EMPIRE REALTY

Empire Realty Team:

- Gavin Grosswald
- Carolina Sepulveda
- Hannah Redmond
- Marium Hamdani

Introduction:

Welcome to Empire Realty ("Empire"), the premier real estate firm situated in the heart of New York City ("NYC"). With a steadfast commitment to excellence and innovation, Empire is positioned for dynamic growth as it continues to execute its strategic approach to property acquisition in the rental market.

The NYC housing and rental market, characterized by its diversity across five distinct boroughs — Manhattan, Brooklyn, Queens, the Bronx, and Staten Island — provides a rich opportunity for investment. Over recent years, this market has witnessed fluctuations driven by economic trends, population dynamics, and evolving consumer preferences. As a global economic hub and cultural epicenter, NYC remains a beacon for individuals and families in pursuit of diverse opportunities and experiences.

In parallel, the growth of platforms like Airbnb has revolutionized the travel industry, offering travelers a multitude of accommodation choices beyond traditional hotels. This transformative shift not only enhances convenience and flexibility for travelers, but also propels economic growth and cultural exchange within the tourism sector.

Equipped with cutting-edge data analytics and machine learning techniques, Empire is strategically positioned to navigate the nuances of market trends and identify lucrative investment opportunities. By leveraging data-driven insights, Empire Realty remains dedicated to guiding its strategic agenda through the complexities of the NYC real estate landscape, ensuring transparency, and maximizing returns for its stakeholders.

Welcome to Empire Realty — where innovation meets opportunity in NYC real estate.

About the Data:

Airbnb is an online marketplace connecting hosts who offer short-term lodging accommodations with travelers seeking unique stays. Founded in 2008, it allows users to book accommodations ranging from apartments to villas based on location, price, and amenities. Airbnb's platform facilitates cultural exchange and community connections, transforming travel experiences globally.

Empire Realty has access to a comprehensive dataset which offers a view into the Airbnb landscape across New York City's five boroughs as of January 5, 2024. The dataset encompasses information crucial for evaluating and understanding the dynamic real estate market in NYC.

The dataset includes 22 variables and 20.5k observations for a total of just over 540k datapoints. The variables included in the dataset are as follows:

- ID
- Listing Name
- Host ID
- Host Name
- Borough
- Neighborhood
- Latitude
- Longitude
- Room Type
- Price
- Minimum Nights

- Number of Reviews
- Date of Last Review
- Number of Reviews per Month
- Number of Reviews (LTM)
- Rating
- Listing Count per Host
- Days Available (365)
- License
- Bedrooms
- Beds
- Baths

Data Mining Problem:

Empire Realty is focused on addressing the challenges within the NYC rental market. The firm aims to leverage Airbnb data and advanced analytics to gain a comprehensive understanding of the real estate market dynamics and to identify lucrative investment opportunities, optimize portfolio acquisition, and maximize returns for stakeholders.

Through strategic portfolio acquisition, the firm intends to select properties offering optimal returns while minimizing risks associated with market volatility and regulatory changes. Ultimately, Empire Realty's goal is to maximize returns for stakeholders by making informed investment decisions, enhancing portfolio performance, and ensuring transparency in its operations within the competitive NYC real estate landscape.

NYC Market Overview:

	Manhattan	Brooklyn	Queens	Bronx	Staten Island
Population	19,034,000	2,679,620	2,177,805	1,297,660	492,925
Avg. Income	$79,205	$53,626	$56,592	$56,723	$54,960
Avg. Home Price	$1,355,955	$949,782	$564,270	$564,270	$670,000

Data Analysis:

Associative Rule Mining

```
      lhs                                                                    rhs             support confidence coverage lif
t count
[1] {name=Rental unit in New York · ★4.60 · 1 bedroom · 1 bed · 1 bath}   => {rating=4.60 }  0.0013          1    0.0013   9
2    26
[2] {name=Rental unit in New York · ★4.60 · 1 bedroom · 1 bed · 1 bath,
     neighbourhood_group=Manhattan}                                       => {rating=4.60 }  0.0013          1    0.0013   9
2    26
[3] {name=Rental unit in New York · ★4.60 · 1 bedroom · 1 bed · 1 bath,
     room_type=Entire home/apt}                                           => {rating=4.60 }  0.0012          1    0.0012   9
2    24
[4] {name=Rental unit in New York · ★4.60 · 1 bedroom · 1 bed · 1 bath,
     beds=[1,2)}                                                          => {rating=4.60 }  0.0013          1    0.0013   9
2    26
[5] {name=Rental unit in New York · ★4.60 · 1 bedroom · 1 bed · 1 bath,
     bedrooms=1}                                                          => {rating=4.60 }  0.0013          1    0.0013   9
2    26
>
```
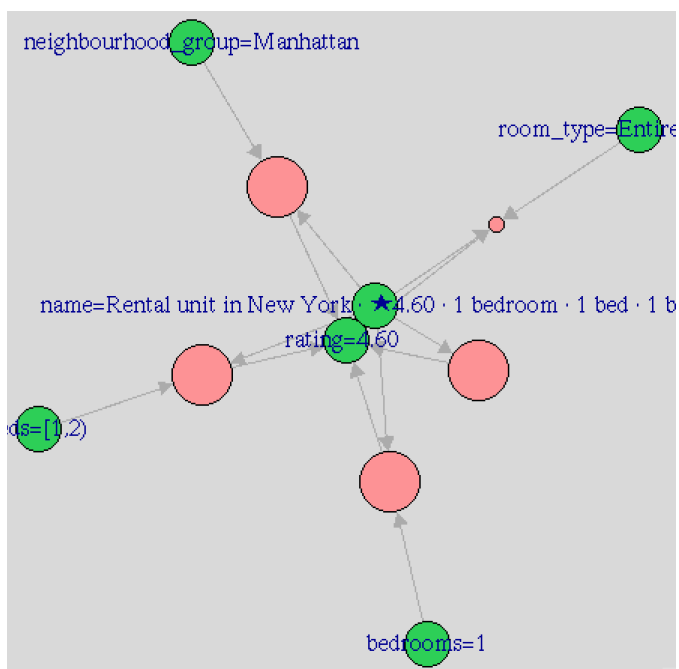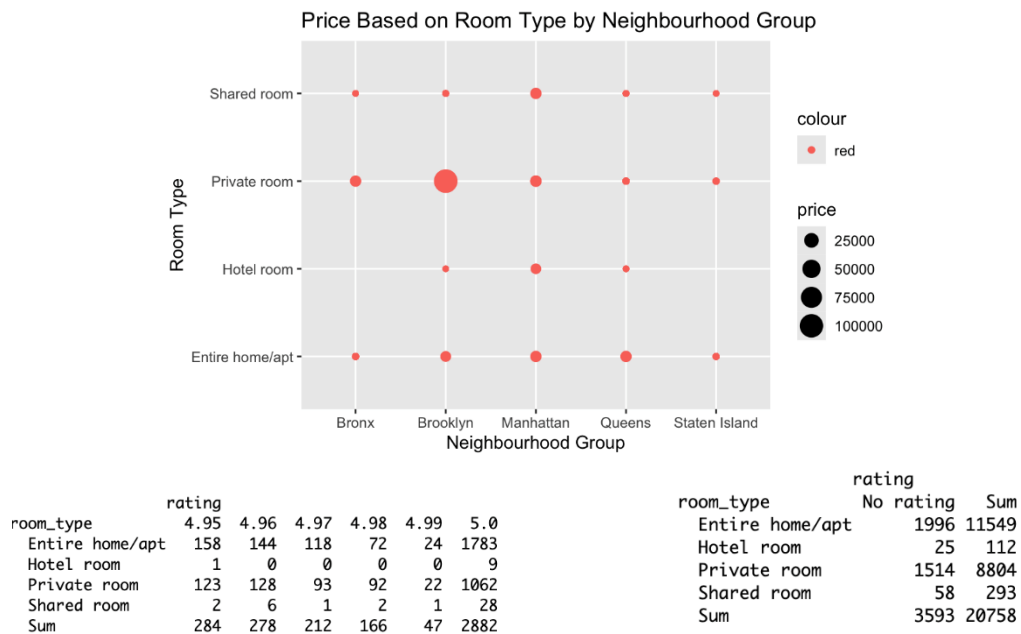

