Airbnb & Zillow Data Analysis

Visualizations, Key Insights and Conclusion

Problem Statement

We are to consult a real estate company that works on the business model of purchasing properties to rent out for short-term. Working specifically within New York City they have concluded that two bedroom properties are most profitable. Their ask from us is to further the scope and extent of this analysis to help them know which zip codes are the best to invest in.

Objective

Based on available data for New York City, find out zip codes that would yield maximum return on investment?

Assumptions

1. Assume an occupancy rate of 75% and it is constant throughout the year irrespective of the fact that it can change due to holidays.
2. The investor will pay for the property in cash
3. The time value of money discount rate is 0%
4. Assume that there is seasonality in the price and that values depend not only on previous values. Cost of property after June 2017 can be predicted using time series forecasting. Factors like holiday effects, superhost, location, amenities etc. could effect the price variation for short term rentals and hence prediction of rental price can be ignored.
5. Assume extraneous variables like natural calamities, slowdown in economy which can affect the cost of the property can be ignored.
6. Any operating costs such as maintenance etc. are covered by cleaning fee.
7. Tax ramifications of owning an investment property can be ignored.

Approach

My approach is to devise and observe the breakeven point for popular area in New York City using rental return and house appreciation. This is the key metric to serve the objective aforementioned.

$$\text{ROI} = \frac{\text{Expected Revenue}}{\text{Cost}} = \frac{\text{Annual Rent} + \text{House Appreciation}}{\text{Cost}}$$

$$\text{Breakeven Time} = \frac{\text{Cost}}{\text{Expected Revenue}} = \frac{\text{Cost} - \text{Appreciation}}{\text{Daily Rent} * \text{Occupancy} * \text{Time Period}}$$

Where,

Cost	: Purchasing price of a property at that zipcode
Rent price	: Rent payable at that zip code
Occupancy	: Occupancy rate (assumed to be constant at 75%)
Time period	: measured in year (365 days)

Exploratory Data Analysis

A host of new variables make it possible to paint more vivid pictures of a location's future risks and opportunities.

Property value appreciation across NYC over time

We start by looking into property prices across zip codes and how it has been appreciating over time. For investing purposes, we look for zip codes that have higher appreciation in property values suggesting increase in popularity of the neighborhood. This also safeguards our investment, for if the decision goes south we always have the option of resale available to us.

