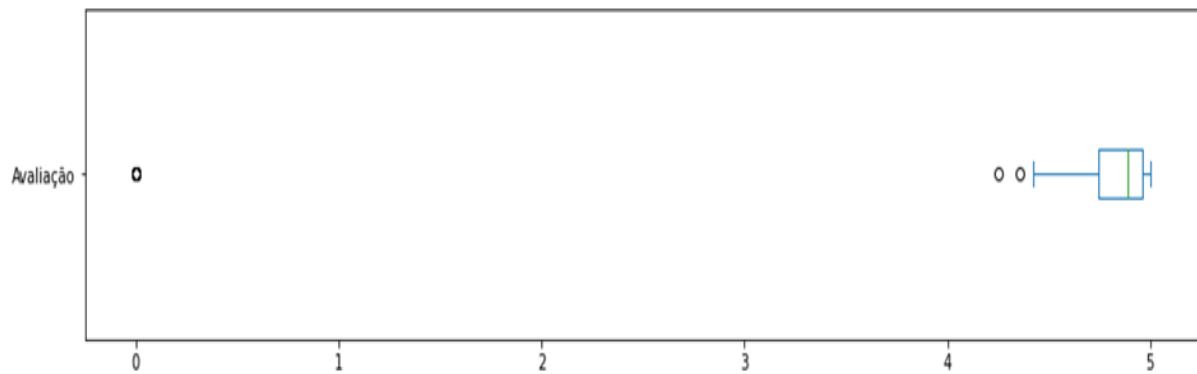
	Cama	Preço_por_noite	Avaliação	Quantidade_de_avaliações
count	216.000000	216.000000	216.000000	216.000000
mean	2.046296	207.384259	4.674537	55.412037
std	1.005883	83.801346	0.931183	59.719590
min	1.000000	51.000000	0.000000	0.000000
25%	1.000000	151.000000	4.742500	15.750000
50%	2.000000	191.000000	4.900000	36.000000
75%	3.000000	267.000000	4.970000	74.250000
max	6.000000	513.000000	5.000000	331.000000

Como é possível observar, a coluna `preço_por_noite` possui uma média (mean) de 207.38 e o seu valor máximo (max) está em 513.00. Sendo que 75% dos imóveis estão localizados na faixa de valores até 267.00. De maneira semelhante temos a coluna `avaliação` que possui uma média de 4.67, e um valor máximo de 5.00.

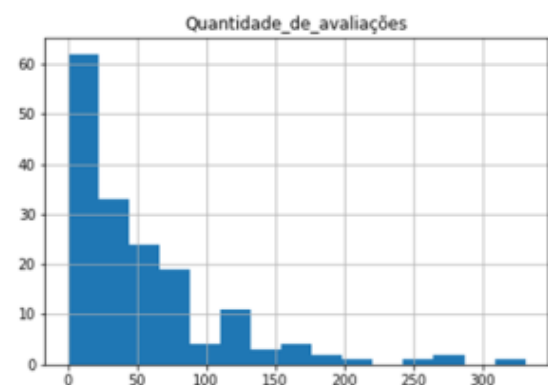
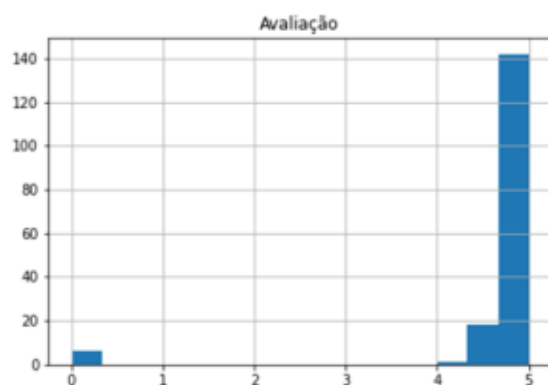
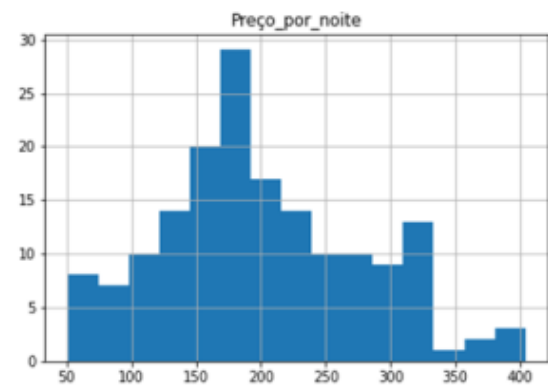
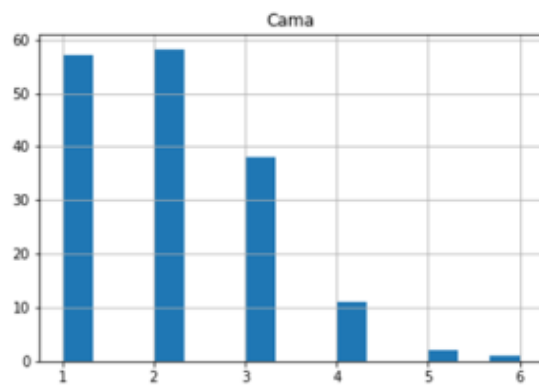
Tratando Outliers

