ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ανοικτάμαθήματα
opencourses

Démographie spatiale/Spatial Demography

Session 2: Pitfalls of spatial data & analysis

Michail Agorastakis

Department of Planning & Regional Development

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Θεσσαλίας**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ecological Fallacy (1)

"*The ecological fallacy is an error in the interpretation of statistical data in an ecological study, whereby inferences about the nature of specific individuals are based solely upon aggregate statistics collected for the group to which those individuals belong. This fallacy assumes that individual members of a group have the average characteristics of the group at large*." **Source**: Huhua & Ruibo (2012)

"*A term used when spatially aggregated data are analysed and the results assumed to apply to relationships at the individual level. In most cases analyses based on area level means give conclusions very different from those that would be obtained from an analysis of unit level data*." **Source:** Everitt (2006)

Ecological Fallacy (2)

"The ecological fallacy is a situation that can occur when a researcher or analyst makes an inference about an individual based on aggregate data for a group."

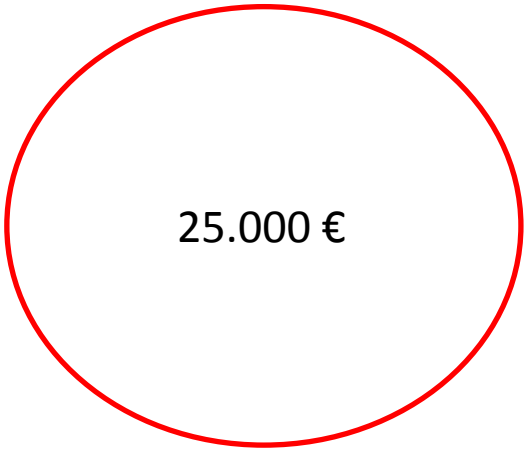
"Assumptions made about individuals based on aggregate data are vulnerable to the ecological fallacy."

"This does not mean that identifying associations between aggregate figures is necessarily defective, and it doesn't necessarily mean that any inferences drawn about associations between the characteristics of an aggregate population and the characteristics of sub-units within the population are absolutely wrong either. What it does say is that the process of aggregating or disaggregating data may conceal the variations that are not visible at the larger aggregate level, and researchers, analysts and crime mappers should be careful."

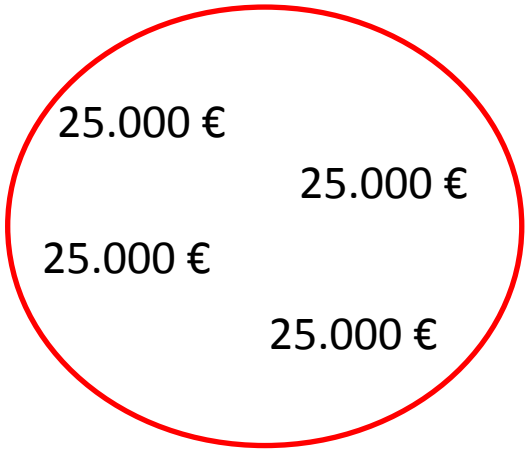
Source: http://www.jratcliffe.net/blog/the-ecological-fallacy/

Ecological Fallacy (3)

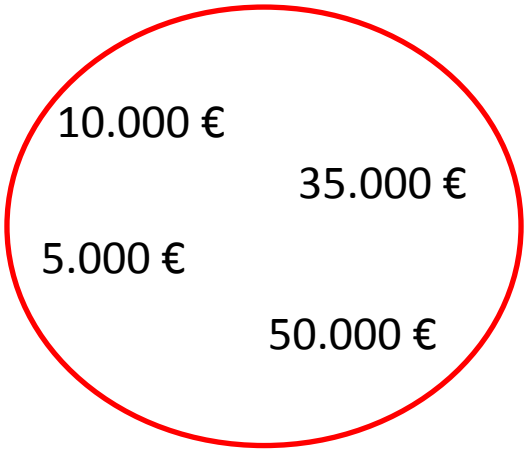
Average (perceived) income of individuals/municipalities etc.



