

Χώρας αειχώρος

ΚΕΙΜΕΝΑ ΠΟΛΕΟΔΟΜΙΑΣ, ΧΩΡΟΤΑΞΙΑΣ ΚΑΙ ΑΝΑΠΤΥΞΗΣ

ΤΟΜΟΣ 3
VOLUME 3

ΤΕΥΧΟΣ 2
ISSUE 2

ΝΟΕΜΒΡΙΟΣ 2004
NOVEMBER 2004



ΣΥΝΤΑΚΤΙΚΗ ΕΠΙΤΡΟΠΗ - Πανεπιστήμιο Θεσσαλίας
*Τμήμα Μηχανικών Χωροταξίας, Πολεοδομίας
και Περιφερειακής Ανάπτυξης*

ΚΟΚΚΩΣΗΣ ΧΑΡΗΣ
ΜΠΕΡΙΑΤΟΣ ΗΛΙΑΣ
ΟΙΚΟΝΟΜΟΥ ΔΗΜΗΤΡΗΣ
ΠΕΤΡΑΚΟΣ ΓΙΩΡΓΟΣ
ΓΟΥΣΙΟΣ ΔΗΜΗΤΡΗΣ
ΔΕΦΝΕΡ ΑΛΕΞΗΣ
ΨΥΧΑΡΗΣ ΓΙΑΝΝΗΣ

ΣΥΜΒΟΥΛΟΙ ΣΥΝΤΑΞΗΣ

Αραβαντινός Αθανάσιος	- ΕΜΠ
Ανδρικόπουλος Ανδρέας	- Οικονομικό Πανεπιστήμιο Αθηνών
Βασενχόβεν Λουδοβίκος	- ΕΜΠ
Γιαννακούρου Τζίνα	- Πανεπιστήμιο Θεσσαλίας
Γιαννιάς Δημήτρης	- Πανεπιστήμιο Θεσσαλίας
Δελλαδέτσιμας Παύλος	- Χαροκόπειο Πανεπιστήμιο
Δεμαθιάς Ζαχαρίας	- Πάντειο Πανεπιστήμιο
Ιωαννίδης Γιάννης	- Tufts University, USA
Καλογήρου Νίκος	- ΑΠΘ
Καρυίδης Δημήτρης	- ΕΜΠ
Κοσμόπουλος Πάνος	- ΔΠΘ
Κουκλέλη Ελένη	- University of California, USA
Λαμπριανίδης Λόης	- Πανεπιστήμιο Μακεδονίας
Λουκάκης Παύλος	- Πάντειο Πανεπιστήμιο
Λουρή Ελένη	- Οικονομικό Πανεπιστήμιο Αθηνών
Μαντουβάλου Μαρία	- ΕΜΠ
Μελαχροινός Κώστας	- University of London, Queen Mary, UK
Μοδινός Μιχάλης	- Εθν. Κέντρο Περιβ. και Δειφ. Ανάπτυξης (ΕΚΠΙΑΑ)
Μπριασούλη Ελένη	- Πανεπιστήμιο Αιγαίου
Παπαθεοδώρου Ανδρέας	- University of Surrey, UK
Παπτάς Βασίλης	- Πανεπιστήμιο Πατρών
Πρεβελάκης Γεώργιος-Στυλ.	- Universite de Paris I, France
Φωτόπουλος Γιώργος	- Πανεπιστήμιο Πατρών
Χαστάογλου Βίλμα	- ΑΠΘ

Διεύθυνση:

Πανεπιστήμιο Θεσσαλίας
Τμήμα Μηχανικών Χωροταξίας, Πολεοδομίας
και Περιφερειακής Ανάπτυξης
Περιοδικό ΔΕΙΧΩΡΟΣ

Πεδίον Άρεως, 38334 ΒΟΛΟΣ

<http://www.prd.uth.gr/aeihoros> e-mail: aeihoros@prd.uth.gr

τηλ.: 24210 – 74456 fax: 24210 – 74380



ΠΑΝΕΠΙΣΤΗΜΙΑΚΕΣ ΕΚΔΟΣΕΙΣ ΘΕΣΣΑΛΙΑΣ

Ειδικό τεύχος - Αφιέρωμα
Γεωπληροφορική

Επιμέλεια

Μαρίνος Κάβουρας
Ελευθερία Καρνάβου

Επιστημονικό Περιοδικό

αειχώρος



Επιμέλεια έκδοσης : Άννα Σαμαρίνα — Παναγιώτης Πανταζής
Λαγούτ : Παναγιώτης Πανταζής
Σχεδιασμός εξωφύλλου : Γιώργος Παρασκευάς
Εκτύπωση : Αλέκος Ξουράφας
Κεντρική διάθεση : Πανεπιστημιακές Εκδόσεις Θεσσαλίας

Περιεχόμενα

Εισαγωγή	4
Κόκλα Μ., Κάβουρας Μ.	8
Προσδιορισμός σημασιολογικών ιδιοτήτων και σχέσεων για την επίλυση οντολογικής ετερογένειας	
Τομαή Ε., Κάβουρας Μ.	24
Απαραίτητα συστατικά για την ανάπτυξη Θεματικών Γεωγραφικών Οντολογιών	
Γραϊκούσης Γ., Φώτης Γ. και Κουτσόπουλος Κ.	40
Χωροχρονική πρόβλεψη σημειακών προτύπων ζήτησης στοχαστικών χωροθετικών προβλημάτων με χρήση Νευρωνικών Δικτύων	
Παντελέλης Μ., Σουλακέλλης Ν.	62
Διαχείριση και ανάλυση χωροχρονικών δεδομένων αέριας ρύπανσης αστικών περιοχών: Η περίπτωση της Αττικής	
Μανέτος Π., Φώτης Γ.	76
Εφαρμογή μεθόδων εξαγωγής χωρικής γνώσης για τον προσδιορισμό και την απόδοση Χωρικών Προτύπων	
Μελιδόνη Μ., Χατζηχρήστος Θ.	94
Σχεδιασμός και ανάπτυξη περιβάλλοντος διεπαφής με το χρηστή για την ανάλυση χωρικών σημειακών προτύπων με την τεχνική "Ανάλυση Καννάβου" σε περιβάλλον ΓΣΠ	
Σαραφίδης Δ., Παρασκάκης Ι.	112
Ένα χρηστικό περιβάλλον επικοινωνίας για την διάθεση των κτηματολογικών δεδομένων στο διαδίκτυο	
Ντόκου Αικ.	134
Χωρικές διαδικασίες τεκμηρίωσης τρισδιάστατης πληροφορίας ιδιοκτησιακών αντικειμένων	
Καμαριανάκης Γ., Κοντός Δ.	154
Ταξινόμηση των δήμων της Ελλάδας σύμφωνα με τα κοινωνικά και οικονομικά χαρακτηριστικά τους όπως προκύπτουν από την Απογραφή του 2001	

Ταξινόμηση των δήμων της Ελλάδας σύμφωνα με τα κοινωνικά και οικονομικά χαρακτηριστικά τους όπως προκύπτουν από την Απογραφή του 2001¹

Γιάννης Καμαριανάκης

Ίδρυμα Τεχνολογίας και Έρευνας και

Τμήμα Οικονομικών Επιστημών, Πανεπιστήμιο Κρήτης

Δημήτρης Κοντός

Ίδρυμα Τεχνολογίας και Έρευνας

Περίληψη

Η μελέτη που ακολουθεί έχει ως στόχο την ανάλυση των κοινωνικοοικονομικών δομών των δήμων και κοινοτήτων της Ελλάδας και την ταξινόμηση τους σε ομάδες με κοινά χαρακτηριστικά. Η ανάλυση γίνεται με τεχνικές πολυμεταβλητής στατιστικής ανάλυσης που επιτρέπουν τη ταξινόμηση-ομαδοποίηση δεδομένων αλλά και τεχνικές που στοχεύουν στη μείωση του πλήθους των δεδομένων με τέτοιο τρόπο ώστε να μη χάνεται η πληροφορία. Σε πρώτο στάδιο η εφαρμογή της τεχνικής των κυρίων συνιστωσών (principal components analysis) οδηγεί στη βέλτιστη μείωση του όγκου των δεδομένων. Αρχίζοντας με 348 μεταβλητές από την Απογραφή Πληθυσμού του 2001, με την εφαρμογή αυτής της τεχνικής προκύπτουν 37 κύριες συνιστώσες που είναι γραμμικοί συνδυασμοί των αρχικών μεταβλητών και διατηρούν το μεγαλύτερο μέρος της πληροφορίας που δείγματος. Στη συνέχεια, οι κύριες συνιστώσες χρησιμοποιούνται για την εφαρμογή τεχνικών ομαδοποίησης (clustering) που κατηγοριοποιούν το αρχικό σύνολο των Καποδιστριακών δήμων σε ένα μικρό αριθμό ομάδων με

¹ Ευχαριστούμε θερμά τους Ν. Λουλάκη και Ε. Κοκοσάλη για την βοήθειά τους στην χαρτογραφική αναπαράσταση των αποτελεσμάτων καθώς και τον Π. Πραστάκο για τα επικοινωνιακά σχόλιά του.

