# Housing Affordability Across England and Wales And Its Relationship to amenities

## Introduction

The UK is generally an expensive place in which to live and work– for England and Wales on average, full-time workers could expect to pay an estimated 7.8 times their annual workplace-based earnings on purchasing a home in 2018. But despite this, housing affordability varies quite a lot across the country. When deciding on where to live in the UK, a potential homeowner needs to take into consideration a variety of factors, including house prices, wage levels, availability of employment, and other factors such as access to leisure activities for instance. The UK Office for National Statistics publish data on house prices and average wages by region, and the data provider FourSquare provide access to an API whereby one can obtain information on local amenities. For this reason I thought it would be instructive to plot out the ONS data alongside amenity information from FourSquare in order to try and see what areas provide a balance of affordability and access to amenities.

The target audience for this study might be a person or family who has to move to a different area of the country, who would like to know more about the area they were moving to. Another audience would be someone who was deciding on a place in which to live. This sort of information could help them in their decision.

## Datasets

### House Price Affordabilty

The Office for National Statistics provide a range of datasets regarding housing affordability across the UK, by region and local areas. These datasets go back over a number of years meaning it is possible it to use for trend analysis. However we will focus only on the data for 2018. The nice thing about these datasets is that they have combined House Prices and Earnings together in order to arrive at a single ‘Affordability Index’ which denotes whether an area is less or more affordable. This Affordability Index is defined as the ratio of House Prices to Earnings. Therefore a lower value of the index denotes a more affordable location.

Earnings data used in the calculation of housing affordability are gross full-time annual earnings where available. Annualised weekly earnings are used when annual earnings are not available.

The house price statistics report the prices paid for residential properties referring to a 12-month period with April in the middle (year-ending September). The earnings data from ASHE provide a snapshot of earnings at April in each year.

The Affordability dataset can be downloaded from the ONS website in Excel format, and comprises several sheets showing median, and lower quartile affordability ratios by areas of varying sizes covering England and Wales. Of particular interest are the sheets showing median ratios by County and by Local Authority District. I wanted to use the County level data for this report, but for reasons that will be discussed later, that approach had to be abandoned and the data at Local Authority level used instead.

An example of the data is shown below:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Local authority name** | **2002** | **2003** | **2004** | **2005** | **2006** | **2007** | **2008** | **2009** | **2010** | **2011** | **2012** | **2013** | **2014** | **2015** | **2016** | **2017** | **2018** |
| County Durham | : | : | : | : | : | : | 5.12 | 4.68 | 4.84 | 4.38 | 4.20 | 4.36 | 4.46 | 4.45 | 4.44 | 4.44 | 4.38 |
| Darlington | 3.42 | 4.06 | 5.12 | 5.97 | 5.94 | 5.49 | 5.88 | 5.46 | 5.45 | 5.16 | 5.26 | 5.34 | 4.76 | 5.18 | 4.84 | 4.96 | 5.29 |
| Hartlepool | 2.86 | 3.02 | 2.82 | 2.86 | 4.14 | 4.63 | 4.70 | 3.98 | 4.75 | 4.32 | 3.98 | 4.79 | 4.69 | 4.74 | 4.62 | 4.82 | 4.57 |
| Middlesbrough | 2.83 | 2.65 | 3.74 | 4.23 | 5.17 | 5.40 | 4.58 | 5.17 | 5.12 | 4.98 | 4.62 | 4.98 | 5.28 | 5.05 | 5.25 | 5.18 | 5.21 |

We will use just the column for 2018, which shows affordability for each area. We could have also used the timeseries data to look at trends over time, but that is outside the scope of this article.

### local amenity data

To obtain information on local facility and amenities I used the FourSquare Places API, making use of the Explore endpoint. This API returns a list of venue recommendations near to a requested location. The API can be called repeatedly to build up a set of venue recommendations for a list of locations. An example of the response returned by the API is shown below:

{

"meta": {

"code": 200,

"requestId": "5ac51ef86a607143de8eg5cb"

},

"response": {

"warning": {

"text": "There aren't a lot of results near you. Try something more general, reset your filters, or expand the search area."

},

"suggestedRadius": 600,

"headerLocation": "Lower East Side",

"headerFullLocation": "Lower East Side, New York",

"headerLocationGranularity": "neighborhood",

"totalResults": 230,

"suggestedBounds": {

"ne": {

"lat": 40.724216906965616,

"lng": -73.9896507407283

},

"sw": {

"lat": 40.72151724718017,

"lng": -73.98693222860872

}

},

"groups": [

{

"type": "Recommended Places",

"name": "recommended",

"items": [

{

"reasons": {

"count": 0,

"items": [

{

"summary": "This spot is popular",

"type": "general",

"reasonName": "globalInteractionReason"

}

]

},

"venue": {

"id": "49b6e8d2f964a52016531fe3",

"name": "Russ & Daughters",

"location": {

"address": "179 E Houston St",

"crossStreet": "btwn Allen & Orchard St",

"lat": 40.72286707707289,

"lng": -73.98829148466851,

"labeledLatLngs": [

{

"label": "display",

"lat": 40.72286707707289,

"lng": -73.98829148466851

}

],

"distance": 130,

"postalCode": "10002",

"cc": "US",

"city": "New York",

"state": "NY",

"country": "United States",

"formattedAddress": [

"179 E Houston St (btwn Allen & Orchard St)",

"New York, NY 10002",

"United States"

]

},

"categories": [

{

"id": "4bf58dd8d48988d1f5941735",

"name": "Gourmet Shop",

"pluralName": "Gourmet Shops",

"shortName": "Gourmet",

"icon": {

"prefix": "https://ss3.4sqi.net/img/categories\_v2/shops/food\_gourmet\_",

"suffix": ".png"

},

"primary": **true**

}

],

"popularityByGeo": 0.9999983845502491,

"venuePage": {

"id": "77298563"

}

}

}

]

}

]

}

}

To use the FourSquare API you need to supply a set of geospatial coordinates (latitude-longitude), for which the Nominatim OpenStreetMap API was used. I made a series of calls to Nominatim supplying the name of the location to obtain coordinates for each Local Authority and attached the results to the Affordability dataset.

### geographic boundary data

I also needed access a set of geographical boundary map data in geojson format, in order to construct Choropleth maps using the Folium library in python. From data.gov.uk I obtained the ‘Counties and Unitary Authorities Generalised (clipped) Boundaries in England and Wales’ dataset. This data comprises boundaries for Local Authority areas across England and Wales. However, when I tried to use it I found that some areas did not exactly match up with the areas used for the affordabilty dataset as some regions were mapped at the higher county level rather than at the administrative authority level, resulting in some areas within the choropleth map being missed. Luckily I came across a useful resource in the form of a Github repository of UK boundary data maintained by Martin Chorley of the Cardiff University School of Computer Science and Infomatics. This is an archive online containing GeoJSON and topoJSON for UK boundary data. It’s all stored on [Github](https://github.com/martinjc/UK-GeoJSON), with a viewer and download site hosted on [Github pages](http://martinjc.github.io/UK-GeoJSON/). In there I found geojson files for the UK, areas broken down into Local Authority Districts (LAD). These files matched the district names used in the Affordability dataset. The link to the boundary data is here:

<http://martinjc.github.io/UK-GeoJSON/>