

Relationship between venues and housing prices

Table of Contents

Introduction	1
Data	2
Property prices.....	2
Latitude and longitude of districts	2
Venue data from Foursquare	3
Method	3
Cleaning the data	3
Using API to request venue data.....	3
One-hot encode venues for correlation analysis	4
Perform correlation analysis	4
Results	4
Correlation of property prices.....	4
Venues correlated with higher prices.....	5
Venues correlated with lower prices	5
Difference for 4-bed rentals.....	5
Discussion.....	6
Conclusion.....	7

Introduction

This project will be aimed at home buyers, sellers and property professionals in London, UK, who want to know what are the indicators of cheaper or more expensive housing in a neighbourhood.

Currently, this audience may not be sure of the effect of certain venues on property prices in their neighbourhood. For example, they may not be sure whether a new pizza restaurant opening up nearby will be beneficial or detrimental to property prices in the area. The convenience of having a new restaurant nearby is good, but the additional noise from patrons may be a negative. The same with parks – the green space is good, but in some urban areas, parks may be seen as a security risk for nearby properties.

So we will aim to clarify what effect – positive or negative – different types of venues being present in a neighbourhood have on property rents and prices.

We will identify the correlations between different types of venue being present in a neighbourhood, and average rental and sales prices in that area. Correlation is by no means causation, but this information should enable homeowners and potential buyers to better understand the effect of new venues opening up or closing down nearby, and how venues in the neighbourhood correlate with the rental or sales prices.

Data

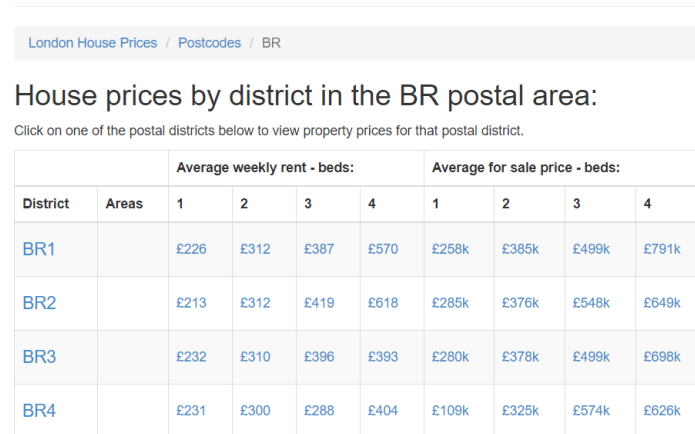
In this section, we describe the data sources that have been selected:

Property prices

The list of London district postcodes, and their associated average rental & sales prices will be scraped from findproperly.co.uk. For example as seen in Figure 1, the average weekly rent for a 1-bed property in the BR1 district is £226. We used the average weekly rent price, and average sale price, of 1, 2, 3 and 4-bed properties in each postcode.

Figure 1: Screenshot from findproperly.co.uk

House Prices in the BR Postal Area



London House Prices / Postcodes / BR									
House prices by district in the BR postal area:									
Click on one of the postal districts below to view property prices for that postal district.									
District	Areas	Average weekly rent - beds:				Average for sale price - beds:			
		1	2	3	4	1	2	3	4
BR1		£226	£312	£387	£570	£258k	£385k	£499k	£791k
BR2		£213	£312	£419	£618	£285k	£376k	£548k	£649k
BR3		£232	£310	£396	£393	£280k	£378k	£499k	£698k
BR4		£231	£300	£288	£404	£109k	£325k	£574k	£626k

Latitude and longitude of districts

The latitude and longitude of the postcodes will be pulled from a CSV file, downloaded from freemaptools.com. For example as seen in Figure 2, the AB10 district has latitude 57.13514, longitude -2.11731.

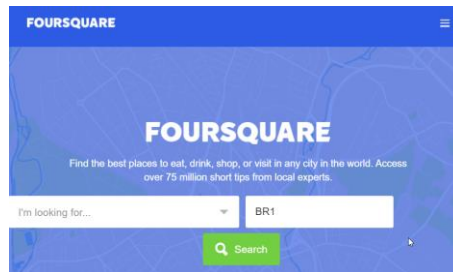
Figure 2: Example of postcode location data from freemaptools.com

id	postcode	latitude	longitude
2	AB10	57.13514	-2.11731
3	AB11	57.13875	-2.09089
4	AB12	57.101	-2.1106

Venue data from Foursquare

The number of venues by category (Bakery, Library etc.) will be pulled from Foursquare API. Foursquare is a crowdsourced location data platform, with a user interface as shown in Figure 3. The company make their data available to developers via an API, which we will be using.

