Explaining Property Value by Zip Code in New York City's Five Boroughs

Lily Davis and Ezey Duru

10 December 2025

Introduction

This paper aims to evaluate the relationship between neighborhood characteristics and property value in New York City's Five Boroughs. Understanding how these factors affect property value is essential for informing policy decisions aimed at improving New York City's housing crisis. In order to evaluate this relationship, a multiple linear regression model was used to examine the relationship between a number of predictors (number of subway stations, grocery stores, new businesses, public and non-public schools, crime rate, population, income, education, and unemployment) and property value. The same regression was also used in each Borough to identify which features affect property values in different areas.

New York City is currently facing its worst housing affordability crisis in a century. Rent is increasing faster than income is increasing, and people are having to spend a major portion of their income on rent leaving very little money for basic necessities. More jobs are being created than new housing, and generally there is a lack of available-to-rent housing. In October of 2025, 86,000 people on average were staying in shelters per night (Gray, 2025).

This housing crisis is constantly debated in New York City's political landscape. A crucial step towards finding effective solutions is understanding the key factors that drive property values across different neighborhoods. Not much previous research on this topic looks at a variety of neighborhood characteristics. The previous work commonly examines one neighborhood feature and then uses property factors as other variables.

One study examined the effect of Business Improvement Districts (BIDs) on commercial property value. A BID is a designated area where property owners pay a required fee to fund extra services like cleaning, security, and beautification that improve the local business environment. A difference-in-differences hedonic model shows that for commercial property values BIDs have, on average, a positive effect. In residential properties, findings show that values increase during the process of BID formation (possibly due to anticipation) but fall once the BIDs are actually formed. The study uses property characteristics as other independent factors such as building age, lot size, odd shape, and others (Ellen et al., 2007).

Another analysis looked into how the number of grocery stores in a census tract affects non-commercial property value. It was found that the increased number of retail food stores increases this value. This study was done at the census tract level and other factors in the model included income, education level, and unemployment rate within census tracts (Kole, 2021).

One study investigated how crime affects commercial property values. 2SLS models were used in an attempt to mitigate any reverse causality that could bias results.

The final results showed that crime has a negative impact on property value especially in lower income and higher minority neighborhoods. This study used building features such as number of floors, number of units, lot location and others as other independent variables (Lens & Meltzer, 2016).

Methods and Data

Data Sources

Desired predictors were sourced from multiple datasets. NYC Open Data was used for crime rate, number of new businesses, and a shapefile identifying zip codes. New York State Government data was used for grocery stores, schools, and borough identification. Data.gov provided subway station data. Social Explorer was used for income, education, and unemployment. Population data was sourced from New York Demographics. Zillow was used for property value data.

Data Cleaning & Preparation

R Studio was used to do all cleaning and data processing. The first cleaning step involved loading and looking at all datasets. Since the goal was to look at all features by zip code, any datasets without zip codes had to be geocoded. This was done with a shapefile that included all zip code tabulation areas in New York City and thus provided the exact list of zip codes that needed to be analyzed. The datasets without zip codes still contained latitude and longitude coordinates, so using the shapefile and the sf or special features library in R (commonly used for geographic data) the coordinates could be converted to spatial points. A spatial join then made it possible to identify what zip code each spatial point was located within. Each observation in the dataset was now assigned its corresponding zip code. This method was used for the subway station and crime rate datasets. The subway data listed all MTA subway and Staten Island Railway stations, with information on their locations. The crime rate dataset included all valid felony, misdemeanor, and violation crimes reported to the New York City Police Department (NYPD).

All other datasets included a zip code variable. For any dataset including data for more than just New York City, the shapefile was used to filter the source to only NYC zip codes. No major cleaning was needed for the population, income, education, unemployment, and grocery store datasets. The population dataset was simply the population of each zip code in NYC. Income was median household income. The education variable was the percentage of the population 25 years and older with a bachelor’s degree or more. The unemployment rate was the percentage of the civilian population in the labor force 16 years and older who are unemployed. The grocery store dataset contained all retail food stores which are licensed by the Department of Agriculture and Markets.

The school dataset included all NY Schools, kindergarten through 12th grade within New York State. Institutions that were not truly schools (i.e. district buildings, afterschool programs) were dropped. Institutions were separated into public and non-public schools (private and independent schools).

New businesses were obtained from a licensing dataset with all licenses issued by the NYC Department of Consumer and Worker Protection (DCWP). By selecting businesses with an initial issuance date in 2024 who had never been issued a license before, the new businesses in 2024 were identified.

