The Best Medium Sized Urban Areas in the UK

1. Introduction

The purpose of this study was to compare and rank medium sized UK cities as the most suitable places to live, looking at a variety of socio-economic factors: cost of living, living environment type of population and amenities.

This was to be determined using readily available data. The UK ONS (Office for National Statistics) holds a considerable amount of socio-economic data, house prices are available from many property websites and the four square application can show what amenities are available in the areas.

The definition of a city in the UK is a place which has been granted city status by the monarch. This means such places St. Asaph and St. Davids with respective populations of 3,335 and 1,841 are considered cities. While London itself is not considered a city but contains the city of London with a population of 7,375 and the city of Westminster with a population of 219,396, while the urban area of London has a population 9,787,426.

Therefore this study will compare mid sized urban conurbations as defined by the ONS. A number of these areas consist of several recognised towns for instance Farnborough / Aldershot consists of both Farnborough and Aldershot plus Camberley, Frimley, Sandhurst and Farnham.

A number of dimensions were chosen to score the areas and produce a total comparison score will that be used to rank the areas. The first step was to identify dimensions to score each area. Using the guidelines of economic, environment, population type, health and amenities.

Ideally the data available will not be homogenous and successfully differentiate the areas. There is a slight concern that some of the data would not be independent and thus count twice, however if it is an important differentiator then perhaps it should count twice. Some initial analysis of the data was done in order to chose suitable dimensions. The data also be weighted as for instance we would expect a larger city to have more facilities than a smaller city. These were the final choices.

1. Housing costs are a major expense, therefore cheap house prices would be considered important but not only as a cost in themselves but also as an overall indicator of the cost of living in the area.

2. Salaries are also an important economic indicator and could mitigate against high house prices.

3. Amount of natural space.

4. Life Expectancy: It was difficult to identify a general measure of health but life expectancy could be considered an overall measure of health.

5. Percentage of 16-64 year olds in the population: Cities are affected by the population so a young vibrant population would make a young vibrant town.

6. Amenities: There were a number of different types of facilities that might be of interest. A broad spectrum was required including different sport, restaurants and culture. Therefore a number of the following facilities were chosen:

a. Golf courses: golf courses are also social clubs with wide open spaces, so they would be a measure of the social scene and the environment.

b. Gyms: while gyms in themselves are important amenities they are probably used by quite different sections of the community.

c. Sushi Restaurants: while there are a variety of types and certainly quality of restaurants, Sushi restaurants in the UK are not as common as in other countries. Their presence does suggest a level of sophistication and quality.

d. Music Venues: in the four square app the category of music venue is very broad, so would cover the needs of a big section of the community.

2. Data Collection

2.a Identifying Urban Areas

The Wiki page ‘https://en.wikipedia.org/wiki/List_of_urban_areas_in_the_United_Kingdom’ has a list of urban areas in the UK, which was supplied from the ONS; as this was readily available the data was taken direct from this page. This data was used to identify a suitable set of areas. The filter was set for Welsh and English contiguous areas with a population of between 200,00 and 400,000. A number of these areas consist of several recognised towns for instance Farnborough / Aldershot consist of both Farnborough and Aldershot plus Camberley, Frimley, Sandhurst and Farnham.

2.b House Prices

A number of prices could be collected as there are numerous types of properties but for simplicity the average house purchase price was chosen. The data was going to be collected as a csv from https://www.statista.com/statistics/1006395/average-house-price-in-the-uk-by-city but because of the prohibitive cost were manually collected from the site www.rightmove.co.uk, and then typed manually into a CSV, to be picked up by Watson.

2.c Salaries

The Ons has a number of datasets available but in the spreadsheet “PROV - Home Travel To Work Area Table 12.1a Weekly pay - Gross 2019.xls” . Even within this dataset there were a lot of different measures available (gross/net/weekly/monthly/with benefits etc.) and even a breakdown into deciles. The average gross weekly wage was chosen.

2.d The amount of natural space and blue space

The amount of natural space, that is any land cover classified as being natural in type, for example, grassland, heath, scrub, orchards, coniferous trees and so on but does not include inland water bodies, should be considered important. However we should also inlcude Blue space, all inland water bodies, for example, rivers, lakes, ponds, canals and so on. This data can be found in: ‘Extent of natural land cover and blue space in 50 built up areas in Great Britain’

2.e Life Expectancy

The ONS has measures of life expectancy from birth and from the age of 65 in “Life Expectancy at Birth and at Age 65 by Local Areas in England and Wales: 2012 to 2014”

2.f The age profile of the population

The ONS has Life Expectancy at Birth and at Age 65 by Local Areas in England and Wales: 2012 to 2014 that can be used. It identifies the number of both male and female 16-64 year olds and the total number of people in that area.

2.g Amenities

The Four Square application gives the addresses of various amenities by location. We can calculate the longtitude and latitude of each urban area by using the geolocator application and then look for various amenities within 10km of that location, the number of each facility can then be calculated.

The lookup was used and the codes for each facility identified, before calling the data. While it would be good to measure quality of each facility a raw number of facilities was used.

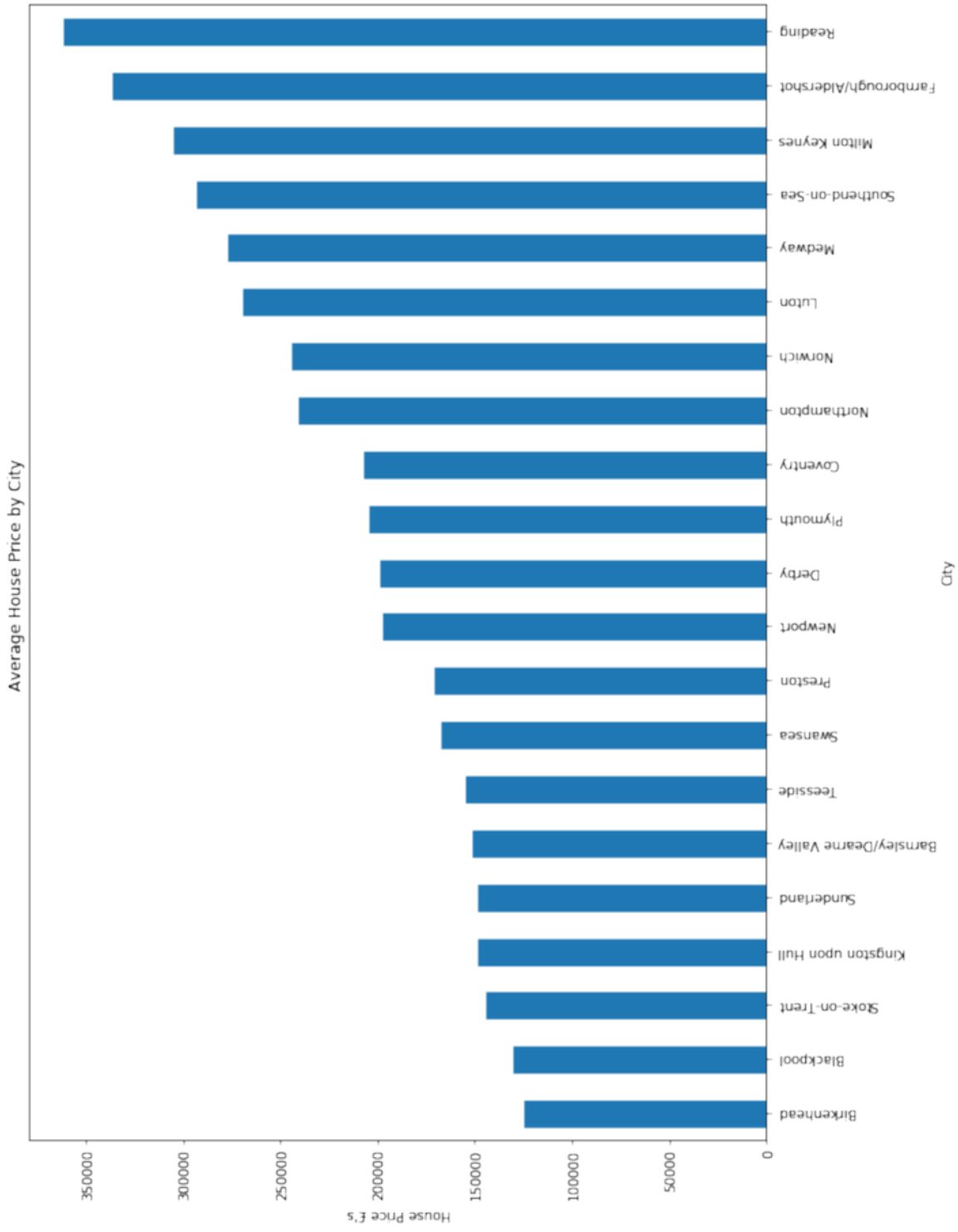
3. Data Collection and Cleaning

Although much of the data was directly taken from ONS websites, some of the data was hard to combine as there were slight differences in naming conventions. The data needed to be cleaned prior to combination as the names of the areas. Many of the spreadsheets held many worksheets and many columns within those worksheets. In order to move this data into CSV further manipulation was required.

