**Relatório de análise estatística**

1. **Resposta a Perguntas de negócio:**
2. **Supondo que uma pessoa esteja pensando em investir em um apartamento para alugar na plataforma, onde seria mais indicada a compra?**

Como primeira opção, o local mais adequado seria Manhattan, apesar de possuir os imóveis com aluguéis mais caros, a cidade também conta com uma das menores porcentagens de disponibilidade maior que 100 dias ( que é quando o imóvel não fica alugado por mais de 100 dias).

Como segunda opção, o local mais adequado seria o Brooklyn, apesar de não ter aluguéis tão caros quanto Manhattan, o Brooklyn também conta com uma das menores porcentagens de disponibilidade maior que 100 dias.

1. **O número mínimo de noites e a disponibilidade ao longo do ano interferem no preço?**

De acordo com análises gráficas, não foi encontrado relações significativas de preço com número mínimo de noites ou disponibilidade ao longo do ano (mais informações na seção 8.

1. **Existe algum padrão no texto do nome do local para lugares de mais alto valor?**

Existem algumas palavras que são mais recorrentes em anúncios com preços mais caros, essas são apresentadas na tabela 3 e 4 na seção 9.

1. **Objetivo:**

Analisar os dados de aluguéis temporários de imóveis e quartos em Nova York, identificando padrões e tendências que afetem o preço.

1. **Tratamento de Dados:**

O método do **intervalo interquartil** foi utilizado para identificar outliers, estes foram removidos quando ultrapassam (quartil 3 + (1,5 \* **intervalo interquartil**)), a tabela 1 apresenta a quantidade de outliers removidos.

Interface gráfica do usuário, Tabela

Descrição gerada automaticamente

Tabela

Os outliers da coluna “price” são, em maior parte, provenientes da região de Manhattan (Figura 2), indicando serem dados reais, porém para realizar a analisar de forma mais simplificada, por enquanto, estes dados serão desconsiderados.

1. Gráfico, Histograma

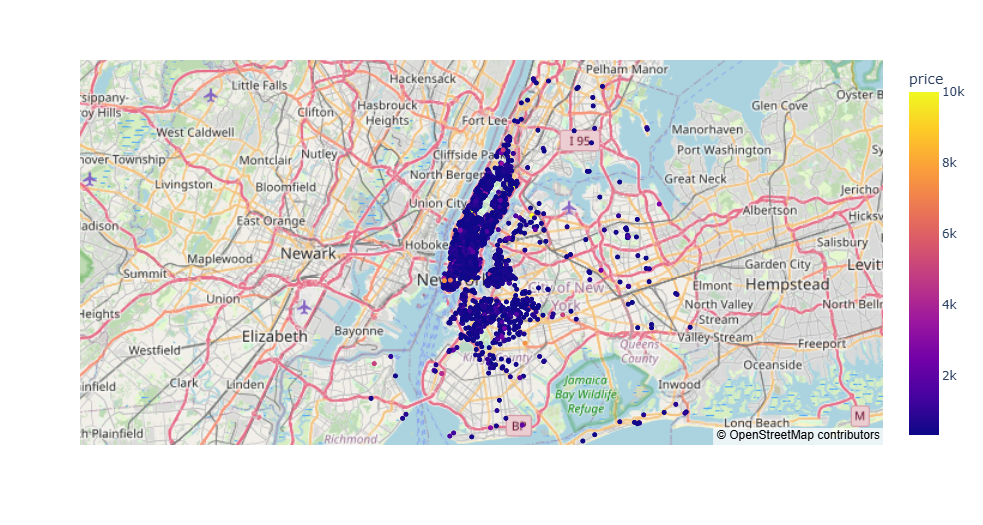
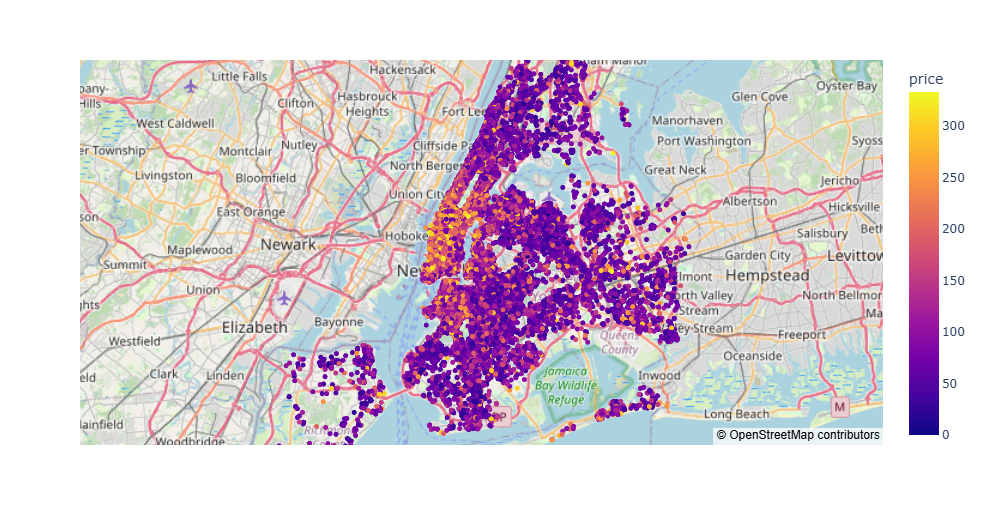
   Descrição gerada automaticamente**Análise de price por bairro\_group:**

