

Chicago Neighborhood Exploration: Venue and Home Price Comparison

McKay Rytting

Introduction

- Gentrification is the process of wealthy individuals displacing long-time residents in a neighborhood that is undergoing economic growth
- As neighborhoods gentrify, typically property values increase significantly, making it more difficult for many residents to remain
- It may be that in addition, the types and number of businesses also change significantly, and begin to resemble wealthier neighborhoods
 - The wealthier neighborhood used in this exploration is Lincoln Park, one of the traditionally wealthiest communities in Chicago

Data

- The data used in this analysis come from Foursquare and Zillow
- Zillow provides home sale price data from every month going back to 2008. These home sale prices aren't perfect because they only show the properties that actually are purchased by those moving into a neighborhood.
- Foursquare provides business venue data, and historical snapshots of that data can be accessed by using different version numbers
- The Zillow and Foursquare data will be restricted to Chicago zip codes only

Methodology

- First, a list of Chicago zip codes and a table of home sale prices in all US zip codes are acquired
- Next, the Chicago zip codes are connected with their latitude and longitude values, which are fed into the Foursquare API to gather venues within a certain radius. A function is written to automatically perform this retrieval for a series of version numbers going back to 2008
- The venues are encoded using a one hot method and then venue frequencies in the neighborhoods are calculated
- These venue frequencies are compared between each zip code and the zip code for Lincoln Park at each time stamp to see if the zips are becoming more like Lincoln Park over time

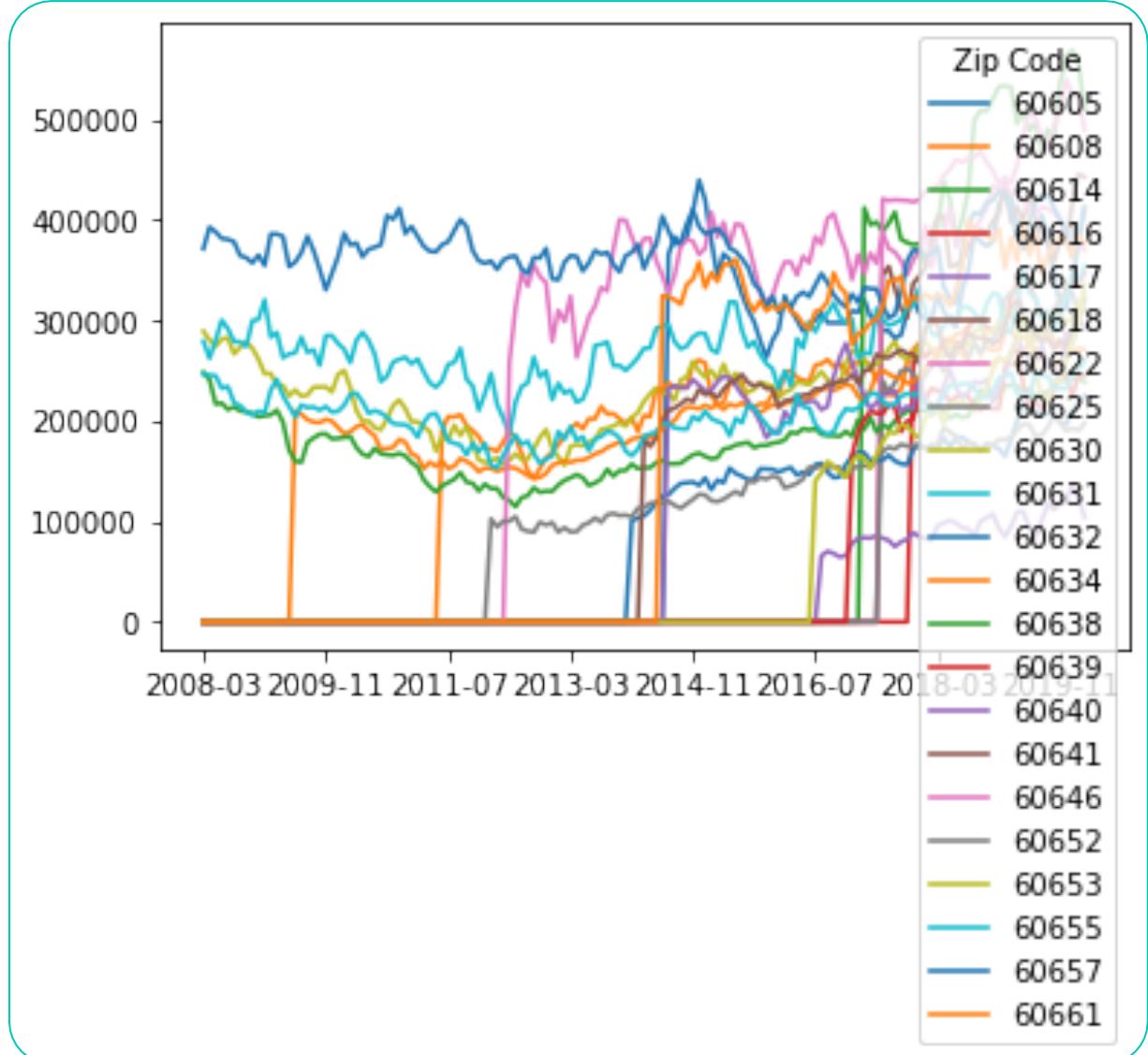
Zip Code	aveSim	devSim	maxSim	minSim	rangeSim	
0	60601	0.007890	0.000000e+00	0.007890	0.007890	0.0
1	60602	0.005891	0.000000e+00	0.005891	0.005891	0.0
2	60603	0.007143	0.000000e+00	0.007143	0.007143	0.0
3	60604	0.011754	0.000000e+00	0.011754	0.011754	0.0
4	60605	0.010842	0.000000e+00	0.010842	0.010842	0.0
5	60606	0.008214	0.000000e+00	0.008214	0.008214	0.0
6	60607	0.008242	0.000000e+00	0.008242	0.008242	0.0
7	60608	0.012897	4.513898e-36	0.012897	0.012897	0.0
8	60609	0.017857	0.000000e+00	0.017857	0.017857	0.0
9	60610	0.013305	4.513898e-36	0.013305	0.013305	0.0
10	60611	0.007500	0.000000e+00	0.007500	0.007500	0.0

