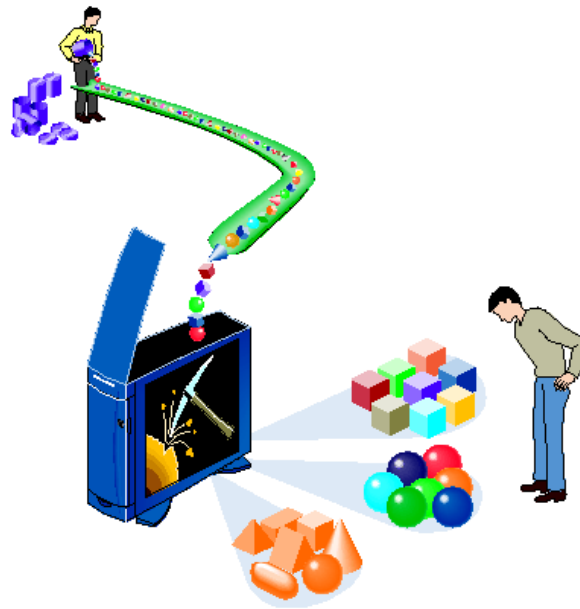




ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΒΙΟΜΗΧΑΝΙΚΗΣ ΔΙΟΙΚΗΣΗΣ
& ΕΠΙΧΕΙΡΗΣΙΑΚΗΣ ΕΡΕΥΝΑΣ

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές
Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής
Αποτελεσμάτων και Συμπεράσματα



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
ΤΟΥ
ΚΩΝΣΤΑΝΤΙΝΟΥ Α. ΧΑΡΑΛΑΜΠΟΥΣ

Υπεύθυνος Καθηγητής : Τασιόπουλος Ηλίας
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2009

Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Μηχανολόγων Μηχανικών

Τομέας Βιομηχανικής Διοίκησης & Επιχειρησιακής Έρευνας

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Του Κωνσταντίνου Α. Χαραλάμπους, Α.Μ. : 02103611

Υπεύθυνος Καθηγητής : Τασιόπουλος Ηλίας, καθηγητής Ε.Μ.Π.

ΤΙΤΛΟΣ :

**Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές
Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής
Αποτελεσμάτων και Συμπεράσματα**

Αθήνα, Ιούλιος 2009

Αφιερώνεται στην οικογένειά μου,

για το ενδιαφέρον και τη στήριξη που μου προσέφερε, κατά τη διάρκεια των μακροχρόνιων σπουδών μου στο Ε.Μ.Π.

ΕΥΧΑΡΙΣΤΙΕΣ :

Θα ήθελα να ευχαριστήσω τον υπεύθυνο καθηγητή μου, κ. Ηλία Τατσιόπουλο, για την ευκαιρία που μου έδωσε να ασχοληθώ με ένα τόσο ενδιαφέρον και συνάμα σύγχρονο θέμα, καθώς και για την καθοδήγηση και τη βοήθεια που μου παρείχε, καθ' όλη τη διάρκεια εκπόνησης της εργασίας.

Ιδιαίτερα, ευχαριστώ πολύ τον επιβλέποντα της εργασίας, υποψήφιο διδάκτορα Γιώργο Τζιραλή, για την ουσιαστική και καθοριστική συμβολή του και την αμέριστη βοήθειά του στην εκπόνηση της εργασίας. Οι μακρές και γεμάτες κέφι συζητήσεις μας οδήγησαν σε μία άριστη συνεργασία και συνέβαλαν στη συνέχιση εκτύλιξης του μίτου της Αριάδνης ή τομής του γόρδιου δεσμού.

0] Έποψη.

Οι Ολυμπιακοί Αγώνες αποτελούν ένα από τα πιο σημαντικά αθλητικά γεγονότα στον παγκόσμιο χάρτη. Δεν αποτελούν απλούς αγώνες, αλλά ένα σπουδαίο κοινωνικό - πολιτισμικό, πολιτικό, ακόμα και οικονομικό γεγονός. Τι γίνεται όμως με τη διοργανώτρια κάθε φορά πόλη - χώρα; Πώς επηρεάζεται η οικονομία της χώρας από τη διοργάνωση ενός τέτοιου σπουδαίου γεγονότος (αρνητικά ή θετικά); Η παρούσα διπλωματική εργασία αποπειράται να απαντήσει σ' αυτά τα καίρια ερωτήματα που αφορούν τη χώρα στην οποία ζούμε, την Ελλάδα, ειδικά μετά τη διεξαγωγή των Ολυμπιακών Αγώνων «Αθήνα 2004». Με την ανάθεση των Ολυμπιακών Αγώνων στην Ελλάδα άλλοι πανηγύρισαν για καθαρά εθνικούς λόγους, αφού οι αγώνες γύρναγαν στη κοιτίδα τους, άλλοι πάλι, γιατί πίστευαν ότι οι Ολυμπιακοί Αγώνες είναι μία μεγάλη ευκαιρία για ανάπτυξη της Ελλάδας όχι μόνο στον τομέα των έργων υποδομής, αλλά και σ' άλλους τομείς της επιχειρηματικής δραστηριότητας. Τα προβλήματα που τέθηκαν στο μικροσκόπιο της έρευνας που διεξήχθη ήταν λίγο πολύ αναμενόμενα. Η πολυπλοκότητα ενός τέτοιου ζητήματος έγκειται κυρίως στο γεγονός ότι ο οικονομικός αντίκτυπος από τη διοργάνωση-διεξαγωγή των Ολυμπιακών Αγώνων στην Αθήνα χρήζει και θα χρήζει για πολλά χρόνια ακόμα μακροοικονομικής και επιχειρηματικής ανάλυσης. Χρησιμοποιώντας στην εργασία αυτή ένα σύνολο οικονομικών και άλλων δεδομένων (data set) σχετικά μικρό - από το 1996 μέχρι το 2007- αντιλαμβάνεται κανείς τη δυσκολία απάντησης σ' ένα ερώτημα της μορφής: «Βοήθησαν ή όχι οι Ολυμπιακοί Αγώνες την Ελλάδα και σε ποιους τομείς είδαμε αποτελέσματα;».

Το πρόγραμμα «Συνολικές Επιπτώσεις των Ολυμπιακών Αγώνων Αθήνα 2004» (Athens 2004 Olympic Games Global Impact - OGGI) είναι ένα ερευνητικό έργο που δημιουργήθηκε στα πλαίσια συμβατικής υποχρέωσης της Οργανωτικής Επιτροπής Ολυμπιακών Αγώνων της Αθήνας, «Αθήνα 2004» προς τη Διεθνή Ολυμπιακή Επιτροπή (ΔΟΕ), ώστε να διερευνηθεί και να αξιολογηθεί η εκτίμηση των επιπτώσεων των Ολυμπιακών Αγώνων του 2004 στους τρεις βασικούς τομείς της αειφόρου ανάπτυξης της Ελλάδας : Περιβάλλον, Οικονομία και Κοινωνία / Πολιτισμός. Στόχος του προγράμματος ήταν ο υπολογισμός των σημαντικότερων δεικτών επίδρασης των Ολυμπιακών Αγώνων, με βάση στοιχεία που αφορούν την περίοδο 1996 - 2007. Όλοι αυτοί οι δείκτες αποτελούν τη «ραχοκοκαλιά» της ελληνικής Οικονομικής Ανάπτυξης και πάρθηκαν από επίσημα στοιχεία της Εθνικής Στατιστικής Υπηρεσίας της Ελλάδας (Ε.Σ.Υ.Ε.), της Τράπεζας της Ελλάδας και της Ευρωπαϊκής Στατιστικής Υπηρεσίας Eurostat.

Η μεθοδολογία που χρησιμοποιείται στο παρόν έντυπο βασίζεται στην εξόρυξη δεδομένων (data mining), μέσω κατάλληλων αλγορίθμων και αποδοτικών τεχνικών. Συχνά υπάρχουν πληροφορίες στα δεδομένα, οι οποίες είναι «κρυμμένες», δεν είναι δηλαδή προφανείς. Οι ανθρώπινοι αναλυτές μπορεί να χρειάζονται εβδομάδες για να ανακαλύψουν χρήσιμες πληροφορίες, ενώ πολλά δεδομένα δεν αναλύονται ποτέ. Σκοπός της μελέτης που επισυνάπτεται παρακάτω είναι η διαδικασία ανακάλυψης (discovery) προτύπων (patterns) που πριν δεν ήταν γνωστά, αλλά ισχύουν και είναι πιθανόν χρήσιμα από ένα σύνολο δεδομένων (data set) και η περαιτέρω ανάλυσή τους, ούτως ώστε να βρούμε μη αναμενόμενες συσχετίσεις μεταξύ των όλων δεδομένων και να συνοψίσουμε τα τελευταία με νέους τρόπους, πιο κατανοητούς και χρήσιμους στον οποιοδήποτε χρήστη τους.

Στο παρόν σύγγραμμα επιχειρείται η αποτίμηση των οικονομικών επιπτώσεων των Ολυμπιακών Αγώνων, μέσω τεχνικών εξόρυξης πληροφορίας. Αρχικά (κεφάλαιο 2), παρουσιάζονται συνοπτικά οι σύγχρονες ολυμπιακές εκδηλώσεις, εστιάζοντας στις επιπτώσεις που έφεραν, κατά αλλά και μετά τη διεξαγωγή τους, στην οικονομία κυρίως της διοργανώτριας κάθε φορά χώρας και δίνεται μια γενική περιγραφή της διοργάνωσης «Αθήνα 2004», καθώς επίσης και της ελληνικής οικονομίας. Στη συνέχεια (κεφάλαιο 3), γίνεται μια συνολική περιγραφή του προγράμματος «Συνολικές Επιπτώσεις των Ολυμπιακών Αγώνων» (OGGI) και των δεικτών ανάλυσης που αυτό επεξεργάζεται. Ακολούθως, σκιαγραφούνται θεωρητικά - κεφάλαιο 4 - οι διάφορες τεχνικές εξόρυξης πληροφορίας μέσα από δεδομένα (data mining) που χρησιμοποιούνται και δίνονται πληροφορίες - κεφάλαιο 5 - σχετικά με το λογισμικό που χρησιμοποιείται, το WEKA. Κατόπιν, αναλύονται τα στάδια της μεθοδολογίας που ακολουθήθηκε, προκειμένου να εξάγουμε τους διάφορους κανόνες που προκύπτουν από το σύνολο των δεδομένων που έτυχαν επεξεργασίας data mining (κεφάλαιο 6). Τελικά, όλα τα αποτελέσματα των κανόνων αυτών παρουσιάζονται αναλυτικά στο κεφάλαιο 7.

Η παρούσα εργασία αποσκοπεί στη δημιουργία ενός μικρού βοηθήματος, που να λειτουργεί σαν ένα είδος δυναμικής ώθησης για φοιτητές ή και άλλους ερευνητές που καταπιάνονται σε επίπεδο επιστημονικό με το προαναφερθέν ζήτημα. Το αν επηρεάστηκε η Ελλάδα θετικά ή αρνητικά στον τομέα της Οικονομίας, από τη διεξαγωγή του μεγάλου αθλητικού γεγονότος αποτελεί ένα πολυσύνθετο πρόβλημα, το οποίο σαφώς δε δύναται να απαντηθεί μονολεκτικά, ούτε και να διερευνηθεί μονολιθικά (μονομερώς), μα χρήζει μακρόχρονων ερευνητικών μελετών και επιστημονικών διαδικασιών.

Πίνακας Περιεχομένων.

0] Έποψη.	5
1] Υπόθεση Εργασίας.	12
1.1] Γενική Οριοθέτηση του Προβλήματος.	12
1.2] Τρόπος Προσέγγισης του Προβλήματος.	14
1.3] Δυσκολίες κατά την προσέγγιση του προβλήματος.	15
2] Ολυμπιακοί Αγώνες, επιπτώσεις των Αγώνων στις διοργανώτριες πόλεις-χώρες, «Αθήνα 2004» και Ελληνική Οικονομία.	17
2.1] Συνοπτική ανασκόπηση της ιστορίας των Σύγχρονων Ολυμπιακών Αγώνων και των επιπτώσεών τους.	17
2.2] Ιδιαίτερα χαρακτηριστικά γνωρίσματα της Ελλάδας, ως διοργανώτριας χώρας.	24
2.3] Οικονομικά στοιχεία της ελληνικής οικονομίας πριν, κατά και μετά το «Αθήνα 2004».	25
2.4] Αναπτυξιακά έργα στην Ελλάδα κατά την περίοδο των Ολυμπιακών Αγώνων, στους τομείς των Δημοσίων Έργων και των Μεταφορών.	26
3] Πρόγραμμα «Συνολικές Επιπτώσεις Ολυμπιακών Αγώνων» (OGGI).	29
3.1] Τι είναι το πρόγραμμα OGGI ; - Γενικά.	29
3.2] Το πρόγραμμα OGGI του «Αθήνα 2004».	31
3.3] Μέθοδος Ανάλυσης του OGGI.	33
3.4] Δείκτες Ανάλυσης.	36
4] Εξόρυξη Δεδομένων (Data Mining).	37
4.1] Ανακάλυψη Γνώσης μέσα από τα Δεδομένα.	37
4.2] Λίγα λόγια για την Τεχνητή Νοημοσύνη και τη Μηχανική Μάθηση.	41
4.3] Η τεχνική Data Mining.	43
4.3.1. Επιστημονικό πεδίο της τεχνικής Data Mining – Βασικότεροι Στόχοι.	43
4.3.2. Συστατικά Στοιχεία Αλγορίθμων της τεχνικής Data Mining.	44
4.3.3. Κυριότεροι τύποι εργαλείων που χρησιμοποιούνται στην τεχνική Data Mining.	44
4.3.4. Τα στάδια της διαδικασίας Data Mining.	46
4.3.5. Οι μέθοδοι της τεχνικής Data Mining.	48

5] Το λογισμικό WEKA.....	52
5.1] Τι είναι το λογισμικό WEKA;.....	52
5.2] Το περιβάλλον του WEKA.....	53
5.3] Αποθήκευση δεδομένων σε αρχεία .arff.....	54
6] Μεθοδολογία.....	56
6.1] Επιλογή Δεδομένων.....	57
6.1.1. Συγκέντρωση όλων των οικονομικών δεικτών ανάλυσης του «OGGI - Αθήνα 2004».	57
6.1.2. Μορφοποίηση Πίνακα Δεδομένων.....	59
6.2] Προεπεξεργασία Δεδομένων.....	61
6.3] Επιλογή των πιο σχετικών, μεταξύ τους, χαρακτηριστικών (select attributes). ..	63
6.4] Επιλογή κατάλληλων αλγορίθμων για περαιτέρω επεξεργασία των δεδομένων.	64
6.5] Παράδειγμα επεξεργασίας ενός χαρακτηριστικού μέσα από τα διάφορα στάδια της μεθοδολογίας.....	65
7] Παρουσίαση και Ανάλυση Αποτελεσμάτων.	72
7.1] Gross Domestic Product (G.D.P.) - Ακαθάριστο Εγχώριο Προϊόν (Α.Ε.Π.).....	73
7.1.1. Ορισμός.....	73
7.1.2. Επιλογή των πιο σχετικών, με το Α.Ε.Π., χαρακτηριστικών (selected attributes).	73
7.1.3. Κανόνες Ταξινόμησης.....	74
7.2] Gross National Product (G.N.P.) - Ακαθάριστο Εθνικό Προϊόν (Α.ΕΘ.Π.).....	76
7.2.1. Ορισμός.....	76
7.2.2. Επιλογή των πιο σχετικών, με το Α.ΕΘ.Π., χαρακτηριστικών (selected attributes). ..	76
7.2.3. Κανόνες Ταξινόμησης.....	76
7.3] Gini Index - Δείκτης Εισοδηματικής Κατανομής (Δ.Ε.Κ.).....	77
7.3.1. Ορισμός.....	77
7.3.2. Επιλογή των πιο σχετικών, με το Δ.Ε.Κ. Gini, χαρακτηριστικών (selected attributes).	77
7.3.3. Κανόνες Ταξινόμησης.....	77
7.4] Consumer Price Index (C.P.I.) - Δείκτης Τιμών Καταναλωτή (Δ.Τ.Κ.).....	80
7.4.1. Ορισμός.....	80
7.4.2. Επιλογή των πιο σχετικών, με το Δ.Τ.Κ., χαρακτηριστικών (selected attributes).....	80

7.4.3. Κανόνες Ταξινόμησης.	81
7.5] Gross Fixed Capital Formation - Ακαθάριστος Σχηματισμός Παγίου Κεφαλαίου.	84
7.5.1. Ορισμός.	84
7.5.2. Επιλογή των πιο σχετικών, με τον Ακαθάριστο Σχηματισμό Παγίου Κεφαλαίου, χαρακτηριστικών (selected attributes).	84
7.5.3. Κανόνες Ταξινόμησης.	85
7.6] Fiscal Revenue - Φορολογικό Εισόδημα.....	88
7.6.1. Ορισμός.	88
7.6.2. Επιλογή των πιο σχετικών, με το Φορολογικό Εισόδημα, χαρακτηριστικών (selected attributes).....	88
7.6.3. Κανόνες Ταξινόμησης.	90
7.7] Inflation - Πληθωρισμός.	95
7.7.1. Ορισμός.	95
7.7.2. Επιλογή των πιο σχετικών, με τον Πληθωρισμό, χαρακτηριστικών (selected attributes).....	95
7.7.3. Κανόνες Ταξινόμησης.	96
7.8] Exchange Rate Euro / U.S.Dollar - Ισοτιμία Συναλλάγματος Ευρώ / Αμερικανικού Δολαρίου.	99
7.8.1. Ορισμός.	99
7.8.2. Επιλογή των πιο σχετικών, με την Ισοτιμία Συναλλάγματος Ευρώ / Αμερικανικού Δολαρίου, χαρακτηριστικών (selected attributes).	99
7.8.3. Κανόνες Ταξινόμησης.	100
7.9] Low Wage Population - Χαμηλόμισθος Πληθυσμός.....	101
7.9.1. Ορισμός.	101
7.9.2. Επιλογή των πιο σχετικών, με το Χαμηλόμισθο Πληθυσμό, χαρακτηριστικών (selected attributes).	101
7.9.3. Κανόνες Ταξινόμησης.	101
7.10] Minimum Daily Wage of Workers - Ελάχιστος Ημερήσιος Μισθός των Εργαζόμενων.	102
7.10.1. Ορισμός.	102
7.10.2. Επιλογή των πιο σχετικών, με τον ελάχιστο ημερήσιο μισθό των Εργαζόμενων, χαρακτηριστικών (selected attributes).	102
7.10.3. Κανόνες Ταξινόμησης.	102

7.11] Per Capita Gross Dept - Κατά Κεφαλή Ακαθάριστο Χρέος.	104
7.11.1. Ορισμός.	104
7.11.2. Επιλογή των πιο σχετικών, με το κατά κεφαλή Ακαθάριστο Χρέος, χαρακτηριστικών (selected attributes).	104
7.11.3. Κανόνες Ταξινόμησης.	104
7.12] Public Spending - Δημόσιες Δαπάνες.	105
7.12.1. Ορισμός.	105
7.12.2. Επιλογή των πιο σχετικών, με τις Δημόσιες Δαπάνες, χαρακτηριστικών (selected attributes).....	105
7.12.3. Κανόνες Ταξινόμησης.	105
7.13] Ratio of Services (Exports to Imports) - Αναλογία των Υπηρεσιών (Εξαγωγές προς Εισαγωγές).	106
7.13.1. Ορισμός.	106
7.13.2. Επιλογή των πιο σχετικών, με την αναλογία των Υπηρεσιών (Εξαγωγές προς Εισαγωγές), χαρακτηριστικών (selected attributes).	106
7.13.3. Κανόνες Ταξινόμησης.	107
7.14] Total Energy Gross Inland Consumption - Συνολική Ακαθάριστη Ενεργειακή Εγχώρια Κατανάλωση.....	109
7.14.1. Ορισμός.	109
7.14.2. Επιλογή των πιο σχετικών, με τη συνολική Ακαθάριστη Ενεργειακή Εγχώρια Κατανάλωση, χαρακτηριστικών (selected attributes).....	109
7.14.3. Κανόνες Ταξινόμησης.	109
7.15] Unemployment Rate - Ποσοστό Ανεργίας.....	110
7.15.1. Ορισμός.	110
7.15.2. Επιλογή των πιο σχετικών, με το ποσοστό ανεργίας, χαρακτηριστικών (selected attributes).....	110
7.15.3. Κανόνες Ταξινόμησης.	110
8] Σύνοψη.	112
8.1] Απολογισμός.....	112
8.2] Προτεινόμενες Κατευθύνσεις Μελλοντικής Ενασχόλησης.	114
8.3] Αντί επιλόγου.	115
9] Βιβλιογραφία.	116
9.1] Συγγράμματα.	116

9.2] Εφημερίδες - Περιοδικά.....	118
9.3] Ηλεκτρονικές Πηγές.	118
9.4] Ηλεκτρονικές Διευθύνσεις για Εικόνες.....	120
10] Παραρτήματα.....	121
10.1] Παράρτημα Α - Δείκτες ΟGGI πλαισίου των Αγώνων (context indexes).....	121
10.2] Παράρτημα Β - Τυποποίηση σε αρχείο .arff, για εισαγωγή των δεδομένων στο λογισμικό WEKA.	135
10.3] Οπτικοποίηση χαρακτηριστικών.	143
10.4] Παράρτημα Δ - Πινακοποίηση των πιο σχετικών, μεταξύ τους, χαρακτηριστικών (select attributes).	158
10.5] Σημαντικότεροι κανόνες ταξινόμησης.....	163

1] Υπόθεση Εργασίας.

