

Visualizing NYC 2017 House Sales Data

James Chen

01/31/2018

Outline

- ① Project Objectives
- ② Data (Source, Pre-Processing, Transformation)
- ③ Shiny App Demonstration
- ④ Additional Observations

Project Objectives

- To visualize NYC house sales pattern, *borough-wise*
In particular:
 - Transaction volumes (number of houses sold)
 - House prices
- To find additional insights from the above visualization
- To practice using R, Shiny, Leaflet

Part II. Data

Data Source and Pre-Processing

- Source
 - NYC Dept. of Finance, 2017 properties rolling sales for all tax classes
 - Raw data set: five files (M, BX, B, Q and SI)
<http://www1.nyc.gov/site/finance/taxes/property-rolling-sales-data.page>
- Pre-Processing
 - All the usual data cleaning
 - For simplicity, all sales < \$100,000 filtered out
 - Grid granularity: Zipcode-wise ((*Lat.*, *Long.*) added to dataset)

Data Transformationation

Boro-wise Normalization Normalize a quantity taken in zip code $\in Boro$ with respect to μ_{Boro} and σ_{Boro} . Essentially z -scores.

(Two examples next)

Transformationation: Boro-wise Normalization (1/2)

Example 1 Normalized volume ZV in Brooklyn, January

Say Zipcode = 11234 ∈ Brooklyn

$$ZV(Zip == 11234) = \frac{V(Zip == 11234) - \bar{V}_{\text{Bklyn., Jan.}}}{\sigma_{\text{Bklyn., Jan.}}}$$

Note

- $\bar{V}_{\text{Bklyn., Jan.}}, \sigma_{\text{Bklyn., Jan.}}$ from

$$\{ V_{\text{Jan.}}(Zip_i) : Zip_i \in \text{Brooklyn} \}$$

- $ZV(Zip_i) == 0$: $V(Zip_i)$ at average level in Brooklyn, January
 - ⊕ $ZV(Zip_i)$: $V(Zip_i)$ above average
 - ⊖ $ZV(Zip_i)$: $V(Zip_i)$ below average

Transformationation: Boro-wise Normalization (2/2)

Example 2 Normalized price ZP in Manhattan, April

Say Zipcode = 10023 ∈ Manhattan

$$ZP(Zip == 10023) = \frac{P(Zip == 10023) - \bar{P}_{\text{Manh., Apr.}}}{\sigma_{\text{Manh., Apr.}}}$$

Note

- $\bar{P}_{\text{Manh., Apr.}}, \sigma_{\text{Manh., Apr.}}$ from

$$\{P_{\text{Apr.}}(Zip_i) : Zip_i \in \text{Manhattan}\}$$

- $ZP(Zip_i) == 0$: $P(Zip_i)$ at average level in Manhattan, April
 - ⊕ $ZP(Zip_i)$: $P(Zip_i)$ above average
 - ⊖ $ZP(Zip_i)$: $P(Zip_i)$ below average

Boro-wise Normalization: Why? (1/2)

Philosophy: Everything is relative

- ① Normalized volume ZV : sharp anomaly pattern *within* a borough

Consider: $(\dots, -1, 0, 1, \dots)$ sharper contrast than $(\dots, \mu - \sigma, \mu, \mu + \sigma, \dots)$

(Example: next slide)

- ② Normalized price ZP : fair price comparisons *across* boroughs

- House price \$1M implies very differently in Manhattan than Staten Island
- So does a \$100K fluctuation in Manhattan than Staten Island

Data Transformationation: Why? (2/2)

Heat maps: House sales volumes, Brooklyn, May

(Heat map: intensity \propto volume)