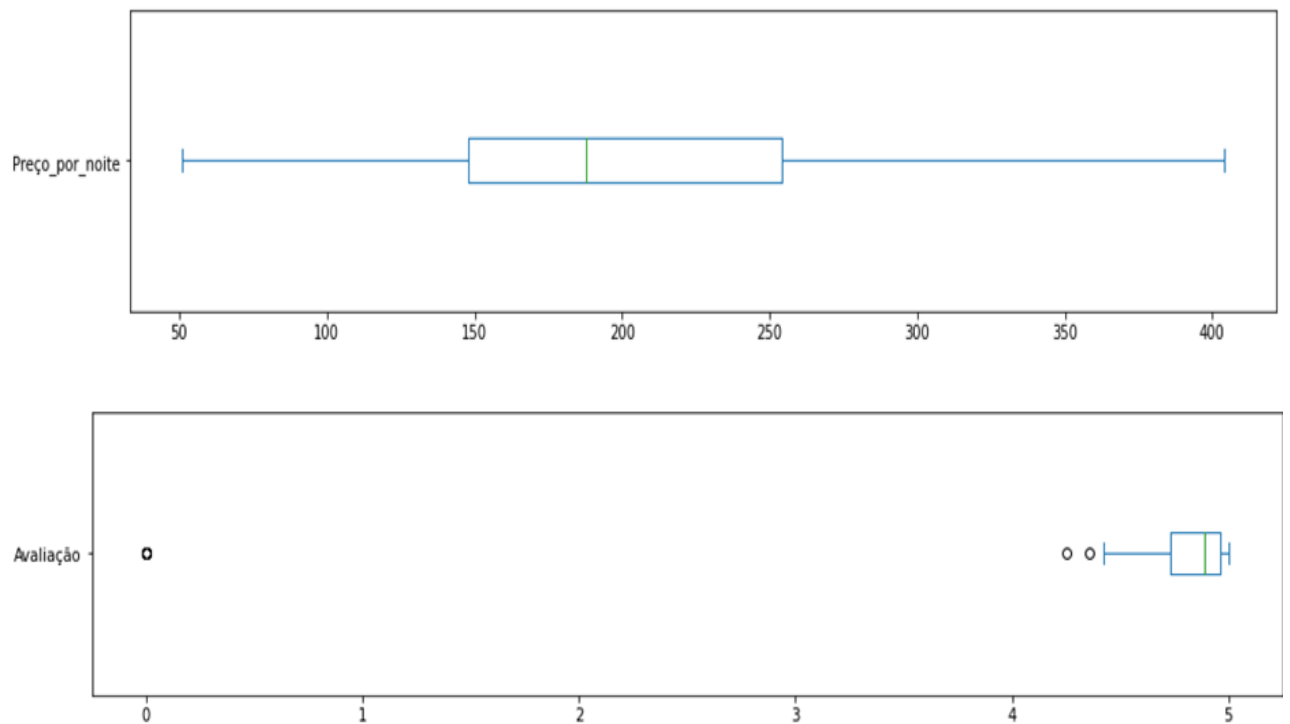
O boxplot nos ajuda a identificar de forma visual os outliers.





Para tratar os outliers identificados escolhemos a [Variação Interquartil](#) para determinar um limite superior para os dados. Após aplicar a regra, encontramos um limite superior de 413,5 para a variável preço_por_noite e um limite superior de 5,2 para a variável avaliação. Após eliminar os outliers ficamos com a seguinte distribuição dos dados:





Identificando a média de preço dos imóveis

Já eliminamos os outliers, agora vamos verificar quais as categorias de imóveis e seus devidos preços:

```
# mostrar quantas vezes a descrição do imóvel se repete
df_clean.Tipo.value_counts()
```

