Assessing Urban Livability: A Multifactor Analysis of Rent, Noise and Crime in New York City

STATS SQUAD: Abhay Shah, David Skorodinsky, Javier Miguelez, Sarah Doctor, Tian Lan

Introduction

With roughly 69% of households renting in NYC (Siegel, J., & Bram, J., 2024) it's clear that rental units are in demand. However, with so many different neighborhoods to choose from across NYC's 5 boroughs, it can be difficult to gain a deeper understanding of what quality of life challenges an area may be facing. In particular, crime & noise levels could impact a tenant's enjoyment of their neighborhood but neighborhood-specific data is not always easily available. Due to this, we have conducted a multi-factor analysis on rent, noise & crime across NYC's 5 boroughs to benefit both the tenants who are looking to live in the city as well as city officials who can use our analysis to influence policy decisions for their communities.

Literature Review

Extensive research in urban studies and economics provides evidence that high levels of crime & noise may be interpreted as signals of overall neighborhood decline.

Studies have shown that increases in these variables have a negative correlation with property/rent values. For instance, Ihlanfeldt and Mayock (2010), found that robbery and assault crimes significantly impacted housing values, while Azad and Ghandehari (2021) found that excessive noise levels negatively impacted residential satisfaction during the pandemic in NYC.

With the existing literature supporting the impact of these variables on quality of life, we have set the foundation for why these variables were selected for further analysis.

Research Questions and Objectives

The objective of this study is to identify neighborhoods in New York City that offer a high quality of life, taking into account the rent prices, the variation in noise complaints, and crime rates across different areas. By analyzing these factors, the livability conditions of various areas can be recognized. This study will benefit potential renters and policymakers who are looking to either move within NYC or allocate resources effectively in the city's diverse neighborhoods.

The research will be conducted through the following questions:

1) How are rent prices expected to change in the next two years across New York City?
2) How do the noise complaint types vary across neighborhoods in New York City?
3) How do the crimes vary across neighborhoods in New York City?

Research Question 1

How are rent prices expected to change in the next two years across New York City? To answer this, we will utilize a time-series analysis of the median rent prices for NYC from 2010 to 2023. This will help discern patterns and trends that may impact future prices and explore the variability of this across the city.

Data and Methodology

- Data: This study utilizes StreetEasy's dataset containing monthly data on the median asking rent in New York City.
- Cleaning: The initial data processing employs R's dplyr, lubridate, and tidyr packages to restructure the dataset into a long format that is suitable for further analysis. The dates are standardized and filtered to retain values from 2010 to 2023. Missing data points are imputed using the mice package, which employs random forest methodology to ensure a robust analysis.
- Preparation: The data is converted into a time series object and split into a training and testing set with training across 2010 to 2022 by months, and testing of 2023 by months to detect the accuracy of our predictions. This facilitates the use of various time-series forecasting techniques, including naive models, exponential smoothing, ETS and ARIMA models to determine which is the most appropriate forecasting method. Following this, the best method will be used to forecast rent prices in each borough in 2024 and 2025.

Time Series Analysis

1) Initial Observations: By taking a descriptive analysis on the median rental prices, we can observe a structural break in the data post-2020, which can be attributed to the macroeconomic impact of Covid-19, causing a drop in rental prices. However, we can observe that this was overset by an upward spike in the past 2 years. The seasonal plot shows how rental prices fluctuate in seasons, with minor peaks in the summer months and troughs in the winter months.

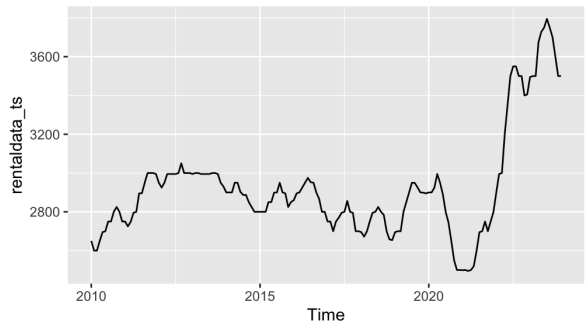

Figure 1.

