

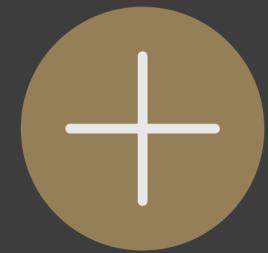


Riyadh AirBnb Analysis - Sep 2024

Noura Abdullah

Contents

- o1 - Clean Data
- o2 - Adding New Columns
- o3 - Data Analysis and Correlation
- o4 - Questions Answered
- o5 - Conclusions



01 - Clean Data



In this step, I focused on cleaning the data to ensure its quality and reliability. This included:

- I replaced repeated values in the **city** column from Arabic to English, such as: الرياض Riyadh.
- Cleaning the Address Column: I unified similar addresses into a standardized format:
 1. الرياض, Riyadh Province, Saudi Arabia Riyadh, Riyadh Province, Saudi Arabia
 2. الرياض, Saudi Arabia Riyadh, Riyadh Province, Saudi
- Cleaning the **Amenities** Column: I removed unnecessary characters from the **previewAmenities** column to ensure the data is organized.

02 - Adding New Columns

In this step, I added new columns to facilitate analysis.

- Neighborhood: Identifies the nearest neighborhood based on the property's latitude and longitude.

```
▶ neighborhood_df = pd.DataFrame(riyadh_neighborhoods)

def find_neighborhood(lat, lon):
    distances = np.sqrt((neighborhood_df['latitude'] - lat) ** 2 + (neighborhood_df['longitude'] - lon) ** 2)
    nearest_index = distances.idxmin()
    return neighborhood_df.iloc[nearest_index]['name']

df['neighborhood'] = df.apply(lambda row: find_neighborhood(row['latitude'], row['longitude']), axis=1)

print(df[['latitude', 'longitude', 'neighborhood']].head())
```

	latitude	longitude	neighborhood
0	24.707180	46.671690	Al Olaya
1	24.887580	46.590042	Al Sahafah
2	24.694460	46.814579	Al Manar
3	24.766002	46.626816	Hittin
4	24.7776365	46.629321	Al Ghadeer

- City Quarter: Categorizes the property by its urban quarter (e.g., Central Riyadh, North Riyadh, East Riyadh).

```
▶ city_quarters = {
    "Central Riyadh": [
        "Al Olaya",
        "Al Malaz",
        "Al Murabba",
        "Al Sulaymaniyah",
        "Al Nasiriyah",
        "Al Rabwah"
    ],
    "North Riyadh": [
        "Hittin",
        "Al Yasmin",
        "Al Malqa",
        "Al Sahafah",
        "Al Muruj",
        "Al Nafeel",
        "Al Ghadeer",
        "An Nakheel",
        "Al Wurud",
        "Qurtubah"
    ],
    "East Riyadh": [
        "Al Manar",
        "Al Rawdah",
        "Al Khaleej",
        "Al Munsiyah",
        "Al Hamra",
        "Al Qadisiyah",
        "Al Fayha",
        "Al Izdihar"
    ]
}

[15] def get_city_quarter(neighborhood):
    for quarter, neighborhoods in city_quarters.items():
        if neighborhood in neighborhoods:
            return quarter
    return None

df['city_quarter'] = df['neighborhood'].apply(get_city_quarter)

▶ print(df[['neighborhood', 'city_quarter']].head())
```

	neighborhood	city_quarter
0	Al Olaya	Central Riyadh
1	Al Sahafah	North Riyadh
2	Al Manar	East Riyadh
3	Hittin	North Riyadh
4	Al Ghadeer	North Riyadh

02 - Adding New Columns

- Nearby Landmarks: I added the Nearby Landmarks column, which lists landmarks within a specified distance from the property. The landmarks included are:



```
▶ landmarks = {
    'Historical Diriyah': (24.7335, 46.5758),
    'Mamlaka Tower': (24.7311, 46.6701),
    'King Fahd National Library': (24.6858, 46.6870),
    'National Museum': (24.6477, 46.7102),
    'Hall of King Abdulaziz': (24.6438, 46.7115),
    'Murabba Palace': (24.6464, 46.7092),
    'Boulevard': (24.7694, 46.6046),
    'King Khalid Airport': (24.9596, 46.7024)
}
```

```
▶ def find_landmarks(row, landmarks, threshold=12):
    listing_location = (row['latitude'], row['longitude'])
    nearby_landmarks = []
    distances = {}

    for landmark, loc in landmarks.items():
        distance = geodesic(listing_location, loc).kilometers
        if distance <= threshold:
            nearby_landmarks.append(landmark)
            distances[landmark] = distance

    if nearby_landmarks:
        sorted_landmarks = sorted(distances.items(), key=lambda x: x[1])
        nearby_landmarks = [landmark for landmark, _ in sorted_landmarks[:2]]

    return nearby_landmarks

df['Nearby Landmarks'] = df.apply(find_landmarks, axis=1, landmarks=landmarks)

print(df[['neighborhood', 'city_quarter', 'Nearby Landmarks']].head())
```

	neighborhood	city_quarter	Nearby Landmarks
0	Al Olaya	Central Riyadh	[Mamlaka Tower, King Fahd National Library]
1	Al Sahafah	North Riyadh	[Boulevard, Historical Diriyah]
2	Al Manar	East Riyadh	[National Museum, Hall of King Abdulaziz]
3	Hittin	North Riyadh	[Boulevard, Mamlaka Tower]
4	Al Ghadeer	North Riyadh	[Boulevard, Mamlaka Tower]

03 - Data Analysis and Correlation

In this step, I performed various analyses to uncover insights from the data, including:

- Distance Analysis: I calculated the distances to nearby landmarks for each neighborhood, allowing for a better understanding of location appeal.

