MGT 6203 Group Project Proposal Template

TEAM INFORMATION (1 point)

Team #: 10. **Team Members:**

Juan David Rodriguez; jrodriguez381. I have wide experience in the Telecom sector in engineering, operations, sales, and analytics. I currently work as Director of Business Analytics in one of the largest Telecom Operators in North Latin America. I have led projects to enhance churn, commercial and operations processes using ML models. I am an Electronic Engineer with postgraduate studies in marketing, and an MBA from IE Business School.

Evgenia Jane Bugai; ebugai3. I have 10+ years of diversified expertise in finance, change management, and data intelligence. Currently, I work as a data engineer for an investment bank where I primarily focus on automation and process improvement. I hold a BBA in Accounting from Baruch College, CUNY and currently pursuing a MS in Data Analytics from Georgia Tech.

Muhammad Ahmad: mahmad67. I have experience in embedded systems, fraud analytics, and econometrics. I am currently working as a Business and Economic advisor for fleek Network. I graduated from Simon Fraser University with honors in Economics.

Nickolas Tyler Reinig; nreinig3. I have over 5 years of experience as a data analyst/specialist in the field of global health metrics. I've received degrees in chemical engineering and molecular biology from the University of Washington, and I am currently working towards an MS in Data Analytics form Georgia Tech.

Hisashi Chris Kominami; hkominami3. My professional background is in natural resources conservation and planning and environmental consulting. Over the past 14 years, I have managed federal conservation program projects in northern Vermont, delineated wetlands for regulatory purposes, investigated pesticide drift and volatility throughout the U.S., and collaborated on a variety of research projects. I have a B.A. in biological sciences, graduate certificates in ecological design and ecological economics, and a M.S. in plant and soil sciences.

OBJECTIVE/PROBLEM (5 points)

Project Title: ***Investigate the Relationship between Noise Complaints and Housing Cost in New York City***

Background Information on chosen project topic:
New York City is an exceptionally vibrant city that appeals to individuals from all backgrounds due to its wide range of job opportunities, rich culture, and educational institutions. This is why numerous people choose to reside and work in NYC, which inevitably impacts the cost of housing. According to Zillow Inc., the average cost of buying a home in NYC is $630,000, while the average rental price is $3,000 per month, both of which have increased steadily over the past ten years (Zillow Inc, Broakerage, 2023).

With over $3M in total housing units in NYC, the real estate market recorded around 26,000 transactions worth over $26 billion in sales and $84 billion in rental income in 2022 (New York City Department of Housing, 2022). This certainly provides lucrative opportunities for various players in the industry, such as property owners, investors, and real estate firms, including rental platforms such as Zillow and Redfin who are attempting to predict the housing market. While traditional price prediction methods have been widely used, there is a trend towards investigating less obvious factors to understand and forecast prices (McKinsey & Co., 2018).

One of such less obvious factors that can influence the New York City real estate market is noise pollution, which has been a growing pain for many New Yorkers. A study conducted by the NYC Health Department in 2012 found that the average levels of outdoor noise in many areas of the city are higher than both federal and international guidelines. Additionally, 20% of New Yorkers report being frequently disturbed by noise in their homes. This issue not only has a negative impact on people's health and quality of life but could also affect housing values. Therefore, our proposal seeks to investigate how noise pollution influences housing prices in NYC, which could potentially be used in price prediction.

Problem Statement (clear and concise statement explaining purpose of your analysis and investigation)

NYC is famous for its vibrancy, but it can also be a costly and loud place to live. It has become increasingly difficult for people to find quiet places to live that are also affordable. Our objective is to better understand the relationship between noise complaints (a proxy for noise pollution) and housing sales and rent prices in various neighborhoods throughout New York City.

State your Primary Research Question (RQ)

What is the relationship between noise complaints and housing cost across New York City neighborhoods?

Add some possible Supporting Research Questions (2-4 RQs that support problem statement):

1. How do noise complaint rates vary across different neighborhoods in New York City, and what factors might contribute to these variations?
2. Are neighborhoods with higher housing sales and rent prices more likely to have a higher or lower number of noise complaints?
3. Are neighborhoods with higher noise complaint rates associated with higher or lower housing sales and rent prices?
4. What is the geographical and frequency distribution of noise complaints by type in New York City?
5. How do noise complaints and types of noise complaints (e.g., construction, traffic, nightlife) impact housing sales and rent prices in different neighborhoods?
6. How do noise complaints and housing sales and rent prices vary by time of year, such as during tourist season or winter months?
7. Are there any notable differences in the relationship between noise complaints and housing sales and rent prices in residential versus commercial areas of the city?

Business Justification:

According to McKinsey's article, "*Getting ahead of the market: How big data is transforming real estate"*, using traditional independent variables to predict property market values (i.e., year built, rooms, location) can be very limited and exclude many of the non-evident factors that can significantly impact the price (McKinsey & Co., 2018).

One of such factors that can affect the quality of life and possibly account for the housing prices in New York City is the level of noise pollution. Our hypothesis is that the intensity of noise complaints could negatively affect both the value of houses and the cost of renting, leading to lower prices.

