ENV-444

Exploratory data analysis in environmental health

Guessous Idris, Joost Stéphane

Cursus	Sem.	Type
Digital Humanities	MA1, MA3	Opt.
Environmental Sciences and Engineering	MA1, MA3	Opt.
Mineur STAS Russie	H	Opt.

Language of teaching	English
Credits	4
Session	Winter
Semester	Fall
Exam	Oral
Workload	120h
Weeks	14
Hours	**3 weekly**
Courses	1 weekly
Project	2 weekly
Number of positions	

Summary

This course teaches how to apply exploratory spatial data analysis to health data. Teaching focuses on the basics of spatial statistics and of epidemiology, and proposes a context to analyse geodatasets making it possible to study the relationship between health and the environment.

Content

This course consists of a theoretical and of a practical part. Each week 45 minutes of ex-catedra teaching are used to present the history of the discipline (exploratory data analysis and spatial epidemiology), of the main theoretical framework and of the main software used.
The theory is followed by 1h30 of practical work during which the notions covered by the theory are implemented in the computer lab. The data used are provided by current research case studies related to environmental health (e.g. sleep troubles and road traffic noise). Data exploration, geovisualization, geocomputation, communication (semiology of graphics) and representation (thematic mapping) are implemented in the context of exercises. In parallel, the data provided and the approaches taught have to be used to write scientific papers (1 short and 1 regular).

Keywords

Exploratory spatial data analysis; environmental health; spatial epidemiology; Exposome; Geocomputation; EDA; ESDA; Geovisualization; GIS; Geoda; Thematic mapping; Semiology of graphics; Spatial statistics; Scientific paper writing

Learning Prerequisites

Recommended courses

- Systèmes d'Information Géographique (SIG), ENV-342, bachelor, 3ème année
- MOOCs Systèmes d'Information Géographique 1 et 2, Coursera

Important concepts to start the course

Statistics; Geographic Information Systems; Epidemiology

Learning Outcomes

By the end of the course, the student must be able to:

- Investigate the variation of attributes according to the change of the location of a set of spatial units
- Elaborate a research project based on the characteristics of a georeferenced data set available
- Formulate hypotheses to be validated in the context of a research project

- Report on the main results obtained in the context of a research project
- Interpret the main results obtained based on the spatial distribution of the objects under investigation
- Structure ideas and arguments in the context of the writing of short scientific papers
- Produce adequate geospatial data sets for the processing of spatial statistics and association models

Transversal skills

- Set objectives and design an action plan to reach those objectives.
- Use a work methodology appropriate to the task.
- Identify the different roles that are involved in well-functioning teams and assume different roles, including leadership roles.
- Take feedback (critique) and respond in an appropriate manner.
- Make an oral presentation.
- Write a scientific or technical report.
- Summarize an article or a technical report.
- Negotiate effectively within the group.

Teaching methods

- Ex-cathedra teaching,
- Exercises (practicals in geolab)
- Writing of short scientific articles
- MOOC

Expected student activities

Attend and participate in theoretical courses, apply the instructions proposed during the practicals, write short articles, present a project orally, show initiative

Assessment methods

Continuous control during the semester:

- 2 exercises (individual) = 10%
- 1 short paper (individual) = 20%
- 1 long paper (group) = 30%
- 1 oral exam (individual) = 40%

Supervision

Office hours	Yes
Assistants	Yes
Forum	Yes

Resources

Bibliography

- Anselin L, McCann M (2009) OpenGeoDa, Open Source Software for the Exploration and Visualization of Geospatial Data. In: Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems GIS Â¿09., pp. 550Â¿551. ACM, New York, NY, USA.

• Cui Yuxia, Balshaw David M., Kwok Richard K., Thompson Claudia L., Collman Gwen W., & Birnbaum Linda S. (2016). The Exposome: Embracing the Complexity for Discovery in Environmental Health. *Environmental Health Perspectives*, *124*(8), A137â##A140. doi: 10.1289/EHP412

• Harris TM (2017) Exploratory Spatial Data Analysis: Tight Coupling Data and Space, Spatial Data Mining, and Hypothesis Generation. In: Regional Research Frontiers - Vol. 2, pp. 181Â¿191. Springer, Cham.

• Morgenthaler, Stephan (2009)Â Exploratory data analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 1, 33-44

• Siroux, V., Agier, L., & Slama, R. (2016). The exposome concept: a challenge and a potential driver for environmental health research. *European Respiratory Review*, *25*(140), 124â##129. doi: 10.1183/16000617.0034-2016

• Tukey JW (1980) We Need Both Exploratory and Confirmatory. The American Statistician, **34**, 23-25.

Ressources en bibliothèque

• Anselin L, McCann M (2009) OpenGeoDa, Open Source Software for the Exploration and Visualization of Geospatial Data

• Morgenthaler, Stephan (2009) Exploratory data analysis

• Tukey JW (1980) We Need Both Exploratory and Confirmatory

Notes/Handbook

Lecture notes are gradually distributed to students during the semester.

Websites

• https://en.wikipedia.org/wiki/Environmental_health

Moodle Link

• https://moodle.epfl.ch/course/view.php?id=6351