English Indices of Deprivation

Student Name: *Judith Grieves*
Date: December 2019

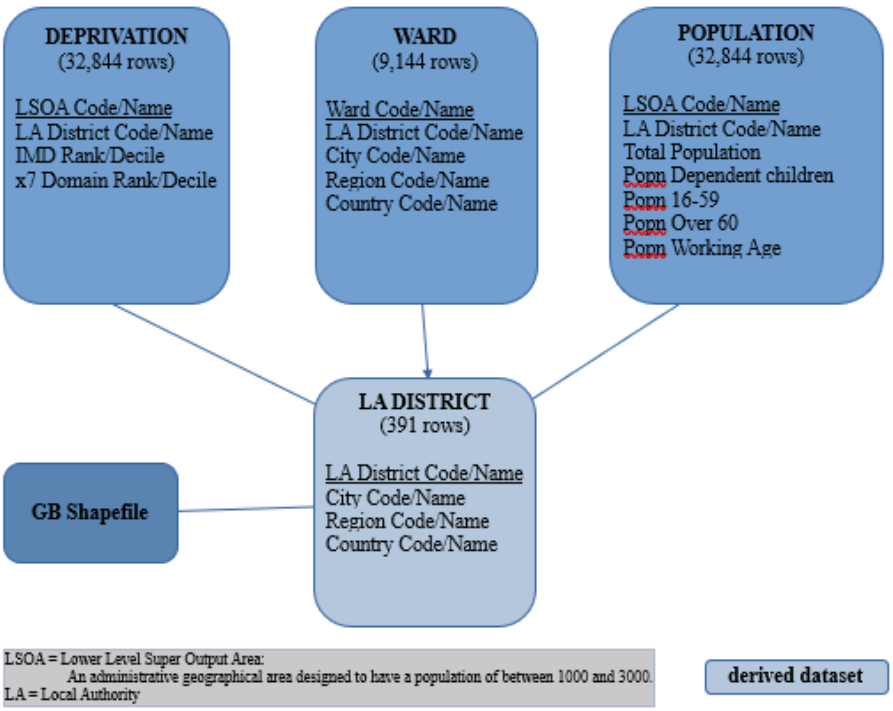
Application domain – Deprivation

This project investigates the English Indices of Deprivation 2019, released by the UK Ministry of Housing, Communities and Local Government and published on the UK government website [1].

The main Deprivation dataset contains measures of deprivation for 32,844 Lower Level Super Output Areas (LSOA) of England in the form of rankings (1=most deprived, 32844=least deprived) and an associated decile (1=most deprived) for each ranking.

A summary measure, 'Index of Multiple Deprivation' (IMD) rank, is derived from 7 individual 'deprivation domains': Income, Employment, Education, Health, Crime, Housing and Environment. These deprivation domains are themselves derived from 39 independently gathered statistics (see Appendix A).

Additional datasets are loaded to provide supporting information (figure 1):



Wards: a geographical hierarchy, allowing the small area information to be analysed by Region

Population: provides population totals and age-related population values for each LSOA (Lower Level Super Output Area).

GB shapefile: to allow the visualisation of the data on a map.

Figure 1: Overview of main datasets

Objectives and Analytical questions

The deprivation of an area is important as 'the best evidence suggests that … there is an effect of area-level deprivation above and beyond … individual socio-economic position' (Nettle, 2015) [3]. The objective of this work is to provide useful guidance on the 2019 Deprivation dataset. It will show overall trends and outliers and understand the individual deprivation indices and their relationships to one another:

1. How does Deprivation vary across England - what variation is there between and within Regions?
2. How do the separate Deprivation domains relate to each other and the IMD?
3. Can Health deprivation be predicted in a population, based on the other indicators?

Expected Output

The summary IMD is derived from the individual domains so we would expect it to accurately identify areas with any type of deprivation.

Health deprivation would be expected to be multifactorial and correlated to other types of deprivation.

Plan

Data Transformations¶

The data quality is good with few missing values.