Figura 3 – Valor de Aluguel por Bairro

Figura 1 - Localização de price sem outliers

Figura 2 - Localização de price com outliers

Manhattan e Brooklyn são os bairros com maior quantidade de imóveis/quartos para alugar, além disso apresentam maior diversidade de preço (Figura 3), possuindo mais imóveis/quartos de aluguél mais caros se comparado ao Queens, Staten Island e Bronx, Isso pode demonstrar que Manhattan e Brookly podem ser mais receptíveis a aluguéis mais caros se comparado aos outros bairros.

1. Gráfico, Histograma

   Descrição gerada automaticamente**Analisando price por room\_type:**

Figura 3 - Preço do Aluguel por Tipo de Quarto

A maioria dos aluguéis são Entire home/apt e Private room. os Entire home/apt possuem boa diversificação de preço, porém quando se trata de Private room e Shared room observa-se que esses são a minoria quando o valor tende a ser maior que 200 (Figura 4), isso pode dificultar a inclusão de aluguéis mais caros para esses tipos de quarto devido a concorrência dos aluguéis mais baratos.

1. **Analisando price por room\_type e bairro\_group:**

Quando observamos o preço dos aluguéis por room\_type e bairro\_group, observamos que Manhattan tem os aluguéis mais caros, independentemente do tipo de imóvel/quarto (Figura 5).

Diagrama

Descrição gerada automaticamente

Figura - Preço por Bairro e Tipo de Quarto

1. Tabela

   Descrição gerada automaticamente**Relação de Disponibilidade de Reserva com room\_type e bairro\_group:**

Tabela 2 - Quantidade de imóveis que ficaram alugados por mais de 100 dias e menos de 100 dias ao ano

Avaliando as disponibilidades (Tabela 2), é possível observar que Manhattan e Brooklyn tem os imóveis/quartos com a menor porcentagem de mais de 100 dias disponíveis, isso indica que comparado aos outros imóveis/quartos eles possuem uma recorrência maior de pessoas alugando, em contra partida State Island possuem a maior porcentagem de disponibilidade maior que 100 dias, demostrando não ser uma região com muita recorrência de pessoas alugando.

1. Tabela

   Descrição gerada automaticamenteGráfico, Gráfico de dispersão

   Descrição gerada automaticamente**Relação de Preço com número mínimo de noites e disponibilidade ao longo do ano:**

Figura 6 - Preço por Número mínimo de noites

Figura 6 - Preço por Disponibilidade

Analisando a relação de Preço com número mínimo de noites e disponibilidade ao longo do ano não foi encontrado relações significativas.

1. **Palavras recorrentes em anúncios de valor mais caro:**

Para encontrar as palavras associadas aos anúncios com aluguéis mais caros, foi utilizada a coluna nome da fonte de dados fornecido para gerar bigramas e trigramas, as palavras mais recomentes são apresentadas abaixo.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Bigramas | | |  | Trigramas | | | |
| Sequência de palavras | | Frequência |  | Sequência de palavras | | | Frequência |
| 2 | bedroom | 501 |  | sonder | stock | exchange | 158 |
| 1 | bedroom | 444 |  | upper | east | side | 153 |
| central | park | 438 |  | lower | east | side | 114 |
| east | village | 419 |  | upper | West | side | 108 |
| west | village | 379 |  | near | central | park | 97 |
| east | side | 281 |  | 2 | bedroom | apartment | 82 |
| one | bedroom | 266 |  | 1 | bedroom | apt | 73 |
| times | square | 256 |  | 2 | bedroom | apt | 68 |
| bedroom | apartment | 252 |  | new | york | city | 65 |
| 3 | bedroom | 206 |  | 1 | bedroom | apartment | 61 |
| bedroom | apt | 201 |  | hell | "s" | kitchen | 59 |
| upper | east | 191 |  | wyndham | midtown | 45 | 58 |
| new | york | 175 |  | bed | 2 | bath | 53 |
| apt | . | 172 |  | near | times | square | 50 |
| 2 | bed | 158 |  | guest | service | fee | 49 |
| sonder | stock | 158 |  | 2 | bed | 2 | 44 |
| stock | exchange | 158 |  | one | bedroom | apartment | 39 |
| 1 | br | 152 |  | heart | east | village | 38 |
| park | slope | 150 |  | east | village | apartment | 37 |
| 2 | br | 148 |  | heart | west | village | 37 |

Tabela 4 – Bigramas menos recorrentes

Tabela 3 – Bigramas mais recorrentes