

EDINBURGH AIRBNB

Scenario

I am working as a junior data analyst for a business intelligence consultant in Edinburgh, UK. I have been on the job for six months, and my boss has now asked me to lead a project for a brand-new client.

The new client is interested in market research of Airbnb listings in the city of Edinburgh. The client currently owns a rental in the Leith area and wants to buy a second property.

The client has the following questions:

Questions to answer:

1. What are the most common room types and how many people do they accommodate?
2. Which five neighbourhoods have the highest number of listings? What is the average rental pricing in each neighbourhood?
3. What is the average availability in 30 days for each one of the five neighbourhoods? What is the average occupancy rate and what is the gross monthly income?
4. Which are the most profitable neighbourhoods to invest in for a property that can accommodate 4 people? Ideally neighbourhoods with more than \$3000 of gross monthly income.
5. If we were to look at each neighbourhood from a broader view, what information can we get for the listings in the broader areas?
6. How many competitors exist in the Leith area? Find entire homes or apartments that have a BBQ grill and parking in their amenities in the Leith area.
7. What is the price trend per night of the accommodations over time? When should I focus on advertisements?
8. Is the price following a similar trend for the entire home/apartment type and private rooms?

Data Cleaning

All data cleaning was performed on Jupyter Notebook.

1. Listings

I downloaded the listings data which was released on the 10th of June 2022 on the website <http://insideairbnb.com/get-the-data> as a primary data source. I kept fields 'id', 'listing_url', 'name', 'host_response_time', 'host_total_listings_count', 'neighbourhood', 'neighbourhood_cleansed', 'latitude', 'longitude', 'property_type', 'room_type', 'accommodates', 'bedrooms', 'beds', 'amenities', 'price', 'minimum_nights', 'maximum_nights', 'availability_30', 'reviews_per_month'. Then, I filled null values in the 'reviews_per_month' field with 0 and saved the data frame as listing_cleansed.csv.

2. Calendar

I downloaded the calendar file which was released on the same date along with the calendar files from the archived 14th of March 2022, 12th of December 2021, and 16th of September 2021, thus four calendar files in total.

I opened all calendar CSV files and checked the size and dates of each one. I then compiled all the calendar CSV files into a data frame. I deleted the adjusted_price field, deleted null rows and changed datatypes of listing_id, minimum_nights, maximum_nights and date fields. Following that, I removed duplicates based on the combination of listing_id and date fields. The price field contained commas and dollar symbols; thus, I removed them and changed the datatype to float. I saved the data frame as all_calendar_data.csv.

3. Broader areas

I downloaded 17 CSV files from <https://statistics.gov.scot/atlas/resource?uri=http%3A%2F%2Fstatistics.gov.scot%2Fid%2Fstatistical-geography%2FS12000036>. Each file represents an electoral ward of Edinburgh and the neighbourhoods that belong to each ward. I chose electoral wards to represent the broader areas of Edinburgh as it is easier to visualize 17 areas on a map. I concatenated all 17 CSV files in a dataset and deleted the first column as I didn't intend to use it. Then, I trimmed the 'Feature_Name' rows and removed duplicates. Finally, I saved the data frame as broader_areas.csv.