1.1] Γενική Οριοθέτηση του Προβλήματος.

«Οι καλύτεροι Αγώνες που έγιναν ποτέ», υποστήριξε ο πρόεδρος της Διεθνούς Ολυμπιακής Επιτροπής, Ζακ Ρογκ, κατά την τελετή λήξης των 28^{ωv} Ολυμπιακών Αγώνων «Αθήνα 2004». Ωστόσο, κάθε νοήμων άνθρωπος διερωτάται εύλογα τα εξής: Πώς μπορεί κάποιος να μετρήσει την επιτυχία ή αποτυχία που επέφεραν οι Ολυμπιακοί Αγώνες; Μια δήλωση του τύπου «οι Αγώνες στέφθηκαν από απόλυτη επιτυχία» τι μπορεί να σημαίνει ακριβώς; Η επιτυχία αυτή είναι οικονομική, κοινωνική, οργανωτική ή απλά αθλητική; Και όλοι αυτοί που επιχειρούν να απαντήσουν στα ερωτήματα αυτά, από ποια οπτική γωνία το κάνουν;¹ Όλα αυτά τα ζητήματα, ακόμα και αν απαντηθούν, σκιαγραφούν μια χαοτική κατάσταση, ανάγοντας την πολυπλοκότητα του προβλήματος των «επιπτώσεων των Ολυμπιακών Αγώνων Αθήνα 2004» σε μείζον θέμα. Πιο συγκεκριμένα, στον Τομέα της Ελληνικής Οικονομίας, τίθενται κάποια θεμελιώδη ερωτήματα: ποια τα οικονομικά οφέλη ή ποιες οι αρνητικές επιπτώσεις που προέκυψαν από τους Ολυμπιακούς Αγώνες της Αθήνας; Σε τι επίπεδα κυμαίνεται η τουριστική ανάπτυξη τα επόμενα χρόνια; Με ποιους τρόπους θα μπορούσαν να αξιοποιηθούν καλύτερα οι Ολυμπιακές εγκαταστάσεις;

Χιλιάδες πολιτικοί, δημοσιογράφοι-αρθρογράφοι, επιστημονικοί αναλυτές, ακόμα και απλοί άνθρωποι της καθημερινότητας έκαναν τη δική τους αποτίμηση των επιπτώσεων που επέφεραν οι Ολυμπιακοί Αγώνες της Αθήνας και εξέφρασαν την άποψή τους, επισήμως ή ανεπίσημως. Πολλές επιστημονικές μελέτες στηρίχθηκαν σε στατιστικές αναλύσεις οικονομικών και άλλων δεικτών, προκειμένου «να ξετυλίξουν το μίτο της Αριάδνης και να βρουν την έξοδο του λαβύρινθου». Μέσω των στατιστικών αναλύσεων, επιχειρήθηκε πολλές φορές να γίνει μία άμεση ή έμμεση σύγκριση των οικονομικών επιπτώσεων των Ολυμπιακών Αγώνων της Αθήνας με τις αντίστοιχες προηγούμενων Ολυμπιάδων του παρελθόντος, ώστε να απαντηθεί το ερώτημα του κατά πόσον οι Ολυμπιακοί Αγώνες της Αθήνας θεωρούνται «οικονομικά» επιτυχημένοι. Άλλες φορές, στα πλαίσια μακροοικονομικών επιχειρηματικών αναλύσεων, επιχειρήθηκε μια μακροσκελής σύγκριση των δεικτών της ελληνικής οικονομίας των χρόνων που προηγήθηκαν και ακολούθησαν των Ολυμπιακών Αγώνων και βγήκαν αποτελέσματα, τα οποία προκάλεσαν διχογνωμίες (θετικές ή αρνητικές οι επιπτώσεις;). Αρκετοί από όλους αυτούς τους πολιτικούς, δημοσιογράφους, επιστημονικούς αναλυτές κ.ά., που ασχολήθηκαν με το

¹ Βλ. Holger Preuss, *The Economics of Staging the Olympics*, σ. 25, εκδ. MPG Books Ltd, Bodmin, Cornwall, Great Britain, 2004.

ζήτημα των επιπτώσεων των Ολυμπιακών Αγώνων της Αθήνας, υποστήριξαν τελικά ότι οι Αγώνες αποτέλεσαν ένα άλμα προόδου για την ελληνική οικονομία. Η Υπουργός Εξωτερικών, Ντόρα Μπακογιάννη, σε ομιλία της στο Πανεπιστήμιο του Πεκίνου το 2007, δήλωσε: «Οι Ολυμπιακοί μας κληροδότησαν μια πιο όμορφη και πιο ανθρώπινη πόλη. Η πόλη της Αθήνας έχει ένα νέο πρόσωπο, το πρόσωπο μιας σύγχρονης, συνεχώς εξελισσόμενης μητρόπολης. Είναι ένα ζωντανό πολιτιστικό και οικονομικό κέντρο της Νοτιο-ανατολικής Ευρώπης. Είναι επίσης μια πιο προσβάσιμη πόλη με 75 χιλιόμετρα πεζοδρομίων, 250 χιλιόμετρα πεζοδρομημένων οδών και πεζοδρομίων, καθώς και νέων δικτύων ηλεκτροδότησης, τα οποία ξεπερνούν τα 300 χιλιόμετρα. Η Αθήνα σήμερα διαθέτει επίσης μια εντυπωσιακή πολιτιστική υποδομή, καθώς και ένα δίκτυο ασφάλειας με ευρεία ζήτηση. Ο τουρισμός στη χώρα μας έχει αυξηθεί εκθετικά, επηρεάζοντας θετικά την οικονομία μας».² Από την άλλη, πολλοί τάσσονται κατά των Ολυμπιακών Αγώνων, ισχυριζόμενοι ότι οι Αγώνες επιφέρουν σημαντική οικονομική ύφεση στη διοργανώτρια χώρα. Ο Γρηγόρης Παπανίκος, μέλος του Πανελληνίου Συνδέσμου Αθλητικών Οικονομολόγων και Στελεχών, σε ερώτηση δημοσιογράφου της εφημερίδας «Ριζοσπάστης» εάν «οι Ολυμπιακοί Αγώνες αποτελούν ευκαιρία για οικονομική άνθηση μιας χώρας», απάντησε κατηγορηματικά : «Σίγουρα όχι. Εξάλλου, αυτό το ίδιο το γεγονός δε δημιουργεί εκείνες τις οικονομικές επιπτώσεις που θα μπορούσαν να αλλάξουν τις αναπτυξιακές προοπτικές μιας χώρας. Το αντίθετο μάλιστα, σε πολλές περιπτώσεις, οι Ολυμπιακοί Αγώνες μπορούν να έχουν αρνητικές επιπτώσεις στην οικονομική ανάπτυξη μιας περιοχής».³

Λαμβάνοντας υπόψη όλα τα πιο πάνω, διερωτάται κάποιος: Τελικά ποια είναι η απάντηση στο πρόβλημα που τίθεται σχετικά με το εάν το σύνολο των επιπτώσεων των Ολυμπιακών Αγώνων προσμετράται με θετικό ή αρνητικό πρόσημο; Υπάρχει σαφής και τεκμηριωμένη απάντηση ή το να επιχειρηθεί κάτι τέτοιο είναι υποκειμενικό;

² Βλ. Μπακογιάννη Ντόρα, «Οι Ολυμπιακοί : Κάτι παραπάνω από απλοί αγώνες, μια παρακαταθήκη για το μέλλον - Πανεπιστήμιο του Πεκίνου», <<http://www.dorabak.gr/default.asp?pid=10&riD=1016&la=1>>, 6/6/2009.

³ Βλ. Πασακυριάκος Κώστας, «Πιθανότερες οι αρνητικές επιπτώσεις, Συνέντευξη με το Γρηγόρη Παπανίκο, μέλος του Πανελληνίου Συνδέσμου Αθλητικών Οικονομολόγων και Στελεχών», *Ριζοσπάστης*, 1/7/2001, σ.19.

1.2] Τρόπος Προσέγγισης του Προβλήματος.

Το πιο πάνω ερώτημα αποτέλεσε την αφορμή συγγραφής της παρούσας διπλωματικής εργασίας. Εξετάζοντας ήδη υπάρχουσες μελέτες σχετικά με το πρόβλημα που οριοθετείται πιο πάνω, διαπιστώθηκε ότι κανείς μέχρι σήμερα δεν επιχείρησε να αποτιμήσει τις επιπτώσεις των Ολυμπιακών Αγώνων της Αθήνας λίγο πιο βαθιά. Είναι φορές που σε κάποια οικονομικά ή άλλα στοιχεία υπάρχουν, πέραν των αριθμών, πληροφορίες «κρυμμένες». Τις πληροφορίες αυτές είναι δύσκολο - έως αδύνατο πολλές φορές - να τις αντιληφθεί ο ανθρώπινος νους, μιας και δεν είναι προφανείς. Έτσι αποφασίστηκε όπως η διπλωματική αυτή εργασία να μην αποτελέσει μια κοινότυπη μελέτη επίπτωσης των Αγώνων της Αθήνας στον τομέα της Οικονομίας της Ελλάδας, αλλά μία καινοτομική εισήγηση, στα πλαίσια της οποίας να εξετάζονται οι διάφοροι δείκτες υπό το πρίσμα της μηχανικής μάθησης (machine learning). Σκοπός του παρόντος συγγράμματος είναι να αποτελέσει ένα είδος ώθησης, ένα μικρό βοήθημα για φοιτητές ή άλλους ερευνητές, οι οποίοι στο μέλλον θα ασχοληθούν με το επίμαχο ζήτημα των επιπτώσεων των Ολυμπιακών Αγώνων της Αθήνας.

Έχοντας ως γνώμονα τις τεχνικές της τεχνητής νοημοσύνης και της στατιστικής ανάλυσης, επιχειρείται μία διαφορετική ανάλυση οικονομικών δεικτών της Ελλάδας, μέσω της μεθοδολογίας Data Mining⁴ και πιο συγκεκριμένα του λογισμικού WEKA.⁵ Στόχο αποτελεί η διαδικασία ανακάλυψης (discovery) προτύπων (patterns) που πριν δεν ήταν γνωστά, αλλά ισχύουν και είναι πιθανόν χρήσιμα, από ένα σύνολο δεδομένων (data set), καθώς και η περαιτέρω ανάλυσή τους, ούτως ώστε να βρούμε μη αναμενόμενες συσχετίσεις μεταξύ των όλων δεδομένων και να συνοψίσουμε τα τελευταία με νέους τρόπους, πιο κατανοητούς και χρήσιμους για τον οποιοδήποτε χρήστη τους. Το σύνολο των δεδομένων που χρησιμοποιείται στην εργασία αποτελεί μια γκάμα των σημαντικότερων δεικτών της ελληνικής οικονομίας, όπως αυτοί υπολογίστηκαν στα πλαίσια του ερευνητικού προγράμματος «OGGI (Olympic Games Global Impact) Athens 2004».⁶

⁴ Βλ. σχετικά το κεφάλαιο 4 της παρούσας εργασίας.

⁵ Βλ. σχετικά το κεφάλαιο 5 της παρούσας εργασίας.

⁶ Σημειωτέον ότι το πρόγραμμα «OGGI» κατέληξε στον υπολογισμό των δεικτών, δανειζόμενο επίσημα στοιχεία της Γενικής Γραμματείας της Εθνικής Στατιστικής Υπηρεσίας Ελλάδας (Ε.Σ.Υ.Ε.) - <<http://www.statistics.gr/>> - και της Ευρωπαϊκής Στατιστικής Υπηρεσίας Eurostat - <<http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home>>.

1.3] Δυσκολίες κατά την προσέγγιση του προβλήματος.

Κατά την προσπάθεια εξόρυξης γνώσης μέσα από το σύνολο δεδομένων των οικονομικών δεικτών που χρησιμοποιήσαμε ως εργαλείο, παρουσιάστηκαν - αναμενόμενα εκ των προτέρων - κάποια προβλήματα. Τα προβλήματα αυτά εστιάζονταν, κυρίως, στο γεγονός ότι η τεχνική Data Mining και οι μέθοδοι οι οποίες χρησιμοποιεί αποτελούν πειραματική επιστήμη και συνεχώς αναβαθμίζονται ή μεταβάλλονται. Παρακάτω παρατίθενται τα δύο σημαντικότερα προβλήματα που έπρεπε να αντιμετωπιστούν:

(α) Στο σύνολο των δεδομένων, το οποίο είχαμε από τους δείκτες ΟGGI, ο αριθμός των υποδειγμάτων ήταν αρκετά μικρός. Στην τεχνική data mining, με τον όρο υποδείγματα (instances) εννοούμε τα ξεχωριστά και ανεξάρτητα παραδείγματα (examples) της όλης αντίληψης (concept) που εξετάζουμε. Τα υποδείγματα μετρούνται από τα χαρακτηριστικά (attributes), τα οποία διακρίνονται σε αριθμητικές και ονομαστικές συνιστώσες.⁷ Στην παρούσα διπλωματική εργασία εισήχθησαν στο λογισμικό WEKA περί τα 400 attributes⁸ (αριθμός πολύ ικανοποιητικός), για να προσδιορίσουν ή να περιγράψουν 12 instances⁹ (σχετικά μικρός αριθμός). Αυτό προκάλεσε και μια σύγχυση στην αποτίμηση των τελικών αποτελεσμάτων και στη γενικότερη εξαγωγή συμπερασμάτων.

(β) Η τεχνική data mining, όντας αντικείμενο της Μηχανικής Μάθησης (Machine Learning), στηρίζεται στο ευρύτερο φάσμα της επιστήμης της Τεχνητής Νοημοσύνης (Artificial Intelligence), η οποία υστερεί σε σχέση με την ανθρώπινη νοημοσύνη, αφού πολλές φορές παρουσιάζει αποτελέσματα που υστερούν φυσικής σημασίας και χρησιμότητας και οδηγούν σε αμφίβολα συμπεράσματα. Αυτό είχε ως αποτέλεσμα τη δυσκολία κατανόησης ορισμένων συσχετίσεων των attributes μεταξύ τους, όπως παρουσιάστηκαν στο WEKA, λόγω του ότι πολλές φορές οι συσχετίσεις έγιναν στα πλαίσια της «τεχνητής» λογικής και όχι στο φάσμα της ανθρώπινης.

Ωστόσο, παρά τις επιμέρους δυσκολίες, το είδος της προσέγγισης των οικονομικών επιπτώσεων των αθλητικών ολυμπιακών αγώνων, που επιχειρείται στην παρούσα μελέτη, καθίσταται ιδιαίτερα ενδιαφέρον, βάσει κυρίως του καινοτομικού χαρακτήρα του. Τα διάφορα

⁷ Βλ. Τζιραλής Γιώργος, «Αλγόριθμοι Εξόρυξης Πληροφορίας, Διάλεξη 2, Συνιστώσες Δεδομένων–Οπτικοποίηση–Εξερεύνηση», <http://gtziralis.googlepages.com/Lecture02_DataComponentsVisualizationExploration.pdf>, 6/6/2009.

⁸ Σημειωτέον ότι με τον όρο attributes εννοούνται όλα τα αποτελέσματα των οικονομικών δεικτών ΟGGI που χρησιμοποιήθηκαν στο πρόγραμμα WEKA. Βλ. σχετικά με τους δείκτες το κεφάλαιο 6.1.1. της Μεθοδολογίας.

⁹ Σημειωτέον το γεγονός ότι τα 12 instances στα οποία αναφερόμαστε είναι οι χρονολογίες της περιόδου 1996-2007, των οποίων χρησιμοποιήθηκαν αριθμητικά χαρακτηριστικά.

εμπόδια που συναντήθηκαν, ήταν λίγο πολύ αναμενόμενα, όχι μόνο από την άποψη του τρόπου αντιμετώπισης του προβλήματος, αλλά και από το ποιόν του ίδιου του υπό εξέταση ζητήματος, της αποτίμησης δηλαδή των οικονομικών συνεπειών των ολυμπιακών παιχνιδιών του 2004. Το θέμα του τρόπου και του βαθμού επίδρασης των ολυμπιακών διοργανώσεων, αναφορικά με την φιλοξενούσα κάθε φορά χώρα, αποτελεί ανέκαθεν ένα κομβικό σημείο αναφοράς των εκάστοτε οικονομικών συζητήσεων, ένα πολυδύναμο γεγονός, που ακυρώνει εκ των προτέρων προσεγγίσεις μονολιθικού τύπου. Παρακάτω, πραγματοποιείται μια συνοπτική παρουσίαση των σύγχρονων ολυμπιακών εκδηλώσεων, εστιάζοντας στις επιπτώσεις που έφεραν, κατά αλλά και μετά τη διεξαγωγή τους, στην οικονομία κυρίως της διοργανώτριας κάθε φορά χώρας.

2] Ολυμπιακοί Αγώνες, επιπτώσεις των Αγώνων στις διοργανώτριες πόλεις-χώρες, «Αθήνα 2004» και Ελληνική Οικονομία.

«Η πόλη που θα έχει την τιμή και, ταυτόχρονα, την υπευθυνότητα να διοργανώσει τους Ολυμπιακούς Αγώνες του 2004 είναι... η Αθήνα». Είναι Παρασκευή, 5 Σεπτεμβρίου του 1997, όταν, στη Λωζάνη της Ελβετίας, ο τότε πρόεδρος της Διεθνούς Ολυμπιακής Επιτροπής (ΔΟΕ), Χουάν Αντόνιο Σάμαρανκ, ανακοινώνει επίσημα την επιστροφή του σπουδαίου αθλητικού -και όχι μόνο- γεγονότος, των Ολυμπιακών Αγώνων, στην κοιτίδα του, στη χώρα που το γέννησε, στην Ελλάδα.¹⁰ Μια ανακοίνωση που σήμανε την απαρχή άπειρων συζητήσεων και πολιτικών αντιπαραθέσεων, τη συγγραφή πληθώρας άρθρων που τάχθηκαν υπέρ ή κατά της διεξαγωγής των αγώνων στην Ελλάδα και γενικώς τη σύγκρουση απόψεων στο εάν οι Ολυμπιακοί αυτοί Αγώνες -που τόσο πολύ θέλαμε ή δε θέλαμε- θα αποτελούσαν ένα άλμα προόδου για τη χώρα ή ένα πισωγύρισμα της ελληνικής οικονομίας σε χρόνια χαλεπά. Πριν, κατά την διάρκεια, αλλά και μετά την διεξαγωγή των 28^{ων} Ολυμπιακών Αγώνων, ποικίλες και αντικρουόμενες απόψεις διατυπώθηκαν γι' αυτό το τεράστιο πολιτικό, κοινωνικό, οικονομικό, αθλητικό και πολιτιστικό εγχείρημα. Στα πλαίσια όλης αυτής της κατάστασης, τέθηκαν -δικαιολογημένα- κάποια σημαντικά ερωτήματα. Ποια τα οικονομικά οφέλη που θα προέκυπταν από τους τυχόν επιτυχημένους Ολυμπιακούς Αγώνες της Αθήνας; Σε τι επίπεδα θα κυμαινόταν η τουριστική ανάπτυξη τα επόμενα χρόνια; Με ποιους τρόπους θα μπορούσαν να αξιοποιηθούν καλύτερα οι Ολυμπιακές εγκαταστάσεις στο άμεσο μέλλον;

2.1] Συνοπτική ανασκόπηση της ιστορίας των Σύγχρονων Ολυμπιακών Αγώνων και των επιπτώσεών τους.

