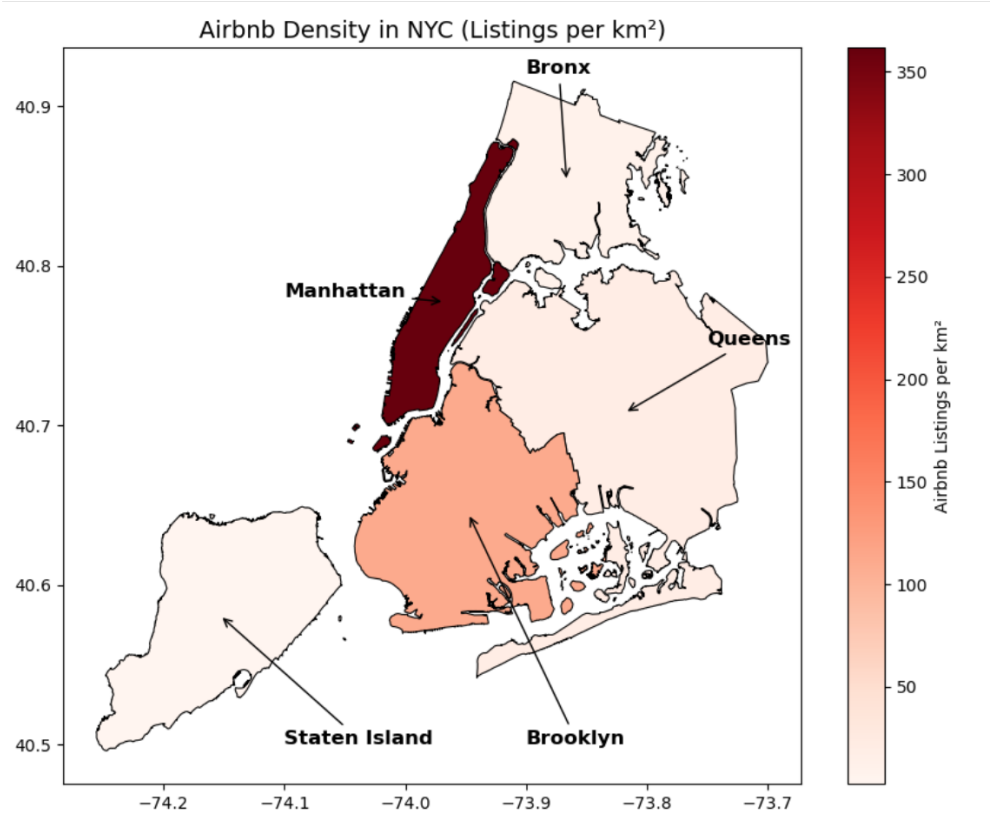
Report 2

Effect of Short-Term Rental on Long-Term Housing Market

The study examines whether short-term rentals, such as Airbnb, affect housing market in New York City. It combines Airbnb data (*AB_NYC_2019.csv*) on neighborhood listings and prices with housing data (*NewYorkHousing.csv*) on property prices and size. The key outcome is housing price, with short term rental (STR) density as the main variable of interest.

