P.PORTO

ESCOLA
SUPERIOR
DE SAÚDE
POLITÉCNICO
DO PORTO

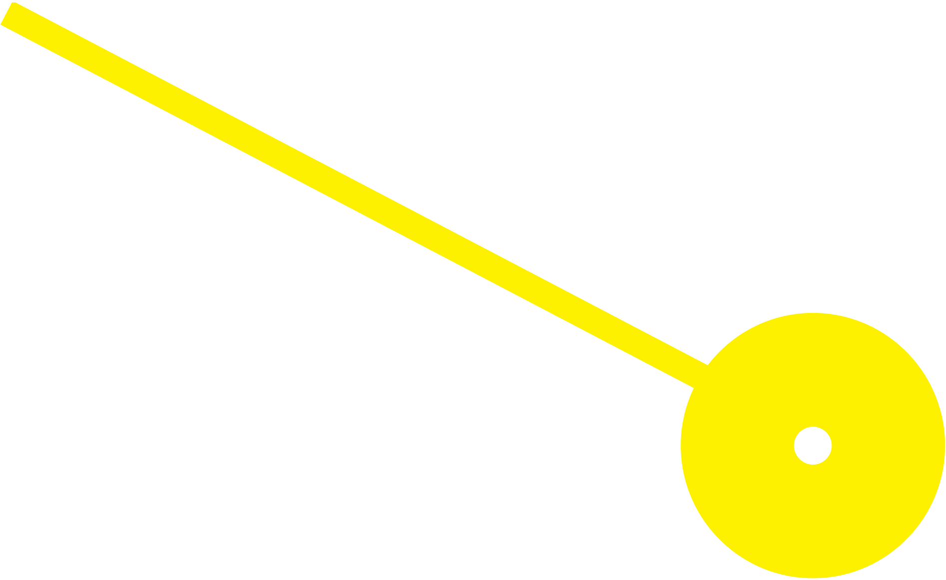
M

MESTRADO
EM BIOESTATÍSTICA E BIOINFORMÁTICA APLICADAS À SAÚDE

Comparing Time Series Forecasting Models for Health Indicators: A Clustering Analysis Approach

Cláudia Vinhal

09/**2024**



Comparing Time Series Forecasting Models for Health Indicators: A Clustering Analysis Approach

Autor
Cláudia Vinhal
Orientadores
Professora Doutora Alexandra Oliveira, Professora Adjunta Convidada da Escola Superior de Saúde do Politécnico do Porto e LIACC-Lab. Inteligência Artificial e Ciência de Computadores da Universidade do Porto
Professora Doutora Brígida Mónica Faria, Professora Coordenadora da Escola Superior de Saúde do Instituto do Porto, e LIACC-Lab. Inteligência Artificial e Ciência de Computadores da Universidade do Porto
Professor Doutor Rui Pimenta, Professor Coordenador Principal da Escola Superior de Saúde do Instituto do Porto e Centro de Estudos e Investigação em Saúde da Universidade de Coimbra

Dissertação apresentada para cumprimento dos requisitos necessários à obtenção do grau de Mestre em **Bioestatística e Bioinformática Aplicadas à Saúde** pela Escola Superior de Saúde do Instituto Politécnico do Porto.

Acknowledgments

I would like to express my gratitude to my supervisors, Alexandra Oliveira, Brígida Faria, and Rui Pimenta, for their guidance, support, and patience throughout the course of this dissertation. Their expertise and mentorship have been fundamental in my academic growth. I want to give a special thanks to Ana Paula Nascimento for her expertise and insightful feedback. Thank you for the encouragement and support that were essential during this period.

To my friends, whose shared knowledge and engaging discussions gave needed inspiration and motivation. Your support and encouragement were vital for this project.

Lastly, to my family, my deepest appreciation for their constant and unwavering support and patience. Their belief in me has been a constant source of motivation.

Resumo

As séries temporais podem ser definidas como a sequência de observações ordenadas por intervalos de tempo iguais, sendo assim fundamentais para abordar questões de causalidade, tendências e previsão. Os dados temporais e a sua análise podem ser aplicados a diversas áreas, como a engenharia, as finanças e a saúde. Com o estudo contínuo de séries temporais surgem vários problemas, um dos quais ao nível da análise classificatória, que tem como objetivo identificar semelhanças entre as séries. Este aspeto é particularmente relevante quando as séries temporais são modeladas por Modelos Autorregressivos Integrados de Médias Móveis (ARIMA), o que torna a compreensão dos seus parâmetros essencial para a sua análise.

Uma das principais aplicações das séries temporais em saúde pública e biomedicina tem sido em estudos epidemiológicos de doenças infeciosas e crónicas, estudos sobre a previsão da procura dos serviços de saúde e estudos sobre a avaliação dos resultados de saúde através de dados sobre a mortalidade e da morbilidade. Estes indicadores são considerados medidas diretas das necessidades em cuidados de saúde, refletindo a carga global da doença na população, sendo então indispensáveis para o estudo e vigilância da saúde pública e para os processos de organização e intervenção dos serviços de saúde. A soma da mortalidade e morbidade é referida como "Carga da Doença" e pode ser medida por uma métrica denominada "Anos de Vida Ajustados por Incapacidade" (DALYs). A análise destes dados é fundamental para identificar padrões geográficos e demográficos, o que permite uma melhor perceção das disparidades na população a nível da saúde.

Os principais objetivos desta dissertação são modelar indicadores de saúde através de processos Médias Móveis (MA), Autorregressivos de Médias Móveis (ARMA) ou Autorregressivos Integrados de Médias Móveis; avaliar a qualidade de ajuste dos modelos aos dados; e comparar as distâncias entre processos quanto à sua eficácia na identificação de grupos naturais.

O estudo começa por explorar as características temporais dos DALYs de cinco doenças não transmissíveis (doenças cardiovasculares, doenças respiratórias crónicas, doenças neurológicas, doenças renais crónicas e diabetes), destacando padrões e tendências subjacentes. De seguida, utilizando um algoritmo automatizado, são aplicados Modelos Autorregressivos Integrados de Médias Móveis para representar e descrever as séries temporais. O ajuste do modelo foi avaliado com métricas de precisão da previsão, como o Erro Absoluto Médio (MAE), Raiz do Erro Quadrático Médio (RMSE), Erro Quadrático Médio (MSE) e Erro Percentual Médio Absoluto (MAPE). É nesta representação de séries temporais que foram aplicadas as medidas de distância Piccolo, Maharaj e LPC para utilizar técnicas de agrupamento (*clustering*) e identificar grupos com perfis semelhantes (*clusters*). Foram utilizados seis métodos diferentes de agrupamento hierárquico: Ward, Complete, Average, Single, Median e Centroid. Além disso, o desempenho do algoritmo de agrupamento foi avaliado através de métricas de avaliação, como o *Silhouette score, CIndex, McClain e Dunn index.*

Os resultados sobre dados de DALYs de doenças não transmissíveis específicos para 48 países europeus, mostram que a escolha da medida de distância influencia os resultados do *clustering* e o número de *clusters* formados. Embora certos métodos revelem padrões geográficos, outros fatores, como as semelhanças culturais ou económicas, podem também influenciar a formação de *clusters*. Além disso, alguns países foram frequentemente isolados no seu próprio cluster em diferentes de métodos de agrupamento e medidas de distância, sugerindo que o seu Modelo Autorregressivo Integrado de Médias Móveis era diferente dos restantes. Um exemplo é a Letónia, que formou *clusters* isolados nas doenças cardiovasculares. Países como a Albânia, Bielorrússia, Lituânia e Suécia, foram agrupados no mesmo *cluster* para vários métodos de *clustering* quando a distância de Piccolo foi aplicada a doenças neurológicas. Para as doenças respiratórias crónicas, são formados 15 *clusters* com a distância LPC, entre 8 e 15 *clusters* com a distância de Piccolo e entre 9 e 15 *clusters* com a distância Maharaj. Estes resultados não só contribuem para o avanço no campo da vigilância e intervenção em saúde pública, com o objetivo final de aliviar o cargo global da doença, mas também contribuem para a nossa compreensão sobre o *clustering* dos Modelos Autorregressivos Integrados de Médias Móveis e como uso de diferentes medidas de distância influenciam os resultados dos *clusters*.

Palavras-chave: Medidas de Distância, Piccolo, Maharaj, LPC, Clustering, DALYs, modelos ARIMA

Abstract

Time series can be defined as the sequence of observations ordered by equal time intervals, thus being fundamental to address questions of causality, trends, and forecast. Temporal data and its analysis can be applied to several areas, such as engineering, finance, and health. With the constant study of time series, several problems arise, one of which is at the level of clustering, which aims to identify similarities between the series. This aspect is particularly relevant when time series are modeled by Autoregressive Integrated Moving Average (ARIMA) models, which makes understanding their parameters essential for their analysis.

One of the main applications of time series in public health and biomedicine has been in epidemiological studies of infectious and chronic diseases, studies on the prediction of demand for health services, and studies on the assessment of health outcomes through data on mortality and morbidity. These indicators are direct measures of health care needs, reflecting the global burden of disease in the population, and are therefore crucial for the study and surveillance of public health, and for the processes of organization and intervention of health services. The sum of mortality and morbidity is referred to as "Burden of Disease" and can be measured by a metric called "Disability Adjusted Life Year" (DALYs). The analysis of this type of data is essential to identify geographic patterns, which allows a better perception of health disparities in the population.

The main objectives for this dissertation are to model health indicators through Moving Average (MA), Autoregressive Moving Average (ARMA) or Autoregressive Integrated Moving Average processes; evaluate the quality of fit of the models to the data; and compare the distances between processes regarding their effectiveness in identifying natural groups.

The study begins by exploring the temporal characteristics of DALYs of five non-communicable diseases (cardiovascular diseases, chronic respiratory diseases, neurological disorders, chronic kidney diseases, and diabetes), highlighting underlying patterns and trends. Then, using an automated algorithm, Autoregressive Integrated Moving Average models are applied to represent and describe the time series. The fit of the model was assessed with forecast accuracy metrics, such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE). It is on this representation of time series that the Piccolo, the Maharaj, and the LPC distance measures were applied to use clustering techniques and identify clusters. Six different hierarchical clustering methods were used, the Ward, the Complete, the Average, the Single, the Median, and the Centroid linkage. Additionally, the performance of the clustering algorithm was weighed through evaluation metrics, such as the Silhouette score, CIndex, McClain Index, and Dunn Index.

The results on non-communicable diseases DALYs data specific to 48 European countries, show that the choice of distance measure greatly influences clustering outcomes, and the number of clusters formed. While certain methods revealed geographic patterns, other factors, such as, cultural or economic similarities can also influence cluster formation. Furthermore, some countries were frequently isolated in their own cluster across clustering methods and distance measures, suggesting that their Autoregressive Integrated Moving Average model was significantly different from the rest. For example, Latvia, which formed isolated clusters in cardiovascular diseases. Other countries, such as Albania, Belarus, Lithuania, and Sweden were grouped into the same cluster across various clustering methods when the Piccolo distance was applied to neurological disorders. For chronic respiratory diseases, 15 clusters were formed with the LPC distance, between 8 and 15 clusters with the Piccolo distance, and between 9 and 15 clusters with the Mahara distance. These insights, not only contribute to advancing the field of public health surveillance and intervention, ultimately aiming to alleviate the global burden if disease, but also contribute to our understanding of clustering Autoregressive Integrated Moving Average models and how the use of different distance measures influences clusters outcomes.

Keywords: Distance Measures, Piccolo, LPC, Maharaj, Clustering, DALYs, ARIMA models

Abbreviations and Acronyms

DALYs – Disability Adjusted Life Years
NCD – Non-communicable Diseases
ARIMA – Autoregressive Integrated Moving Average
MA – Moving Average
ARMA - Autoregressive Moving Average
GBD – Global Burden of Disease
YLL – Years of Life Lost from Premature Death
YLD – Years Lived with Disability
IHME – institute for Health Metrics and Evaluation
WHO – World Health Organization
EU – European Union
ADF – Augmented Dickey-Fuller
KPSS – Kwiatkowski-Phillips-Schmidt-Shin
ACF – Autocorrelation Function
PACF – Partial Autocorrelation Function
MLE – Maximum Likelihood Estimation
AIC – Akaike Information Criterion
BIC – Bayesian Information Criterion
MAE – Mean Absolute Error
MSE - Mean Squared Error
RMSE – Root Mean Squared Error
MAPE – Mean Absolute Percentage Error
AR - Autoregressive

Table of Contents

Figures

Tables

1. Introduction

Time series analysis holds a significant role in public health and biomedicine. One of the main applications of time series in these fields has been in epidemiological studies of infectious and chronic diseases, studies on the prediction of demand for health services and studies on the assessment of health outcomes through data on mortality and morbidity (1). These indicators are considered direct measures of healthcare needs, reflecting the overall burden of disease within populations, and are therefore essential for the study and surveillance of public health, organizational processes, and the intervention of health services (2). The sum of mortality and morbidity metrics is referred to as the “Burden of Disease” and can be quantified by the metric “Disability Adjusted Life Years” (DALYs) (3). The analysis of this data is fundamental to identifying geographic and demographic patterns, which offers a better perception of health disparities within populations.

Non-communicable diseases (NCDs) are responsible for a substantial proportion of global morbidity and mortality. The burden of NCDs, as measured by DALYs, represents both premature mortality and disability attributable to these conditions. Regardless of the efforts made to address the burden of NCDs through various interventions and policies, the prevalence and impact of these diseases continues to rise (4). Factors such as population ageing, urbanization, unhealthy lifestyles, and ineffective healthcare infrastructure present significant challenges to the prevention and control of NCDs, highlighting the need to make them a priority area for public health action (5).

Temporal data of health indicators show unique time-dependent patterns that can be represented by Autoregressive Integrated Moving Average (ARIMA) models (6). When it comes to clustering these models, the choice of distance measure is crucial since standard measures can fail to capture effectively the time patterns (7). Some specialized distance measures take into account the underlying ARIMA model parameters. This ensures to reflect both the temporal structure and the statistical properties of the time series, which is essential to accurately show the dynamics of DALYs over time. This way the choice of the appropriate distance measure is a critical step that influences the outcomes of time series clustering and the subsequent public health analysis.

1.1. Objectives

The main objectives for this dissertation are to model health indicators through MA (Moving Average), ARMA (Autoregressive Moving Average) or ARIMA processes; evaluate the quality of fit of the models to the data; and compare the distances between processes regarding their effectiveness in identifying natural groups. Furthermore, it is expected to obtain homogeneous and separate clusters and carry out a comparative analysis of different classification methods using the various distances, such as the Piccolo distance (8), the Maharaj distance (9,10), and the LPC distance (11), applied to ARIMA models.

1.2. Document Structure

To a better understanding of the research topic and to frame the importance of analyzing time series data related to health indicators, Chapter 2 focuses on the global burden of disease with particular emphasis on non-communicable diseases. Chapter 3 presents a comprehensive review of the theoretical background and existing literature on the topic. It begins with an introduction to the concept of time series data, followed by an explanation of ARIMA models and their importance in representing time series data for analysis. The chapter then provides an overview of clustering techniques, including distance measures and methods to evaluate clustering effectiveness. Finally, it concludes with a review of the state of the art in clustering ARIMA models. Chapter 4 outlines the methodological framework, detailing the approach used for data analysis. The chapter begins with an overview of the data sources and the selection criteria applied. This is followed by an explanation of how the data is pre-processed. It then presents how the time series were modelled and how their accuracy and effectiveness are evaluated. Finally, the chapter concludes with a description of the specific clustering techniques and evaluation metrics that were used to group the ARIMA models. Chapter 5 presents the findings from the data analysis, beginning with an exploratory data analysis. This is followed by a detailed representation of the time series for each disease. The chapter concludes by presenting the clustering results using different distance measures for each disease. In Chapter 6 the results are discussed, exploring the implications of the findings for future research and public health policies. Lastly, Chapter 7 summarizes this dissertation key findings and contributions, and discusses potential directions for future work.

2. Burden of Disease

The burden of disease is a fundamental concept in public health, summarizing the overall impact that diseases, injuries, and health conditions exert on individuals and populations (12). Studying the burden of disease allows a better understanding of the scale and severity of health challenges faced by communities. Evaluating factors such as incidence, prevalence, mortality, and morbidity of various health conditions, gives insights into the complex relations of genetic, social, and environmental factors that impact the public health sector (12,13). This knowledge is crucial to inform healthcare policies, for better resource allocation, and intervention strategies to improve public health outcomes. The burden of a disease is tied to a multitude of risk factors, that can be modifiable and non-modifiable, meaning they have changeable or non-changeable conditions, respectively (5). Modifiable risk factors, such as lifestyle choices, environmental exposures, and healthcare access, play a significant role in shaping the burden of disease. Understanding and addressing these factors are key to preventive strategies. Non-modifiable risk factors, including genetic predispositions and age, also contribute to the burden of disease but require different approaches in terms of management and mitigation (5). By identifying and mitigating risk factors, public health efforts can effectively reduce the overall burden of diseases on communities.

The Global Burden of Disease (GBD) study began 30 years ago with the goal of providing timely, valid, and relevant assessments of critical health outcomes (14). The GBD provides annual estimates from 1990 to the present for 371 diseases and injuries, as well as over 3400 clinical outcomes, for 204 countries and territories and for subnational units in more than 20 countries. This scientific effort allows comparisons over time, across populations and between health problems (14). One of the study's strengths is the common health metrics for all diseases. The GBD primary health metric is the disability-adjusted life year (DALY) (15). This metric was introduced as a single measure to quantify the burden of diseases, injuries, and risk factors (2). The DALY is a composite measure that adds years of life lost from premature death (YLL), and the years lived with disability (YLD), providing a comprehensive measure of the overall health loss attributed to a specific condition. DALYs serve as a valuable bridge between mortality and morbidity data and offer a more general perspective on the real burden a disease imposes on a population (13). Currently, the Institute for Health Metrics and Evaluation (IHME), is the coordinating center for the GBD study and presents an online data catalogue for all data sources used in the GBD (14).

The burden of disease extends its influence beyond individual health to the broader healthcare system. High disease burden places strains on healthcare infrastructure, needing effective resource management and healthcare delivery strategies. The common consequences are the increasing demand for medical services, diagnostics, and treatment modalities (13). Additionally, the economic burden associated with healthcare expenses, both at individual and community levels, emphasizes the urgency of effective disease prevention and management (4). Consequently, understanding the burden of disease is indispensable for healthcare planning, policy formulation, and the establishment of resilient and responsive healthcare systems.

While the GBD study represents a valuable effort to quantify the magnitude of all major diseases and injuries, it does have certain limitations, such as the data gaps for important health outcomes (12). In some countries, there is an incomplete vital registration with medical certification, and/or have not had a census in a long time, which makes the population number uncertain. Furthermore, the large array of different data sources, combined with missing data, makes the GBD analysis for a single outcome complex (14).

In conclusion, over the last three decades, the Global Burden of Disease collaboration has had significant growth, meeting a pivotal need in global health (14). As demands from decision-makers for timely information increase, the GBD's role in providing such information and translating it into policy-relevant scenarios has proven to be crucial. With the use of metrics like DALYs, exploration of risk factors, and assessment of healthcare system impacts, it is possible to develop a nuanced understanding of health challenges, paving the way for informed and effective interventions to improve population health and well-being (4).

2.1. The burden of non-communicable diseases

Non-communicable diseases (NCDs) are a group of medical conditions that are characterized by their chronic nature. These persist over an extended period of time and often have a slow progress. Unlike communicable diseases, NCDs are not caused by infectious agents, such as bacteria or viruses, and cannot be transmitted from person to person. They can be caused by a combination of non-infectious factors, which include genetic predisposition, lifestyle choices, environmental exposures, and metabolic abnormalities (16). Examples of medical conditions that have their place in the NCDs group include asthma, cancer, diabetes, Alzheimer's disease, and osteoarthritis. These conditions notably impact people's health and quality of life, and often require long-term management and care. According to the World Health Organization (WHO), NCDs are responsible for 74% of the total number of deaths globally each year (41 million people) (17). Annually, 17 million people die before the age of 70 from NCDs, of which 86% occur in low- and middle-income countries. The top four killers among NCDs with the highest number of deaths are cardiovascular diseases (17.9 million deaths each year), cancers (9.3 million), chronic respiratory diseases (4.1 million), and diabetes (2.0 million including kidney disease deaths caused by diabetes). These four groups are responsible for over 80% of all premature deaths from NCDs (17). In the past, some non-communicable diseases were associated with economic growth and were referred to as "diseases of the rich". However, there has been a shift with the burden of NCDs in developing countries increasing. Moreover, the mortality rates of these diseases in low and middle-income countries have doubled the burden of NCDs (16). Recently, there has been a growing interest in population well-being and economic development based on Gross National Happiness. This approach to social progress considers factors outside of just economic indicators. Nevertheless, the epidemic of NCDs holds back the improvement of GNH, because good health is a fundamental requirement to achieve happiness and overall well-being (16).

The GBD study has documented that shifts to more non-communicable diseases deaths and a larger proportion of deaths attributable to NCD are due to an ageing population and to sharper declines in deaths due to communicable, maternal, neonatal, and nutritional causes (15). As indicated before, the majority of preventable NCD burden falls on low and middle-income countries, simultaneously, NCD are progressively more diseases of low-income individuals living in high-income countries, who are more often exposed to risk factors, such as, unhealthful diets and alcohol consumption, tobacco smoke, indoor and ambient air pollution, and environments that discourage physical activity (15). Non-communicable diseases impose a significantly high burden on the healthcare system, both financially and structurally, of developed and developing countries, due to their long duration and slow progression (4,15). The four major NCDs in the European Union (EU) are cardiovascular diseases, cancer, diabetes mellitus and chronic respiratory disease. These diseases claim a significant share of the total healthcare budget - at least 25% of health spending in the EU (4). Additionally, NCDs represent a challenge for national economic development, resulting in production and productivity losses due to disability in daily work and life for patients, informal care offered by family and more (18).

In addressing the rising burden of Non-Communicable Diseases, effective management of risk factors plays a pivotal role. The most common causes of non-communicable diseases are metabolic and behavioral risk factors, such as, tobacco and alcohol consumption, physical inactivity, poor diet, and lack of dental care (16). These risk factors can be generally preventable through individual self-management and health education efforts led by healthcare specialists. Public health sector interventions, including partnerships between different sectors, are crucial for NCD prevention strategies. It is worth mentioning the environmental factors (air pollution, climate changes, sunlight) and their impact on the development of NCDs, with emerging technologies like microchips expected to play a role in air quality monitoring (16).

The prevention strategies for NCDs require collaboration at multiple levels, from individuals making healthier lifestyle choices to global policy initiatives. At the individual level, lifestyle modifications and improvements play a fundamental role in raising awareness and managing risk factors (16). Addressing the NCD burden will entail countries training or retraining healthcare workers to prevent and treat these diseases; educating the community about the causes, signs, and symptoms of NCD; working across government agencies and industry to minimize nutritional and environmental risks; and providing infrastructure for screening of pre-clinical and treatment of acute and chronic phase NCD (15).

Furthermore, governments must design strategic plans based on their economies, implementing low-priced yet effective policies such as promoting physical activity and allocating resources to primary healthcare, meeting the challenge of the growing burden of NCD need not place overwhelming demands on healthcare spending (15). The collaboration between governments, schools, and research institutions is essential to provide education on lifestyle modifications and early detection methods. Additionally, research focusing on food biotechnology, agriculture, and medical tests, also contributes to NCD prevention and management. Ultimately, encouraging healthy lifestyles at the individual and family levels is pivotal in achieving optimal health outcomes and effectively addressing the burden of NCDs (16).

To conclude, the burden of non-communicable diseases presents a complex challenge that requires efforts at individual, societal, and global levels. NCDs not only impact individual health but also pose significant economic and social implications, affecting productivity and straining healthcare systems. However, by prioritizing prevention strategies, promoting healthier lifestyles, and encouraging collaboration across various sectors, it is possible to lessen the burden of NCDs and enhance the overall quality of life, and pave the way for a healthier future for generations to come.

3. Background and State of the Art

This chapter explores the foundational principles of time series analysis, focusing on their representation through ARIMA models, and dives into various clustering methodologies. It further discusses evaluation metrics and distance measures crucial for time series clustering. It provides an overview of their evolution and establishes a robust framework for their practical application in diverse domains.

3.1. Time Series

Time series can be defined as the sequence of observations ordered by equal time intervals, usually days, weeks, months, quarters, or years (19). If a sequence of values is created through a mathematical function, it is called deterministic. On the contrary, when a random value is introduced into the series using a mathematical time function, the series becomes stochastic or non-deterministic (20). Time series are therefore fundamental for addressing questions of causality, trends, and prediction. Temporal data and its analysis can be applied to fields as diverse as engineering, finance, economics, sociology, climatology, energy, and public health. Examples in public health include weekly admissions to the emergency department, annual expenditures, epidemiological studies of infectious and chronic diseases, and studies on the assessment of health outcomes through mortality and morbidity data (19). Time series analysis is crucial to extract important information from temporal data, offering a multitude of benefits across several areas, such as pattern recognition, forecasting, risk management, and resource planning (20). This allows organizations across different industries to make informed decisions by unveiling patterns and trends that are not straightaway apparent in raw data. The primary objective of descriptive analysis is to generate visual representations and summary statistics that contribute to a better understanding of the variability in responses over time (19). To do that, it is essential to understand the four time series components (21) (Figure 1):

- Trend – the long-term movement in time series. Trends can be upward, downward, or flat.
- Seasonality – Regular periodic occurrences within regular intervals (daily, monthly, or quarterly). It often corresponds to external factors.
- Cyclic behavior – Repeated oscillations around the trend that do not follow a strict time frame.
- Random noise – Short-term unpredictable and irregular variations that cannot be credited to trend, seasonality, or cyclic patterns.

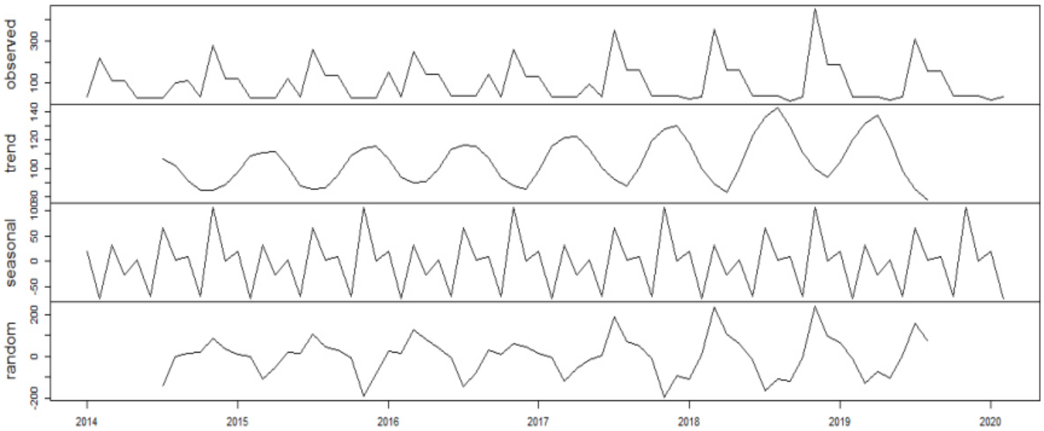


Figure 1 Time series components. Adapted from (21)

Methods and techniques for time series analysis can be generally classified into parametric and non-parametric methods. Both methods are viable choices, but it is crucial to know beforehand which one is best suited for a specific time series problem (22). Non-parametric approaches explicitly assess the spectrum or covariance of the stochastic process without assuming a specific structure for the stationary process and do not require a predefined number of parameters to predict future values of time series data. Even though non-parametric models are computationally slower when compared to parametric models, they make fewer assumptions about the time series data. Over the last decade, non-parametric methods have gained importance, for their ability to learn through trial and error and improve accuracy over iterations. Non-parametric approaches include models such as neural networks, support vector regression, regression trees, Gaussian process, long short-term memory, and others, collectively known as machine learning algorithms. These techniques have been applied to time series forecasting (22). On the other hand, parametric models assume that the underlying stochastic stationary process has a specific structural formation that can be described using a limited number of parameters. The main task of this approach is to estimate these parameters of the time series model that describes the stochastic process. Furthermore, when using parametric models, only parameters are required from the historic data to forecast future values of the time series (22). A few common parametric methods are exponential smoothing, Autoregressive Moving Averages (ARMA) and Autoregressive Integrated Moving Averages (ARIMA).

The Exponential Smoothing was originated with the work of Holt, Brown, and Winters in the late 1950s (23). This method includes a selection of models such as Simple Exponential Smoothing, Double Exponential Smoothing, and Triple Exponential Smoothing. This method assigns exponentially decreasing weights to past observations in a time series. The more recent observations have higher associated weight. This approach is valuable for short-term forecasting, particularly in situations with trend and seasonality (23).

3.2. Representation of time series through ARIMA models

Like other traditional statistical time series models, the ARIMA model relies on a time series being stationary. A stationary series, also known as a "white noise series", has three key properties: a constant mean, constant variance, and constant covariance, which depends solely on the time interval between values (24). In other words, a stationary time series means that its statistical characteristics do not change over time. This series can be modelled with fewer parameters and are, therefore, easier to analyze. Stationarity is a key assumption in time series analysis, facilitating statistical inferences and prediction (25). Non-stationarity in time series data typically stems from two main sources. Firstly, changing variance over time, known as heteroscedasticity, can often be diminished with the use of a log transformation. Secondly, an increasing or decreasing trend can usually be addressed by the differencing technique, which is the process of taking the differences between consecutive observations ($Y_t - Y_{t-1}$). Differencing helps stabilize the mean of a time series by removing changes in its level, which in turn eliminates trend and seasonality (24). One way to determine whether differencing is required is to use a unit root test. These tests are statistical assessments of stationarity specifically designed to determine whether differencing is warranted. Various unit root tests are available, each based on different assumptions and potentially leading to conflicting outcomes (23). The most used unit root tests are the Augmented Dickey-Fuller (ADF) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test (23). The ADF test examines if the autoregressive coefficient in a model is 1, which means there's a unit root, indicating non-stationarity. Its null hypothesis is that the series is non-stationary, while the alternative is that it is stationary. The KPSS test directly tests if the series is stationary around a trend. Its null hypothesis is stationarity, while the alternative is non-stationarity (23).

Time series observations are frequently correlated with past observations and consequently not independently distributed. This is known as autocorrelation and is often associated with non-stationarity. Differencing the data is regularly used to remove autocorrelation, hence any required data transformations should be conducted before testing for autocorrelation. Autocorrelation functions (ACFs) are useful for checking stationarity and autocorrelation in time series data. An ACF plot shows the correlation between each observation and its previous values at different time lags. The number of time points between an observation and its previous values is referred to as the lag (24). The partial autocorrelation function (PACF) complements the ACF by showing the correlation between an observation and its past values that cannot be explained by correlations at lower order lags. In a stationary time series, the ACF decreases to zero relatively quickly, whereas in non-stationary data, it decreases slowly (23).

The Autoregressive Integrated Moving Averages (ARIMA) model is an extension of the ARMA model that includes an integration stage. This integration stage refers to differentiating the data to make the time series stationary, with constant average variation over time. The ARMA model combines characteristics of the Autoregressive model (AR(p)) and Moving Average models (MA(q)) to portray the temporal dependencies in a time series. In the autoregressive model (AR(p)), the variable of interest is predicted using a linear combination of its past values. An AR(p) model, or an autoregressive model of order p can be written as:

$$y_t = C + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \tag{1}$$

where ε_t is white noise. This type of model is very flexible at handling a broad range of different time series patterns.

In the moving average model (MA(q)), the forecast of a point in time depends linearly on the errors (residuals) of previous observations. An MA(q) model, or a moving average model of order q can be written as:

$$y_t = C + \varepsilon_1 + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \tag{2}$$

Therefore, an ARIMA model with parameters (p, d, q) integrates three statistical techniques: autoregression (AR(p)), which delineates the relationship amongst observations; differencing (d), which indicates the number of differences required for stationarity; and moving average (MA(q)), which incorporates additional factors. The full model ARIMA (p, d, q) can be written as:

$$y'_t = C + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{3}$$

where y'_t is the differenced series (23).

For example, an ARIMA (0,0,0) is a white noise (stationary model). An AR model with p lags is ARIMA (p, 0, 0), and an MA model with q lags is ARIMA (0, 0, q) (24).

From a time plot alone, determining suitable values for p and q is not possible. However, the ACF plot and the PACF plot can sometimes be used for this intention. As was said before, the ACF plot shows how the current value of a time series is correlated with its past values. However, it is important to note that some correlations could happen simply because certain values are linked through intermediary values, rather than due to any new information. The PACF plot, on the other hand, reveals the partial correlation between the series and its lagged values. Unlike the ACF, the first partial autocorrelation is the same as the first autocorrelation because there is no intermediary value to eliminate the correlation (23).

If the data follow an ARIMA (p, d, 0) or ARIMA (0, d, q) model, the ACF and PACF plots can aid in determining the value of p or q. However, if both p and q are positive, these plots may not be helpful. For an ARIMA (p, d, 0) model, the ACF and PACF plots of the differenced data may exhibit the following patterns: the ACF shows exponential decay or sinusoidal behavior; there is a significant spike at lag p in the PACF, with no spikes beyond lag p. On the other hand, for an ARIMA (0, d, q) model, the PACF demonstrates exponential decay or sinusoidal patterns and there is a significant spike at lag q in the ACF, with no spikes beyond lag q (23). Table 1 shows how to determine p and q parameters from the ACF and PACF plots.

Table 1 Tips for selecting p and q parameters from ACF and PACF. Adapted from (23)

Model Type	**Characteristics of ACF and PACF**	
	ACF	**PACF**
ARIMA (p,d,0)	Tails off or is sinusoidal	Cuts off lag p
ARIMA (0,d,q)	Cuts off lag q	Tails off or is sinusoidal
ARIMA (p,d,q)	Tails off or is sinusoidal	Tails off or is sinusoidal

Once the model order, represented by the values of p, d, and q, has been identified in ARIMA modelling, the next step involves estimating the model parameters, denoted as c, $\phi_1,\ldots,\phi_p$, $\theta_1,\ldots,\theta_q$. To do so, the maximum likelihood estimation (MLE) is commonly used. This technique seeks to find the parameter values that maximize the probability of observing the given data. In the context of ARIMA models, MLE is alike to the least squares estimation used in regression models, where it minimizes the sum of squared residuals. However, estimating ARIMA models is more complex compared to regression models due to the nature of time series data (23).

To evaluate the goodness of fit of each model, it is frequently to use metrics such as Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The AIC can be written as

$$AIC=-2log(L)+2(p+q+k+1), \tag{4}$$

being L the likelihood of the data, $k=1$ if $c\neq0$ and $k=0$ if $c=0$. For ARIMA models, the corrected AIC can be written as

$$AICc = AIC + \frac{2(p+q+k+1)(p+q+k+2)}{T-p-q-k-2}, \tag{5}$$

and the Bayesian Information Criterion can be written as

$$BIC = AIC + [\log(T) - 2](p+q+k+1). \tag{6}$$

It is possible to obtain a good model by minimizing the AIC, AICc and BIC (23).

The next step is to check whether the residuals of the chosen model show white noise behaviour. To do so, it is necessary to look for any patterns or trends in the residual plots, or to formally test for the presence of autocorrelation in the residuals using the Ljung-Box test for white noise (a small p-value (< 0.05) suggests the presence of autocorrelation in the residuals, indicating that the model may need adjustment). The orders of the AR and/or MA should be adjusted if autocorrelation continues or if the model does not adequately fit the data (24). Figure 2 shows the flow chart for selecting an ARIMA model.

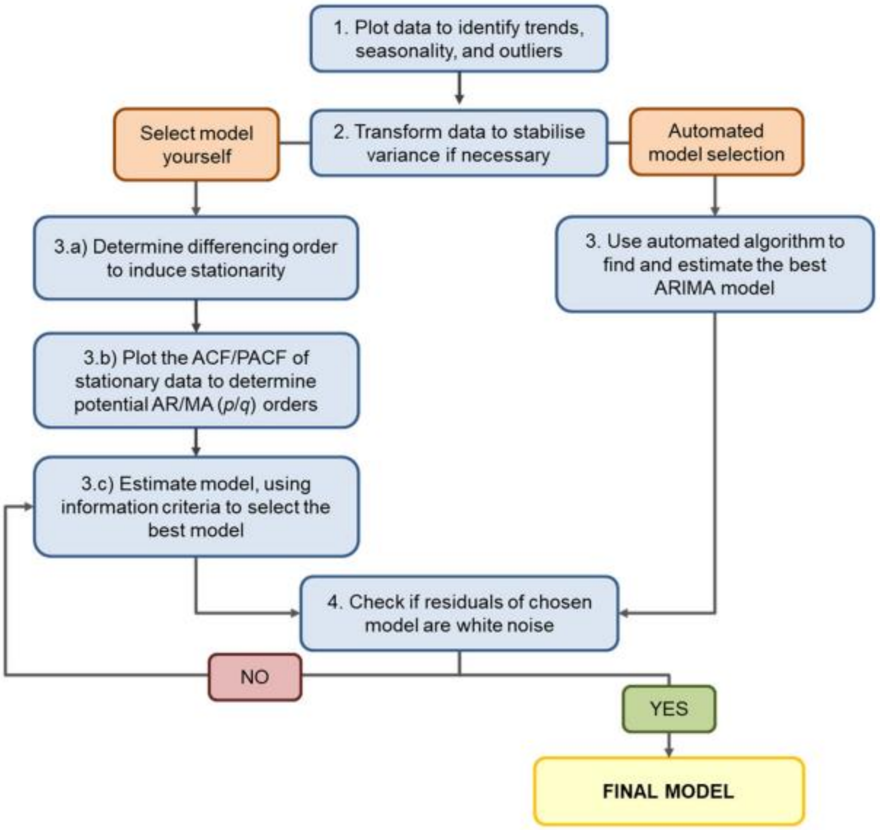


Figure 2 Flow chart for ARIMA model selection. Adapted from (23)

After selecting the final model, we can use ARIMA time series for forecasting. Forecasting means predicting future events using past data and statistical methods. It can be short-term forecast, just few days, weeks, or months ahead, medium-term forecast, extending over 1 to 2 years, or long-term forecast, which goes beyond a couple of years into the future. (23).

The summary of the forecast errors is often used to measure the forecast accuracy. The difference between a real observed value and its forecast is called the forecast error. When the errors are in the same scale as the data, they are called "scale-dependent errors", thus are not used to make comparisons between series with different units (23). The most common measures are the Mean absolute error (MAE), Mean squared error (MSE), and the Root mean squared error (RMSE):

$$MAE = mean\ (\ |e_t|\ ) \tag{7}$$

$$RMSE = \sqrt{mean\ (e_t^2)} \tag{8}$$

Where $e_t = y_t - y^t$ is the error at time *t.*

"Percentage errors" are unit-free and are frequently used to compare data sets forecast performances (23). The most used is the Mean absolute percentage error (MAPE):

$$MAPE = mean\ (|\,p_t\,|) \tag{9}$$

Where $p_t = 100e_t/y_t$ is the percentage error.

3.3. Clustering: methods, distance measures and evaluation

Clustering consists of a set of exploratory multivariate data analysis methods for identifying natural groups (clusters) in data. Clusters are formed by grouping objects that have maximum similarity within the same group, and minimum similarity with objects in other groups (7). Given a temporal dataset *D,* defined as $D = \{F_1,F_2,\ldots,F_n\}$, the time series clustering involves the process of unsupervised partitioning of D into clusters Ci, where $C = \{C_1,C_2,\ldots,C_k\}$, $k \leq n$. The process is based on a specific similarity measure, with the goal of grouping homogeneous time series together without any prior knowledge of group association. Here, each C_i represents a cluster, where

$$D = \mathrm{U}^k{}_{i=1}\, C_i, \text{ and } C_i \cap C_j = \emptyset \text{ for } I \neq j$$

(7). This is crucial for understanding, analyzing, and extracting insights from temporal data across various domains, playing a fundamental role in exploratory data analysis, predictive modeling, and decision-making processes. Time series databases are valuable sources of information, particularly for pattern discovery. To uncover these patterns, clustering is a frequent method used within time series datasets.

Given their large size, human inspection of time series databases can be difficult. Therefore, many users prefer structured datasets rather than large ones. As a result, time series data are often represented as sets of similar time series by grouping in non-overlapping clusters or by a hierarchical taxonomy of abstract concepts (7). Time series clustering is a widely used exploratory technique and is also a subroutine in more complex data mining algorithms, including rule discovery, indexing, classification, and anomaly detection (26). Visual representation of time series cluster structures helps users to quickly understand the data's structure, clusters, anomalies, and other regularities within the datasets.

However, time series clustering can be challenging due to a number of factors. Firstly, the data size is often larger than the memory capacity, slowing down the clustering process. Secondly, the data is usually high-dimensional, which poses challenges for clustering algorithms and further slows down the process. Lastly, selecting an appropriate similarity measure greatly affects the effectiveness and accuracy of the clusters generated.

According to Aghabozorgi in (7), time series clustering can be classified into three categories:

- Whole time series clustering – a technique that involves categorizing complete time series into groups, rather than focusing only on individual points in those series. This approach is valuable when you want to understand general patterns and similarities between entire time series;
- Subsequence clustering – focuses on identifying and grouping small segments of a time series;
- Time point clustering – a technique that groups time points by considering the temporal proximity and the similarity of the corresponding values.

The focus of this dissertation is on whole time series clustering. For any time series clustering approach, the main points to be considered are: how to measure the similarity between time series; how to compress the series or reduce dimension, and what algorithm to use for cluster.

Data representation stands significant challenges for time series clustering due to the large size of time series data, which exceeds memory capacity, leading to exponential increases in processing time. Thus, it is crucial to represent this type of data efficiently to ease algorithm execution slowdowns and minimize data loss. Effective representation methods should primarily reduce data size and preserve both local and global shape characteristics (27). Data representations are usually considered in four sections:

- Data adaptive methods – adjust parameters based on time series data being processed. These methods use segments of unequal lengths to minimize global reconstruction errors.
- Non-data adaptive methods – use fix-size parameters for the representing time series data.

- Model-based methods – assume that the time series was generated by an underlying model. The challenge is determining the parameters that define this model. Time series produced by the same parameters using the underlying model are considered similar.
- In data dictated methods – the dimension reduction rate is automatically determined without user intervention.

The similarity measure is essential for time series clustering, determining how closely two time series are similar to each other. Unlike traditional methods, time series clustering depend on approximate calculations based on distance functions, where larger estimated distances suggest lower similarity (7). The similarity between two time series T and U, each of size "n", is defined as the length of the path connecting pairs of points. This distance serves as a measure of similarity, D (T, U), where "d" is the function that calculates the distance between the two time series (T, U) (27). Various methods exist for measuring similarity, grouped into three categories: similarity in time, shape, and change (7).