Associative rule mining emerged as an effective tool for Empire Realty, enabling the identification of frequent patterns and relationships within the dataset crucial for making informed investment decisions. By evaluating the level of support, confidence, and lift, Empire Realty gained valuable insights into the dataset's underlying dynamics. Support, indicating the frequency of occurrence of an itemset, allowed for the identification of commonly recurring patterns, while confidence measured the strength of the relationship between items in the rule. Lift provided a measure of how much more often items appeared together in transactions, offering a deeper understanding of the dataset's associations.

In the assessment of properties, Empire Realty considered several key characteristics, including property reviews, minimum nights stayed, and number of beds. Through the analysis of lift, support, and confidence metrics, consistent conclusions emerged, pointing towards a focused strategy. Specifically, the findings highlighted the importance of targeting rental units in New York,

particularly in Manhattan, with ratings exceeding 4.60 and featuring at least 1 bed and 1 bathroom. Notably, the utilization of a support level of 0.001, coupled with a confidence of 1 and lift of 9, confirmed these conclusions, providing support for strategic decision-making.
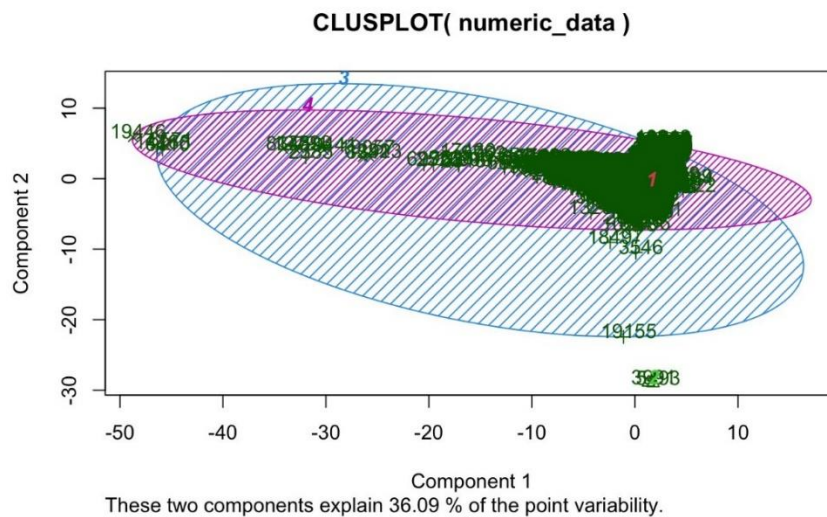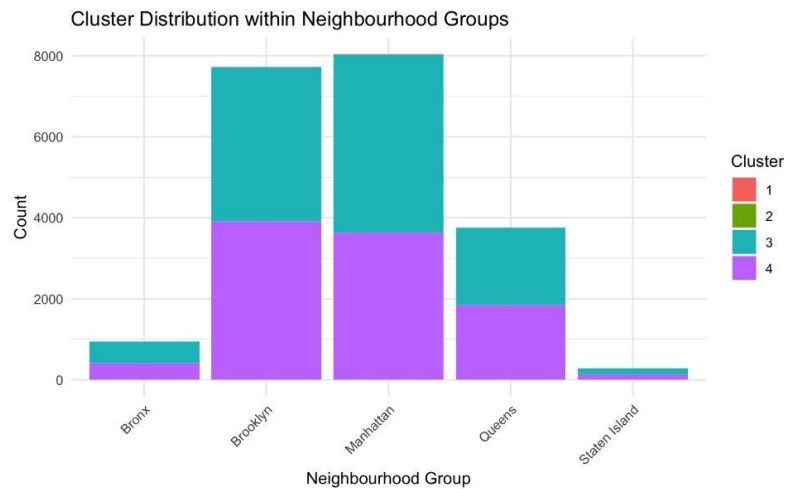
The associative rule mining analysis not only reaffirmed but also aligned with the overarching investment strategy. It became evident that directing investments towards Manhattan, the more populous borough with a higher volume of listings, was a logical choice.