Property value was obtained using the Zillow Home Value Index (ZHVI): A measure of the typical home value and market changes across a given region for all housing types. It reflects the typical value for homes in the 35th to 65th percentile range as a smoothed and seasonally adjusted measure. The selected region was New York City, and the data was presented monthly, so all months in 2024 were averaged to identify the yearly average in each zip code.

Lastly, the borough dataset reported all zip codes with their respective counties. NYC's five boroughs are technically counties, but three of them use different common names from their county names. Kings county is Brooklyn, Richmond is Staten Island, and New York is Manhattan. These counties were coded to be listed as their more commonly used borough names.

Dataset Construction

Once all datasets had been cleaned and had a zip code associated with every observation, the master dataset could be compiled. The variables fell into two categories. Count-based variables include features that are measured by the number of instances in each zip code (i.e. subway stations, number of crimes). Continuous attribute-based variables are features with a specific value for each zip code (i.e. average property value, population). Datasets for count-based variables were grouped by zip code and then the summarise command was used to count the number of observations in each zip code. All features were added to the master dataset with a series of joins. After repeating this process for all variables, the final dataset had an observation for each zip code and each column was a different feature. The master dataset was compiled and ready for modeling. It included two identifier columns (zip code and borough), ten numeric predictors, and one numeric dependent variable.

Modeling Approach

With all variables being numeric and continuous, linear regression was selected for modeling. This was a logical decision for a few reasons. The goal of this analysis was to quantify how much property values change when neighborhood characteristics change, and linear regression gives us direction (+/–), magnitude, and significance for each variable. This project was aimed at explanation rather than prediction. New zip codes are unlikely to be created, so property value doesn't need to be predicted, but being able to understand which factors matter and how strongly they influence property values across zip codes could help inform future policy decisions. Lastly, aggregated socioeconomic variables (income, education, business counts) typically show stable, monotonic relationships with housing prices. A linear functional form is reasonable and widely used in urban economics and real estate analysis.

The primary model includes all ten predictors and can be written as follows:

$$\text{PropertyValue}_i = \beta_0 + \beta_1(\text{NewBusinesses}_i) + \beta_2(\text{PublicSchools}_i) + \beta_3(\text{NonPublicSchools}_i) + \beta_4(\text{Population}_i) + \beta_5(\text{TransitAccess}_i) + \beta_6(\text{CrimeRate}_i) + \beta_7(\text{GroceryStores}_i) + \beta_8(\text{Education}_i) + \beta_9(\text{Unemployment}_i) + \beta_{10}(\text{Income}_i) + \epsilon_i$$

*i represents each individual zip code. The data is at zip-code-level.

After running this regression, we wanted to identify if property values in different areas of the city are impacted by different features. The same model was repeated five times with data from one borough each time.

Results

Primary Model Results

In the Model 1 of multiple linear regression, six out of the ten features; *num_subway_stations, num_grocery_stores, num_new_businesses_2024, num_crimes_2024, income, and bachelor,* have statistical significance impacts on property value. Significance being determined by $p > 0.5$. Two features in particular, *income* and *num_grocery_stores,* had the highest level of significance. *Income* with a p-value less than 0.001 and t-statistics of 4.949. Income's large t-value and extremely small p-value show that it is a highly significant predictor, meaning higher-income areas are strongly associated

```
Residuals:
    Min      1Q  Median      3Q     Max
-726439 -139513    9440  116980 1817198

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)            8799.524 143887.335   0.061  0.95131
num_subway_stations   24000.948  12079.447   1.987  0.04864 *
num_grocery_stores     2986.848   1054.710   2.832  0.00522 **
num_new_businesses_2024 -4227.651  2037.929  -2.074  0.03964 *
population               -3.674      1.920  -1.913  0.05750 .
num_public_schools    -7072.377   4089.403  -1.729  0.08566 .
num_non_public_schools  6510.169  3682.768   1.768  0.07901 .
num_crimes_2024        2599.043   1312.630   1.980  0.04942 *
income                    5.373      1.086   4.949 1.87e-06 ***
bachelor               4150.589   2079.110   1.996  0.04759 *
unemployed             9851.600   9375.580   1.051  0.29495
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 284800 on 160 degrees of freedom
Multiple R-squared:  0.5813,    Adjusted R-squared:  0.5551
F-statistic: 22.21 on 10 and 160 DF,  p-value: < 2.2e-16
```

Model 1: Primary Model

with higher property values, and this effect is not due to random change. With the positive coefficient of 5.373, as income increases by $1, we would expect to see a $5.37 increase in property value. As well as for *num_grocery_stores,* with a p-value less than 0.01 and t-statistic of 2.832, this also indicates a high level of significance not due to random chance. The coefficient of 2986.848 indicates a positive correlation between grocery stores and property values. As the number of grocery stores increases by one, property value increases by $2986.85. With these highly positive correlated results, we would reject the null-hypothesis that income and number of grocery stores do not have an impact on property value. This remains true for the other significant factors named earlier.

