Attempting to identify correlates of population loneliness in the UK

David Ireland   Gavin Leech   Grant Stevens   Stefan Radic Webster   Dan Whettam

Abstract

Attention has been paid to a putative 'loneliness epidemic' across the Western world. However, direct data on population loneliness is unavailable. We extend an ONS analysis using the prevalence of GP prescriptions as a proxy for regional loneliness, and track the relationship between this proxy and university attendance, socioeconomic deprivation, educational attainment, and migration.

Index Terms

medical prescriptions, loneliness, official statistics, geostatistics, education, migration, university

I. Introduction

The Data Science Campus in collaboration with the Social Analysis Team in the Office for National Statistics (ONS) produced a loneliness prescription index as an indicator of loneliness levels in the UK [7]. The outcome-based loneliness index was constructed using open prescription data from General Practices (GPs) to help understand loneliness disparity in the UK by creating maps to identify geographical clusters for loneliness.

They used the relative prevalence of depression, Alzheimer's, hypertension, insomnia, addiction, and social anxiety as a proxy for loneliness. The first four of these are heavily confounded with age, so analyses using age as a feature will be double counting.

A key limitation of the loneliness prescription index is that it is a double proxy: it uses the prevalence of certain illnesses as a proxy for loneliness, and uses prescription data as a proxy for these illnesses. Moreover, as many medications are prescribed for different conditions, including 'off-label' uses, it is difficult to formally capture. This could be controlled for with distributions over the target conditions of each medicine but such data is unavailable.

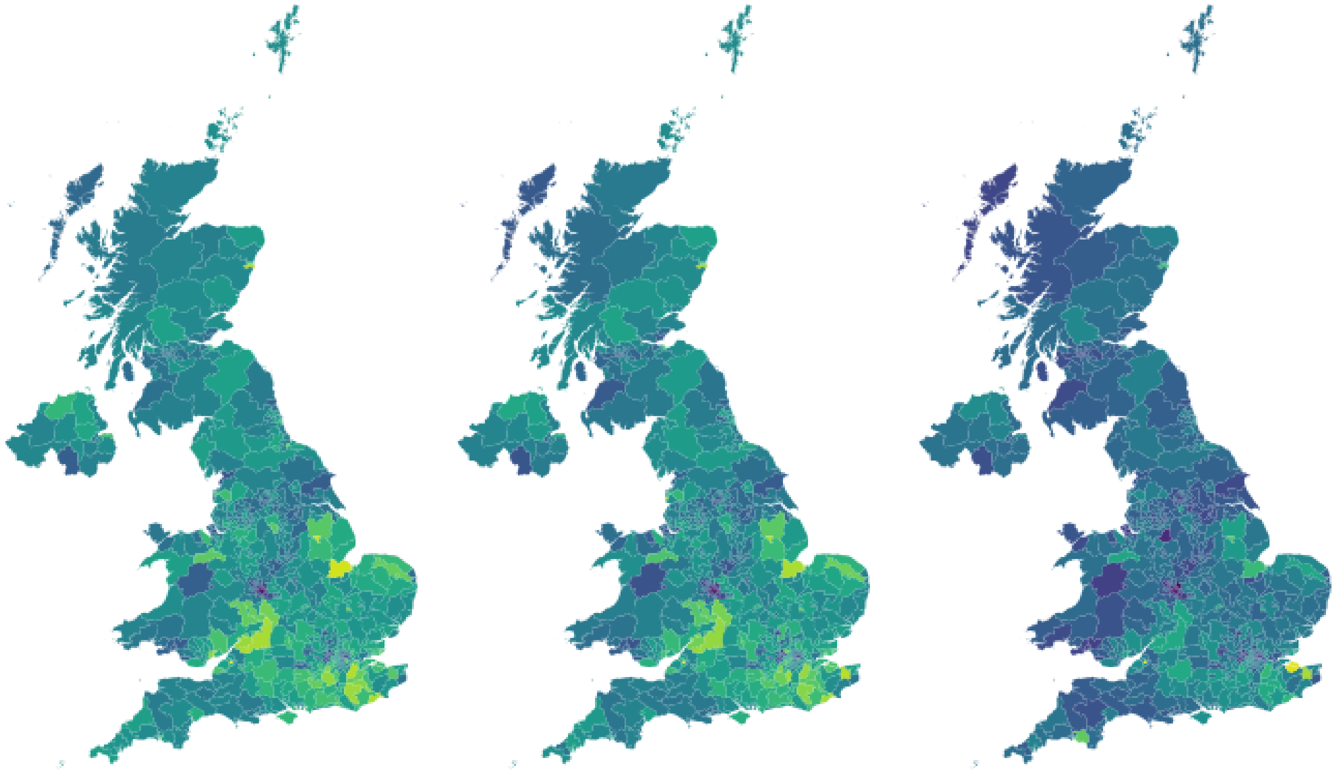
In light of the drawbacks of the loneliness prescription index, we investigated other indicators of loneliness to determine whether they backup the identified loneliness hot spots. As highlighted by [7] loneliness arises due to a lack of desired social interaction or emotional support. Known indicators of loneliness are areas where a higher proportion of people live alone, in poor living conditions or have barriers to housing and services as well as lower incomes and higher crime rates. We are investigating whether deprivation, educational attainment and internal migration are also indicators of loneliness by proposing the following research questions:

- *School quality and attainment*. The OFSTED ratings, attendance and attainment at schools vary geographically which may impact on social cohesion causing loneliness. Can school data indicate loneliness in specific regions?
- *University offers by region*. University acceptance rates differ across the country. Does this variation in acceptance rate have any correlation with loneliness in different geographical locations?
- *Net migration*. A large part of community cohesion and social interaction are the bonds that are formed in communities over time. However, with the increase in people moving longer distances for work and university, communities can be involved in a constant reshuffle of residents. Does the inflow and outflow of population indicate regional loneliness?
- *Socioeconomic deprivation*. Many forms of social participation require resources (for instance transport costs and flexible working) and so economic deprivation might be expected to correlate with loneliness. Does the regional Index of Multiple Deprivation indicate loneliness?

All code for this analysis can be found on the team Github repository.

II. Data Preparation

The basis for our work in this project was formed from the original code found in this Github repository. We obtained prescription data for 2016-9, matched the medicine codes to the target set of 'lonely' illnesses, matched the prescriptions to GP practices in order to obtain a postcode, and then matched the postcode to larger geographical features (constituency and MSOA). The prescriptions per region were then z-score standardised. (See Appendix A for full listing of data sources). The loneliness index was then simply the sum of standardised scores of prescriptions for the 6 target illnesses.