Neighborhood	Historical Diriyah	Mamlaka Tower	King Fahd National Library	National Museum	Hall of King Abdulaziz	Murabba Palace	Boulevard	King Khalid Airport
Al Olaya	10.1305	2.65447	2.83	7.65507	8.0946	7.72914	9.67284	28.1325
Al Sahafah	8.33124	11.7187	16.6781	21.4806	21.9117	21.5427	4.47808	20.3092
Al Manar	24.5443	15.1719	12.9469	11.7663	11.8469	11.921	22.8075	31.483
Hittin	6.29224	5.84091	10.7703	15.5855	16.0205	15.6527	2.27826	22.7648
Al Ghadeer	7.29082	6.49265	11.6955	16.4345	16.875	16.5998	2.61646	21.5989
Al Hamra	23.4897	15.2421	17.6055	20.0761	20.4192	20.2488	19.4659	18.5219
Al Khaleej	27.028	17.5854	17.4574	17.8719	18.0663	18.0434	24.0561	26.6768
Al Yasmin	13.4537	6.18965	10.5699	14.7758	15.2174	14.9116	9.56545	19.826
Al Maqta	7.40958	8.48192	13.5544	18.3793	18.8171	18.4505	2.49741	20.776
Al Rawdah	16.0928	7.20532	9.79798	13.1105	13.5169	13.2692	12.8189	21.8088
Al Murabba	17.9772	10.2419	5.57297	2.70211	2.82477	2.8604	17.8809	33.4254
Al Nafel	10.4119	3.71026	8.83142	13.4742	13.9257	13.5848	6.9367	21.8159
Al Rayha	19.1761	9.66877	8.61722	9.52622	9.80941	9.70224	17.158	27.5003
Qurtubah	21.6683	13.5631	16.2261	19.0016	19.3662	19.1713	17.624	18.1297
Al Izdihar	13.7567	5.35474	9.07244	13.0637	13.4992	13.2057	10.383	21.4945
An Nakheel	8.22941	1.6221	6.67326	11.5013	11.9448	11.582	6.24641	24.8216
Al Muruj	10.987	2.38247	7.01804	11.5579	12.0088	11.6745	8.25022	23.4039
Al Wurud	12.2103	3.63031	2.25471	6.64951	7.10099	6.76236	11.3664	28.1624

- Average Price Analysis: I analyzed the average price based on host status:

1. Superhosts : Higher average prices.
2. Non-Superhosts : Lower average prices.

Average Price by Super Host Status:

isSuperhost	TotalPrice
False	2705.6
True	2852

Correlation between super host status and price: 0.04

03 - Data Analysis and Correlation

- City Quarter Analysis: I examined how prices and ratings vary across different city quarters, highlighting the influence of location on guest experiences.

city_quarter	TotalPrice	rating	PricePerDay
Central Riyadh	2615.96	4.86154	211.462
East Riyadh	2837.72	4.9027	202.581
North Riyadh	2793.25	4.95019	265.364

- Correlation Analysis: I calculated the correlation coefficient to identify relationships between the number of guests and pricing. I analyzed how the number of guests affects daily pricing.

	persons	PricePerDay
persons	1	0.472108
PricePerDay	0.472108	1

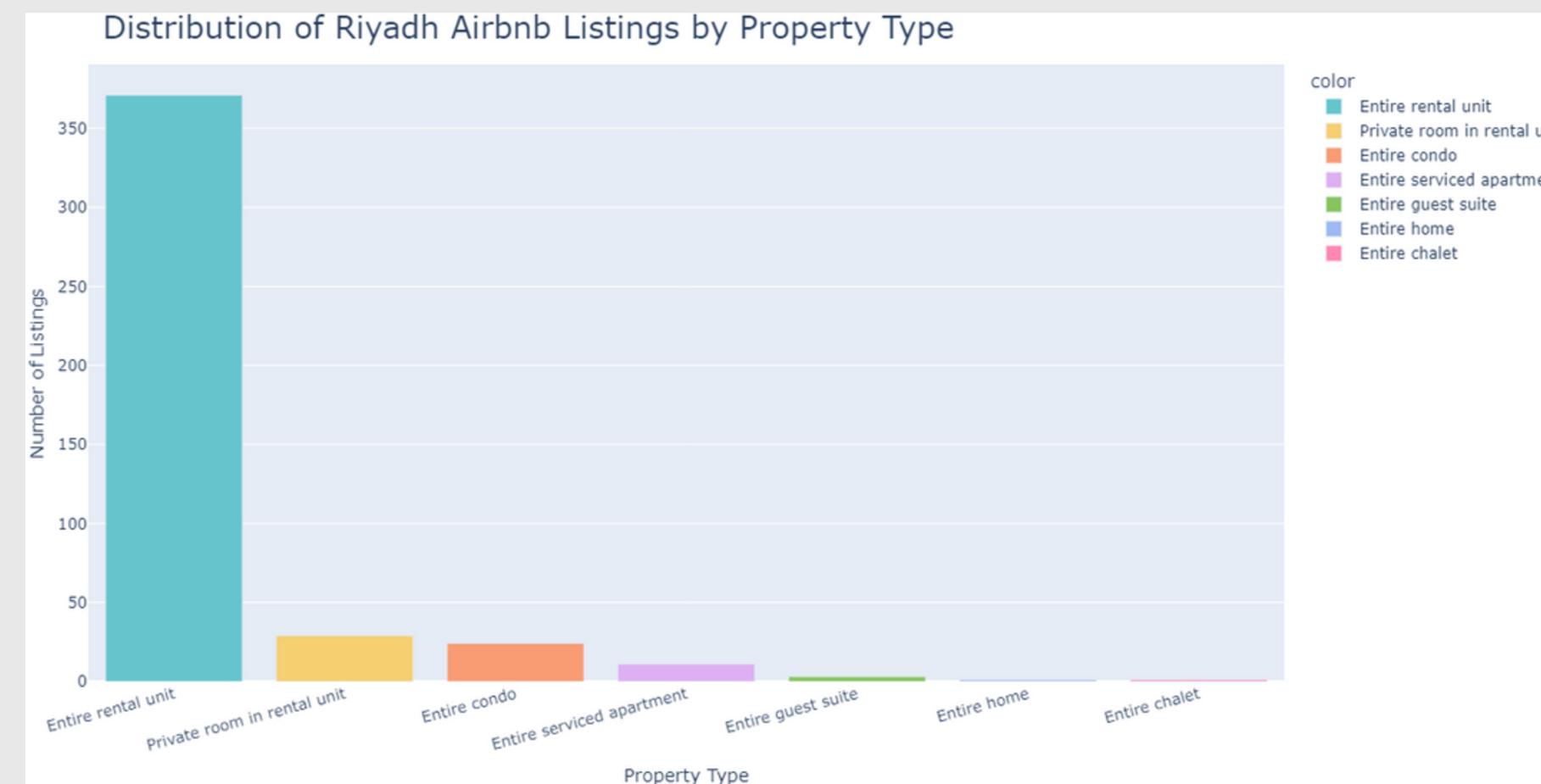
Correlation between persons and PricePerDay: 0.47210762060751366