1. Ward data will be grouped to create a lookup of LA District to LA and Region.
2. There are area code mismatches between datasets – corrections are made in the most appropriate data.
3. Percentages of Population age bands will be derived – these features may be correlated to deprivation.
4. Join the Population, Deprivation and LA district datasets on LSOA in order to create a single master dataset and group this data by LA District. Join the mapping data.

Analysis strategy¶

1. Perform the analysis at LSOA and LA district level, depending on which is most appropriate.
2. Analyse the correlation between the individual domain rankings.
3. Determine what domains Health is most closely correlated with, model the relationship and analyse the fit of any predictive Health model.

Findings and Reflections

How does Deprivation vary in the Data?

This data covers 9 English regions, a total of 54,728,204 people. The regions vary in size between approximately 9 million people and 2.5 million (figure 2). Small areas (LSOA) are approximately 1700 people and Local Authority Districts between 1-200,000.

Region	Total Population
South East	8,940,404
London	8,662,622
North West	7,166,543
East of England	6,068,635
West Midlands	5,747,778
South West	5,467,722
Yorkshire and The Humber	5,383,626
East Midlands	4,670,229
North East	2,620,645

Figure 2 Total Population of England by Region

Aggregation of the Deprivation data to Region level shows that the North East and North West of England are the most deprived areas (figure 3). Nettle (2015) [3] describes how Tyneside, in the North East, has a number of areas in the 1% most deprived.