2017 2018 2019

Fig. 1: Map of our loneliness score across 2017, 2018, and 2019

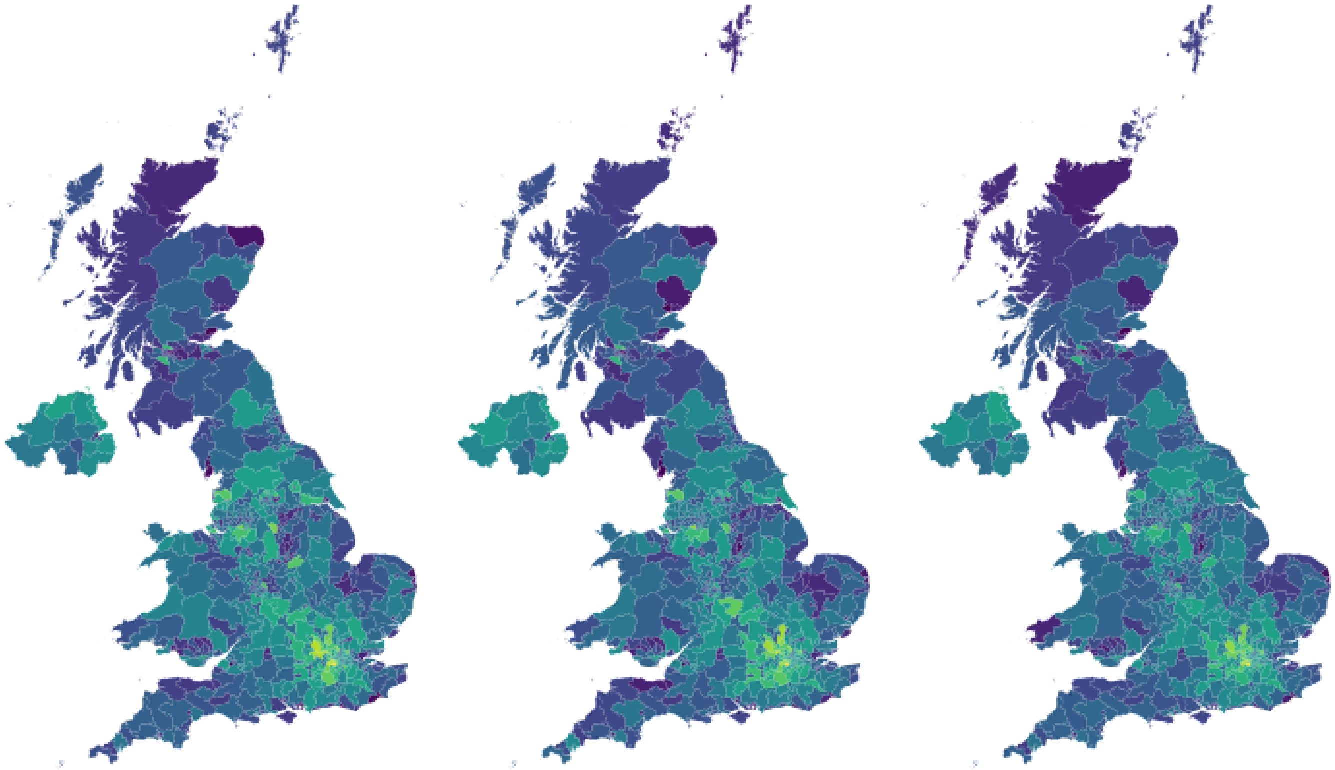
III. Data Exploration

First we explore the loneliness proxy itself. Figure 1 shows a UK-wide visualisation of the proxy, where larger values are a brighter colour. Because the data is z-score standardised, these figures show relative loneliness, not absolute values. Because of this relativity, we cannot make conclusions by directly comparing these figures across the three years without any additional statistics (e.g. averages). We can, however, use these figures to identify hot-spots relative to the rest of the UK, as well as compare these figures to some of the measures we explore in this report (education, migration etc.), as long as we are not comparing across different years. What we can directly observe from this Figure is that our loneliness proxy is much higher in the south of the UK than elsewhere for 2017 and 2018. In 2019, however, our loneliness proxy becomes more constant across the entire country. Investigating this further, we looked at the mean and standard deviation for the entire UK for each year.

Year	Mean	Standard Dev.
2017	0.05	1.22
2018	0.30	1.25
2019	0.64	1.39

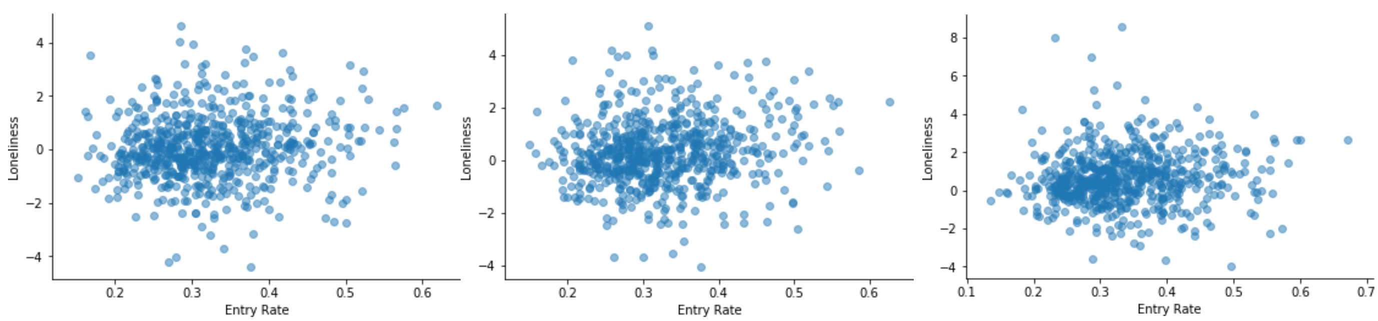
TABLE I: Mean and standard deviation of the loneliness proxy across 2017, 2018 and 2019. Higher values correspond to a brighter colour.

Table I shows that there is a significant increase in the mean loneliness proxy over time, as well as a more gradual increase in the standard deviation. This increase in mean over time demonstrates that whilst there hot-spots in the south of the country appear to be less distinct, the overall trend is an increase in our loneliness proxy. The gradual increase in standard deviation is not trivial from observing the figures - 2019 appears to have fewer hot-spots, so why is the standard deviation increasing? The answer to this is not entirely certain, however, it seems that for 2019 there is a large amount of variation between the dark sections in Scotland and Wales, and the few bright spots in the South East of England. Whilst the country as a whole appears to be more uniform, there are more extreme values in 2019. This sort variation may have caused the increase in standard deviation between years



2017 2018 2019

Fig. 2: Map of university acceptance rates across 2017, 2018, and 2019



2017 2018 2019

Fig. 3: Scatter plots of university entrance rate and loneliness across 2017,2018, and 2019. Each data-point corresponds to a UK constituency. Higher values correspond to a brighter colour.

A. *Universities*

There is a large amount of data available on University attendance and acceptance rates[1] through UCAS. This data covers a wide range of areas related to university acceptance including clearing data, geographic data, acceptance rates and types of offers. For this project we have explored the data on university acceptance rates within UK parliamentary constituencies. University acceptance rate provides a measure of the number of people from an area attending university. It seems probable that the number of people from an area that attend university may correlate with their loneliness - if someone does not have any friends or family that are at university, or have attended university, attending university themselves will feel much more foreign and lonely.

[1] https://www.ucas.com/data-and-analysis/undergraduate-statistics-and-reports/ucas-undergraduate-sector-level-end-cycle-data-resources-2019

The UCAS data we've used consists of acceptance rates for 18 year-olds within each constituency. We can use this to visualise university acceptance rates across the country and gain an understanding of if and how acceptance rates may correlate with loneliness.

Year	Pearson's r	p-value
2017	0.10	$p < 0.016$
2018	0.13	$p < 0.0007$
2019	0.08	$p < 0.049$

TABLE II: Correlation of university acceptance rates with loneliness score

We began by plotting the university acceptance rates with our loneliness index for 2017, 2018 and 2019. Figure 3 shows that there is little to no correlation between university entrance rate and our loneliness index. In 2019 there is perhaps a slight positive correlation, but it is minimal, and multiple outliers with a high loneliness score likely obfuscate any correlation. To confirm the lack of correlation between the two variables we performed a Pearson's rank test (Table II).

The results in Table II confirm what we see in Figure 3, with little to no correlation seen across the three years. Figure 2 shows maps of acceptance rates across the UK for 2017, 2018 and 2019. A comparison between figures 1 and 2 supports our correlation analysis. Whilst there are hot-spots of high acceptance rates around London and the south in both figures, university acceptance seems to be more wide-spread and seems to correlate with locations of large student cities (e.g. London, Birmingham, Manchester).