- Similarity in time – This category highlights the temporal dependence of time series data. Correlation based distances and the Euclidean distance measure are commonly used for this purpose. However, calculating similarity directly on raw time series can be computationally expensive, so preprocessing or transformation steps are often required.
- Similarity in shape – The focus is on identifying time series with similar patterns, regardless of the specific time they occur. Time series with similar patterns are assigned to the same cluster. Elastic measures, such as, the Dynamic Time Warping (DTW), are used to find similar time series in shape.
- Similarity in change – In this approach, the data represented through a model, such as ARIMA, and then the distance measure is applied on the parameters of the fitted model. This metric results in time series clusters with similar autocorrelation structures, reflecting similarities in how the time series change over time.

Time series clustering algorithms can be categorized in five sections: partition-based, hierarchical-based, density-based, grid-based, and model-based (27). The reason why there is multiple clustering approaches and techniques that have been proposed for time series data, is due to the fact that there's no exact definition to the idea of "cluster". Each algorithm uses a different inclusion principle (28). Some adapt static data algorithms directly or incorporate preprocessing steps to transform raw time series data into feature vectors or model parameters before clustering (27).

The partitional clustering approach makes *k* groups, defined *a priori*, in a way that each groups has at least one object. The K-means is a frequently used algorithm that divides data into predefined sets, each cluster has its prototype (cluster center). The main goal of this algorithm is to minimize the total distance between all objects in a cluster from their prototype (7). The K-means algorithm is suitable for large data sets, however, it depends on initial conditions, such as, the number of clusters (29). Another partitional approach to clustering is the Fuzzy C-means, also known as Fuzzy K-means, algorithm. The advantage is that, contrarily of K-means, this algorithm allows each object to be part of two or more clusters (7).

In hierarchical clustering methods, clusters are formed by using agglomerative or divisive algorithms. The agglomerative method starts with each observation as an independent cluster and gradually merges them based on similarity, this is also known as the bottom-up approach. Divisive methods start with all observations as one cluster and separate them step by step, otherwise known as the top-down approach (7). Hierarchical clustering gives a graphical representation called dendrograms or cluster trees, where the distance between clusters is represented by the dendrogram's height (30). Hierarchical clustering methods can be further grouped based on linkages (Figure 3):

- Single linkage – the distance between two clusters is determined by the minimum distance from any member of a cluster to any member of other cluster (28,30).
- Complete linkage – the distance between two clusters is the maximum distance between two members from the clusters (28,30).
- Average linkage – the distance between two clusters is the average between all members of the two clusters (28,30).
- Centroid linkage – consists of the distance between the centroids of the members in the clusters (30).
- Ward linkage – also known as Ward's minimum variance method, the distance between two clusters is calculated based on the deviation of an element from the mean (31).

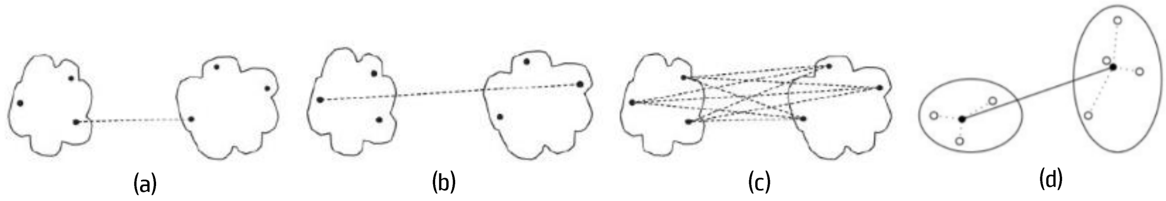


Figure 3 Hierarchical clustering linkages: (a) Single; (b) Complete; (c) Average; (d) Centroid. Adapted from (30)

The advantage of hierarchical clustering is that it is not required to specify the number of clusters *a priori*, as an initial parameter (29). However, with this approach it is not possible to change the clusters after they have been formed, either through divisive or agglomerative methods, meaning that hierarchical clustering algorithms are not able to correct possible misclassifications (28).

Density-based clustering identifies clusters based on dense multi-dimensional areas where objects attract each other. This approach is useful for noisy environments and handling outliers. DBSCAN (32) and

OPTICS (33) are common examples of algorithms that used density-based clustering (27). This type of clustering can be computationally complex and highly sensitive to input parameters (7).

Model based clustering assumes a model for each cluster and finds the best fit of data to that model. This approach starts by selecting centroids randomly and adding noise to the with a normal distribution. Model-based methods can use statistical approaches, such as COBWED, or neural network approaches, for example, ART or Self-Organization Map. Overall, this type of clustering method aims to identify clusters by fitting data to predefined models, making this approach flexible to various types of data. However, it has some limitations, for example, it requires setting parameters and relies on user assumptions, which can lead to inaccurate cluster results, and it often has slow processing time especially when using the neural networks approach on large data sets (7).

Lastly, Grid-based clustering involves clustering the space surrounding objects rather than the objects themselves. This method partitions data space using grids made of square cells, with STING being a typical example. However, the clustering results may lack significance and reliability, despite of having a fast processing time (7,27).

Overall, each clustering method has its advantages and limitations, and ultimately, the most suitable clustering method depends on the specific goals of the analysis and the characteristics of the data being clustered.

Determining the optimal number of clusters for each method and evaluating the performance of a clustering algorithm is a very important task. It provides a quantitative measure of how well the clusters created by an algorithm align with the underlying structure of the data and the clustering objectives. The evaluation metrics should consider if the clustering accurately defines the separations of the data, similar to a ground truth set of classes or satisfying specific assumptions, such as that objects belong to the same class are more similar than objects of different classes according to some similarity measure. Some evaluation metrics used are the Silhouette score (34), the Dunn index, the Cindex, and the McClain index.

The Silhouette score (34), assesses the cluster quality by evaluating the cohesion within clusters and the separation between clusters, it measures the proximity of each data point to other members of its own cluster compared with its proximity to data points in the nearest cluster. The silhouette score is defined as (35):

$$s(i) = \frac{b(i) - a(i)}{\max\left(a(i), b(i)\right)} \tag{10}$$

where $a(i)$ is the mean distance between a data point and all other points in the same class, and $b(i)$ is the mean distance between a data point and all other points in the nearest cluster. The Silhouette score takes values between -1 and 1. When the value is close to 1, indicates that the data point is closer to points in its own cluster then to points in the next neighbor cluster, in other words, it means the data point is well clustered. On the other hand, if the value is close to -1, indicates that the data point is closer to points in the neighboring cluster than the points in its owns cluster, meaning the data point is in the wrong cluster. If the value is close to 0 means that the data point might belong to some other cluster (36). Therefore, the score is higher when clusters are dense and well separated (35).

The Dunn Index (37) calculates the ratio of the minimum inter-cluster distance to the maximum intra-cluster distance to evaluate the compactness and separation of clusters. Higher values indicate well separated and compact clusters. The Dunn index is defined as (38):

$$Dunn = \frac{\min\limits_{1 \le i < j \le q} d(C_i, C_j)}{\max\limits_{1 \le k \le q} diam\ (C_k)} \tag{11}$$

The Cindex (39) calculates the sum of the smallest and largest distances between all pairs of points (*Smin* and *Smax*), and the sum of distances within clusters (*Sw*). the index is through the formula (38):

$$Cindex = \frac{Sw - Smin}{Smax - Smin} \tag{12}$$

Lower values indicate a better clustering, as it implies that the points within each cluster are close compared to the overall distances in the dataset.

The McClain Index (40) compares the average intra-cluster distance to the average inter-cluster distance, a lower value indicates better clustering. This index can be calculated through the following formula (38):

$$McClain = \frac{\bar{S}w}{\bar{S}b} = \frac{Sw/Nw}{Sb/Nb} \tag{13}$$

Where Sw is the sum of the within-cluster distances, Sb is the sum of the between-cluster distances, Nw is the total number of pairs of observations belonging to the same cluster, and Nb is the total number of pairs of observations belonging to different clusters.

3.4. State of the Art of Clustering Methods for ARIMA Time Series

The process for measuring the similarity between two time series, X_T and Y_T, requires fitting an ARIMA model to each series and then compare the fitted models to determine the level of similarity. The first step is to estimate the structure and the parameters of the models. The structure of ARIMA models can either be assumed or automatically estimated using the AIC and BIC values, while the parameters are usually estimated using the least squares estimators (41). Some of the dissimilarity measures that have been developed over the years are the Piccolo distance (8), the Maharaj distance (9,10), and the Cepstral-based distance (11).

Piccolo (8), introduced a dissimilarity measure for invertible ARIMA processes as the Euclidean distance between the AR (∞) operators approximating their respective ARIMA structures. Piccolo argued that autoregressive expansions capture all pertinent information about the stochastic structure of such processes, except for initial values (41). To address non-stationarity, differencing is performed to achieve stationarity, while any seasonality present is removed before further analysis. The approach involves fitting truncated AR (∞) models of orders k_1 and k_2 to approximate the generating processes of the time series X_T and Y_T, respectively.

If $\widehat{\Pi}_{XT} = (\hat{\pi}_{1,XT}, \dots, \hat{\pi}_{k1,XT})^T$ and $\widehat{\Pi}_{YT} = (\hat{\pi}_{1,YT}, \dots, \hat{\pi}_{k2,YT})^T$ represent the vectors of AR(k_1) and AR(k_2) parameter estimations for X_T and Y_T, respectively, then the Piccolo's distance is given by

$$d_{PIC}\,(XT, YT) = \sqrt{\sum_{j=1}^{k}\left(\hat{\pi}'_{j,XT} - \hat{\pi}'_{j,YT}\right)^2} \tag{14}$$

where $k = \max(k_1, k_2)$, $\hat{\pi}'_{j,XT} = \hat{\pi}_{j,XT}$, if $j \leq k_1$, and $\hat{\pi}'_{j,XT} = 0$ otherwise, and analogously $\hat{\pi}'_{j,YT} = \hat{\pi}_{j,YT}$, if $j \leq k_2$, and $\hat{\pi}'_{j,YT} = 0$ (41).

Maharaj (9,10) introduced, for the class of ARMA processes, two discrepancy measures based on hypotheses testing to determine whether two time series have significantly different generating processes. The first of these metrics is given by the test statistic:

$$d_{MAH}\,(XT, YT) = \sqrt{T}\left(\widehat{\Pi}'_{XT} - \widehat{\Pi}'_{YT}\right)^T \hat{V}^{-1}\left(\widehat{\Pi}'_{XT} - \widehat{\Pi}'_{YT}\right) \tag{15}$$

where $\widehat{\Pi}'_{XT}$ and $\widehat{\Pi}'_{YT}$ are the AR (k) parameter estimations, with k selected as in the Piccolo's distance, and $\hat{V} = \sigma^2_{XT} R^{-1}_{XT}(k) + \sigma^2_{YT} R^{-1}_{YT}(k)$, being σ^2_{XT} and σ^2_{YT} the variances of the white noise processes, and the sample covariance matrices are represented by R_{XT} and R_{YT} (41).

Maharaj also showed that under the null hypothesis of equal generating processes ($\Pi_{XT} = \Pi_{YT}$), the d_{MAH} asymptotically follows a $\mathcal{X}^2$ distribution. Therefore, the dissimilarity between $\widehat{\Pi}'_{XT}$ and $\widehat{\Pi}'_{YT}$ can be measured using the associated p value,

$$d_{MAH,p}\ (XT, YT) = P\left(\mathcal{X}_k^2 > d_{MAH}(XT, YT)\right) \tag{16}$$

A hierarchical algorithm based on the pairwise matrix of p-values $d_{MAH,p}$ can be developed to cluster time series data. By setting a threshold significance level α (e.g., 5% or 1%), a criterion for clustering homogeneity is established. Series with p-values exceeding α are grouped together, indicating similarity in their dynamic structures, while those with p-values below α are considered significantly different and placed in separate groups.

While both d_{MAH} and d_{PIC} evaluate dissimilarity by comparing autoregressive approximations of two series, they differ significantly. The d_{PIC} overlooks the variance of the white noise processes associated with the observed series, whereas d_{MAH} incorporates these variances in its calculation. This distinction is crucial, particularly in clustering applications, as d_{MAH} values are influenced by the scale unit. Therefore, understanding this difference is important when utilizing these dissimilarity measures (41).

Kalpakis, Gada, and Puttagunta (11) proposed the Cepstral-based distance, a measure based on the linear predicting coding (LPC) cepstrum. This is a kind of signal processing analysis (42). The cepstrum is obtained by taking the inverse Fourier transform of the short-time logarithmic amplitude spectrum. Specifically, the LPC cepstrum is derived from the autoregression coefficients of the linear model of the signal. This approach is chosen because LPC cepstral coefficients demonstrate strong discriminatory abilities among ARIMA time series. Notably, only a select few LPC cepstral coefficients are needed to effectively capture and retain a significant amount of information regarding the underlying ARIMA model. The LPC coefficients for AR (p) time series are defined by:

$$\psi_h = \begin{cases} \varphi_1 & if\ h = 1 \\ \varphi_h + \sum_{m=1}^{h-1} (\varphi_m - \psi_{h-m}) & if\ 1 < h \le p \\ \sum_{m=1}^{p} (1 - \frac{m}{h}) \varphi_m \psi_{h-m} & if\ p < h \end{cases} \tag{17}$$

where X_T is a time series that follows an AR (p) structure, $X_T = \sum_{r=1}^{p} \varphi_r X_{t-r} + \varepsilon_t$, being φ_r the autoregression coefficients and ε_t a white noise process with zero mean and zero variance (41).

Kalpakis et al. (2001) consider using the Euclidean distance between the estimated LPC cepstral coefficients of two time series X_T and Y_T as a measure of their distance,

$$d_{LPC.Cep}(X_T, Y_T) = \sqrt{\sum\nolimits_{i=1}^{T} \left(\psi_{i,XT} - \psi_{1,YT}\right)^2} \tag{18}$$

A study by Nascimento et al. (43), introduced and extension of the affinity coefficient for the autoregressive expansions of the invertible autoregressive moving average models to measure their similarity. This research findings indicate that processes with similar forecast functions are grouped together as expected concerning the affinity coefficient. Moreover, the coefficient is very sensitive to behavior changes of the forecast functions. Another study (44), by the same author, with the goal to identify a hierarchy of clusters of Portuguese mortality by non-communicable diseases, used the Euclidean distance along with the complete and average linkage. Six clusters were identified with both criteria, indicating some order of disease severity in the way clusters are grouped.

Various authors implemented the Piccolo distance to measure the similarity between time series in their work. In their study, Hendrawati et al. (45) aimed to cluster rainfall data from West Java using the Piccolo distance at both individual and cluster levels. The dataset included average monthly rainfall intensity (mm) spanning from 1981 to 2014, obtained from the Meteorology Climatology and Geophysics Agency. They used ARIMA models for data modeling and the Piccolo distance to quantify dissimilarity between time series. Clustering was carried out using the Ward method, with the Silhouette index serving as the evaluation metric. Their findings demonstrated that the cluster model was more effective and can be used to represent individual models.

In a similar study (46), with the purpose of grouping national inflation data in Indonesia, Hendrawati et al. modelled the data using ARIMA models and the Piccolo distance along with the Ward method and the Silhouette index to cluster the time series. The study concluded that the cluster model was more efficient than the individual models.

In the 2020 study, Vatansever et al (47) aimed to identify homogeneous housing market areas among 196 districts across five major Turkish cities based on house sale price indices. Additionally, the study sought to forecast these indices. Applying an autoregressive (AR) model and the Piccolo distance measure, a fuzzy clustering approach, specifically fuzzy c-means, was used for clustering. Evaluation methods including Silhouette, Davies–Bouldin, Calinski–Harabasz, Dunn, R^2, C-index, Krzanowski–Lai, Hartigan, and weighted inter/intra, homogeneity, and separation criteria were used to determine the optimum number of clusters. Lastly, the experimental findings revealed that AR model-based fuzzy time series clustering approaches outperformed AR models for 140 out of 196 sale price indices based on evaluation metrics such as MAE, MAPE, and RMSE. The AR model-based fuzzy time series clustering approach granted more successful results for clustering and forecasting purposes, and identified three distinct clusters among the 196 districts of the five major Turkish cities in terms of house sale price trends.

4. Methodology

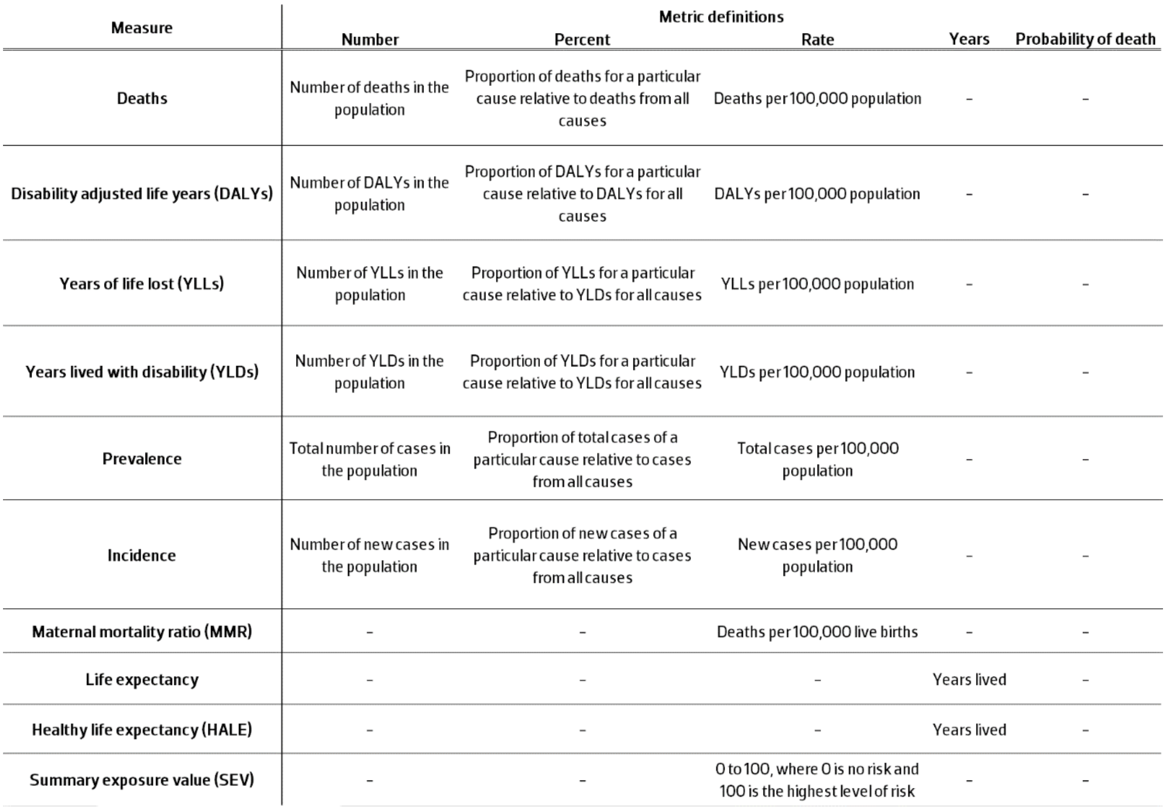
This chapter is going to focus on the methodology employed in this study, related to the description of the database and the process of data selection. The first section refers to an overview of the database, highlighting its structure, content, and relevance to research studies. Then the process for data selection is detailed, outlining the criteria and considerations used when selecting the dataset. Finally, the adopted methods are explained. This includes the processes for data preparation, the representation of the time series through ARIMA models, and the clustering methods.

This work was submitted to the ethics committee of E2S-P.PORTO on January 18th, 2024, with the process number CE0069E, and approved on May 27th, 2024.

4.1. Description of database and data selection

The most complete and recent Global Burden of Disease study was developed by the Institute for Health Metrics and Evaluation (IHME) at the University of Washington in the USA (14). The dataset chosen was obtained through the IHME website (3). The IHME allows to compare and analyze estimates of health indicators within a country, compare countries with other countries, world regions, or the entire world, and explore trends by country, age, and sex using the GBD Compare (48) and the GBD Results Tool (49) available on the website. The GBD Compare is an interactive tool that uses treemaps, arrow diagrams, and many other charts to explore health data and trends. The data can be filtered by cause, measure, year, age, sex, and metric. Table 2 provides definitions for each metric by measure included in the GBD Results Tool (48▯).

Table 2 Definitions for metric by measure in the GBD Results Tool (49)



Measure	Metric definitions				
	Number	**Percent**	**Rate**	**Years**	**Probability of death**
Deaths	Number of deaths in the population	Proportion of deaths for a particular cause relative to deaths from all causes	Deaths per 100,000 population	-	-
Disability adjusted life years (DALYs)	Number of DALYs in the population	Proportion of DALYs for a particular cause relative to DALYs for all causes	DALYs per 100,000 population	-	-
Years of life lost (YLLs)	Number of YLLs in the population	Proportion of YLLs for a particular cause relative to YLDs for all causes	YLLs per 100,000 population	-	-
Years lived with disability (YLDs)	Number of YLDs in the population	Proportion of YLDs for a particular cause relative to YLDs for all causes	YLDs per 100,000 population	-	-
Prevalence	Total number of cases in the population	Proportion of total cases of a particular cause relative to cases from all causes	Total cases per 100,000 population	-	-
Incidence	Number of new cases in the population	Proportion of new cases of a particular cause relative to cases from all causes	New cases per 100,000 population	-	-
Maternal mortality ratio (MMR)	-	-	Deaths per 100,000 live births	-	-
Life expectancy	-	-	-	Years lived	-
Healthy life expectancy (HALE)	-	-	-	Years lived	-
Summary exposure value (SEV)	-	-	0 to 100, where 0 is no risk and 100 is the highest level of risk	-	-

Figure 4, taken from the GBD Compare page, represents the treemap of the burden of disease for the European Region, for both sexes. In blue is represented the non-communicable diseases, in red the communicable, maternal, neonatal, and nutritious diseases, and lastly in green the injuries. A darker shade of the color indicates a higher annual percentual change of DALYs from 1990 to 2019. When analyzing the health data for the European Region it is clear to see that non-communicable diseases play a big role in the overall burden of disease. Among the NCDs, the neoplasms, cardiovascular diseases, chronic respiratory diseases, digestive diseases, mental and neurological disorders, and diabetes and kidney diseases are the ones that stand out the most. This is in line with what was previously said in Chapter 2. For this reason, five of the non-communicable diseases mentioned were chosen for this study, the cardiovascular diseases, the chronic respiratory diseases, the neurological disorders, the chronic kidney diseases, and diabetes mellitus.

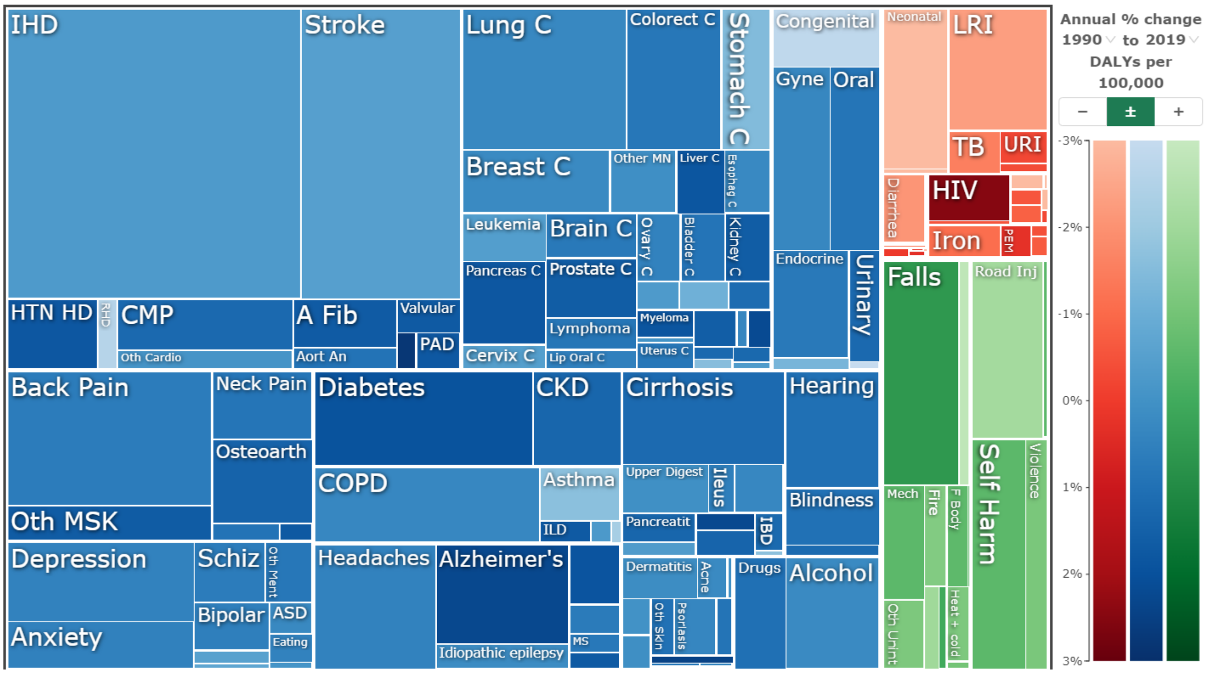


Figure 4 Tree map of the Burden of Disease for the European Region (48)

It is possible to download the data in the GBD Results page by selecting a few parameters:

- GBD estimate – was selected "Cause of death or injury"
- Measure – this parameter is the indicator for which estimates are produced. "DALYs" was selected
- Metric – is the unit by which a measure is expressed. Was picked "Rate"
- Cause – it is possible to choose just a single disease or injury or multiple. Here the five diseases mentioned above were selected
- Location – here 48 countries of the European Region were selected (Albania, Andorra, Armenia, Austria, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Republic of Moldova, Romania, Russia, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, and United Kingdom)
- Age – it is possible to choose different age groups from 0-6 days to 95+ years. We selected all ages
- Sex - it's possible to choose just male, female, or both. Here "Both" was selected
- Year Range – all years from 1990 to 2019 were selected.

After was stored in a file with a set of columns, which contain the information downloaded:

- measure_id – represents a unique numeric identifier for each GBD measure.
- measure_name – name of the GBD measure.
- location_id – numeric identifier for locations of various location types.
- location_name – name of the location, in this case the countries from the European Region.
- sex_id - unique numeric identifier for each sex.
- sex_name – name of the sex.
- age_id - unique numeric identifier for each age group.
- age_name – name of the age group.
- cause_id – represents a unique numeric identifier for each GBD cause.
- cause_name – name of the GBD cause.
- metric_id – unique numeric identifier for each metric.
- metric_name – name of the metric.
- year – from 1990 to 2019.
- val – the mean values of the estimate.
- upper – 95% uncertainty interval.
- lower - 95% uncertainty interval.

The columns used for the analysis of the data set were "measure_id", "location_name", "cause_id", "metric_id", and "val".

4.2. Data Pre-processing, ARIMA Model Evaluation, and Clustering Techniques for Time Series Analysis

After selecting the data from a public database, an exploratory data analysis was carried out, this allows evaluating data quality and identifying trends, which can provide initial information about the dynamics and behavior of DALYs over time. In a next phase, the ARIMA models that best fit each time series are estimated through automated model selection algorithms. The fit of the model was assessed with forecast accuracy metrics, such as MAE, RMSE, MSE, and MAPE. Regarding the steps for clustering, various hierarchical clustering methods were implemented, such as, complete, average, single, centroid, median, and ward. Alongside the different distance measures mentioned in Chapter 3, the Piccolo distance, Maharaj distance, and LPC distance. The performance of the clustering algorithm was weighed through evaluation metrics, such as the Silhouette score, CIndex, McClain Index, and Dunn Index. For the practical implementation of all these steps, software such as R or Python were used. Finally, the results obtained from this analytical approach, will be discussed highlighting significant trends identified in the data and their implications for public health.

4.2.1. Data Pre-processing and Exploratory Data Analysis

The data preparation and cleaning process is crucial for a proper exploratory data analysis (EDA). This process helps discovering patterns, detect anomalies, and identify relationships between the variables, which is crucial for obtaining insights and making informed decisions. After collecting the raw data, we examined the dataset through diverse visualization techniques. Line plots to visualize the time series for each country across NCDs, boxplots to explore data distribution, and additionally the time series components were isolated. This part of the analysis was carried out in Python using multiple packages for data manipulation and visualization, such as, NumPy, Pandas, Matplotlib, and Seaborn.

4.2.2. Representation and evaluation of time series with ARIMA models

In this subchapter, the process to obtain the ARIMA models that best represent the data and the steps to evaluate those models will be explained. This part of the process was carried in Python using the *pmdarima* package and *TimeSeriesSplit* from the scikit-learn library. ARIMA models are an effective tool for forecasting and understanding time series patterns. To find the best model for each country, the cross-validation technique was applied along with the autoARIMA function, after ensuring the data is appropriately formatted for analysis.

Cross-validation is a technique used to identify the best performing model and its parameters by iteratively training and testing the model on different subsets of the data (50). The simplest form of this method is the classical train-test split, where the dataset is divided into a training set, used to fit the model, and a test set, used to evaluate its performance. To enhance this approach, the train-test split can be performed multiple times, with the training and test data varying in each iteration. By making sure that every data point is used for both training and evaluation at different stages, cross-validation helps in selecting the most robust model, ensuring its performance is consistent across the entire dataset (50). Classical cross-validation techniques assume that the samples are independent, however, time series data is characterized by its temporal dependency, and so the model should be evaluated on the future observations to those that are used to train the model. This can be done using the *TimeSeriesSplit*, which returns the first *k* folds as a training set and the *(k+1)*th fold as a testing set. Figure 5 represents a visualization of this cross-validation behavior.

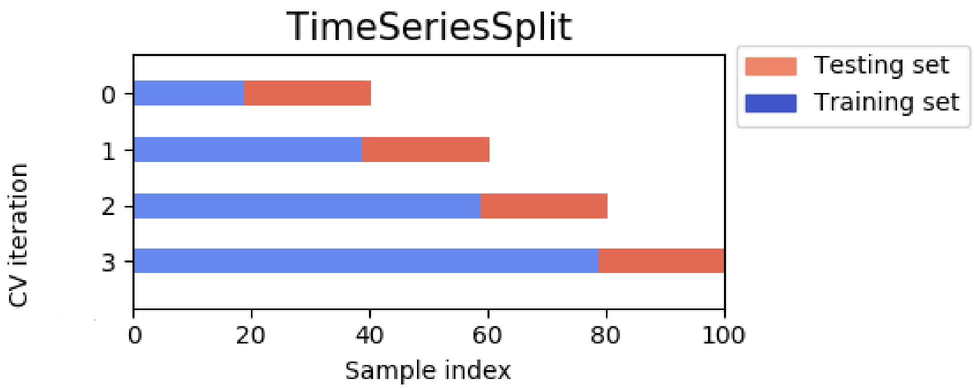


Figure 5 TimeSeriesSplit Cross-Validation behaviour

Wrapping different statistical and machine learning libraries, the *pmdarima* allows to generalize ARIMA models into a single class. The autoARIMA function performs a grid search over several possible values of the parameters *p* and *q*. The model is chosen based on the combination of parameters that results in the lowest value of AIC or BIC. For the parameter *d*, the function uses a test of stationarity and seasonality for seasonal models (51). This process is known for its efficiency and comprehensiveness since it saves time in model selection by automating the search for the best parameters.

At this stage the main goal is to obtain the best ARIMA model for each country, to do so, the dataset was split into training and testing sets, with 80% of the data being used for training and 20% for testing. The *TimeSeriesSplit* was defined with 5 splits for cross-validation, ensuring that the training data is sequentially split into different folds for model validation and maintaining the order of the time series. The code iterates over each country's time series data and over each cross-validation split. Within each split, the training and testing data are extracted, an autoARIMA model is fitted to the training set, the fitted model forecasts for the test set and the MAPE, MSE, MAE and RMSE values are calculated to validate the forecast accuracy of the model of the last fold. The autoARIMA models were saved in several dictionaries. Additionally, the indices of the train and test sets for each split were verified to provide a clear understanding of how the data was divided during cross-validation. With this approach, it's hoped for a better insight of the model's variability across different folds.

4.2.3. Clustering methods for ARIMA models

Clustering is a powerful analytical tool used to identify patterns and group objects in data. In this subchapter will be detailed the process of clustering ARIMA models employing different distance measures and experimenting with several hierarchical clustering methods. Each method was evaluated using multiple clustering indices to determine the optimal number of clusters and assess the quality of clustering. This part of the process was implemented using RStudio, where multiple libraries were used: *TSclust*, *tidyverse*, *gt*, *NbClust*, *gglopt2*, *gridExtra*, and *rworldmap*. These packages are required for time series clustering, visualization, and manipulation.

Firstly, based on the ARIMA models parameters obtained in Python, time series data was simulated with the arima.sim function for each country. Then the distance matrix for the Piccolo distance, the Maharaj distance, and the LPC distance, was calculated using the diss function from the *TSclust* package (41). This package in R provides a versatile set of tools for time series clustering. It supports multiple distance measures and with the function diss it is possible to compute a dissimilarity matrix for a set of time series data. After calculating the dissimilarity matrix for the three distance measures chosen, hierarchical clustering was performed using the *NbClust* package (38). This package allows to perform clustering using various linkage methods, the ones chosen for this analysis were Ward, Complete, Single, Average, Centroid, and Median. Using the same package, the best number of clusters and the clustering quality for each method was evaluated using four different indices, Silhouette, Dunn, Cindex, and Mcclain.

The results were stored in lists for further analysis. The following step was to visualize the clusters formed, to do so dendrograms were plotted and for better visualization of potential geographical patterns, a world map with the cluster membership between countries was made with the package *rworldmap*.

5. Results

In this chapter, the results of the analysis conducted on the time series are presented, focusing on three key aspects: the initial data analysis, the representation of time series and the clustering outcomes. The first section details the exploratory analysis that was performed on the dataset to discover patterns and associations between the variables. The second section explores time series modeling, highlighting the performance and accuracy of the ARIMA models in describing temporal patterns and forecasting trends. The third section focuses on the clustering results, where various clustering methodologies are applied to group the time series data based on underlying similarities. This chapter aims to provide a clear and comprehensive interpretation of the results, showcasing the chosen models and techniques.

5.1. Exploratory Data Analysis

Figure 6 represents a boxplot for each non-communicable disease chosen. Boxplots, also known as box-and-whisker plots, are useful to compare distributions of multiple sets of data side by side. For an easier interpretation of the data, it's important to understand that the median line inside the box shows the central tendency of the data, if the line is not centered then the data may be skewed. Also, the length of the box, or the interquartile range (IQR), shows the variability or spread of the middle 50% of the data. Another important point to notice, is the outliers, which are unusually high or low values in the dataset. As shown in figure 6 the cardiovascular diseases have the biggest interquartile range from 5000 to around 10000, the whiskers show a wide range of values and there are several outliers above the upper whisker which indicates unusually high values above 20000. For the remaining diseases the data is more well distributed, having a much narrower range and IQRs when compared to the cardiovascular diseases and the medians are significantly lower as well. The chronic kidney disease is the NCD with lower DALYs rate values. All diseases, except chronic respiratory diseases, have outliers.

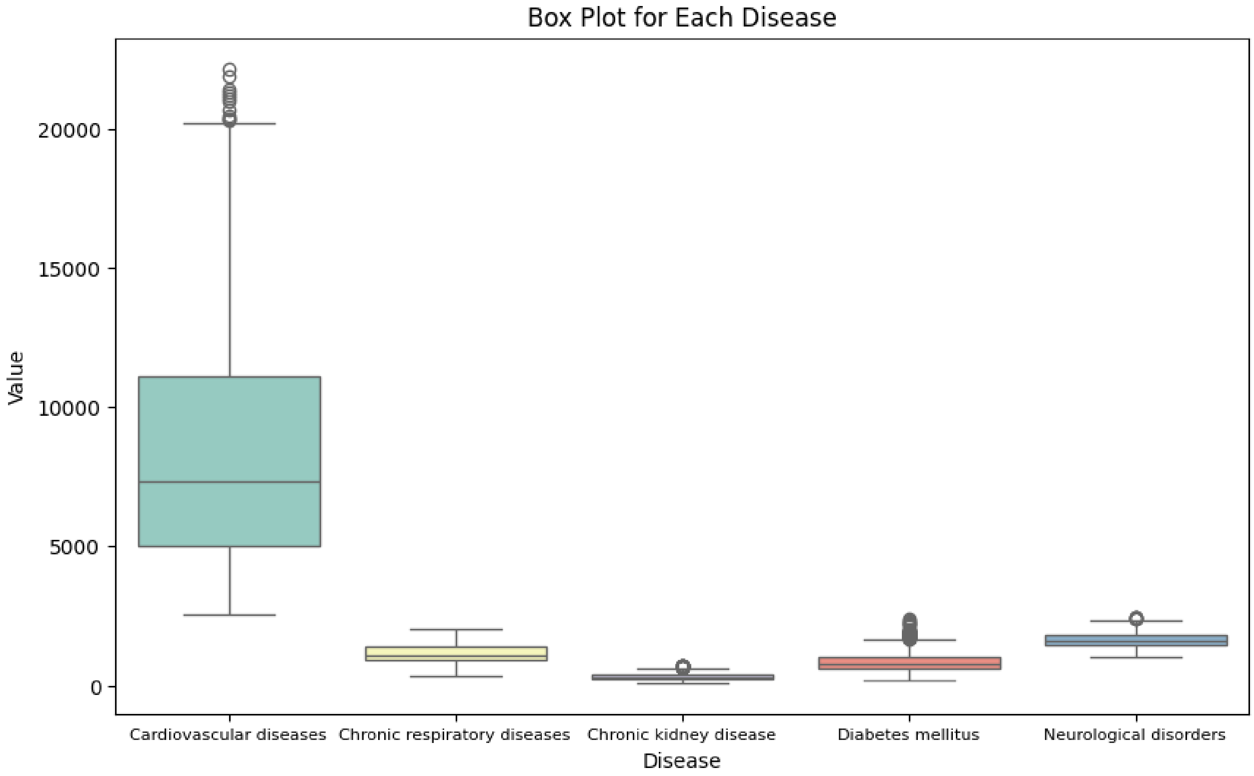


Figure 6 Boxplot of the DALYs values for each disease

Now analyzing each disease individually and starting with the cardiovascular diseases, Figure 7 shows a line graph representing the time series of the DALYs rate value from 1990 to 2019 for all selected countries for both sexes and all ages. We can see that some countries have a similar pattern to each other. For example, Latvia, Russia, Ukraine, and Belarus showed a sharp increase of the values of DALYs from 1992 to 1994 and then experienced a fall until around 1997 where the values started to rise gradually until 2005, and from that point onward has shown a steady decrease. Other countries such Italy, Andorra, Iceland, France, and Portugal remained fairly stable having values between 2500 and 10000 DALYs per 100.000 and showing a gentle downward trend. Bulgaria is the country with a higher value of DALYs, reaching its peak around 1997 and then has seen a slight fall over the years. In a general way, although for some countries the values of DALYs have significant changes over the years, the values from 1990 are close to the values from 2019, and there are clearly trends that allow to identify possible clusters.

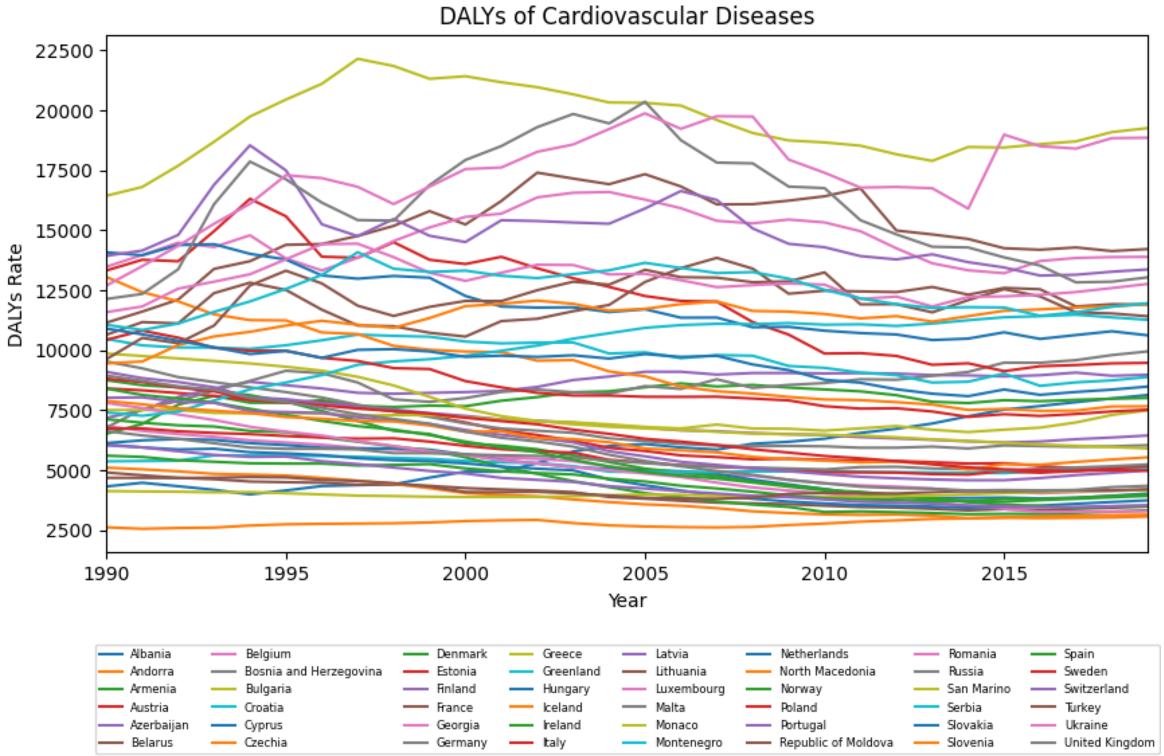


Figure 7 Time series of each country for Cardiovascular Diseases from 1990 to 2019

The time series represented in Figure 8 show the values of DALYs rate from 1990 to 2019 for all selected countries for both sexes and all ages for Chronic Respiratory Diseases. Firstly, it is clear the values of DALYs rates are significantly lower when comparing to the cardiovascular diseases values. For some countries the progression has been more linear over the years while for others, from 1990 to 2000, the values have shown more significant yearly changes. Ukraine and Belarus were the countries that showed a more dramatic fluctuation, having experienced an increase from 1990 to 1995 and then an abrupt decline until 2015 and since then the values flattened out. Denmark and the United Kingdom have the highest values of DALYs over the years and show a slight downward trajectory since 2000. The country with the lowest values is Montenegro with values ranging from 200 to 500 and has shown an upward trend since 1990.

