

Amenitization vs. Rents vs. Walkscore

Introduction:

The purpose of this study is to understand the relationship that exists (if any) between the amenitization of an urban environment and the prices (rents) associated with real estate in that vicinity. One would believe that, because amenities such as cafes and schools offer accretive value to their local residents, the value of these amenities would be priced into the surrounding real estate.

Data used:

- *NYC neighborhoods*: For this particular analysis, we have focused on New York City, and specifically on its 4 largest boroughs (Manhattan - or New York County, Brooklyn - or Kings County, Queens - Queens County, and The Bronx - Bronx County). We are using Zillow's classifications of these neighborhoods, along with Zillow's categorization of the borough for where each of these neighborhoods belong.
- *OpenStreetMap*: To track amenities, we have used OpenStreetMap, an open-source web application that populates amenities across urban environments. Specifically, we have used the Python applications OSMx and Pandana to track the specific geographic locations of these amenities.
- *Zillow rent values*: We have used Zillow's publicly available rent values to track real estate values in each neighborhood. This data is available online on the neighborhood level. Specifically, we have focused on March 2019 rent values for 1-bedroom units.
- *Walkscore*: To cross-validate distance to amenities, we have used Walkscore, a publicly available metric which measures the walkability of particular neighborhoods across the U.S.

Unit of analysis: For simplicity, our unit of analysis is the NYC neighborhood (e.g. Williamsburg, Lower East Side, Jamaica)

Technical methods used:

- *Data Collection*: All sources except OpenStreetMap were manually uploaded via pandas `read_csv`. OpenStreetMap data was pulled directly using OSMx and Pandana, Python applications.
- *Analysis*: We built functions looping through each neighborhood to calculate average distance to nearest amenity across each neighborhood, focusing on schools and cafes as the two amenities to focus for this particular study.
- *Visualization*: In addition to building heat maps visualizing the distance to amenities across each borough, we also built scatterplots highlighting the relationship between average amenity distance, Walkscore, and Zillow rent values. We grouped these scatterplots by borough.

Results:

- As expected, we notice a clear inverse relationship between distance to amenities and rental values. Specifically, the smaller the average distance to an amenity, the greater the average neighborhood rental value. This is cross-validated against walkscore -- neighborhoods with higher walkscore have been found to have a lower average distance to cafe, and also have higher average rents.

Limitations:

- The main limitation for this study is that the current analysis relies on average distance to nearest amenity in the same neighborhood. As outlined in next steps below, future code would solve this issue by sourcing distances to the nearest amenity, even if that amenity is beyond the immediate neighborhood. For example, there are certain neighborhoods (such as in Queens) that do not have any cafe -- the current code does not calculate the nearest cafe for these neighborhoods.
- Another limitation is that we are relying on OpenStreetMap's categorizations of and aggregation of amenities. It is likely that OpenStreetMap does not include all cafes and schools present in New York.

Next steps:

- *Compare density of amenities to quality of amenities:* This current analysis only considered the nearest cafe/school, without taking into consideration how many other cafes or schools are nearby or the quality of these cafes/schools. A continued analysis would both measure density (via calculating average distance to, for example, the third nearest cafe/school) and quality (via using a database, for example, of hottest new cafes/bars in New York or the top-rated schools in New York). Another way to incorporate quality would be to consider, for example, Yelp review scores.
- *Expand average distance to amenity function to consider amenities beyond neighborhood:* Per limitation mentioned above.
- *Use time-sensitive amenity data:* It would be useful to incorporate the date when amenities were introduced, as this would allow us to measure how real estate values change over time with the introduction of certain new amenities. More broadly, time-sensitive amenity data would allow us to understand whether a neighborhood is undergoing gentrification, which would be measured by increasing real estate values alongside new, highly-rated amenity additions.
- *Compare to demographics-related data:* In addition to cross-validating against Walkscore, it would be useful for next steps to cross-validate against demographics data (e.g. household income) to consider whether there exists a relationship between distance to quality amenities and household wealth.