EKOIST Journal of Econometrics and Statistics,
EKOIST 2025, 43: 91–102

https://doi.org/10.26650/ekoist.2025.43.1651865

Submitted 05.03.2025
Accepted 03.06.2025
Published Online 03.11.2025

EKOIST Journal of Econometrics and Statistics

Research Article

Open Access

Visualisation of the populations of the European Union countries by age groups using correspondence analysis

Mehmet Hakan Özdemir [1]

[1] Turkish-German University, Faculty of Economic and Administrative Sciences, Department of Business Administration, İstanbul, Türkiye

Abstract

Particularly since the early 2000s, the European Union (EU) has undergone significant expansion through the accession of numerous countries, and it presently comprises 27 member states. There exists a multifaceted demographic change occurring within the member states. The primary factors influencing this demographic change include the ageing of the population, intra-member state population mobility, and immigration from countries outside the Union. These components exert varying effects on each member state, and the 27 member states exhibit significantly diverse demographic structures in relation to one another. This study aims to visually illustrate the similarities and differences in the populations of EU countries by age groups. To that end, four age groups (0-14, 15-24, 25-64, and 65+) were determined in the study. The populations of the EU countries by age groups as of January 1, 2023, were sourced from EUROSTAT. The populations were subsequently analysed, followed by the establishment of the appropriate number of dimensions and the visualisation of the populations of the EU member states categorised by age groups using correspondence analysis, a dimensionality reduction technique that allows researchers to visualise the results graphically without losing significant information. The study's findings indicate that the demographic structures of the EU member states differ from one another, revealing various groupings among them. The study thus assists decision makers in comprehending the demographic composition of member states and developing a strategic roadmap, as well as in taking various measures or steps if required.

Keywords Correspondence Analysis · European Union · Population · Age Group

Citation: Özdemir, M. H. (2025). Visualisation of the populations of the European Union countries by age groups using correspondence analysis. *EKOIST Journal of Econometrics and Statistics*, 43, 91-102. https://doi.org/10.26650/ekoist.2025.43.1651865

Corresponding author: Mehmet Hakan Özdemir hakan.ozdemir@tau.edu.tr

Visualisation of the populations of the European Union countries by age groups using correspondence analysis

The term "demography," first introduced by Guillard in 1855, refers to the statistical study of human populations. This primarily entails the assessment of the population size and its growth or decline. The factors influencing changes in these figures are births, deaths, and migration, which the demographers examine through the functions of fertility, mortality, and population transfer. Populations can also be analysed for their characteristics at a specific moment, such as age distribution or genetic composition. Such characteristics seldom undergo rapid change (Cox, 1976).

Since the early 2000s, the EU has significantly expanded with the accession of numerous countries, currently comprising 27 member states. As of January 1, 2023, the population of the EU was recorded at 448.8 million people. Germany was the most populous country in the EU, with 84.4 million inhabitants (19% of the EU total), followed by France with 68.2 million (15%), Italy with 59.0 million (13%), Spain with 48.1 million (11%), and Poland with 36.8 million (8%). Collectively, these five EU member states comprised 66% of the EU population. The least populous EU countries were Malta (542,000 inhabitants, representing 0.1% of the EU total), Luxembourg (661,000, also 0.1%), and Cyprus (921,000, 0.2%). From January 1, 2003 to January 1, 2023, the EU's population rose from 431.2 million to 448.8 million, reflecting a growth of 4%. During this period, 18 EU countries experienced population growth, while nine reported declines. The most significant relative increases occurred in Luxembourg (47%), Malta (36%), Ireland (33%), and Cyprus (29%), whereas the most substantial relative decreases were noted in Latvia (−18%), Lithuania and Bulgaria (both −17%), and Romania (−12%). The most significant absolute increases occurred in Spain and France (both 6.3 million), whereas the most substantial absolute decreases were noted in Romania (−2.6 million), Poland (−1.5 million), and Bulgaria (−1.4 million) (EUROSTAT, 2024a).

The EU, as a collective entity, confronts a dynamic demographic landscape marked by ageing populations, immigration complexities, and varied population trends among member states. The recognition of these demographic subtleties provides a fundamental basis for the subsequent examination of the challenges encountered by the EU. These challenges include the consequences of an ageing population as well as the complex dynamics of immigration, urbanisation, and socio-economic inequalities (Stupariu, 2023).

The demographic shift in the EU is characterised by a growing proportion of older age groups. The increasing percentage of elderly individuals poses significant implications for healthcare systems and social security, labour markets, and economic stability. As life expectancy increases and birth rates decline, EU member states must address the pressures on their pension programs and healthcare systems. This demographic transition although involving challenges, also offers opportunities for economic growth and innovation. Senior employees can enhance labour productivity, and diverse age groups can distinctly contribute to the economy. To effectively address these demographic changes, comprehensive strategies that consider regional and national differences are essential (Szűcs, 2024).

