MSc Data Science

NCG616 Dissertation - August 2023

Mapping Mental Health: The Role of Urban Living and Socioeconomic Conditions in Antidepressant Prescriptions

Ernestas Kriauciunas

Supervisors: Prof. Christopher Brunsdon and Prof. Kevin Credit

Abstract

The objective of this dissertation is to assess whether urban environments have a higher prevalence of antidepressant prescriptions compared to rural environments. This research is carried out using publicly available 2014 National Health Service (NHS) prescription data, National Postcode Lookup (NPL) and 2011 Output Area Classification (OAC) data for analysis. Prescription numbers are identified for each of the different Supergroups which define whether an Output Area is under the urban or rural label. An increased rate of antidepressant prescriptions can be identified for urban areas when compared to rural ones. The Poisson regression model that is implemented provided statistically significant coefficient values for prescription rates. Time series analysis of monthly prescription counts showed an identifiable seasonal trend for all Supergroups. The trends, when compared for rural and urban Supergroups, showed no difference in trend or seasonality. This suggested that while urban and rural prescription counts differed significantly, their overall rate counts followed a similar trend throughout the year.

Contents

1 Introduction

The world's population has been rapidly increasing since the 1900s. Due to human innovation and scientific breakthroughs, global birth rates have been exceeding mortality rates in recent history (Bongaarts, 2009). This population growth momentum is aided by improvements in education, reproductive health and child survival (Bavel, 2013). With the population number reaching 8.0 billion in 2022, the United Nations future projection predicts that this number will reach 10 billion in 2050 (UN, 2023). Historical population trends indicate that humans generally gravitate towards more densely populated areas or communities. Interestingly, before the year 1850, no society could be categorized as urban. Humans, as a whole, lived predominantly in rural areas. Although there is evidence of some village/town areas having bigger populations than other rural areas, this difference in population density is marginal compared to the immense numbers of a modern city nowadays. In recent times, as nations have industrialized, there has been a swift shift towards urbanization. This concept pertains to the phenomenon in which urban areas experience a more pronounced relative rise in population compared to the overall population growth of a nation (Davis, 1965).

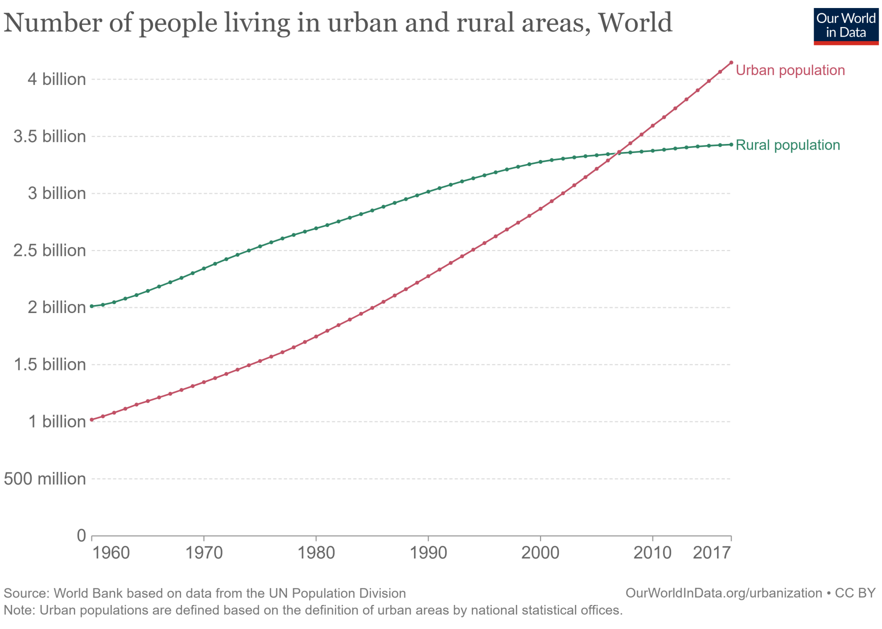

Figure 1: Number of people living in urban and rural areas, World

Current trends show us that urban populations have overtaken, and are outgrowing their counterpart every year (Figure 1) (Ritchie and Roser, 2019). Recent studies have projected global urban populations to increase by 2.5 billion over the next 30 years (Mahtta et al., 2022). The future extent of urbanization implies that urban living will predominantly be the standard for individuals in the future, surpassing its current prevalence.

People choose to live in urban areas for a variety of beneficial reasons. Unlike rural areas, there are plenty of better job opportunities, access to healthcare and education, better transport and overall infrastructure (Davis, 1965). Modern urban cities are hubs of economic activity and culturally diverse groups of people. Such an environment lays the groundwork for the inception and growth of businesses, and the diversity of its inhabitants fosters a rich tapestry of cultural events. The fast-paced, bustling environments provide a greater amount of services and amenities to its residents. Although people are motivated to reside in urban settings owing to their numerous advantages, the challenges inherent to such environments can adversely impact both the physical and mental well-being of individuals.

Urban living in most cases, has a myriad of challenges and inherent issues associated with it. The most glaring difference when compared to rural living is the environment. To accomodate the necessary housing, business and service space, the environment itself is transformed to what some would call a "concrete jungle" (McDonald et al., 2018). The economic and health consequences of the resulting air pollution, carbon emissions, and water quality are well documented (Remoundou and Koundouri, 2009; Liang and Gong, 2020). As green open spaces become limited, an often overlooked issue of overcrowding becomes present. Household crowding can lead to depressive symptoms in people (Ruiz-Tagle and Urria, 2022). Communities that might be vulnerable to this include migrants, students, and economically disadvantaged ethnic groups. Looking at overcrowding as a whole, a recent study conducted in Salzburg, Austria found a link between heightened human physiological stress response when individuals are present in highly crowded environments (Zhang et al., 2023). In contrast to rural regions, mega-urban centers have embraced a round-the-clock lifestyle. Take New York City, often dubbed "the city that never sleeps," as an example. The natural sleep-wake cycle of individuals is disrupted by heightened noise and light disturbances at night. Research indicates that such interruptions can adversely affect people's mental well-being (Chepesiuk, 2009). Perhaps the greatest stressor of life in urban environments is crime. Studies tell us that crime rates correlate with city size (Glaeser and Sacerdote, 1999). While crime rates can differ across cities, living in areas with elevated crime levels can induce anxiety issues among the residents. There is research that suggests that the consequences of fear and anxiety "include distrusting staff, cooperating reluctantly, learning reticence, delaying help-seeking, avoiding services, feeling unsafe in the community and avoiding exposure" all which can lead to mental health problems (Sweeney et al., 2015).

