Leveraging artificial intelligence to analyze the COVID-19 distribution pattern based on socio-economic determinants

Mohammadhossein Ghahramani *, Francesco Pilla

Spatial Dynamics Lab, University College Dublin, Ireland

ABSTRACT

The spatialization of socioeconomic data can be used and integrated with other sources of information to reveal valuable insights. Such data can be utilized to infer different variations, such as the dynamics of city dwellers and their spatial and temporal variability. This work focuses on such applications to explore the underlying association between socioeconomic characteristics of different geographical regions in Dublin, Ireland, and the number of confirmed COVID cases in each area. Our aim is to implement a machine learning approach to identify demographic characteristics and spatial patterns. Spatial analysis was used to describe the pattern of interest in electoral divisions (ED), which are the legally defined administrative areas in the Republic of Ireland for which population statistics are published from the census data. We used the most informative variables of the census data to model the number of infected people in different regions at ED level. Seven clusters detected by implementing an unsupervised neural network method. The distribution of people who have contracted the virus was studied.

1. Introduction

In March 11th, 2020, the Republic of Ireland's government launched a national action plan in response to COVID-19, a widespread lock-down in order to minimize the risk of illness. The impacts of pandemics such as the current COVID-19 should be explored extensively. To mitigate and recover from the negative repercussions, it is of paramount importance to study the effects on the social tissue in cities. It seems that various research is needed to thoroughly investigate, understand, mitigate and recover from the effect of this pandemic. Some studies have been focused on providing risk assessment frameworks based on artificial intelligence and leveraging data generated from heterogeneous sources such as disease-related data, demographic, mobility, and social media data (Beria & Lunkar, 2021; Ge et al., 2020; Sannigrahi, Pilla, Basu, Basu, & Molter, 2020; Shokouhyar, Shokoohyar, Sobhani, & Gorizi, 2021; Silva et al., 2021). The exposure risk of the pandemic in different environments has been assessed. Many researchers are exploring the dynamics of the pandemic in urban areas to mitigate effects and understand the impacts of COVID-19 on cities (Das et al., 2021; Rumpler, Venkataraman, & Göransson, 2020; Silva et al., 2021). In this area of research, four distinctive categories have received significant attention: environmental quality, socio-economic impacts, management and governance, and transportation and urban design (Sharifi & Khavarian-Garmsir, 2020). As far as the socio-economic impacts are concerned, pandemics can substantially negatively affect people at the bottom of the socio-economic hierarchy, those with low education, low income, and low-status jobs. For instance, it has been discussed that the Black and Latino people's mortality rate is twice that of the Whites in the US (Wade & Khavarian-Garmsir, 2020). The pandemics can also hit vulnerable groups of people in poor sanitary conditions. Moreover, various factors such as high density, inadequate access to health services and infrastructure facilities can exacerbate the situation (Duffey & Zio, 2020; Rahman et al., 2020). Different inequality issues can also make it difficult to maintain social distancing (Sun & Zhai, 2020). Hence, it is essential to understand the existed relation between socio-economic inequalities and the pandemic. As discussed, such inequalities can threaten public health by making it difficult to enforce protective measures such as social distancing.

Artificial Intelligence technologies such as neural networks and deep learning can play a significant role during a pandemic. They can be used to provide different platforms for social distance tracking (Ahmed, Ahmad, Rodrigues, Jeon, & Din, 2021; Ghahramani, Galle, Duarte, Ratti, & Pilla, 2021; Nagrath et al., 2021), monitor and control the spread of COVID-19 (Bhattacharya et al., 2021; Zivkovic et al., 2021).
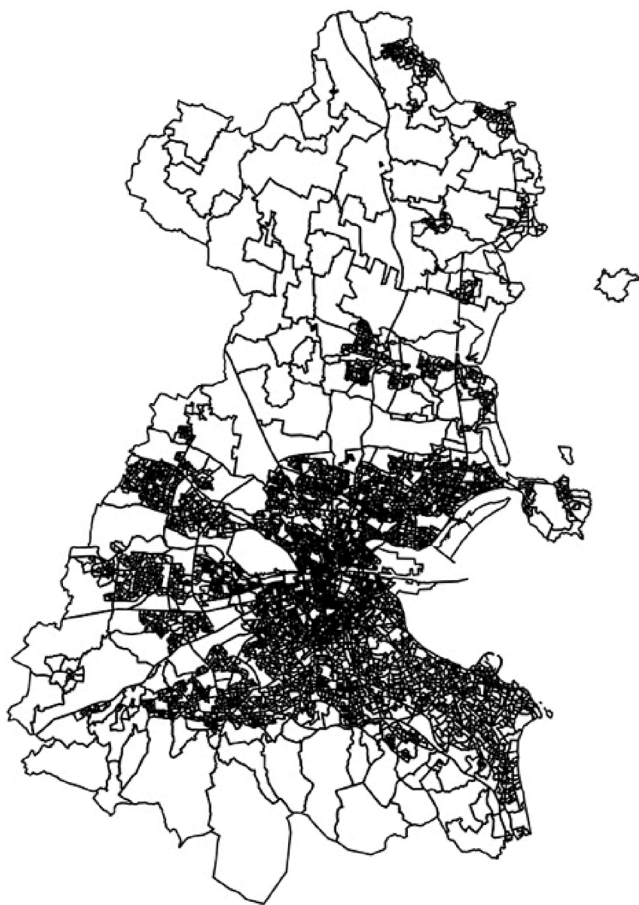

Fig. 1. Dublin shapefile including different polygons of the administrative boundary and attributes of geographic features.

Such technology has been used in this study. We assess the association between the demographic features and the number of confirmed cases at electoral divisions (i.e., ED) in Dublin, Ireland based on an optimized self-organizing neural network. It should be mentioned that the number of cases until September 10, 2020, have been considered in this work. Our aim is to understand the impacts of the pandemic on Dublin city given associated characteristics and study the related patterns in different clusters obtaining from demographic information, i.e., census data. We used a machine learning method based on an unsupervised learning approach to group spatial data into meaningful clusters (Hu, O'Hagan, Sweeney, & Ghahramani, 2020). In doing so, the similarities among spatial objects were taken into account. Given the implemented model, the implicit information about different EDs were extracted, and all associated relations were examined. Such data exploration can help us extract demographic information related to various clusters. First, a feature selection method was used to extract the most relevant variables since the census data includes over 700 features, and redundant features can significantly affect the model accuracy. Feature extraction aims to project high-dimensional data sets into lower-dimensional ones in which relevant features can be preserved. These features, then, were used to distinguish patterns. Dimensionality reduction and feature selection/extraction methods (Ghahramani, Qiao, Zhou, O'Hagan, & Sweeney, 2020), e.g., principal component analysis (PCA), linear discriminant analysis (LDA), and canonical correlation analysis (CCA), play a critical role in dealing with noise and redundant features. These methods were used as a pre-processing phase of data analysis and helped us obtain better insights and robust decisions.

Broadly speaking, dimensionality reduction is considered as a method to remove redundant variables. This technique can be regarded

as two distinctive approaches, i.e., feature extraction and feature selection. Feature extraction refers to those techniques that project original variables to a new latent space with lower dimensionality, while feature selection methods aim to choose a subset of variables such that a trained model minimizes redundancy and maximizes relevance to the target feature. In this work, we deal with a clustering problem and high-dimensionality issue; hence, a feature extraction technique was used. Since interpreting associated patterns in feature extraction methods can be a subjective process, different tests were implemented to deal with related issues such as readability and interpretability. PCA is a classic approach to dimensionality reduction (feature extraction) and has been implemented in various research studies. However, it suffers from a global linearity issue. Thus, to address this concern, a nonlinear technique (i.e., kernel PCA (Kim & Klabjan, 2020)) was used in this work.

