What are your best options in Boston Airbnb listings?

Introduction

I came across a dataset that includes interesting details about Airbnb listings located in Boston City, MA. If you are planning a visit to Boston soon, you might be interested to read some insights about listings there!

What was I looking for?

I had 3 questions in mind, and I was exploring the data to find answers for them, here are my questions:

1- Extracting the listings with the highest # of reviews (top 10). Which one got the highest % of positive reviews?
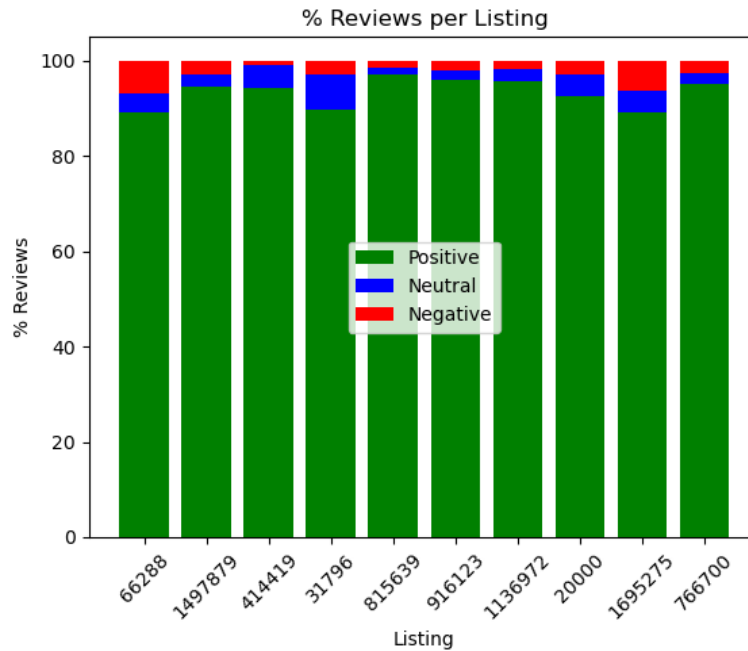
2- What are the top 3 most frequent neighborhoods or zip codes in Boston? (zip codes with highest number of listings).

3- How the listing price correlates to the property_type? What types are the most expensive?
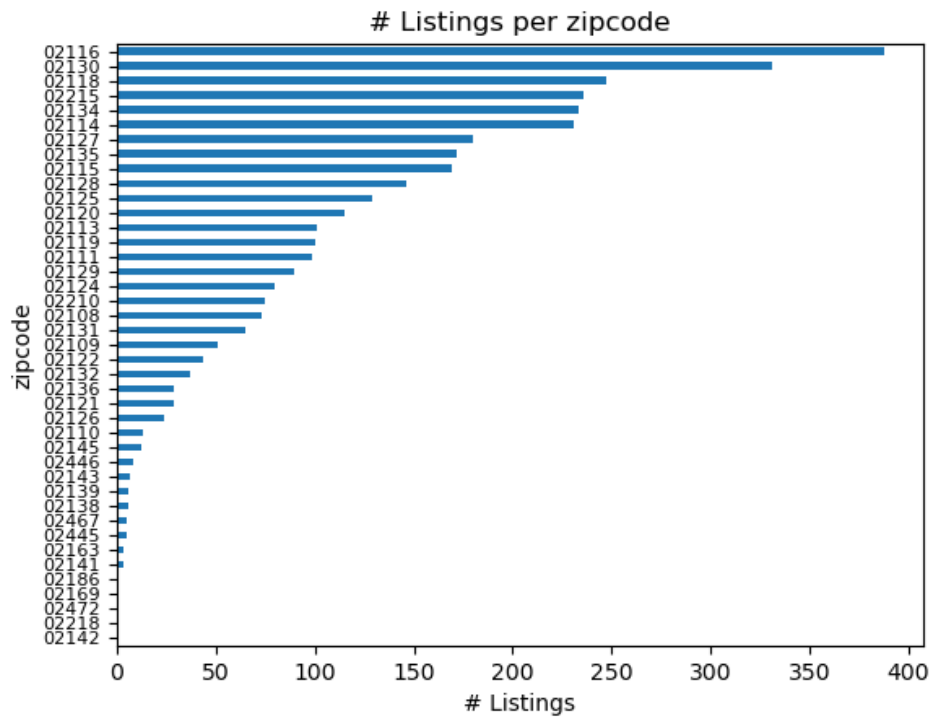
Let's see what I found out!

