

Contents lists available at [ScienceDirect](#)

# Software Impacts

journal homepage: [www.journals.elsevier.com/software-impacts](http://www.journals.elsevier.com/software-impacts)

Original software publication

## eudistance: Distance calculator for the different levels of European NUTS regions

Marcell T. Kurbucz <sup>a,b,\*</sup>, Attila I. Katona <sup>c</sup><sup>a</sup> Department of Computational Sciences, Wigner Research Centre for Physics, 29-33 Konkoly-Thege Miklós Street, H-1121 Budapest, Hungary<sup>b</sup> Department of Statistics, Corvinus University of Budapest, 8 Fővám Square, H-1093, Hungary<sup>c</sup> Department of Quantitative Methods, Faculty of Business and Economics, University of Pannonia, 10 Egyetem Street, H-8200 Veszprém, Hungary

### ARTICLE INFO

#### Keywords:

Nomenclature of Territorial Units for Statistics (NUTS)  
 European Union  
 Distance matrix  
 Spatial analysis  
 Regional studies  
 Geography

### ABSTRACT

The presented R function called `eudistance` calculates the centroid distance matrix for the different levels of the European Union (EU) Nomenclature of Territorial Units for Statistics (NUTS) regions in meters and complements it by their code, name, level, and country identifier, as well as their geometric data and centroids. Optionally, this R function displays centroids on a map of Europe to ease the verification of their positions. The `eudistance` R package (containing the aforementioned function) with the distance dataset it generates is available in this article and publicly on GitHub (Kurbucz and Katona, 2022) and Mendeley Data (Kurbucz and Katona, 2022).

### Code metadata

Current code version	v0.1.0
Permanent link to code/repository used for this code version	<a href="https://github.com/SoftwareImpacts/SIMPAC-2022-61">https://github.com/SoftwareImpacts/SIMPAC-2022-61</a>
Permanent link to Reproducible Capsule	<a href="https://codeocean.com/capsule/9747417/tree/v1">https://codeocean.com/capsule/9747417/tree/v1</a>
Legal Code License	GNU General Public License v3.0
Code versioning system used	Git
Software code languages, tools, and services used	R
Compilation requirements, operating environments & dependencies	R 4.1.3 or later. Dependencies: <code>eurostat</code> , <code>sf</code> , <code>rworldmap</code> , <code>SimDesign</code> . These packages are installed automatically when the function is called (if they have not been previously installed).
If available Link to developer documentation/manual	NA
Support email for questions	<a href="mailto:marcell.kurbucz@wigner.hu">marcell.kurbucz@wigner.hu</a>

## 1. Introduction

In Economics, regional analysis is commonly used to model and understand interactions or flows between geospatial locations. Several researches aimed to conduct regional analysis related to the European Union (EU) Nomenclature of Territorial Units for Statistics (NUTS) regions (<https://ec.europa.eu/eurostat/web/nuts/background>, retrieved: 23 May 2022) in different contexts such as regional development [1], government quality and trade [2], innovation effectiveness [3], road

transport [4], and mobility in higher education [5], to state a few. Obtaining distance-related data among the regions is crucial in such regional studies. Although there are a few existing R packages that can be applied to calculate centroid distances over NUTS regions, they are limited to specific regions and NUTS levels (e.g., the `SDPmod` package [6] only covers the NUTS-3 regions of Germany). To the best of our knowledge, a package that covers the calculation of distances for all NUTS regions or levels is not yet available in the R software. To eliminate this gap, in this paper, the authors provide the `eudistance`

The code (and data) in this article has been certified as Reproducible by Code Ocean: (<https://codeocean.com/>). More information on the Reproducibility Badge Initiative is available at <https://www.elsevier.com/physical-sciences-and-engineering/computer-science/journals>.

\* Corresponding author at: Department of Computational Sciences, Wigner Research Centre for Physics, 29-33 Konkoly-Thege Miklós Street, H-1121 Budapest, Hungary.

E-mail address: [kurbucz.marcell@wigner.hu](mailto:kurbucz.marcell@wigner.hu) (M.T. Kurbucz).