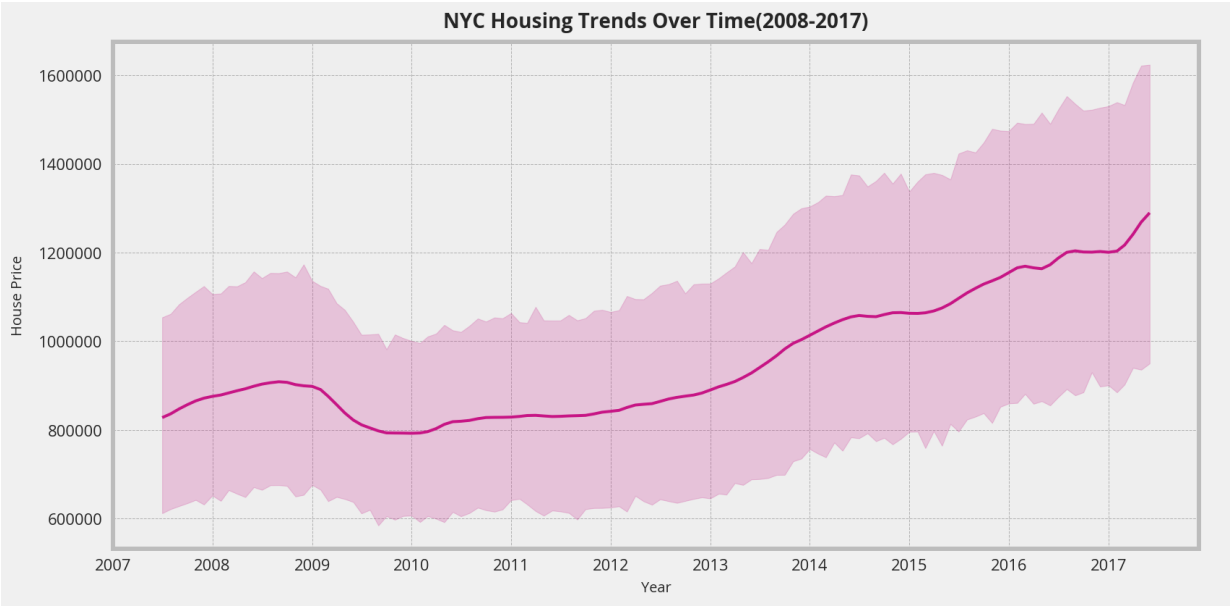

Fig1. NYC Median Housing Price Over Time (2008-2017)

Shading area shows price range within zip codes

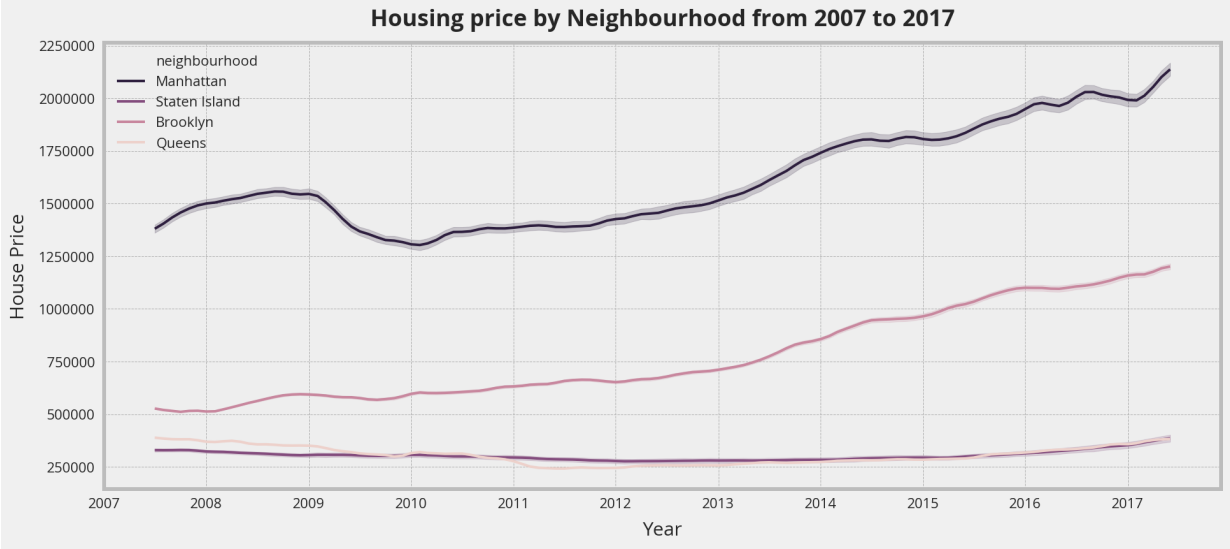

Fig2. NYC Median Housing Price by Neighborhood Over Time (2008-2017)

Housing price Manhattan and Brooklyn have seen major property price increase in the last 7-10 years. Staten Island and Queens have been stagnant for over a decade. If the trends were to continue, Manhattan and Brooklyn also make a sound choice as the re-sale value can be expected to be higher and will protect the investement if things went south in the years to come.

Returns through rental

After securing one aspect of our investment, we then move on to calculate profitability. Our business model of renting property over a short term would be called successful only if the majority of revenue comes from renting. We hence, look at the density and rent distribution of properties across NYC.

The rent distribution in NYC has a mean of approximately 200$/night. We see that Manhattan has a significantly higher number of options and also their rentals have a wider spread. The properties in Brooklyn, however less dense than Manhattan, have a relatively tight spread with their mean around the same value. For more clarity let's look at rent distribution for each neighborhood.

