NYC Housing Investment Analysis

Chuchen Xiong

Business Goal -> Assumption -> Recommendations -> Analysis Roadmap -> Next Step

Agenda

Business Goal

Assumption

Recommendations

Analysis Roadmap

Next Step

Business Goal

1 **Key Question**

Which zip codes would generate the most profit on short term rentals within New York City

2 **Question Breakdown**

Who		Why		How	After	When	In	Where
A real estate company	Wants to	Generate profit	By	Purchasing properties to rent out short-term		07/2019		New York City

3 **Factors**

- Cash flow of the company
- Long term or short term profit
- Which type of properties to invest
- When to invest
- Which zip codes to invest

Assumption

New Assumptions

- People rent Airbnb for one night (Use daily price for analysis)
- Availability / Occupancy will not change dramatically in the future
- There is no seasonality of the price per night in the Airbnb

Assumptions in the Problem Statements

- The investor will pay for the property in cash
- The time value of money discount rate is 0%
- All properties and all square feet within each locale can be assumed to be homogeneous (The more square feet, the more revenue and cost)
- Two bedroom properties are the most profitable

Recommendations

Questions		Recommendations		Evidences
Who	Cash flow of the company	Does not have much cash		
Why	Long term or short term profit	Has a large amount of cash	Short term	**10309, 10312, 10307, 10308, and 10310 in Staten Island**
Where	Which zip code to invest		Long term	**11215, 10036, 10003, and 10025 in Manhattan or Brooklyn**
When	When to invest	Current time (July 2019) is the best to invest		
How	Which type of properties to invest	Select 2-bedrooms type Properties to invest		

- Least breakeven period (~20 y)
- The cost ($ 500,000) is five times less than Manhattan

- Decent breakeven period (~30 y), many choices (200+)
- Increase brand awareness
- Low risk

- July usually has lowest cost historically
- Cost will keep increasing

- Two bedroom properties are the most profitable

Analysis Roadmap

Metrics Design

Cost	Cash needed to invest properties
Breakeven Period (Year)	Year to breakeven in certain zip code
Number of Properties	The number of properties in certain zip code

Evaluate

- Ease to Invest
- Profitability
- Potential Expansion
- Investment Risk

Data Preprocessing Highlight

Airbnb Price Data → Price Modification

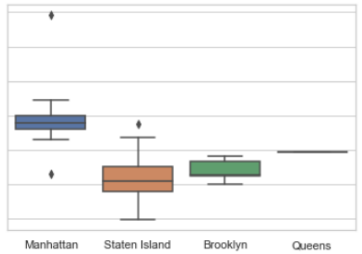
Zillow Cost Data → Cost Forecasting

Price Modification, Cost Forecasting → Data Merging → Cost Imputation → Cleaned Data

Visualization

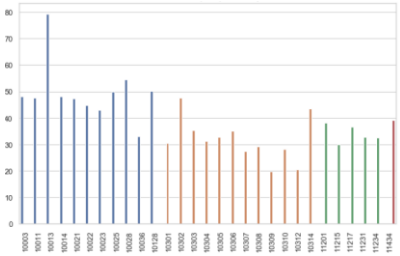
Metrics Pyramid

Borough Analysis



Map Analysis

Zip Code Analysis



Metrics Design

New Metrics

Occupancy = 1 - availability_30 / 30

Breakeven Period Year = Property Cost / (Price per Night * 365 * Occupancy)

Cleaned Data

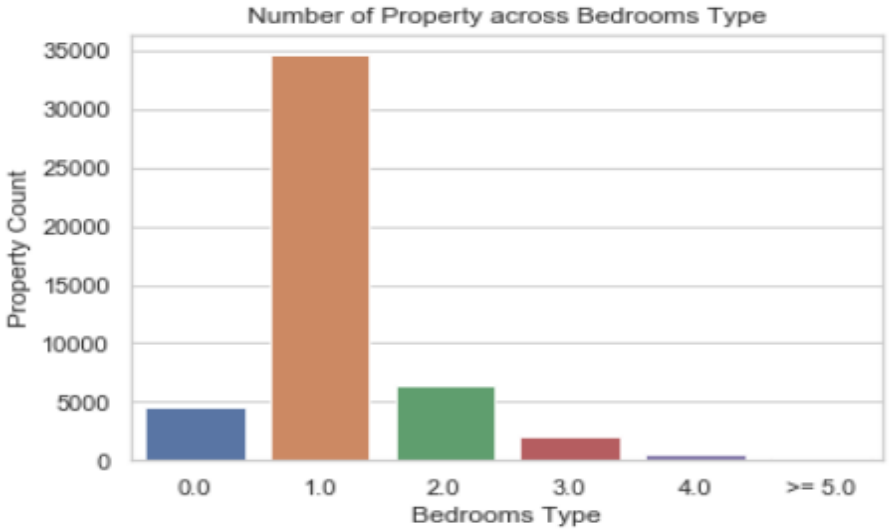
	zipcode	count	price	current_cost	occupancy	borough	breakeven_period_year
2	10003	887	175.0	2482707.0	0.806238	Manhattan	48.209336
9	10011	649	200.0	2855777.0	0.823369	Manhattan	47.512378
11	10013	444	171.5	3613811.0	0.728228	Manhattan	79.275855
12	10014	599	200.0	2838864.0	0.810351	Manhattan	47.989782
18	10021	239	145.0	2059279.0	0.822455	Manhattan	47.308846