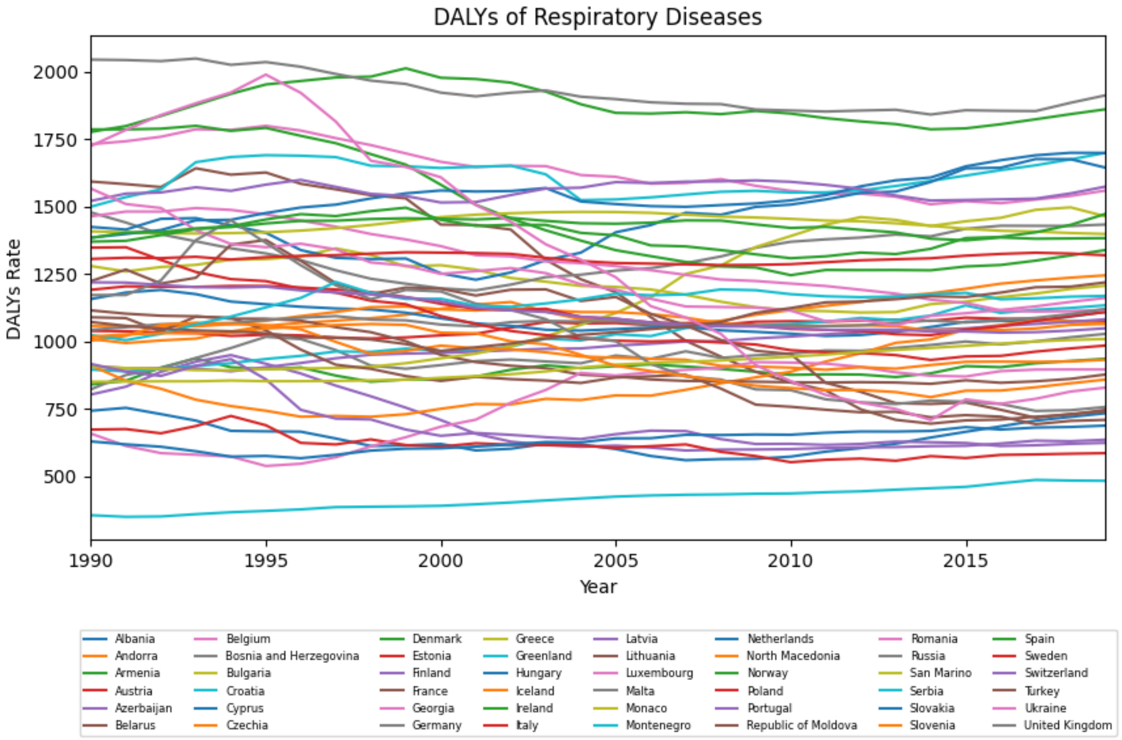


Figure 8 Time series of each country for Chronic Respiratory Diseases from 1990 to 2019

Figure 9 shows the line graph representing the time series of the DALYs rate value from 1990 to 2019 for all selected countries for both sexes and all ages for neurological disorders. The values range from 1000 and around 2500, and for the big majority of countries in the European Region the DALYs values have

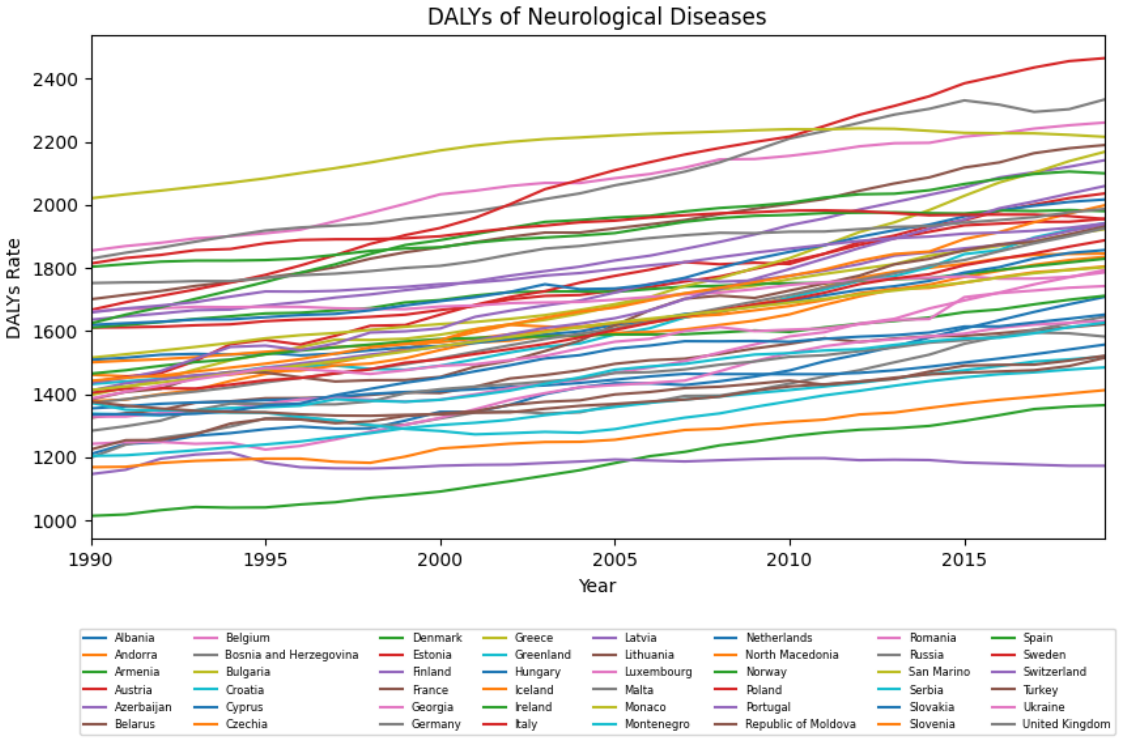


Figure 9 Time series for each country for Neurological Diseases from 1990 to 2019

been rising over the years, following a steady upward trend. Italy experienced a more significant rise of the values over the years, being the country with the highest value of DALYs for neurological disorders in 2019. Armenia started out as the country with the lowest values of DALYs, however it showed a gradual upward trajectory since 1990 and in 2005 surpassed Azerbaijan, which has a steady downward trend and was the country with the lowest DALYs rate value in 2019.

Figure 10 shows the line graph representing the time series of the DALYs rate value from 1990 to 2019 for all selected countries for both sexes and all ages for chronic kidney diseases. The values range from 100 to 700 DALYs per 100000. The majority of countries showed an upward trend, having some a more significant increase in the value of DALYs over the years, while for others it has been a steady rise. Portugal, Georgia, and Bulgaria experienced a rapidly upward trajectory since 1997. Greece is the country with the highest values of DALYs since 2005.

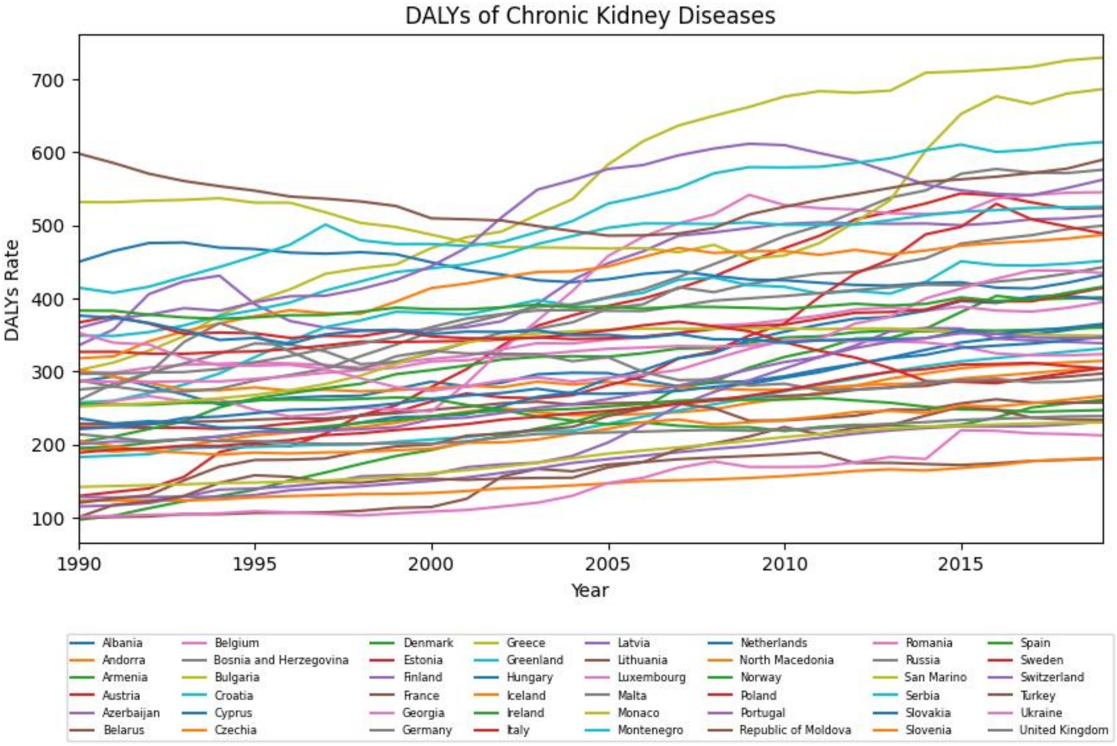


Figure 10 Time series for each country for Chronic Kidney Diseases from 1990 to 2019

Figure 11 shows the line graph representing the time series of the DALYs rate value from 1990 to 2019 for all selected countries for both sexes and all ages for diabetes mellitus. The values range from 200 to 2500 DALYs per 100000. Most countries show an upward trend over the years, having countries such as Bosnia and Herzegovina, Armenia, Georgia, and Croatia experienced a more significant variation. Iceland, Italy, Ireland, and Luxembourg are some of the countries that remained with a more stable fluctuation over the years. Bosnia and Herzegovina has the highest values since 2006.

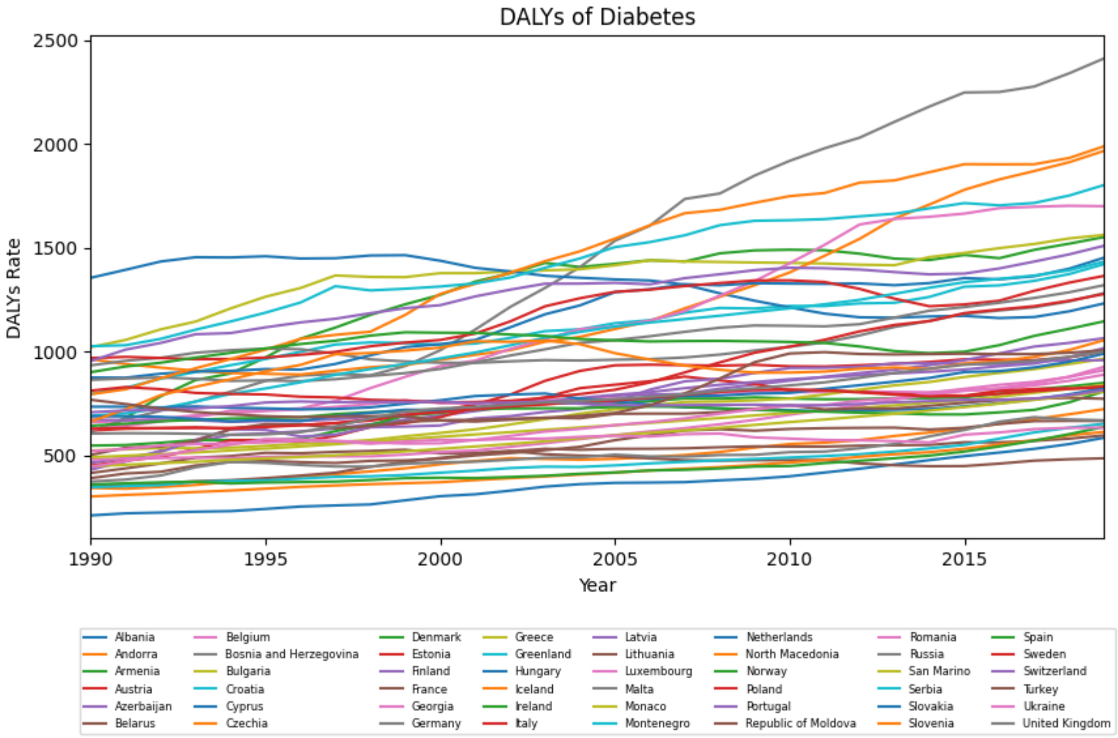
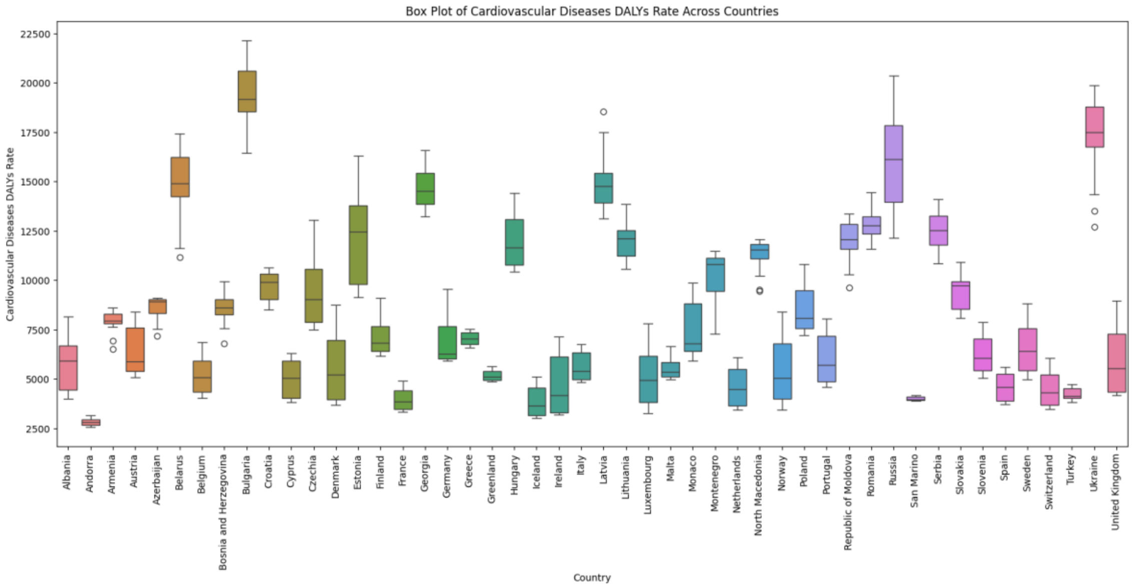
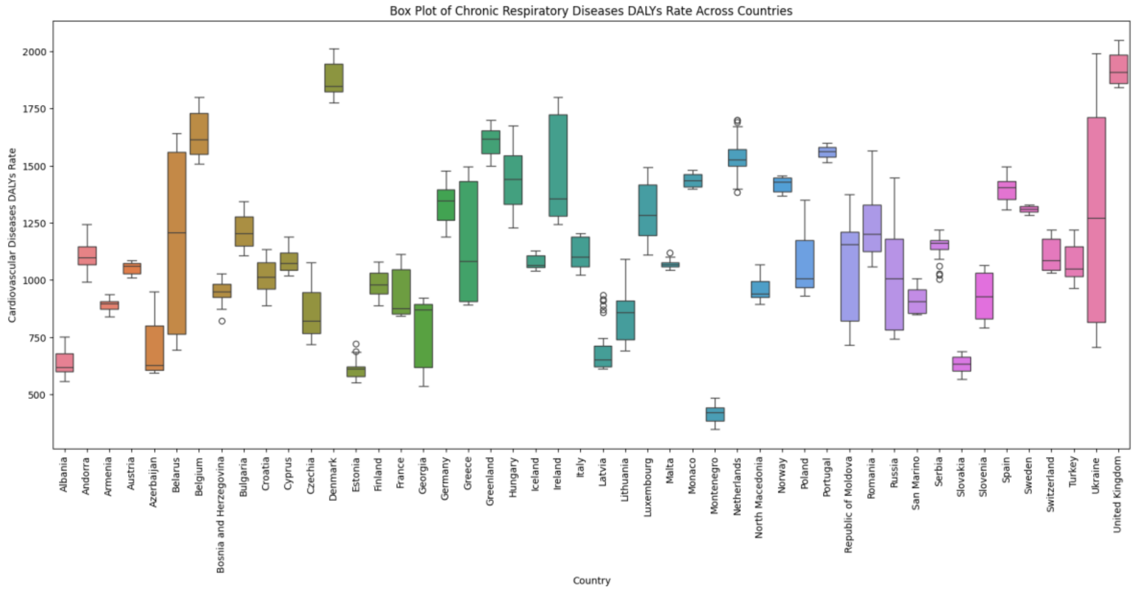


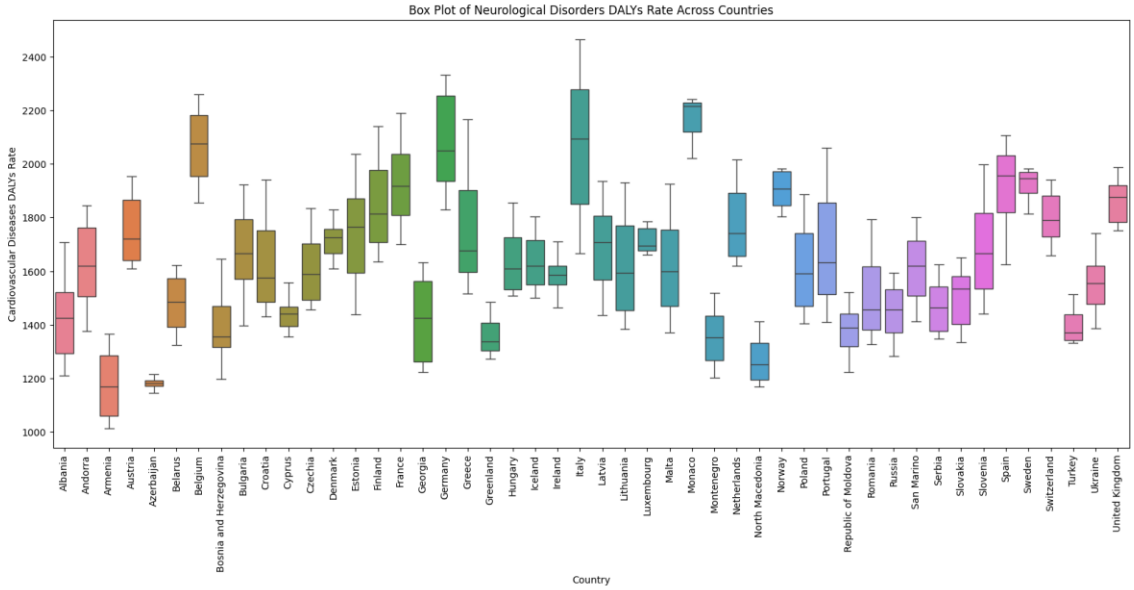
Figure 11 Time series for each country for Diabetes Mellitus from 1990 to 2019

To further analyze and understand the dataset, a graph of boxplots across countries for each disease was made. This way is possible to analyze each country individually and compare data distributions. As was mentioned before, when comparing box plots, it is useful to have in consideration the medians of the box plot, the interquartile ranges and whiskers, potential outliers, and lastly, signs of skewness. In Figure 12 are represented five graphs, one for each disease, of the boxplots across countries.

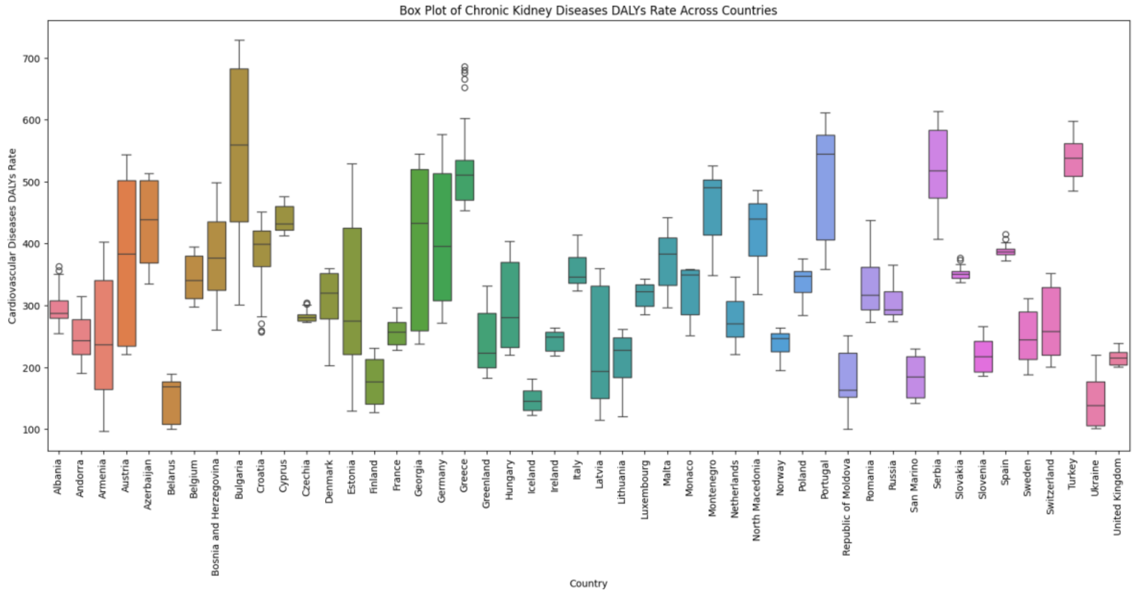




Box Plot of Chronic Respiratory Diseases DALYs Rate Across Countries



Box Plot of Neurological Disorders DALYs Rate Across Countries



Box Plot of Chronic Kidney Diseases DALYs Rate Across Countries

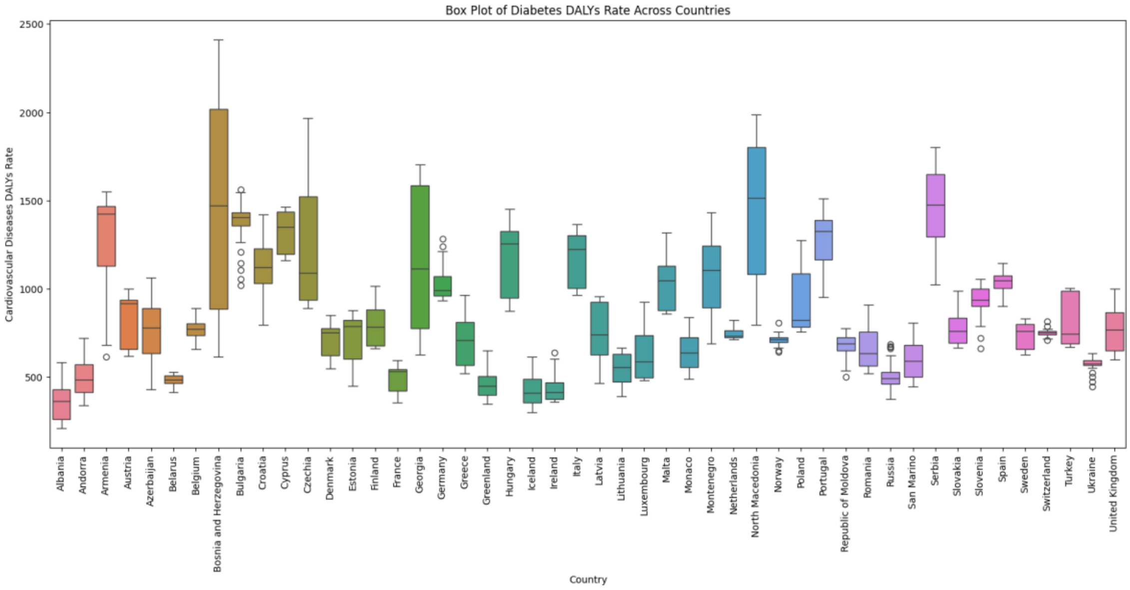
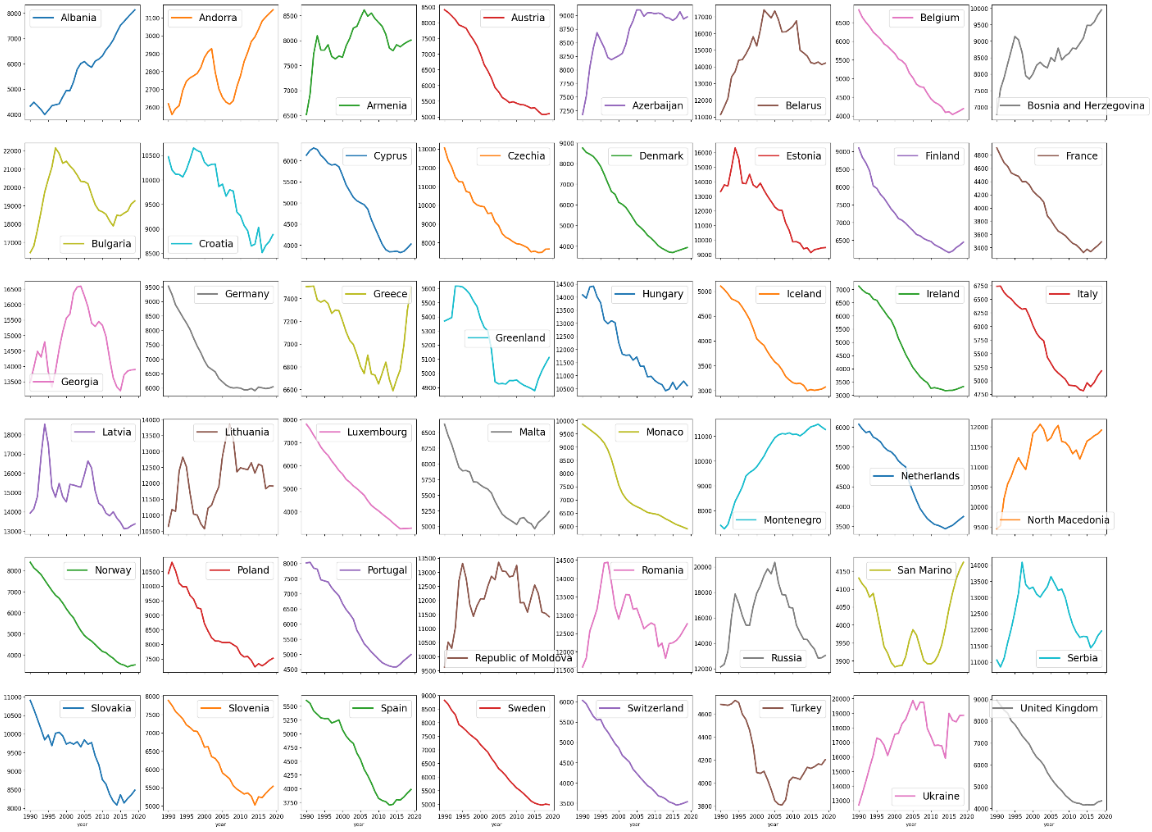


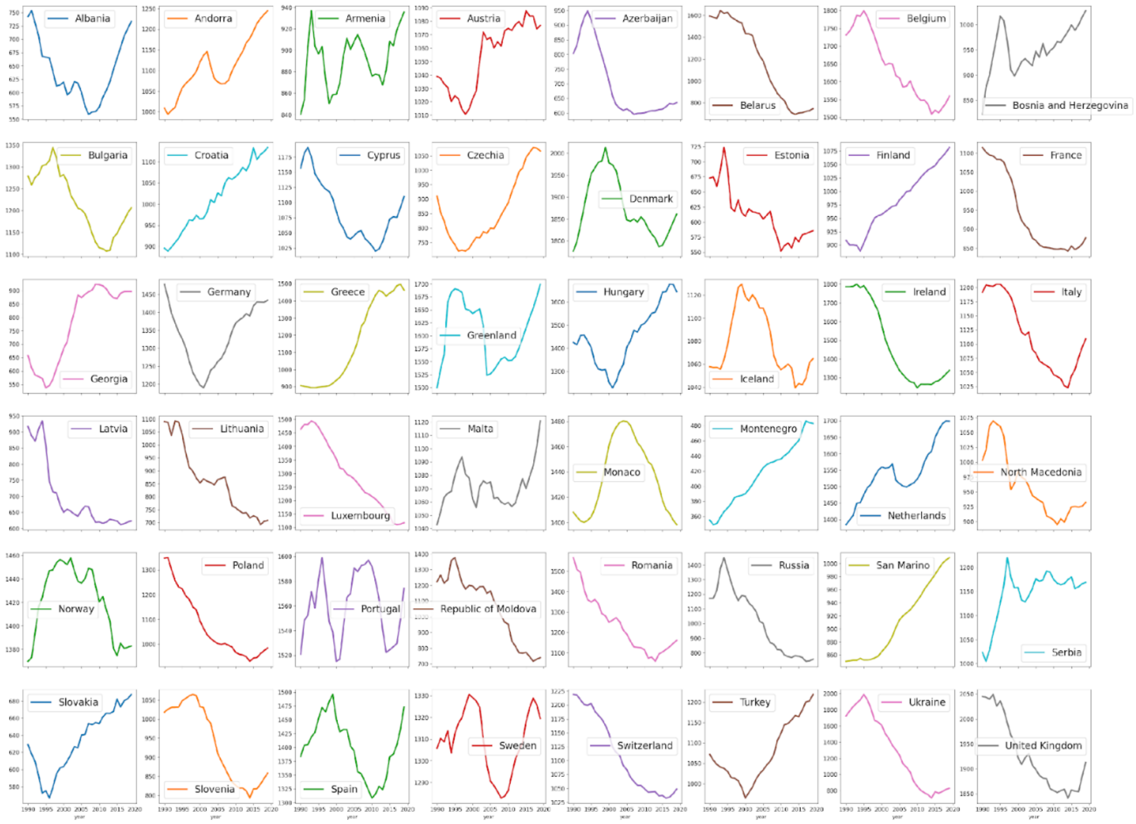
Figure 12 Graphs of boxplots across countries for each NCD

To a better understanding of the variability in responses over time, the Figure 13 illustrates the trend component of the time series across countries for each NCD. This demonstrates how the values of each country evolved over time, revealing upward or downward movement. By focusing on the trend, it is possible to better understand the fundamental trajectory of the data, which is crucial to identify potential long-term changes.

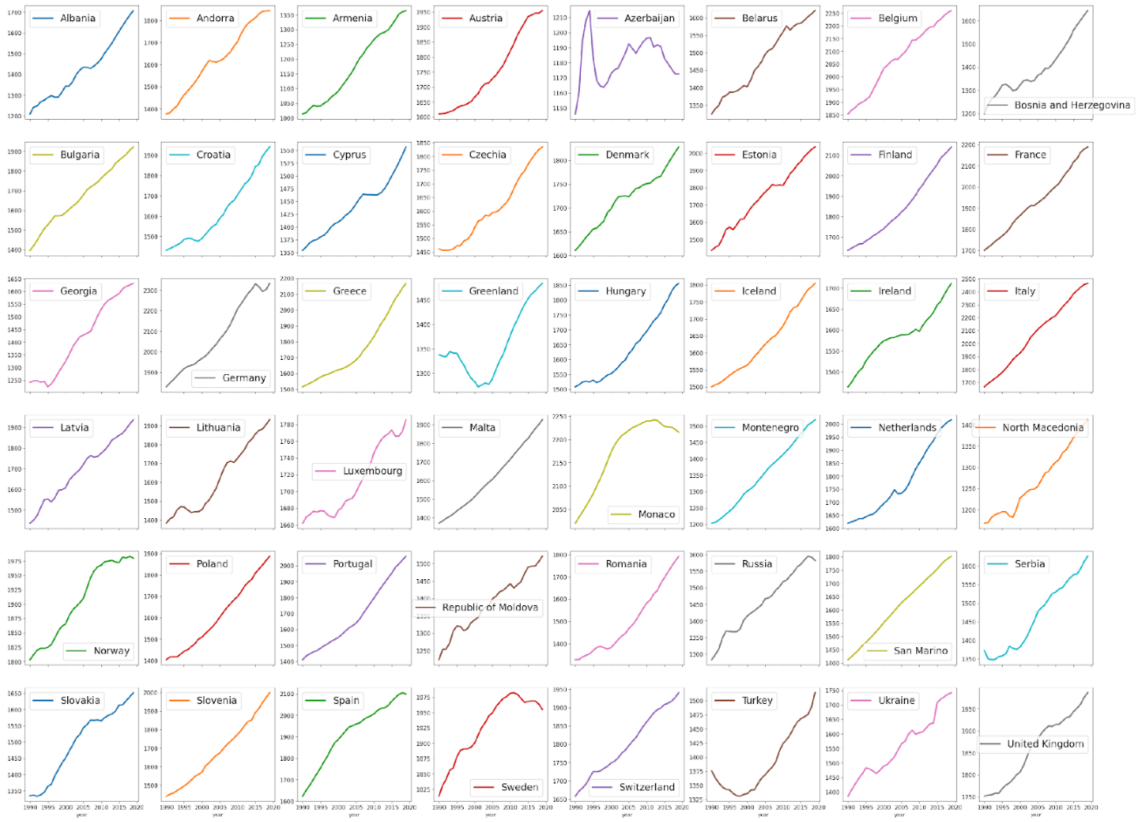
Trend in Cardiovascular Diseases



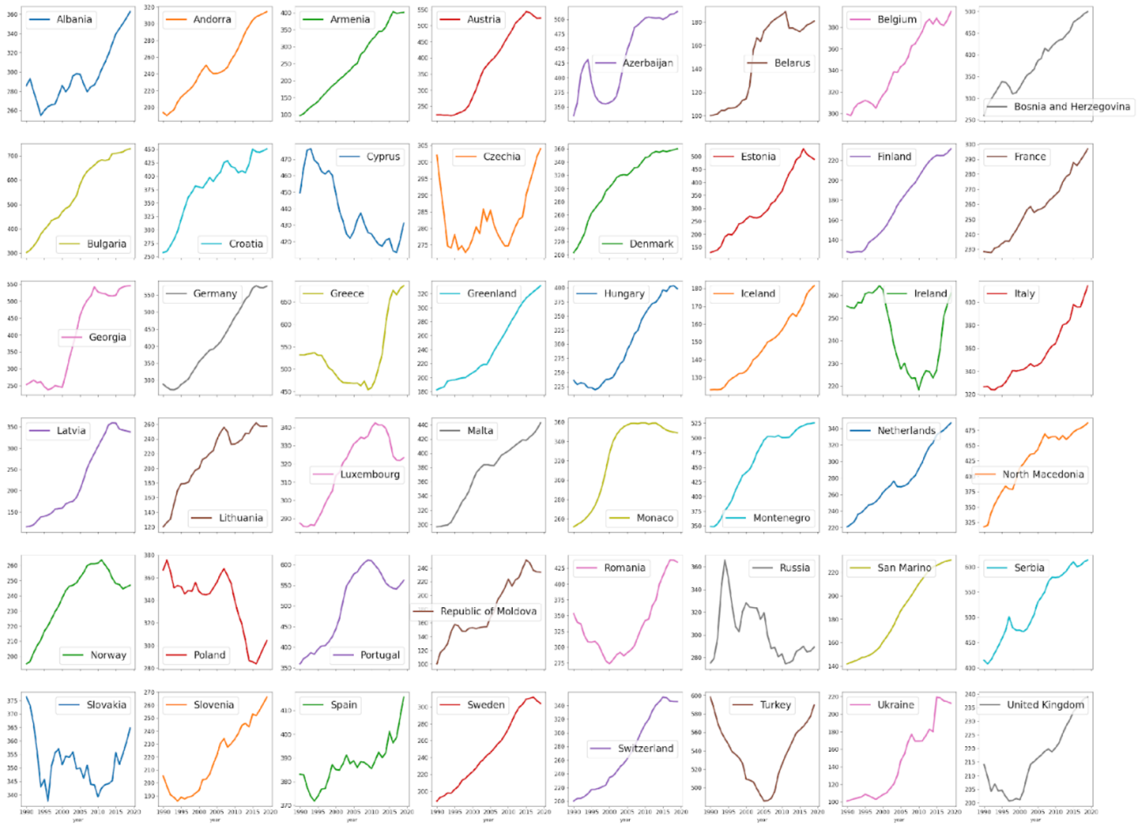
Trend in Respiratory Diseases



Trend in Neurological Disorders



Trend in Chronic Kidney Diseases



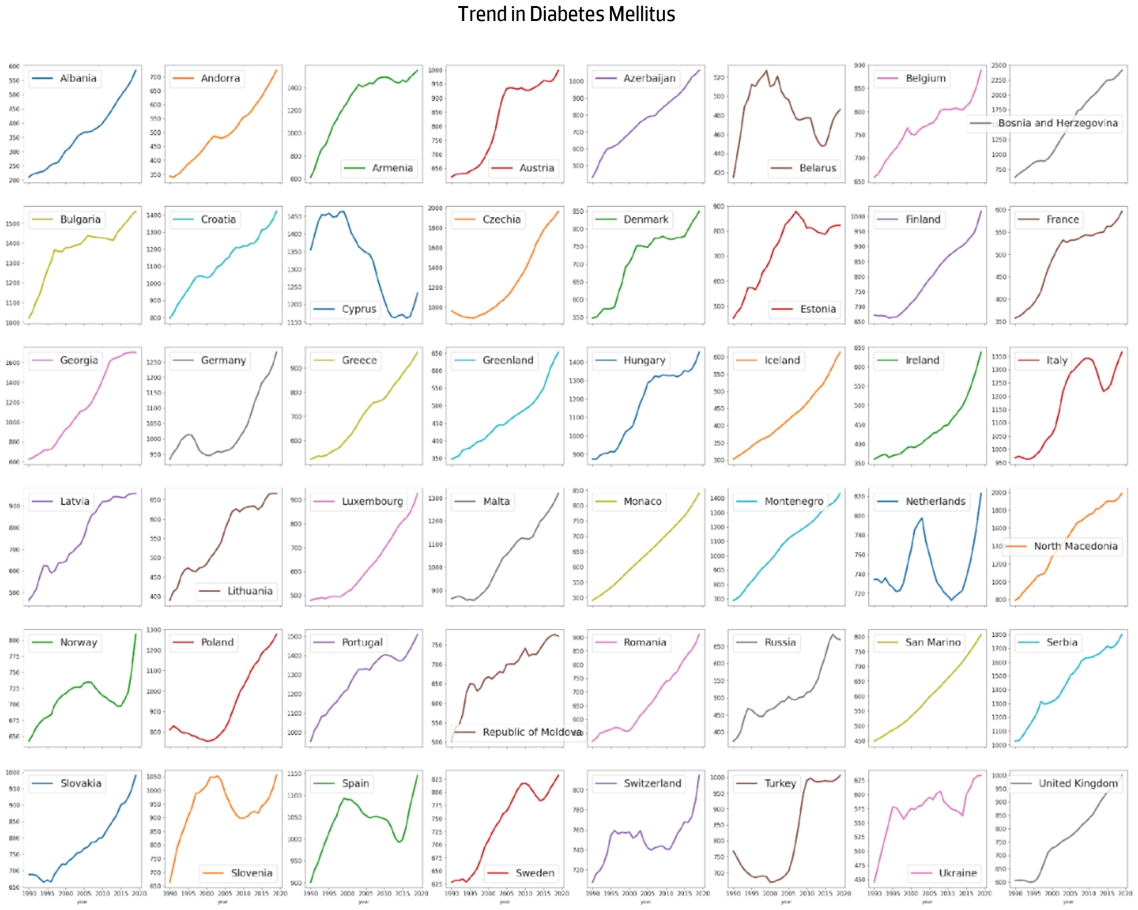


Figure 13 Trend component across countries for each NCD

5.2. Representation of the time series

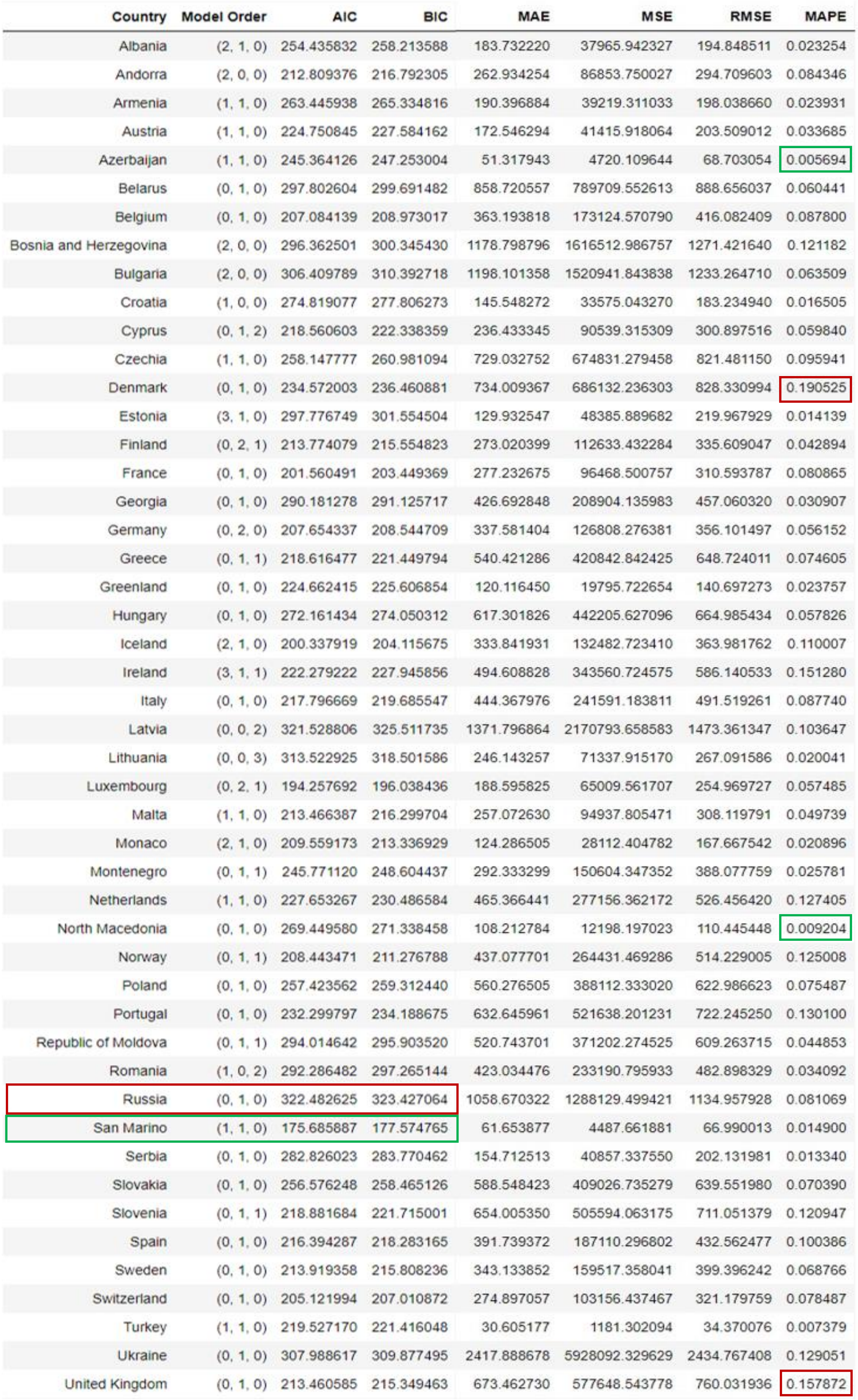
As it was mentioned in the previous chapter, the data of the DALYs rate for the 48 European countries was modelled using the autoARIMA function, automating the search for the best parameters. The next subchapters detail the results obtained from this process, for each non-communicable disease.

