Innovative Enterprises in Europe – a Country Cluster Analysis using R Shiny

Tudorel Andrei[1], Andreea Mirică[2] and Ionela-Roxana Petcu[3]
[1)2)3)] *Bucharest University of Economic Studies, Bucharest, Romania*
E-mail: andreitudorel@yahoo.com; E-mail: miricaandreea89@gmail.com
E-mail: roxana.glavan10@yahoo.com

Please cite this paper as:

Andrei, T., Mirică, A. and Petcu, I.R.,2023. Innovative Enterprises in Europe – a Country Cluster Analysis using R Shiny. In: R. Pamfilie, V. Dinu, C. Vasiliu, D. Pleșea, L. Tăchiciu eds. 2023. *9th BASIQ International Conference on New Trends in Sustainable Business and Consumption.* Constanța, Romania, 8-10 June 2023. Bucharest: ASE, pp. 260-266

DOI: 10.24818/BASIQ/2023/09/031

Abstract

For EU countries, keeping and growing innovative business is vital in order to be globally competitive. Innovation in enterprises, as a cornerstone for future production, is of great significance for all countries depending more and more on services. Due to the way innovation is understood within enterprises, such process no longer is bound to single entity. It easily exceeds the sphere of one organisation rapidly evolving to an ecosystem of many actors and 3rd party suppliers and different shareholders across different countries. All these aspects reveal the link between business development, effective created value for the customers and the potential to innovate. The role of information technology both from operational perspective and in the development of innovative potential and new products is currently acknowledged both by enterprises and by small start-up companies. In such context, Eurostat introduced a new tool, "Innovation profiles", to monitor enterprise innovation level.

Thus, our paper explores such data analysing the similarities between 17 European countries. The source is Eurostat and latest data available from 2020, is used. The results of the Cluster Analysis performed in R Shiny app are obtained employing Pearson Correlation coefficient, measuring the Squared Euclidian distance, and the Ward linkage method. Positive moderate to low correlation is obtained in terms of innovation for the three obtained clusters, emphasizing the link between Romania, Poland, and Bulgaria.

Keywords

Cluster analysis, R Shiny app, Innovative enterprises, Romania.

DOI: 10.24818/BASIQ/2023/09/031

Introduction

The term "innovation" originates from the Latin noun *innovates* and has its basic interpretation sourced to the work of Schumpeter (1934) acknowledging it to "new combinations" of knowledge, resources, and other factors. Most of what is seen today in the field of innovation is the rooted work across generations starting in the industrial revolution moving forward to the age of information technology. Recent advances in the field see the innovation relying on scientific effort, technology modernization and social development with responsible innovation benefiting of anticipation, reflexivity, inclusion and responsiveness (Stilgoe et al., 2013).

Innovation is widely acknowledged as the main process which drives economic growth and offers sustainable competitivity leverages for both enterprises and countries. In such terms it stimulates continuous economic growth in developed countries and improves competitiveness in developing countries (Audretsch et al., 2017). At the enterprise level the process of innovation aims to develop innovative technologies based on the results of research and development, scientific and technical advances. Furthermore, Wu (2021) highlights the link between education and enterprise innovation as well as their favoring effect on employment.

As a consequence of the utmost importance of monitoring innovation, Eurostat has developed the "Innovation profiles", a new tool for assessing innovation in the business sector at the European level. It provides insights on the innovation style at enterprise level for each country (Eurostat 2023). Our paper aims to explore similarities between countries considering these new data together with variables concerning innovative business strategies.

1. Review of the scientific literature

Recent advances in the field of innovation from European researchers discussed topics on social innovation (Nicholls and Murdock, 2012), public innovation (Swann, 2014), design-driven innovation (Verganti, 2009) and responsible innovation (Stilgoe et al., 2013), and focus more on the aspects of integrating technological innovation with the social and humanitarian value attributes. Goel and Nelson, (2021) analyse the impact of innovation and Research and development on employment for 127 countries highlighting a positive effect on job creation both for foreign and government owned enterprises.