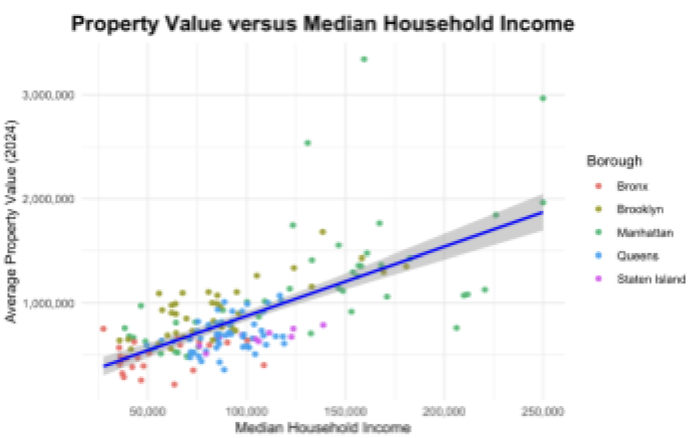


Figure 1

Figure 2

The predictive neighborhood characteristics such as population, number of public schools, number of non-public schools, and unemployment, did not indicate statistical significance.

Model Fit & Error Metrics

For the model's error metrics, the residual standard error indicates the model's prediction differs from the true values roughly around $285K on average. With a Multiple R-squared of 0.5813 explaining 58.13% of the variation in property values. The predictors explain a little more than half of what drives property value differences across New York zip codes. Additionally with the Adjusted R-squared, the model explains 55.5% of variations after adjusting and penalizing for the number of variables. Finally, with a large F-statistic of 22.21 and small p-value of 2.2e-17, this linear model is highly statistically significant indicating this model and some of the predictors are truly associated with property value and not because of random chance.

Borough-Level Results

Additionally, we ran multiple linear regression individually on each of the five New York City boroughs to analyze how the variety of variables impact the differing boroughs. The discovery was that variable importances changes as each borough is analyzed on its own, signifying each borough has its own unique drivers that shape property value. In both Model 2 and Model 3 the boroughs of Staten Island and Bronx, of the ten, there were no predictive features that indicated a statistically significant relationship with property value. In Model 4, Queens had one variable, *population,* that had a negative correlation, where property value decreases by $6.46 as population increases by one. In Brooklyn, *num_grocery_stores*, *num_new_businesses_2024*, had a positive and negative correlation respectively which can be investigated in Model 5. Lastly, in Model 6 in the appendix, Manhattan individually had the highest amount of statistically significant predictors being; *num_subway_stations(positive increase of $117,093.32)*, *num_new_businesses_2024 (decrease of $24,275.04)*, *Income (increase of $7.16)*, and *num_public_schools (decrease of $31,830.29).*

Discussion

In light of the literature review, we saw interesting insights. There were particular variables in our primary Model 1, we were surprised to see factors such as how the number of new businesses and crime affected property value. We expected new business to increase and crime to decrease property value.

For some variables like number of grocery stores per zip code, we expected a positive correlation based on a previously discussed study which finds that an additional grocery store in a census tract increased non-commercial property value about 6.8% (Kole, 2021). This finding aligns with our models predictions: an additional grocery store per zip code is associated with higher property value by $2986.85. We deemed these results reasonable, as New York residents may desire proximity to grocery stores and be willing to pay more due to convenience. Conversely, we expect crime to reduce property value, consistent with prior literature showing that higher crime reduces commercial property value (Lens and Meltzer, 2016). However, our model did identify crime as a strong predictor, showing that one unit increase in reported crime is associated with a $2,599.04 increase in property value, an unexpected result that contrasted with our initial assumptions and existing literature.

As we investigated the individual boroughs, we found it interesting how the significance and direction of certain predictors changes across boroughs. Our initial assumption is that the number of zip codes within each borough may contribute to the variation. Boroughs like the Bronx and Staten Island had no significant predictors and they also have the least amount of zip codes of 23 and 12, respectively, compared to Manhattan (43), Queens (56), and Brooklyn (37). There is possibly not enough zip code observation to find significance between the features in these two boroughs. We believe due to the small sample size, there may be lower statistical power to detect any real effect. With limited data, random variation can dominate the actual relationship making it harder to distinguish real impact among the noise. Understanding property value drivers allows policymakers to make well informed decisions on zoning policies, infrastructure investment, and affordable housing strategies, ensuring that future development supports equitable growth across neighborhoods.

