GSCM 521
Abdi Alamir

Analysis Assignment 1 - Part B

Linking & Cleaning Data via Excel

- Linking Reviews and Listings: Using the listing ID, we linked the AirBnB *Listings* data file and the *Reviews* data file. We started by transforming the Reviews data, to get the reviews by year. A pivot table was created to summarize the count of reviews, grouped by year, for each listing. This created new columns for "review_count_[year]" going from years 2010 to 2023. This new formatted data was linked to the listing data by "listing id" using the Vlookup function.
- Linking Listings and Associations dataset: In the Neighborhood associations data file, we used "split text to column" to generate two columns "Neighborhood" and "Association". Then, we linked this dataset to the listing data using the "neighborhood" variable using the Xlookup function.
- Linking Listings and Portland Neighborhood 2020 Data Profiles: We converted the *Portland Neighborhood 2020 Data Profiles* file from a PDF into an Excel. Each sheet contains the data for a KPI, it includes the "Neighborhood", "percentage/index", "Rank".
  To join the different KPIS to the listings data, we used the Vlookup function to pull in the percentage/index and rank from the *Neighborhood 2020 Data Profiles* file onto the AirBnB *Listings* data file by the 'neighborhood' column.
- Data Warehouse: All the joined data files are put in one consolidated sheet. The file now includes all the data sources along with a sheet that contains the joined datasets together and is now acting as a data warehouse.

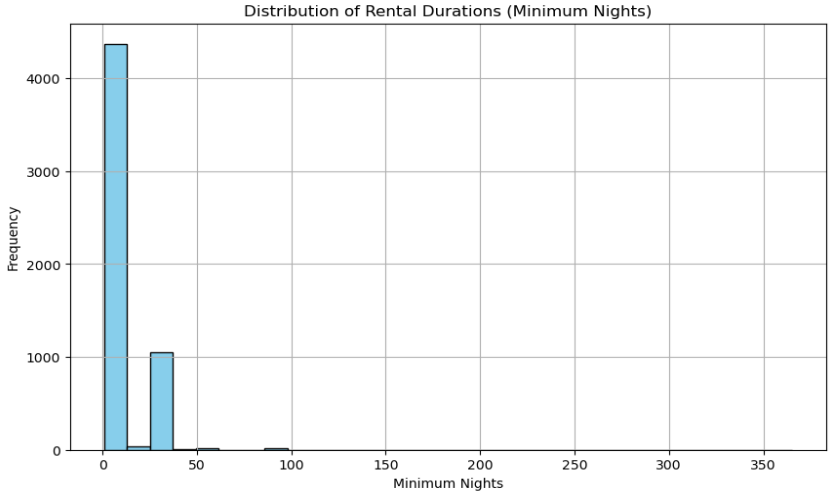
Assumptions

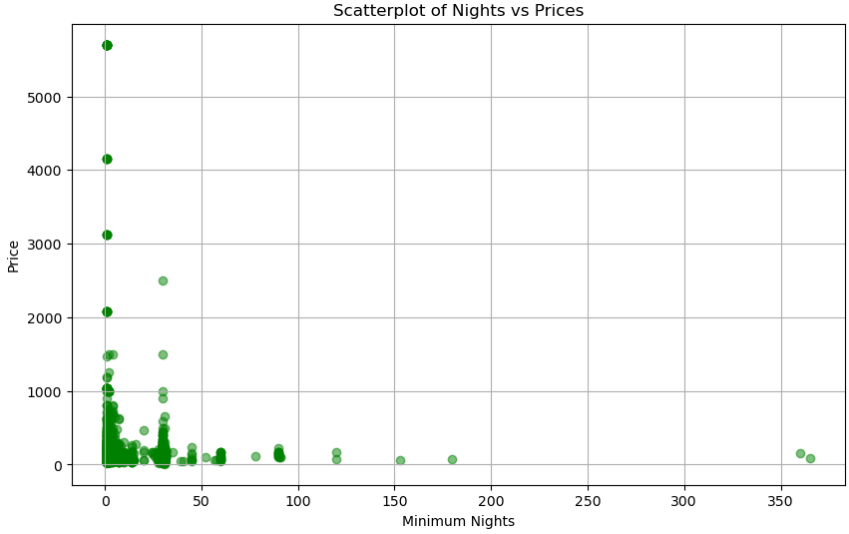
For the purpose of our research, our team has made the assumption that the number of reviews is directly correlated to the popularity of a listing. Without data on review ratings, we cannot accurately determine if a review is positive or negative and therefore have assumed all reviews are positive and equally weighted.

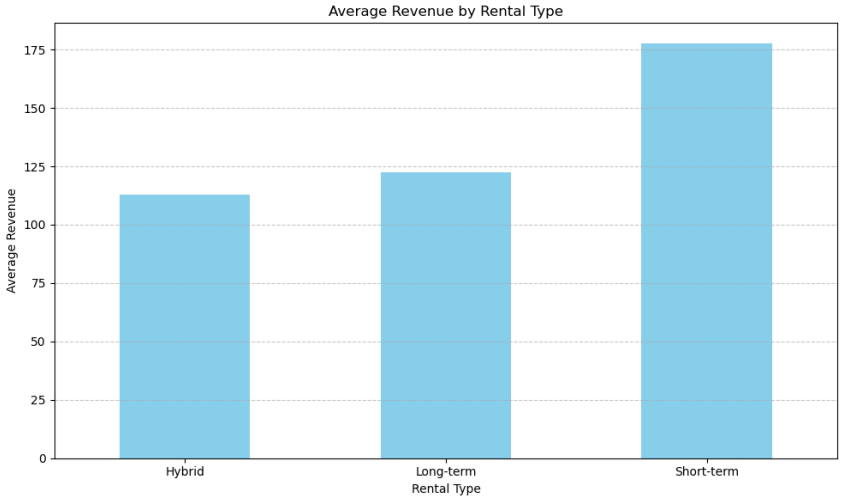
Two assumptions to consider:

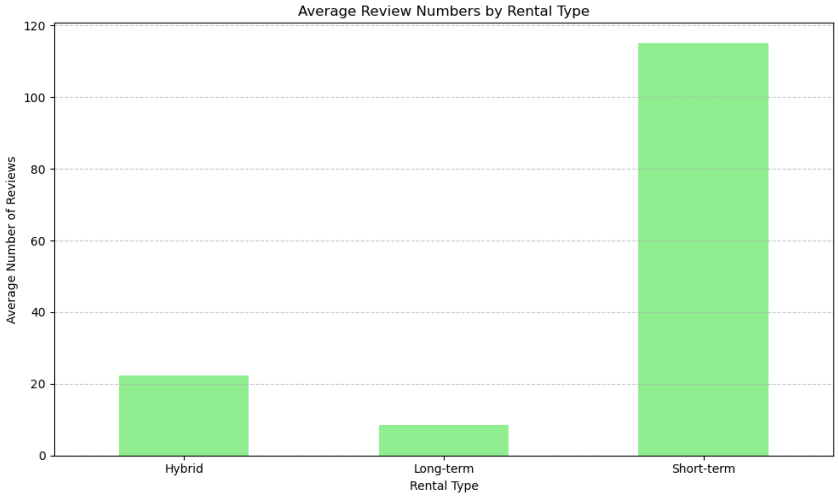
- The number of reviews describes the popularity of a listing.
- Without an information on the ratings, we assume all reviews are positive and equal.

Question 1: Should the company focus on being a short-term, long-term, or a hybrid rental company, considering the market dynamics and potential revenue?









Classifying Our Rental Units:

In our quest for operational insights and strategic planning, we've divided our rental properties into three groups: Short-term, Long-term, and Hybrid. This categorization is based on the duration guests can stay, giving us a closer look at the temporal aspect of our diverse portfolio.