ομογενή χαρακτηριστικά. Για την εξαγωγή ορθότερων συμπερασμάτων συγκρίνονται τα αποτελέσματα διαφόρων παραλλαγών της τεχνικής clustering. Τα αποτελέσματα παρουσιάζονται σε διδιάστατα γραφήματα αλλά και σε χάρτη.

Λέξεις-κλειδιά

Γεωδημογραφία, ανάλυση κυρίων συνιστωσών, ομαδοποίηση, ταξινόμηση.

Classification of greek municipalities according to their socio-economic characteristics as derived from the recent (2001) census

This article aims at exploring the socio-economic structures of Greek municipalities and classifying them into homogeneous groups. The analysis is based on multivariate statistical analysis techniques applied to variables that were measured during the recent (2001) census. First, we perform principal components analysis; starting with 348 variables we get 37 principal components –linear combinations of the original variables- that represent the majority of information contained in the sample. Next, the principal components are used in the application of clustering techniques so that municipalities are classified into a relatively small number of homogeneous groups. We compare results related to various clustering methods; the most informative ones are presented into graphs and maps.

Keywords

Geo-demography, principal components analysis, clustering, classification.

ΕΙΣΑΓΩΓΗ

Η μελέτη που ακολουθεί έχει ως στόχο την ανάλυση των κοινωνικοοικονομικών δομών των δήμων και κοινοτήτων της Ελλάδας και την ταξινόμηση τους σε ομάδες με κοινά χαρακτηριστικά. Η ανάλυση γίνεται με στατιστικές τεχνικές που επιτρέπουν τη ταξινόμηση-ομαδοποίηση δεδομένων αλλά και τεχνικές που στοχεύουν στη μείωση του πλήθους των δεδομένων με τέτοιο τρόπο ώστε να μη χάνεται η πληροφορία. Και οι δύο βασικές τεχνικές που χρησιμοποιήθηκαν (ανάλυση ομαδοποίησης και ανάλυση κυρίων συνιστωσών αντίστοιχα) ανήκουν στο πεδίο της πολυμεταβλητής στατιστικής ανάλυσης και χρησιμοποιούνται ευρύτατα από τις αρχές της δεκαετίας του 1980, οπότε και έγινε πολύ πιο εύκολη η αποθήκευση και ανάλυση τεράστιων όγκων δεδομένων μέσω της προόδου της πληροφορικής, σε γεωδημογραφικά προβλήματα.

Η ανάγκη για την χρησιμοποίηση τεχνικών για την ελαχιστοποίηση του πλήθους των δεδομένων γίνεται εύκολα κατανοητή αν ληφθεί υπ' όψη ότι τα δεδομένα που αναλύθηκαν σε αυτή την μελέτη περιλαμβάνουν 348 μεταβλητές για τα ποιοτικά χαρακτηρι-

στικά του πληθυσμού 862 Καποδιστριακών δήμων με πληθυσμό άνω των 2000 κατοίκων. Η πραγματική πληροφορία που "περιέχουν" αυτές οι μεταβλητές, μπορεί να εκφρασθεί με πολύ λιγότερες, λόγω της ύπαρξης εσωτερικών δομών συσχέτισης. Αυτό επιτυγχάνεται με την ανάλυση κυρίων συνιστωσών (principal components analysis). Στόχος λοιπόν είναι κατ'αρχάς μεν να προσδιοριστούν οι δομές-σχέσεις μεταξύ των μεταβλητών έτσι ώστε να μειωθεί ο όγκος των αριθμητικών πράξεων και να φανούν καθαρά οι εσωτερικές δομές που διέπουν το σύστημα που αναλύεται. Στη συνέχεια, δε να γίνει μία ταξινόμηση των δήμων-κοινοτήτων σε ομοειδείς κατηγορίες (με παραπλήσιο μορφωτικό, οικονομικό κ.λπ. επίπεδο) χρησιμοποιώντας τους παράγοντες που επήλθαν με μετασχηματισμό των αρχικών μεταβλητών μέσω της μεθόδου κυρίων συνιστωσών.

Εφαρμογές παρόμοιου χαρακτήρα είναι συνήθεις εδώ και αρκετά χρόνια στο εξωτερικό. Ενδεικτικά αξίζει να αναφέρουμε αυτές των Imber (1977), Jolliffe κ.α. (1986), Moser και Scott (1961), Prastacos και Hootkins (1984) και Webber και Craig (1978). Σχετικές εργασίες που αφορούν περιοχές του Ηνωμένου Βασιλείου έγιναν από τους Charlton, M., Openshaw, S. και Wymer, C. (1985) και τους Wallace, M. και Denham, C. (1996). Τα συμπεράσματα από μελέτες τέτοιου είδους μπορούν να χρησιμοποιηθούν σε ένα ευρύ φάσμα πεδίων όπως: σχεδιασμός πολιτικής, δειγματοληπτικές μελέτες, προγνώσεις πωλήσεων, προώθηση προϊόντων, επιλογή χώρου εγκατάστασης και ποιοτικών χαρακτηριστικών καταστημάτων κ.ά. Στην Ελλάδα τέτοιου είδους τεχνικές δεν είχαν χρησιμοποιηθεί καθόλου μέχρι πολύ πρόσφατα (βλ. Καμαριανάκης και Πραστάκος, 2001) και κύριος λόγος είναι η έλλειψη στοιχείων. Η Εθνική Στατιστική Υπηρεσία της Ελλάδας (ΕΣΥΕ) δημοσίευσε τα πλήρη αποτελέσματα της απογραφής του 1981 το 1987 και μόνο σε βιβλία, ενώ τα στοιχεία της απογραφής του 1991 σε ολοκληρωμένη μορφή διατέθηκαν ηλεκτρονικά για πρώτη φορά το 1998 (*Αθηνά-91, 1998*). Αλλά και τα στοιχεία που δημοσιεύονται στις απογραφές παρουσιάζουν ελλείψεις που δεν επιτρέπουν την ανάλυση σε επίπεδο μικρότερο από αυτό του δήμου. Αυτό οφείλεται στο γεγονός ότι όλα τα στοιχεία είναι διαθέσιμα μόνο σε επίπεδο περιφερειών, νομών και δήμων/κοινοτήτων. Κοινωνικο-οικονομικά στοιχεία για απογραφικούς τομείς ή οικοδομικά τετράγωνα δεν υπάρχουν αν και το βιοτικό επίπεδο διαφέρει σημαντικά από συνοικία σε συνοικία. Μ' αυτόν τον τρόπο αποκρύπτονται σημαντικές λεπτομέρειες και οι δήμοι τείνουν να φαίνονται ίδιοι.

