Hackathon Three – Ecosystem services

1 Introduction

The aim of this hackathon is to determine if the way we interact with land, and the ecosystem that is present, has impact on the spread of covid-19 and the outcome of getting infected with covid-19.

2 Methodology

Data was drawn from different sources [1], [2] and [3] to create the dataset for this hackathon. The numbers of deaths, by local authority area, in week 16 from all causes and from Covid-19 were added to the dataset from [1], population information from [2] and land usage from [3]. During week 16 the UK population was in lock-down and interpersonal contact was being kept to a minimum.

The land usage data provided broke land usage down into percentages as follows.

- Developed Land
    - Community Service
    - Defence
    - Industry and Commerce
    - Minerals and Landfill
    - Residential
    - Transport and utilities
    - Unknown developed use
    - Total
- Non-Developed Land
    - Agriculture
    - Forest open land and water
    - Outdoor recreation
    - Residential garden
    - Undeveloped land
    - Total
- Vacant land
- Total

Death rates per 100,000 were calculated for each local authority area. To remove outliers only local authority areas with a population >= 100,000 were used. This removed some of the smaller and lower population density local authority areas from the dataset, but it left 215 more consistent records. Outliers often provide researchers with insight, but in this case, they were deemed not to. Where populations were low any deaths would affect the death rates per 100,000 calculation disproportionately. Correlation, using the Pearson's Correlation test, was then tested for between each variable.

3 Results

No statistically significant correlations were found between deaths from all causes per 100,000 population and any of the land usages.

Statistically significant correlations were found between deaths from covid-19 per 100,000 population and four of the land usages, these are shown in Table 1.

Land Usage	Pearson's Correlation Coefficient	Significance Level (2-tailed)
Unknown developed use	0.145	0.034
Agriculture	-0.148	0.030
Outdoor recreation	0.210	0.002
Undeveloped land	0.179	0.009

Table 1: Statistically significant correlations between covid-19 deaths per 100,000 population and land usage

One of these correlations (Agriculture) is negative, indicating that as the percentage of land used for agriculture increases the deaths from covid-19 per 100,000 population decreases. This could be taken to indicate that going back to an agrarian past might create a society that is less vulnerable to covid-19. However, it might simply indicate that farmers live more isolated lives than their urban counterparts and therefore are less likely to contract covid-19.

Three of the correlations (Unknown developed use, Outdoor recreation and Undeveloped land) are positive, indicating that as the percentage of land put aside for each of these uses increases so does the deaths from covid-19 per 100,000 population. Care must be taken though as correlation does not indicate causation, see Vigen (2015) for examples of spurious correlations [4], or simply visit the website http://www.tylervigen.com/spurious-correlations. In this case however there may be causation as well as correlation. In week 16 the UK population was in lock-down and so contact between people should have been minimal. It happened to coincide though with unexpectedly pleasant weather; the sun shone, and the temperature was very hot. It may be that people went outside to enjoy this weather; they may have gone to outdoor recreation land, or where that was unavailable other deserted pieces of land (unknown developed use and undeveloped land). Congregating in these areas enhanced the spread of the disease leading to an increased death rate from covid-19. These positive correlations may be evidence that the breaking of lock-down caused extra unnecessary deaths. These correlations are very small; for example, the coefficient of determination for outdoor recreation land is only 4.41%, indicating that its contribution to the deaths from covid-19 per 100,000 population is quite small. However, these are unnecessary deaths and will cause pain to the victims loved ones.

4 Conclusion

The data used in this investigation is a snapshot when the UK population was in lock-down. Consequently, this investigation only considers land usage while the UK population is in lock-down and can't make a comparison with not being in lock-down. Comparing data from a pre-lock-down week is not feasible as many other variables will have been in play.

The data may show that when people broke lock-down to enjoy the sunny weather they inadvertently increased the spread of covid-19 leading to an increased death rate.

5 Finally

This investigation has not found any relationship between land usage and the death rate from covid-19, rather it may have found a relationship between people's behaviour and the death rate from covid-19. As we move beyond lock-down it may be people's behaviour that is the important factor to consider. Publicising how behaviour has cost lives in the past may help save lives in the future.

References

[1] Weekly provisional figures on deaths occurring and registered in England and Wales by Local Authority, UK Statistics Authority

[2] Table KS101EW, 2011 Census: Usual resident population, local authorities in England and Wales, Office for National Statistics

[3] Land Use Statistics England 2017, Ministry of Housing, Communities and Local Government

[4] Vigen, T., 2015. *Spurious correlations*. Hachette books.