Data Preprocessing - Price Modification

Issue

Small Sample Size

The number of two bedrooms only accounts for a small number of rooms



Solution

Enlarge Sample Size

Calculate a modified price using accommodates column for all type of bedrooms

	bedrooms	accommodates	price	price_factor
0	1.0	2.243540	117.136048	0.492211
1	2.0	4.558089	232.614084	1.000000
2	3.0	6.442383	308.177246	1.413396
3	4.0	8.073359	472.727799	1.771216

price_factor = accommodates / 4.558
modified_price = price * price_factor

Data Preprocessing - Cost Forecasting

Issue

Lack of Future Data

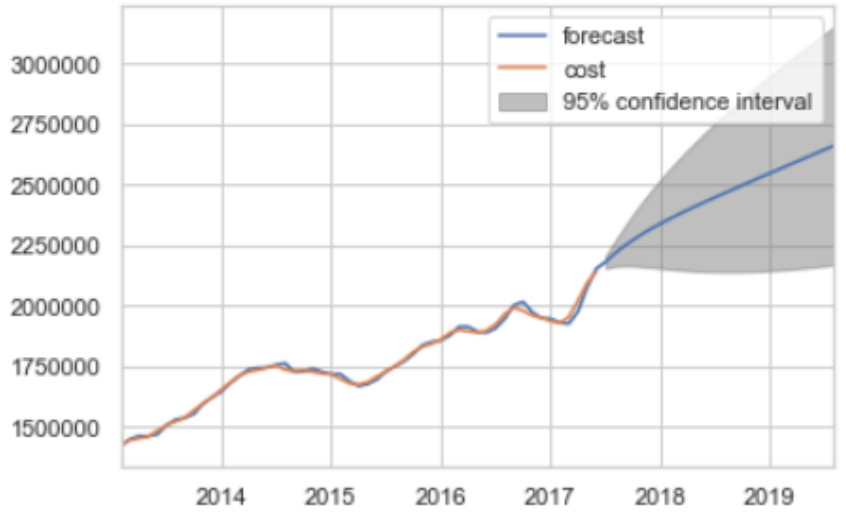
Zillow dataset is up to June 2017, does not have cost data in July 2019

2016-12	2017-01	2017-02	2017-03	2017-04	2017-05	2017-06
1354800.0	1327500	1317300	1333700	1352100	1390000	1431000
1951300.0	1937800	1929800	1955000	2022400	2095000	2142300
1541600.0	1557800	1582900	1598900	1646100	1720500	1787100
2332100.0	2313300	2319600	2342100	2365900	2419700	2480400
1935100.0	1915700	1916500	1965700	2045300	2109100	2147000

Solution

Time Series Forecasting

Use ARIMA model to forecast the cost of 07/2019 for each zip code



Data Preprocessing - Cost Imputation

Issue

Missing Value

Many zip codes do not have cost data (186 zip codes in price data but only 25 in cost data)

```
len(airbnb_cleaned['zipcode'].unique())
```

```
186
```

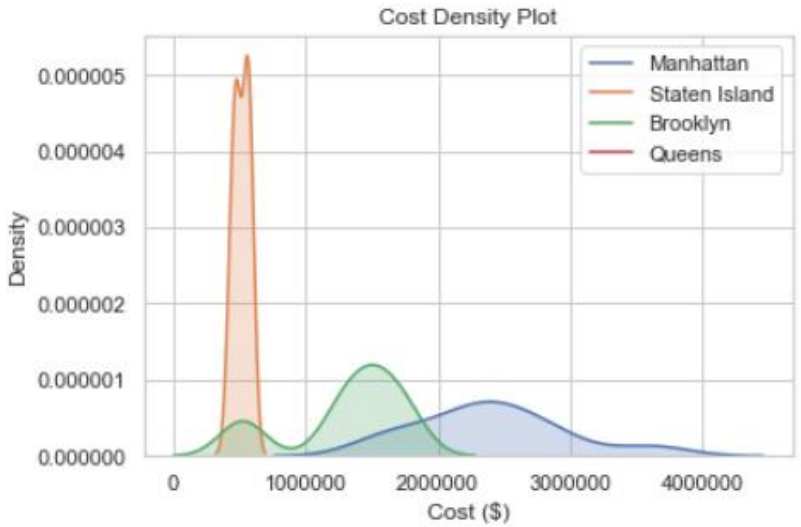
```
len(zillow_cleaned['zipcode'].unique())
```

```
25
```

Solution

Mean Imputation

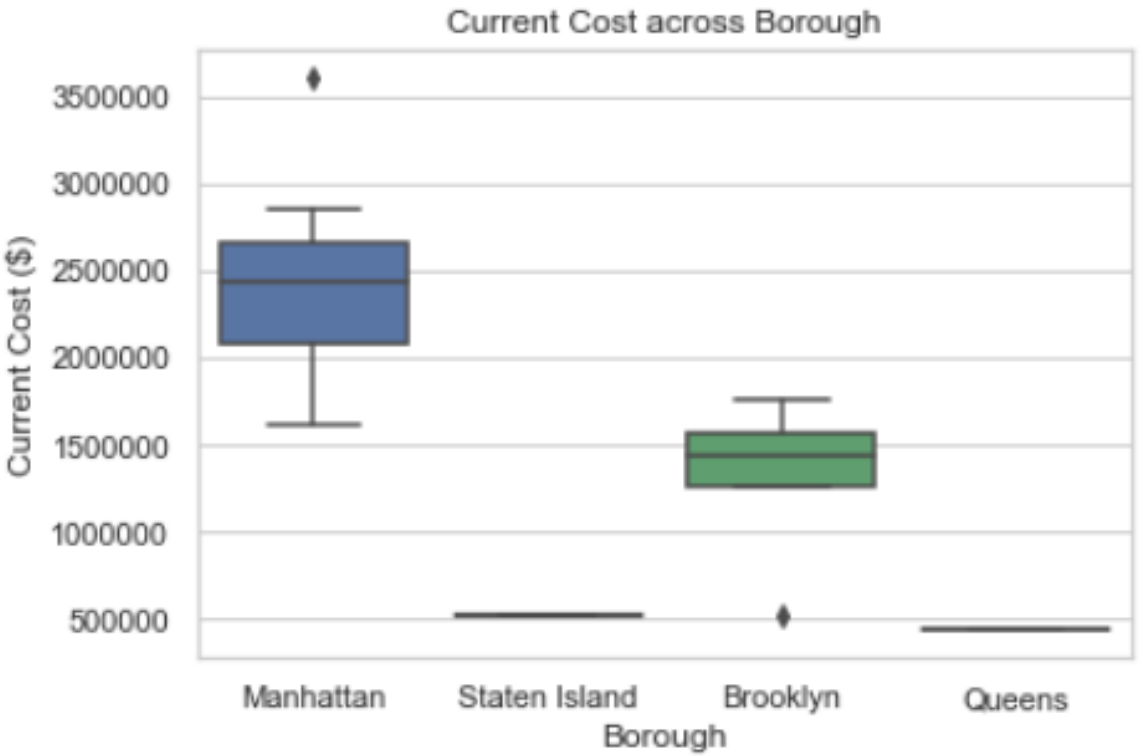
Impute missing Staten Island cost data with the mean of Staten Island

