

Airbnb Bangkok Listing Data Analytics

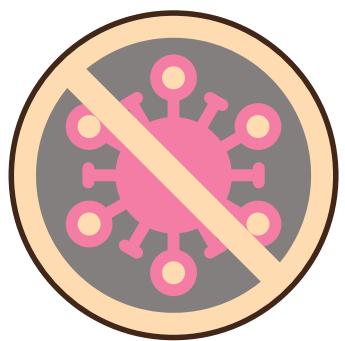
By: Harish Muhammad

Outlines

- Business understanding
- problem Statement
- Data Understanding
- Data Cleaning
- Data Analysis
- Insights
- Recommendations
- References



Business understanding



Japan leads the way and Thailand in Top 10: Airbnb's fastest growing countries

David Barrett
08.01.2024 7637

42 Shares A row of small social media sharing icons for Facebook, LinkedIn, Twitter, and others.

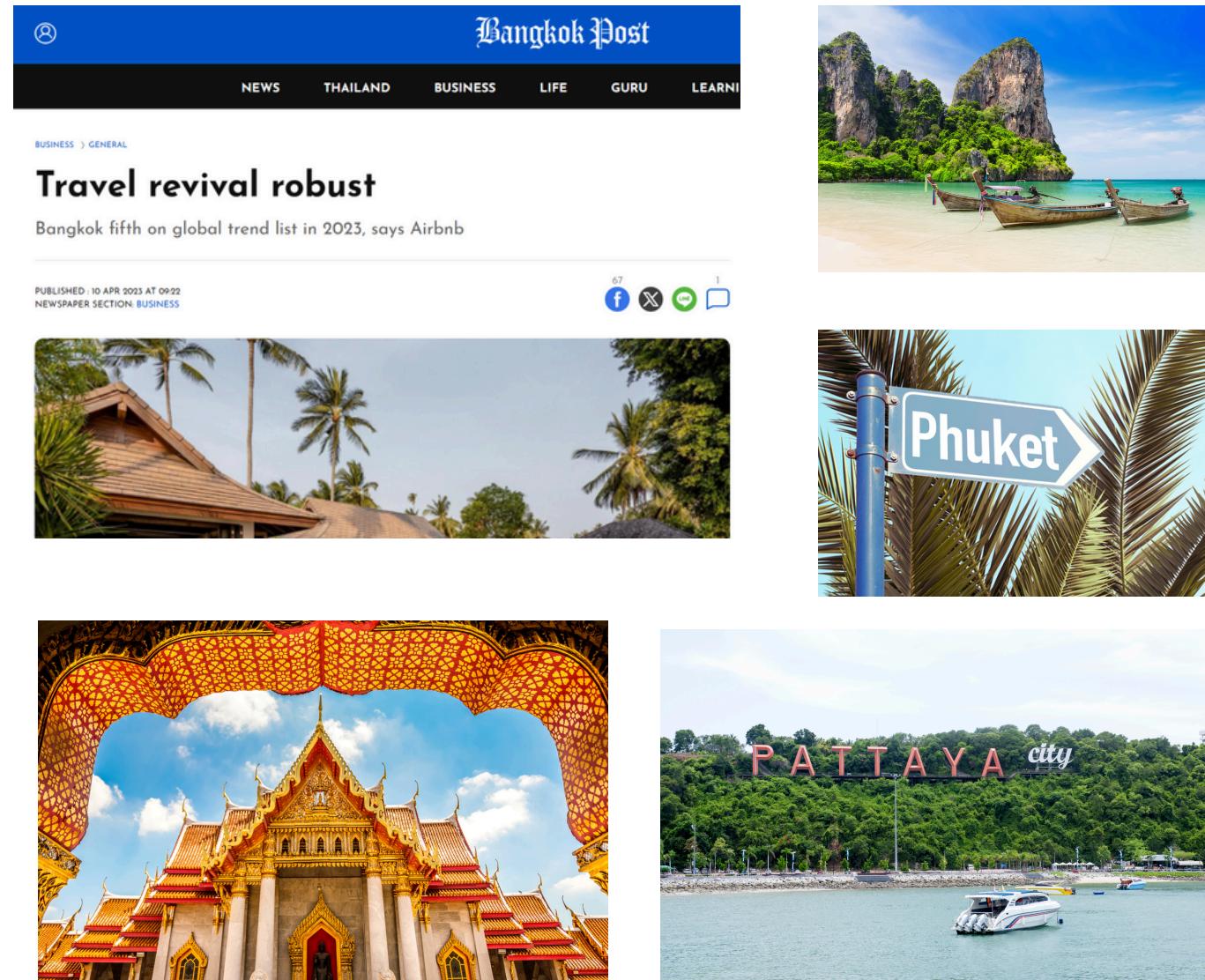
Top 10 Destinations 2024



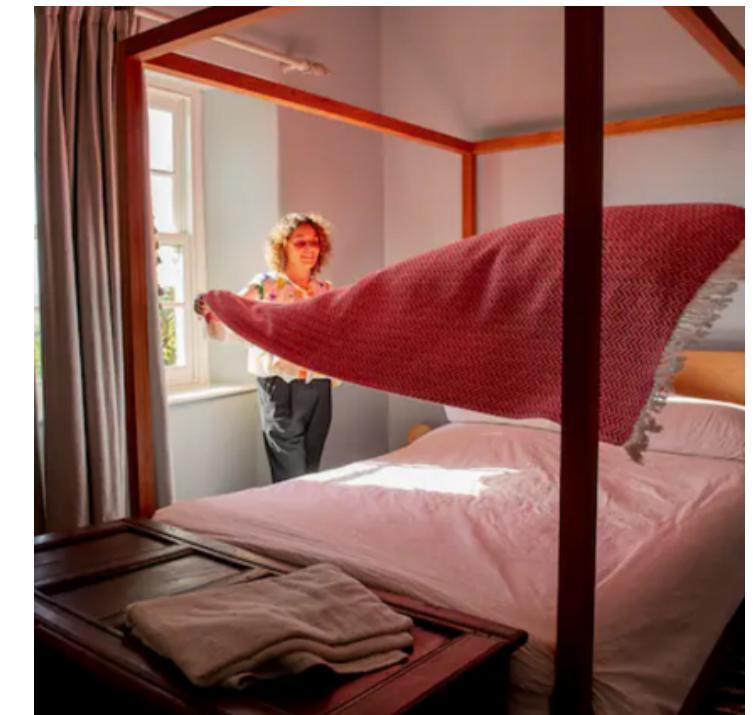
Post pandemic, global tourism is recovering

Thailand is one of top 10 destinations

Business understanding



In Thailand, Bangkok was the most booked on Airbnb in 2023



More opportunities are coming to Airbnb partners



Problem statement

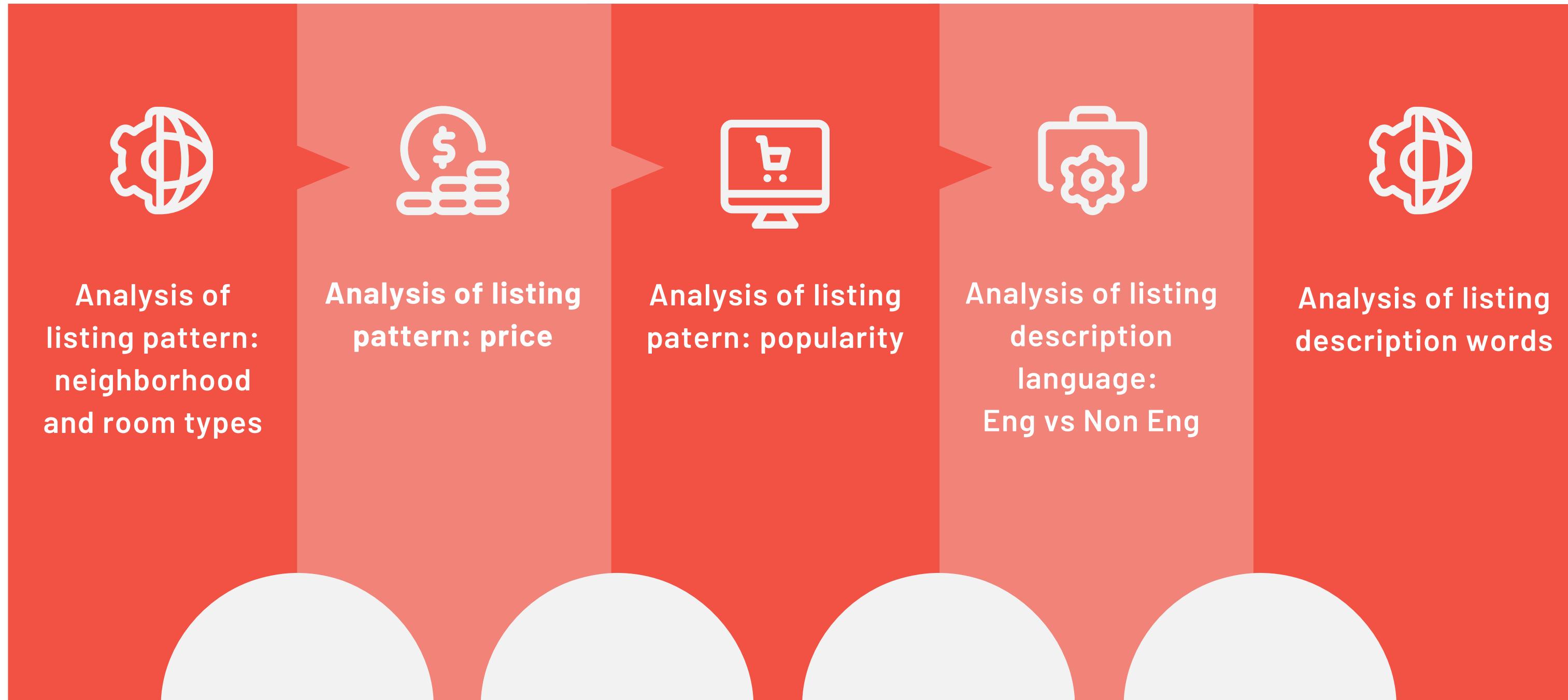


Question:

**What are the key recommendations
that can be provided to new listing
providers embarking on their
journey as Airbnb partners in
Bangkok?**



ANALYTICAL APPROACH



Data understanding

Dataset summary

```
# Displaying the summary of the dataset  
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 15854 entries, 0 to 15853  
Data columns (total 16 columns):  
 #   Column           Non-Null Count  Dtype     
---  --  
 0   id               15854 non-null   float64  
 1   name              15846 non-null   object  
 2   host_id            15854 non-null   int64  
 3   host_name          15853 non-null   object  
 4   neighbourhood       15854 non-null   object  
 5   latitude            15854 non-null   float64  
 6   longitude           15854 non-null   float64  
 7   room_type           15854 non-null   object  
 8   price               15854 non-null   int64  
 9   minimum_nights      15854 non-null   int64  
 10  number_of_reviews    15854 non-null   int64  
 11  last_review          10064 non-null   object  
 12  reviews_per_month    10064 non-null   float64  
 13  calculated_host_listings_count 15854 non-null   int64  
 14  availability_365     15854 non-null   int64  
 15  number_of_reviews_ltm 15854 non-null   int64  
dtypes: float64(4), int64(7), object(5)  
memory usage: 1.9+ MB
```

12 Numerical columns (data type: int64 & float64)

4 categorical columns (data type: object)

New additional columns:

'price_category', 'listing_in_english', & 'name_no_emoji'

Data Cleaning

Handling duplicated data

```
# Finding duplicated data of overall data  
df.duplicated().sum()
```

0

Handling Anomalies

```
# Checking and filtering Listing where price == 0  
df[df['price'] == 0]
```

	id	name	host_id	host_name	neighbourhood	latitude	longitude	room_type	price	
1	11103	44563108.0	Somerset Maison Asoke Bangkok	360620448	Somerset Maison Asoke	Vadhana	13.73815	100.5642	Hotel room	0

Handling Missing values

```
# Calculating the overall missing values  
df.isnull().sum().sum()
```

11587

```
# Replace the missing value in the column 'name' with 'unknown'  
df['name'].fillna('unknown', inplace=True)
```

```
# Replace the missing value in the column 'host_name' with 'unknown'  
df['host_name'].fillna('unknown', inplace=True)
```

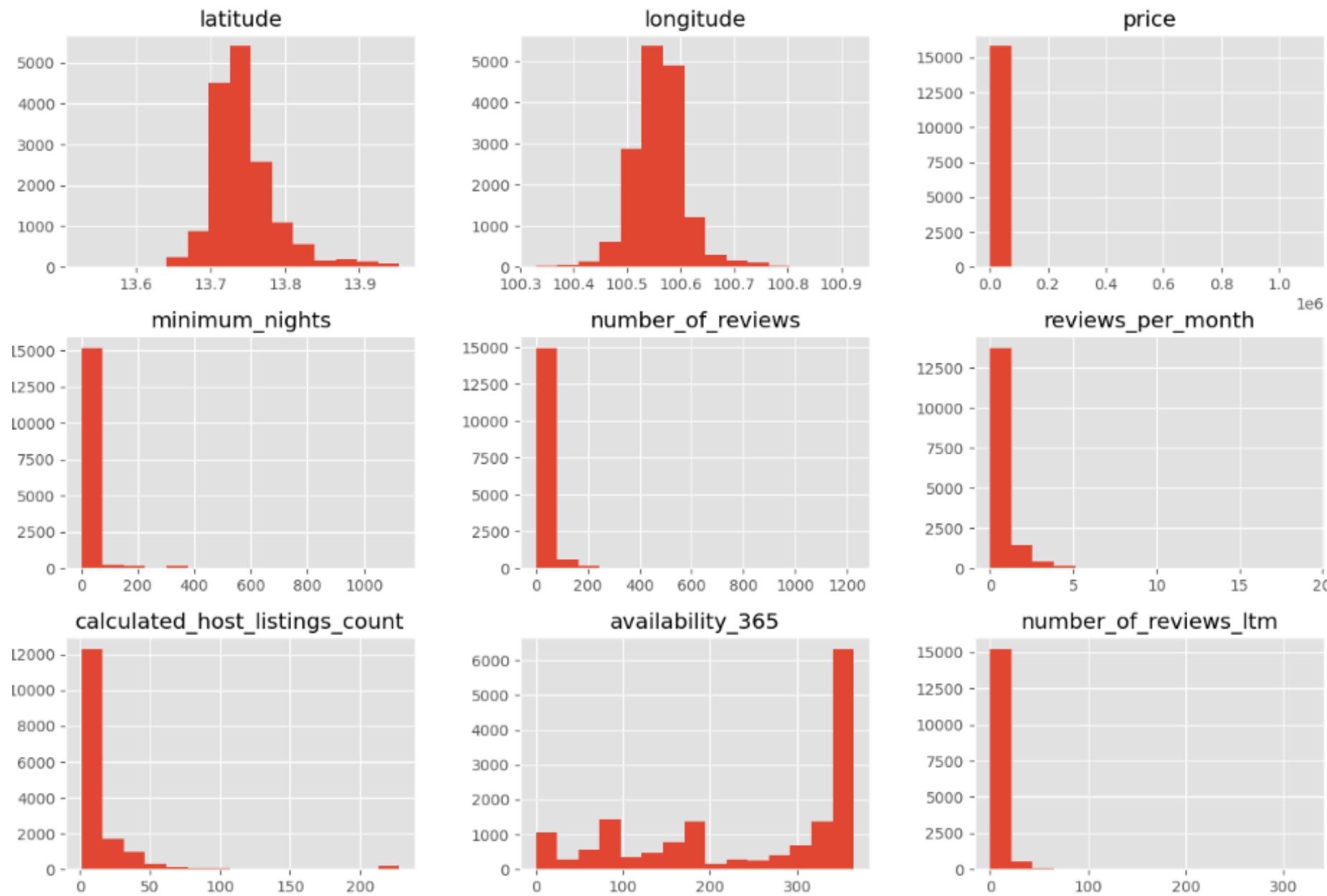
```
# Imputation method for 'reviews_per_month'  
df['reviews_per_month'] = df['reviews_per_month'].fillna(value=0)
```

```
# Dropping the columns of 'last_review'  
df.drop(columns=['last_review', 'id', 'host_id', 'host_name'], inplace=True)
```