Data Analysis

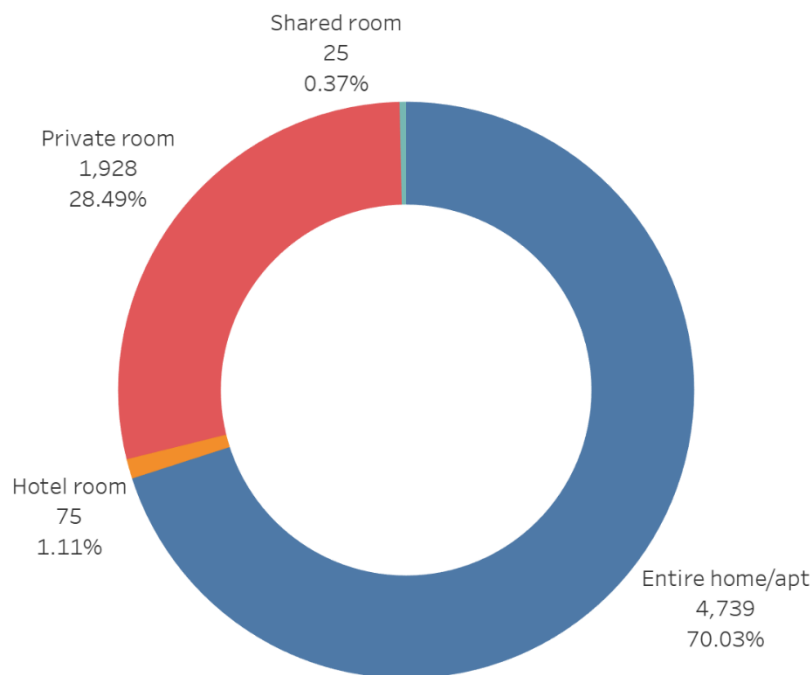
For exploring the data and answering the questions, I used SQL. For visualizing data, I used Tableau.

The Edinburgh Airbnb dashboard can be found at:

<https://public.tableau.com/app/profile/andreas.zachariou4878>

I first started with the listings dataset. I wanted to check the number and percentage of each room type.

1. What are the most common room types and how many people do they accommodate?

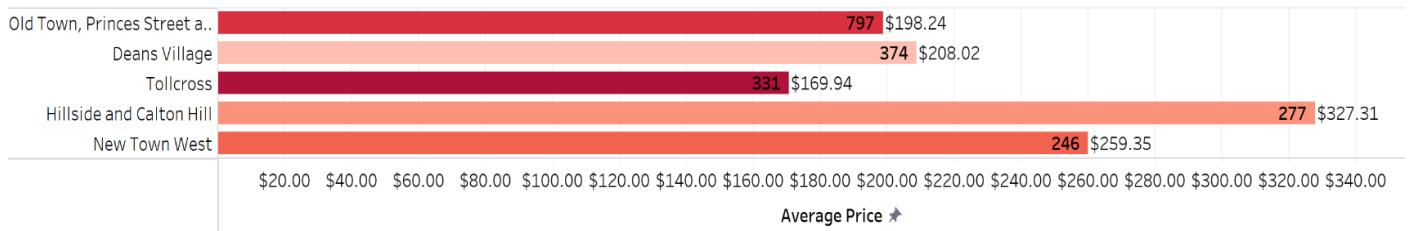


The entire home/apartment has the largest proportion of room types with 70% and an average accommodate number of 4.32. The second highest proportion belongs to private rooms with approximately 28.5% and can accommodate 2 people. Hotel rooms have only 75 listings and a 1.1% proportion that can accommodate 2.39 people on average. Shared rooms were the least popular type with only 25 rooms and an average of 5.28 accommodates. The average accommodation number for shared rooms is not a representative result as a higher sample size is required.

Then, I searched for the top five neighbourhoods with the highest listing number and the average price for each neighbourhood.

2. Which five neighbourhoods have the highest number of listings? What is the average rental pricing in each neighbourhood?

Highest Accommodation Number Neighbourhoods



Old Town, Princes Street and Leith Street neighbourhood has the highest number of listings since it is located in the heart of the city. However, the most expensive average price per night out of the five neighbourhoods is in Hillside and Calton Hill.

3. What is the average availability in 30 days for each one of the five neighbourhoods? What is the average occupancy rate and what is the gross monthly income?

Neighbourhoods	Average Availability in 30 days	Occupancy Rate	Gross Monthly Income
Hillside and Calton Hill	5.25	82.51	8819.24
New Town West	6.78	77.39	5542.97
Old Town, Princes Street and Leith Street	5.86	80.46	4755.4
Deans Village	7.54	74.88	4161.26
Tollcross	6.29	79.03	3960.79

The occupancy rate of a rental is the number of days that the property is rented out to guests during a specific period. At this point, we can observe that all the neighbourhoods have an increased occupancy rate since anything over 65% can be considered high according to a report by AllTheRooms.