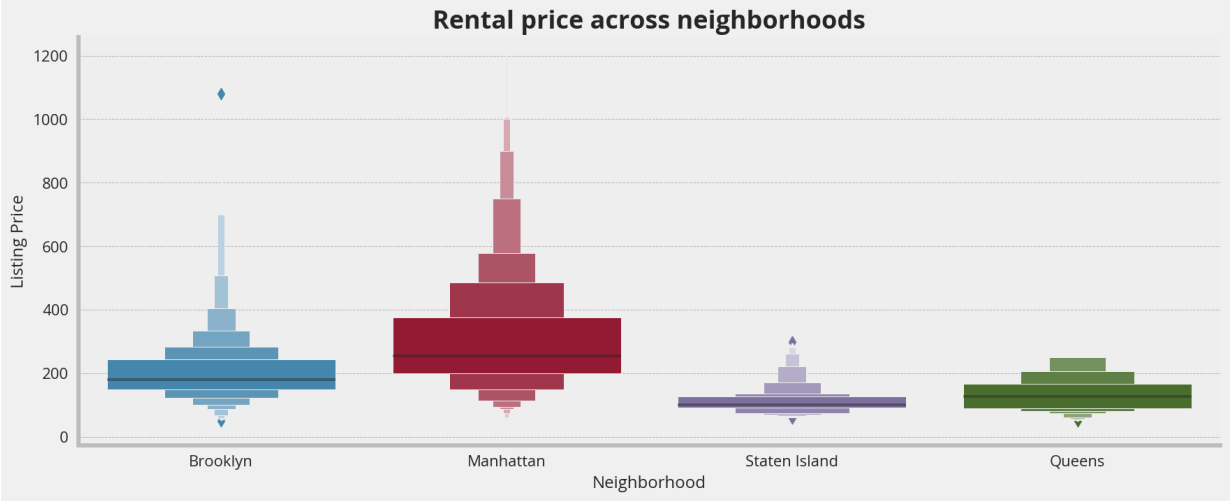

Fig3 Airbnb rental price within neighborhood in 2019

Manhattan, owing to its number of offerings, has a distribution with even its lower quartile range exceeding the mean rental of others. It should be noted at this point that Brooklyn has the same rent distribution as Queens. This means that although Queens has a significantly lesser options for rental as compared to Manhattan and Brooklyn, the return in the neighborhood is decent. Let's get more granular in this and look at the exact zip codes for accurate estimates.