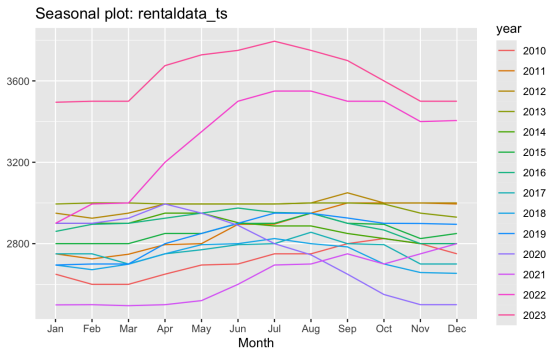

Figure 2.

2) Model Selection: To obtain predictions with the highest accuracy, we compare the performance of simple forecasting, exponential smoothing, ETS and ARIMA models on the testing data. As observed in Figure 4, the Holt Winter multiplicative method closely mimics the testing data. This model is effective as the data exhibits both seasonal and trend variations. It provides us with the lowest RMSE score at 96.63. Additionally, the diagnostic checks show the residuals behaving as white noise, with no apparent autocorrelations, affirming the model's fitness. The model's forecast aligns well with the actual data when visualized, offering credibility for future rental price predictions.

TS_model	RMSE
seasonal_naive_model	348.701042
drift_model	219.9423463
ses_model	247.7619202
holt_model	437.8064078
holt_damped_model	318.7642899
hw_additive_model	197.8411018
hw_multiplicative	96.63298701
ets_auto	321.3442813
arima	144.3624561

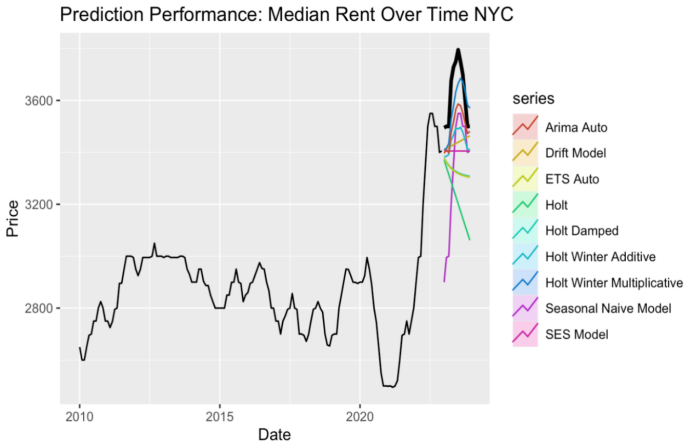

Figure 3.

3) Forecasting: By observing Figure 4, We can observe that there is an upward trend in the median rent prices. The forecast for January 2023 starts at $3,410, and by December 2025, it rises to $3,682. The method incorporates seasonality and a damping factor that gradually reduces the trend component over time, providing a point forecast with a confidence interval, as indicated by the shaded region. The multiplicative nature of the method suggests that seasonal changes are proportionally related to the level of the time series. This steady increase could be attributed to growth in the rental market due to inflation, housing demand and other macroeconomic factors.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2024	3564.45	3559.97	3569.98	3656.35	3699.56	3751.00	3775.43	3788.51	3768.74	3726.02	3665.37	3647.04
2025	3636.35	3627.19	3633.10	3716.89	3756.93	3805.48	3826.80	3836.80	3813.75	3767.72	3703.81	3682.89

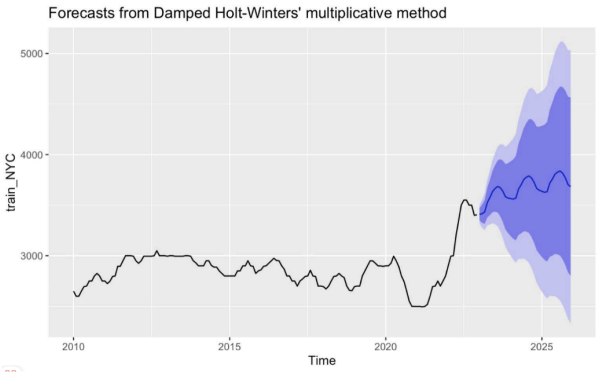

Figure 4.