Price Based on Room Type by Neighbourhood Group

room_type	rating					
	4.95	4.96	4.97	4.98	4.99	5.0
Entire home/apt	158	144	118	72	24	1783
Hotel room	1	0	0	0	0	9
Private room	123	128	93	92	22	1062
Shared room	2	6	1	2	1	28
Sum	284	278	212	166	47	2882

room_type	rating	
	No rating	Sum
Entire home/apt	1996	11549
Hotel room	25	112
Private room	1514	8804
Shared room	58	293
Sum	3593	20758

A bubble graph was crafted to illustrate the comparative costs of room types across various neighborhood groups. This insightful visualization equips Empire Realty with the means to devise a targeted portfolio strategy. By leveraging the data presented in the graph, Empire Realty can curate a diverse selection of properties, tailoring room types to specific neighborhood demographics.

The introduction of Airbnb Rooms highlights the growing trend towards private room accommodations. This knowledge facilitates seamless interactions for both guests seeking private room listings and hosts aiming to provide enhanced guest experiences. The emergence of Airbnb rooms signals a shift in consumer preferences towards personalized and comfortable lodging options. As more individuals gravitate towards private room experiences when booking accommodations, Empire Realty stands poised to capitalize on this trend, adapting its portfolio offerings to align with evolving consumer demands.

Cluster Analysis

Cluster Distribution within Neighbourhood Groups

CLUSPLOT(numeric_data)

These two components explain 36.09 % of the point variability.

The cluster analysis conducted on this dataset encountered significant challenges in extracting meaningful insights about the real estate market. With the dataset encompassing diverse property characteristics, clustering such heterogeneous data into coherent groups proved to be a difficult task. The analysis focused on utilizing neighborhood and price variables, with initial observations suggesting that Brooklyn and Manhattan had the highest property prices.

Similarly, the Hierarchical Agglomerative Clustering (HAC) algorithm provided unsatisfactory and inconclusive results. This can be attributed to the dataset's complexity, comprising numerous features ranging from property attributes to pricing information. The HAC algorithm struggled to effectively handle high-dimensional data, leading to poor outcomes. Attempts to visualize the results were hindered by the extensive dataset size, even when utilizing only 1% of the data. This posed challenges in interpreting the graphical representations of the clusters.

Despite testing of different distance metrics, no significant improvement in the results was observed. The intricacies of the dataset and the limitations of the clustering algorithm showed the need for alternative approaches to uncover meaningful patterns within the real estate market data.

HAC Analysis of Price and Neighbourhood Groups
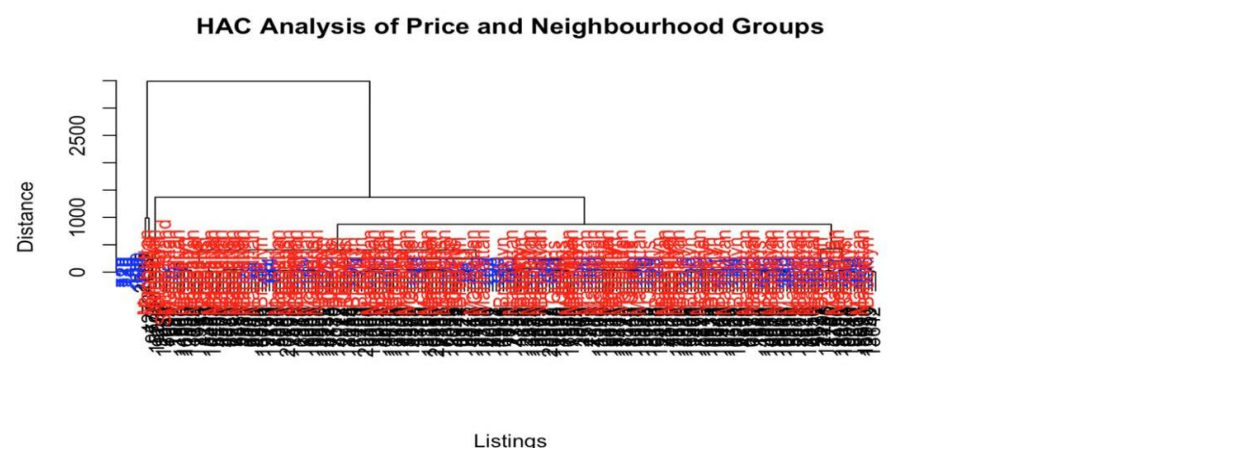
Decision Tree