How We Classified the Rentals:

We decided on this categorization by looking at the number of nights guests stay:

- Short-term rentals have a stay of fewer than 30 nights.
- Long-term rentals are for stays longer than 180 nights.
- Hybrid rentals fall in between these two, covering stays within the intermediary range.

This approach helps us understand the time-related patterns in our properties more thoroughly.

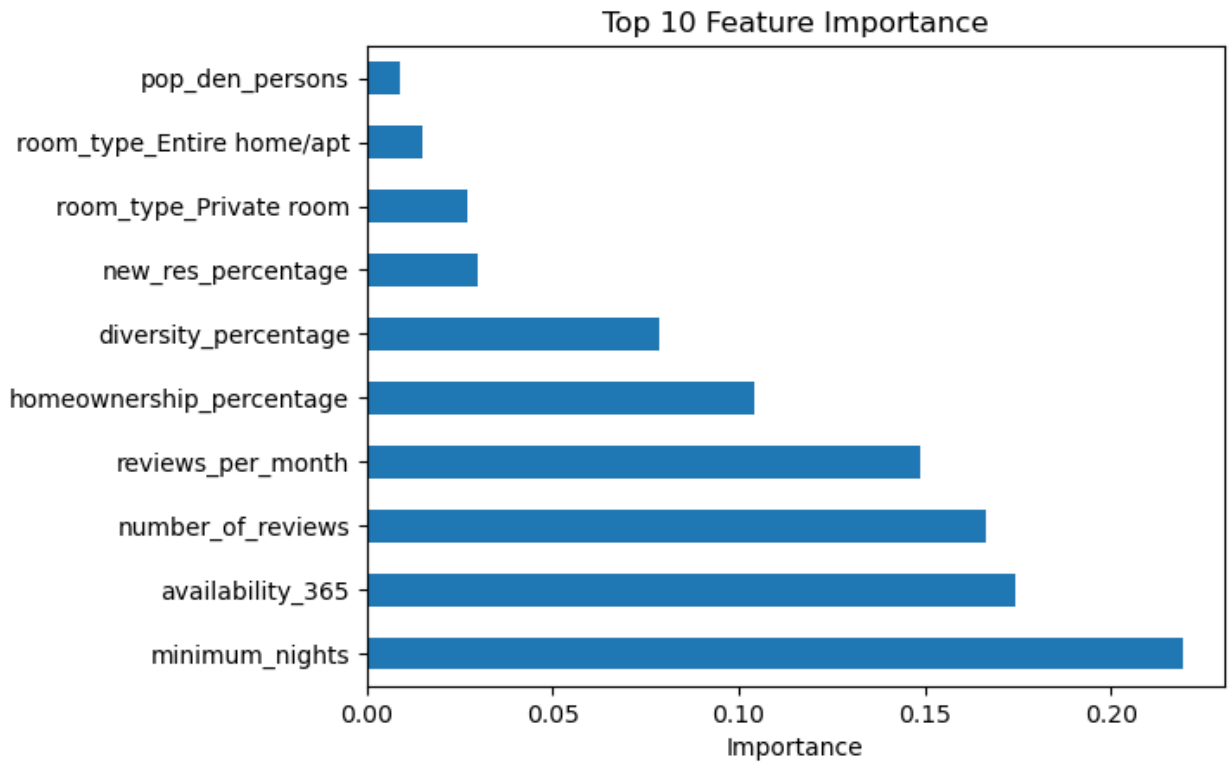
Understanding the Categories:

1. Histogram of Minimum Nights: The histogram shows us that most of our rentals have a stay duration ranging from 0 to 50 nights. It's a visual summary of how long guests typically stay.

2**. Scatterplot of Nights vs Prices:** The scatterplot gives us a visual insight into how the duration of stay relates to the price. In a nutshell, it shows that shorter stays generally bring in more money.
3. Average Revenue by Rental Type (Bar Chart): This bar chart tells us that Short-term rentals make the most money on average. Hybrid rentals come next, and Long-term rentals, while contributing, don't bring in as much revenue.
4. Average Review Numbers by Rental Type (Bar Chart): Looking at reviews, it's clear that guests leave more reviews for Short-term stays, creating a positive buzz. Hybrid rentals also get good reviews, while Long-term rentals get fewer reviews on average.

What This Means for Our Strategy: Understanding these categories helps us plan our expansion:
1. Invest in Short-Term Rentals: Since short-term rentals bring in the most revenue and positive reviews, focusing our expansion efforts here makes sense.
2. Highlight Short-Term Properties: In marketing, we can showcase what makes our short-term rentals special – the location, amenities, and overall experience.
3. Consider Hybrid Variety: While short-term is a start, we can also diversify our hybrid category, this can attract a broader audience looking for something in between.
4. Improve Long-Term Offerings: For long-term rentals, there's room for improvement. We can explore ways to make them more attractive, whether through pricing or added amenities. In essence, these insights guide our strategic decisions for expansion, ensuring we capitalize on the strengths of each rental type.

Question 2: What features significantly influence property prices, and how can the company optimize these factors for competitive pricing?



In our exploration of property pricing determinants, the strategic application of Random Forest Regression unfolded meaningful insights. Going a bit deeper into the pivotal features identified by the model and how their nuances contribute to the pricing dynamics.

Analytical Deeper Dive:
1. Feature Preprocessing: Commencing our analysis, we ensured consistency by transforming percentage variables for a standardized dataset.
2. Why Random Forest? The deliberate choice of Random Forest stems from its ability to navigate non-linear relationships, offering a comprehensive understanding of real estate dynamics.
3. Model Training and Evaluation: The trained Random Forest Regressor exhibited a an R-squared score of 0.72 and a Mean Squared Error (MSE) of 59409, attesting to its proficiency in unraveling nuanced patterns influenced by property prices.

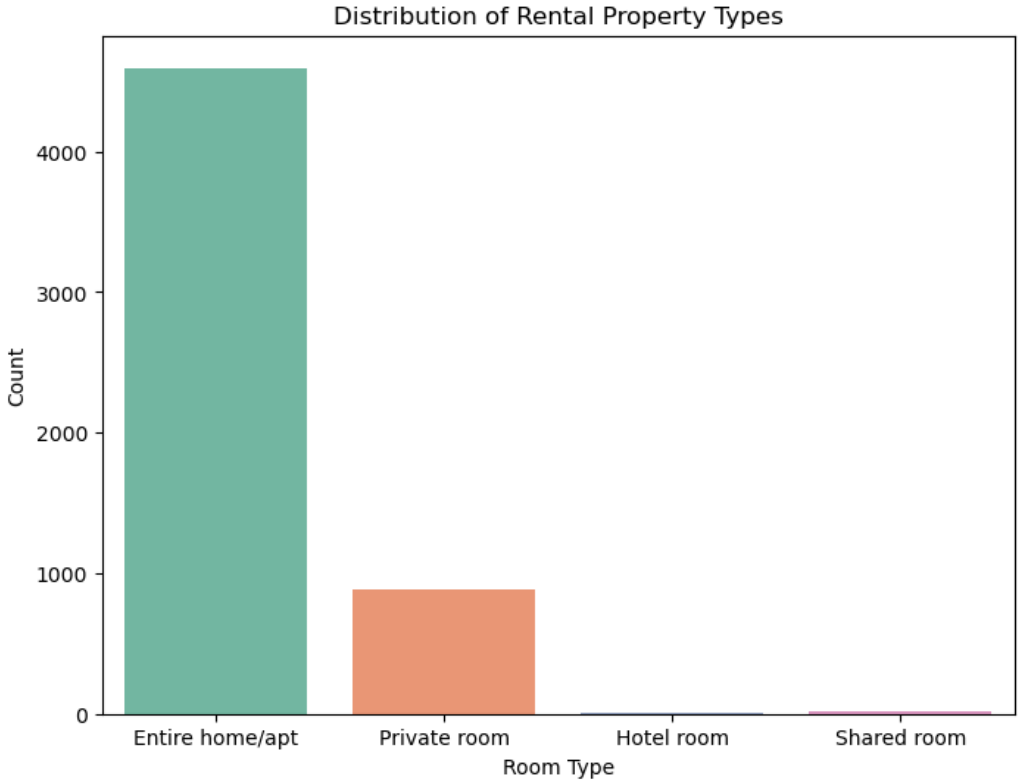
4. In-Depth Exploration of Key Features:

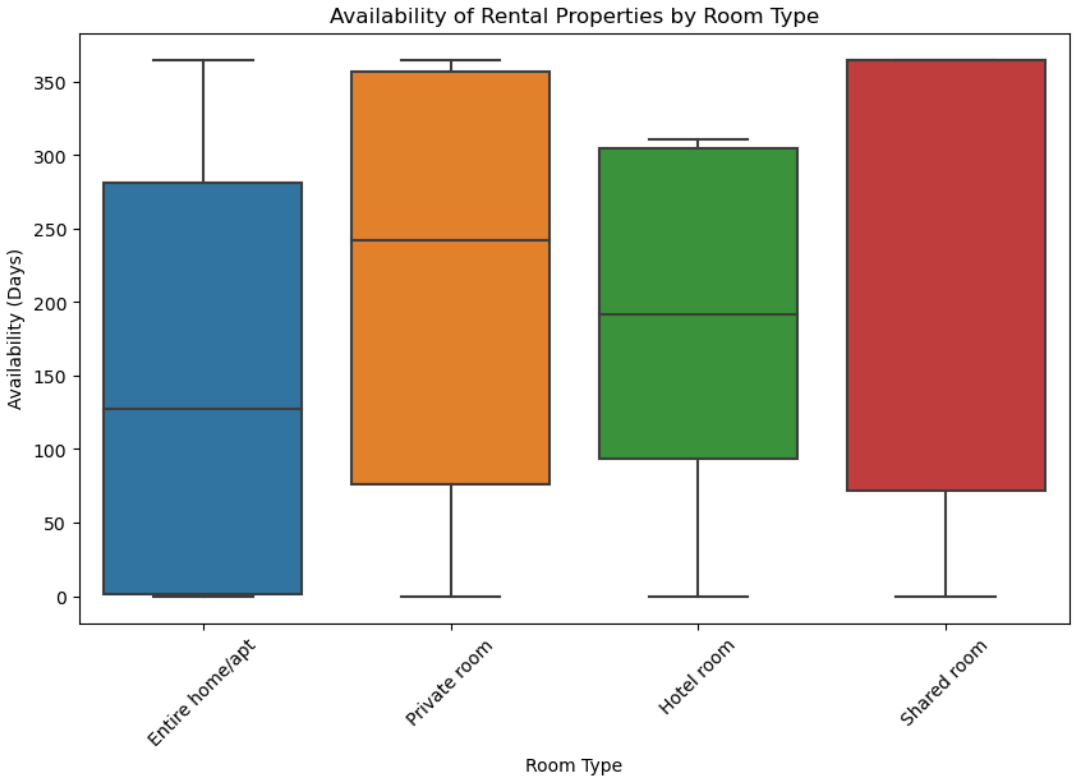
- Availability:
- Impact: Increased availability aligns with more competitive pricing.
-Recommendation: Strategically manage availability for optimal competitiveness.
- Number of Reviews:
- Impact: Positive reviews significantly elevate property valuations.
- Recommendation: Foster positive guest experiences to enhance property value.
- Room Type:
- Impact: Distinct room types distinctly influence pricing dynamics.
- Recommendation: Tailor marketing strategies to highlight unique room attributes.
- Minimum Nights:
- Impact: Stay duration bears a discernible influence on property prices.
- Recommendation: Adapt minimum nights policies dynamically to market trends.
- Diversity Percentage:
- Impact: Socio-economic diversity substantially contributes to pricing dynamics.
- Recommendation: Leverage neighborhood diversity to enhance property value.

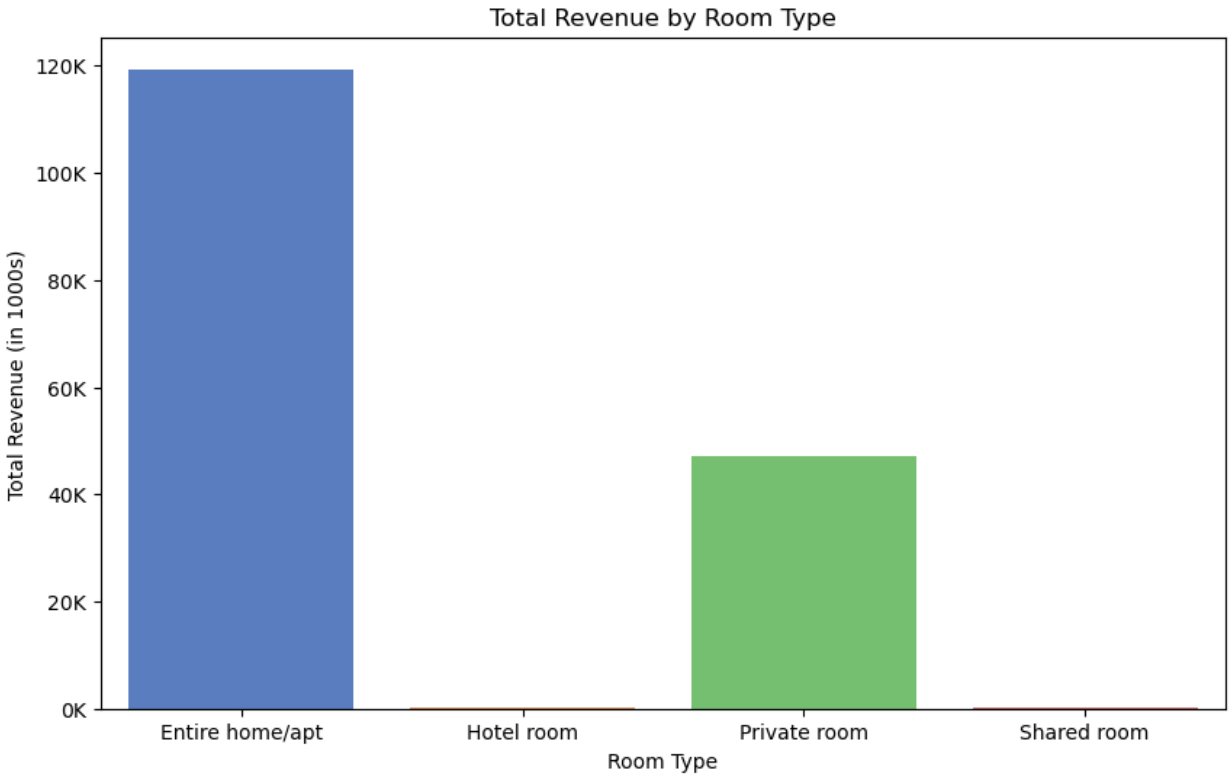
Understanding Feature Interactions:
- When Availability Increases: Competitive pricing tends to follow, highlighting the strategic management of availability.
- As the number of Positive Reviews Rises: Property valuation sees a positive impact, emphasizing the significance of positive guest experiences.
- Differentiating Room Types: This practice directly influences pricing dynamics, emphasizing the need for strategic marketing differentiation.
- Adjusting Minimum Nights: The duration of stay plays a role in pricing, necessitating adaptable policies.
- Socio-economic Diversity Influence: Neighborhood diversity significantly contributes to pricing, offering a unique selling proposition.