Then, the extracted features from the census data were fed into a clustering model, and different clusters were identified. The goal in this phase is to cluster EDs (including various demographic variables) such that similarities among them within each group are maximized. The model is based on an advanced spatial clustering technique and can deal with non-linear relationships between features of a high dimensional data set. To do so, we implemented an unsupervised approach based on an artificial neural network (ANN) that can properly transform geo-referenced data into information. The main property of ANNs is their ability to learn and model nonlinear and complex relationships. The model employs a competition-based learning mechanism to generate insights from unlabelled data. It leverages a multi-layer clustering approach, i.e., a self-organizing neural network (Díaz Ramos, López-Rubio, & Palomo, 2020; Yu, Lu, & Zhang, 2020), to transform a complex high-dimensional input space into low dimensional output space while preserving the topology of the data. Given a set of EDs, the model groups together different spatial objects that are similar with other (i.e., the distance among observations is minimized in a given cluster). Different validity measures were also applied and the results are illustrated. For visualization, we use the shapefile of Dublin. Fig. 1 demonstrates the Dublin shapefile, including different districts.

The contributions of this work are as follows:

1. The link between the number of confirmed Covid cases and socio-economic determinants at electoral division level in Dublin, Ireland is analyzed based on an AI-based spatial clustering method.
2. A topology-preserving model is implemented to explore nonlinear relationship among electoral divisions given the census data to characterize the spatial distribution of city dwellers.

The remainder of this paper is organized as follows: some related work on application of machine learning and artificial intelligence to deal with concerns related to the pandemic is described in Section 2; data pre-processing operations including feature extraction is explained in Section 3; the proposed approach with its associated discussions is presented in Section 4; Section 5 shows the experimental settings and the clustering results; and the future work and conclusions are presented in Section 6.

2. Related work

Due to the global spread of coronavirus, many researchers across the world are working to understand the underlying patterns of the pandemic from different perspectives. They are looking for effective ways to manage the flow of people and prevent new viral infections. As expected, numerous research has been undertaken as to medical concerns (e.g., diagnosis and treatment of the disease like lung disease, lung nodules, chronic inflammation, chronic obstructive pulmonary diseases) to ensure all required measures are in place. Different strategies, such as chest computed tomography imaging (Xie et al., 2020) and polymerase chain reaction (Hu, Gao, et al., 2020), have been discussed for detecting and classifying COVID-19 infections. Artificial intelligence (AI)
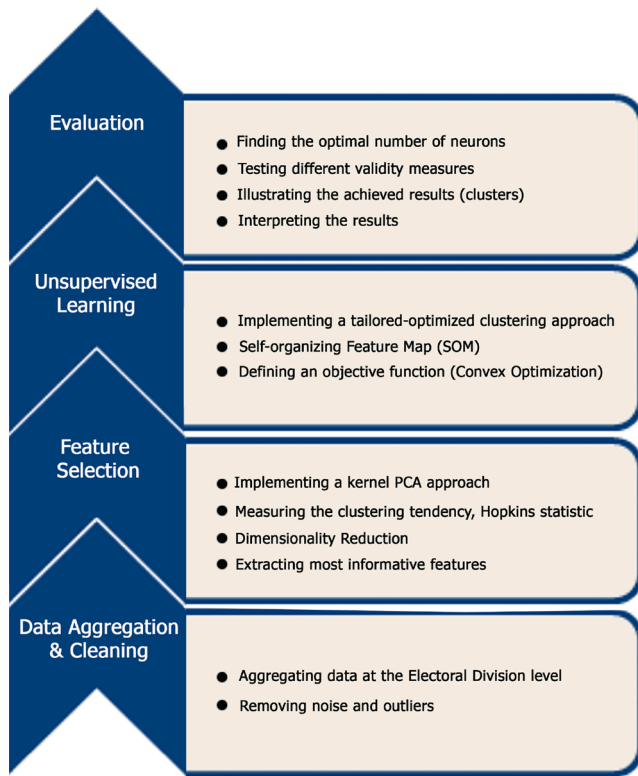
Fig. 2. Different phases of the analysis model used in this work.

approaches have also been used in the field of medical data analysis (Bhattacharya et al., 2021), and different algorithms have been implemented for such analysis and patients' classification. Different neural network techniques have been utilized for diagnosis based on identified clinical characteristics such as cough, fever, sputum development, and pleuritic chest pain (Li et al., 2020; Ouyang et al., 2020). Various impacts of the pandemic on urban areas have also attracted the attention of researchers. In Alsaeedy and Chong (2020), the authors have introduced a novel method to identify regions with high human density and mobility, which are at risk for spreading COVID-19 by exploiting cellular-network functionalities. In doing so, they have used the frequency of handover and cell selection events to identify the density of congestion. Several visualization techniques like class activation mapping (CAM) (Sun et al., 2020), class-specific saliency map, and gradient-weighted class activation mapping (Grad-CAM) (He et al., 2020) has been used to generate localization heatmaps in order to highlight crucial areas that are closely associated with the pandemic. Rustam et al., have implemented four Machine Learning models, such as linear regression, least absolute shrinkage, and selection operator, support vector machine, and exponential smoothing to understand the threatening factors of COVID-19 (Rustam et al., 2020). Different features, such as the number of newly infected cases, the number of deaths, and the number of recoveries have been taken into account in their model.

Network analysis, as a set of integrated techniques, can be used to

provide direct visualization of the pandemic risk. By illustrating the degree of similarity among various areas given confirmed cases, So et al. have demonstrated that network analysis can provide a relatively simple yet powerful way to estimate the pandemic risk (So, Tiwari, Chu, Tsang, & Chan, 2020). Such analysis can also supplement traditional modelling techniques to improve global control and prevention of the disease and provide more timely evidence to inform decision-making in crisis zones. In Montes-Orozco et al. (2020), the authors have presented a methodology to identify spreaders using the analysis of the relationship between socio-cultural and economic characteristics with the number of infections and deaths caused by the virus in different countries. The authors have explored the effect of socioeconomics, population, gross domestic product, health, and air connections by solving a vertex separator problem in multiplex complex networks.

Targeting policy responses to crises such as the current pandemic and interventions exclusively on people who live in deprived areas requires insights such as which clusters in society are most affected. In this work, we explore demographic and socioeconomic factors and investigate the role of socioeconomic factors in the spread of COVID-19. Our aim is to analyze underlying features obtained from census data and describe such demographic information concerning the geolocation of patients. We study the link of the pandemic with such factors. Fig. 2 illustrates different phases of the proposed model.

3. Data processing

Geodemographic is referred to as the study of spatial patterns and socio-economic characteristics of different areas. Associated demographic databases, such as census data, can be used to understand population diversity better since they include characteristics of a country's inhabitants. Generally speaking, spatio-temporal datasets can be divided into different categories, such as geo-referenced data points, geo-referenced time series, moving objects, and trajectories. The estimation of a region's population has been a critical application of geospatial science in demography. In this sense, geodemographic clustering can be considered as a tool to understand spatially dependent datasets. This kind of clustering is unsupervised learning that groups spatial data into meaningful clusters based on similarities among various areas. The learning procedure is correlated to the tendency of people to associate themselves with others who have common characteristics. Census data can be considered as a reference for overall population estimation. It includes information about individuals who have been counted within households in different regions. Such data sets have some special characteristics such as geospatial features. They consist of measurements or observations taken at specific locations, referenced by latitude and longitude coordinates and/or associated within specific regions (in this work electoral divisions). Census data for the population living in the Republic of Ireland are available at a different level, i.e., small area and electoral division (ED), from a survey taken in 2016. However, since the number of confirmed cases are available at EDs, the census data at such administrative areas were incorporated.

3.1. Dataset

Demographic information is available at the local population level via censuses carried out by countries. In Ireland, a census is conducted at

Table 1
Some observations of the census data at electoral divisions level consisting of 764 variables.

GEOGID	GEOGDESC	T1-1AGE0M	T1-1AGE1M	T1-1AGE2M	T1-1AGE3M	T1-1AGE4M	...	T15-3-N	T15-3-NS	Covid cases
E02008	Ayrfield	33	33	34	31	37	...	341	43	133
E02012	Ballygall B	10	10	5	8	11	...	266	27	109
E02022	Beaumont B	29	26	35	24	21	...	270	38	75
E02006	Ashtown A	100	84	70	66	49	...	626	111	99
E02093	Whitehall D	11	15	12	11	5	...	258	16	150

Table 2
Summary information on a subset of summarized variables from the Irish census data across all EDs.