Οι Ολυμπιακοί αγώνες δεν έχουν διεξαχθεί με τον ίδιο τρόπο παντού. Οι πόλεις που τους φιλοξένησαν βίωσαν τον αντίκτυπο των Αγώνων με τον ένα ή τον άλλο τρόπο. Ωστόσο, η έκταση και ο τύπος του αντίκτυπου υπήρξε διαφορετικός από πόλη σε πόλη [1]. Μια συνοπτική ανασκόπηση της ιστορίας των Ολυμπιακών Αγώνων των τελευταίων χρόνων, ίσως βοηθήσει στο σχηματισμό μιας γενικής εικόνας του ζητήματος της οικονομικής επίδρασής τους, στη διοργανώτρια κάθε φορά χώρα.¹¹

¹⁰ Βλ. Γωγουβίτης Ξένος, «Ολοκληρωμένα άρθρα : Ολυμπιάδα 2004 (η ανάθεση)», *KybernoGRAFOI.gr*, τχ.10, Οκτώβριος 1998, < <http://www.kybernoGRAFOI.gr/modules.php?name=News&file=article&sid=124>>, 9/5/2009.

¹¹ Βλ. Ίδρυμα Μείζονος Ελληνισμού, «Ολυμπιακοί Αγώνες και σύγχρονη ιστορία, οι επικερδείς Αγώνες : από το Λος Άντζελες στο Σίδνεϊ», <<http://www.fhw.gr/olympics/modern/gr/history/h302.html>>, 9/5/2009.

Το αρνητικό παράδειγμα των Ολυμπιακών Αγώνων του Μόντρεαλ (1976), δηλαδή το μεγάλο οικονομικό έλλειμμα της διοργάνωσης και η επιβάρυνση της οικονομίας του Καναδά με ένα χρέος για την αποπληρωμή του οποίου χρειάστηκε να περάσουν πολλά χρόνια, σήμανε το τέλος της αποκλειστικά κρατικής χρηματοδότησης των Αγώνων [1]. Σε μια εποχή οικονομικής ύφεσης, που διατηρήθηκε σε ολόκληρη τη δεκαετία του 1970, ελάχιστα κράτη είχαν τη βούληση και κυρίως τη δυνατότητα να δαπανήσουν τόσα χρήματα για ένα αναμφισβήτητο σημαντικό γεγονός, το οποίο ωστόσο διαρκούσε μόνο δύο εβδομάδες. Για τη Σοβιετική Ένωση και τις ΗΠΑ η ανάληψη της διοργάνωσης, το 1980 και το 1984 αντίστοιχα, αποσκοπούσε μάλλον σε μια επιβεβαίωση και ανάδειξη κύρους, παρά στον εμπλουτισμό του εθνικού ταμείου. Για την απόλυτα ελεγχόμενη από το κράτος οικονομία της Σοβιετικής Ένωσης, το συμβολικό κύρος που πρόσφερε η διοργάνωση των Αγώνων μπορούσε να δικαιολογήσει και συνεπώς να καλύψει την επιβάρυνση του κρατικού σχεδιασμού. Άλλωστε, τόσο η κατασκευή του Ολυμπιακού Χωριού, όσο και των αθλητικών εγκαταστάσεων εντάχθηκαν σε πολυετή προγράμματα οικονομικής ανάπτυξης [1]. Κάτι τέτοιο δεν ίσχυσε ασφαλώς και για τις Ηνωμένες Πολιτείες. Εδώ, η άρνηση της κυβέρνησης να καλύψει έστω και μέρος του κόστους των Αγώνων οδήγησε την οργανωτική επιτροπή στην αναζήτηση εναλλακτικών τρόπων χρηματοδότησης. Η μεταβολή στους όρους πώλησης των δικαιωμάτων τηλεοπτικής κάλυψης των Ολυμπιακών Αγώνων και η προσέλκυση μεγάλων εταιρειών ως χορηγών, όχι μόνο πέτυχε την κάλυψη του κόστους των Αγώνων, αλλά επέφερε και κέρδη πολλών εκατομμυρίων δολαρίων [1]. Παρά τις επικρίσεις που δέχθηκαν οι διοργανωτές για «εμπορευματοποίηση» των Ολυμπιακών Αγώνων, η οργανωτική επιτροπή οργάνωσε την πρώτη οικονομικά επιτυχημένη διοργάνωση μετά το 1932 (Λος Άντζελες). Έτσι, ο θεσμός των Ολυμπιακών αγώνων αρχίζει και πάλι να προσελκύει τα κράτη, τα οποία, κατά τις δύο προηγούμενες δεκαετίες, δίσταζαν να αναλάβουν, μια τέτοια διοργάνωση. Ενδεικτική της αλλαγμένης αυτής στάσης των κρατών, καθίσταται και η προοδευτική αύξηση του αριθμού των υποψηφιοτήτων για κάθε διοργάνωση των αγώνων, κατά τα χρόνια που ακολουθούν το Β' Παγκόσμιο πόλεμο μέχρι και το 1960: 6 υποψηφιοτήτες υποβάλλονται για το 1948, 9 για το 1952, 10 για το 1956 και 16 για το 1960. Έκτοτε, η πτώση του ενδιαφέροντος υπήρξε δραματική, καθώς ο αριθμός των υποψήφιων πόλεων μειώθηκε στις 4 για τους Αγώνες του 1964, του 1968 και του 1972, στις 3 για του 1976, στις 2 για του 1980, ενώ το 1984 το Λος Άντζελες ήταν η μοναδική πόλη που διεκδικούσε την ανάληψη των Ολυμπιακών Αγώνων [1]. Κάτι τέτοιο είχε να συμβεί από τους Αγώνες του 1932, όπου και πάλι μια μόνο πόλη, και μάλιστα η ίδια, διεκδικούσε τους Αγώνες: το Λος Άντζελες. Η οικονομική επιτυχία του Λος Άντζελες το 1984 συνδέεται με την ανάδειξη των Ολυμπιακών Αγώνων σε τηλεοπτικό υπερθέαμα. Την εξέλιξη αυτή, που γίνεται φανερή στις αρχές της δεκαετίας του 1980,

εκμεταλλεύτηκε με τον πλέον αποδοτικό οικονομικά τρόπο ο Peter Ueberroth, ένας αυτοδημιούργητος 40χρονος επιχειρηματίας που ενσάρκωνε το πρότυπο του «αμερικανικού ονείρου» και είχε επιλεγεί ως πρόεδρος της οργανωτικής επιτροπής. Τα έσοδα που αποκομούνται το 1984 από την τηλεοπτική εκμετάλλευση των Αγώνων, σε σχέση με τα συνολικά έσοδα της διοργάνωσης, είναι εντυπωσιακά αυξημένα, συγκρινόμενα και με τα αντίστοιχα προηγούμενων διοργανώσεων. Είναι χαρακτηριστικό ότι, ενώ το 1972 (Μόναχο) τα τηλεοπτικά έσοδα κάλυπταν λιγότερο από το 3% των συνολικών εσόδων, τα αντίστοιχα έσοδα των Αγώνων του 1984 κάλυπταν σχεδόν το 50% [1]. Το υπόλοιπο ποσοστό προερχόταν από τις πωλήσεις εισιτηρίων, γραμματοσήμων, νομισμάτων και άλλων αναμνηστικών, κυρίως όμως από τις χορηγίες. Βασική καινοτομία της οργανωτικής επιτροπής του 1984 ήταν η εξής: για πρώτη φορά τα τηλεοπτικά δίκτυα θα έπρεπε να πληρώσουν ένα χρηματικό ποσό έτσι ώστε, να είναι σε θέση να "διαπραγματευτούν" την αγορά των δικαιωμάτων μετάδοσης των Αγώνων. Η εξέλιξη αυτή ευνόησε τον πολλαπλασιασμό των χρηματικών εισροών. Ταυτόχρονα, η οργανωτική επιτροπή επαναπροσδιόρισε την πολιτική της προς τους χορηγούς, επιλέγοντας να προσελκύσει μεγάλες πολυεθνικές εταιρείες που, έναντι ενός πολύ υψηλότερου ποσού συγκριτικά με τις προηγούμενες διοργανώσεις, αποκτούσαν το δικαίωμα χρησιμοποίησης των Αγώνων για διαφήμιση των εταιρειών και των προϊόντων τους.

Παρ' ότι η πολιτική προσέλκυσης μεγάλων χορηγών επικρίθηκε σφοδρά, υιοθετήθηκε πολύ σύντομα από τη ΔΟΕ. Η τομή που επέφερε η διοργάνωση του 1984 γίνεται φανερή, αν συγκριθούν οι πηγές χρηματοδότησης των Ολυμπιακών Αγώνων πριν και μετά το Λος Άντζελες. Έτσι, ενώ από τη δεκαετία του 1930 οι κρατικοί πόροι συνιστούσαν τη σχεδόν αποκλειστική πηγή χρηματοδότησης των Αγώνων, μετά το 1984 ο ρόλος του κράτους περιορίζεται κατά πολύ ή εκμηδενίζεται. Έτσι, από τις τέσσερις διοργανώσεις που ακολούθησαν, στη μία (Ατλάντα 1996) τα έσοδα καλύφθηκαν, όπως και το 1984, δίχως κρατική χρηματοδότηση, ενώ στις άλλες τρεις η συμμετοχή του κράτους ήταν ισότιμη με εκείνη των άλλων πηγών. Η πλέον κερδοφόρα διοργάνωση υπήρξε εκείνη της Ατλάντας [1]. Θα πρέπει, ωστόσο, να σημειωθεί ότι τόσο στο Λος Άντζελες (1984), όσο και στην Ατλάντα (1996) η οργανωτική επιτροπή είχε τη δυνατότητα να διοργανώσει τους Αγώνες δίχως οικονομική βοήθεια από το κράτος των ΗΠΑ, καθώς στις δύο αυτές περιπτώσεις κατασκευάστηκαν ελάχιστα νέα έργα. Προτιμήθηκε δηλαδή να χρησιμοποιηθούν οι ήδη υπάρχουσες αθλητικές εγκαταστάσεις, ενώ για τη φιλοξενία των αθλητών, αντί της κατασκευής Ολυμπιακού Χωριού επιλέχθηκε η χρησιμοποίηση φοιτητικών εστιών.

Αντίθετα, στη Σεούλ (1988), τη Βαρκελώνη (1992) και το Σύννεϋ (2000) οι δαπάνες για την κατασκευή των αθλητικών και των άλλων υποδομών επαύξησε το κόστος των Αγώνων, καθιστώντας απαραίτητη την κρατική χρηματοδότηση [1]. Ωστόσο, στις περιπτώσεις αυτές στον υπολογισμό των εσόδων θα πρέπει να προσμετρηθεί το κέρδος για τη διοργανώτρια πόλη από τα έργα αυτά, καθώς και οι θετικές τους επιπτώσεις στην τοπική και εθνική οικονομία. Το παράδειγμα της Βαρκελώνης είναι το πλέον χαρακτηριστικό. Η κατασκευή των ολυμπιακών έργων σήμανε μια μεγάλης κλίμακας παρέμβαση στο «τοπίο» και τις λειτουργίες της πόλης, καθιστώντας την ταυτόχρονα, έναν από τους πλέον ελκυστικούς τόπους προσέλευσης τουριστών στην Ευρώπη [1]. Έτσι, η τετράχρονη επιβάρυνση της ισπανικής οικονομίας για την κάλυψη του κόστους των Αγώνων, μακροπρόθεσμα δεν αποτέλεσε μια οικονομικά λανθασμένη απόφαση. Στη συνέχεια παρουσιάζονται αναλυτικότερα κάποια σημαντικά οικονομικά στοιχεία των τριών πιο «επιτυχημένων» Ολυμπιάδων,¹² ώστε να γίνει μια πρόχειρη σύγκριση με το «Αθήνα 2004».

Οικονομικά Οφέλη από τους Αγώνες :

Σεούλ: Οι Αγώνες συνέβαλαν στην αύξηση της εμπιστοσύνης των καταναλωτών και των επιχειρηματιών για τη δυνατότητα της χώρας να επιτύχει υψηλούς αναπτυξιακούς ρυθμούς και σε μεγάλο βαθμό έδωσαν νέα κίνητρα για την εντατικοποίηση της προσπάθειας για ανάπτυξη. Συνέβαλαν επίσης στην ανάπτυξη βιομηχανιών που σχετίζονται με τον αθλητισμό, όπως βιομηχανίες αθλητικών ειδών, ψυχαγωγίας, αλλά και ηλεκτρονικής και τηλεπικοινωνιακής τεχνολογίας. Επίσης, βοήθησαν να γίνουν γνωστά παγκοσμίως διάφορα εμπορικά σήματα της χώρας, δίνοντας νέα ώθηση στις εξαγωγές. Οι εξαγωγές αγαθών αυξήθηκαν από 46,6 δισ.δολάρια το 1987 σε 162,6 δισ.δολάρια το 2002 [1].

Βαρκελώνη: Η συνολική θετική επίπτωση των Αγώνων στο ΑΕΠ της Ισπανίας, στη χρονική περίοδο 1986-1993 ανήλθε στα 16,6 δισ.δολάρια, ενώ πρόσθετες ευνοϊκές επιπτώσεις προέκυψαν και μετά το 1993 [1]. Ο βασικός στόχος της ανάληψης των Ολυμπιακών Αγώνων από τη Βαρκελώνη ήταν η προώθηση των αναγκαίων επενδύσεων σε έργα υποδομής, τα οποία ακόμη δε θα είχαν υλοποιηθεί, εάν δεν είχαν γίνει οι Αγώνες στην πόλη αυτή.

Σύννεϋ: Οι Ολυμπιακοί Αγώνες είχαν σημαντικές θετικές επιπτώσεις σε πολλούς τομείς οικονομικής δραστηριότητας, ιδιαίτερα στη δεκαετία του 1990, και εξακολουθούν να αποτελούν το βασικό εργαλείο για την προβολή του Σύννεϋ και της ευρύτερης περιοχής της Νότιας Νέας

¹² Βλ. Αθηναϊκό Πρακτορείο Ειδήσεων (ΑΠΕ), «Οφέλη από τη διοργάνωση των Ολυμπιακών Αγώνων», 10/10/2004, < <http://www.forthnet.gr/templates/newsPosting.aspx?p=109413>>, 9/5/2009.

Ουαλίας (NNO) και για την προσπάθεια ανάπτυξης τους σε τόπους τουριστικού προορισμού υψηλού επιπέδου [1]. Επίσης, οι επενδύσεις για τη διοργάνωση των Ολυμπιακών Αγώνων δημιούργησαν κύκλο εργασιών, ύψους 3 δισ.ευρώ, που προήλθε κατά 600 εκατ. δολάρια, από τη διενέργεια νέων επιχειρηματικών επενδύσεων, κατά 288 εκατ. δολάρια από τις νέες εργασίες μέσω της Έκθεσης Τεχνολογίας της Αυστραλίας και κατά 2 δισ.δολάρια από μετα-ολυμπιακά συμβόλαια εργασιών σε υποδομές και υπηρεσίες (PricewaterhouseCoopers, 2005). Από το σύνολο των εργασιών των 3 δισ. ευρώ, άνω των 500 εκατ. ευρώ αφορούν εργασίες, πωλήσεις και επενδύσεις που αναλήφθηκαν από επιχειρήσεις στην περιοχή της NNO. Οι Ολυμπιακοί Αγώνες συμβάλλουν στη δημιουργία εσόδων από διοργάνωση συνεδρίων ύψους 1,2 δισ.ευρώ στην περίοδο 1993-2007 και στην είσπραξη εσόδων από τον εισερχόμενο τουρισμό άνω των 6 δισ. ευρώ, μόνο το 2001. Σε διάφορες εκτιμήσεις για τα οφέλη των Ολυμπιακών Αγώνων, που είχαν γίνει πριν από τους Αγώνες, από το υπουργείο Οικονομικών της NNO, το καθαρό οικονομικό αποτέλεσμα των Αγώνων σε όρους αύξησης του ΑΕΠ ανερχόταν στα 6,2 - 8,0 δισ.δολάρια.

Τουριστική Ανάπτυξη :

Σεούλ: Κατά τη διάρκεια των δύο, πριν από τους Αγώνες, χρόνων, η προσφορά ξενοδοχειακών δωματίων αυξήθηκε κατά 64% , δημιουργώντας φόβους υπερπροσφοράς που αποδείχθηκαν υπερβολικοί, καθώς οι αφίξεις τουριστών αυξήθηκαν σημαντικά [1]. Το 1986, ο αριθμός των τουριστών ήταν μόνο 1,7 εκατ. , ενώ σήμερα η Κορέα δέχεται περίπου 5,5 εκατ. τουρίστες ετησίως.

Βαρκελώνη: Η προβολή της πόλης της Βαρκελώνης διεθνώς αυξήθηκε εντυπωσιακά κατά τη διάρκεια και μετά τους Ολυμπιακούς Αγώνες του 1992, ενώ η σημαντικά βελτιωμένη υποδομή της πόλης συνέβαλε στην ουσιαστική αύξηση της προσβασιμότητάς της από το εσωτερικό (της Ισπανίας) και από το εξωτερικό και στη βελτίωση της ανταγωνιστικότητας της Βαρκελώνης, ως πόλης τουριστικού προορισμού. Έτσι, ενώ πριν τους Ολυμπιακούς του 1992 η Βαρκελώνη εθεωρείτο η 16η πιο δημοφιλής τουριστική περιοχή της Ευρώπης, μέχρι το 1999 είχε φθάσει στην 3η θέση [1]. Ειδικότερα, το 2001 η ευρύτερη περιοχή της Καταλονίας προσέλκυσε 10,1 εκατ. ξένους τουρίστες, έναντι των 49,5 εκατ. τουριστών που αφίχθησαν από το εξωτερικό σε ολόκληρη την Ισπανία. Πάντως και η Βαρκελώνη δεν απέφυγε την εμφάνιση κάποιου πλεονάσματος ξενοδοχειακών δωματίων κατά τα πρώτα έτη μετά τους Ολυμπιακούς Αγώνες. Έτσι, το 1993 το επίπεδο πληρότητας έπεσε στα επίπεδα του 50%. Ωστόσο, στη συνέχεια, η κατάσταση άρχισε να βελτιώνεται σταδιακά, με αποτέλεσμα το 2003-2004 τα ξενοδοχεία της

πόλης να απολαμβάνουν επίπεδο πληρότητας που πλησιάζει το 75% [1]. Εκτιμάται ότι η πόλη φιλοξενεί περί τα 700 μεγάλα συνέδρια ετησίως.