The World Health Organisation (WHO) states that depression is one of the leading causes of disability (WHO, 2023a). This condition can alter a persons life drastically. Some profound effects include: physical health decline (Shen et al., 2011), emotional suffering (Sheng et al., 2017), reduced cognitive ability (Lam et al., 2014), substance abuse (Conner et al., 2009) and social isolation (Ge et al., 2017). With the WHO reporting that mental health problems are increasing worldwide, it is important to understand what the underlying causes of it are so that necessary steps in treatment can be implemented effectively. While mental health issues primarily affect the individual experiencing them, there's a ripple effect that impacts their close associates and the broader communities they belong to. There is evidence to suggest that those who suffer from depression will have reduced work-productivity and discipline. With the help of medical treatment, the depressive symptoms can be alleviated (Hammer-Helmich et al., 2018). Awareness and education about this mental health condition are of high importance. Not only does depression impact someones day-to-day life, in its worst case scenario, if left untreated, can lead to suicide (Bradvik, 2018).

Significant advancements in modern medical practice have been achieved in utilizing mental health services and medications to effectively treat individuals dealing with depression. Precise diagnosis enables people to access the targeted assistance they require (Cuijpers et al., 2020). As of 2023 , it is estimated that 3.8% of the world population experiences depression. However, it is stated that the main barriers associated with receiving treatment are: lack of investment in mental health care, untrained health-care providers and the social stigma associated with mental disorders (WHO, 2023b). In particular, due to the costs and lack of awareness, as much as 75% of people in low-middle income countires receive no treatment.

The purpose of this paper is to analyse antidepressant prescription data for a modern developed country that is rich in both urban and rural environments. Looking at counts of prescriptions for both settings can provide insight into potential disparities and trends associated with depression in those areas. The key objectives will be to i) Compare prescription counts for urban and rural environments and identify if the differences are statistically significant and ii) Perform a time series analysis of prescription data for both settings to identify if seasonal patterns are present.

2 Method

2.1 Data Sources

The country of choice for this study was the United Kingdom (UK). Publicly available drug prescription data was sourced from the UK National Health Service (NHS) for the year 2014. Due to the dataset being quite large, it was filtered for England only. The data has been accessible to the public since January 2014 and receives monthly updates, following a consistently refreshed release schedule.

UK postcode data was gathered from the Office for National Statistics. Their Open Geography Portal provides publicly available Postcode Lookup (NSPL) data that connects both active and discontinued postcodes to various current statutory geographies using a "best-fit" assignment based on the 2011 Census Output Areas.

Output Area Classification (OAC) data was also sourced from the Office for National Statistics. This data categorizes communities in England into distinct community groups based on the 2011 census. Clusters are contained within this data for all the local authorities that are present. Descriptions and characteristics of each cluster are divided into labels called 'Supergroups', 'Groups' and 'Subgroups'.

2.2 Prescription data

Anonymised details of every prescription issued via the NHS are available, including the location of the practice (or hospital) where the prescription was issued, the drug that was prescribed, and the cost of the medicine. The large dataset (15Gb) can be worked on by splitting it by the type of drug prescribed.

The prescribed drugs are described using the British National Formulary (BNF) codes. This code specifies the nature of the prescribed drugs (or appliances). A typical code may be something like 0202030C0BEAAAC - this can be interpreted by splitting the code into one, two or four-character chunks. For this example 02|02|03|0|C0|BE |AAAC is interpreted as this:

BNF Code		
Code	Section	Meaning
02	BNF Chapter	Cardiovascular System
02	BNF Section	Diuretics
03	BNF Paragraph	Potassium-sparing diuretics and aldosterone antagonists
0	BNF Subparagraph	Potassium-sparing diuretics and aldosterone antagonists
C0	BNF Chemical Substance	Amiloride hydrochloride
BE	BNF Product	Amoride
AAAC	BNF Presentation	Amoride 5mg tablets

Table 1: BNF Code section and meaning.

This coded data allows for all different type of prescribed drugs to be filtered out. For the purposes of this paper antidepressant prescriptions are filtered out using the BNF Chapter and BNF Section codes 0405. Each row of the corresponding dataset contains the details of a single prescription. The different columns of information that are available are:

- Postcode: The postcode of the prescribing practice or hospital
- BNF Code: The BNF code of the prescription as set out above
- Year Month: The year and month the prescription was issued.
- Quantity: The quantity of a medicine, dressing or appliance for which an individual item was prescribed and dispensed, for each BNF Presentation. This represents a pseudo pack size, to illustrate the typical range of prescribed quantities of a given presentation
- Items: the total number of times that the medicine, dressing or appliance appeared on prescription forms that were prescribed and dispensed
- Total Quantity: the total quantity of a medicine, dressing or appliance prescribed and dispensed for each BNF Presentation. This is calculated by multiplying Quantity by Items
- NIC: The 'Net Ingredient Cost' (NIC) of the items normally based on the price given in the Drug Tariff or published wholesale prices
- Actual Cost: The 'Actual Cost' that accounts for the national average discount and some payments to dispensers
- ADQUSAGE: ADQ Usage - Average Daily Quantity (ADQ) is the typical daily dose of a medication, prescribed to adult patients by GP Practices

Of particular note, the 'Postcode' and 'Year Month' columns will allow for geolocation and time series analysis of the prescriptions.

2.3 NSPL Data

The NSPL dataset contains the longitude and latitude point locations for every postcode in the UK. It supports the production of area based statistics from postcoded data. In addition to this, each unique postcode also has the associated country, city and region codes listed. Utilising this information will allow the linkage of prescription data postcodes to their respective geographical locations. The prescribing practices can now be geolocated and information can be deduced as to the number of prescriptions given out.

2.4 OAC Classification

The Output Area Classification utilizes data from the 2011 UK population census to categorize geographic regions into clusters, based on shared population attributes or characteristics within each group. These classifications are hierarchichal, which consist of three tiers of groupings. The labels assigned to supergroups, groups, and subgroups, along with their descriptions, aim to depict the attributes of areas concerning demographic makeup, household structure, housing, socio-economic features, and employment trends.

Clusters assigned to the Supergroup classification provide the most generic descriptions of the population in the UK. There are eight distinct supergroups with descriptions provided below:

The labels placed on these cluster groups provide greater insight than just the statistical outputs of cluster analysis alone. There may be variation in the exact characteristics of each clusters population, however it is safe to assume that they represent the general characteristics of the local authorities in the clusters.