If the hypothesis is true, this information can be used to enhance existing market research and predictive toolkits. By better understanding the factors that impact price and improving its predictability, city planners, developers, and property owners can make better business decisions. They can identify novel areas with potential value and avoid investing in areas where the model predicts value erosion. The model can also improve pricing by considering factors that are relevant to tenants and buyers: **Assuming a 10% addressable market, property owners and developers selling $2.6 billion a year in houses and receiving $8.4 billion a year in housing rentals could benefit from our model. If the additional information results in even a 1% improvement in their performance, the delivered value could surpass $100 million per year.**

The model will also enable the identification of types of noise complaints that impact prices and the magnitude of their effects. The betas from the model can be utilized to develop a noise rating system for each geographical area by considering the distribution and types of noise complaints. This noise rating system could be monetized by online rental platforms like Zillow or Redfin, enhancing the value offered to property owners listing their properties, as well as improving the search experience for their users, and potentially driving more traffic to their websites, resulting in increased revenue.

Besides the financial benefits for real estate professionals, New York residents can use the noise rating system to choose neighborhoods that suit their preferences and reduce their exposure to noise, ultimately leading to a better quality of life. Additionally, property owners and policymakers can collaborate to enhance the current noise regulations in New York City.

DATASET/PLAN FOR DATA (4 points)

Data sources, data Description and key variables: The project will use five sizable datasets that will have to be grouped and transformed to be joined. Some of them will require the development of programmatic ETLs using available APIs:

311 NYC calls (Appendix 4) (New York City Open Data, 2023) (https://data.cityofnewyork.us/Social-Services/311-Service-Requests-from-2010-to-Present/erm2-nwe9). Noise complaint rate, type and frequency of complaints will be tested as independent variables. This is an extensive open database that includes all 311 New York City Calls from 2010 to present. 311 service in New York provides access to non-emergency services and city information. The dataset is comprised of more than 32 million records and 41 columns. It records geographical information (zip code, address, geographical coordinates), status and classification for each 311 calls. Given the size of the dataset, an ETL would have to be developed using the API from NYC Open Data so that we can limit our extraction to noise complaints in the period of interest. Most relevant variables from this dataset will be *Created Date, Complaint Type, Descriptor, Borough and Incident Zip Code.* Dummy variables representing the type of noise complaint will be developed in the raw dataset and then grouped by geography create a monthly rate per 1000 inhabitants. Additional transformations that will be explored during the project include event-frequency variables, and time series gradients to understand how the month over month change in the complaint rate may change the sale or rent price of houses.

Zillow Housing Research Data - Zillow Home Value Index (ZHVI) (Appendix 5) (Zillow, Inc., 2023) (https://www.zillow.com/research/data/)**:** Housing values will be tested as a dependent variable. Zillow provides an index that represents the typical home value for a given geographical area. The index is provided on a value per month for each zip code. It is calculated using the value of homes between 35 and 65 percentiles. The dataset can be downloaded in csv format from the Zillow housing research website, the raw dataset has 27022 rows and 286 columns. Relevant variables are the RegionName (zip code), and the ZHVI value per month (header in the dataset is the last day of the month). Dataframe will have to be transformed into a Tibble per zip code and month to join the data with the transformed 311 complaints' dataset. Month over month changes will be calculated for each zip code to understand the impact of the change in noise complaints in the housing sales prices.

Zillow Observed Rent Index (ZORI) (Appendix 6) (Zillow, Inc., 2023) (https://www.zillow.com/research/data/)**:** Housing rent prices will be also tested as a dependent variable. Zillow provides an index that represents the typical rental housing value in each geographical area. The index is provided on a value per month for each zip code. It is calculated by computing the mean of the rent value from houses between 40 and 60 percentiles. The dataset can be downloaded in csv format from the Zillow housing research website, the raw dataset has 6064 rows and 104 columns. Relevant variables are the RegionName (zip code), and the ZORI value per month (header in the dataset is the last day of the month). The dataframe will have to be transformed into a Tibble per zip code and month to join the data with the transformed 311 complaints' dataset. Month over month changes will be calculated for each zip code to understand the impact of the change in noise complaints in the housing rent prices.

US census estimated population series by zip code (Appendix 7) (United States Census Bureau, 2023): (https://data.census.gov/table?g=0100000US$8600000&tid=ACSST5Y2021.S0101). Due to the wide range of population sizes in different zip codes, complaints cannot be analyzed in absolute volumes, they must be calculated as a rate per 1000 inhabitants. This dataset by the US Census Bureau provides the population estimation for each zip code, allowing the complaint rate calculation. The dataset can be downloaded in csv format from the US Census Bureau website, the raw dataset has 33776 rows and 914 columns. Relevant variables are the *Geographical Area Name* and the *Total Population.* Different regex and casting techniques will have to be used to create usable join keys and data values.

IRS returns statistics. Income by zip code (Appendix 8) (United States Internal Revenue Service - IRS, 2023) (https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-statistics-2020-zip-code-data-soi)**:** One of our hypotheses is that the relationship between house sales or rent values and noise complaints may differ depending on the income level of the zip code. IRS publishes comprehensive geolocated data (by zipcode) of the statistics of the tax returns. This information can be used to classify zip codes by estimated income level. The dataset can be downloaded in csv format from the IRS returns statistics website, the raw dataset has 27745 rows and 165 columns. Relevant variables are the *Zip Code, Total Income (A02650), and Total Returns (N02650).* An income per return rate will be calculated per zip code using these variables as a proxy of the incomer per capita standard metric. Values will then have to be categorized using ntile distributions to divide zip codes per income category (low, mid, high).

