



HELLENIC  
STATISTICAL  
AUTHORITY -  
GREECE



NATIONAL  
INSTITUTE OF  
STATISTICS



ECOLOGICAL  
UNIVERSITY OF  
BUCHAREST



STATISTIK  
AUSTRIA



CENTRAAL  
BUREAU  
VOOR DE  
STATISTIEK  
NETHERLANDS



THE UNIVERSITY  
OF BUCHAREST

# The 12th International Conference The Use of R in Official Statistics uRos2024

## PROGRAMME

On-site event

Hosted by Hellenic Statistical Authority - Greece

Piraeus, Greece

27-29 November 2024



<http://urosconf.org/>

## SCIENTIFIC BOARD

*Athanasios C. Thanopoulos (President, Hellenic Statistical Authority)*  
*Tudorel Andrei (President, National Institute of Statistics - Romania)*  
*Matthias Templ (ZHAW school of engineering, Switzerland)*  
*Adrian Dusa (University of Bucharest, Romania)*  
*Kazumi Wada (Ministry of Internal Affairs and Communications, Japan)*  
*Marcello d’Orazio (Italian National Institute of Statistics - Rome, Italy)*  
*Valentin Todorov (UNIDO, Vienna Austria)*  
*Alina Matei (University of Neuchatel, Switzerland)*  
*Bernhard Meindl (Statistics Austria, Austria)*  
*Edwin de Jonge (Statistics Netherlands, The Netherlands)*  
*Gergely Daróczi (RX Studio, USA/Hungary)*  
*Kamarul Ariffin Mansor (MARA University of Technology, Malaysia)*  
*Marius Nicolae Jula (University of Bucharest, Romania)*  
*Matyas Meszaros (Eurostat, Luxemburg)*  
*Mihaela Paun (University of Bucharest)*  
*Roxana Adam (University of Bucharest, Romania)*  
*Silvia Pisciă (Statistics Romania)*  
*Daniela Ștefănescu (Statistics Romania)*  
*Bogdan Oancea (University of Bucharest, Statistics Romania)*  
*Raluca-Mariana Drăgoescu (Statistics Romania)*

## ORGANIZING COMMITTEE

*Katerina Kalioti (Hellenic Statistical Authority - ELSTAT)*  
*Georgia Panagopoulou (Hellenic Statistical Authority – ELSTAT)*  
*Ana Maria Ciuhu (Statistics Romania and Institute of National Economy, Romanian Academy)*  
*Bogdan Oancea (University of Bucharest, Statistics Romania)*  
*Marian Necula (Statistics Romania)*  
*Silvia Pisciă (Statistics Romania)*  
*Daniela Ștefănescu (Statistics Romania)*  
*Ciprian Alexandru (Ecological University of Bucharest)*  
*Alexander Kowarik (Statistics Austria)*  
*Mark van der Loo (Statistics Netherlands)*  
*Raluca-Mariana Drăgoescu (Statistics Romania)*

## SECRETARIATE

*Katerina Kalioti, Hellenic Statistical Authority (ELSTAT)*  
*Ana Maria Ciuhu, National Institute of Statistics, Romania*

## CONTACT

*uros2024 [at] statistics.gr*

## PROGRAMME

### November 27, Wednesday

9<sup>00</sup> – 9<sup>45</sup> **Registration**

9<sup>45</sup> - 10<sup>00</sup> **Welcome message**

10<sup>00</sup> – 12<sup>00</sup> **Tutorial session I**



**“nonprobsvy: an R package for modern methods for non-probability surveys”**

Łukasz Chrostowski and Maciej Beręsewicz, Statistics Poland

12<sup>00</sup> – 13<sup>00</sup> **Lunch break**

13<sup>00</sup> – 15<sup>00</sup> **Tutorial session II**



**“R packages around JDemetra+: a versatile toolbox for time series analysis”**

Anna Smyk and Tanguy Barthelemy, INSEE France

15<sup>00</sup> – 15<sup>30</sup> **Coffee break**

15<sup>30</sup> – 17<sup>30</sup> **Tutorial session III**



**“Data validation”**

Mark Van Der Loo, Statistics Netherlands (CBS)

### November 28, Thursday

8<sup>30</sup> – 8<sup>45</sup> **Walk In**

8<sup>45</sup> – 9<sup>00</sup> **Welcome message**

9<sup>00</sup> – 10<sup>00</sup> **Keynote Speaker 1:** Sandra Barragan, Statistics Spain (INE)



*Chairman:* Alexander Kowarik, Statistics Austria

Presentation - “Why and for what purpose R in Official Statistics?”