Across multiple forms of analysis and visualisation it is clear that there is little to no correlation between university acceptance rates and the prescription loneliness proxy. This lack of correlation suggests that university acceptance is unlikely to have any sort of systemic, widespread impact on loneliness across the UK. There are, however, some subtleties within the data that may be contributing to the lack of correlation. Primarily, the constituencies given by the UCAS acceptance data are of the student's location prior to attending university. If, instead, UCAS provided the same data upon moving to university, there may be some indication of students attending less prestigious universities being more lonely. Universities lower in the league tables are likely to have higher numbers of students admitted through clearing, due to not receiving the required grades to study at their first choice of institution, and therefore may go on to experience higher levels of loneliness. Further analysis of university clearing data may offer more insight into the topic. Unfortunately, whilst UCAS provide large amounts of data it is very difficult to obtain the precise information you are looking for - their website is focused on providing summaries of the data, not providing direct access to the data itself.

B. Deprivation

The Index of Multiple Deprivation (IMD) is a ranking of regions in terms of 7 socioeconomic indicators: income, employment, health, education, access to services, crime rate, and housing [5]. Low numbers (1,2, ...) denote high relative deprivation.

It is unclear what we should expect the prior relationship of IMD and loneliness to be. One reason to expect a positive relationship is that many forms of social participation require resources (for instance, transport costs and flexible working) and good health, and so deprivation might be expected to correlate with loneliness [2].

Component of IMD	Pearson's r	p-value
SIMD_Rank	0.11	$p < 0.001$
SIMD_Income_Domain_Rank	0.15	$p < 0.001$
SIMD_Employment_Domain_Rank	0.12	$p < 0.001$
SIMD_Health_Domain_Rank	0.13	$p < 0.001$
SIMD_Education_Domain_Rank	0.08	$p < 0.001$
SIMD_Access_Domain_Rank	0.03	$p < 0.001$
SIMD_Crime_Domain_Rank	-0.01	$p = 0.051$
SIMD_Housing_Domain_Rank	-0.06	$p < 0.001$

TABLE III: Correlation of IMD rank with loneliness score

Overall we see weak correlations of each domain with our loneliness proxy (Table III, and in a surprising direction: our loneliness proxy is *inversely* correlated with deprivation, $r = 0.11$ of composite IMD rank against our 'loneliness' score. (Note that, in the social sciences, $r > 0.4$ denotes an unusually strong relationship, and the median correlation in preregistered studies is 0.16 [15].) The p-values denote the probability, under the null hypothesis that r = 0, of observing results *at least* as extreme under the null hypothesis as those which we observed.

The weakness and heterogeneity of the correlations make interpretation difficult. We can however note the strongest relations, that low income, high unemployment, poor health correlate inversely with loneliness to a roughly average degree (average among reported social science effects).

C. Schools

The hypothesis being investigated that poor attainment at school and high absence rate are indicators of loneliness has limited publications in the literature. [16] found that lack of peer acceptance and friendship quality contributed to withdrawal and loneliness and affected academic functioning. Exam results, financial information and Ofsted reports since 1991 for UK primary and secondary schools are available for download on the government website[2]. Ofsted inspection reports exist for all state schools and give scores covering the following: outcomes for pupils, quality of teaching and personal development, as well as behaviour and welfare, all of which could indicate regional loneliness. However, the data was not used due to the fact the scores are categorical (scores 1-4 for outstanding to inadequate) limiting the ability to generalise and accurately predict loneliness. Attainment scores and absence data for pupils in Key Stage 4 (covering ages 14-16 taking GCSEs) and key stage 5 (covering ages 16-18 taking A-Levels and other equivalent qualifications) was downloaded for the past 3 full academic years from 2016/17 to 2018/19. For KS4, the average attainment 8 score per pupil (average grades across 8 subjects) was used whilst for KS5 the average point score per A level entry was chosen as the performance metric. The absence data consists of the percentage of overall absence for the academic year for every school.

After removing the schools which had no exam result data left data for approximately 5600 KS4 schools and 3100 KS5 schools presented as an average attainment for every school. The data also included the school postcode and a school unique reference number (URN). The pupil absence data covered all schools identified by the URN.

The exam result data was standardised to have a mean of zero and standard deviation of 1 and then combined. The next step included mapping the data for England to highlight which areas had higher and lower than average exam results and school absence rates for the school years ending in 2017, 2018 and 2019. To map the data for the UK, Westminster parliamentary constituency codes (PCON) were used which divide the UK in to 1219 areas each identified by a unique 9-digit code. The PCON code, area names and area coordinate for plotting were found for each postcode using the Find that Postcode API [9]. For schools in the same parliamentary constituency, an average of grades or absence rates were used for plotting.

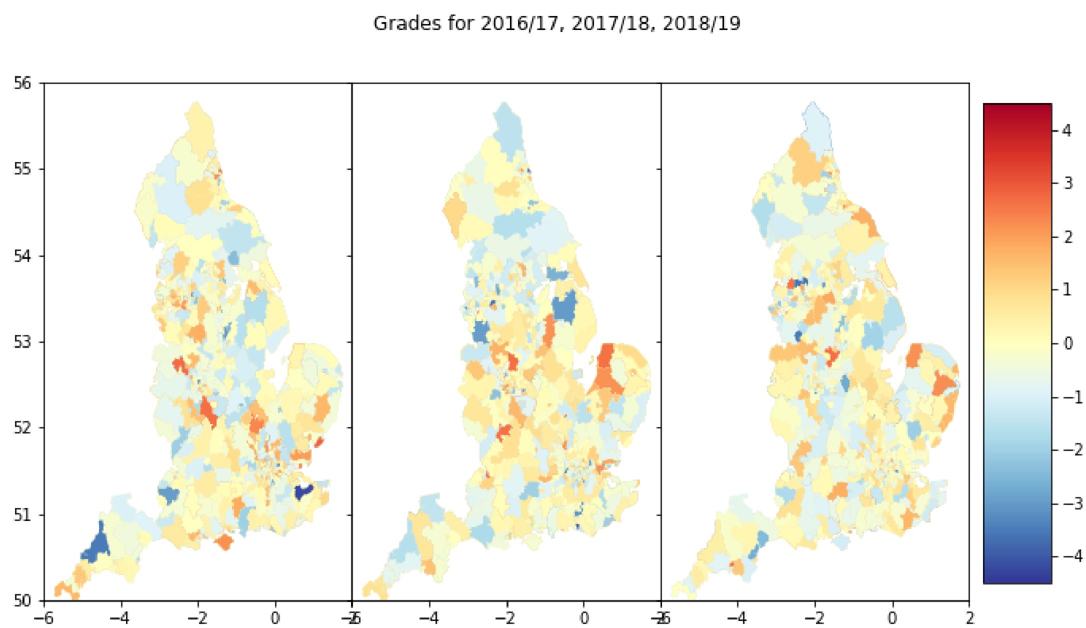


Fig. 4: School grades scaled for the last 3 full academic years

Grades and absence rates were plotted for England in figures 4 and 5 and analysed to highlight regional disparity. Particular regions of the country show far higher or lower grades (dark red and dark blue, respectively) for different academic years from 2016 through to 2019. Regions that have consistently lower than average grades for the past 3 years are parts of the south west as well as certain areas of the north east and north west of England. The absence rates were shown as a percentage of time pupils were absence with the majority of the country falling under 5% whilst a few regions showing very high absence rates.