In Conclusion: This exploration not only enhances our grasp of property pricing dynamics but also provides some actionable insights for strategic decision-making. The application of Random Forest, with its nuanced understanding of feature impacts, reaffirms its value in navigating the complexities of real estate dynamics for informed decision-making in a competitive market.

Question 3: Should the company invest in single-unit (house) or multi-unit (apartment) rentals to maximize profitability?


Distribution of Rental Property Types
Count
4000
3000
2000
1000
0
Entire home/apt
Private room
Hotel room
Shared room
Room Type


Availability of Rental Properties by Room Type
Availability (Days)
350
300
250
200
150
100
50
0
Entire home/apt
Private room
Hotel room
Shared room
Room Type



1. Distribution of Property Types:

The bar plot illustrates that the majority of room rental types are Entire homes/apartments (83.5%) and Private rooms (16.1%), collectively constituting 99.6% of the total listings.

2. Future Listings Availability Across Room Types:

Examining the availability of future listings across room types, the data shows that Entire homes/apartments have the lowest average availability at approximately 147 days, indicating a significant variability. Despite a notable range, the quartiles for Entire homes/apartments are comparatively lower than other room types. Shared rooms exhibit the highest average availability at 251 days, with a significant portion showing availability for 365 days. Hotel rooms have an average availability of 185 days, slightly less than Private rooms at 206 days.

3. Average Revenue by Room Type:

The third chart presents the visualization of average revenue by room type. Entire homes/apartments have the highest revenue at $119.3 million, followed by Private rooms with $47 million. Conversely, Hotel rooms ($219,000) and Shared rooms ($193,000) significantly trail behind in terms of total revenue.

Recommendations:

1. Consider Investing in Entire Homes/Apartments:

Given their dominance in both listings and revenue, Entire homes/apartments present a potentially high-return investment choice. Stakeholders should explore opportunities within this category, considering its strong performance in terms of revenue generation and market demand.

2. Evaluate Private Rooms as a Secondary Option:

Private rooms also demonstrate considerable revenue potential and could serve as a complementary

investment choice alongside Entire homes/apartments. While not as dominant in revenue as Entire homes/apartments, Private rooms still present a viable opportunity for stakeholders to diversify their portfolios.

3. Caution Regarding Hotel and Shared Rooms:

The comparatively lower revenue figures for Hotel and Shared rooms suggest a less lucrative investment opportunity in these categories. Stakeholders should exercise caution when considering investments in these types of accommodations, as they may not offer the same level of return as Entire homes/apartments or Private rooms.

4. Analyze Availability Trends:

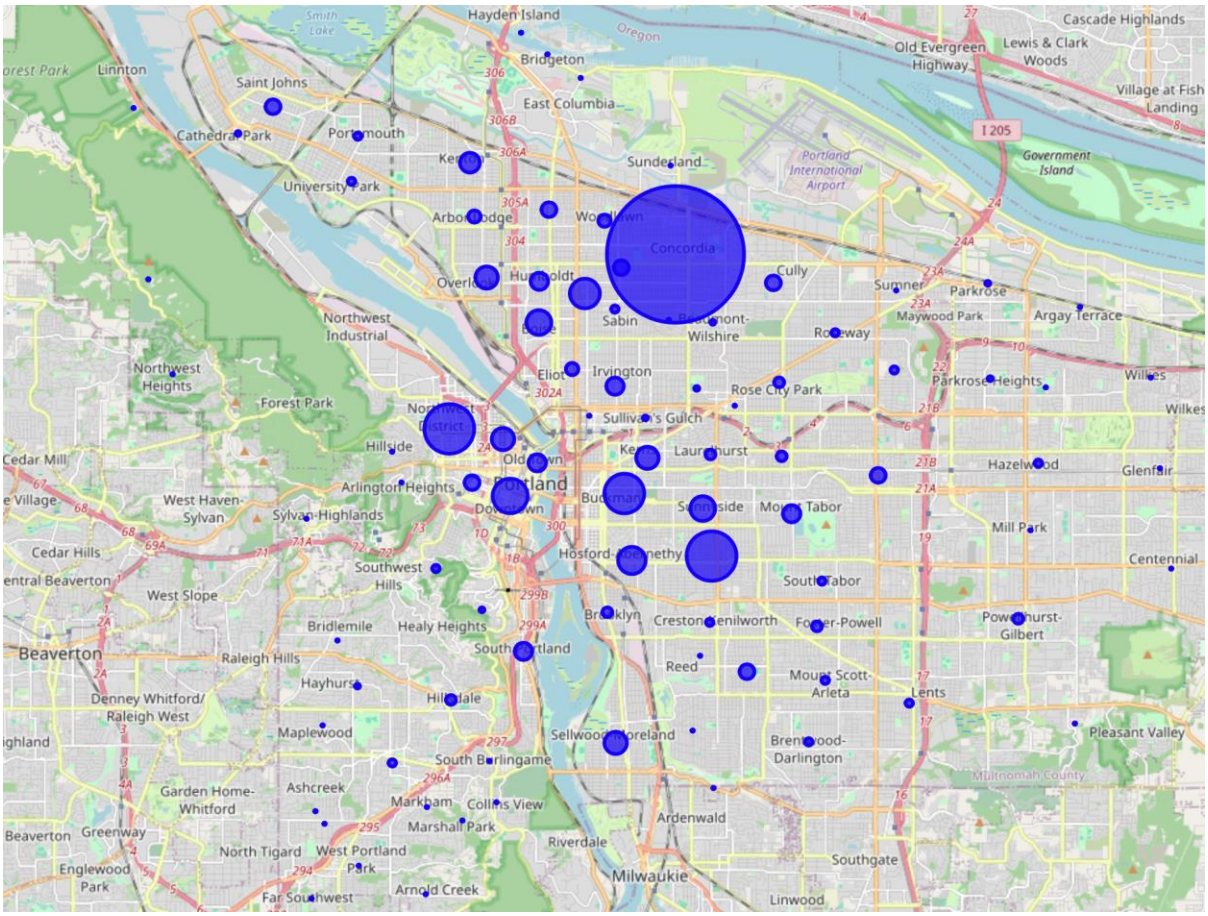
It's essential for stakeholders to delve deeper into the availability trends across different room types. Understanding fluctuations in availability can help inform pricing strategies and investment decisions, ensuring optimal utilization of resources and maximizing revenue potential.