Features	Statistics				
	Mean	Std deviation	Median absolute deviation	IQR	Median
Percentage of population aged 0–4	7.298	2.168	1.425	[5.797, 8.638]	7.238
Percentage of population aged 5–14	14.053	3.379	1.964	[12.272, 16.228]	14.313
Percentage of population aged 65 and over	13.580	4.413	2.620	[10.721, 16.071]	13.243
Percentage of single population	56.157	4.881	2.432	[53.146, 58.103]	55.468
Percentage of house-share household	4.254	4.147	1.389	[3.112, 5.984]	4.347
Percentage with higher education degrees	20.471	9.131	4.292	[14.908, 23.724]	18.501
Percentage of professional social class	4.981	3.816	1.863	[2.511, 6.417]	4.098
Percentage of unemployed population	11.015	3.938	2.436	[8.241, 13.249]	10.526

five-year periods by the government, with the most recent census prior to this work occurring in 2016. The census of Ireland is disseminated by the Central Statistics Office (CSO) and provides a vast amount of information. Spatial data like a census typically involves a large number of observations, meaning analysis of this nature tends to involve complex multivariate analysis and machine learning methods (Ghahramani, Zhou, & Hon, 2019a, 2019b; Ghahramani, Zhou, & Wang, 2020). There are 322 EDs in Dublin, and the census consists of 764 features (relating to, for example, age, household size, marriage status, and education levels, etc.) for each of 322 EDs. The census reports the features as a count of people. We converted these features to percentages of the population within each ED. Some sample records are presented in Table 1. The number of Covid cases are also aggregated in this table. There are no missing values or outliers in the census data. The dataset were normalized; the variables were scaled and transformed so that they each make an approximately equal contribution to the results. For example, there are about 100 variables relating to age information in the raw census data that they are summarized into percentages of different age bands; and there are about 40 variables relating to education levels that are converted to percentages of people holding a third-level higher education degree and above for each area. Take some variables demonstrated in Table 1 as an example. The variables T1-1AGE0M, T1-1AGE1M, T1-1AGE2M, T1-1AGE3M, and T1-1AGE4M, which refer to the number of people in different age bands (infants to four years old) have been merged, and a new feature Age0-4 has been created. In total, we extracted 53 variables that are synthesized from the census data, and a subset of these variables is presented in Table 2. For the sake of brevity, not all summarized census variables are presented and discussed in detail. All the features created in this phase are used in a dimensionality reduction phase to be explained later. It should be mentioned that spatial features cannot be illustrated or modelled in a simple way due to their complex characteristics, e.g., size, boundaries, direction and connectivity. Hence, spatial analysis is more sophisticated than relational data processing in terms of algorithmic efficiency and the complexity of possible patterns because interrelated information at a spatial scale has

to be considered. Therefore, spatial or geodemographic clustering is used for grouping and labelling geographical neighbourhoods in terms of their social and economic characteristics. Such an approach can be used to understand our spatially dependent data and the potential underlying associations between this data and confirmed number of Covid cases. Such applications allow similarities between patient structures in different EDs to be highlighted, geodemographically speaking.

Each observation (EDs consisting of demographic information) can be defined as an m-tuple (m is the number of features).

Let matrix $X \in \mathbf{R}^{n \times m}$ as:

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \quad (1)$$

where $\mathbf{R}$ is the real number set, X_i is the ith region and its corresponding variables (m-tuple), and n is the number of all areas. As stated earlier, we deal with high dimensionality in this work. Such datasets can pose serious challenges, such as model overfitting. The more the number of variables increases, the more the chance of overfitting.

3.2. Dimensionality reduction

Dimensionality reduction is the process of eliminating redundant variables. To handle such concerns, different approaches have been considered in the literature. Generally speaking, feature extraction and feature selection techniques are applied to reduce data dimensionality. In the former approach, original features are mapped to a new feature space with lower dimensionality. The latter refers to those methods that identify and select a subset of features such that the trained model (based on the selected features) minimizes redundancy and maximizes relevance to the target feature. PCA is the most common dimensionality reduction approach; however, the transformation applied is linear. But when data follow a nonlinear structure, as in our case, approximating the model by a linear method like PCA will not perform well on the original data. Likewise, multidimensional scaling (Saeed, Nam, Al-Naffouri, & Alouini, 2019) and independent component analysis (ICA) (Feng & Li, 2020; Shi, Yang, Xu, Zhang, & Farahani, 2019) suffer from the linearity issue. To address this shortcoming, nonlinear techniques such as kernel PCA, Laplacian eigenmaps (Sun, 2019), and semidefinite embedding (Xiang, Nie, Zhang, & Zhang, 2009) can be used. The two first-mentioned methods have been applied in this work. The result of the Kernel PCA is illustrated to save space. We can define the variance-covariance matrix as

$$S = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^T (X_i - \overline{X}) \quad (2)$$

The aim is to maximize the trace of the covariance matrix (i.e., $A^* = \arg\max_A tr(S)$) given a weighted covariance eigen decomposition approach (Chan, Wu, & Tsui, 2012), where A is a set of eigenvectors (unitary matrices that can represent rotations of the space). A nonlinear transformation $\phi(X)$ from the original m-dimensional space has been considered, and the covariance matrix of the projected features has been measure as

$$S = \frac{1}{n} \sum_{i=1}^{n} \phi(X_i) \phi(X_i)^T \quad (3)$$

The eigenvalues and eigenvectors are given by

$$S\nu_k = \lambda \nu_k \quad (4)$$

The eigenvectors have been measured ($\nu_k = \sum_{i=1}^{n} a_{ki} \phi(X_i)$), where k is the new number of dimensions.
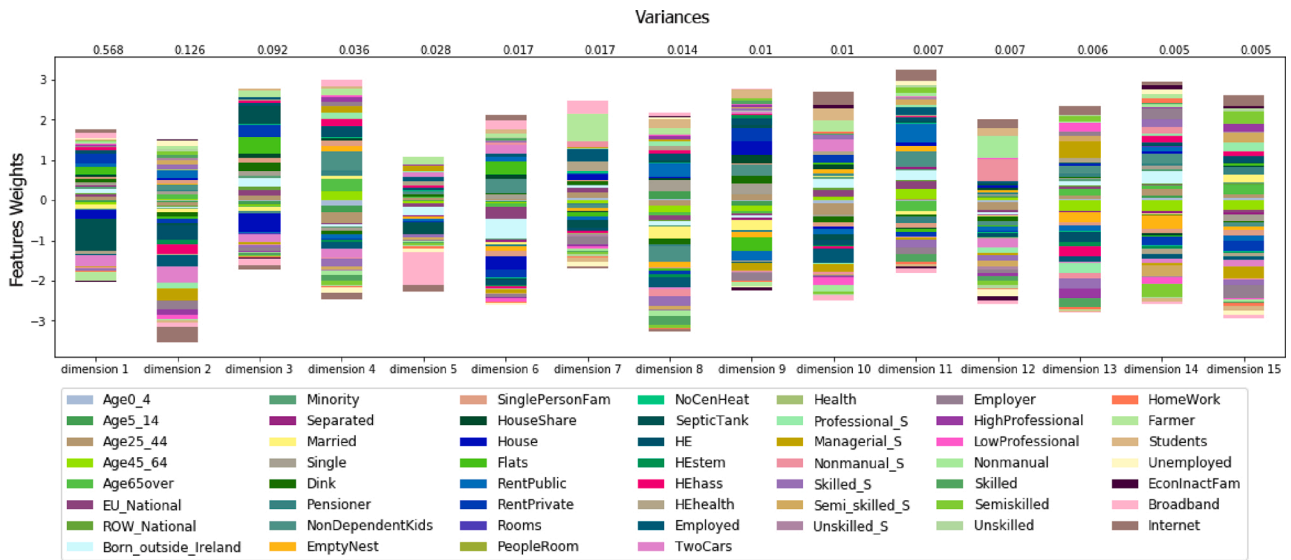
Fig. 3. Result of the dimensionality reduction phase implemented for feature extraction based on Kernel PCA.