5.2.1. Cardiovascular Diseases

Table 3 summarizes the ARIMA model order of the last cross-validation fold for each country along with the AIC and BIC values and the values of various error metrics, MAE, MSE, RMSE and MAPE, for model predictions. From this table it is possible to observe varying model orders across countries, indicating the differences in the time series patterns observed in each region. The AIC and BIC are metrics used for model selection where lower values indicate a better fit of the model with respect to the complexity of the data. In this context these values are crucial when comparing ARIMA models to determine the best performing models in terms of predictive accuracy and model parsimony. The country with the lowest AIC and BIC values is San Marino, 175.685 and 177.57 respectively, suggesting a highly suitable ARIMA model, while the country with the highest values is Russia with an AIC value of 322.48 and a BIC value of 323.43, indicating a more complex or less efficient model fit.

Lower values of error metrics indicate a better model performance. Azerbaijan, Greenland, and North Macedonia are some of the countries that exhibit lower error values, in contrast to the United Kingdom and Denmark which have much higher values of MAPE.

Table 3 Selected Arima model for each country for Cardiovascular Diseases



Country	Model Order	AIC	BIC	MAE	MSE	RMSE	MAPE
Albania	(2, 1, 0)	254.435832	258.213588	183.732220	37965.942327	194.848511	0.023254
Andorra	(2, 0, 0)	212.809376	216.792305	262.934254	86853.750027	294.709603	0.084346
Armenia	(1, 1, 0)	263.445938	265.334816	190.396884	39219.311033	198.038660	0.023931
Austria	(1, 1, 0)	224.750845	227.584162	172.546294	41415.918064	203.509012	0.033685
Azerbaijan	(1, 1, 0)	245.364126	247.253004	51.317943	4720.109644	68.703054	0.005694
Belarus	(0, 1, 0)	297.802604	299.691482	858.720557	789709.552613	888.656037	0.060441
Belgium	(0, 1, 0)	207.084139	208.973017	363.193818	173124.570790	416.082409	0.087800
Bosnia and Herzegovina	(2, 0, 0)	296.362501	300.345430	1178.798796	1616512.986757	1271.421640	0.121182
Bulgaria	(2, 0, 0)	306.409789	310.392718	1198.101358	1520941.843838	1233.264710	0.063509
Croatia	(1, 0, 0)	274.819077	277.806273	145.548272	33575.043270	183.234940	0.016505
Cyprus	(0, 1, 2)	218.560603	222.338359	236.433345	90539.315309	300.897516	0.059840
Czechia	(1, 1, 0)	258.147777	260.981094	729.032752	674831.279458	821.481150	0.095941
Denmark	(0, 1, 0)	234.572003	236.460881	734.009367	686132.236303	828.330994	0.190525
Estonia	(3, 1, 0)	297.776749	301.554504	129.932547	48385.889682	219.967929	0.014139
Finland	(0, 2, 1)	213.774079	215.554823	273.020399	112633.432284	335.609047	0.042894
France	(0, 1, 0)	201.560491	203.449369	277.232675	96468.500757	310.593787	0.080865
Georgia	(0, 1, 0)	290.181278	291.125717	426.692848	208904.135983	457.060320	0.030907
Germany	(0, 2, 0)	207.654337	208.544709	337.581404	126808.276381	356.101497	0.056152
Greece	(0, 1, 1)	218.616477	221.449794	540.421286	420842.842425	648.724011	0.074605
Greenland	(0, 1, 0)	224.662415	225.606854	120.116450	19795.722654	140.697273	0.023757
Hungary	(0, 1, 0)	272.161434	274.050312	617.301826	442205.627096	664.985434	0.057826
Iceland	(2, 1, 0)	200.337919	204.115675	333.841931	132482.723410	363.981762	0.110007
Ireland	(3, 1, 1)	222.279222	227.945856	494.608828	343560.724575	586.140533	0.151280
Italy	(0, 1, 0)	217.796669	219.685547	444.367976	241591.183811	491.519261	0.087740
Latvia	(0, 0, 2)	321.528806	325.511735	1371.796864	2170793.658583	1473.361347	0.103647
Lithuania	(0, 0, 3)	313.522925	318.501586	246.143257	71337.915170	267.091586	0.020041
Luxembourg	(0, 2, 1)	194.257692	196.038436	188.595825	65009.561707	254.969727	0.057485
Malta	(1, 1, 0)	213.466387	216.299704	257.072630	94937.805471	308.119791	0.049739
Monaco	(2, 1, 0)	209.559173	213.336929	124.286505	28112.404782	167.667542	0.020896
Montenegro	(0, 1, 1)	245.771120	248.604437	292.333299	150604.347352	388.077759	0.025781
Netherlands	(1, 1, 0)	227.653267	230.486584	465.366441	277156.362172	526.456420	0.127405
North Macedonia	(0, 1, 0)	269.449580	271.338458	108.212784	12198.197023	110.445448	0.009204
Norway	(0, 1, 1)	208.443471	211.276788	437.077701	264431.469286	514.229005	0.125008
Poland	(0, 1, 0)	257.423562	259.312440	560.276505	388112.333020	622.986623	0.075487
Portugal	(0, 1, 0)	232.299797	234.188675	632.645961	521638.201231	722.245250	0.130100
Republic of Moldova	(0, 1, 1)	294.014642	295.903520	520.743701	371202.274525	609.263715	0.044853
Romania	(1, 0, 2)	292.286482	297.265144	423.034476	233190.795933	482.898329	0.034092
Russia	(0, 1, 0)	322.482625	323.427064	1058.670322	1288129.499421	1134.957928	0.081069
San Marino	(1, 1, 0)	175.685887	177.574765	61.653877	4487.661881	66.990013	0.014900
Serbia	(0, 1, 0)	282.826023	283.770462	154.712513	40857.337550	202.131981	0.013340
Slovakia	(0, 1, 0)	256.576248	258.465126	588.548423	409026.735279	639.551980	0.070390
Slovenia	(0, 1, 1)	218.881684	221.715001	654.005350	505594.063175	711.051379	0.120947
Spain	(0, 1, 0)	216.394287	218.283165	391.739372	187110.296802	432.562477	0.100386
Sweden	(0, 1, 0)	213.919358	215.808236	343.133852	159517.358041	399.396242	0.068766
Switzerland	(0, 1, 0)	205.121994	207.010872	274.897057	103156.437467	321.179759	0.078487
Turkey	(1, 1, 0)	219.527170	221.416048	30.605177	1181.302094	34.370076	0.007379
Ukraine	(0, 1, 0)	307.988617	309.877495	2417.888678	5928092.329629	2434.767408	0.129051
United Kingdom	(0, 1, 0)	213.460585	215.349463	673.462730	577648.543778	760.031936	0.157872

Figure 14 helps to visually understand model's performance, with the boxplot's outliers showing in which countries the models performed worse. Latvia and Ukraine are outliers across three metrics, MAE, MSE and RMSE. While Bosnia and Herzegovina is an outlier in the boxplots for MSE and RMSE. And Bulgaria and Russia are outliers only in the boxplot for MSE. The boxplot for the MAPE metric shows no outliers and its median value is considered low, indicating that even though there are outliers for other metrics, the model generally performed well across most countries.

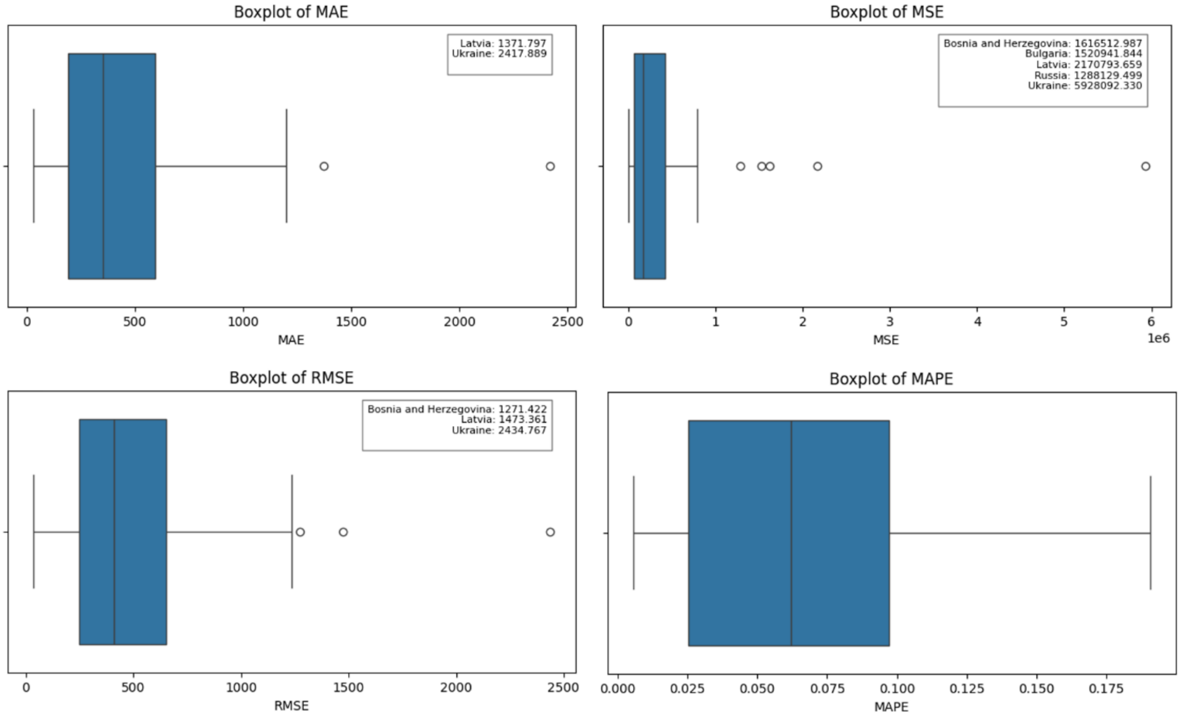


Figure 14 Boxplots for MAE, MSE, RMSE, and MAPE for Cardiovascular Diseases

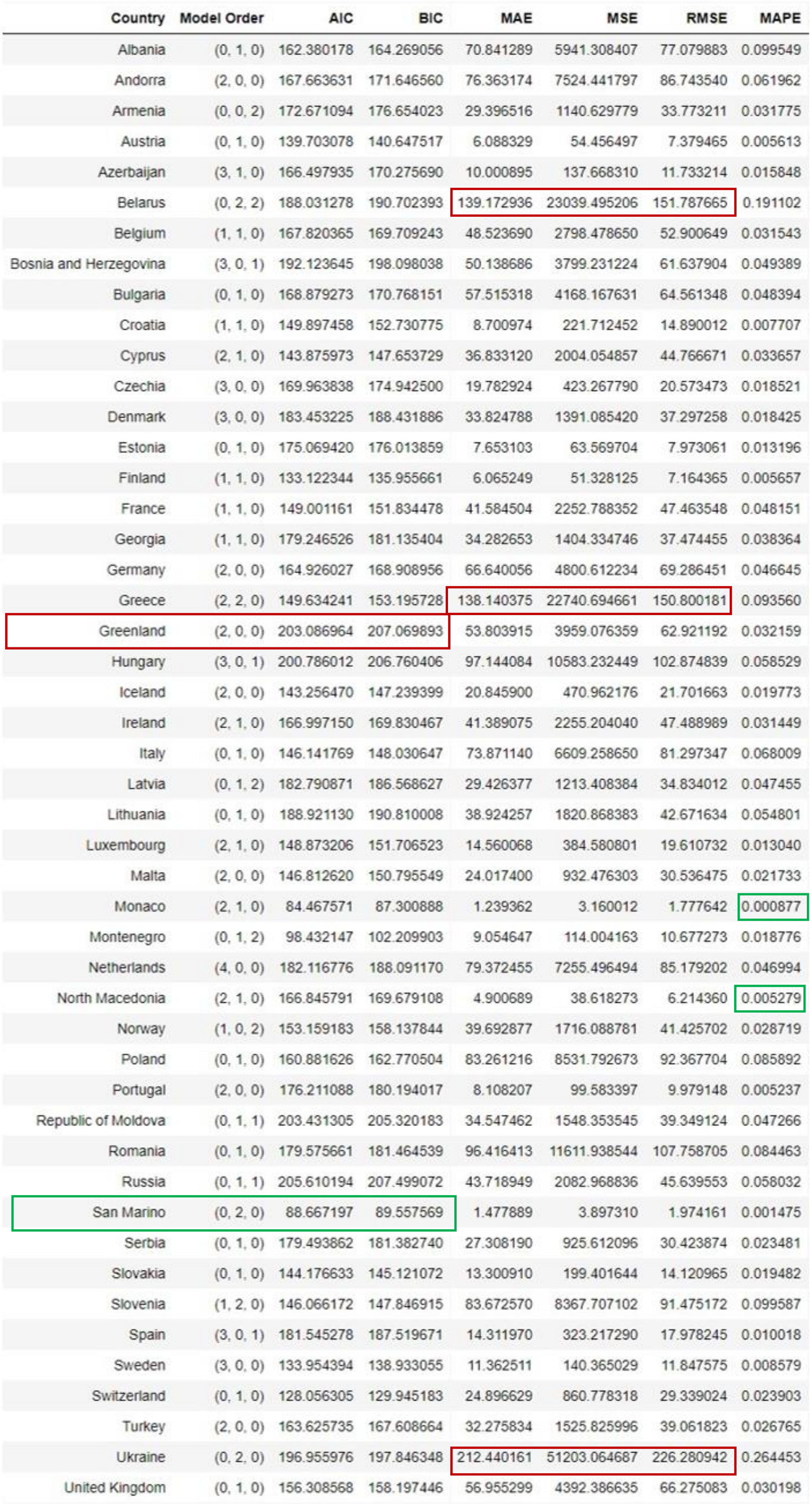
Combining insights from the previous Table 3 and Figure 14, it is possible to provide a deeper analysis of each model's performance. For instance, when analyzing AIC/BIC and error metrics values, San Marino has very low AIC and BIC values suggesting a better fit of the model, this is reflected by relatively low errors across all metrics. On the other hand, countries like Russia and Latvia, which had higher values of AIC and BIC, also show larger prediction errors. When comparing model complexity and its predictive accuracy, the results show that countries with more complex models, meaning with higher p, d, q order, for example Ireland with a (3, 1, 0) model, not always exhibit lower prediction errors. This emphasizes the importance of balancing model complexity and performance.

5.2.2. Chronic Respiratory Diseases

Table 4 summarizes the ARIMA model order of the last cross-validation fold for each country along with the AIC and BIC values and the values of various error metrics, MAE, MSE, RMSE and MAPE, for model predictions. The different ARIMA model orders indicate differences in the patterns and trends among countries. High AIC and BIC values, for example, Greenland with a value of AIC of 203.08 and BIC of 207.06, might indicate that the data is more difficult to model. Countries like Belarus, Greece, and Ukraine have higher MAE, MSE, and RMSE values which indicates that the models struggle to accurately forecast the data. This might be due to unstable trends, or inconsistencies in reporting data. Monaco is one of the countries that have lower error metric values, with a MAPE value of 0.0008 indicating a better model performance.

Figure 15 helps to visually understand model's performance, with the boxplot's outliers showing in which countries the models performed worse. Belarus and Ukraine are outliers across all metrics. While Greece is an outlier in the boxplots for MAE, MSE and RMSE. And Romania is an outlier only in the boxplot for MSE. Although the boxplot for the MAPE metric shows two outliers, its median value is considered very low, indicating that even though there are outliers for all four metrics, the models generally performed well across most countries.

Table 4 Selected Arima model for each country for Chronic Respiratory Diseases



Country	Model Order	AIC	BIC	MAE	MSE	RMSE	MAPE
Albania	(0, 1, 0)	162.380178	164.269056	70.841289	5941.308407	77.079883	0.099549
Andorra	(2, 0, 0)	167.663631	171.646560	76.363174	7524.441797	86.743540	0.061962
Armenia	(0, 0, 2)	172.671094	176.654023	29.396516	1140.629779	33.773211	0.031775
Austria	(0, 1, 0)	139.703078	140.647517	6.088329	54.456497	7.379465	0.005613
Azerbaijan	(3, 1, 0)	166.497935	170.275690	10.000895	137.668310	11.733214	0.015848
Belarus	(0, 2, 2)	188.031278	190.702393	139.172936	23039.495206	151.787665	0.191102
Belgium	(1, 1, 0)	167.820365	169.709243	48.523690	2798.478650	52.900649	0.031543
Bosnia and Herzegovina	(3, 0, 1)	192.123645	198.098038	50.138686	3799.231224	61.637904	0.049389
Bulgaria	(0, 1, 0)	168.879273	170.768151	57.515318	4168.167631	64.561348	0.048394
Croatia	(1, 1, 0)	149.897458	152.730775	8.700974	221.712452	14.890012	0.007707
Cyprus	(2, 1, 0)	143.875973	147.653729	36.833120	2004.054857	44.766671	0.033657
Czechia	(3, 0, 0)	169.963838	174.942500	19.782924	423.267790	20.573473	0.018521
Denmark	(3, 0, 0)	183.453225	188.431886	33.824788	1391.085420	37.297258	0.018425
Estonia	(0, 1, 0)	175.069420	176.013859	7.653103	63.569704	7.973061	0.013196
Finland	(1, 1, 0)	133.122344	135.955661	6.065249	51.328125	7.164365	0.005657
France	(1, 1, 0)	149.001161	151.834478	41.584504	2252.788352	47.463548	0.048151
Georgia	(1, 1, 0)	179.246526	181.135404	34.282653	1404.334746	37.474455	0.038364
Germany	(2, 0, 0)	164.926027	168.908956	66.640056	4800.612234	69.286451	0.046645
Greece	(2, 2, 0)	149.634241	153.195728	138.140375	22740.694661	150.800181	0.093560
Greenland	(2, 0, 0)	203.086964	207.069893	53.803915	3959.076359	62.921192	0.032159
Hungary	(3, 0, 1)	200.786012	206.760406	97.144084	10583.232449	102.874839	0.058529
Iceland	(2, 0, 0)	143.256470	147.239399	20.845900	470.962176	21.701663	0.019773
Ireland	(2, 1, 0)	166.997150	169.830467	41.389075	2255.204040	47.488989	0.031449
Italy	(0, 1, 0)	146.141769	148.030647	73.871140	6609.258650	81.297347	0.068009
Latvia	(0, 1, 2)	182.790871	186.568627	29.426377	1213.408384	34.834012	0.047455
Lithuania	(0, 1, 0)	188.921130	190.810008	38.924257	1820.868383	42.671634	0.054801
Luxembourg	(2, 1, 0)	148.873206	151.706523	14.560068	384.580801	19.610732	0.013040
Malta	(2, 0, 0)	146.812620	150.795549	24.017400	932.476303	30.536475	0.021733
Monaco	(2, 1, 0)	84.467571	87.300888	1.239362	3.160012	1.777642	0.000877
Montenegro	(0, 1, 2)	98.432147	102.209903	9.054647	114.004163	10.677273	0.018776
Netherlands	(4, 0, 0)	182.116776	188.091170	79.372455	7255.496494	85.179202	0.046994
North Macedonia	(2, 1, 0)	166.845791	169.679108	4.900689	38.618273	6.214360	0.005279
Norway	(1, 0, 2)	153.159183	158.137844	39.692877	1716.088781	41.425702	0.028719
Poland	(0, 1, 0)	160.881626	162.770504	83.261216	8531.792673	92.367704	0.085892
Portugal	(2, 0, 0)	176.211088	180.194017	8.108207	99.583397	9.979148	0.005237
Republic of Moldova	(0, 1, 1)	203.431305	205.320183	34.547462	1548.353545	39.349124	0.047266
Romania	(0, 1, 0)	179.575661	181.464539	96.416413	11611.938544	107.758705	0.084463
Russia	(0, 1, 1)	205.610194	207.499072	43.718949	2082.968836	45.639553	0.058032
San Marino	(0, 2, 0)	88.667197	89.557569	1.477889	3.897310	1.974161	0.001475
Serbia	(0, 1, 0)	179.493862	181.382740	27.308190	925.612096	30.423874	0.023481
Slovakia	(0, 1, 0)	144.176633	145.121072	13.300910	199.401644	14.120965	0.019482
Slovenia	(1, 2, 0)	146.066172	147.846915	83.672570	8367.707102	91.475172	0.099587
Spain	(3, 0, 1)	181.545278	187.519671	14.311970	323.217290	17.978245	0.010018
Sweden	(3, 0, 0)	133.954394	138.933055	11.362511	140.365029	11.847575	0.008579
Switzerland	(0, 1, 0)	128.056305	129.945183	24.896629	860.778318	29.339024	0.023903
Turkey	(2, 0, 0)	163.625735	167.608664	32.275834	1525.825996	39.061823	0.026765
Ukraine	(0, 2, 0)	196.955976	197.846348	212.440161	51203.064687	226.280942	0.264453
United Kingdom	(0, 1, 0)	156.308568	158.197446	56.955299	4392.386635	66.275083	0.030198

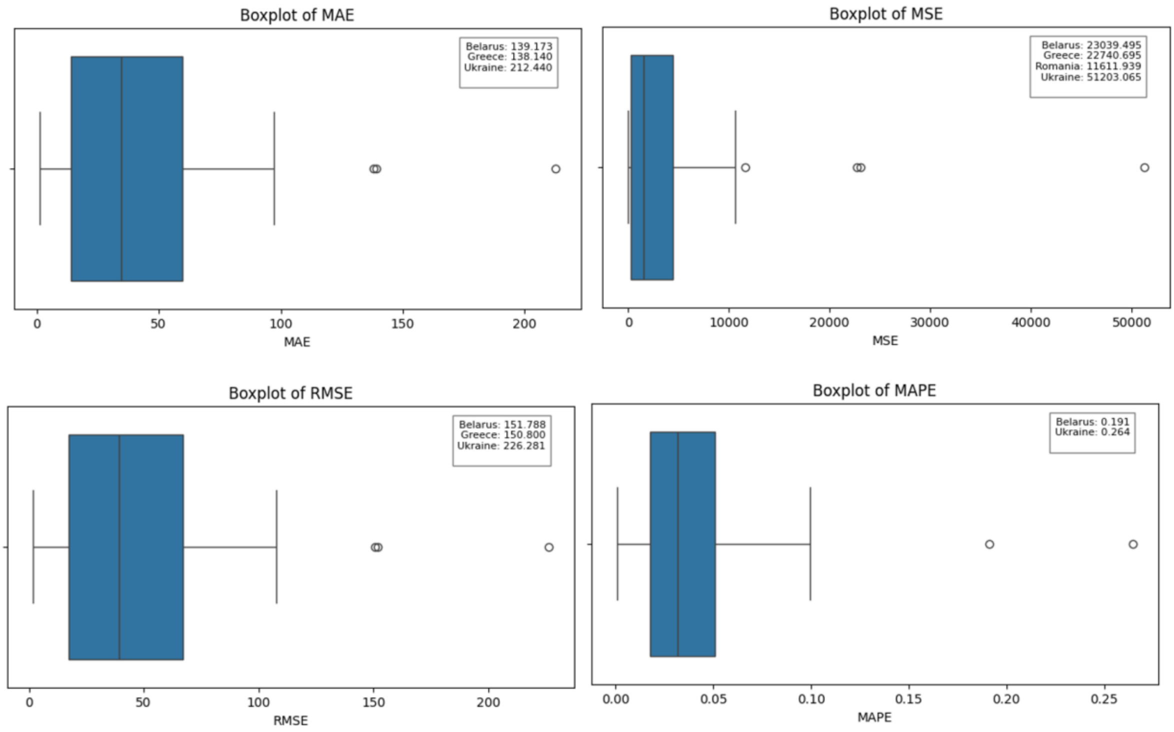
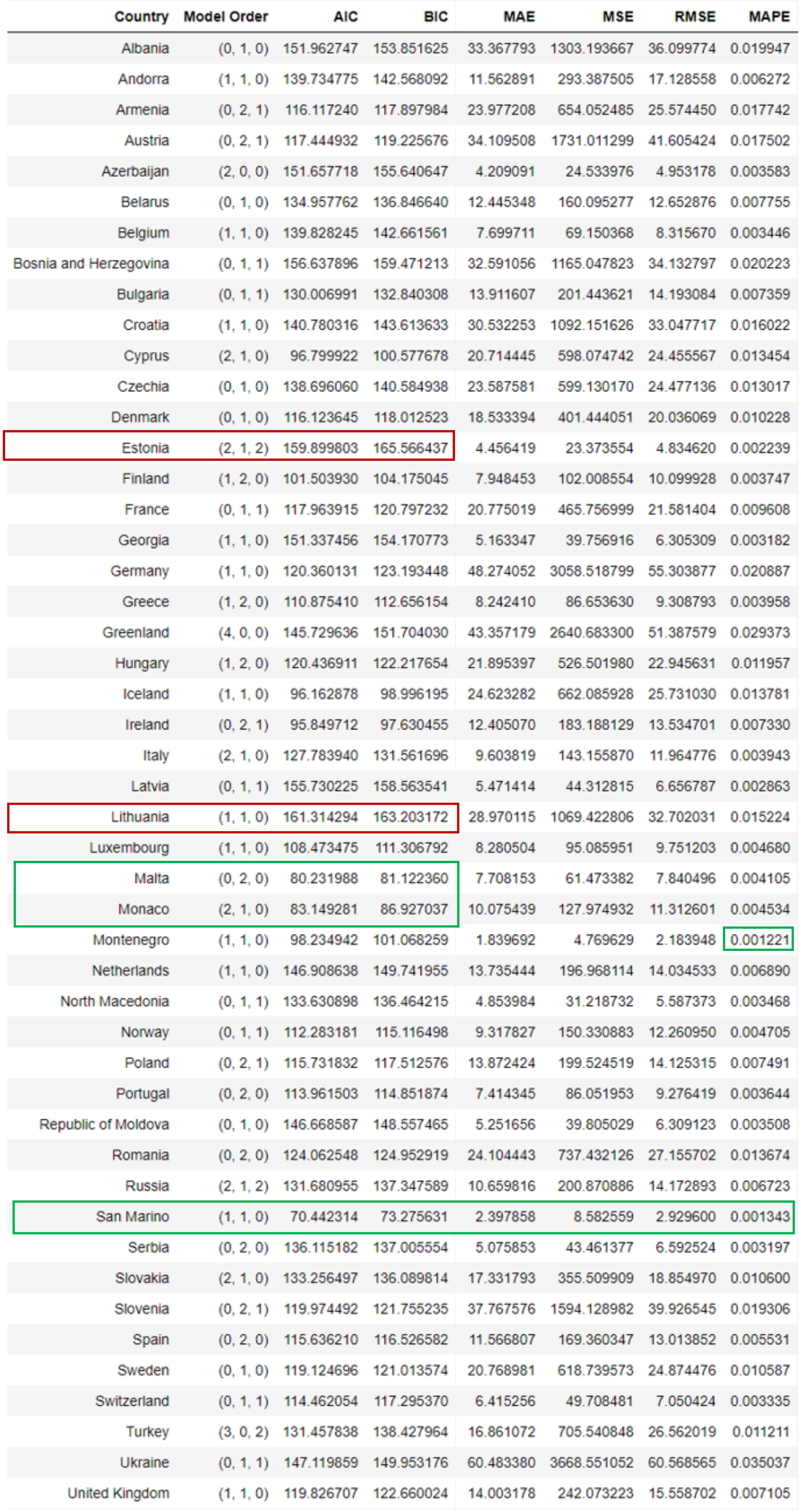


Figure 15 Boxplots for MAE, MSE, RMSE, and MAPE for Chronic Respiratory Diseases

5.2.3. **Neurological Disorders**

Table 5 summarizes the ARIMA model order of the last cross-validation fold for each country along with the AIC and BIC values and the values of various error metrics, MAE, MSE, RMSE and MAPE, for model predictions. The different ARIMA model orders indicate differences in the patterns and trends among countries. Low AIC and BIC values, for example, Malta 80.23 and 81.12 respectively, might suggest a good balance between model fit and complexity. In the other hand, countries like Estonia and Lithuania have higher AIC and BIC values indicating that the data is more challenging to model accurately. Germany and Ukraine have high MAE, MSE, and RMSE values which implies that the models struggle to accurately forecast the data. Montenegro and San Marino are some of the countries that have lower error metric values, with a MAPE value of 0.0012 and 0.0013, respectively, indicating that the data is likely more consistent and easier to forecast.

Table 5 Selected Arima model for each country for Neurological Disorders



Country	Model Order	AIC	BIC	MAE	MSE	RMSE	MAPE
Albania	(0, 1, 0)	151.962747	153.851625	33.367793	1303.193667	36.099774	0.019947
Andorra	(1, 1, 0)	139.734775	142.568092	11.562891	293.387505	17.128558	0.006272
Armenia	(0, 2, 1)	116.117240	117.897984	23.977208	654.052485	25.574450	0.017742
Austria	(0, 2, 1)	117.444932	119.225676	34.109508	1731.011299	41.605424	0.017502
Azerbaijan	(2, 0, 0)	151.657718	155.640647	4.209091	24.533976	4.953178	0.003583
Belarus	(0, 1, 0)	134.957762	136.846640	12.445348	160.095277	12.652876	0.007755
Belgium	(1, 1, 0)	139.828245	142.661561	7.699711	69.150368	8.315670	0.003446
Bosnia and Herzegovina	(0, 1, 1)	156.637896	159.471213	32.591056	1165.047823	34.132797	0.020223
Bulgaria	(0, 1, 1)	130.006991	132.840308	13.911607	201.443621	14.193084	0.007359
Croatia	(1, 1, 0)	140.780316	143.613633	30.532253	1092.151626	33.047717	0.016022
Cyprus	(2, 1, 0)	96.799922	100.577678	20.714445	598.074742	24.455567	0.013454
Czechia	(0, 1, 0)	138.696060	140.584938	23.587581	599.130170	24.477136	0.013017
Denmark	(0, 1, 0)	116.123645	118.012523	18.533394	401.444051	20.036069	0.010228
Estonia	(2, 1, 2)	159.899803	165.566437	4.456419	23.373554	4.834620	0.002239
Finland	(1, 2, 0)	101.503930	104.175045	7.948453	102.008554	10.099928	0.003747
France	(0, 1, 1)	117.963915	120.797232	20.775019	465.756999	21.581404	0.009608
Georgia	(1, 1, 0)	151.337456	154.170773	5.163347	39.756916	6.305309	0.003182
Germany	(1, 1, 0)	120.360131	123.193448	48.274052	3058.518799	55.303877	0.020887
Greece	(1, 2, 0)	110.875410	112.656154	8.242410	86.653630	9.308793	0.003958
Greenland	(4, 0, 0)	145.729636	151.704030	43.357179	2640.683300	51.387579	0.029373
Hungary	(1, 2, 0)	120.436911	122.217654	21.895397	526.501980	22.945631	0.011957
Iceland	(1, 1, 0)	96.162878	98.996195	24.623282	662.085928	25.731030	0.013781
Ireland	(0, 2, 1)	95.849712	97.630455	12.405070	183.188129	13.534701	0.007330
Italy	(2, 1, 0)	127.783940	131.561696	9.603819	143.155870	11.964776	0.003943
Latvia	(0, 1, 1)	155.730225	158.563541	5.471414	44.312815	6.656787	0.002863
Lithuania	(1, 1, 0)	161.314294	163.203172	28.970115	1069.422806	32.702031	0.015224
Luxembourg	(1, 1, 0)	108.473475	111.306792	8.280504	95.085951	9.751203	0.004680
Malta	(0, 2, 0)	80.231988	81.122360	7.708153	61.473382	7.840496	0.004105
Monaco	(2, 1, 0)	83.149281	86.927037	10.075439	127.974932	11.312601	0.004534
Montenegro	(1, 1, 0)	98.234942	101.068259	1.839692	4.769629	2.183948	0.001221
Netherlands	(1, 1, 0)	146.908638	149.741955	13.735444	196.968114	14.034533	0.006890
North Macedonia	(0, 1, 1)	133.630898	136.464215	4.853984	31.218732	5.587373	0.003468
Norway	(0, 1, 1)	112.283181	115.116498	9.317827	150.330883	12.260950	0.004705
Poland	(0, 2, 1)	115.731832	117.512576	13.872424	199.524519	14.125315	0.007491
Portugal	(0, 2, 0)	113.961503	114.851874	7.414345	86.051953	9.276419	0.003644
Republic of Moldova	(0, 1, 0)	146.668587	148.557465	5.251656	39.805029	6.309123	0.003508
Romania	(0, 2, 0)	124.062548	124.952919	24.104443	737.432126	27.155702	0.013674
Russia	(2, 1, 2)	131.680955	137.347589	10.659816	200.870886	14.172893	0.006723
San Marino	(1, 1, 0)	70.442314	73.275631	2.397858	8.582559	2.929600	0.001343
Serbia	(0, 2, 0)	136.115182	137.005554	5.075853	43.461377	6.592524	0.003197
Slovakia	(2, 1, 0)	133.256497	136.089814	17.331793	355.509909	18.854970	0.010600
Slovenia	(0, 2, 1)	119.974492	121.755235	37.767576	1594.128982	39.926545	0.019306
Spain	(0, 2, 0)	115.636210	116.526582	11.566807	169.360347	13.013852	0.005531
Sweden	(0, 1, 0)	119.124696	121.013574	20.768981	618.739573	24.874476	0.010587
Switzerland	(0, 1, 1)	114.462054	117.295370	6.415256	49.708481	7.050424	0.003335
Turkey	(3, 0, 2)	131.457838	138.427964	16.861072	705.540848	26.562019	0.011211
Ukraine	(0, 1, 1)	147.119859	149.953176	60.483380	3668.551052	60.568565	0.035037
United Kingdom	(1, 1, 0)	119.826707	122.660024	14.003178	242.073223	15.558702	0.007105

Figure 16 helps to visually understand model's performance, with the boxplot's outliers showing in which countries the models performed worse. Ukraine is an outlier across all metrics. While Germany is an outlier in the boxplots for MAE, MSE and RMSE. Greenland is an outlier in the boxplots for MSE and RMSE. Austria and Slovenia are outliers for the MSE metric. Similarly to the Chronic Respiratory Diseases, the boxplot for the MAPE metric shows two outliers, but its median value can be considered low, which indicates that even though there are outliers for all four metrics, the models generally performed well across most countries.

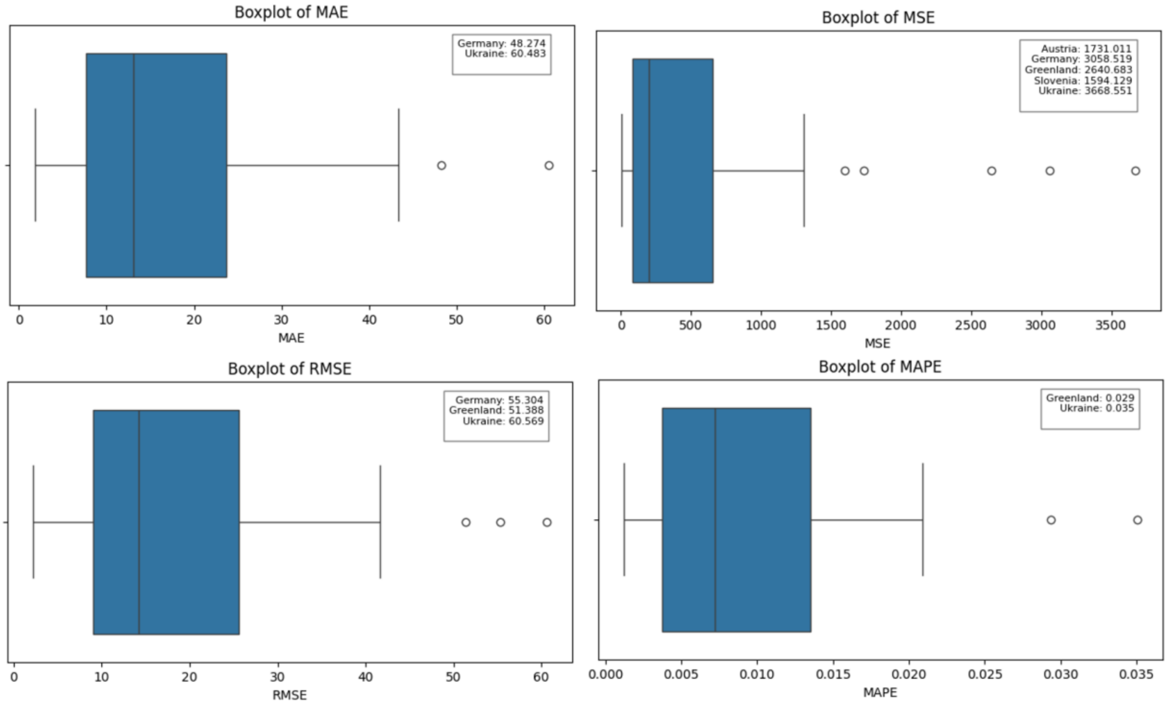


Figure 16 Boxplots for MAE, MSE, RMSE, and MAPE for Neurological Disorders

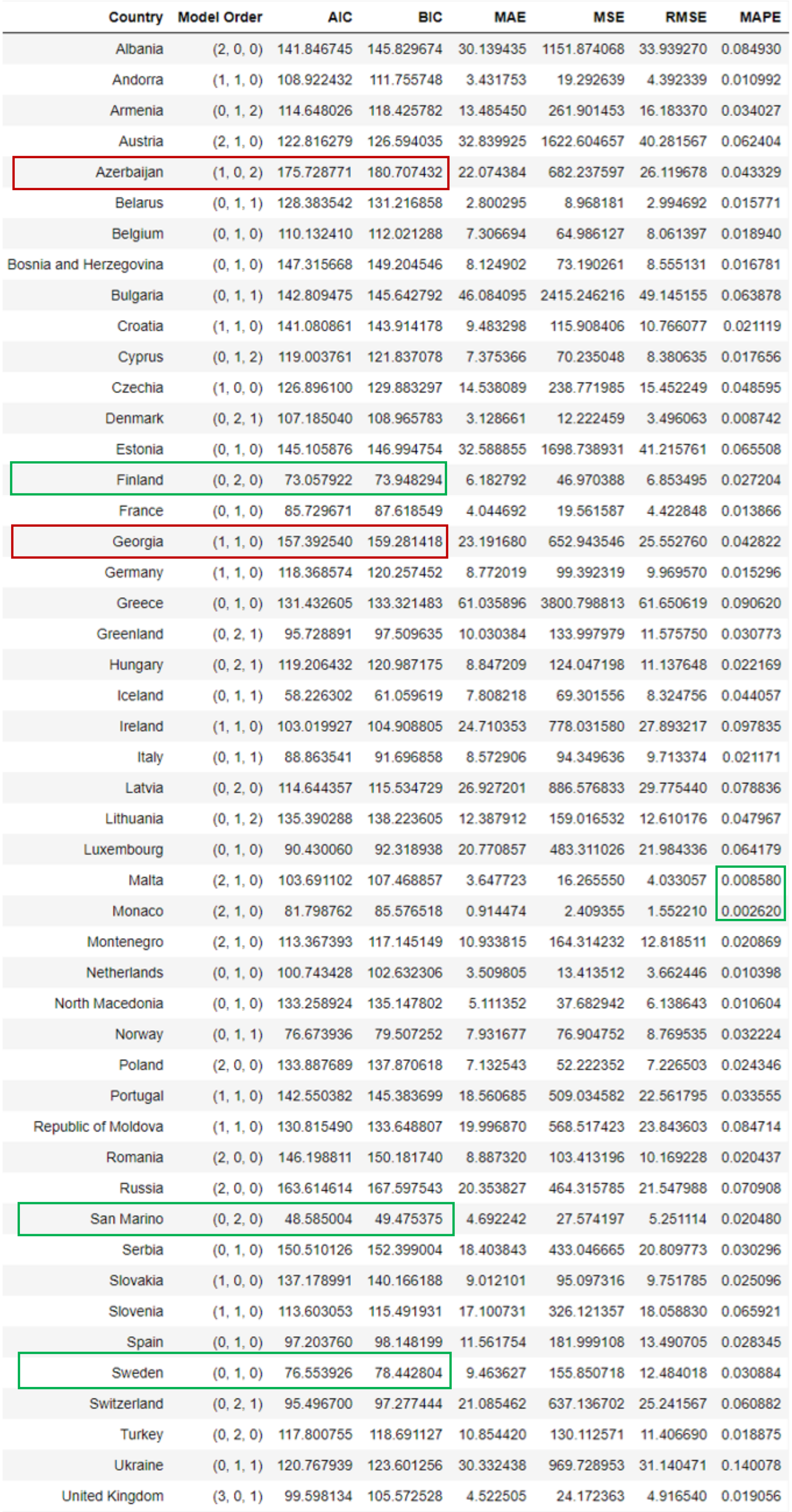
5.2.4. **Chronic Kidney Diseases**

Table 6 summarizes the ARIMA model order of the last cross-validation fold for each country along with the AIC and BIC values and the values of various error metrics, MAE, MSE, RMSE and MAPE, for model predictions. The different ARIMA model orders indicate differences in the patterns and trends among countries. High AIC and BIC values, for example, Azerbaijan with AIC value of 175.73 and BIC of 180.77, and Georgia AIC of 157.39 and BIC of 159.28 suggest their models are nor as optimal as the other countries possible due to more complex data trends. In the contrast, countries like Sweden have lower AIC and BIC values, 76.55 and 78.44 respectively, indicating a more effective model. Greece and Bulgaria show high MAE, MSE, RMSE, and MAPE values which implies that the data is more difficult to accurately forecast.

Malta and Monaco are some of the countries that have lower error metric values, with a MAPE value of 0.0085 and 0.0026, respectively, indicating highly accurate model forecasts.

Figure 17 helps to visually understand model's performance, with the boxplot's outliers showing in which countries the models performed worse. Bulgaria and Greece are outliers in the boxplots for MAE, MSE and RMSE. Austria and Estonia are outliers for the MSE metric. Ukraine is the only outliers in the boxplot for the MAPE metric, but similarly to the NCDs discussed before, its median value can be considered low, which indicates that even though there are outliers for all four metrics, the models generally performed well across most countries.

