

New York Property Market. 2017 Insights

Tomasz Szmidt

Abstract

The dataset used in this research, [NYC Property Sales](#), was downloaded from kaggle.com. The research is aimed at answering the question: how does the distribution of NYC housing market look like across city boroughs? The methods of descriptive statistics, statistical analysis were used to come up with findings. The research shows that Manhattan is the most luxurious part of the city, Queens the most popular, Brooklyn hold steady prices and number, Bronx and Staten Island are the most affordable boroughs.

Motivation

When somebody wants to sell a house the price is often a rough guess. One may pay for detailed quotation as an alternative. A year worth of all property sales in NYC is **\$78bln**. One family dwellings make almost **11%** of that sum. The aim of this research is to provide broader market recognition to house owners and property agents. These insights will help make better decisions, to avoid selling at too low or too high prices. The overall distribution of housing market in NYC is a source of great value for property agents. Although this research is limited only to New York, the scheme would work with data from any other city.

Dataset

- The chosen dataset, [NYC Property Sales](#), comes from kaggle.com
- This is a .csv record of all real estates transactions in New York City
- It holds over 80000 entries starting from 09.2016 till 09.2017
- 21 columns contain text data on: property characteristics, price, address

Data Preparation and Cleaning

Problem: Since investigating transactions is the main goal of the project. Numerous \$0 or \$1 donations became an issue.

Solution: The most common donations (\$0, \$1, \$10) were dropped, as well as transactions below \$1000.

Problem: Missing values and duplicates

Solution: `dropna()` and `drop_duplicates()` operations were run.

The cleaned dataset consists of **43589** rows instead of initial **84548**

Problem: Data sorting

Solution: For best chronologic sorting, sale date was set to date time and set as index ID. The dataset was divided into monthly subsets.

Research Questions

PRIMARY QUESTION:

How does the distribution of NYC housing market look like across city boroughs?

SECONDARY QUESTIONS (answers in findings):

- *Is the NYC property market subject to seasonality?*
- *What are average prices across different boroughs?*

Methods

The structured and statistical nature of chosen data set implies the use of **quantitative analysis**. Therefore, methods of **descriptive statistics** were used:

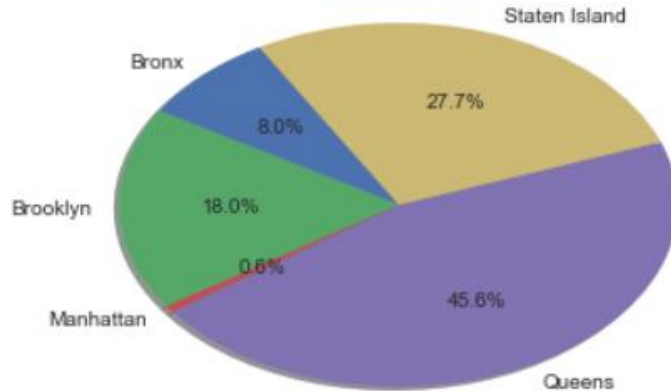
- Analysis of numerical and categorical variables
- Swarm plots, bar, line and pie charts to visualize findings

Findings

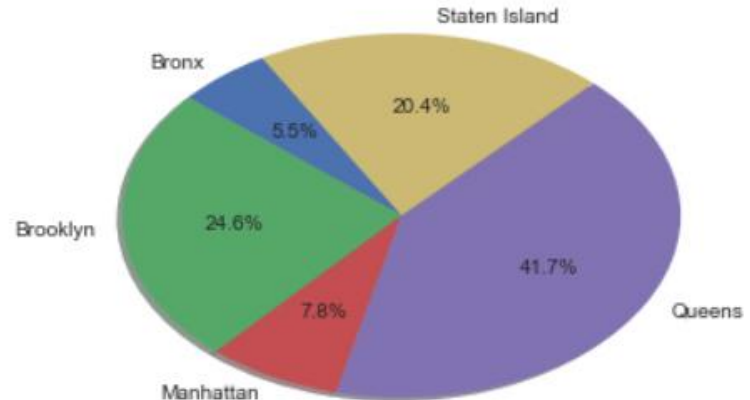
The summary transactions data from Sep 2016 to Aug 2017 indicates that:

- Manhattan is highly expensive borough. Only 0.6% of transactions make 7.8% of market worth
- Queens is by all means the most popular borough
- Transactions in Bronx and Staten Island provide less cash despite large volume

Share of number of transactions in NYC boroughs.



