

AIRBNB AND BOOKING.COM DATA ANALYSIS

Goal of the Project.

- Assess the level of popularity between Airbnb and Booking.com during the festive season.
- Conduct a comparative analysis of listing prices in the United States and the United Kingdom on both platforms.
- Compare the price disparities of listings in different locations before and after Christmas, specifically on the dates of December 1-3 and December 23-27.
- Analyze the variation in prices based on locations, identifying the areas that exhibit the highest costs, and examine how listing types impact pricing.
- Investigate the most favored types of listings based on their popularity.

Introduction to Airbnb and Booking.com

Airbnb and Booking.com are online platforms that serve as marketplaces, connecting individuals who own properties and wish to rent them out as holiday accommodations with travelers in search of temporary lodging.

As of September 2021, Airbnb boasted a remarkable presence in the market, with over 5.6 million active listings available worldwide in more than 220 countries and regions. Since its inception, Airbnb has facilitated over 750 million guest arrivals and has a vast network of over 4 million hosts spanning across the globe.

On the other hand, Booking.com is a substantial online platform valued at \$96 billion. It operates in over 230 countries and territories, positioning itself as one of the largest online travel agencies with an extensive global reach.

Investing in rental properties can be a lucrative method for generating additional income, provided you are situated in a favorable location. Nevertheless, contrary to expectations, managing rental properties requires substantial commitment. It necessitates an unwavering dedication to comprehensively understanding the market and remaining actively involved throughout the entire rental period.

Currently, one widely embraced rental strategy is centered around vacation rental properties. These properties are acquired with the sole purpose of listing them as short-term rentals on online platforms like Airbnb.

The primary reason why many real estate investors opt for vacation rentals is due to their potential to generate higher income compared to long-term rentals. However, it is crucial to note that although they tend to be more profitable, they may not automatically be the ideal choice for everyone. Several factors must be taken into consideration before making a decision.

Obtaining the Data.

Why Web Scraping?

Web scraping is the process of extracting data from various websites and parsing it. In other words, it's a technique to extract unstructured data and store that data either in a local file or in a database.

The data used in the data analysis was extracted using BeautifulSoup, a Python web scraping library that allows the parsing and scraping of HTML and XML pages. One can search, navigate, and modify data using a parser. It's versatile and saves a lot of time.

This process is well documented in python files in a [GitHub repository](#); '[Booking UK.py](#)', '[Bnb UK.py](#)', '[Bnb Scraping.py](#)'. Here's a snippet of the web scraping code:

```
8 headers = {'user-agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/104.0.0.0 Safari/537.36'}
9
10 url = "https://www.booking.com/searchresults.en-gb.html?ss=United+Kingdom&sb_entire_place=1&ssne=United+Kingdom&ssne_untouched=United+Kingdom"
11 with open('Listings UK(23-27).csv', 'w', encoding='utf8', newline='') as file:
12     thewriter = writer(file)
13
14     header = ['Name', 'Roominfo', 'Beds', 'Location', 'DiscountPrice', 'ActualPrice', 'Reviews', 'Host']
15     thewriter.writerow(header)
16
17     def scrape_page(url):
18         page = requests.get(url, headers=headers)
19         soup = BeautifulSoup(page.content, 'lxml')
20         get_data(soup)
21
22         for page in range(25, 925, 25):
23             next_url = "https://www.booking.com/searchresults.en-gb.html?ss=United+Kingdom&sb_entire_place=1&ssne=United+Kingdom&ssne_untouched=United+Kingdom"
24
25             page = requests.get(next_url, headers=headers)
26             soup = BeautifulSoup(page.content, 'lxml')
27             get_data(soup)
28             #scrape_page(next_url)
29
30     def get_data(soup):
31         for list in soup.select('[data-testid=property-card]'):
32             name = list.find('div', class_='fcab3ed991').text
33             roominfo = list.find('div', class_='cb5b4b68a4').text
34             beds = list.find('span', class_='bb58e7a787').text
35             location = list.find('span', class_='f4bd0794db b4273d69aa').text
36             discountprice = list.find('span', class_='fcab3ed991 bd73d13072').text
37             actualprice = list.find('span', class_='a0c113411d e293a04099').text
38             reviews = list.find('div', class_='b5cd09854e d10a6220b4').text
39             host = list.find('div', class_='b5cd09854e d10a6220b4').text
```

Data Cleaning and Preprocessing.