4. Scoring

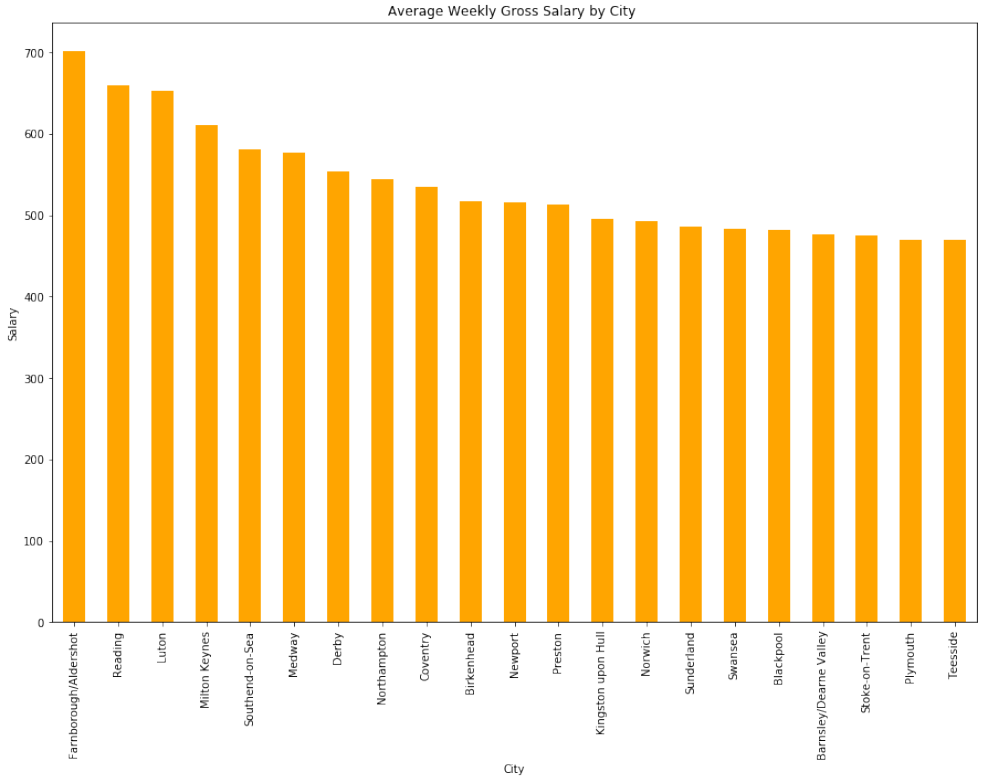
4.a Houseprices

Mean house prices were collected from the csv file and as we can see from below house prices seems quite heterogeneous. The data was then ranked and scores from 0-20 were given from the lowest to the highest, as cheaper housing is a good thing. This data was collected for a final comparison.

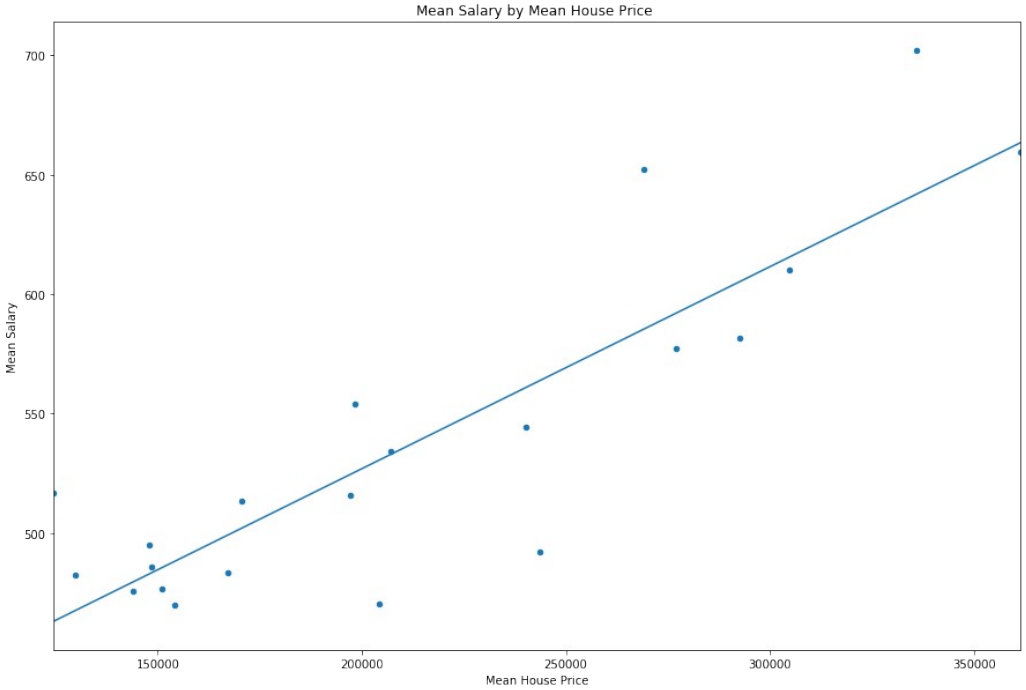


4.b Salaries

Average weekly gross salaries were collected from the ONS csv file and as we can see from below this data is again very different for each city.