<https://doi.org/10.1016/j.simpa.2022.100327>

Received 27 April 2022; Received in revised form 23 May 2022; Accepted 26 May 2022

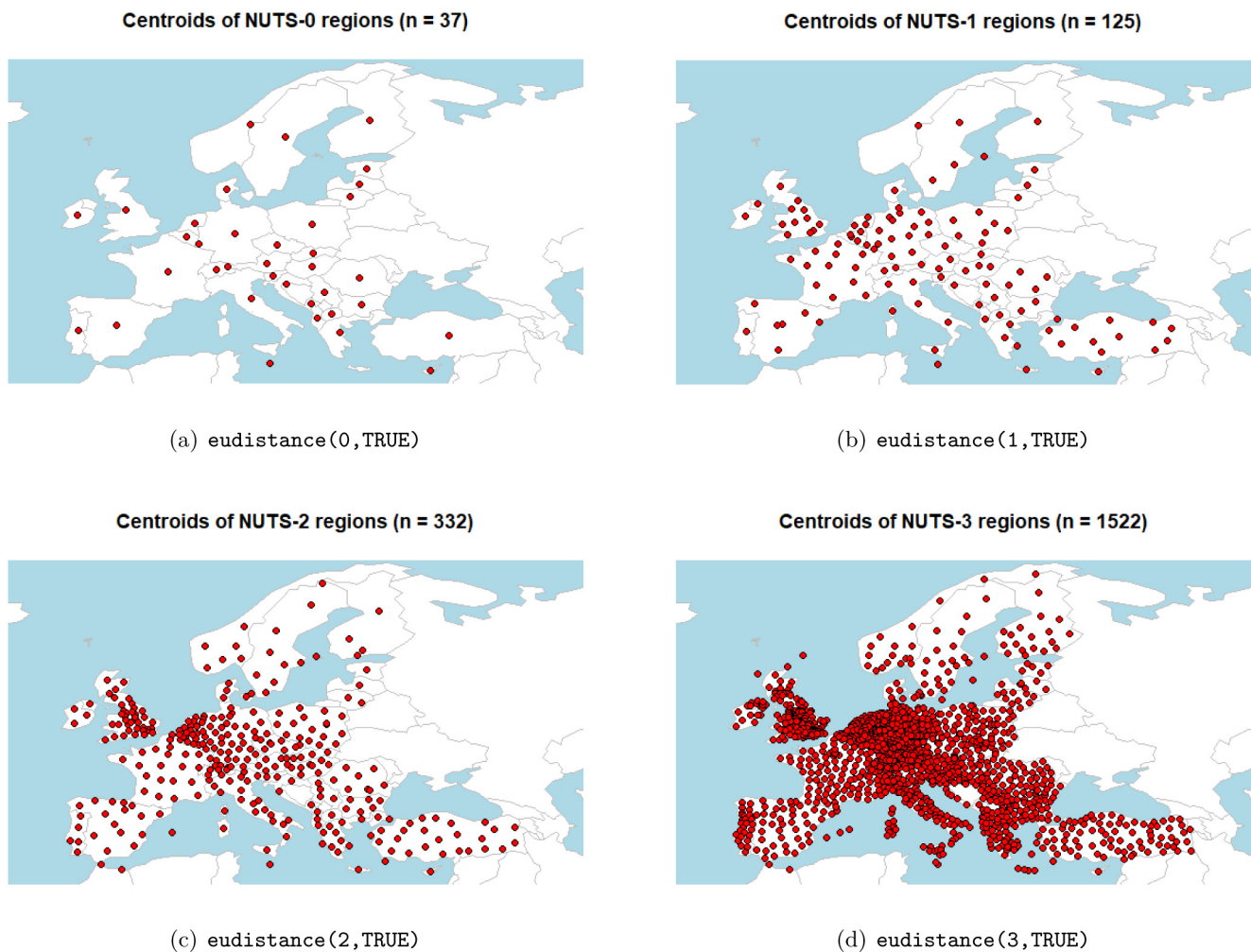


Fig. 1. Centroids of different NUTS regions.

package supporting the researchers in calculating regional distances and visualizing centroids for all NUTS regions and levels.

### Description

The presented R function called `eudistance` calculates the centroid distance matrix for the different levels of NUTS regions in meters and complements it by their code, name, level, and country identifier, as well as their geometric data and centroids. Optionally, this R function displays centroids on a map of Europe to ease the verification of their positions.

To obtain the geometric data for different levels of NUTS regions, `eurostat` (version: 3.7.10) R package [7] was employed. The centroids of these regions, as well as their distances, are calculated by using the `sf` (version: 1.0-7) R package [8]. Centroids are calculated based on the largest contiguous shape of the regions. For visualization, `rworldmap` (version: 1.3-6) R package [9] is applied. These packages are installed automatically when the `eudistance` function is called (if they have not been previously installed).

The `eudistance` R package (containing the aforementioned function) with the dataset it generates is available in this article and publicly on GitHub [10] and Mendeley Data [11].