One may also find a broad scientific literature exploring innovation in enterprises in various contexts. For example, Ionescu et al. (2020) analyse the potential for innovation and entrepreneurship in the EU using a hierarchical cluster analysis and variables such as the Global Innovation Index, the Global Entrepreneurship Index as well as other variables characterising the economic environment. Scutariu et al (2021) focus on e-commerce enterprises in the context of the COVID for 31 countries. Szabo and Herman analyse the productive entrepreneurship in the EU by performing a cluster analysis using several variables including the Summary Innovation Index. Hollenstein (2003) focusses on the innovation modes in the Swiss service sector though the integration of innovation indicators as well as other variables on a firms' activity in a cluster analysis. Laureti et al. (2022) performed a cluster analysis with the k-Means algorithm and Silhouette Coefficient found four clusters among the European countries with a clear division between the Center-Northern countries, that have the higher levels of employment in innovative enterprises compared to the Southern-Eastern counterpart countries.

2. Research Methodology

For the purpose of this paper data on innovative enterprises for 2020 (latest available year) was downloaded from Eurostat. Then, 7 variables representing the share of enterprises by innovative profile in the total number enterprises, regardless of class size (NACE Rev 2 activity: Innovation core activities) have been considered for further analysis as follows:

- V1. The share of enterprises that have developed product innovations themselves, with market novelties (profile I) in the total number enterprises
- V2. The share of enterprises that have developed product innovations themselves, without market novelties (profile II) in the total number enterprises
- V3. The share of enterprises that have developed business process innovations themselves, without product innovation (profile III) in the total number enterprises
- V4. The share of enterprises that have introduced but not themselves developed innovations (profile IV) in the total number enterprises
- V5. The share of enterprises that have innovation activities but not introduced any innovation (ongoing or abandoned innovation activities) (profile V) in the total number enterprises
- V6. The share of enterprises that have no innovation activities but potential to innovate (profile VI) in the total number enterprises
- V7. The share of enterprises that have no innovation activities and no potential to innovate (profile VII) in the total number enterprises.

Also, 6 other variables regarding the business strategy applied by the enterprises were selected (the respective strategy is considered of a high importance):

- V8. Share of innovative enterprises with focus on improving existing goods or services in the total number of innovative enterprises

● V9. Share of innovative enterprises with focus on introducing entirely new goods or services in the total number of innovative enterprises

● V10. Share of innovative enterprises with focus on reaching out to new customer groups in the total number of innovative enterprises

● V11. Share of innovative enterprises with focus on customer specific solution in the total number of innovative enterprises

● V12. Share of innovative enterprises with focus on low-price in the total number of innovative enterprises

● V13. Share of innovative enterprises with focus on high quality in the total number of innovative enterprises

Due to the availability of the data, 17 countries were included in the analysis are: Bulgaria, Czechia, Germany, Estonia, Greece, Spain, Italy, Latvia, Lithuania, Hungary, Malta, Poland, Portugal, Romania Slovakia, Sweden and Türkiye. In order to address the research objective, a cluster analysis is performed.

In order to compute the cluster analysis, the Cluster Analysis R Shiny app developed by Mizumoto (2015) was used. As all variables are interval-ratio type, the Pearson Correlation coefficient as computed in order to analyse the extent the chosen variables are related. In order to group countries into clusters, variables were standardised and the Squared Euclidean distance measure was employed as this measure accounts best for outliers so that countries are grouped mainly due to similarities (Sørensen and Puigvert Gutiérrez, 2006). Also, the Ward linkage method was used as it is the appropriate for quantitative variables (The Pennsylvania State University, 2023) and it is widely used in other similar scientific papers (see for example Pelau and Chinie, 2018; Scutariu et al, 2021).

3. Results

The paper Figure 1 shows the correlation coefficients and the scatter-plot matrices between the selected variables. Most of the variables are relatively low or moderate correlated.