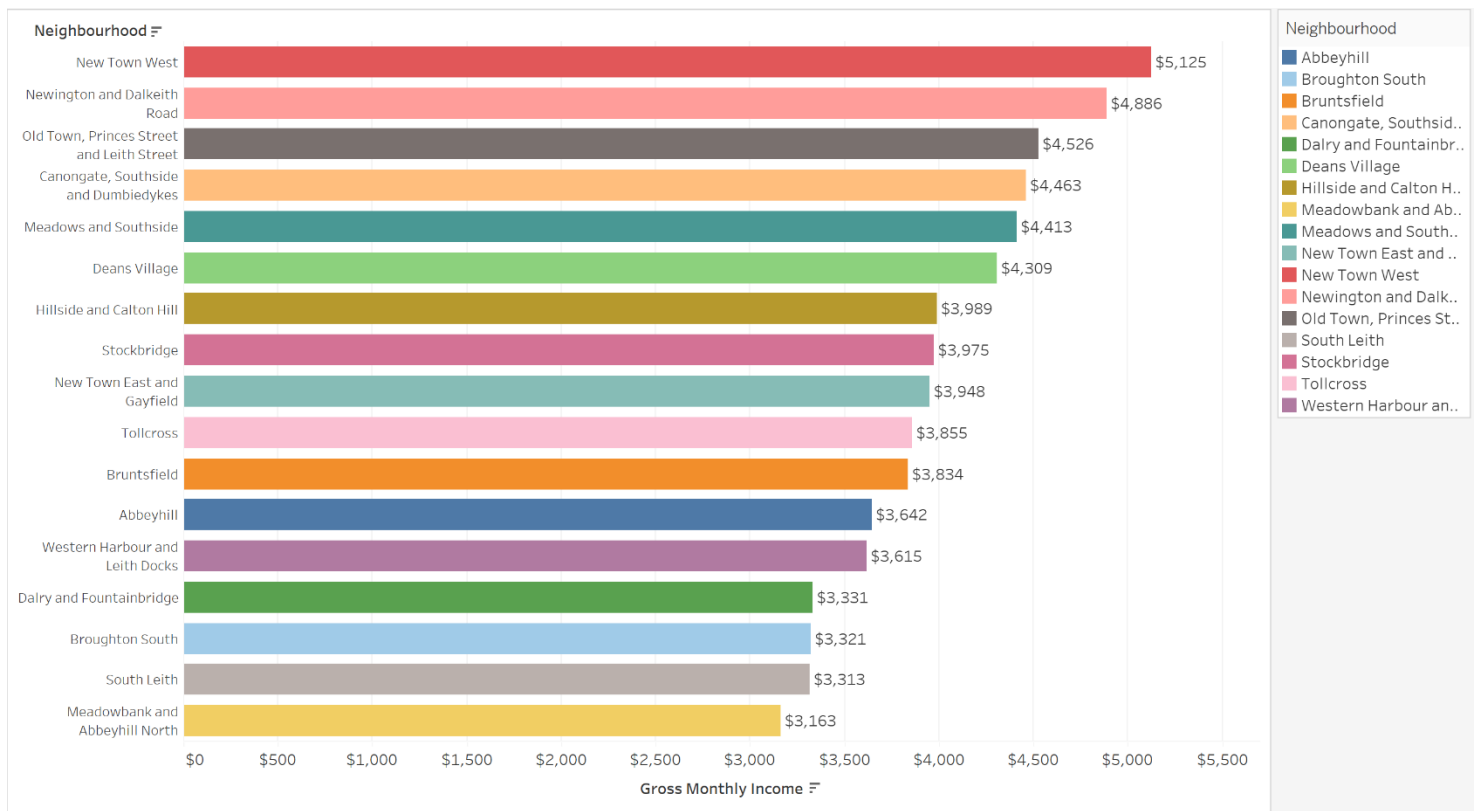
Source: <https://www.alltherooms.com/analytics/average-airbnb-occupancy-rates-by-city/>

The higher the occupancy rate the greater the income. If you are having a rental in Hillside and Calton Hill you can expect on average a gross monthly income of \$8819.24.