04 - Questions Answered

In this section, I addressed key questions by visualizing the data through charts and graphs:

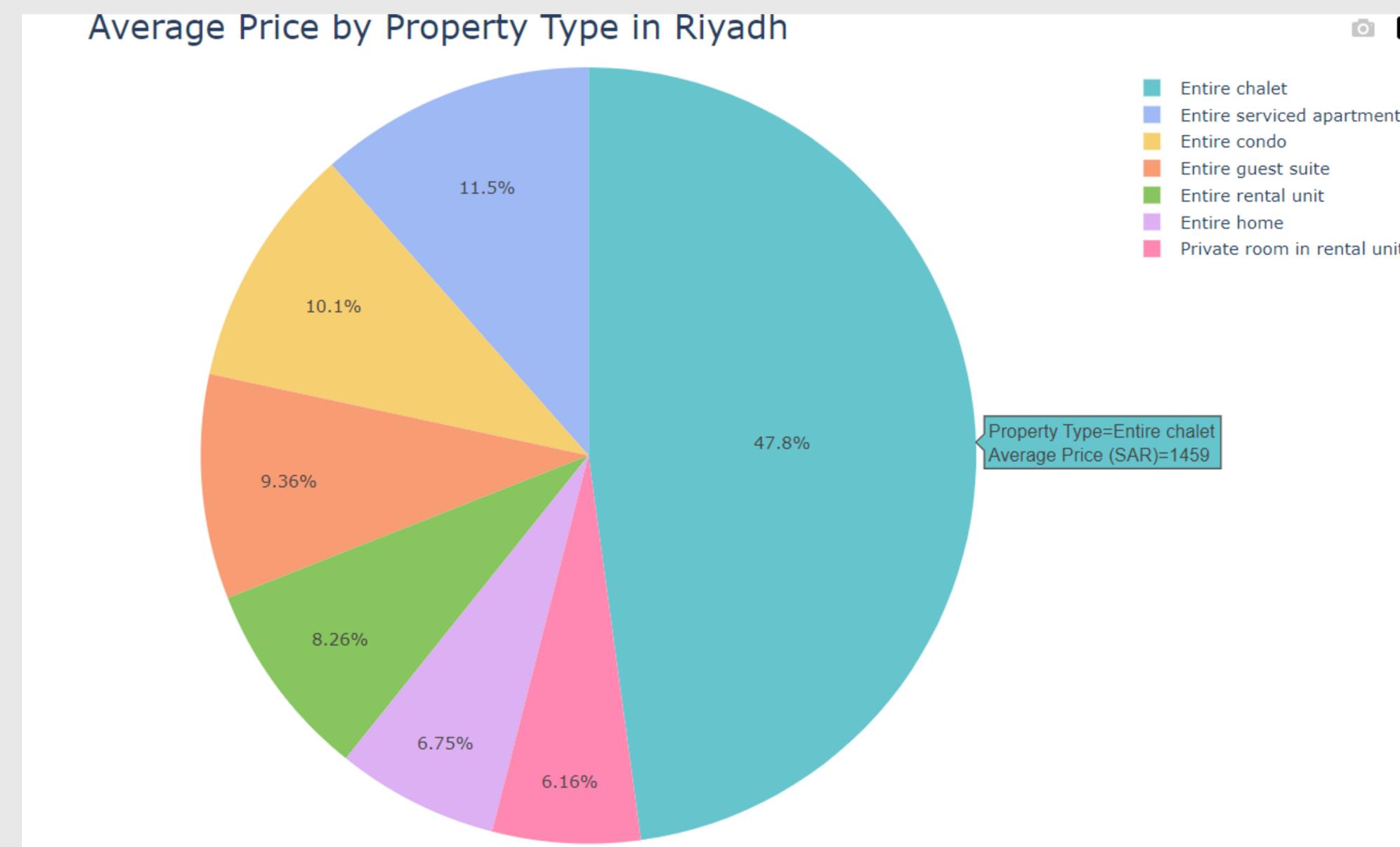
- What are the most sought-after properties in the Riyadh Airbnb market?

The analysis revealed that the most common property type is Entire rental unit, with 371 listings.