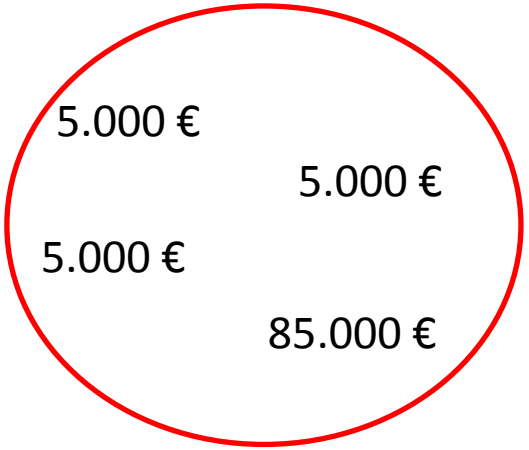
Real income



or



or



Modifiable Areal Unit Problem (1)

"The MAUP is "a problem arising from the imposition of artificial units of spatial reporting on continuous geographical phenomena resulting in the generation of artificial spatial patterns" (Heywood, 1988). In other words, artifacts or errors are created when one groups data into units for analysis. …..However, this grouping may distort or exaggerate the actual data pattern."

***Source**: http://gispopsci.org/maup*

"It has been argued that the MAUP is a fundamental geographical problem that is endemic to all studies of spatially aggregated data. It is a geographical fact of life that the results of spatial study will always depend on the areal units that are being studied."

***Source**:* Openshaw (1983), available at http://www.uio.no/studier/emner/sv/iss/SGO9010/openshaw1983.pdf

Modifiable Areal Unit Problem (2)

"There are two distinct types of MAUP: Scale (or aggregation) and zone (or grouping).

Scale MAUP

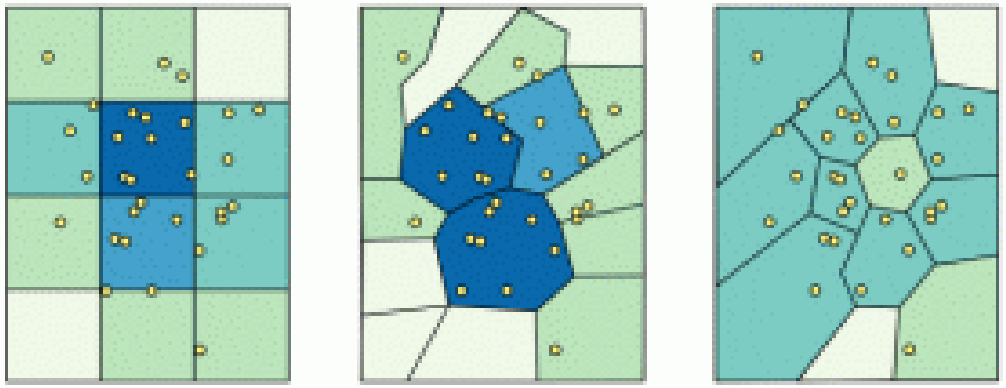
The scale at which one chooses to analyze information, be it for the entire United States, by state, by county, or even block-by-block, can produce different results […] As with all analyses, it is important to choose your scale to match your research question: Investigating the effects of a new county hospital by looking at state-wide mortality rates is not ideal. However, research data often comes in prepackaged sizes. If you are unable to choose the scale of your data, be aware of its implications and, if possible, choose a finer scale than you think is necessary. Finer-scale data can be aggregated, while coarser scale data cannot easily be divided."