The numbers above are not completely accurate, as hosts sometimes can set their listings unavailable due to different adjustments to their page. Therefore, these numbers are just a snapshot of the specific time that scraping took place. Rentals that had availability_365 as 0 were excluded.

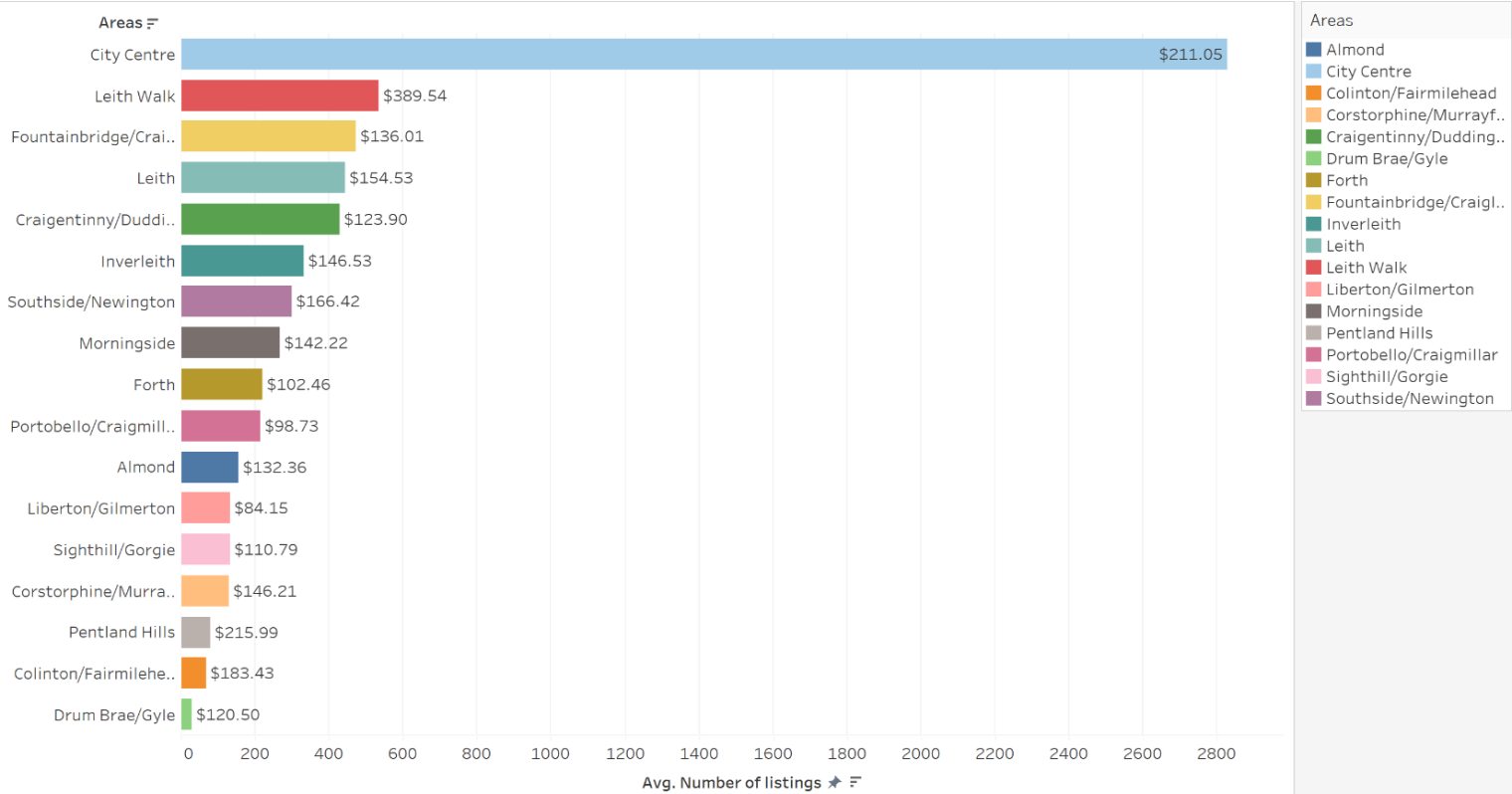
- Which are the most profitable neighbourhoods to invest in for a property that can accommodate 4 people? Ideally neighbourhoods with more than \$3000 of gross monthly income.