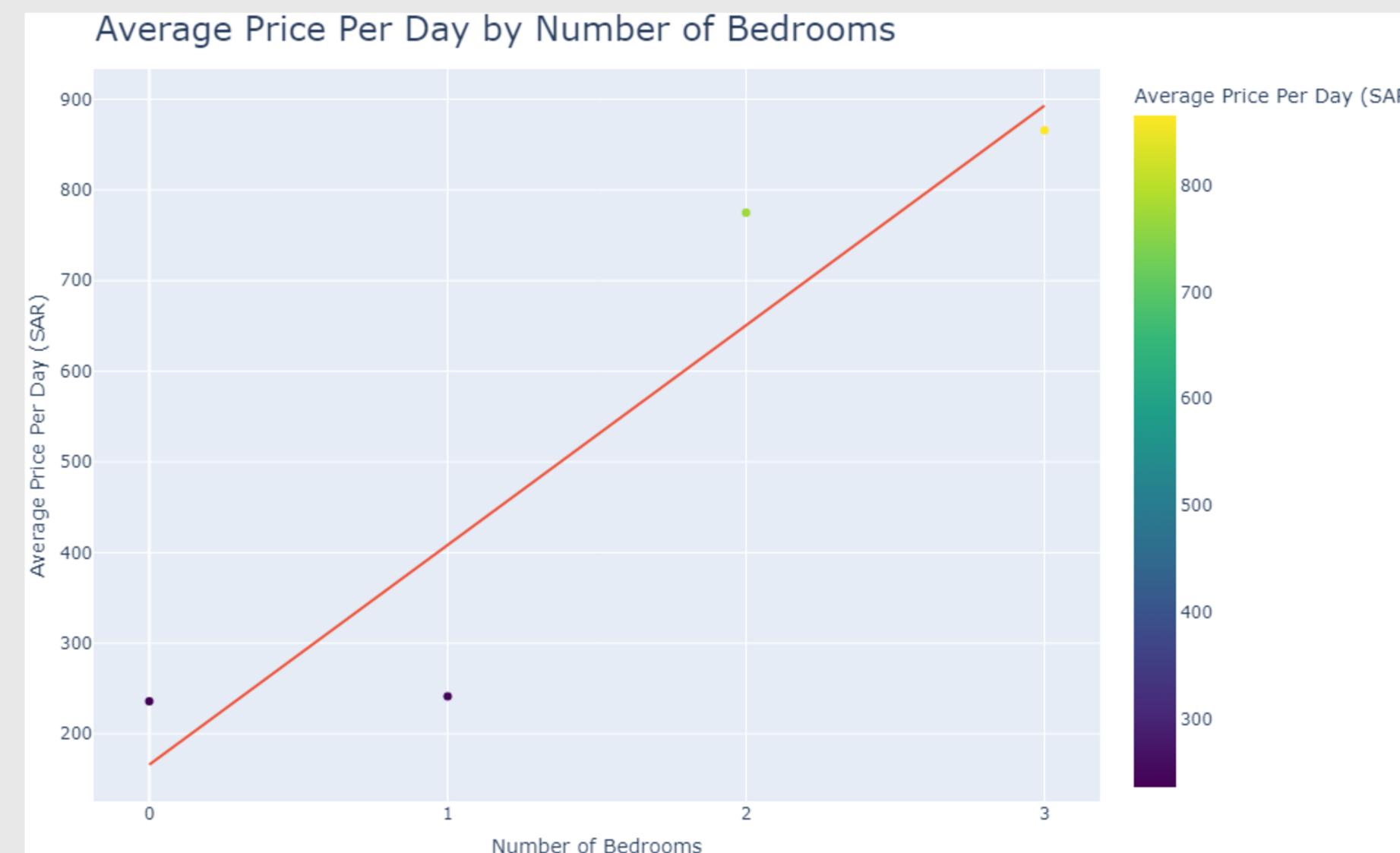
04 - Questions Answered

- What is the average daily rental price for each type of property in Riyadh? The average daily rental prices vary significantly across property types. For instance, Entire chalet has the highest average price at 1459 SAR, while Private room in rental unit is approximately 188 SAR.



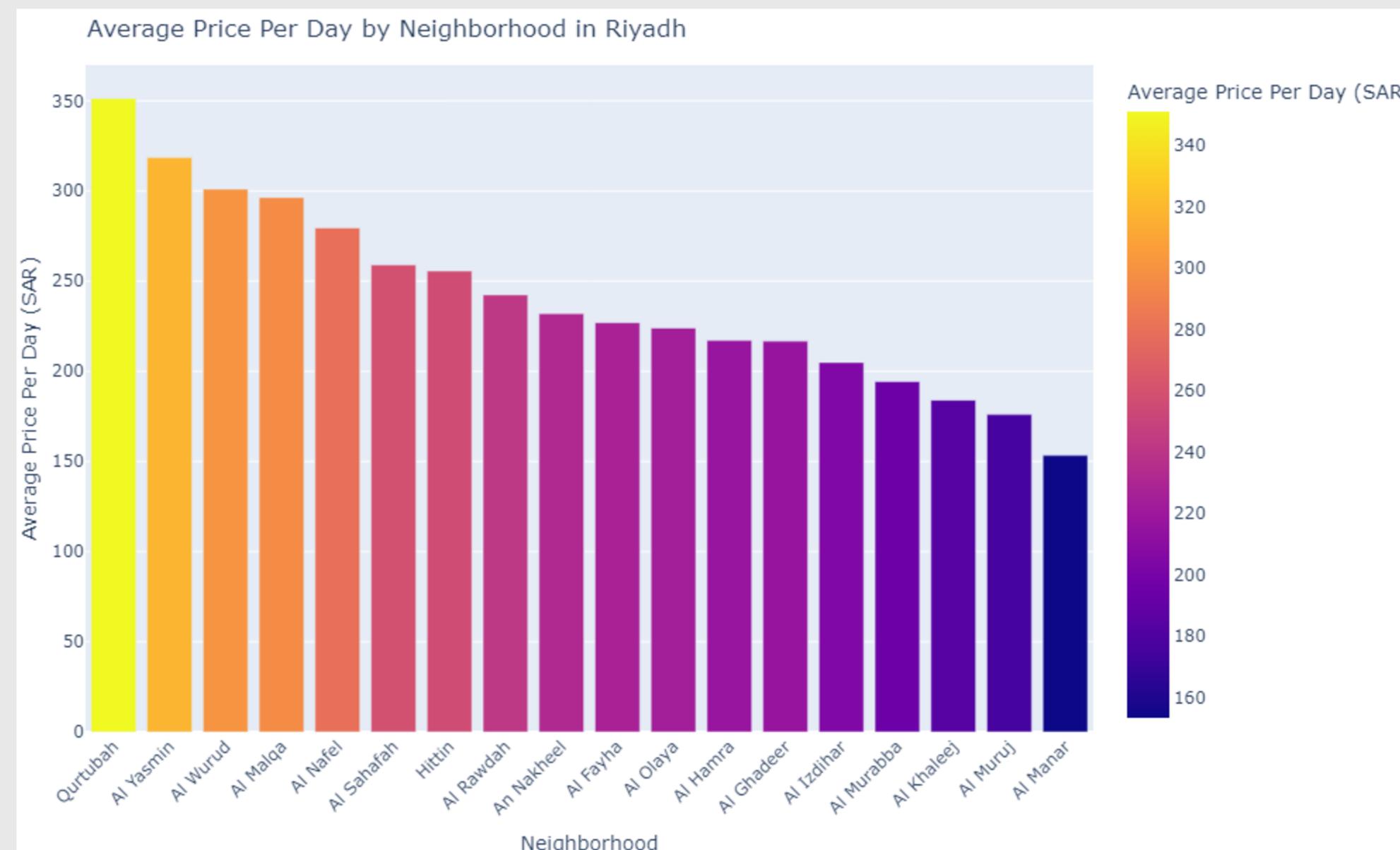
04 - Questions Answered

- What is the impact of the number of bedrooms on the average daily rental price in Riyadh? The analysis showed that as the number of bedrooms increases, the average price per day also tends to increase significantly. For instance, properties with two bedrooms have an average price of approximately 775 SAR, while those with three bedrooms average around 866 SAR. This indicates that larger properties generally command higher rental prices.