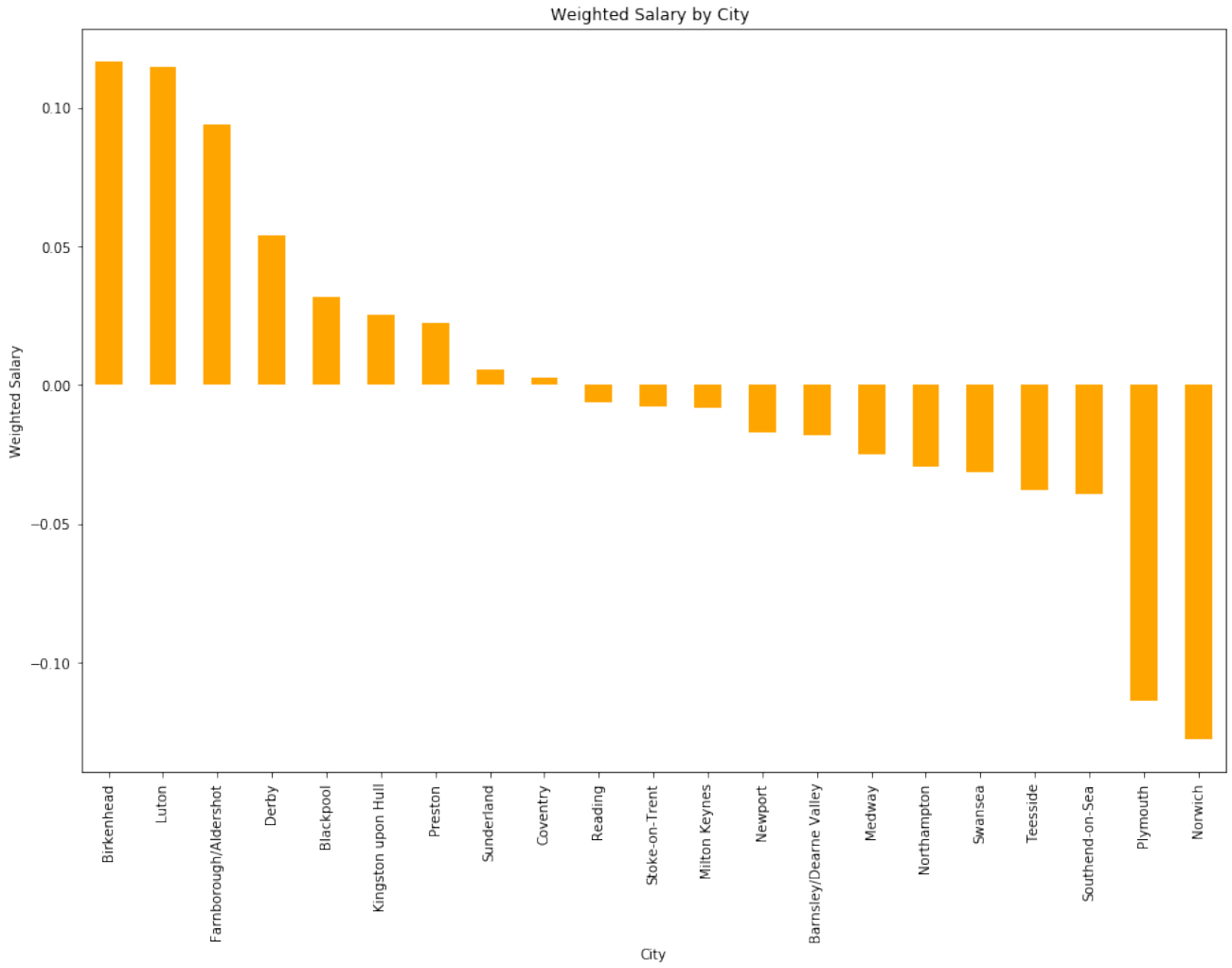
If the cost of housing is a good measure of expenses of the area then to understand if salaries are high compared to the cost of living, we need to understand the relationship between salaries and house prices. Plotting Salaries against House Prices, we get this graph.



By including a regression line on the graph, we can see there is a clear relationship between the data. This makes sense because in order to buy higher price properties people would require a higher salary. Therefore the question could be asked for each city what is the difference between the salary and the expected salary determined by the house price. In order to weight the difference this should be divided by the expected salary determined by the house price.

Weighted Salary = (Observed Salary – Expected Salary) / Expected Salary

Using this new data cities were then ranked and scores from 0-20 were given from highest to the lowest; this data was collected for a final comparison. This gave a good spread of data that could be used to rank cities. This data was collected for a final comparison.

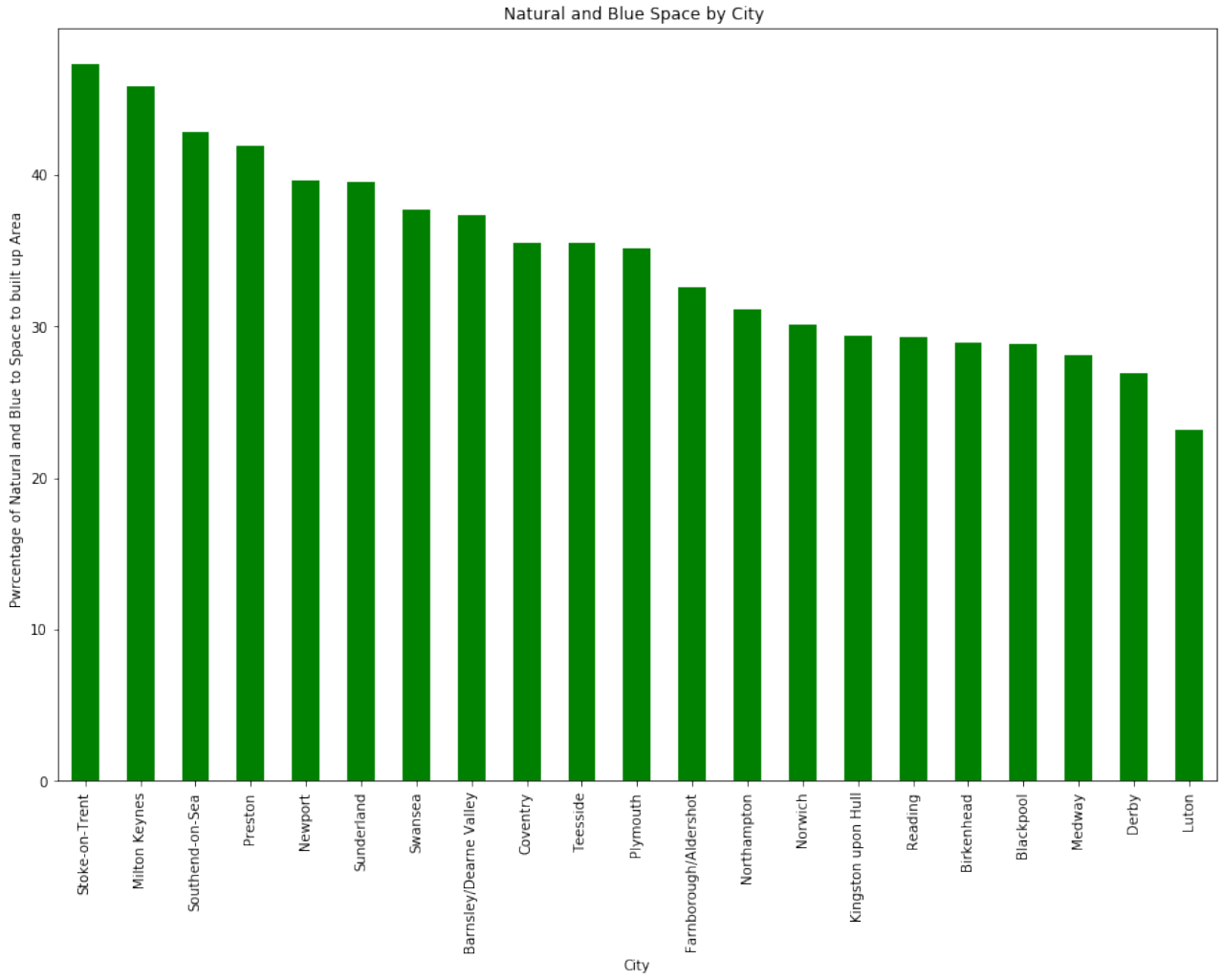


4.c The amount of natural space and blue space

The amount of natural space and blue space was collected from the csv file supplied from the ONS, in order to weight the data this was divided by the built up area of the city, to understand what percentage of the city was natural space.

‘Weighted Space’ = (‘Natural Space’ + ‘Blue Space’) / ‘Built Up Area’

There is good variation of the data suggesting this is a good differentiator for the areas. This data was collected for a final comparison.