Select Results of Foursquare Venue Data

Foursquare Venue Data Insights

- Frustratingly, the business venue profiles generated from the Foursquare data did not appear to change over time, at least with respect to the similarity to Lincoln Park
- If there were changes in the business profiles of the neighborhoods, they were not significant enough to register changes in this exploration
- This could indicate that neighborhoods take longer to change significantly than the timeframe of these observations (~12 years)
- This could also indicate that neighborhood identities are less malleable than I had thought

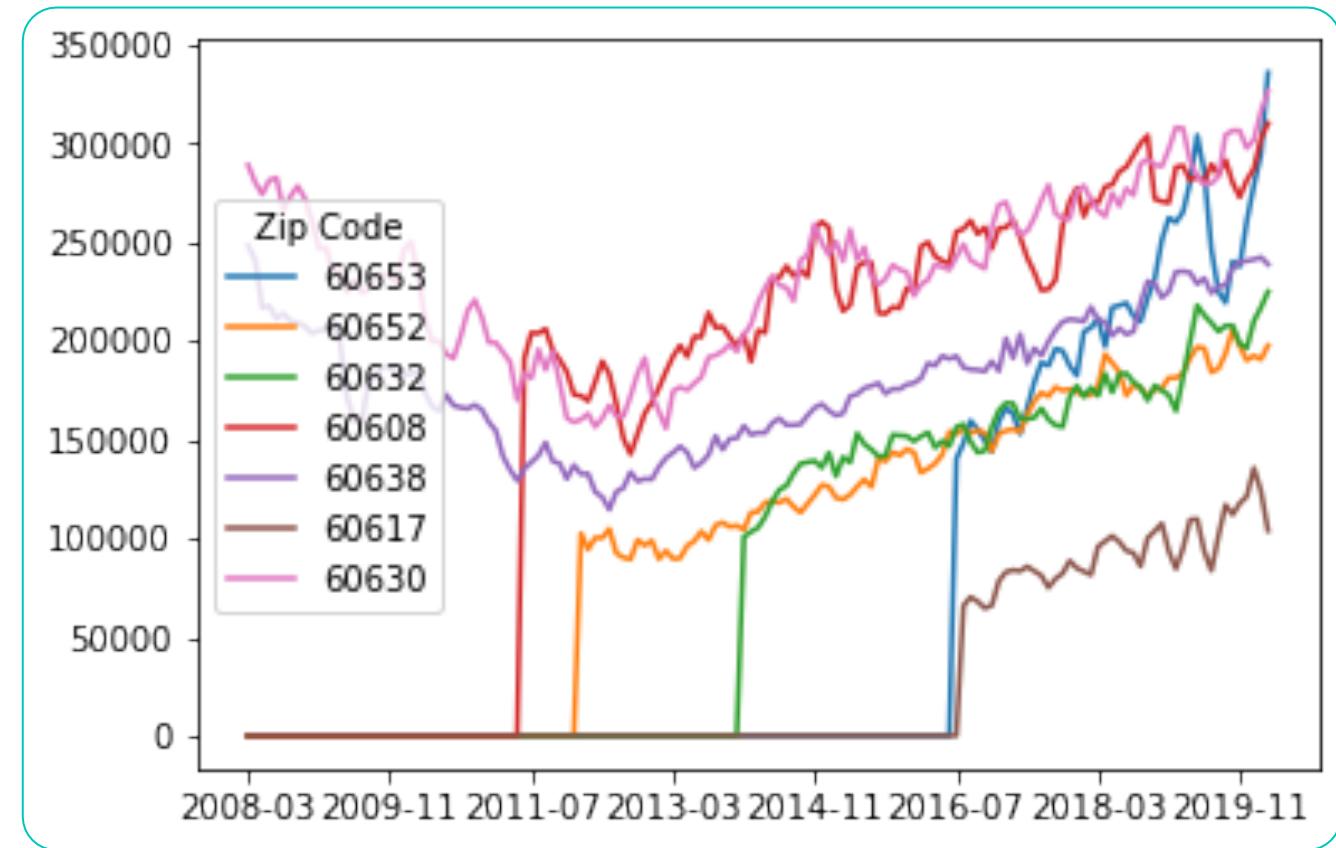
Zillow Home Data Insights