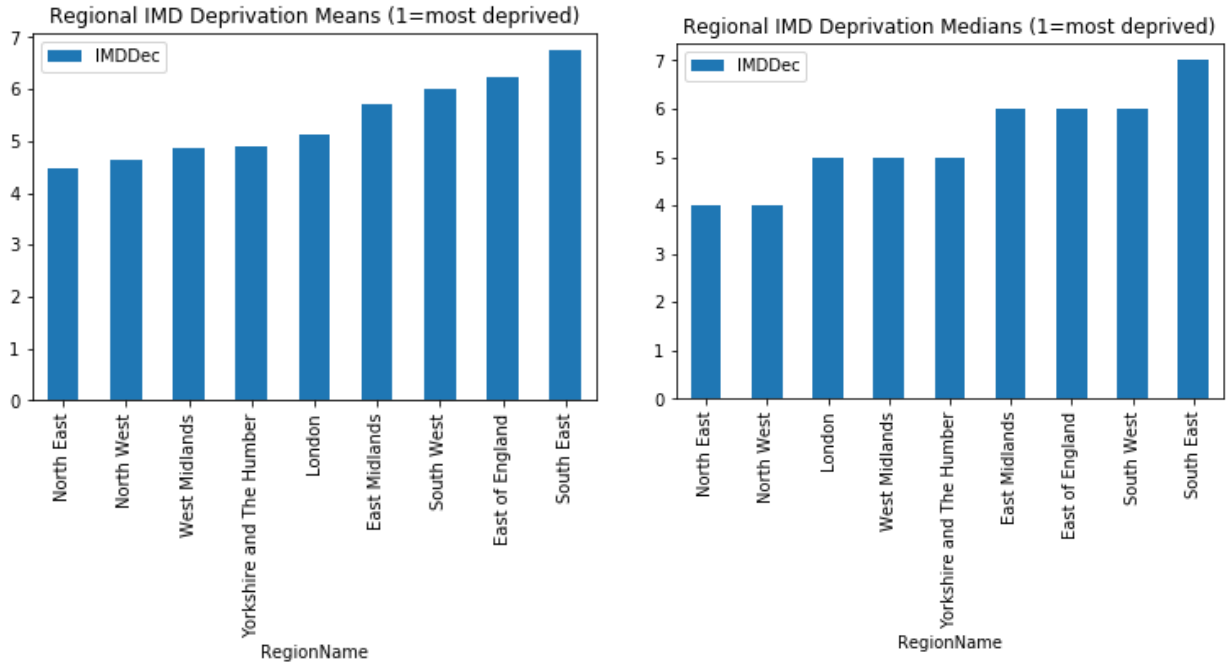
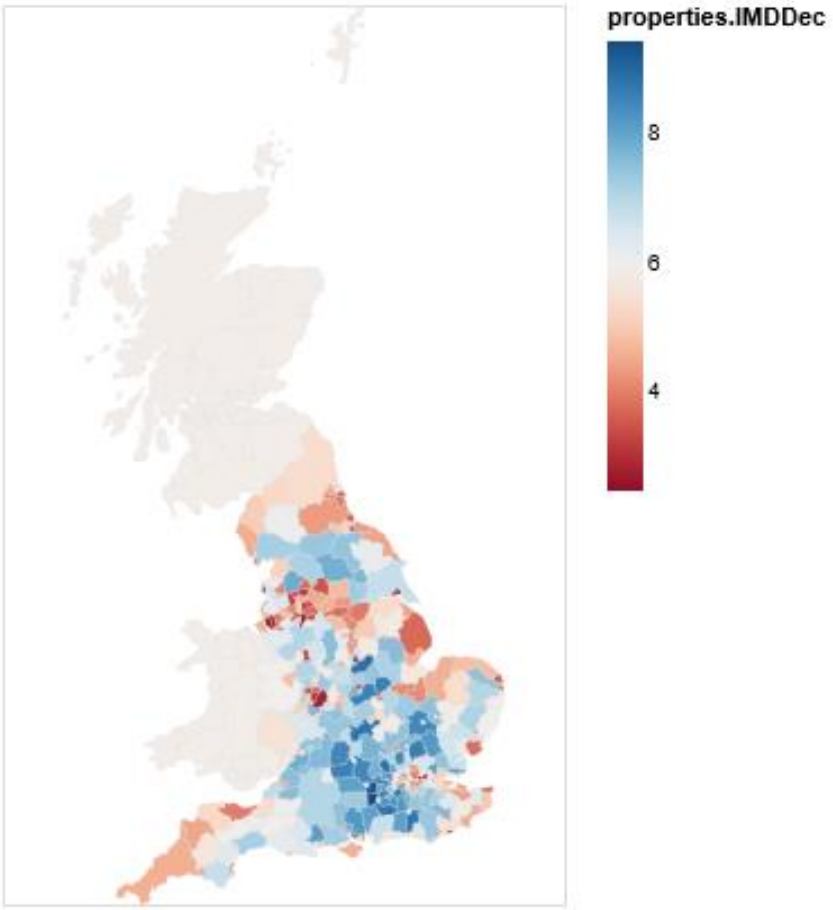


Figure 3: Aggregated Deprivation by Region

The least deprived area is the South East of England; with a relatively narrow range of deprivation within the upper and lower quartiles.