$$\frac{1}{n} \sum_{i=1}^{n} \phi(X_i)\{\phi(X_i)^T \nu_k\} = \lambda_k \nu_k \tag{5}$$

By substituting ν_k in above equation

$$\frac{1}{n} \sum_{i=1}^{n} \phi(X_i)\phi(X_i)^T \sum_{j=1}^{n} a_{ki}\phi(X_j) = \lambda_k \sum_{i=1}^{n} a_{ki}\phi(X_i) \tag{6}$$

The kernel function ($\Psi(X_i, X_j) = \phi(X_i)^T \phi(X_j)$) is, then, multiply both sides of Eq. (6) and the kernel principal components can be calculated as:

$$\phi(X)^T \nu_k = \sum_{i=1}^{n} a_{ki}\Psi(X, X_i) \tag{7}$$

It should be mentioned that we have constructed the kernel matrix from the census data. To that end, a Gaussian kernel ($\Psi(X_i, X_j) = \exp(- ||X_i - X_j||^2/2\sigma^2)$) has been used, where c is a constant. Given the measured variance for each feature, the associated weight can be measured

$$\sigma_X^2 = \frac{\sum_{i=1}^{n} \omega_i^2 (X_i - \overline{X})^2}{\sum_{i=1}^{n} \omega_i^2} \tag{8}$$

We have also examined the relevance of all features using the coefficient of determination. In doing so, the proportion of the variances have been tested. A supervised learner has been used, and iteratively one feature of the dataset has been considered as the dependent variable and others as the independent variables. The Hopkins statistic, which is a way of measuring the clustering tendency of a data set, has been calculated for both scenarios with the value of 0.59 before dimensionality reduction and 0.67 after that phase. A value close to 1 indicates that the data is highly clustered. Fig. 3 illustrates the result of the dimensionality reduction given the Kernel PCA approach. Given the fraction of variances measured in this phase and also given all the weights associated to each feature, 21 features, such as percentage of population aged 65 and over, percentage of house-share household, and percentage of the unemployed population, have been selected. All these features have been integrated with two additional variables, i.e., the population of each ED and the number of confirmed covid cases in each of those areas. The final dataset is then used in the second phase (i.e., clustering) of the model.

4. Clustering approach

After performing all the data preprocessing operations explained above, a clustering method can be implemented to find underlying patterns. Due to characteristics of this work, i.e., non-linear dynamics, an unsupervised learning mechanism based on a vector quantization technique (Xie, Chen, Lewis, & Xie, 2018) has been considered. It should be mentioned that most neural network approaches operate based on the non-linear optimization of a criterion, which may result in the local minimum issue and/or the convergence may take a long time. It has been discussed that self-organizing maps are less sensitive to such concerns. This approach is motivated by retina-cortex mapping and considered as an optimal technique for vector quantization problems. The topographic mechanism used in this method can enable us to study relationships among spatial and non-spatial features and identify associated patterns. The model is self-organized and operates based on learning rules and neuron interactions. The learning process is based on cooperation and competition among neurons. Moreover, neurons maintain proximity relationships during the learning process. The idea is to quantize the input space into a finite number of vectors. All observations in the input space (census vectors, together with the number of Covid cases in each spatial area) are projected to post-synaptic neurons in the latent space. The implemented model can transform all the census features in the input space into a low-dimensional discrete output space while preserving the relationships among variables. To do so, all vectors are mapped to neurons based on synaptic connections, each of which is assigned with weights. These weights are updated such that adjacent neurons on the lattice have similar values. The clustering procedures consists of different phases, i.e., competition, collaboration, and weight updating.

In the competition phase of the algorithm, a predefined number of neurons are initialized by randomly setting their weights using census features. Neurons compete for each input vector's ownership, and the most similar neuron (given the distance measure between an ED object together with all relevant features and all neurons) to a given observation is detected. The winning neuron is called the best matching unit (BMU). There are different distance measures to find the similarity between neurons and an input vector, such as the Euclidian distance, correlation tests, and cosine similarity. However, the squared Euclidean distance is often used in a real application. Let X_i be the ith input vector (i.e., ith ED's features) and W_j the associated weights of the jth neuron. Then, the distance matrix $D_{ij} = \frac{1}{n}\sum_{i=1}^{n} \sum_{j=1}^{k} (X_i - W_j)^2$ can be defined as:

$$D_{ij} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & \ldots & d_{1k} \\ d_{21} & d_{22} & d_{23} & \ldots & d_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & d_{n3} & \ldots & d_{nk} \end{bmatrix} \tag{9}$$

The BMU can be measured according to

$$\Psi = \arg\min_j ||X_i - W_j||_2 \tag{10}$$

In the collaboration phase, the adjacent neurons of a given BMU are updated. The aim is to find out which of the non-winning neurons are within the BMU's neighbourhood detected in the previous phase. To do so, the spatial location of a topological neighbourhood of the excited neuron is detected. Several neighbourhood functions can be used to calculate the neighbourhood radius, i.e., rectangular, Mexican hat, and Gaussian functions. The latter (i.e., Gaussian function) is the most commonly used one and has been utilized in this work. The cooperative process in this phase starts with defining an initial neighbourhood radius, which shrinks throughout different iterations based on the neighbourhood function. For each neuron j (N_j) in the neighbourhood of the ith winning neuron (N_i), the algorithm updates all the weights associated with the jth neuron based on a learning rate. It should be mentioned that the weights of other neurons outside of N_i neighbourhood are not adjusted (in a given iteration). The procedure can be defined by the function below:

$$\lambda(\xi_{ij}) = exp\left(-\frac{\xi_{ij}^2}{2\sigma^2}\right) \tag{11}$$

where $\lambda(\xi_{ij})$ is the topological neighbourhood value of the ith winning neuron (N_i), ξ_{ij} is a lateral distance (the distance between Ψ_i and its adjacent neurons N_j), and σ is a function of the number of iterations and starts with an initial value (σ_o). A decay function ($-\frac{n}{T}$) is also employed, $\sigma(n) = \sigma_o.exp(-\frac{n}{G})$, where n is the number of iterations, and G is a constant. By defining the distance function formulated above, the neighbourhood territory for updating all adjacent neurons is explored. Two different connections, i.e., short-range excitatory connections and long-range inhibitory interconnections, are used during the projection process. The former is utilized at the presynaptic layer and the latter at the postsynaptic one. The process can be expressed as:

$$\frac{\partial Y_j(n)}{\partial n} + \tau Y_j(n)$$
$$= \sum_j W_{ij}(n)X_i(n) + \sum_k \eta_k Y_k^*(n) - \sum_{k'} \gamma_{k'} Y_{k'}^*(n)$$

where τ is a constant, $W_{ij}(n)$ is the synaptic strength between input vectors at the presynaptic layer and neurons at the postsynaptic layer, η_k and γ_k are connection weights at the presynaptic and postsynaptic layers, respectively, and Y^* is an active neuron at the postsynaptic layer.

In the third phase, two methods (i.e., Hebb's rule (Martins & de Lima Neto, 2020; Wickramasinghe, Amarasinghe, & Manic, 2019) and Forgetting rule (Chushig-Muzo, Soguero-Ruiz, Engelbrecht, De Miguel Bohoyo, & Mora-Jiménez, 2020)) for adjusting weights of neurons are considered. Based on the Hebb's rule, the change of the synaptic weight (ΔW) is a function of relative neuron spike timing and is proportional to the correlation between an input (X) and an output (Y) of a network, i.e.,

$$\Delta W = \frac{\partial W_{ij}(n)}{\partial t} = \Theta Y_j(n)X_i(n) \tag{12}$$

where Θ is the learning rate ($0 < \Theta < 1$). A sigmoid function has been applied during the learning process on the outputs to make sure that they are not negative.

$$Y_j(n+1) = \Phi\left[W_j^T X(n) + \sum_j \eta Y_j(n)\right] \tag{13}$$

Table 3
Summary of the notations.