APPROACH/METHODOLOGY (8 points)

In this section, we outline our approach and methodology for addressing the problem. The steps involved are as follows:

1. Data extraction and preparation: We will extract noise complaints from the 311-dataset using the sodapy API.
    a) These complaints will be summarized by geographical area, and features will be developed to understand changes in noise complaints by type.
    b) We will also extract and group data on house values, rent prices, population, and income per geographical area.
2. Datasets will be joined using the geographical area and month as keys.
3. We will perform an analysis of the prices time series to detrend the series, exclude outliers or zip codes without residential units, and identify other influential variables such as income that must be controlled.
4. Understanding the relationship between noise complaints and housing and rent prices:
    a) We will develop a linear regression to understand the relationship between total noise complaint ratios and the house value index. Since the relationship is unlikely to be linear, we will test various transformations such as linear-log, log-linear, and log-log.
    b) We will then use diagnostic plots to determine which transformation provides the best fit for the data.
    c) We will also test a regression model by comparable income groups to determine if the resulting parameters (Adjusted R-Square, diagnostic plots) improve.
    d) A multivariate regression will be tested by separating the noise complaints by type and timeframe.
    e) We will test the change in complaints' ratio as a predictor of the change in t+1, t+2, and t+3 house and rent prices.

Regression model:

General Idea for the Models:

Note: These might change in final draft, this is only meant for proposal purposes.

In our model, we conduct a regression analysis of the average rental price and residential property prices in a given year within a Zip Code while accounting for the impact of population growth. We consider population growth as a suitable proxy to account for natural changes in rental income that are not related to noise complaints. Additionally, we will explore detrending the average rental income and using it as an independent variable. However, this approach may result in a loss of information that could potentially weaken the impact of noise.

To avoid the over-representation of income in our analysis, we will employ separate regression models for each income level, recognizing that the relationship between noise and rental income may differ across income levels. Income levels will be categorized based on the number of standard deviations from the average income level in New York City (i.e., threshold of three standard deviations below the average income level will define the low-income group). We will also consider incorporating higher dimensions of income as an explanatory variable to account for non-linear relationships at the extreme ends of the income spectrum.

We will include Zip Code as a dummy variable in our regression to account for potential variations among Zip Codes, such as differing laws and regulations. However, we acknowledge that this approach may adversely affect the performance of our model. From an economic research standpoint, controlling for these variables is still recommended, but our primary focus is on prediction performance. Therefore, we will also seek access to data on rent control and rent stabilization policies to control for the variation due to different rental policies.

The underlying assumption for our models is that we perceive no reason for endogeneity in our model as the fundamental premise is that fluctuations in noise complaints are exogenous and are not influenced by changes in rental prices. This assumption is reasonable because property owners, in general, do not alter rental rates to manipulate the number of noise complaints, nor do changes in rental rates lead to a variation in the amount of noise generated.

Important Note: Geographical references or Zip Codes will solely serve as controls, as we intend to employ an index to obfuscate any reference to Zip Codes in relation to geographical locations. It is important to note that our study does not concern itself with potential differences among Zip Codes.

We are aware that depending on the income level, the relationship may be affected. Therefore, we plan to run separate models for different income levels.

Below are some examples of our OLS strategy-based models:

Regression model with Zip Code as a proxy for each income level:

$$Rnt_{iym} = \beta_0 + \beta_1 TNC_{iym} + \sum_{2}^{j} \beta_j NC_{ijym} + \gamma Pop_y + \omega AvgInc_{iy} + \tau {AvgInc^2}_{iy} + \rho_i I + \epsilon_i$$

Regression model with percentage or share of rent control/stabilization properties or units:

$$Rnt_{iy} = \beta_0 + \beta_1 TNC_{imy} + \sum_{2}^{j} \beta_j NC_{ijmy} + \gamma Pop_y + \omega AvgInc_{iy} + +\tau {AvgInc^2}_{iy} + \rho_i SRC_i + \epsilon_i$$

Gradient (change) model in rent prices:

$$\Delta Rnt_{iym} = \beta_0 + \beta_1 \Delta TNC_{iym} + \sum_{2}^{j} \beta_j \Delta SNC_{ijym} + \gamma SRC_i + \in_i$$

Gradient (change) model in sales prices:

$$\Delta Pr_{iym} = \beta_0 + \beta_1 \Delta TNC_{iym} + \sum_{2}^{j} \beta_j \Delta SNC_{ijym} + \gamma SRC_i + \in_i$$

Example of a probabilistic model:

$$P(+ve\Delta Rnt_{iy}) = \frac{1}{1 + e^{-\beta_0 + \beta_1 TNC_{imy} + \sum_2^j \beta_j NC_{ijmy} + \gamma Pop_y + \omega AvgInc_{iy} + \tau AvgInc^2{}_{iy} + \rho_i I_i + \epsilon_i}}$$

- β_1: Coefficient that represents the change in rental or sales prices given change in noise complaints.
- β_j: Coefficient that represents the change in rental or sales prices given change in noise complaints of a j type