The process of population ageing actually began several decades ago in Europe. This dynamic is marked by an increasing percentage of elderly individuals in the total population, alongside a decreasing proportion of the working-age demographic. As of January 1, 2021, individuals aged 65 and older represented 20.8% of the EU's population, reflecting a 0.2% increase from 2020 and a 0.6% rise from 2019. By 2050, the elderly population in Europe will exceed 30%, resulting in an increased old-age dependency ratio, with fewer than two working-age individuals for each elderly person. This demographic transition is expected to continue in the future (Szűcs, 2024).

In addition to the ageing population, immigration plays a significant role in the demographic development of the EU as a destination region for international immigration. Immigration policy is regularly balanced with the best interests of member states. Maintaining a long-term understanding of demography is essential. Important factors to consider include the number of immigrants from third countries entering the member state, their education levels, their integration into the labour market and society, and the effectiveness of the immigration policy enforcement. Immigration levels significantly impact population and labour force size. Without third-country immigration, the EU population would decline to 466 million by 2060 due to lower fertility, similar to the 1980s. Immigration levels have a limited impact on the EU age structure, as immigrants settle and age similarly to the native population. Regardless of immigration levels, a trend towards continued demographic ageing in the EU is likely (Lutz et al., 2019).

The EU's expansion in the early 2000s, as aforementioned, was another factor influencing substantial intra-migration flows from Eastern to Western Europe. The enlargement of the EU has influenced immigration patterns within the region, particularly East-West mobility (Cozma & Pricop, 2024). In 2022, Luxembourg recorded the highest proportion of immigrants originating from another member state, accounting for 75.0% of its total immigrant population, followed by Slovakia at 65.6%. In 2021, Luxembourg (91.0%), Slovakia (67.9%), and Austria (56.0%) exhibited the highest proportions of immigrants originating from another member state (EUROSTAT, 2024b).

In conclusion, the demographic landscape of the EU is characterised by an ageing population, declining birth rates, and significant immigration flows, all of which have different impacts. This study aims to visually illustrate the similarities and differences in the populations of EU countries by age groups. The populations of the EU countries by age groups as of January 1, 2023, were sourced from EUROSTAT and subsequently analysed and visualised using correspondence analysis.

Literature Review

This section will discuss various studies that apply correspondence analysis to demographic data.

The study by Özer et al. (2011) employed multiple correspondence analysis to examine the relationships between bullying behaviour and various individual demographic variables. This study clarifies how demographic characteristics may influence the dynamics of bullying within educational settings, thereby providing essential insights for educators and policymakers.

Silva et al. (2015) used multiple correspondence analysis to examine the relationships between the prevalence of overweight and diverse demographic and socioeconomic indicators in Brazil. This study is important as it examines a crucial public health concern by elucidating the correlation between various factors and adult overweight status, which is vital for developing targeted interventions.

Chatzipetrou and Moschidis (2016) used multiple correspondence analysis to examine data gathered from a survey of 149 Greek supermarkets, concentrating on the correlation between diverse demographic and operational factors and the adoption of quality costing practices.

Lana et al. (2017) offered an extensive examination of the relationship between socioeconomic and demographic factors and malaria prevalence in Brazil's endemic areas. The study uses multiple correspondence analysis to clarify the intricate relationships among diverse categorical variables, thus providing insights that can guide public health strategies for malaria control.

Yates and Lockley (2018) employed multiple correspondence analysis to delineate the relationships between various types of social media usage and demographic factors associated with social class. The results demonstrate that social media behaviours are not solely personal decisions but are profoundly shaped by overarching social, economic and cultural contexts.

Diéguez et al. (2020) used multiple correspondence analysis to examine the complex relationships between demographic variables and public perceptions of bullfighting in Portugal. The results demonstrate that demographic factors significantly influence the perceptions of bullfighting, revealing distinct profiles among various demographic groups.

The research conducted by Idrissou et al. (2023) seeks to address this gap by evaluating the impact of adaptation strategies employed by cattle farmers in Benin on livestock productivity. Surveys were administered to 360 cattle farmers, and 30 farms were monitored in the dry and sub-humid tropical regions of Benin. The authors gathered information on the socio-demographic traits of cattle farmers, their adaptation strategies, and herd productivity. They conducted a multiple correspondence analysis, succeeded by an ascending hierarchical classification to determine the typology of the adaptation strategies employed by cattle farmers.

Methodology

Correspondence analysis enables researchers to simplify a complex data matrix while preserving significant information. A notable benefit of correspondence analysis is its ability to present results graphically. Variables with analogous frequencies are represented as closely clustered points in space, while those with disparate frequencies are depicted as points that are distanced from one another (Kakai, 2003). The three fundamental concepts of correspondence analysis are profiles, masses, and χ^2 distances (Greenacre, 1994). By converting the frequencies of variables into percentages relative to the row or column totals, known as profiles, the analysis enables direct comparison of occurrence magnitudes among categories. Profiles, regarded as mathematical vectors, elucidate the proportional frequencies of categories solely in relation to the group total, whereas masses are values that modify the various numbers within the group totals, derived by dividing each group total by the grand total. Distances between points in space are calculated as χ^2 distances, which correspond to Euclidean distances with an adjustment for the proportions in the frequencies of categories using profiles (Kakai, 2003).