Σύδνεϋ: Για να γίνει δυνατή η σύγκριση με την περίοδο πριν από τους Ολυμπιακούς Αγώνες, σημειώνεται ότι το 1998 επισκέφθηκαν την Αυστραλία 3,86 εκατ. τουρίστες και το 1999 4,1 εκατ. τουρίστες, εκ των οποίων τα 2,28 εκατ. κατευθύνθηκαν στην περιοχή του Σύδνεϋ (NNO). Το 2000, έτος των Ολυμπιακών Αγώνων, ο αριθμός των τουριστών της Αυστραλίας έφτασε τα 4,57 εκατ., για να υποχωρήσει ελαφρώς τα επόμενα δύο χρόνια, κυρίως λόγω των επιπτώσεων από την τρομοκρατική επίθεση της 11ης Σεπτεμβρίου 2001 και λόγω του πολέμου στο Αφγανιστάν και μετά στο Ιράκ και στη συνέχεια λόγω της εμφάνισης της επιδημίας SARS. Η θετική εικόνα που εκπέμπεται από τη διοργάνωση των Αγώνων θεωρείται ως ο βασικός παράγοντας που έδωσε τη δυνατότητα στο Σύδνεϋ να κερδίσει τη διοργάνωση 45 μεγάλων συνεδρίων, από το 2004 έως το 2012. Η πληρότητα των ξενοδοχείων του Σύδνεϋ ήταν πολύ υψηλή το 1999 και το 2000 (76,9% και 74,8% αντίστοιχα, κατά μέσο όρο, κατά τη διάρκεια του έτους), ενώ περιορίστηκε στο 67,7% το 2001, με ταυτόχρονη πτώση των τιμών κατά 5% περίπου [1].

Αξιοποίηση των Ολυμπιακών Εγκαταστάσεων μετά τους Αγώνες :

Σεούλ: Ο υφυπουργός Αθλητισμού της Κορέας υποστήριξε πρόσφατα στην Αθήνα ότι όλες οι Ολυμπιακές Εγκαταστάσεις της Κορέας χρησιμοποιούνται σήμερα στο 100%. Το υψηλό επίπεδο της αθλητικής υποδομής της Κορέας της επέτρεψε να διεκδικήσει και να επιτύχει την ανάληψη της διοργάνωσης του Παγκόσμιου Κυπέλλου στο Ποδόσφαιρο το 2002 (μαζί με την Ιαπωνία).

Βαρκελώνη: Η εκμετάλλευση των εγκαταστάσεων των λιγότερο δημοφιλών αθλημάτων έγινε με εκχώρησή τους στις αρμόδιες Ομοσπονδίες και με θέσπιση κρατικών επιχορηγήσεων, μέσω των προγραμμάτων του υπουργείου Αθλητισμού, που συνέβαλαν στη μύηση των κατοίκων της Βαρκελώνης και της ευρύτερης περιοχής στα νέα αθλήματα. Εκθέσεις και μεγάλες αθλητικές διοργανώσεις παγκόσμιου βεληνεκούς φιλοξενούνται έκτοτε στη Βαρκελώνη, προσφέροντας μια σημαντική πηγή εσόδων για τις επιχειρήσεις της περιοχής [1].

Σύδνεϋ: Μεγάλη δραστηριότητα σημειώνεται στη διοργάνωση συνεδρίων και άλλων εκδηλώσεων που συμβάλλουν στη διατήρηση της πληρότητας του ξενοδοχειακού συγκροτήματος του Ολυμπιακού Πάρκου του Σύδνεϋ, στο 76%. Αυτό οδηγεί τη μητρική εταιρεία του συγκροτήματος αυτού, να ολοκληρώνει το χτίσιμο ενός τρίτου ξενοδοχείου στο συγκρότημα

αυτό και να προγραμματίζει ένα τέταρτο ξενοδοχείο μέχρι το τέλος του 2006. Τα νέα ξενοδοχεία αποτελούν τα πρώτα στάδια του νέου 20ετούς επενδυτικού προγράμματος των 5 δισ. δολαρίων σε εμπορικές και οικοδομικές κατασκευές στο Ολυμπιακό Πάρκο [1]. Μέσα σε αυτά τα 20 χρόνια, το μέγεθος της οικοδομημένης επιφάνειας στην περιοχή του Ολυμπιακού Πάρκου αναμένεται να έχει αυξηθεί είκοσι φορές σε σχέση με το σημερινό επίπεδο του και το αναβαθμισμένο Ολυμπιακό Πάρκο αναμένεται να έχει πληθυσμό κατά τη διάρκεια της ημέρας 70.000 ατόμων (27.000 μόνιμους κάτοικους, 17.000 εργαζομένους, 9.500 φοιτητές πανεπιστημίου και ένα μεταβαλλόμενο αριθμό επισκεπτών). Τα μεγαλύτερα προβλήματα παρουσιάζονται στις εγκαταστάσεις, που είναι διασκορπισμένες στα Προάστια του Σύδνεϋ. Για παράδειγμα, οι εγκαταστάσεις για το άθλημα της Ορεινής Ποδηλασίας (κόστους 360 χιλ δολαρίων περίπου) έκλεισαν λόγω ελάχιστης χρήσης τους. Επίσης, το Ολυμπιακό Κέντρο Σκοποβολής και το Ολυμπιακό Κέντρο Ιππασίας προσελκύουν ελάχιστους επισκέπτες και η συνέχιση της λειτουργίας τους τίθεται υπό συζήτηση.

2.2] Ιδιαίτερα χαρακτηριστικά γνωρίσματα της Ελλάδας, ως διοργανώτριας χώρας.

Οι Ολυμπιακοί Αγώνες μπορούν και έχουν άμεση επίδραση στα βασικά οικονομικά μεγέθη μίας οικονομίας και ειδικότερα μίας περιφερειακής οικονομίας όπως η ελληνική, τόσο σε παρόντα χρόνο (περίοδος προετοιμασίας και αγωνιστική περίοδος), όσο και σε μέλλοντα χρόνο (περίοδος μετά τη λήξη των αγώνων) [2]. Η Ελλάδα έχει ορισμένα ιδιαίτερα χαρακτηριστικά, τα οποία επιτάσσουν - πολύ περισσότερο από ό,τι θα ίσχυε σε οποιαδήποτε άλλη χώρα - να εξετάσουμε προσεκτικά την οικονομική άνοδο ή ύφεση της χώρας από τη διοργάνωση των Ολυμπιακών Αγώνων.¹³

Ένα πρώτο χαρακτηριστικό είναι ότι η Ελλάδα αποτελεί, κατά τη σύγχρονη εποχή, την πιο μικρή χώρα που έχει αναλάβει τη διοργάνωση των Ολυμπιακών Αγώνων. Η προηγούμενη μικρή χώρα και κατ' επέκταση μικρή οικονομία, η οποία ανέλαβε τους Ολυμπιακούς Αγώνες, ήταν το 1952 η Φινλανδία, όταν βέβαια ήταν πάρα πολύ διαφορετική η κατάσταση και των αγώνων και των οικονομικών τους απαιτήσεων. Το δεύτερο χαρακτηριστικό είναι ότι η διοργάνωση των Ολυμπιακών Αγώνων, όλα τα έργα, οι υποδομές και οι υπηρεσίες, χρηματοδοτήθηκαν σχεδόν εξολοκλήρου από τον κρατικό προϋπολογισμό. Η προηγούμενη περίπτωση τόσο εκτεταμένης κρατικής συμμετοχής και ευθύνης ήταν στη Σοβιετική Ένωση. Το τρίτο γνώρισμα είναι ότι οι Ολυμπιακοί Αγώνες στην Ελλάδα χαρακτηρίστηκαν από μια ένταση και πύκνωση της χρηματοδότησης, η οποία πραγματοποιήθηκε πρακτικά σε δυο χρόνια, το 2003 και το 2004. Το τέταρτο χαρακτηριστικό είναι ότι στη διάρκεια της προετοιμασίας των Ολυμπιακών Αγώνων προστέθηκαν σημαντικά νέα έργα, όπως για παράδειγμα το έργο της ασφάλειας των Ολυμπιακών Αγώνων, το οποίο έγινε καθοριστικότερης σημασίας, μετά τα τραγικά γεγονότα της 11ης Σεπτεμβρίου στη Νέα Υόρκη. Και το πέμπτο είναι ότι η διοργάνωση και πολύ περισσότερο η χρηματοδότηση των Ολυμπιακών Αγώνων συνέπεσε με μια συγκυρία στη χώρα μας, όπου υπήρχαν πάρα πολλές απαιτήσεις για ανάπτυξη της περιφέρειας, για χρηματοδότηση υποδομών, για μια επιτάχυνση όλων των δραστηριοτήτων.

¹³ Βλ. Χριστοδουλάκης Νίκος, «Η ελληνική οικονομία μετά τους Ολυμπιακούς Αγώνες», ομιλία σε συνέδριο της International Herald Tribune, 6/11/2004, Αθήνα, <<http://www.christodoulakis.gr/%289909C513EF7D8FB4372FAAB3986228FDB6283DFD8CFC47F5%29/ecportal.asp?id=4480&nt=18&lang=1>>, 10/5/2009.

2.3] Οικονομικά στοιχεία της ελληνικής οικονομίας πριν, κατά και μετά το «Αθήνα 2004».

Στο σημείο αυτό παρουσιάζονται ορισμένα οικονομικά στοιχεία της Ελλάδας για το σχηματισμό μιας πιο εποπτικής εικόνας του τι εστί «Ολυμπιακοί Αγώνες-Αθήνα 2004». Η βασικότερη χρηματοδότηση των Ολυμπιακών Αγώνων γίνεται το 2003 και το 2004, όπου συνολικά διατίθενται, μαζί με το 2002, 4,5 δισ. ευρώ για τις δαπάνες κατασκευής ολυμπιακών εγκαταστάσεων, αλλά φυσικά και των κατάλληλων υποδομών που απαιτούνται για την εξυπηρέτηση ενός τέτοιου εγχειρήματος, το οποίο δεν αφορά μόνο το Λεκανοπέδιο της Αττικής, αλλά ολόκληρη τη χώρα και κάθε περιφέρεια [2]. Αν στο ανωτέρω ποσό συνυπολογίσουμε και τους πόρους ύψους περίπου 2 δισ. ευρώ, οι οποίοι διατέθηκαν από τον προϋπολογισμό της Οργανωτικής Επιτροπής για την προετοιμασία των αγώνων, το συνολικό ποσό επένδυσης σε έργα, εξοπλισμό, υπηρεσίες και νέες τεχνολογίες, φθάνει περίπου τα 6,5 δισ. ευρώ, δηλαδή αποτελεί το 4,5% του ετήσιου ΑΕΠ [2]. Ένα ύψος επένδυσης το οποίο είναι εξαιρετικά μεγάλο για το μέγεθος της ελληνικής οικονομίας. Τα έργα, τα οποία αφορούν τους Ολυμπιακούς Αγώνες, αριθμούν περίπου 480 προγράμματα, 1.760 υποπρογράμματα και περιλαμβάνουν μια πολύ μεγάλη γκάμα δραστηριοτήτων και υποδομών για μεταφορές, ασφάλεια, μέσα μαζικής ενημέρωσης, υγεία, πυροσβεστικές υπηρεσίες και άλλα πολλά.¹⁴ Γίνεται, πλέον, αντιληπτή η βελτίωση της εικόνας της χώρας, όχι μόνο της Αθήνας, αλλά και ολόκληρης της Περιφέρειας, με αφορμή τα έργα, τις υπηρεσίες, τις κύριες και υποστηρικτικές υποδομές που δημιουργούνται για τους Ολυμπιακούς Αγώνες. Την εποχή που η Ελλάδα ανέλαβε τη διοργάνωση των Ολυμπιακών Αγώνων, το βιοτικό επίπεδο των κατοίκων, υπολογισμένο ως κατά κεφαλή ΑΕΠ με βάση την αγοραστική δύναμη, ισοδυναμούσε περίπου με το 65% του μέσου όρου της Ευρωπαϊκής Ένωσης. Η Ελλάδα πέτυχε το 2001 πραγματικό ρυθμό ανάπτυξης 4%, ξεπερνώντας τον ευρωπαϊκό μέσο όρο (1,6%) κατά 2,4%. Το 2002, η Ελλάδα σημείωσε ρυθμό ανάπτυξης 3,8% έναντι 1,1% του ευρωπαϊκού μέσου όρου, το 2003 4,1% έναντι μηδενικής ανάπτυξης για την Ευρωπαϊκή Ένωση, το 2004 4,2% και το 2005 3,4%. Βασικές κινητήριες δυνάμεις της ελληνικής ανάπτυξης, η οποία βασίστηκε σε μεγάλο βαθμό στα ολυμπιακά έργα, ήταν η εγχώρια κατανάλωση και οι επενδύσεις, οι οποίες παρέμειναν σε υψηλά επίπεδα (7,7% το 2003) έναντι των πολύ χαμηλών αντίστοιχων ποσοστών της ευρωζώνης (-1,8% για το 2003) [2].

¹⁴ Βλ. Χριστοδουλάκης Νίκος, «Η ελληνική οικονομία μετά τους Ολυμπιακούς Αγώνες», ομιλία σε συνέδριο της International Herald Tribune, 6/11/2004, Αθήνα, <<http://www.christodoulakis.gr/%289909C513EF7D8FB4372FAAB3986228FDB6283DFD8CFC47F5%29/ecportal.asp?id=4480&nt=18&lang=1>>, 10/5/2009.

2.4] Αναπτυξιακά έργα στην Ελλάδα κατά την περίοδο των Ολυμπιακών Αγώνων, στους τομείς των Δημοσίων Έργων και των Μεταφορών.

Τέλος, αξίζει να αναφερθεί και η ανάπτυξη που παρουσιάστηκε στην Ελλάδα κατά την περίοδο των Ολυμπιακών Αγώνων στους τομείς των Δημοσίων Έργων και των Μεταφορών. Πολλά έργα απ' αυτά που επισυνάπτονται στη συνέχεια, είτε ολοκληρώθηκαν, είτε συνεχίζονται μέχρι σήμερα.¹⁵

Δημόσια Έργα

Η ανάληψη των Ολυμπιακών Αγώνων αντιμετωπίστηκε, σε συνδυασμό με το Γ' Κοινωνικό Πλαίσιο Στήριξης, από τη χώρα ως ο μοχλός ανάπτυξης με τον οποίο τέθηκαν τα γερά θεμέλια για την αντιμετώπιση των προκλήσεων της νέας χιλιετηρίδας, δημιουργώντας τις απαραίτητες υποδομές. Στα έργα αυτά συγκαταλέγονται:

Ολυμπιακό Χωριό :

Το Ολυμπιακό Χωριό κατασκευάστηκε κάτω από την Πάρνηθα, στο Δήμο Αχαρνών, σε συνολική έκταση 1240 στρεμμάτων. Το Ολυμπιακό Χωριό φιλοξένησε 17.500 αθλητές των Ολυμπιακών Αγώνων «Αθήνα 2004» και μέχρι 6.000 αθλητές των Παραολυμπιακών, σε κτίρια ειδικά εξοπλισμένα για ανθρώπους με ειδικές ανάγκες. Μετά τους Ολυμπιακούς Αγώνες, το Ολυμπιακό Χωριό λειτουργεί ως πρότυπος οικισμός για τους δικαιούχους του Οργανισμού Εργατικής Κατοικίας.

Χωριά Δημοσιογράφων :

Σε επτά σημεία της Αττικής δημιουργήθηκαν τα Χωριά Τύπου που φιλοξένησαν, κατά τη διάρκεια των Αγώνων, τους 11.290 εκπροσώπους των ΜΜΕ. Από αυτά, τα τέσσερα ήταν νέοι οικισμοί, δυναμικότητας 1.000-1.600 κλινών, που στη μετα-ολυμπιακή χρήση κάλυψαν συγκεκριμένες στεγαστικές ανάγκες κατοίκων και υπηρεσιών. Σε κάθε χωριό δημιουργήθηκαν καταστήματα και χώροι εστίασης, αναψυχής και ψυχαγωγίας.

Εγκαταστάσεις Ελληνικού :

Στο αεροδρόμιο του Ελληνικού, σε έκταση 1.900 στρεμμάτων κατασκευάστηκαν Ολυμπιακές Εγκαταστάσεις (οι περισσότερες προσωρινές) για τη διεξαγωγή των αθλημάτων

¹⁵ Βλ. media info 2004, «Ολυμπιακά θέματα : Δημόσια έργα - Μεταφορές»,
< <http://www.mediainfo2004.gr/cgi-bin/hweb?-A=46&-V=olympicissues&-w=>>, 11/5/2009.

του μπίτζμπολ, του σόφτμπολ, του χόκεϋ, της καλαθοσφαίρισης και της χειροσφαίρισης. Το έργο για το άθλημα του κανό-σλάλομ, περιλαμβάνει τεχνητή λίμνη 25.000 τ.μ. με κανάλια και γέφυρα και αποτελεί βασικό στοιχείο αναψυχής και αισθητικής αναβάθμισης κατά την μετα-ολυμπιακή χρήση.

Ενοποίηση αρχαιολογικών χώρων :

Το πρόγραμμα ενοποίησης των αρχαιολογικών χώρων ανέλαβε η «Ενοποίηση Αρχαιολογικών Χώρων Αθήνας ΑΕ» που συστήθηκε από τα Υπουργεία ΠΕΧΩΔΕ και Πολιτισμού. Η υλοποίηση των πολυάριθμων έργων που συνέθεσαν το Πρόγραμμα της Ενοποίησης, αλλά και των άλλων παρεμβάσεων που εντάχθηκαν στη γενικότερη αναβάθμιση του ιστορικού κέντρου της πόλης, είχε ως στόχο την ανάπτυξη και την ιστορική ανάκτηση του κέντρου της Αθήνας.

Από τα υπόλοιπα δημόσια έργα, ενδεικτικά αναφέρονται τα εξής:

- Κέντρο Κωπηλασίας
- Κέντρο πολλαπλών χρήσεων
- Ολυμπιακό κέντρο Beach Volley Φαλήρου
- Ολυμπιακό κέντρο Ιστιοπλοΐας
- Ολυμπιακό κέντρο Ιππασίας
- Ολυμπιακό κέντρο Σκοποβολής
- Έργα σε υφιστάμενες αθλητικές εγκαταστάσεις
- Λιμενικές εγκαταστάσεις
- Τουριστικά έργα
- Προγράμματα για την αρχαία Ολυμπία

Μέσα Μαζικής Μεταφοράς

Οι μεταφορές κατά τη διάρκεια των Αγώνων περιλάμβαναν την καθημερινή μεταφορά χιλιάδων αθλητών, συνοδών ομάδων, Κριτών-Διαιτητών, Εκπροσώπων των ΜΜΕ, θεατών, Επισήμων, μελών της Ολυμπιακής Οικογένειας, χορηγών, εργαζομένων και εθελοντών στους προορισμούς τους, καθορίζοντας κατά πολύ την επιτυχία των Αγώνων. Ο ολοκληρωμένος σχεδιασμός του Συστήματος των Ολυμπιακών Μεταφορών και Υποδομών προέβλεπε σημαντικά έργα από το Υπουργείο Μεταφορών και Επικοινωνιών και το Υ.ΠΕ.ΧΩ.Δ.Ε. Τα έργα αυτά επέφεραν μεγάλες αλλαγές στην υποδομή των μεταφορών στην πρωτεύουσα, βελτιώνοντας τις κυκλοφοριακές συνθήκες και τις μετακινήσεις, ενώ ταυτόχρονα καθιέρωσαν

νέες συνήθειες μεταφοράς για τον πληθυσμό της πόλης. Πραγματοποιήθηκαν σημαντικές παρεμβάσεις τόσο από πλευράς κατασκευαστικών έργων (Προαστιακός, Μετρό, ΗΣΑΠ, Τραμ), όσο και από πλευράς προμηθειών (ανανέωση στόλου ΕΘΕΛ, ΗΛΠΑΠ). Ανάμεσα στα σημαντικότερα Συγκοινωνιακά Έργα που αφορούν την ελληνική πρωτεύουσα συγκαταλέγονται:

- Ολυμπιακός Δακτύλιος
- Οδικά έργα υποστήριξης - 120 χιλιόμετρα νέων, σύγχρονων δρόμων και 90 χιλιόμετρα αναβαθμισμένων δρόμων - 40 γέφυρες ανισόπεδης διάβασης
- Αττικό Μετρό - 7,7 χιλιόμετρα νέων γραμμών του Μετρό
- Εκσυγχρονισμός Σταθμών - Συρμών ΗΣΑΠ
- Προαστιακός Σιδηρόδρομος - 32 χιλιόμετρα Προαστιακού Σιδηρόδρομου (καταλήγει στο Διεθνές Αεροδρόμιο Ελ. Βενιζέλος)
- Δίκτυο Τραμ - 23,7 χιλιόμετρα δικτύου του Τραμ
- Νέος Αερολιμένας - Αεροδρόμιο Ελ. Βενιζέλος
- Αττική Οδός
- Έργα διαχείρισης και ελέγχου Κυκλοφορίας και Νέο υπερσύγχρονο Κέντρο Διαχείρισης Κυκλοφορίας
- Σύγχρονοι σιδηροδρομικοί σταθμοί
- Χώροι Στάθμευσης σε κεντρικά σημεία, με νέα συστήματα διαχείρισης Συγκοινωνιακά έργα στους ολυμπιακούς πόλους

Όλα τα οικονομικά στοιχεία και τα αναπτυξιακά έργα υποδομής που προαναφέρθηκαν στα πλαίσια βελτίωσης της διοργανώτριας πόλης των Ολυμπιακών Αγώνων του 2004, της Αθήνας, παρουσιάζονται ενδεικτικά στους δείκτες του ερευνητικού προγράμματος «Συνολικές Επιπτώσεις Ολυμπιακών Αγώνων» (πρόγραμμα ΟGGI), το οποίο αναλύεται λεπτομερώς στην επόμενη ενότητα.