The main aim of this step is to identify and fix issues with quality and structure in the data and transform the scraped data into a more usable format for analysis.

- The data was inspected by checking duplicated listings. Null values and issues with the data types were also corrected. Each content of every column was inspected too.
- The Location and Number of beds columns contained tags similar to that of a HTML code tag. The lstrip and rstrip functions were used to get rid of the unwanted characters of the strings in the dataset.
- Geopy (used to locate the coordinates of addresses across the globe), was used to generate the coordinates from the scraped locations for easier data visualization. I used this code to extract

the coordinates:

```
#Initialize Nominatim API
geolocator = Nominatim(timeout=10, user_agent= "geoapiExercises")
```

```
#To avoid getting the Attribute error, create functions that make sure there are no Null values being passed in the geolocator
def get_latitude(x):
    if hasattr(x, 'latitude') and (x.latitude is not None):
        return x.latitude

def get_longitude(x):
    if hasattr(x, 'longitude') and (x.longitude is not None):
        return x.longitude
```

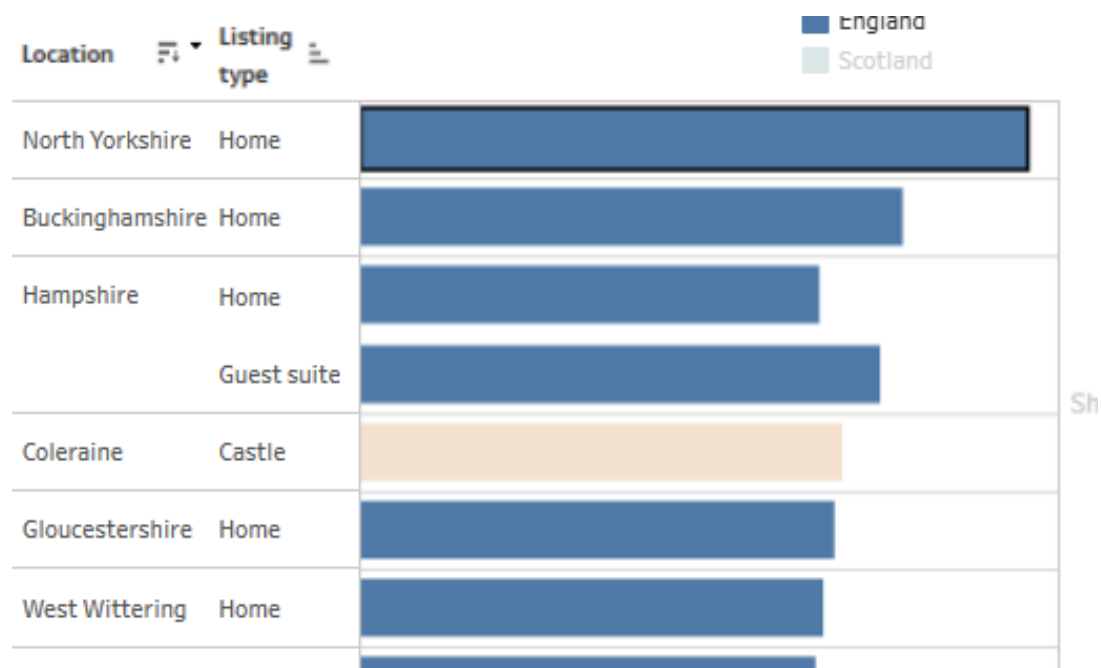
```
#Use the geolocator function to get the Latitudes and Longitudes

geolocate_column = df_clean['Location'].apply(geolocator.geocode)
df_clean['Latitude'] = geolocate_column.apply(get_latitude)
df_clean['Longitude'] = geolocate_column.apply(get_longitude)
```