10<sup>00</sup> – 10<sup>30</sup> **Coffee Break**

10<sup>30</sup> – 12<sup>00</sup> **Scientific Session *R in Organization***

*Chairman:* Marian Necula, Romanian National Institute of Statistics



1. *Using R and SAS Viya for Statistical Production – a Multiple-Tool*

*Approach* Andreas Jahn, Federal Statistical Office Germany (Destatis)

Benedikt Ellinger, Federal Statistical Office Germany (Destatis)

Bernhard Fischer, Federal Statistical Office Germany (Destatis)



2. *Managing a Central Server for R Programming at Destatis: Practical Insights and Challenges*

Benedikt Ellinger, Federal Statistical Office Germany (Destatis)  
Andreas Jahn, Federal Statistical Office Germany (Destatis)  
Bernhard Fischer, Federal Statistical Office Germany (Destatis)



3. *Romania's experience in using R in the official statistics*

Bogdan Oancea, University of Bucharest; National Institute of Statistics  
Marian Necula, National Institute of Statistics; University of Bucharest  
Ana-Maria Ciuhu, National Institute of Statistics; Institute of National Economy, Romanian Academy  
Ciprian Alexandru, Ecological University of Bucharest  
Raluca-Mariana Drăgoescu, National Institute of Statistics



4. *Encouraging use of R at NSTAC Japan*

Ichiro Murata, National Statistics Center, Japan

**12<sup>00</sup> – 13<sup>00</sup> Lunch Break & Group Photo**

**13<sup>00</sup>– 14<sup>15</sup> Scientific Session *Machine Learning and AI***

*Chairman:* Bogdan Oancea, University of Bucharest



1. *Supervised statistical (machine) learning for domain estimation with business survey data*

Vasilis Chasiotis, Department of Statistics, Athens University of Economics and Business, Athens, Greece



2. *The Synergy of R and Generative AI in Statistics*

Vytas Vaiciulis, Central Statistics Office, Ireland



3. *Automating Data Validation on SQL Server Using R and the Machine Learning package*

Paula Hartung, Statistics Iceland

4. Lightning talk: *RMLUtils: an Official Statistics oriented common interface for Machine Learning*

Luis Sanguiao Sande, Statistics Spain (INE)  
Carlos Sáez Calvo, Statistics Spain (INE)  
Sandra Barragán Andrés, Statistics Spain (INE)  
Ester Puerto Sanz, Statistics Spain (INE)  
María Novás Filgueira, Statistics Spain (INE)  
Javier Villaescusa Almagro, Statistics Spain (INE)  
Sergio Pardina Quirós, Statistics Spain (INE)  
Álvaro García Tenorio, Statistics Spain (INE)



**14<sup>15</sup> - 15<sup>00</sup> Scientific Session *Data cleaning***

*Chairman:* Ana-Maria Ciuhu, Romanian National Institute of Statistics;  
Institute of National Economy, Romanian Academy



1. *A process for making imputations in official statistics*  
Claude Lamboray, STATEC Luxembourg  
Johann Neumayr, STATEC Luxembourg



2. *Automatic validation in R for EU-SILC microdata*  
Margherita Zupardo, Statistics Iceland

**15<sup>00</sup> - 15<sup>30</sup> Coffee Break**

**15<sup>30</sup> - 16<sup>30</sup> Scientific Session *NLP in Official Statistics and Time series***

*Chairman:* Mark Van Der Loo, Statistics Netherlands (CBS)



1. *A Neural Network Approach to Text Classification for International Standardized Codes*  
Nina Niederhametner, Statistics Austria  
Alexander Kowarik, Statistics Austria  
Johannes Gussenbauer, Statistics Austria



2. *Semantic address matching using Keras for R*  
Paula Cruz, Statistics Portugal; NOVA IMS  
Leonardo Vanneschi, NOVA IMS  
Marco Painho, NOVA IMS  
Filipa Ribeiro, Statistics Portugal