ΣΤΑΤΙΣΤΙΚΗ ΜΕΘΟΔΟΛΟΓΙΑ: ΑΝΑΛΥΣΗ ΚΥΡΙΩΝ ΣΥΝΙΣΤΩΣΩΝ ΚΑΙ ΤΕΧΝΙΚΕΣ ΟΜΑΔΟΠΟΙΗΣΗΣ

Η μέθοδος κυρίων συνιστωσών²

Ο σκοπός της ανάλυσης κυρίων συνιστωσών είναι ο σχηματισμός ενός μικρού σχετικά αριθμού γραμμικών συνδυασμών ενός συνόλου μεταβλητών (κύριες συνιστώσες), που θα διατηρεί όσο το δυνατόν περισσότερη από την πληροφορία των αρχικών μεταβλητών. Έτσι γίνονται πιο εύκολες και κατανοητές αναλύσεις περιγραφικού τύπου, παλινδρόμησης, ομαδοποίησης κ.α. όταν πρέπει να εφαρμοστούν σε δείγματα που περιέχουν ένα μεγάλο αριθμό αρχικών μεταβλητών. Η ανάλυση κυρίων συνιστωσών δύναται να ειπωθεί και ως προσπάθεια προσέγγισης των γραμμικών σχέσεων σε ένα σύνολο μεταβλητών. Η πρώτη παρουσίαση της μεθόδου έγινε από τον Pearson (1901), ενώ ο Hotelling (1933) την εξέλιξε σε σημαντικό βαθμό. Ο ενδιαφερόμενος αναγνώστης μπορεί να βρει λεπτομερειακή περιγραφή της ανάλυσης κυρίων συνιστωσών στα βιβλία των Jolliffe (1986) και Mardia, Kent και Bibby (1979).

Οι κύριες συνιστώσες είναι διανύσματα που σχηματίζονται σαν γραμμικοί συνδυασμοί των μεταβλητών του συνόλου των δεδομένων και κατασκευάζονται έτσι ώστε να είναι κάθετα μεταξύ τους (στον πολυδιάστατο χώρο που ορίζουν ανάλογα με το πλήθος τους) και να αντιπροσωπεύουν κατά φθίνουσα τάξη ποσοστά της αρχικής μεταβλητότητας των δεδομένων. Το πλήθος τους είναι ίσο με το πλήθος των αρχικών μεταβλητών: όμως, από κάποιο σημείο και μετά πολλά από αυτά δεν είναι χρήσιμα γιατί δεν επεξηγούν κάποιο ποσοστό από την αρχική μεταβλητότητα. Αυτό συμβαίνει στην περίπτωση που οι μεταβλητές είναι έντονα συσχετισμένες μεταξύ τους.

Τα αποτελέσματα της τεχνικής των κυρίων συνιστωσών προκύπτουν μετά από ανάλυση του πίνακα συνδιακύμανσης ή του πίνακα συσχέτισης των δεδομένων. Με αυτόν τον τρόπο δεν χρησιμοποιούνται τα ίδια τα δεδομένα για την εξαγωγή συμπερασμάτων αλλά η "εσωτερική δομή συσχέτισής" τους. Τις περισσότερες φορές χρησιμοποιείται ο πίνακας συσχέτισης στον οποίο οι μεταβλητές με μεγάλη διασπορά δεν έχουν βαρύτητα μεγαλύτερη

² Η ανάλυση κυρίων συνιστωσών συχνά συγχέεται με την παραγοντική ανάλυση (factor analysis). Βασικοί στόχοι της παραγοντικής ανάλυσης είναι: α) η δημιουργία νέων μεταβλητών, των παραγόντων, στις οποίες μπορούμε με ένα υποκειμενικό τρόπο να αναγνωρίσουμε κάποιες όχι εύκολα μετρήσιμες μεταβλητές, και β) η επεξήγηση των συσχετίσεων που υπάρχουν στα δεδομένα για τις οποίες έχουμε υποθέσει ότι οφείλονται αποκλειστικά στην ύπαρξη κάποιων κοινών παραγόντων που δημιούργησαν τα δεδομένα. Σε αυτή την εργασία το βασικό ζητούμενο είναι η μείωση του πλήθους των αρχικών μεταβλητών ώστε να είναι εφικτή η εφαρμογή των μεθόδων ομαδοποίησης γι'αυτό και χρησιμοποιήθηκε η ανάλυση κυρίων συνιστωσών.

από τις υπόλοιπες. Αυτό συμβαίνει γιατί η συσχέτιση είναι ουσιαστικά ένα τυποποιημένο μέτρο συνδιακύμανσης. Έτσι, η επιλογή των μονάδων μέτρησης των μεταβλητών δεν παίζει κανένα ρόλο όταν τα δεδομένα δεν περιέχουν ομοειδείς μεταβλητές. Όταν όλες οι μεταβλητές που περιέχονται στα δεδομένα εκφράζονται με την ίδια μονάδα μέτρησης, τότε συνήθως χρησιμοποιείται ο πίνακας συνδιακύμανσης. Όσον αφορά την επιλογή του πίνακα συνδιακύμανσης ή του πίνακα συσχέτισης για την ανάλυση του προβλήματος, η δεύτερη περίπτωση αντιστοιχεί στην παραδοχή ότι οι μεταβλητές είναι της ίδιας βαρύτητας. Αντίθετα, η χρησιμοποίηση του πίνακα συνδιακύμανσης οδηγεί στη μεροληπτική σε σχέση και με το μέγεθος της διασποράς της κάθε μεταβλητής (πέρα από τα μεγέθη συνδιακύμανσης) επιλογή του βάρους που έχει κάθε μεταβλητή στον σχηματισμό της κάθε κύριας συνιστώσας.

Τεχνικές ομαδοποίησης

Οι τεχνικές ταξινόμησης-ομαδοποίησης (clustering) έχουν ως σκοπό το διαχωρισμό ενός συνόλου υποκειμένων-αντικειμένων σε ομάδες. Υπό ιδανικές συνθήκες και όταν οι υπό μελέτη μεταβλητές είναι λίγες, για την ταξινόμηση των υποκειμένων-αντικειμένων αρκεί η οπτική εξέταση ενός γραφήματος. Στην πράξη όμως τα πράγματα δεν είναι τόσο απλά και για αυτό το λόγο τα τελευταία 25 χρόνια εξελίχθηκε μεγάλο πλήθος τέτοιων τεχνικών. Σήμερα, έχουν επικρατήσει τρεις διαφορετικές μέθοδοι εφαρμογής τεχνικών ομαδοποίησης.

- α) **Ιεραρχικές τεχνικές:** στις οποίες η ταξινόμηση είναι πολυβηματική και μπορεί να γίνει είτε 'συσσωρευτικά', δηλαδή αρχικά το κάθε υποκειμένο-αντικείμενο αποτελεί από μόνο του μια ομάδα, σε δεύτερο στάδιο ενώνονται οι δύο πλησιέστερες ομάδες, στη συνέχεια οι επόμενες δύο κ.ο.κ., είτε 'διαιρετικά' όπου ένα σύνολο παρατηρήσεων διαμερίζεται σε ομοιογενή υποσύνολα. Το πιο βασικό στοιχείο διαφοροποίησης των ιεραρχικών τεχνικών μεταξύ τους είναι ο ορισμός της απόστασης μεταξύ των ομάδων. Π.χ. η μέθοδος single linkage θεωρεί ως απόσταση μεταξύ δύο ομάδων την απόσταση των πλησιέστερων στοιχείων τους, ενώ η μέθοδος group average χρησιμοποιεί το μέσο όρο των αποστάσεων.
- β) **Βελτιστοποιητικές τεχνικές:** στις οποίες στόχος είναι η μεγιστοποίηση κριτηρίων (προκαθορισμένα μέτρα) ταξινόμησης. Κύρια διαφορά τους σε σχέση με τις προηγούμενες είναι το ότι επιτρέπεται η επανατοποθέτηση των υπό ομαδοποίηση μεταβλητών.
- γ) **Τεχνικές πυκνότητας:** οι οποίες αναζητούν περιοχές μεγάλης πυκνότητας στο μετρικό χώρο που ορίζουν οι μεταβλητές υπό ομαδοποίηση.

Όσον αφορά την επιλογή του 'ιδανικού' αριθμού ομάδων έχουν επικρατήσει ως κριτήρια οι στατιστικές ψευδό-F, ψευδό-t₂ και το Cubic Clustering Criterion. Για εμβάθυνση στις τεχνικές ομαδοποίησης κλασικά θεωρούνται τα βιβλία των Anderberg (1973), Kaufman και Rousseeuw (1990) και Everitt (1993). Για μια πιο πρόσφατη προσέγγιση ο ενδιαφερόμενος αναγνώστης μπορεί να ανατρέξει στο βιβλίο των Everitt, Landau και Leese (2001).