Table 6 Selected Arima model for each country for Chronic Kidney Diseases



Country	Model Order	AIC	BIC	MAE	MSE	RMSE	MAPE
Albania	(2, 0, 0)	141.846745	145.829674	30.139435	1151.874068	33.939270	0.084930
Andorra	(1, 1, 0)	108.922432	111.755748	3.431753	19.292639	4.392339	0.010992
Armenia	(0, 1, 2)	114.648026	118.425782	13.485450	261.901453	16.183370	0.034027
Austria	(2, 1, 0)	122.816279	126.594035	32.839925	1622.604657	40.281567	0.062404
Azerbaijan	(1, 0, 2)	175.728771	180.707432	22.074384	682.237597	26.119678	0.043329
Belarus	(0, 1, 1)	128.383542	131.216858	2.800295	8.968181	2.994692	0.015771
Belgium	(0, 1, 0)	110.132410	112.021288	7.306694	64.986127	8.061397	0.018940
Bosnia and Herzegovina	(0, 1, 0)	147.315668	149.204546	8.124902	73.190261	8.555131	0.016781
Bulgaria	(0, 1, 1)	142.809475	145.642792	46.084095	2415.246216	49.145155	0.063878
Croatia	(1, 1, 0)	141.080861	143.914178	9.483298	115.908406	10.766077	0.021119
Cyprus	(0, 1, 2)	119.003761	121.837078	7.375366	70.235048	8.380635	0.017656
Czechia	(1, 0, 0)	126.896100	129.883297	14.538089	238.771985	15.452249	0.048595
Denmark	(0, 2, 1)	107.185040	108.965783	3.128661	12.222459	3.496063	0.008742
Estonia	(0, 1, 0)	145.105876	146.994754	32.588855	1698.738931	41.215761	0.065508
Finland	(0, 2, 0)	73.057922	73.948294	6.182792	46.970388	6.853495	0.027204
France	(0, 1, 0)	85.729671	87.618549	4.044692	19.561587	4.422848	0.013866
Georgia	(1, 1, 0)	157.392540	159.281418	23.191680	652.943546	25.552760	0.042822
Germany	(1, 1, 0)	118.368574	120.257452	8.772019	99.392319	9.969570	0.015296
Greece	(0, 1, 0)	131.432605	133.321483	61.035896	3800.798813	61.650619	0.090620
Greenland	(0, 2, 1)	95.728891	97.509635	10.030384	133.997979	11.575750	0.030773
Hungary	(0, 2, 1)	119.206432	120.987175	8.847209	124.047198	11.137648	0.022169
Iceland	(0, 1, 1)	58.226302	61.059619	7.808218	69.301556	8.324756	0.044057
Ireland	(1, 1, 0)	103.019927	104.908805	24.710353	778.031580	27.893217	0.097835
Italy	(0, 1, 1)	88.863541	91.696858	8.572906	94.349636	9.713374	0.021171
Latvia	(0, 2, 0)	114.644357	115.534729	26.927201	886.576833	29.775440	0.078836
Lithuania	(0, 1, 2)	135.390288	138.223605	12.387912	159.016532	12.610176	0.047967
Luxembourg	(0, 1, 0)	90.430060	92.318938	20.770857	483.311026	21.984336	0.064179
Malta	(2, 1, 0)	103.691102	107.468857	3.647723	16.265550	4.033057	0.008580
Monaco	(2, 1, 0)	81.798762	85.576518	0.914474	2.409355	1.552210	0.002620
Montenegro	(2, 1, 0)	113.367393	117.145149	10.933815	164.314232	12.818511	0.020869
Netherlands	(0, 1, 0)	100.743428	102.632306	3.509805	13.413512	3.662446	0.010398
North Macedonia	(0, 1, 0)	133.258924	135.147802	5.111352	37.682942	6.138643	0.010604
Norway	(0, 1, 1)	76.673936	79.507252	7.931677	76.904752	8.769535	0.032224
Poland	(2, 0, 0)	133.887689	137.870618	7.132543	52.222352	7.226503	0.024346
Portugal	(1, 1, 0)	142.550382	145.383699	18.560685	509.034582	22.561795	0.033555
Republic of Moldova	(1, 1, 0)	130.815490	133.648807	19.996870	568.517423	23.843603	0.084714
Romania	(2, 0, 0)	146.198811	150.181740	8.887320	103.413196	10.169228	0.020437
Russia	(2, 0, 0)	163.614614	167.597543	20.353827	464.315785	21.547988	0.070908
San Marino	(0, 2, 0)	48.585004	49.475375	4.692242	27.574197	5.251114	0.020480
Serbia	(0, 1, 0)	150.510126	152.399004	18.403843	433.046665	20.809773	0.030296
Slovakia	(1, 0, 0)	137.178991	140.166188	9.012101	95.097316	9.751785	0.025096
Slovenia	(1, 1, 0)	113.603053	115.491931	17.100731	326.121357	18.058830	0.065921
Spain	(0, 1, 0)	97.203760	98.148199	11.561754	181.999108	13.490705	0.028345
Sweden	(0, 1, 0)	76.553926	78.442804	9.463627	155.850718	12.484018	0.030884
Switzerland	(0, 2, 1)	95.496700	97.277444	21.085462	637.136702	25.241567	0.060882
Turkey	(0, 2, 0)	117.800755	118.691127	10.854420	130.112571	11.406690	0.018875
Ukraine	(0, 1, 1)	120.767939	123.601256	30.332438	969.728953	31.140471	0.140078
United Kingdom	(3, 0, 1)	99.598134	105.572528	4.522505	24.172363	4.916540	0.019056

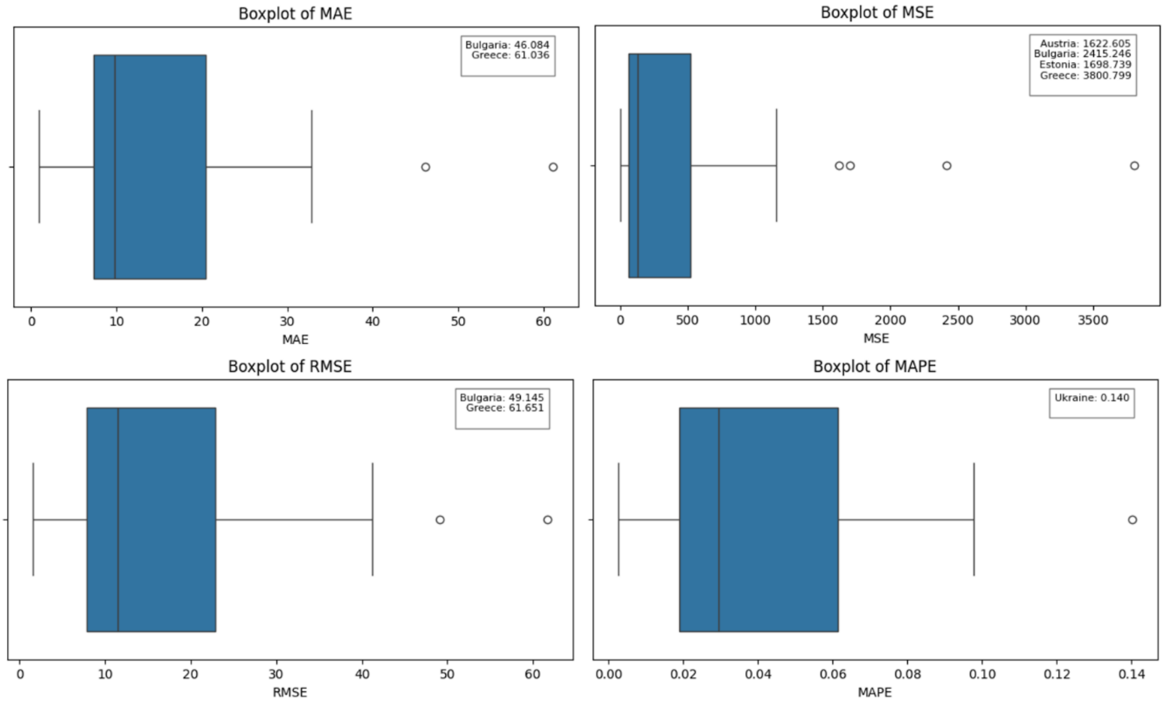
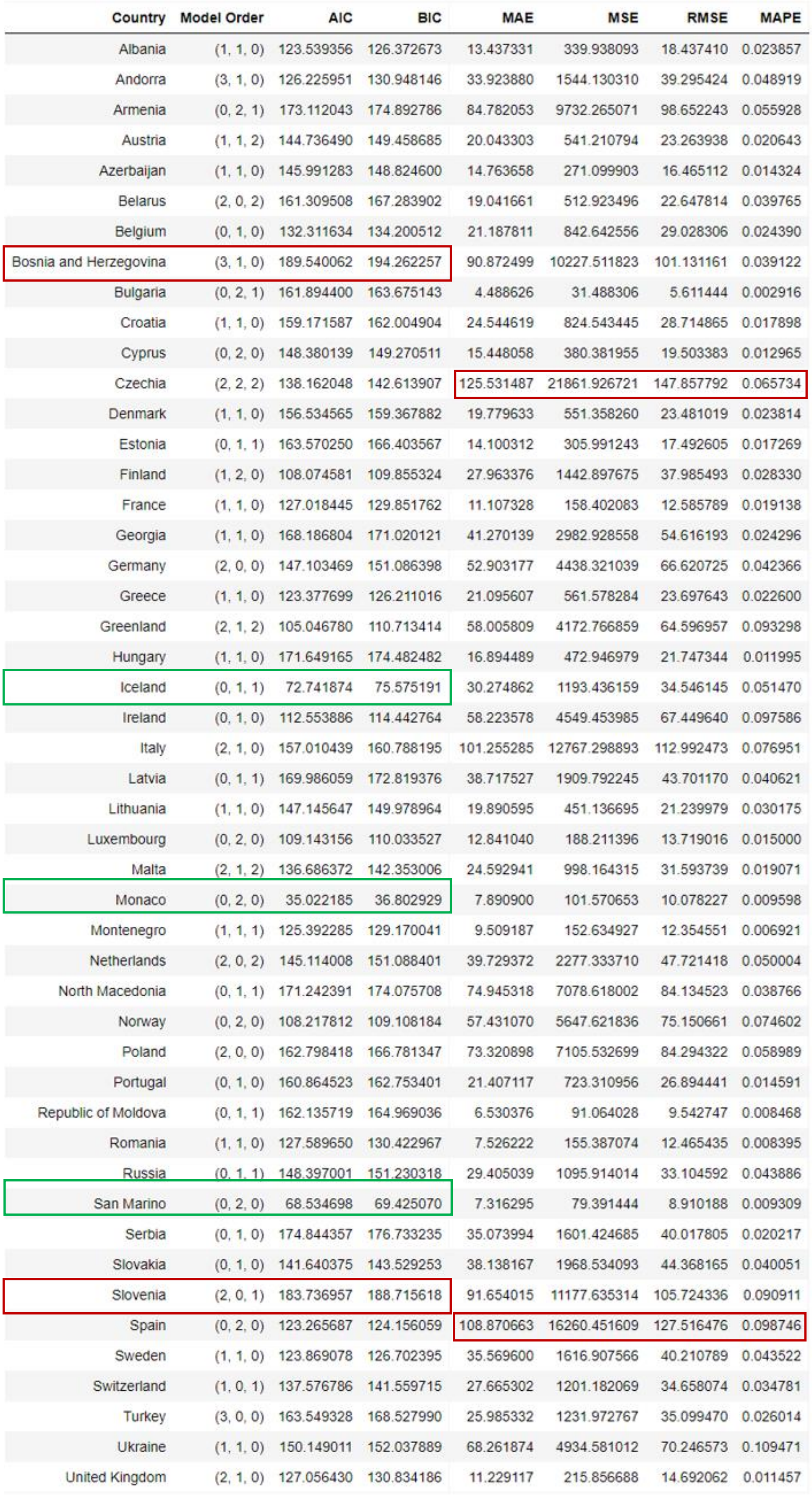


Figure 17 Boxplots for MAE, MSE, RMSE, and MAPE for Chronic Kidney Diseases

5.2.5. **Diabetes Mellitus**

Table 7 summarizes the ARIMA model order of the last cross-validation fold for each country along with the AIC and BIC values and the values of various error metrics, MAE, MSE, RMSE and MAPE, for model predictions. The different ARIMA model orders indicate differences in the patterns and trends among countries. High AIC and BIC values, for example, Slovenia with AIC value of 183.73 and BIC of 188.71, and Bosnia and Herzegovina with AIC of 189.54 and BIC of 194.26 suggest their models are nor as optimal as the other countries. In the contrast, countries like Monaco, San Marino, and Iceland have lower AIC and BIC values, indicating more effective models. Monaco and San Marino stand out again with low MAE, MSE, RMSE, and MAPE values, along with Bulgaria, which implies that the forecasts are highly accurate. Czechia and Spain are some of the countries that have higher error metric values, which indicates that the models struggle to make accurate forecasts.

Table 7 Selected Arima model for each country for Diabetes Mellitus



Country	Model Order	AIC	BIC	MAE	MSE	RMSE	MAPE
Albania	(1, 1, 0)	123.539356	126.372673	13.437331	339.938093	18.437410	0.023857
Andorra	(3, 1, 0)	126.225951	130.948146	33.923880	1544.130310	39.295424	0.048919
Armenia	(0, 2, 1)	173.112043	174.892786	84.782053	9732.265071	98.652243	0.055928
Austria	(1, 1, 2)	144.736490	149.458685	20.043303	541.210794	23.263938	0.020643
Azerbaijan	(1, 1, 0)	145.991283	148.824600	14.763658	271.099903	16.465112	0.014324
Belarus	(2, 0, 2)	161.309508	167.283902	19.041661	512.923496	22.647814	0.039765
Belgium	(0, 1, 0)	132.311634	134.200512	21.187811	842.642556	29.028306	0.024390
Bosnia and Herzegovina	(3, 1, 0)	189.540062	194.262257	90.872499	10227.511823	101.131161	0.039122
Bulgaria	(0, 2, 1)	161.894400	163.675143	4.488626	31.488306	5.611444	0.002916
Croatia	(1, 1, 0)	159.171587	162.004904	24.544619	824.543445	28.714865	0.017898
Cyprus	(0, 2, 0)	148.380139	149.270511	15.448058	380.381955	19.503383	0.012965
Czechia	(2, 2, 2)	138.162048	142.613907	125.531487	21861.926721	147.857792	0.065734
Denmark	(1, 1, 0)	156.534565	159.367882	19.779633	551.358260	23.481019	0.023814
Estonia	(0, 1, 1)	163.570250	166.403567	14.100312	305.991243	17.492605	0.017269
Finland	(1, 2, 0)	108.074581	109.855324	27.963376	1442.897675	37.985493	0.028330
France	(1, 1, 0)	127.018445	129.851762	11.107328	158.402083	12.585789	0.019138
Georgia	(1, 1, 0)	168.186804	171.020121	41.270139	2982.928558	54.616193	0.024296
Germany	(2, 0, 0)	147.103469	151.086398	52.903177	4438.321039	66.620725	0.042366
Greece	(1, 1, 0)	123.377699	126.211016	21.095607	561.578284	23.697643	0.022600
Greenland	(2, 1, 2)	105.046780	110.713414	58.005809	4172.766859	64.596957	0.093298
Hungary	(1, 1, 0)	171.649165	174.482482	16.894489	472.946979	21.747344	0.011995
Iceland	(0, 1, 1)	72.741874	75.575191	30.274862	1193.436159	34.546145	0.051470
Ireland	(0, 1, 0)	112.553886	114.442764	58.223578	4549.453985	67.449640	0.097586
Italy	(2, 1, 0)	157.010439	160.788195	101.255285	12767.298893	112.992473	0.076951
Latvia	(0, 1, 1)	169.986059	172.819376	38.717527	1909.792245	43.701170	0.040621
Lithuania	(1, 1, 0)	147.145647	149.978964	19.890595	451.136695	21.239979	0.030175
Luxembourg	(0, 2, 0)	109.143156	110.033527	12.841040	188.211396	13.719016	0.015000
Malta	(2, 1, 2)	136.686372	142.353006	24.592941	998.164315	31.593739	0.019071
Monaco	(0, 2, 0)	35.022185	36.802929	7.890900	101.570653	10.078227	0.009598
Montenegro	(1, 1, 1)	125.392285	129.170041	9.509187	152.634927	12.354551	0.006921
Netherlands	(2, 0, 2)	145.114008	151.088401	39.729372	2277.333710	47.721418	0.050004
North Macedonia	(0, 1, 1)	171.242391	174.075708	74.945318	7078.618002	84.134523	0.038766
Norway	(0, 2, 0)	108.217812	109.108184	57.431070	5647.621836	75.150661	0.074602
Poland	(2, 0, 0)	162.798418	166.781347	73.320898	7105.532699	84.294322	0.058989
Portugal	(0, 1, 0)	160.864523	162.753401	21.407117	723.310956	26.894441	0.014591
Republic of Moldova	(0, 1, 1)	162.135719	164.969036	6.530376	91.064028	9.542747	0.008468
Romania	(1, 1, 0)	127.589650	130.422967	7.526222	155.387074	12.465435	0.008395
Russia	(0, 1, 1)	148.397001	151.230318	29.405039	1095.914014	33.104592	0.043886
San Marino	(0, 2, 0)	68.534698	69.425070	7.316295	79.391444	8.910188	0.009309
Serbia	(0, 1, 0)	174.844357	176.733235	35.073994	1601.424685	40.017805	0.020217
Slovakia	(0, 1, 0)	141.640375	143.529253	38.138167	1968.534093	44.368165	0.040051
Slovenia	(2, 0, 1)	183.736957	188.715618	91.654015	11177.635314	105.724336	0.090911
Spain	(0, 2, 0)	123.265687	124.156059	108.870663	16260.451609	127.516476	0.098746
Sweden	(1, 1, 0)	123.869078	126.702395	35.569600	1616.907566	40.210789	0.043522
Switzerland	(1, 0, 1)	137.576786	141.559715	27.665302	1201.182069	34.658074	0.034781
Turkey	(3, 0, 0)	163.549328	168.527990	25.985332	1231.972767	35.099470	0.026014
Ukraine	(1, 1, 0)	150.149011	152.037889	68.261874	4934.581012	70.246573	0.109471
United Kingdom	(2, 1, 0)	127.056430	130.834186	11.229117	215.856688	14.692062	0.011457

Figure 18 helps to visually understand model's performance, with the boxplot's outliers showing in which countries the models performed worse. Czechia is an outlier in the boxplots for MAE, MSE and RMSE. Italy, Slovenia, Spain, and Bosnia and Herzegovina are outliers for the MSE metric. Additionally, Ukraine and Spain are outliers in the boxplot for the MAPE metric, but its median value is considered low, which indicates that even though there are outliers for all four metrics, the models generally performed well across most countries.

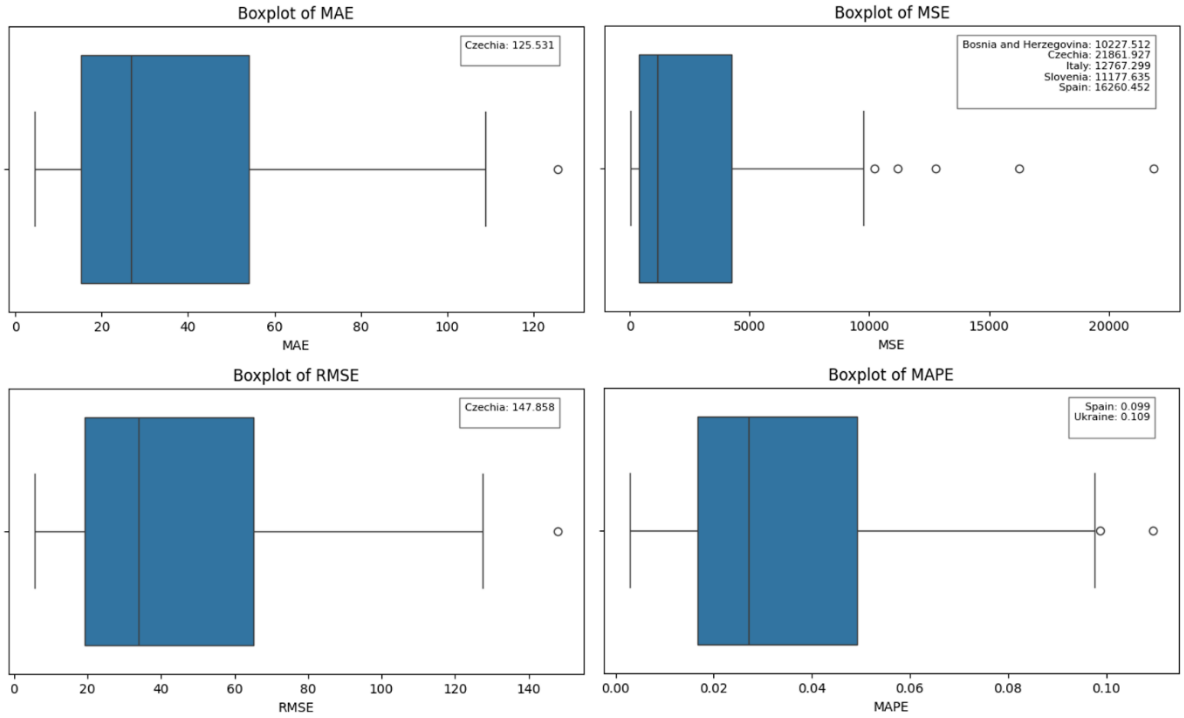


Figure 18 Boxplots for MAE, MSE, RMSE, and MAPE for Diabetes Mellitus

5.3. Clusters results

After finding the best model for each country, they were grouped together using three different distance measures and several hierarchical clustering methods. Each method was evaluated using four clustering indices to determine the optimal number of clusters and assess the quality of clustering. The hierarchical clustering methods used were Ward (ward.D), Single, Median, Complete, Centroid, and Average. The NbClust package in R was used to determine the number of clusters for each method. The minimum number of clusters possible was 2 and the maximum was 15. The ideal number of clusters was determined based on the majority rule among four clustering indices. In cases where there was a tie or a lack of consensus between the indices, the highest suggested number of clusters was selected.

5.3.1. Cardiovascular Diseases – Piccolo Distance

Table 8 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. Across most methods, Mcclain metric suggests 2 clusters except for the Complete method, which suggests 3. According to the Cindex metric 15 clusters are ideal across all methods. The Silhouette score shows most methods favoring 15 clusters, but for the Average and Centroid methods the optimal number is 2, and for Complete is 11 clusters. The Dunn index prefers smaller cluster numbers, especially 2 clusters, except for Complete and Median methods, with 6 and 10 clusters respectively.

Table 8 Cluster evaluation metrics: values and optimal number of clusters with Piccolo Distance for Cardiovascular Diseases

Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	2	3	2	2	2	2
	Value	0,2536	0.2023	0,0143	0,0143	0,0143	0,0848
Cindex	Clusters	15	15	15	15	15	15
	Value	0,0352	0.0486	0,1226	0,0542	0,0984	0,0756
Silhouette	Clusters	15	11	2	2	15	15
	Value	0,6471	0.6496	0,658	0,658	0,6709	0,6638
Dunn	Clusters	2	6	2	2	2	10
	Value	0,239	0.2820	0,4946	0,4946	0,4946	0,2389

Figure 19 shows the representation of the cluster's membership in geographical maps across countries. For the Ward, Complete, Centroid and Median methods, the countries were grouped in 15 clusters, and for the Single and Average methods in 2 clusters. A detailed description of each cluster's members can be found in annex A.

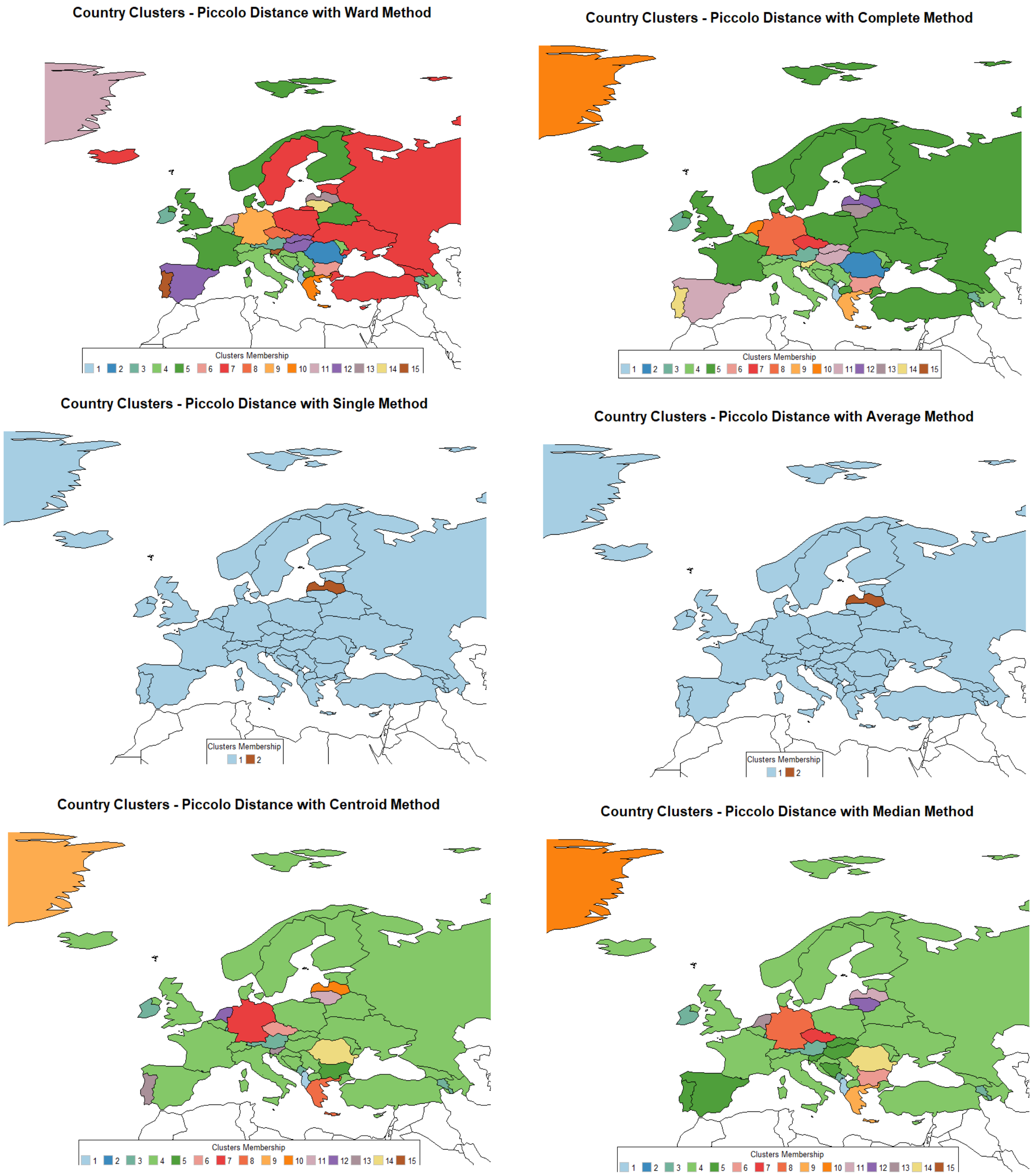


Figure 19 Geographical maps of the clustering results with Piccolo Distance for Cardiovascular Diseases

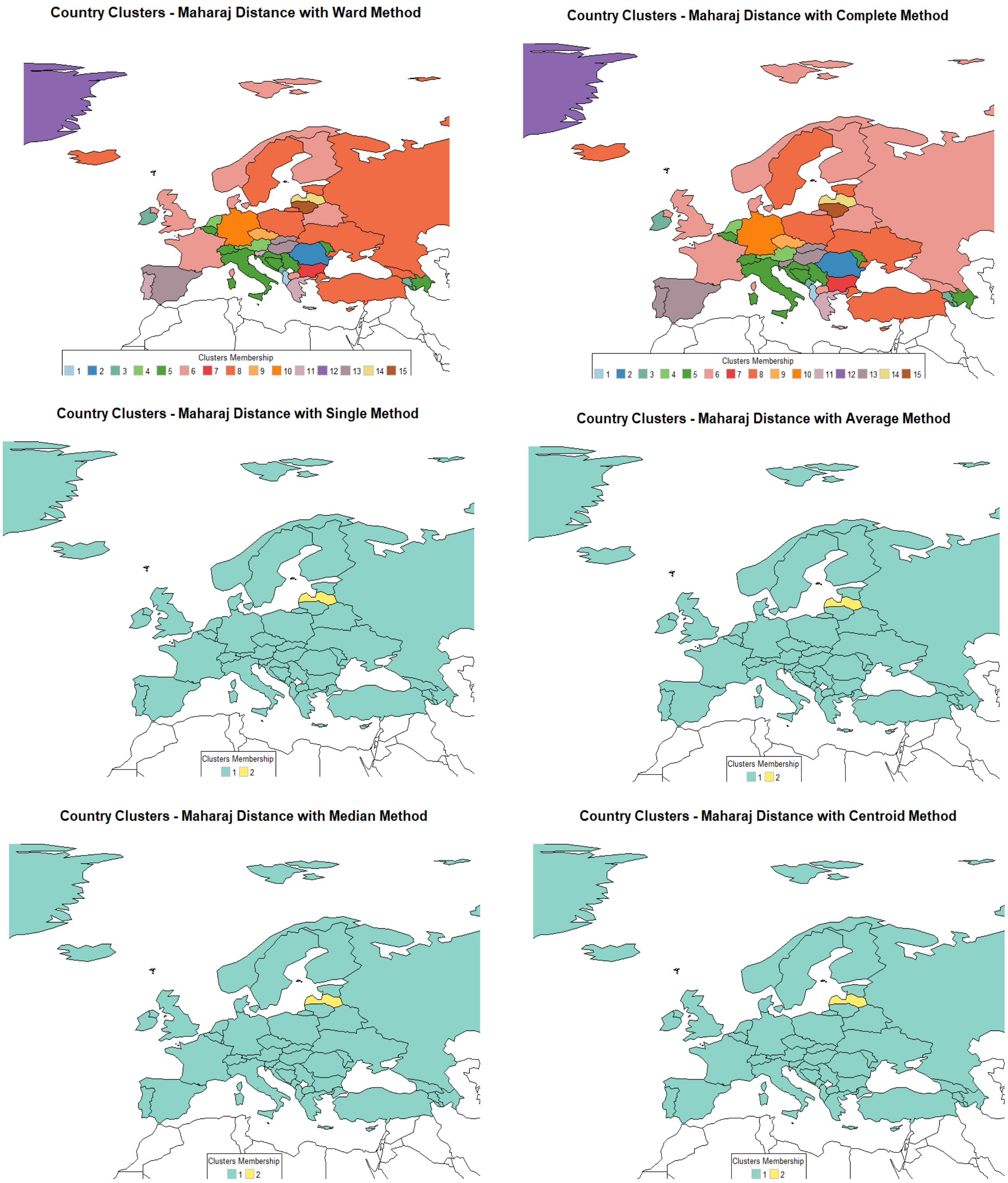
5.3.2. **Cardiovascular Diseases – Maharaj Distance**

Table 9 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. Across most methos, the Mcclain index suggests 2 clusters, except for the Complete method, where the optimal number is 15. With this index, the methods that suggested 2 clusters have the lowest Mcclain values meaning they form more compact clusters. With Cindex, the Single method suggests 6 clusters, while the rest of the methods prefer 15 clusters. The Ward method has the lowest value for Cindex, which means it provides the best balance between compactness and separation between clusters. The Silhouette score suggests 15 cluster for Ward and Complete and for the other methods 2 clusters, which achieve the highest Silhouette score indicating well separated clusters. Lastly, the Dunn index suggests 2 clusters for all methods except Complete in which 15 clusters and average which gives 5 clusters.

Table 9 Cluster evaluation metrics: values and optimal number of clusters with Maharaj Distance for Cardiovascular Diseases

Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	2	15	2	2	2	2
	Value	0,1678	0,232	0,0072	0,0072	0,0072	0,0072
Cindex	Clusters	15	15	6	15	15	15
	Value	0,0164	0,0205	0,0974	0,0712	0,0712	0,027
Silhouette	Clusters	15	15	2	2	2	2
	Value	0,7115	0,6172	0,8123	0,8123	0,8123	0,8123
Dunn	Clusters	2	15	2	5	2	2
	Value	0,0252	0,0049	0,2652	0,2652	0,2652	0,2652

Figure 20 shows the representation of the cluster's membership in geographical maps across countries. For the Ward and Complete methods, the countries were grouped in 15 clusters, and for the Single, Average, Centroid, and Median methods in 2 clusters. A detailed description of each cluster's members can be found in annex A.



Country Clusters - Maharaj Distance with Ward Method

Country Clusters - Maharaj Distance with Complete Method

Country Clusters - Maharaj Distance with Single Method

Country Clusters - Maharaj Distance with Average Method

Country Clusters - Maharaj Distance with Median Method

Country Clusters - Maharaj Distance with Centroid Method

Figure 20 Geographical maps of the clustering results with Maharaj Distance for Cardiovascular Diseases

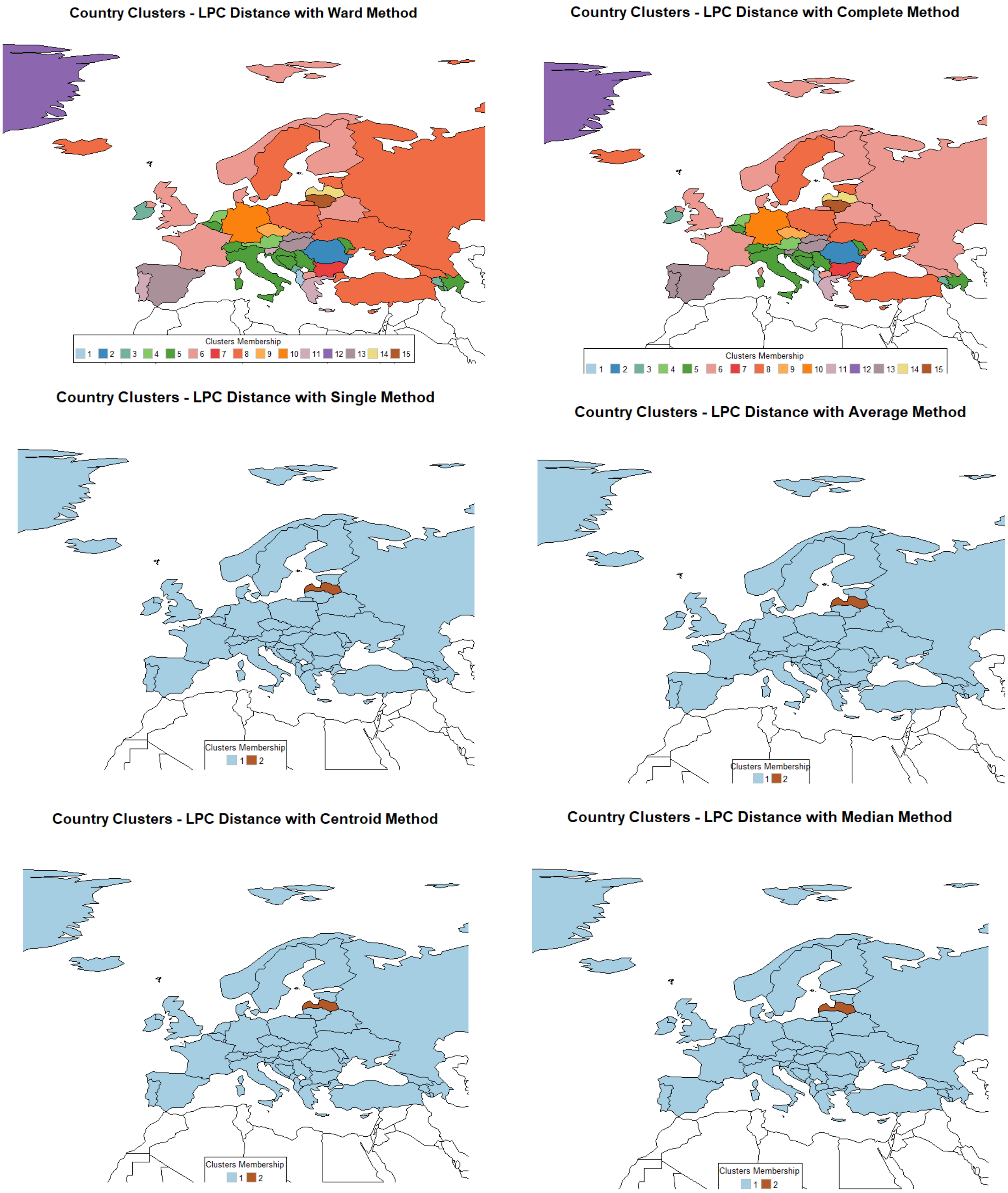
5.3.3. Cardiovascular Diseases – LPC Distance

Table 10 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. The Mcclain index suggests 2 clusters for most methods, except Complete where it recommends 3 clusters. The Cindex suggests 15 clusters for all six clustering methods. The Silhouette score and the Dunn index recommend the same optimal number of clusters, with 15 clusters for Ward and Complete, and 2 clusters for the rest of the methods.

Table 10 Cluster evaluation metrics: values and optimal number of clusters with LPC Distance for Cardiovascular Diseases

Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	2	3	2	2	2	2
	Value	0,6591	0,4992	0,0152	0,0152	0,0152	0,0152
Cindex	Clusters	15	15	15	15	15	15
	Value	0,0594	0,0704	0,2021	0,1386	0,1718	0,1011
Silhouette	Clusters	15	15	2	2	2	2
	Value	0,5701	0,503	0,6465	0,6465	0,6465	0,6465
Dunn	Clusters	15	15	2	5	2	2
	Value	0,1441	0,0686	0,5881	0,5881	0,5881	0,5881

Figure 21 the representation of the cluster's membership in geographical maps across countries. For the Ward and Complete methods, the countries were grouped in 15 countries. For the Single, Average, Median, and Centroid methods, the countries formed 2 clusters. A detailed description of each cluster's members can be found in annex A.



Country Clusters - LPC Distance with Ward Method

Country Clusters - LPC Distance with Complete Method

Country Clusters - LPC Distance with Single Method

Country Clusters - LPC Distance with Average Method

Country Clusters - LPC Distance with Centroid Method

Country Clusters - LPC Distance with Median Method

Figure 21 Geographical maps of the clustering results with LPC Distance for Cardiovascular Diseases

5.3.4. Chronic Respiratory Diseases – Piccolo Distance

Table 11 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. The Mcclain metric shows that for all methods the ideal number of clusters is 2. The Cindex suggests 15 clusters across the six chosen clustering methods. The Silhouette score recommends 8 clusters for Ward method, 14 clusters for Complete method, and 15 clusters for the Single, Average, Median, and Centroid methods where it achieves the best score. For the Dunn index, 8 clusters are suggested for Ward, 9 for Complete, 15 for Single and Median methods, 7 clusters for Average, and lastly 12 cluster for the Centroid method. According to the rule of majority, for Ward the countries were grouped in 8 clusters, and for the rest of the methods in 15 clusters.

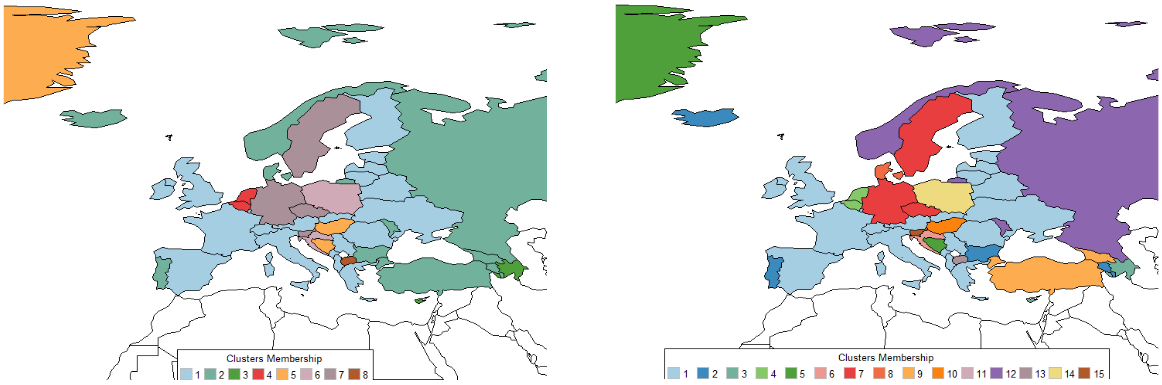
Table 11 Cluster evaluation metrics: values and optimal number of clusters with Piccolo Distance for Chronic Respiratory Diseases

Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	2	2	2	2	2	2
	Value	0,3678	0,3373	0,0279	0,0382	0,0222	0,3678
Cindex	Clusters	15	15	15	15	15	15
	Value	0,072	0,0827	0,1068	0,0935	0,1068	0,0935
Silhouette	Clusters	8	14	15	15	15	15
	Value	0,5599	0,6388	0,6985	0,6923	0,6985	0,6923
Dunn	Clusters	8	9	15	7	12	15
	Value	0,4285	0,5593	0,5593	0,3978	0,5593	0,3856

Figure 22 shows the representation of the cluster's membership in geographical maps across countries. A detailed description of each cluster's members can be found in annex B.

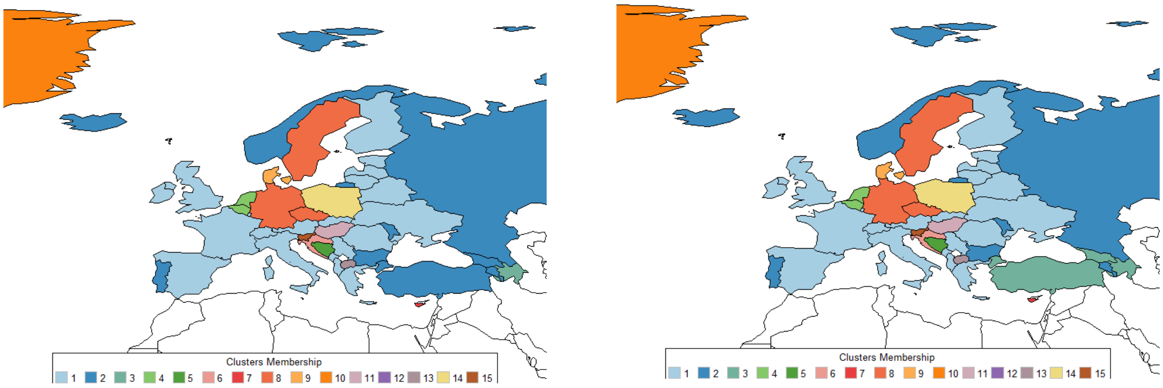
Country Clusters - Piccolo Distance with Ward Method

Country Clusters - Piccolo Distance with Complete Method



Country Clusters - Piccolo Distance with Single Method

Country Clusters - Piccolo Distance with Average Method



Country Clusters - Piccolo Distance with Median Method

Country Clusters - Piccolo Distance with Centroid Method

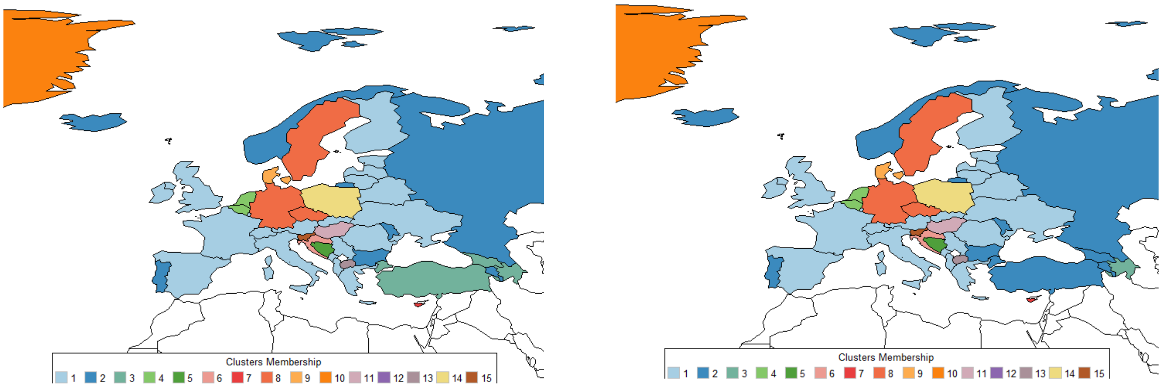


Figure 22 Geographical maps of the clustering results with Piccolo Distance for Chronic Respiratory Diseases

5.3.5. Chronic Respiratory Diseases – Maharaj Distance

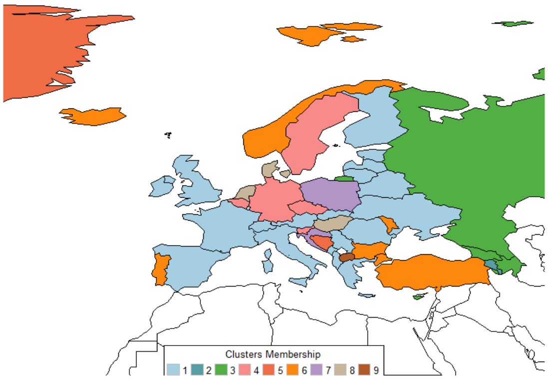
Table 12 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. The Mcclain index suggests 2 clusters for all methods, except for Ward where it recommends 9 clusters. The Cindex suggests 15 clusters except for the Single linkage where the optimal number of clusters is 6. The Silhouette score recommends 14 clusters for the Ward and Median methods, 13 for Average method, and 15 clusters for Complete, Single, and Centroid. Lastly, the Dunn index suggests 9 clusters for the Ward method, 13 for Complete and Centroid, 15 clusters for Single and Average linkages, and 14 for the Median method. Following the rule of majority, the Ward method groups the countries into 9 clusters, the Median into 14 clusters, and finally the Complete, Single, Average, and Centroid linkages groups the countries in 15 clusters.