The decision tree model was trained using 10% of the original dataset due to its large volume. The objective of this model was to predict rental property prices based on two key features: the number of reviews per month and the neighborhood group.

The initial split in the Decision Tree model delineated between neighborhood groups and the number of reviews. This indicates that neighborhood location plays a significant role in determining rental property prices, regardless of review frequency. Interestingly, the model revealed that properties with higher numbers of reviews tended to have lower price value, suggesting a trend where more affordable properties attract greater popularity among renters.
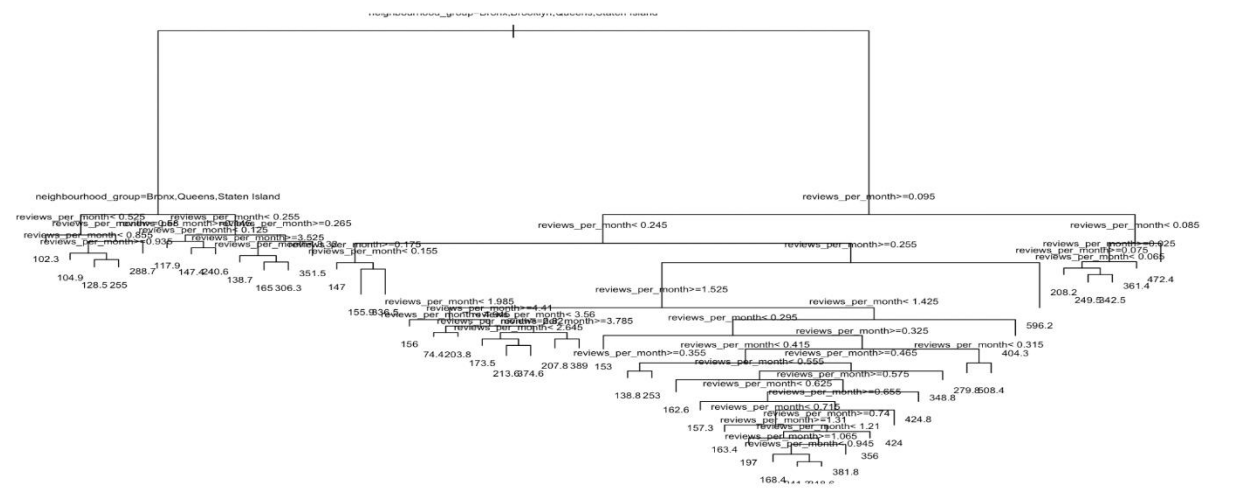
Recommendations:

Based on the comprehensive analysis of the NYC rental market, Empire Realty is strategically positioned to capitalize on investment opportunities in Manhattan and Brooklyn. These boroughs stand out as the most promising areas for maximizing returns and portfolio growth due to several key factors.

The findings from associative rule mining show the significance of targeting rental units in Manhattan, particularly those with ratings exceeding 4.60 and featuring at least 1 bed and 1 bathroom. Similarly, insights from the decision tree model emphasized the importance of neighborhood location in determining rental property prices. The model revealed that properties with higher numbers of reviews tended to have lower price values, indicating a trend where more affordable properties attract greater popularity among renters.

Additionally, both Manhattan and Brooklyn boast significant volumes of properties and high population densities, indicating robust demand for rental accommodations. With a population of 19M and 2.6M respectively, these boroughs offer ample opportunities for Empire Realty to cater to diverse housing needs. The ease of transportation in Manhattan and Brooklyn, facilitated by extensive subway networks and bus routes, also enhances the desirability of rental properties in these areas, making them attractive options for tenants seeking convenient commuting options.