17 neighbourhoods are potentially ideal for buying a property in and can provide a gross income of more than \$3000 per month. New Town West is the best neighbourhood that provides \$5125 per month and it is located in the city centre, while in Meadowbank and Abbeyhill North you can expect on average \$3163 per month, probably because it is a bit far from the centre and most attractions.



After joining my listings data with broader areas data, I have assigned the 111 neighbourhoods to 17 areas.

5. If we were to look at each neighbourhood from a broader view, what information can we get for the listings in the broader areas?



Sorting the data by the number of accommodations, we can notice that almost half (49.71%) of the listings are located in the City Centre and Leith Walk. If we include the accommodations from Fountainbridge/Craiglockhart, Leith, Craigentinny/Duddingston and Inverleith we have approximately three-quarters of the listings. This shows us that the majority of the listings are either in the city centre or surrounding the centre.

Area	Number of Accommodations	Rolling_Sum	Area Percentage	Rolling Percentage
City Centre	2830	2830	41.82%	41.82%
Leith Walk	534	3364	7.89%	49.71%
Fountainbridge/Craiglockhart	473	3837	6.99%	56.7%
Leith	443	4280	6.55%	63.25%
Craigentinny/Duddingston	429	4709	6.34%	69.59%
Inverleith	332	5041	4.91%	74.49%
Southside/Newington	299	5340	4.42%	78.91%
Morningside	266	5606	3.93%	82.84%
Forth	221	5827	3.27%	86.11%
Portobello/Craigmillar	213	6040	3.15%	89.26%
Almond	156	6196	2.31%	91.56%
Liberton/Gilmerton	133	6329	1.97%	93.53%
Sighthill/Gorgie	131	6460	1.94%	95.46%
Corstorphine/Murrayfield	130	6590	1.92%	97.38%
Pentland Hills	78	6668	1.15%	98.54%
Colinton/Fairmilehead	69	6737	1.02%	99.56%
Drum Brae/Gyle	30	6767	0.44%	100%

A map was also created for more comprehensive visualization. I first selected some fields from listings data and joined them with the broader areas data based on the neighbourhood_cleansed field. Then, in Tableau, I created a relationship between my joined table and a GeoJSON file which contained the geographical locations of each broader area.

6. How many competitors exist in the Leith area? Find entire homes or apartments that have a BBQ grill and parking in their amenities in the Leith area.

There are currently 443 listings in the Leith area with an average price of \$154.53 per night.