Conclusion

Our analysis is a comprehensive attempt at answering our research question: How do number of new businesses, schools (public, private, and independent), subway stations, grocery stores, population, crime rate, unemployment rate, education rate, and median household income affect property value by zip code in New York City's 5 boroughs?

We discovered that subway stations, grocery stores, new businesses, crime rate, income, and education level, have statistically significant impacts on property value. An increase in subway stations, grocery stores, crime rate, income, and education are all associated with an increase in property value. An increase in new businesses is associated with a decrease in property value.

In the borough specific regression models, only three boroughs had significant variables that impact property value. In Queens, population increase has a negative effect on property value. In Brooklyn an increase in grocery stores increases value while an increased number of new businesses decreases it. In Manhattan, subway stations and income have a positive correlation with property value while new businesses and public schools have a negative correlation with property value.

Our research identified significant findings, however we were limited by potential errors such as sourcing data from multiple platforms which can introduce zip code variability in sampling and location coverage. In future research, we will look to incorporate time-series data that will allow us to build predictive models to predict future property values by existing zip code, as these predictive characteristics change over time.

It is important to capture what drives property value in order to forecast future shifts. The model includes a limited set of ten neighborhood control. Future versions could incorporate measures such as accessibility, school quality, business mix, and environmental amenities including parks, waterfronts, green roofs, and other neighborhood profiles. Spatial statistics techniques like Geospatial Information Systems (GIS) could also be used to measure proximity and identify patterns.

References

Duncombe, W. D., Yinger, J., & Zhang, P. (2014). How Does School District Consolidation Affect Property Values? A Case Study of New York. *Public Finance Review*, *44*(1), 52-79. https://doi.org/10.1177/1091142114524617 (Original work published 2016)

Ellen, Ingrid Gould, et al. “The Impact of Business Improvement Districts on Property Values: Evidence from New York City [with Comments].” *Brookings-Wharton Papers on Urban Affairs*, 2007, pp. 1–39. *JSTOR*, http://www.jstor.org/stable/25067439. Accessed 5 Dec. 2025.

Gray, A. (2025, September 17). New York’s Housing Crisis: Self-Inflicted and Solvable. Vital City. Retrieved November 20, 2025, from https://www.vitalcitynyc.org/articles/new-yorks-housing-crisis-self-inflicted-and-solvable

Ichihara, K., Cohen, J.P. New York City property values: what is the impact of green roofs on rental pricing?. Lett Spat Resour Sci 4, 21–30 (2011). https://doi.org/10.1007/s12076-010-0046-4

Kole, Kyle. *Grocery Stores Raise Property Values: Evidence from FRESH*. University of California, Irvine, 2021. ResearchGate, https://www.economics.uci.edu/files/docs/2021/gradjobmarket/kole_kyle.pdf

Lens, Michael C., and Rachel Meltzer. "Is Crime Bad for Business? Crime and Commercial Property Values in New York City." *Journal of Regional Science*, vol. 56, no. 3, 2016, pp. 442–470. Wiley Online Library, https://doi.org/10.1111/jors.12254.

Data Sources

New York City Department of Consumer and Worker Protection. Issued Licenses. NYC Open Data. https://data.cityofnewyork.us/Business/Issued-Licenses/w7w3-xahh

New York City Department of Health and Mental Hygiene. Modified Zip Code Tabulation Areas (MODZCTA) map. NYC Open Data. https://data.cityofnewyork.us/Health/Modified-Zip-Code-Tabulation-Areas-MODZCTA-Map/5fzm-kpwv

New York City Police Department. NYPD Complaint Data - Current Year to Date. NYC Open Data. https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Current-Year-To-Date-/5uac-w243

NewYork-Demographics.com. ZIP Codes by Population. https://www.newyork-demographics.com/zip_codes_by_population

New York State Department of Agriculture & Markets. Retail Food Stores. data.ny.gov. https://data.ny.gov/Economic-Development/Retail-Food-Stores/9a8c-vfzj

New York State Education Department. K–12 Schools (GIS) dataset. New York State GIS. https://data.gis.ny.gov/datasets/sharegisny::schools-k-12
New York State. New York State ZIP Codes – County FIPS Cross-Reference. data.ny.gov. https://data.ny.gov/Government-Finance/New-York-State-ZIP-Codes-County-FIPS-Cross-Referen/juva-r6g2
Social Explorer. *American Community Survey 5-Year Estimates*. U.S. Census Bureau, accessed 2024. https://www.socialexplorer.com/.
United States General Services Administration. MTA Subway Stations. Data.gov. https://catalog.data.gov/dataset/mta-subway-stations
Zillow. (2025). ZHVI All Homes, Time Series, Seasonally Adjusted, Smoothed ($) for ZIP Code. Zillow Research. https://www.zillow.com/research/data/