3] Πρόγραμμα «Συνολικές Επιπτώσεις Ολυμπιακών Αγώνων» (OGGI).

3.1] Τι είναι το πρόγραμμα OGGI ; - Γενικά.

Η φιλοξενία των Ολυμπιακών Αγώνων έχει σημαντικό αντίκτυπο στη διοργανώτρια κάθε φορά πόλη και την κοινότητά της. Από τα έργα υποδομής, όπως οι αθλητικές εγκαταστάσεις και τα έργα μεταφορών, μέχρι τη βελτίωση της εικόνας της διοργανώτριας πόλης, το εγχείρημα «Ολυμπιακοί Αγώνες» ενεργεί ως καταλύτης, που αφήνει στο διάβα του ένα μόνιμο σημάδι στην πόλη, στη χώρα που τους φιλοξενεί και τους ανθρώπους της, μέσω της οικονομικής, αστικής, κοινωνικής ή ιστορικής επιρροής του. Τα ερωτήματα μέχρι ποιο βαθμό επηρεάζουν οι Αγώνες μια χώρα και με τι αποτελέσματα (αρνητικά ή θετικά) αποτελούσαν πάντα ζήτημα ποικίλων συζητήσεων. Για να απαντηθούν -στο βαθμό του εφικτού- τα ερωτήματα αυτά και για να βοηθηθούν οι εκάστοτε οργανωτικές επιτροπές Ολυμπιακών Αγώνων των πόλεων που φιλοξένησαν ή θα φιλοξενήσουν τους Αγώνες, η Διεθνής Ολυμπιακή Επιτροπή (ΔΟΕ) προώθησε τα τελευταία χρόνια το πρόγραμμα «Συνολικές Επιπτώσεις Ολυμπιακών Αγώνων» (Olympic Games Global Impact - OGGI).¹⁶ Το OGGI είναι ένα ερευνητικό έργο, που αποσκοπεί στην καλύτερη αξιολόγηση, διαχείριση, εποπτεία του κόστους και των αποτελεσμάτων των Ολυμπιακών Αγώνων από την εκάστοτε Οργανωτική Επιτροπή της διοργανώτριας πόλης-χώρας. Με τον τρόπο αυτό τίθεται ως μελλοντικός στόχος του προγράμματος OGGI η ανάδειξη της επίδρασης των Ολυμπιακών Αγώνων στη διοργανώτρια κάθε φορά πόλη, ώστε να βοηθηθούν οι μελλοντικοί διοργανωτές των Αγώνων.

Το 2001 η ΔΟΕ ξεκίνησε τη συνεργασία της με τοπικά πανεπιστήμια της πόλης που διοργανώνει τους επικείμενους κάθε φορά Ολυμπιακούς Αγώνες, προκειμένου να αναπτύξει ένα συγκεκριμένο μεθοδολογικό πλαίσιο και να επιλέξει ένα σύνολο μετρήσιμων ερευνητικών δεικτών, για την όσο το δυνατόν καλύτερη διερεύνηση και αξιολόγηση της συνολικής επίπτωσης της διεξαγωγής των Ολυμπιακών Αγώνων στις χώρες που τους διοργανώνουν σε τρεις συγκεκριμένους τομείς: Οικονομία, Κοινωνία και Περιβάλλον. Έτσι, τα πανεπιστήμια των πόλεων αναλαμβάνουν να φέρουν εις πέρας την αποτίμηση των επιπτώσεων των Ολυμπιακών Αγώνων της χώρας τους στους τρεις συγκεκριμένους τομείς, μέσα στα πλαίσια του προγράμματος OGGI. Η ΔΟΕ, με βάση την αποτίμηση αυτή, ενσωματώνει κάθε φορά τα καινούρια στοιχεία-συμπεράσματα που προκύπτουν στην ατζέντα της και καθοδηγεί καλύτερα

¹⁶ Βλ. «What is the Olympic Games Global Impact?», Olympic Review - Official Publication of the Olympic Movement, June 2006, <http://multimedia.olympic.org/pdf/en_report_1077.pdf>, 16/5/2009.

τους εκάστοτε υποψήφιους διοργανωτές Ολυμπιακών Αγώνων, υπό το πρίσμα συμβουλευτικών οδηγιών.

Η Ολυμπιακή Επιτροπή της κάθε διοργανώτριας χώρας υποχρεούται να παραδώσει έτοιμη τη μελέτη της από το πρόγραμμα ΟGGI μέσα στο χρονικό ορίζοντα των 11 χρόνων από την ημέρα της επίσημης θέσης υποψηφιότητας για ανάληψη των Ολυμπιακών Αγώνων από την Επιτροπή. Καθ' όλο το χρονικό διάστημα των 11 χρόνων η Επιτροπή οφείλει να συλλέξει τα κατάλληλα στοιχεία που πρέπει και να τα καταγράψει σε 4 εκθέσεις-αναφορές, παραδίδοντας εν τέλει τις αναφορές αυτές στη ΔΟΕ σε ισάριθμα τακτά διαστήματα, τα οποία προκαθορίζονται [1].

3.2] Το πρόγραμμα OGGI του «Αθήνα 2004».

Στην περίπτωση της Αθήνας γίνεται για πρώτη φορά συνολική αποτίμηση της επίδρασης των Ολυμπιακών Αγώνων με επιστημονικές μεθόδους. Το πρόγραμμα OGGI για τους Ολυμπιακούς Αγώνες «Αθήνα 2004» εκπονήθηκε για λογαριασμό της ΔΟΕ και πιο συγκεκριμένα, της AISTS (Academie Internationale des Sciences et Techniques du Sport), που εδρεύει στη Λωζάνη.¹⁷

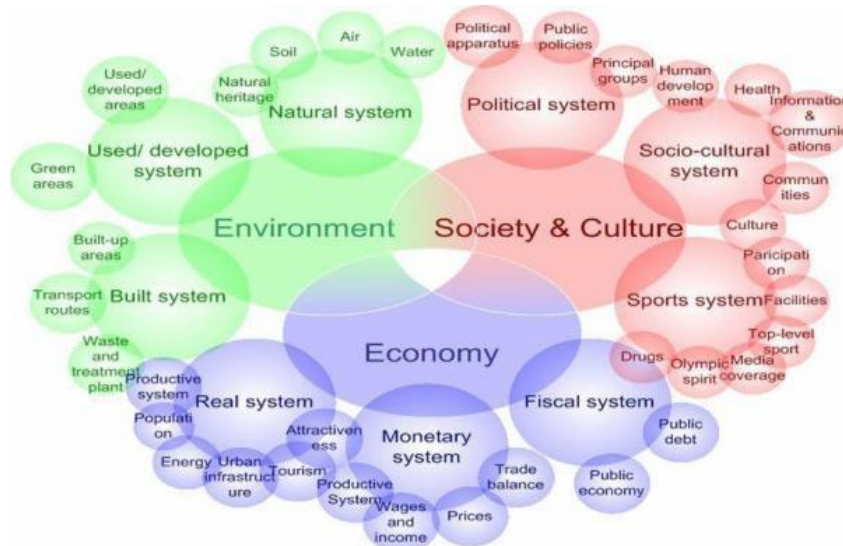


The AISTS Takes a Trans-disciplinary Approach

Σχήμα 1.2.2.α : Η γενική φιλοσοφία της AISTS (<http://www.aists.org/en/aists/the-aists-thinking-beyond-the-score.html>).

Το πρόγραμμα ξεκίνησε το Σεπτέμβριο του 2004 και ολοκληρώθηκε το Μάρτιο του 2007. Στόχος του προγράμματος ήταν ο υπολογισμός των σημαντικότερων δεικτών επίδρασης των Ολυμπιακών Αγώνων, με βάση στοιχεία που αφορούν την περίοδο 1996 - 2007. Οι δείκτες αυτοί χρησιμοποιήθηκαν για τη συνολική εκτίμηση των επιπτώσεων των Ολυμπιακών Αγώνων του 2004 στους τρεις βασικούς τομείς της αειφόρου ανάπτυξης της Ελλάδας: Περιβάλλον, Οικονομία και Κοινωνία / Πολιτισμός.

¹⁷ Βλ. OGGI Athens 2004, «Athens 2004, Olympic Games Global Impact», <http://147.102.46.88/OGGI/front-page/view?set_language=el>, 16/5/2009.



Σχήμα 1.2.2.β : Το τρίπτυχο Περιβάλλον-Οικονομία-Κοινωνία/Πολιτισμός, το οποίο αξιολογείται και μελετάται από το πρόγραμμα OGGI (<http://147.102.46.88/OGGI/methodos-analysis>).

3.3] Μέθοδος Ανάλυσης του OGGI.

«Η AISTS καθορίζει την επίδραση (impact) ως αλληλοσυσχέτιση μεταξύ του Ολυμπιακού γεγονότος και του περιβάλλοντος, στο οποίο έλαβε χώρα το γεγονός αυτό. Ο απαιτούμενος υπολογισμός των επιπτώσεων πραγματοποιείται στην ιδεώδη περίπτωση με τη συγκριτική αντιπαράθεση της εξελικτικής πορείας του περιβάλλοντος με και χωρίς το υπό μελέτη γεγονός. Καθώς όμως ένας τέτοιος υπολογισμός είναι ουτοπικός και επομένως αδύνατος, αναδύονται δύο διαφορετικές προσεγγίσεις. Η πρώτη εστιάζει στο ίδιο το γεγονός και αποπειράται να αναδείξει τις επιπτώσεις μετρώντας τη διεγερτική του επιρροή κατά τη διάρκεια της περιόδου προετοιμασίας και διεξαγωγής των Αγώνων, ενώ η δεύτερη υιοθετεί αντιδιαμετρική οπτική γωνία, εστιάζοντας στο περιβάλλον φιλοξενίας των Αγώνων και στην εξελικτική του διαδικασία.

Υιοθετώντας αυτή τη διφυή προσέγγιση, η AISTS προτείνει την εισαγωγή δύο ομάδων δεικτών, η πρώτη εκ των οποίων θα απεικονίζει το γεγονός, ενώ η δεύτερη θα σκιαγραφεί παράλληλα το περιβάλλον. Η σφαιρική επίδραση ανακλύπτει εν τέλει με έμμεσο τρόπο, ως η περιοχή στην οποία διαδραματίζεται η περίπλοκη αλληλοσυσχέτιση ανάμεσα στο γεγονός των Αγώνων και το περιβάλλον συντέλεσής του.

Η επακόλουθη οροθέτηση της περιοχής αυτής καθορίζει τα αναγκαία όρια του πλαισίου της σφαιρικής επίδρασης των Αγώνων. Η χαρτογράφηση της δεδομένης πλέον περιοχής, με τις κατάλληλες χρωματικές κλίμακες για την απεικόνιση των παραμέτρων του χώρου, του χρόνου και του επιπέδου λεπτομέρειας, επιλέχθηκε να υλοποιηθεί με τη χρήση κατάλληλα επιλεγμένων δεικτών. Ο χαρακτηρισμός των χρωματικών παραμέτρων περιγράφεται στις ακόλουθες γραμμές. Ωστόσο πρέπει εξ' αρχής να αναφερθεί ότι, η ικανότητα αναγνώρισης των επιπτώσεων με το συγκεκριμένα επιλεγμένο εργαλείο είναι εκ φύσεως περιορισμένη.

Ο πρώτος βαθμός ελευθερίας που χαρακτηρίζει τη διαδικασία προσδιορισμού των δεικτών είναι ο χώρος. Το πολυδιάστατο περιβάλλον πραγματοποίησης των Αγώνων αποκαλύπτει την αναγκαιότητα υιοθέτησης μιας συμβατά ποικιλόμορφης χωρικής διασποράς. Η μελέτη αναφέρεται σε τέσσερις διαφορετικές γεωγραφικές οντότητες (τοπικό, περιφερειακό, εθνικό και παγκόσμιο επίπεδο). Ωστόσο, τα τοπικά και περιφερειακά όρια συμπίπτουν και ταυτίζονται με την ευρύτερη περιοχή της Αττικής.

Ο χρόνος συνιστά το δεύτερο βαθμό ελευθερίας του σχεδιασμού των δεικτών. Οι Ολυμπιακοί Αγώνες τείνουν να αποτελέσουν αναπόσπαστο τμήμα της ιστορίας της

φιλοξενούσας περιοχής. Κατά συνέπεια ο αυστηρός προσδιορισμός της χρονικής έκτασης των επιδράσεων τους είναι πιθανά μη εφικτός. Ως υποκειμενικά όρια της έκτασης αυτής επιλέχθηκαν η επίσημη ανακοίνωση της υποψηφιότητας της πόλης (αρχική κατάσταση) και η συμπλήρωση δύο ετών από την επιχείρηση των Αγώνων (τελική κατάσταση). Το μεθοδολογικό πλαίσιο της AISTS διακρίνει τις εξής χαρακτηριστικές βαθμίδες:

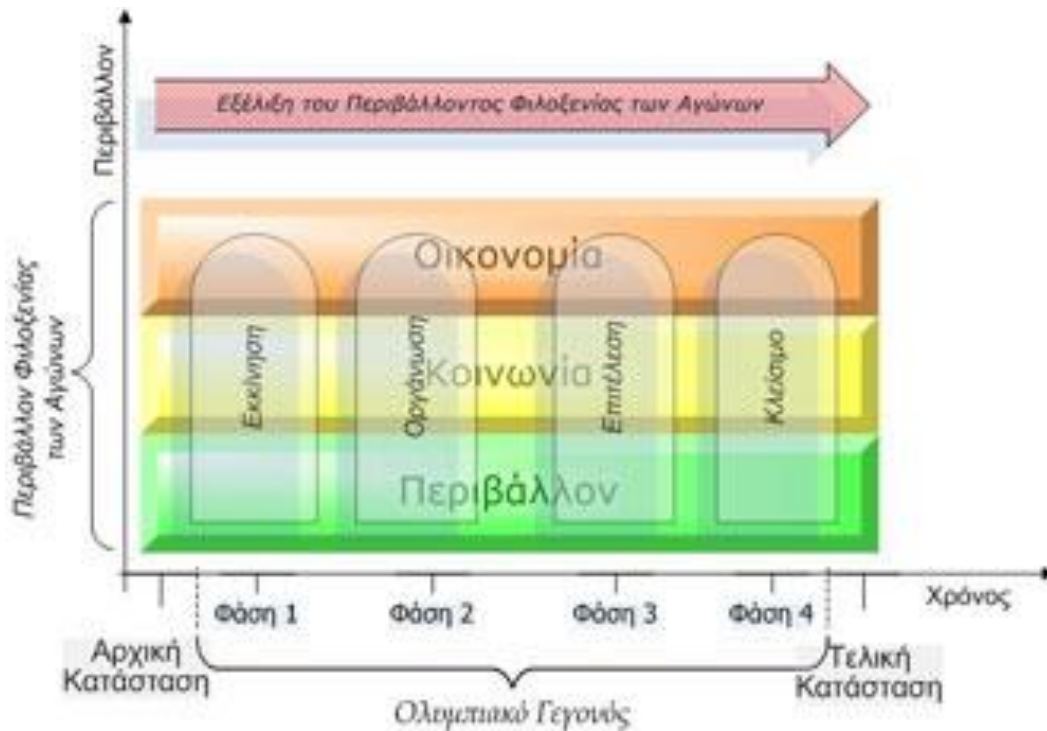
- Εκκίνηση (Conception, Φάση 1): η περίοδος μεταξύ της κατάθεσης της υποψηφιότητας και της εκλογής ως οικοδέσποινα πόλη
- Οργάνωση (Organization, Φάση 2): η περίοδος μεταξύ εκλογής και έναρξης των Αγώνων
- Επιτέλεση (Staging, Φάση 3): η περίοδος μεταξύ έναρξης λειτουργίας του Ολυμπιακού χωριού και περάτωσης λειτουργίας του Παραολυμπιακού χωριού
- Κλείσιμο (Closure, Φάση 4): περίοδος δύο ετών, επακόλουθη των Αγώνων». ¹⁸

Οι σφαίρες δραστηριοτήτων, με βάση τις οποίες επιχειρείται μέσα από το πρόγραμμα OGGI να περιγραφεί η ανάπτυξη του περιβάλλοντος ενός γεγονότος, αποτελούν τις τρεις βασικές διαστάσεις της αειφόρου ανάπτυξης της Ελλάδας

- Οικονομική διάσταση (Economic, Ec)
- Κοινωνική διάσταση (Social, So)
- Περιβαλλοντική διάσταση (Environmental, En)

Εν κατακλείδι, η σφαιρική επίδραση των Ολυμπιακών Αγώνων προσδιορίζεται λεπτομερώς μέσα από τέσσερις γεωγραφικές οντότητες, τέσσερις χρονικές φάσεις και τρεις σφαίρες δραστηριοτήτων, όπως φαίνεται και στο πιο κάτω σχήμα.

¹⁸ Βλ. Τζιραλής Γιώργος, OGGI Athens 2004, «Athens 2004, Olympic Games Global Impact - Μέθοδος Ανάλυσης», <<http://147.102.46.88/OGGI/methodos-analysis>>, 17/5/2009.



Διάγραμμα 1.2.3.α : Το πλαίσιο ανάλυσης της σφαιρικής επίδρασης των Ολυμπιακών Αγώνων. (<http://147.102.46.88/OGGI/methodos-analysis>)

3.4] Δείκτες Ανάλυσης.

Οι δείκτες ενός προγράμματος OGGI είναι το αποτέλεσμα της μελέτης που πραγματοποιήθηκε από την AISTS και τη ΔΟΕ.¹⁹ Έτσι, οι δείκτες ομαδοποιούνται με δύο τρόπους:

- δείκτες πλαισίου (context) - δείκτες γεγονότος (event).
- ανά σφαίρα ενδιαφέροντος (περιβαλλοντικοί, κοινωνικοπολιτιστικοί, οικονομικοί).

Οι δείκτες πλαισίου απεικονίζουν τις αντιπροσωπευτικότερες πτυχές των συνεπειών που προκαλούνται από το Ολυμπιακό γεγονός στη χώρα που φιλοξενεί τους Αγώνες. Οι δείκτες γεγονότος αναφέρονται σε μετρήσιμα μεγέθη αποκλειστικά κατά την περίοδο των Αγώνων.

Το σύνολο δεικτών συνολικά προορίζεται να παρέχει την καλύτερη δυνατή αντιπροσώπευση του σφαιρικού αντίκτυπου των Ολυμπιακών Αγώνων. Η σφαίρα της Οικονομίας αποτέλεσε κύριο πυροδότη αλλά και αποδέκτη των επιπτώσεων των Ολυμπιακών Αγώνων της Αθήνας. Σε πιστή συμφωνία με τις προδιαγραφές, ο εντοπισμός και η ανάλυση των επιπτώσεων αυτών πραγματοποιείται σε δύο επίπεδα, αυτό των δεικτών πλαισίου και εκείνο των δεικτών γεγονότος.