Symbol	Meaning
X	Census features
$p = |X|$	The number of observations
k	Size of the lattice
σ	The neighbourhood parameter
Θ	The learning rate
Ψ	The lateral distance
ξ	Best Matching Unit
l^{N_i}	Position of the ith neuron on the lattice

where Φ means a sigmoid function. Since adopting Hebbe's rule for weight updating can make weights saturated, the Forgetting rule ($\beta Y_j(n) W_{ij}(n)$) is also used in the model. Given (12) and the Gaussian neighbourhood function defined by (11), let $\Theta = \beta$, then

$$\beta Y_j(n) = \Theta Y_j(n) = \Theta\lambda(\xi_{ij})$$

we can formulate the synaptic learning rule as:

$$\begin{aligned}\frac{\partial W_{ij}(n)}{\partial t} &= \Theta Y_j(n)X_i(n) - \beta Y_j(n)W_{ij}(n) \\ &= \Theta[X_i(n) - W_{ij}(n)]Y_j(n)\end{aligned} \tag{14}$$

With the above discussions, the weight updating process can be defined as

$$\begin{aligned}W_j(n+1) &= W_j(n) + \Delta W_j \\ &= W_j(n) + \Theta(n)\lambda(\xi_{ij})[X(n) - W_j(n)]\end{aligned} \tag{15}$$

where $\Theta(n)$ is the learning rate for the nth iteration, $W_j(t)$ is the weight vector of the jth neuron, and λ is a neighbourhood function. The learning rate is also a function of time and decreases monotonically, i.e.,

$$\Theta(n) = \Theta_0 exp\left(\frac{n}{-G^2}\right)$$

where Θ_0 is an initial value, G is a constant, and n is the number of iterations.

After the weights for all the input vectors are calculated, both the learning rate and the radius are diminished. The postsynaptic weights are adjusted to resemble the census features and reflect its properties as closely as possible. To sum up the procedures, the pseudo-code of the implemented Self-organizing map is presented in Algorithm 1. The summary of notations used is also given in Table 3. Two quantization and organization criteria have been utilized to measure the reliability of the model. Given such validity measures, the sensitive parameters of the algorithm have been adjusted. A discussion regarding the settings of the algorithm such as the learning rate, the size of lattice (the number of neurons), and level of similarities among neurons are presented next.

4.1. Algorithm convergence and parameter settings

The learning rate and the number of units needed should be set in the algorithm, while the level of similarities among units and the proper number of clusters are designated thereafter. Different techniques can be utilized to explore the convergence of the algorithm, such as quantization error (QE) (Fan, Yang, & Ye, 2018), topographic error, weight-value convergence, and probabilistic models. It should be noted that there is no exact cost function that a self-organizing map (SOM) follows precisely. As explained before, two criteria (i.e., QE and topology preservation metric) have been taken into account to ensure that the output of the model is reliable. The quantization metric was used to assess the required number of neurons. The squared distance between an observation X_i and its corresponding neuron was calculated. In other words, an optimization problem was solved based on the similarity between vectors at presynaptic and postsynaptic layers. The ultimate synaptic
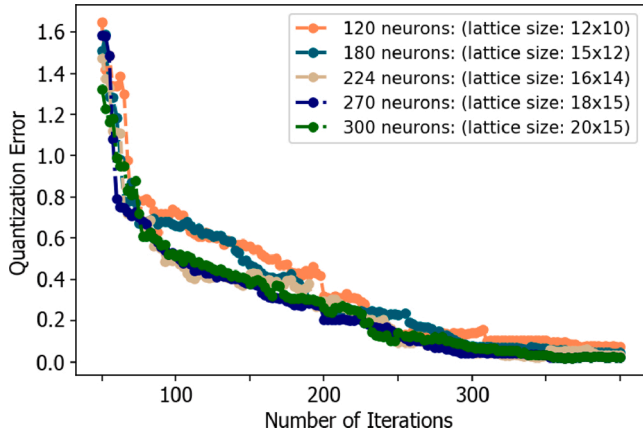
Fig. 4. Comparing the quantization error given different lattice size.

weights of neurons were achieved after running Algorithm 1. The metric calculates the variance associated with neurons' synaptic weights by measuring the average distance between each observation and its corresponding BMU, i.e.,

$$Q_E = \frac{1}{p} \sum_{i=1}^{p} ||X_i - \Psi(i)|| \tag{16}$$

where p is the number of observations at the presynaptic layer, summing all the errors can be expressed as:

$$\Omega = \sum_{i=1}^{k} \sum_{X^j \in V^i} \xi^2(X^j, \Psi^i) = \arg\min_{X^j} \xi^2(X^j, \Psi^i) \tag{17}$$

where k is the size of the lattice (the number of neurons at the post-synaptic layer) and V^i is the Voronoi areas associated with the ith BMU (Ψ^i). Therefore, by using such a metric for determining the convergence of the algorithm, the proper number of neurons was detected. The learning rate of the algorithm is a value between 0 and 1. Different initial values for the learning rate of the algorithm were tested, and the results are illustrated in Fig. 4. The initial learning rate has been set to 0.57, and 270 neurons have been considered.

Algorithm 1. Pseudo-code for the SOM model

Input : $X \leftarrow$ Census features, $p \leftarrow |X|$, $k \leftarrow k_0$,
$\sigma \leftarrow \sigma_0$, $\Theta \leftarrow \Theta_0$
$\{N_1, N_2, \cdots, N_k\}$: k neurons;
$\{l^{N_1}, l^{N_2}, \cdots, l^{N_k}\}$: position set;
$\{w^{N_1}, w^{N_2}, \cdots, w^{N_k}\}$: initial weights;
Output: neurons' weight vectors
1 Set N = $\{N_1, N_2, \cdots, N_k\}$;
2 **for** $i \leftarrow 1, \cdots, p$ **do**
3      $\sigma(i) = \sigma_0 \cdot exp(-\frac{i}{T_2})$
4      $\Theta(i) = \Theta_0 \cdot exp(-\frac{i}{T_2})$
5   Select the ith observation (ED) $x_i \in X$;
6      $\Psi = \arg\min_{n \in N} ||x_i - w^N||_2$;
7   **for** $j \leftarrow 1, \cdots, k$ **do**
8          $\xi = ||l^{x_j} - l^{\Psi}||_2$;
9   **if** $\xi < \sigma$ **then**
10             $w^{N_j} = w^{N_j} + \Theta \cdot \lambda(\xi, \sigma, n) \cdot (x_i - w^{N_j})$;
11   **end**
12   **end**
13 **end**
14 Output the result.

5. Results

5.1. Optimal number of clusters

Given the implemented model, the algorithm leads to an organized representation of activation patterns and prototypes that well represent the census features are obtained. The next step is determining the level of similarity among neurons. We have performed different validity measures to divide neurons at the postsynaptic layer into clusters where inter-cluster similarities are minimized while the intracluster similarities are maximized. Let $C = \{C_1, C_2, \ldots, C_m\}$ be a set m clusters' centroids, $N = \{N_1, N_2, \ldots, N_k\}$ be k neurons at the postsynaptic layer and $\varphi(x_i, x_j)$ be the similarity measure between two EDs x_i and x_j. $|N_i|^{\{m\}}$ is the number of neurons in the mth cluster. The first validity measure used in this work, Davies–Bouldin index (DBI), operates based on the inter-cluster and intra-cluster variance. The similarities among all ED's features projected into neurons are considered. Let denote the mean distance of all neurons belonging to cluster C_m to their centroid as:

$$\delta_m = \frac{1}{|N|^{\{m\}}} \sum_{N_i \in CI^{\{m\}}} ||N_i^{\{m\}} - C_m|| \tag{18}$$

Let Δ_{ij} be the distance between two centroids (C_i and C_j). The Davies–Bouldin index can be formulated as:

$$\text{DBI}(p) = \frac{1}{p} \sum_{i=1}^{p} max\left(\frac{\delta_i + \delta_j}{\Delta_{ij}}\right) \tag{19}$$

The number of clusters, i.e., p in (18) which minimizes the index can be considered as an optimal value.