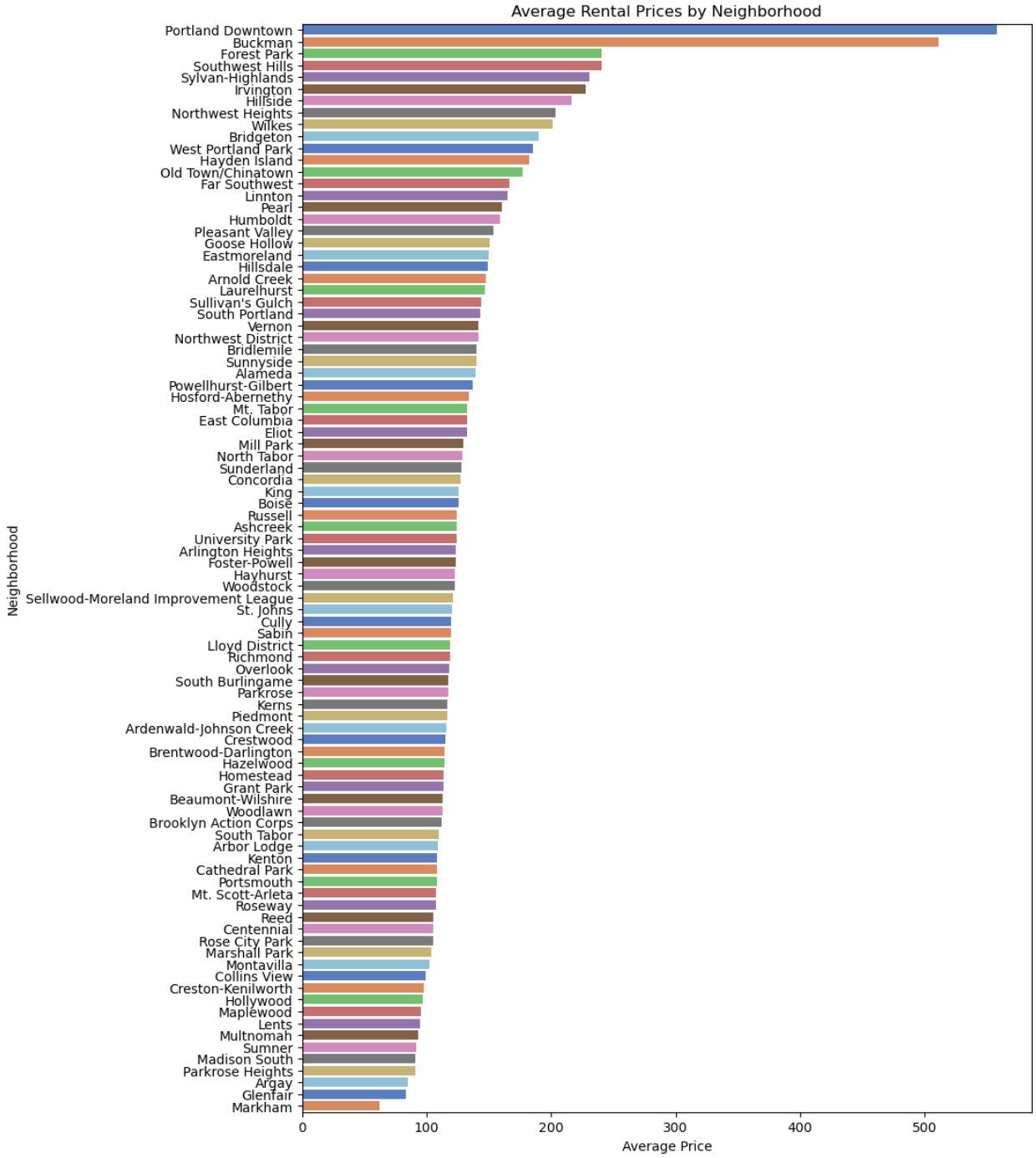
5. Explore Market Dynamics:

Continual monitoring of market dynamics, including demand patterns and competitor strategies, is crucial for informed decision-making. Stakeholders should stay attuned to shifts in consumer preferences and market trends to adapt their investment strategies accordingly.

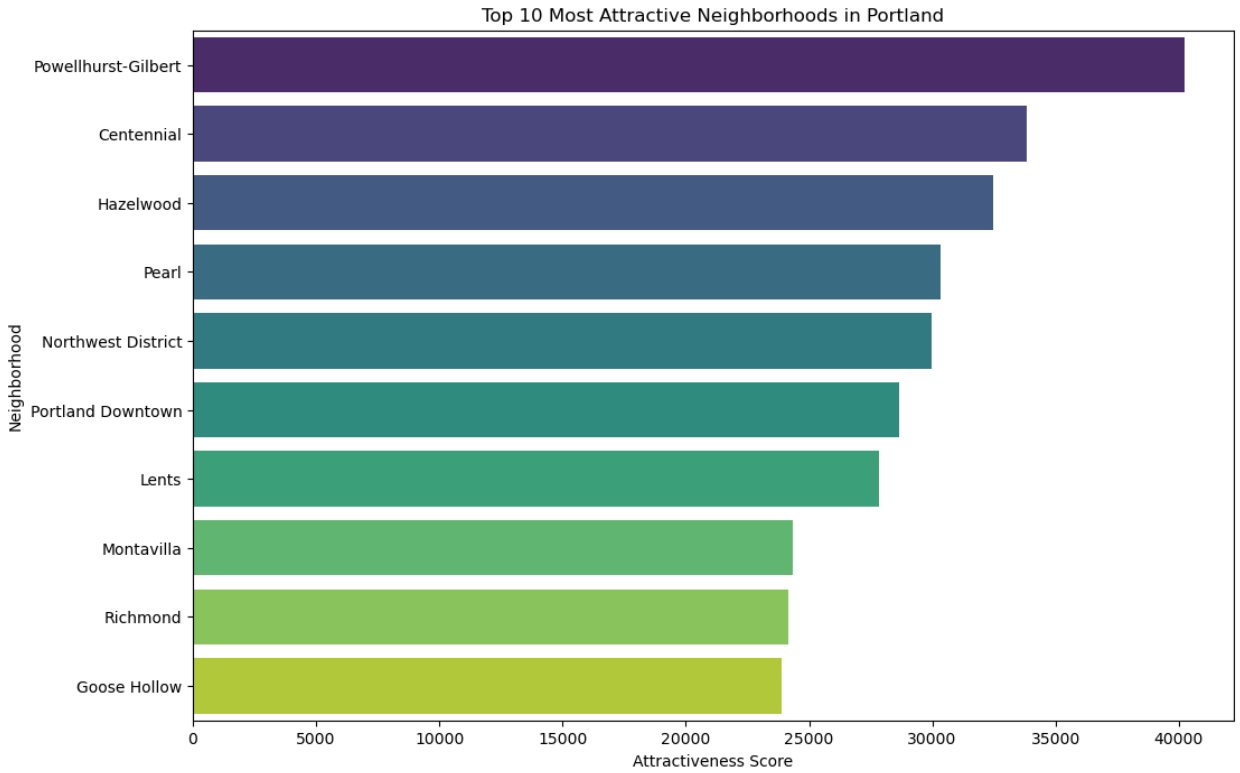
In summary, stakeholders should leverage insights from the distribution of property types, future listing availability, and average revenue by room type to make informed decisions about expanding and selecting specific property types within the rental market. A strategic approach that capitalizes on lucrative opportunities while mitigating risks associated with less profitable options will pave the way for sustainable growth and success in the rental industry.

Question 4: Are there specific locations in Portland with notably higher demand for rentals?