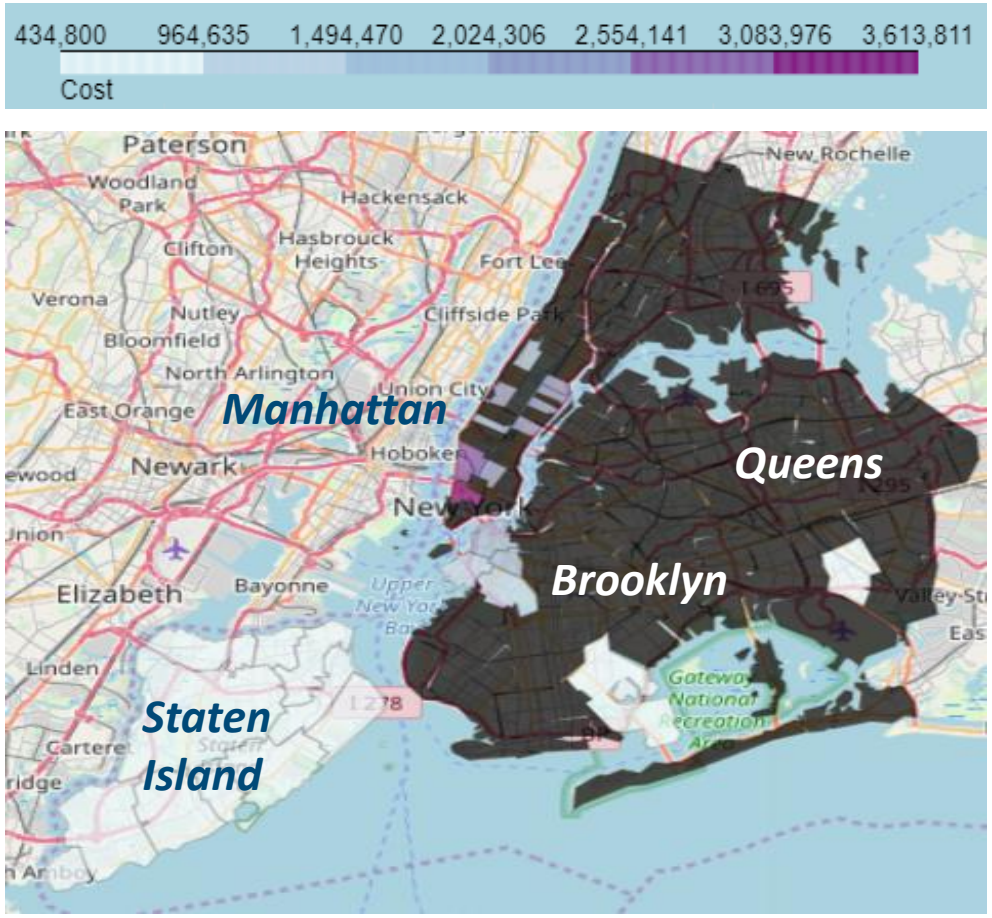


Only cost in Staten Island is highly centered

Visualization - Cost

Borough Analysis ▶ **Map Analysis** ▶ Zip Code Analysis

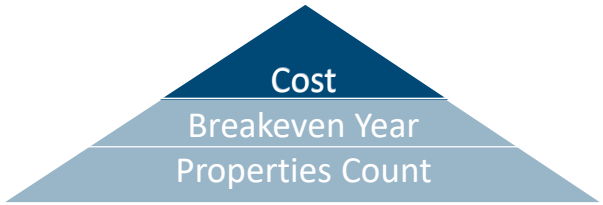




INSIGHT

- ***Lowest cost in Staten Island and Queens***
- ***Highest cost in Manhattan***