Table 12 Cluster evaluation metrics: values and optimal number of clusters with Maharaj Distance for Chronic Respiratory Diseases

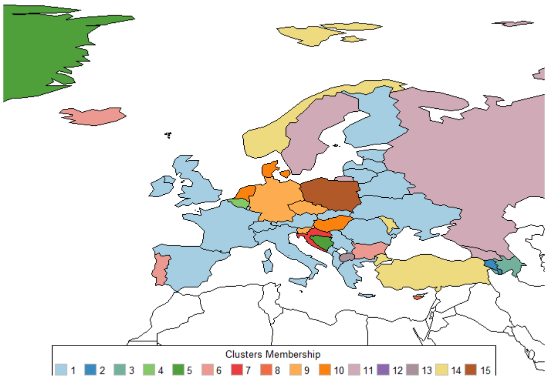
Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	9	2	2	2	2	2
	Value	0,2509	0,0856	0,0204	0,0364	0,0364	0,0364
Cindex	Clusters	15	15	6	15	15	15
	Value	0,0473	0,0797	0,0861	0,0815	0,0838	0,0817
Silhouette	Clusters	14	15	15	13	15	14
	Value	0,7067	0,7609	0,7182	0,7838	0,8088	0,7741
Dunn	Clusters	9	13	15	15	13	14
	Value	0,2126	0,3045	0,1609	0,447	0,425	0,3045

Figure 23 shows the representation of the cluster's membership in geographical maps across countries. A detailed description of each cluster's members can be found in annex B.

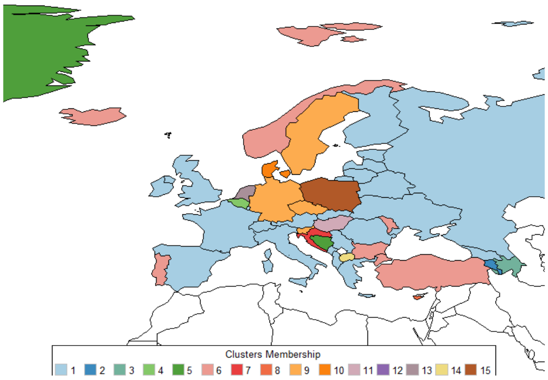
Country Clusters - Maharaj Distance with Ward Method



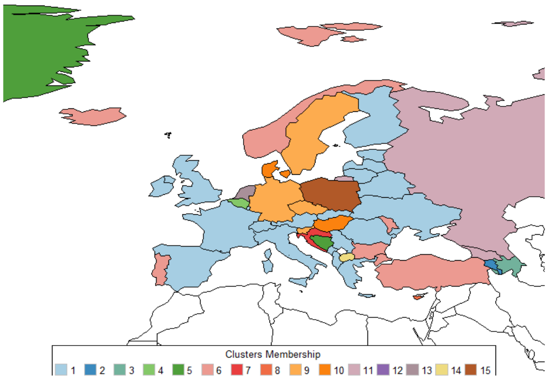
Country Clusters - Maharaj Distance with Complete Method



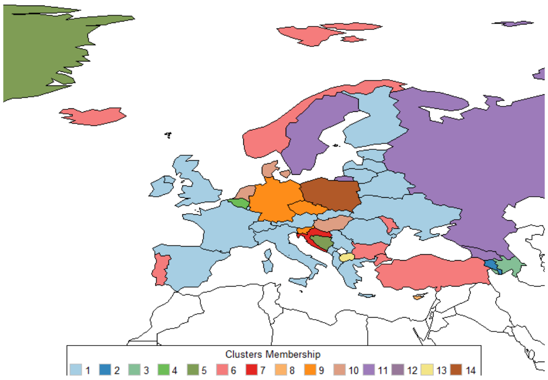
Country Clusters - Maharaj Distance with Single Method



Country Clusters - Maharaj Distance with Average Method



Country Clusters - Maharaj Distance with Median Method



Country Clusters - Maharaj Distance with Centroid Method

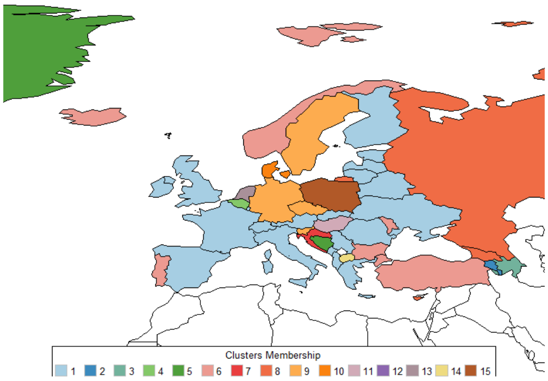


Figure 23 Geographical maps of the clustering results with Maharaj Distance for Chronic Respiratory Diseases

5.3.6. **Chronic Respiratory Diseases – LPC Distance**

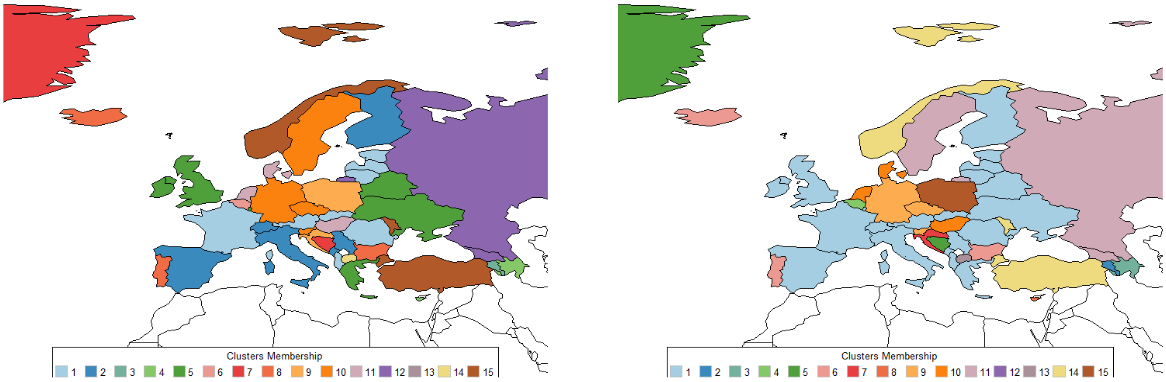
Table 13 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. The Mcclain index suggests only 2 clusters for all clustering methods. The Cindex and Silhouette score, both suggests 15 clusters for all methods. The Dunn index recommends 5 clusters for Ward method, 13 for Complete linkage, and 15 for Single, Average, Centroid, and Median methods. According to the rule of majority, all six clustering methods grouped the countries into 15 clusters.

Table 13 Cluster evaluation metrics: values and optimal number of clusters with LPC Distance for Chronic Respiratory Diseases

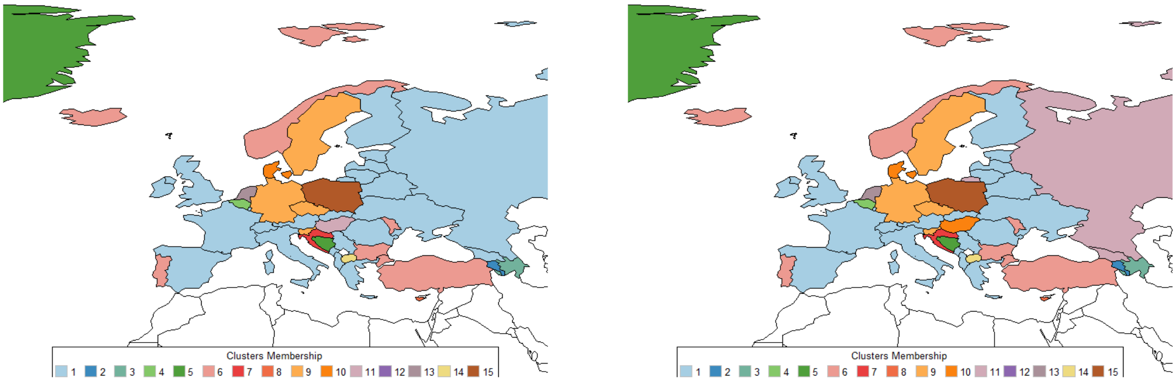
Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	2	2	2	2	2	2
	Value	0,7998	0,3903	0,0331	0,0541	0,0541	0,1144
Cindex	Clusters	15	15	15	15	15	15
	Value	0,0913	0,146	0,2154	0,1481	0,1601	0,1601
Silhouette	Clusters	15	15	15	15	15	15
	Value	0,5448	0,6038	0,6084	0,6251	0,6474	0,6474
Dunn	Clusters	5	13	15	15	15	15
	Value	0,3089	0,5786	0,4069	0,6602	0,436	0,436

Figure 24 shows the representation of the cluster’s membership in geographical maps across countries. A detailed description of each cluster’s members can be found in annex B.

Country Clusters - LPC Distance with Ward Method **Country Clusters - LPC Distance with Complete Method**



Country Clusters - LPC Distance with Single Method **Country Clusters - LPC Distance with Average Method**



Country Clusters - LPC Distance with Median Method **Country Clusters - LPC Distance with Centroid Method**

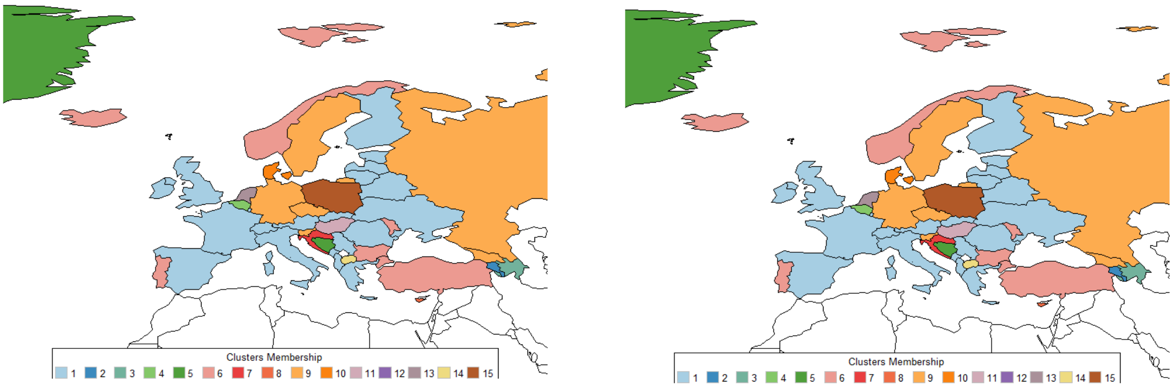


Figure 24 Geographical maps of the clustering results with LPC Distance for Chronic Respiratory Diseases

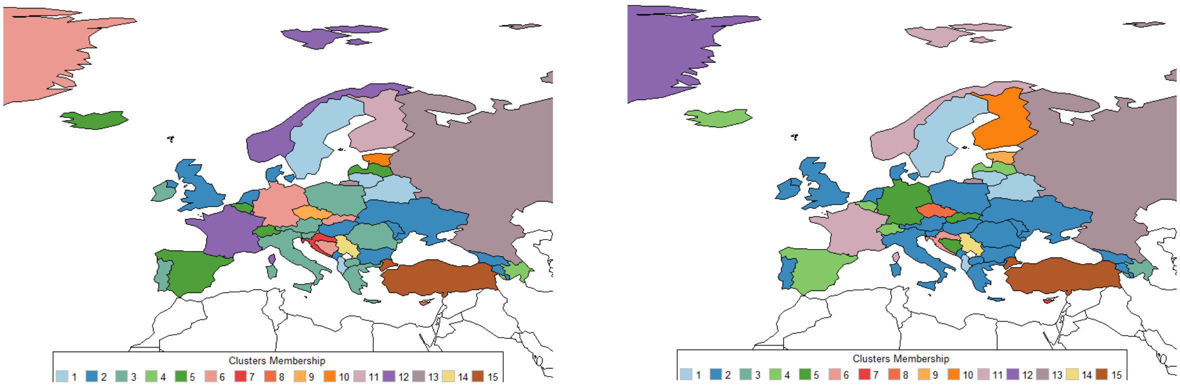
5.3.7. Neurological Disorders – Piccolo Distance

Table 14 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. The Mcclain metric suggests 2 clusters for all methods, except for Ward, where the optimal number of clusters is 8. The Cindex is consistent across the chosen methods and recommends 15 clusters. The Silhouette score suggests 15 clusters for the Ward and Median, 13 for Complete, Average, and Centroid methods, and 11 clusters for the Single linkage. The Dunn index recommends 6 clusters for Ward, 9 for Complete, and 10 clusters for the remaining methods. With the rule of majority, the countries were grouped in 15 clusters for all methods.

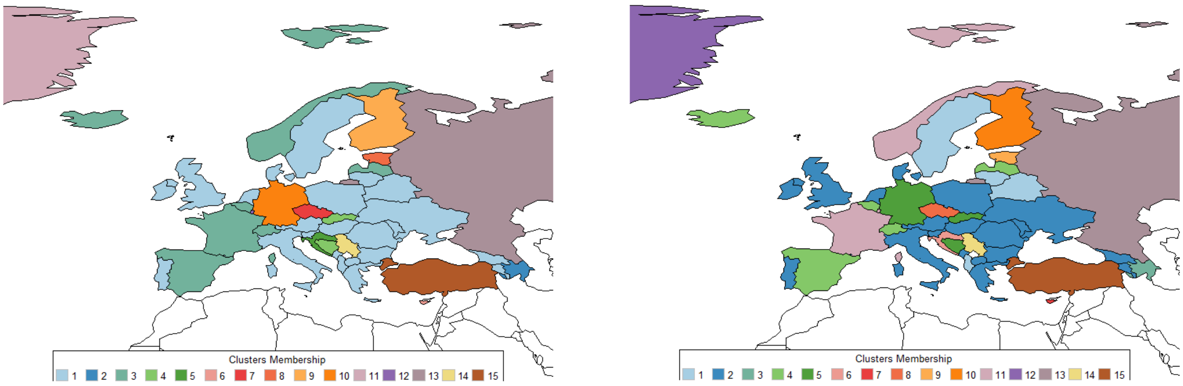
Table 14 Cluster evaluation metrics: values and optimal number of clusters with Piccolo Distance for Neurological Disorders

Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	8	2	2	2	2	2
	Value	0,2354	0,0498	0,0156	0,0498	0,034	0,0498
Cindex	Clusters	15	15	15	15	15	15
	Value	0,033	0,0445	0,0701	0,0445	0,0638	0,0718
Silhouette	Clusters	15	13	11	13	13	15
	Value	0,6772	0,7943	0,7943	0,7943	0,7943	0,786
Dunn	Clusters	6	9	10	10	10	10
	Value	0,4479	0,6293	0,6293	0,6293	0,6293	0,6293

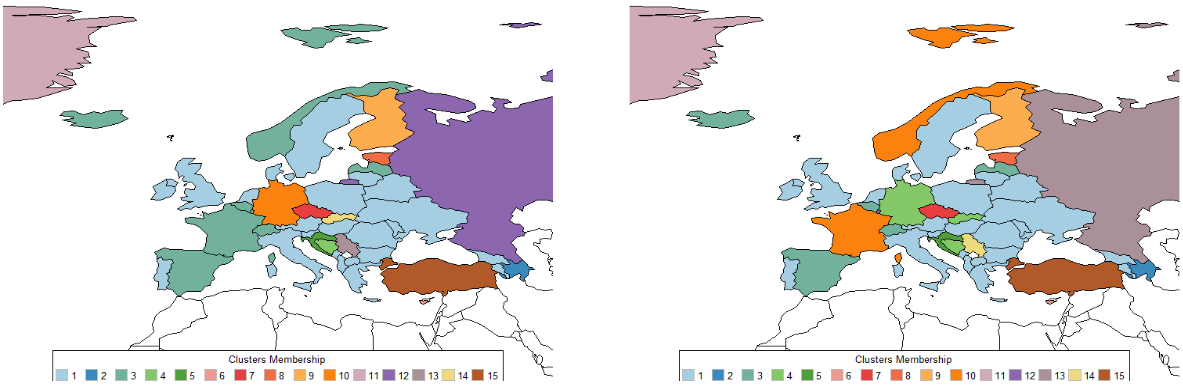
Figure 25 shows the representation of the cluster's membership in geographical maps across countries. A detailed description of each cluster's members can be found in annex C.


Country Clusters - Piccolo Distance with Ward Method

Country Clusters - Piccolo Distance with Complete Method


Country Clusters - Piccolo Distance with Single Method

Country Clusters - Piccolo Distance with Average Method


Country Clusters - Piccolo Distance with Median Method

Country Clusters - Piccolo Distance with Centroid Method

Figure 25 Geographical maps of the clustering results with Piccolo Distance for Neurological Disorders

5.3.8. Neurological Disorders – Maharaj Distance

Table 15 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. The Mcclain index suggests 2 clusters for Ward, Single, Average, and Centroid methods, and 3 clusters for Complete and Median linkages. The Cindex recommends 14 cluster for all methods. Lastly, both the Silhouette score and Dunn index, recommend 3 clusters across all methods. Following the rule of majority, for all methods 3 clusters were formed to groups the countries.

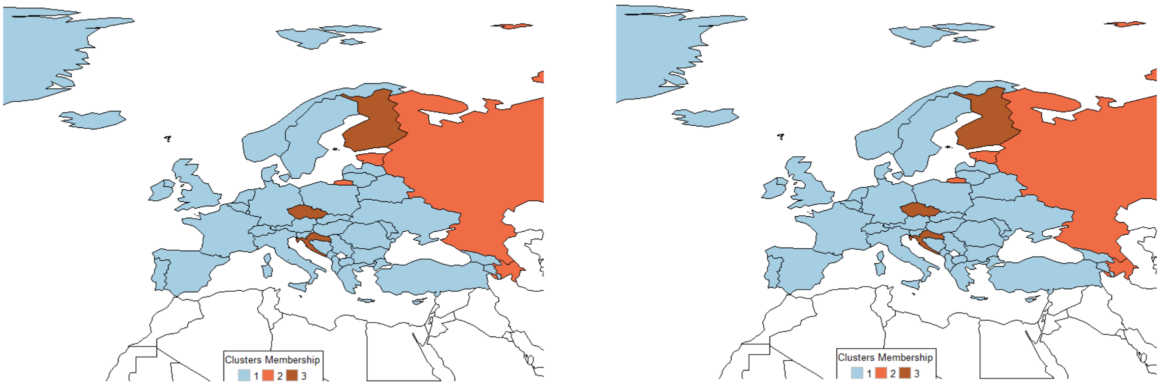
Table 15 Cluster evaluation metrics: values and optimal number of clusters with Maharaj Distance for Neurological Disorders

Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	2	3	2	2	2	3
	Value	0,0253	0,0339	0,0253	0,0253	0,0253	0,0339
Cindex	Clusters	15	15	15	15	15	15
	Value	0,0175	0,0332	0,0677	0,0341	0,0341	0,0372
Silhouette	Clusters	3	3	3	3	3	3
	Value	0,8324	0,8324	0,8324	0,8324	0,8324	0,8324
Dunn	Clusters	3	3	3	3	3	3
	Value	0,384	0,384	0,384	0,384	0,384	0,384

Figure 26 shows the representation of the cluster's membership in geographical maps across countries. A detailed description of each cluster's members can be found in annex C.

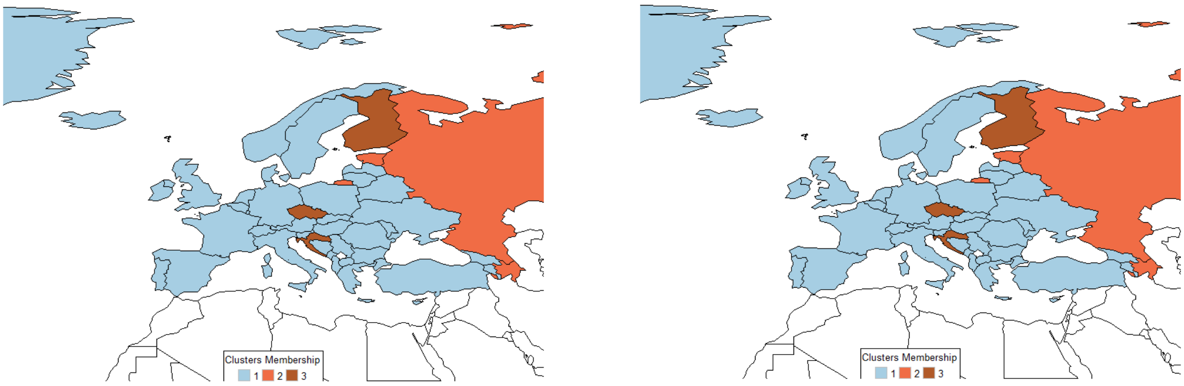
Country Clusters - Maharaj Distance with Ward Method

Country Clusters - Maharaj Distance with Complete Method



Country Clusters - Maharaj Distance with Single Method

Country Clusters - Maharaj Distance with Average Method



Country Clusters - Maharaj Distance with Median Method

Country Clusters - Maharaj Distance with Centroid Method

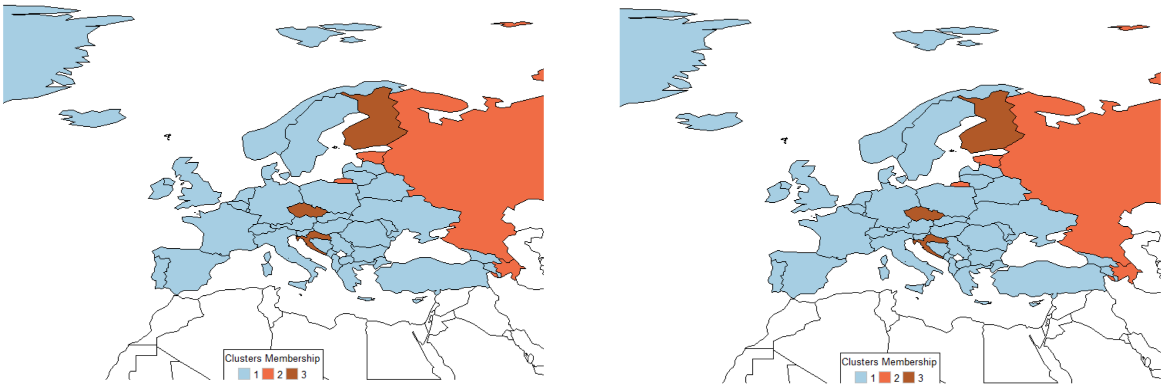


Figure 26 Geographical maps of the clustering results with Maharaj Distance for Neurological Disorders

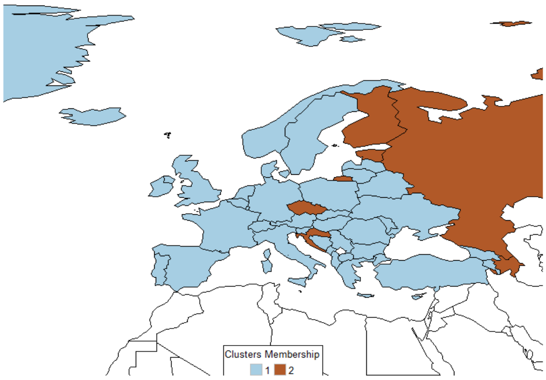
5.3.9. **Neurological Disorders – LPC Distance**

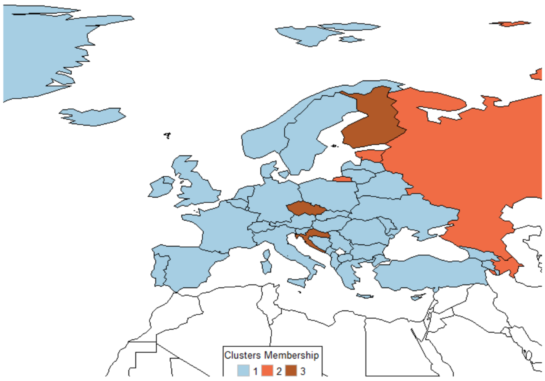
Table 16 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. The Mcclain metric suggests 2 clusters across all methods. The Cindex recommends 15 clusters for all methods. The Sillhouette score recommended 2 clusters for Ward, 3 for Complete, 14 clusters for Single, and 15 clusters for Average, Centroid, and Median methods. Lastly, the Dunn index recommended 3 clusters for all methods except the Ward methods where 2 clusters were the optimal number. According to the rule of majority, the ideal number of clusters for Ward were 2, for Complete linkage were 3 , and for the remaining methods the countries were grouped in 15 clusters.

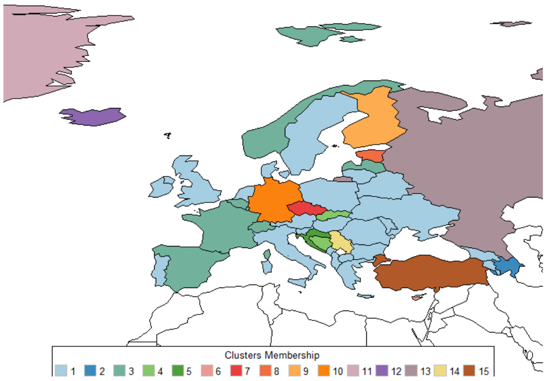
Table 16 Cluster evaluation metrics: values and optimal number of clusters with LPC Distance for Neurological Disorders

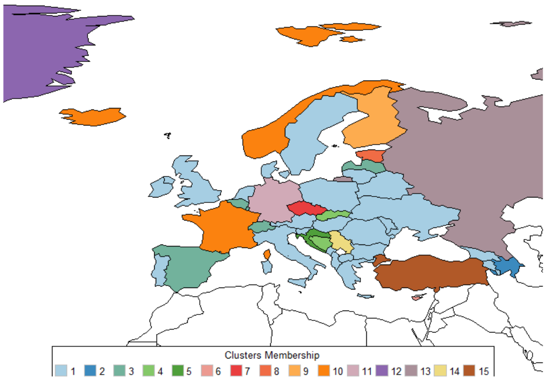
Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	2	2	2	2	2	2
	Value	0,0936	0,0551	0,053	0,053	0,053	0,0551
Cindex	Clusters	15	15	15	15	15	15
	Value	0,0601	0,0936	0,1325	0,1264	0,1264	0,1285
Silhouette	Clusters	2	3	14	15	15	15
	Value	0,6106	0,6362	0,6602	0,679	0,679	0,6626
Dunn	Clusters	2	3	3	3	3	3
	Value	0,3842	0,6392	0,6392	0,6392	0,6392	0,485

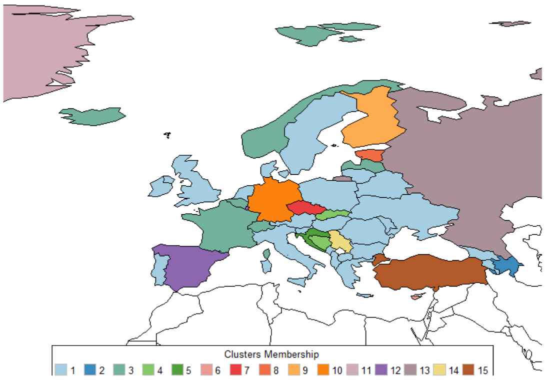
Figure 27 shows the representation of the cluster's membership in geographical maps across countries. A detailed description of each cluster's members can be found in annex C.


Country Clusters - LPC Distance with Ward Method


Country Clusters - LPC Distance with Complete Method


Country Clusters - LPC Distance with Single Method


Country Clusters - LPC Distance with Average Method


Country Clusters - LPC Distance with Median Method

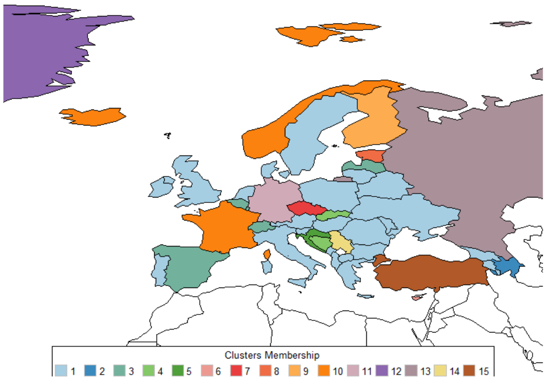

Country Clusters - LPC Distance with Centroid Method

Figure 27 Geographical maps of the clustering results with LPC Distance for Neurological Disorders

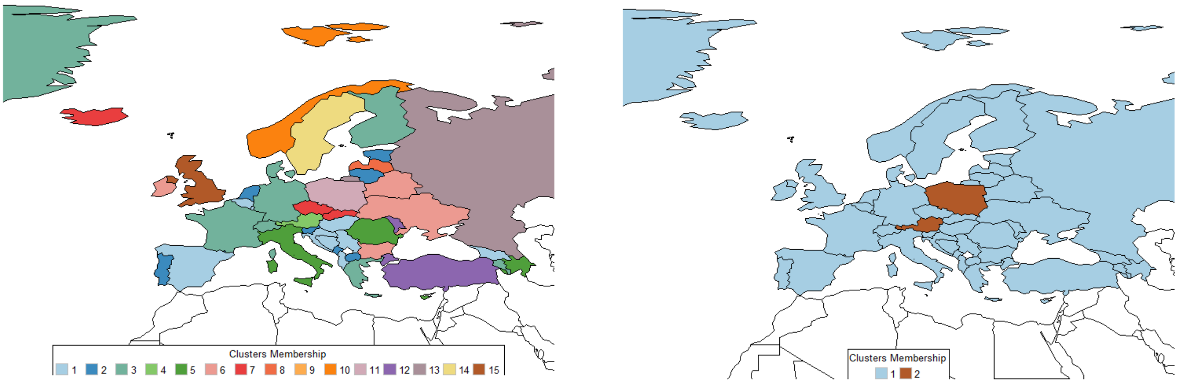
5.3.10. **Chronic Kidney Diseases – Piccolo Distance**

Table 17 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. The Mcclain and Dunn indexes recommend 2 clusters for all methods except for the Ward method where 3 clusters were suggested. The Cindex suggests 15 clusters across all clustering methods. The Silhouette score recommended 15 clusters for Ward, 2 for Complete, and 3 clusters for the remaining methods. Following the rule of majority, the ideal number of clusters for the Ward method is 15 and for the remaining methods is 2 clusters.

Table 17 Cluster evaluation metrics: values and optimal number of clusters with Piccolo Distance for Chronic Kidney Diseases

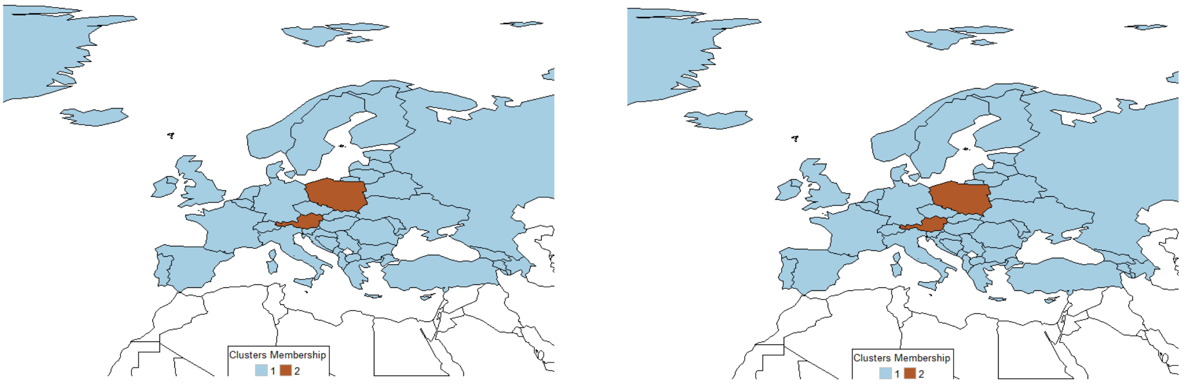
Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	3	2	2	2	2	2
	Value	0,4403	0,0249	0,0249	0,0249	0,0249	0,0249
Cindex	Clusters	15	15	15	15	15	15
	Value	0,0556	0,086	0,1095	0,0968	0,0974	0,086
Silhouette	Clusters	15	2	3	3	3	3
	Value	0,6386	0,7048	0,7183	0,7183	0,7183	0,7183
Dunn	Clusters	3	2	2	2	2	2
	Value	0,2955	0,8132	0,8132	0,8132	0,8132	0,8132

Figure 28 shows the representation of the cluster's membership in geographical maps across countries. A detailed description of each cluster's members can be found in annex D.



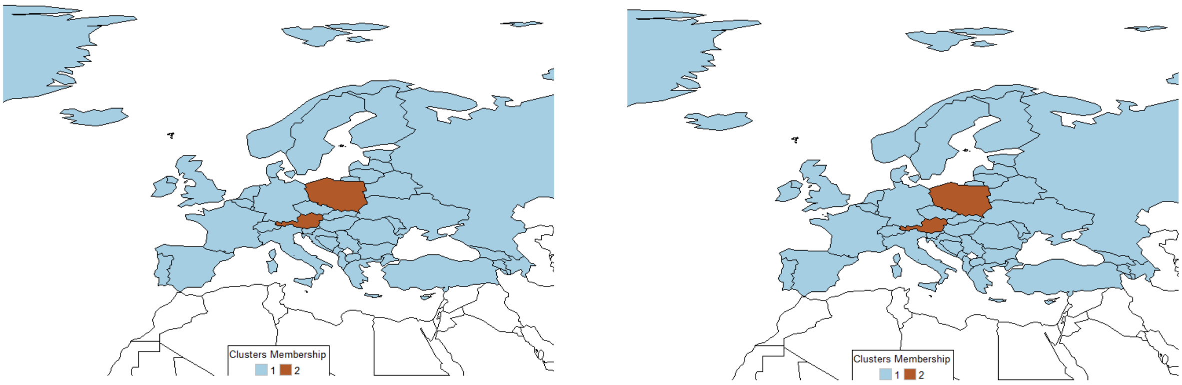
Country Clusters - Piccolo Distance with Ward Method

Country Clusters - Piccolo Distance with Complete Method



Country Clusters - Piccolo Distance with Single Method

Country Clusters - Piccolo Distance with Average Method



Country Clusters - Piccolo Distance with Median Method

Country Clusters - Piccolo Distance with Centroid Method

Figure 28 Geographical maps of the clustering results with Piccolo Distance for Chronic Kidney Diseases

5.3.11. **Chronic Kidney Diseases – Maharaj Distance**

Table 18 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. The Mcclain index and Silhouette score recommend 2 clusters across all methods. The Cindex suggested 15 clusters for all methods except the Single linkage where 3 was the ideal number of clusters. Lastly the Dunn index suggested 2 clusters for Ward and Complete, 5 clusters for Single, Average, and Centroid, and 4 clusters for Median methods. According to the rule of majority, 2 clusters was the optimal number for all clustering methods.

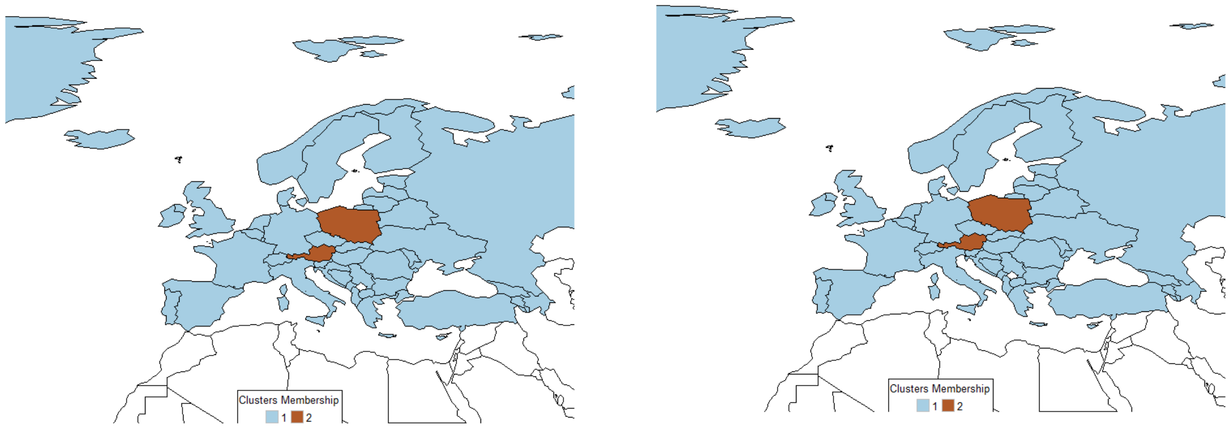
Table 18 Cluster evaluation metrics: values and optimal number of clusters with Maharaj Distance for Chronic Kidney Diseases

Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	2	2	2	2	2	2
	Value	0,0099	0,0099	0,0099	0,0099	0,0099	0,0099
Cindex	Clusters	15	15	3	15	15	15
	Value	0,0438	0,0526	0,1081	0,0435	0,0435	0,0416
Silhouette	Clusters	2	2	2	2	2	2
	Value	0,8783	0,8783	0,8783	0,8783	0,8783	0,8783
Dunn	Clusters	2	2	5	5	5	4
	Value	0,3698	0,3698	0,3945	0,3945	0,3945	0,3945

Figure 29 shows the representation of the cluster's membership in geographical maps across countries. A detailed description of each cluster's members can be found in annex D.

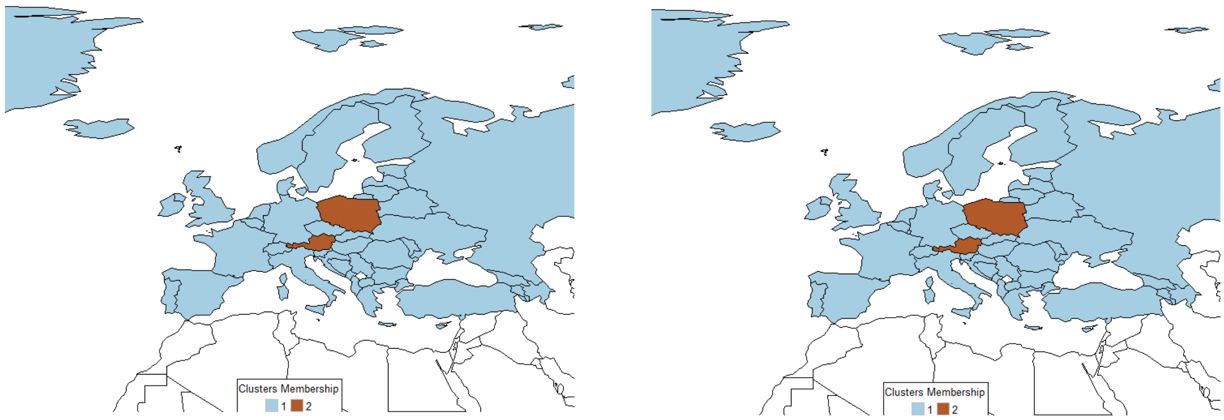
Country Clusters - Maharaj Distance with Ward Method

Country Clusters - Maharaj Distance with Complete Method



Country Clusters - Maharaj Distance with Single Method

Country Clusters - Maharaj Distance with Average Method



Country Clusters - Maharaj Distance with Median Method

Country Clusters - Maharaj Distance with Centroid Method

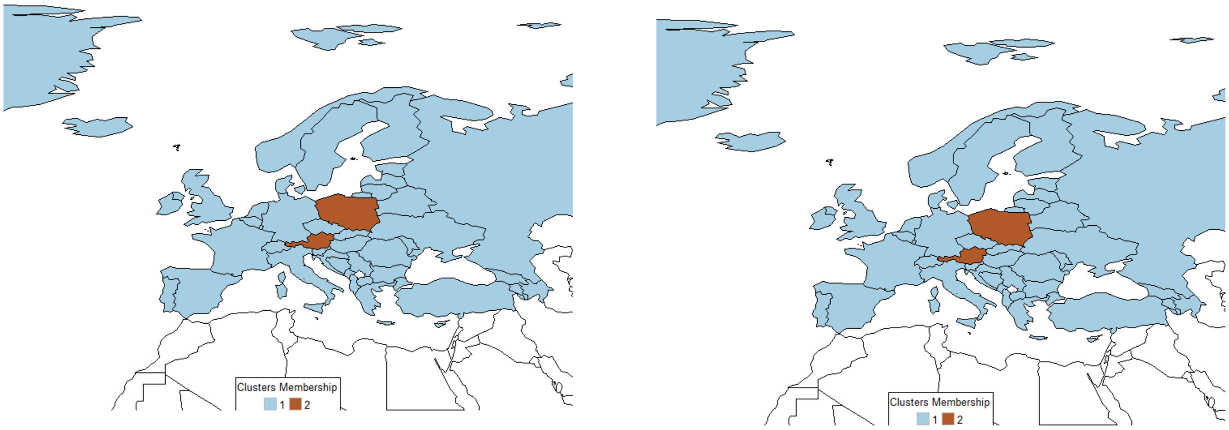


Figure 29 Geographical maps of the clustering results with Maharaj Distance for Chronic Kidney Diseases

5.3.12. **Chronic Kidney Diseases – LPC Distance**

Table 19 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. The Mcclain and Dunn index suggested 2 clusters for all methods except the Ward method where 3 clusters were recommended. Cindex recommended 15 clusters across all clustering methods. The Silhouette score suggested 2 clusters for all methods except for the Ward method where 14 clusters were recommended. According to the rule of majority, the optimal number of clusters was 3 for the Ward method, and for the Complete, Single, Average, Centroid, and Median methods were 2 clusters.