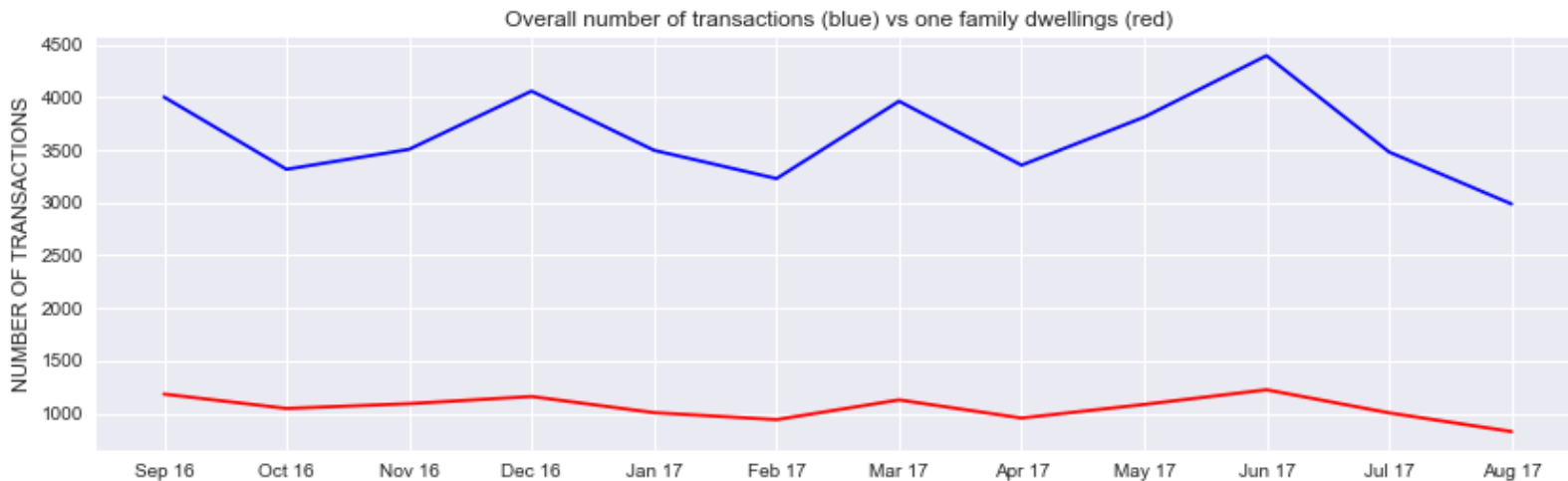
Worth of transactions in NYC boroughs.



