

Exploratory Analysis of Data for Airbnb Listings in NYC

Youngsoo Baek, Irene Yi Ji, Phuc Nguyen

Exploratory Analysis

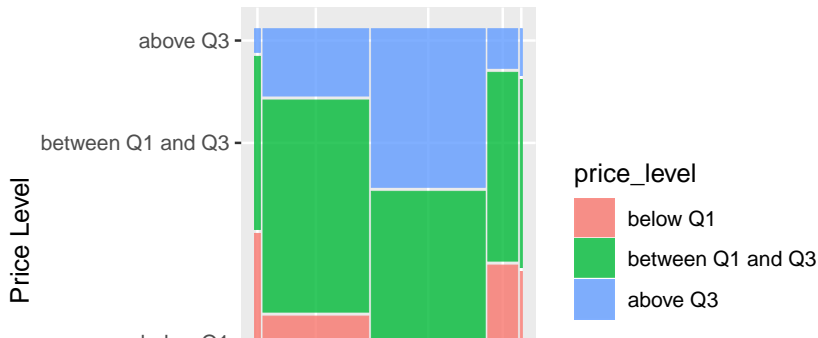
Heterogeneity among Boroughs

	Bronx	Brooklyn	Manhattan	Queens	Staten
below Q1	574	6572	2514	2549	
between Q1 and Q3	455	10274	10872	2577	
above Q3	62	3258	8275	540	

- Pearson's Chi-squared test

X-squared = 6196.4, df = 8, p-value < 2.2e-16

Price Level by Boroughs



Response of Interest: Price and Popularity

Choosing a Meaningful Measure of Popularity

Heterogeneity across Neighbourhoods/Boroughs

Spatial Correlation

Predictors of Interest

Possibly Unreliable Predictors

Modeling

Price and Popularity: Bivariate Mixed Effects Regression

Did We Miss Spatial Correlation Within Neighbourhoods?

Text Analysis for Listing Names

Further Work