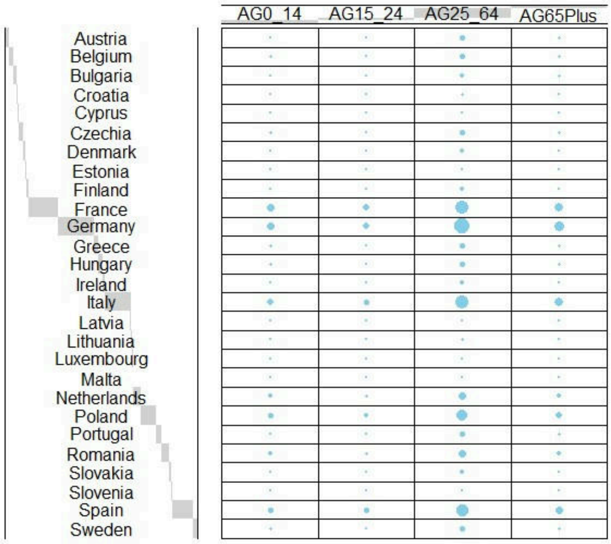
Findings

This study aims to visually illustrate the similarities and differences in the populations of EU countries by age groups using correspondence analysis. The analysis was performed using R. The populations of the EU countries by age groups as of January 1, 2023, were sourced from EUROSTAT (EUROSTAT, 2025). Four age groups (0-14, 15-24, 25-64 and 65+) were determined in the study as the categorisation of various age demographics has traditionally been represented as "child - youth - adult - elderly." The demographic encompassing the years of fundamental education from birth to age fourteen is referred to as the child group. The interval following basic education until individuals start employment (ages 15-24) constitutes the youth demographic. The interval between working age and retirement age (25-64 years) is termed the adult period, while the post-retirement phase (65+) is designated as the elderly group (Bilir, 2018).

The contingency table can be visualised as in Figure 1 showing a graphical matrix in which each cell contains a dot, the size of which represents the relative magnitude of the corresponding component (Kassambara, 2017). Accordingly, Figure 1 illustrates that the countries with the largest populations are France, Germany, Italy, Poland, and Spain, as indicated in the introduction.

Figure 1

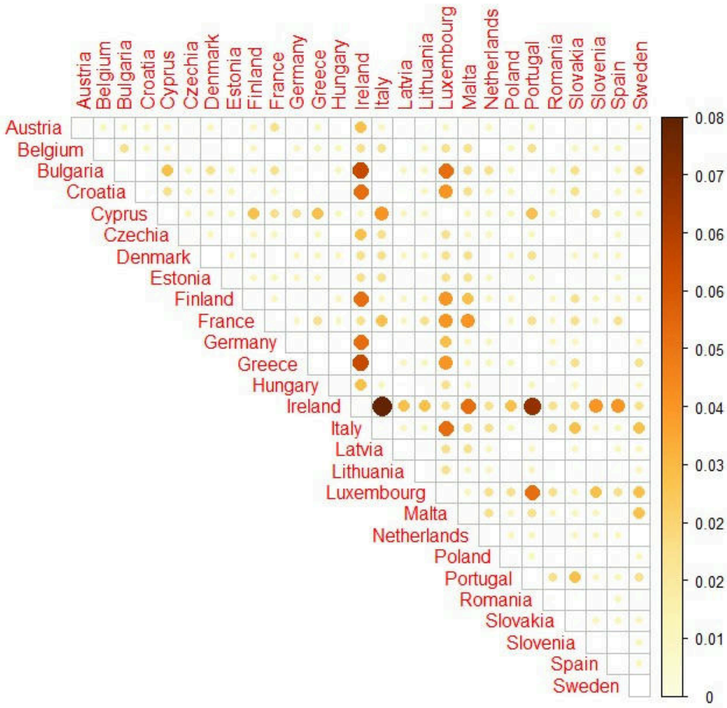
Population of the EU countries by age groups



Furthermore, Figure 2 may be consulted to examine the row profiles and identify both similarities and differences. The larger the size of the dot, the more dissimilar the countries are (Kassambara, 2017). Consequently, it is evident that Ireland, followed by Luxembourg, possesses a distinct profile compared to other countries. This is due to the higher proportion of individuals in the 0–14 age group in Ireland compared to other EU countries and the lower proportion of individuals in the 65+ age group in the total population in both Luxembourg and Ireland compared to other EU countries.