Research Question 2

Noise Complaint Analysis: How do noise complaints vary across various neighborhoods in the city? To analyze this, we can implement cluster analysis to group neighborhoods on the basis of the various types of noise complaints. This will help identify the characteristics of a given neighborhood and determine whether it is suitable for a residents requirements.

Data and Methodology

- Data: This study utilizes NYC Open Data's dataset containing roughly 6.5 million complaints. It contains information on the type of complaint, description, date, and location is categorized by the Zipcode and Community Board.
- Cleaning: The study uses 10 years of data from 2012 to 2022, which is filtered using dplyr. As only 0.5% of the data contains missing values, these entries are deleted. Lastly, mismatched boroughs and community boards are readjusted. The dataset is summarized and arranged by borough, community board, zip code, complaint type and descriptor, to provide a structured view for analysis.
- Preparation: The required columns are scaled to standardize the data, ensuring each variable has equal weight in the clustering process. This scaling is essential to perform clustering algorithms like Hierarchical clustering, K-means and Model-based clustering, which are sensitive to the scale of the data.

Cluster Analysis

1) Initial observations: Upon visualizing the data, the number of complaints from each Community Board can be observed. 12 MANHATTAN (including the neighborhoods of Inwood and Washington Heights) and 12 BRONX (including neighborhoods of Edenwald, Wakefield, Williamsbridge, Woodlawn, Fish Bay, Eastchester, Olinville, and Baychester) appear to have a significantly high number of complaints while Staten Island has the least number of complaints.

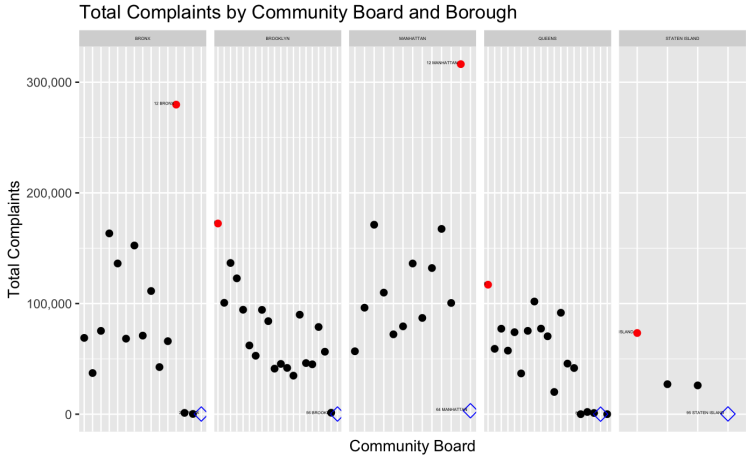

Figure 5.

2) Model Selection: A K-means model is selected for this analysis, as it is scalable and easy to implement. A model with 3 clusters is selected as it has the lowest Silhouette Width as viewed in Figure 7. Hierarchical clustering and model-based clustering are also tested out in both 3 and 4 clusters, however, the results from such are less easily interpretable when looking at the clustering profile. One noticeable community board stood out when setting the clusters between 3 and 4 in both hierarchical or K-means clustering, which is 08 Queens (including the neighborhoods of Jamaica Estates, Holliswood, and Flushing South). This areas shows high complaints from residential, vehicle and truck noise while holding the least complaints from the commercial category, indicating a noisy area with a less vibrant lifestyle near the JFK airport.

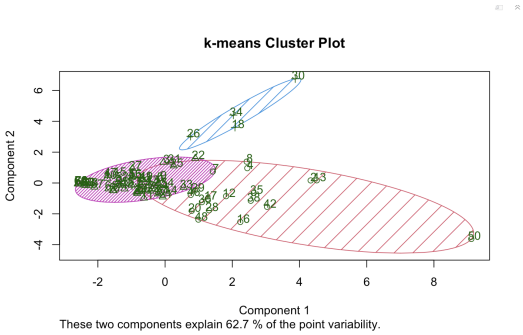

Figure 6.