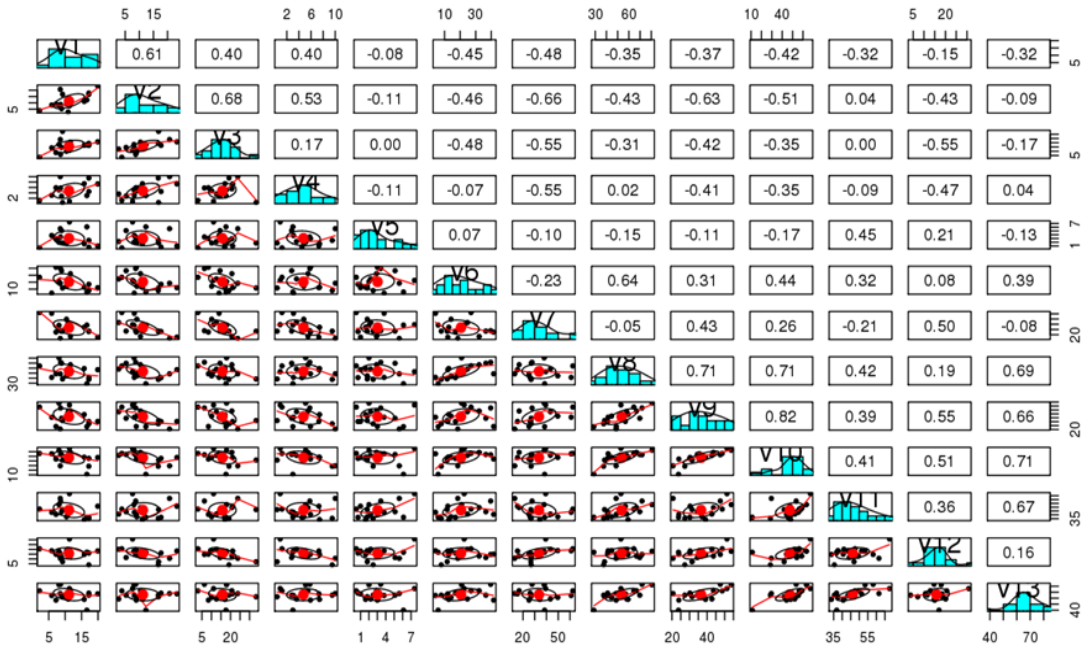

Figure no. 1. Scatter-plot matrices and correlation coefficients between the selected variables
Source: Designed by the authors using the Cluster analysis R Shiny app: Mizumoto, A. (2015). Langtest (Version 1.0) [Web application]. Retrieved from http://langtest.jp

Figure 2 shows the clusters that emerged based on the analysis: cluster 1 (Romania, Poland, Bulgaria and Latvia); cluster 2 (Italy, Sweden, Greece, Czechia, Lithuania, Germany and Estonia); cluster 3 (Hungary,

Malta, Spain, Portugal and Slovakia). Turkey was not included in any cluster, although it is close to the first one.

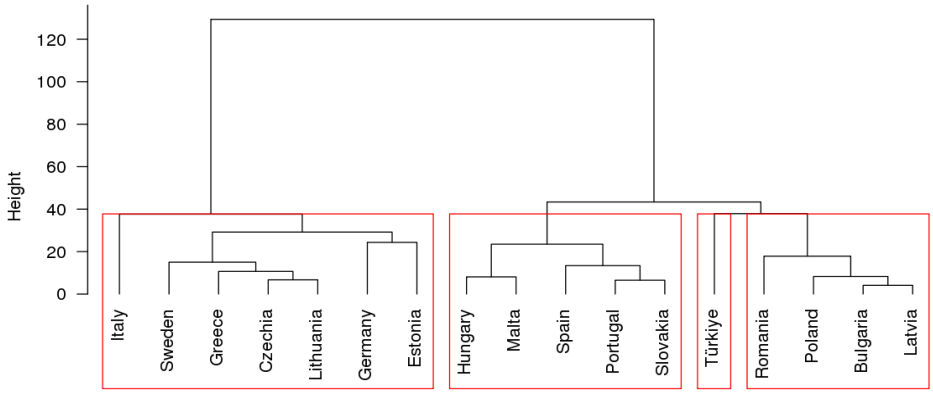

Figure no. 2. Cluster dendrogram

Source: Designed by the authors using the Cluster analysis R Shiny app: Mizumoto, A. (2015). Langtest (Version 1.0) [Web application]. Retrieved from http://langtest.jp

Descriptive statistics for variables in each cluster are included in the Table no. 1, while figure 3 displays a profile for each group. Countries in cluster 1 are characterised by low values for variables V1 to V6 and high values for V7. Countries in cluster 2 present high values for variables V1 to V3 and low values of the rest of the variables. Cluster 3 has a profile approximately symmetric to the profile of cluster 2. Turkey has high values for variables V5, and V9 to V13 and low for the rest.