Πώς όμως μπορούν να μελετηθούν οι οικονομικοί δείκτες OGGI και να ανιχνευτούν πληροφορίες μέσα από αυτούς, οι οποίες δεν είναι άμεσα ορατές; Στην παρούσα διπλωματική εργασία χρησιμοποιείται μεθοδολογία που βασίζεται στην εξόρυξη γνώσης μέσα από δεδομένα (data mining), μέσω κατάλληλων αλγορίθμων και αποδοτικών τεχνικών. Οι τεχνικές Data Mining χρησιμοποιούν αλγόριθμους, οι οποίοι είναι ικανοί να συνεισφέρουν στην εξαγωγή χρήσιμων κανόνων-προτύπων, κατά την προσέγγιση πολύπλοκων ζητημάτων, όπως είναι αυτά του τομέα της οικονομίας.

¹⁹ Βλ. OGGI Athens 2004, «Athens 2004, Olympic Games Global Impact – Οι δείκτες», <<http://147.102.46.88/OGGI/oi-deiktes-1>>, 17/5/2009.

4] Εξόρυξη Δεδομένων (Data Mining).

Ο 21^{ος} αιώνας έχει χαρακτηριστεί από πολλούς, ειδικούς και μη, ως ο αιώνας της πληροφορίας. Πέραν του ανθρώπινου δυναμικού, η πληροφορία αποτελεί τον πλέον πολύτιμο πόρο των σύγχρονων επιχειρήσεων. Επομένως, συνιστά καθοριστικό παράγοντα για την επίτευξη των διάφορων επιχειρησιακών στόχων. Εκείνοι οι οποίοι είναι σε θέση και έχουν την ικανότητα να συλλέγουν πληροφορίες και δεδομένα και έπειτα να τα αναλύουν και να τα αξιοποιούν, αναπόφευκτα βρίσκονται σε πρωταγωνιστική θέση [3]. Η πληροφορία και η αξιοποίησή της, καθώς και η ανάλυση διάφορων δεδομένων, τα οποία μπορούν να συγκεντρωθούν, δίνουν τη δυνατότητα σε κάθε ενδιαφερόμενο να αποκτήσει ένα ανταγωνιστικό πλεονέκτημα στο χώρο στον οποίο δραστηριοποιείται και να πάρει τις βέλτιστες αποφάσεις σε θέματα και διλήμματα που τον αφορούν. Αυτού του είδους οι αναλύσεις, που λαμβάνουν χώρα σε ποιοτικά²⁰ αλλά και αριθμητικά δεδομένα, πραγματοποιούνται, μεταξύ άλλων, και με την βοήθεια της επιστήμης της Τεχνητής Νοημοσύνης και πιο συγκεκριμένα με τη χρήση των τεχνικών Εξόρυξης Δεδομένων (Data Mining), οι οποίες δίνουν τη δυνατότητα εξαγωγής κανόνων μέσω των ηλεκτρονικών υπολογιστών.

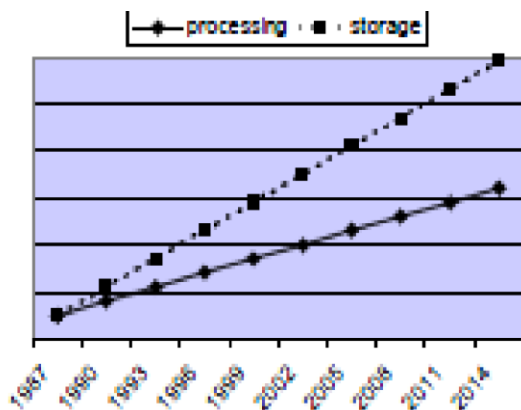
4.1] Ανακάλυψη Γνώσης μέσα από τα Δεδομένα.

Σύμφωνα με το νόμο του Moore περί ανάπτυξης της πληροφορίας, η ταχύτητα επεξεργασίας των υπολογιστών διπλασιάζεται κάθε 18 μήνες και τα δεδομένα που αποθηκεύονται διπλασιάζονται κάθε 9 μήνες, ενώ ο χρόνος που δαπανάται για την κατανάλωση των πληροφοριών παραμένει σταθερός.²¹ Κατά συνέπεια, πολύ λίγα από αυτά τα δεδομένα μπορεί να παρακολουθήσει και να αναλύσει ο άνθρωπος. Μόνο ένα μικρό ποσοστό (της τάξης του 5-10 %) των συλλεγόμενων δεδομένων τυγχάνει πλήρους ανάλυσης. Αν σκεφτούμε ότι μία τυπική επιχειρησιακή βάση δεδομένων των ημερών μας περιέχει μεγάλο αριθμό εγγραφών (10^8 - 10^{12}) δεδομένων πολλών διαστάσεων (10 - 10^4), αντιλαμβανόμαστε πόσο χαοτική γίνεται η ανάλυση των δεδομένων από τον άνθρωπο. Θα έλεγε κανείς ότι «πνιγόμαστε στα δεδομένα,

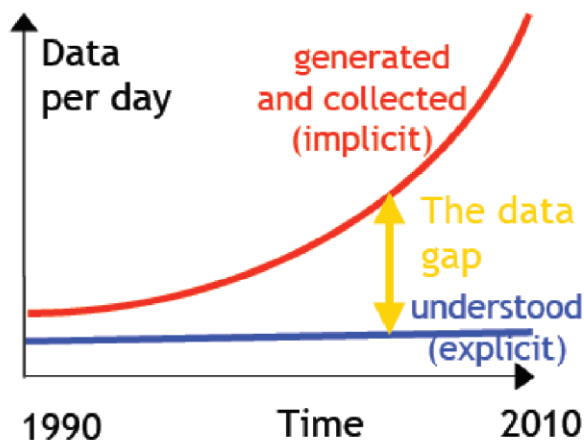
²⁰ Σημειώτεον το γεγονός ότι, για να μπορέσει κάποιος να πραγματοποιήσει μια ποιοτική ανάλυση, οφείλει να έχει πρώτα ποσοτικοποιήσει όλα τα - προς ανάλυση - δεδομένα.

²¹ Βλ. Θεοδωρίδης Γιάννης, Πελέκης Νίκος, «Αποθήκες Δεδομένων και εξόρυξη γνώσης, Εισαγωγή», σ.1.2., <<http://infolab.cs.unipi.gr/courses/dwdm/slides/1-DWDM-intro.pdf>>, 19/5/2009.

αλλά λιμοκτονούμε στη γνώση μέσα από αυτά» («We are drowning in data, but starving for knowledge» - John Naisbett).²²



Διάγραμμα 1.3.1.α : Νόμος του Moore – Η αποθήκευση των δεδομένων (storage) είναι η διπλάσια από την ταχύτητα επεξεργασίας των υπολογιστών (processing), στον ίδιο χρόνο (<http://infolab.cs.unipi.gr/courses/dwddm/slides/1-DWDDM-intro.pdf>).

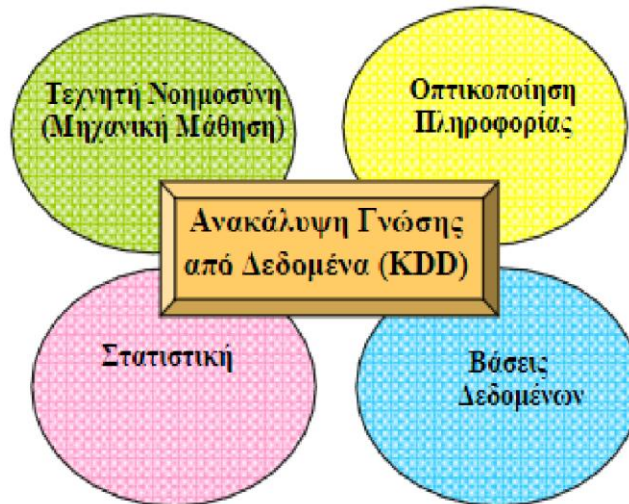


Διάγραμμα 1.3.1.β : Ο χρόνος που δαπανάται για την κατανάλωση των πληροφοριών παραμένει σταθερός, σε αντίθεση με τα δεδομένα (data) (http://gtziralis.googlepages.com/Lecture01_Introduction.pdf)

Πώς είναι, λοιπόν, δυνατόν να εξερευνηθούν εκατομμύρια εγγραφών από εκατοντάδες μεταβλητές, ούτε ώστε να ευρεθούν κατάλληλα πρότυπα (patterns); Για να γίνει κάτι τέτοιο χρειάζεται μία εκ βαθέων μελέτη, η οποία να οδηγεί στην ανακάλυψη γνώσης μέσα από τα δεδομένα (Knowledge Discovery in Data – KDD), μια γνώση που θα δίνει χρησιμότητα και πραγματικό νόημα στα δεδομένα. Η KDD αποτελεί μία μη τετριμμένη διαδικασία εύρεσης

²² Βλ. Τζιραλής Γιώργος, «Αλγόριθμοι Εξόρυξης Πληροφορίας, Διάλεξη 1, Εισαγωγή», σ.5, <http://gtziralis.googlepages.com/Lecture01_Introduction.pdf>, 19/5/2009.

έγκυρων, πρωτότυπων, πιθανώς χρήσιμων και, προπαντός, κατανοητών προτύπων (patterns) μέσα στα δεδομένα.²³ Χρησιμοποιεί κάποια επιστημονικά πεδία συγγενικά, προκειμένου να εξάγει τα κατάλληλα πρότυπα, τα οποία φαίνονται στο παρακάτω σχήμα:



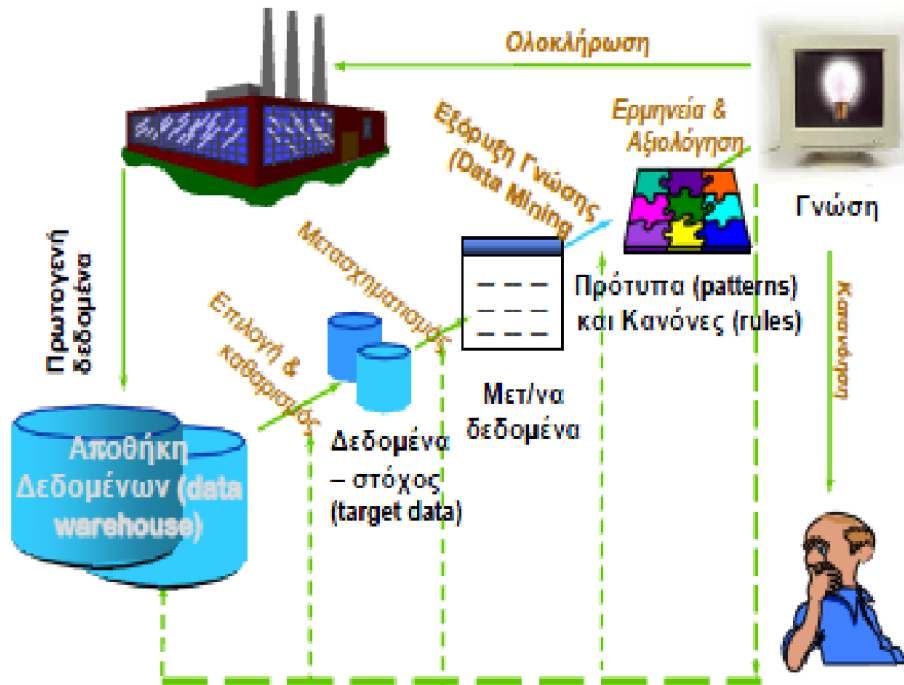
Σχήμα 1.3.1.α : Η Ανακάλυψη Γνώσης από Δεδομένα (KDD) και τα σχετικά με αυτήν επιστημονικά πεδία (<http://infolab.cs.unipi.gr/courses/dwdm/slides/1-DWDM-intro.pdf>)

Η διαδικασία ανακάλυψης γνώσης μέσα από δεδομένα αποτελείται από διαδοχικά δομημένα στάδια [4] :

- Αποθήκευση δεδομένων (data warehouses).
- Επιλογή και καθαρισμός δεδομένων (target data).
- Μετασχηματισμός δεδομένων (preprocessed data).
- Εξόρυξη γνώσης (data mining) από μετασχηματισμένα δεδομένα και εύρεση προτύπων (patterns) και κανόνων (rules).
- Ερμηνεία και αξιολόγηση των προτύπων και των κανόνων.
- Γνώση και κατανόηση από τον ανθρώπινο παράγοντα.

Όλα τα πιο πάνω βήματα φαίνονται σχηματικά παρακάτω.

²³ Βλ. Θεοδωρίδης Γιάννης, Πελέκης Νίκος, «Αποθήκες Δεδομένων και εξόρυξη γνώσης, Εισαγωγή», σ.1.3., <<http://infolab.cs.unipi.gr/courses/dwdm/slides/1-DWDM-intro.pdf>>, 19/5/2009.



Σχήμα 1.3.1 β : Η διαδικασία Ανακάλυψης Γνώσης από Δεδομένα
(<http://infolab.cs.unipi.gr/courses/dwdm/slides/1-DWDM-intro.pdf>)

4.2] Λίγα λόγια για την Τεχνητή Νοημοσύνη και τη Μηχανική Μάθηση.

Ο όρος «Τεχνητή Νοημοσύνη» δεν είναι εύκολο να οριστεί. Πολλοί επιστήμονες προσπάθησαν να εξηγήσουν τον όρο αυτό, ο οποίος αποτελεί μία από τις πιο πολύπλοκες επιστήμες στον κόσμο, χρησιμοποιώντας διάφορες εκφράσεις. Ο καθηγητής του Massachusetts Institute of Technology (MIT) στον τομέα της Τεχνητής Νοημοσύνης (Artificial Intelligence) Marvin Minsky, υποστηρίζει πως «Τεχνητή Νοημοσύνη είναι η επιστήμη που επιχειρεί να κάνει τις μηχανές να δημιουργούν πράγματα, τα οποία θα απαιτούσαν τον ανθρώπινο παράγοντα για να πραγματοποιηθούν».²⁴ Από την άλλη, η λέκτορας του University of Texas at Austin στον τομέα της Τεχνητής Νοημοσύνης, Elaine Rich, δηλώνει πως «Τεχνητή Νοημοσύνη είναι η μελέτη του πώς να κάνουμε τους ηλεκτρονικούς υπολογιστές να κάνουν πράγματα για τα οποία, προς το παρόν, οι άνθρωποι είναι καλύτεροι» (Elaine Rich, 1983).²⁵ Επίσης, διάσταση απόψεων υπάρχει και στο εάν η Τεχνητή Νοημοσύνη αποτελεί επιστήμη ξεχωριστή, ή απλά έναν κλάδο της επιστήμης των υπολογιστών [5].

Η Τεχνητή Νοημοσύνη γεννήθηκε στις αρχές της δεκαετίας του 1940, με το ενδιαφέρον να περιστρέφεται στην κατασκευή προγραμμάτων για παίξιμο παιχνιδιών και αποδείξεις θεωρημάτων. Η κυρίως έμφαση δόθηκε στην κατασκευή συστημάτων τα οποία είχαν κάποιο βαθμό γενικής ευφυΐας ή ικανότητας, για την επίλυση προβλημάτων από διαφορετικά πεδία και χώρους. Βασική κατεύθυνση της επιστήμης της Τεχνητής Νοημοσύνης και βασικός της σκοπός ήταν η δημιουργία μηχανικών δομών που παριστάνουν τον τρόπο με τον οποίο ο ανθρώπινος εγκέφαλος αποθηκεύει δεδομένα, πληροφορία και γνώση και η αντιστοιχία των δομών αυτών με τον ανθρώπινο συλλογισμό. Οι προσπάθειες όλες αφιερώθηκαν στην κατασκευή συστημάτων με περιορισμένο πεδίο εφαρμογής που κατείχαν σημαντικό όγκο γνώσης για το συγκεκριμένο πεδίο (το πεδίο προσδιορίζει το ανάλογο αντικείμενο, επιστημονικό χώρο). Στο σημείο αυτό θα αναφερθούν επιγραμματικά οι σπουδαιότεροι τομείς της επιστήμης «Τεχνητή Νοημοσύνη», ενώ θα γίνει εκτενέστερη αναφορά στη «Μηχανική Μάθηση» (Machine Learning), η οποία αποτελεί κύριο εργαλείο της τεχνικής «Εξόρυξης Δεδομένων» (Data Mining), που θα δούμε στη συνέχεια. Τομείς της «Τεχνητής Νοημοσύνης» [5] :

²⁴ Βλ. Παγουρόπουλος Απόστολος, «Διπλωματική εργασία με θέμα : Data Mining στη Χρηματοοικονομική Ανάλυση», Πάτρα, 17/10/2006, σ.4, <http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/867/1/Nimertis_Pagouropoulos.pdf>, 19/5/2009

²⁵ Βλ. Παγουρόπουλος Απόστολος, «Διπλωματική εργασία με θέμα : Data Mining στη Χρηματοοικονομική Ανάλυση», Πάτρα, 17/10/2006, σ.4, <http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/867/1/Nimertis_Pagouropoulos.pdf>, 19/5/2009

- ✓ Παίξιμο Παιχνιδιών (Game Playing).
- ✓ Αυτοματοποιημένος Συλλογισμός (Automated Reasoning).
- ✓ Έμπειρα Συστήματα (Expert Systems).
- ✓ Κατανόηση Φυσικής Γλώσσας (Natural Language Understanding).
- ✓ Κατανόηση Εικόνας (Image Understanding).
- ✓ Μηχανική Μάθηση (Machine Learning).

Μηχανική Μάθηση :

Η Μηχανική Μάθηση αποτελεί και αυτή ένα από τα πολυπλοκότερα ζητήματα. Αναφέρεται στο κατά πόσο μπορεί ένα πρόγραμμα να είναι ικανό για αυτομάθηση, είτε μέσω εμπειρίας, είτε μέσω παραδειγμάτων αλλά και περιγραφών σε φυσική γλώσσα. Αυτός ο τομέας εξελίσσεται δυναμικά και σε αυτόν βασίζονται και οι τεχνικές Data Mining [6]. Η Μηχανική Μάθηση αποσκοπεί στη διερεύνηση των μηχανισμών και των υπολογιστικών διαδικασιών, μέσω των οποίων είναι δυνατή η εξαγωγή και οργάνωση της γνώσης από την υπάρχουσα εμπειρία. Μεταξύ των διαφόρων μορφών συστημάτων μάθησης, η επαγωγική μάθηση μέσω παραδειγμάτων (inductive learning) έχει γνωρίσει τη μεγαλύτερη διάδοση. Στο συγκεκριμένο πεδίο σημαντικό ρόλο παίζει και η ποιότητα των κανόνων που εξάγονται από την υπάρχουσα γνώση, καθώς και η βελτίωσή της.

4.3] Η τεχνική Data Mining.

Η τεχνική Data Mining, η ελληνική απόδοση της οποίας είναι «Εξόρυξη ή Ανεύρεση Γνώσης από Δεδομένα»,²⁶ είναι η αναλυτική διαδικασία η οποία έχει σχεδιαστεί για να αναλύει και να εξερευνεί δεδομένα σε μεγάλες ποσότητες, αποσκοπώντας στη δημιουργία προτύπων (patterns), κανόνων (rules) και συσχετίσεων (correlations) μεταξύ των μεταβλητών που ερευνούνται κάθε φορά. Γενικά, ο όρος Data Mining αναφέρεται σε υψηλού επιπέδου εφαρμογές και μεθόδους, που χρησιμοποιούνται για να παρουσιάσουν και να αναλύσουν δεδομένα σε πεδία λήψης αποφάσεων. Η επιστήμη Data Mining αποτελεί σύγχρονη και, συνεπώς, πειραματική επιστήμη, αφού το πρώτο σχετικό συνέδριο πραγματοποιήθηκε το 1995.

4.3.1. Επιστημονικό πεδίο της τεχνικής Data Mining – Βασικότεροι Στόχοι.