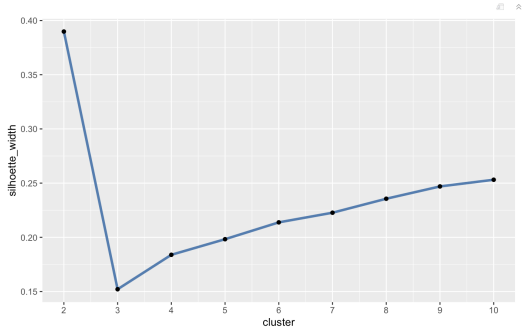

Figure 7.

3) Findings: By observing Figure 8, we can see that the city has been categorized into 3 distinct neighborhood types.

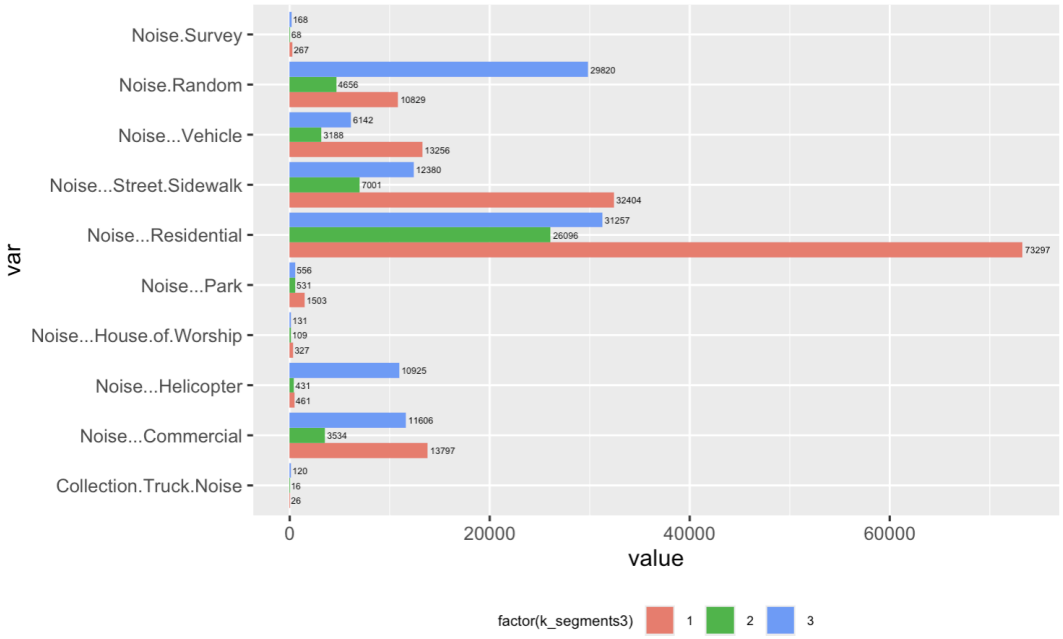

Figure 8.

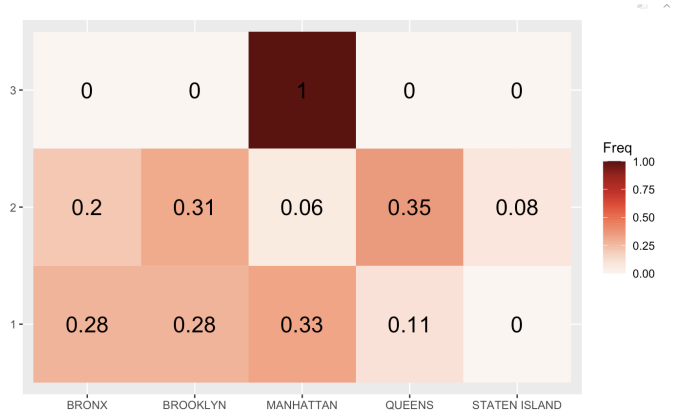

Figure 9.