Left: Volume; Right: Normalized Volume



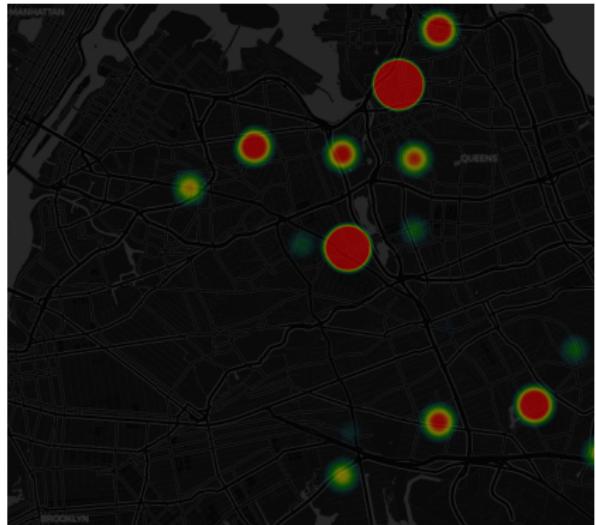
Part III. Shiny Demonstration

https://jamesczq.shinyapps.io/NYC_House_Sales_2017/

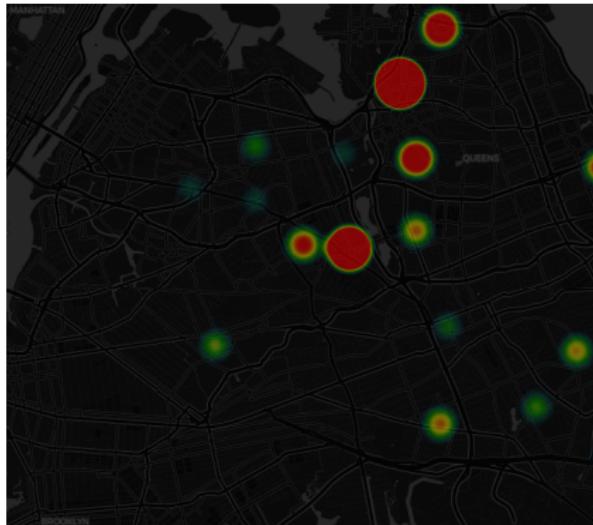
Shiny Demonstration (1/2)

Queens Transaction Pattern, May versus June, 2017

May 2017

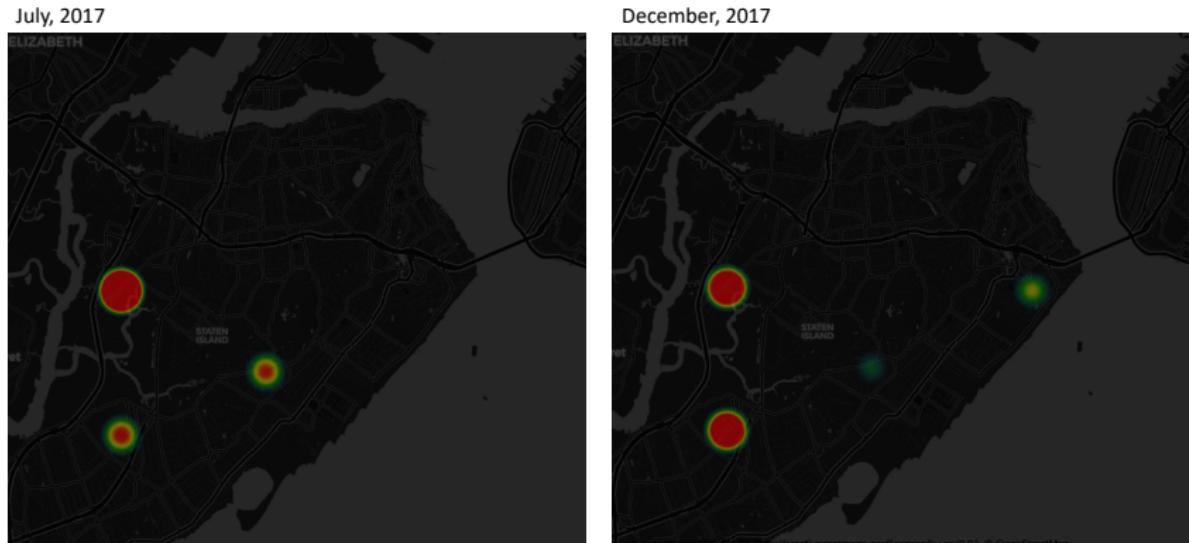


June 2017



Shiny Demonstration (2/2)

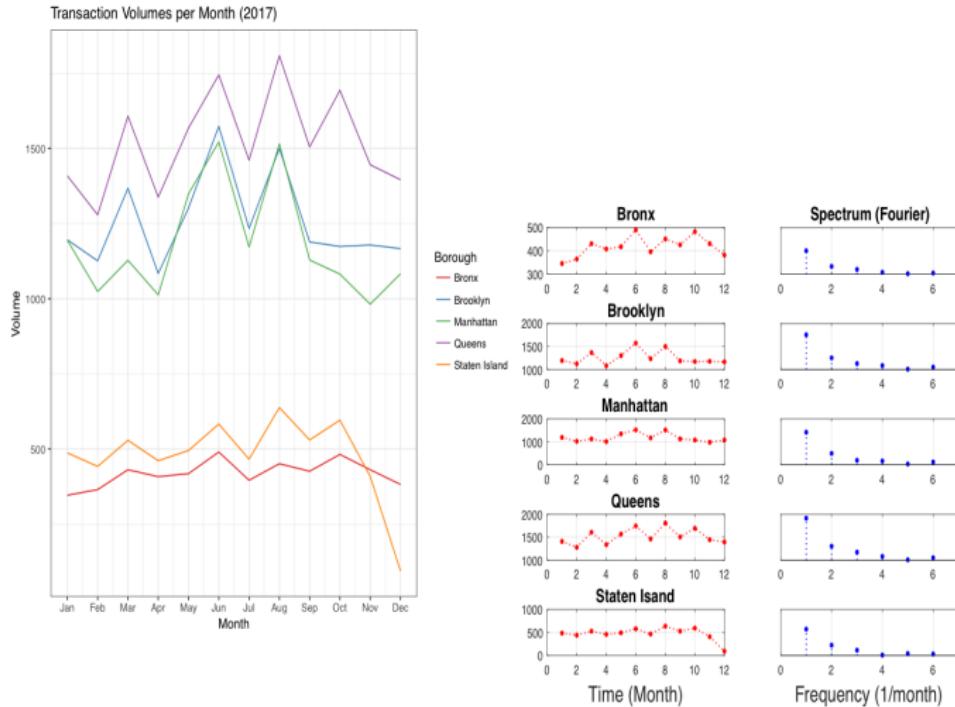
Staten Island Transaction Pattern, July versus December, 2017