Findings

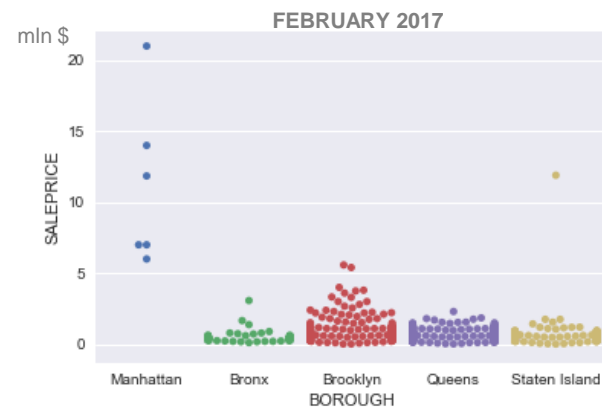
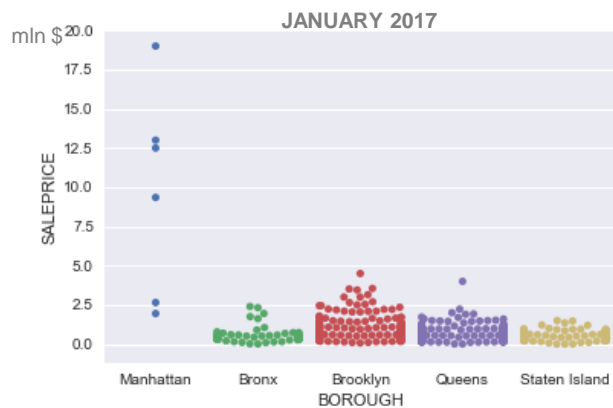
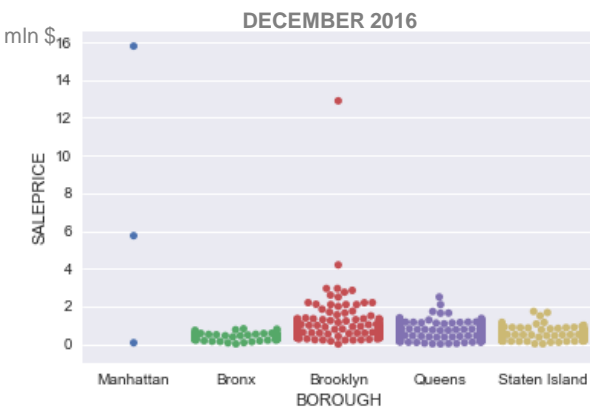
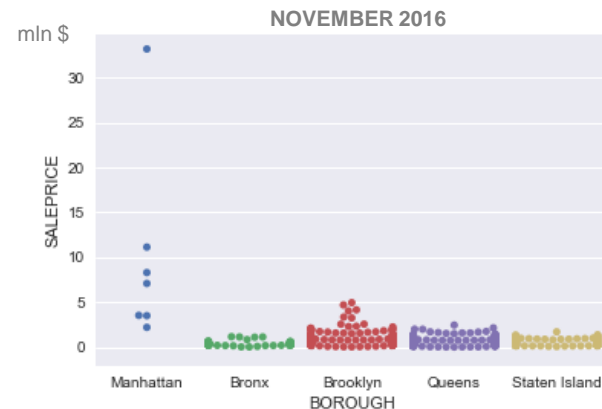
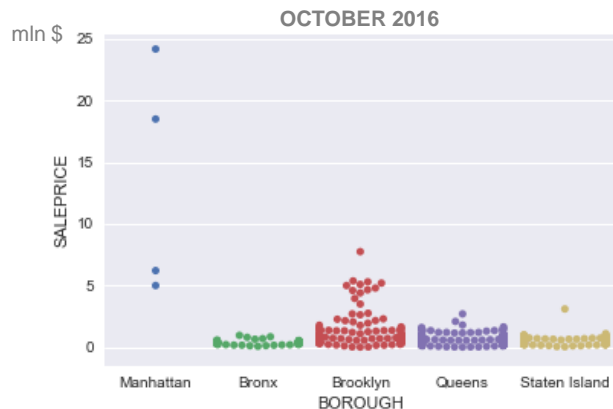
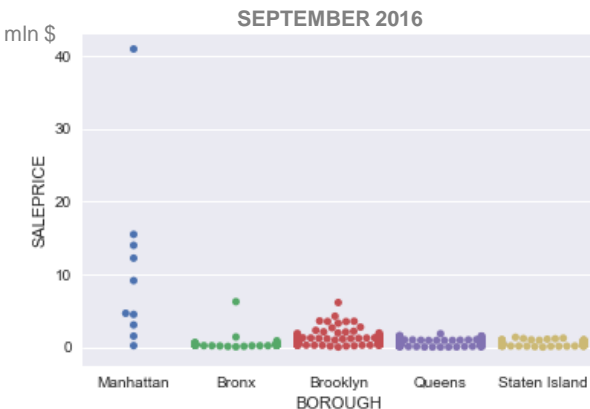
The NYC property market is subject to seasonality changes