Comparing the highlighted regions showing higher than average absence or lower than average grades with the loneliness prescription index failed to identify specific regions with high levels on prescription loneliness index. However, the analysis

[2]https://www.compare-school-performance.service.gov.uk/download-data

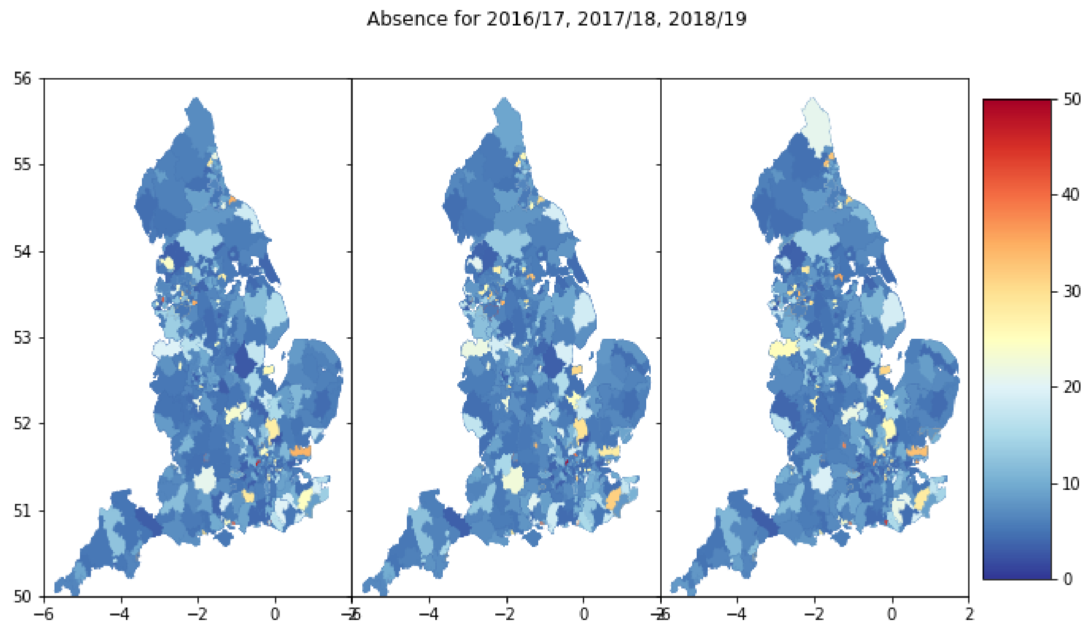


Fig. 5: School absence rates scaled for the last 3 full academic years

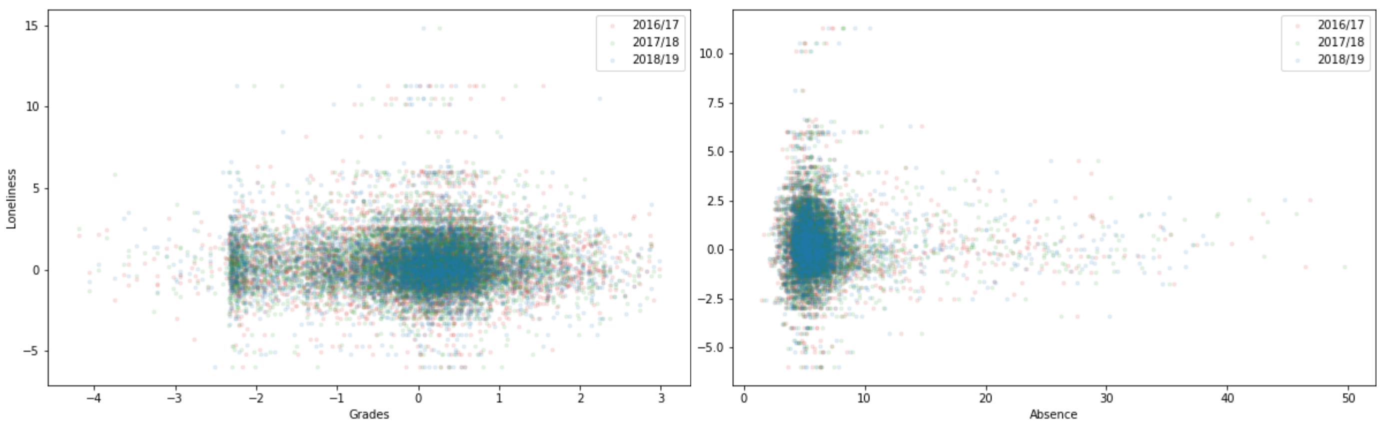


Fig. 6: Scatter plot for loneliness and grades (left) and loneliness and absence (right)

of the plots is very difficult and key information can easily be missed. One reason could be due to the disparity of population densities with cities with very high population densities such as London having small parliamentary areas. A better plot would be a population cartogram which could show the data on plots with equal population densities.

The scatter plots of loneliness against grades and absence (figure 6) for the past 3 years show dense area where most of the data sits. The loneliness-grade scatter plot shows little correlation which was backed up with a correlation coefficients of 0.017, 0.026 and -0.023 for 2016/17, 2017/18 and 2018/19, respectively. Furthermore, the loneliness-absence scatter plot show little correlation with correlation coefficients of -0.014, -0.006 and -0.010 for 2016/17, 2017/18 and 2018/19 respectively. One key reason for the lack of correlation is that the medical conditions used for constructing the GP prescription loneliness index are typically associated with older age such as Alzheimer's disease, high blood pressure and insomnia. The gathered school grades and absence data gave no indication of loneliness when using the prescription loneliness index.

D. Migration

Regardless of country, research has found that people who move from one city to another within their country are likely to be unhappier than locals. [1], [3], [4], [8], [11] This includes all age groups: children [12], young adults [8] and the elderly. [10], [17] We wanted to see if the loneliness index used correlated with the findings above. Using the ONS: Internal migration datasets (2016-2018) [14], the inflow, outflow and net changes of population of each location was found. Due to difference in data available for each country, only England & Wales are present in this section.