ΤΑ ΔΕΔΟΜΕΝΑ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ

Τα δεδομένα της εφαρμογής προέρχονται από την γενική απογραφή του πληθυσμού η οποία έλαβε χώρα το έτος 2001 και διατέθηκαν σε ηλεκτρονική μορφή από την Εθνική Στατιστική Υπηρεσία. Τα στοιχεία αυτά είναι επίσης διαθέσιμα από την ηλεκτρονική βάση δεδομένων ΑΘΗΝΑ 2001 (Πραστάκος και Κογχυλάκης, 2003). Συγκεκριμένα οι αρχικές μεταβλητές που χρησιμοποιήθηκαν είναι 348 τον αριθμό και αφορούν κυρίως ποιοτικά χαρακτηριστικά για τους 862 δήμους και κοινότητες που επιλέχθηκαν. Η ανάλυση αφορά δήμους και κοινότητες με πληθυσμό πάνω από 2000 κατοίκους. Οι Κατοδιστριακοί δήμοι και οι κοινότητες με πληθυσμό μικρότερο των 2000 κατοίκων αποκλείστηκαν από την εφαρμογή, μιας και το πλήθος τους σε συνδυασμό με την περιορισμένη σχετικά πληροφορία που φέρουν τα δεδομένα της απογραφής του 2001 (παλαιότερες απογραφές περιείχαν και περισσότερα στοιχεία οικονομικού χαρακτήρα όπως π.χ. θέση στο επάγγελμα) θα έκανε αδύνατη την επιτυχή εφαρμογή των τεχνικών ομαδοποίησης. Επίσης απορρίφθηκαν μεταβλητές που παρουσίαζαν μηδενική διασπορά και επομένως δεν μας πρόσφεραν καμία ουσιαστική πληροφορία, ενώ σε άλλες περιπτώσεις πολλές από αυτές ενώθηκαν (π.χ. ηλικίες 65 και άνω). Όλες οι μεταβλητές εκφράζονται σε ποσοστά και όχι με απόλυτα μεγέθη για να αποφεύγονται προβλήματα που θα ανέκυπταν καθώς ο πληθυσμός διαφέρει σημαντικά στους διάφορους δήμους, επιπλέον δε με αυτόν τον τρόπο λαμβάνονται υπ' όψη για την αξιολόγηση κάθε δήμου μόνο ποιοτικά χαρακτηριστικά και όχι το μέγεθός του.

Τα διαθέσιμα δεδομένα είναι κυρίως κοινωνικού χαρακτήρα. Δυστυχώς ο οικονομικός παράγοντας εμφανίζεται έμμεσα και αυτό δυσκολεύει την εφαρμογή τεχνικών ομαδοποίησης καθώς οι διαφορές μεταξύ δήμων-κοινοτήτων είναι λιγότερο ευδιάκριτες. Οι μεταβλητές μπορούν να χωρισθούν νοηματικά στις εξής 6 ομάδες:

1. Ποσοστά επί του πληθυσμού κάθε ηλικιακής ομάδας (για κάθε δήμο-κοινότητα), του πληθυσμού κατά φύλο και ομάδες ηλικιών: π.χ. Θήλεις 65 ετών και άνω/Σύνολο πληθυσμού ηλικίας 65 ετών και άνω.
2. Ποσοστά επί του πληθυσμού κάθε ηλικιακής ομάδας, του πληθυσμού κατά φύλο, ομάδες ηλικιών και οικογενειακή κατάσταση: π.χ. διαζευγμένοι άρρενες ηλικίας 55-59 ετών/Σύνολο πληθυσμού ηλικίας 55-59 ετών.

3. Ποσοστά επί του πληθυσμού κάθε ηλικιακής ομάδας, του πληθυσμού ηλικίας 10 ετών και άνω ανά φύλο, ομάδες ηλικιών και επίπεδο εκπαιδευσεως: π.χ. Θήλεις ηλικίας 40-44 ετών με μεταπτυχιακό/Σύνολο πληθυσμού ηλικίας 40-44 ετών.
4. Ποσοστά επί του πληθυσμού κάθε ηλικιακής ομάδας, του οικονομικώς ενεργού πληθυσμού κατά φύλο και ομάδες ηλικιών: π.χ. Οικονομικώς ενεργοί εργαζόμενοι άρρενες ηλικίας 15-19 ετών/Σύνολο πληθυσμού ηλικίας 15-19 ετών.
5. Ποσοστά επί του συνόλου των οικονομικά ενεργών, του οικονομικώς ενεργού πληθυσμού κατά φύλο, ομάδες ηλικιών και ομάδες ατομικών επαγγελμάτων: π.χ. Οικονομικώς ενεργοί θήλεις ηλικίας 60-64 ετών απασχολούμενοι στην παροχή υπηρεσιών/Σύνολο οικονομικά ενεργών.
6. Ποσοστά επί του συνόλου των νοικοκυριών κάθε νοικοκυριού κατά αριθμό δωματίων και μελών του νοικοκυριού: π.χ. Νοικοκυριά με 3 δωμάτια και 2 μέλη/Σύνολο νοικοκυριών

Σε σχέση με τα δεδομένα που χρησιμοποιήθηκαν στην προηγούμενη παρόμοια μελέτη που έγινε για την Ελλάδα με δεδομένα της απογραφής του 1991 (Καμαριανάκης και Πραστάκος, 2001), εμφανίζεται μια νέα κατηγορία μεταβλητών, η έκτη. Δυστυχώς όμως από την πρόσφατη απογραφή απουσιάζουν οι δύο παρακάτω κατηγορίες μεταβλητών:

Ποσοστά επί του συνόλου των οικονομικά ενεργών, του οικονομικά ενεργού πληθυσμού κατά φύλο, ομάδες κλάδων οικονομικής δραστηριότητας και θέση στο επάγγελμα: π.χ. Οικονομικώς ενεργοί άρρενες σε ξενοδοχεία και εστιατόρια, εργοδότες/Σύνολο οικονομικά ενεργών.

Ποσοστά επί του συνόλου των οικονομικά ενεργών, του οικονομικά ενεργού πληθυσμού κατά φύλο, ομάδες ατομικών επαγγελμάτων και θέση στο επάγγελμα: π.χ. Οικονομικώς ενεργοί θήλεις ασκούντες επιστημονικά και ελευθέρια επαγγέλματα, μισθωτοί/Σύνολο οικονομικά ενεργών.

Σε σχέση με τη μελέτη που βασίστηκε στα δεδομένα της απογραφής του 1991 το πλήθος των μεταβλητών που χαρακτηρίζουν κάθε δήμο είναι αισθητά μικρότερο. Στο άρθρο των Καμαριανάκη και Πραστάκου (2001) το σύνολο των διαθέσιμων μεταβλητών έφτανε τις 1076 ενώ στην παρούσα μελέτη το πλήθος τους δεν ξεπερνά τις 470³. Επιπλέον, ο αριθμός των "Καποδιστριακών" δήμων με πληθυσμό μεγαλύτερο των 2000 κατοίκων φτάνει τους 862, 301 περισσότεροι σε σχέση με τον αριθμό των δήμων που χρησιμοποι-

³ Τελικά χρησιμοποιήθηκαν 348 μεταβλητές που προέκυψαν μετά από διαγραφή αυτών με μηδενική διασπορά (άρα και πληροφορία) και την ένωση μεταβλητών που αφορούν τα μεγάλα ηλικιακά στρώματα.

ήθηκαν στην προηγούμενη μελέτη⁴. Ο μειωμένος αριθμός των επεξηγηματικών μεταβλητών σε συνδυασμό με την αύξηση στον αριθμό των δήμων δυσχεραίνει την εξαγωγή ουσιαστικών αποτελεσμάτων από τις μεθόδους ομαδοποίησης.

ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

Οι κύριες συνιστώσες των μεταβλητών της απογραφής του 2001

Αρχικά η ανάλυση κυρίων συνιστωσών πραγματοποιήθηκε με την χρήση του πίνακα συσχέτισεων. Σε αυτήν την περίπτωση ο αριθμός των κυρίων συνιστωσών που χρειάζονται για να επεξηγήσουν το 70% της μεταβλητότητας των αρχικών δεδομένων είναι 37, ενώ αντίστοιχα για την ερμηνεία του 90% της μεταβλητότητας απαιτούνται 114 συνιστώσες. Από την άλλη η χρησιμοποίηση του πίνακα συνδιασποράς για την εξαγωγή των συνιστωσών έχει σαν αποτέλεσμα να χρειάζονται μόνο 8 για την ερμηνεία του 70% της αρχικής μεταβλητότητας και 29 για το 90% αντίστοιχα. Αυτό συμβαίνει επειδή στην ανάλυση του πίνακα συσχέτισης όλες οι μεταβλητές ανεξάρτητα από το μέγεθος της διασποράς τους έχουν το ίδιο βάρος στο σχηματισμό των κυρίων συνιστωσών. Στην συνέχεια παραθέτουμε δύο γραφήματα στα οποία απεικονίζονται οι σχετικές θέσεις των δήμων-κοινοτήτων με βάση τις 2 πρώτες κύριες συνιστώσες τα οποία προέκυψαν από τον πίνακα συσχέτισης και τον πίνακα συνδιασποράς αντίστοιχα.