- There are several monthly peaks throughout the year: September, December, March, June
- One family dwellings transactions make around 10% of the entire market,
- It can be seen that one family dwellings curve follows the same seasonality patterns as the overall curve



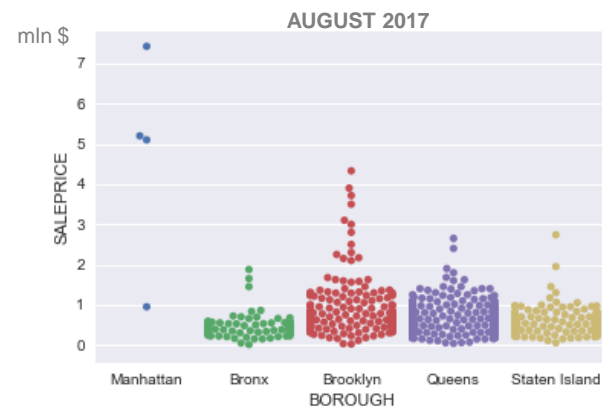
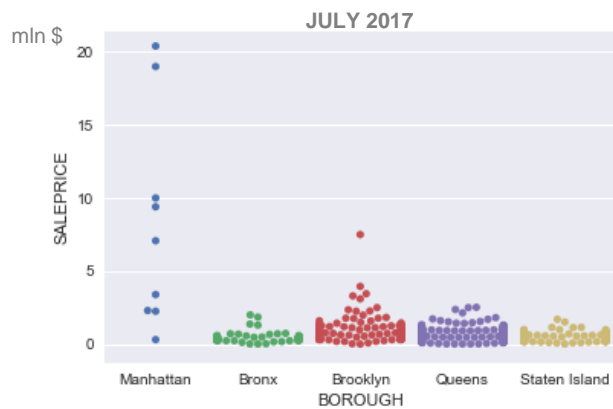
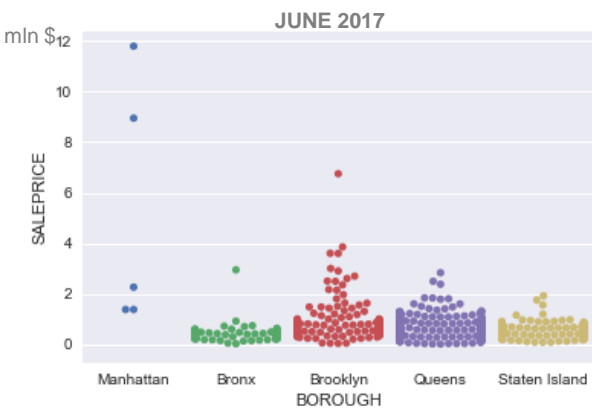
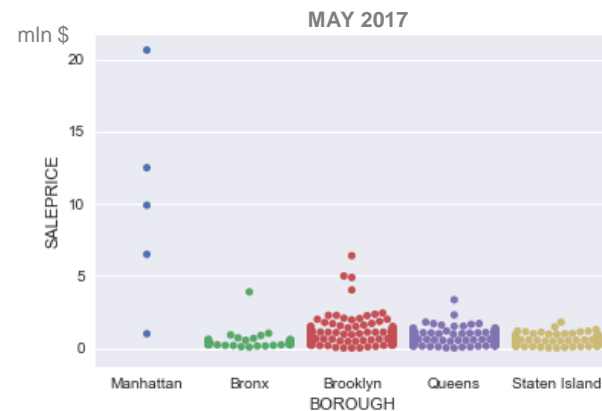
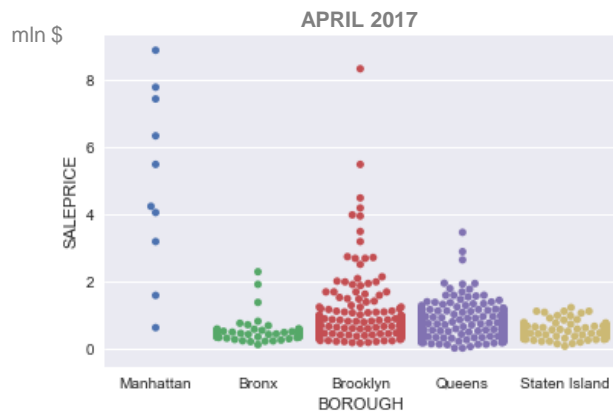
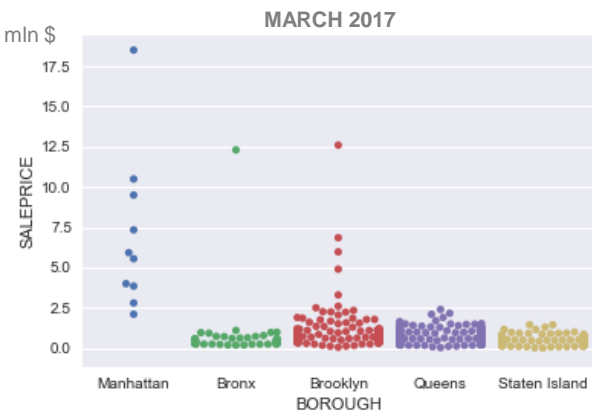
Findings: monthly split of transactions