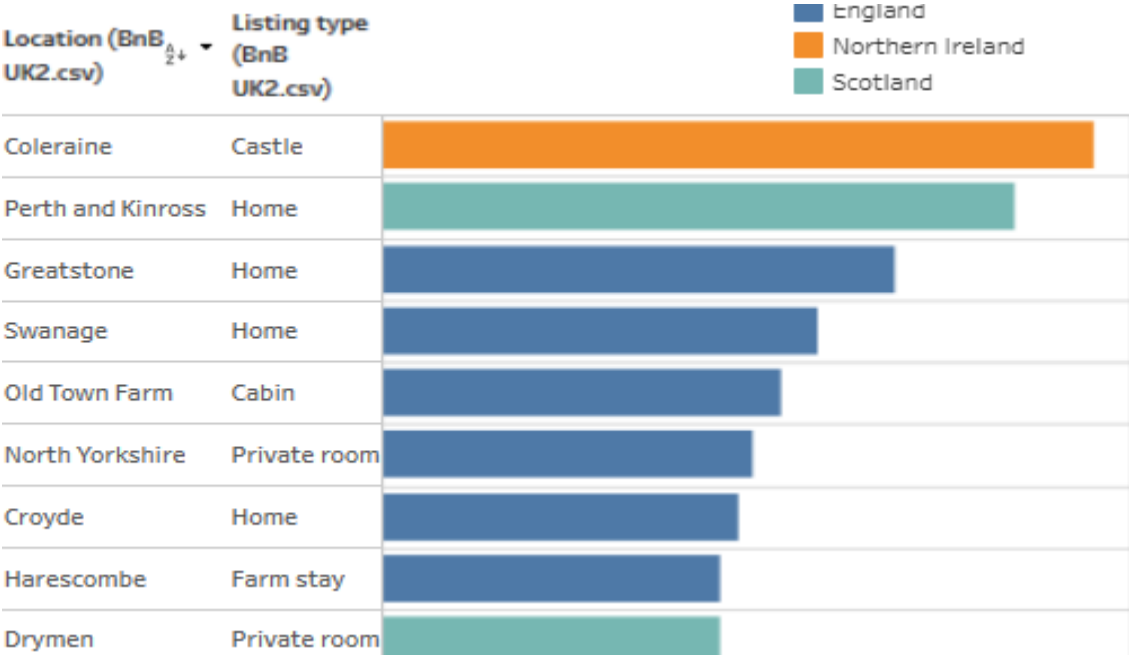
The full data cleaning steps and procedures are documented on my [GitHub repository](#).

Price per Location (Airbnb).

Between the dates of December 1st and 3rd, in the United Kingdom, England emerged as the region with the highest number of listings on Airbnb, consequently making it the most expensive. For instance, in North Yorkshire, a home was listed at an average rate of \$980 per night, while in Hampshire, a guest suite had an average listing price of \$760 per night. As shown [below](#):

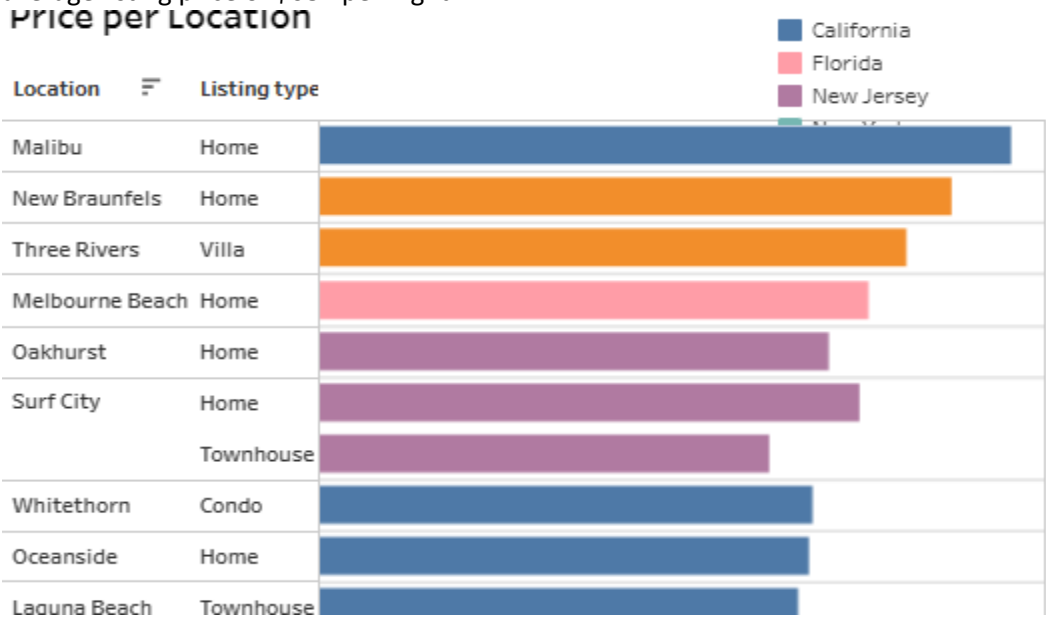


During the period of December 23rd to 27th, North Ireland featured the highest-priced listing, with a castle in Coleraine being listed at an average rate of \$865 per night. On the other hand, England had the highest number of listings, with a home in Greatstone being listed at an average price of \$620 per night.