Όπως αναφέρθηκε στη δεύτερη ενότητα του άρθρου, οι κύριες συνιστώσες αποτελούν γραμμικούς συνδυασμούς των αρχικών μεταβλητών. Κάθε συνιστώσα διαφοροποιείται από τις υπόλοιπες ανάλογα με τα βάρη των αρχικών μεταβλητών που την χαρακτηρίζουν. Με τον τρόπο αυτό η μέθοδος μας επιτρέπει να τις αναγνωρίσουμε δίνοντας ονόματα στις καινούργιες μεταβλητές (συνιστώσες) απλώς παρατηρώντας ποιες από τις αρχικές μεταβλητές παρουσιάζουν μεγαλύτερη επίδραση πάνω σε αυτές. Με τον τρόπο αυτό μπορούμε να ποσοτικοποιήσουμε μη μετρήσιμες ποσότητες που το μέγεθος τους διαφοροποιείται στους διάφορους δήμους.

Χαρακτηριστικά αναφέρουμε την περίπτωση των τριών πρώτων συνιστωσών όπως αυτές προέκυψαν από τον πίνακα συσχέτισης:

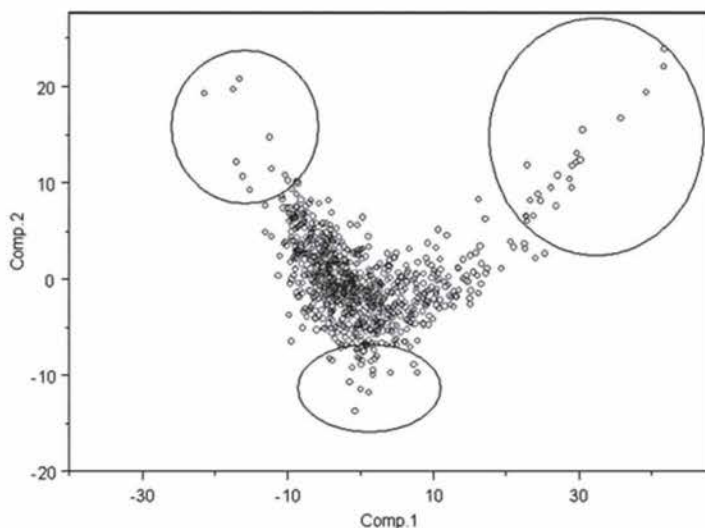
- Η πρώτη συνιστώσα χαρακτηρίζεται από την παρουσία μεταβλητών που έχουν να κάνουν με την ύπαρξη ατόμων από όλα τα ηλικιακά στρώματα τα οποία διαθέτουν ανώτερη και ανώτατη μόρφωση. Οι μεταβλητές αυτές έχουν πολύ μεγάλα θετικά βάρη και δημιουργούν αντίθεση μέσα στην πρώτη αυτή

⁴ Στο άρθρο των Καμαριανίκη και Πραστάκου (2001) χρησιμοποιήθηκαν οι δήμοι με πληθυσμό μεγαλύτερο των 2000 κατοίκων σύμφωνα με τον παλαιότερο (προ Καποδίστρια) ορισμό τους.

συνιστώσα, με μεταβλητές που χαρακτηρίζονται από μεγάλα αρνητικά βάρη και δηλώνουν χαμηλό ή στοιχειώδες επίπεδο εκπαίδευσης.

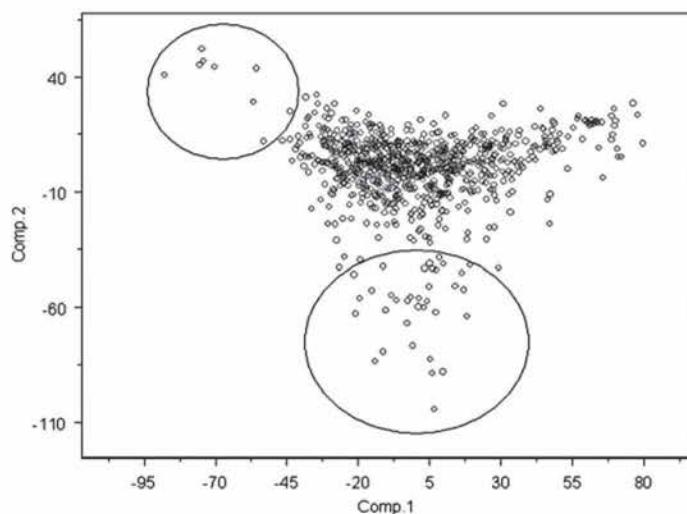
- Στη δεύτερη κύρια συνιστώσα είναι εμφανής η παρουσία με θετικά βάρη μεταβλητών που αφορούν ομάδες οικονομικά ενεργών αρρένων και θηλέων ηλικίας 30-64 ετών οι οποίες παρουσιάζουν αντίθεση με μεταβλητές που αφορούν αντίστοιχες κατηγορίες ατόμων οικονομικά μη ενεργών και χαρακτηρίζονται από αρνητικά βάρη.
- Στην τρίτη κύρια συνιστώσα προβάλλεται η ύπαρξη μεταβλητών που αφορούν το μέγεθος των νοικοκυριών και των αριθμό των μελών που τα αποτελούν . Επίσης είναι εμφανής η αντίθεση ανάμεσα σε έγγαμους και άγαμους άρρενες ηλικίας 20-30 ετών, όπως επίσης και αντίθεση ανάμεσα σε έγγαμες και άγαμες θήλυνς ηλικίας 10-20 ετών.

Σχήμα 1: Διοδιάστατη απεικόνιση των δήμων-κοινοτήτων με βάση τις δύο πρώτες κύριες συνιστώσες όπως προέκυψαν από τον πίνακα συσχέτισης



τονίζονται τα στοιχεία των πιο ακραίων ομάδων που προέκυψαν με χρήση της μεθόδου ομαδοποίησης του Ward

Σχήμα 2: Διοδιάστατη απεικόνιση των δήμων-κοινοτήτων με βάση τις δύο πρώτες κύριες συνιστώσες όπως προέκυψαν από τον πίνακα συνδιακύμανσης

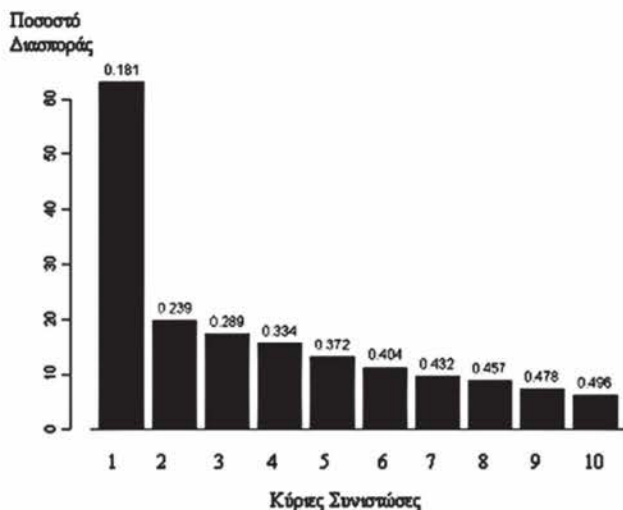


τονίζονται τα στοιχεία των πιο ακραίων ομάδων που προέκυψαν με χρήση της μεθόδου ομαδοποίησης του Ward

Η επιλογή του αριθμού των συνιστωσών που επαρκούν για την ερμηνεία της μεταβλητότητας των αρχικών δεδομένων βασίζεται σε εμπειρικά κριτήρια. Ο αριθμός αυτός είναι δυνατόν να επιλεγεί μέσω ενός γραφήματος στο οποίο απεικονίζεται το ποσοστό της διασποράς κάθε συνιστώσας ως προς τον αριθμό των συνιστωσών.