Variables:

- Rnt_{iym}: Average rent price on i[th] geographical area, month m and year y
- Pr_{iym}: Average sales price on i[th] geographical area, month m and year y
- TNC_{iym}: Total number of noise complaints in i[th] geographical area at month m and year y
- SNC_{ijm}: Share of total number of noise complaints in i[th] geographical area and type j at month m and year y
- SRC_i: Share of rentals in the geographical area that are rent controlled.
- I_i: Index for i^{th} zip-code or geographical area
- Pop_y: Population at month y
- $AvgInc_{iy}$: Average income of i^{th} geographical area at year y

Anticipated Conclusions/Hypothesis:

We hypothesize that housing and rent prices in New York City neighborhoods are correlated with the number of noise complaints in those neighborhoods, and that the relationship is statistically significant. We expect this to be the case especially since the pandemic began, since the increased number of remote workers and increased time spent at home has likely made individuals more sensitive to noise.

We also predict that there could be a different relationship between noise complaints and housing/rent prices among different populations (characterized by different income levels, or by different regions of the city). Thus, we will likely want to control for other variables in our regression model and/or stratify our data by income levels, for example, to fully explain the relationship between noise complaints and housing and rent prices.

Even if our hypothesized correlation is found to exist, it may be useful to further explore the relative timing of large changes in noise complaints and changes in housing prices. If, for example, noise complaint increases precede housing/rent price decline (instead of occurring after those declines), then it may be possible to use noise complaints as a predictor or early warning sign of price deterioration in that region. Additionally, we are interested in looking at different time periods within the data and hypothesize that a stronger relationship between noise complaints and rent/housing price exists during the pandemic than pre-pandemic, given the unprecedented rise in remote work (and the necessity of maintaining a quiet environment that goes along with that).

We will also analyze types of noise complaints and the time of day when each complaint was made. Certain types of noise may have a stronger correlation with housing/rent prices in neighborhoods, and other types may be less correlated or even correlated in a different direction. The time of day/night a noise complaint occurs may be similar, with certain times more strongly correlated to housing prices. For example, noise late at night may indicate higher crime rates compared to noise during the afternoon, with crime likely having a strong correlation to housing/rent prices.

What business decisions will be impacted by the results of your analysis? What could be some benefits?

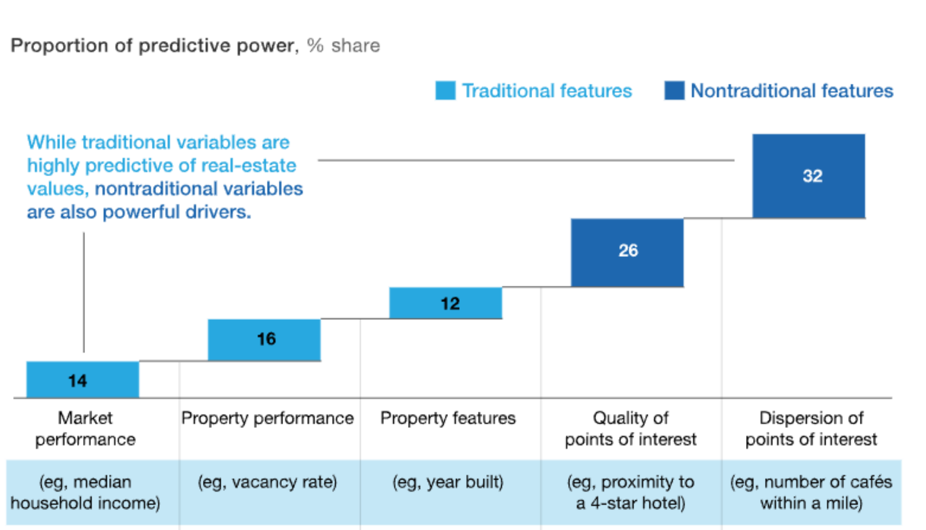
- Landlords and property managers: If our hypothesis is correct and noise complaints and housing rents are correlated, then this knowledge could be very useful for landlords and property managers in terms of how they market and price their units to prospective tenants and how they approach dealing with excessive noise in a specific area. For instance, landlords may discover untapped value in properties where there is little noise and where rents are low compared to other areas. Additionally, and depending on the source and type of noise pollution, landlords and property managers could proactively use the results from our analysis to develop strategies to mitigate noise.
- NYC Tenants: With noise pollution being the top civic complaint by NYC residents, we suspect that the results from our analysis, specifically a noise rating system, could be valuable to prospective tenants in helping them to make more informed decisions about rental units based on their tolerance for noise and preferences.
- Policy makers and governmental agencies: If our hypothesis is true, public policy-makers should consider noise complaints as a proxy for neighborhood healthiness, track its trend, and act to improve healthy and tolerant coexistence among neighbors. The same is true on a private level. Normally noise complaints are answered effectively mostly in affluent areas, while they might be ignored in other areas.