Number of Listings	Average Price	broader_area
443	154.53	Leith

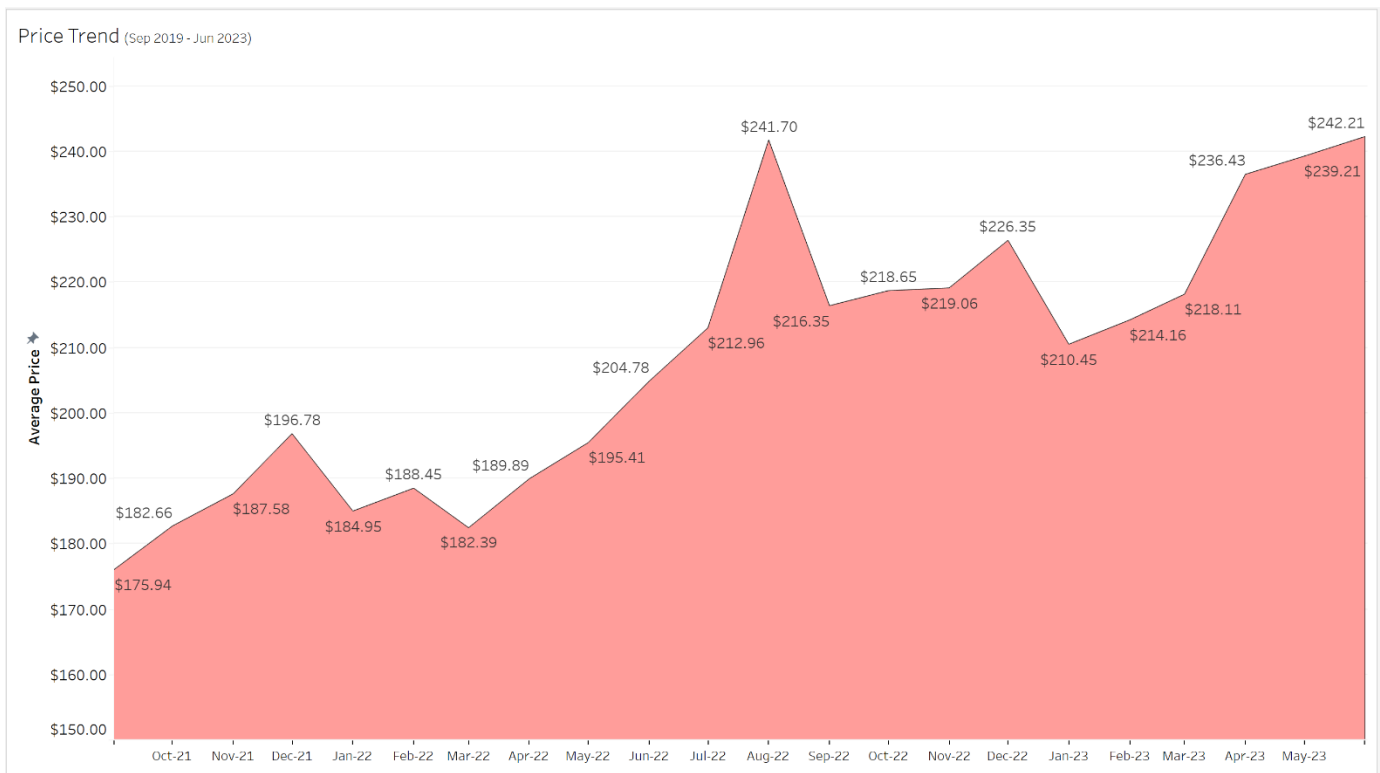
After searching for entire homes or apartments in the Leith area with a BBQ grill and coffee maker included in their amenities, I found six rentals.

id	listing_url	Price
50714986	https://www.airbnb.com/rooms/50714986	216
18006548	https://www.airbnb.com/rooms/18006548	181
43660414	https://www.airbnb.com/rooms/43660414	167
622809716778999272	https://www.airbnb.com/rooms/622809716778999272	152
37380314	https://www.airbnb.com/rooms/37380314	120
16423945	https://www.airbnb.com/rooms/16423945	115

The client will use the table above to compare each listing with her rental in order to investigate prices, amenities, design etc. and become more competitive.

7. What is the price trend per night of the accommodations over time? When should I focus on advertisements?

From the calendar data, I found the average price per month for all the listings from September 2021 to June 2023. I visualized that data as an area chart.