My findings

First question: To answer it, I extracted the data of the 10 listings that has the highest number of reviews and calculated the % of reviews for each category [positive, neutral, negative]. Then I plotted a chart that shows the % of each category of review for these listings. Let's see the outcome
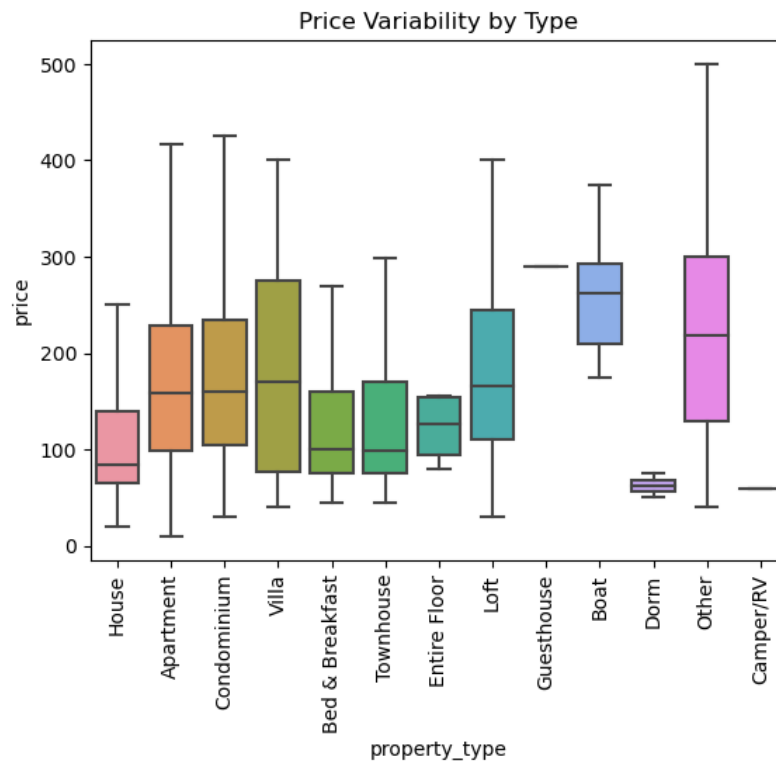
% Reviews per Listing

As you can see, the listing with ID starting with 815 did an amazing job! It has the highest % of positive reviews among other listings.

For the **second question** regarding the most frequent neighborhoods or zip codes, I started by exploring the neighborhood column, but it had around 15% missing values with no other source to impute the missing values, so I used the zip code column to help me answer my question. Let's see the zip codes count I found

Listings per zipcode

It seems that the first 2 zip codes are the most popular ones in terms of # listings.

Finally, I moved to my **third question** to understand the correlation between property type and the listing price. I used a boxplot to see the variability in the price data for each type of listing

Generally, we can use the horizontal line in the middle of the boxplot to compare between median price of different categories.

It was found that Guesthouse has the highest average price, but we only have 1 record with this type. That's why we find a flat small line in the boxplot. So, we moved to the next highest type which is "Boat".

Conclusion

By this, we have obtained answers to our 3 questions to help you navigate different options next time you consider using Boston Airbnb listings. We have retrieved the listing with the highest % of positive reviews among the top 10 frequent listings. We also extracted the most frequent zip codes in our data, and finally, we explored different types of listings and their price variability.

Resources

The dataset was obtained from Kaggle https://www.kaggle.com/datasets/airbnb/boston

You can view my code here https://github.com/nesmasilman/Data_Scientist_Nano_Degree-Project-1/tree/main