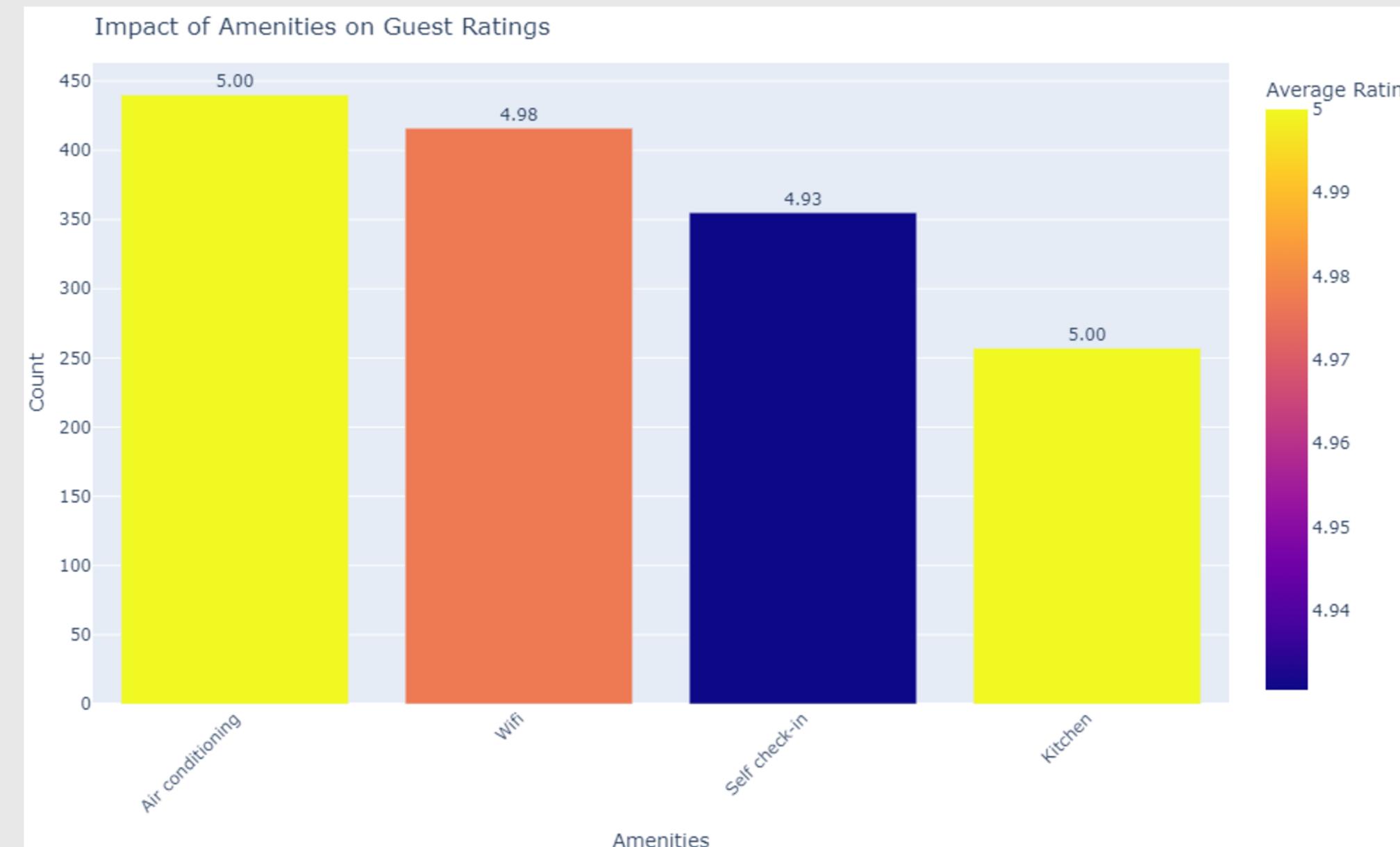
04 - Questions Answered

- How does the geographical location (neighborhood) affect the average rental prices in Riyadh? The analysis revealed notable variances in average prices across different neighborhoods. For example, Qurtubah has the highest average price of 351 SAR, while Al Muruj has an average price of 176 SAR. This highlights the significant impact of location on rental pricing.