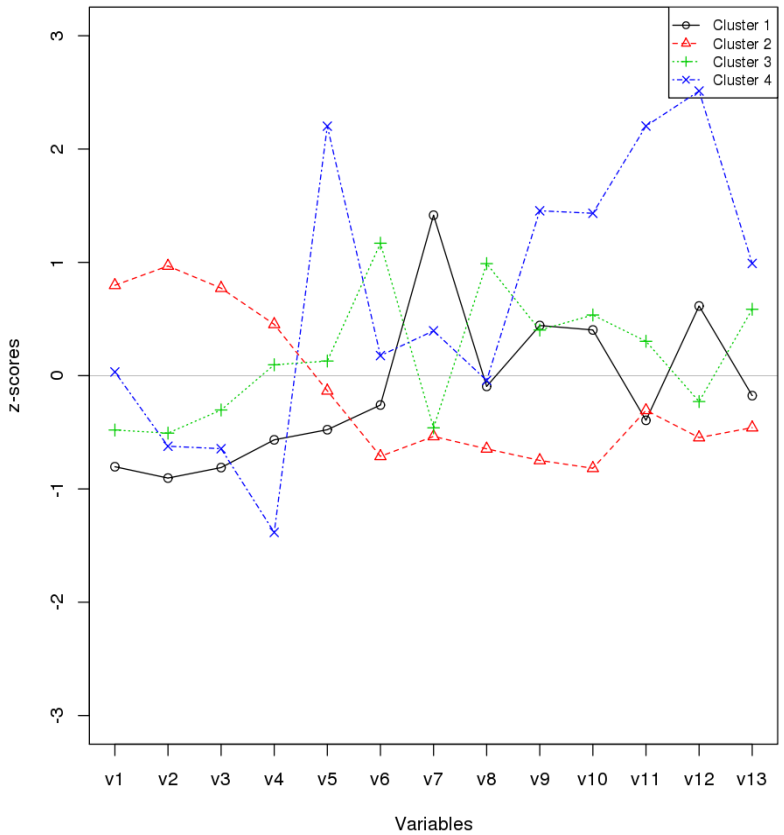

Figure no. 3. Profile plot

Source: Designed by the authors using the Cluster analysis R Shiny app: Mizumoto, A. (2015). Langtest (Version 1.0) [Web application]. Retrieved from http://langtest.jp