Similarly, the higher average income and home prices in Manhattan and Brooklyn compared to other boroughs signal a greater willingness and ability to pay premium rental rates. This presents an opportunity for Empire Realty to acquire properties with favorable rental yield potential, thus maximizing returns on investment.

Finally, Manhattan's status as a cultural epicenter and global business hub, along with Brooklyn's creative scene, make both boroughs prime locations for rental investments.

While Staten Island presents potential opportunities, its comparatively lower population density, income levels, and cultural amenities may pose challenges for achieving optimal returns. Therefore, Empire Realty should exercise caution when considering investments in Staten Island, conducting thorough market research and risk assessments to mitigate potential downsides.

Future Considerations:

To ensure comprehensive analysis and foster future education and growth, it is imperative to confront the challenges encountered in this report head-on.

One significant hurdle arose from the size of the Airbnb dataset utilized, comprising 22 variables and over 20.5k observations, totaling just over 540k data points. The sheer magnitude of the dataset often led to timeouts in R during analyses such as Random Forest, SVM, and K-means. To mitigate and alleviate the pressure put on R, a random subset of the data was utilized (as little as 2.0%). This decision resulted in a loss of valuable insights and potentially skewed results that may not accurately reflect the true dynamics of the NYC rental market. It diminished the integrity of the data.

Furthermore, it is essential to acknowledge the limitations of using Airbnb data, which predominantly comprises short-term rentals, to predict long-term lease dynamics. The fundamental differences in rental durations, tenant demographics, and regulatory frameworks between short-term and long-term rentals pose significant challenges in extrapolating insights accurately. Therefore, future analyses should incorporate a more diverse and representative dataset, including data from traditional long-term rental platforms like Zillow or StreetEasy, to provide a comprehensive understanding of the NYC rental market landscape.

In addition, conducting in-depth research to understand the unique characteristics and drivers of both short-term and long-term rental markets within each borough of NYC would be beneficial. This approach can help identify distinct trends, preferences, and investment opportunities specific to each segment, enabling Empire Realty to tailor its strategies and recommendations accordingly.

Overall, addressing the challenges related to dataset size and the limitations of Airbnb data while adopting a more tailored approach to market analysis will be pivotal in ensuring Empire Realty remains at the forefront of the NYC real estate landscape, delivering accurate insights and maximizing value for its stakeholders.

Conclusion:

Empire Realty is a key player in the New York City real estate market, strategically positioned to capitalize on promising investment opportunities. With a keen eye for market trends and a commitment to leveraging data-driven insights, Empire Realty identifies Manhattan and Brooklyn as prime areas for portfolio growth.

Manhattan, with its populous streets and global allure, offers a lucrative market for rental accommodations. Its dense population, thriving business environment, and cultural richness make it a magnet for residents and tourists alike. Empire Realty recognizes the potential to meet the demand for diverse housing needs in Manhattan, capitalizing on its vibrant atmosphere and economic strength.

Similarly, Brooklyn presents a compelling case for investment, characterized by its diverse neighborhoods and creative scene. With a growing population and robust transportation infrastructure, Brooklyn offers a range of rental opportunities tailored to various lifestyles. Empire Realty aims to tap into Brooklyn's potential, aligning its portfolio with the borough's unique charm and spirit.

While Manhattan and Brooklyn serve as primary focuses, Empire Realty remains attentive to opportunities across all five boroughs. Queens, the Bronx, and Staten Island each offer distinct investment opportunities. By adopting a tailored approach and conducting thorough market research, Empire Realty aims to navigate the complexities of the NYC rental market effectively.

Empire Realty's strategic vision and commitment to excellence position it for success in the competitive NYC real estate landscape. By prioritizing investments in Manhattan and Brooklyn while remaining adaptable to market shifts, Empire Realty aims to maximize returns and deliver value to its stakeholders in the ever-evolving industry of New York City real estate.