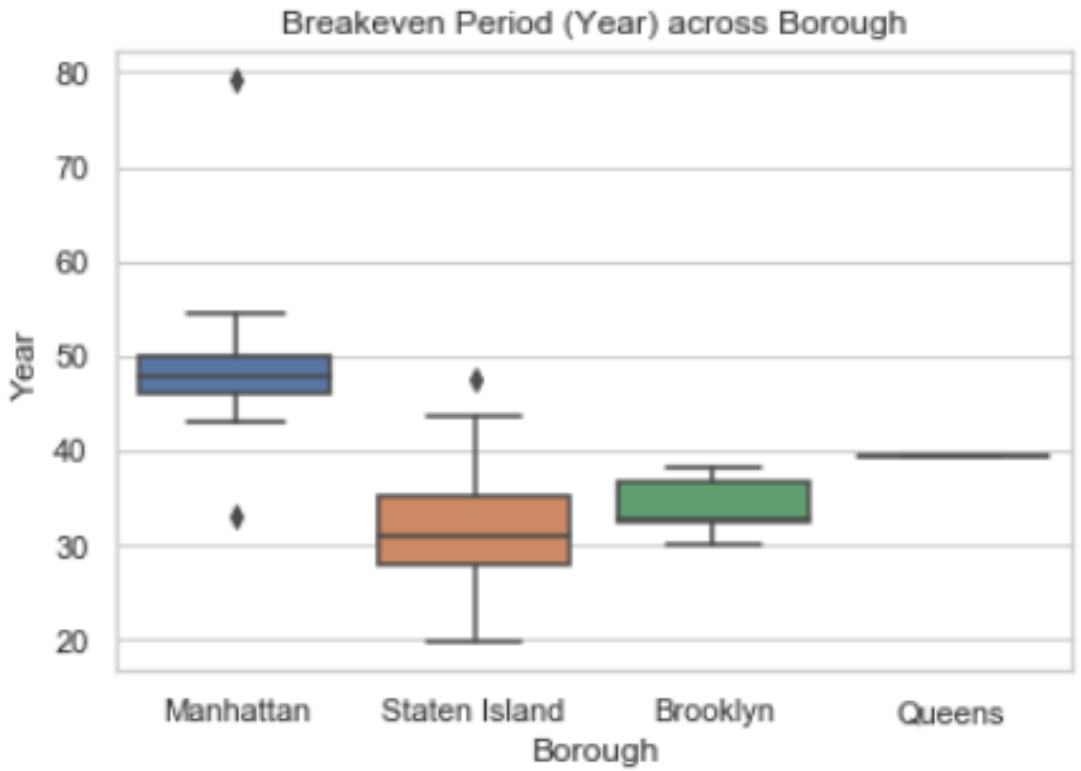
Visualization - Cost



Borough Analysis ▶ Map Analysis ▶ **Zip Code Analysis**

Visualization - Breakeven

Borough Analysis ▶ **Map Analysis** ▶ Zip Code Analysis

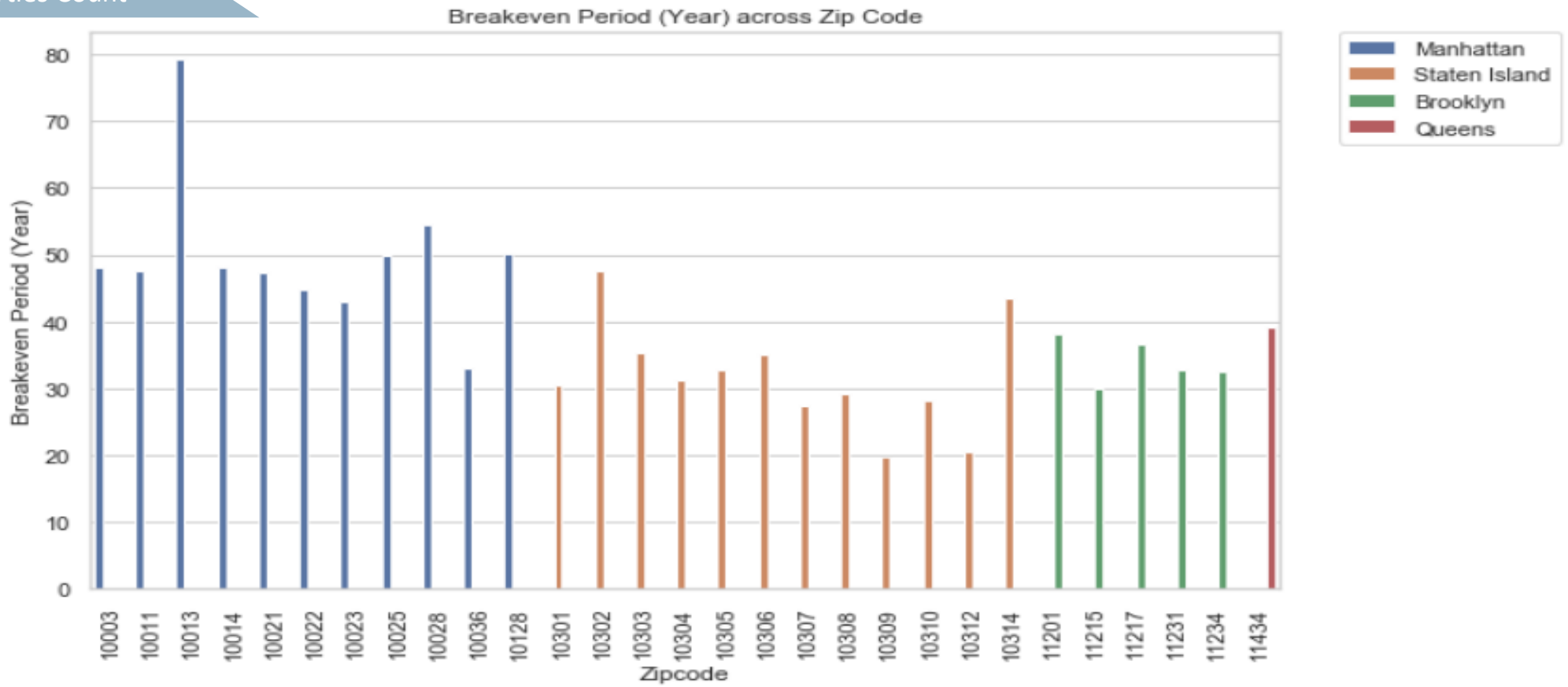


INSIGHT

- ***Least breakeven period in Staten Island***
- ***Longest breakeven period in Manhattan***