Η βασική ιδέα πίσω από τον όρο Data Mining είναι η ανεύρεση εκείνης της μη μηδενικής λύσης, η οποία δίνει τη δυνατότητα εξαγωγής χρήσιμων και ουσιαστικών κανόνων σε δεδομένα. Η όλη διαδικασία βασίζεται στη χρησιμοποίηση αλγορίθμων, οι οποίοι αναζητούν κανόνες μεταξύ των μεταβλητών των δεδομένων και έπειτα καταχωρούν τα δεδομένα σε νέες βάσεις δεδομένων [7]. Τεχνικώς, οι αλγόριθμοι αυτοί είναι τα συστατικά της διαδικασίας η οποία βρίσκει συσχετισμούς ή κανόνες μέσα από τεράστιες βάσεις αποθηκευμένων δεδομένων / πληροφοριών. Επίσης, η διαδικασία Data Mining αναφέρεται συχνά και ως Computerized Technology (Πληροφοριακή Τεχνολογία),²⁷ η οποία χρησιμοποιεί πολύπλοκους αλγόριθμους που δημιουργούν κανόνες και σχέσεις μέσα σε τεράστιες βάσεις δεδομένων, αναλύοντας τις βάσεις αυτές, με σκοπό τη λήψη στρατηγικών αποφάσεων. Το ερευνητικό πεδίο της τεχνικής Data Mining αποτελεί τομή μεθόδων και εργαλείων που πηγάζουν από διάφορα άλλα επιστημονικά πεδία, όπως αυτά της Μηχανικής Μάθησης (Machine Learning), της Αναγνώρισης Κανόνων (Pattern Recognition), των Βάσεων Δεδομένων (Databases), της Στατιστικής (Statistics), της Τεχνητής Νοημοσύνης (Artificial Intelligence) και των Έμπειρων Συστημάτων (Expert Systems).

Οι κυριότεροι στόχοι της τεχνικής Data Mining είναι η Πρόβλεψη (Prediction) και η Περιγραφή (Description). Η Πρόβλεψη χρησιμοποιεί υπάρχουσες μεταβλητές στις βάσεις

²⁶ Σημειωτέον ότι η ελληνική απόδοση του όρου «Data Mining» δεν αποτελεί «πιστή» μετάφρασή του.

²⁷ Βλ. σχετικά Παγουρόπουλος Απόστολος, «Διπλωματική εργασία με θέμα : Data Mining στη Χρηματοοικονομική Ανάλυση», Πάτρα, 17/10/2006, σ.6, <http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/867/1/Nimertis_Pagouropoulos.pdf>, 19/5/2009

δεδομένων, έτσι ώστε να προβλέπει άγνωστες ή μελλοντικές αξίες ενδιαφέροντος. Από την άλλη μεριά, η Περιγραφή επικεντρώνεται στο να βρίσκει κανόνες, περιγράφοντας δεδομένα [8].

4.3.2. Συστατικά Στοιχεία Αλγορίθμων της τεχνικής Data Mining.

Οι αλγόριθμοι της κάθε τεχνικής Data Mining αποτελούνται από τρία συστατικά στοιχεία [9], τα οποία αναλύονται παρακάτω :

- **Model Representation** (Αναπαράσταση Μοντέλου) :

Είναι η γλώσσα L, η οποία περιγράφει ανακαλυφθείσες μεθόδους. Εάν η αναπαράσταση είναι μικρή και περιορισμένη αρκετά, τότε τίποτε από παραδείγματα ή από χρόνο εκπαίδευσης δε θα παράγει ένα μοντέλο για τα δεδομένα.

- **Model Evaluation** (Αποτίμηση Μοντέλου) :

Υπολογίζει κατά πόσο μια συγκεκριμένη μέθοδος, ένα μοντέλο με τις παραμέτρους του, πληρούν τα κριτήρια της διαδικασίας Knowledge Discovery Databases (Ανεύρεση γνώσης σε βάσεις δεδομένων).

- **Search Method** (Μέθοδος Αναζήτησης) :

Αποτελείται από δύο επιμέρους στοιχεία. Το πρώτο είναι η parameter search, όπου ο αλγόριθμος ψάχνει για τυχόν παραμέτρους που συνοψίζουν τα μοντέλα αποτίμησης που έλαβαν δεδομένα, ενώ το δεύτερο είναι η model research, όπου γίνεται μία διορθωτική έρευνα του όλου μοντέλου.

4.3.3. Κυριότεροι τύποι εργαλείων που χρησιμοποιούνται στην τεχνική Data Mining.

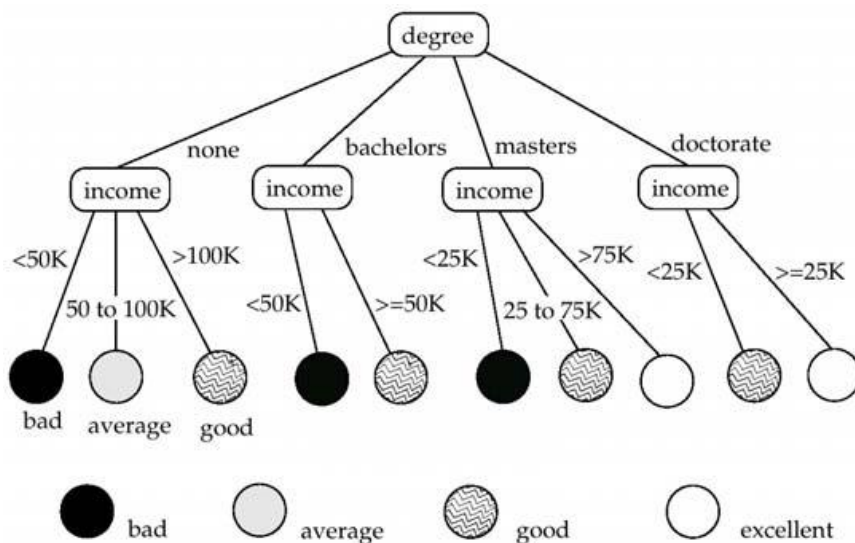
Τα εργαλεία εξόρυξης (Mining Tools or Search Engines) είναι συνήθως «έξυπνα» από το πεδίο της Τεχνητής Νοημοσύνης και σχετικά με βάσεις δεδομένων. Ποικίλα είναι τα είδη / τύποι των εργαλείων που χρησιμοποιούνται στην τεχνική Data Mining, τα κυριότερα εκ των οποίων αναφέρουμε στη συνέχεια, κάνοντας μια μικρή ανάλυση για το κάθε ένα [7].

- **Τεχνητά Νευρωνικά Δίκτυα** (Artificial Neural Networks) :

Τα Τεχνητά Νευρωνικά Δίκτυα είναι μη γραμμικά μοντέλα πρόβλεψης για λήψη αποφάσεων, τα οποία χρησιμοποιούν υπάρχοντα δεδομένα που θεωρούνται γνωστά αποτελέσματα (outcomes), για να εκπαιδεύσουν ένα μοντέλο το οποίο μπορεί μετά να χρησιμοποιηθεί για να κάνει προβλέψεις.

- **Δέντρα Αποφάσεων** (Decision Trees) :

Τα Δέντρα Αποφάσεων είναι μοντέλα υποστήριξης λήψης αποφάσεων, τα οποία δημιουργούν κάποιους κανόνες, ώστε να ταξινομήν ένα σύνολο δεδομένων. Το κάθε δέντρο αναπαριστά σύνολα από αποφάσεις (decisions). Στο παρακάτω σχήμα βλέπουμε ένα παράδειγμα Δέντρου Αποφάσεων, σύμφωνα με το οποίο ακολουθείται μια σειρά κανόνων, ώστε να διακριθούν οι πελάτες μιας επιχείρησης σε καλούς και κακούς, ανάλογα με το επίπεδο μόρφωσης και το εισόδημά τους.



Σχήμα 1.3.3.α : Παράδειγμα Decision Tree

(<http://infolab.cs.unipi.gr/courses/dwdm/slides/1-DWDM-intro.pdf>).

- **Εισαγωγή Κανόνων** (Rule Induction) :

Με τον όρο εισαγωγή κανόνων εννοούμε τη χρησιμοποίηση if-then κανόνων σε σύνολα από δεδομένα, τα οποία βασίζονται κυρίως σε στατιστικά μοντέλα. Στο κάτωθι σχήμα βλέπουμε ένα παράδειγμα if-then κανόνα, όπου ο ενδιαφερόμενος για τένις, προτρέπεται από τον κανόνα να μην παίξει το εν λόγω παιχνίδι, εφόσον ο καιρός είναι ηλιόλουστος και η υγρασία υψηλή.

	attributes				
	Outlook	Temperature	Humidity	Windy	Play
instances	Sunny	85	85	False	No
	Sunny	80	90	True	No
	Overcast	83	86	False	Yes
	Rainy	75	80	False	Yes

concept description:

If outlook = sunny and humidity = high
then play = no

Σχήμα 1.3.3.β : Παράδειγμα Κανόνα if-then

(http://qtziralis.googlepages.com/Lecture02_DataComponentsVisualizationExploration.pdf).

- **Γενετικοί Αλγόριθμοι** (Genetic Algorithms) :

Οι Γενετικοί Αλγόριθμοι είναι μία τεχνική η οποία βασίζεται σε στοιχεία / συστατικά φυσικής εξέλιξης (natural evolution), χρησιμοποιώντας γενετικούς συνδυασμούς.

- **Μέθοδος Γειτνίασης** (Nearest Neighbor Method) :

Η Μέθοδος Γειτνίασης είναι μία τεχνική ταξινόμησης δεδομένων, η οποία βασίζεται σε ομοειδείς εγγραφές, μέσα από προηγούμενες (historical) βάσεις δεδομένων.

- **Απεικόνιση Δεδομένων** (Data Visualization) :

Όπως εξυπακούεται, η Απεικόνιση Δεδομένων είναι μια εικονική αναπαράσταση των διαφόρων σχέσεων μεταξύ των δεδομένων.

4.3.4. Τα στάδια της διαδικασίας Data Mining.

Προτού αναφερθούν τα στάδια της διαδικασίας Data Mining, είναι χρήσιμο να επισημανθεί πως πολλοί αναφέρουν ότι η διαδικασία Data Mining αποτελεί μία στατιστική διαδικασία η οποία αναλύει δεδομένα αποθηκευμένα μέσα σε Data Warehouses. Τα Data Warehouses αποτελούν τεράστια πεδία με αποθηκευμένα δεδομένα, όπου εκεί περιέχονται ποικίλες πληροφορίες για όλες τις λειτουργίες ενός οργανισμού (ιδιωτική επιχείρηση, νοσοκομείο, δημόσια υπηρεσία κ.ά). Τα δεδομένα αυτά είναι επεξεργασμένα κατά τέτοιο τρόπο, ώστε να διευκολύνουν τα εργαλεία της μεθόδου Data Mining να τα διαχειρίζονται καλύτερα.

Μικρότερες σε μέγεθος εκδόσεις των Data Warehouses είναι τα Data Marts, τα οποία επικεντρώνονται και εστιάζονται σε συγκεκριμένες περιοχές / ομάδες ενδιαφέροντος, είναι οικονομικότερα και λειτουργούν σε μικρότερο περιβάλλον. Ακολουθεί ένα παράδειγμα σύγκρισης μεταξύ Data Warehouse και Data Mart σε Πίνακα, ώστε να γίνει ευκρινέστερη η διάκρισή τους.

DATA WAREHOUSE	DATA MART
Total company sales information Πληροφορίες για τις συνολικές πωλήσεις	Sales for a simple location Πωλήσεις για μία τοποθεσία, εστία της αγοράς
Store sales / Hourly sales Συνολικό ποσοστό πωλήσεων	Credit card sales only Πωλήσεις / αγορές που έγιναν με πιστωτική κάρτα
Customer profiles Εικόνα / Προφίλ πελατών	Credit card customer profiles Εικόνα πελατών σχετικά με την χρησιμοποίηση της πιστωτικής κάρτας
	Credit card purchasing history Πορεία πιστωτικής κάρτας

Πίνακας 1.3.3.α : Σύγκριση μεταξύ ενός Data Warehouse και ενός Data Mart

Η διαδικασία Data Mining αποτελείται από τρία βασικά στάδια, το στάδιο της «αναζήτησης» και «εξερεύνησης» των δεδομένων (initial exploration), το στάδιο δημιουργίας και κατασκευής του κατάλληλου μοντέλου πρόβλεψης (model building and validation) και, τέλος, το στάδιο της περαιτέρω ανάπτυξης της διαδικασίας (deployment) [10].

- **«Αναζήτηση» και «Εξερεύνηση» των δεδομένων** (initial exploration) :

Αυτό το στάδιο συνήθως ξεκινά με την προετοιμασία των δεδομένων (data preparation), που περιλαμβάνει καθαρισμό από πιθανούς θορύβους που υπάρχουν μέσα στα δεδομένα, μετασχηματισμούς δεδομένων και επιλογή των κατάλληλων υποσυνόλων εγγραφών και καθορισμό ομάδων δεδομένων με μεγάλους αριθμούς μεταβλητών και πεδίων (fields), διεξάγοντας μία επιλογή χαρακτηριστικών (feature selections). Έπειτα, ανάλογα με τη φύση του προβλήματος, αυτό το στάδιο της διαδικασίας μπορεί να περιλαμβάνει μία απλή επιλογή από μεθόδους πρόβλεψης για ένα μοντέλο ανάλυσης, ώστε να βγάλει κάποια συμπεράσματα χρησιμοποιώντας μία μεγάλη ποικιλία από στατιστικές μεθόδους (EDA - Exploratory Data Analysis). Αυτό γίνεται, για να αναγνωριστούν οι πιο σχετικές μεταβλητές και να αμβλυνθεί η πολυπλοκότητα στο επόμενο στάδιο υλοποίησης της διαδικασίας.

- **Δημιουργία του κατάλληλου μοντέλου πρόβλεψης** (model building and validation) :

Το στάδιο αυτό περιλαμβάνει την επεξεργασία πολλών και ποικίλων μοντέλων και την επιλογή εκείνου που βασίζεται στην άριστη / βέλτιστη απόδοση της πρόβλεψης. Αυτό ίσως ακούγεται σαν μια απλή διαδικασία, αλλά είναι γεγονός ότι μερικές φορές αποτελεί μία πολύπλοκη και δύσκολη διαδικασία. Μία μεγάλη γκάμα τεχνικών αναπτύχθηκαν για να επιτύχουν το στόχο αυτού του σταδίου και έχουν την προσωνομία «Competitive Evaluation Models». Σκοπός είναι η σύγκριση διαφορετικών μοντέλων με το ίδιο σύνολο δεδομένων, ώστε (ανάλογα με την απόδοση της μεθόδου κάθε φορά) να επιλέγεται το βέλτιστο. Αυτές οι τεχνικές, οι οποίες αποτελούν τη βάση του «Predictive Data Mining» είναι οι εξής: Bagging (Voting, Averaging), Boosting, Stacking (Stacked Generalizations) και Meta-Learning.

- **Ανάπτυξη** (Deployment) :

Στο τελικό στάδιο χρησιμοποιούμε το μοντέλο που επιλέχθηκε σαν το καλύτερο από τα στάδια της δημιουργίας του μοντέλου της κάθε διαδικασίας και το εφαρμόζουμε σε νέα δεδομένα, ώστε να παράγει προβλέψεις ή να εκτιμά το νέο αποτέλεσμα, να εξάγει δηλαδή νέους κανόνες.

4.3.5. Οι μέθοδοι της τεχνικής Data Mining.

Οι κυριότερες μέθοδοι της τεχνικής Data Mining, οι οποίες χρησιμοποιούν κανόνες Μηχανικής Μάθησης, είναι η ταξινόμηση (classification), η συσχέτιση (association) και η ομαδοποίηση (clustering). Υπάρχουν βέβαια και άλλες μέθοδοι, όπως η στατιστική ανάλυση (statistical analysis), η αριθμητική πρόβλεψη (numerical prediction), η παλινδρόμηση (regression analysis) και άλλες πολλές [11]. Στη συνέχεια αναλύονται ο τρόπος λειτουργίας και δομής των κανόνων των τριών σπουδαιότερων μεθόδων, ταξινόμησης, συσχέτισης και ομαδοποίησης [3].

Ταξινόμηση (Classification) :

Η ταξινόμηση (classification) αποτελεί, ίσως την πιο δημοφιλή και αποτελεσματική μέθοδο. Οι αλγόριθμοι ταξινόμησης εφαρμόζονται σε δεδομένα, τα οποία έχουν προταξινομηθεί σε συγκεκριμένες κλάσεις, με στόχο την εξαγωγή κανόνων οι οποίοι μπορεί μετέπειτα να χρησιμοποιηθούν για ταξινόμηση νέων δεδομένων στις ίδιες κλάσεις. Ένα σύνολο εξαγόμενων κανόνων ονομάζεται ταξινομητής (classifier). Η λειτουργία των αλγορίθμων ταξινόμησης αποτελείται από τα εξής στάδια [12] :

(α) Εισάγουμε ένα σύνολο από δεδομένα (data set) σε έναν αλγόριθμο ταξινόμησης.

(β) Ο αλγόριθμος έπειτα, με τη σειρά του, «κατανοεί» τους κανόνες βάση των οποίων ταξινομήθηκαν τα δεδομένα.

(γ) Ακολούθως, βάση των συγκεκριμένων κανόνων, ο αλγόριθμος έχει την ικανότητα να ταξινομήσει νέα δεδομένα.

Υπάρχουν δύο βασικοί τύποι αλγορίθμων ταξινόμησης, ανάλογα με το είδος του ταξινομητή που παράγουν :

(i) Οι αλγόριθμοι που παράγουν λίστες αποφάσεων.

(ii) Οι αλγόριθμοι που παράγουν δέντρα αποφάσεων.

Οι κανόνες που παράγουν λίστες αποφάσεων έχουν τη μορφή λογικών κανόνων που βγάζουν ανάλογα συμπεράσματα και είναι μία σχετικά νέα μορφή αλγορίθμων [13]. Από την άλλη μεριά, οι αλγόριθμοι που παράγουν δέντρα αποφάσεων είναι εκείνοι που αποτελούν ίσως την πιο παλιά μορφή και έκφραση της τεχνικής Data Mining. Οι αλγόριθμοι αυτοί έχουν στη ρίζα τους και στους ενδιάμεσους κόμβους τιμές των διάφορων πεδίων και στα φύλλα τους τιμές του πεδίου κλάσης. Ο κάθε κόμβος διακλαδώνεται προς τα κάτω, έχοντας ένα κλαδί για κάθε διακριτή τιμή του πεδίου. Σε περίπτωση συνεχούς αριθμητικού πεδίου, χωρίζεται το εύρος του πεδίου σε διαστήματα και ο κόμβος διακλαδώνεται με βάση αυτά. Τέλος ένας αλγόριθμος που παράγει δέντρα αποφάσεων ακολουθεί συνήθως αναλυτική προσέγγιση (top-down), δημιουργεί δηλαδή το δένδρο από τη ρίζα και προχωράει προς τα κάτω [14].

✓ **Συσχέτιση (Association) :**

Η συσχέτιση είναι μία από τις σημαντικότερες μεθόδους της τεχνικής Data Mining. Σκοπός της είναι η εύρεση των σημαντικών αλληλεξαρτήσεων μεταξύ των διάφορων πεδίων / χαρακτηριστικών του συνόλου εκπαίδευσης. Η πιο συνηθισμένη εφαρμογή της μεθόδου της συσχέτισης είναι η ανάλυση του «καλαθιού της νοικοκυράς» [12]. Σκοπός είναι να αναγνωρισθούν τα αγαθά που αγοράζονται μαζί. Έστω, για παράδειγμα, ο παρακάτω πίνακας με τα δεδομένα, όπου σε κάθε γραμμή του απεικονίζει τα αγαθά που αγοράστηκαν σε ένα καλάθι μέσα σε μία υπεραγορά :

Coca-cola	Γάλα	Τυρί	Φρυγανιές	Κρέας
Τυρί	Νερό	Γάλα	Κρέας	

Κρέας	Γάλα	Ψωμί		
...

Πίνακας 1.3.3.β : Το «καλάθι της νοικοκυράς» (κάθε γραμμή αποτελεί ξεχωριστό «καλάθι»).