Data understanding

Data distribution and outliers

Data distribution (Histogram)



Data distribution - (Normality test)

	Feature	D'Agostino-Pearson test	P-value	Distribution
0	latitude	4356.959558	0.0	Not Normally Distributed
1	longitude	2077.433465	0.0	Not Normally Distributed
2	price	43046.717756	0.0	Not Normally Distributed
3	minimum_nights	22190.129178	0.0	Not Normally Distributed
4	number_of_reviews	19163.172470	0.0	Not Normally Distributed
5	reviews_per_month	15621.035696	0.0	Not Normally Distributed
6	calculated_host_listings_count	16474.501173	0.0	Not Normally Distributed
7	availability_365	12885.350519	0.0	Not Normally Distributed
8	number_of_reviews_ltm	24895.698981	0.0	Not Normally Distributed

Outlier presence

	Column	Total Outliers	Percentage (%)	Lower Bound	Upper Bound
0	latitude	1094	6.90	13.660975	13.818615
1	longitude	602	3.80	100.446500	100.668340
2	price	1403	8.85	-1393.500000	4722.500000
3	minimum_nights	3168	19.98	-8.000000	16.000000
4	number_of_reviews	2240	14.13	-19.500000	32.500000
5	reviews_per_month	1471	9.28	-1.005000	1.675000
6	calculated_host_listings_count	1832	11.56	-17.000000	31.000000
7	availability_365	0	0.00	-195.000000	693.000000
8	number_of_reviews_ltm	2219	14.00	-4.500000	7.500000



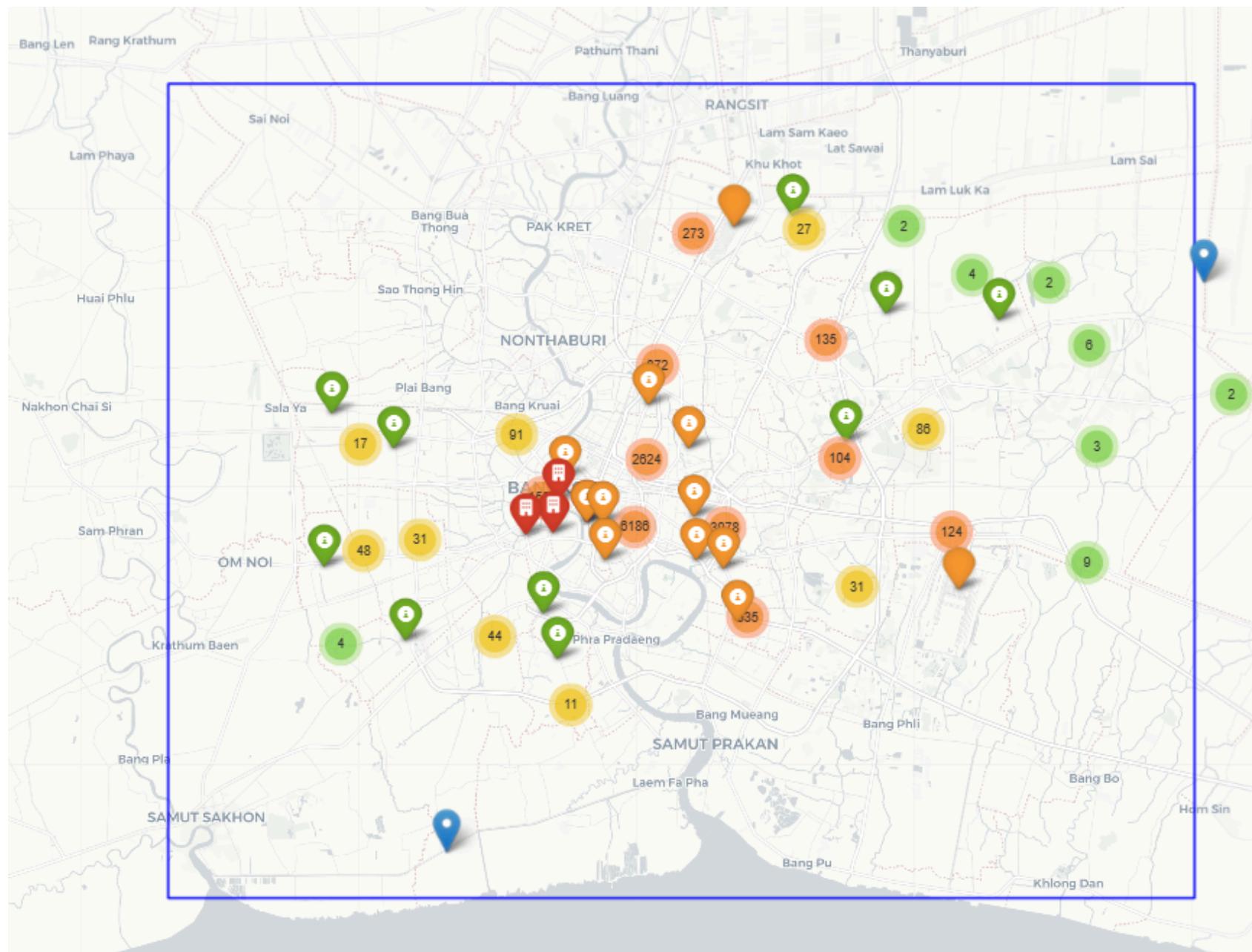
Analysis of listing pattern: neighborhood and room types

- **What is the current listing distribution of Airbnb in Bangkok? which area of Bangkok has the highest number of listings and which does not?**
- **What are the compositions of room types of Airbnb in Bangkok? which one is the most common?**
- **Is there any association between neighborhood location and listing room type?**



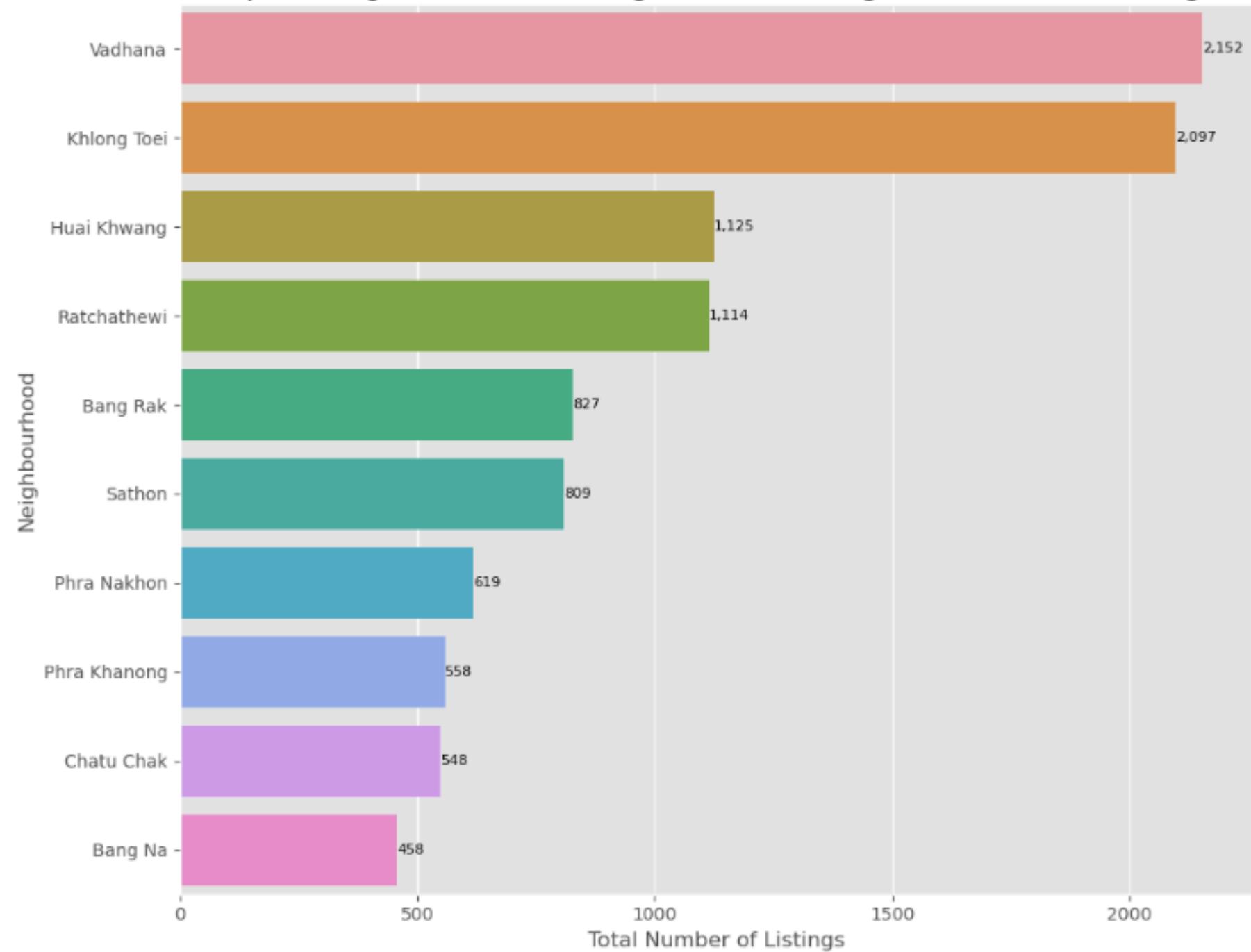
Analysis of listing pattern: neighborhood

Airbnb listing distribution in Bangkok



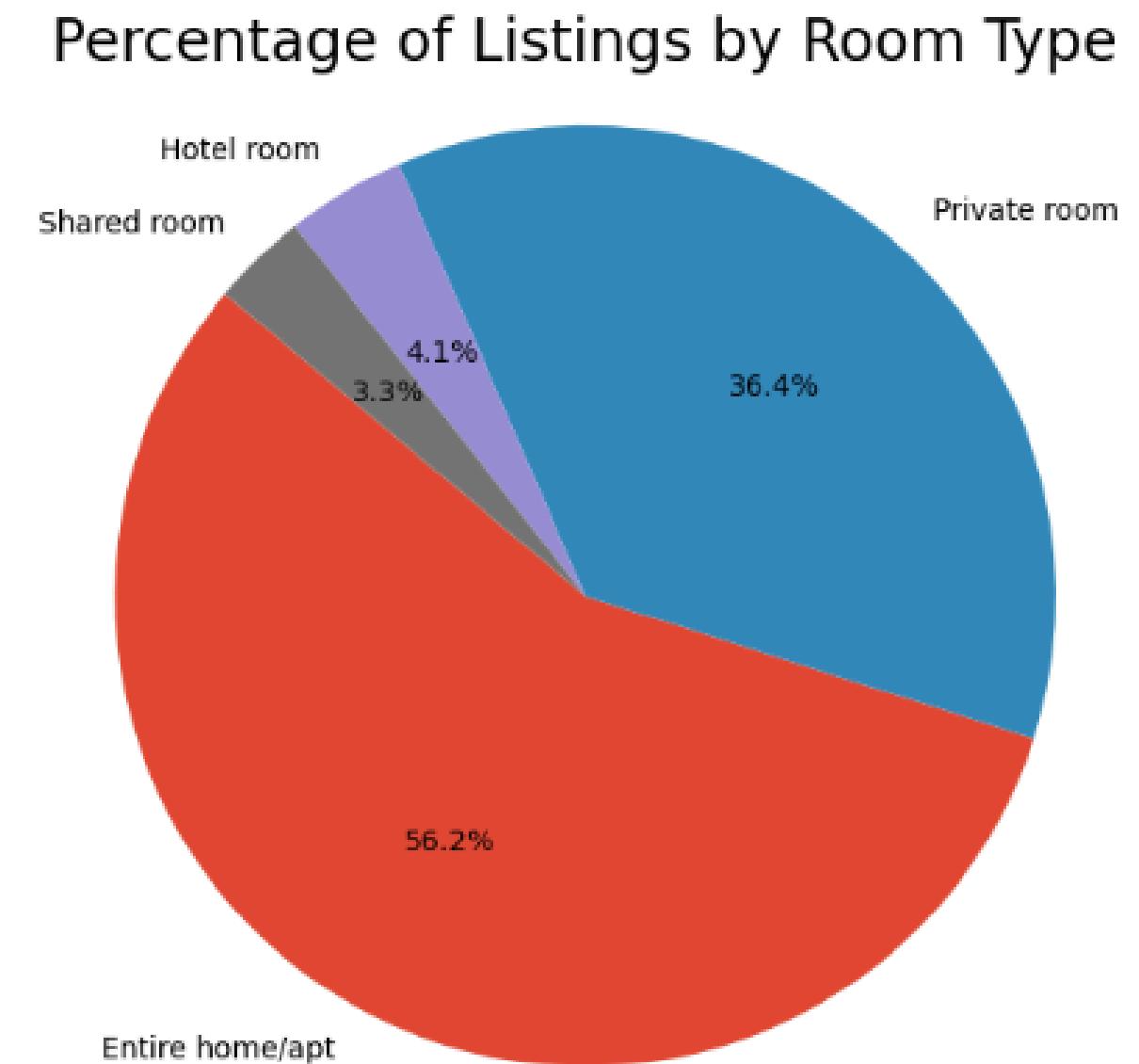
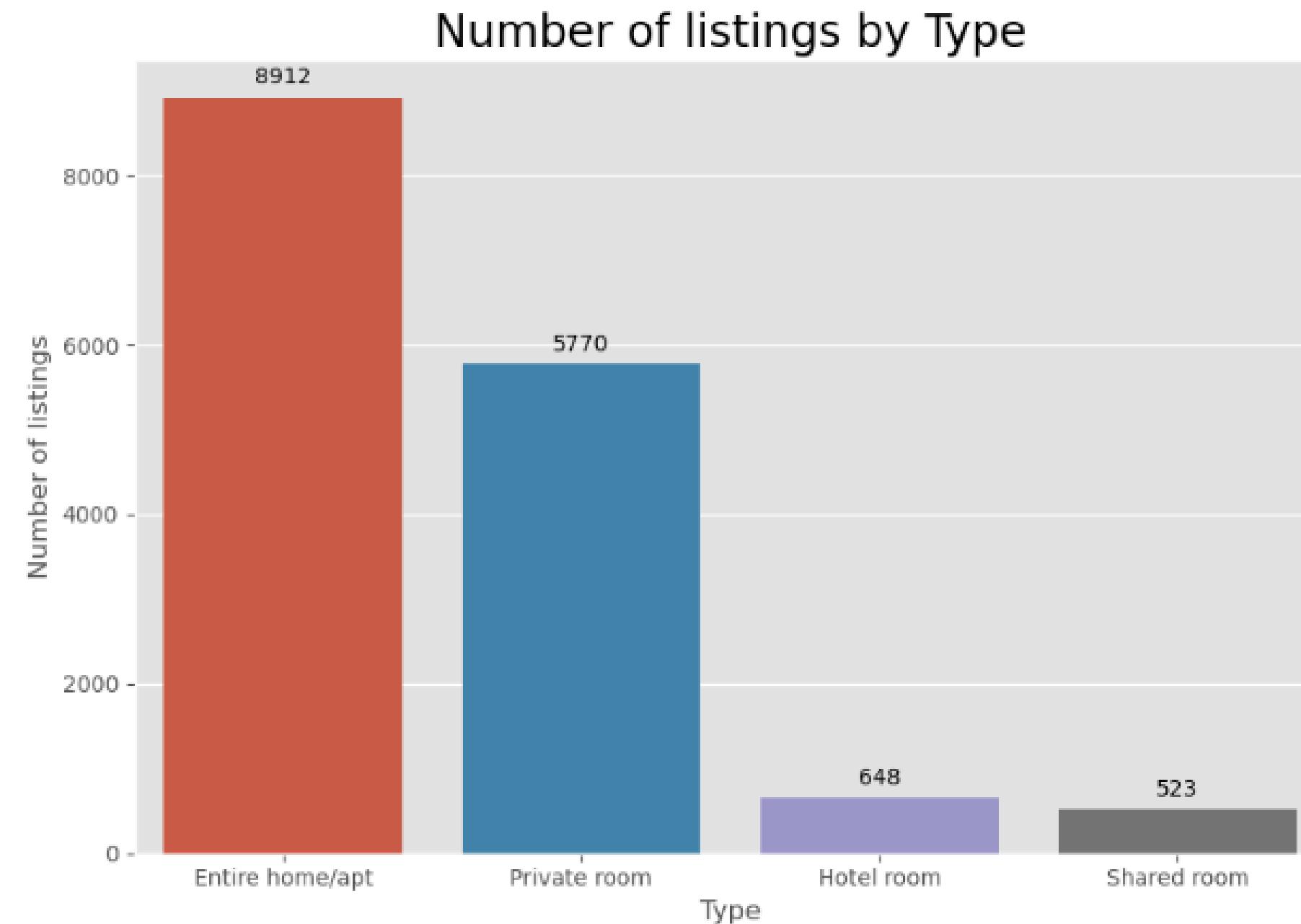
- Orange icons:** Neighbourhoods with a high number of listings
- Green icons:** Neighborhoods with a low number of listings
- Red icons:** Major tourism area