It is important to understand how an Output Area (OA) is defined. When census data is published, there exists the smallest geographical unit as the key geography. Each unit roughly contains 125 households per that area. When census data is analysed for each unit, the characteristics that are looked into are 60 different variables

Supergroup		
Number	Name	Description
1	Rural Residents	Population lives in far less densely populated areas
2	Cosmopolitans	Population lives in densely populated urban areas
3	Ethnicity Central	Population lives in denser central areas of London
4	Multicultural Metropolitans	Population concentrated in transitional areas between urban and suburban areas
5	Urbanites	Population lives in urban areas
6	Suburbanites	Population lives on the outskirts of urban areas
7	Constrained City Dwellers	Population lives in urban areas
8	Hard-pressed Living	Population lives in urban surroundings

Table 2: OAC Supergroups.

derived from this data. Each variable belongs to one of 5 different 'domains'. They are: demographic structure, composition of household, housing type, socio-economic indicators and employment.

The Supergroup classifications provide a broad categorization of an OA, but the Group and Subgroup classifications further break it down into one of 26 groups or one of 76 subgroups, respectively. Each subsequent group category expands on the Supergroup classification via different socio-economic factors.

2.5 Urban-Rural Comparison

The OAC Supergroup classifications will be the basis for prescription rate comparison in urban and rural environments. The classification numbers that represent rural settings most accurately will be Supergroup 1 and 6. For urban settings the remaining groups contain the rest of population. Joining the OAC data with the NSPL will provide the relevant geographical coordinates for each OA. Using spatial joins, the prescribing locations can be linked to their relevant OA's and assigned their corresponding Supergroup.

The prescription data is dated by month and year. This allows for the overall trend of population prescription to be analysed. Making use of the Supergroup classifications, urban vs rural time series can be produced and the corresponding monthly trends, if any, can be visually represented.

2.6 Model Approach

The prescription counts will act as the response variable with categorical predictors of Supergroup (1,2...,8). Since the Supergroups are on a categorical scale, each distinct group variable will be recoded to ones and zeros in order to calculate expected count. Each prescription can only take on a single value of one, denoting the Supergroup that the prescribing practice belongs to, with the remaining Supergroups taking on a count of zero. Summarising the total yearly precription counts and population size for each Supergroup will provide the basis for a General Linear Model (glm) creation. To adjust for the fact that population sizes will differ between different Supergroups, number of cases per population will be the rate. The Poisson regression model will produce the Rate Ratio for each Supergroup's slope values. Finally, a hypothesis test should be constructed to assess the statistical significance of each estimate value to see if they describe the rate of prescriptions for each group.

3 Results

3.1 Total Prescription rates

Summarising the Prescription and Population counts for Supergroups (1,2,...,8):

Supergroup	Prescriptions	Total Population
1	515,652	7,352,746
2	728,703	3,158,655
3	766,159	3,590,174
4	1,527,888	8,229,373
5	2,548,627	11,064,122
6	865,917	13,119,920
7	1,417,911	4,997,068
8	1,303,451	11,670,120

Table 3: Supergroup Prescriptions and Population.

With some table manipulation, the prescription numbers against population are visually represented:

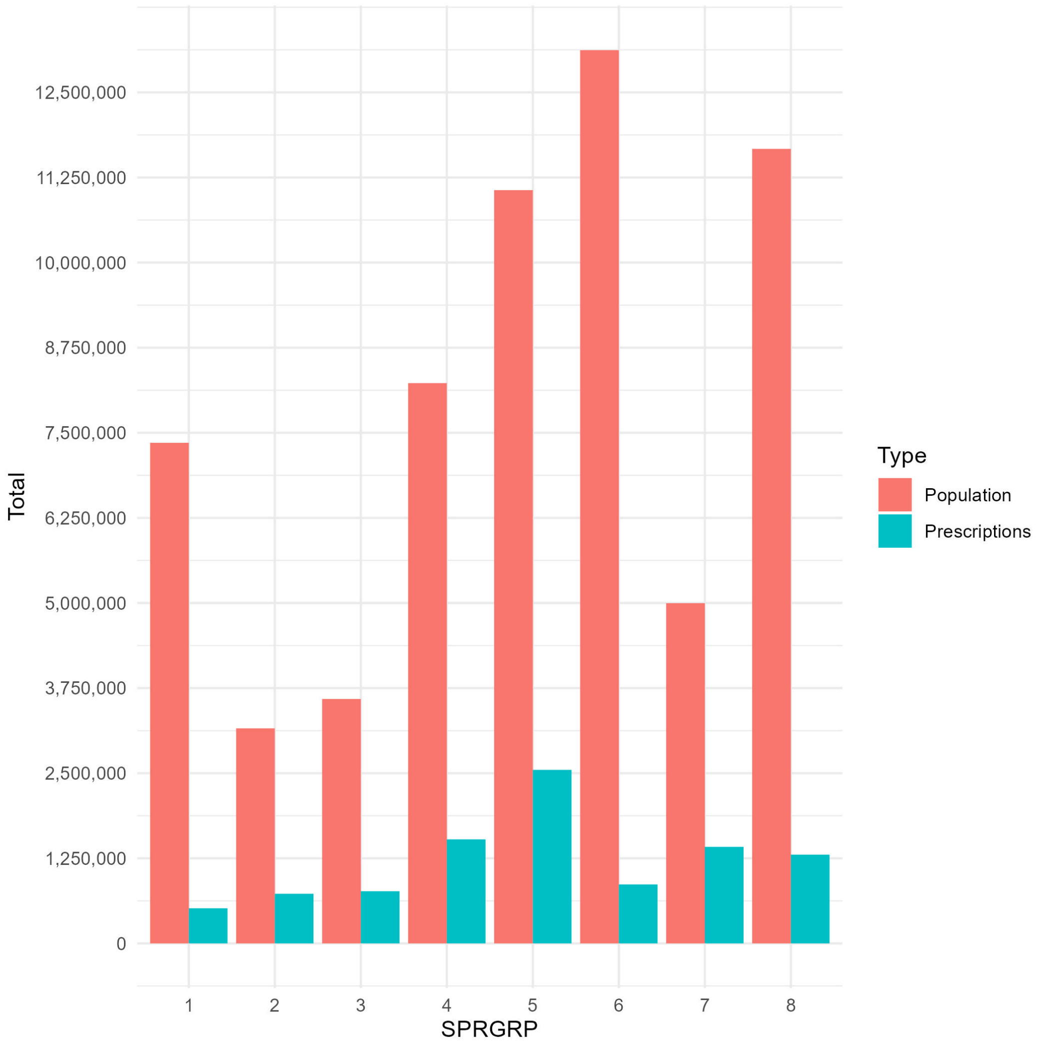

Figure 2: Barchart comparison of Prescription counts compared to Population

The OAC had a total population of 63,182,178 people living in England during the census data collection in 2011. The NHS data resulted in 9,619,611 prescriptions of antidepressants for the year 2014. Looking at figure 2, there is a clear indication that population number is not an indicator of prescription rates. A Poisson regression model is implemented in form of:

$$log(y) = b_0 + b_1 \cdot Supergroup_2 + ... + b_n + \cdot Supergroup_{1+n} + log(Population) \quad (1)$$

where:

$n = 0, 1, 2, ..., 7$

$log(y)$ = log count of prescription rate

b_0 = the intercept (in this case Supergroup 1)

b_{1+n} = slope values for Supergroups $1 + n$

$log(Population)$ = offset to adjust for difference in population size per Supergroup

After plugging the values into this model, the coefficients values with their significance were gathered:

Supergroup	Estimate	Std. Error	z value	p-val
Intercept	-2.68	0.0014	-1908	$<<< 0.05$
2	1.19	0.0018	654.34	$<<< 0.05$
3	1.11	0.0018	617.81	$<<< 0.05$
4	0.97	0.0016	604.51	$<<< 0.05$
5	1.19	0.0015	778.82	$<<< 0.05$
6	-0.06	0.0018	-34.51	$<<< 0.05$
7	1.4	0.0016	859.50	$<<< 0.05$
8	0.465	0.0016	282.88	$<<< 0.05$

Table 4: Poisson Regression coefficients.

The estimates generated from this model are on the logarithmic scale. To calculate the rate ratios (RR), they need to be exponentiated and a 95% confidence interval will be added to measure accuracy.

The Intercept (RR=0.07 95% CI 0.07,0.07) in table 5 for Supergroup 1 is used as a reference level. It can be interpreted as the rate of prescriptions per population when no other predictors (Supergroups) are in the model. It tells us that there are 0.07 prescriptions per individual in the population. The narrow Confidence Interval indicates high precision.

Supergroup	RR	2.5%	97.5%
Intercept	0.07	0.07	0.07
2	3.29	3.28	3.30
3	3.04	3.03	3.05
4	2.65	2.64	2.66
5	3.28	3.27	3.29
6	0.94	0.94	0.94
7	4.05	4.03	4.06
8	1.59	1.59	1.60

Table 5: Rate Ratios with 95% CI's for estimates.

Going down the list, RR values greater than 1 suggest a higher rate of prescription while values less than 1 indicate lower rates of prescription. The Supergroup that has the highest rate is 7 (RR=4.05 95% CI 4.03, 4.06) and the lowest rate is 6 (RR=0.94 95% CI 0.94,0.94). Intepreting them as such:

Supergroup 7 rate: $0.07 \cdot 4.05 = 0.2835$

Supergroup 6 rate: $0.07 \cdot 0.94 = 0.0658$

For the significance test in Table 4. The p-value is associated with the z-statistic. It's a measure of evidence against a null hypothesis. Here, the null hypothesis for each predictor is that its coefficient is zero (or that the predictor has no effect). In this table, all of the predictors have very small p-values (less than 0.05), which suggests that each of them is significantly different from zero.

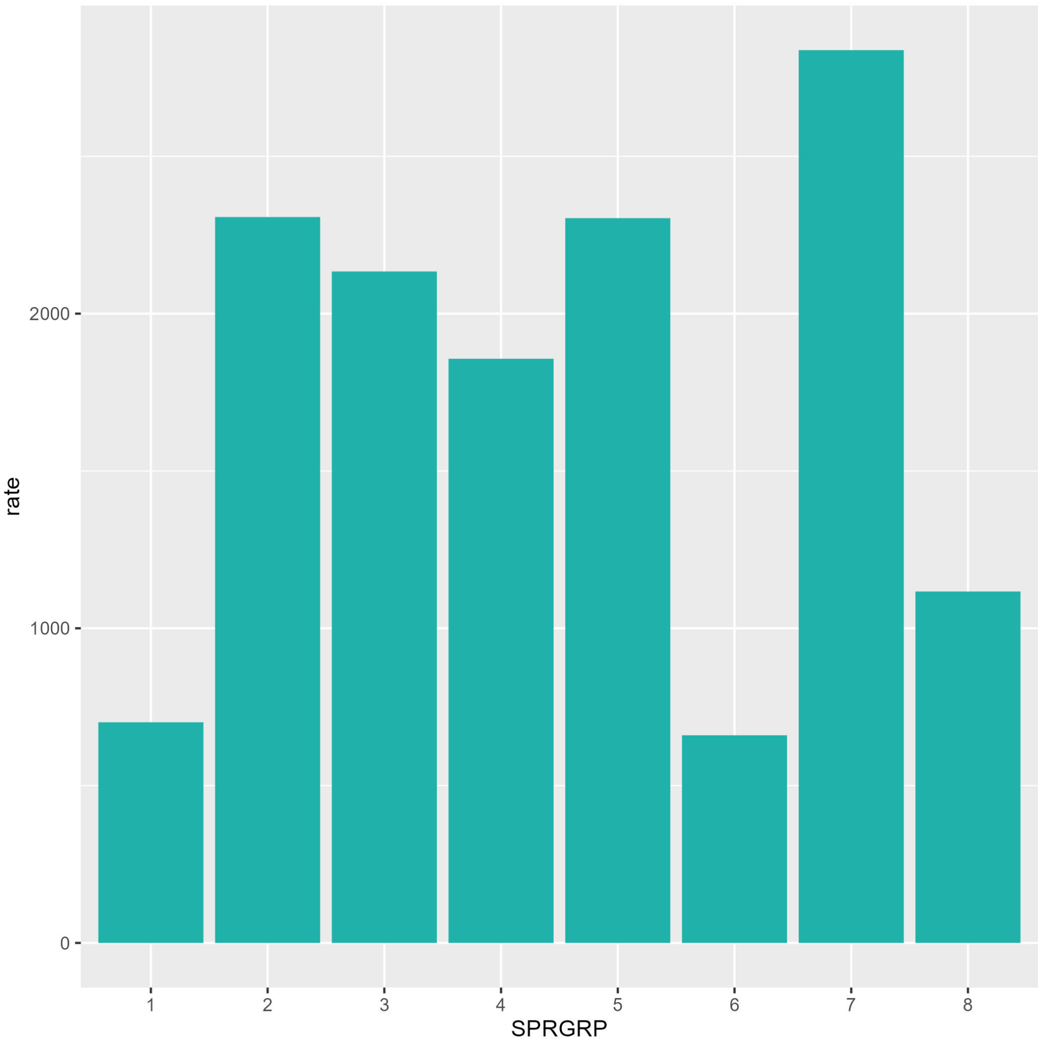

Figure 3: Barchart comparison of Prescription rates per 10,000 population

In Figure 3 above, the different rates are plotted on a barchart per 10,000 population. The two lowest Supergroup rates are 1 and 6 which fit the rural description most accurately while the remaining Supergroups that fit the urban description show the highest rates.