Figure 2

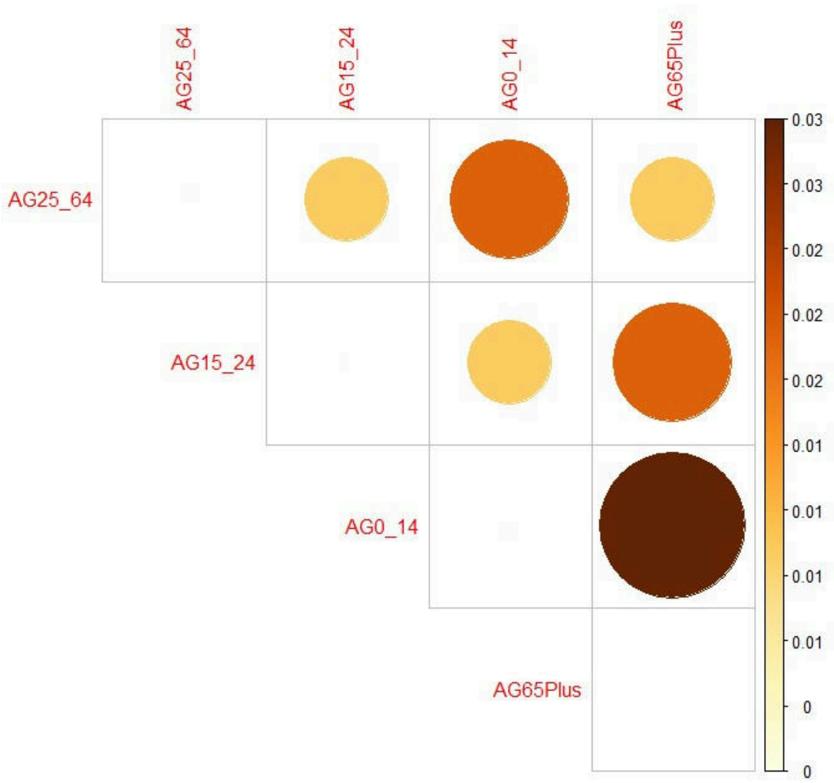
Dissimilarity between the row profiles



Comparably, Figure 3 reveals the similarities and dissimilarities between the column profiles (Kassambara, 2017). The most significant disparity was observed between the age group 65+ and the 0-14 age group.

Figure 3

Dissimilarity between the column profiles



The initial step in interpreting the correspondence analysis is to determine if there is a substantial dependency between the rows and columns (Kassambara, 2017). To that end, χ^2 test is computed. Table 1 shows that the row and the column variables are statistically significantly associated.

Table 1

χ^2 test results

χ^2 = 2085758	Degrees of freedom: 78	p-value: 2.2e-16

The eigenvalues are computed in the second step to ascertain the number of dimensions that should be considered. Eigenvalues represent the quantity of information preserved by each respective dimension (Kassambara, 2017). Table 2 shows the eigenvalue, the ratio of explained variance, and the cumulative explained variance attributed to each dimension, respectively.

Table 2

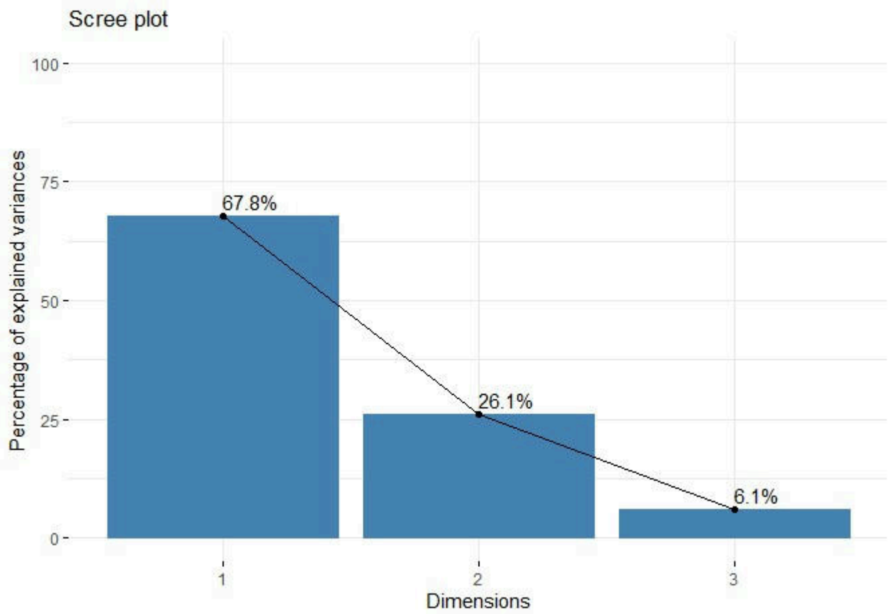
Eigenvalues and explained variance ratios

Dimension	Eigenvalue	Explained variance ratio	Cumulative explained variance ratio
1	0,0031512924	0,67807949	0,6780795
2	0,0012139154	0,26120429	0,9392838
3	0,0002821713	0,06071622	1

Table 2 indicates that the first two dimensions account for most variance, approximately 93.93%, signifying a good dimensionality reduction. Alternatively, the scree plot in Figure 4 can also be used to determine the number of dimensions (Kassambara, 2017). The scree plot indicates that after the second dimension, the subsequent eigenvalue is comparatively minor.

Figure 4

Scree plot



After determining the number of dimensions, a biplot can be drawn as in Figure 5, where the blue points denote the rows and the red triangles the columns. The distance between any row or column points serves as an indicator of their similarity or dissimilarity. Row points exhibiting similar profiles are located close to each other on the biplot. The same applies to column points (Kassambara, 2017).