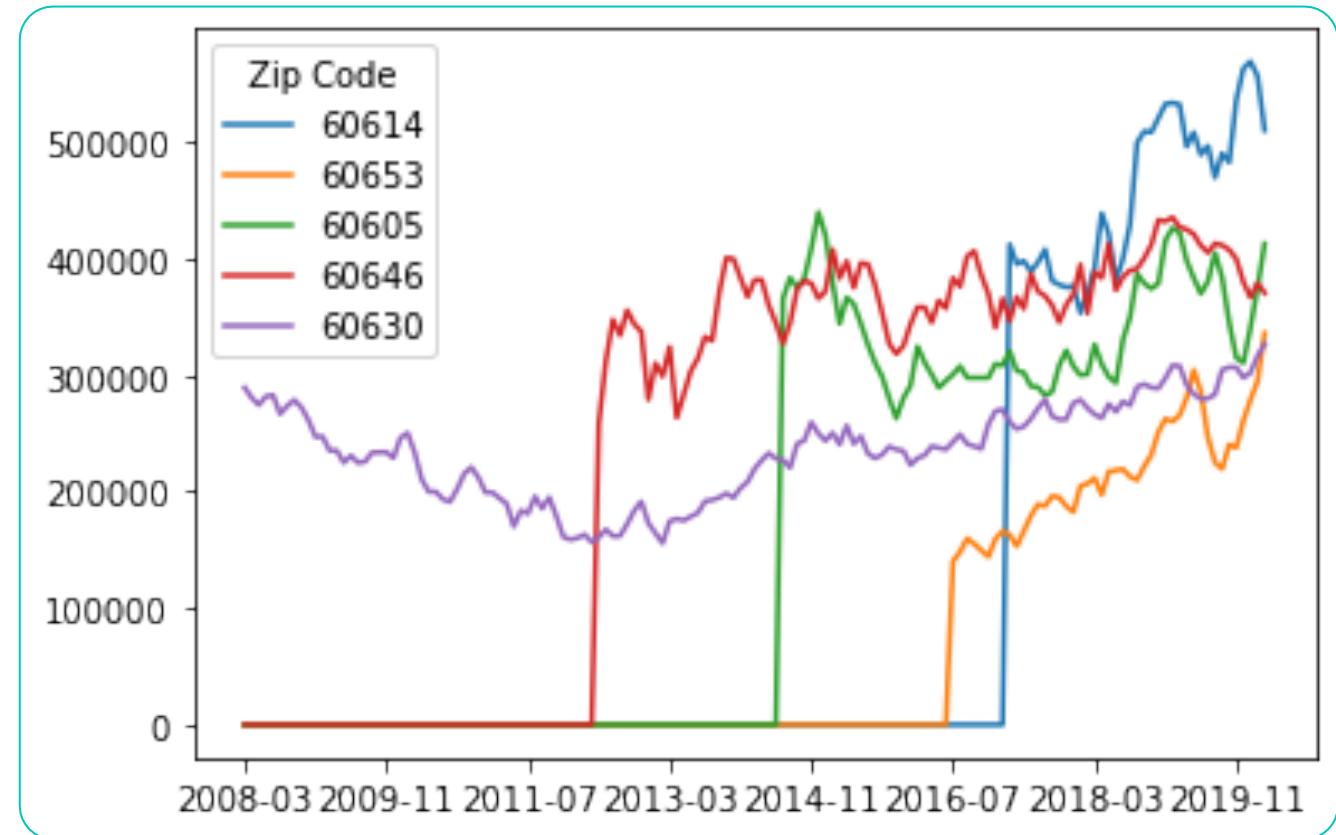
Zillow Home Data Insights

- All of the Zillow data together was a little hard to parse visually
- I calculated the change in home prices over the span of up to 9 years from 2011 to 2020 and plotted the neighborhoods with the largest proportional increases and the largest raw increases from their post-2008 troughs
 - Some neighborhoods did not have data going this far back, so I used the first available data points
- The neighborhoods with the largest raw increases were all neighborhoods with already high home values
- Many of the neighborhoods with the largest proportional increases began with very low home values