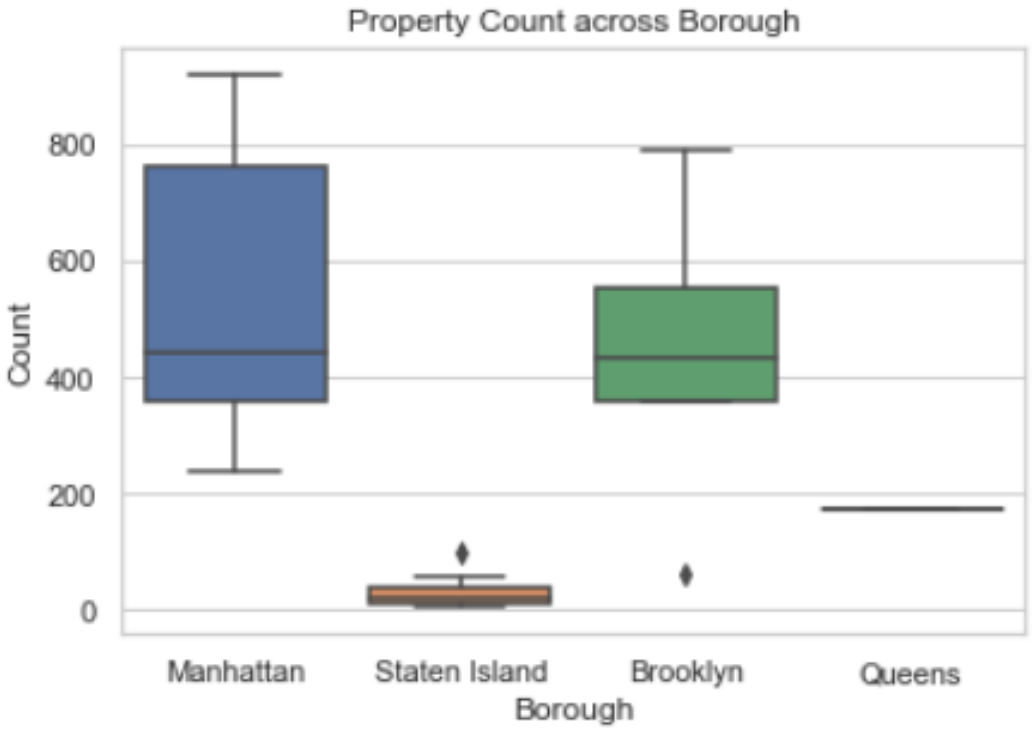
Visualization - Breakeven

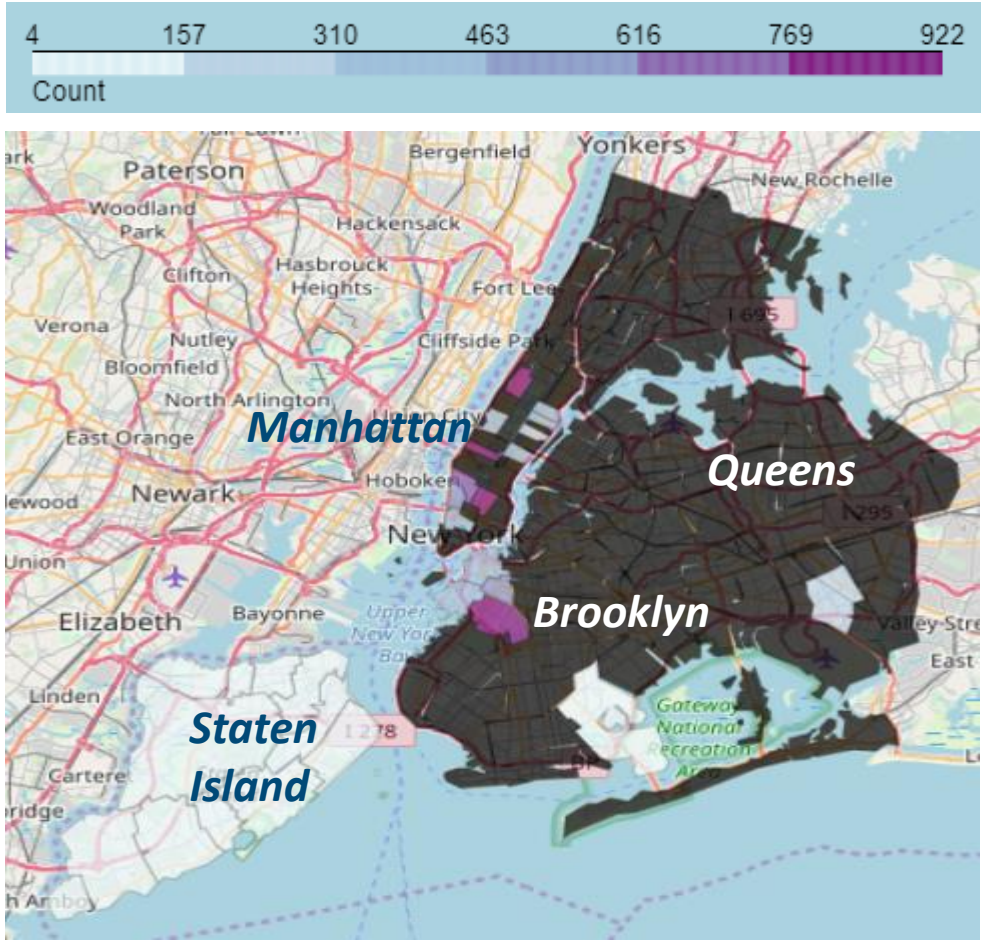
Borough Analysis ▶ Map Analysis ▶ **Zip Code Analysis**