The distribution of transactions in boroughs including the sale price (in mln \$)



Findings: monthly split of transactions

The distribution of transactions in boroughs including the sale price (in mln \$)



Findings: monthly split of transactions

The look into one year of one family dwellings sales gives following insights:

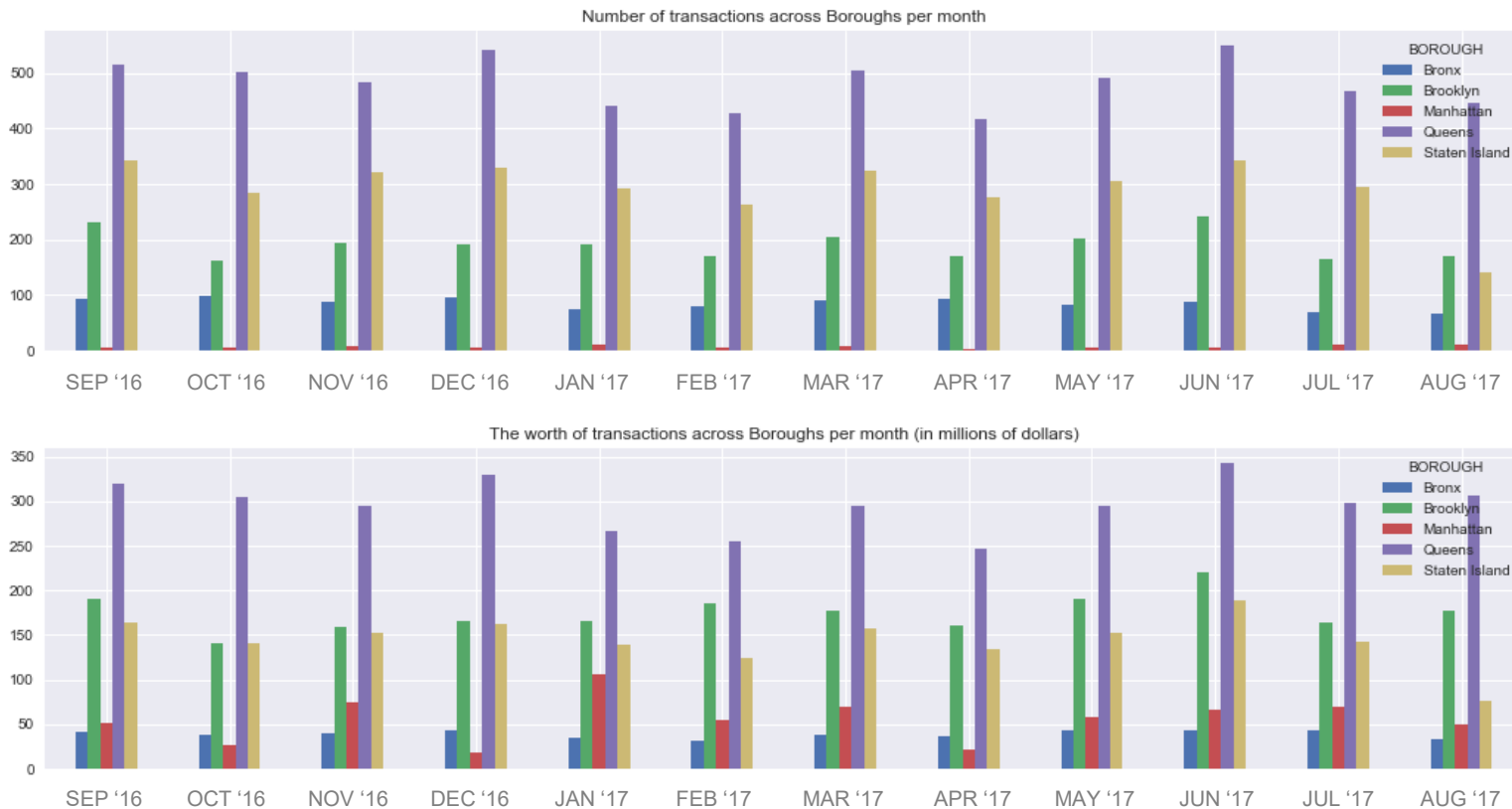
- Although the number of transactions in Manhattan is low, the property prices easily pass \$10 mln price tag
- All other boroughs are way behind, as the most expensive properties hardly reach \$5 million price tag
- The properties in Bronx are usually sold at average prices. Few transactions go above the mean.
- Queens makes the biggest market share due to number of transactions
- Brooklyn is the borough where properties are quite often sold at above average prices