Supplementary Information

Model 1: Primary Linear Regression with All Predictors

```
Residuals:
    Min      1Q  Median      3Q     Max 
-726439 -139513    9440  116980 1817198 

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)    
(Intercept)             8799.524 143887.335   0.061  0.95131    
num_subway_stations    24000.948  12079.447   1.987  0.04864 *  
num_grocery_stores      2986.848   1054.710   2.832  0.00522 ** 
num_new_businesses_2024 -4227.651  2037.929  -2.074  0.03964 *  
population                -3.674      1.920  -1.913  0.05750 .  
num_public_schools     -7072.377   4089.403  -1.729  0.08566 .  
num_non_public_schools  6510.169   3682.768   1.768  0.07901 .  
num_crimes_2024         2599.043   1312.630   1.980  0.04942 *  
income                     5.373      1.086   4.949 1.87e-06 ***
bachelor                4150.589   2079.110   1.996  0.04759 *  
unemployed              9851.600   9375.580   1.051  0.29495    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 284800 on 160 degrees of freedom
Multiple R-squared:  0.5813,	Adjusted R-squared:  0.5551 
F-statistic: 22.21 on 10 and 160 DF,  p-value: < 2.2e-16
```

Model 2-6: Linear Regression by Borough

Model 2: Staten Island

```
Call:
lm(formula = avg_property_value_2024 ~ ., data = curr_borough_df)

Residuals:
    1     2     3     4     5     6     7     8     9    10    11    12 
-5551 -1462 -2673 -4224 -4666 -5786 14034  3627 15103 -6814 -2842  1254 

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)             -9.941e+05  8.677e+05  -1.146    0.457
num_subway_stations      9.029e+04  5.056e+04   1.786    0.325
num_grocery_stores       1.168e+04  7.190e+03   1.624    0.351
num_new_businesses_2024  7.046e+02  3.916e+03   0.180    0.887
population               3.651e+00  3.431e+00   1.064    0.480
num_public_schools       4.864e+04  3.700e+04   1.314    0.414
num_non_public_schools  -4.328e+04  3.676e+04  -1.177    0.448
num_crimes_2024         -1.356e+04  1.048e+04  -1.294    0.419
income                   1.565e+01  9.346e+00   1.675    0.343
bachelor                -2.074e+04  1.839e+04  -1.128    0.462
unemployed               5.030e+04  3.855e+04   1.305    0.416

Residual standard error: 24650 on 1 degrees of freedom
Multiple R-squared:  0.9895,	Adjusted R-squared:  0.8845 
F-statistic: 9.424 on 10 and 1 DF,  p-value: 0.2487
```

Model 3: Bronx

```
Call:
lm(formula = avg_property_value_2024 ~ ., data = curr_borough_df)

Residuals:
    Min      1Q  Median      3Q     Max 
-229203  -69904   15047   81088  205575 

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)            819702.933 494482.717   1.658    0.123
num_subway_stations     11584.237  29828.379   0.388    0.705
num_grocery_stores      -1235.763   1720.721  -0.718    0.486
num_new_businesses_2024   830.806   5803.253   0.143    0.889
population                 -2.474      2.931  -0.844    0.415
num_public_schools       3350.229   6019.451   0.557    0.588
num_non_public_schools    605.923  11831.940   0.051    0.960
num_crimes_2024         -1335.293   1982.618  -0.673    0.513
income                      3.852      5.793   0.665    0.519
bachelor               -11178.904   7469.783  -1.497    0.160
unemployed              -4577.955  20088.018  -0.228    0.824

Residual standard error: 148300 on 12 degrees of freedom
Multiple R-squared:  0.4092,	Adjusted R-squared:  -0.08318 
F-statistic: 0.8311 on 10 and 12 DF,  p-value: 0.6099
```