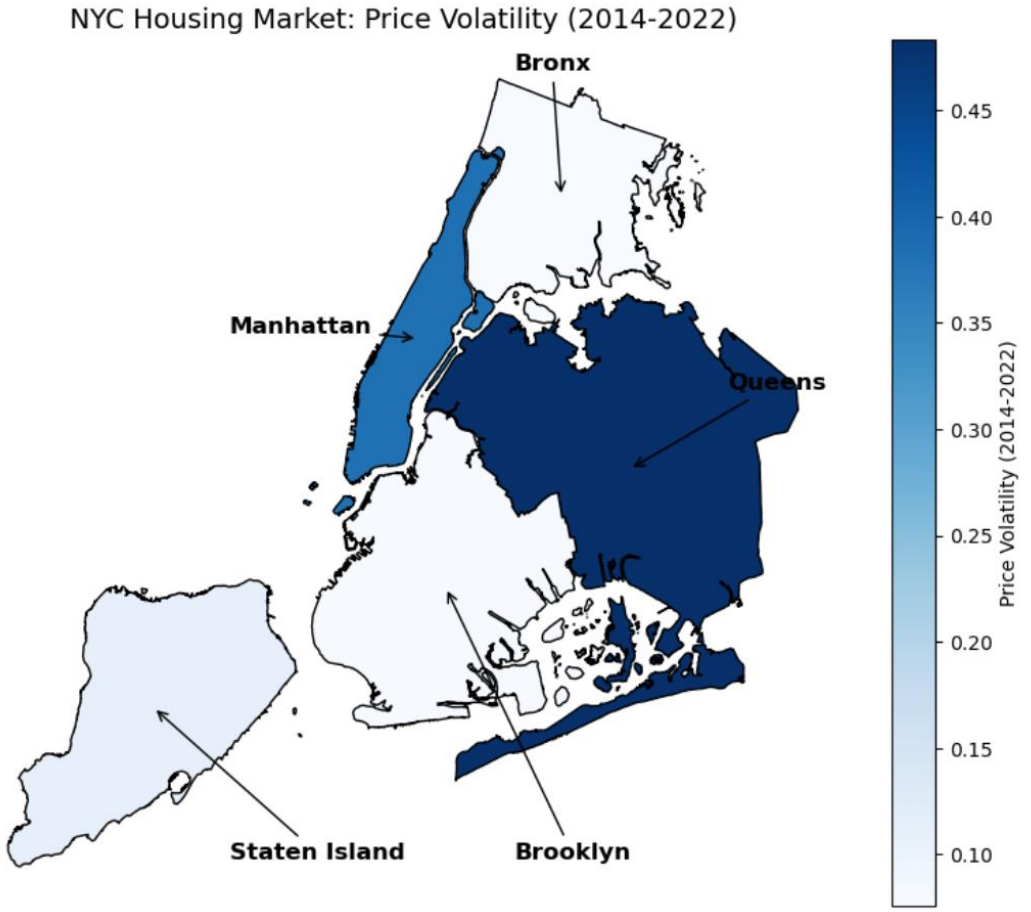
Map Visualization



$$\text{Density}_i = \frac{\text{Number of Airbnb listings}_i}{\text{Area of borough}_i}$$

This map illustrates the Airbnb density per km^2 in New York City. In the supply and demand theory in Economics, if a house is used as short term rental, it is no longer in the on-sale housing market. A decrease in overall supply of housing market will lead to increase in price. We use this map as a base to other maps to see if higher density led to a higher volatility and price growth rate. We also use that data to perform future linear

regression analysis.



$$\sigma_i = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(\Delta P_{i,t} - \overline{\Delta P_i}\right)^2}$$

where

$\Delta P_{i,t}$: annual percentage change in median housing price for borough i in year t;

$\overline{\Delta P_i}$: mean annual price change across all years for borough i ;

T : total number of years in the analysis.

This map shows the price volatility (standard deviation) in different boroughs of New York City. We observe that Queens has the highest volatility, followed by Manhattan, while Manhattan have the higher STR density in previous map. This provides suggestions on testing significance between STR density and price volatility in further linear regression models.

Regression Table

OLS Regression Results: STR Density and Housing Price

				Dependent variable: price
	Baseline	Property Controls	Market Trends	Full Model
	(1)	(2)	(3)	(4)
C(city)[T.Brooklyn]				-83182.917*
				(43796.073)
C(city)[T.Manhattan]				30569.601***
				(11252.755)
C(city)[T.Queens]				84453.663***
				(27343.053)
C(city)[T.Staten Island]				58978.749
				(53334.122)
CAGR			-863066.447*	-36254.436**
			(450457.251)	(14143.483)
Intercept	657573.723***	-209802.272***	-298386.406***	-296130.199***
	(21273.127)	(39134.974)	(43655.630)	(60879.686)
STR_density_per_km2	3326.348***	4387.310***	4311.132***	4794.358***
	(201.722)	(202.576)	(204.244)	(248.717)
Volatility_Weighted_Change			96277.967***	60696.314***
			(18249.983)	(16599.461)
bath		312930.027***	311963.196***	308123.587***
		(18152.506)	(18386.566)	(18420.752)
bed		18858.938	23928.590*	26604.750**
		(12186.486)	(12330.660)	(12356.301)
Observations	4288	3872	3872	3872
R^2	0.060	0.206	0.212	0.213
Adjusted R^2	0.059	0.205	0.211	0.212
Residual Std. Error	1047659.720 (df=4286)	984098.215 (df=3868)	980806.387 (df=3866)	979930.101 (df=3865)
F Statistic	271.912*** (df=1; 4286)	334.618*** (df=3; 3868)	207.722*** (df=5; 3866)	174.731*** (df=6; 3865)
Note:				*p<0.1; **p<0.05; ***p<0.01

Model 1 establish a base relationship between housing price and STR density per km^2, and it shows a high significance to the housing price with $p < 0.01$. Model 2 incorporate housing statistics like number of beds and number of bathrooms. STR density remains significant indicates that STR density correlates with housing price even when controlling some house attributes. Model 3 add compound annual growth rate (CAGR) and price volatility, and it shows that both CAGR and volatility shows a

significance to housing prices. Finally, model 4 introduces dummy variables, and use Bronx as the base. We observe that STR density is still significant to housing price. Manhattan and Queens have significantly higher price than the Bronx, while Staten Island don't.

Generally, even we introducing control variables and dummy variables, there is still a high significance between STR density and housing price, with a p value smaller than 0.01. This implies that short term rentals density difference has a relationship with housing price change, which fit our main research topic.

Next Step

Fitting more linear regression models using different response variables, such as CAGR or volatility, to further investigate the relationship between them and STR density.

Then, apply same models to different cities to prove the universality.