Source: http://qispopsci.org/maup

Modifiable Areal Unit Problem (3)

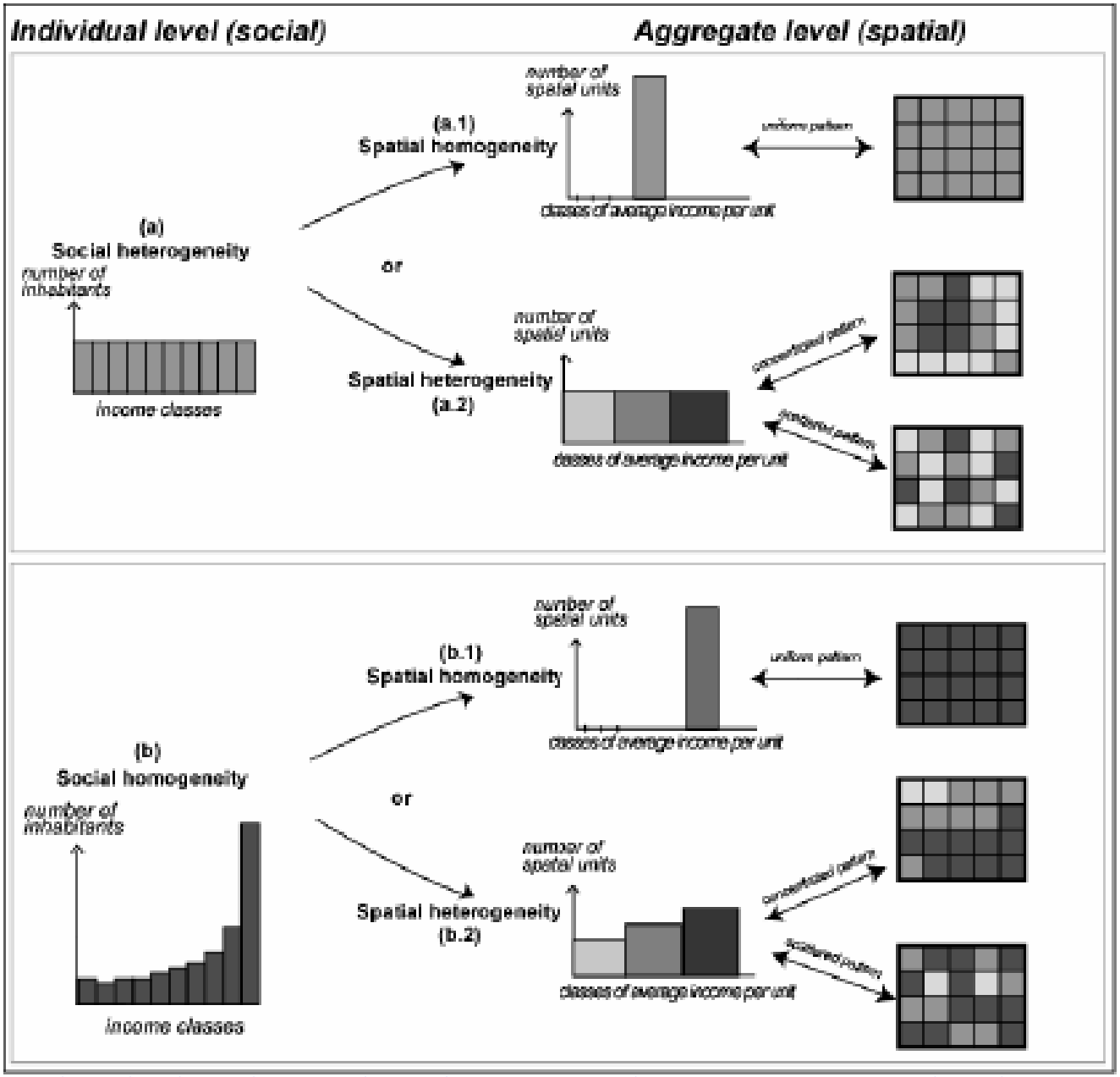
Zone MAUP

The zones or grouping schemes that one uses for data analysis can also be an issue, even if the units are all of the same scale. […] While this effect may accidentally add sources of error or misnterpretation into a research project, it may also be used to intentionally manipulate the results. […] However, you should be aware that any set of rules creates bias; thus, you should compare the results of different schemes. ***Source***: *http://gispopsci.org/maup*