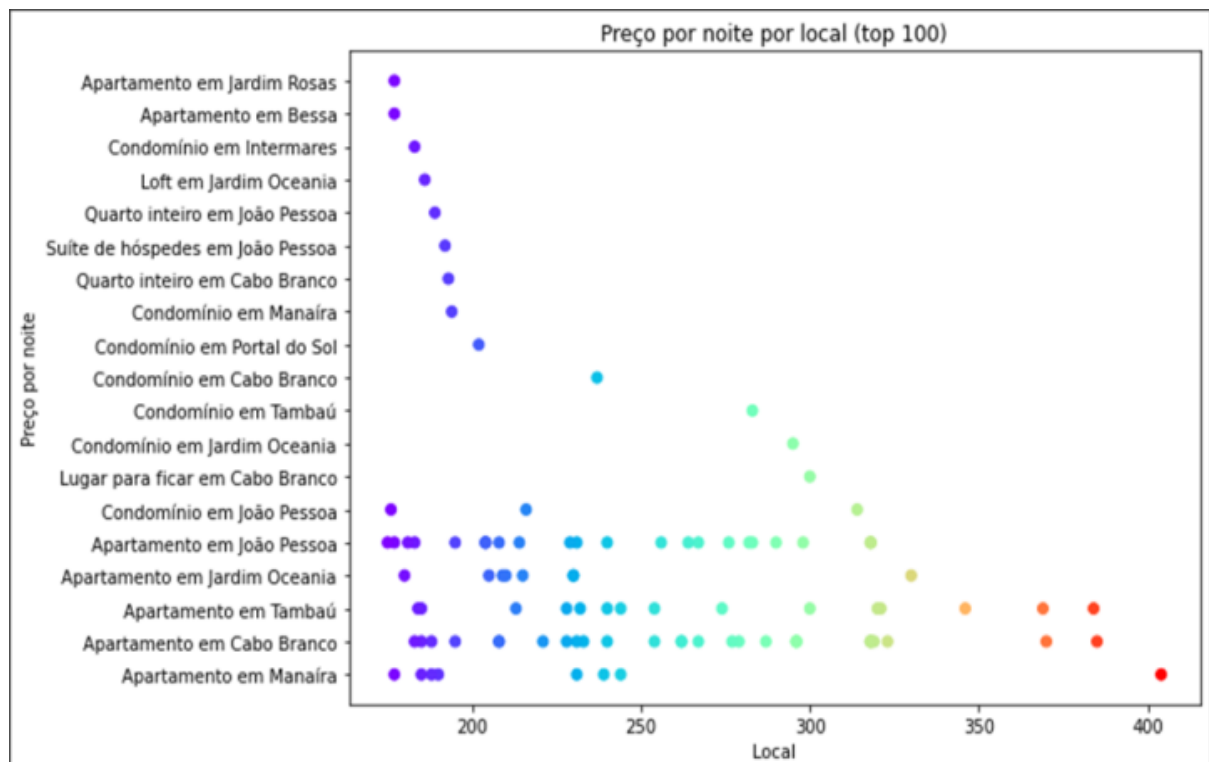
Flat a Beira Mar do Bessa	2
Apartamento aconchegante com 01 quarto, próximo a Orla de Manaíra - João Pessoa - PB	2
Flat a beira mar	2
Apto Completo proximo a praia- Manaíra	2
Quarto aconchegante	2
...	..
Apto incrível na melhor localização de Tambaú	1
Flat em João Pessoa - Urbani Manaíra (Sem Toalhas)	1
Maravilhoso Flat com vista para o mar F311	1
Manaíra Flat 206	1
Sunrise Blue Suites Luxo / com vista do mar e varanda	1

Name: Tipo, Length: 159, dtype: int64

```
# mostrar quantos imóveis há em cada localização
df_clean.Local.value_counts()
```

Apartamento em João Pessoa	37
Apartamento em Cabo Branco	36
Apartamento em Manaíra	21
Apartamento em Tambaú	19
Apartamento em Jardim Oceania	18
Quarto inteiro em João Pessoa	7
Quarto inteiro em Jardim Oceania	4
Condomínio em João Pessoa	4
Quarto inteiro em Manaíra	3
Apartamento em Bessa	3
Apartamento em Intermares	2
Apartamento em Loteamento Bela Vista	2
Quarto inteiro em Cabedelo	1
Condomínio em Tambaú	1
Condomínio em Jardim Oceania	1
Microcasa em João Pessoa	1
Condomínio em Portal do Sol	1
Apartamento em Gramame	1
Suíte de hóspedes em João Pessoa	1
Condomínio em Intermares	1
Loft em Cabo Branco	1
Quarto inteiro em Cabo Branco	1
Casa em João Pessoa	1
Loft em Jardim Oceania	1
Lugar para ficar em Cabo Branco	1
Apartamento em Jardim Rosas	1
Condomínio em Manaíra	1
Condomínio em Cabo Branco	1
Apartamento em Bancários	1
Quarto inteiro em Loteamento Bela Vista	1
Apartamento em Cabedelo	1