For the second validity metric (i.e., Silhouette index), the within-cluster distance (Eq. (20)), the mean distance among neurons in each cluster (Cl_i), and the intra-cluster similarity (Eq. (21)) between the cluster to which N_i belongs and its nearest cluster are calculated.

$$\alpha(i) = \frac{1}{|N|^{\{m\}} - 1} \sum_{N_i, N_j \in CI^{\{m\}}} d(N_i, N_j) \tag{20}$$

Table 4

Two validity measures tested for selecting an appropriate number of clusters.

Number of clusters	Silhouette index	Davies–Bouldin index
3	0.4212	0.1721
4	0.4961	0.1281
5	0.5007	0.0998
6	0.6741	0.0954
7	0.8311	0.0704
8	0.8019	0.0731
9	0.7702	0.0782

Table 5

The number of confirmed Covid cases across seven clusters; the corresponding values of the cases/population metric for clusters 5, 6, and 7 are higher than those of others.

Clusters	Number of cases	Population	Cases/Pop
Cluster 1	788	97,014	0.0081
Cluster 2	1034	157,018	0.0065
Cluster 3	901	129,784	0.0069
Cluster 4	1077	180,540	0.0059
Cluster 5	2540	271,128	0.0093
Cluster 6	1824	171,103	0.0106
Cluster 7	3635	350,772	0.0103

Table 6

Some characteristics of clusters.

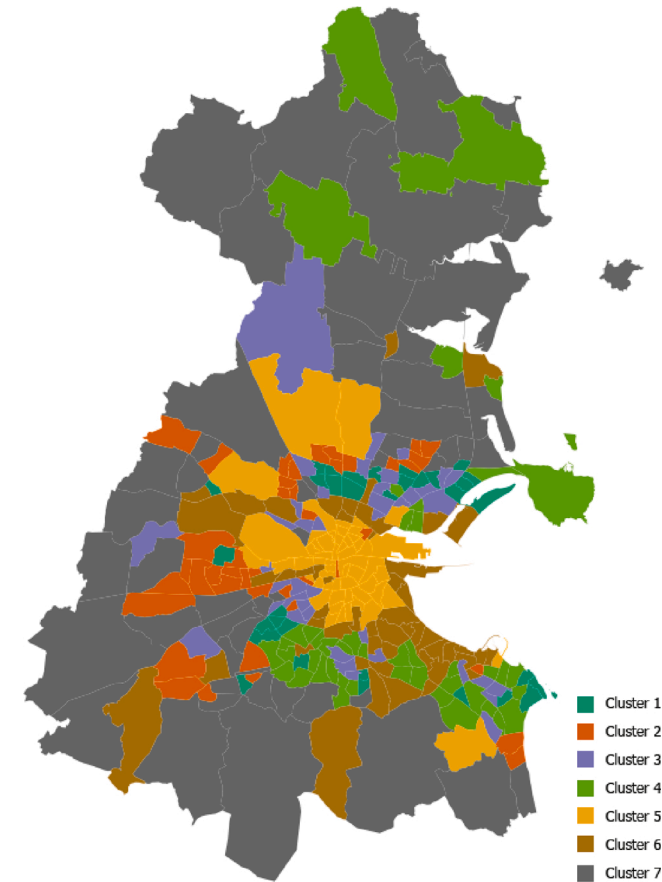| Clusters	Some characteristics of three clusters with high number of cases
Cluster 5	• High percentage of house share• High number of couples with no child• High proportion of aged 25–44
Cluster 6	• High percentage of house share• High proportion of dink family• High employment rate
Cluster 7	• High percentage of house share • High employment rate• High proportion of aged 0–14

Fig. 5. Clustering result of the implemented method for electoral divisions based on the census data, in which 7 clusters are detected; due to the fact that the small areas are dense in the city centre area.

$$\Lambda(N_i, C_p) = \frac{1}{|N|^{\{p\}}} \sum_{N_j \in Cl^{\{p\}}} d(N_i, N_j) \qquad (21)$$

The smallest intra-cluster distance is then calculated, $\beta(i) = \arg\min_{m \neq p} \Lambda(N_i, C_p)$. The Silhouette index ($\check{S}$) for each neuron ($N_i$) at the postsynaptic layer can be defined as

$$\check{S} = \frac{\beta(i) - \alpha(i)}{\max(\alpha(i), \beta(i))} \qquad (22)$$

The mean of the index defined above for a given cluster is then calculated. Silhouette values fall between −1 and 1, and a value close to 1 indicates that the corresponding number of clusters is optimal. Considering the DBI measure, the average distance among clusters should be minimized. Hence, the minimum values for this validity index are considered. According to the results achieved from the validity measures presented in Table 4, we choose seven as the optimal number of clusters. The results achieved in this work show that the algorithm converges

appropriately, and the generated neural network units have been decently grouped into super-clusters. Finally, the results of the clustering method are illustrated in Fig. 5.

We have aggregated the number of confirmed COVID cases in each electoral division given the identified clusters, and the results are demonstrated in Table 5. As shown, the number of confirmed COVID cases in clusters 5, 6, and 7 are higher comparing with others. Given the result of the clustering model and the visualizations in Fig. 5, we can identify different characteristics of each cluster. The detailed features are presented in Table 6. We have found that those clusters with a high number of cases have the lowest proportions of the population with age over 65, high percentage of employment, high percentage of private rent, and high percentage of the population aged 25–44 (young professionals). At the same time, they have the highest proportion of house shares. The boxplots illustrated in Fig. 6 correspond to the cluster characteristics in the seven detected clusters.

6. Conclusions and future work

In this work, we have proposed a multiple-level approach to study the association between geodemographic clustering and the number of confirmed Covid cases in Dublin, Ireland. This work suggests that by incorporating and clustering the publicly available census data, we can obtain valuable insights regarding the spatial variations of people who have contracted the virus. The proposed method includes various phases. As the census data used in this work consists of numerous features, and such characteristics can make a predictive modelling task challenging, a feature selection approach has been implemented based on a non-linear method. Different tests have also been applied to make sure the most relevant features are selected. Then, an advanced geodemographic clustering algorithm was implemented based on a self-organizing feature map to extract clusters given the selected features. The quality of the generated map was analyzed. It should be noted that there is no universal definition of what is good clustering, and this notion is relative. As discussed throughout the paper, an SOM was considered in this work due to the inherent non-linear characteristics of the spatial dataset. Different validity measures were employed to make sure the results of the method used are reliable. We demonstrated that the algorithm has converged properly.