Visualization – Property Count

Borough Analysis ▶ **Map Analysis** ▶ Zip Code Analysis



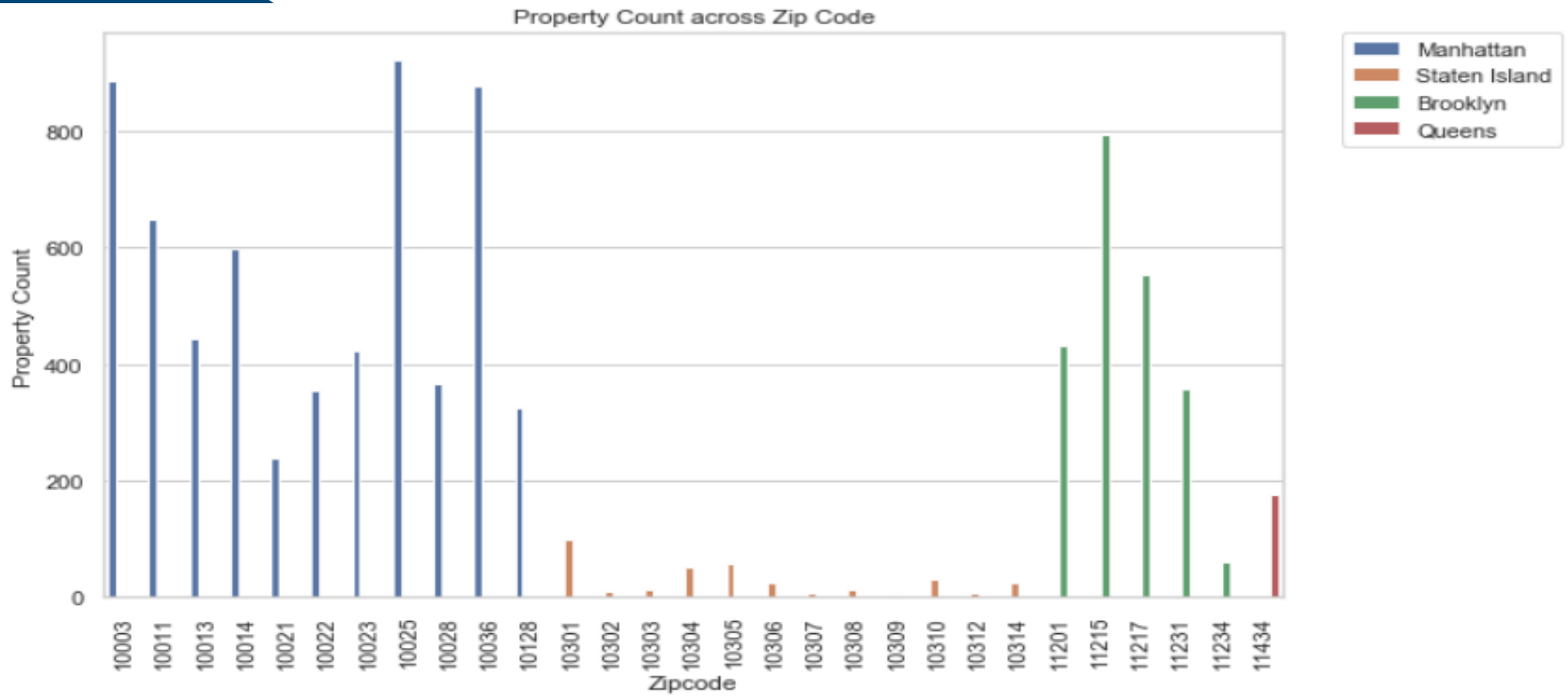


INSIGHT

- ***Low property count in Staten Island***
- ***High property count in Manhattan and Brooklyn***

Visualization – Property Count

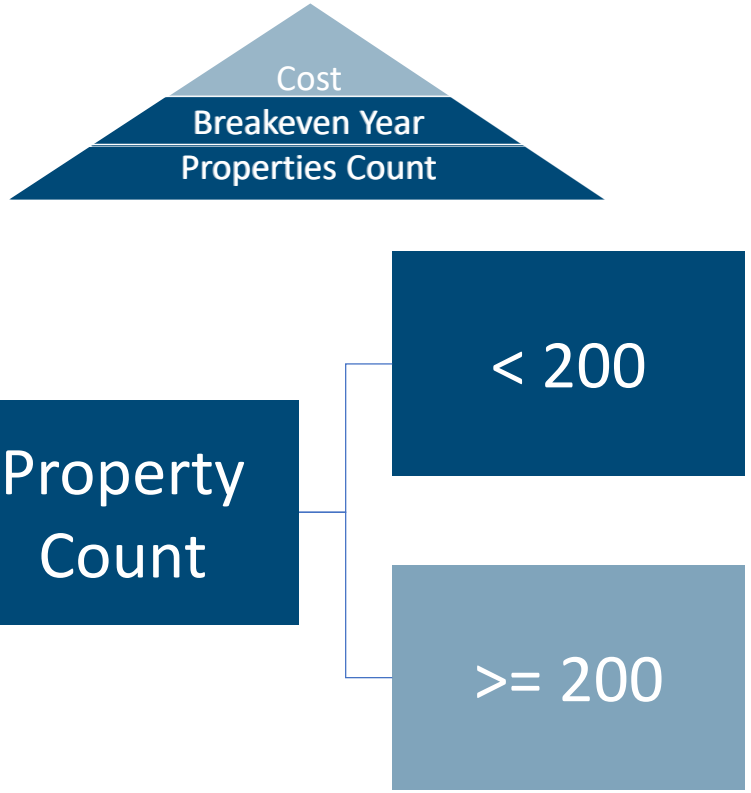
Borough Analysis ▶ Map Analysis ▶ **Zip Code Analysis**

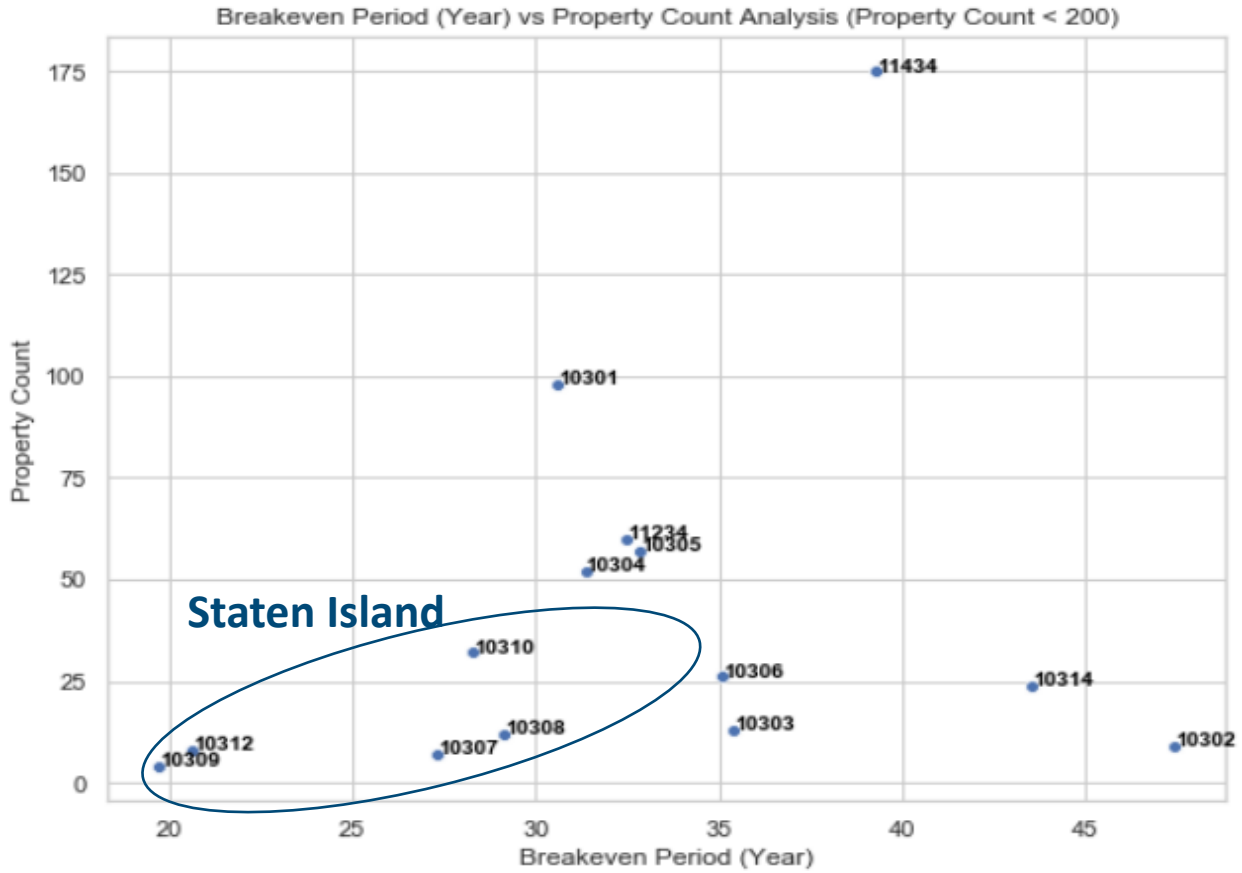


Visualization – Breakeven & Property Count

INSIGHT

- ***10309, 10312, 10307, 10308, and 10310 in Staten Island have least breakeven period***
- ***Best investment for short-term profitability, but less property count means higher risk***