- Cluster 1: High Levels of Noise Complaints.
  These neighborhoods have the highest levels of noise complaints, particularly for residential, street/sidewalk, and vehicle noises. This indicates high population density and commercial zones with significant street activity. 18 areas fall under this cluster, including the East and West Village, Park Slope, and the center of Flushing. Individuals who enjoy vibrant, lively environments and prioritize convenience over noise could choose to live here.

- Cluster 2: Moderate to Low levels of noise complaints.
  These neighborhoods may have a small mix of residential and commercial areas. All areas within Staten Island fall under this category, indicating that it is the quietest borough overall. The majority of the Bronx, the Financial District, and Jackson Heights also fall under this category. These neighborhoods are ideal for families and individuals who prioritize peace and tranquility and live in a commutable distance from urban conveniences.

- Cluster 3: High to Moderate Levels of Noise Complaints.
  These neighborhoods have relatively moderate levels from streets and sidewalks indicating that they are not very densely populated. However commercial, helicopter and random noises are significantly higher here than in the other two clusters. This could be due to proximity to helipads, offices, etc. As observed in Figure 9, only Manhattan neighborhoods lie in this cluster, including Chelsea, the Upper West Side and the Upper East Side. These areas can be suitable for residents and families that seek to be within short commutable distance from restaurants, commercial spaces, and attractions.

Research Question 3

Crime Level Analysis: How does crime quantity & type differ between neighborhoods? To understand how crime quantity differs between neighborhoods we will use spatial analysis to understand the distribution of crime across New York City. Secondly, we will also use spatial analysis to understand the most complaints per precinct.

Data and Methodology

- Data: We will use NYC open data for the crime and noise dataset.
- Cleaning: The initial data processing employs R's dplyr and tidy packages to restructure the dataset into a long format suitable for analysis. The dates are standardized and filtered to retain values from 2012 to 2022.
- Procedure: We used the library ggmap to load the ggmap package, which provides functions for visualizing spatial data with Google Maps API. We then used get_map() to fetch maps from Google Maps to serve as a base layer for the spatial visualization. It sets the location, zoom level, and scale for the map of New York City. ggmap() then initializes the plotting of the map object obtained from get_map(). Finally, we used functions such as theme(), and guides() to modify the appearance of the plot, including legend text size and the size of color points in the legend and labs() function to add labels and titles to the plot, which in this case adds the title "Crime Complaints Distribution in New York City 2012 - 2022".

Spatial Analysis

Crime Complaints Distribution Map:

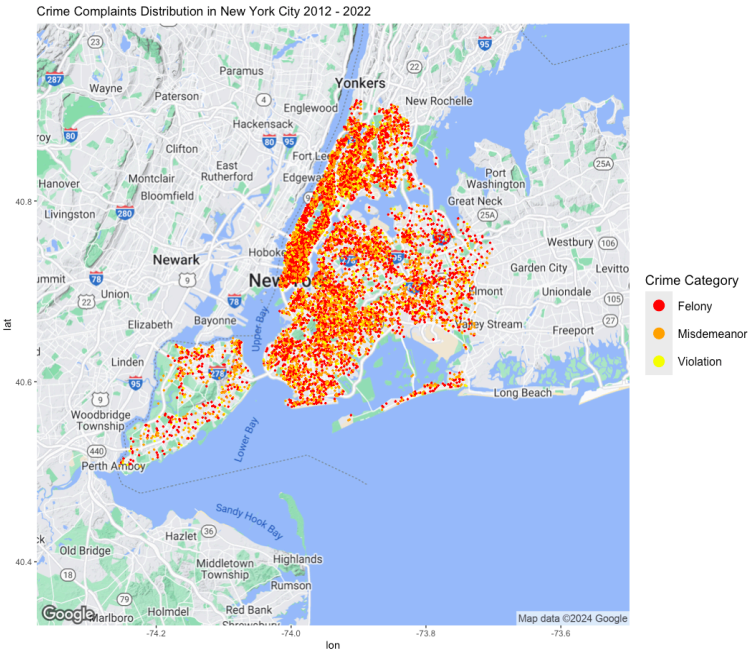

Figure 10.