Model 4: Queens

```
Call:
lm(formula = avg_property_value_2024 ~ ., data = curr_borough_df)

Residuals:
    Min      1Q  Median      3Q     Max 
-346192  -85109  -31729   81904  323333 

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)  
(Intercept)            455575.146 247944.762   1.837   0.0728 .
num_subway_stations     -7594.209  14543.668  -0.522   0.6041  
num_grocery_stores       2146.758   1769.833   1.213   0.2315  
num_new_businesses_2024   386.936   3947.079   0.098   0.9223  
population                 -6.458      2.709  -2.383   0.0214 *
num_public_schools       6117.570   7905.315   0.774   0.4431  
num_non_public_schools  11842.662  11971.417   0.989   0.3278  
num_crimes_2024          1425.110   1580.921   0.901   0.3722  
income                      3.069      1.854   1.655   0.1048  
bachelor                 -870.815   2255.997  -0.386   0.7013  
unemployed               5635.761  12560.104   0.449   0.6558  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 150000 on 45 degrees of freedom
Multiple R-squared:  0.2867,	Adjusted R-squared:  0.1282 
F-statistic: 1.809 on 10 and 45 DF,  p-value: 0.08648
```

Model 5: Brooklyn

```
Call:
lm(formula = avg_property_value_2024 ~ ., data = curr_borough_df)

Residuals:
    Min      1Q  Median      3Q     Max 
-289508  -87175  -12710  106844  334319 

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)  
(Intercept)            387501.079 288030.142   1.345   0.1901  
num_subway_stations     -7104.137  12989.372  -0.547   0.5891  
num_grocery_stores       2575.497   1106.647   2.327   0.0280 *
num_new_businesses_2024 -4637.711   1925.911  -2.408   0.0234 *
population                 -3.303      3.034  -1.089   0.2863  
num_public_schools      -1370.332   5392.147  -0.254   0.8014  
num_non_public_schools   5818.917   2935.794   1.982   0.0581 .
num_crimes_2024           613.338   2004.271   0.306   0.7620  
income                      2.991      2.700   1.108   0.2780  
bachelor                 6937.871   5047.664   1.374   0.1810  
unemployed               7710.727  22606.085   0.341   0.7358  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 159400 on 26 degrees of freedom
Multiple R-squared:  0.7459,	Adjusted R-squared:  0.6482 
F-statistic: 7.633 on 10 and 26 DF,  p-value: 1.496e-05
```

Model 6: Manhattan

```
Call:
lm(formula = avg_property_value_2024 ~ ., data = curr_borough_df)

Residuals:
    Min      1Q  Median      3Q     Max 
-582639 -231363  -30964  193460 1186789 

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)   
(Intercept)            286790.023 590625.443   0.486  0.63058   
num_subway_stations    117093.317  38457.600   3.045  0.00463 **
num_grocery_stores       7542.387   4401.676   1.714  0.09629 . 
num_new_businesses_2024 -24275.040   8738.846  -2.778  0.00908 **
population                 -7.854      6.481  -1.212  0.23444   
num_public_schools     -31830.924  14308.880  -2.225  0.03329 * 
num_non_public_schools  27780.568  25381.609   1.095  0.28190   
num_crimes_2024          9447.968   4780.275   1.976  0.05678 . 
income                      7.160      2.874   2.491  0.01810 * 
bachelor                -5777.819   9888.402  -0.584  0.56311   
unemployed              13459.667  22881.444   0.588  0.56050   
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 415100 on 32 degrees of freedom
Multiple R-squared:  0.6651,	Adjusted R-squared:  0.5604 
F-statistic: 6.355 on 10 and 32 DF,  p-value: 2.682e-05
```

Figure 1:

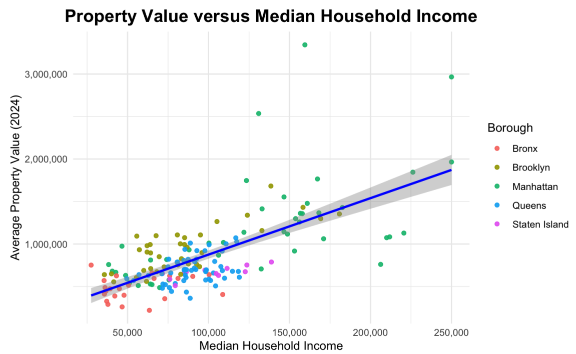
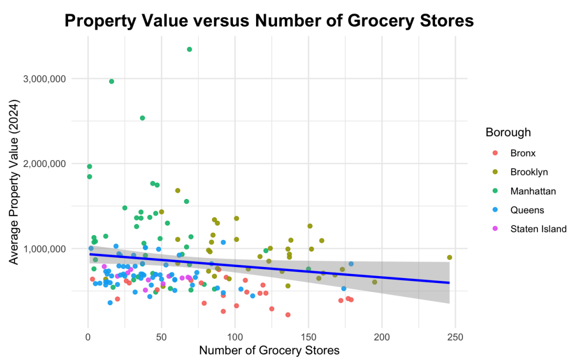
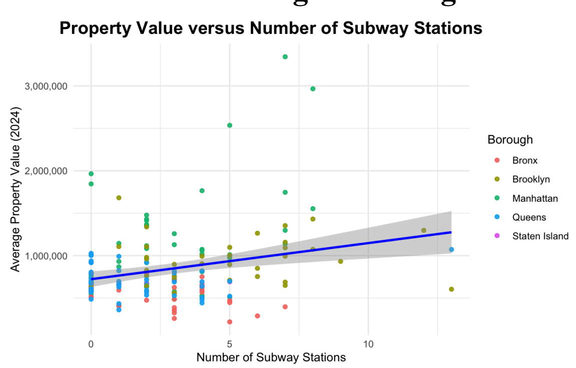
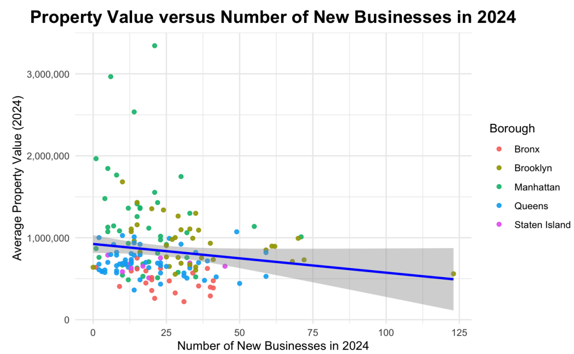