### Attached files:

- **R function called `eudistance` (`eudistance.R`):** The mandatory input parameter is the NUTS level (0, 1, 2, 3, 'all'). The second input parameter, `map_centroids` (FALSE, TRUE), is

optional and it displays centroids on a map of Europe only if its value is TRUE. By default, its value is FALSE. The output data frame object contains the following variables:

- `NUTS_X_code`: Code of the selected regions.
- `NUTS_X_name`: Name of the selected regions.
- `Level_code`: NUTS level of the selected regions.
- `Country_code`: Code of the countries related to the selected regions.
- `Geometry`: Geospatial information of the selected regions.
- `Centroid`: Centroids of the selected regions.
- `Distance_matrix` (multiple columns): Centroid distances in meters between selected regions.

- **Dataset generated by the `eudistance` function (`distance_data.xlsx`):** Microsoft Office spreadsheet with 5 worksheets as follows:

- NUTS-0: Distances between EU countries (e.g., Germany).
- NUTS-1: Distances between the major socio-economic regions of the EU (e.g., Bundesländer).
- NUTS-2: Distances between the EU basic regions for the application of regional policies (e.g., Regierungsbezirke).
- NUTS-3: Distances between the EU small regions for specific diagnoses (e.g., Kreise).
- All: Distances between all regions mentioned above.

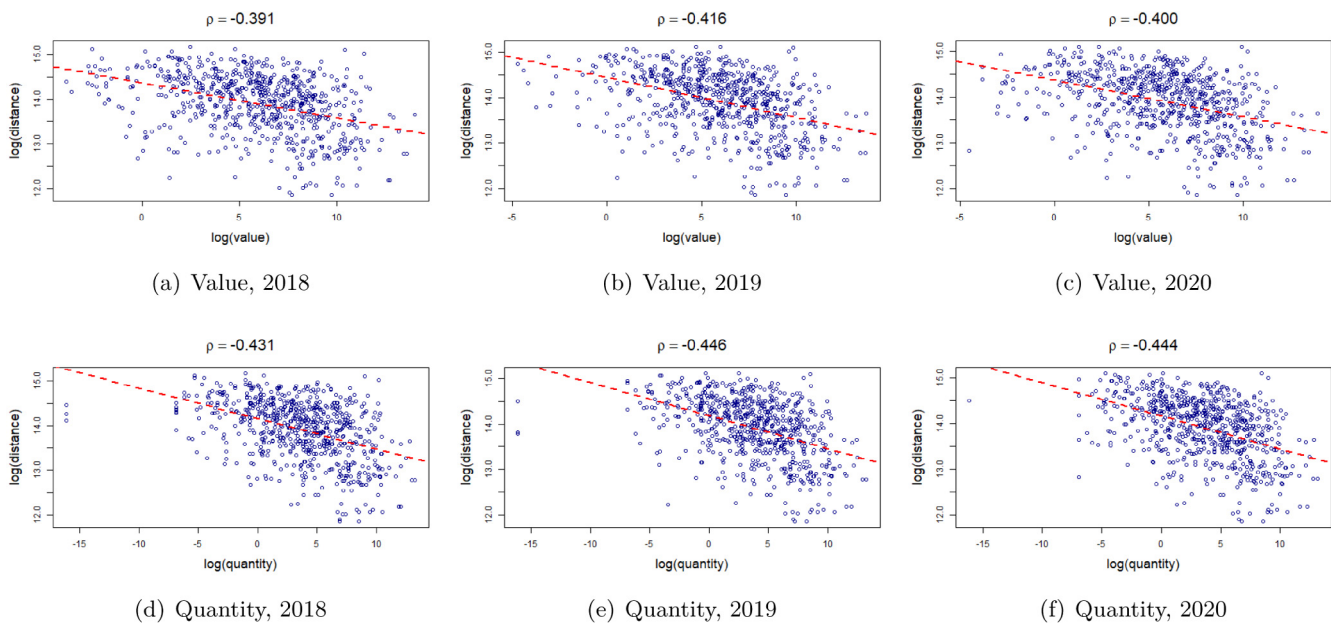


Fig. 2. The relationship between geographical distance and intra-European trade of live animals. (Remark:  $\rho$  denotes the Pearson correlation coefficient.)

### Visualization of centroids

The `eudistance` R function is able to display the centroids of the selected NUTS regions on a map of Europe to ease the verification of their positions. To do this, its second (optional) parameter `map_centroids` must be set `TRUE`. For different NUTS levels, it results in the following figures (see Fig. 1).

### Impact overview

The dataset that can be compiled by the `eudistance` R function is primarily useful for EU-related spatial analysis, regional studies, geography, and international trade research. In contrast to the available R functions (see, e.g., the `gN3dist` function of the `SDPDmod` R package [6], that calculates the centroid distances between the NUTS-3 regions of Germany), the proposed function provides distance and spatial data for all regions of the selected (or all) NUTS levels. To make the function easy to use, required R packages (`eurostat`, `sf`, `rworldmap`, `SimDesign`) are installed automatically when the function is called (if they have not been previously installed). Furthermore, this function displays centroids on a map of Europe optionally to ease the verification of their positions.