Σύμφωνα με τον Cattell (1966) ενδείκνυται να επιλεγεί ο αριθμός εκείνος των συνιστωσών στον οποίο εμφανίζεται μεγάλη εναλλαγή στην κλίση του γραφήματος. Όπως φαίνεται και στο σχήμα 3 κάτι τέτοιο μεταφράζεται στην επιλογή δύο μόνο συνιστωσών γεγονός που δεν είναι αποδεκτό μιας και η ομαδοποίηση θα βασιζόταν μόνο στο 40% της μεταβλητότητας των αρχικών δεδομένων. Παρόμοιοι γραφικοί τρόποι αναζήτησης του βέλτιστου αριθμού των κυρίων συνιστωσών που βασίζονται στις ιδιοτιμές του πίνακα συσχέτισης οδηγούν επίσης σε μη αποδεκτά αποτελέσματα. Στην συγκεκριμένη περίπτωση επιλέχθηκε ένας αριθμός 37 principal components που καλύπτει ένα προκαθορισμένο και κοινώς αποδεκτό ποσοστό της αρχικής μεταβλητότητας της τάξης του 70%.

Σχήμα 3: Ποσοστό μεταβλητότητας που αντιστοιχεί σε κάθε κύρια συνιστώσα με βάση τον πίνακα συσχέτισης



Ομαδοποίηση των Κατοδιαστριακών δήμων

Για την ομαδοποίηση των δήμων χρησιμοποιήθηκαν ιεραρχικές τεχνικές ταξινόμησης⁵. Στις συγκεκριμένες τεχνικές η διαδικασία ομαδοποίησης είναι πολυβηματική και εξελίσσεται είτε προσθετικά, είτε διαιρετικά. Στη περίπτωση της προσθετικής διαδικασίας θεωρούμε αρχικά το κάθε αντικείμενο ότι αποτελεί μόνο του μια αυτόνομη ομάδα. Στην συνέχεια με την χρησιμοποίηση του κατάλληλου κριτηρίου ενώνονται τα αντικείμενα εκείνα τα οποία υποδεικνύονται από το κριτήριο και σχηματίζουν μια καινούργια ομάδα που αποτελείται από επιπλέον μέλη. Στο επόμενο βήμα επαναλαμβάνεται η διαδικασία για τις καινούργιες ομάδες, έως ότου καταλήξουμε τελικά μετά την πάροδο αρκετών βημάτων σε μια τελική ομάδα που να περιέχει όλα τα αντικείμενα. Στην περίπτωση της διαιρετικής διαδικασίας αντίθετα θεωρούμε σε πρώτη φάση ότι τα αντικείμενα ανήκουν στην ίδια ομάδα η οποία στα επόμενα βήματα διαμερίζεται σε ομοιογενή σύνολα. Η συγκεκριμένη μέθοδος έχει το πλεονέκτημα ότι λόγω του πολυβηματικού χαρακτήρα με τον οποίο εξελίσσεται μας επιτρέπει να παρατηρούμε άμεσα τις μεταβολές που υφίστανται οι ομάδες κατά την εξέλιξη της διαδικασίας στα διάφορα στάδια και να επεμβαίνουμε δυναμικά ερμηνεύοντας τα αίτια της συγχώνευσης ή διάσπασης των εκάστοτε ομάδων.

⁵ Στο άρθρο των Καμαριανάκη και Πραστάκου (2001) χρησιμοποιήθηκαν και τεχνικές πυκνότητας χωρίς όμως να δώσουν ικανοποιητικά αποτελέσματα.

Στην παρούσα εφαρμογή χρησιμοποιήθηκαν τα 37 πρώτα principal components (70% της αρχικής μεταβλητότητας) τα οποία προέκυψαν από τον πίνακα συσχέτισης ενώ οι ιεραρχικές μέθοδοι που εφαρμόστηκαν ήταν οι: average linkage, complete linkage, single linkage, ward linkage, weighted linkage, με την ward linkage⁶ να δίνει και τα πιο ικανοποιητικά αποτελέσματα τα οποία παρουσιάζονται παρακάτω. Επίσης χρησιμοποιήθηκε και το σύνολο των κυρίων συνιστωσών του πίνακα συσχέτισης που ελεξηγούν το 90% της αρχικής μεταβλητότητας των δεδομένων, καθώς και οι κύριες συνιστώσες που προέκυψαν από την ανάλυση του πίνακα συνδιασποράς και ελεξηγούσαν, είτε το 70%, είτε το 90% της αρχικής μεταβλητότητας. Σε όλες όμως αυτές τις περιπτώσεις τα αποτελέσματα δεν ήταν ικανοποιητικά καθώς δεν υπήρχε κάποια λογική εξήγηση στον τρόπο με τον οποίο σχηματιζόνταν οι ομάδες.

Ως εκ τούτου επιλέχθηκε η ομαδοποίηση που προέκυψε από την μέθοδο ward linkage χρησιμοποιώντας τις 37 πρώτες κύριες συνιστώσες. του πίνακα συσχέτισης και υπέδειξε 14 ομάδες. Όπως φαίνεται στα σχήματα 4 και 5 και στην ανάλυση των αποτελεσμάτων που ακολουθεί οι ομάδες παρουσιάζουν ικανοποιητική ομοιογένεια ως προς τα χαρακτηριστικά τους. Το γεγονός αυτό οφείλεται εν μέρει στο χαρακτηριστικό το οποίο παρουσιάζει η μέθοδος ward να δημιουργεί ομάδες με μικρό σχετικά αριθμό στοιχείων και σε αρκετές περιπτώσεις ισομεγέθη.

Οι πιο βασικές από τις δεκατέσσερις ομάδες και τα γενικά χαρακτηριστικά τους παρουσιάζονται συνοπτικά παρακάτω. Ο ενδιαφερόμενος αναγνώστης μπορεί να βρει τα πλήρη αποτελέσματα στην ιστοσελίδα <http://www.iacm.forth.gr/regional/people/kamarianakis.html>.

Ομάδα 1: Αστικές περιοχές με υψηλό βιοτικό επίπεδο

Στην ομάδα αυτή ανήκουν αστικοί δήμοι κυρίως της Αθήνας και της Θεσσαλονίκης που παρουσιάζουν υψηλό βιοτικό επίπεδο. Χαρακτηριστικά παραδείγματα που ανήκουν σε αυτή την ομάδα είναι οι δήμοι Κηφισιάς, Εκάλης, Ψυχικού, Αγίας Παρασκευής, Χαλανδρίου για την περιοχή της Αθήνας και ο δήμος Πανοράματος για την περιοχή της Θεσσαλονίκης. Εδώ εμφανίζονται πολύ υψηλά ποσοστά σε πτυχιούχους ανωτάτων σχολών με τον μέσο όρο να είναι σχεδόν τριπλάσιος σε σχέση με τον μέσο όρο των υπόλοιπων δήμων για τις νέες ηλικίες 20-30 ετών ενώ αυξάνεται στις μεγαλύτερες ηλικίες φτάνοντας ακόμη και οχταπλάσιο σε ηλικίες πάνω από 65 ετών. Επίσης τα ποσοστά κατόχων μεταπτυχιακών και διδακτορικών διπλωμάτων εμφανίζονται να είναι τα υψηλότερα σε όλη την χώρα για τις νέες ηλικίες ενώ τα υψηλά αυτά ποσοστά διατηρούνται και στις μεγάλες ηλικίες ακόμα και για άτομα που βρίσκονται σε συντάξιμη ηλικία.

⁶ Η βασική διαφορά των ιεραρχικών μεθόδων ομαδοποίησης που εφαρμόστηκαν βρίσκεται στη μετρική της απόστασης μεταξύ δύο ομάδων.