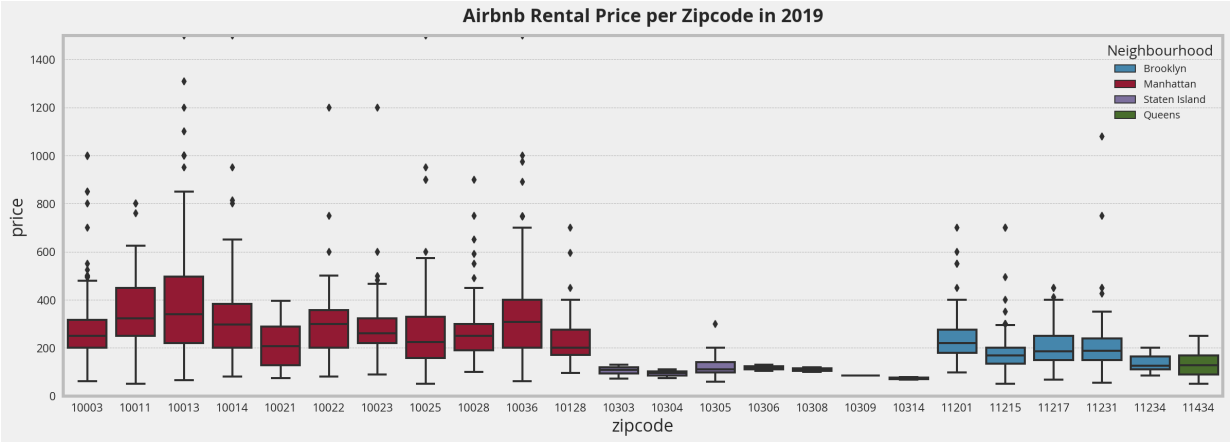

Fig4 Airbnb rental price aross zip code in 2019

As expected, in addition to expected zip codes from Manhattan which was expected, we see here some zip codes from Brooklyn and Queens with impressive rental distributions. This boils our analysis down to some certain zip codes.