3.2 Time Series of prescriptions

Plotting the total prescription values over the whole population gave a time series depiction of monthly prescription figures.

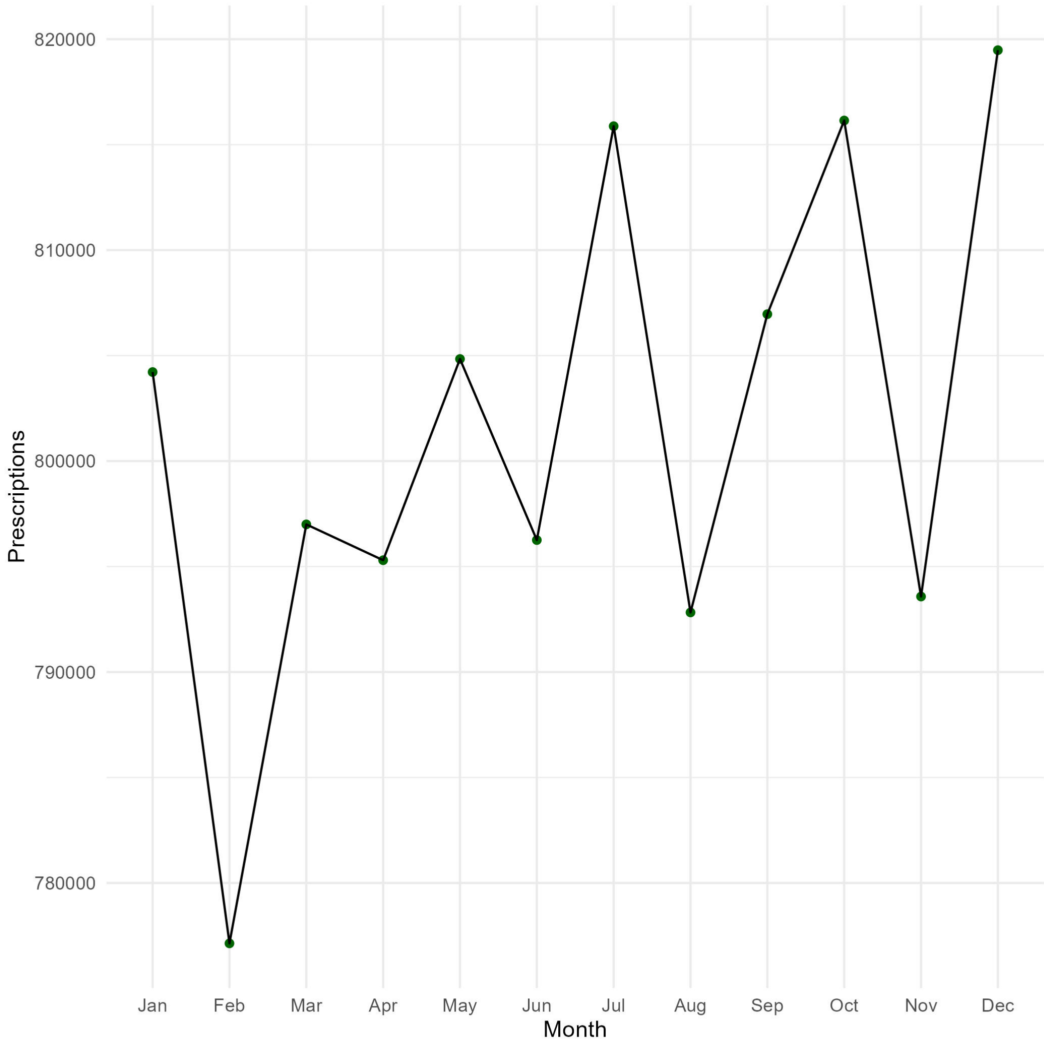

Figure 4: Scatterplot of total prescriptions by month

The variation in prescription numbers is quite small when taking the total numbers into account. The highest prescription rates occured in December (819,478) with lowest in February (777,136). The variation between min and max values is 42,342 which is 5.1% of the max December figure, this tells us that during the year, rates do not decrease below $\sim 95\%$ of the yearly max. While the change in prescription rates is quite small, there exists a seasonal pattern with dates before June and after July where the average prescription values differ.

The same analysis is done for the different Supergroups and their monthly prescriptions.