Hayden Island
Bridgeton
East Columbia
Sunderland
Concordia
Cully
Sumner
Parkrose
Maywood Park
Argay Terrace
Roseway
Rose City Park
Irvington
Sullivan's Gulch
Laurelhurst
Hazelwood
Glenfair
Mill Park
Centennial
Powellhurst-Gilbert
Foster-Powell
Mount Tabor
Buckman
Kerns
Old Town
Portland
Downtown
Arlington Heights
Hillside
Northwest District
Northwest Industrial
Northwest Heights
Forest Park
Linnton
Saint Johns
Cathedral Park
Portsmouth
University Park
Arbor Lodge
Overlook
Humboldt
Sabin
Eliot
Beaumont-Wilshire
Cascade Highlands
Old Evergreen Highway
Lewis & Clark Woods
Village at Fish Landing
I 205
Government Island
Portland International Airport
Willes
Wilkes
Parkrose Heights
Hosford-Abernethy
Southeast Hills
Southwest Hills
West Slope
West Haven-Sylvan
Sylvan Highlands
Cedar Hills
Cedar Mill
Bridlemile
Healy Heights
South Portland
Brooklyn
Creston-Kenilworth
Reed
Mount Scott-Arleta
Lents
Foster-Powell
Beaverton
Raleigh Hills
Hayhurst
Hillsdale
Denney Whitford/Raleigh West
Maplewood
Sellwood-Moreland
Brentwood-Darlington
Pleasant Valley
South Burlingame
Garden Home-Whitford
Ashcreek
Markham
Collins View
Marshall Park
Ardenwald
Riverdale
Milwaukie
Southgate
Linwood
North Tigard
West Portland Park
Arnold Creek
Far Southwest
Englewood Park
Greenway
Beaverton
Multnomah County


Average Rental Prices by Neighborhood
Neighborhood
Portland Downtown
Buckman
Forest Park
Southwest Hills
Sylvan-Highlands
Irvington
Hillside
Northwest Heights
Wilkes
Bridgeton
West Portland Park
Hayden Island
Old Town/Chinatown
Far Southwest
Linnton
Pearl
Humboldt
Pleasant Valley
Goose Hollow
Eastmoreland
Hillsdale
Arnold Creek
Laurelhurst
Sullivan's Gulch
South Portland
Vernon
Northwest District
Bridlemile
Sunnyside
Alameda
Powellhurst-Gilbert
Hosford-Abernethy
Mt. Tabor
East Columbia
Eliot
Mill Park
North Tabor
Sunderland
Concordia
King
Boise
Russell
Ashcreek
University Park
Arlington Heights
Foster-Powell
Hayhurst
Woodstock
Sellwood-Moreland Improvement League
St. Johns
Cully
Sabin
Lloyd District
Richmond
Overlook
South Burlingame
Parkrose
Kerns
Piedmont
Ardenwald-Johnson Creek
Crestwood
Brentwood-Darlington
Hazelwood
Homestead
Grant Park
Beaumont-Wilshire
Woodlawn
Brooklyn Action Corps
South Tabor
Arbor Lodge
Kenton
Cathedral Park
Portsmouth
Mt. Scott-Arleta
Roseway
Reed
Centennial
Rose City Park
Marshall Park
Montavilla
Collins View
Creston-Kenilworth
Hollywood
Maplewood
Lents
Multnomah
Sumner
Madison South
Parkrose Heights
Argay
Glenfair
Markham
0
100
200
300
400
500
Average Price


Availability of Rentals by Neighborhood
Neighborhood
Availability (Days)
0
50
100
150
200
250
300
Bridgeton
Lloyd District
Portland Downtown
Old Town/Chinatown
Goose Hollow
Hayden Island
Marshall Park
Bridlemile
Parkrose
Sylvan-Highlands
Pearl
Argay
Arnold Creek
West Portland Park
Hayhurst
Ashcreek
Glenfair
Kenton
Northwest District
Northwest Heights
Powellhurst-Gilbert
South Burlingame
Wilkes
Pleasant Valley
Homestead
Hazelwood
Beaumont-Wilshire
Kerns
Grant Park
Buckman
Brentwood-Darlington
South Portland
Far Southwest
Markham
Laurelhurst
Foster-Powell
Southwest Hills
Roseway
Montavilla
Eastmoreland
Centennial
East Columbia
Madison South
Humboldt
Arbor Lodge
Maplewood
Brooklyn Action Corps
Forest Park
Parkrose Heights
Cully
Woodstock
Portsmouth
Sellwood-Moreland Improvement League
Crestwood
Sunnyside
Hosford-Abernethy
Reed
Vernon
Overlook
Boise
Multnomah
Concordia
Russell
University Park
Mill Park
Piedmont
Hillsdale
Sullivan's Gulch
Sabin
Richmond
Rose City Park
Alameda
North Tabor
Collins View
St. Johns
Hollywood
Hillside
South Tabor
Mt. Scott-Arleta
Cathedral Park
Creston-Kenilworth
Irvington
Woodlawn
Eliot
King
Mt. Tabor
Lents
Sumner
Linnton
Arlington Heights
Ardenwald-Johnson Creek
Sunderland



1. Distribution of Rental Listings:

Concordia, with an impressive 731 listings, stands out as the neighborhood saturating the market, surpassing the average neighborhood count by a remarkable 1118%. Northwest District and Richmond follow closely with 269 and 265 listings, respectively, well above the average count. This saturation suggests high rental activity and potential demand in these areas.

2. Pricing Dynamics Across Neighborhoods:

Portland Downtown emerges as the costliest neighborhood with an average price per night of $566.74, starkly contrasting with the overall average of $118 across all neighborhoods. Understanding the pricing dynamics per neighborhood is crucial for investors and renters alike to make informed decisions based on their budgets and preferences.

3. Availability Assessment:

Examining the availability of rentals across neighborhoods reveals interesting insights. Bridgeton, Lloyd District, and Portland Downtown exhibit high availability, indicating ample future days for new bookings. Conversely, Concordia records the lowest availability at only 10%, highlighting potential challenges for prospective renters in securing accommodations in this high-demand neighborhood.

4. Attractiveness Score Calculation:

To identify the top ten neighborhoods, we considered a range of socio-economic, demographic, and environmental factors, these include:

- Homeownership Percentage
- Diversity Percentage
- Poverty Percentage
- Trees Percentage
- College Percentage
- SVT Index (an index reflecting certain environmental factors)

- Population Density
- Total Residential Population
- Percentage of New Residents
- Voting Percentage

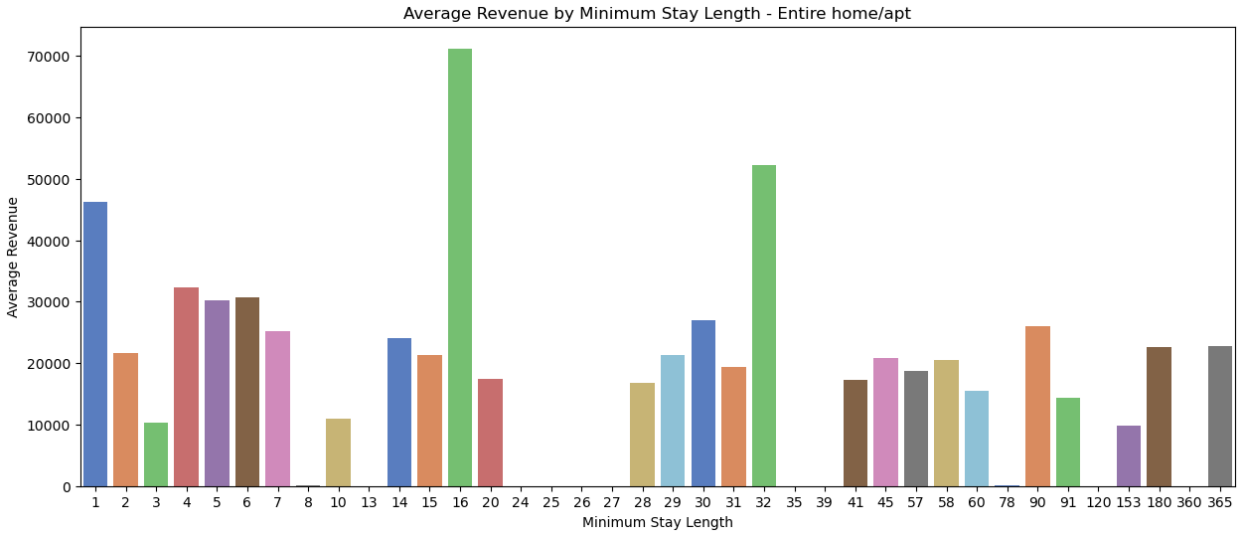
By aggregating and analyzing these factors, we computed an "attractiveness score" for each neighborhood, allowing us to rank them based on their overall desirability and investment potential.