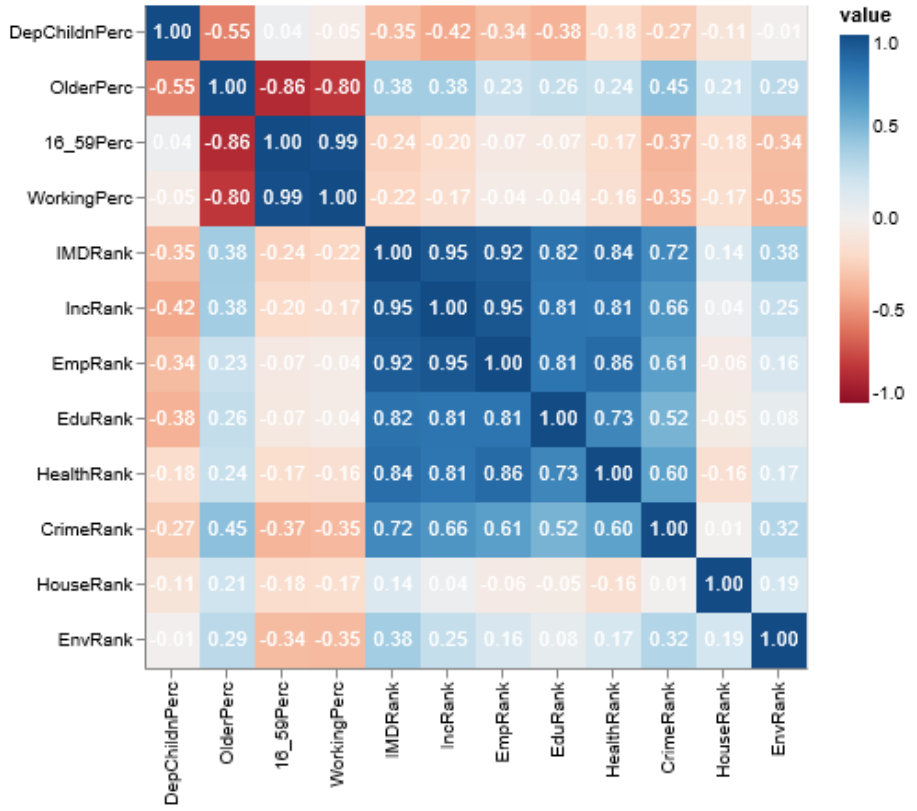


A spatial visualisation of Local Authority Districts (figure 4) shows that, in general, the most deprived districts are in the North of Englan and in coastal areas elsewhere: around the Humber, the Kent coast and Cornwall. Extreme deprivation is seen in the smaller districts which correspond to densely populated cities and conurbations.

However, note also that deprived LA districts are spread throughout the country and there is variation within regions; as Nettle (2015) [3] points out, the most and least deprived areas can often be in very close proximity.

Figure 4: Spatial Deprivation by Local Authority District

How do the Deprivation Domains vary relative to the Summary index?



In order to investigate how the *separate* Deprivation Domains relate to each other and the IMD, the correlation between pairwise attributes was analysed (figure 5) and a ranking of most correlated features compiled.

The IMD is strongly correlated to Income and Employment, and slightly less strongly correlated to Education, Health and Crime. This is a reasonable result as these domains comprise over 80% of the weight of the Index.

These findings are validated by scatterplots of IMD vs domains (figure 6).