- This map shows the distribution of crime by type (felony, misdemeanor, violation) across the city's geography.
- Crime is more concentrated in Manhattan, Southern Bronx & Central Brooklyn. Felonies were the most common category of crime
- Displaying the frequency of crimes as color-coded points allows for immediate visual identification of areas with higher crime densities. For spatial analysis, this map can help

identify hotspots for specific types of crime and observe any geographic patterns or clusters. For example, a high concentration of red points in a particular area would indicate a region with a high incidence of felonies.

Most Complaints per Precinct Map:

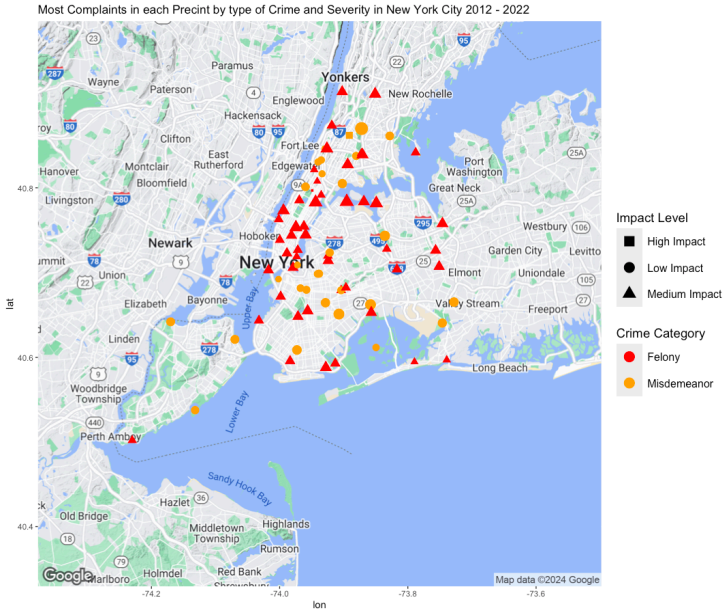

Figure 11.

- This map takes a different approach by focusing on the most common complaints per precinct, categorized by the severity of the crime.
- Distribution of the most complaints by crime category across NYC precincts
- Most complaints in Midtown & Lower Manhattan were medium-impact felonies.
- Low-impact misdemeanors were the category with the most complaints in Staten Island
- It provides a summarized view where each precinct is represented by a single symbol that characterizes the most frequent type of complaint. This helps in quickly identifying which precincts have higher levels of certain types of crimes and how the impact level of crimes varies across precincts.

TOTAL COMPLAINTS					
FELONY		MISDEMEANOR		VIOLATION	
HIGH (BROOKLYN)	LOW (STATEN)	HIGH (BROOKLYN)	LOW (STATEN)	HIGH (BROOKLYN))	LOW (STATEN)
1546	224	1499	256	319	67

Discussions and Recommendations

When choosing a neighborhood, tenants can consider the balance between rental costs, noise levels, and safety to find an area that aligns with their lifestyle and budgetary requirements.

- **Informed Selection:** Tenants can consider the rent forecast trends and varying noise levels and crime rates while considering their rental prices. Additionally, they can choose to begin agreements in winter instead of summer. For the year 2024, the median rental price in June is expected to be $3775, whereas it is $3647 in December.
- **Engagement and Advocacy:** Residents can engage with community boards, particularly in Cluster 1, for example, neighborhoods such as the East and West Village, Park Slope, and the center of Flushing to advocate for improved conditions and enhance neighborhood liveability.
- **Safety and Security:** Prospective renters wary of personal safety, can opt for areas such as Park Slope, Kew Gardens, or Staten Island with lower crime rates.

City Officials use this study for the following:

- **Resource Distribution:** City officials should utilize the study's insights to direct resources effectively toward noise hotspots in Midtown Manhattan.
- **Emergency Service Optimization:** Spatial data can be used to optimize the placement and response strategies of emergency services, improving response times in high-risk areas and potentially reducing crime-related damages.