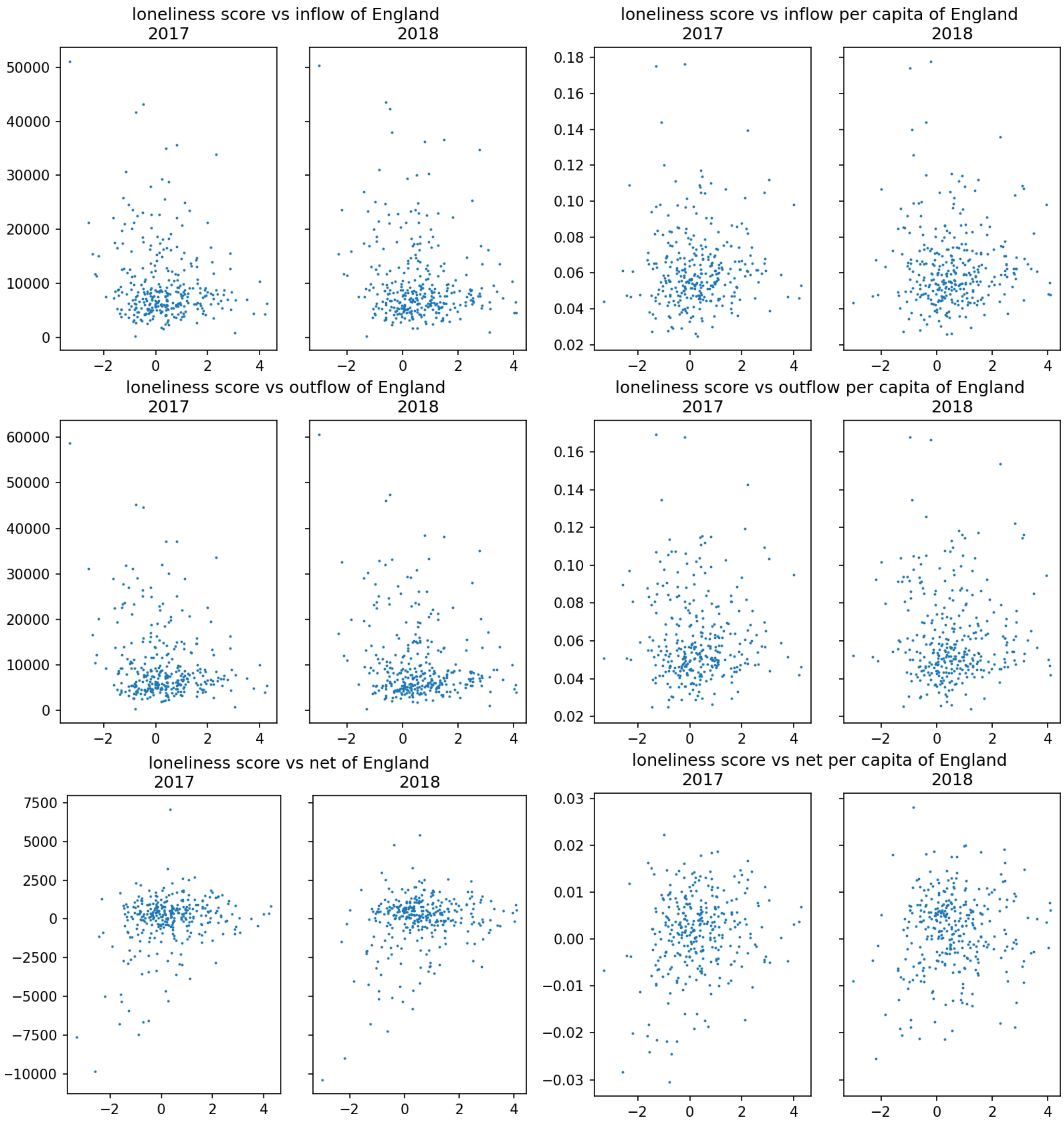
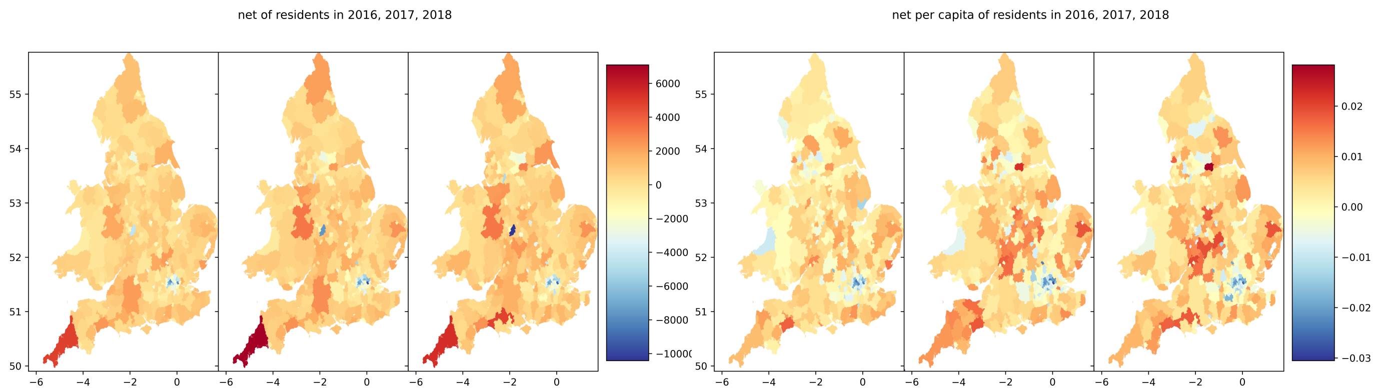


Fig. 7: Scatter plots of the 3 migration statistics (inflow, outflow, net and their per capita counterparts) of areas of England plotted against loneliness scores.

To account for regional population density disparity, population data was found from ONS: Population estimates for the UK datasets (2016-2018) [13]. These were then used to generate inflow, outflow and net based on the proportion of the total population. One problem that we experienced was that some of the location codes found in the original loneliness dataset were



classified as out-of-date by the *findthatpostcode* API, resulting in manual investigation and changes for those codes.

For plotting, the *findthatpostcode* API [9] was used to get the required geojson data for the boundary locations of each area code. This was needed as, unfortunately, due to the data coming from different datasets, the locations used in each set were of different scales (postcodes, constituencies, MSOA, LSOA, etc.). This meant that one location on the loneliness data may overlap into multiple areas in the migration data and vice versa. To combat this we took each postcode from the loneliness data and, once again using the *findthatpostcode* API, found the corresponding area code from the migration data. The data was then grouped together by area code and the loneliness scores averaged where required. This allowed us to have a direct comparison of migration statistics and loneliness index scores.

Migration Measure	Pearson's r	p-value
inflow	-0.16	$p < 0.001$
inflow per capita	0.03	$p = 0.43$
outflow	-0.2	$p < 0.001$
outflow per capita	-0.02	$p = 0.55$
net	0.24	$p < 0.001$
net per capita	0.15	$p < 0.001$

TABLE IV: Correlation of Migration measurements with loneliness score

As with the research listed above, loneliness and depression can often follow moving to a new place. As shown in Table IV, we did find some small evidence for this with a weak correlation (0.24) with net population change and loneliness.

Although loneliness per capita was not a metric we were able to produce, it would be interesting to see if that along with net per capita, would produce a stronger link.

IV. Conclusion

A. Findings

There was little to no correlation between university entrance rate and the loneliness index for 2016, 2017 and 2018. Hot-spots of high acceptance rates are more widespread than loneliness and seem to correlate with locations of large student cities. The constituencies given by the UCAS acceptance data are of the student's location prior to attending university as opposed to the location of the universities where the students ended up attending. This omits the ability to analyse internal migration for university which may shed light on loneliness caused by students attending less prestigious university through clearing.

Weak correlations were also found between the 7 indicators that make up the Index of Multiple Deprivation and the prescription loneliness proxy which made interpretation difficult. The strongest relations of low income, high unemployment and poor health correlate inversely with loneliness.

Neither school attainment nor absence rates yielded a strong correlation with the loneliness prescription index. However, this can be put down to difficultly of analysing geographical data and the averaging of school data within the constituencies.

Migration, in the case of England and Wales atleast, only showed a weak positive correlation in the net population change each year. Although we expected the result to be higher, it is encouraging to see that this metric (out of inflow, outflow and their population normalised counterparts) produced the strongest relationship.

The investigation into whether deprivation, educational attainment and internal migration yielded no positive result as an indicator of loneliness according to the prescription loneliness index. This is not surprising as we are using data for the entire country as a double proxy for loneliness. Because our response data is so indirect, we have no way to check external validity. As such, it's inadvisable to take any of the above as strong evidence about loneliness in the UK.

Instead, we could do the converse inference. Each of the "determinants" can be treated as a different view on the validity of the loneliness-prescription measure. To the extent that we believe that each other domain actually does entail loneliness, then weak correlations with loneliness index reflect more on the index.