PROJECT TIMELINE/PLANNING (2 points)

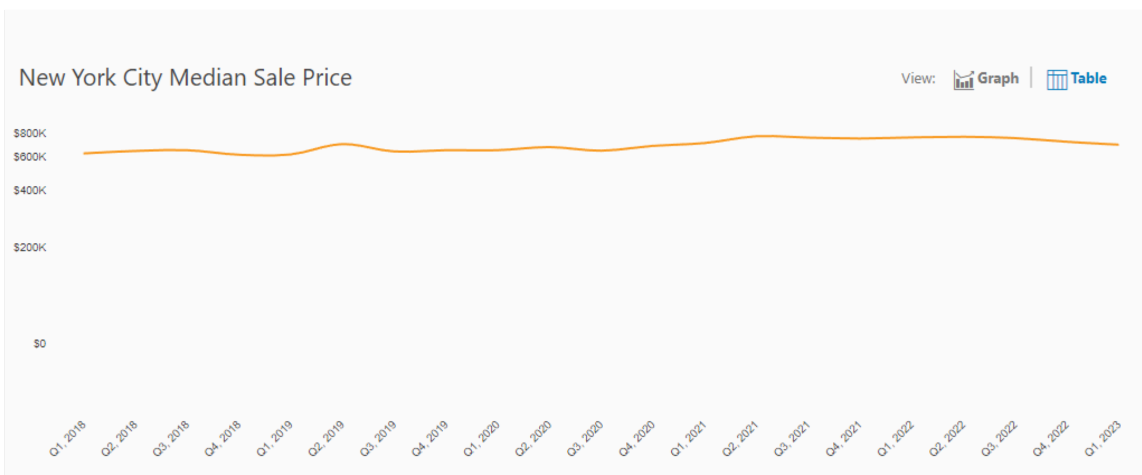
Assignment	**Date**
Complete draft of project proposal	Mon 3/6
Group meeting to revise proposal draft	Mon 3/6
Meet with TAs to discuss proposal	Tues 3/7
Work on revising/finalizing proposal	Weds 3/8 – Sun 3/12
Turn in project proposal	Sun 3/12
Work on project proposal video	Mon 3/13 – Sun 3/26
Turn in proposal video (4-5 min)	Sun 3/26
Work on project progress report	Mon 3/27 – Sun 4/2
Turn in progress report (4-5 pgs)	Sun 4/2
Work on final report	Sun 4/2 – Sun 4/16
Work on final presentation video	Sun 4/2 – Sun 4/19
Turn in final report (8-10 pgs)	Sun 4/16
Turn in final presentation video (10-12 min)	Weds 4/19
Turn in final slides, code, data, and any other materials	Weds 4/19
Work on outside-of-group peer review	Thurs 4/20 – Weds 4/26
Turn in peer review	Weds 4/26
Complete and turn in within-group performance evaluation	Weds 4/26

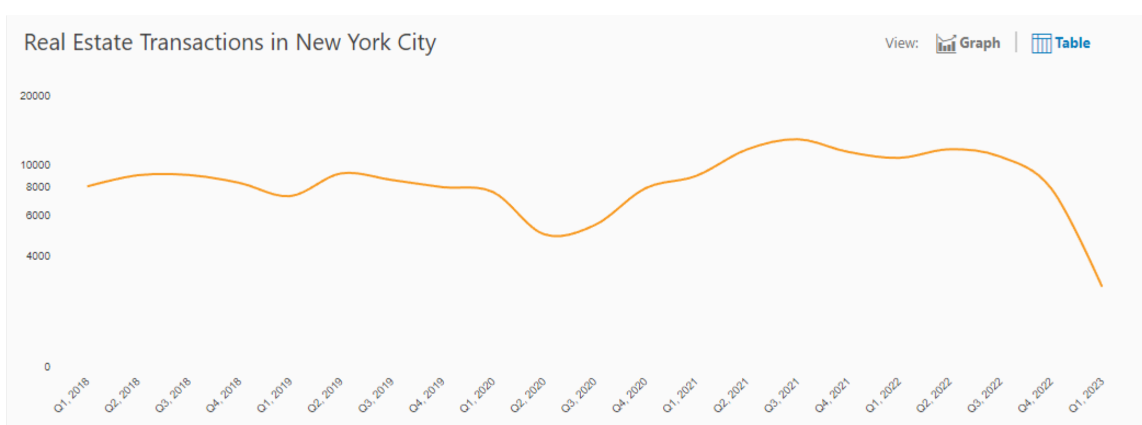
Appendix

Appendix 1: The value of nontraditional datasets to predict real state pricing (McKinsey & Co., 2018)