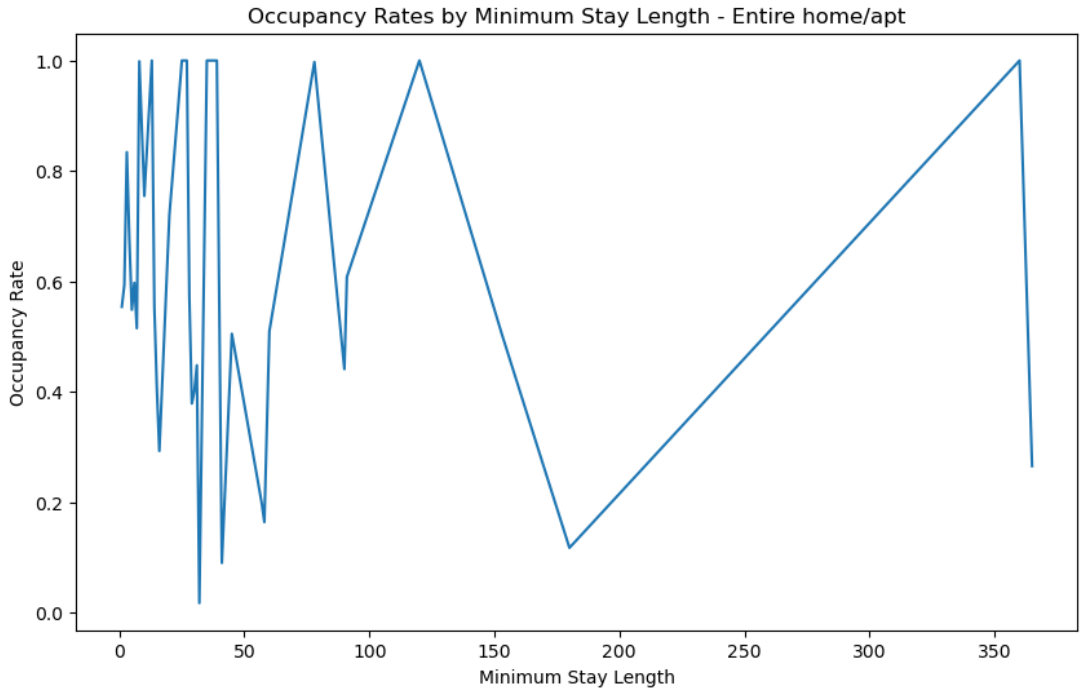
5. Top 10 Most Attractive Neighborhoods:

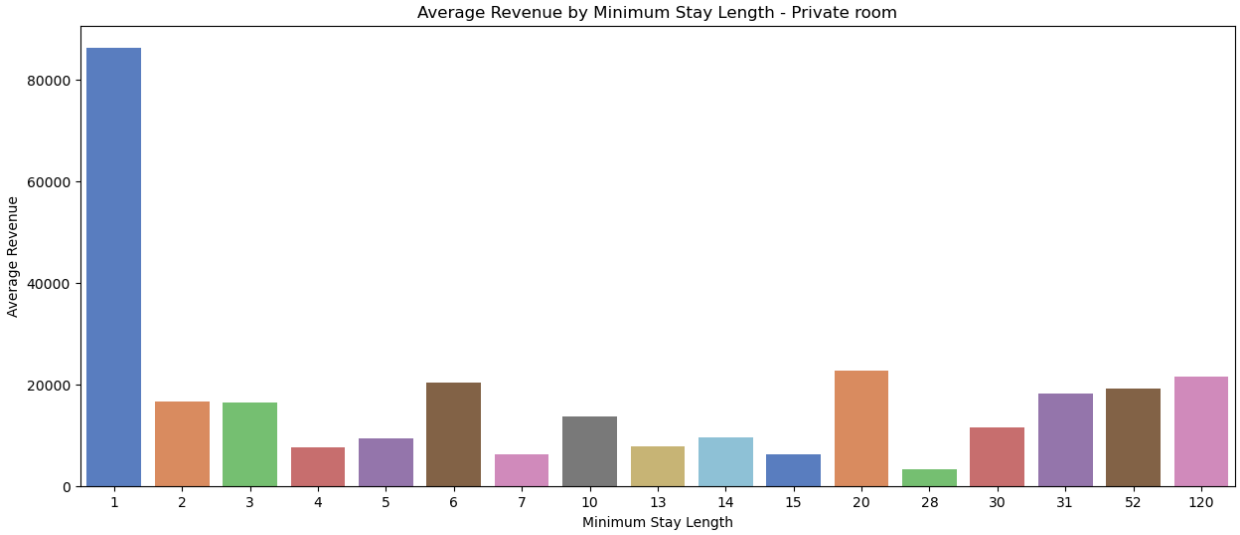
The top 10 neighborhoods, including Powellhurst Gilbert, Centennial, Hazelwood, Parkrose, Northwest District, Portland Downtown, Lents, Montavilla, Richmond, and Goose Hollow, exhibit diverse characteristics contributing to their overall attractiveness. These neighborhoods present promising opportunities for investors and renters alike, offering a blend of amenities, community features, and economic vitality.

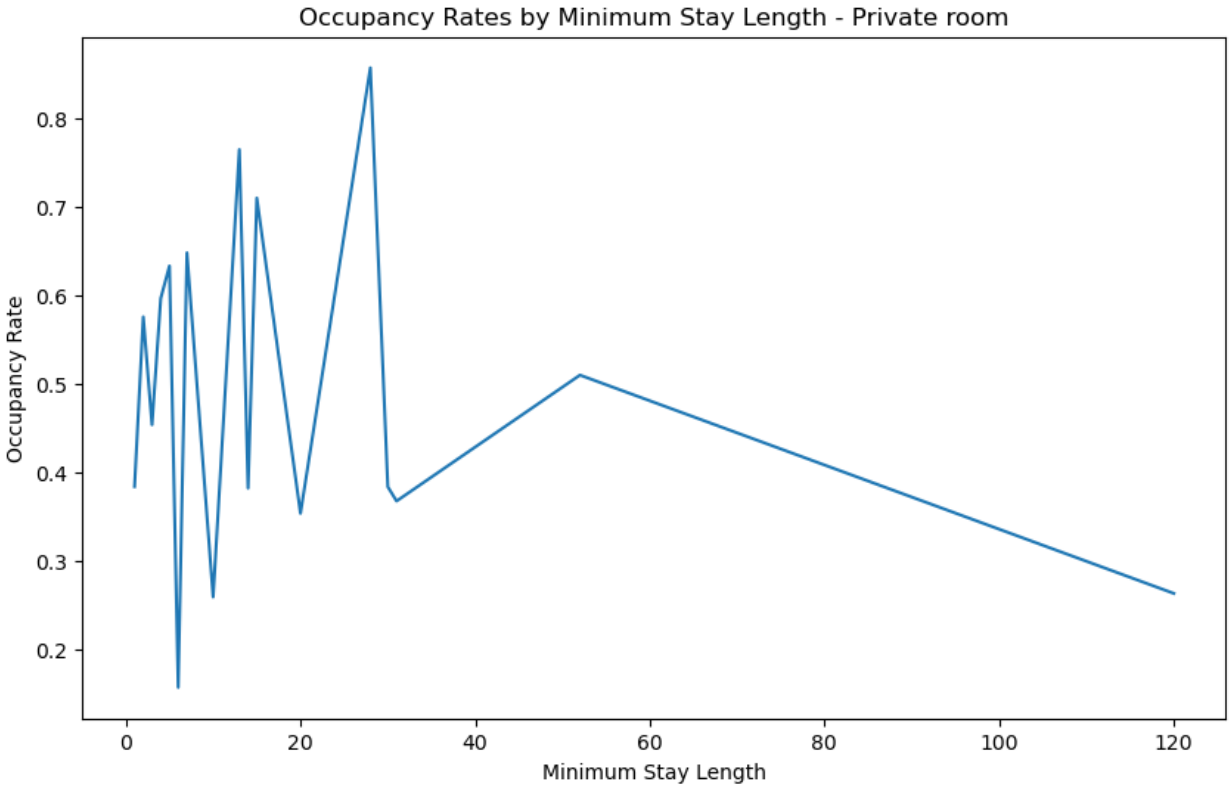
In summary, our analysis considers a holistic view of Portland's rental landscape, encompassing rental distribution, pricing dynamics, availability trends, and neighborhood attractiveness. By understanding these factors, stakeholders can make informed decisions to navigate the dynamic rental market effectively and capitalize on emerging opportunities for growth and investment.