Top 10 Neighborhoods in Bangkok with the Highest Number of Listings



Analysis of listing pattern: Room type

Airbnb listing: room type compositions

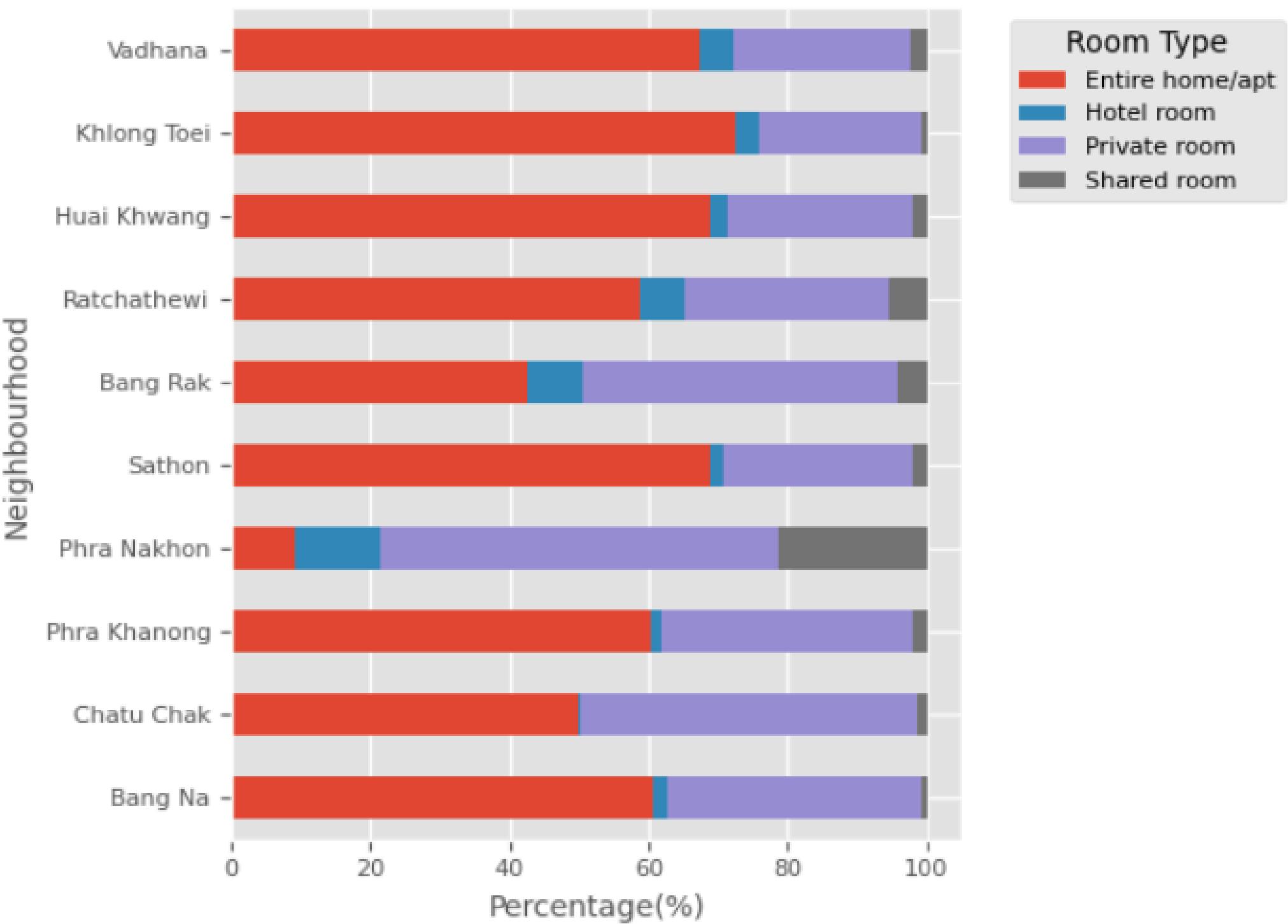


Analysis of listing pattern: Room type

Association among neighbourhood and room type

There is an association between neighborhoods and room types (pvalue ≤ 0.05)

Chi-square: Neighbourhood vs Room Type





Analysis of listing price pattern

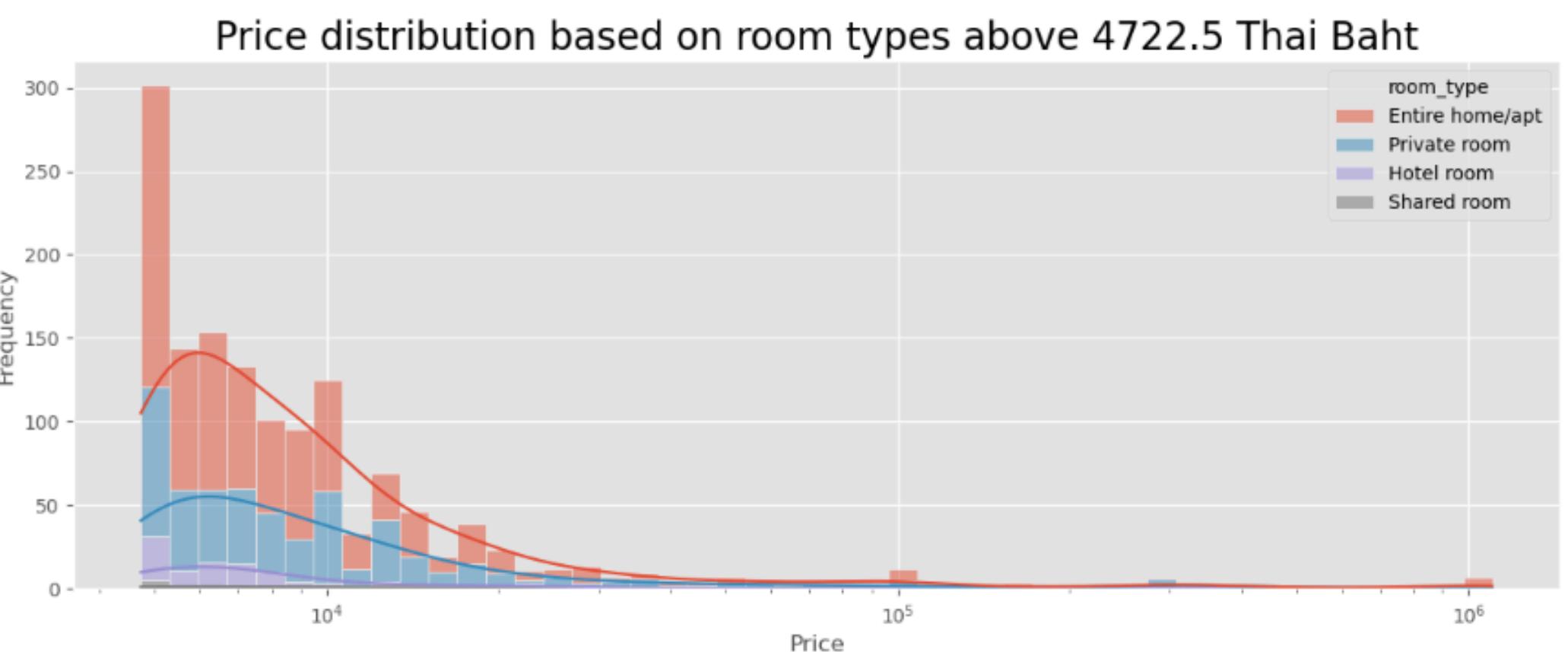
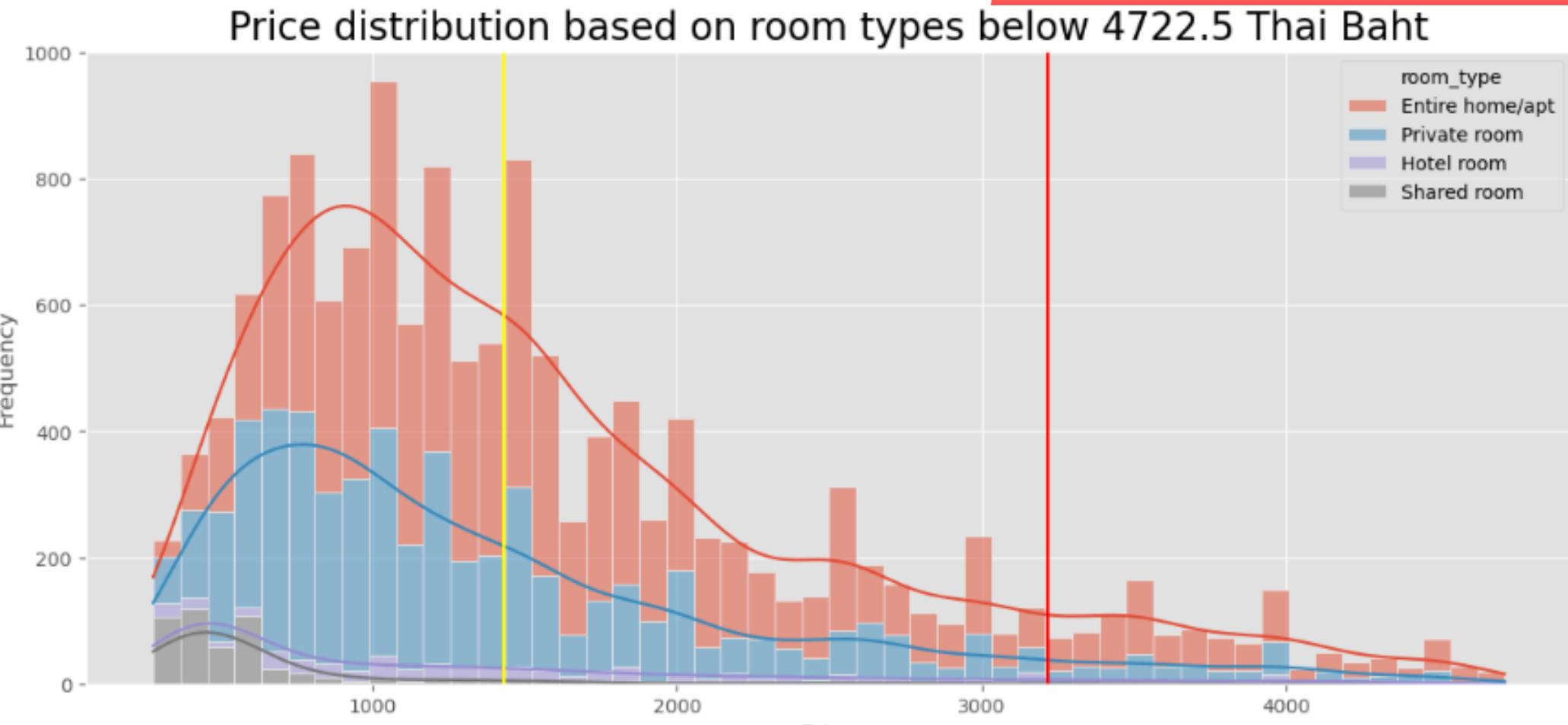
- **What are the listing price ranges of Airbnb listings in Bangkok?**
- **Which neighborhoods have the highest price of listing? Which neighborhoods have the lowest price of listing?**
- **Which room type has the highest price? Which room type that has the lowest price?**