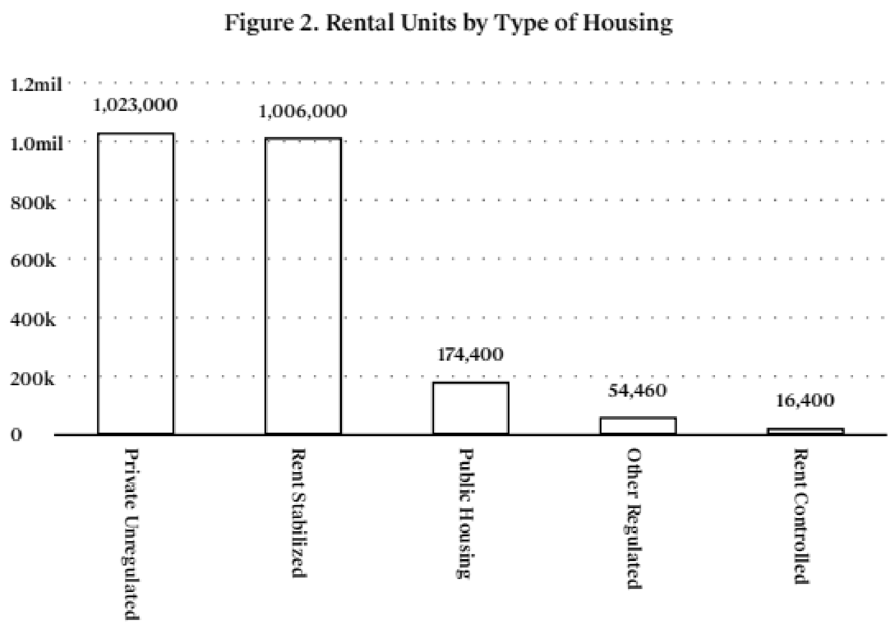


Appendix 2: Median Sale Price and number of transactions in New York (Zillow Inc, Broakerage, 2023)

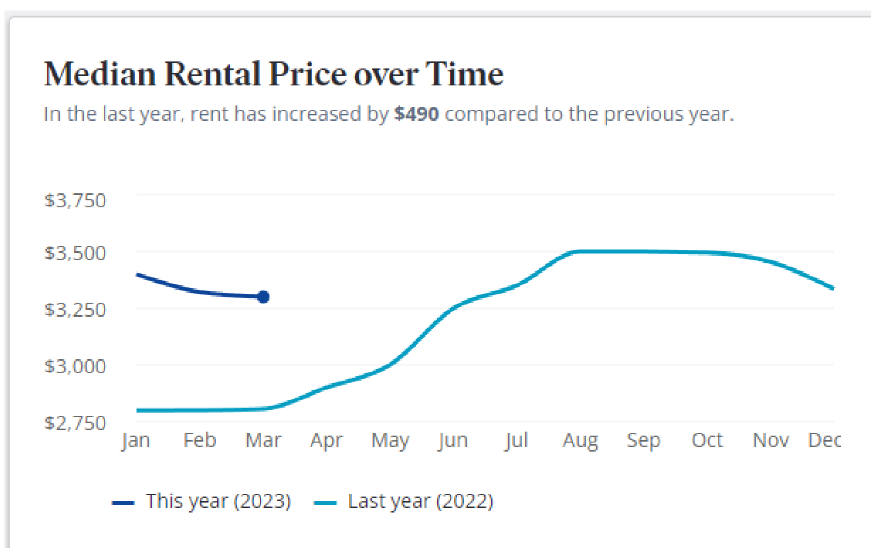




Appendix 3: Number of rental units by type in New York (New York City Department of Housing, 2022)



Appendix 4: Median Rental Price New York (Zillow Inc, Brokerage, 2023)



Appendix 5: 311 Calls Preview (New York City Open Data, 2023)

Uniqu...	Cre	Close...	Agen...	Agen...	Complaint Type	Descriptor	Locat...	Incident Zip
56976998	03/08/20...		DSNY	Departme...	Derelict Vehicles	Derelict Vehicles	Street	10463
56979063	03/08/20...		DSNY	Departme...	Derelict Vehicles	Derelict Vehicles	Street	11208
56985254	03/08/20...		NYPD	New York...	Illegal Parking	Blocked Hydrant	Street/Si...	11214
56979766	03/08/20...		DSNY	Departme...	Sanitation Worker or Vehicle Co...	Noise	Street	11216
56979855	03/08/20...		NYPD	New York...	Noise - Street/Sidewalk	Loud Music/Party	Street/Si...	11230
56983802	03/08/20...		NYPD	New York...	Noise - Residential	Loud Music/Party	Residenti...	10469

Appendix 6: Zillow Home Value Index (ZHVI) (Zillow, Inc., 2023)

RegionID	SizeRank	RegionName	RegionType	StateName	State	City	Metro	CountyName	1/31/2000	2/29/2000
91940	0	77449	zip	TX	TX		Houston-The Woodlands-Sugar Land, TX	Harris County	111025.3645	111044.18
91982	1	77494	zip	TX	TX		Houston-The Woodlands-Sugar Land, TX	Fort Bend County	217709.4609	217896.5096
93144	2	79936	zip	TX	TX	El Paso	El Paso, TX	El Paso County	93297.84602	93280.38057
62080	3	11368	zip	NY	NY	New York	New York-Newark-Jersey City, NY-NJ-PA	Queens County	75454.47444	76217.95511
62093	4	11385	zip	NY	NY	New York	New York-Newark-Jersey City, NY-NJ-PA	Queens County	237548.7708	238572.171
95992	5	90011	zip	CA	CA	Los Angeles	Los Angeles-Long Beach-Anaheim, CA	Los Angeles County	118733.4148	119187.4044
84630	6	60629	zip	IL	IL	Chicago	Chicago-Naperville-Elgin, IL-IN-WI	Cook County	90099.23606	90061.79477
91733	7	77084	zip	TX	TX	Houston	Houston-The Woodlands-Sugar Land, TX	Harris County	110385.5353	110340.0181
96361	8	91331	zip	CA	CA	Los Angeles	Los Angeles-Long Beach-Anaheim, CA	Los Angeles County	128621.7919	128796.5043
96193	9	90650	zip	CA	CA	Norwalk	Los Angeles-Long Beach-Anaheim, CA	Los Angeles County	166267.3394	166405.5852
61148	10	8701	zip	NJ	NJ	Lakewood	New York-Newark-Jersey City, NY-NJ-PA	Ocean County	132697.8686	133245.2363
62046	11	11236	zip	NY	NY	New York	New York-Newark-Jersey City, NY-NJ-PA	Kings County	184431.4075	185150.6957