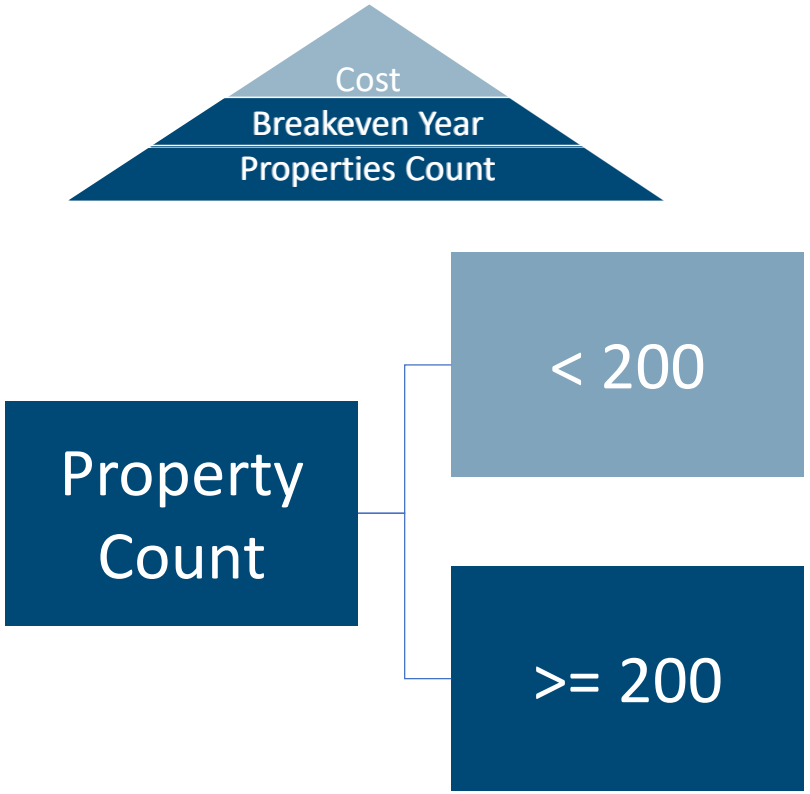


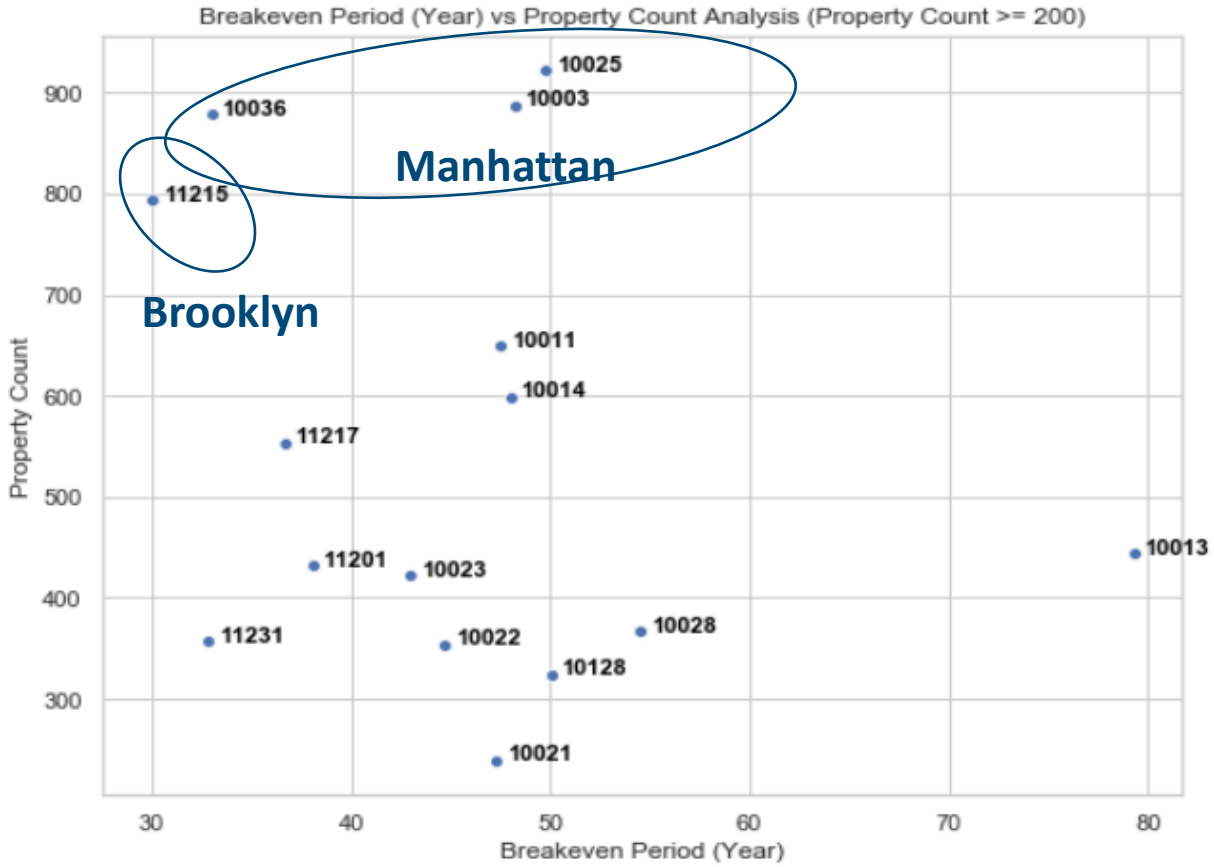


Visualization – Breakeven & Property Count

INSIGHT

- ***11215, 10036, 10025, and 10003 have decent breakeven period and high property count***
- ***Best investment for long-term profitability, but also need high cash flow***