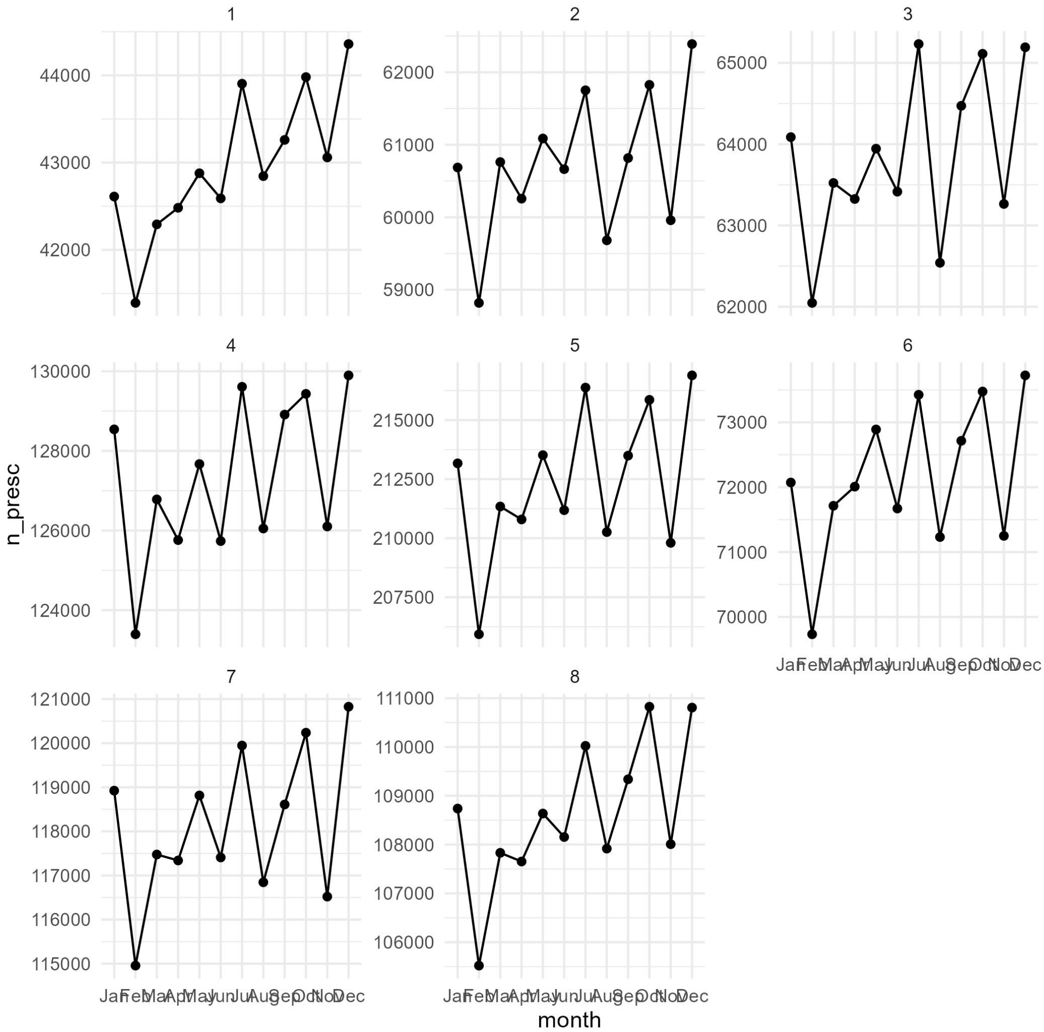

Figure 5: Scatterplot of total prescriptions by month per Supergroup

The trend for each Supergroup is very similar to the overall total prescription trend in Figure 4. Each pattern has a trend that is linearly increasing overall from January to December. There is some seasonality patterns as the peaks and troughs appear somewhat regularly throughout the year. There are three different months that exhibit the highest peaks for all Supergroups: July, October and December. February is consistently the least prescribed month for each graph with the rest of the months in year fall somewhere in the middle.

Interestingly, the most rural Supergroup (1) looks to have the most distinct pattern of prescriptions, particularly in the first 6 months of the year. The difference between peaks and troughs is consistently the lowest in the year, whereas the rest of the Supergroups show large jumps by month.

Lastly, Supergroup 1 and 6 are categorized as rural and the rest are grouped as urban. Superimposing them into two separate categories yielded the comparison plots Figure 6 and 7 below. There are no strong difference indicators for the two categories. Both groups follow a similar trend and seasonality.

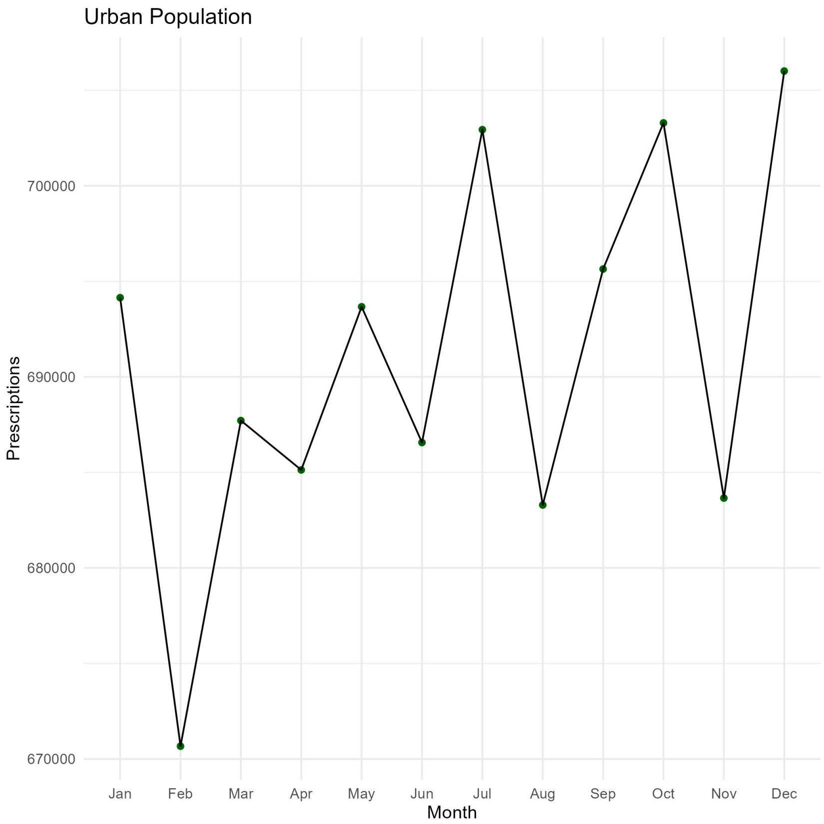

Figure 6: Scatterplot of total prescriptions by month for Urban Population

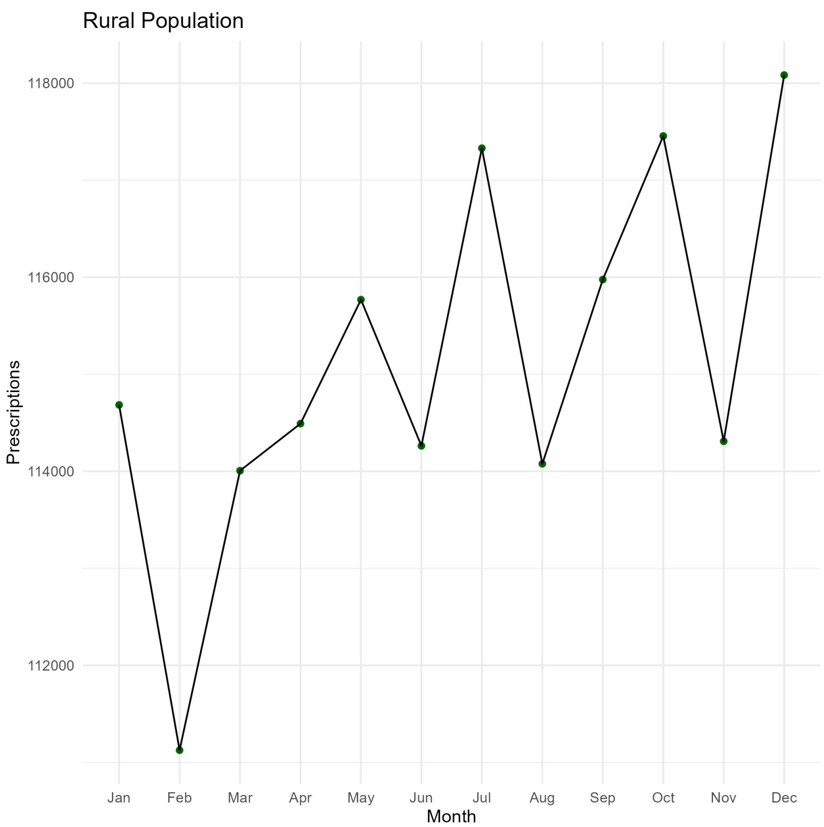

Figure 7: Scatterplot of total prescriptions by month for Rural Population

4 Discussion

This project was carried out to assess whether living in urban areas is worse for mental health compared to living in rural areas. Antidepression prescription counts for England is a good indicator to measure the rates of depression for this population. The publicly available OAC classifications data was used to broadly characterise distinct populations into 8 supergroups. The group characterisations came solely from 2011 census data which proved to be very beneficial. Without the specificity of census style questions submitted by the public, it would prove to be very difficult to define urban and rural areas as the debate is still ongoing as to how urban environments should be defined (Ritchie and Roser, 2019).