Appendix 7: Zillow Observed Rent Index (ZORI) (Zillow, Inc., 2023)

RegionID	SizeRank	RegionNa	RegionType	StateNam	State	City	Metro	CountyName	3/31/2015	4/30/2015
91940	0	77449	zip	TX	TX		Houston-The Woodlands-Sugar Land, TX	Harris County	1328.456694	1342.566883
91982	1	77494	zip	TX	TX		Houston-The Woodlands-Sugar Land, TX	Fort Bend County	1559.889116	1561.690931
93144	2	79936	zip	TX	TX	El Paso	El Paso, TX	El Paso County		
62093	4	11385	zip	NY	NY	New York	New York-Newark-Jersey City, NY-NJ-PA	Queens County	2155.301037	2193.199603
95992	5	90011	zip	CA	CA	Los Angeles	Los Angeles-Long Beach-Anaheim, CA	Los Angeles County		
84630	6	60629	zip	IL	IL	Chicago	Chicago-Naperville-Elgin, IL-IN-WI	Cook County		
91733	7	77084	zip	TX	TX	Houston	Houston-The Woodlands-Sugar Land, TX	Harris County	1272.693935	1285.905748
96193	9	90650	zip	CA	CA	Norwalk	Los Angeles-Long Beach-Anaheim, CA	Los Angeles County		
62046	11	11236	zip	NY	NY	New York	New York-Newark-Jersey City, NY-NJ-PA	Kings County		
96083	12	90201	zip	CA	CA	Bell	Los Angeles-Long Beach-Anaheim, CA	Los Angeles County		
96816	13	92335	zip	CA	CA	Fontana	Riverside-San Bernardino-Ontario, CA	San Bernardino County		
62019	14	11208	zip	NY	NY	New York	New York-Newark-Jersey City, NY-NJ-PA	Kings County		
61807	15	10467	zip	NY	NY	New York	New York-Newark-Jersey City, NY-NJ-PA	Bronx County	1383.923519	1380.87561
62037	16	11226	zip	NY	NY	New York	New York-Newark-Jersey City, NY-NJ-PA	Kings County	1983.45186	2008.605562

Appendix 8: US census estimated population series by zip code. (United States Census Bureau, 2023)

GEO_ID	NAME	S0101_C01_001E	S0101_C01_001M	S0101_C01	S0101_C01	S0101_C01	S0101_C01	S0101_C01	S0101_C01
Geography	Geographic Area Name	Estimate!!Total!!Total population	Margin of Error!!Total!!Total population	Annotatio	Annotatio	Estimate!	Margin of	Annotatio	Annotatio
860Z200US00601	ZCTA5 00601	17126	429	null	null	643	24	null	null
860Z200US00602	ZCTA5 00602	37895	279	null	null	1267	32	null	null
860Z200US00603	ZCTA5 00603	49136	841	null	null	1873	105	null	null
860Z200US00606	ZCTA5 00606	5751	355	null	null	223	24	null	null
860Z200US00610	ZCTA5 00610	26153	382	null	null	838	28	null	null
860Z200US00611	ZCTA5 00611	1283	358	null	null	85	80	null	null
860Z200US00612	ZCTA5 00612	64090	1466	null	null	2649	191	null	null
860Z200US00616	ZCTA5 00616	10186	1247	null	null	233	130	null	null
860Z200US00617	ZCTA5 00617	22803	283	null	null	856	57	null	null
860Z200US00622	ZCTA5 00622	7751	1105	null	null	180	115	null	null
860Z200US00623	ZCTA5 00623	39652	1105	null	null	1297	115	null	null
860Z200US00624	ZCTA5 00624	21912	412	null	null	1151	119	null	null
860Z200US00627	ZCTA5 00627	32885	*****	*****	null	1194	*****	*****	null
860Z200US00631	ZCTA5 00631	1098	380	null	null	13	21	null	null
860Z200US00636	ZCTA5 00636	1174	525	null	null	70	92	null	null

Appendix 9: IRS returns statistics. Income by zip code. (United States Internal Revenue Service - IRS, 2023)