B. Individual takeaways

1) Grant (Migration): The mismatch of the both the geographical location of the data as well as the availability was an issue that took way too much tweaking by hand. This is the reason only England and Wales are shown in the migration analysis, as the data came together in one document. Adding Scotland and Northern Ireland would have only made the problem worse.
   I believe that having a loneliness score that takes into account population would help show more of a link between any of our metrics and loneliness but due to reasons above it would have been too time consuming to get it working for all the countries.
2) Dan and David (University): Whilst it seems likely that some aspects of university acceptance and attendance, such as whether someone got into their first choice university, distance of the university from home, how many of their friends got into university would have an impact on an individuals loneliness, this does not appear to have been captured through our analysis. It is likely that either university acceptance rate is a poor measure for the number of people attending university in an area, or that our loneliness proxy is not an accurate measure of loneliness. To better understand our university data we need more fidelity in the data. Breakdowns of acceptance rates by age, universities attended, type of offer (first choice, second choice, clearing) and others, would allow a much more in-depth analysis of the problem. Unfortunately, the UCAS data, whilst extensive, is very hard to search through and obtain what you are looking for. This data may, in fact, be available on their website, however, after several hours of searching the simple acceptance rate by constituency was the closest we could find.
   As mentioned, it is also possible that the loneliness proxy is not very information. Given that most of the other analyses performed also had weak results, this seems plausible.
3) Stefan (School): The ability to find corrections from geographical data was very difficulty and did not yield a conclusion on whether attainment and/or absence rates at school can be used to predict loneliness for specific regions. The disparity in the population density of the UK made the interpretation even harder, especially around cities such as London and Manchester. One solution to this would be to construct a cartogram with equal population density.
   To get around the issues of analysing geographical data, scatter plots were constructed to identify any correlations between loneliness and grades or absence. However, taking the entire data for country removed the ability to highlight hotspots or local correlations and resulted in near to zero correlation coefficients. An increase in the average absence rate for a school did correlate with a reduction to grades but had no effect on the loneliness levels. One interesting finding in figure 6 was the concentration of grades at around 2.2% below than the average and then very little data beyond which could indict the pass-fail boundary.
4) Gavin (Deprivation): I am struck by the underdetermination of weak results; it could be due to our analysis choices, to silent bugs, or to our double proxy simply not containing enough information about the real variable, or any variable but itself. In the case of my question, deprivation, it's unclear what the prior effect should be: poverty and crime could even promote solidarity and community links, and correlate with housing density, which you'd expect to increase interactions and, maybe, fellowship.

C. Future Work

The parliamentary constituency areas used from the *findthatpostcode* API divide the UK into 1,219 areas (even though there are only 650 constituencies represented in the House of Commons) and are divided roughly on population (although the size of the electorate ranges from 55,000 to 113,000) as opposed to area. The limited number could lose vital variations within constituencies as the variable values were averaged for the mapping. Possible solutions could involve using Middle Super Output Area with 7,355 areas to improve the resolution of the maps. This then results in the need to infer from observed facilities to geographical areas and such data transformation data is known as a 'change in support problem' [6]. The value for an area without any data could be estimated using a statistical model such as 'inverse distance weighting' which is based on the distance-weighted combination of data values at known points. This method could be used to compare directly with the loneliness prescription index which was based on MSOAs.

We would also like to have a broader range of university data. For instance, the data for acceptance only included people that applied at the age of 18. It would have been interesting to see if age was a factor in determining loneliness coupled with acceptance rate (e.g. using some kind of linear model) as getting rejected might effect people more the older they are.

As explained, the original score of loneliness is a sum of metrics of illness values. One potential next step could be to combine all the metrics we have discussed in this report and form a secondary loneliness metric, and see how that relates to the original loneliness score.

Its hard to say whether a particular metric is a good indicator of loneliness as it's hard to test such a subjective thing. However, one could try and predict trends for future years given enough data from previous years. However, this kind of time-series analysis would only be good at predicting the proxy metric for loneliness, which, if it turned out was not good at representing loneliness, would not be useful.

REFERENCES

[1] AKSEL, Ş., GÜN, Z., IRMAK, T. Y., AND CENGELCI, B. Migration and psychological status of adolescents in turkey. *Adolescence 42*, 167 (2007).

[2] ALGREN, M. H., EKHOLM, O., NIELSEN, L., ERSBØLL, A. K., BAK, C. K., AND ANDERSEN, P. T. Social isolation, loneliness, socioeconomic status, and health-risk behaviour in deprived neighbourhoods in denmark: A cross-sectional study. *SSM-population health 10* (2020), 100546.

[3] APPLETON, S., AND SONG, L. Life satisfaction in urban china: Components and determinants. *World Development 36*, 11 (2008), 2325–2340.

[4] CHENG, Z., WANG, H., AND SMYTH, R. Happiness and job satisfaction in urban china: A comparative study of two generations of migrants and urban locals. *Urban Studies 51*, 10 (2014), 2160–2184.

[5] EXECUTIVE, S. Scottish index of multiple deprivation 2006 technical report. *Office of the Chief Statistician, Scottish Executive* (2006), 10.

[6] GELFAND, A. E., ZHU, L., AND CARLIN, B. P. On the change of support problem for spatio-temporal data . *Biostatistics 2*, 1 (03 2001), 31–45.

[7] GRIMSLEY, J., HOPKINS, K., TIZORA, E., AND LEWIS, D. Developing a loneliness prescription index.

[8] HENDRIKS, M., LUDWIGS, K., AND VEENHOVEN, R. Why are locals happier than internal migrants? the role of daily life. *Social Indicators Research 125*, 2 (2016), 481–508.

[9] KANE, D. Find that postcode api.

[10] KING, R., CELA, E., FOKKEMA, T., AND VULLNETARI, J. The migration and well-being of the zero generation: Transgenerational care, grandparenting, and loneliness amongst albanian older people. *Population, Space and Place 20*, 8 (2014), 728–738.

[11] KNIGHT, J., AND GUNATILAKA, R. Great expectations? the subjective well-being of rural–urban migrants in china. *World development 38*, 1 (2010), 113–124.

[12] LU, Y., AND ZHOU, H. Academic achievement and loneliness of migrant children in china: School segregation and segmented assimilation. *Comparative education review 57*, 1 (2013), 85–116.

[13] PARK, N. Dataset population estimates for the uk, england and wales, scotland and northern ireland, Jun 2018.

[14] PARK, N. Dataset internal migration: by local authority and region, age and sex, Jun 2019.

[15] SCHÄFER, T., AND SCHWARZ, M. A. The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology 10* (2019), 813.

[16] SHIN, Y. Peer relationships, social behaviours, academic performance and loneliness in Korean primary school children. *School Psychology International* (2007).