Figure 5

Corresponding analysis biplot

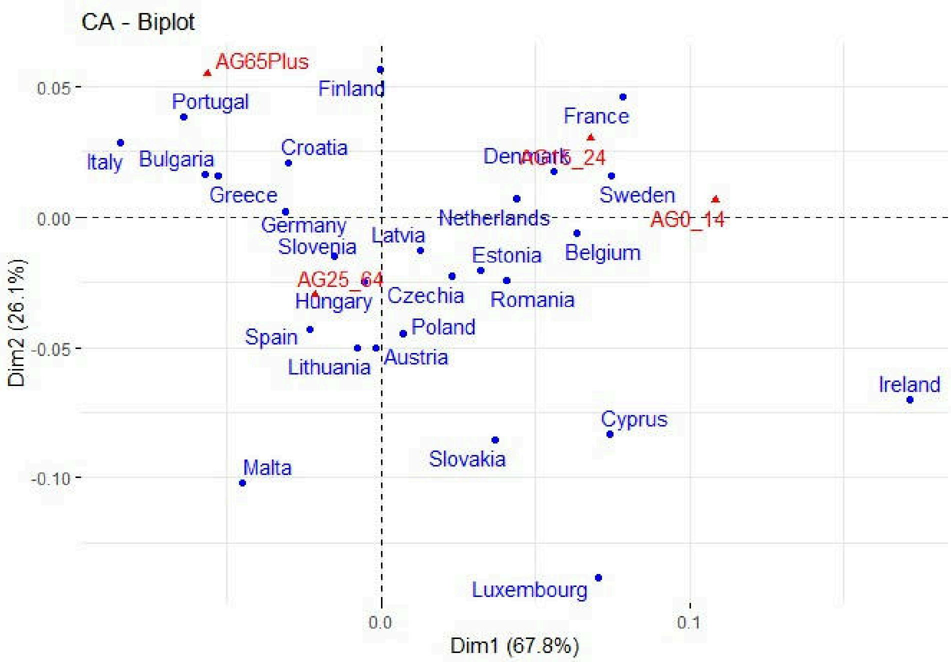
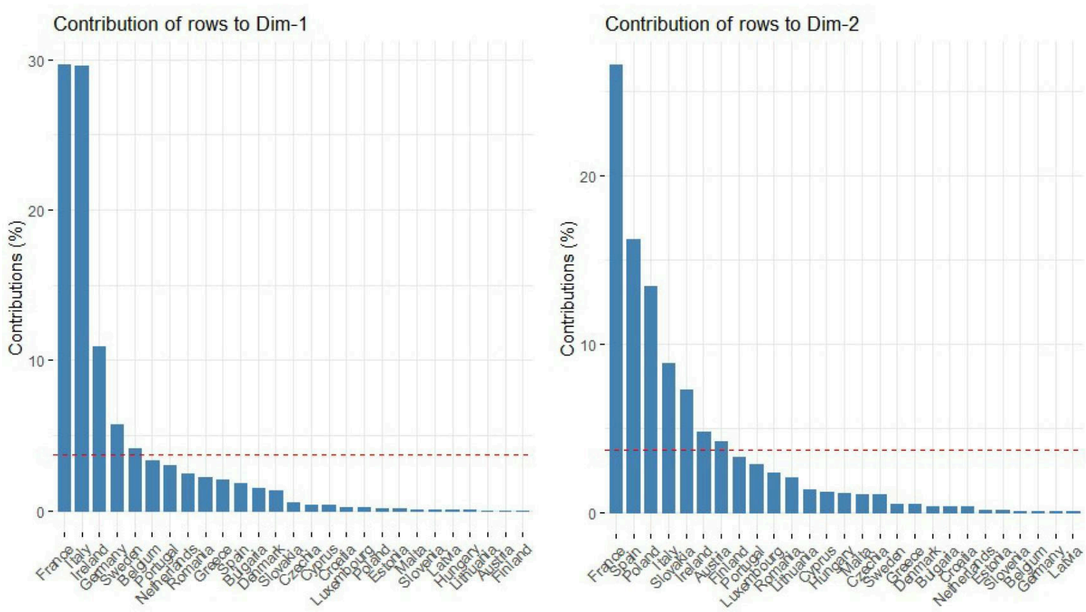


Figure 5 demonstrates again that the first dimension (Dim1 or Dim-1) accounts for 67.8% of the variance, while the second dimension (Dim2 or Dim-2) accounts for 26.1% of the variance. Figure 5 shows that Italy, Portugal, Bulgaria and Finland are close to the 65+ age group. In these countries, the proportion of this age group in the total population is higher than that in other EU countries. In France, Denmark, and the Netherlands, the proportion of the 15-24 age group in the total population is higher compared to other EU countries, except Ireland, which ranks first in this age group. These countries are therefore centred around the 15-24 age group. In Malta, the proportion of the 25-64 age group in the total population is the highest compared to other EU countries. In addition, Malta has the lowest proportion of the 0-14 age group in the total population after Italy. Furthermore, Figure 5 demonstrates that Ireland stands apart from the other EU countries. This finding agrees with the finding from Figure 2 that Ireland has a distinct profile. There are two reasons for this: First, Ireland has the highest proportion of the 0-14 age group in the total population compared to other EU countries. Second, Ireland has the lowest proportion of the 65+ age group in the total population after Luxembourg. In terms of proportion, Sweden ranks second and France third in the 0-14 age group. Luxembourg, similar to Ireland, is situated at a significant distance from other EU countries, corroborating the conclusions drawn from Figure 2. This can be attributed to the fact that in Luxembourg, the proportion of the 65+ age group is the lowest.