Neighborhoods with Largest Proportional Increase



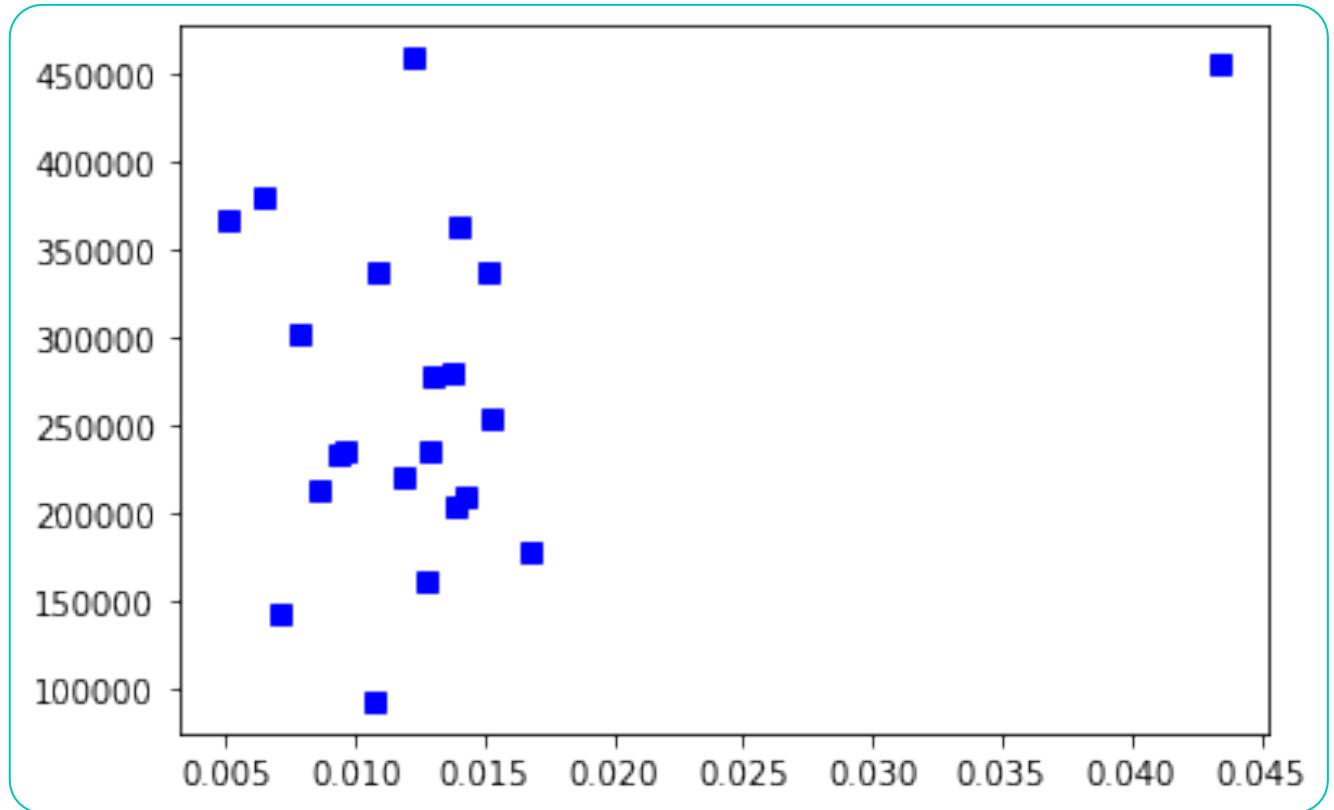
Neighborhoods with Largest Raw Increase



Does Foursquare Venue Data Correlate with Wealth

- Finally, I compared the calculated similarity score to Lincoln Park with home prices and compared the two
- I thought that neighborhoods which were more similar to Lincoln Park in terms of home value would also have similar types of businesses, but the correlation appears to be basically nonexistent
- It could be that you could find the right sets of businesses, maybe group categories better before the one hot encoding, etc and find a way to correlate these datasets better, but the entire set of data from Foursquare does not appear to be predictive of neighborhood wealth

Does Foursquare Venue Data Correlate with Wealth?



Conclusions

- Neighborhoods do change, but slowly and often less than we might think
- Even when neighborhoods do have increased wealth, they may not develop a very uniform set of business venues
- The diversity between similarly wealthy neighborhoods is much greater than the similarity between similarly wealthy neighborhoods