Analysis of listing price pattern

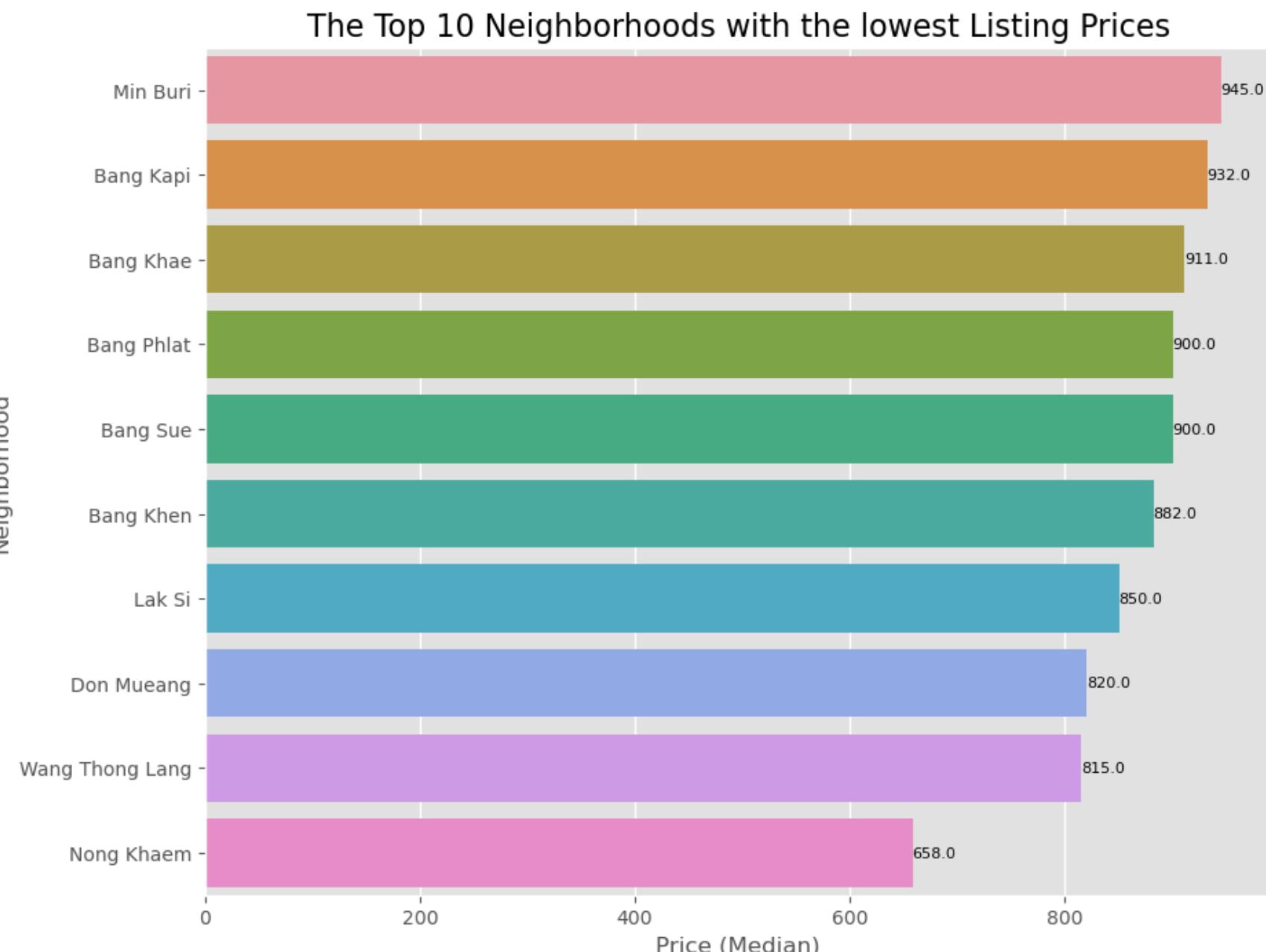
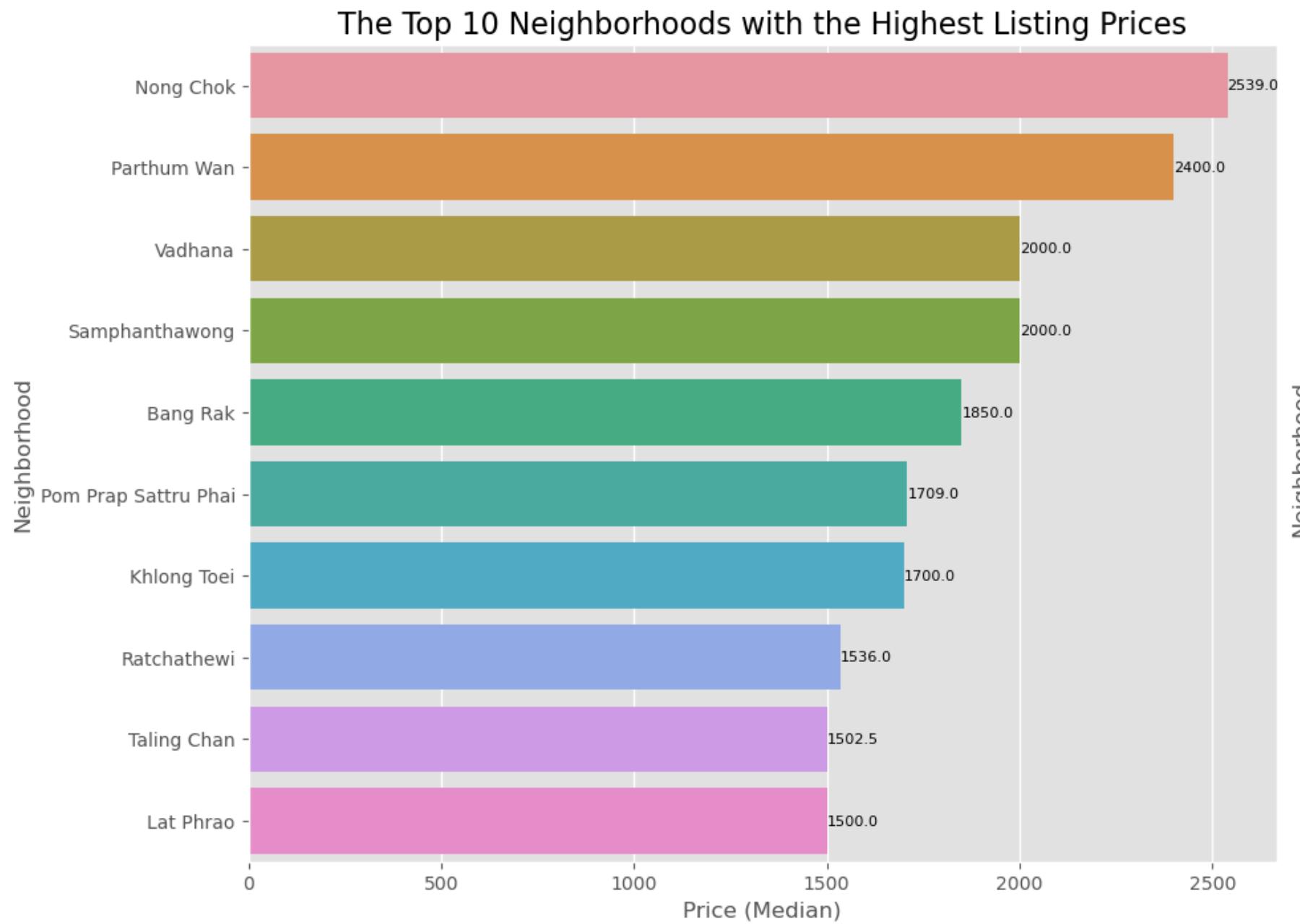
Price ranges of Airbnb listing in Bangkok

Min price	Max Price	IQR Price	Lower Bound	Upper Bound	Total Outliers
278	1100000	1529.0	-1393.5	4722.5	1403



Analysis of listing price pattern

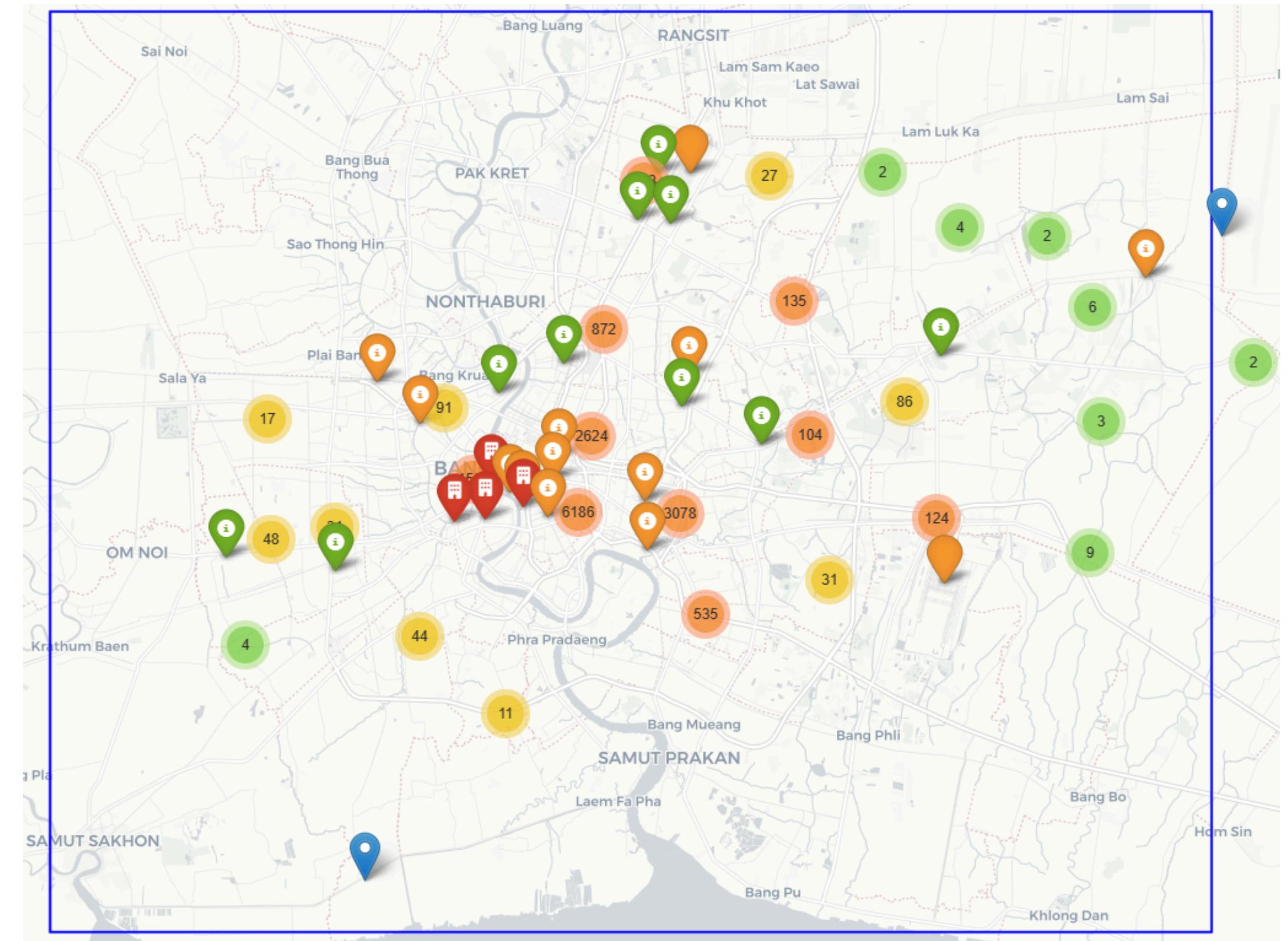
The neighborhoods that have the highest and the lowest price of listing



Analysis of listing price pattern

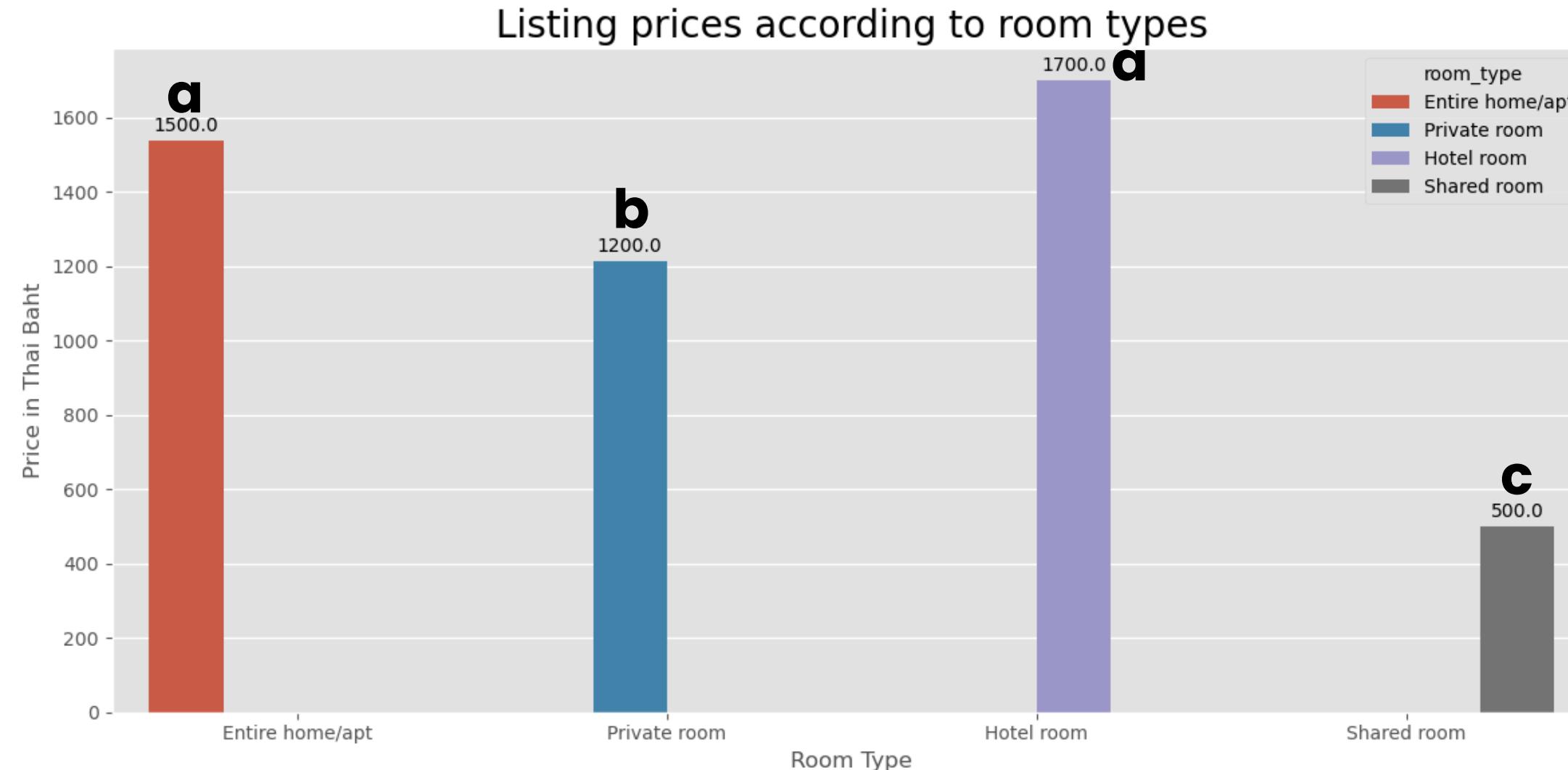
Airbnb listing price distribution in Bangkok

- Orange icons:** Neighbourhoods with the highest listing price
- Green icons:** Neighborhoods with the lowest listing price
- Red icons:** Major tourism area



Analysis of listing price pattern

Listing prices according to room types



The result of Kruskal-Wallis test: p-value = 0.00000. p-value <= 0.05

Dunn's test results:

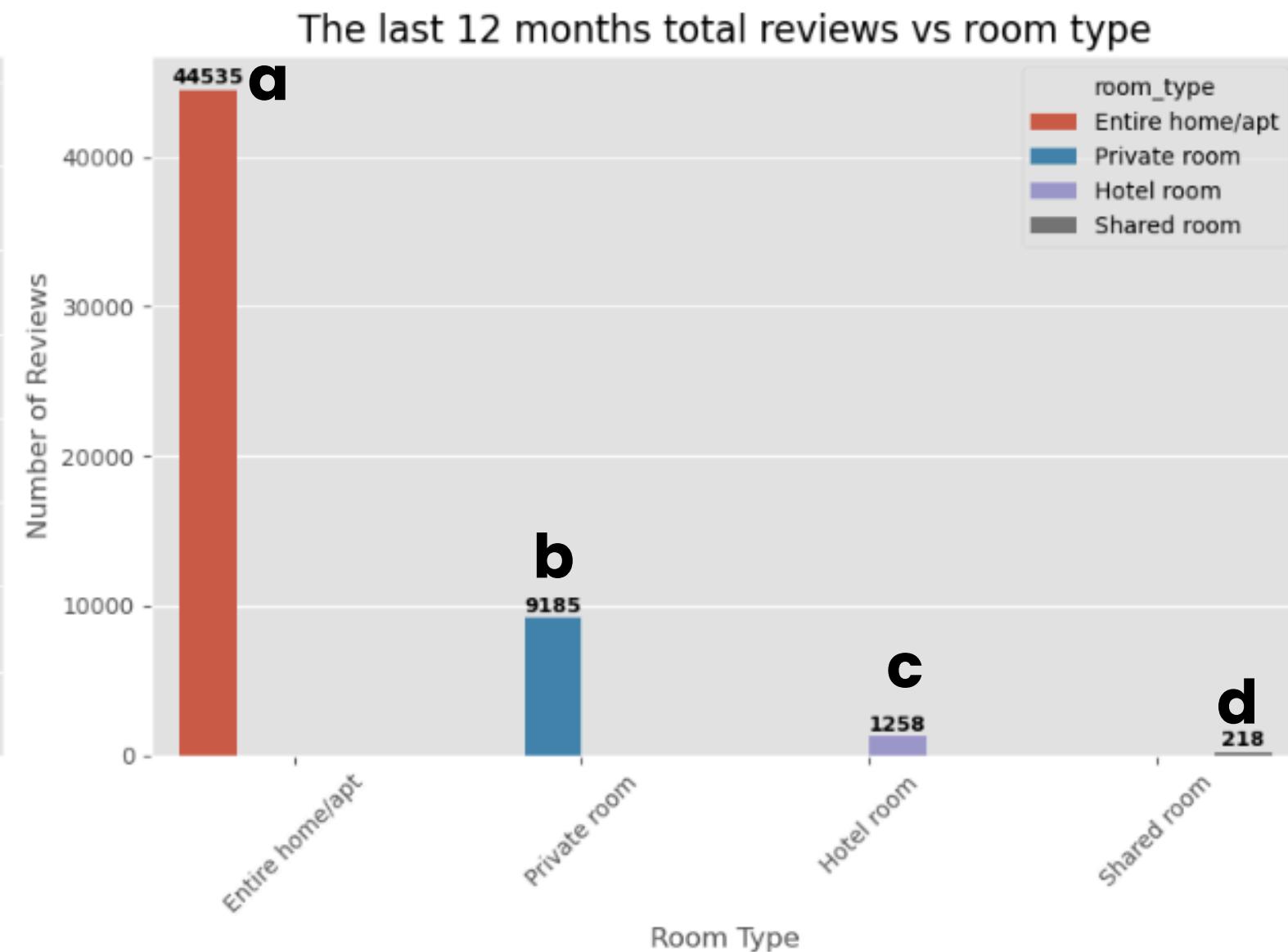
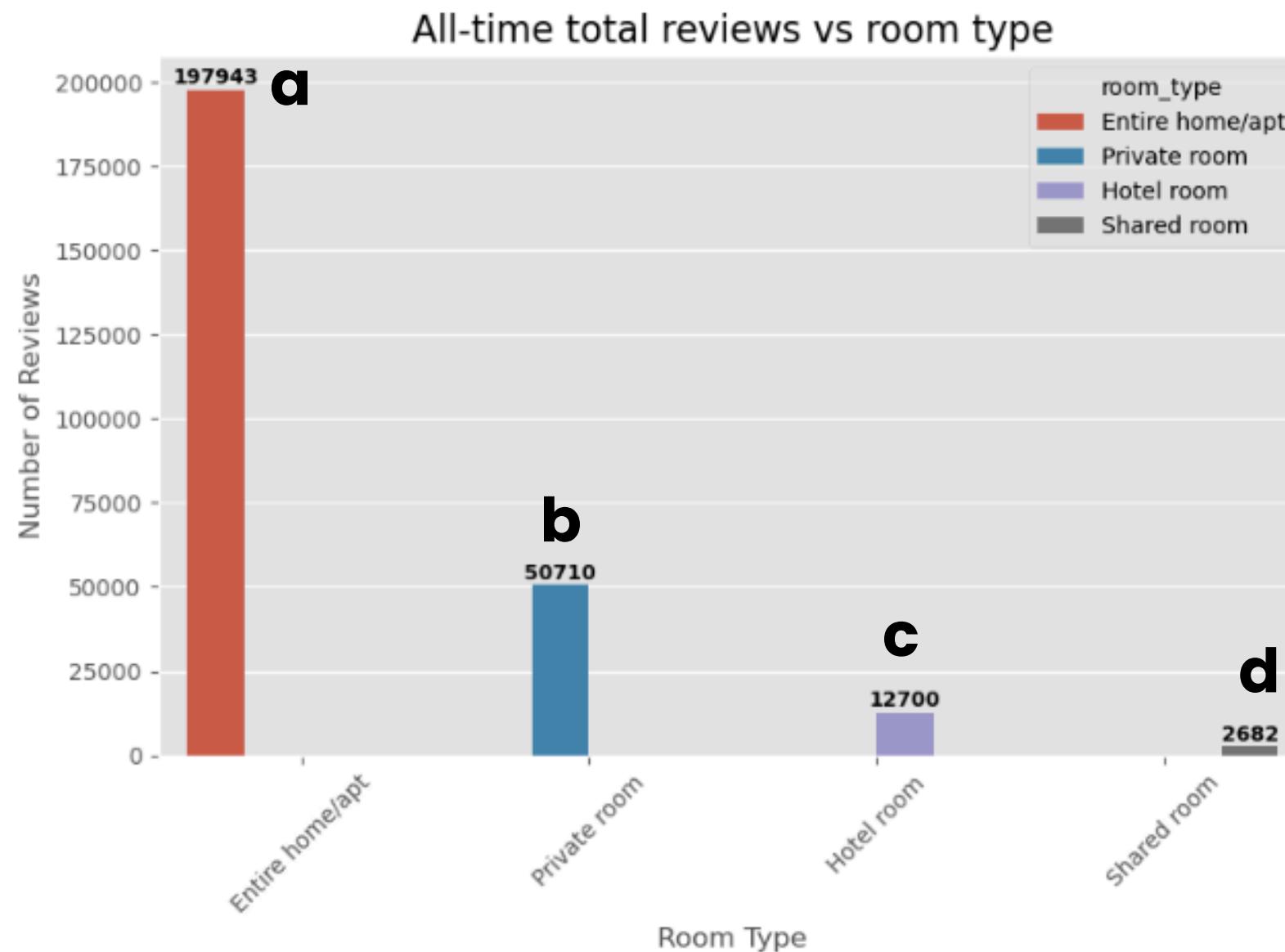
	Entire home/apt	Hotel room	Private room	Shared room
Entire home/apt	1.000000e+00	4.408613e-01	6.754733e-72	2.328538e-216
Hotel room	4.408613e-01	1.000000e+00	7.034041e-16	3.076910e-133
Private room	6.754733e-72	7.034041e-16	1.000000e+00	2.093437e-130
Shared room	2.328538e-216	3.076910e-133	2.093437e-130	1.000000e+00



Analysis of Listing Popularity

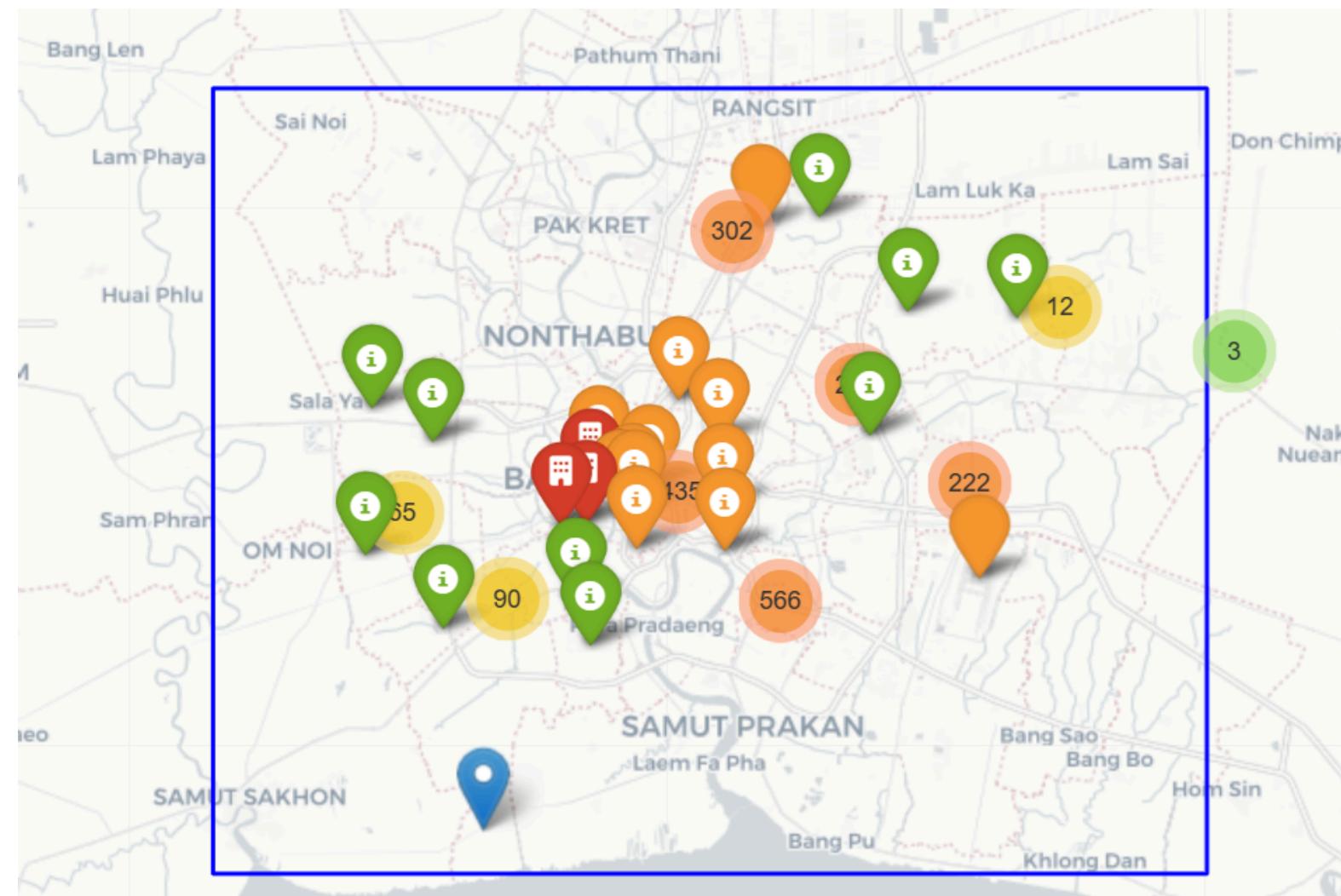
- 📢 **What is the most popular room type in Airbnb Bangkok?**
- 💡 **What are the most popular neighborhoods for Airbnb listing in Bangkok which have the highest number of reviews?**
- 🌐 **What is the price range or price category that is most popular or has the highest total reviews?**

Listing Popularity Pattern Analysis



Listing Popularity Pattern Analysis

Total Reviews vs listing distributions in neighbourhood



Orange icons:

Neighbourhood with the high number of reviews

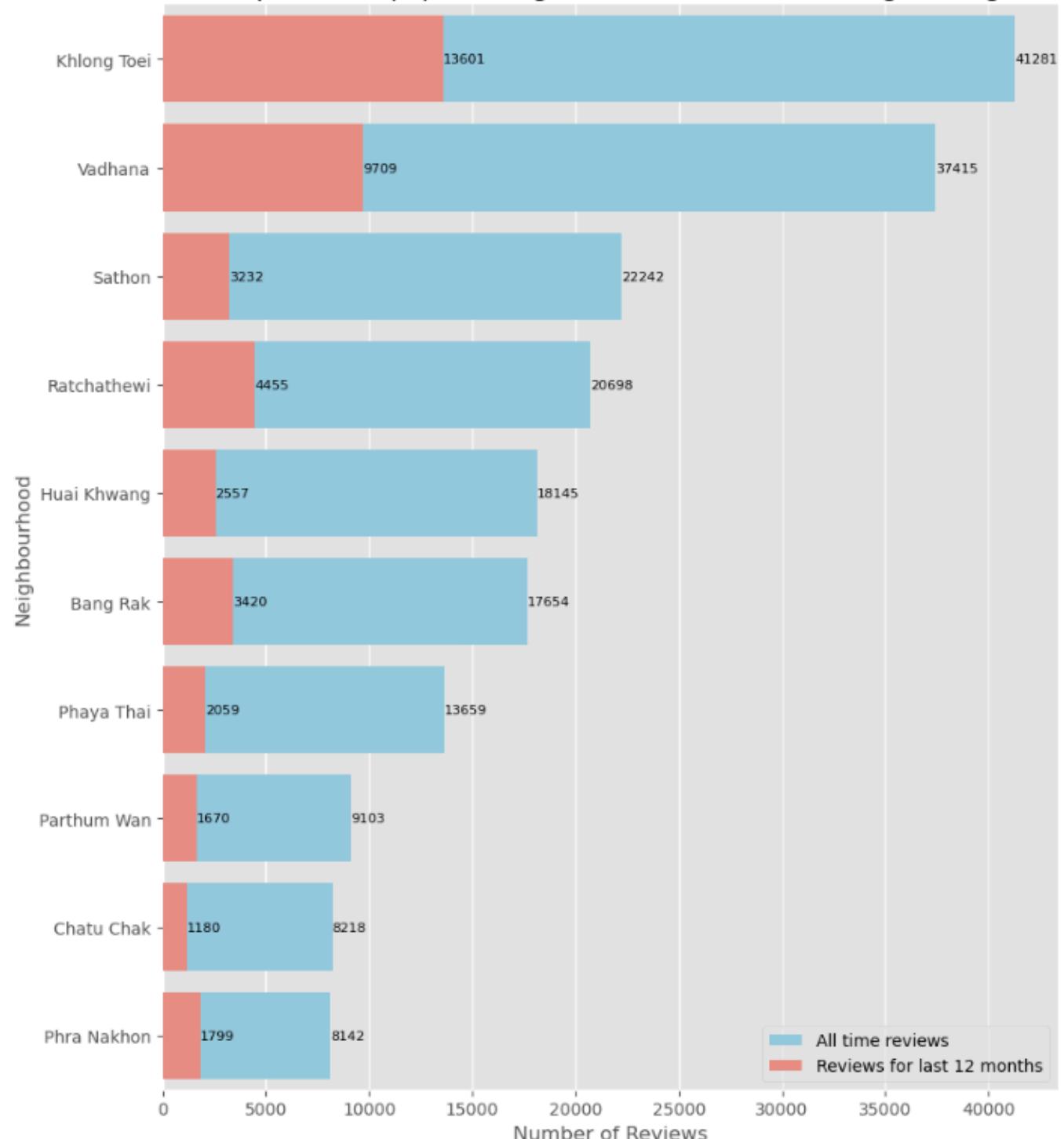
Green icons:

Neighborhood with the low number of reviews

Red icons:

Major tourism area

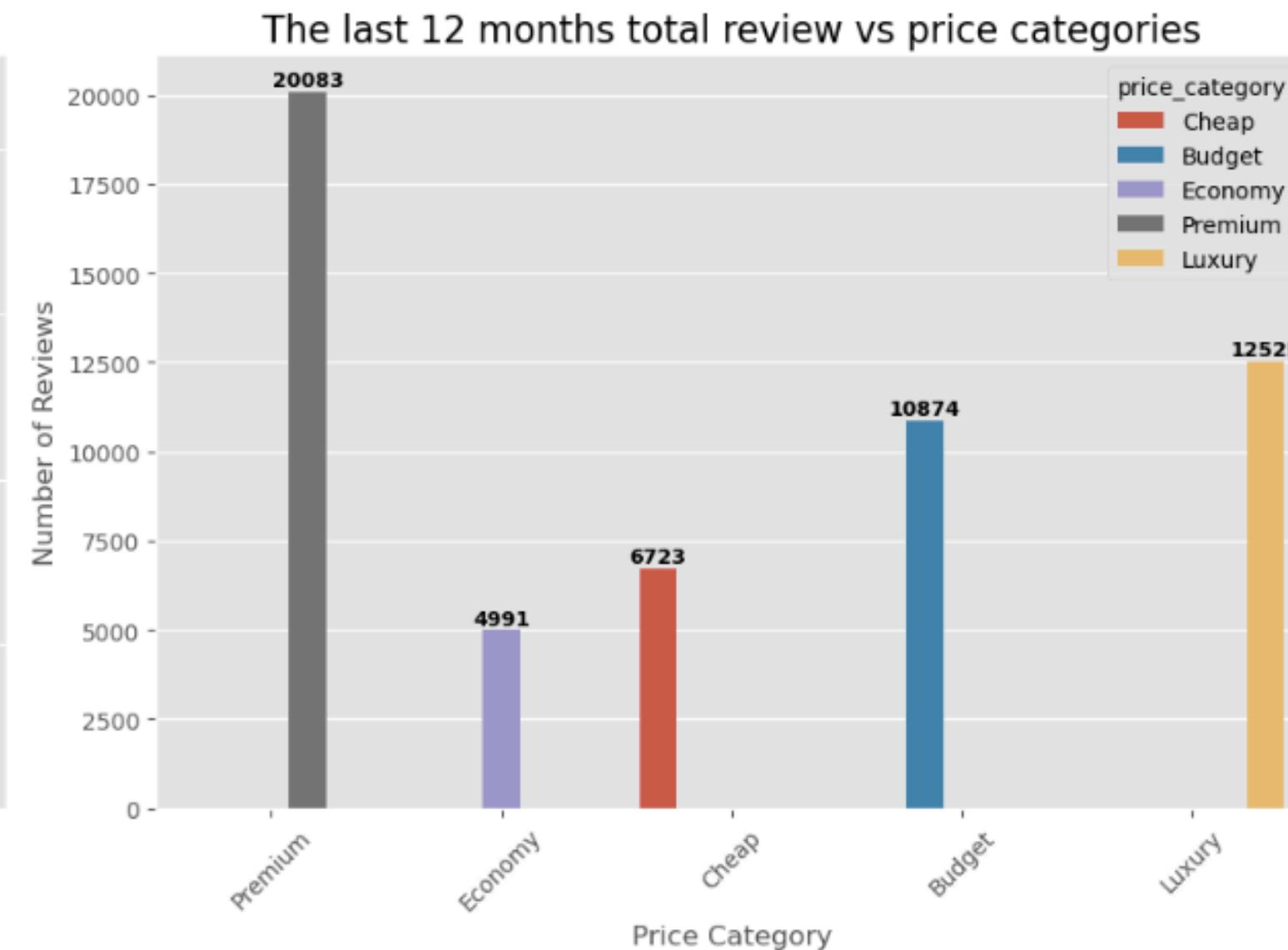
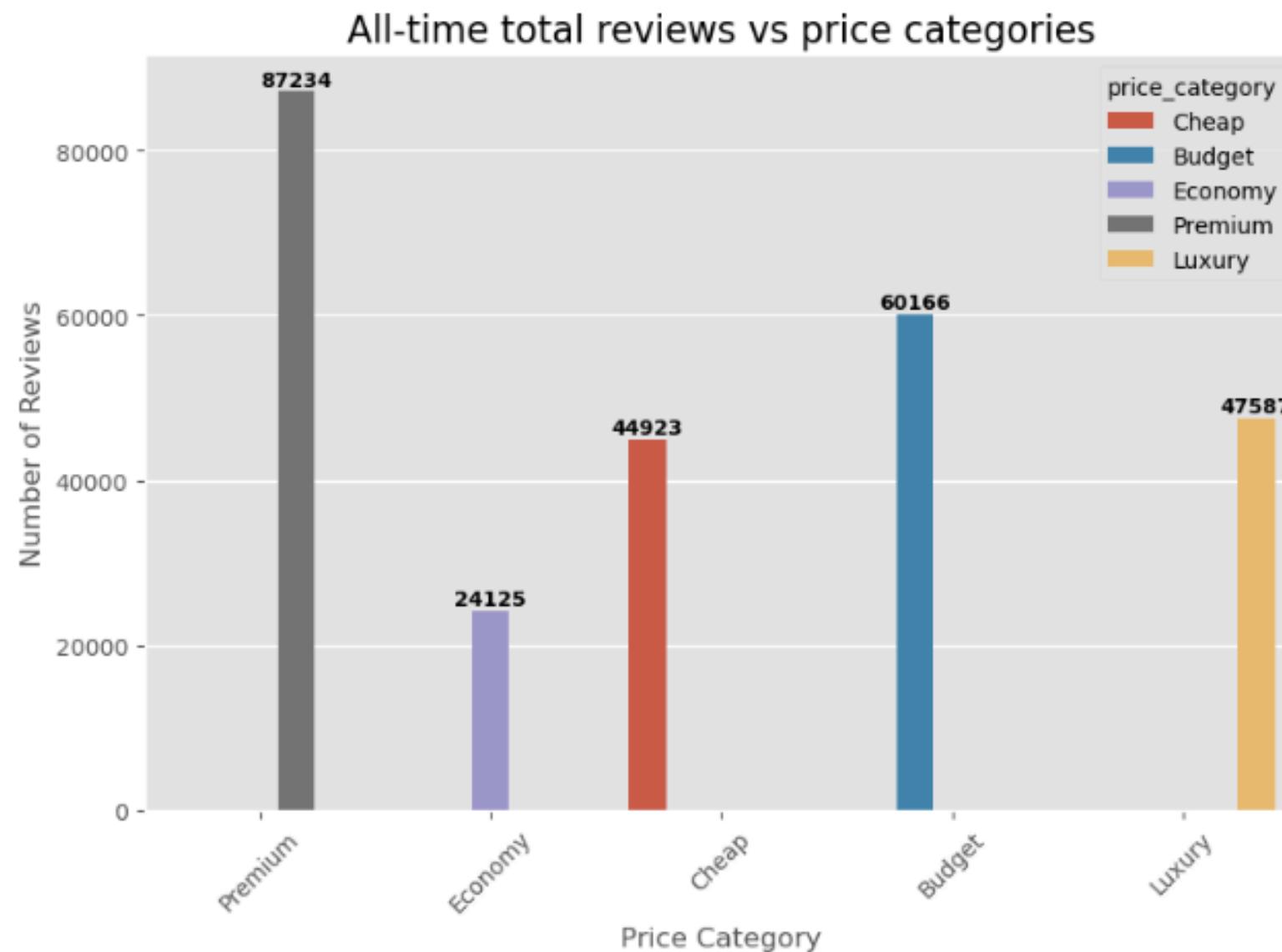
The Top 10 most popular neighbourhood for airbnb listing in Bangkok



Listing Popularity Pattern Analysis

Total Reviews vs Price category

No	Price category	Price range (in Thai Baht)
1	Cheap	= < 800
2	Budget	801 - 1200
3	Economy	1201 - 1429
4	Premium	1430 - 2800
5	Luxury	> 2800



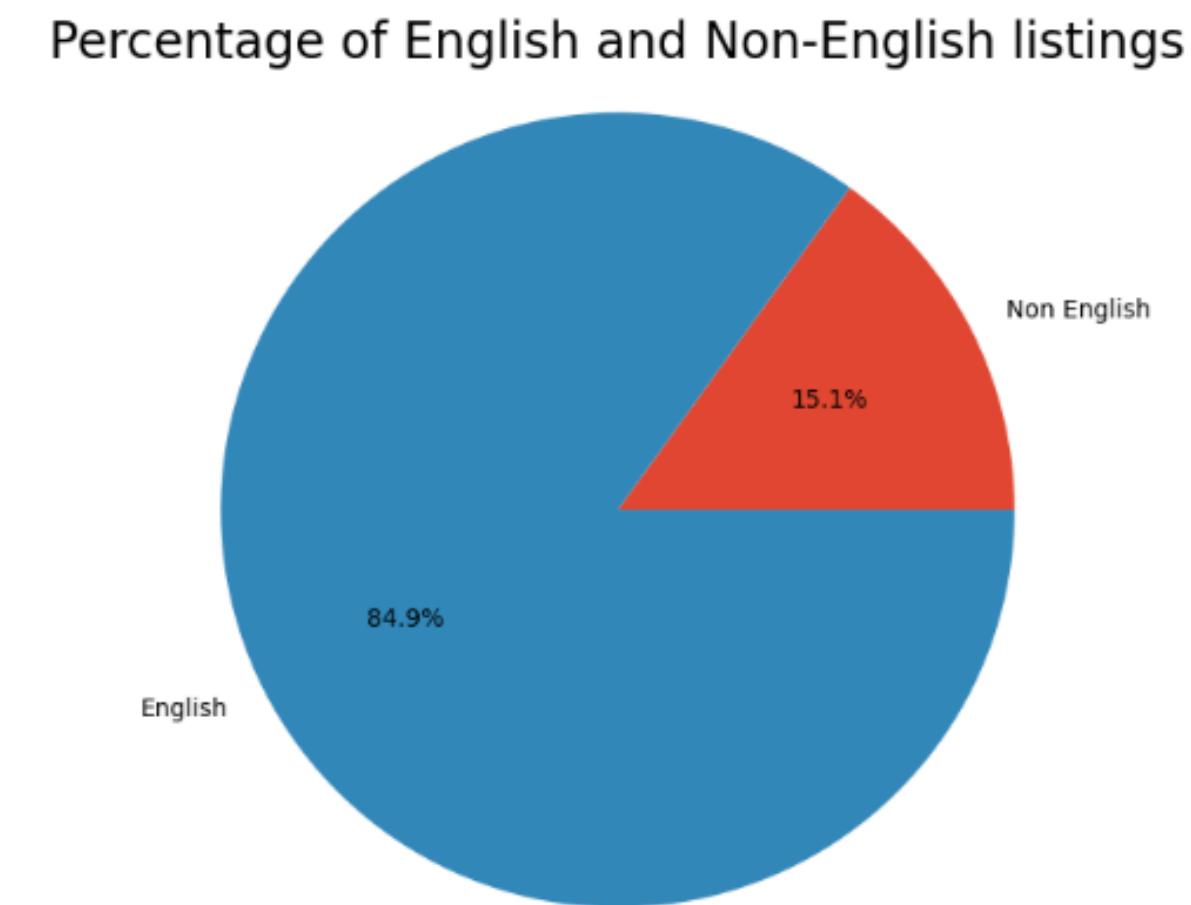
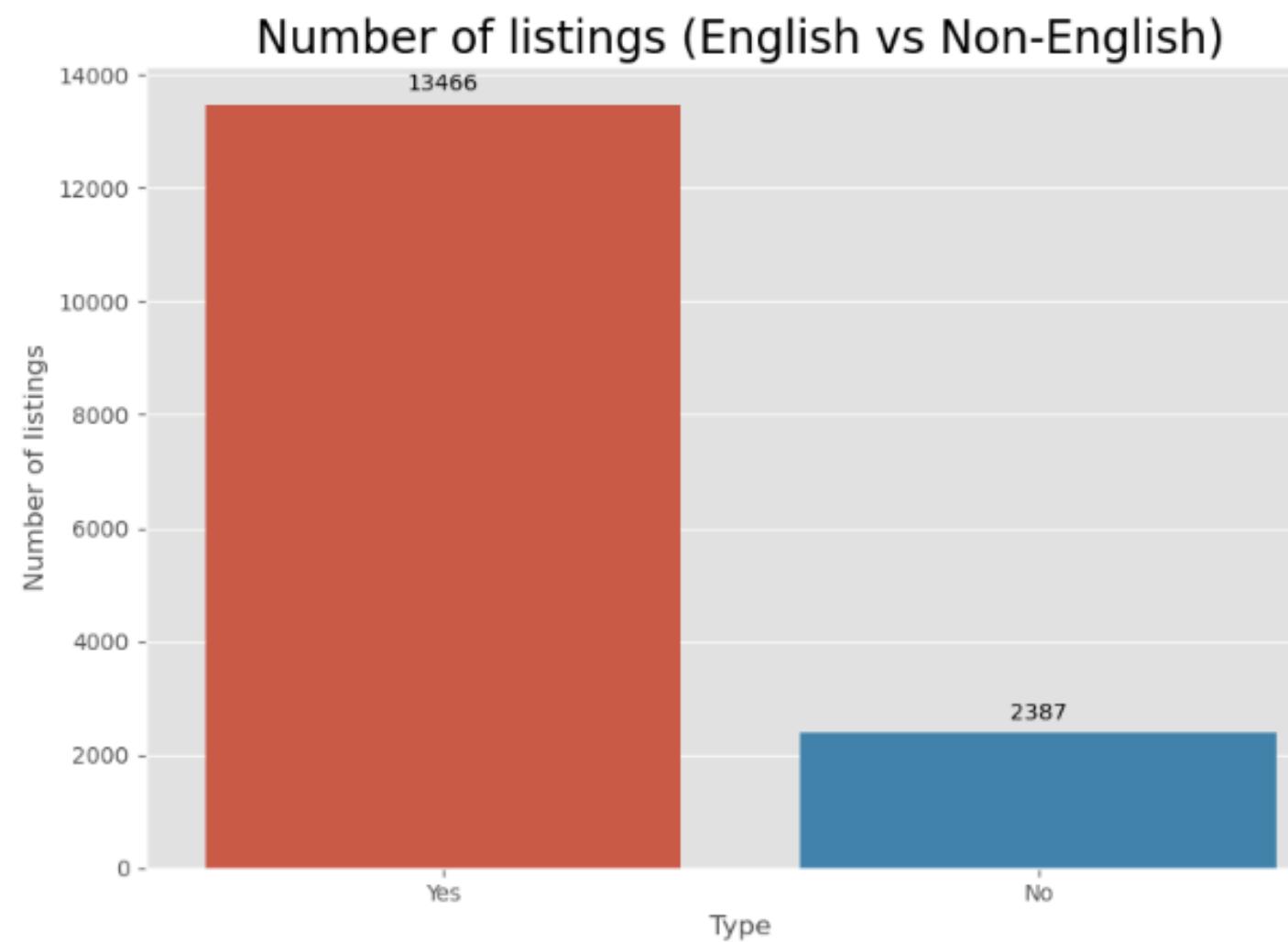