Name: Local, dtype: int64




```
df_clean.groupby('Local').Preço_por_noite.mean().sort_values(ascending=False)
```

Local	
Loft em João Pessoa	414.000000
Lugar para ficar em Cabo Branco	300.000000
Condomínio em Jardim Oceania	295.000000
Condomínio em Tambaú	283.000000
Apartamento em Cabo Branco	251.723404
Apartamento em Tambaú	251.541667
Condomínio em Cabo Branco	237.000000
Apartamento em Jardim Oceania	218.083333
Apartamento em João Pessoa	215.000000
Condomínio em João Pessoa	211.500000

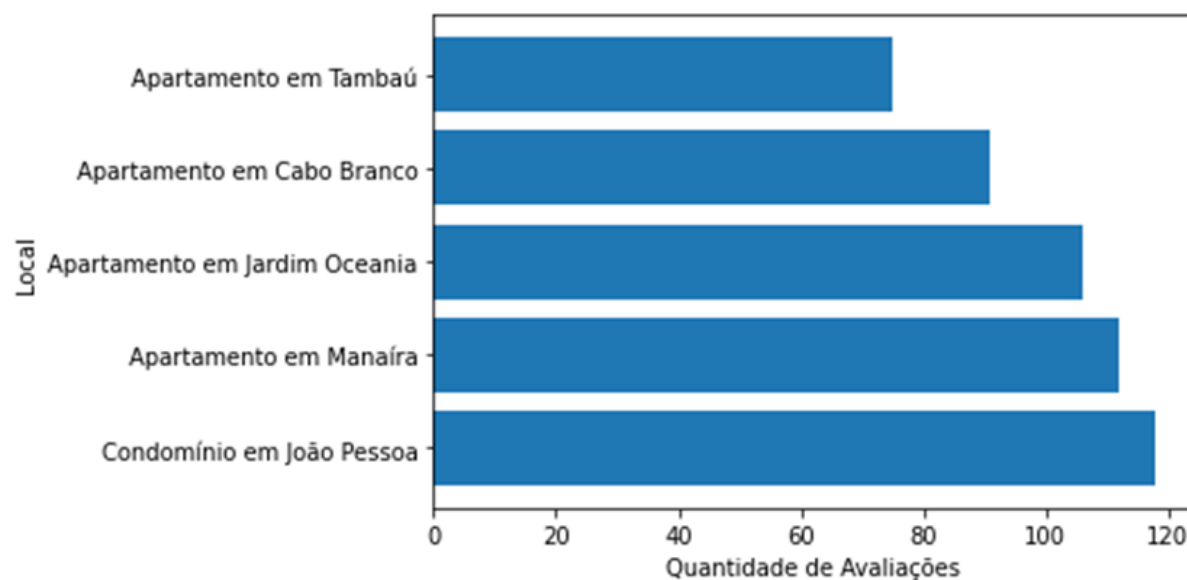
Observe na tabela acima, a qual apresenta a média de preço das acomodações mais caras, que a acomodação que possui a média de aluguéis mais cara é o loft em João Pessoa (R\$414,00), já a que tem a menor média de preço é o condomínio em João Pessoa (R\$211,50).

Apartamento em Cabedelo	140.000000
Apartamento em Loteamento Bela Vista	138.000000
Casa em João Pessoa	128.000000
Apartamento em Intermares	113.000000
Quarto inteiro em João Pessoa	101.285714
Quarto inteiro em Cabedelo	98.000000
Apartamento em Bancários	97.000000
Microcasa em João Pessoa	84.000000
Quarto inteiro em Manaíra	71.333333
Apartamento em Gramame	64.000000
Quarto inteiro em Jardim Oceania	55.250000

No que diz respeito a média de preços das acomodações mais baratas, como é possível ver na tabela acima, o apartamento em Cabedelo é o mais caro, com uma média de R\$140,00, já o quarto inteiro em Jardim Oceania é o mais barato, na média de R\$55,25.