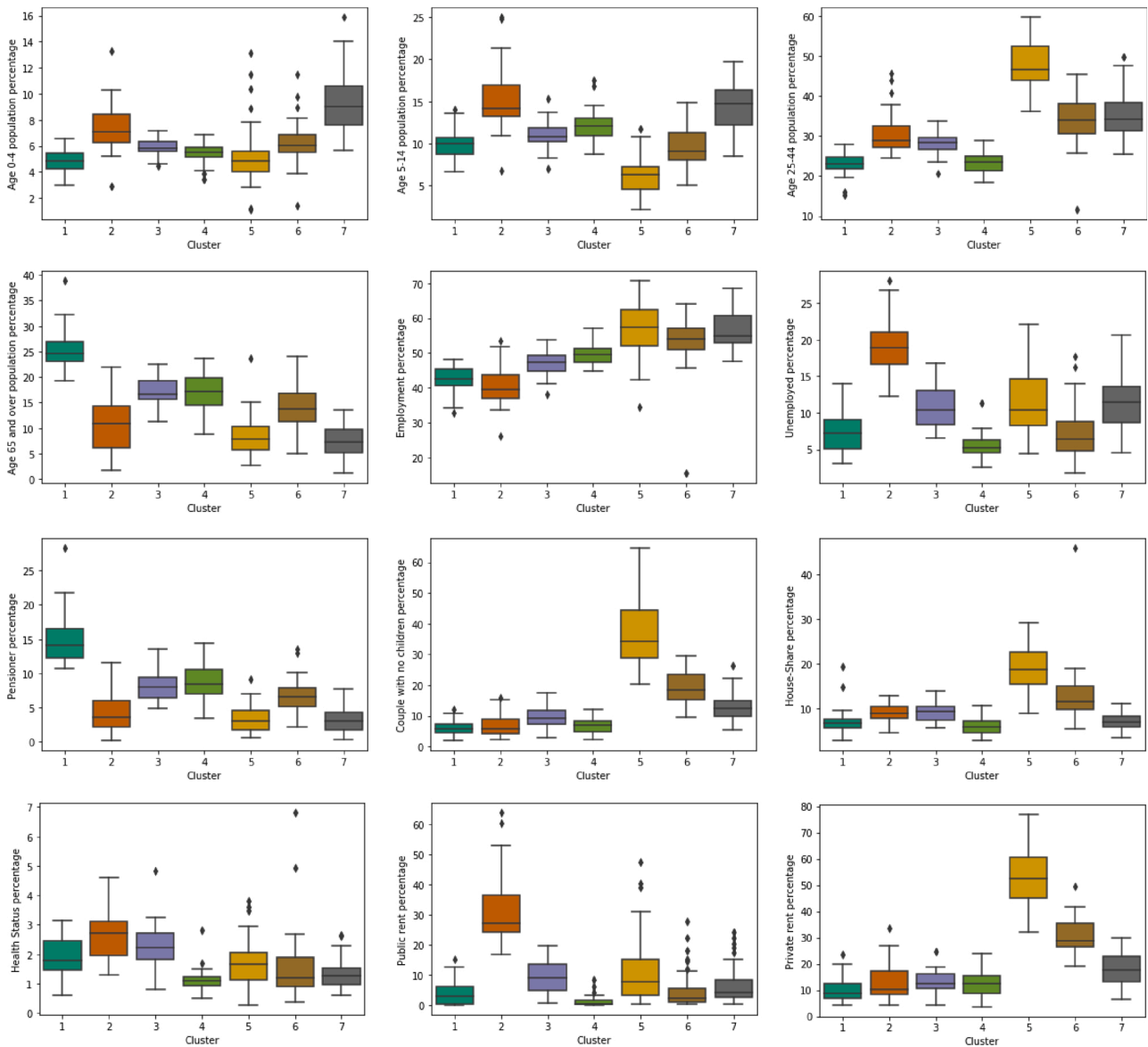According to the analysis, we have detected seven clusters based on

Fig. 6. Boxplots of census data on percentage of different variables given 7 detected clusters.

the census data and the spatial distribution of the people were explored using the unsupervised neural network method. The distribution of people who have contracted the virus was studied. The use of the proposed geodemographic approach incorporating spatial data of a geodemographic nature means that clusters can be interpreted in terms of real-life infected people attributes.

Declaration of Competing Interest

The authors report no declarations of interest.

References

Ahmed, I., Ahmad, M., Rodrigues, J. J., Jeon, G., & Din, S. (2021). A deep learning-based social distance monitoring framework for covid-19. *Sustainable Cities and Society, 65*, 102571. https://doi.org/10.1016/j.scs.2020.102571

Alsaeedy, A. A. R., & Chong, E. K. P. (2020). Detecting regions at risk for spreading covid-19 using existing cellular wireless network functionalities. *IEEE Open Journal of Engineering in Medicine and Biology, 1*, 187–189. https://doi.org/10.1109/OJEMB.2020.3002447

Beria, P., & Lunkar, V. (2021). Presence and mobility of the population during the first wave of covid-19 outbreak and lockdown in italy. *Sustainable Cities and Society, 65*, 102616. https://doi.org/10.1016/j.scs.2020.102616

Bhattacharya, S., Reddy Maddikunta, P. K., Pham, Q. V., Gadekallu, T. R., Krishnan, S., Chowdhary, S. R., et al. (2021). Deep learning and medical image processing for coronavirus (covid-19) pandemic: A survey. *Sustainable Cities and Society, 65*, 102589. https://doi.org/10.1016/j.scs.2020.102589

Chan, S. C., Wu, H. C., & Tsui, K. M. (2012). Robust recursive eigendecomposition and subspace-based algorithms with application to fault detection in wireless sensor networks. *IEEE Transactions on Instrumentation and Measurement, 61*, 1703–1718. https://doi.org/10.1109/TIM.2012.2186654

Chushig-Muzo, D., Soguero-Ruiz, C., Engelbrecht, A. P., De Miguel Bohoyo, P., & Mora-Jiménez, I. (2020). Data-driven visual characterization of patient health-status using electronic health records and self-organizing maps. *IEEE Access, 8*, 137019–137031. https://doi.org/10.1109/ACCESS.2020.3012082

Das, A., Ghosh, S., Das, K., Basu, T., Dutta, I., & Das, M. (2021). Living environment matters: Unravelling the spatial clustering of covid-19 hotspots in Kolkata megacity, India. *Sustainable Cities and Society, 65*, 102577. https://doi.org/10.1016/j.scs.2020.102577

Díaz Ramos, A., López-Rubio, E., & Palomo, E. J. (2020). The forbidden region self-organizing map neural network. *IEEE Transactions on Neural Networks and Learning Systems, 31*, 201–211. https://doi.org/10.1109/TNNLS.2019.2900091

Duffey, R. B., & Zio, E. (2020). Analysing recovery from pandemics by learning theory: The case of covid-19. *IEEE Access, 8*, 110789–110795. https://doi.org/10.1109/ACCESS.2020.3001344

Fan, Q., Yang, G., & Ye, D. (2018). Quantization-based adaptive actor-critic tracking control with tracking error constraints. *IEEE Transactions on Neural Networks and Learning Systems, 29*, 970–980. https://doi.org/10.1109/TNNLS.2017.2651104

Feng, Y., & Li, H. (2020). Dynamic spatial-independent-component-analysis-based abnormality localization for distributed parameter systems. *IEEE Transactions on Industrial Informatics, 16*, 2929–2936. https://doi.org/10.1109/TII.2019.2900226

Ge, X. Y., Pu, Y., Liao, C. H., Huang, W. F., Zeng, Q., Zhou, H., et al. (2020). Evaluation of the exposure risk of sars-cov-2 in different hospital environment. *Sustainable Cities and Society, 61*, 102413. https://doi.org/10.1016/j.scs.2020.102413

Ghahramani, M., Galle, N. J., Duarte, F., Ratti, C., & Pilla, F. (2021). Leveraging artificial intelligence to analyze citizens' opinions on urban green space. *City and Environment Interactions, 10*, 100058. https://doi.org/10.1016/j.cacint.2021.100058

Ghahramani, M., Qiao, Y., Zhou, M. C., O'Hagan, A., & Sweeney, J. (2020a). Ai-based modeling and data-driven evaluation for smart manufacturing processes. *IEEE/CAA Journal of Automatica Sinica, 7*, 1026–1037. https://doi.org/10.1109/JAS.2020.1003114

Ghahramani, M., Zhou, M., & Hon, C. T. (2019a). Extracting significant mobile phone interaction patterns based on community structures. *IEEE Transactions on Intelligent Transportation Systems, 20*, 1031–1041. https://doi.org/10.1109/TITS.2018.2836800

Ghahramani, M., Zhou, M., & Hon, C. T. (2019b). Mobile phone data analysis: A spatial exploration toward hotspot detection. *IEEE Transactions on Automation Science and Engineering, 16*, 351–362. https://doi.org/10.1109/TASE.2018.2795241

Ghahramani, M., Zhou, M., & Wang, G. (2020b). Urban sensing based on mobile phone data: Approaches, applications, and challenges. *IEEE/CAA Journal of Automatica Sinica, 7*, 627–637. https://doi.org/10.1109/JAS.2020.1003120

He, T., Guo, J., Chen, N., Xu, X., Wang, Z., Fu, K., et al. (2020). Medimlp: Using grad-cam to extract crucial variables for lung cancer postoperative complication prediction. *IEEE Journal of Biomedical and Health Informatics, 24*, 1762–1771. https://doi.org/10.1109/JBHI.2019.2949601

Hu, S., Gao, Y., Niu, Z., Jiang, Y., Li, L., Xiao, X., et al. (2020a). Weakly supervised deep learning for covid-19 infection detection and classification from ct images. *IEEE Access, 8*, 118869–118883. https://doi.org/10.1109/ACCESS.2020.3005510

Hu, S., O'Hagan, A., Sweeney, J., & Ghahramani, M. (2020b). A spatial machine learning model for analysing customers' lapse behaviour in life insurance. *Annals of Actuarial Science, 10*, 1–27. https://doi.org/10.1017/S1748499520000329

Kim, C., & Klabjan, D. (2020). A simple and fast algorithm for l1-norm kernel pca. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 42*, 1842–1855. https://doi.org/10.1109/TPAMI.2019.2903505

Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., et al. (2020). Artificial intelligence distinguishes covid-19 from community acquired pneumonia on chest ct. *Radiology, 19*.