4.1 OAC Limitations

While the 2011 OAC data is the most valuable public classification data accessible, it does come with certain limitations. The most glaring problem is that the data comes from the year 2011. As of now, it is 12 years old and the information contained within the dataset is limited to that year. The other data sources (NHS and NSPL) used in this project are updated more regularly. Selecting the 2014 drug prescription data to pair with the OAC dataset is feasible, given that the 3-year difference still allows the data to be usable. Confidence in the OAC's capacity to effectively distinguish between areas and pinpoint those areas with relative socioeconomic advantages will diminish over time. As urbanization continues and new postcodes are added, they will not show up in the 2011 OAC. For current trends to be analysed and updated, a new census with its respective OAC will have to be released to the public. It's also vital to highlight that while OAC categorizes individuals by their postcode, individuals may not always exhibit similar traits as others in their postal code.

4.2 Model Insights

After taking the total counts of prescriptions recorded in the NHS data, postcodes of the prescribing practices or hospitals were geolocated to the defined Supergroups coming from OAC. For example, if a prescription was given out in an Output Area belonging to Supergroup 1(rural), it can be assumed that the individual receiving it resided in that same area. Even though there might be outliers, one could contend that according to the law of large numbers, any outlier from a Supergroup receiving a prescription in a different Supergroup would be balanced out when considering large datasets.

A Poisson model (1) was used to calculate the different rates of prescription counts, accounting for population size of their respective Supergroups. Using Supergroup 1 as the reference level it was found that the "urban" Supergroups (2,3,4,5,7 and 8) had a positive multiplicative effect (denoted by rate ratio RR) on prescription rates (Table 5). Interestingly, Supergroup 6 (Suburbanites) had a small reduction affect on the reference level. Individuals who are classified as suburbanites live on the outskirts of urban environments. Their description has more similarities to a rural classification as opposed to an urban type. It would be a lengthy endeavour to provide the rigorous descriptions of each Classification and their corresponding Groups and Subgroups. This study was done to get a general overview of prescription data for the broad categories of urban and rural environments, for a deeper look into specifics, the CDRC open data portal contains all the necessary information at https://data.cdrc.ac.uk/dataset/output-area-classification-2011.

4.3 Seasonal Trends

The NHS prescription data is dated monthly. Therefore, the time series was carried out for distinct count variables that summarised the prescriptions numbers from the beginning to the end of each month. The overall population and the split Supergroup time series graphs reveal some seasonality in the distribution of prescriptions. December consistently showed up to be the most prescribed month with July and October coming in a close 2nd and 3rd. The least prescribed months are contained in the first half of the year January to June. The peak in December could be explained due to "holiday stress". The financial stresses and the extra workloads associated with the holiday season is well documented by Psychologists (Greenberg and Berktold, 2006). Individuals may not be necessarily suffering clinical depression, but may show temporary symptoms stemming from holiday stress, which would prompt the prescription of antidepressants. Similarly, the higher prescription numbers in July might be explained by factors like the financial strain of holidays, disruptions to regular routines, and body image concerns, all of which can contribute to depressive symptoms in individuals (Griffin, 2021).

Finally, when grouping the Supergroups into Urban = (2,3,4,5,7,8) and Rural = (1,6). The corresponding time series trend comparisons (Figures 6 and 7) yielded little to no difference in seasonality. While it has been previously demonstrated that prescription rates are higher in urban areas, this elevated prevalence is not apparent when examining monthly data trends. One can infer that when directing resources for mental health services and depression awareness on a monthly basis, there shouldn't be a distinction between urban and rural areas.

4.4 Related work

Previous work has been done to look at the differences in depression levels for urban and rural areas. In particular, a study consisting of 4.4 million people in Sweden found that those living in the most densely populated areas had up to 20% more chance of developing depression. It concluded that in densely populated areas, urbanisation was associated with depression increases (Sundquist et al., 2004). Another study conducted in the Brussels capital region examined the long-term environmental effects such as, air pollution, noise and building morphology on depressive disorders. The study discovered a positive correlation between exposure to traffic-related air pollution and an increased likelihood of depressive disorders (Pelgrims et al., 2021). One study compared depression rates for similar age groups of residents living in urban versus rural environments. It concluded that rural residence appeared to be a "buffer" against major depression (Jr et al., 1986).

Although the findings of this paper show that urban mental health problems are more prevalent, there is a need to improve mental health services for rural populations. One paper argues that there is a stigma associated seeking help for mental illness due to the "effect of rural culture" (Nicholson, 2018). Due to the geographical properties of remote, rural locations, providing mental health services may be more challenging. While rural communities may have less antidepressant prescriptions, a practical reason to explain that may be the increased distance needed to travel to medical practices. A study on Australian rural communities concluded that residing in a rural or isolated area was linked to a decreased likelihood of using psychiatrist services, while living in a remote location showed a reduced association with the utilization of psychologist services (Fraser et al., 2002).

References

J. Van Bavel. The world population explosion: causes, backgrounds and -projections for the future. *Facts Views Vis Obgyn*, 5(4):281–91, 2013. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3987379/.

John Bongaarts. Human population growth and the demographic transition. *Philos Trans R Soc Lond B Biol*, 364(1532):22985–90, 2009. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2781829/.

Louise Bradvik. Suicide risk and mental disorders. int j environ res public health. *Int J Environ Res Public Health*, 15(9):2028, 2018. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6165520/.

Ron Chepesiuk. Missing the dark: Health effects of light pollution. *Environ Health Perspect*, 1(117):A20–A27, 2009. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2627884/.

Kenneth R. Conner, Martin Pinquart, and Stephanie A. Gamble. Meta-analysis of depression and substance use among individuals with alcohol use disorders. *J Subst Abuse Treat*, 37(2):127–37, 2009. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4864601/.