Analysis of Listing promotional language: English vs non-English

- 📢 **What is the current proportion of listings that use English and listings that use Non-English as promotional language in the airbnb Bangkok?**
- 💡 **Which listings are more popular, listings that use English as promotional language or not, based on all time and recent (last 12 months) total reviews?**
- 🌐 **Which listings are higher in price, listings that use English as promotional language or not?**

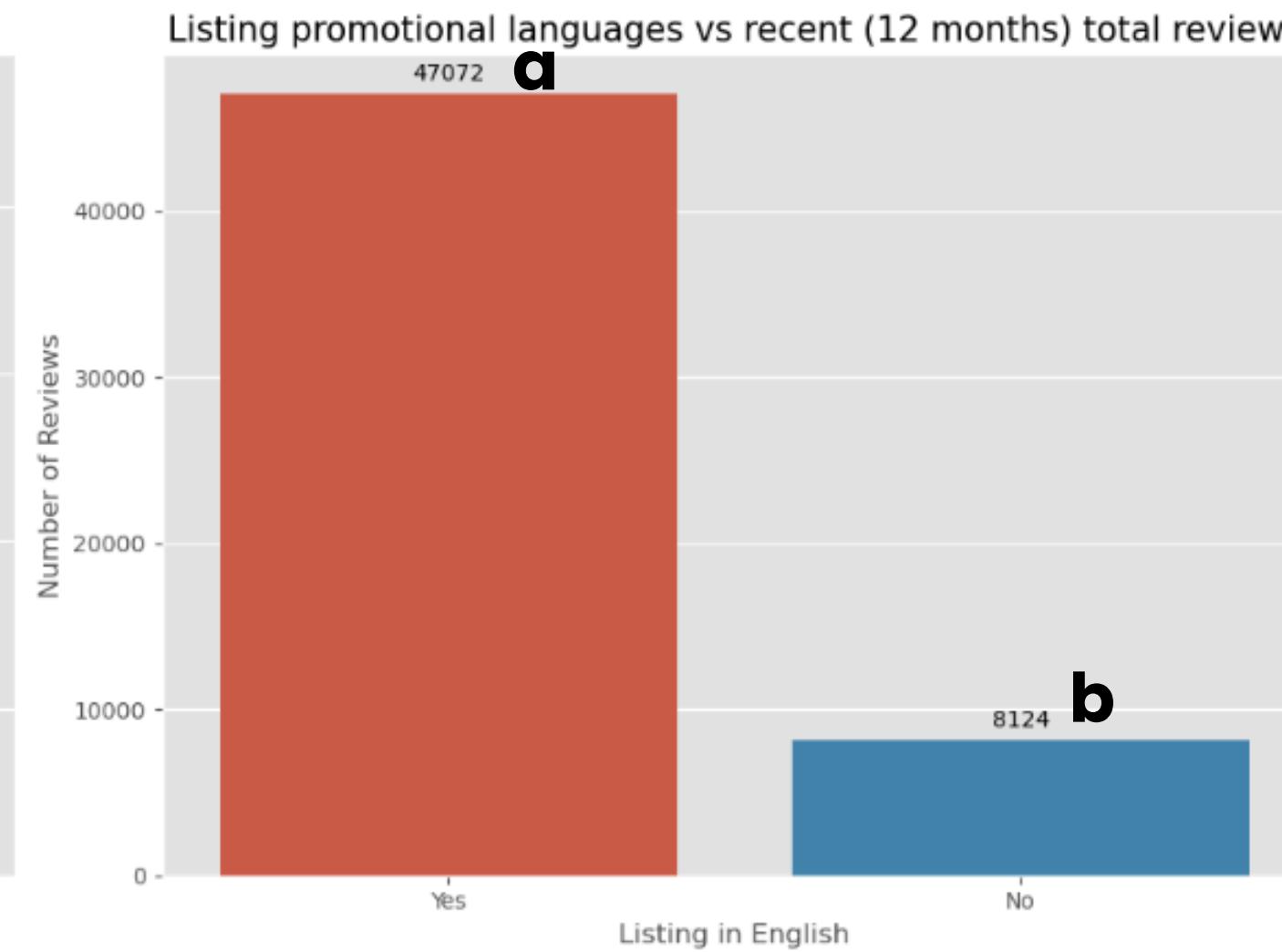
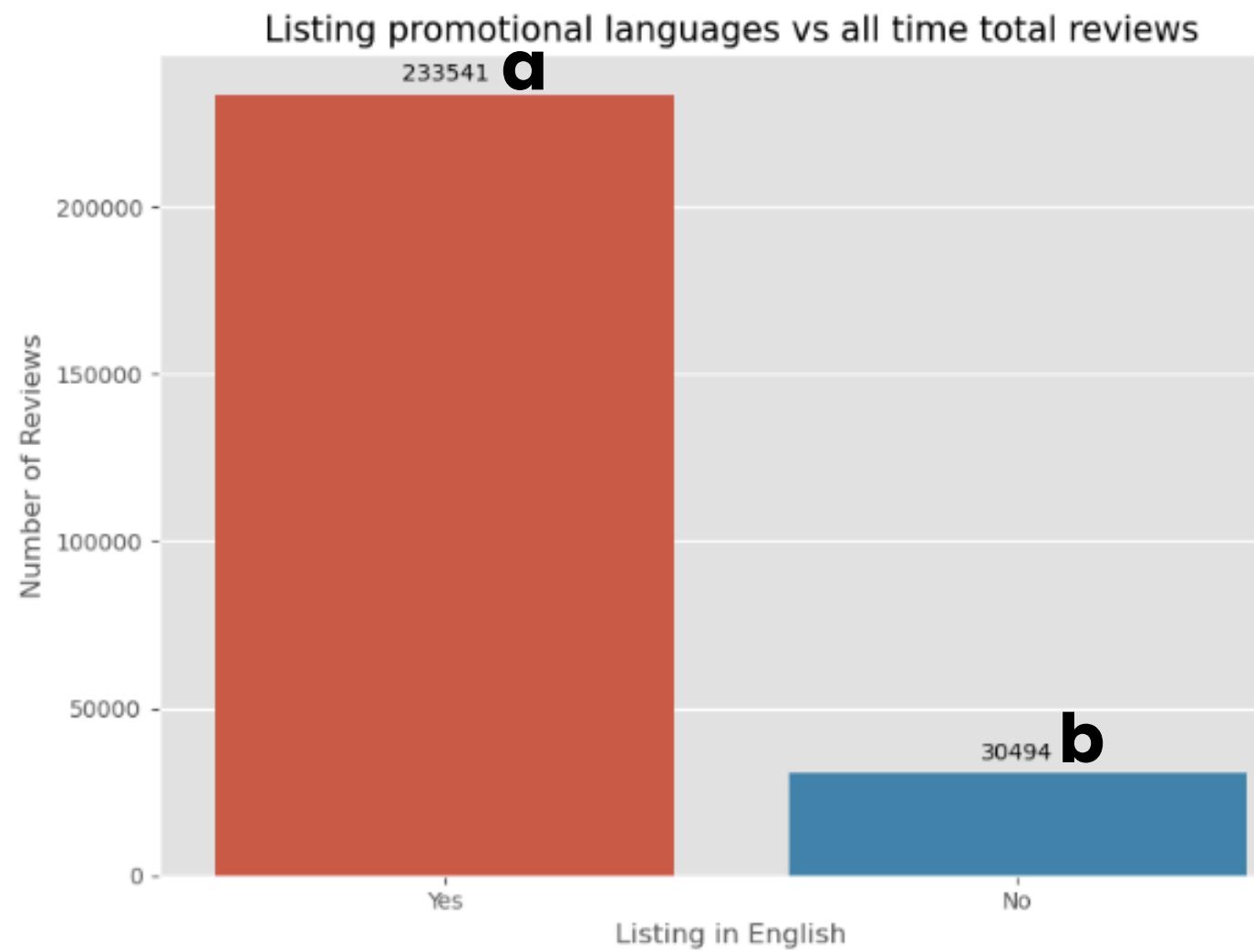
Analysis of Listing promotional language: English vs non-English

The proportion of listings that use English and listings that use Non-English



Listing Popularity Pattern Analysis

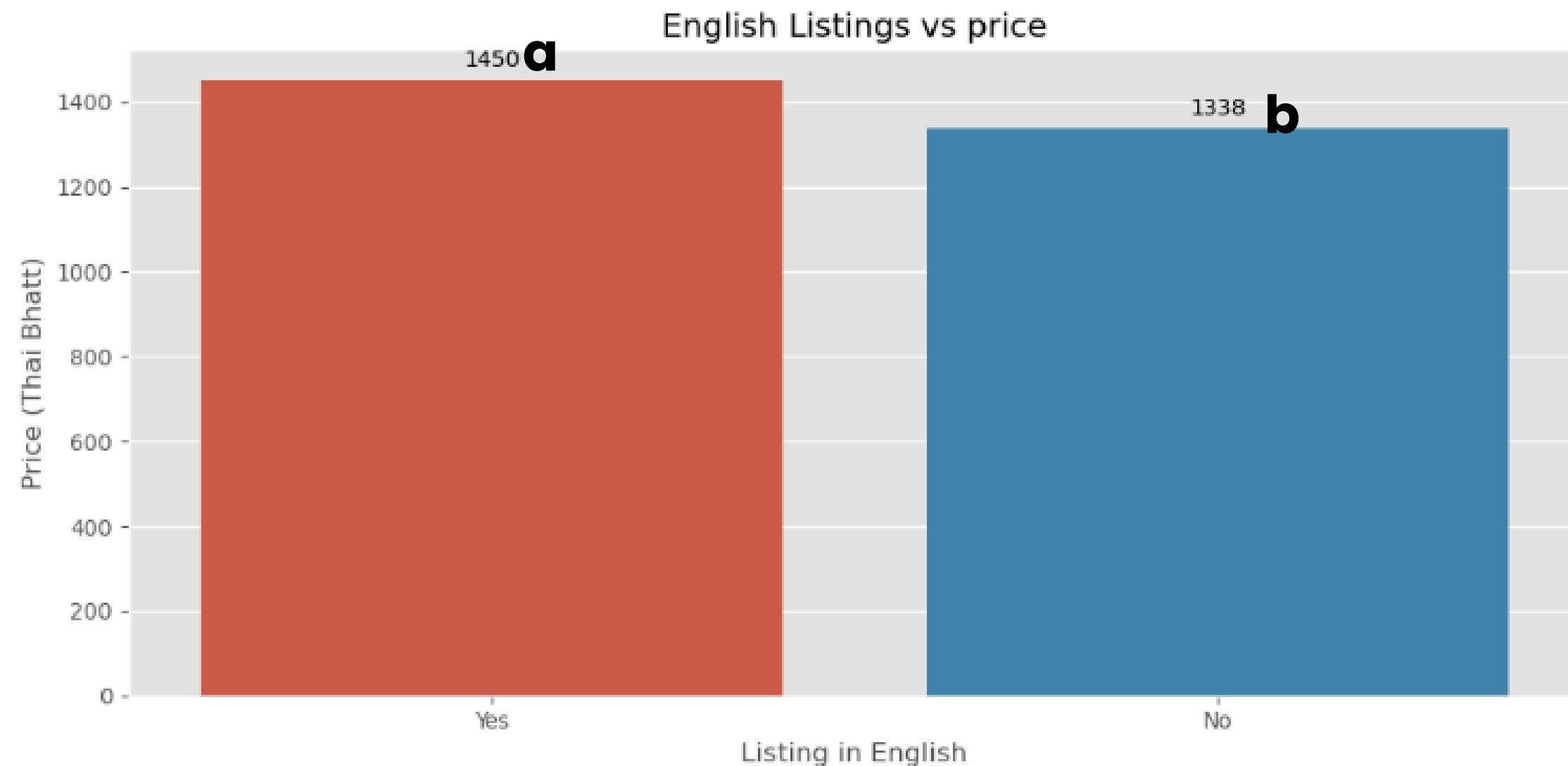
Which listings are more popular? listings that use English vs Non English?



Listings that use English are more popular/significantly higher in number of reviews than Non-English listings ($p\text{-value} < 0.05$). There is no difference in trend between all-time reviews and recent reviews.

Listing Popularity Pattern Analysis

Which listings are higher in price?