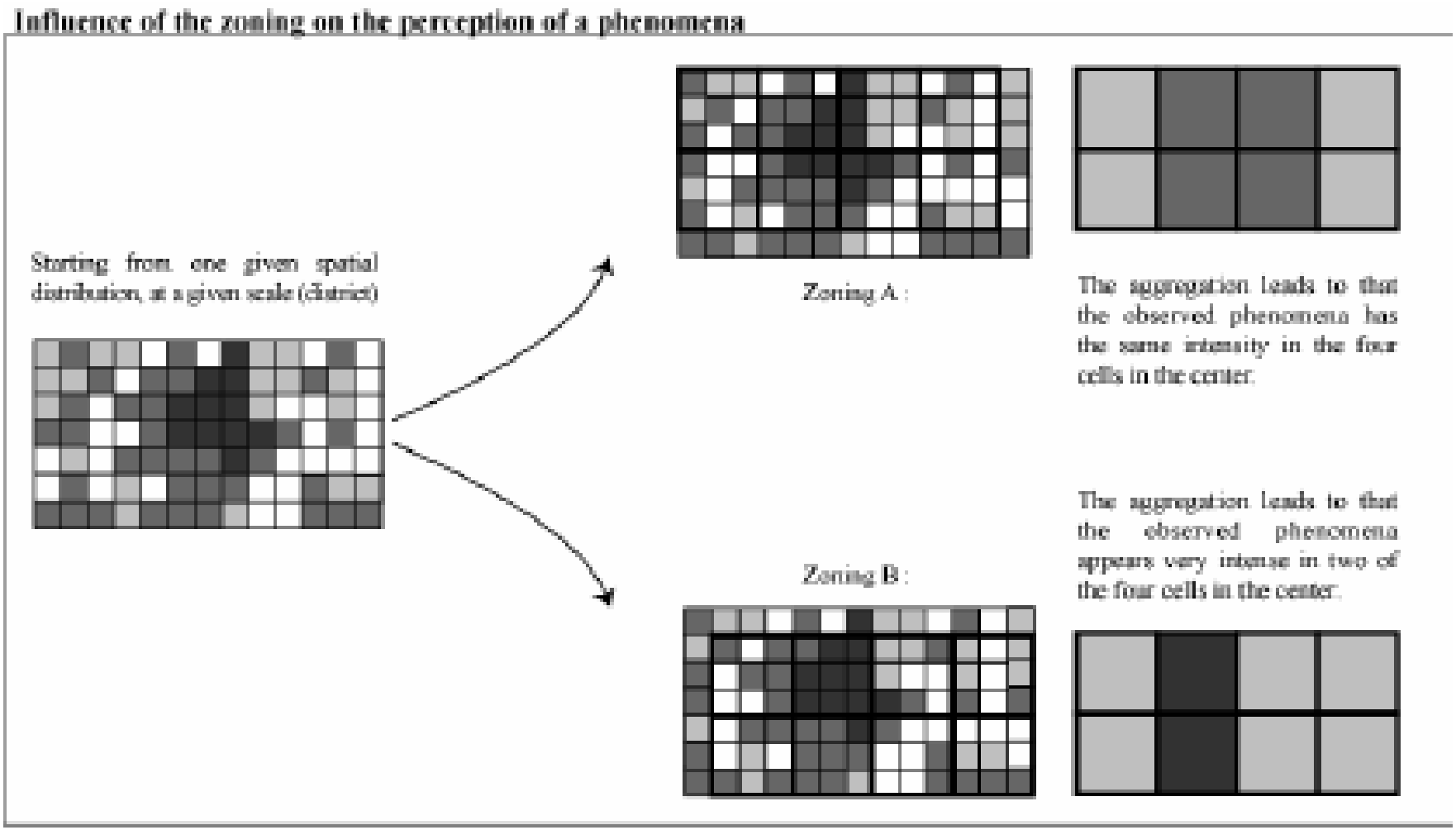
Source: *http://gispopsci.org/maup, image transposed*