Once the biplot is drawn, the contributions of the rows to Dim1 and Dim2 in percentages can be seen in Figure 6. The row variables with higher values make the most significant contributions to the delineation of the dimensions. The rows that exert the greatest influence on Dim1 and Dim2 are the most significant in elucidating the variability within the dataset. The red dashed line in Figures 6 and 7 signifies the expected value, assuming uniform contributions (Kassambara, 2017). According to Figure 6, France makes the highest contribution both to Dim1 and to Dim2.

Figure 6

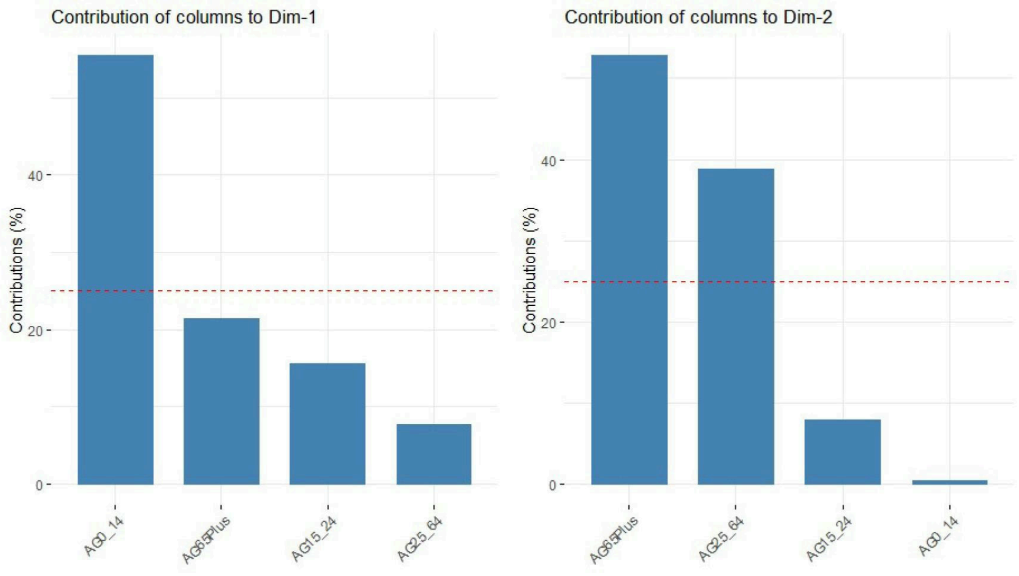
Contributions of the rows to Dim1 and Dim2



The contributions of the columns to Dim1 and Dim2 in percentages can be seen in Figure 7. Accordingly, age group 0-14 makes the highest contribution to Dim1, whereas age group 65+ contributes the most to Dim2.

Figure 7

Contributions of the columns to Dim1 and Dim2



Finally, it may be interesting to examine the quality of the representation, which is called the squared cosine (cos2). The value of cos2 is between 0 and 1. A cos2 value close to 1 indicates that the column/row variables are well represented on the biplot (Kassambara, 2017). Figure 8 shows, for example, that Belgium and Germany are well represented in Dim1, while Finland and Austria are well represented in Dim2. Figure 9 illustrates that the age group 0-14 is well represented in Dim1, while the age group 65+ is well represented in Dim2.