Property density across NYC

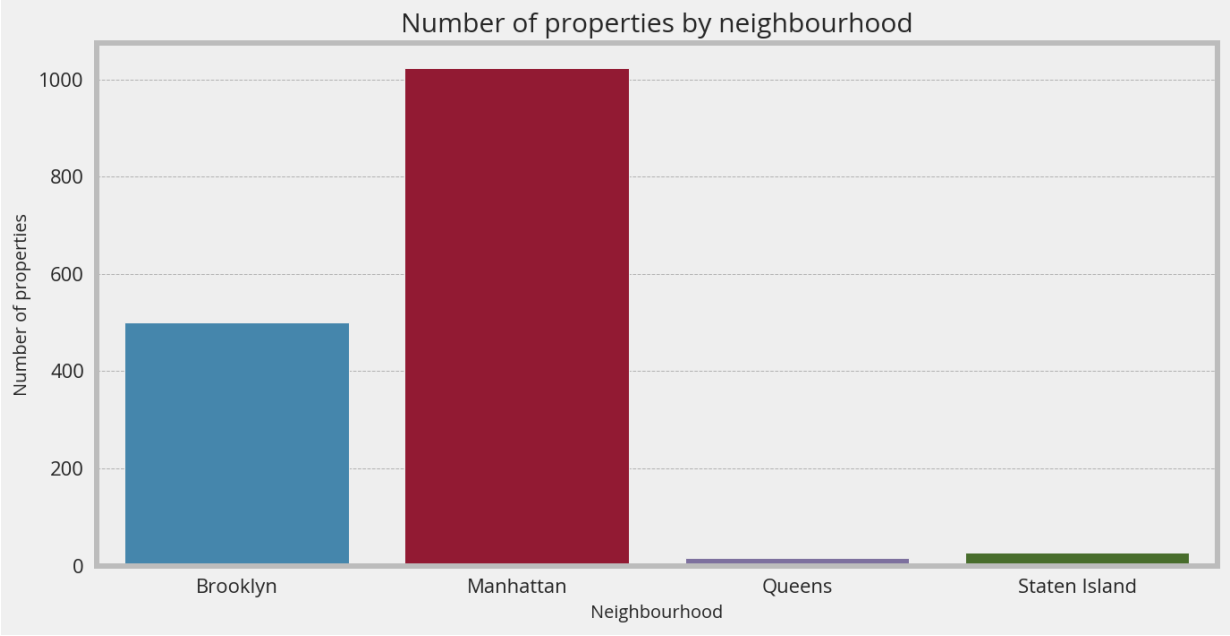

Fig5 Number of 2-bedroom Listings on Airbnb across Neighborhoods

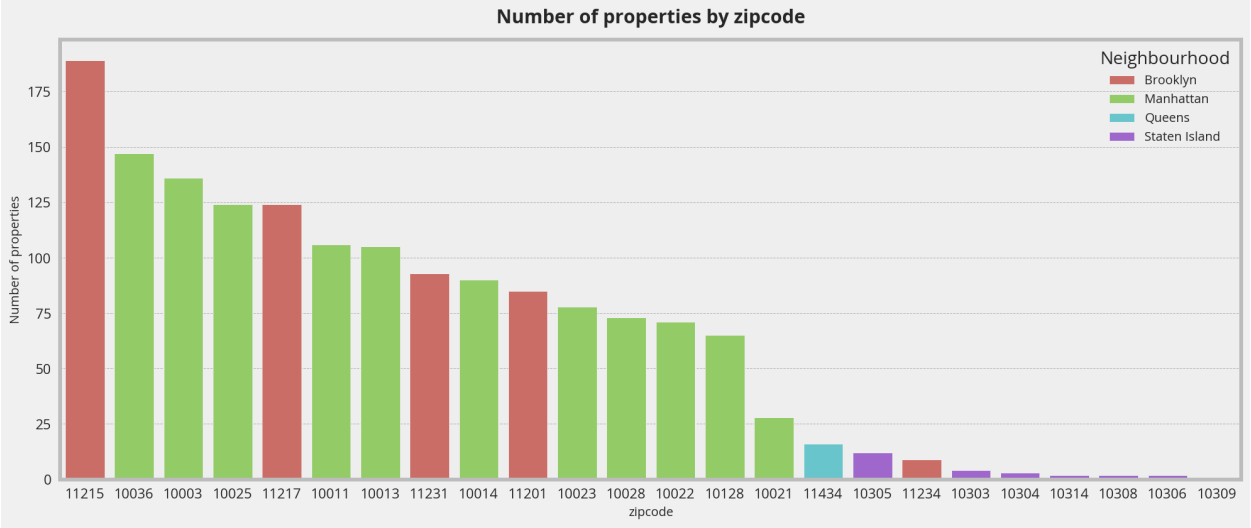

Fig6 Number of 2-bedroom Listings on Airbnb by zip codes

We revisit neighborhood density to confirm our judgment until now. To prove our theory of zip codes worth recommending based on rental return expected, we should confirm if there are sufficient options for us to consider buying.

Brooklyn having higher mean density gives an indication of more options to consider. This is healthy given the price and rent balance it holds as compared to Manhattan. Queens and Staten Island also show-up with a zip code each of significant offerings.

Break Even Period Analysis

Until now we have been looking at the options qualitatively. We now substantiate our judgment for recommending exact zip codes with quantitative backing.

$$\text{ROI} = \frac{\text{Expected Revenue}}{\text{Cost}} = \frac{\text{Annual Rent} + \text{House Appreciation}}{\text{Cost}}$$

$$\text{Breakeven Time} = \frac{\text{Cost}}{\text{Expected Revenue}} = \frac{\text{Cost} - \text{Appreciation}}{\text{Daily Rent} * \text{Occupancy} * \text{Time Period}}$$

Where,

Cost	: Purchasing price of a property at that zipcode
Rent price	: Rent payable at that zip code
Occupancy	: Occupancy rate (assumed to be constant at 75%)
Time period	: A year (365 days)

Using the above mentioned formula, we first calculate the break-even time without accounting for house appreciation. Following are the observations