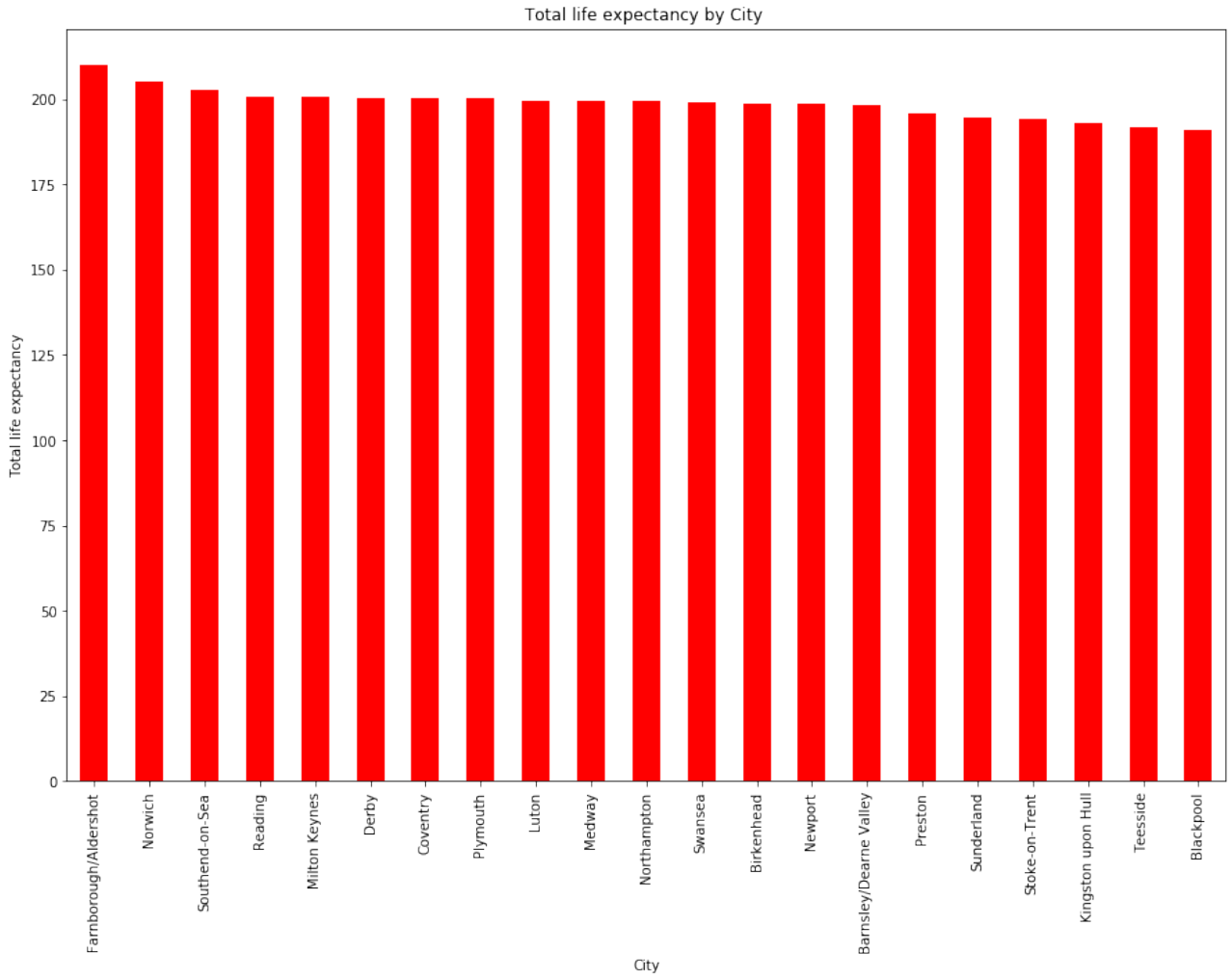


4.d Life Expectancy

Life expectancy data was collected from the csv file from the ONS. There were several figures available and these were combined to get a general measure of life expectancy and therefore health.

‘Total Life Expectancy’= ‘Male life expectancy from birth’ + ‘Male life expectancy from 65’+
‘Female life expectancy from birth’ + ‘Female life expectancy from 65’

The results seem suite homogenous with a range of 190 to 210, however the results range nationally from a total of 190 to 215, suggesting that is not to be unexpected. This data was collected for a final comparison.

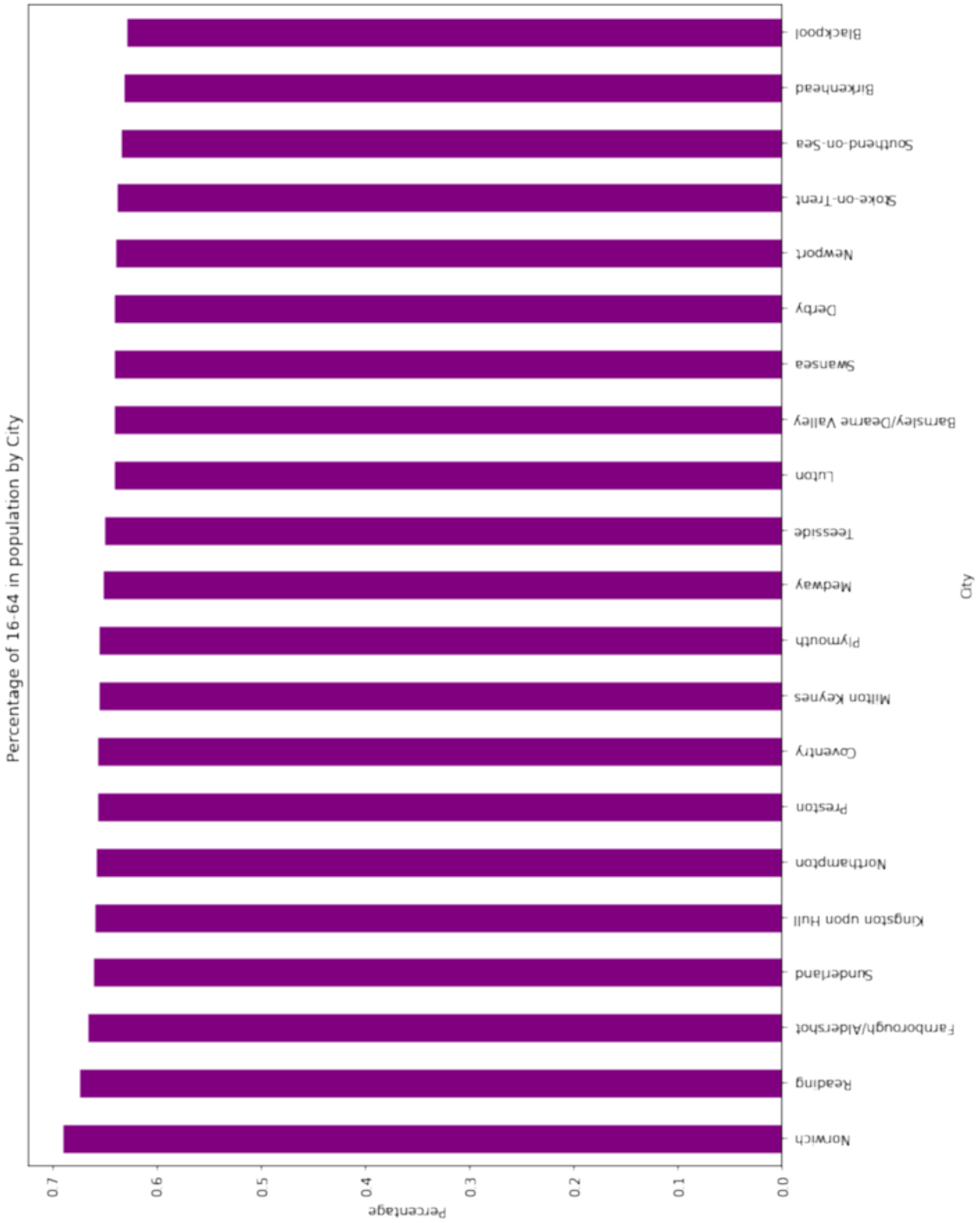


4.e Life Expectancy

Life expectancy data was collected from the csv file. There were several figures available and these were combined. This was the formula used:

‘Percentage of 16-64 yr olds’= (‘No of male of 16-64 yr olds’ + ‘No. of Female of 16-64 yr olds’)/
(‘Total no. of males’ + ‘Total no. of females’)

This data also seems quite homogenous, with a range from 0.61 to 0.68 with that of the general UK population 0.53 to 0.75. This data was collected for a final comparison.