Figure 8

Quality of representation of the rows in Dim1 and Dim2

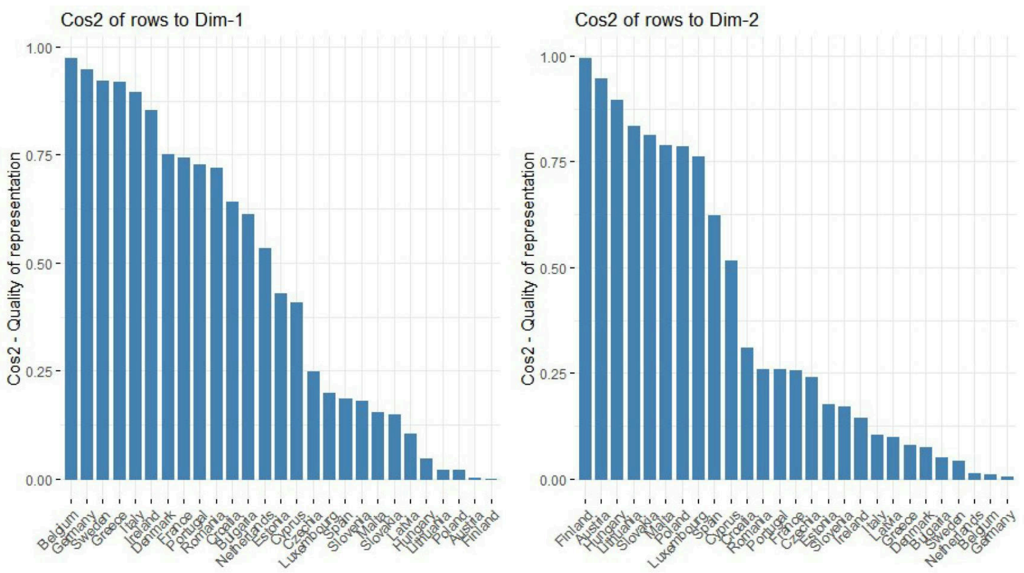
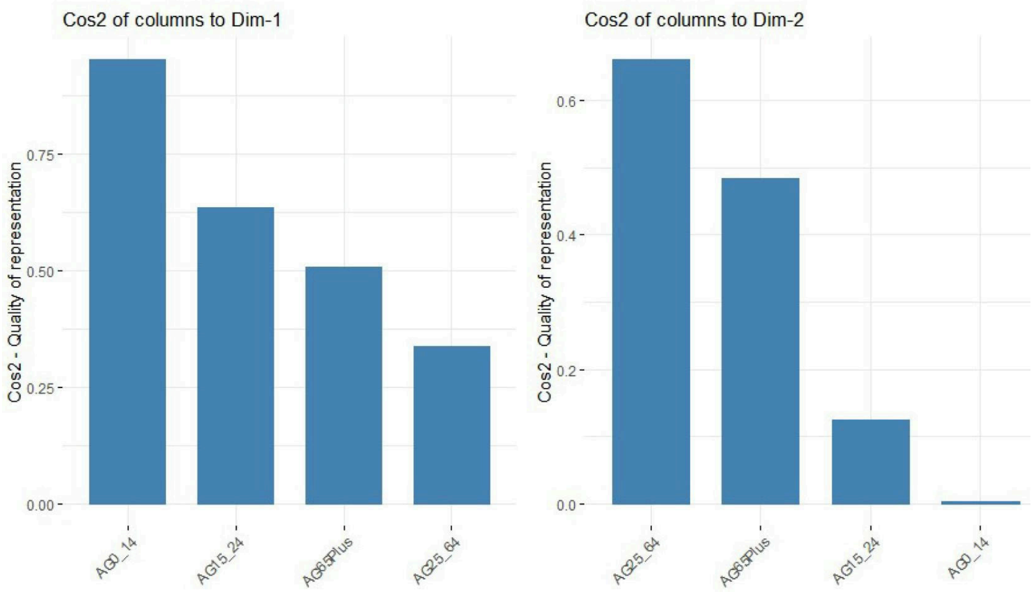


Figure 9

Quality of representation of the columns in Dim1 and Dim2



Discussion and Conclusion

The EU faces ageing demographics, immigration challenges, and diverse population trends. Growing numbers of older age groups define the main demographic change in the EU. The growing proportion of elderly people has important consequences for labour markets, social security and healthcare systems, as well as for economic stability. Birth rates drop while life expectancy rises. Apart from the ageing population, immigration significantly influences the demographic development in the EU. The number of immigrants from third countries into the EU is rising, and in the absence of this demographic trend, the EU's population will decline due to ageing population. Furthermore, the expansion of the EU over the years has enhanced mobility between Western and Eastern countries. There has been significant immigration of people from various Eastern EU countries to Western EU countries. In summary, the demographic composition of EU member states is notably intricate.

Particularly in Italy and Portugal, and subsequently in Finland, the age group 65+ constitutes a more significant proportion of the total population compared to other EU member states. Bulgaria is also among these countries. The reason for this, as mentioned in the introduction, is the immigration of Bulgaria's young population to Western EU countries as a result of the enlargement of the EU by incorporating new member states since the early 2000s and the consequent increase in East-West mobility within the Union. As also indicated in the introduction, France and Spain rank among the member states of the Union with the highest absolute increase in population. France ranks fourth in terms of the high proportion of the 14-25 age group in the total population. It is also the country with the highest total proportion of the 0-14 and 15-24 age groups in its total population, following Ireland. This suggests that France's population will continue to grow in the foreseeable future. As also indicated in the introduction, Ireland and Luxembourg are among the countries with the most significant relative population growth. This is due to the fact that Ireland has a very young population, as mentioned earlier, and Ireland and Luxembourg are the two countries with the lowest proportion of the 65+ age group in the total population. Therefore, the growth can be expected to continue in these two countries.

This study aims to shed light on the demographic structure of the EU countries by using correspondence analysis and help decision makers to draw a roadmap accordingly and take appropriate measures or steps if necessary. Correspondence analysis is a dimensionality reduction technique that allows researchers to visualise the results graphically without losing significant information. The appropriate number of dimensions was initially established, followed by the visualisation of the populations of the EU member states categorised by age groups.

In future research, multiple correspondence analysis can be performed by incorporating additional categories. Alternatively, cluster analysis may be conducted, and the findings can be compared with the findings of this study.

Peer Review Externally peer-reviewed.
Conflict of Interest The author has no conflict of interest to declare.
Grant Support The author declared that this study has received no financial support.