STATEFIPS	STATE	ZIPCODE	AGI_STUB	N1	MARS1	MARS2	MARS4	ELF	CPREP	PREP	DIR_DEP	VRTCRIND	N2	TOTAL_VI	VITA	TCE	VITA_EIC
1	AL	0	0	2161370	982450	749300	379640	2007020	48290	1161780	1529920	18330	4123620	26020	18840	7180	2670
1	AL	35004	0	5420	2450	2060	790	5060	130	2470	3950	50	10400	20	0	20	0
1	AL	35005	0	3440	1650	800	900	3150	90	1820	2610	30	6210	30	30	0	0
1	AL	35006	0	1230	520	550	150	1170	30	800	920	0	2500	0	0	0	0
1	AL	35007	0	12600	5380	5080	1770	11500	380	5980	8620	130	25270	150	80	70	0
1	AL	35010	0	8280	3580	2670	1860	7860	170	4080	5900	40	15690	590	250	340	90
1	AL	35014	0	1730	770	530	410	1620	50	1000	1250	0	3110	0	0	0	0
1	AL	35016	0	7700	3250	3350	930	7230	170	4390	5630	70	15200	220	190	30	30
1	AL	35019	0	1000	430	440	100	960	20	610	740	0	2030	0	0	0	0
1	AL	35020	0	9570	4920	950	3500	8620	340	5690	7450	50	17240	140	140	0	40
1	AL	35022	0	10890	5050	3300	2160	9940	300	5820	7550	120	19570	120	90	30	0
1	AL	35023	0	10660	4850	3470	2150	9910	250	6100	8000	90	19970	50	50	0	0
1	AL	35031	0	3080	1200	1440	360	2890	80	1890	2350	20	6530	0	0	0	0
1	AL	35033	0	1520	570	770	130	1450	20	930	1050	0	3100	0	0	0	0

Reference List

Ariel Property Advisors. (2023, January). *Multifamily Year In Review 10+ Residential Units: New York City | 2022.* Retrieved from https://arielpa.com/report/report-MFYIR-2022

Forbes. (2023, March 5th). *Partying Like It's 2015: Multifamily Housing In New York City.* Retrieved from https://www.forbes.com/sites/shimonshkury/2023/02/09/partying-like-its-2015-multifamily-housing-in-new-york-city/?sh=16bc2efe2736

Hammer, M. S., Swinburn, T. K., & Neitzel, R. L. (2014). Environmental noise pollution in the United States: developing an effective public health response. *Environmental health perspectives*, *122*(2), 115-119.Retrieved from HYPERLINK "https://ehp.niehs.nih.gov/doi/pdf/10.1289/ehp.1307272"https://ehp.niehs.nih.gov/doi/pdf/10.1289/ehp.1307272

McKinsey & Co. (2018, October 8th). *Getting ahead of the market: How big data is transforming real estate*. Retrieved from https://www.mckinsey.com/industries/real-estate/our-insights/getting-ahead-of-the-market-how-big-data-is-transforming-real-estate#

McKinsey & Company. (2022, March). *McKinsey Global Private Markets Review 2022.* Retrieved from https://www.mckinsey.com/~/media/mckinsey/industries/private%20equity%20and%20principal%20investors/our%20insights/mckinseys%20private%20markets%20annual%20review/2022/mckinseys-private-markets-annual-review-private-markets-rally-to-new-heights-vf.pdf

New York City Department of Health and Mental Hygiene (2014). *Ambient Noise Disruption in New York City.* Retrieved from https://www.nyc.gov/assets/doh/downloads/pdf/epi/databrief45.pdf

New York City Department of Health and Mental Hygiene (2013). *Preventing noise-induced hearing loss among young people*. Retrieved from https://www.nyc.gov/assets/doh/downloads/pdf/epi/databrief45.pdf

New York City Department of Housing. (2022, May 16th). *2021 New York City Housing and Vacancy Survey*. Retrieved from https://www.nyc.gov/assets/hpd/downloads/pdfs/services/2021-nychvs-selected-initial-findings.pdf

New York City Open Data. (2023, March 5th). *311 Service Requests from 2010 to Present*. Retrieved from https://data.cityofnewyork.us/Social-Services/311-Service-Requests-from-2010-to-Present/erm2-nwe9

New York City Open Data. (2023, March 5th). *DOF: Summary of Neighborhood Sales by Neighborhood Citywide by Borough*. Retrieved from https://data.cityofnewyork.us/City-Government/DOF-Summary-of-Neighborhood-Sales-by-Neighborhood-/5ebm-myj7

PropertyShark - Yardi Systems, Inc. (2023, March 9th). *Market Trends*. Retrieved from https://www.propertyshark.com/mason/market-trends/residential/nyc-all

United States Census Bureau. (2023, March 5th). *S0101 ACS 5 Year Estimates Subject Tables*. Retrieved from https://data.census.gov/table?g=0100000US$8600000&tid=ACSST5Y2021.S0101

United States Internal Revenue Service - IRS. (2023, March 5th). *SOI Tax Stats - Individual Income Tax Statistics - 2020 ZIP Code Data (SOI)*. Retrieved from https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-statistics-2020-zip-code-data-soi

World Health Organization. (2011). *Burden of disease from environmental noise: Quantification of healthy life years lost in Europe*. World Health Organization. Regional Office for Europe. Retrieved from HYPERLINK "https://apps.who.int/iris/bitstream/handle/10665/326424/9789289002295-

eng.pdf?sequence=l&isAUowed=y"https://apps.who.int/iris/bitstream/handle/10665/326424/9789289002295-eng.pdf?sequence=l&isAUowed=y

Zillow, Inc. (2023, March 5th). *Zillow Housing Research Data*. Retrieved from https://www.zillow.com/research/data/

Zillow Inc, Broakerage. (2023, March 9th). *New York, NY Rental Market*. Retrieved from https://www.zillow.com/rental-manager/market-trends/new-york-ny/