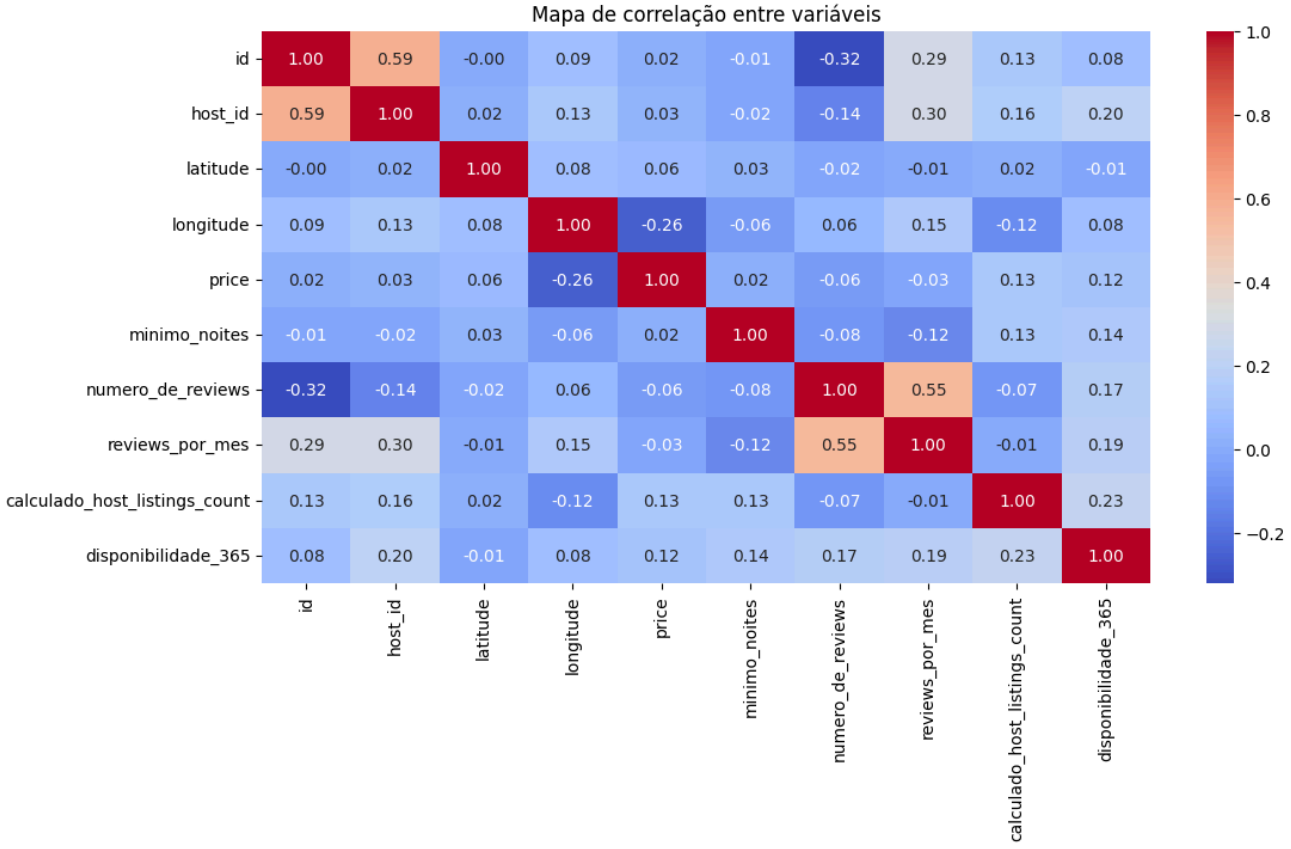
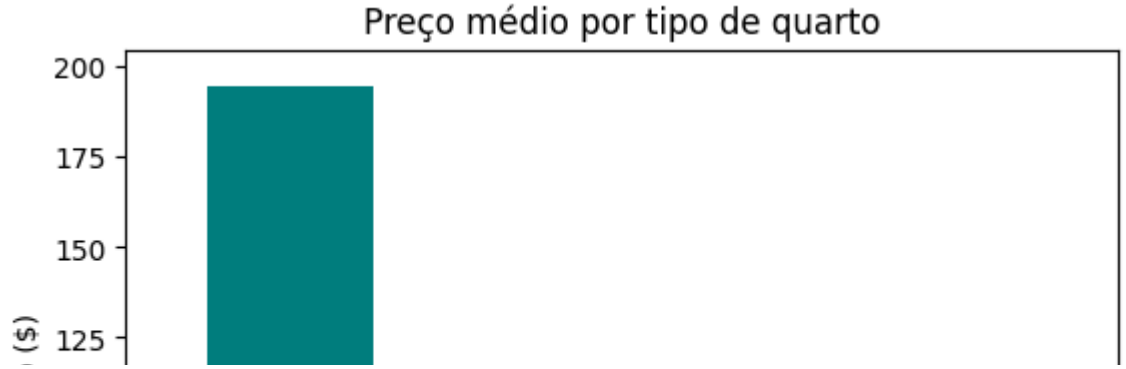
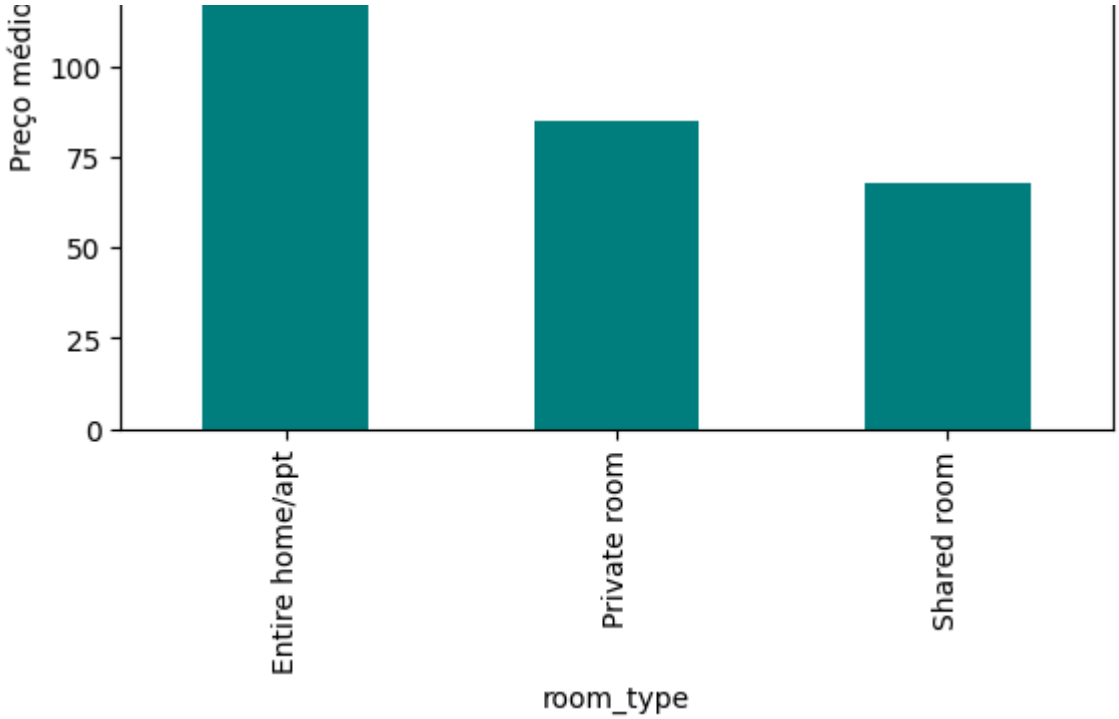
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Carregar os dados
url = "https://raw.githubusercontent.com/caiooooooo3/Desafio_Lighthouse/main/teste_indiciu
data = pd.read_csv(url)