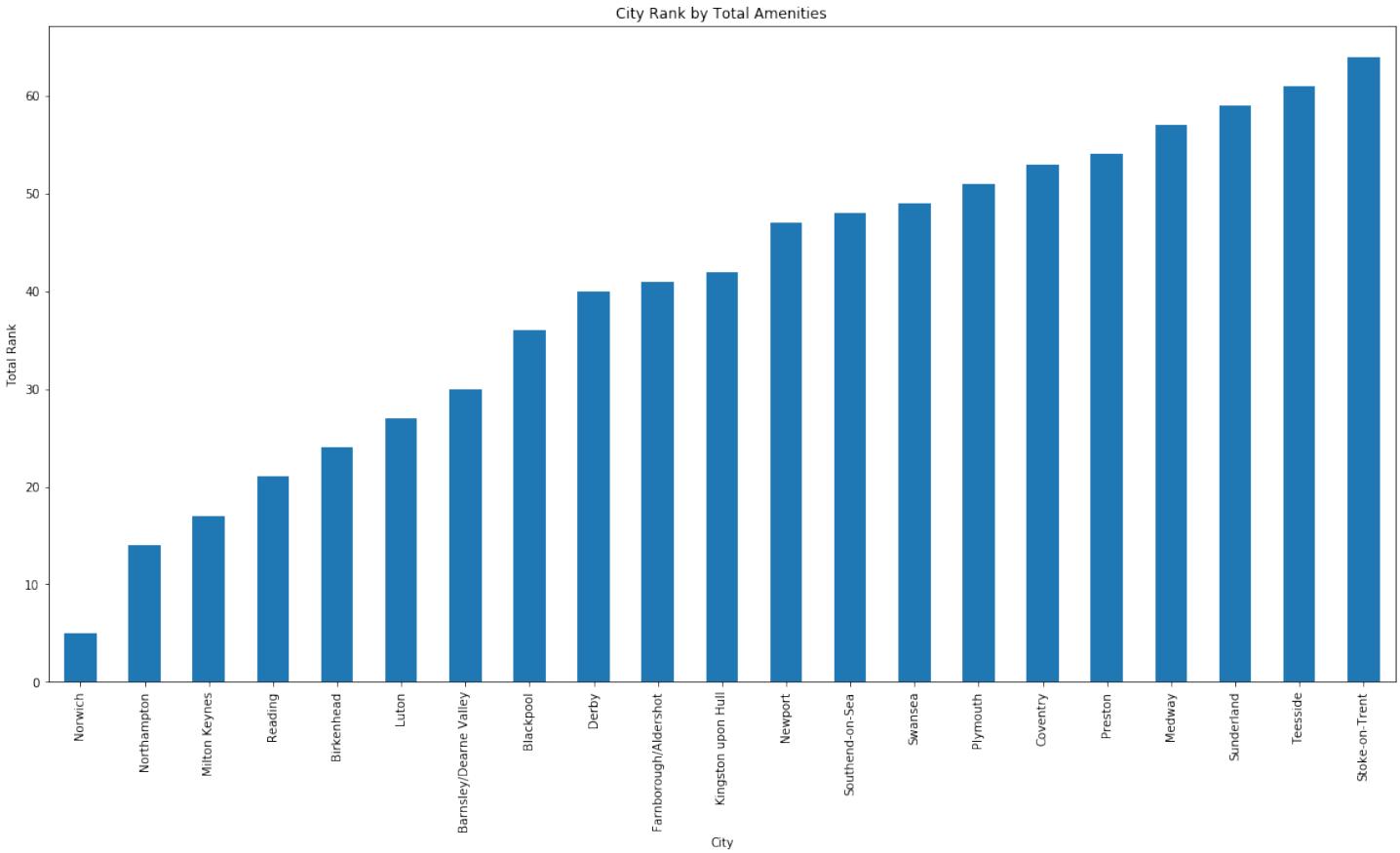


4.f Amenities

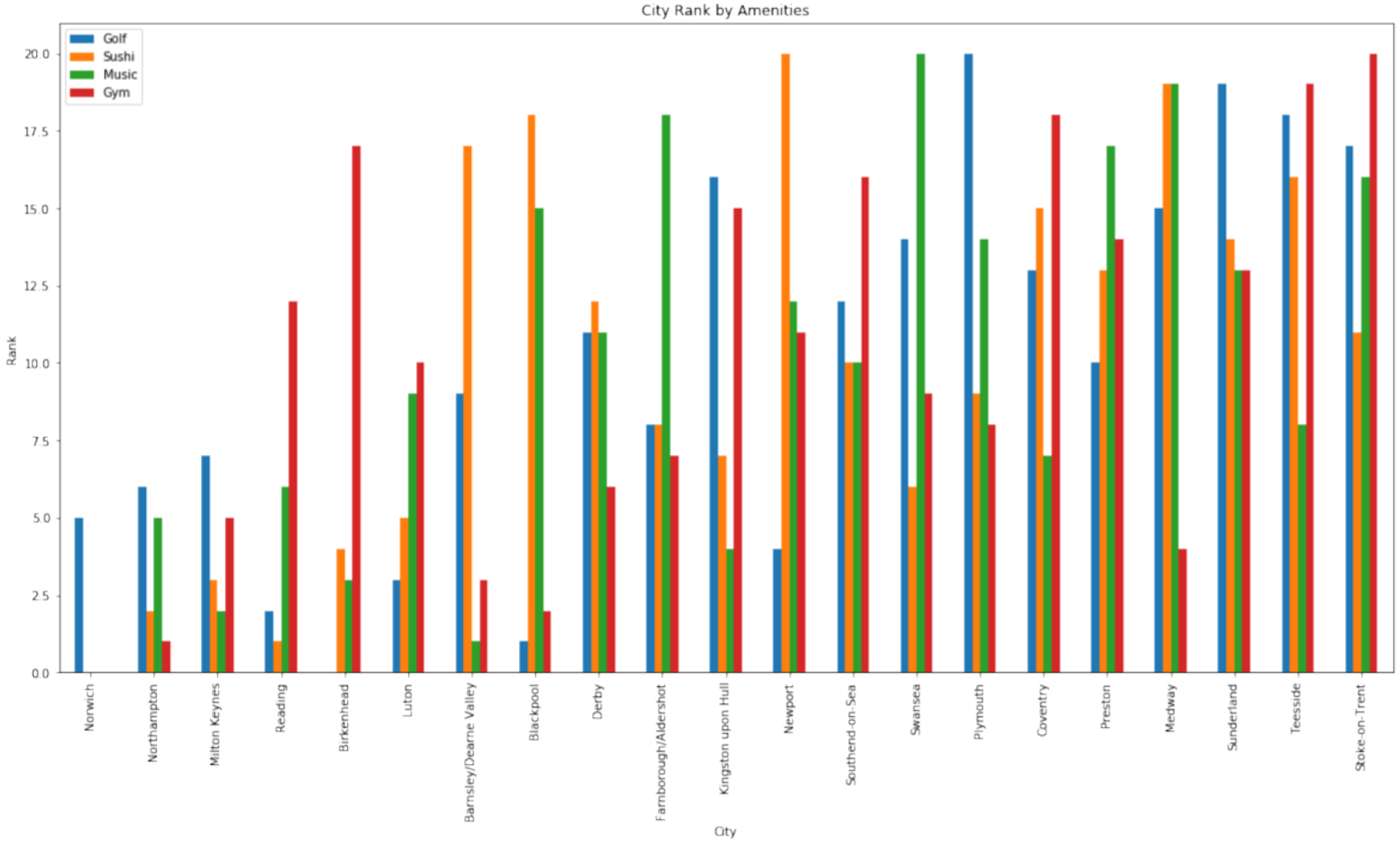
There was a search using the Four Square application to identify the number of different facilities within the urban area. These were number of: golf courses, sushi restaurants, music venues and gyms. In order to weight the data in terms of population the number of each facility was divided by the overall population for the area. The areas were then ranked in terms of the most number of facilities by head of population. The sum of the ranking wee then combined to give one score for all am

Some cities had few Sushi restaurants in fact the maximum number was 5 with several having 0. As we are dividing by the overall population there were no cities having the same score if they had at least one sushi restaurant. To create a better distributed ranking, a score of 0.1 was added to all the number of sushi restaurants; thus if two cities both had no sushi restaurants the ciy with the lower population would be ranked higher.

