Where should I live? An analysis of areas in major UK cities

Applied Data Science Capstone Project – The Battle of Neighbourhoods

# Introduction

It is well known that it is becoming increasingly difficult to afford a house in the UK.[1](https://www.ons.gov.uk/peoplepopulationandcommunity/housing/bulletins/housingaffordabilityinenglandandwales/1997to2016) This is particularly noticeable in southern cities such as Oxford (my current hometown), where the cost of living has been driven up significantly above the national average by a number of factors.[2](https://abcfinance.co.uk/blog/the-true-cost-of-living-in-uk-cities/) Part of the reason that these places are desirable are the easy access to venues such as museums, cafes and restaurants, and nature reserves.

### Problem:

Where could someone move that would offer a similar lifestyle to their current location?

In this analysis, we will use Foursquare to analyse the access to venues in different regions of major cities across the UK, and then compare the cost of housing in similar areas. This will allow people who are unable to afford to buy a house in their current area to identify similar areas around the country that might be more affordable.

### Target Audience:

Young people seeking to step onto the property ladder.

# Data

To characterise neighbourhoods, we will use Foursquare to obtain information on the types of venue in each postcode zone (i.e. number of coffee shops, open spaces, museums in EC4). The neighbourhoods will then be clustered using an appropriate algorithm such as K-means to match them up with neighbourhoods in other cities.

This will then be compared with the history of house-price sales in those areas in the last 2 years (see Appendix 1 for the SPARQL query used to obtain these data).[3](https://landregistry.data.gov.uk/app/ppd) Overall average house price, as well as house price for different property types (e.g. terrace, semi-detached) will be the main outcome.

So as an example, if we currently lived in EC4, we would look to see which cluster EC4 resides in. We could then look at average house prices in other postcodes in that cluster to identify a similar area that has a lower average house price.

# References

1. Housing affordability in England and Wales: 2016 <https://www.ons.gov.uk/peoplepopulationandcommunity/housing/bulletins/housingaffordabilityinenglandandwales/1997to2016>
2. The True Cost of Living in UK Cities   
   <https://abcfinance.co.uk/blog/the-true-cost-of-living-in-uk-cities/>
3. HM Land Registry Open Data Price Paid Dataset  
   <https://landregistry.data.gov.uk/app/ppd>

# Appendices

## Appendix 1: SPARQL query to obtain price data

PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

PREFIX text: <http://jena.apache.org/text#>

PREFIX ppd: <http://landregistry.data.gov.uk/def/ppi/>

PREFIX lrcommon: <http://landregistry.data.gov.uk/def/common/>

SELECT ?item ?ppd\_propertyAddress ?ppd\_transactionCategory ?ppd\_transactionDate ?ppd\_estateType ?ppd\_hasTransaction ?ppd\_pricePaid ?ppd\_transactionId ?ppd\_newBuild ?ppd\_propertyAddressCounty ?ppd\_propertyAddressDistrict ?ppd\_propertyAddressLocality ?ppd\_propertyAddressPaon ?ppd\_propertyAddressPostcode ?ppd\_propertyAddressSaon ?ppd\_propertyAddressStreet ?ppd\_propertyAddressTown ?ppd\_propertyType ?ppd\_recordStatus

WHERE

{ { ?ppd\_propertyAddress

text:query ( lrcommon:town "( CITY )" 3000000 ) .

?item ppd:propertyAddress ?ppd\_propertyAddress ;

ppd:estateType lrcommon:freehold ;

ppd:transactionCategory ppd:standardPricePaidTransaction ;

ppd:transactionDate ?ppd\_transactionDate ;

ppd:hasTransaction ?ppd\_hasTransaction ;

ppd:pricePaid ?ppd\_pricePaid ;

ppd:transactionId ?ppd\_transactionId

FILTER ( ?ppd\_transactionDate >= "2018-01-01"^^xsd:date )

FILTER ( ?ppd\_transactionDate <= "2020-02-22"^^xsd:date )

}

OPTIONAL

{ ?item ppd:newBuild ?ppd\_newBuild }

OPTIONAL

{ ?ppd\_propertyAddress

lrcommon:county ?ppd\_propertyAddressCounty

}

OPTIONAL

{ ?ppd\_propertyAddress

lrcommon:district ?ppd\_propertyAddressDistrict

}

OPTIONAL

{ ?ppd\_propertyAddress

lrcommon:locality ?ppd\_propertyAddressLocality

}

OPTIONAL

{ ?ppd\_propertyAddress

lrcommon:paon ?ppd\_propertyAddressPaon

}

OPTIONAL

{ ?ppd\_propertyAddress

lrcommon:postcode ?ppd\_propertyAddressPostcode

}

OPTIONAL

{ ?ppd\_propertyAddress

lrcommon:saon ?ppd\_propertyAddressSaon

}

OPTIONAL

{ ?ppd\_propertyAddress

lrcommon:street ?ppd\_propertyAddressStreet

}

OPTIONAL

{ ?ppd\_propertyAddress

lrcommon:town ?ppd\_propertyAddressTown

}

OPTIONAL

{ ?item ppd:propertyType ?ppd\_propertyType }

OPTIONAL

{ ?item ppd:recordStatus ?ppd\_recordStatus }

BIND(lrcommon:freehold AS ?ppd\_estateType)

BIND(ppd:standardPricePaidTransaction AS ?ppd\_transactionCategory)

}