Table no. 1. Basic statistics of each cluster

```
Descriptive statistics by group
group: 1
    vars n  mean    sd median trimmed   mad  min  max range  skew kurtosis   se
v1     1 4  7.35  4.10   8.15    7.35  3.71  2.2 10.9   8.7 -0.24    -2.16 2.05
v2     2 4  5.95  3.17   6.00    5.95  3.93  2.7  9.1   6.4 -0.01    -2.36 1.59
v3     3 4  9.90  4.94  11.35    9.90  2.89  2.9 14.0  11.1 -0.53    -1.85 2.47
v4     4 4  3.17  1.50   3.35    3.17  1.26  1.2  4.8   3.6 -0.25    -1.90 0.75
v5     5 4  2.08  1.36   2.00    2.08  1.26  0.5  3.8   3.3  0.12    -1.90 0.68
v6     6 4 17.52 10.49  19.45   17.52  9.49  4.2 27.0  22.8 -0.26    -2.12 5.24
v7     7 4 54.08  9.19  55.30   54.08  9.49 43.3 62.4  19.1 -0.14    -2.26 4.60
v8     8 4 52.15 10.89  51.35   52.15 12.31 41.5 64.4  22.9  0.08    -2.29 5.45
v9     9 4 41.03  5.89  39.90   41.03  5.19 35.7 48.6  12.9  0.28    -2.10 2.94
v10   10 4 52.88  3.96  53.80   52.88  3.11 47.6 56.3   8.7 -0.36    -2.02 1.98
v11   11 4 42.33  4.19  42.15   42.33  4.30 37.5 47.5  10.0  0.08    -1.98 2.10
v12   12 4 19.48  1.03  19.40   19.48  1.19 18.5 20.6   2.1  0.06    -2.34 0.52
v13   13 4 64.15  5.87  62.85   64.15  4.30 58.7 72.2  13.5  0.41    -1.91 2.93
------------------------------------------------------------
group: 2
    vars n  mean    sd median trimmed   mad  min  max range  skew kurtosis   se
v1     1 7 14.91  4.33   16.6   14.91  1.19  8.5 19.9  11.4 -0.52    -1.61 1.64
v2     2 7 16.94  4.30   16.8   16.94  5.78 11.1 23.3  12.2  0.02    -1.55 1.62
v3     3 7 21.30  5.89   19.2   21.30  3.11 16.9 33.4  16.5  1.10    -0.34 2.22
v4     4 7  6.01  3.52    7.3    6.01  4.00  0.4 10.0   9.6 -0.32    -1.62 1.33
v5     5 7  2.76  2.19    1.8    2.76  1.19  0.7  6.1   5.4  0.59    -1.64 0.83
v6     6 7 12.26  4.50   12.1   12.26  6.23  5.9 18.5  12.6  0.02    -1.60 1.70
v7     7 7 25.80  9.22   26.7   25.80 11.27 12.9 38.4  25.5 -0.13    -1.61 3.48
v8     8 7 44.34 11.05   41.1   44.34 11.12 28.8 59.2  30.4  0.11    -1.70 4.17
v9     9 7 28.94  5.95   30.2   28.94  4.30 20.6 39.1  18.5  0.23    -1.14 2.25
v10   10 7 36.01 13.66   41.8   36.01  9.19 11.1 48.0  36.9 -0.66    -1.20 5.16
v11   11 7 43.19 10.40   36.8   43.19  4.60 33.7 61.5  27.8  0.62    -1.38 3.93
v12   12 7 12.54  5.48   14.7   12.54  5.34  3.8 18.3  14.5 -0.43    -1.67 2.07
v13   13 7 61.26 10.10   65.0   61.26  4.15 39.9 68.4  28.5 -1.23    -0.06 3.82
------------------------------------------------------------
group: 3
    vars n  mean    sd median trimmed   mad  min  max range  skew kurtosis   se
v1     1 5  8.88  1.65    8.6    8.88  0.74  7.2 11.6   4.4  0.66    -1.30 0.74
v2     2 5  8.28  1.87    8.7    8.28  0.89  5.3 10.4   5.1 -0.51    -1.37 0.84
v3     3 5 13.56  6.12   14.5   13.56  9.04  6.3 20.7  14.4 -0.06    -2.08 2.74
v4     4 5  5.02  1.28    5.5    5.02  0.59  2.8  5.9   3.1 -0.92    -1.11 0.57
v5     5 5  3.28  1.14    3.0    3.28  0.59  2.0  5.1   3.1  0.52    -1.37 0.51
v6     6 5 34.14  8.19   36.5   34.14  6.38 20.9 41.5  20.6 -0.64    -1.47 3.66
v7     7 5 26.90  7.60   27.0   26.90  2.82 16.7 38.0  21.3  0.14    -1.44 3.40
v8     8 5 67.50 11.70   69.7   67.50  4.30 48.7 80.6  31.9 -0.54    -1.35 5.23
v9     9 5 40.62 11.68   43.6   40.62 13.64 23.9 53.7  29.8 -0.29    -1.80 5.22
v10   10 5 54.70  7.15   54.0   54.70  6.97 44.1 63.3  19.2 -0.27    -1.60 3.20
v11   11 5 49.14  7.02   49.5   49.14  8.01 40.6 58.1  17.5  0.03    -1.95 3.14
v12   12 5 14.44  2.44   14.7   14.44  2.52 11.5 18.0   6.5  0.23    -1.64 1.09
v13   13 5 71.96 11.31   72.5   71.96 14.53 56.8 83.2  26.4 -0.18    -1.99 5.06
------------------------------------------------------------
group: 4
    vars n mean sd median trimmed mad  min  max range skew kurtosis se
v1     1 1 11.3 NA   11.3    11.3   0 11.3 11.3     0   NA       NA NA
v2     2 1  7.6 NA    7.6     7.6   0  7.6  7.6     0   NA       NA NA
v3     3 1 11.1 NA   11.1    11.1   0 11.1 11.1     0   NA       NA NA
v4     4 1  0.9 NA    0.9     0.9   0  0.9  0.9     0   NA       NA NA
v5     5 1  7.4 NA    7.4     7.4   0  7.4  7.4     0   NA       NA NA
v6     6 1 22.6 NA   22.6    22.6   0 22.6 22.6     0   NA       NA NA
v7     7 1 39.3 NA   39.3    39.3   0 39.3 39.3     0   NA       NA NA
v8     8 1 52.9 NA   52.9    52.9   0 52.9 52.9     0   NA       NA NA
v9     9 1 51.3 NA   51.3    51.3   0 51.3 51.3     0   NA       NA NA
v10   10 1 67.1 NA   67.1    67.1   0 67.1 67.1     0   NA       NA NA
v11   11 1 67.7 NA   67.7    67.7   0 67.7 67.7     0   NA       NA NA
v12   12 1 30.8 NA   30.8    30.8   0 30.8 30.8     0   NA       NA NA
v13   13 1 76.1 NA   76.1    76.1   0 76.1 76.1     0   NA       NA NA
```