Based on the comparison of listing prices between the two date ranges on Airbnb, the findings indicate that England had the highest number of listings, predominantly consisting of homes. This observation can be attributed to the approaching festive season, which resulted in an increased availability of vacation rentals.

During the period of December 1st to 3rd, in the United States (US), California and New Jersey emerged as the states with the most significant number of listings. Notably, in Malibu, California, a home was listed at an average price of \$984 per night, while in Laguna Beach, California, a townhouse had an average listing price of \$682 per night.



During the period of December 23rd to 27th in the United States, California maintained its position as the state with the highest number of listings. However, there was a notable shift in the types of listings, with cabins becoming more prevalent compared to earlier in the month, while the number of homes decreased. The most expensive listing during this period was a townhouse in Laguna Beach, California, with an average price of \$979 per night. Interestingly, this price was approximately \$300 higher than what the townhouse was listed for between December 1st and 3rd. It is important to note that this price difference is not unique to California but is observed in other locations as well. The price hike can be attributed to the festive season, specifically the Christmas holidays, which leads to an increase in prices across various listings. The change can be shown [here](#).

Price per Location (Booking.com)

During the period of December 1st to 3rd, England maintained its position as the region with the highest number of listings on Booking.com in the UK. Notably, the majority of these listings were apartments. For example, in Camden, London, an apartment was listed at an average price of \$895 per night, while a private suite in the same area had an average price of \$614 per night. In Scotland, an apartment was listed at an average rate of \$263 per night. These figures demonstrate that England tends to have higher-priced listings, which may explain its prominence on Booking.com.

From December 23rd to 27th, England continued to lead in terms of the number of listings on Booking.com, with holiday homes being the most prevalent. In Bridlington, a holiday home was listed at an average price of \$1,421, while in Chichester, it cost \$1,086. This price disparity is consistent with the trends observed on Airbnb, where listings tend to be more expensive during the Christmas holidays compared to the beginning of the month.

In the United States on Booking.com, Florida comprised the highest number of listings during the period of December 1st to 3rd. The predominant listing types were private suites, with a few apartments also available. For example, in Miami, a private suite was listed at an average price of \$619 per night, while in Fort Lauderdale, an apartment had an average listing price of \$370 per night. These prices were comparatively more reasonable when compared to the rates on Airbnb during the same period.

From December 23rd to 27th in the US, the majority of listings were concentrated in Florida, with a few also present in Texas. For instance, in North Miami Beach, Florida, an apartment was listed at an average price of \$834 per night, while in Avalon, California, an apartment had an average price of \$700. Based on the price comparison, it appeared that Florida had a higher concentration of higher-priced listings compared to Texas and California.

Different types of listings.

The United Kingdom offers a diverse selection of listing types, encompassing castles, guest suites, homes, domes, cottages, and even farm stays. The range of available listings undergoes changes during the festive season (December 23rd to 27th), with additional types of accommodation becoming available, such as lofts and bungalows.

Similarly, in the United States, there is a wide array of listing types, surpassing the variety found in the UK. These include townhouses, villas, towers, homes, apartments, yurts, bungalows, chalets, domes, condos, and cabins. The listing options in the US offer a diverse selection to cater to various preferences and needs.

CONCLUSION.

In conclusion, the analysis of Airbnb and Booking.com listings during the periods of December 1st to 3rd and December 23rd to 27th in the United Kingdom and the United States reveals interesting trends.

In the United Kingdom, England consistently had the highest number of listings on both platforms, with a significant presence of homes and apartments. The prices tended to be higher during the festive season, particularly in North Yorkshire and North Ireland. Booking.com showcased England's prominence with higher-priced listings, while Airbnb had a wider variety of listing types available, including castles, guest suites, homes, and more.

In the United States, California and Florida emerged as the leading states with a substantial number of listings. Both platforms showed an increase in cabin listings during the Christmas holidays, with California maintaining its position as the state with the highest number of listings. The price comparison between the two periods indicated an overall price hike during the festive season, with significant differences observed in Laguna Beach, California. Florida, particularly Miami, had more reasonable prices compared to Airbnb.

Overall, the analysis highlights the impact of the festive season on listing availability and prices, with England and California being prominent in terms of both the number of listings and higher-priced options. The diversity of listing types in both countries provides a wide range of choices for travelers.