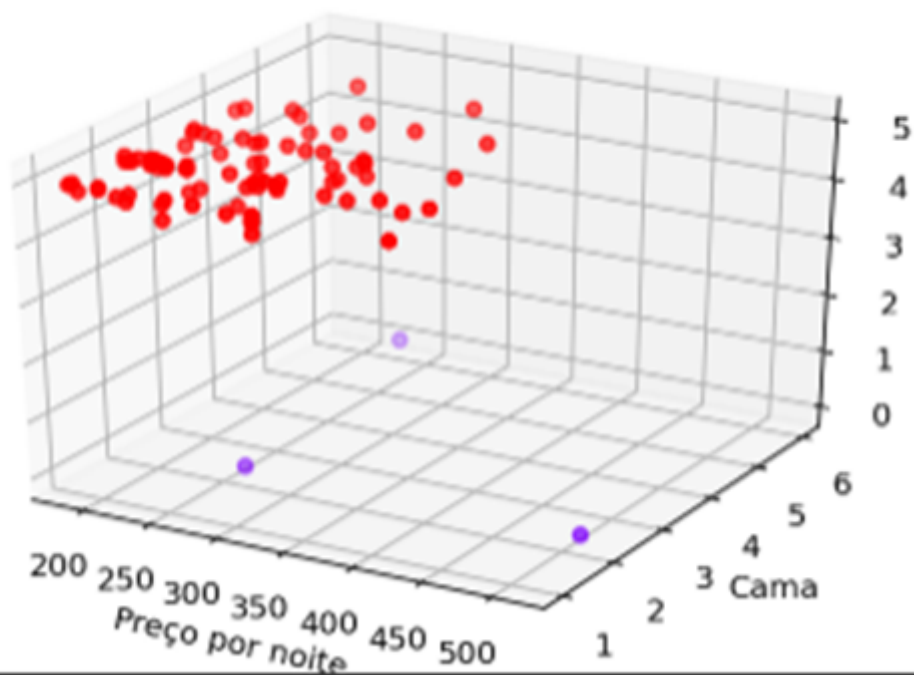
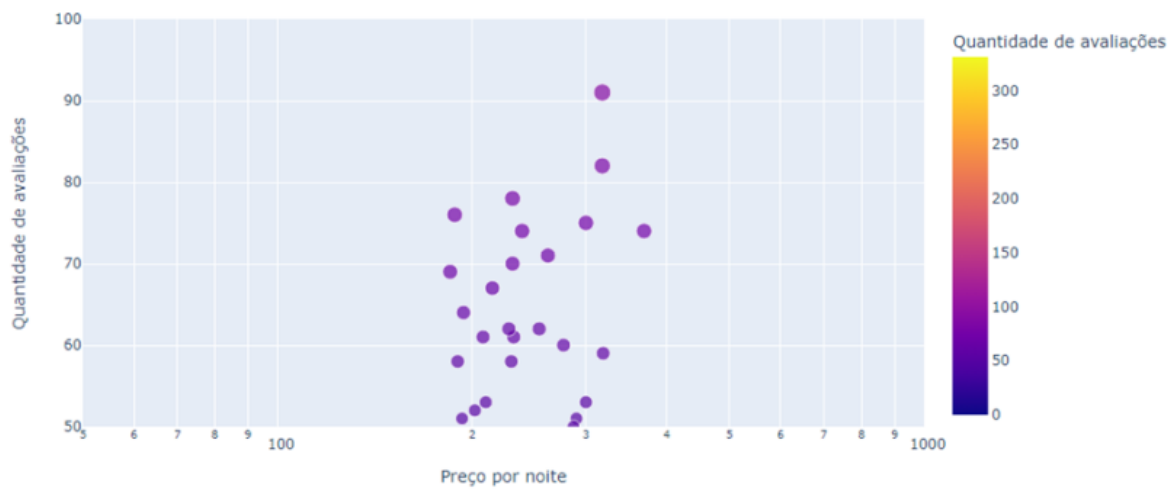


No gráfico acima, que mostra a porcentagem de preços das 13 primeiras acomodações, é possível ver que o apartamento em João Pessoa possui o maior custo financeiro, enquanto o quarto inteiro em Loteamento Bela Vista é o mais barato.



No gráfico acima, que mostra as acomodações que receberam mais avaliações, o condomínio em João Pessoa tem aproximadamente 120 avaliações, podendo ser boas ou ruins, e o apartamento em Tambaú possui cerca de 80 avaliações.

Relação entre preço por noite e quantidade de avaliações



Conclusão

Pode-se concluir que João Pessoa, sendo uma cidade turística, possui opções de imóveis considerados em conta, levando em consideração que a média de aluguel mais caro é de R\$414,00, e a média de aluguel mais barato é de R\$55,25.