In conclusion, the value of the `eudistance` function can be summarized as follows:

- The `eudistance` R function calculates the centroid distance matrix for the different levels of the EU NUTS regions.
- It also collects the code, name, level, and country identifier of the selected regions, as well as their geometrical data and centroids.
- Optionally, this function displays centroids on a map of Europe to ease the verification of their positions.
- Researchers and policymakers in different fields of knowledge can employ the dataset generated by this function for EU-related spatial analysis (e.g., performed with a gravity model).

### An example for the application

Finally, we present a simple example of using the centroid distance matrix compiled by the `eudistance` R function. In this example, we illustrate relationship between the geographical distance of European countries (NUTS-0 regions) and the annual intra-European trade of live animals [12] from 2018 to 2020 (see Fig. 2).

As Fig. 2 shows, both the value and the quantity of intra-European live animal trade highly depend on the distance of the trading countries. Since the welfare of animals can be heavily impaired during their

transportation (see, e.g., [13,14]), the tougher regularization and the restriction of this area is currently one of the European Commission's priorities [15,16]. The effect of efforts to reduce livestock transport would be reflected in a reduction of the correlation coefficients ( $\rho$ ) between distances and live animal trade over time. As shown in Fig. 2, these coefficients were relatively stable between 2018 and 2020.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

Marcell T. Kurbucz thanks the support of Eötvös Loránd Research Network.

### References

- [1] M. Stanickova, L. Melecký, Understanding of resilience in the context of regional development using composite index approach: the case of European union NUTS-2 regions, *Reg. Stud. Reg. Sci.* 5 (1) (2018) 231–254.
- [2] J. Barbero, G. Mandras, E. Rodríguez-Crespo, A. Rodríguez-Pose, Quality of government and regional trade: evidence from European Union regions, *Reg. Stud.* 55 (7) (2021) 1240–1251.
- [3] A. Zóttaszek, A. Olejnik, Regional effectiveness of innovation: leaders and followers of the EU NUTS 0 and NUTS 2 regions, *Innov. Euro. J. Soc. Sci. Res.* (2021) 1–22.
- [4] D. Persyn, J. Díaz-Lanchas, J. Barbero, Estimating road transport costs between and within European union regions, *Transp. Policy* (2020).
- [5] Z.T. Kosztyán, V.V. Csányi, Z. Banász, A. Jakobi, I. Neumanné-Virág, A. Telcs, The role of higher education in spatial mobility, *Appl. Netw. Sci.* 6 (1) (2021) 1–30.
- [6] S. Rozeta, Spatial dynamic panel data models (SDPDmod package), 2021, GitHub, URL: <https://github.com/rozetasimonovska/sdpdmod/>.
- [7] L. Lahti, J. Huovari, M. Kainu, P. Biecek, Retrieval and analysis of eurostat open data with the `eurostat` package, *R J.* 9 (1) (2017) 385.
- [8] E.J. Pebesma, et al., Simple features for R: Standardized support for spatial vector data, *R J.* 10 (1) (2018) 439.
- [9] A. South, `Rworldmap`: A new R package for mapping global data, *R J.* 3 (1) (2011).
- [10] M.T. Kurbucz, A. Katona, `Eudistance`: Distance calculator for the different levels of European NUTS regions, 2022, GitHub, URL: <https://Github.Com/Mtkurbucz/Eudistance/>.

- [11] M.T. Kurbucz, A. Katona, Distance calculator for the different levels of European NUTS regions, 2022, Mendeley Data, URL: <https://data.mendeley.com/datasets/hvjzvpfgbp/>.
- [12] G. Gaulier, S. Zignago, BACI: International Trade Database at the Product-Level. The 1994–2007 Version, CEPII, Working Paper, 2010-23, 2010, URL: <http://www.cepii.fr/cepii/fr/publications/wp/abstract.asp?nodoc=2726>.
- [13] J. Schrama, W. Van der Her, J. Gorssen, A. Henken, M. Verstegen, J. Noordhuizen, Required thermal thresholds during transport of animals, *Vet. Q.* 18 (3) (1996) 90–95.
- [14] N. Bachelard, Animal transport as regulated in europe: a work in progress as viewed by an NGO, *Animal Front.* 12 (1) (2022) 16–24.
- [15] European Union, European parliament resolution of 14 february 2019 on the implementation of council regulation (EC) no 1/2005 on the protection of animals during transport within and outside the EU (2018/2110 (INI)), 2019.
- [16] European Commission, Farm to fork strategy for a fair, healthy and environmentally-friendly food system, in: Communication from the Commission To the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, European Commission Brussels, Belgium, 2020.