Pim Cuijpers, Argyris Stringaris, and Miranda Wolpert. Treatment outcomes for depression: challenges and opportunities. *The Lancet Psychiatry*, 7(11): 925–927, 2020. URL https://www.thelancet.com/journals/lanpsy/article/PIIS2215-0366(20)30036-5/fulltext.

Kingsley Davis. The urbanization of the human population. *Scientific American*, 213(3):40–53, 1965. URL http://www.jstor.org/stable/24931110.

Caitlin Fraser, Fiona Judd, Henry Jackson, Greg Murray, John Humphreys, and Gene A. Hodgins. Does one size really fit all? why the mental health of rural australians requires further research. *The Australian Journal of Rural Health*, 10 (6):288–295, 2002. URL https://doi.org/10.1046/j.1440-1584.2002.00463.x.

Lixia Ge, Chun Wei Yap, Reuben Ong, and Bee Hoon Heng. Social isolation, loneliness and their relationships with depressive symptoms: A population-based study. *PLoS One*, 12(8), 2017. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5568112/.

Edward L. Glaeser and Bruce Sacerdote. Why is there more crime in cities? *Journal of Political Economy*, 107(S6):S225–S258, 1999. URL https://www.jstor.org/stable/10.1086/250109?seq=10.

Anna Greenberg and Jennifer Berktold. Holiday stress. *American Psychology Association*, 2006. URL https://www.apa.org/news/press/releases/2006/12/holiday-stress.pdf.

Morgan R. Griffin. Tips for summer depression. *WebMD*, 2021. URL https://www.webmd.com/depression/summer-depression.

Lene Hammer-Helmich, Josep Maria Haro, Bengt Jonsson, Audrey Tanguy Melac, Sylvie Di Nicola, Julien Chollet, Dominique Milea, Benoit Rive, and Delphine Saragoussi. Functional impairment in patients with major depressive disorder: the 2-year perform study. *Neuropsychiatr Dis Treat*, (14), 2018. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5767094/.

Bradford A. Crowell Jr, Linda K. George, Dan Blazer, and Richard Landerman. Psychosocial risk factors and urban/rural differences in the prevalence of major depression. *The British Journal of Psychiatry*, 149(3):307–314, 1986. URL https://doi.org/10.1192/bjp.149.3.307.

Raymond W Lam, Sidney H Kennedy, Roger S McIntyre, and Atul Khullar. Cognitive dysfunction in major depressive disorder: effects on psychosocial functioning and implications for treatment. *Can J Psychiatry*, 59(12):649–54, 2014. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4304584/.

Lu Liang and Peng Gong. Urban and air pollution: a multi-city study of long-term effects of urban landscape patterns on air quality trends. *Scientific Reports*, 10 (18618), 2020. URL https://doi.org/10.1038/s41598-020-74524-9.

Richa Mahtta, Michail Fragkias, Burak Guneralp, Anjali Mahendra, Meredith Reba, Elizabeth A. Wentz, and Keren C. Seto. Urban land expansion: the role of population and economic growth for 300+ cities. *npj Urban Sustainability*, (5), 2022. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3987379/.

Robert Ian McDonald, Timothy Beatley, and Thomas Elmqvist. The green soul of the concrete jungle: the urban century, the urban psychological penalty, and the role of nature. *Sustainable Earth*, 1(3):281–91, 2018. URL https://doi.org/10.1186/s42055-018-0002-5.

Laura Anne Nicholson. Rural mental health. *Advances in Psychiatric Treatment*, 14(4):302–311, 2018. URL https://www.cambridge.org/core/journals/

advances-in-psychiatric-treatment/article/rural-mental-health/C83E79A4B4C8C04357C7B5BC1A0B43D5.

Ingrid Pelgrims, Brecht Devleesschauwer, Madeleine Guyot, Hans Keune, Tim S. Nawrot, Roy Remmen, Nelly D. Saenen, Sonia Trabelsi, Isabelle Thomas, Raf Aerts, and Eva M. De Clercq. Association between urban environment and mental health in brussels, belgium. *BMC Public Health*, 21(635):302–311, 2021. URL https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-021-10557-7.

Kyriaki Remoundou and Phoebe Koundouri. Environmental effects on public health: an economic perspective. *Int J Environ Res Public Health*, 6(8):2160–78, 2009. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2738880/.

Hannah Ritchie and Max Roser. Urbanization, 2019. URL https://ourworldindata.org/urbanization.

Jaime Ruiz-Tagle and Ignacio Urria. Household overcrowding trajectories and mental well-being. *Social Science and Medicine*, 296(114051), 2022. URL https://doi.org/10.1016/j.socscimed.2021.114051.

Biing-Jiun Shen, Stacy A Eisenberg, Uta Maeda, Kristin A Farrell, Ernst R Schwarz, Frank J Penedo, Eugene J Bauerlein, and Stephen Mallon. Depression and anxiety predict decline in physical health functioning in patients with heart failure. *Ann Behav Med*, 41(3):373–82, 2011. URL https://pubmed.ncbi.nlm.nih.gov/21181518/.

Jiyao Sheng, Shui Liu, Yicun Wang, Ranji Cui, and Xuewen Zhang. The link between depression and chronic pain: Neural mechanisms in the brain. *Neural Plast*, (9724371), 2017. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5494581/.

Kristina Sundquist, Golin Frank, and Jan Sundquist. Urbanisation and incidence of psychosis and depression. *British Journal of Psychiatry*, 184:293–298, 2004. URL https://www.cambridge.org/core/journals/the-british-journal-of-psychiatry/article/urbanisation-and-incidence-of-psychosis-and-depression/AF3FDF51E9DA192097BEF153D9A02148.

Angela Sweeney, Steve Gillard, Til Wykes, and Diana Rose. The role of fear in mental health service users' experiences: a qualitative exploration. *Soc Psychiatry Psychiatr Epidemiol*, 50(7):1079–87, 2015. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4463981/.

UN. Global issues population, 2023. URL https://www.un.org/en/global-issues/population#:~:text=Our%20growing%20population&text=The%20world's%20population%20is%20expected,billion%20in%20the%20mid%2D2080s.

WHO. Mental health, 2023a. URL https://www.who.int/health-topics/mental-health#tab=tab_1.

WHO. Depressive disorder(depression), 2023b. URL https://www.who.int/news-room/fact-sheets/detail/depression#:~:text=An%20estimated%203.8%25%20of%20the,world%20have%20depression%20(1).

Zhaoxi Zhang, Kristyna Mechurova, Bernd Resch, Prince Amegbor, and Clive E. Sabel. Assessing the association between overcrowding and human physiological stress response in different urban contexts: a case study in salzburg, austria. *International Journal of Health Geographics*, 22(15), 2023. URL https://doi.org/10.1186/s12942-023-00334-7.