Looking at the data it looks like there is some dependence between the variables; if an area has many sushi restaurants they would generally have more facilities.



Once the cities were all ranked for the number of amenities per head of population, on overall posiion was calculated using the simple um of all positions. This data fed into the final comparison.


City Rank by Amenities
Golf
Sushi
Music
Gym
Rank
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
Norwich
Northampton
Milton Keynes
Reading
Birkenhead
Luton
Barnsley/Dearne Valley
Blackpool
Derby
Farnborough/Aldershot
Kingston upon Hull
Newport
Southend-on-Sea
Swansea
Plymouth
Coventry
Preston
Medway
Sunderland
Teesside
Stoke-on-Trent
City

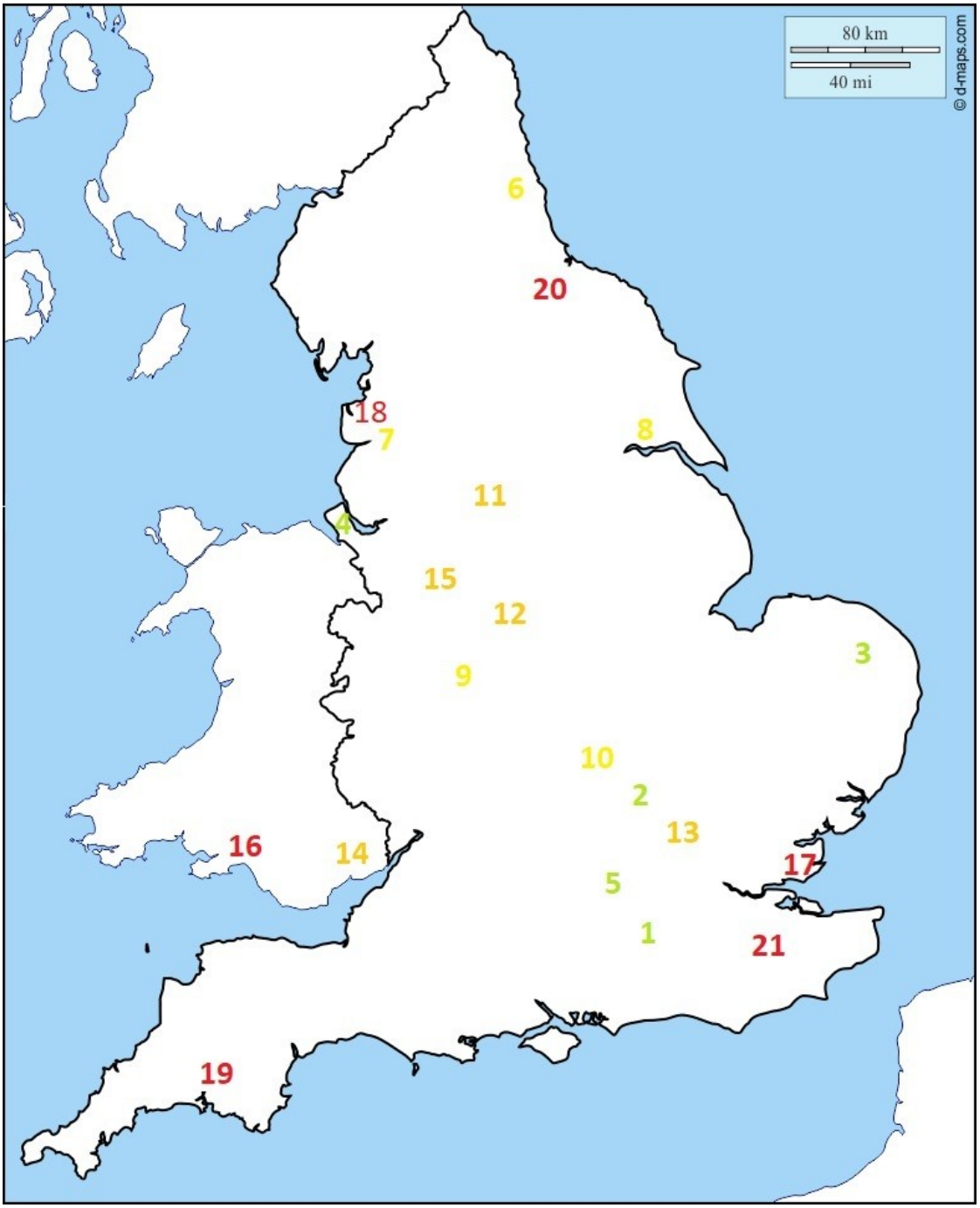
5. Results

Once all he ranking were completed, they were combined to give a number that again could be used to rank the data. All the ranking are presented below with the total ranking added and then sorted by that ranking.

Rank	City	House Prices	Salary	Natural Space	Life Expectancy	Percentage of 16-65	Amenities	Total
1	Farnborough/ Aldershot	19	2	11	0	2	9	43
2	Milton Keynes	18	11	1	4	8	2	44
3	Norwich	14	20	13	1	0	0	48
4	Birkenhead	0	0	16	12	19	4	51
5	Reading	20	9	15	3	1	3	51
6	Sunderland	4	7	5	16	3	18	53
7	Preston	8	6	3	15	6	16	54
8	Kingston upon Hull	3	5	14	18	4	10	54
9	Coventry	12	8	8	6	7	15	56
10	Northampton	13	15	12	10	5	1	56
11	Barnsley/ Dearne Valley	5	13	7	14	13	6	58
12	Derby	10	3	19	5	15	8	60
13	Luton	15	1	20	8	12	5	61
14	Newport	9	12	4	13	16	11	65
15	Stoke-on-Trent	2	10	0	17	17	20	66
16	Swansea	7	16	6	11	14	13	67
17	Southend-on-Sea	17	18	2	2	18	12	69
18	Blackpool	1	4	17	20	20	7	69
19	Plymouth	11	19	10	7	9	14	70
20	Teesside	6	17	9	19	11	19	81
21	Medway	16	14	18	9	10	17	84

These cities are put on the UK map below, with the index being their ranking and using the following colour codes for numbers.