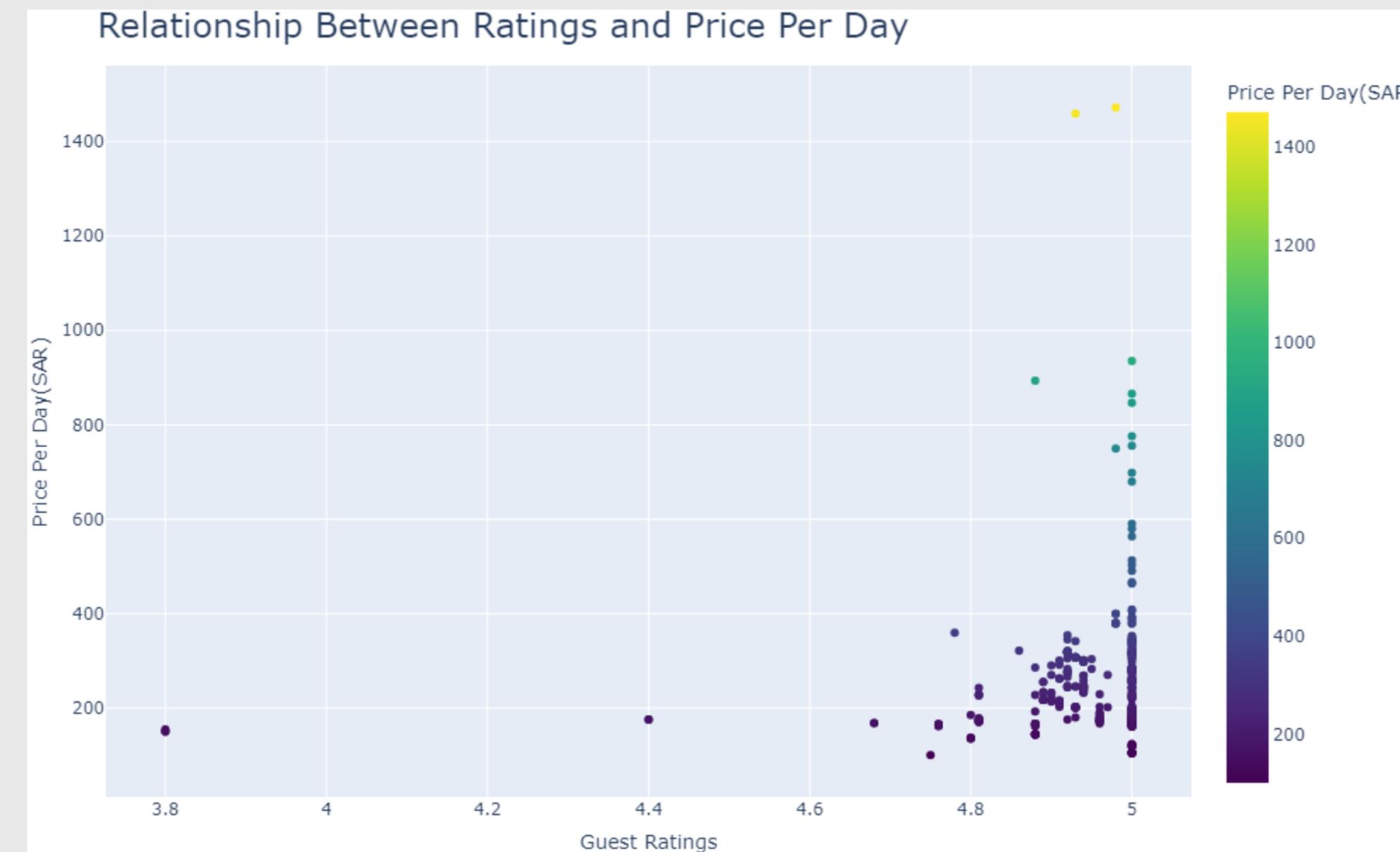
04 - Questions Answered

- What amenities do guests prefer in Riyadh, and how do these amenities affect guest ratings of properties? The analysis showed that the most preferred amenities include Air conditioning and Wifi, which are linked to high average ratings of 5.00 and 4.98, respectively.



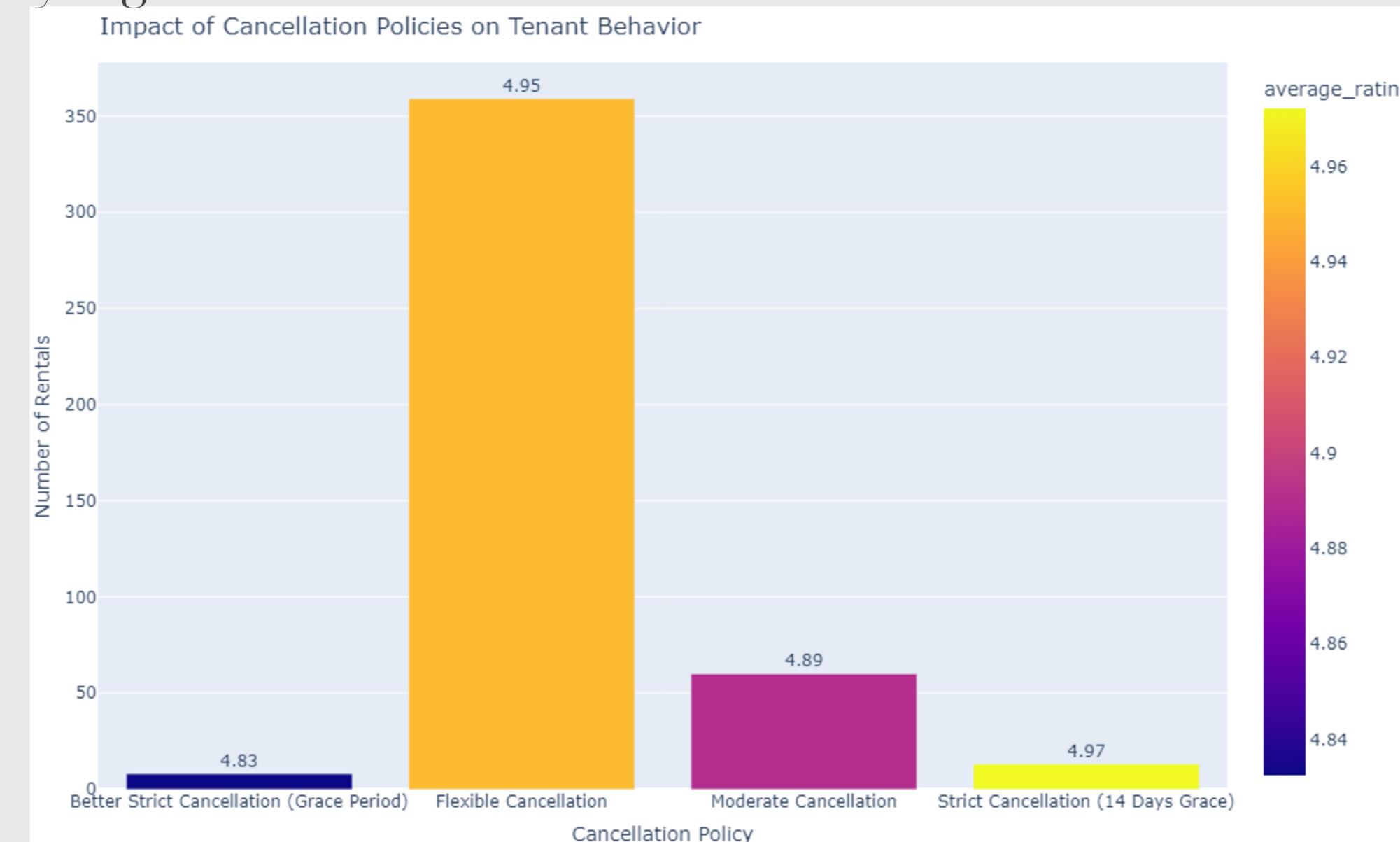
04 - Questions Answered

- How does the daily rental price relate to guest ratings in the Riyadh Airbnb market? The scatter plot demonstrates a positive relationship between guest ratings and daily rental prices. Higher-rated properties tend to have higher prices.