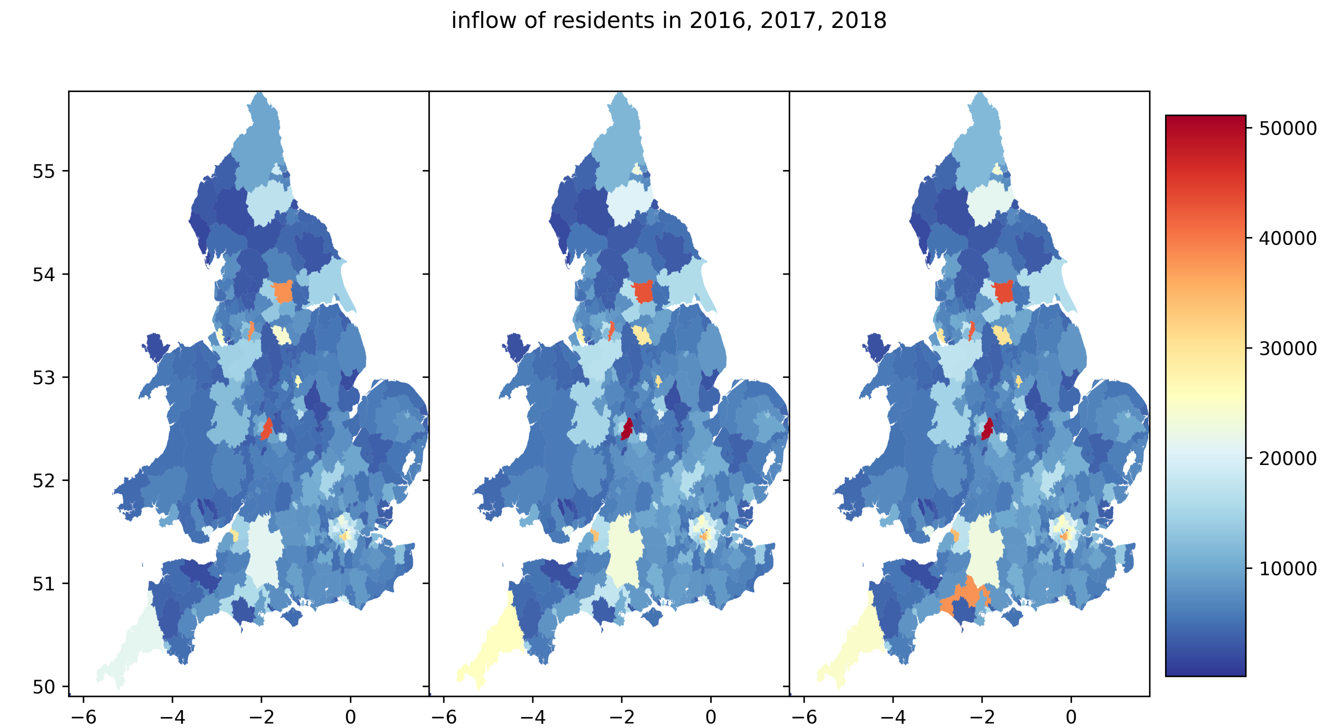
[17] WU, Z., AND PENNING, M. Immigration and loneliness in later life. *Ageing & Society 35*, 1 (2015), 64–95.

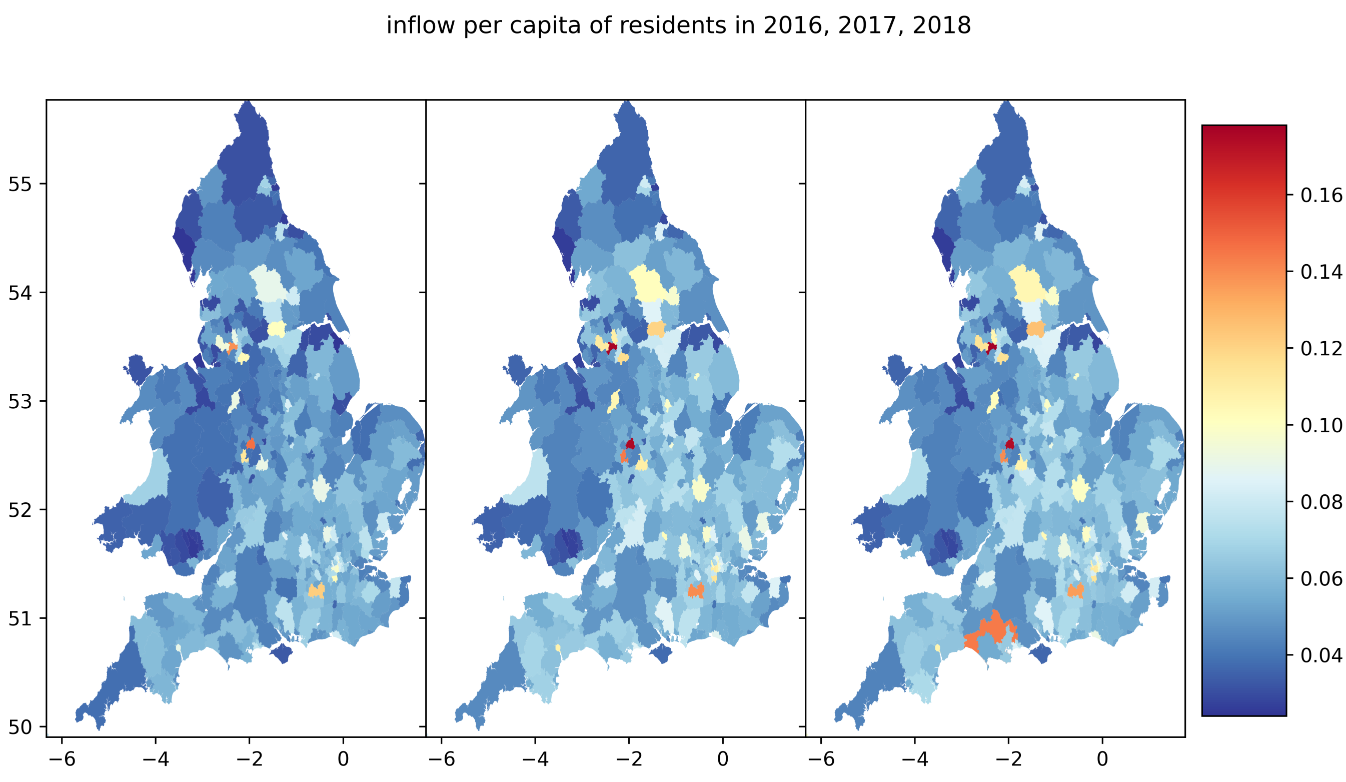
APPENDIX

A. Data

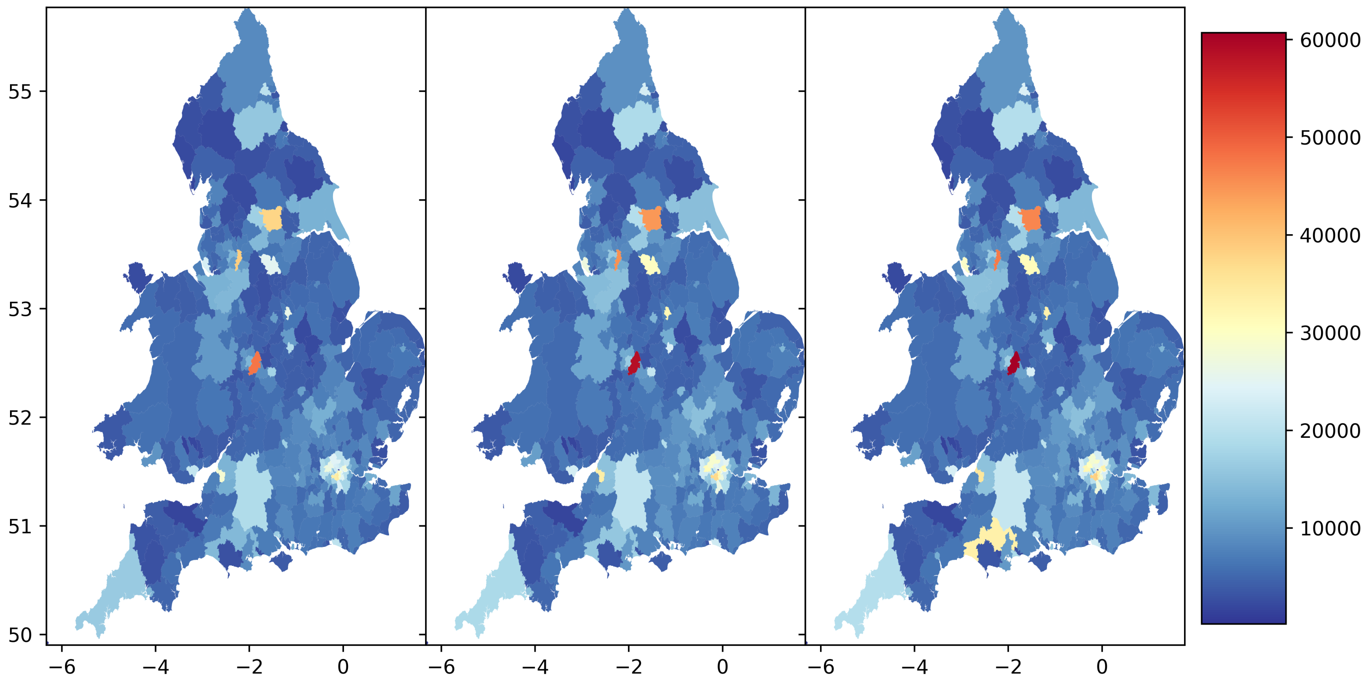
1) 'Prescriptions in the Community'
2) 'GP Practice Contact Details and List Sizes'
3) National Statistics Postcode Lookup
4) Scottish Index of Multiple Deprivation 2020
5) MSOAs and Constituencies
6) Find and compare schools in England