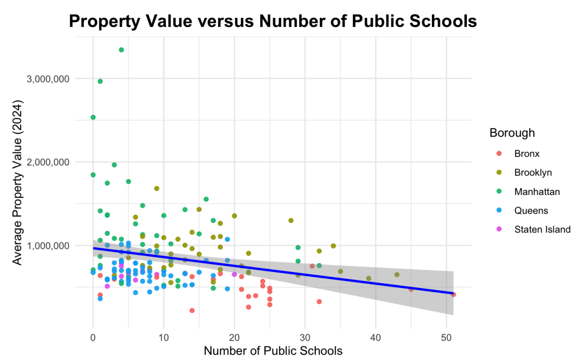
Figure 2:

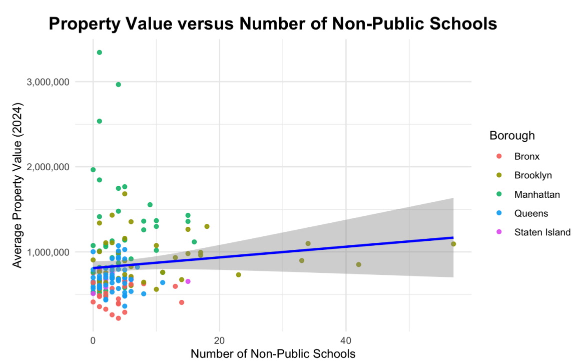


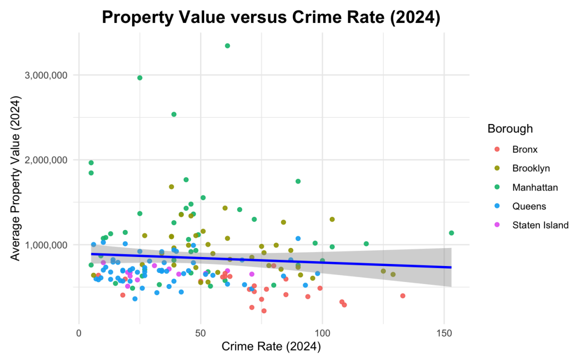
Additional Linear Regression Figures

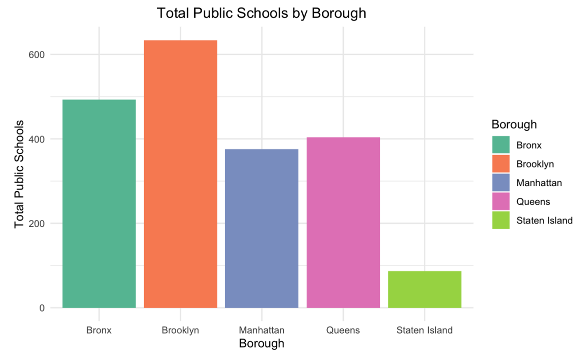


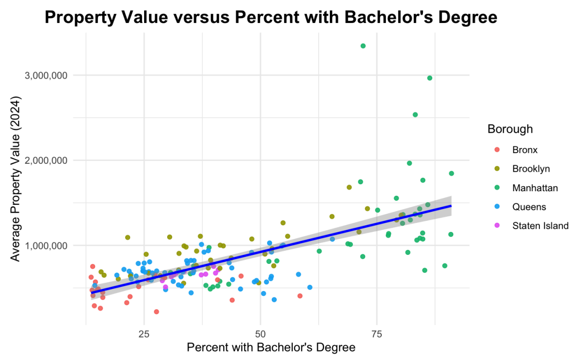


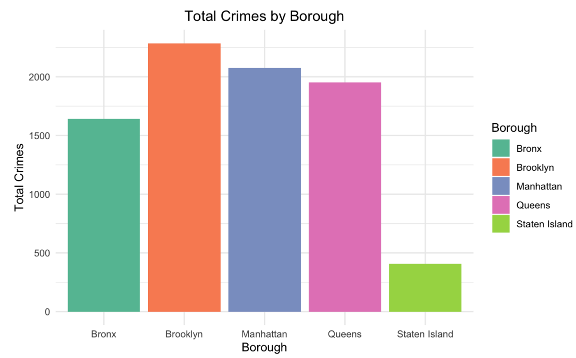


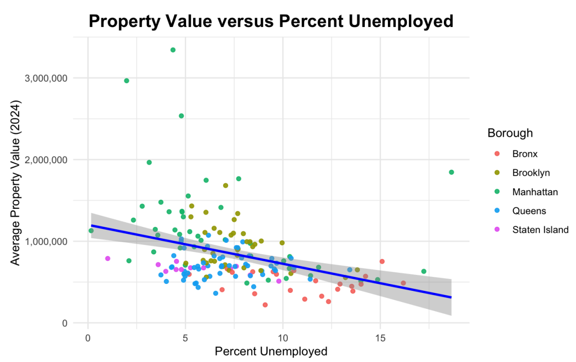