3. *TEAM: an R package for time series model identification*  
C. Sáez Calvo, S.G. for Methodology and Sampling Design, Statistics Spain  
L. Sanguiao Sande, S.G. for Methodology and Sampling Design, Statistics Spain  
Félix Aparicio Pérez, S.G. for Methodology and Sampling Design, Statistics Spain  
María Teresa Vázquez Gutiérrez, S.G. for Information Technologies and Communications, Statistics Spain  
José Fernando Arranz Arauzo, S.G. for Information Technologies and Communications, Statistics Spain



4. *MEDOS: A Shiny Application for Seasonal Adjustment of Short-Term Statistics*  
Muhammed Fatih Tüzen, Turkish Statistical Institute (TURKSTAT)

**16<sup>30</sup> – 17<sup>45</sup> Scientific Session *Using R in New Data Sources & Report and GUIs***

*Chairman:* Ciprian Alexandru, Ecological University of Bucharest



1. *New data sources and Official Statistics: An exemplary research workflow in R using different API wrapper packages*

Yannik Buhl, Federal Statistical Office Germany (Destatis)

2. *Use of R tools in the statistical disclosure control of census data in Poland*

Kamil Wilak, Statistical Office in Poznań, Centre for Small Area Estimation; Poznań University of Economics and Business, Chair of Statistics



Tomasz Józefowski, Statistical Office in Poznań, Centre for Small Area Estimation; Poznań University of Economics and Business, Chair of Statistics

Andrzej Młodak, Statistical Office in Poznań, Centre for Small Area Estimation; The University of Kalisz, Inter-faculty Department of Mathematics and Statistics

Tomasz Klimanek, Statistical Office in Poznań, Deputy Director; Poznań University of Economics and Business, Chair of Statistics



3. *Streamlining Web Paradata Retrieval and Analysis at Statistics Sweden: A Semi-Automated Approach for Enhanced Efficiency and Consistency*

Gustaf Andersson, Statistics Sweden

Emma Stavås, Statistics Sweden



4. *Rmarkdown report in R Shiny*

Olesja Nikoluskina, The Central Statistical Bureau of Latvia



5. *Transforming Excel reports into SMDX compatible datasets with R*

Athanasios Stavrakoudis, Applied Informatics and Computational Economics Lab, Department of Economics, University of Ioannina, Greece

19<sup>30</sup> **Official Dinner**

---

November 29, Friday

8<sup>30</sup> – 9<sup>00</sup> Walk In

9<sup>00</sup> – 10<sup>00</sup> **Keynote Speaker 2:** Romain Lesur, INSEE France



Presentation - Moving away from SAS: an opportunity to modernise the practices of statisticians

10<sup>00</sup> – 10<sup>30</sup> Coffee Break

10<sup>30</sup> – 12<sup>00</sup> **Scientific Session *Methods for Official Statistics***

*Chairman: Ioanna Markaki, ELSTAT*



1. *Access to official statistics from R: part II*

Olav ten Bosch, CBS Netherlands  
Edwin de Jonge, CBS Netherlands



2. *Detection and visualization of outliers in establishment surveys*

Valentin Todorov, UNIDO (retired)



3. *Is statistical matching feasible?*

Marcello D'Orazio, Statistics Italy (ISTAT)



4. *Testing for the bias in the estimation of business structure indexes from different data sources*

Michaela Balkoudi, Aristotle University of Thessaloniki, Department of Mathematics



5. *Regional food purchasing behavior characteristics contained in the Household Income and Expenditure Survey data*

Atsushi Kimura, National Statistics Center, Japan

12<sup>00</sup> – 13<sup>30</sup> **Scientific Session *Sampling and Estimation***

*Chairman: Marius Nicolae Jula, University of Bucharest*



1. *Joint calibration estimators for totals and quantiles for probability and nonprobability samples*

Maciej Beręsewicz, Poznań University of Economics and Business; Statistical Office in Poznań  
Marcin Szymkowiak, Poznań University of Economics and Business; Statistical Office in Poznań

2. *nonprobsvy: an R package for modern methods for non-probability surveys*

Łukasz Chrostowski, Faculty of Mathematics and Informatics, Adam Mickiewicz University



Piotr Chlebicki, Faculty of Mathematics and Informatics, Adam Mickiewicz University; Centre for the Methodology of Population Studies, Statistical Office in Poznań

Maciej Beręsewicz, Poznań University of Economics and Business, Department of Statistics; Centre for the Methodology of Population Studies, Statistical Office in Poznań, Poland