Listings that use English are significantly higher in price than Non-English listing ($p\text{-value} \leq 0.05$)



Analysis of the most frequent words in the listing description

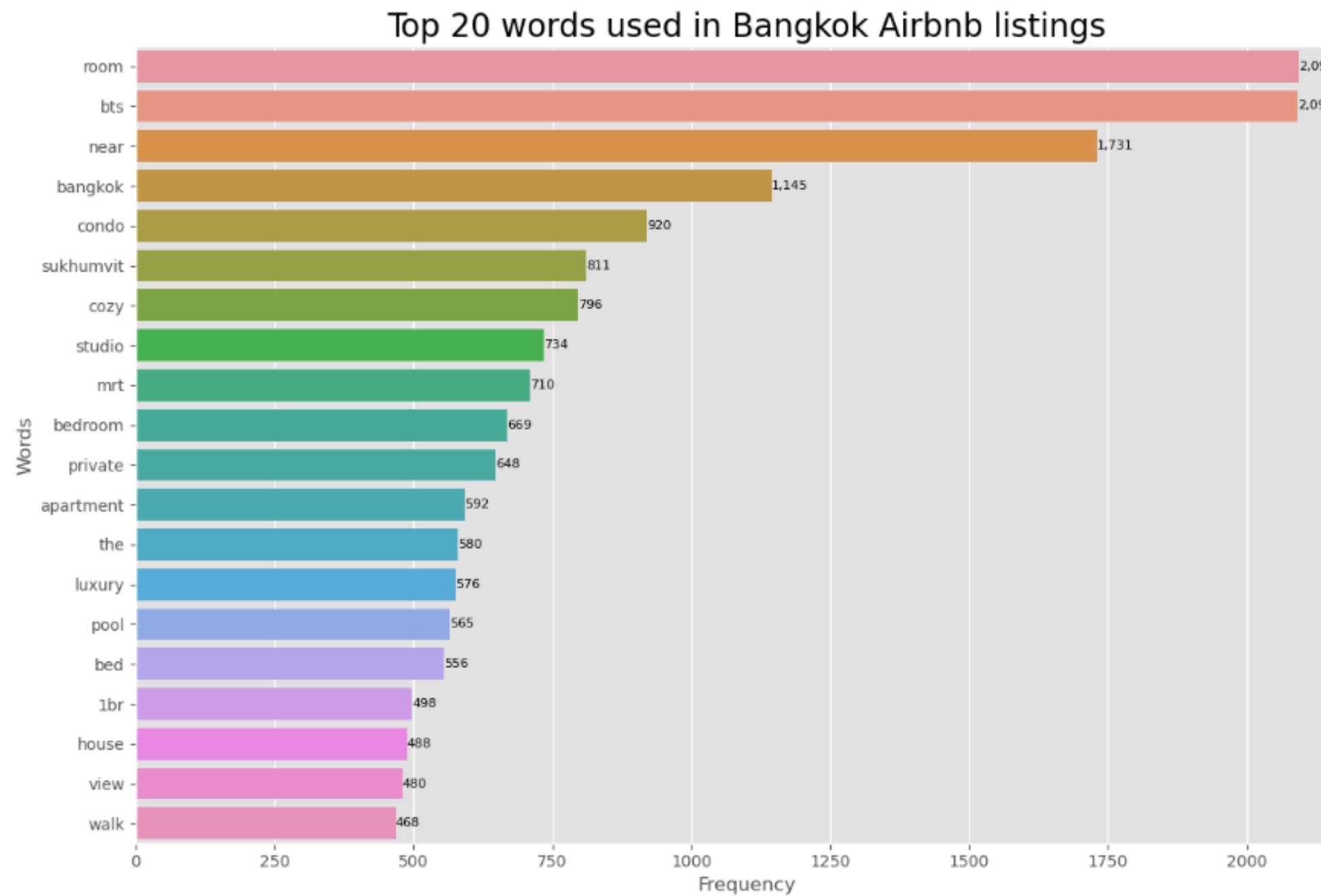


What are the top 20 most frequent words or terms used in listing description from the current airbnb Bangkok dataset?



Analysis of the most frequent words in the listing description

Most frequent words or terms

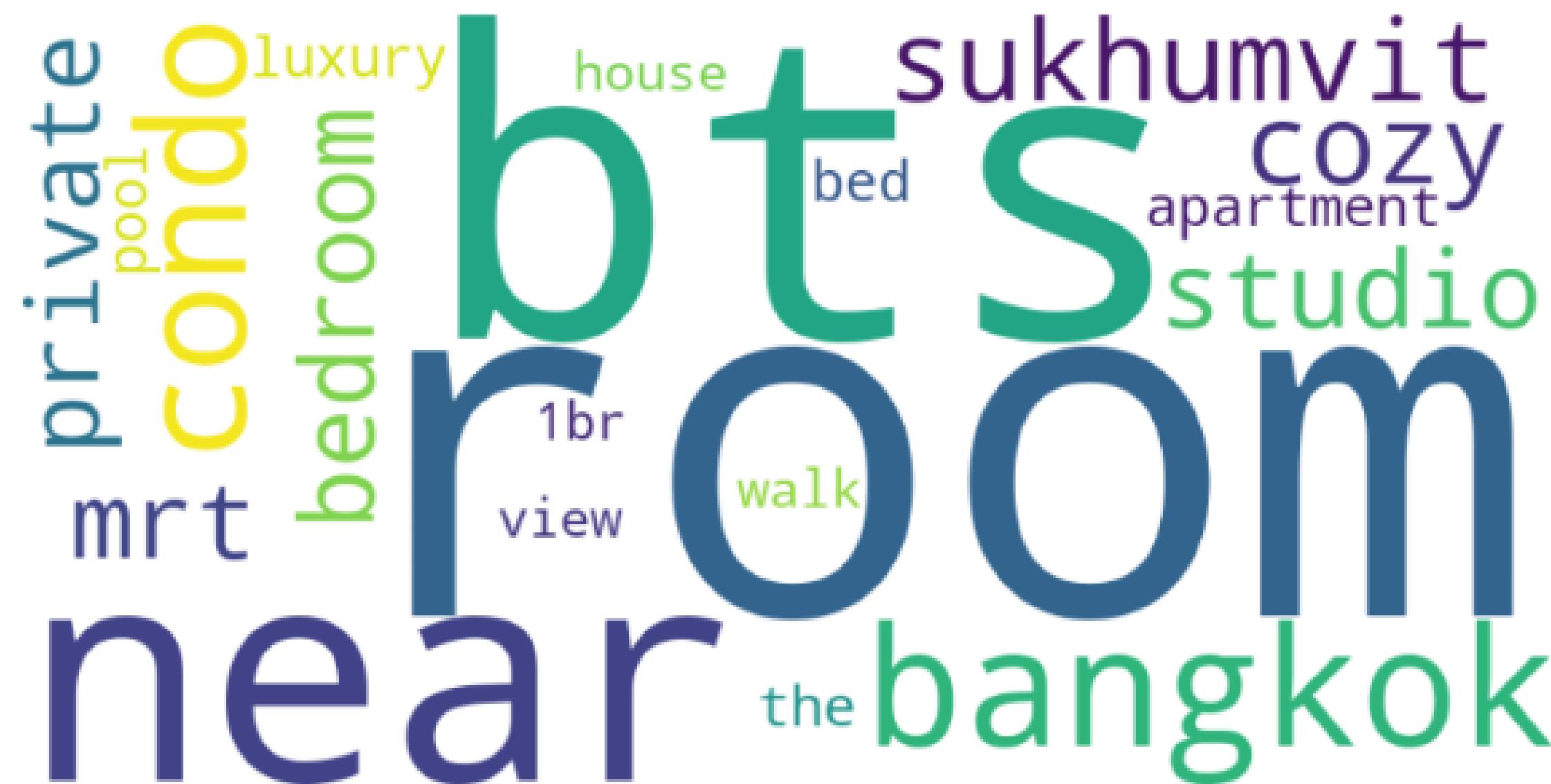


the current listing providers are trying to highlight 3 major terms:

- **Room types & facilities:** room, condo, studio, apartment, 1br, bed, & house
- **Nearby attractions & access:** Sukhumvit, BTS, MRT, walk, & near
- **State of comfort:** cozy, private, luxury, & view

Analysis of the most frequent words in the listing description

Word cloud



Insights

Listing pattern analysis: neighborhood & room types

- The highest number of listings are located near the attraction area
- The most common room type of listing in Bangkok is the entire home/apt.

Listing pattern analysis: price

- Listings with the highest prices are located near the attraction area
- The highest price of listing type is the hotel room

Listing popularity analysis: Total reviews

- Listings that generated the highest number of reviews are concentrated near the attraction area
- The entire home/apt is the most popular listing

Insights

Listing description language Eng vs Non-Eng

- Listings that use English for listing description are significantly higher in number of reviews and price than Non-English language.**

Listing most frequent words

- The top 20 words that are the most commonly used in the listing description consist of room, bts, near, condo, Sukhumvit, cozy, studio, art, bedroom, private, apartment, luxury, pool, bed, 1br, house, view, walk.**

Recommendations

- In highly competitive locations, the new listing providers need to find their own additional uniqueness that can be promoted in the listing description and used as a competitive advantage to attract customers.
- If the new listing providers have resources, consider organizing the listing into the entire homes/apt type as this room type is most popular
- The new listing providers need to place their listing in the premium price category (1480-2800), as this category has had the highest number of reviews recently and at all times.
- Considering the use of English for listing promotion to aim for higher prices and generate more reviews.
- In formulating a listing description, consider highlighting the nearby major attraction and access as selling points, providing an accurate description of room type and listing facilities, and adding the element of comfort



THANK YOU

My Contact

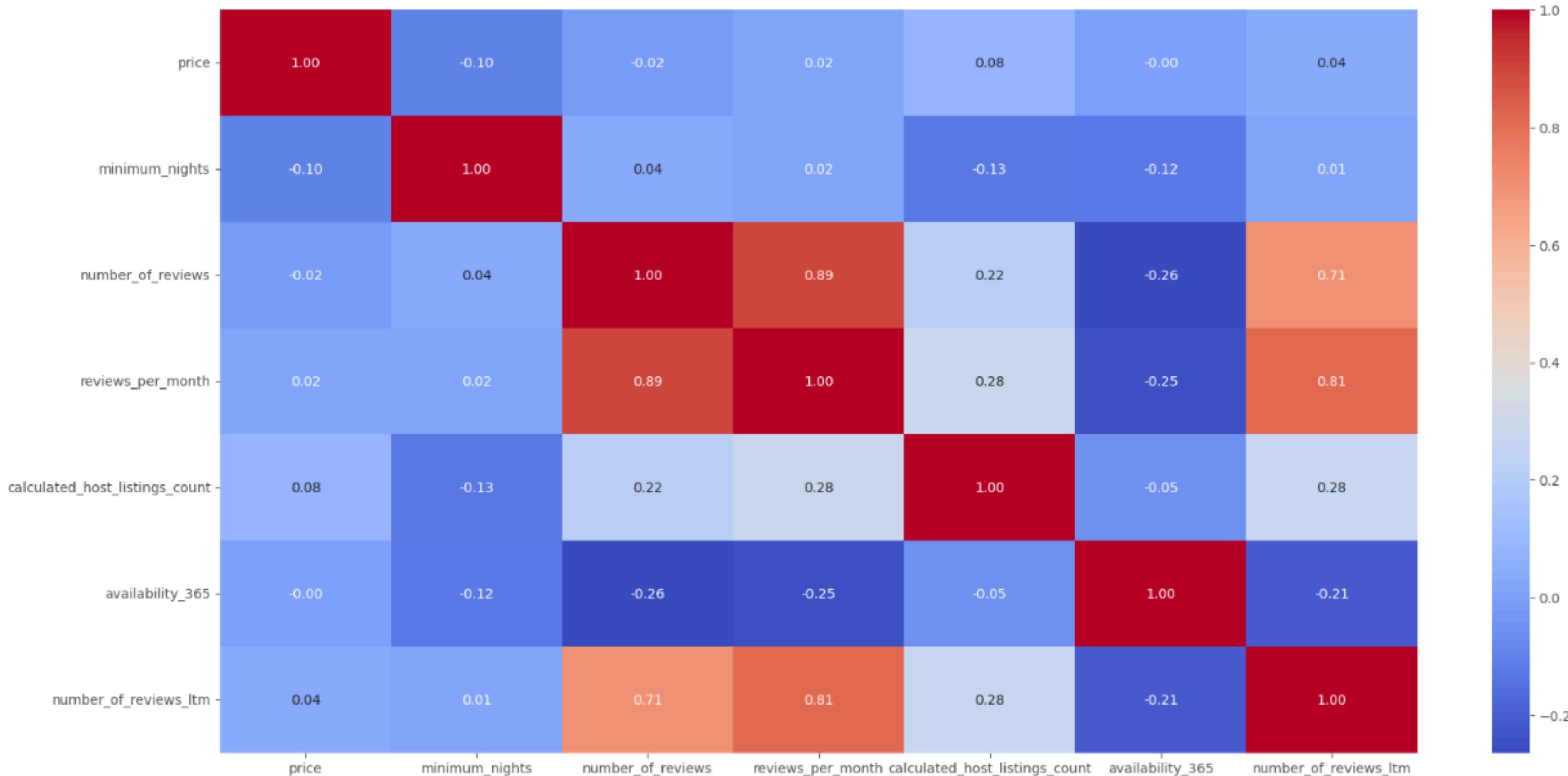
 <https://github.com/harishmuh>

 www.linkedin.com/in/harish-muhammad-7b600b102/

 harishmuh@gmail.com

Appendix

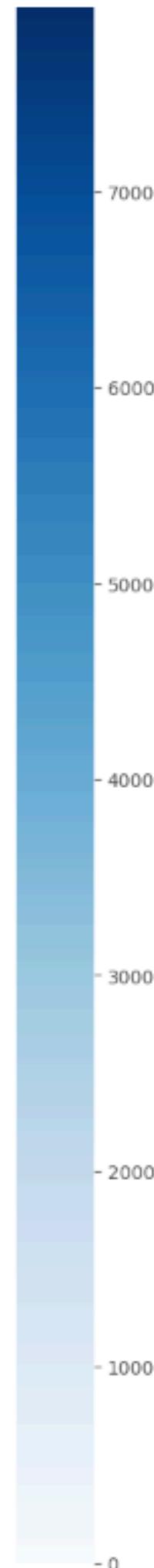
Heatmap correlation (Spearmann)



Listing price recommendation

Median Price by Neighbourhood and Room Type

Neighbourhood	Entire home/apt	Hotel room	Private room	Shared room
				Room Type
Bang Bon	1600	-1	900	-1
Bang Kapi	900	600	995	346
Bang Khae	870	750	975	-1
Bang Khen	892	1286	850	550
Bang Kho laen	1100	7765	1700	-1
Bang Khun thain	2057	-1	1000	7276
Bang Na	980	2200	1188	1000
Bang Phlat	900	2100	900	450
Bang Rak	2228	1690	1757	400
Bang Sue	853	-1	1000	519
Bangkok Noi	950	1800	1180	460
Bangkok Yai	1490	1700	1143	400
Bueng Kum	1085	-1	1120	-1
Chatu Chak	1100	4944	1102	829
Chom Thong	1300	-1	1000	1500
Din Daeng	1347	1000	1000	390
Don Mueang	1200	965	800	384
Dusit	1940	-1	1400	600
Huai Khwang	1375	2599	1290	350
Khan Na Yao	915	-1	1500	-1
Khlong Sam Wa	1750	-1	922	-1
Khlong San	1474	2486	1200	440
Khlong Toei	1732	1816	1497	560
Lak Si	923	608	800	-1
Lat Krabang	1400	1288	849	390
Lat Phrao	3900	-1	900	5000
Min Buri	1374	-1	850	800
Nong Chok	3832	-1	822	-1
Nong Khaem	3160	-1	501	-1
Parthum Wan	2590	3000	2000	500
Phasi Charoen	1000	1106	875	600
Phaya Thai	1512	850	1400	650
Phra Khanong	1004	1000	1000	470
Phra Nakhon	2331	1349	1350	550
Pom Prap Sattru Phai	2078	745	1851	514
Pra Wet	1200	1311	900	525
Rat Burana	1048	-1	850	-1
Ratchathewi	1650	1504	1429	470
Sai Mai	2650	-1	570	-1
Samphanthawong	2300	2900	1586	620
Saphan Sung	2199	-1	790	313
Sathon	1400	2500	1200	490
Suanluang	1004	1000	1000	388
Taling Chan	2500	-1	1381	775
Thawi Watthana	800	-1	1600	-1
Thon buri	1116	1400	1140	320
Thung khru	1057	-1	700	-1
Vadhana	2100	2440	1796	565
Wang Thong Lang	800	-1	810	1600
Yan na wa	1668	7943	1090	1135



Bangkok Tourist maps



Bangkok Rail maps