Table 19 Cluster evaluation metrics: values and optimal number of clusters with LPC Distance for Chronic Kidney Diseases

Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	3	2	2	2	2	2
	Value	0,6462	0,0264	0,0264	0,0264	0,0264	0,0264
Cindex	Clusters	15	15	15	15	15	15
	Value	0,0967	0,1203	0,191	0,1078	0,1318	0,1038
Silhouette	Clusters	14	2	2	2	2	2
	Value	0,5905	0,6907	0,6907	0,6907	0,6907	0,6907
Dunn	Clusters	3	2	2	2	2	2
	Value	0,2174	0,6781	0,6781	0,6781	0,6781	0,6781

Figure 30 shows the representation of the cluster's membership in geographical maps across countries. A detailed description of each cluster's members can be found in annex D.

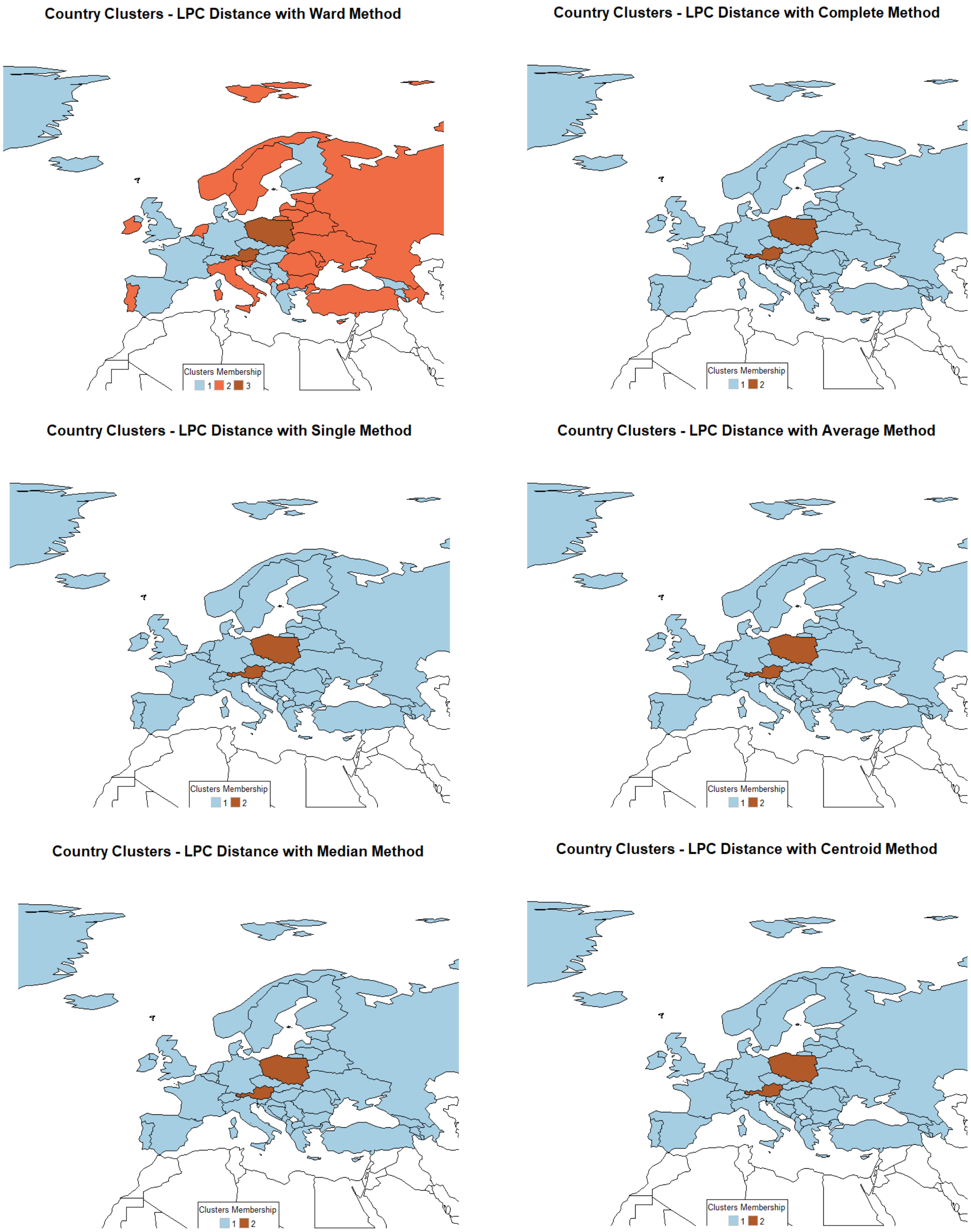


Figure 30 Geographical maps of the clustering results with LPC Distance for Chronic Kidney Diseases

5.3.13. **Diabetes Mellitus– Piccolo Distance**

Table 20 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. The Mcclain index recommended 2 clusters for all methods except for Complete where the ideal number was 14 clusters. Cindex suggested 15 clusters across all clustering methods. The Silhouette score recommended 13 clusters for Ward, 10 for Complete, and 12 for Single, Average, Centroid, and Median methods. Lastly the Dunn index recommended 9 clusters for all methods except the Ward method where 4 clusters were suggested. Since all metrics suggested different number of clusters for each method, the highest number of clusters given for each method was chosen, meaning the countries where grouped into 15 clusters across all methods.

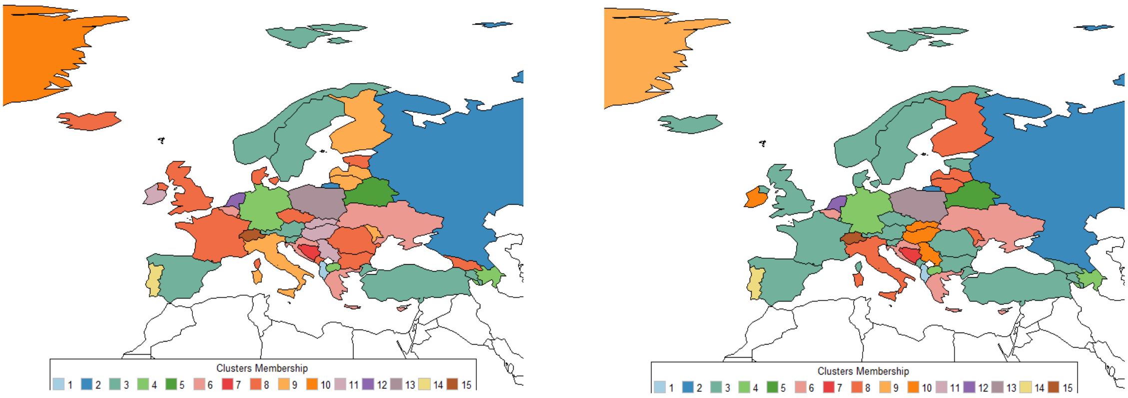
Table 20 Cluster evaluation metrics: values and optimal number of clusters with Piccolo Distance for Diabetes Mellitus

Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	2	14	2	2	2	2
	Value	0,2421	0,1989	0,012	0,012	0,012	0,012
Cindex	Clusters	15	15	15	15	15	15
	Value	0,0239	0,0301	0,0802	0,0459	0,0459	0,0459
Silhouette	Clusters	13	10	12	12	12	12
	Value	0,6825	0,7531	0,8021	0,8021	0,8067	0,8012
Dunn	Clusters	4	9	9	9	9	9
	Value	0,3818	0,8263	0,8263	0,8263	0,9263	0,8263

Figure 31 shows the representation of the cluster's membership in geographical maps across countries. A detailed description of each cluster's members can be found in annex E.

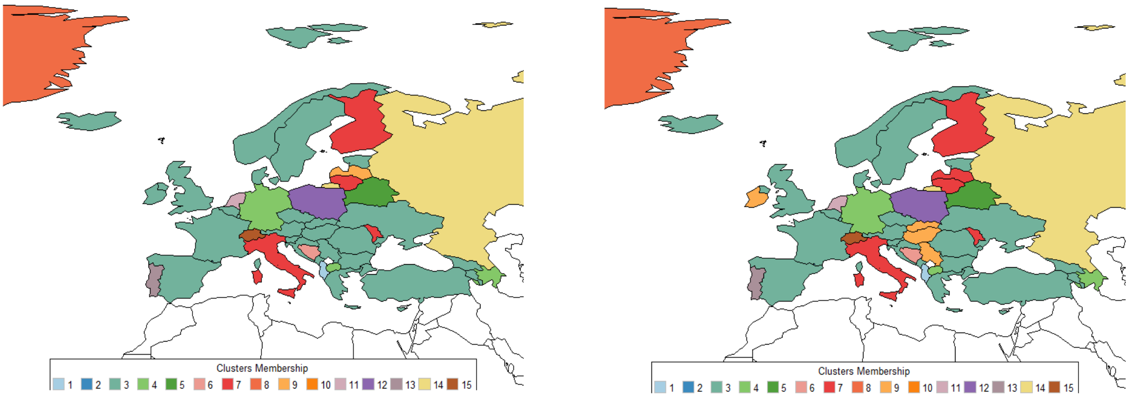
Country Clusters - Piccolo Distance with Ward Method

Country Clusters - Piccolo Distance with Complete Method



Country Clusters - Piccolo Distance with Single Method

Country Clusters - Piccolo Distance with Average Method



Country Clusters - Piccolo Distance with Median Method

Country Clusters - Piccolo Distance with Centroid Method

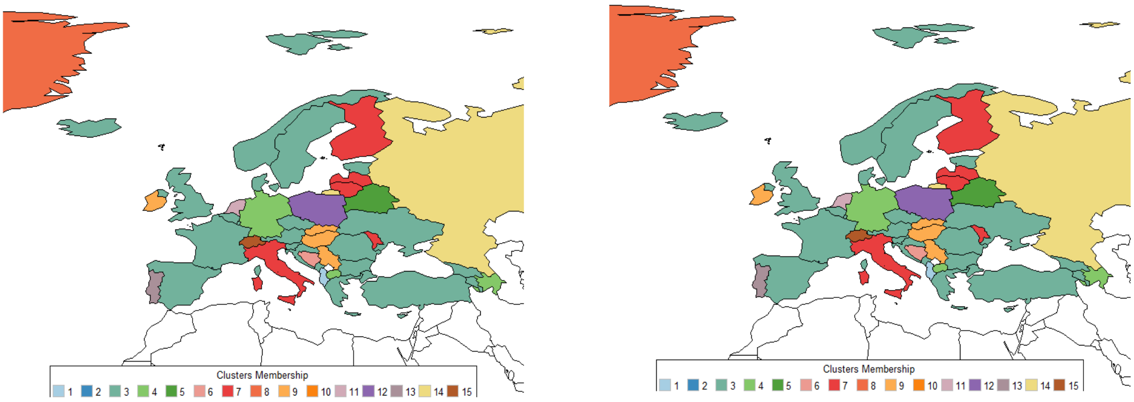


Figure 31 Geographical maps of the clustering results with Piccolo Distance for Diabetes Mellitus

5.3.14. **Diabetes Mellitus– Maharaj Distance**

Table 21 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. The Mcclain index suggested 12 clusters for the Ward method, 15 for Complete, and 2 clusters for the remaining clustering methods. The Cindex recommended 15 clusters for all methods except the Single method in which 13 clusters were suggested. The Silhouette score suggested 15 clusters for Ward and Complete. 11 for Single linkage, and 13 clusters for Average, Centroid, and Median methods. Lastly, the Dunn index recommended 3 clusters for Ward, 15 for Complete, 11 for Single linkage and 7 for the remaining clustering methods. According to the rule of majority, the countries were into 15 clusters for Ward, Complete, Average, Centroid, and Median methods, and 11 clusters for the Single method.

Table 21 Cluster evaluation metrics: values and optimal number of clusters with Maharaj Distance for Diabetes Mellitus

Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	12	15	2	2	2	2
	Value	0,1159	0,0972	0,0132	0,0104	0,0104	0,0104
Cindex	Clusters	15	15	13	15	15	15
	Value	0,0111	0,0179	0,0958	0,0179	0,0179	0,0179
Silhouette	Clusters	15	15	11	13	13	13
	Value	0,7718	0,7762	0,8099	0,8092	0,8092	0,8092
Dunn	Clusters	3	15	11	7	7	7
	Value	0,0795	0,0491	0,2891	0,2891	0,2891	0,2891

Figure 32 shows the representation of the cluster's membership in geographical maps across countries. A detailed description of each cluster's members can be found in annex E.

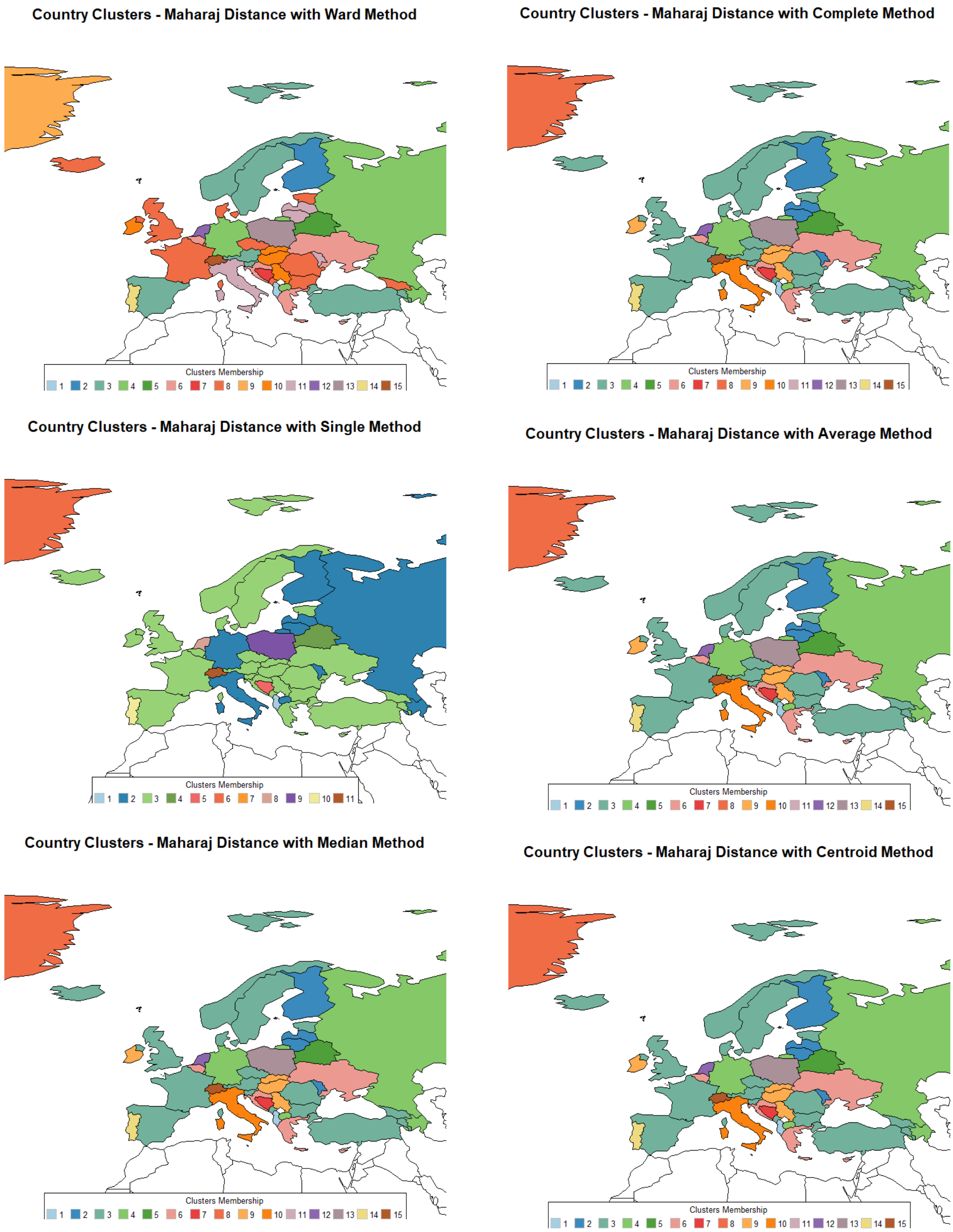


Figure 32 Geographical maps of the clustering results with Maharaj Distance for Diabetes Mellitus

5.3.15. **Diabetes Mellitus– LPC Distance**

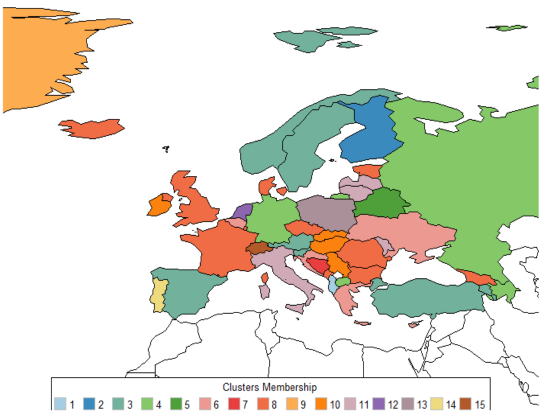
Table 22 shows the values corresponding to each clustering index used to determine the optimal number of clusters across six clustering methods. The Mcclain metric and the Cindex suggest 2 and 15 clusters, respectively, for all six clustering methods. The Silhouette score suggests 15 clusters for Ward, Complete, and Centroid methods, 11 for Single linkage, 13 clusters for Average, and 14 clusters for Median method. The Dunn index clusters the countries into 11 cluster for Single, Centroid, and Median methods, 3 for Ward, 15 for Complete, and 7 clusters for Average linkage. Following the rule of majority, all methods formed 15 clusters, except the Single method which created 11 clusters.

Table 22 Cluster evaluation metrics: values and optimal number of clusters with LPC Distance for Diabetes Mellitus

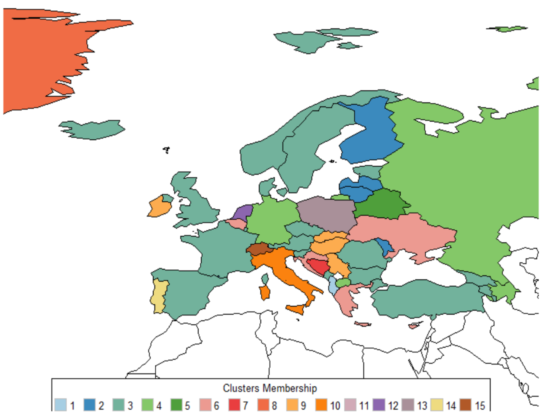
Evaluation Metrics		Clustering Methods					
		Ward	Complete	Single	Average	Centroid	Median
Mcclain	Clusters	2	2	2	2	2	2
	Value	0,4675	0,5715	0,0192	0,0186	0,0186	0,0186
Cindex	Clusters	15	15	15	15	15	15
	Value	0,0464	0,0683	0,1513	0,0976	0,0975	0,0683
Silhouette	Clusters	15	15	11	13	15	14
	Value	0,6194	0,6492	0,6724	0,6836	0,6847	0,6834
Dunn	Clusters	3	15	11	7	11	11
	Value	0,2959	0,2201	0,531	0,531	0,531	0,531

Figure 33 shows the representation of the cluster's membership in geographical maps across countries. A detailed description of each cluster's members can be found in annex E.

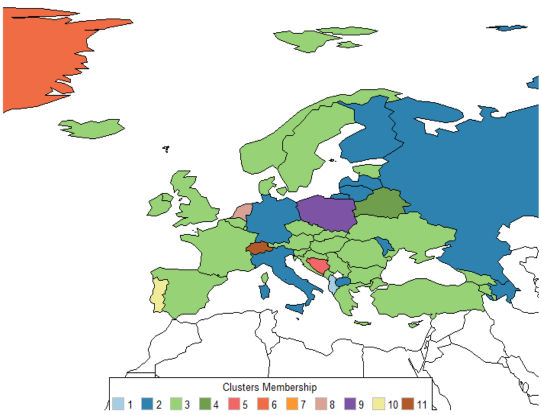
Country Clusters - LPC Distance with Ward Method



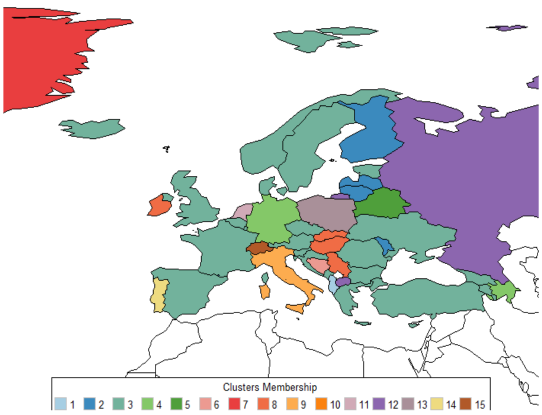
Country Clusters - LPC Distance with Complete Method



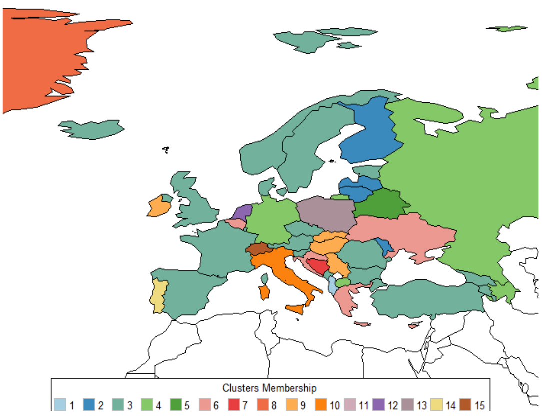
Country Clusters - LPC Distance with Single Method



Country Clusters - LPC Distance with Average Method



Country Clusters - LPC Distance with Median Method



Country Clusters - LPC Distance with Centroid Method

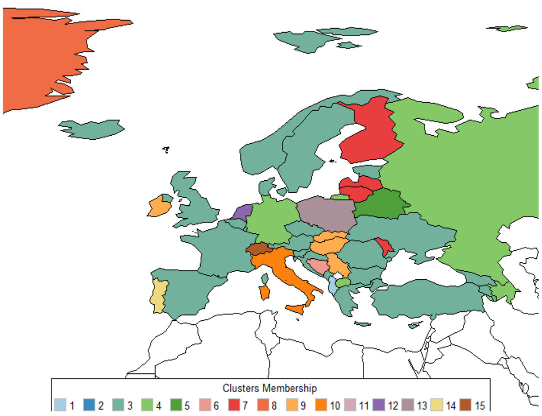


Figure 33 Geographical maps of the clustering results with LPC Distance for Diabetes Mellitus

6. Discussion

The purpose of this study was to model health indicators through MA, ARMA or ARIMA processes; evaluate the quality of fit of the models to the data; and compare the distances between processes regarding their effectiveness in identifying natural groups. Additionally, it was expected to obtain homogeneous and separate clusters and carry out a comparative analysis of different classification methods using the Piccolo distance the Maharaj distance, and the LPC distance, applied to ARIMA models. Firstly, the results of the fit of ARIMA models to the data will be discussed, then the clusters results, considering the clustering methods and the distance measures used, and how these factors can influence clusters results. To the best of our knowledge, these three distance measures were not used to cluster DALYs data before. However, the Piccolo distance, along with the Ward method, was previously applied to cluster rainfall data (45). In another study, the Piccolo distance was used to identify homogeneous housing market areas among districts across five Turkish cities based on house sale price indices (47). Additionally, a study compared 10 distance measures, among them the Piccolo, Maharaj, and LPC distances, using hierarchical clustering and K-means clustering for time series and applied them to power consumption time series as power consumption profiling (52).

Through the exploratory data analysis conducted, it was observed that for the majority of European countries, the DALYs rates for neurological disorders, chronic kidney diseases, and diabetes mellitus have demonstrated a consistent upward trend since 1990. In contrast, the trends for cardiovascular diseases and respiratory diseases appear to be more variable, exhibiting periods of both increase and decline. This suggests that existing health policies may need to be reassessed, and new strategies should be considered to address the growing burden of these non-communicable conditions.

Some important insights obtained from the analysis of ARIMA model fit and performance indicate that higher model complexity, meaning higher values of p, d, and q parameters, does not necessarily correspond to better predictive accuracy. Countries with more complex models did not always exhibit lower prediction errors. More complex ARIMA models are designed to capture more intricate patterns in the time series data, but the risk of overfitting increases with model complexity, which can lead to poor forecasting performance. On the other hand, simpler models might generalize better and give more reliable predictions, especially if there are limitations on the data. This highlights the importance of balancing model complexity with performance, by choosing models based not only on their fit to the data but also on their forecast results, and avoiding overfitting. Additionally, some countries consistently achieved lower prediction errors, such as San Marino and Monaco, while others, like Ukraine, demonstrated poorer performance across most diseases. The varying performance of ARIMA models across countries might also reflect underlying epidemiological, socioeconomic, and political differences. Factors such as data quality, healthcare systems with reliable tracking of diseases, population health behaviors, and disease prevalence could all affect how well a model fits a country's data. For instance, countries with better control of certain diseases or more effective healthcare interventions might show more predictable disease trends, while others with more unstable health dynamics may cause challenges for accurate modeling.

The differences in predictive accuracy between countries suggest that a single approach to health data modeling may not work well for all. Instead, models should be adapted to consider specific factors for each country, like health systems, population changes, and policy actions, to improve forecasting accuracy. Additionally, better data collection and consistent reporting at the national level could make predictive models more reliable, helping to inform health policies more effectively.

Now regarding the clustering results, it is important to emphasize that the function arima.sim in R was used to simulate time series data from a specified ARIMA model. This function allows to generate random time series data based on their parameters (p, d, q). It uses these inputs to simulate a series of data points as if they were generated from the defined ARIMA model. When simulating an ARIMA model in R using the arima.sim function, setting a seed ensures that the random number generation is reproducible, meaning it gives the same results every time the code is run. This is useful for consistency, especially when comparing different models. Using the same seed across multiple ARIMA models ensures that any differences in results are due to the model itself, not random variations.

The results suggest that hierarchical clustering using ARIMA model-based time series, combined with different distance measures, reveals patterns in cardiovascular disease trends across European countries. Across all methods (Ward, Complete, Average, Single, Median, Centroid), Latvia consistently forms its own cluster when using Maharaj, Piccolo, and LPC distances. This indicates that Latvia's cardiovascular disease time series may have unique characteristics that set it apart from other countries. This could be an outlier behavior or a time series with distinct ARIMA parameters. Countries like Armenia, Ireland, Luxembourg, Montenegro frequently group together across the distances and methods. This indicates they share similar cardiovascular time series patterns, perhaps due to similar healthcare systems, population structures, or geographic proximity. The Piccolo distance exhibits more variability in the clustering patterns. It shows that small country clusters emerge, and in many methods, we see clusters that group geographically close or socioeconomically similar countries, for example, Western Europe or Eastern Europe, together. The Piccolo distance might be sensitive to differences in ARIMA model parameters reflecting distinct cardiovascular trends. The Maharaj distance also shows a tendency for small country clusters, but the cluster formation is generally more consistent across methods compared to Piccolo. Countries like Hungary, Slovakia, and Spain consistently form a group, while countries like

Greece, Portugal, and Slovenia also group consistently, suggesting similar temporal dynamics in these countries' cardiovascular diseases. The LPC distance results are similar to Maharaj, though some cluster compositions differ slightly. For instance, Greenland occasionally forms its own cluster, indicating that LPC may capture subtle differences in time series similarities compared to other distances.

The clustering analysis using ARIMA models for chronic respiratory diseases reveals several key insights across different hierarchical clustering methods (Ward, Complete, Average, Single, Median, and Centroid) and distance measures (Piccolo, Maharaj, and LPC). While each distance measure emphasizes different aspects of the time series data, several consistent patterns emerge, alongside notable outliers. Using the Piccolo distance, the Ward method forms large clusters that group many Western and Central European countries, such as France, Greece, Ireland, Switzerland, and the United Kingdom, indicating similar chronic respiratory disease trends across these nations. However, smaller clusters (e.g., Azerbaijan, Cyprus, and Malta) indicate that certain countries display distinct time series patterns. Across other clustering methods, the main clusters remain largely stable, reinforcing the similarities in disease trends among countries. The Piccolo distance also captures subtle differences in smaller or geographically isolated countries, such as Iceland and Greenland, suggesting its sensitivity to small variations in ARIMA model parameters. With the Maharaj distance, the Ward method also identifies a dominant group of countries that resemble the main Piccolo clusters. However, smaller clusters consistently emerge for countries like Armenia and Malta, suggesting these nations experience unique respiratory disease patterns. The Complete and Average methods break these larger groups into smaller clusters, isolating countries like Belgium, Azerbaijan, Denmark, and Hungary, which indicates different temporal dynamics in these countries. Maharaj reveals more consistent clustering across methods, implying it captures broader disease trends while highlighting distinct country-level variations. The LPC distance shares similarities with Maharaj, particularly in the Ward method, where large clusters dominate. However, LPC frequently forms unique small clusters, especially for geographically diverse countries like Greenland, Norway, and the Republic of Moldova. These smaller, isolated clusters appear consistently across different methods, suggesting that LPC effectively captures slight variations in respiratory disease trends, particularly for more remote or less connected nations. This emphasizes the ability of LPC to reflect regional nuances in disease dynamics. Across all distance measures, a core group of countries, including Austria, France, Greece, and Ireland, consistently clusters together, pointing to common trends in chronic respiratory diseases. These similarities may be attributed to shared healthcare systems, lifestyle factors, or environmental conditions. In contrast, countries like Armenia, Malta, and Greenland frequently form distinct clusters, highlighting potential outlier behavior or unique disease patterns. The Piccolo distance exhibits more variability in clustering, reflecting its sensitivity to differences in ARIMA parameters, while Maharaj and LPC provide more stable clustering patterns, capturing both general trends and country-specific variations. Overall, this analysis sheds light on both broad and localized patterns in chronic respiratory disease trends across Europe.

The clustering analysis of neurological disorders using ARIMA models offers some interesting patterns when looking at how countries group together based on their time series trends. First, with the Piccolo distance, it is possible to see that some countries, like Albania, Belarus, Lithuania, and Sweden, often appear in the same group across various clustering methods. This suggests they have similar patterns in neurological disorders. Larger clusters, such as those seen with countries like Andorra, Armenia, the United Kingdom, and Denmark, hint that these countries might share common factors in disease trends. On the other hand, countries like Azerbaijan and Croatia frequently stand out in their own clusters, indicating their patterns are different from others. Looking at the Maharaj distance, the groups tend to be larger and more stable across different methods. For instance, a large cluster includes most European countries, showing that many of these countries share similar disease trends. However, countries like Croatia, Czechia, and Finland regularly form separate groups, suggesting their trends are different from the majority. This distance measure helps highlight broad similarities but still picks up on some unique differences. The LPC distance also groups most countries together in larger clusters, but again, Croatia, Czechia, Estonia, and Azerbaijan often form their own smaller groups, showing distinct differences in their neurological disorder trends. This suggests that while many countries follow similar patterns, a few have unique trends that set them apart. Across all the distance measures, some countries like Azerbaijan, Russia, Croatia, and Finland consistently appear in separate clusters. This indicates that these countries have distinct trends in neurological disorders, possibly due to differences in healthcare systems or other factors. On the other hand, countries like the United Kingdom, France, and Germany often group together, likely reflecting similarities in how they manage and report these disorders. The analysis shows that while many European countries share similar patterns of neurological disorders, some countries consistently stand out with unique trends. This could be due to differences in healthcare, reporting, or regional factors.

In this study, clustering of chronic kidney disease time series data across different European countries using ARIMA models revealed significant insights into how these countries group together based on similarities in their disease patterns. Using the Piccolo distance, clusters varied across methods. Notably, Austria, Poland, and Russia often formed their own distinct clusters, suggesting that these countries have unique patterns in chronic kidney disease trends. In contrast, many countries consistently formed large, similar clusters, such as Albania, Belgium, Croatia, and Spain, indicating similar time series patterns in disease progression. This distance measure was particularly effective at highlighting both large, cohesive groups and smaller, isolated clusters. The Maharaj distance yielded more uniform results across methods, with the majority of countries grouped into a single large cluster, implying that most European countries share similar time series characteristics for chronic kidney diseases. However, Austria and Poland consistently appeared as a separate cluster, suggesting that they exhibit significantly different trends compared to the rest. This distance measure seems to emphasize broad similarities, which can obscure finer differences. The LPC distance, which focuses on pattern similarity, provided more granular distinctions. For example, under the Ward method, countries like Albania, Belgium, and Croatia consistently clustered together, while others, such as Austria and Poland, remained distinct. This distance measure was useful for identifying countries with slightly more unique time series behaviors, such as Azerbaijan, Italy, and Ireland, which sometimes formed smaller, distinct clusters. Across all three distance measures and clustering methods, some countries consistently stood out as having unique chronic kidney disease patterns, particularly Austria and Poland. In contrast, other countries frequently grouped into large, cohesive clusters, such as those in Western and Southern Europe, pointing to common trends in chronic kidney disease across these regions. Ultimately, the choice of distance measure and clustering method influenced the granularity of the clustering results, with Piccolo and LPC offering more specific groupings and Maharaj highlighting broader similarities.

The clustering results of ARIMA modeled diabetes mellitus time series across European countries using various hierarchical clustering methods and distance measures reveal several key insights about the structure of the data. Across different methods like Ward, Complete, and Average, as well as the distance measures used, certain countries tend to consistently cluster together. For instance, Albania frequently appears as its own cluster, which may suggest that the ARIMA model for diabetes in Albania is significantly distinct compared to other countries. Similarly, small groups of countries such as Armenia, Austria, and Belgium, often cluster together, indicating similar ARIMA characteristics for diabetes across these nations, possibly due to shared socio-economic factors or health policies. The choice of distance measure—Piccolo, Maharaj, or LPC—impacts the clustering but not drastically. For example, in Ward's method, the clusters remain relatively stable, though slight changes are observed, such as Latvia and Lithuania being clustered differently under Piccolo compared to Maharaj. This suggests that the structural differences in time series are somewhat robust across different distance measures, though certain specific relationships are sensitive to the measure applied. The use of various clustering methods like Ward (which tends to form compact clusters) versus Single (which forms long, chain-like clusters) affects how countries are grouped. In Single linkage, we observe larger, more expansive clusters, such as when Andorra and Russia cluster with Finland under Maharaj distance. Ward, on the other hand, creates more precise clusters, indicating its suitability when more defined groupings are desired. This suggests that researchers need to carefully choose the method depending on whether they prioritize clear, compact clusters or larger, more inclusive groupings. Clusters such as Belgium, Croatia, Cyprus, and Greece tend to stay together across most clustering methods and distance measures. This stability suggests strong underlying similarities in the diabetes trends across these countries, potentially driven by shared cultural, dietary, or healthcare-related factors. On the contrary, countries like Russia and North Macedonia tend to shift clusters, pointing to more variable time-series behavior. Countries like Greenland, Malta, and Switzerland often form their own clusters or very small groups, regardless of method or distance measure. This suggests that their time series patterns are unique and not closely aligned with other European nations, which could be due to geographic isolation or distinct healthcare systems. The results demonstrate that while there is some consistency in how countries are grouped based on diabetes time-series models, the clustering is sensitive to the method and distance measure used. This highlights the importance of methodological choice in clustering analysis. Closer examination of the consistently clustered countries may reveal shared health policies, environmental factors, or similar social conditions that influence diabetes trends. At the same time, the distinct clusters for countries like Albania and Greenland suggest the importance of local factors in shaping their diabetes profiles, which may warrant country-specific healthcare interventions.

7. Conclusions and Future Work

The analysis of time series data on various diseases (cardiovascular, respiratory, neurological, kidney diseases, and diabetes) across European countries reveals several broad insights into disease trends and healthcare systems. Many countries consistently cluster together, indicating similar patterns in disease progression. For instance, countries like Austria, France, Greece, Ireland, and Spain frequently group, which suggests shared factors such as healthcare systems, environmental conditions, or socio-economic similarities driving these trends. On the other hand, countries like Latvia, Poland, and Greenland often form distinct clusters, indicating unique health patterns that set them apart from the rest of Europe.

Geographic proximity also plays a role in clustering. Countries within the same region, such as Western or Eastern Europe, often group together due to shared healthcare systems, lifestyle factors, or regional disease trends. However, outlier countries that consistently form their own clusters, reflect distinct ARIMA models behavior and consequently disease dynamics, possibly due to isolation, differences in healthcare quality, or specific public health challenges.

The different sensitivity of distance measures highlights how methodological choices impact the clustering results. Therefore, to deepen our understanding of how clustering results are influenced by distance measures, it is important to carry out further studies involving a variety of distance metrics. In previous studies on non-communicable diseases mortality, the use of Euclidean distance and hierarchical clustering in identifying disease clusters in Portuguese mortality data provides a foundation for further exploration of clustering methods in time series data. Building on these insights, future studies could explore the integration of time series ARIMA models with Euclidean and other distance metrics to investigate temporal patterns in DALYs. Additionally, examining gender-specific patterns and other risk factors, as in Nascimento's study, may further improve the understanding of disease burden trends and contribute to more targeted public health interventions.

The choice of distance measure can significantly affect the way time series are grouped, as each measure captures different aspects of the temporal data, such as trend, seasonality, or noise. By exploring alternative distance measures, we can better understand how these subtle variations in data similarity impact the formation and interpretation of clusters. Collaboration with experts, particularly in the field of public health, is essential to ensure that the clusters formed reflect meaningful patterns in disease trends, assisting in the development of targeted health intervention and policies. Additionally, the selection of clustering algorithms, whether hierarchical, partition-based, or density-based methods, also plays a crucial role in determining the clustering outcomes. Different algorithms may have varying sensitivities to time series characteristics such as noise, outliers, or time lags. Hence, testing a broader range of methods will help reveal the strengths and limitations of each approach when applied to time series data. Moreover, employing a variety of time series datasets, representing different domains such as financial markets, environmental data, or health data, would provide a more comprehensive view of how well certain distance measures and clustering methods generalize across different types of data. Some distance measures might perform well in specific contexts, such as short-term or periodic series, while others may be more effective in capturing long-term trends. A fundamental takeaway from this study is that countries with similar ARIMA patterns may benefit from shared public health strategies and collaborative research efforts. Meanwhile, countries that consistently appear in distinct clusters might require more tailored healthcare interventions due to their unique disease dynamics. Overall, this study shows that time series clustering can uncover important regional and country-specific trends in disease progression. It emphasizes the need for careful selection of clustering methods and distance measures to capture both broad patterns and subtle differences in health trends across Europe.

References

1. Zeger SL, Irizarry R, Peng RD. On time series analysis of public health and biomedical data. Vol. 27, Annual Review of Public Health. 2006. p. 57–79.
2. Mathers CD. History of global burden of disease assessment at the World Health Organization. Archives of Public Health. 2020 Aug 24;78(1).
3. Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 (GBD 2019). Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2020.
4. Vandenberghe D, Albrecht J. The financial burden of non-communicable diseases in the European Union: A systematic review. Vol. 30, European Journal of Public Health. Oxford University Press; 2020. p. 833–9.
5. Choudhury MJH, Chowdhury TI, Nayeem A, Jahan WA. Modifiable and Non-Modifiable Risk Factors of Stroke: A Review Update. Vol. 1, J Natl Inst Neurosci Bangladesh. 2015.
6. Jian Y, Zhu D, Zhou D, Li N, Du H, Dong X, et al. ARIMA model for predicting chronic kidney disease and estimating its economic burden in China. BMC Public Health. 2022 Dec 1;22(1).
7. Aghabozorgi S, Seyed Shirkhorshidi A, Ying Wah T. Time-series clustering - A decade review. Inf Syst. 2015 May 30;53:16–38.
8. Piccolo D. A Distance Measure For Classifying Arima Models. J Time Ser Anal. 1990 Mar 28;11(2):153–64.
9. Maharaj EA. A significance test for classifying ARMA models. J Stat Comput Simul. 1996 Jun;54(4):305–31.
10. Maharaj EA. Cluster of Time Series. J Classif. 2000 Jul 1;17(2):297–314.
11. Kalpakis K, Gada D, Puttagunta V. Distance measures for effective clustering of ARIMA time-series. In: Proceedings 2001 IEEE International Conference on Data Mining [Internet]. IEEE Comput. Soc; 2001. p. 273–80. Available from: https://ieeexplore.ieee.org/document/989529
12. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet. 2020 Oct;396(10258):1204–22. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673620309259
13. von der Lippe E, Devleesschauwer B, Gourley M, Haagsma J, Hilderink H, Porst M, et al. Reflections on key methodological decisions in national burden of disease assessments. Archives of Public Health. 2020 Dec 1;78(1).
14. Murray CJL. The Global Burden of Disease Study at 30 years. Nat Med. 2022 Oct 1;28(10):2019–26.
15. Benziger CP, Roth GA, Moran AE. The Global Burden of Disease Study and the Preventable Burden of NCD. Vol. 11, Global Heart. Elsevier B.V.; 2016. p. 393–7.

16. Budreviciute A, Damiati S, Sabir DK, Onder K, Schuller-Goetzburg P, Plakys G, et al. Management and Prevention Strategies for Non-communicable Diseases (NCDs) and Their Risk Factors. Vol. 8, Frontiers in Public Health. Frontiers Media S.A.; 2020.
17. World Health Organization. Noncommunicable Diseases (NCD). (2019). Available online at: https://www.who.int/gho/ncd/mortality_morbidity/en/.
18. Muka T, Imo D, Jaspers L, Colpani V, Chaker L, van der Lee SJ, et al. The global impact of non-communicable diseases on healthcare spending and national income: a systematic review. Vol. 30, European Journal of Epidemiology. Kluwer Academic Publishers; 2015. p. 251–77.
19. Nielsen A. Practical Time Series Analysis. 1st Edition. Sebastopol, CA: O'Reilly Media Inc.; 2019.
20. Montgomery DC, Jennings CL, Kulahci M. Introduction to Time Series Analysis and Forecasting. 2015.
21. Bhatti UA, Yan Y, Zhou M, Ali S, Hussain A, Qingsong H, et al. Time Series Analysis and Forecasting of Air Pollution Particulate Matter (PM2.5): An SARIMA and Factor Analysis Approach. IEEE Access. 2021;9:41019–31.
22. Gautam A, Singh V. Parametric versus non-parametric time series forecasting methods: A review. Vol. 13, Journal of Engineering Science and Technology Review. Eastern Macedonia and Thrace Institute of Technology; 2020. p. 165–71.
23. Hyndman RJ, Athanasopoulos G. Forecasting: principles and practice. OTexts; 2018.
24. Schaffer AL, Dobbins TA, Pearson SA. Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions. BMC Med Res Methodol. 2021 Dec 1;21(1).
25. Palma W. Time Series Analysis. Hoboken, New Jersey: John Wiley & Sons, Inc.; 2016.
26. Chiş M, Banerjee S, Hassanien AE. Clustering Time Series Data: An Evolutionary Approach. In 2009. p. 193–207. Available from: http://link.springer.com/10.1007/978-3-642-01091-0_9
27. Ergüner Özkoç E. Clustering of Time-Series Data. In: Data Mining - Methods, Applications and Systems. IntechOpen; 2021.
28. Saxena A, Prasad M, Gupta A, Bharill N, Patel OP, Tiwari A, et al. A review of clustering techniques and developments. Neurocomputing. 2017 Dec;267:664–81. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0925231217311815
29. Omran MGH, Engelbrecht AP, Salman A. An overview of clustering methods. Intelligent Data Analysis. 2007 Nov 9;11(6):583–605.
30. Jarman AM. Hierarchical Cluster Analysis: Comparison of Single linkage,Complete linkage, Average linkage and Centroid Linkage Method. Int J Georg Soutern Univ. 2020; Available from: https://www.researchgate.net/publication/339443595
31. Müllner D. Modern hierarchical, agglomerative clustering algorithms. 2011 Sep 12;

32. Deng D. DBSCAN Clustering Algorithm Based on Density. In: 2020 7th International Forum on Electrical Engineering and Automation (IFEEA). IEEE; 2020. p. 949–53.