Ομάδα 2: Πρωτεύουσες νομών

Αποτελείται από 91 πρωτεύουσες (κυρίως) νομών, μεγάλων νησιών και κάποιους δήμους Αθηνών και Θεσσαλονίκης. Χαρακτηριστικοί αντιπρόσωποι της ομάδας είναι τα Χανιά, η Λάρισα, η Δράμα, η Καλαμάτα, το Ηράκλειο και η Πάτρα. Στους δήμους αυτής της ομάδας το ποσοστό των αγράμματων είναι ελαττωμένο στο μισό σε σχέση με όλους τους υπόλοιπους δήμους, το ποσοστό πτυχιούχων είναι αυξημένο κατά το ένα τέταρτο, ενώ σημαντικά μικρότερα είναι τα ποσοστά των οικονομικά ενεργών νέων και των απασχολούμενων μεγάλης ηλικίας.

Ομάδα 5: Παραθαλάσσια θέρετρα

Περιέχει με μεγάλο ποσοστό παραθαλάσσια θέρετρα και παριστάνεται στο σχήμα 6 με μωβ χρώμα. Έχει 45 δήμους και μερικά από αυτά είναι η Κως, η Νάξος, η Μύκονος, η Σκιάθος, ο δήμος Ρόδου, η Πάρος και η Χερσόνησος Ηρακλείου. Στους δήμους αυτής της ομάδας το ποσοστό αυτών που εργάζονται σε ξενοδοχεία και εστιατόρια είναι αυξημένο κατά 50% σε σχέση με τον υπόλοιπο μέσο όρο ενώ αντίθετα το ποσοστό αυτών που ασχολούνται με τη γεωργία και την κτηνοτροφία είναι ελαττωμένο κατά το ένα τρίτο.

Ομάδα 10: Αστικές περιοχές με σχετικά χαμηλό βιοτικό επίπεδο

Σε αυτή την ομάδα ανήκουν πολλοί αστικοί δήμοι με χαμηλό σχετικά βιοτικό επίπεδο. Χαρακτηριστικά παραδείγματα δήμων σε αυτή την ομάδα είναι το Περιστερί, Άνω Λιόσια, Αιγάλεω, Ταύρος, Κερατσίνι, Πέραμα, Άγιος Δημήτριος του Πολεοδομικού Συγκροτήματος Αθηνών-Πειραιά, το Κορδελιό, ο δήμος Συκεών, η Νεάπολη στη Θεσσαλονίκη και η Νέα Αλικαρνασός στο πολεοδομικό συγκρότημα Ηρακλείου Κρήτης. Στην ομάδα ανήκουν 39 δήμοι. Οι μέσοι όροι των μετρήσεων της ομάδας 1 σε σχέση με όλους τους υπόλοιπους δήμους-κοινότητες παρουσιάζουν σημαντικά χαμηλότερο επίπεδο εκπαίδευσης ειδικά στους νέους όπου ένα μεγάλο ποσοστό έχει μόνο τη στοιχειώδη εκπαίδευση. Χαμηλά (περίπου 10% κάτω από τον υπόλοιπο μέσο όρο) βρίσκονται τα επίπεδα απασχόλησης που αφορούν νέους 20-29 ετών, ενώ ιδιαίτερα χαμηλά (περίπου 20% κάτω από τον υπόλοιπο μέσο όρο) είναι η απασχόληση στις ηλικίες κοντά στη συνταξιοδότηση (60-64 ετών). Υψηλή αναλογία (περίπου 10% πάνω από τον υπόλοιπο μέσο όρο) έχουν τα τεχνικά επαγγέλματα (τεχνίτες, εργάτες εκτός γεωργίας, χειριστές μεταφορικών μέσων κ.λπ.).

Ομάδα 13: Μικροί Δήμοι της Ξάνθης και της Ροδόλης- Μειονότητες

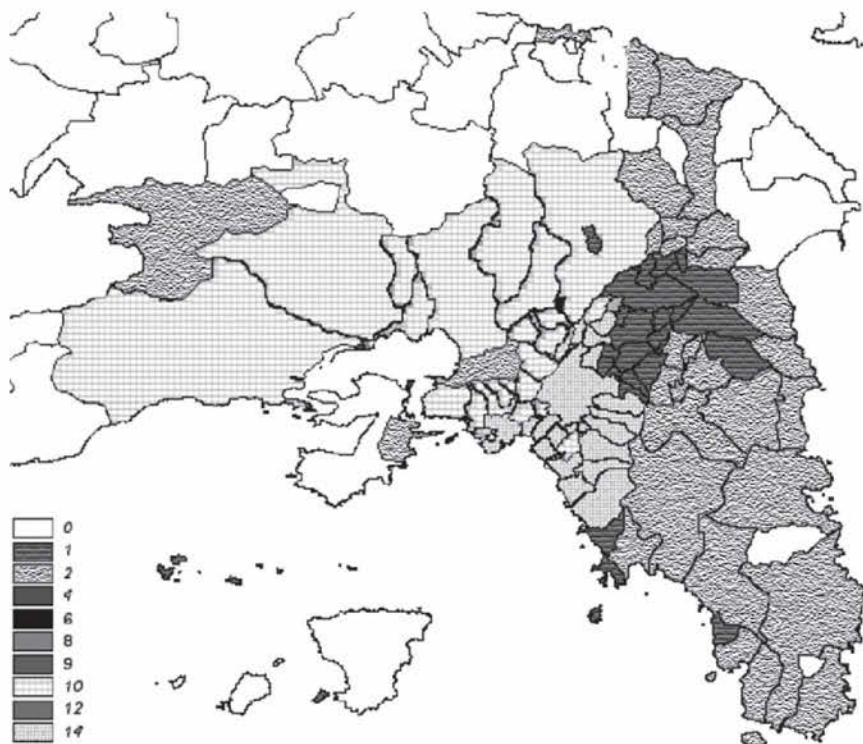
Εδώ βρίσκονται 7 κοινότητες της Ξάνθης και της Ροδόλης που μόλις ξεπερνούν τους 2000 κατοίκους. Έχουν το καφέ χρώμα. Το ποσοστό των έγγαμων κοριτσιών ηλικίας 10-14 ετών είναι επταπλάσιο ενώ το ποσοστό αυτών που ασχολούνται με τη γεωργία και την κτηνο-

τροφία είναι σχεδόν πενταπλάσιο σε αυτές τις κοινότητες σε σχέση με τον μέσο όρο των υπολοίπων δήμων-κοινοτήτων.

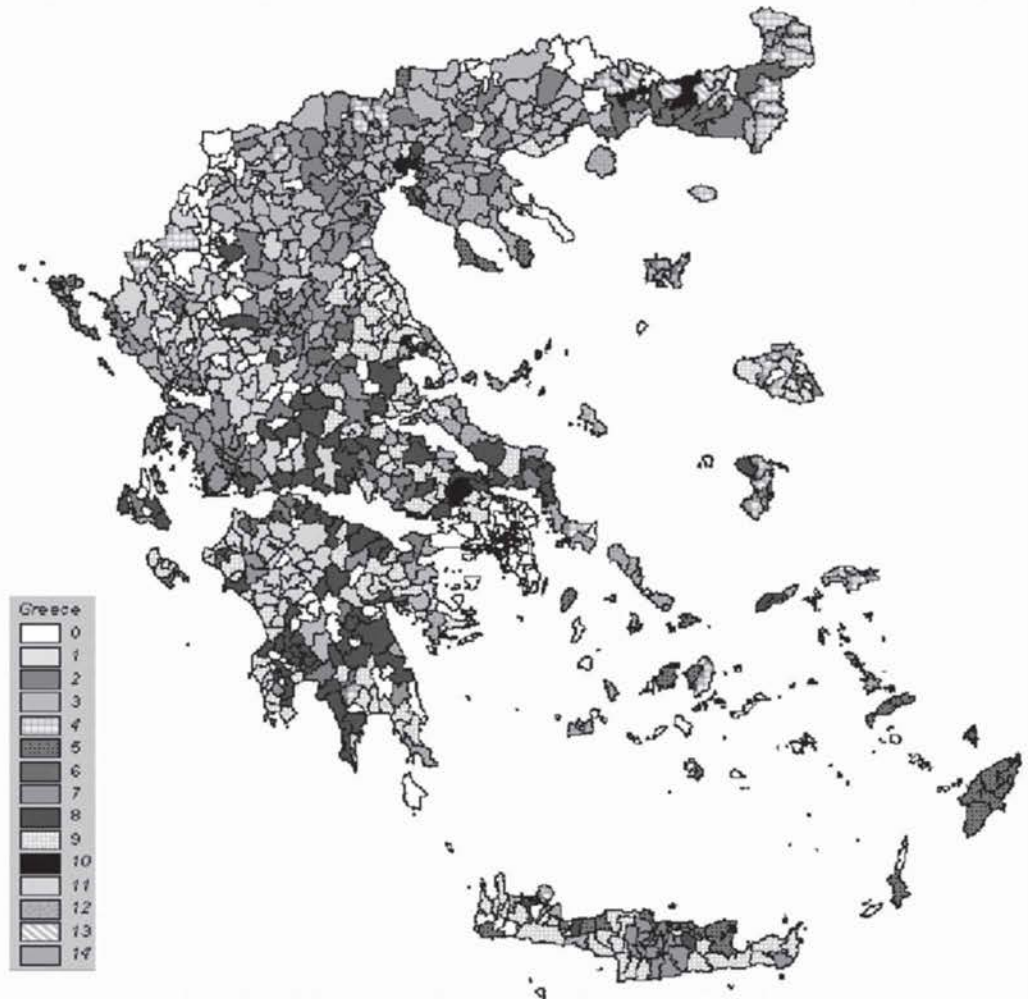
Ομάδα 14: Μέσες Αστικές Περιοχές – Προάστια Αθήνας και Θεσσαλονίκης