Figure 5: Correlation Heatmap of pairwise attributes

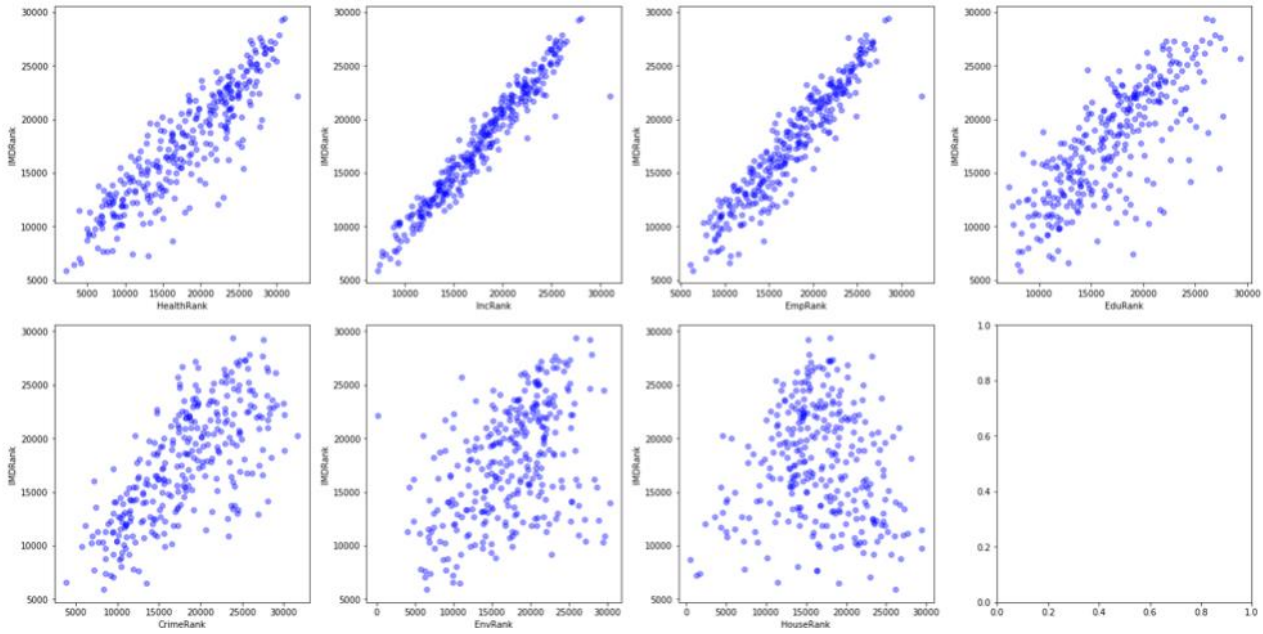


Figure 6: Deprivation Domains plotted against Index of Multiple Deprivation

More surprisingly, despite being a component of the Index, Environment (0.38) and Housing (0.14) are so weakly correlated to IMD.
A spatial analysis of IMD vs Environment and Housing (figure 7) shows very different geographical distributions.

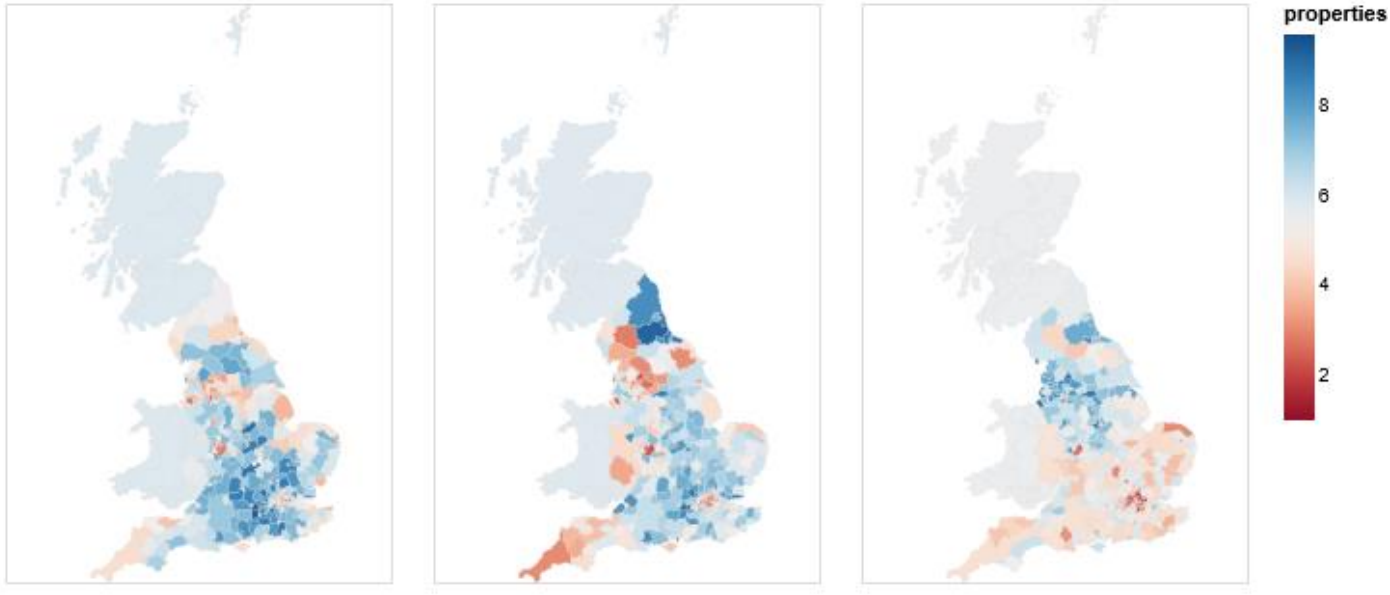


Figure 7 Deprivation by LA District: (a) IMD (b) Environment (c) Housing (1=most deprived)

We can also see the most deprived LA districts for these domains (figure 8) are very different to those for IMD.