Revenue Outlook by Short-term Rent Out in 2019

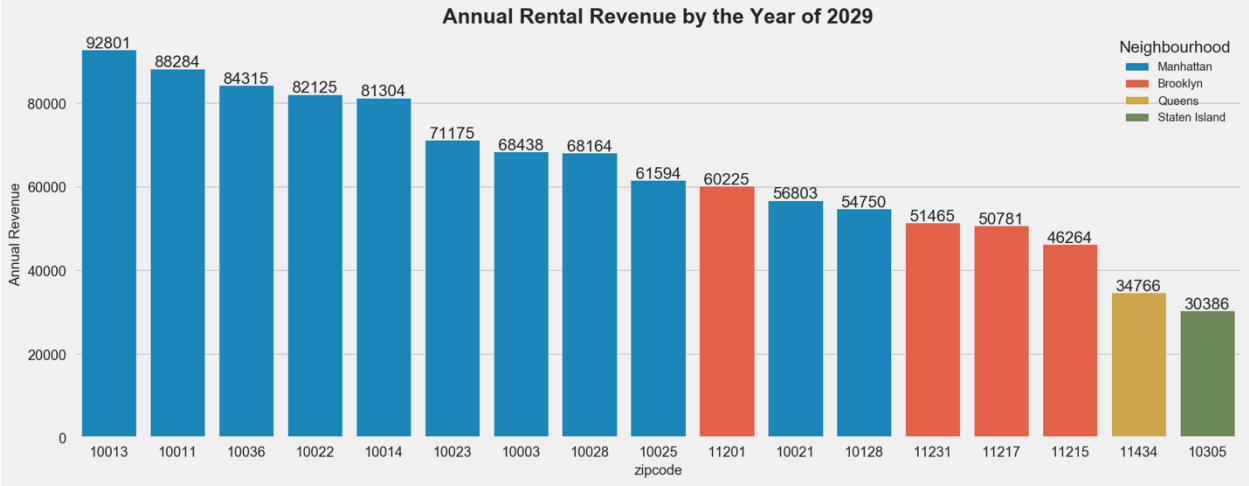

Fig7 Annual Gross Rent Income across Zip codes

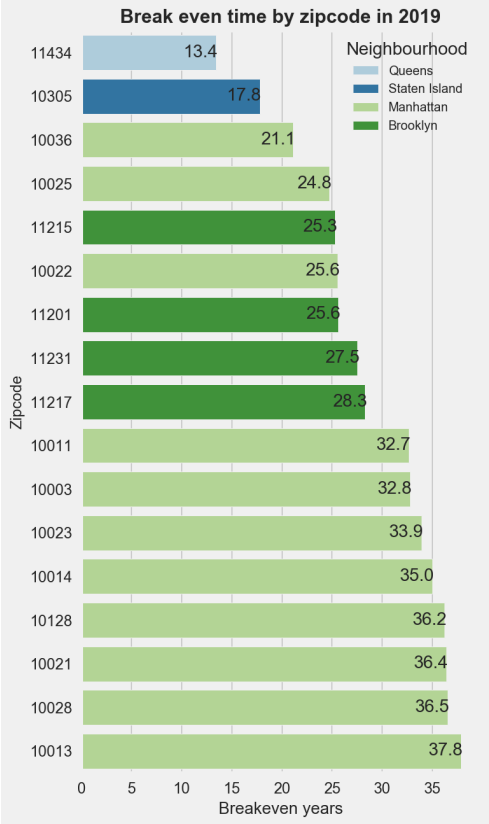

Fig8 Breakeven time in 2019

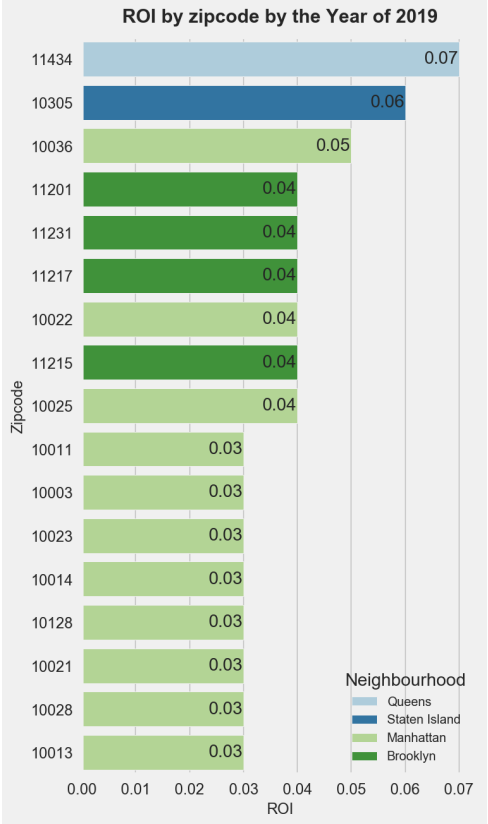

Fig9 ROI vs zip codes

We see from the above graph the zip codes that have the lowest breakeven time duration. 11434 (Queens) and 10305 (Staten Island) top the chart with one zip code each. However, those two zip codes yield the lowest annual revenue. Considering as real estate company, we must ascertain what is the realistic income the target property can generate on a sustainable basis year in and year out.