Εδώ βρίσκονται 22 δήμοι. Χαρακτηριστικοί αντιπρόσωποι της ομάδας είναι το Μοσχάτο, η Ν. Ιωνία και η Νέα Φιλαδέλφεια στο πολεοδομικό συγκρότημα της πρωτεύουσας και ο δήμος Θεσσαλονίκης.

Σχήμα 4: Παρουσίαση των αποτελεσμάτων της ομαδοποίησης με την μέθοδο ward linkage στο χάρτη του λεκανοπεδίου Αττικής.



Σχήμα 5: Παρουσίαση των αποτελεσμάτων της ομαδοποίησης με την μέθοδο ward linkage στο χάρτη της Ελλάδας



ΣΥΜΠΕΡΑΣΜΑΤΑ

Τα αποτελέσματα τα οποία προέκυψαν από την εφαρμογή των παραπάνω μεθόδων κρίνονται σε μεγάλο βαθμό ικανοποιητικά. Εμφανίζονται ομάδες που αποτελούνται από δήμους με παρόμοια χαρακτηριστικά γεγονός που συμβαδίζει με την κοινή αντίληψη για αυτές τις περιοχές. Στην πλειοψηφία τους οι ομάδες διαφοροποιούνται αρκετά αν και εμφανίζονται και περιπτώσεις στις οποίες είναι δύσκολο να επισημανθούν κάποιες διαφορές. Το γεγονός αυτό οφείλεται εν μέρει στα χαρακτηριστικά των ίδιων των μεταβλητών που χρησιμοποιήθηκαν. Όπως αναφέρθηκε οι μεταβλητές είχαν κυρίως κοινωνικό χαρακτήρα και λιγότερο οικονομικό, ενώ οι πληροφορίες που μας έδιναν δεν ήταν ιδιαίτερα αναλυτικές. Πιθανόν έχοντας περισσότερα στοιχεία για την οικονομική ζωή της κάθε περιοχής, καθώς και στοιχεία για την παρουσία οικονομικών μεταναστών θα είχε σαν αποτέλεσμα μια ακόμα πιο λεπτομερή ταξινόμηση. Φυσικά πρέπει να αναφέρουμε ότι τα σχετικά αστικοποιημένα αποτελέσματα και το ότι από κάποιο βαθμό και έπειτα οι απομακρυσμένοι ελαρχικοί δήμοι δεν παρουσιάζουν διαφορές μεταξύ τους είναι μια ένδειξη της μειωμένης περιφερειακής ανάπτυξης της χώρας.

Βιβλιογραφία

- Anderberg M.R. (1973) *Cluster Analysis for Applications*, New York: Academic Press, Inc.
- Cattell R.B. (1966) "The screen test for the number of factors", *Journal of Multivariate Behavioral Research*, 1: 245-276.
- Charlton M., Openshaw S. και Wymer C. (1985) "Some new classifications of census enumeration districts in Britain. A poor man's ACORN", *Journal of Economic and Social Measurement*, Vol 13: 69-96.
- Everitt B.S. (1993) *Cluster Analysis*, London: Edward Arnold.
- Everitt B.S., Landau, S. και Leese, M. (2001) *Cluster Analysis*, London: Edward Arnold.
- Hotelling H. (1933) "Analysis of a Complex of Statistical variables into principal components", *Journal of Educational Psychology*, 24:417-441, 498-520.
- Imber V. (1977) "A classification of the English personal social services authorities", DHSS, Statistical and Research Report Series, No. 16 HMSO, London.
- Jolliffe I.T. (1986), *Principal Component Analysis*, Springer-Verlag New York.
- Jolliffe I.T., Jones, B. και Morgan, B.J.T. (1986) "Comparison of cluster analyses of the English personal social services authorities", *Journal of the Royal Statistical Society Series A-Stat. Soc.*, 149: 253-270.
- Καμαριανάκης Γ. και Πραστάκος Π. (2001) "Ταξινόμηση των δήμων της Ελλάδας

- σύμφωνα με τα κοινωνικά και οικονομικά χαρακτηριστικά τους (Απογραφή 1991)", στο Κοτζαμάνης Β. και Παππάς Β. (επ.) *Οι Χωρικές Διαστάσεις των Δημογραφικών Φαινομένων*, Πανεπιστήμιο Θεσσαλίας σσ. 187-208.
- Kaufman L. και Rousseeuw P.J. (1990) *Finding groups in Data*, New York: John Wiley & Sons.
- Lance G.N. και Williams W.T. (1967) "A General Theory of Classificatory Sorting Strategies, I. Hierarchical Systems", *Computer Journal*, 9: 373-380.
- Mardia K.V., Kent J.T. και Bibby J.M. (1979) *Multivariate Analysis*, London: Academic Press.
- Milligan G.W. (1980) "An Examination of the Effect of Six Types of Error perturbation on Fifteen Clustering Algorithms", *Psychometrika*, 45: 325-342.
- Milligan G.W. και Cooper M.C. (1987) "A Study of Variable Standardization", *College of Administrative Science Working Paper Series*, 87-63, The Ohio State University, Columbus, Ohio.
- Moser C.A. και Scott W. (1961) *British Towns*, Edinburgh: Oliver and Boyd.
- Pearson K.(1901) "On lines and planes of closest fit to systems of points in space", *Philosophical Magazine* (6), 2: 559-572.
- Prastacos P. και Hootkins S. (1984) *Social Area Analysis of the San Francisco Area*, United Way, San Francisco, California.
- Πραστάκος Π. και Κογχυλάκης Γ. (1998) *Το σύστημα "Αθηνά" για την απογραφή του 1991*, Ίδρυμα Τεχνολογίας και Έρευνας, Ηράκλειο, Κρήτη.
- Πραστάκος Π. και Κογχυλάκης Γ. (2003) *Το σύστημα "Αθηνά" για την απογραφή του 2001*, Ίδρυμα Τεχνολογίας και Έρευνας, Ηράκλειο, Κρήτη.
- Wallace M. και Denham C. (1996) *The ONS classification of local and health authorities of Great Britain. Studies on Medical and Population Subjects*, ONS. Number 59.
- Ward J.H. (1963) "Hierarchical Grouping to Optimize an Objective Function", *Journal of the American Statistical Association*, 58: 236-244.
- Webber R. και Craig J. (1978) *Socio-economic classification of local authority areas*, OPCS Studies on Medical and Population Subjects, No. 35, HMSO, London.

Γιάννης Καμαριανάκης

Ίδρυμα Τεχνολογίας και Έρευνας, Ινστιτούτο Υπολογιστικών Μαθηματικών, Τομέας Περιφερειακής Ανάλυσης, Βασιλικά Βουτών, Τ.Θ. 1527, 71110 Ηράκλειο, e-mail: kamarian@iacm.forth.gr

Δημήτρης Κοντός

Ίδρυμα Τεχνολογίας και Έρευνας, Ινστιτούτο Υπολογιστικών Μαθηματικών, Τομέας Περιφερειακής Ανάλυσης, Βασιλικά Βουτών, Τ.Θ. 1527, 71110 Ηράκλειο, e-mail: manikasdkp@yahoo.com