Modifiable Areal Unit Problem (4)



Source: http://www.espon.eu/export/sites/default/Documents/Projects/ESPON2006Projects/StudiesScientificSupportProjects/MAUP/espon343_maup_final_version2_nov_2006.pdf

Modifiable Areal Unit Problem (5)



Source: *http://www.espon.eu/export/sites/default/Documents/Projects/ESPON2006Projects/StudiesScientificSupportProjects/MAUP/espon343_maup_final_version2_nov_2006.pdf*

Spatial autocorrelation (1)

The first law of geography according to Waldo Tobler (1970) is "***Everything is related to everything else, but near things are more related than distant things.***"

"Spatial autocorrelation is a complicated name for the obvious fact that data from locations near one another in space are more likely to be similar than data from locations remote from one another."

Source: *O'Sullivan & Unwin (2010) , emphasis added*

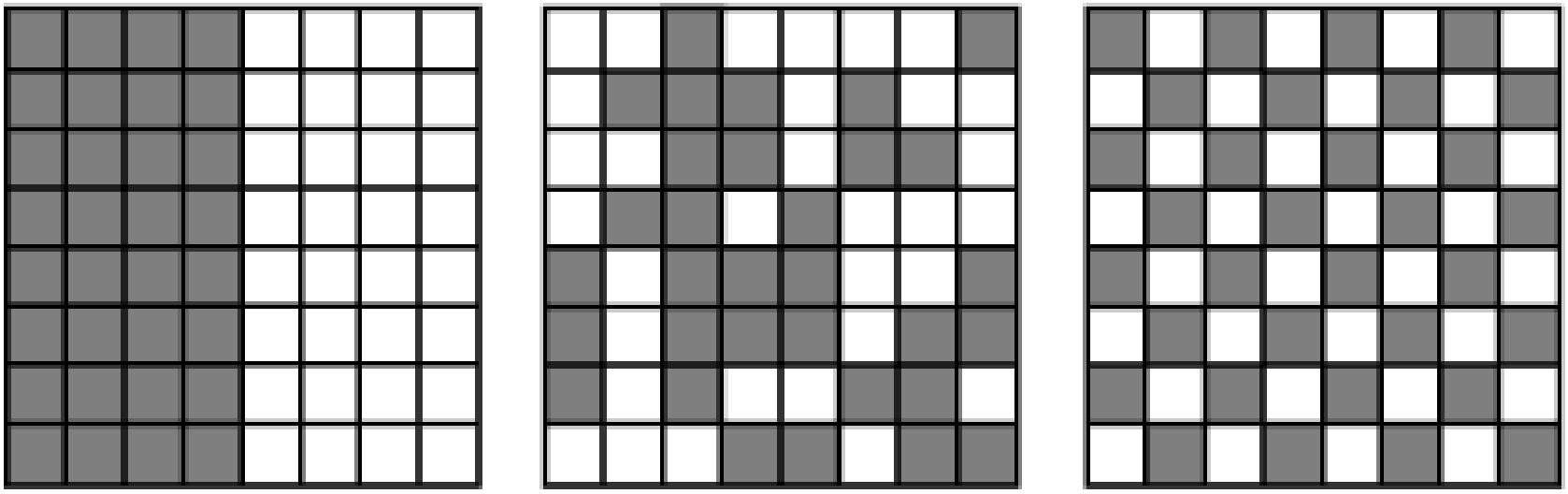
Spatial autocorrelation - spatial dependency

Spatial autocorrelation (2)

Positive: nearby locations are likely to be similar to one another, namely high values tend to be located near high values, medium values near medium values, and low values near low values.

Zero: observations seem to vary randomly through space

Negative: observations from nearby observations are likely to be different from one another.



Source: *http://www.hypergeo.eu/spip.php?article223*

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

End of Session