Conclusion

Through this analysis, we are able to provide easy-to-understand neighborhood data that is important for prospective tenants and city officials. With these insights, we can help tenants make smart decisions, and highlight areas that need improvement so city officials can better serve their constituents. For tenants, we would recommend looking at apartments in Staten Island if they are searching for safe & quiet neighborhoods but they should be aware of the longer commute to Manhattan. We would also suggest that potential residents of Manhattan take into account the higher levels of crime & noise found in the borough when determining what neighborhood is best for them. For city officials, we would urge representatives for the Manhattan Community Board 12 & Bronx Community Board 12 to take action on reducing the significantly higher counts of noise complaints found in those regions. Additionally, since residential noise is a particular pain point for residents in the East Village, West Village, Flushing & Park Slope, officials should see what they can do to cut down on residential-related noise. Last but not least, the NYPD should strategically deploy its resources to Lower Manhattan & Midtown due to the quantity and severity of crimes occurring in those areas.

Limitations

While this study is comprehensive, we encountered several limitations that warrant mention. We found that connecting our datasets together was quite a challenge due to different neighborhood identifiers such as neighborhood name, zip code, precinct and community board. Additionally,

the lack of publicly available data on rental amenities limited the analysis we were able to conduct on rental prices in the city. Further research would include the formation of a directory that would allow us to connect neighborhoods with greater ease and the gathering of unit specific amenities. Additionally, now that we have the data to determine what needs to be improved for the various neighborhoods, we would want to conduct further research into what exact actions can be taken to improve the situations these communities face related to crime & noise.

References

Azad, S., & Ghandehari, M. (2021). Spatio-temporal study of the determinants of residential satisfaction in New York City during COVID-19 using crowdsourced data. In ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume VIII-4/W1-2021: 6th International Conference on Smart Data and Smart Cities, 15–17 September 2021, Stuttgart, Germany.

Ihlanfeldt, K., & Mayock, T. (2010). Panel data estimates of the effects of different types of crime on housing prices. Regional Science and Urban Economics, 40(2–3), 161-172. https://doi.org/10.1016/j.regsciurbeco.2010.02.005.

Siegel, J., & Bram, J. (2024, January 17). Spotlight: New York City's Rental Housing Market. Office of the New York City Comptroller Brad Lander. https://comptroller.nyc.gov/reports/spotlight-new-york-citys-rental-housing-market/

Zabel, J. (2016). A Dynamic Model of the Housing Market: The Role of Vacancies. Journal of Real Estate Finance and Economics, 53, 368–391. https://doi.org/10.1007/s11146-014-9466-z.

Zumper. (2024). Zumper National Rent Report. Zumper. Retrieved February 25, 2024, from https://www.zumper.com/blog/rental-price-data/

Data Sources

StreetEasy:
https://streeteasy.com/blog/data-dashboard

U.S. Census Bureau:
https://data.census.gov/table/ACSDP1Y2022.DP04?q=DP04:%20Selected%20Housing%20Characteristics&g=050XX00US36005,36047,36061,36081,36085_160XX00US3651000

NYPD:
https://www.nyc.gov/site/nypd/stats/crime-statistics/borough-and-precinct-crime-stats.page#manhattan

NYC Open Data 311:
https://data.cityofnewyork.us/Social-Services/311-Noise-Complaints/p5f6-bkga/about_data

Appendix

Figure 1: Community Boards: https://boundaries.beta.nyc/?map=cd

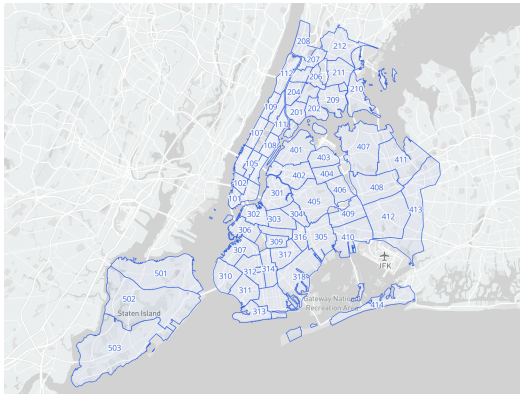

Figure 2: Police Precincts: https://boundaries.beta.nyc/?map=pp