Let's see what the landscape looks like with house appreciation.

ROI Outlook by Rent and House Appreciation After 10 Years

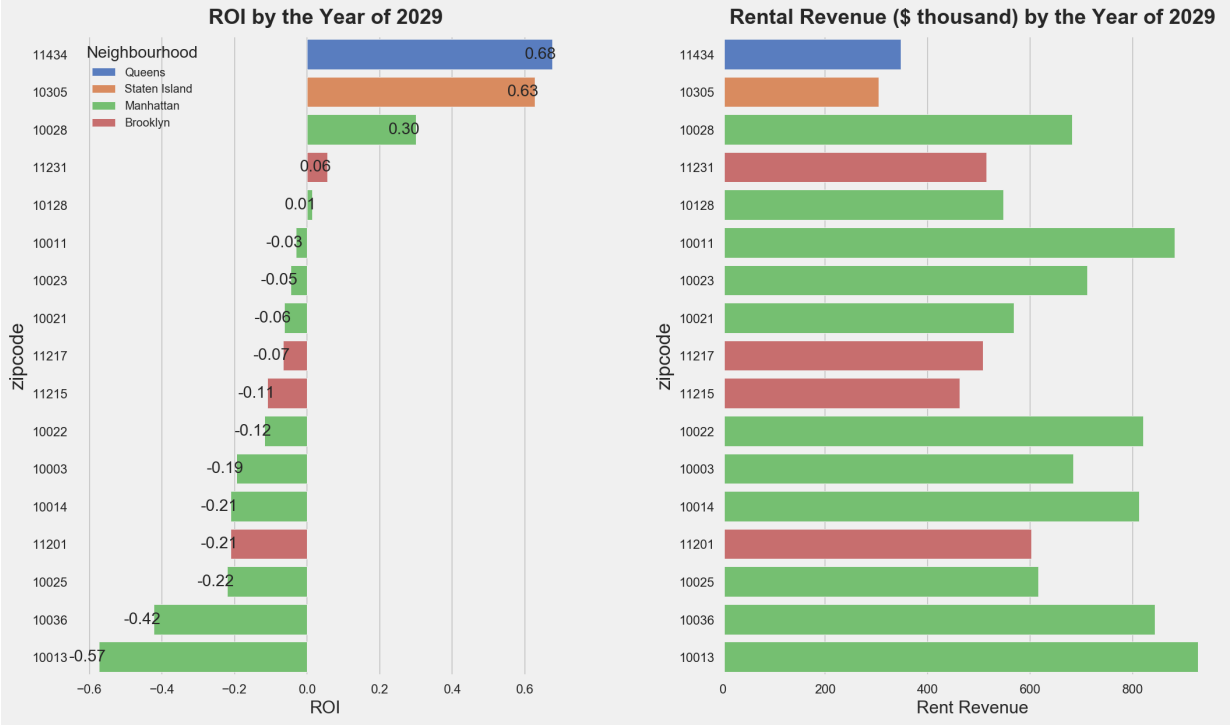

Fig10 ROI & Rental Income after 10 years

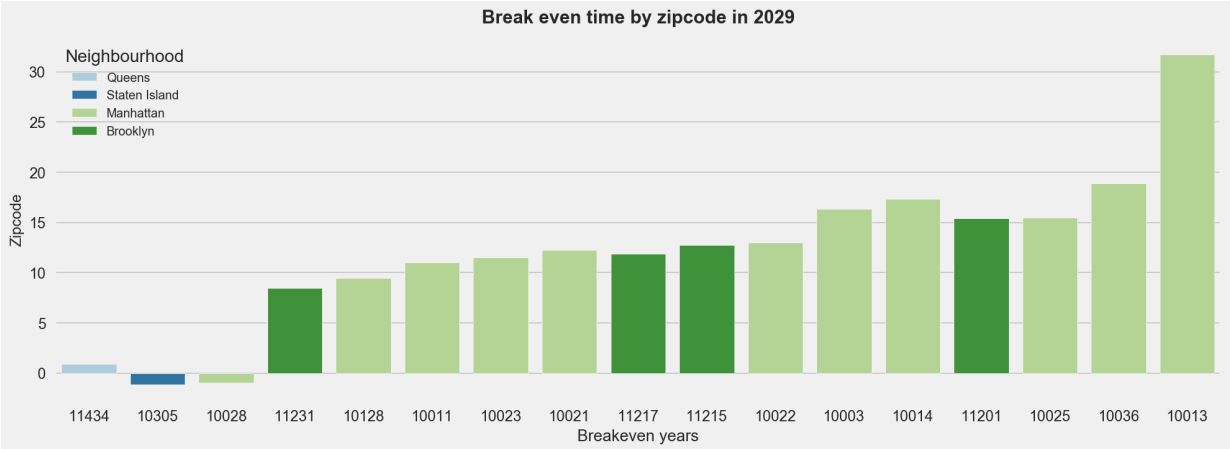

Fig11 Breakeven Years by 2029

Two zip codes from Manhattan (10028, 10128) jump up into top5 because of their property as we include house appreciation, one of them has started to be profitable before 2029. A zip code each from Brooklyn (11231) remains in the top5 with promising annual income. 11434 (Queens) and 10305 (Staten Island) continue being at the top due to relatively low value of their property.

Conclusions drawn from analysis

Top 5 location choices: 10028, 10128, 11231, 10305, 11434

Zip code 10028 in Manhattan is the best zipcodes to purchase 2-bedroom properties. The Average revenue ($68,164/year per property) is one of the highest (see Figure 7). It has 73 two-bedroom listings on Airbnb (not the highest and far from the lowest); and it takes shortest years to break even in Manhattan area.
10128 in Manhattan is also a safe choice, with fast breakeven period and less cost.

Zip code 11231 in **Brooklyn** is also a very good location to purchase 2 bedroom properties.
It is one of the less expensive options (avg. cost of 1.4M). Also, it has a higher number of 2-bedroom properties listed on Airbnb.

11434(Queens) and 10305(Staten Island) are good choices if we focus less on cash flow because their annual rent is at the lower end. Eventhough the Avg. annual revenue obtained by renting through Airbnb is $30K+/property, the cheap purchase cost offsets the disadvantage of reaping lesser revenue and makes those zipcodes it a great location to invest.

Business Limitations