# Visão geral dos dados
data.info()
data.describe()

# Remover outliers
data = data[data['price'] <= 1000]

# Análise de correlações
plt.figure(figsize=(12, 6))
sns.heatmap(data.corr(numeric_only=True), annot=True, cmap="coolwarm", fmt=".2f")
plt.title("Mapa de correlação entre variáveis")
plt.show()

# Análise de preço médio por tipo de quarto
data.groupby('room_type')['price'].mean().plot(kind='bar', color='teal')
plt.title("Preço médio por tipo de quarto")
plt.ylabel("Preço médio ($)")
plt.show()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 48894 entries, 0 to 48893
Data columns (total 16 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   id                              48894 non-null  int64
 1   nome                            48878 non-null  object
 2   host_id                         48894 non-null  int64
 3   host_name                       48873 non-null  object
 4   bairro_group                    48894 non-null  object
 5   bairro                          48894 non-null  object
 6   latitude                        48894 non-null  float64
 7   longitude                       48894 non-null  float64
 8   room_type                       48894 non-null  object
 9   price                           48894 non-null  int64
 10  minimo_noites                   48894 non-null  int64
 11  numero_de_reviews               48894 non-null  int64
 12  ultima_review                   38842 non-null  object
 13  reviews_por_mes                 38842 non-null  float64
 14  calculado_host_listings_count   48894 non-null  int64
 15  disponibilidade_365             48894 non-null  int64
dtypes: float64(3), int64(7), object(6)
memory usage: 6.0+ MB
```








```
from google.colab import files
files.download("EDA.ipynb")
```