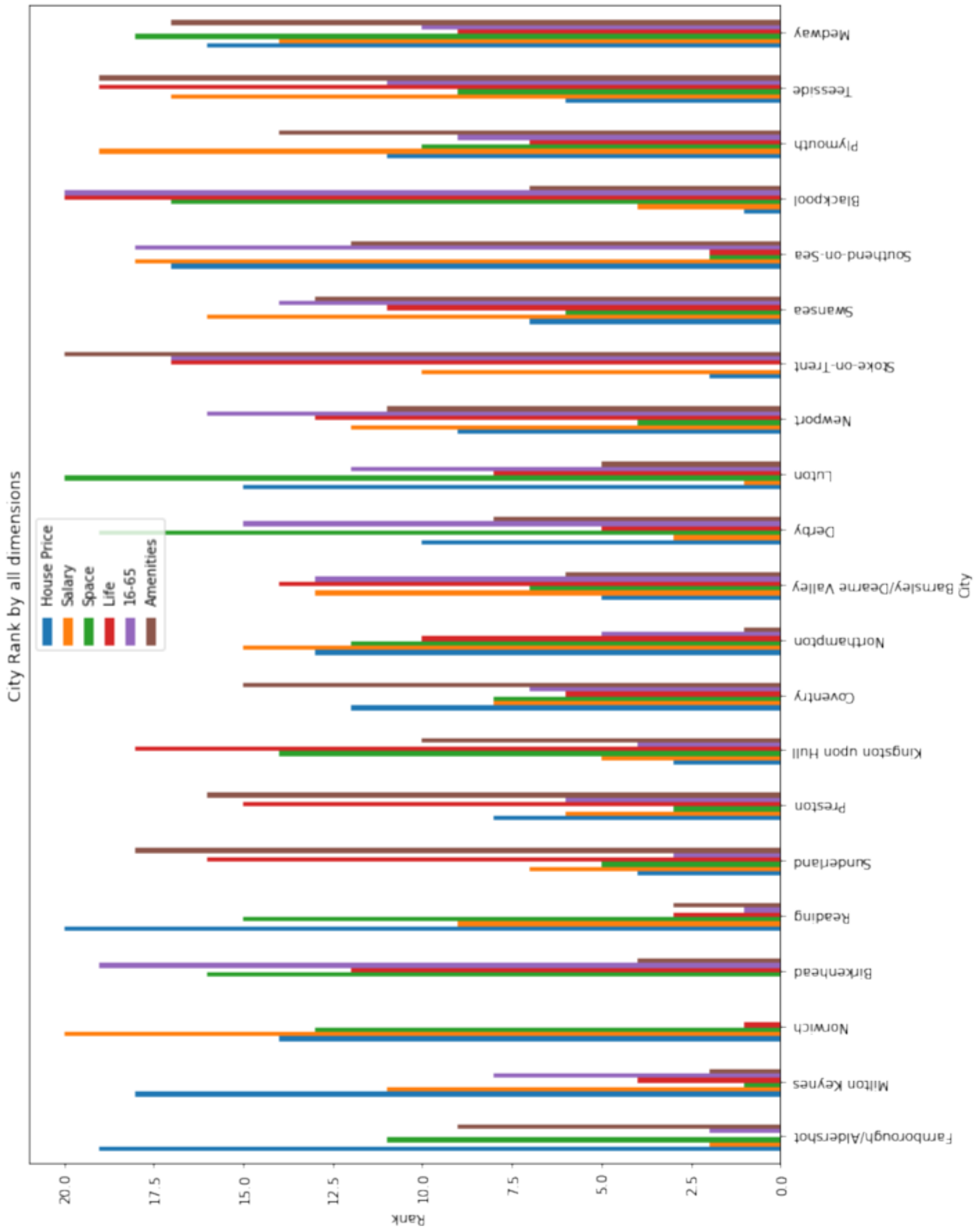
Rank Position	Colour
1-5	Green
6-10	Yellow
11-15	Orange
16-21	Red

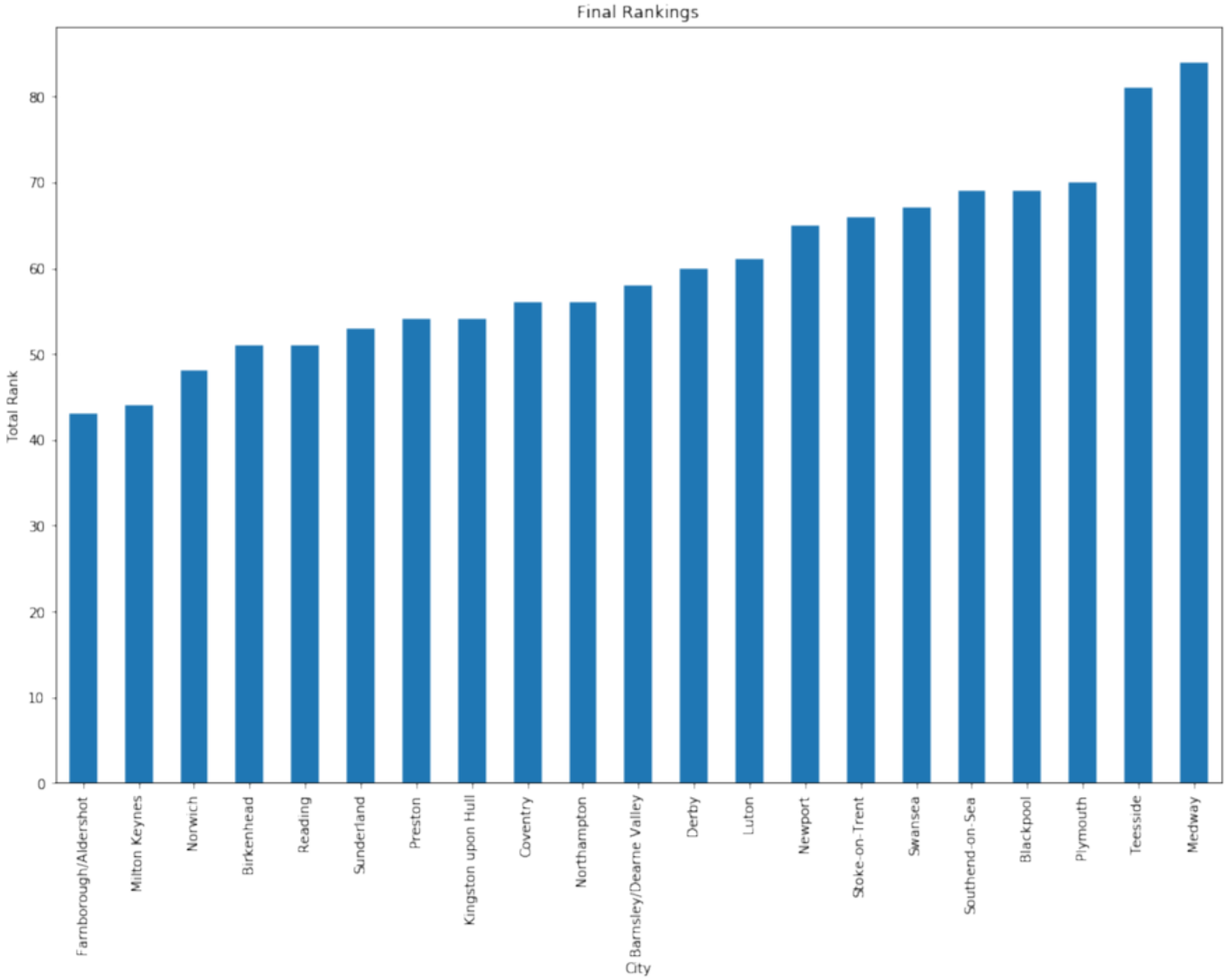


Geographical Mapping of Ranked Urban Areas

Rank	City	Rank	City	Rank	City	Rank	City
1	Farnborough/ Aldershot	6	Sunderland	11	Barnsley/ Dearne Valley	16	Swansea
2	Milton Keynes	7	Preston	12	Derby	17	Southend-on-Sea
3	Norwich	8	Kingston upon Hu	13	Luton	18	Blackpool
4	Birkenhead	9	Coventry	14	Newport	19	Plymouth
5	Reading	10	Northampton	15	Stoke-on-Trent	20	Teesside
						21	Medway

There seems no obvious geographically relationship between their ranking, but four of the top five are in the south east of England, with the majority of the middle ranking cities in the centre of England the worst performing areas were in the extremities of the whole area, with all lying on the coast.




Final Rankings
Total Rank
0
10
20
30
40
50
60
70
80
Farnborough/Aldershot
Milton Keynes
Norwich
Birkenhead
Reading
Sunderland
Preston
Kingston upon Hull
Coventry
Northampton
Barnsley/Dearne Valley
Derby
Luton
Newport
Stoke-on-Trent
Swansea
Southend-on-Sea
Blackpool
Plymouth
Teesside
Medway
City

Discussions

8.1 Report Limitations

There is a limitation to the whole of this analysis due to the choices made in determining how to rank the cities, the choices that have been made are relatively subjective and the following could be reviewed if the analysis were to be progressed:

a. Choice of Scoring Dimensions

The choice of dimension focused on a number of scores that were considered important and were averaged. The price of housing is probably considered important to most people, and average house price has been used to compare between cities. However we have ignored the renting sector and different types of properties, should these dimensions have been used in the comparison.

b. Dimension Independence

While there seemed to an obvious and understandable relationship between house prices and salary, there has only been a cursory investigation of the relationship between other dimensions. If the variables are not independent this could skew data. Further investigation is required.

c. Choice of scoring methods for each dimension

The scoring of each dimension was based on a simple ranking between each city, although this was basec based on a weighted score. Some cities lost a position by only a few points of the score, while other positions were lost by a lot more. For instance the difference in average house price between Preston and Newport was over 26,000 while the difference between Newport and Derby was just over a 1000. However Derby scored 1 less than point than Newport for a score difference of around 1,000 but Newport scored 1 less than point of Preston for a score difference of around 26,000.