3. *Estimation of population parameters in EL-SILC with the R packages: the experience of ELSTAT*



Irene Sarantou, Hellenic Statistical Authority

Anastasia Mnimatidou, National Technical University of Athens

4. *blocking: an R package for blocking of records for probabilistic record linkage based on approximate nearest neighbors algorithms*



Maciej Beręsewicz, Poznań University of Economics and Business, Department of Statistics; Centre for the Methodology of Population Studies, Statistical Office in Poznań, Poland

**13<sup>30</sup> – 14<sup>30</sup> Lunch Break**

**14<sup>30</sup> – 15<sup>30</sup> Scientific Session *Disclosure control, dissemination and visualization***

*Chairmen: Christina Pierrakou and Ioannis Monopolis, ELSTAT*

1. *Use of R to automate Eurostat statistical publications: Insights from the third edition of the EMOS Coding Lab*



Meszaros Matyas, Eurostat

Gallelli Andrea, Eurostat

Steenvoorden Tina, GOPA Worldwide Consultants

Maja Islam, Eurostat



2. *Using R for efficient content production - The case of a monthly Statistics Explained article*

Bogdan Alexandru Micu, European Commission – DG TAXUD, Belgium



3. *Statistical disclosure control for Census data with sdcMicro package: the computational challenges*

Vasiliki Spiliopoulou, Hellenic Statistical Authority

Dionysios Fragkopoulos, Hellenic Statistical Authority



15<sup>30</sup> – 16<sup>30</sup> **Scientific Session *R in Production: Automation***

*Chairman: Dionysis Fragkopoulos, ELSTAT*



1. *Interactive visualizations of very large datasets using {shiny} and Base R*

Rahul Sangole, USA



2. *Statistical Business Registers updates monitoring and dashboards*

Pedro Carrasqueira, Statistics Portugal

António Portugal, Statistics Portugal

Isabel Farinha, Statistics Portugal



3. *Automatic Classification of Economic Activities with NOGAuto: Exploring Rules for Use in Statistical Production*

Athanasia Chalimourda, Swiss Federal Statistical Office (SFSO)

Mathias Constantin, Swiss Federal Statistical Office (SFSO)



4. Lightning talk: *A practical example for embedding R into Eviews-programming*

Michel Geller, STATEC Luxembourg

16<sup>30</sup> – 17<sup>00</sup> Coffee Break

17<sup>00</sup> – 18<sup>00</sup> **Scientific Session *Data analysis***

*Chairman: Adrian Duşa, University of Bucharest*



1. *Leveraging R for Official Statistics: A Case Study Using Traffic Count Data*

Nele van der Wielen, Central Statistics Office Ireland



2. *Utilizing R for Simulating Studies in International Large-Scale Assessments*

Umut Atasever, IEA Hamburg (International Association for the Evaluation of Educational Achievement), Germany



Francis L. Huang, University of Missouri, USA

Leslie Rutkowski, Indiana University, USA



3. *Unemployment, Net Income and Emigration. An Econometric Study*

Marius Nicolae Jula, Faculty of Business and Administration

Dorin Jula, Institute of Economic Forecasting, Romanian Academy and Faculty of Financial Management, Ecological University of Bucharest



4. *Estimating equivalized income for school students in Austria*

Dominik Ernst, Statistics Austria

Johannes Gussenbauer, Statistics Austria



5. Lightning talk: *Time Series Analysis of Well-being and Poverty Using R*  
Tudor-Ștefan Irimia, Bucharest Academy of Economic Studies  
Iliana Caragea, Bucharest Academy of Economic Studies  
Emilia Țițan, Bucharest Academy of Economic Studies

**18<sup>00</sup> Closing Remarks**

## BRIEF AGENDA

### **Wednesday, 27 November 2024**

9:00 – 9:45 Registration

9:45 – 10:00 Welcome message

10:00 - 12:00 Tutorial session I “nonprobsvy: an R package for modern methods for non-probability surveys” by Łukasz Chrostowski and Maciej Beręsewicz, Statistics Poland

12:00 - 13:00 Lunch break

13:00 - 15:00 Tutorial session II “R packages around JDemetra+: a versatile toolbox for time series analysis” by Anna Smyk and Tanguy Barthelemy, INSEE France

15:00-15:30 Coffee break

15:30-17:30 Tutorial session III “Data validation” by Mark van der Loo, CBS Netherlands