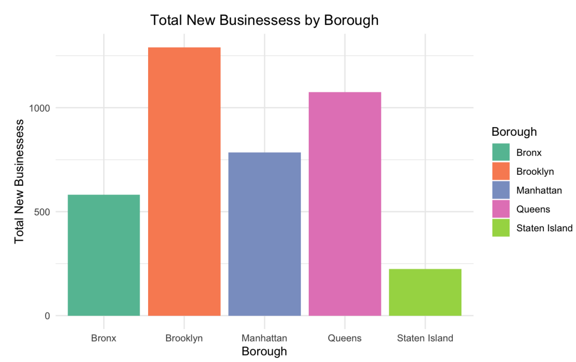




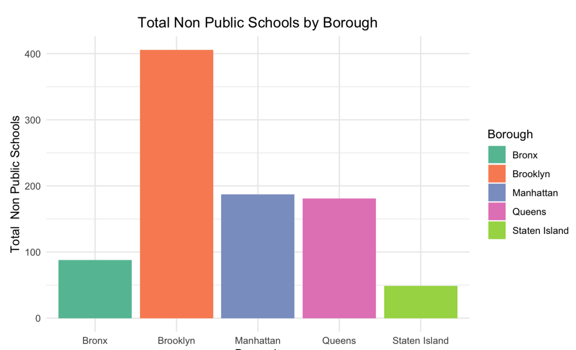


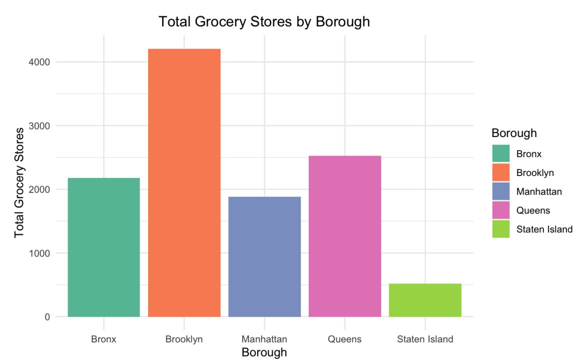


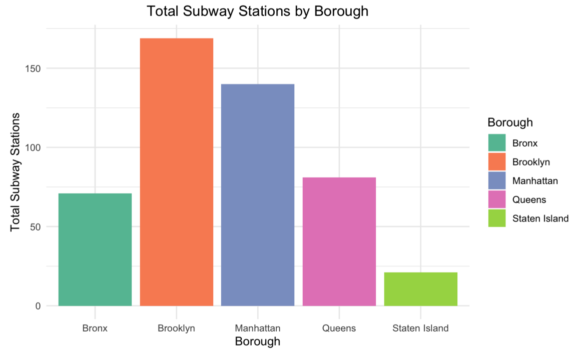




Additional Exploratory Data Analysis







Borough	Number of Zip Codes (Observations)
The Bronx	23
Brooklyn	37
Manhattan	43
Queens	56
Staten Island	12