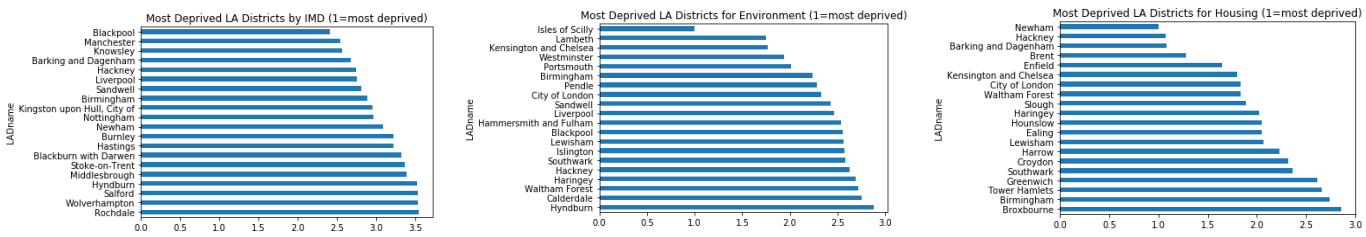


Figure 8: Deprivation by Local Authority District for IMD, Environment & Housing

There does not appear to be any significant relationship between the age profile of an area and the Deprivation domains, although there is a weak correlation showing that Crime (0.45) and Income (0.38) deprivation decreases as the % of older people increases.

What Relationships are there between Individual Domains?

The heatmap (figure 5) also shows that the strongest correlations between domains are:

- Income: Employment (0.95) Health (0.81)
- Employment – Health (0.86)
- Income/Employment – Education (0.81)
- Health – Education (0.73)

Income and Employment are sufficiently correlated to be used as proxies for each other in any analysis and would be suitable candidates to predict Health.

Crime is less correlated to the other domains, suggesting that other factors not present in this data are at play.

- Crime vs Income (0.66), Employment (0.61), Health (0.60), Education (0.52)

Housing and Environment are uncorrelated to each other and the other domains.

Analysis of the individual domains of Deprivation across regions is performed (figure 9).