Source: Designed by the authors

Based on the presented data in Table 1, with basic statistics of each cluster we find that the countries belonging to each cluster in 1st, 2nd, 3rd, and 4th have different mean values. Thus, the average values for variables V1 to V5 get larger from cluster 1 to 4, while for variables V11 to V13 the average value increase from 2nd to 4th cluster. In respect to variables V8 to V10 average values from the descriptive statistics indicate an increase from the 2nd to the 1st followed by the 4th cluster. This means that the 4 clusters show ordering effect depending on the inspected variables.

Conclusions

Innovation is widely acknowledged as the main process which drives economic growth and offers sustainable competitivity leverages for both enterprises and countries. For European countries, keeping and growing innovative business is vital to be globally competitive.

Analysing the data made available by Eurostat on Innovation core activities on the share of enterprises by the innovative profile in the total figures as well on the business strategies applied by the enterprises, our results highlight a positive low to moderate correlation for three clusters.

The results of the Cluster Analysis performed in R Shiny app are obtained employing Pearson Correlation coefficient, measuring the Squared Euclidian distance, and the Ward linkage method. Countries in the first cluster (Romania, Poland, Bulgaria and Latvia) are characterised by low values for V1 to V6 and high values for V7. For these countries in the first cluster this corresponds to having an economic profile which reveals an unexplored capacity and low figures of enterprises with innovation activities and with potential to innovate. Thus, more efforts are needed to increase the total share of innovative enterprises and their potential to innovate with a key focus on the areas where self-innovation with or without market novelties as well as with innovative business development.

The second cluster of countries (Italy, Sweden, Greece, Czechia, Lithuania, Germany and Estonia) resemble high values for variables V1 to V3 and low values for the rest, while the third (Hungary, Malta, Spain, Portugal and Slovakia) shares profile similarities with it. Such aspect for the countries in the second cluster is reflected in high values for the share of enterprises that have developed product innovations themselves, "with or without market novelties (profile I and profile II) in the total number enterprises" as well as in the share of enterprises that have developed business process innovations themselves, without product innovation (profile III) in the total number enterprises and with low values for the remaining variables.

Out of all countries analysed Turkey is placed in an individual cluster as it has high values for V5 and V9 to V13 and low for the rest.