### **Thursday, 28 November 2024**

08:30 – 08:45 Walk in

08:45 – 09:00 Welcome message

09:00 - 10:00 Keynote speaker 1 “Why and for what purpose R in Official Statistics?”  
by Sandra Barragan

10:00-10:30 Coffee break

10:30 - 12:00 Scientific session I R in Organization

12:00 - 13:00 Lunch break and Group photo

13:00 - 14:00 Scientific session II Machine Learning and AI

14:00-15:00 Scientific session III Data cleaning

15:00-15:30 Coffee break

15:30-16:30 Scientific session IV NLP in Official Statistics and Time series

16:30-17:30 Scientific session V Using R in New Data Sources & Report and GUIs

19:30 Official dinner

### **Friday, 29 November 2024**

08:30 – 09:00 Walk in

09:00 - 10:00 Keynote speaker 2 “Moving away from SAS: an opportunity to modernise the practices of statisticians” by Romain Lesur

10:00-10:30 Coffee break

10:30 - 12:00 Scientific session I Methods for Official Statistics

12:00 - 13:30 Scientific session II Sampling and Estimation

13:30-14:30 Lunch break

14:30 - 15:30 Scientific session III Disclosure control, dissemination and visualization

15:30 - 16:30 Scientific session IV R in Production: Automation

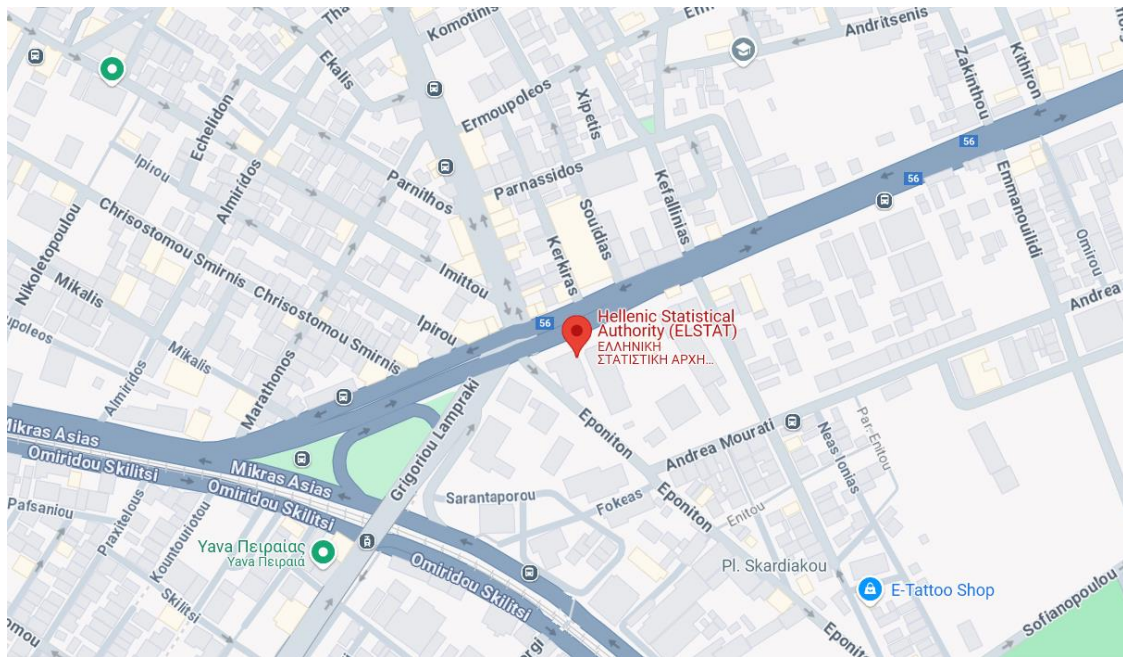
16:30-17:00 Coffee break

17:00-18:00 Scientific session V Data analysis

18:00 Closing remarks

## CONFERENCE VENUE

Hellenic Statistical Authority (ELSTAT)  
46, Pireos and Eponiton street  
18547- Piraeus



**Contact:**

Hellenic Statistical Authority (ELSTAT)  
46, Pireos and Eponiton street  
18547- Piraeus  
[www.statistics.gr](http://www.statistics.gr)  
e-mail: uros2024 [at] statistics.gr