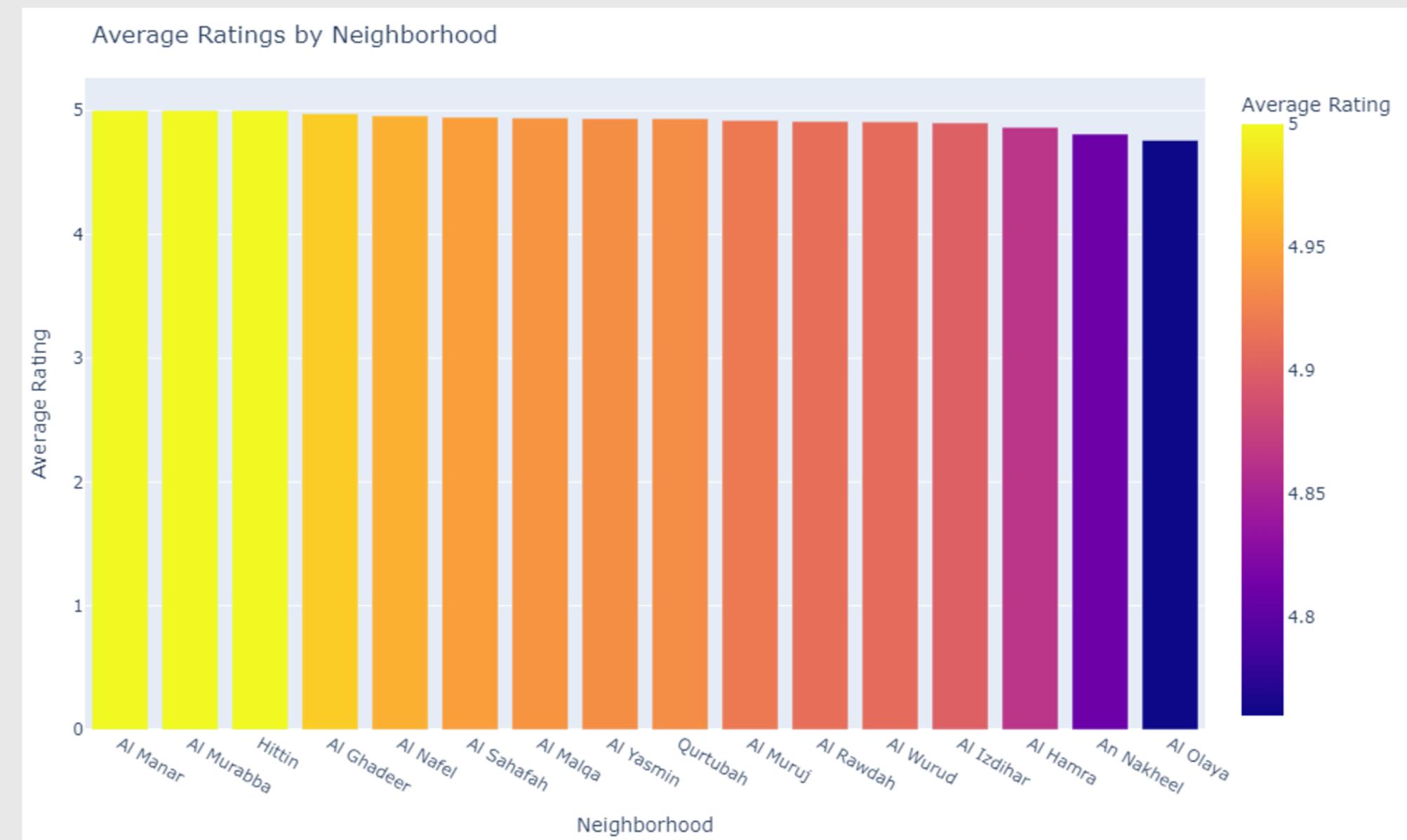
04 - Questions Answered

- How do cancellation policies affect tenant behavior and ratings in Riyadh? The analysis revealed that Flexible Cancellation policies are the most common, with 359 rentals, and they have an average rating of 4.95. In contrast, Better Strict Cancellation policies, while less common (only 8 rentals), have a high average rating of 4.83, indicating that even with stricter policies, guest satisfaction remains relatively high.