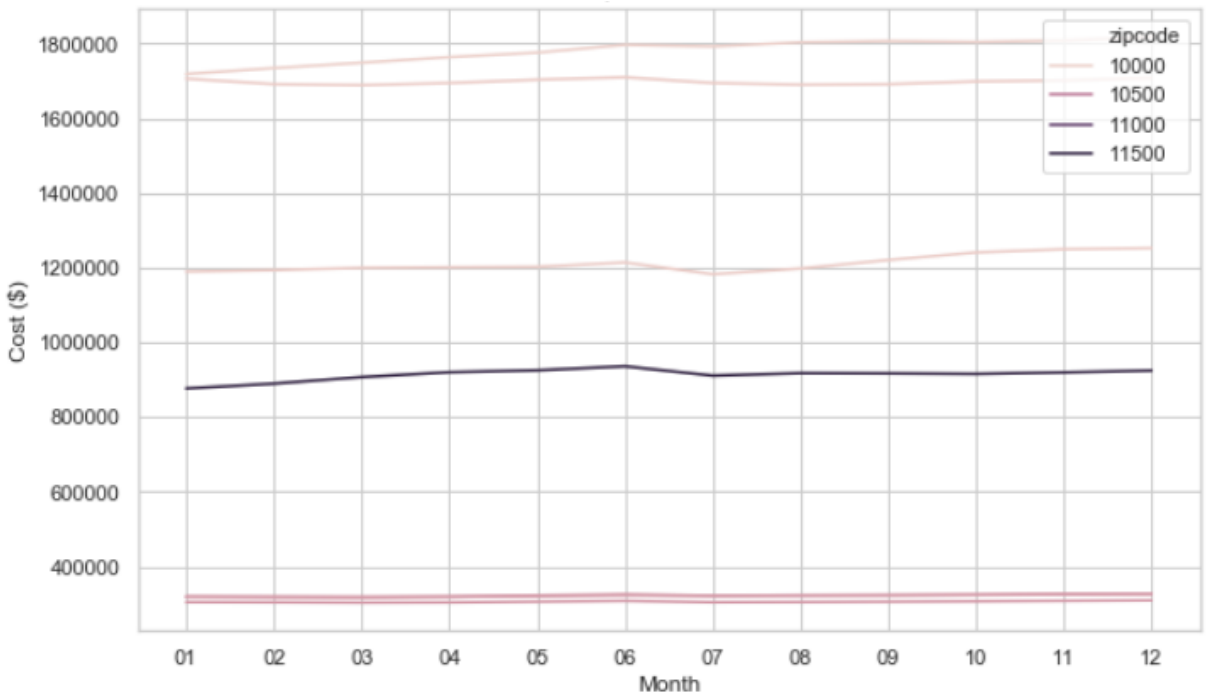


Visualization – When to Invest

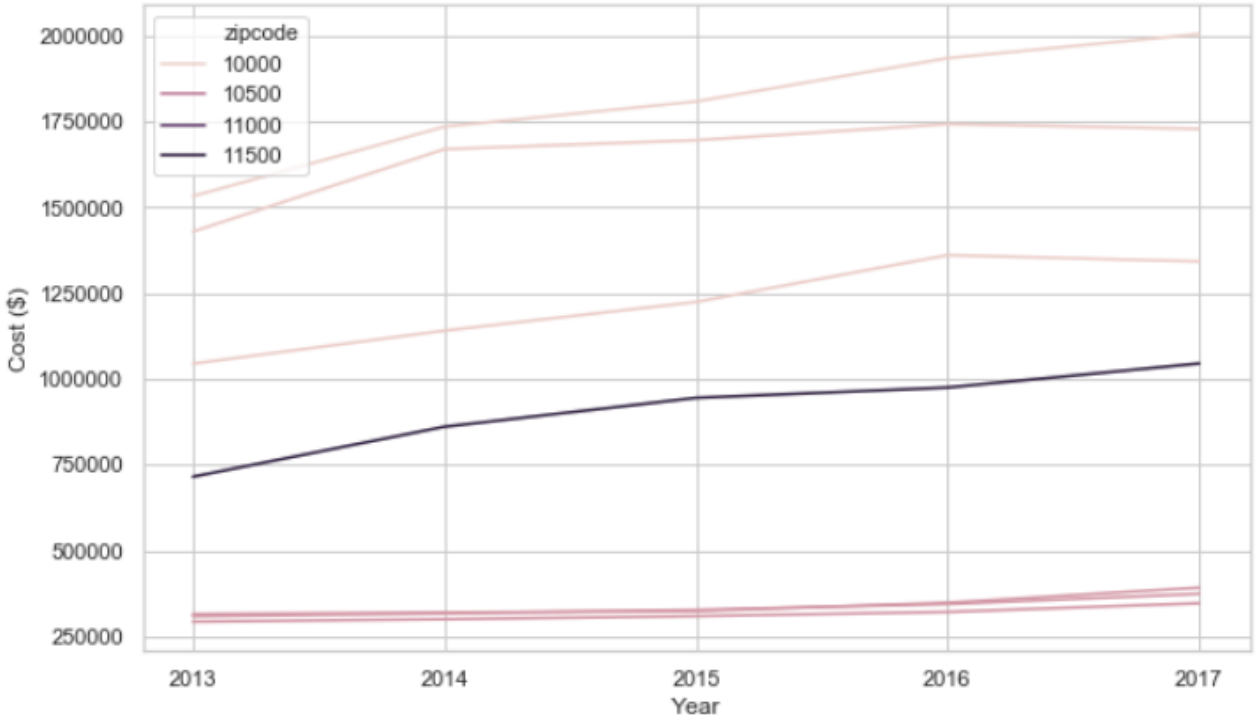
INSIGHT

- ***Property cost drops in July on average and as time goes property cost keeps increasing***
- ***The best time to invest is now (July 2019)***

Cost across Zip Codes in Different Month



Cost across Zip Codes in Different Year



Next Step

Actions	Goal
Consider other business actions such as selling properties in the future	More informative decisions zip code selection
Analyze the description of listings	Better understanding listings feature in different locations
Consider different price for different stay time	More accurate breakeven period