Author Details **Mehmet Hakan Özdemir (Assoc. Prof. Dr.)**
[1] Turkish-German University, Faculty of Economic and Administrative Sciences, Department of Business Administration, İstanbul, Türkiye
0000-0002-7174-9807 hakan.ozdemir@tau.edu.tr

References

Bilir, N. (2018). Yaşlılık Tanımı, Yaşlılık Kavramı, Epidemiyolojik Özellikler. *Yaşlılık ve Solunum Hastalıkları. Ertürk A, Bahadır A, Koşar FSerisi (ed): TÜSAD Eğitim Kitapları, Istanbul*, 13-31.

Chatzipetrou, E., & Moschidis, O. (2016). Quality Costing: A survey in Greek Supermarkets using Multiple Correspondence Analysis. *International Journal of Quality & Reliability Management*, *33*(5). https://doi.org/10.1108/ijqrm-01-2014-0004

Cozma, O. M., & Pricop, A. I. (2024). Economic Security and Migration within the European Union. *Sivas Interdisipliner Turizm Araştırmaları Dergisi*, *7*(2), 18-38.

Cox, P. R. (1976). *Demography*. Cambridge University Press.

Diéguez, F. J., Zau, Y., Viegas, I., Fragoso, S., Turner, P. V., & Da Graça-Pereira, G. (2020). An Evaluation of Portuguese Societal Opinion towards the Practice of Bullfighting. *Animals*, *10*(11), 2065. https://doi.org/10.3390/ani10112065

EUROSTAT (2024a). Demography of Europe–2024 Edition. Retrieved from https://ec.europa.eu/eurostat/web/interactive-publications/demography-2024

EUROSTAT (2024b). EU population diversity by citizenship and country of birth. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_population_diversity_by_citizenship_and_country_of_birth

EUROSTAT (2025). Population by age group. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/TPS00010/default/table

Greenacre, M. (1994). Correspondence analysis and its interpretation. In M. Greenacre, & J. Blasius (Eds.), *Correspondence Analysis in the Social Sciences* (pp. 3–22). San Diego, CA: Academic Press.

Idrissou, Y., Korir, D., Seidou, A. A., Worogo, H. S. S., Baco, M. N., & Traoré, I. A. (2023). Adapting cattle farming to climate change in the dry and sub-humid tropical zones in Benin: how adaptation strategies affect productivity. *Revue D'élevage Et De Médecine Vétérinaire Des Pays Tropicaux*, *76*, 1–10. https://doi.org/10.19182/remvt.37118

Kakai, H., Maskarinec, G., Shumay, D. M., Tatsumura, Y., & Tasaki, K. (2003). Ethnic differences in choices of health information by cancer patients using complementary and alternative medicine: an exploratory study with correspondence analysis. *Social Science & Medicine*, *56*(4), 851–862. https://doi.org/10.1016/s0277-9536(02)00086-2

Kassambara, A. (2017). *Practical Guide To Principal Component Methods in R (Multivariate Analysis II). PCA, M(CA), FAMD, MFA, HCPC, factoextra.* STHDA.

Lana, R. M., Riback, T. I. S., Lima, T. F. M., Da Silva-Nunes, M., Cruz, O. G., Oliveira, F. G. S., Moresco, G. G., Honório, N. A., & Codeço, C. T. (2017). Socioeconomic and demographic characterization of an endemic malaria region in Brazil by multiple correspondence analysis. *Malaria Journal*, *16*(1). https://doi.org/10.1186/s12936-017-2045-z

Lutz, W., Amran, G., Bélanger, A., Conte, A., Gailey, N., Ghio, D., Grapsa, E., Jensen, K., Loichinger, E., Marois, G., Muttarak, R., Potančoková M., Sabourin P., & Stonawski, M. (2019). *Demographic scenarios for the EU: Migration, population and education*. Publications Office of the European Union, Luxembourg. https://doi.org/10.2760/590301

Özer, A., Totan, T., & Atik, G. (2011). Individual Correlates of Bullying Behaviour in Turkish Middle Schools. *Australian Journal of Guidance and Counselling*, *21*(2), 186–202. https://doi.org/10.1375/ajgc.21.2.186

Silva, V. S. D., Souza, I., Silva, D. A. S., Petroski, E. L., & Fonseca, M. D. J. M. D. (2015). Correspondence between overweight and socioeconomic and demographic indicators in the adult Brazilian population. *Revista Brasileira De Epidemiologia*, *18*(2), 476–489. https://doi.org/10.1590/1980-5497201500020015

Stupariu, M. I. (2023). Demographic Policies in the European Union. *Revista Română De Geografie Politică, 25*(2), 97–122. https://doi.org/10.30892/rrgp.252105-368

Szűcs, P. (2024). Demographic Shifts and Economic Challenges in the European Union. *Európai Tükör*, *27*(1), 87–106. https://doi.org/10.32559/et.2024.1.5

Yates, S. & Lockley, E. (2018). Social Media and Social Class. *American Behavioral Scientist*, *62*(9), 1291–1316. https://doi.org/10.1177/0002764218773821