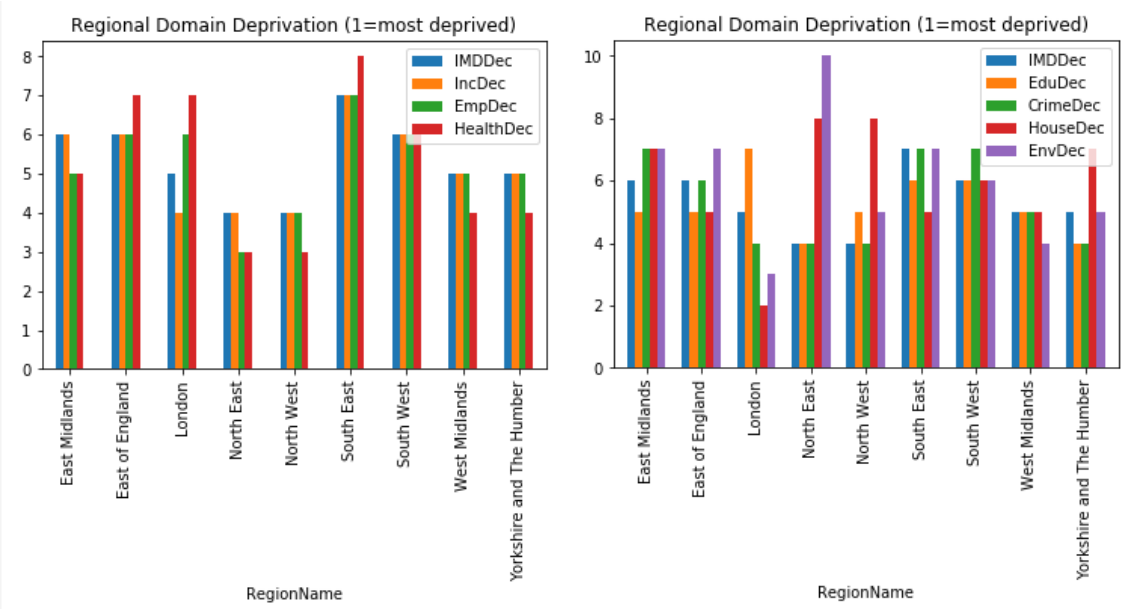


Figure 9: Comparison of Domains of Deprivation by Region

Note that although Income, Employment and Health are similar for most regions, the results show disparities for Health:

1. Less deprivation than the IMD would suggest in London, South East and East of England.
2. Worse than expected in the West Midlands, Yorkshire and the Humber and the North West.
3. Similar to Employment for the North East, East Midlands and South West.

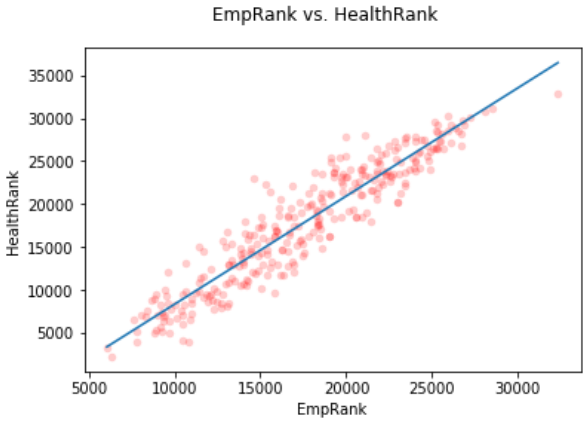
For the remaining domains, most Regions have similar levels of deprivation for each (relative to IMD) – but again there are exceptions:

1. London - less Educational deprivation but worse for Housing

2. North East - less Housing and Environmental deprivation
3. North West/Yorkshire & the Humber - less Housing deprivation

Can Health Deprivation be predicted in a population, based upon other indicators?

We construct a Linear Regression Model to examine whether Health Deprivation can be predicted. Employment deprivation is the most highly correlated of the remaining domains to Health and the other domains are co-correlated so we will exclude them and fit the model with a single predictor variable.




```
Slope:  1.2614341929327784
Intercept:  -4309.530347918717
r_value:  0.9372462293292977
p_value:  1.9497153494791455e-143
std_err:  0.02669586409139484
```

Figure 10: Linear Regression fit of Health Deprivation from Employment

The fitted model (figure 10) returns an r-value of 0.93 and very low p-value which shows a good fit with low residual errors.

Following the regional variations noted above, linear regressions were also performed with subsets of Regions but no better fit was obtained.

Conclusions

We have seen that the IMD is highly correlated to Income, Employment and Health which, in turn, are strongly correlated with each other – this finding should be of interest to policy makers. The IMD is a suitable proxy for policy or decisions related to these domains.

Environment and Housing deprivation are not significantly correlated to the other domains and, consequently, to the IMD. Users of this data with a focus on Housing or Environment should use the individual domain's data in preference to the IMD.

There is significant Regional variation in the data; overall deprivation varies by Region but also the relationship between individual types of deprivation can be localised. Users of the data within regions should consider this when performing further analysis.

Further Analysis

This data is rich and detailed but the data was aggregated for analysis - variance at more granular levels should be investigated further.

Appendix A

Figure 3: Summary of the domains, indicators and data used to create the Indices of Deprivation 2019

References¶

1. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019
2. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/835115/IoD2019_Statistical_Release.pdf
3. Nettle, D. 2015, Tyneside neighbourhoods: deprivation, social life and social behaviour in one British city, 1st edn, Open Book Publishers, Cambridge, England.