Average monthly prices of one family dwellings (in millions of dollars)

BOROUGH	SEPTEMBER 2016	OCTOBER 2016	NOVEMBER 2016	DECEMBER 2016	JANUARY 2017	FEBRUARY 2017	MARCH 2017	APRIL 2017	JUNE 2017	JULY 2017	AUGUST 2017
Bronx	0.44299	0.39768	0.44592	0.44739	0.46392	0.38665	0.41482	0.38927	0.48919	0.62545	0.51306
Brooklyn	0.83172	0.87190	0.82062	0.86700	0.86074	1.09399	0.86409	0.95568	0.91542	1.00088	1.05154
Manhattan	10.11000	5.15200	8.23636	4.66875	10.57888	13.47125	9.85206	7.20750	11.14167	7.00650	4.97458
Queens	0.62087	0.60474	0.61098	0.60694	0.60180	0.59756	0.58373	0.59182	0.62377	0.63847	0.68727
Staten Island	0.47876	0.49847	0.47253	0.49584	0.47683	0.47559	0.48733	0.48351	0.55251	0.48093	0.54594

Findings:

Number of transactions vs the monthly worth of transactions



Limitations

Although the examined dataset is detailed it doesn't even cover one full year. The time span (Sep '16 - Aug '17) is too short to look for trends.

The dataset provides only information on the sale date without giving any details on when the property was put on sale. That limits the opportunity of putting together many interesting comparisons. For instance how price affected the time of finding a buyer.

Conclusions

How does the distribution of NYC housing market look like across city boroughs?

- Queens play vital role (41,7% of the market worth, 45,6% of all transactions)
- Manhattan is the most exclusive borough (average monthly prices spread from \$4mln up to over \$13mln)
- Brooklyn is a popular location with almost 200 monthly transactions and average prices close to \$1mln
- Bronx and Staten Island are the cheapest boroughs

Acknowledgements

The data was taken from open database on kaggle.com

<https://www.kaggle.com/new-york-city/nyc-property-sales/data>

I shared my notebook twice with kaggle.com community as a kernel. My user name: Thoms Shmims

<https://www.kaggle.com/new-york-city/nyc-property-sales/kernels>

References

All the work is my own achievement

I used <http://iquantny.tumblr.com/> as a source to get better understanding of the nature of NYC data in general