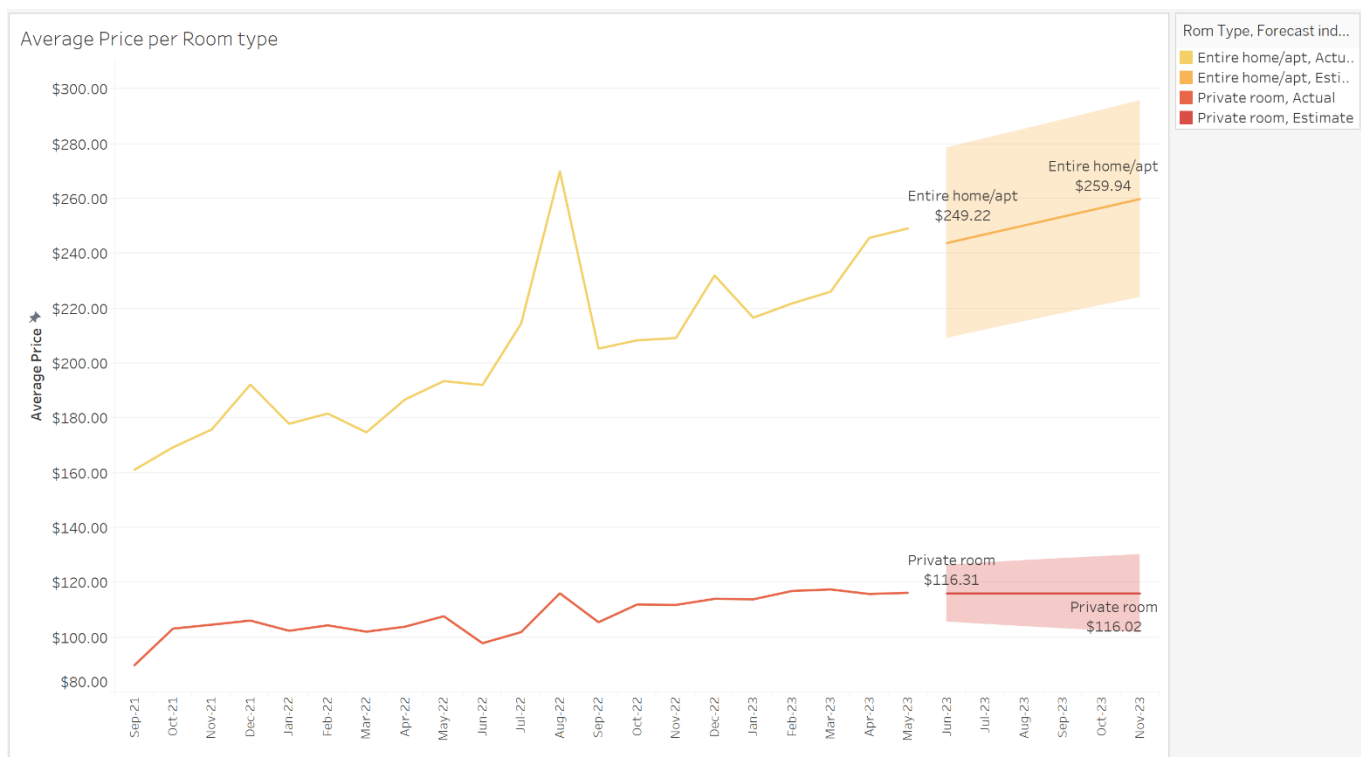
If we compare September 2021 with September 2022, the price per night increased by 23%. The price continues to have an ascending trend, after a decline from the upcoming Christmas season of 2022, until it sets slightly higher than the current year's high to \$242.21 (June 2023). Price has a tendency to increase every March for the two presented years, but this cannot be verified as we require more data. If this hypothesis is correct, the average price rises gradually in the summer, probably because the weather gets "hotter" (11°C - 19°C) and reaches a peak in August when the Edinburgh Festival Fringe takes place.

Source:

[https://www.introducingedinburgh.com/climate#:~:text=Summer%20in%20Edinburgh,C\)%20in%20July%20and%20August.](https://www.introducingedinburgh.com/climate#:~:text=Summer%20in%20Edinburgh,C)%20in%20July%20and%20August.)

If a year-round marketing strategy is followed, then advertisements should mainly take place before summer, in April and May when demand is not as high as in the summer. November and December are also considered good periods to advertise since there are graduation ceremonies and the Christmas season.

8. Is the price following a similar trend for the entire home/apartment type and private rooms?



The two graphs are following a similar trend during the summer of 2022 however, private rooms have a lower peak than entire homes/apartments in August. An explanation might be that private rooms can accommodate on average 2 people, while entire homes/apartments 4.3 people, thus the price change in the latter, has a higher intensity. The entire home/apartment type will take an upward movement the following year and the forecast is supporting that prices will increase at least \$10 from May 2023 to November 2023. In contrast, private rooms will have a stable price from October 2022 until November 2023.