When it comes to investment in real estate, cash flow matters. More factors should be taken into consideration to ascertain what is the realistic income the target property can generate on a sustainable basis year in and year out. The current and historical income figures matter. Often, house price appreciation is not assured. If primarily focus is property appreciation and not income, we will put investment in risk of housing bubble. There is no real value for real estate if it does not generate income or save a person on rent.

Future steps

1. A flat occupancy rate of 75% can be corrected with more data. Neighborhoods such as Staten Island have low rent prices but way fewer properties listed. Considering it's location, it is probable they have low occupancy rate but due to the static rate used in the analysis, we have to assume its occupancy to be equal to a property in Manhattan or Brooklyn.
   Possible solution: Integrate data from public data sources and from related real estate agencies, to build more reliable occupancy rate model.

2. Data, being old, doesn't confirm the actual landscape of market trends. Possible solution: Create a more robust prediction model for current price.
3. Consider seasonality in Airbnb rental prices whenever available to obtain more robust break even estimates. Weekly and monthly rent is cheaper than nightly rent, so considering longer terms would give us the worst-case scenario.
4. Factors such as transportation, proximity to work location can be used to map the area for long-term rentals; Crime score is an important factor in determining the price. Crime score is freely available online and can be used for deeper analysis.
5. Majority of the tourists book short term rentals in NY city hence holidays (variable pricing can be done) and closeness to tourist locations should be accounted for.