B. Migration Maps

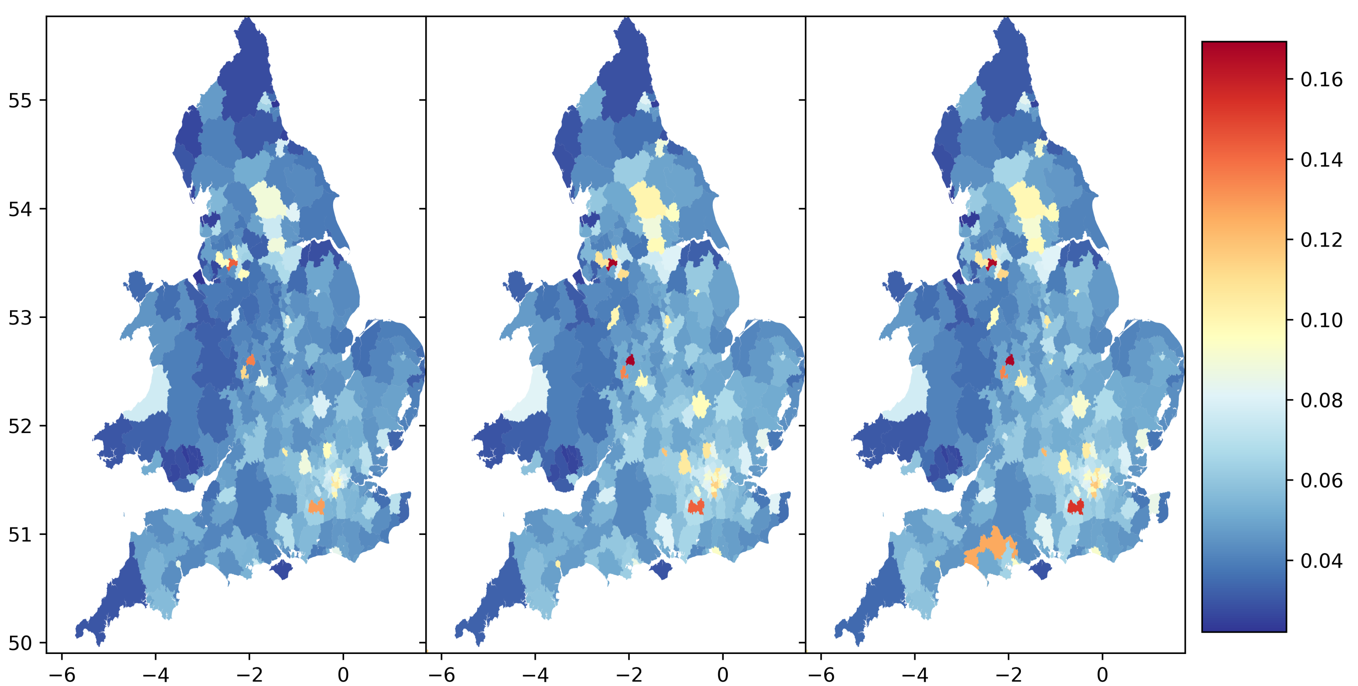




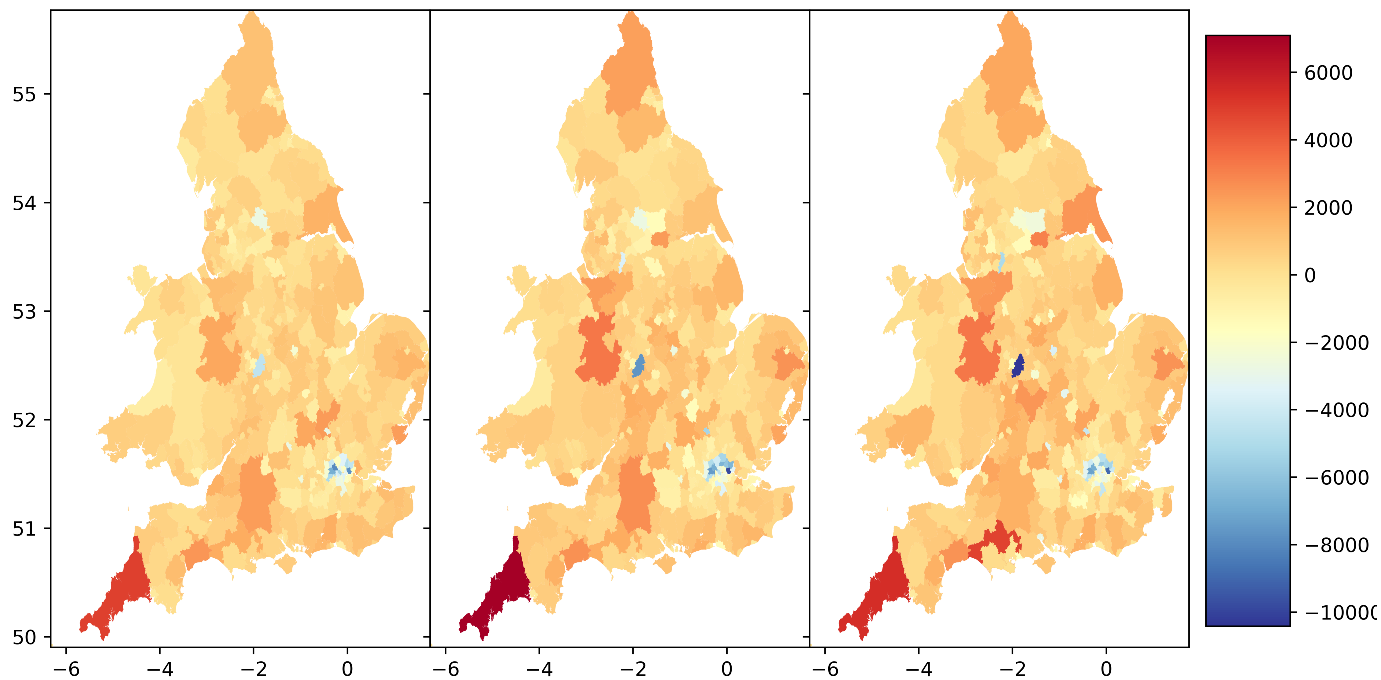
outflow of residents in 2016, 2017, 2018



outflow per capita of residents in 2016, 2017, 2018



net of residents in 2016, 2017, 2018



net per capita of residents in 2016, 2017, 2018