Figure 3: Foursquare search interface



London postcode districts vary significantly in size, so we select 1 mile (~1,600 metres) as the radius for the search. With additional data on the approximate radius of each district, we could assign a different radius to each Foursquare search, but for this model we will keep it to one mile. We select 100 venues for each district, as this is the maximum currently returned by the API with our free account.

Method

The accompanying notebook provides full step-by-step details of the method, and the key stages are outlined below.

Cleaning the data

We cleaned the data imported from third party sources to ensure that all columns were labelled correctly and were populated with numerical data.

Using API to request venue data

We made requests to the Foursquare API, to gather the venue data for each district. The radius was set at one mile, and the number of venues set at the maximum available for the free account. This could be improved in future iterations by breaking the districts down

further, and making multiple requests per district with slightly different coordinates, and then deduplicating the results.

One-hot encode venues for correlation analysis

As the venue data is categorical, we one-hot encoded the venues to ensure each was an individual column, with 1 or 0 as appropriate to indicate the venue type.

Perform correlation analysis

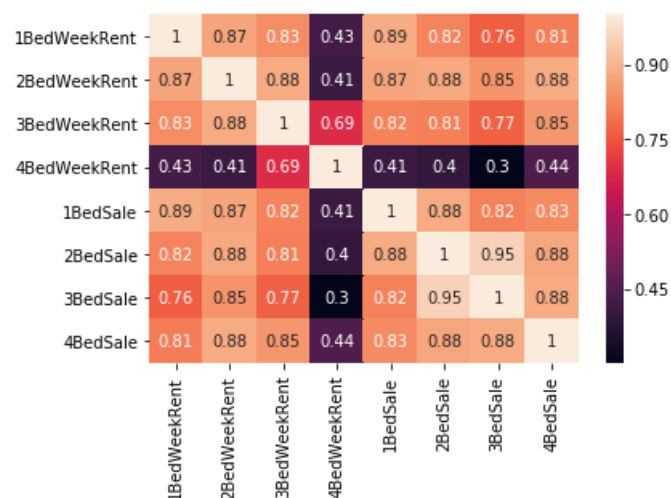
A standard Pearson correlation analysis was then performed, obtaining correlation coefficients and p-values. Correlations with a p-value of higher than 0.05 were discarded to ensure we only keep the most significant correlations. This means that for the results we kept, the probability of these results happening at random is less than 5%.

Results

Correlation of property prices

First we investigated whether the property prices move in the same way by neighbourhood for each of the property types. If they do, this saves us repeating the same analysis for each property type. As shown in Figure 4, prices are mostly very well correlated (all p-values below 0.05), but 4-bed rental properties appear to be less positively correlated with the other property types. Therefore in the next steps, we analyse 4-bed rentals separately from the other property types.

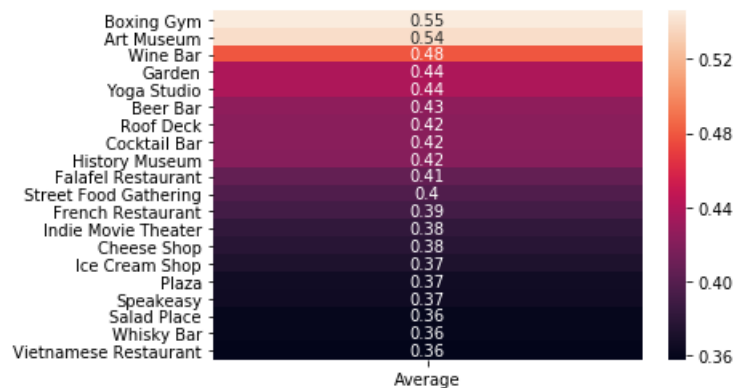
Figure 4. Correlations of prices of different properties, across neighbourhoods



Venues correlated with higher prices

We excluded the 4-bed rentals from the results to start with, and took the average correlation of venue type with property price for the other property types. Figure 5 shows that the expected types of venue – art museums, wine bars – correlate with the more expensive areas. However, there are some unexpected appearances, such as boxing gyms and falafel restaurants.