Martins, D. M. L., & de Lima Neto, F. B. (2020). Hybrid intelligent decision support using a semiotic case-based reasoning and self-organizing maps. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 50*, 863–870. https://doi.org/10.1109/TSMC.2017.2749281

Montes-Orozco, E., Mora-Gutiérrez, R., De-Los-Cobos-Silva, S., Rincón-García, E., Torres-Cockrell, G., Juárez-Gómez, J., et al. (2020). Identification of covid-19 spreaders using multiplex networks approach. *IEEE Access, 8*, 122874–122883. https://doi.org/10.1109/ACCESS.2020.3007726

Nagrath, P., Jain, R., Madan, A., Arora, R., Kataria, P., & Hemanth, J. (2021). Ssdmnv2: A real time dnn-based face mask detection system using single shot multibox detector and mobilenetv2. *Sustainable Cities and Society, 66*, 102692. https://doi.org/10.1016/j.scs.2020.102692

Ouyang, X., Huo, J., Xia, L., Shan, F., Liu, J., Mo, Z., et al. (2020). Dual-sampling attention network for diagnosis of covid-19 from community acquired pneumonia. *IEEE Transactions on Medical Imaging, 39*, 2595–2605. https://doi.org/10.1109/TMI.2020.2995508

Rahman, M. A., Zaman, N., Asyhari, A. T., Al-Turjman, F., Alam Bhuiyan, M. Z., & Zolkipli, M. (2020). Data-driven dynamic clustering framework for mitigating the adverse economic impact of covid-19 lockdown practices. *Sustainable Cities and Society, 62*, 102372. https://doi.org/10.1016/j.scs.2020.102372

Rumpler, R., Venkataraman, S., & Göransson, P. (2020). An observation of the impact of covid-19 recommendation measures monitored through urban noise levels in central Stockholm, Sweden. *Sustainable Cities and Society, 63*, 102469. https://doi.org/10.1016/j.scs.2020.102469

Rustam, F., Reshi, A. A., Mehmood, A., Ullah, S., On, B., Aslam, W., et al. (2020). Covid-19 future forecasting using supervised machine learning models. *IEEE Access, 8*, 101489–101499. https://doi.org/10.1109/ACCESS.2020.2997311

Saeed, N., Nam, H., Al-Naffouri, T. Y., & Alouini, M. (2019). A state-of-the-art survey on multidimensional scaling-based localization techniques. *IEEE Communications Surveys Tutorials, 21*, 3565–3583. https://doi.org/10.1109/COMST.2019.2921972

Sannigrahi, S., Pilla, F., Basu, B., Basu, A. S., & Molter, A. (2020). Examining the association between socio-demographic composition and covid-19 fatalities in the European region using spatial regression approach. *Sustainable Cities and Society, 62*, 102418. https://doi.org/10.1016/j.scs.2020.102418

Sharifi, A., & Khavarian-Garmsir, A. R. (2020). The covid-19 pandemic: Impacts on cities and major lessons for urban planning, design, and management. *Science of The Total Environment*, 142391. https://doi.org/10.1016/j.scitotenv.2020.142391

Shi, X., Yang, H., Xu, Z., Zhang, X., & Farahani, M. R. (2019). An independent component analysis classification for complex power quality disturbances with sparse auto encoder features. *IEEE Access, 7*, 20961–20966. https://doi.org/10.1109/ACCESS.2019.2898211

Shokouhyar, S., Shokoohyar, S., Sobhani, A., & Gorizi, A. J. (2021). Shared mobility in post-covid era: New challenges and opportunities. *Sustainable Cities and Society, 67*, 102714. https://doi.org/10.1016/j.scs.2021.102714

Silva, J. C. S., de Lima Silva, D. F., de Sá Delgado Neto, A., Ferraz, A., Melo, J. L., Ferreira Júnior, N. R., et al. (2021). A city cluster risk-based approach for sars-cov-2 and isolation barriers based on anonymized mobile phone users' location data. *Sustainable Cities and Society, 65*, 102574. https://doi.org/10.1016/j.scs.2020.102574

So, M., Tiwari, A., Chu, A., Tsang, J., & Chan, J. (2020). Visualizing covid-19 pandemic risk through network connectedness. *International Journal of Infectious Diseases, 96*, 558–561.

Sun, C., & Zhai, Z. (2020). The efficacy of social distance and ventilation effectiveness in preventing covid-19 transmission. *Sustainable Cities and Society, 62*, 102390. https://doi.org/10.1016/j.scs.2020.102390

Sun, G., Zhang, S., Zhang, Y., Xu, K., Zhang, Q., Zhao, T., & Zheng, X. (2019). Effective dimensionality reduction for visualizing neural dynamics by Laplacian eigenmaps. *Neural Computation, 31*, 1356–1379.

Sun, K. H., Huh, H., Tama, B. A., Lee, S. Y., Jung, J. H., & Lee, S. (2020). Vision-based fault diagnostics using explainable deep learning with class activation maps. *IEEE Access, 8*, 129169–129179. https://doi.org/10.1109/ACCESS.2020.3009852

Wade, L., & Khavarian-Garmsir, A. R. (2020). An unequal blow. *Science*, 700–770.

Wickramasinghe, C. S., Amarasinghe, K., & Manic, M. (2019). Deep self-organizing maps for unsupervised image classification. *IEEE Transactions on Industrial Informatics, 15*, 5837–5845. https://doi.org/10.1109/TII.2019.2906083

Xiang, S., Nie, F., Zhang, C., & Zhang, C. (2009). Nonlinear dimensionality reduction with local spline embedding. *IEEE Transactions on Knowledge and Data Engineering, 21*, 1285–1298. https://doi.org/10.1109/TKDE.2008.204

Xie, K., Chen, C., Lewis, F. L., & Xie, S. (2018). Adaptive asymptotic neural network control of nonlinear systems with unknown actuator quantization. *IEEE Transactions on Neural Networks and Learning Systems, 29*, 6303–6312. https://doi.org/10.1109/TNNLS.2018.2828315

Xie, X., Zhong, Z., Zhao, W., Zheng, C., Wang, F., & Liu, J. (2020). Chest ct for typical 2019-ncov pneumonia: Relationship to negative rt-pcr testing. *Radiology, 12*.

Yu, H., Lu, J., & Zhang, G. (2020). Online topology learning by a Gaussian membership-based self-organizing incremental neural network. *IEEE Transactions on Neural Networks and Learning Systems, 31*, 3947–3961. https://doi.org/10.1109/TNNLS.2019.2947658

Zivkovic, M., Bacanin, N., Venkatachalam, K., Nayyar, A., Djordjevic, A., Strumberger, I., et al. (2021). Covid-19 cases prediction by using hybrid machine learning and beetle antennae search approach. *Sustainable Cities and Society, 66*, 102669. https://doi.org/10.1016/j.scs.2020.102669