Part IV. Additional Observations

Additional Observations (1/3)

Time-behavior: Periodicity in transactions

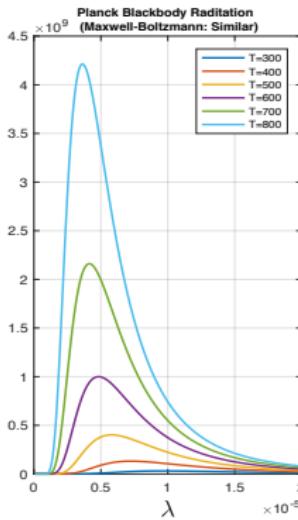
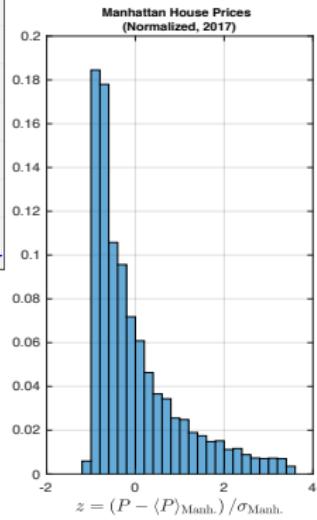
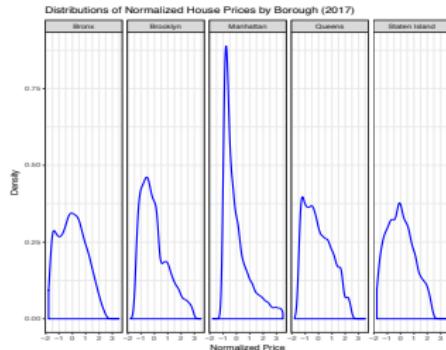


Additional Observations (2/3)

Thermodynamics of NYC real estate business

Postulate: Systems in local thermodynamical equilibria (LTE)

NYC Real Estate Business		Thermodynamics
Location	Temperature	

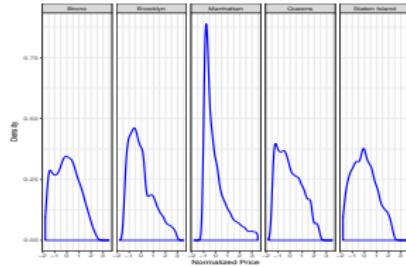


Additional Observations (3/3)

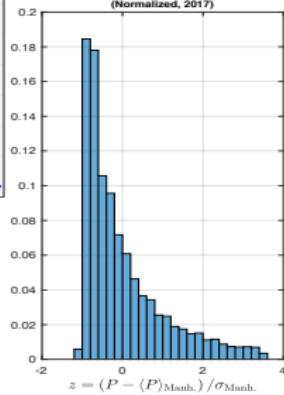
NYC real estate thermodynamics (systems in LTE)

<i>NYC Real Estate Business</i>	<i>Thermodynamics</i>
Location	Temperature
(Volume, Price)	Phase Space (x, \dot{x})
Multiplicity in (Volume, Price)	Entropy

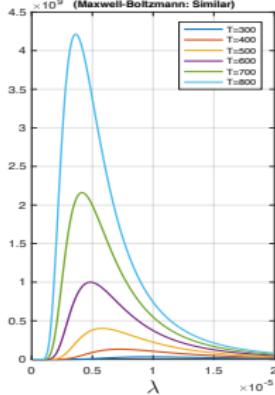
Distributions of Normalized House Prices by Borough (2017)



Manhattan House Prices
(Normalized, 2017)



Planck Blackbody Radiation
(Maxwell-Boltzmann: Similar)



Thank You!

- Possible improvements:
 - 1 Refine plots and Shiny appearance
 - 2 Incorporate data from past years (2016 and prior)
 - 3 Refine study with different house categories
 - 4 ...
- Questions?
- Thanks for the help with Shiny from Aaron, Drace, Kathryn, and Zeyu
- Thank you!