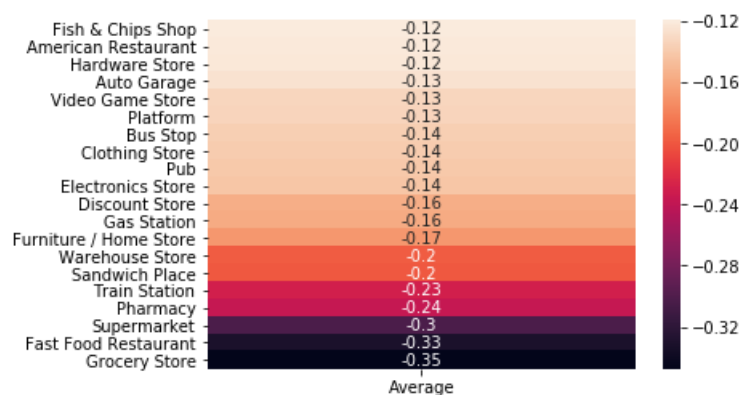
Figure 5. Top 20 venues correlated with higher prices across properties except 4-bed rentals



Venues correlated with lower prices

In the same way, also excluding the 4-bed rentals, we found the venue types with the strongest negative correlations with property prices. Figure 6 shows that supermarkets, fast food restaurants and grocery stores are correlated with lower housing prices.

Figure 6. The 20 venues most negatively correlated with prices across properties except 4-bed rentals



Difference for 4-bed rentals

As shown in Figure 7 and Figure 8, the venues correlating with higher and lower prices are not much different from those for the other property types. However, the correlations are generally less strong, so we can conclude that a different factor - possibly the HMOs - are causing 4-bed rental prices to be less affected by venues.

Figure 7. Top 20 venues correlated with higher prices, for 4-bed rentals

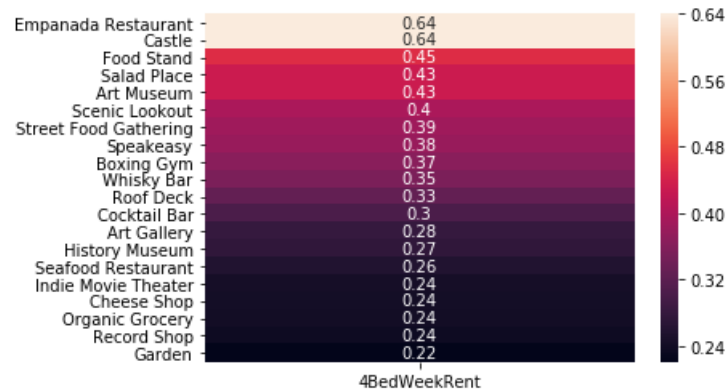
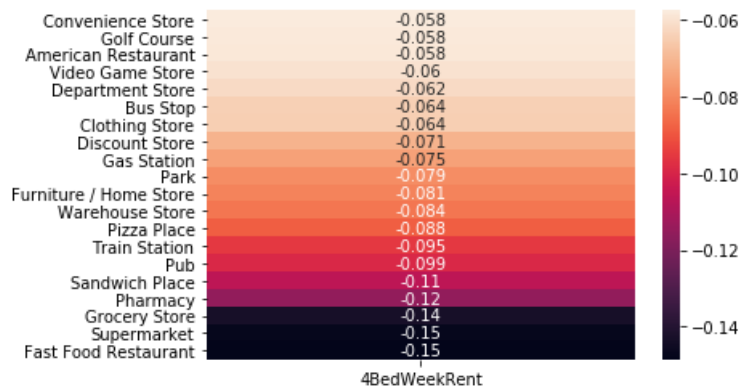


Figure 8. The 20 venues most negatively correlated with prices, for 4-bed rentals



The difference between 4-bed rentals and other properties are greatest for the negatively correlated venues – for example fast food restaurants correlation is half as negative for 4-bed rentals. This indicates that a venue of this type opening up is likely to have a lower negative effect on 4-bed rental prices.

Discussion

From the results, we can see that venues like art museums and wine bars are more common in districts with higher property prices. This is to be expected, but there are a few venues like boxing gyms, beer bars and falafel restaurants that are less obvious.

This may reassure people looking to rent or buy, that new venues of these types opening up are not likely to have a detrimental effect on property prices. However, if a fast food

restaurant or supermarket plans to open up, this may have a negative effect, although the effect may be lower for 4-bed rentals.

There may also be a positive side for prospective buyers or renters who do not mind living near the amenities on the 'negative correlation' list – they may be able to take advantage of lower prices in this area, if they have no personal objection to living near these amenities. However, buyers in particular should note that this may affect their ability and price at which they can sell the property on afterwards.

Conclusion

We have used neighbourhood venue and property price data, to identify the venue types that correlate with higher or lower property prices, and how this changes specifically for 4-bed rentals. We have discussed how this affects homeowners, buyers and renters, and how they can use this information.

To make the analysis more robust, we would need to look at property prices in a neighbourhood before and after introduction of each venue type, which could be analysed in future work.