Question 5: What is the ideal minimum stay length for maximizing occupancy rates and revenue?




Occupancy Rates by Minimum Stay Length - Entire home/apt
Occupancy Rate
Minimum Stay Length
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150
200
250
300
350


Average Revenue by Minimum Stay Length - Private room
Average Revenue
Minimum Stay Length
0
20000
40000
60000
80000
1
2
3
4
5
6
7
10
13
14
15
20
28
30
31
52
120



1.Data Filtering and Aggregation:

We filtered the dataset to specific room types: "Entire home/apt", "Private room", "Hotel room", and "Shared room". Aggregating the data by "minimum_nights", we calculated the occupancy rate using the "availability_365" for each group. This process helped derive the average revenue for distinct minimum stay lengths within each room type.

Revenue Analysis by Room Type and Minimum Stay Length:

- **Entire Home/Apt:**

The highest revenue, $71,208, is associated with a minimum stay of 16 days, yielding an occupancy rate of 29.3% (Figure 1).
The second-highest revenue, $52,279, is generated from a minimum stay of 32 days, despite exhibiting the lowest occupancy rate within this room type at 1.8% (Figure 2). This highlights a potential area for revenue growth by enhancing occupancy.

- **Private Room:**

The peak revenue of $86,400 comes from a minimum stay of 1 day, albeit registering a 10% below-average occupancy rate for this room type (Figure 4).
Hotel Room:

All listings require a minimum stay of 1 day, with an occupancy rate of 49.2%. This rate is notably 8%

higher than the average occupancy rates for other room types with a similar minimum stay requirement.

- **Private Room:**

¨Private rooms also require a one-night minimum stay, with an occupancy rate lagging behind the average for all room types with the same minimum stay, standing at 27.6% below average. Furthermore, their revenue is 68.4% lower compared to the overall average. Identifying strategies to uplift occupancy and revenue in this room type presents another opportunity for investment.

In summary, analyzing revenue trends across different room types and minimum stay lengths provides valuable insights into occupancy rates and revenue potential. Stakeholders can leverage this information to optimize pricing strategies, enhance occupancy rates, and identify investment opportunities tailored to specific room types and minimum stay requirements.

Question 6: To what extent do customer reviews impact rental pricing, and how can the company leverage positive reviews for pricing strategies?

	Feature	Coefficient
0	number_of_reviews	-0.192708
1	reviews_per_month	0.223063

1. Correlation Analysis:

The correlation matrix reveals the relationships between pricing and the number of reviews. The correlation coefficient (-0.091715) suggests a weak negative correlation between pricing and the number of reviews. This implies that as the number of reviews increases, there is a slight tendency for prices to decrease.

2. Regression Analysis:

The regression model coefficient for the number of reviews (-0.192708) further illustrates this impact, it indicates that an increase in the number of reviews is associated with a decrease in pricing, albeit the effect is relatively small.

3. Interpretation and Insights:

Customer reviews, as indicated by the number of reviews, seem to influence rental pricing, albeit moderately. A higher number of reviews may imply a more competitive market or higher turnover of guests, which could potentially lead to more competitive pricing.

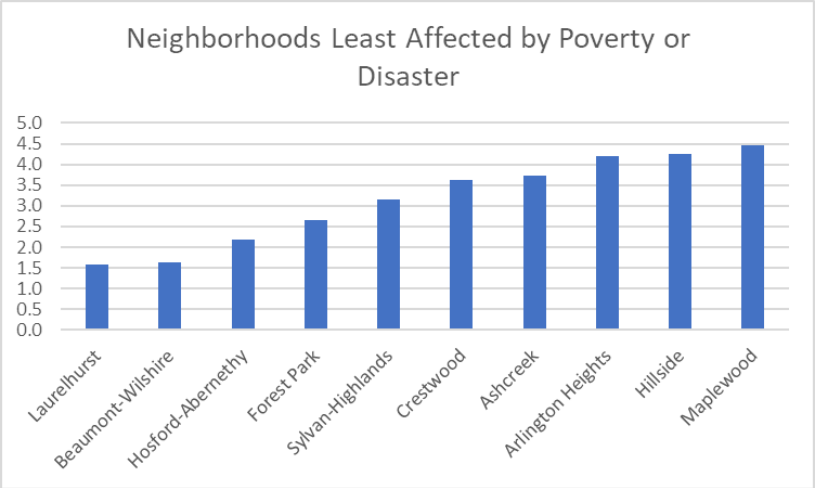
It's important to realize that while customer reviews are essential for reputation management and customer satisfaction, their direct impact on pricing may be influenced by various factors, including market dynamics, competition, and property-specific attributes.

4. Leveraging Insights for Pricing Strategies:

- **Competitive Pricing Analysis:** Monitor pricing strategies of competitors in relation to their review volumes. Understanding how competitors adjust prices based on review volumes can provide insights into market positioning and pricing competitiveness.
- **Dynamic Pricing Adjustments:** Implement dynamic pricing algorithms that take into account review volumes as one of the factors for price adjustments. Higher review volumes could potentially trigger adjustments to maintain competitiveness and optimize revenue.
- **Customer Satisfaction Focus:** While pricing strategies are influenced by review volumes, maintaining a focus on customer satisfaction and experience remains paramount. Positive guest experiences contribute not only to higher review volumes but also to long-term customer loyalty and positive brand perception.

In summary, while the number of reviews does impact rental pricing to some extent, leveraging this insight effectively requires a comprehensive understanding of market dynamics and customer preferences. By integrating review volumes into pricing strategies and maintaining a customer-centric approach, the company can optimize pricing decisions while enhancing guest satisfaction and competitiveness in the rental market.

Do economic and demographic factors of a neighborhood influence rental preferences, and how should the company consider these factors in decision-making?



We examined the key performance indicators (diversity, percentage of voters, etc.) of each neighborhood and determined their correlations to each other by using the CORREL function in Excel. The highest correlations lead to the following observations:

- Neighborhoods with higher poverty are more vulnerable to disaster
- Neighborhoods with a higher percentage of college graduates are less prone to disaster

These observations narrowed down which factors would be most relevant to our investment decision. We have concluded that the most important factors in determining which neighborhood to invest in should be its vulnerability to disaster (natural or manmade) and the percentage of households living below the Federal poverty line. We used a weighted average to rank each neighborhood on their combined exposure to these two factors and have illustrated the 10 neighborhoods least affected in exhibit ___.

Exhibit __: Neighborhoods Least Affected by Poverty or Disaster