City	House Price	Rank	Weighted Score
Preston	170631	8	0.60
Newport	197209	9	0.22
Derby	198245	10	0.21

There are a number of complex algorithms that could be used to score each city and combined to come to a comparison in place of the simple ranking method. If we consider the scores to come from a normal population, we could find an average and standard deviation for the whole population and normalise the score. This assumes we know the nature of the distribution,

Assuming it was a normal distribution it would be possible to provide a weighted score that could be used as a measure.

[weighted score = (score-mean of sample)/ standard deviation of sample. This give thes new following table (See appendix for spreadsheet).

New Rank	Area	TOTAL	Old Rank	Change
1	Luton	3.51	13	2
2	Norwich	2.85	3	2
3	Birkenhead	2.65	4	15
4	Blackpool	2.28	18	1
5	Reading	1.92	5	-4
6	Farnborough/Aldershot	1.86	1	4
7	Northampton	0.69	10	5
8	Derby	0.27	12	1
9	Coventry	0.04	9	5
10	Newport	-0.18	14	1
11	Barnsley/Dearney	-0.19	11	-9
12	Milton	-0.30	2	-5
13	Preston	-0.33	7	8
14	Medway	-0.35	21	3
15	Southend-on-Sea	-0.43	17	-7
16	Kingston/upon/Hull	-1.14	8	4
17	Teesside	-1.19	20	-1
18	Swansea	-2.72	16	1
19	Plymouth	-2.83	19	-13
20	Sunderland	-2.87	6	-5
21	Stoke-on-Trent	-3.54	15	-21

d. Weighing of scores to come to a conclusion

The dimensions have been combined on an equal basis to come to a conclusion. So that the importance of having cheap houses is equivalent to having good facilities as well as all the other factors. But for instance, it might be considered twice as important to have cheap housing as having good facilities, which might be considered three time more important natural space. This algorithin could be easily applied, however we need determine the best approach.

8.2 Possible Solutions

a. Finding more objective dimensions

There are many ways to collect, score and combine data to rank cities. However what might be a good city for a night clubbing, gym obsessed youth who wants to live in a two bedroomed flat might not be a good city for a golf and nature loving pensioner who wants to live in a bungalow in the suburbs. One way would be to collect data on what people think is the best way to score the data to come to a more objective conclusion.

b. Supplying individual ranking

Another way to meet this requirement would be to allow a user to have a choice over their scoring criteria and make the scoring interactive. So one person could rank a city based on the number of gyms, nightclubs but with greater weight to the price of two bedroomed flats and another on the price of Bungalows but with a greater weight applied to number of golf clubs and natural space. This would be similar to many existing applications such as https://www.purina.co.uk/dogs/dog-breeds/breed-selector, which chooses the right dog for you depending on the choices you make.

9. Conclusion

While this study has tried to focus on what is the best mid sized urban area to live in? It has brought the cold world statistics to define in what the end will be an emotional decision. The methodology has been discussed in the previous section.

The study has however answered the question in it’s own limited way. Which is the best place mid-sized Urban area to live in the UK. The answer was Farnborough / Aldershot, it might not be a surprise this area used to have be called ‘the stockbroker belt’. It has also ranked all other cities and it may also not be a surprise to find that 4 of the top 5 are in the affluent south east of England. While the poorest ranked are on the extremities of Britain These areas were ranked despite having the cost of housing working against them.

The data available on the living areas of the UK is ripe for analysis and it would be interesting to understand what makes an area a great place to live and the underlying reasons for that.

Appendix 1

	Weighted	Amenities	Weighted	Houseprice	Weighted	Salary	Weighted	Space	Weighted	16-64	
	Mean	0.00021467	Mean	212843.19	Mean	537.77619	Mean	0.23325402	Mean	0.34585579	
Area	**SD**	4.7344E-05	**SD**	70900.0873	**SD**	68.8941573	**SD**	0.10322246	**SD**	0.30221643	**TOTAL**
Luton	0.00023642	0.45944511	269155	-0.79424175	652.4	1.66376677	0.5547619	3.11470862	0.06516824	-0.9287634	3.51491535
Norwich	0.00033307	2.50101496	243701	-0.4352295	492.1	-0.66299077	0.26400215	0.29788209	0.69357732	1.15057122	2.851248
Birkenhead	0.00026748	1.1154396	124469	1.24646095	516.9	-0.3030183	0.20246769	-0.29825234	0.61422096	0.88799002	2.64861994
Blackpool	0.00024226	0.58291635	129921	1.16956401	482.5	-0.80233495	0.27321631	0.38714718	0.62956129	0.93874941	2.276042
Reading	0.00025156	0.77930885	361464	-2.09620066	659.4	1.7653719	0.26903305	0.34662053	0.68632064	1.1265597	1.92166033
Farnborough/Ald	0.00020206	-0.2662006	336065	-1.73796414	702.2	2.38661472	0.27429688	0.39761551	0.67188208	1.07878415	1.85884964
Northampton	0.00027319	1.23625546	240169	-0.38541292	544.5	0.09759622	0.30199667	0.665966	0.06558586	-0.92738152	0.68702326
Derby	0.00021075	-0.08279426	198245	0.20589806	553.9	0.2340374	0.32054398	0.84564889	0.06424418	-0.93182099	0.2709691
Coventry	0.00018649	-0.59505307	207211	0.07943841	534.4	-0.04900547	0.18491349	-0.46831407	0.66932476	1.07032225	0.03738805
Newport	0.00020206	-0.26631787	197209	0.22051017	515.6	-0.32188783	0.15469547	-0.76106068	0.63347618	0.95170335	-0.17705286
Barnsley/Dearne	0.00025976	0.9525326	151120	0.87056579	476.7	-0.88652206	0.213671	-0.18971671	0.06379537	-0.93330606	-0.18644644
Milton	0.00026529	1.06919104	304728	-1.29597597	610.4	1.05413597	0.21283124	-0.19785213	0.06516768	-0.92876525	-0.29926633
Preston	0.00017554	-0.8264478	170631	0.59537572	513.3	-0.35527237	0.1487395	-0.81876105	0.67090431	1.07554881	-0.3295567
Medway	0.00018448	-0.63761441	276950	-0.9041852	577.2	0.57223734	0.39345603	1.5520073	0.06515995	-0.92879081	-0.34634578
Southend-on-Sea	0.00019302	-0.45725144	292684	-1.12610312	581.6	0.63610343	0.18901527	-0.42857679	0.63084734	0.94300482	-0.4328231
Kingston/upon/Hu	0.00020381	-0.22929311	147987	0.91475473	495.1	-0.61944571	0.2039967	-0.28343949	0.06699563	-0.92271676	-1.14014034
Teesside	0.00016993	-0.94497535	154200	0.82712438	469.7	-0.9881272	0.12225403	-1.07534728	0.64608732	0.99343218	-1.18789327
Swansea	0.00017646	-0.80698692	167156	0.64438835	483.4	-0.78927144	0.14626324	-0.84275058	0.06497819	-0.92939223	-2.72401281
Plymouth	0.0001691	-0.96246331	204261	0.12104626	470.3	-0.97941819	0.22416587	-0.08804433	0.0661104	-0.92564589	-2.83452545
Sunderland	0.00015503	-1.25958454	148342	0.90974769	485.9	-0.75298389	0.14728706	-0.83283201	0.06496294	-0.9294427	-2.86509545
Stoke-on-Trent	0.00015022	-1.36112131	144039	0.97043873	475.8	-0.89958558	0.09672698	-1.32264867	0.064601	-0.93064031	-3.54355715