33. Ankerst M, Breunig MM, Kriegel HP, Sander J. OPTICS: ordering points to identify the clustering structure. ACM SIGMOD Record. 1999 Jun;28(2):49–60.

34. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987 Nov;20:53–65.

35. Lin Z, Laska E, Siegel C. A general iterative clustering algorithm. Statistical Analysis and Data Mining: The ASA Data Science Journal. 2022 Aug 31;15(4):433–46.

36. Shahapure KR, Nicholas C. Cluster Quality Analysis Using Silhouette Score. In: 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA). IEEE; 2020. p. 747–8.

37. Dunn JC. Well-Separated Clusters and Optimal Fuzzy Partitions. Journal of Cybernetics. 1974 Jan 30;4(1):95–104.

38. Charrad M, Ghazzali N, Laval U, Niknafs A. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set Véronique Boiteau. Vol. 61, JSS Journal of Statistical. 2014. Available from: http://www.jstatsoft.org/

39. Hubert LJ, Levin JR. A general statistical framework for assessing categorical clustering in free recall. Psychol Bull. 1976 Nov;83(6):1072–80.

40. Mcclain JO, Rao VR. CLUSTISZ: A Program to Test for the Quality of Clustering of a Set of Objects. Vol. 12, Source: Journal of Marketing Research. 1975. Available from: http://www.jstor.orgURL:http://www.jstor.org/stable/3151097

41. Montero P, Vilar JA. TSclust: An R Package for Time Series Clustering [Internet]. Vol. 62, JSS Journal of Statistical Software. 2014. Available from: http://www.jstatsoft.org/

42. Kardiyen F, Güney H. A Study on Time Series Clustering. Vol. 28, Gazi University Journal of Science GU J Sci. 2015.

43. Nascimento AP, Oliveira A, Faria BM, Pimenta R, Vieira M, Prudêncio C, et al. Affinity Coefficient for Clustering Autoregressive Moving Average Models. Computational and Mathematical Methods. 2024;2024.

44. Nascimento AP, Prudêncio C, Vieira M, Pimenta R, Bacelar-Nicolau H. A typological study of Portuguese mortality from non-communicable diseases. Advances in Science, Technology and Engineering Systems. 2020;5(5):613–9.

45. Hendrawati T, Wigena AH, Sumertajaya IM, Sartono B. Clustering rainfall data using piccolo method. In: AIP Conference Proceedings. American Institute of Physics Inc.; 2020.

46. Hendrawati T, Wigena AH, Sumertajaya IM, Sartono B. Clustering of Commodity Inflation Pattern based on Estimated ARIMA Model. In: Journal of Physics: Conference Series. IOP Publishing Ltd; 2021.

47. Vatansever M, Demir İ, Hepşen A. Cluster and forecasting analysis of the residential market in Turkey: An autoregressive model-based fuzzy clustering approach. International Journal of Housing Markets and Analysis. 2020 Jun 16;13(4):583–600.
48. Institute for Health Metrics and Evaluation (IHME). GBD Compare Data Visualization. Seattle, WA: IHME, University of Washington, 2020. Available from http://vizhub.healthdata.org/gbd-compare. (Accessed 2023-11-28).
49. Institute for Health Metrics and Evaluation (IHME). GBD Results. Seattle, WA: IHME, University of Washington, 2020. Available from https://vizhub.healthdata.org/gbd-results. (Accessed 2023-11-28).
50. Berrar D. Cross-Validation. In: Encyclopedia of Bioinformatics and Computational Biology. Elsevier; 2019. p. 542–5. Available from: https://linkinghub.elsevier.com/retrieve/pii/B978012809633820349X
51. Smith, Taylor G., et al. pmdarima: ARIMA estimators for Python, 2017-, http://www.alkaline-ml.com/pmdarima [Online; accessed 2024-05-06].
52. Kim J, Kim J. Comparison of time series clustering methods and application to power consumption pattern clustering. Commun Stat Appl Methods. 2020;27(6):589–602.

Annex A

Cardiovascular Diseases – Piccolo Distance

For the Ward method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Romania, San Marino;
- Cluster 3: Armenia, Austria, Ireland, Luxembourg, Montenegro;
- Cluster 4: Azerbaijan, Belgium, Bosnia and Herzegovina, Croatia, Italy, Republic of Moldova, Serbia, Switzerland;
- Cluster 5: Belarus, Denmark, Finland, France, Monaco, North Macedonia, Norway, United Kingdom;
- Cluster 6: Bulgaria;
- Cluster 7: Cyprus, Estonia, Georgia, Iceland, Malta, Poland, Russia, Sweden, Turkey, Ukraine;
- Cluster 8: Czechia;
- Cluster 9: Germany;
- Cluster 10: Greece;
- Cluster 11: Greenland, Netherlands;
- Cluster 12: Hungary, Slovakia, Spain;
- Cluster 13: Latvia;
- Cluster 14: Lithuania;
- Cluster 15: Portugal, Slovenia.

For the Complete method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Romania;
- Cluster 3: Armenia, Austria, Ireland, Luxembourg, Montenegro;
- Cluster 4: Azerbaijan, Belgium, Bosnia and Herzegovina, Croatia, Italy, Republic of Moldova, Serbia, Switzerland;
- Cluster 5: Belarus, Cyprus, Denmark, Estonia, Finland, France, Georgia, Iceland, Malta, Monaco, North Macedonia, Norway, Poland, Russia, Sweden, Turkey, Ukraine, United Kingdom;
- Cluster 6: Bulgaria;
- Cluster 7: Czechia;
- Cluster 8: Germany;
- Cluster 9: Greece;
- Cluster 10: Greenland, Netherlands;

- Cluster 11: Hungary, Slovakia, Spain;
- Cluster 12: Latvia;
- Cluster 13: Lithuania;
- Cluster 14: Portugal, Slovenia;
- Cluster 15: San Marino.

For the Average and Single methods, the 2 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Austria, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Lithuania, Luxembourg, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Republic of Moldova, Romania, Russia, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom;
- Cluster 2: Latvia.

For the Median method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra;
- Cluster 3: Armenia, Austria, Ireland, Luxembourg, Montenegro;
- Cluster 4: Azerbaijan, Belarus, Belgium, Croatia, Cyprus, Denmark, Estonia, Finland, France, Georgia, Iceland, Italy, Malta, Monaco, North Macedonia, Norway, Poland, Republic of Moldova, Russia, Serbia, Sweden, Switzerland, Turkey, Ukraine, United Kingdom;
- Cluster 5: Bosnia and Herzegovina, Hungary, Portugal, Slovakia, Slovenia, Spain;
- Cluster 6: Bulgaria;
- Cluster 7: Czechia;
- Cluster 8: Germany;
- Cluster 9: Greece;
- Cluster 10: Greenland;
- Cluster 11: Latvia;
- Cluster 12: Lithuania;
- Cluster 13: Netherlands;
- Cluster 14: Romania;
- Cluster 15: San Marino.

For the Centroid method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra;

- Cluster 3: Armenia, Austria, Ireland, Luxembourg, Montenegro;
- Cluster 4: Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Croatia, Cyprus, Denmark, Estonia, Finland, France, Georgia, Hungary, Iceland, Italy, Malta, Monaco, North Macedonia, Norway, Poland, Republic of Moldova, Russia, Serbia, Slovakia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom;
- Cluster 5: Bulgaria;
- Cluster 6: Czechia;
- Cluster 7: Germany;
- Cluster 8: Greece;
- Cluster 9: Greenland;
- Cluster 10: Latvia;
- Cluster 11: Lithuania;
- Cluster 12: Netherlands;
- Cluster 13: Portugal, Slovenia;
- Cluster 14: Romania;
- Cluster 15: San Marino.

Cardiovascular Diseases – Maharaj Distance

For the Ward method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Romania;
- Cluster 3: Armenia, Ireland, Luxembourg, Montenegro;
- Cluster 4: Austria, Netherlands;
- Cluster 5: Azerbaijan, Belgium, Bosnia and Herzegovina, Croatia, Italy, Republic of Moldova, Serbia, Switzerland;
- Cluster 6: Belarus, Denmark, Finland, France, Monaco, North Macedonia, Norway, United Kingdom;
- Cluster 7: Bulgaria, San Marino;
- Cluster 8: Cyprus, Estonia, Georgia, Iceland, Malta, Poland, Russia, Sweden, Turkey, Ukraine;
- Cluster 9: Czechia;
- Cluster 10: Germany;
- Cluster 11: Greece, Portugal, Slovenia;
- Cluster 12: Greenland;
- Cluster 13: Hungary, Slovakia, Spain;

- Cluster 14: Latvia;
- Cluster 15: Lithuania.

For the Complete method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Romania;
- Cluster 3: Armenia, Ireland, Luxembourg, Montenegro;
- Cluster 4: Austria, Netherlands;
- Cluster 5: Azerbaijan, Belgium, Bosnia and Herzegovina, Croatia, Italy, Republic of Moldova, Serbia, Switzerland;
- Cluster 6: Belarus, Denmark, Finland, France, Georgia, Malta, Monaco, North Macedonia, Norway, Russia, United Kingdom;
- Cluster 7: Bulgaria, San Marino;
- Cluster 8: Cyprus, Estonia, Iceland, Poland, Sweden, Turkey, Ukraine;
- Cluster 9: Czechia;
- Cluster 10: Germany;
- Cluster 11: Greece;
- Cluster 12: Greenland;
- Cluster 13: Hungary, Portugal, Slovakia, Slovenia, Spain;
- Cluster 14: Latvia;
- Cluster 15: Lithuania.

For the Average, Single, Median, and Centroid methods the 2 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Austria, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Lithuania, Luxembourg, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Republic of Moldova, Romania, Russia, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom;
- Cluster 2: Latvia.

Cardiovascular Diseases – LPC Distance

For the Ward method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Romania;
- Cluster 3: Armenia, Ireland, Luxembourg, Montenegro;

- Cluster 4: Austria, Netherlands;
- Cluster 5: Azerbaijan, Belgium, Bosnia and Herzegovina, Croatia, Italy, Republic of Moldova, Serbia, Switzerland;
- Cluster 6: Belarus, Denmark, Finland, France, Monaco, North Macedonia, Norway, United Kingdom;
- Cluster 7: Bulgaria, San Marino;
- Cluster 8: Cyprus, Estonia, Georgia, Iceland, Malta, Poland, Russia, Sweden, Turkey, Ukraine;
- Cluster 9: Czechia;
- Cluster 10: Germany;
- Cluster 11: Greece, Portugal, Slovenia;
- Cluster 12: Greenland;
- Cluster 13: Hungary, Slovakia, Spain;
- Cluster 14: Latvia;
- Cluster 15: Lithuania.

For the Complete method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Romania;
- Cluster 3: Armenia, Ireland, Luxembourg, Montenegro;
- Cluster 4: Austria, Netherlands;
- Cluster 5: Azerbaijan, Belgium, Bosnia and Herzegovina, Croatia, Italy, Republic of Moldova, Serbia, Switzerland;
- Cluster 6: Belarus, Denmark, Finland, France, Georgia, Malta, Monaco, North Macedonia, Norway, Russia, United Kingdom;
- Cluster 7: Bulgaria, San Marino;
- Cluster 8: Cyprus, Estonia, Iceland, Poland, Sweden, Turkey, Ukraine;
- Cluster 9: Czechia;
- Cluster 10: Germany;
- Cluster 11: Greece;
- Cluster 12: Greenland;
- Cluster 13: Hungary, Portugal, Slovakia, Slovenia, Spain;
- Cluster 14: Latvia;
- Cluster 15: Lithuania.

For the Single, Average, Median, Centroid methods the 2 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Austria, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Lithuania, Luxembourg, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Republic of Moldova, Romania, Russia, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom;
- Cluster 2: Latvia.

Annex B

Chronic Respiratory Diseases – Piccolo Distance

For the Ward method the 8 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, Romania, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia, Bulgaria, Denmark, Georgia, Iceland, Monaco, Norway, Portugal, Moldova, Russia, Turkey;
- Cluster 3: Azerbaijan, Cyprus, Malta;
- Cluster 4: Belgium, Netherlands;
- Cluster 5: Bosnia and Herzegovina, Greenland, Hungary;
- Cluster 6: Croatia, Poland;
- Cluster 7: Czechia, Germany, Slovenia, Sweden;
- Cluster 8: North Macedonia.

For the Complete method the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, Romania, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia, Bulgaria, Iceland, Portugal;
- Cluster 3: Azerbaijan, Cyprus;
- Cluster 4: Belgium, Netherlands;
- Cluster 5: Bosnia and Herzegovina, Greenland;
- Cluster 6: Croatia;
- Cluster 7: Czechia, Germany, Sweden;
- Cluster 8: Denmark;
- Cluster 9: Georgia, Turkey;
- Cluster 10: Hungary;
- Cluster 11: Malta;
- Cluster 12: Monaco, Norway, Republic of Moldova, Russia;
- Cluster 13: North Macedonia;
- Cluster 14: Poland;
- Cluster 15: Slovenia.

For the Average and Median methods the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, Romania, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia, Bulgaria, Iceland, Monaco, Norway, Portugal, Republic of Moldova, Russia;
- Cluster 3: Azerbaijan, Georgia, Turkey;
- Cluster 4: Belgium, Netherlands;
- Cluster 5: Bosnia and Herzegovina;
- Cluster 6: Croatia;
- Cluster 7: Cyprus;
- Cluster 8: Czechia, Germany, Sweden;
- Cluster 9: Denmark;
- Cluster 10: Greenland;
- Cluster 11: Hungary;
- Cluster 12: Malta;
- Cluster 13: North Macedonia;
- Cluster 14: Poland;
- Cluster 15: Slovenia.

For the Single and Centroid methods the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, Romania, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia, Bulgaria, Georgia, Iceland, Monaco, Norway, Portugal, Republic of Moldova, Russia, Turkey;
- Cluster 3: Azerbaijan;
- Cluster 4: Belgium, Netherlands;
- Cluster 5: Bosnia and Herzegovina;
- Cluster 6: Croatia;
- Cluster 7: Cyprus;
- Cluster 8: Czechia, Germany, Sweden;
- Cluster 9: Denmark;
- Cluster 10: Greenland;
- Cluster 11: Hungary;
- Cluster 12: Malta;
- Cluster 13: North Macedonia;

- Cluster 14: Poland;
- Cluster 15: Slovenia.

Chronic Respiratory Diseases – Maharaj Distance

For the Ward method the 9 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, Romania, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia, Malta;
- Cluster 3: Azerbaijan, Cyprus, Georgia, Monaco, Russia;
- Cluster 4: Belgium, Czechia, Germany, Slovenia, Sweden;
- Cluster 5: Bosnia and Herzegovina, Greenland;
- Cluster 6: Bulgaria, Iceland, Norway, Portugal, Republic of Moldova, Turkey;
- Cluster 7: Croatia, Poland;
- Cluster 8: Denmark, Hungary, Netherlands;
- Cluster 9: North Macedonia.

For the Complete method the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, Romania, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia;
- Cluster 3: Azerbaijan;
- Cluster 4: Belgium;
- Cluster 5: Bosnia and Herzegovina, Greenland;
- Cluster 6: Bulgaria, Iceland, Portugal;
- Cluster 7: Croatia;
- Cluster 8: Cyprus;
- Cluster 9: Czechia, Germany, Slovenia;
- Cluster 10: Denmark, Hungary, Netherlands;
- Cluster 11: Georgia, Monaco, Russia, Sweden;
- Cluster 12: Malta;
- Cluster 13: North Macedonia;
- Cluster 14: Norway, Republic of Moldova, Turkey;

- Cluster 15: Poland.

For the Average method the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, Romania, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia;
- Cluster 3: Azerbaijan;
- Cluster 4: Belgium;
- Cluster 5: Bosnia and Herzegovina, Greenland;
- Cluster 6: Bulgaria, Iceland, Norway, Portugal, Republic of Moldova, Turkey;
- Cluster 7: Croatia;
- Cluster 8: Cyprus;
- Cluster 9: Czechia, Germany, Slovenia, Sweden;
- Cluster 10: Denmark, Hungary;
- Cluster 11: Georgia, Monaco, Russia;
- Cluster 12: Malta;
- Cluster 13: Netherlands;
- Cluster 14: North Macedonia;
- Cluster 15: Poland.

For the Single method the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Georgia, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Monaco, Montenegro, Romania, Russia, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia;
- Cluster 3: Azerbaijan;
- Cluster 4: Belgium;
- Cluster 5: Bosnia and Herzegovina, Greenland;
- Cluster 6: Bulgaria, Iceland, Norway, Portugal, Republic of Moldova, Turkey;
- Cluster 7: Croatia;
- Cluster 8: Cyprus;
- Cluster 9: Czechia, Germany, Slovenia, Sweden;
- Cluster 10: Denmark;
- Cluster 11: Hungary;
- Cluster 12: Malta;

- Cluster 13: Netherlands;
- Cluster 14: North Macedonia;
- Cluster 15: Poland.

For the Median method the 14 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, Romania, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia;
- Cluster 3: Azerbaijan;
- Cluster 4: Belgium;
- Cluster 5: Bosnia and Herzegovina, Greenland;
- Cluster 6: Bulgaria, Iceland, Norway, Portugal, Republic of Moldova, Turkey;
- Cluster 7: Croatia;
- Cluster 8: Cyprus;
- Cluster 9: Czechia, Germany, Slovenia;
- Cluster 10: Denmark, Hungary, Netherlands;
- Cluster 11: Georgia, Monaco, Russia, Sweden;
- Cluster 12: Malta;
- Cluster 13: North Macedonia;
- Cluster 14: Poland.

For the Centroid method the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, Romania, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia;
- Cluster 3: Azerbaijan;
- Cluster 4: Belgium;
- Cluster 5: Bosnia and Herzegovina, Greenland;
- Cluster 6: Bulgaria, Iceland, Norway, Portugal, Republic of Moldova, Turkey;
- Cluster 7: Croatia;
- Cluster 8: Cyprus, Georgia, Monaco, Russia;
- Cluster 9: Czechia, Germany, Slovenia, Sweden;
- Cluster 10: Denmark;
- Cluster 11: Hungary;

- Cluster 12: Malta;
- Cluster 13: Netherlands;
- Cluster 14: North Macedonia;
- Cluster 15: Poland.

Chronic Respiratory Diseases – LPC Distance

For the Ward method the 15 clusters were the following:

- Cluster 1: Albania, Austria, Estonia, France, Latvia, Lithuania, Romania, Slovakia;
- Cluster 2: Andorra, Finland, Italy, Montenegro, Serbia, Spain, Switzerland;
- Cluster 3: Armenia;
- Cluster 4: Azerbaijan, Cyprus;
- Cluster 5: Belarus, Greece, Ireland, Luxembourg, San Marino, Ukraine, United Kingdom;
- Cluster 6: Belgium;
- Cluster 7: Bosnia and Herzegovina, Greenland;
- Cluster 8: Bulgaria, Iceland, Portugal;
- Cluster 9: Croatia, Poland;
- Cluster 10: Czechia, Germany, Slovenia, Sweden;
- Cluster 11: Denmark, Hungary, Netherlands;
- Cluster 12: Georgia, Monaco, Russia;
- Cluster 13: Malta;
- Cluster 14: North Macedonia;
- Cluster 15: Norway, Republic of Moldova, Turkey.

For the Complete method the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, Romania, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia;
- Cluster 3: Azerbaijan;
- Cluster 4: Belgium;
- Cluster 5: Bosnia and Herzegovina, Greenland;
- Cluster 6: Bulgaria, Iceland, Portugal;
- Cluster 7: Croatia;
- Cluster 8: Cyprus;
- Cluster 9: Czechia, Germany, Slovenia;

- Cluster 10: Denmark, Hungary, Netherlands;
- Cluster 11: Georgia, Monaco, Russia, Sweden;
- Cluster 12: Malta;
- Cluster 13: North Macedonia;
- Cluster 14: Norway, Republic of Moldova, Turkey;
- Cluster 15: Poland.

For the Average method the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, Romania, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia;
- Cluster 3: Azerbaijan;
- Cluster 4: Belgium;
- Cluster 5: Bosnia and Herzegovina, Greenland;
- Cluster 6: Bulgaria, Iceland, Norway, Portugal, Republic of Moldova, Turkey;
- Cluster 7: Croatia;
- Cluster 8: Cyprus;
- Cluster 9: Czechia, Germany, Slovenia, Sweden;
- Cluster 10: Denmark, Hungary;
- Cluster 11: Georgia, Monaco, Russia;
- Cluster 12: Malta;
- Cluster 13: Netherlands;
- Cluster 14: North Macedonia;
- Cluster 15: Poland.

For the Single method the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Georgia, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Monaco, Montenegro, Romania, Russia, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia;
- Cluster 3: Azerbaijan;
- Cluster 4: Belgium;
- Cluster 5: Bosnia and Herzegovina, Greenland;
- Cluster 6: Bulgaria, Iceland, Norway, Portugal, Republic of Moldova, Turkey;
- Cluster 7: Croatia;

- Cluster 8: Cyprus;
- Cluster 9: Czechia, Germany, Slovenia, Sweden;
- Cluster 10: Denmark;
- Cluster 11: Hungary;
- Cluster 12: Malta;
- Cluster 13: Netherlands;
- Cluster 14: North Macedonia;
- Cluster 15: Poland.

For the Median and Centroid methods the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Austria, Belarus, Estonia, Finland, France, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, Romania, San Marino, Serbia, Slovakia, Spain, Switzerland, Ukraine, United Kingdom;
- Cluster 2: Armenia;
- Cluster 3: Azerbaijan;
- Cluster 4: Belgium;
- Cluster 5: Bosnia and Herzegovina, Greenland;
- Cluster 6: Bulgaria, Iceland, Norway, Portugal, Republic of Moldova, Turkey;
- Cluster 7: Croatia;
- Cluster 8: Cyprus;
- Cluster 9: Czechia, Georgia, Germany, Monaco, Russia, Slovenia, Sweden;
- Cluster 10: Denmark;
- Cluster 11: Hungary;
- Cluster 12: Malta;
- Cluster 13: Netherlands;
- Cluster 14: North Macedonia;
- Cluster 15: Poland.

Annex C

Neurological Disorders – Piccolo Distance

For the Ward method the 15 clusters were the following:

- Cluster 1: Albania, Belarus, Lithuania, Sweden;
- Cluster 2: Andorra, Armenia, Bulgaria, Denmark, Georgia, Hungary, Montenegro, Netherlands, Republic of Moldova, Ukraine, United Kingdom;
- Cluster 3: Austria, Greece, Ireland, Italy, Malta, Monaco, North Macedonia, Poland, Portugal, Romania, Slovenia;
- Cluster 4: Azerbaijan;
- Cluster 5: Belgium, Iceland, Latvia, Spain, Switzerland;
- Cluster 6: Bosnia and Herzegovina, Germany, Greenland, Slovakia;
- Cluster 7: Croatia;
- Cluster 8: Cyprus;
- Cluster 9: Czechia;
- Cluster 10: Estonia;
- Cluster 11: Finland;
- Cluster 12: France, Luxembourg, Norway, San Marino;
- Cluster 13: Russia;
- Cluster 14: Serbia;
- Cluster 15: Turkey.

For the Complete and Average methods the 15 clusters were the following:

- Cluster 1: Albania, Belarus, Lithuania, Sweden;
- Cluster 2: Andorra, Armenia, Austria, Bulgaria, Denmark, Georgia, Greece, Hungary, Ireland, Italy, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Poland, Portugal, Republic of Moldova, Romania, Slovenia, Ukraine, United Kingdom;
- Cluster 3: Azerbaijan;
- Cluster 4: Belgium, Iceland, Latvia, Spain, Switzerland;
- Cluster 5: Bosnia and Herzegovina, Germany, Slovakia;
- Cluster 6: Croatia;
- Cluster 7: Cyprus;
- Cluster 8: Czechia;
- Cluster 9: Estonia;
- Cluster 10: Finland;

- Cluster 11: France, Luxembourg, Norway, San Marino;
- Cluster 12: Greenland;
- Cluster 13: Russia;
- Cluster 14: Serbia;
- Cluster 15: Turkey.

For the Single method the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Austria, Belarus, Bulgaria, Denmark, Georgia, Greece, Hungary, Ireland, Italy, Lithuania, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Poland, Portugal, Republic of Moldova, Romania, Slovenia, Sweden, Ukraine, United Kingdom;
- Cluster 2: Azerbaijan;
- Cluster 3: Belgium, France, Iceland, Latvia, Norway, San Marino, Spain, Switzerland;
- Cluster 4: Bosnia and Herzegovina, Slovakia;
- Cluster 5: Croatia;
- Cluster 6: Cyprus;
- Cluster 7: Czechia;
- Cluster 8: Estonia;
- Cluster 9: Finland;
- Cluster 10: Germany;
- Cluster 11: Greenland;
- Cluster 12: Luxembourg;
- Cluster 13: Russia;
- Cluster 14: Serbia;
- Cluster 15: Turkey.

For the Median method the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Austria, Belarus, Bulgaria, Denmark, Georgia, Greece, Hungary, Ireland, Italy, Lithuania, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Poland, Portugal, Republic of Moldova, Romania, Slovenia, Sweden, Ukraine, United Kingdom;
- Cluster 2: Azerbaijan;
- Cluster 3: Belgium, France, Iceland, Latvia, Luxembourg, Norway, San Marino, Spain, Switzerland;
- Cluster 4: Bosnia and Herzegovina;
- Cluster 5: Croatia;
- Cluster 6: Cyprus;
- Cluster 7: Czechia;
- Cluster 8: Estonia;

- Cluster 9: Finland;
- Cluster 10: Germany;
- Cluster 11: Greenland;
- Cluster 12: Russia;
- Cluster 13: Serbia;
- Cluster 14: Slovakia;
- Cluster 15: Turkey.

For the Centroid method the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Austria, Belarus, Bulgaria, Denmark, Georgia, Greece, Hungary, Ireland, Italy, Lithuania, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Poland, Portugal, Republic of Moldova, Romania, Slovenia, Sweden, Ukraine, United Kingdom;
- Cluster 2: Azerbaijan;
- Cluster 3: Belgium, Iceland, Latvia, Spain, Switzerland;
- Cluster 4: Bosnia and Herzegovina, Germany, Slovakia;
- Cluster 5: Croatia;
- Cluster 6: Cyprus;
- Cluster 7: Czechia;
- Cluster 8: Estonia;
- Cluster 9: Finland;
- Cluster 10: France, Norway, San Marino;
- Cluster 11: Greenland;
- Cluster 12: Luxembourg;
- Cluster 13: Russia;
- Cluster 14: Serbia;
- Cluster 15: Turkey.

Neurological Disorders – Maharaj Distance

For the Ward, Complete, Single, Average, Centroid, and Median methods the 3 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Austria, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Cyprus, Denmark, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Republic of Moldova, Romania, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom;
- Cluster 2: Azerbaijan, Estonia, Russia;

- Cluster 3: Croatia, Czechia, Finland.

Neurological Disorders – LPC Distance

For the Ward method the 2 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Austria, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Cyprus, Denmark, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Republic of Moldova, Romania, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom;
- Cluster 2: Azerbaijan, Croatia, Czechia, Estonia, Finland, Russia.

For the Complete method the 3 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Austria, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Cyprus, Denmark, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Republic of Moldova, Romania, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom;
- Cluster 2: Azerbaijan, Estonia, Russia;
- Cluster 3: Croatia, Czechia, Finland.

For the Average and Centroid methods, the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Austria, Belarus, Bulgaria, Denmark, Georgia, Greece, Hungary, Ireland, Italy, Lithuania, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Poland, Portugal, Republic of Moldova, Romania, Slovenia, Sweden, Ukraine, United Kingdom;
- Cluster 2: Azerbaijan;
- Cluster 3: Belgium, Latvia, Luxembourg, Spain, Switzerland;
- Cluster 4: Bosnia and Herzegovina, Slovakia;
- Cluster 5: Croatia;
- Cluster 6: Cyprus;
- Cluster 7: Czechia;
- Cluster 8: Estonia;
- Cluster 9: Finland;
- Cluster 10: France, Iceland, Norway, San Marino;
- Cluster 11: Germany;
- Cluster 12: Greenland;
- Cluster 13: Russia;

- Cluster 14: Serbia;
- Cluster 15: Turkey.

For the Single method the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Austria, Belarus, Bulgaria, Denmark, Georgia, Greece, Hungary, Ireland, Italy, Lithuania, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Poland, Portugal, Republic of Moldova, Romania, Slovenia, Sweden, Ukraine, United Kingdom;
- Cluster 2: Azerbaijan;
- Cluster 3: Belgium, France, Latvia, Luxembourg, Norway, San Marino, Spain, Switzerland;
- Cluster 4: Bosnia and Herzegovina, Slovakia;
- Cluster 5: Croatia;
- Cluster 6: Cyprus;
- Cluster 7: Czechia;
- Cluster 8: Estonia;
- Cluster 9: Finland;
- Cluster 10: Germany;
- Cluster 11: Greenland;
- Cluster 12: Iceland;
- Cluster 13: Russia;
- Cluster 14: Serbia;
- Cluster 15: Turkey.

For the Median method the 15 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Austria, Belarus, Bulgaria, Denmark, Georgia, Greece, Hungary, Ireland, Italy, Lithuania, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Poland, Portugal, Republic of Moldova, Romania, Slovenia, Sweden, Ukraine, United Kingdom;
- Cluster 2: Azerbaijan;
- Cluster 3: Belgium, France, Iceland, Latvia, Norway, San Marino, Switzerland;
- Cluster 4: Bosnia and Herzegovina, Slovakia;
- Cluster 5: Croatia;
- Cluster 6: Cyprus;
- Cluster 7: Czechia;
- Cluster 8: Estonia;
- Cluster 9: Finland;
- Cluster 10: Germany;
- Cluster 11: Greenland;

- Cluster 12: Luxembourg, Spain;
- Cluster 13: Russia;
- Cluster 14: Serbia;
- Cluster 15: Turkey.

Annex D

Chronic Kidney Diseases – Piccolo Distance

For the Ward method the 15 clusters were the following:

- Cluster 1: Albania, Belgium, Bosnia and Herzegovina, Croatia, Georgia, Hungary, Luxembourg, Serbia, Spain;
- Cluster 2: Andorra, Estonia, Lithuania, Montenegro, Netherlands, North Macedonia, Portugal, Slovenia;
- Cluster 3: Armenia, Denmark, Finland, France, Germany, Greece, Greenland, Switzerland;
- Cluster 4: Austria;
- Cluster 5: Azerbaijan, Cyprus, Italy, Romania;
- Cluster 6: Belarus, Bulgaria, Ireland, San Marino, Ukraine;
- Cluster 7: Czechia, Iceland, Slovakia;
- Cluster 8: Latvia, Monaco;
- Cluster 9: Malta;
- Cluster 10: Norway;
- Cluster 11: Poland;
- Cluster 12: Republic of Moldova, Turkey;
- Cluster 13: Russia;
- Cluster 14: Sweden;
- Cluster 15: United Kingdom.

For the Complete, Single, Average, Centroid, and Median methods the 2 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Portugal, Republic of Moldova, Romania, Russia, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom;
- Cluster 2: Austria, Poland.

Chronic Kidney Diseases – Maharaj Distance

For all clustering method the 2 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Portugal, Republic of Moldova, Romania, Russia, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom;
- Cluster 2: Austria, Poland.

Chronic Kidney Diseases – LPC Distance

For the Ward method the 3 clusters were the following:

- Cluster 1: Albania, Armenia, Belgium, Bosnia and Herzegovina, Croatia, Czechia, Denmark, Finland, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Luxembourg, Serbia, Slovakia, Spain, Switzerland, United Kingdom;
- Cluster 2: Andorra, Azerbaijan, Belarus, Bulgaria, Cyprus, Estonia, Ireland, Italy, Latvia, Lithuania, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Portugal, Republic of Moldova, Romania, Russia, San Marino, Slovenia, Sweden, Turkey, Ukraine;
- Cluster 3: Austria, Poland.

For the Complete, Single, Average, Centroid, and Median methods the 2 clusters were the following:

- Cluster 1: Albania, Andorra, Armenia, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Portugal, Republic of Moldova, Romania, Russia, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom;
- Cluster 2: Austria, Poland.

Annex E

Diabetes Mellitus – Piccolo Distance

For the Ward method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Russia;
- Cluster 3: Armenia, Austria, Monaco, Norway, San Marino, Slovenia, Spain, Sweden, Turkey;
- Cluster 4: Azerbaijan, Germany, North Macedonia;
- Cluster 5: Belarus;
- Cluster 6: Belgium, Croatia, Cyprus, Greece, Ukraine;
- Cluster 7: Bosnia and Herzegovina;
- Cluster 8: Bulgaria, Czechia, Denmark, Estonia, France, Georgia, Iceland, Luxembourg, Montenegro, Romania, United Kingdom;
- Cluster 9: Finland, Italy, Latvia, Lithuania, Republic of Moldova;
- Cluster 10: Greenland;
- Cluster 11: Hungary, Ireland, Serbia, Slovakia;
- Cluster 12: Malta, Netherlands;
- Cluster 13: Poland;
- Cluster 14: Portugal;
- Cluster 15: Switzerland.

For the Complete method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Russia;
- Cluster 3: Armenia, Austria, Bulgaria, Czechia, Denmark, Estonia, France, Georgia, Iceland, Luxembourg, Monaco, Montenegro, Norway, Romania, San Marino, Slovenia, Spain, Sweden, Turkey, United Kingdom;
- Cluster 4: Azerbaijan, Germany, North Macedonia;
- Cluster 5: Belarus;
- Cluster 6: Belgium, Croatia, Cyprus, Greece, Ukraine;
- Cluster 7: Bosnia and Herzegovina;
- Cluster 8: Finland, Italy, Latvia, Lithuania, Republic of Moldova;
- Cluster 9: Greenland;
- Cluster 10: Hungary, Ireland, Serbia, Slovakia;
- Cluster 11: Malta;

- Cluster 12: Netherlands;
- Cluster 13: Poland;
- Cluster 14: Portugal;
- Cluster 15: Switzerland.

For the Average, Median, and Centroid methods the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra;
- Cluster 3: Armenia, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, France, Georgia, Greece, Iceland, Luxembourg, Monaco, Montenegro, Norway, Romania, San Marino, Slovenia, Spain, Sweden, Turkey, Ukraine, United Kingdom;
- Cluster 4: Azerbaijan, Germany, North Macedonia;
- Cluster 5: Belarus;
- Cluster 6: Bosnia and Herzegovina;
- Cluster 7: Finland, Italy, Latvia, Lithuania, Republic of Moldova;
- Cluster 8: Greenland;
- Cluster 9: Hungary, Ireland, Serbia, Slovakia;
- Cluster 10: Malta;
- Cluster 11: Netherlands;
- Cluster 12: Poland;
- Cluster 13: Portugal;
- Cluster 14: Russia;
- Cluster 15: Switzerland.

For the Single method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra;
- Cluster 3: Armenia, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, France, Georgia, Greece, Hungary, Iceland, Ireland, Luxembourg, Monaco, Montenegro, Norway, Romania, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Turkey, Ukraine, United Kingdom;
- Cluster 4: Azerbaijan, Germany, North Macedonia;
- Cluster 5: Belarus;
- Cluster 6: Bosnia and Herzegovina;
- Cluster 7: Finland, Italy, Lithuania, Republic of Moldova;
- Cluster 8: Greenland;
- Cluster 9: Latvia;

- Cluster 10: Malta;
- Cluster 11: Netherlands;
- Cluster 12: Poland;
- Cluster 13: Portugal;
- Cluster 14: Russia;
- Cluster 15: Switzerland.

Diabetes Mellitus – Maharaj Distance

For the Ward method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Finland;
- Cluster 3: Armenia, Austria, Monaco, Norway, San Marino, Slovenia, Spain, Sweden, Turkey;
- Cluster 4: Azerbaijan, Germany, North Macedonia, Russia;
- Cluster 5: Belarus;
- Cluster 6: Belgium, Croatia, Cyprus, Greece, Ukraine;
- Cluster 7: Bosnia and Herzegovina;
- Cluster 8: Bulgaria, Czechia, Denmark, Estonia, France, Georgia, Iceland, Luxembourg, Montenegro, Romania, United Kingdom;
- Cluster 9: Greenland;
- Cluster 10: Hungary, Ireland, Serbia, Slovakia;
- Cluster 11: Italy, Latvia, Lithuania, Republic of Moldova;
- Cluster 12: Malta, Netherlands;
- Cluster 13: Poland;
- Cluster 14: Portugal;
- Cluster 15: Switzerland.

For the Complete, Average, Median, and Centroid methods the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Finland, Latvia, Lithuania, Republic of Moldova;
- Cluster 3: Armenia, Austria, Bulgaria, Czechia, Denmark, Estonia, France, Georgia, Iceland, Luxembourg, Monaco, Montenegro, Norway, Romania, San Marino, Slovenia, Spain, Sweden, Turkey, United Kingdom;
- Cluster 4: Azerbaijan, Germany, North Macedonia, Russia;
- Cluster 5: Belarus;

- Cluster 6: Belgium, Croatia, Cyprus, Greece, Ukraine;
- Cluster 7: Bosnia and Herzegovina;
- Cluster 8: Greenland;
- Cluster 9: Hungary, Ireland, Serbia, Slovakia;
- Cluster 10: Italy;
- Cluster 11: Malta;
- Cluster 12: Netherlands;
- Cluster 13: Poland;
- Cluster 14: Portugal;
- Cluster 15: Switzerland.

For the Single method the 11 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Azerbaijan, Finland, Germany, Italy, Latvia, Lithuania, North Macedonia, Republic of Moldova, Russia;
- Cluster 3: Armenia, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, France, Georgia, Greece, Hungary, Iceland, Ireland, Luxembourg, Monaco, Montenegro, Norway, Romania, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Turkey, Ukraine, United Kingdom;
- Cluster 4: Belarus;
- Cluster 5: Bosnia and Herzegovina;
- Cluster 6: Greenland;
- Cluster 7: Malta;
- Cluster 8: Netherlands;
- Cluster 9: Poland;
- Cluster 10: Portugal;
- Cluster 11: Switzerland.

Diabetes Mellitus – LPC Distance

For the Ward method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Finland;
- Cluster 3: Armenia, Austria, Monaco, Norway, San Marino, Slovenia, Spain, Sweden, Turkey;
- Cluster 4: Azerbaijan, Germany, North Macedonia, Russia;
- Cluster 5: Belarus;

- Cluster 6: Belgium, Croatia, Cyprus, Greece, Ukraine;
- Cluster 7: Bosnia and Herzegovina;
- Cluster 8: Bulgaria, Czechia, Denmark, Estonia, France, Georgia, Iceland, Luxembourg, Montenegro, Romania, United Kingdom;
- Cluster 9: Greenland;
- Cluster 10: Hungary, Ireland, Serbia, Slovakia;
- Cluster 11: Italy, Latvia, Lithuania, Republic of Moldova;
- Cluster 12: Malta, Netherlands;
- Cluster 13: Poland;
- Cluster 14: Portugal;
- Cluster 15: Switzerland.

For the Complete and Median methods the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Finland, Latvia, Lithuania, Republic of Moldova;
- Cluster 3: Armenia, Austria, Bulgaria, Czechia, Denmark, Estonia, France, Georgia, Iceland, Luxembourg, Monaco, Montenegro, Norway, Romania, San Marino, Slovenia, Spain, Sweden, Turkey, United Kingdom;
- Cluster 4: Azerbaijan, Germany, North Macedonia, Russia;
- Cluster 5: Belarus;
- Cluster 6: Belgium, Croatia, Cyprus, Greece, Ukraine;
- Cluster 7: Bosnia and Herzegovina;
- Cluster 8: Greenland;
- Cluster 9: Hungary, Ireland, Serbia, Slovakia;
- Cluster 10: Italy;
- Cluster 11: Malta;
- Cluster 12: Netherlands;
- Cluster 13: Poland;
- Cluster 14: Portugal;
- Cluster 15: Switzerland.

For the Average method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Finland, Latvia, Lithuania, Republic of Moldova;

- Cluster 3: Armenia, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, France, Georgia, Greece, Iceland, Luxembourg, Monaco, Montenegro, Norway, Romania, San Marino, Slovenia, Spain, Sweden, Turkey, Ukraine, United Kingdom;
- Cluster 4: Azerbaijan, Germany;
- Cluster 5: Belarus;
- Cluster 6: Bosnia and Herzegovina;
- Cluster 7: Greenland;
- Cluster 8: Hungary, Ireland, Serbia, Slovakia;
- Cluster 9: Italy;
- Cluster 10: Malta;
- Cluster 11: Netherlands;
- Cluster 12: North Macedonia, Russia;
- Cluster 13: Poland;
- Cluster 14: Portugal;
- Cluster 15: Switzerland.

For the Single method the 11 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra, Azerbaijan, Finland, Germany, Italy, Latvia, Lithuania, North Macedonia, Republic of Moldova, Russia;
- Cluster 3: Armenia, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, France, Georgia, Greece, Hungary, Iceland, Ireland, Luxembourg, Monaco, Montenegro, Norway, Romania, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Turkey, Ukraine, United Kingdom;
- Cluster 4: Belarus;
- Cluster 5: Bosnia and Herzegovina;
- Cluster 6: Greenland;
- Cluster 7: Malta;
- Cluster 8: Netherlands;
- Cluster 9: Poland;
- Cluster 10: Portugal;
- Cluster 11: Switzerland.

For the Centroid method the 15 clusters were the following:

- Cluster 1: Albania;
- Cluster 2: Andorra;

- Cluster 3: Armenia, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, France, Georgia, Greece, Iceland, Luxembourg, Monaco, Montenegro, Norway, Romania, San Marino, Slovenia, Spain, Sweden, Turkey, Ukraine, United Kingdom;
- Cluster 4: Azerbaijan, Germany, North Macedonia, Russia;
- Cluster 5: Belarus;
- Cluster 6: Bosnia and Herzegovina;
- Cluster 7: Finland, Latvia, Lithuania, Republic of Moldova;
- Cluster 8: Greenland;
- Cluster 9: Hungary, Ireland, Serbia, Slovakia;
- Cluster 10: Italy;
- Cluster 11: Malta;
- Cluster 12: Netherlands;
- Cluster 13: Poland;
- Cluster 14: Portugal;
- Cluster 15: Switzerland.