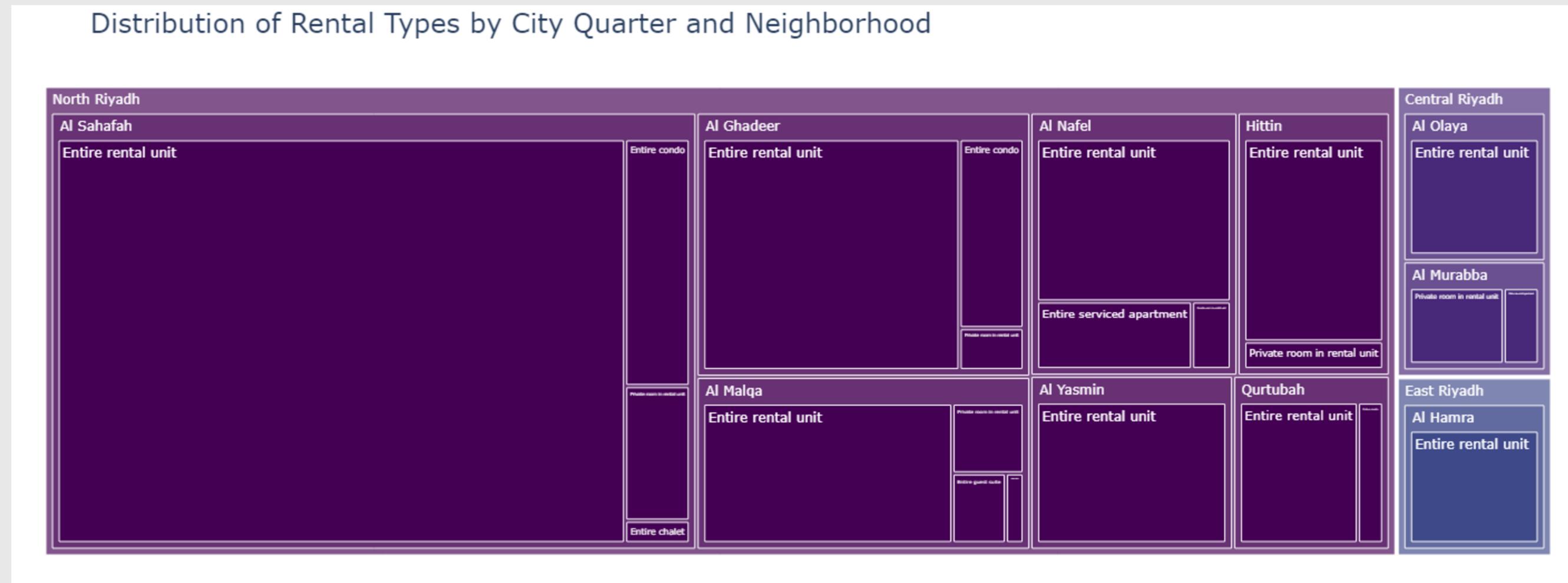
04 - Questions Answered

- Which neighborhoods have the highest ratings in Riyadh? The analysis showed that Al Manar, Al Murabba, and Hittin all have the highest average rating of 5.00, reflecting exceptional guest satisfaction in these areas. This suggests that these neighborhoods are popular among guests for their quality offerings.



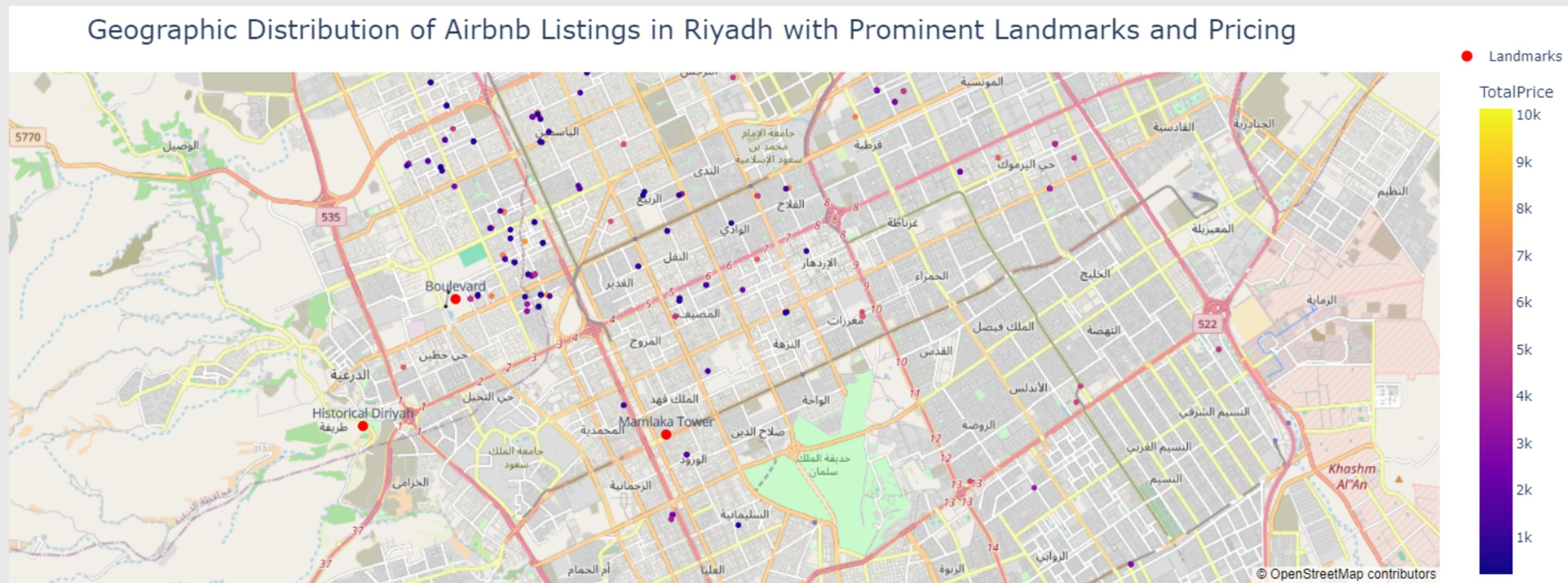
04 - Questions Answered

- How is the distribution of rental types by city quarter and neighborhood? The analysis offers insights into how various rental types are distributed across neighborhoods in Riyadh.



04 - Questions Answered

- What is the distribution of properties and key landmarks in terms of geographical locations? The geographic distribution map illustrates how Airbnb listings are spread across Riyadh, highlighting prominent landmarks.



Conclusions

The analysis of Airbnb listings in Riyadh reveals several key insights: there is a strong demand for entire rental units, indicating that guests prioritize privacy and space.

Neighborhoods such as Al Manar and Al Sahafah demonstrate high ratings and a variety of rental types, with proximity to landmarks significantly influencing pricing.

Essential amenities, like air conditioning and Wi-Fi, positively impact guest ratings and satisfaction.

Additionally, flexible cancellation policies are common and associated with higher ratings, while stricter policies still maintain guest satisfaction.

A positive correlation exists between rental prices and guest ratings, suggesting that guests are willing to pay more for quality accommodations.

Thanks

Noura Abdullah