Although economic advances are obtained regionally, challenges still persist when looking into the progress of innovation across countries moreover in the context of paradigms raised by the Sustainable Development Goals (SDGs) to mobilize efforts to end all forms of poverty, inequality and tackle climate change while ensuring no one is left behind.

References

Audretsch, D. B., Acs, Z. J., Lehmann, E.E., and Licht, G., 2017. National systems of innovation. *The Journal of Technology Transfer*, 42, pp.997-1008.

Eurostat, 2023. *Innovation profiles of enterprises – statistics*, [online] Available at: <https://ec.europa.eu/eurostat/statistics explained/index.php?title=Innovation_profiles_of_enterprises_-_statistics > [Accessed 05 April 2023].

Goel, R.K., and Nelson, M.A., 2022. Employment effects of R&D and process innovation: Evidence from small and medium-sized firms in emerging markets. *Eurasian Business Review*, 12(1), pp.97-123.

Hollenstein, H., 2003. Innovation modes in the Swiss service sector: a cluster analysis based on firm-level data. *Research policy*, 32(5), pp.845-863.

Ionescu, G.H., Firoiu, D., Pîrvu, R., Enescu, M., Rădoi, M.I., and Cojocaru, T.M., 2020. The potential for innovation and entrepreneurship in EU countries in the context of sustainable development. *Sustainability*, 12(18), p.7250.

Laureti, L., Costantiello, A., Matarrese, M. and Leogrande, A., 2022. The Employment in Innovative Enterprises in Europe. *SSRN Electronic Journal*. [online] https://doi.org/10.2139/ssrn.4010098.

Nicholls, A. and Murdock, A., 2012. The Nature of Social Innovation. In: A. Nicholls and A. Murdock, eds. *Social Innovation: Blurring Boundaries to Reconfigure Markets*. London: Palgrave Macmillan UK. pp.1–30.

Pelau, C., and Chinie, A.C., 2018. Cluster Analysis for the Determination of Innovative and Sustainable Oriented Regions in Europe. *Studia Universitatis „Vasile Goldis" Arad–Economics Series*, 28(2), pp.36-47.

R Shiny app: Mizumoto, A., 2015. Langtest (Version 1.0) [Web application]. Retrieved from http://langtest.jp

Schumpeter, J., 1934. *The theory of economic development*. Harvard University Press. Cambridge, MA.

Scutariu, A.L., Șușu, Ș., Huidumac-Petrescu, C.E., and Gogonea, R.M., 2022. A cluster analysis concerning the behavior of enterprises with e-commerce activity in the context of the COVID-19 pandemic. *Journal of Theoretical and Applied Electronic Commerce Research*, 17(1), pp.47-68.

Sørensen, C.K. and Puigvert Gutiérrez, J.M., 2006. *Euro area banking sector integration: using hierarchical cluster analysis techniques*. ECB working paper, 627.

Stilgoe, J., Owen, R. and Macnaghten, P., 2013. Developing a framework for responsible innovation. *Research policy*, 42(9), pp.1568-1580.

Swann, G.P., 2014. *Common innovation: how we create the wealth of nations*. Edward Elgar Publishing.

Szabo, Z.K., and Herman, E., 2014. Productive entrepreneurship in the EU and its barriers in transition economies: A cluster analysis. *Acta Polytechnica Hungarica*, 11(6), pp.73-94.

The Pennsylvania State University, 2023. *Pennsylvania State University* [online] Available at: <https://online.stat.psu.edu/stat505/lesson/14/14.7 > [Accessed 05 April 2023].

Verganti, R., 2008. Design, meanings, and radical innovation: A metamodel and a research agenda. *Journal of product innovation management*, 25(5), pp.436-456.

Wu, X., 2021. Discussion on the Organic Integration of Innovation and Entrepreneurship Education and Employment Guidance Education Under Internet +. In: V. Sugumaran, Z. Xu and H. Zhou, eds. *Application of Intelligent Systems in Multi-modal Information Analytics*, Advances in Intelligent Systems and Computing. Cham: Springer International Publishing. pp.17–22. https://doi.org/10.1007/978-3-030-74811-1_3.