Ένας κανόνας συσχέτισης θα μπορούσε να πει ότι το γάλα πωλείται μαζί με το τυρί, με την προφανή αξιοποίηση της πληροφορίας που είναι η γειννίαση του σημείου πώλησης γάλακτος με αυτό του τυριού.

Ένας κανόνας συσχέτισης είναι μία έκφραση της μορφής $X \Rightarrow Y$, όπου X, Y είναι σύνολα τιμών των πεδίων, όπως για παράδειγμα σύνολα οικονομικών αγαθών. Η σπουδαιότητα ενός κανόνα συσχέτισης καθορίζεται αναλογικά από το ποσοστό εφαρμογής του κανόνα επί του συνόλου εκπαίδευσης. Συγκεκριμένα, οι αλγόριθμοι συσχέτισης που έχουν προταθεί και εφαρμόζονται πρακτικά, εξάγουν κανόνες συσχέτισης της μορφής: «το 98% των πελατών που αγοράζουν γάλα και κρέας αγοράζουν επίσης και τυρί. Αλλά και στο 70% των αγορών έχουν αγορασθεί γάλα, κρέας και τυρί». Το πρώτο ποσοστό αναφέρεται ως αξιοπιστία (confidence) του κανόνα, ενώ το δεύτερο ως επιβεβαίωση (support) [12]. Η επιβεβαίωση αφορά το ποσοστό που εμφανίζονται και τα τρία αγαθά μαζί, επί όλου του συνόλου εκπαίδευσης, ενώ η αξιοπιστία αφορά το ποσοστό που εμφανίζονται τα αγαθά μαζί, επί του αριθμού των αγορών που περιέχουν γάλα και κρέας. Το πρόβλημα της εύρεσης κανόνων συσχέτισης εστιάζεται στην εύρεση όλων των κανόνων που έχουν μία καθορισμένη από το χρήστη ελάχιστη τιμή επιβεβαίωσης και αξιοπιστίας.

✓ **Ομαδοποίηση (Clustering) :**

Οι κανόνες ομαδοποίησης είναι, επίσης, αρκετά διαδεδομένοι. Διαφέρουν από εκείνους της ταξινόμησης στο γεγονός ότι τα δεδομένα που χρησιμοποιούνται για μάθηση δεν είναι προταξινομημένα, λείπει δηλαδή η κλάση του πεδίου. Η τεχνική ομαδοποίησης χωρίζει ουσιαστικά ένα σύνολο εγγραφών σε ομάδες, έτσι ώστε οι εγγραφές που βρίσκονται στην ίδια ομάδα να έχουν περισσότερες ομοιότητες μεταξύ τους, με βάση ορισμένα προκαθορισμένα κριτήρια, απ' ό,τι με εγγραφές άλλων ομάδων [15]. Σήμερα είναι ιδιαίτερα σημαντικό για τις επιχειρήσεις να μπορούν να ομαδοποιούν τους πελάτες τους σε συγκεκριμένες κατηγορίες. Με βάση αυτές τις κατηγορίες μπορούν να αξιολογούν ένα νέο πελάτη, ανάλογα με την ομάδα στην οποία κατατάσσεται αυτός, ή ακόμα να προσδιορίσουν τα χαρακτηριστικά των πελατών που αποφέρουν μεγάλα κέρδη στην εταιρεία. Με τον τρόπο αυτό, οι επιχειρήσεις προσανατολίζουν

τη στρατηγική τους στην εξειδικευμένη εξυπηρέτηση ορισμένων πελατειακών ομάδων. Για παράδειγμα, από την ανάλυση ενός πολύ μεγάλου συνόλου πελατών, μπορεί να μειωθεί το κόστος μίας διαφημιστικής εκστρατείας που βασίζεται στην αποστολή διαφημιστικών φυλλαδίων. Αυτό γίνεται με τον περιορισμό του πλήθους των πελατών στους οποίους απευθύνεται, επιλέγοντας αυτούς με μεγαλύτερη πιθανότητα να αντιδράσουν θετικά. Για να μπορέσει να γίνει η επιλογή του κατάλληλου αλγορίθμου, απαραίτητη προϋπόθεση είναι η μελέτη των δεδομένων που θα χρησιμοποιηθούν για τον προσδιορισμό κυρίως του κριτηρίου ομοιότητας των εγγραφών μίας ομάδας.

Η μέθοδος της ομαδοποίησης μπορεί να είναι είτε στατιστική, είτε αριθμητική (statistical / numerical clustering), οπότε χρησιμοποιούνται διάφορα αριθμητικά κριτήρια ομοιότητας και οι ομάδες που προκύπτουν περιγράφονται από αριθμητικές τιμές, είτε εννοιολογική (conceptual clustering), οπότε σε αυτή την περίπτωση ο προσδιορισμός των ομάδων βασίζεται στο νόημα και στις έννοιες που τα διάφορα αριθμητικά στοιχεία αντιπροσωπεύουν και οι τιμές που προκύπτουν είναι κατηγορικές και όχι αριθμητικές. Πολλοί από τους αλγόριθμους ομαδοποίησης επιτρέπουν και το μικτό σύνολο εκπαίδευσης (αριθμητική και εννοιολογική μέθοδο ταυτόχρονα) [16].

Εν τέλει, η ενασχόληση του παρόντος συγγράμματος με την επιστήμη της εξόρυξης γνώσης μέσα από σύνολο δεδομένων, βρίσκει πρακτική εφαρμογή μέσω του λογισμικού WEKA. Οι βασικές ιδιότητες και λειτουργίες του συγκεκριμένου λογισμικού αναλύονται εκτενώς στα κεφάλαια 5 και 6 που ακολουθούν.

5] Το λογισμικό WEKA.

5.1] Τι είναι το λογισμικό WEKA;

Το λογισμικό (software) WEKA δημιουργήθηκε το 1994 από μία ομάδα καθηγητών και φοιτητών του Πανεπιστημίου του Waikato στη Νέα Ζηλανδία και αναπτύσσεται μέχρι σήμερα. Δανείζεται το όνομά του από ένα ενδημικό πτηνό, που υπάρχει στη Νέα Ζηλανδία και το οποίο δεν έχει καθόλου πτητική ικανότητα. Το πλήρες όνομα του λογισμικού WEKA είναι Waikato Environment for Knowledge Analysis.²⁸ Από το όνομα και μόνο, αντιλαμβάνεται κανείς, πως το λογισμικό WEKA ασχολείται με την ανακάλυψη γνώσης μέσα από δεδομένα (KDD), η οποία αναλύθηκε εκτενώς στην προηγούμενη ενότητα.



Σχήμα 1.4.1.α : Το πτηνό WEKA

(http://gtziralis.googlepages.com/Lecture02_DataComponentsVisualizationExploration.pdf)

Στο σημείο αυτό καλό είναι να δοθούν τα βασικά χαρακτηριστικά του λογισμικού WEKA :

- Είναι γραμμένο σε JAVA.
- Παρέχει πολλούς διαφορετικούς αλγόριθμους, για την εξόρυξη δεδομένων (data mining) και τη μηχανική μάθηση (machine learning).
- Είναι λογισμικό ανοικτής πηγής (open-source) και ελεύθερης πρόσβασης (<http://www.cs.waikato.ac.nz/ml/weka/>).

²⁸ Βλ. Ντούτση Ειρήνη, «Ανακάλυψη Γνώσης από Δεδομένα και Εξόρυξη Γνώσης στο εργαλείο WEKA», σ.4, <<http://infolab.cs.unipi.gr/courses/dwdm/tutorials/WEKA-tutorial-2008.pdf>>, 2/6/2009.

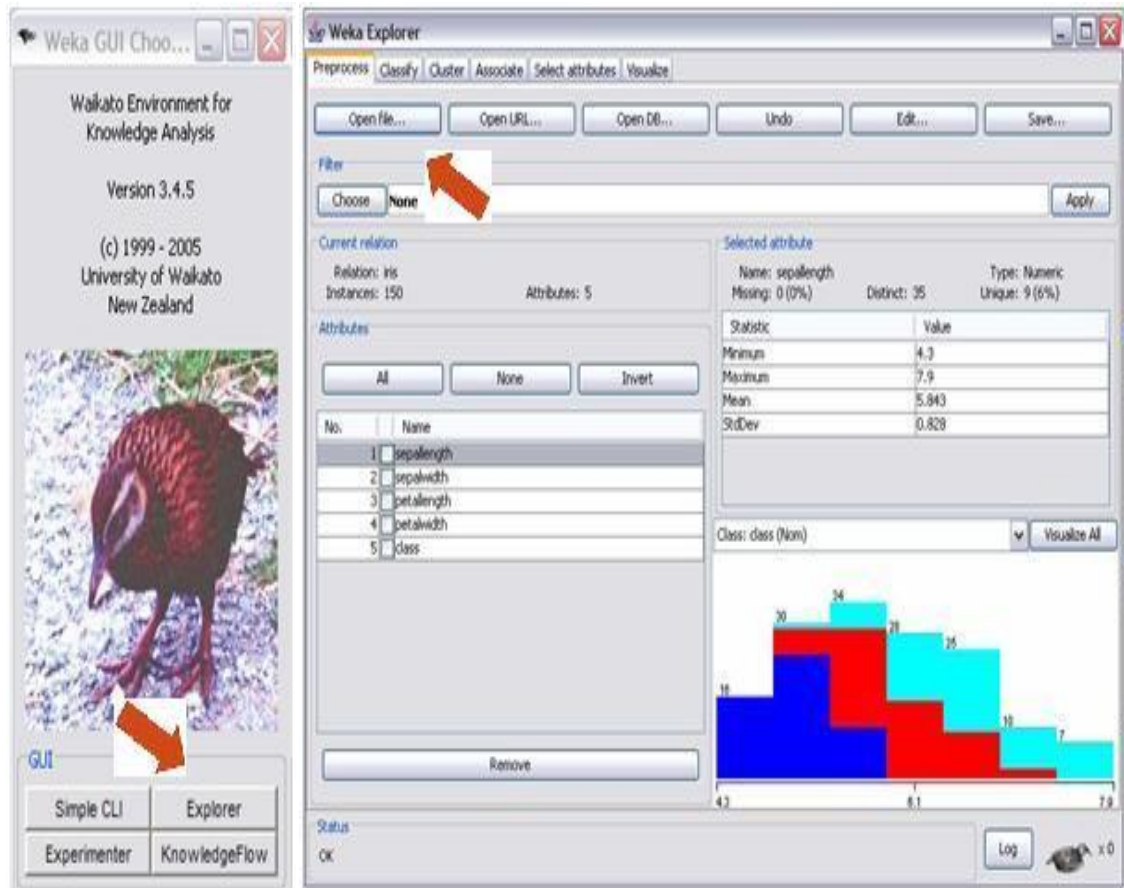
- Είναι πολύ εύχρηστο λογισμικό, ακόμα και σε ανθρώπους που δεν ασχολούνται με τεχνικές Data Mining.
- Ενημερώνεται (updating) συνεχώς για τους νέους αλγόριθμους που εμφανίζονται καθημερινά στον τομέα της Μηχανικής Μάθησης και αναβαθμίζεται.
- Περιέχει υλοποιημένες μεθόδους για προεπεξεργασία (preprocessing), ταξινόμηση (classification), συσχέτιση (association) και ομαδοποίηση (clustering) δεδομένων.

5.2] Το περιβάλλον του WEKA.

Ανοίγοντας το πρόγραμμα, μέσω του μενού «Application → Explorer → Open file», δίνεται η δυνατότητα στο χρήστη να επιλέξει ένα σύνολο δεδομένων, στο οποίο μπορούν να εφαρμοστούν τεχνικές που αφορούν :

- ✓ **Preprocess** (προεπεξεργασία) :
Επιλογή και τροποποίηση των δεδομένων που τυγχάνουν επεξεργασίας.
- ✓ **Classify** (ταξινόμηση) :
«Κατανόηση» προεπεξεργασίας των δεδομένων και ταξινόμησή τους.
- ✓ **Cluster** (ομαδοποίηση) :
Διαχωρισμός των δεδομένων σε συστάδες (ομάδες).
- ✓ **Associate** (συσχέτιση) :
Εύρεση κανόνων συσχέτισης μεταξύ των δεδομένων.
- ✓ **Select Attributes** (επιλογή χαρακτηριστικών) :
Επιλογή των πιο σχετικών μεταξύ τους χαρακτηριστικών (attributes).
- ✓ **Visualize** (οπτικοποίηση) :
Γραφική αναπαράσταση όλων των συνδυασμών μεταξύ των χαρακτηριστικών, με σκοπό την ανίχνευση συσχετίσεων μεταξύ τους.

Επιλέγοντας ένα σύνολο δεδομένων (το οποίο είναι αποθηκευμένο σε αρχείο .arff), εμφανίζονται γραφικά τα δεδομένα για καθένα από τα γνωρίσματα ξεχωριστά, καθώς και στατιστικές πληροφορίες γι' αυτά. Εάν στο σύνολο δεδομένων δίνεται και κάποια κλάση στην οποία ταξινομούνται, τα δεδομένα που ανήκουν στην ίδια κλάση εμφανίζονται με το ίδιο χρώμα.



Σχήμα 1.4.2.α : Η βασική οθόνη του WEKA (GUI Chooser και Explorer)

(http://gtziralis.googlepages.com/Lecture02_DataComponentsVisualizationExploration.pdf)

5.3] Αποθήκευση δεδομένων σε αρχεία .arff.

Η τυποποίηση .arff αποτελεί τη φυσική μέθοδο αποθήκευσης δεδομένων του WEKA.²⁹ Υποστηρίζει αριθμητικά (numeric) και ονομαστικά (nominal) χαρακτηριστικά (attributes). Τα αρχεία που περιέχουν το σύνολο των δεδομένων, πρέπει να έχουν συγκεκριμένη μορφή (format) και να αποθηκεύονται με την επέκταση .arff. Στο φάκελο C:\Program Files\Weka-3-5\data, ο οποίος εμφανίζεται μετά την επιλογή «Open file», περιέχονται κάποια παραδείγματα τέτοιων αρχείων. Δεδομένα μπορούν επίσης να δοθούν από ένα URL ή από μία SQL βάση, αλλά στην παρούσα εργασία έχει χρησιμοποιηθεί, κατά κόρον, η τυποποίηση .arff. Παρακάτω παρουσιάζεται ένα παράδειγμα τυποποίησης δεδομένων σε ένα αρχείο .arff.

²⁹ Βλ. σχετικά Τζιραλής Γιώργος, «Αλγόριθμοι Εξόρυξης Πληροφορίας, Διάλεξη 02 - Συνιστώσες Δεδομένων, Οπτικοποίηση και Εξερεύνηση», <http://gtziralis.googlepages.com/Lecture02_DataComponentsVisualizationExploration.pdf>, 3/6/2009.

% ARFF file for weather data with some numeric features

@relation weather

@attribute outlook {sunny, overcast, rainy}

@attribute temperature numeric

@attribute humidity numeric

@attribute windy {true, false}

@attribute play? {yes, no}

@data

sunny, 85, 85, false, no

sunny, 80, 90, true, no

overcast, 83, 86, false, yes

...

...

...

Στο επόμενο κεφάλαιο γίνονται πιο κατανοητές οι διάφορες τεχνικές data mining, τις οποίες εφαρμόζει το συγκεκριμένο λογισμικό που περιγράφεται πιο πάνω. Συγκεκριμένα, αναλύονται τα στάδια της μεθοδολογίας που ακολουθήθηκε, προκειμένου να εξάγουμε τους διάφορους κανόνες που προκύπτουν από το σύνολο των δεδομένων που έτυχαν επεξεργασίας data mining.

6] Μεθοδολογία.

Ξεκινώντας τη διαδικασία αποτίμησης επιπτώσεων των Ολυμπιακών Αγώνων «Αθήνα 2004» στον τομέα της Οικονομίας, επιλέχθηκε η αποπεράτωση της όλης διαδικασίας μέσα από ένα πρίσμα πρωτοποριακής, καινοτομικής εισήγησης και όχι η συνέχιση παραδοσιακών μεθόδων στατιστικής ανάλυσης. Με τον όρο καινοτομία εννοείται η μελλοντικά άμεσα υλοποιούμενη σε απτά αποτελέσματα, νέα και πρωτοποριακή ιδέα ή η νέα διαδικασία αυτής της υλοποίησης.³⁰ Το να γίνει όμως κάτι τέτοιο δεν είναι απλό ούτε και εύκολο, πόσο μάλλον το να το καταγράψει και να το παρουσιάσει κανείς μέσα σε περιορισμένο αριθμό σελίδων. Από την άλλη, το να παρουσιαστούν αποτελέσματα και να εξαχθούν συμπεράσματα μέσω της τεχνικής Data Mining που επιλέχθηκε, είναι κάτι που κρύβει «παγίδες». Τις «παγίδες» αυτές κληθήκαμε να τις ξεπεράσουμε, ακολουθώντας μια συγκεκριμένη μεθοδολογία. Η όλη μεθοδολογία δεν ακολουθεί πιστά όλους τους αλγόριθμους και όλες τις τεχνικές που χρησιμοποιεί η μεθοδολογία Data Mining, αλλά μέρος αυτών, των σπουδαιότερων ίσως. Τα στάδια της μεθοδολογίας data mining που ακολουθήθηκε, από τη στιγμή επιλογής των δεδομένων - δεικτών OGGI - μέχρι την τελική επιλογή αλγορίθμων κατάλληλου λογισμικού - και δη του WEKA - και «τρέξιμο» αυτών, παρουσιάζονται επιγραμματικά παρακάτω και αναλύονται περιληπτικά στη συνέχεια.

Στάδια μεθοδολογίας που ακολουθήθηκε :

- Επιλογή συνόλου δεδομένων (data set) που εισάγεται στο λογισμικό (WEKA).
- Προεπεξεργασία (preprocessing) δεδομένων.
- Επιλογή των πιο σχετικών μεταξύ τους χαρακτηριστικών (selection of attributes).
- Επιλογή και «τρέξιμο» κατάλληλων αλγορίθμων για περαιτέρω επεξεργασία των δεδομένων και εξαγωγή αποτελεσμάτων - συμπερασμάτων - κανόνων.

³⁰ Βλ. «Καινοτομία, Από τη Βικιπαίδεια, την ελεύθερη εγκυκλοπαίδεια», <<http://el.wikipedia.org/wiki>>, 10/6/2009.

6.1] Επιλογή Δεδομένων.

Το πρώτο στάδιο ξεκινά με την προετοιμασία των δεδομένων (data preparation), που περιλαμβάνει καθαρισμό από πιθανούς θορύβους που υπάρχουν μέσα στα δεδομένα, μετασχηματισμούς δεδομένων και επιλογή των κατάλληλων υποσυνόλων εγγραφών και καθορισμό ομάδων δεδομένων με μεγάλους αριθμούς μεταβλητών και πεδίων (fields), διεξάγοντας μία επιλογή χαρακτηριστικών (feature selections) [17]. Αυτό γίνεται, για να αναγνωριστούν τα πιο σημαντικά χαρακτηριστικά (attributes) και να αμβλυνθεί η πολυπλοκότητα στο επόμενο στάδιο υλοποίησης της διαδικασίας. Κατά την παρούσα εργασία, το στάδιο αυτό χωρίζεται σε δύο επιμέρους βοηθητικά στάδια, τα οποία είναι τα εξής :

- ✓ Συγκέντρωση όλων των δεικτών ανάλυσης του OGGI - Αθήνα 2004 στη σφαίρα της Οικονομίας και ταξινόμησή τους σε ένα Πίνακα Δεδομένων (Data Table).
- ✓ Μετατροπή του Πίνακα Δεδομένων από αρχείο .xls σε αρχείο .csv (αρχικά) και τελικά τυποποίηση σε αρχείο .arff, για εισαγωγή των δεδομένων στο λογισμικό WEKA.

6.1.1. Συγκέντρωση όλων των οικονομικών δεικτών ανάλυσης του «OGGI - Αθήνα 2004».

Στο σημείο αυτό παρατίθενται επιγραμματικά παρακάτω όλοι οι δείκτες του Τομέα της Οικονομίας που υπολογίστηκαν στα πλαίσια του προγράμματος OGGI-Αθήνα 2004. Αξίζει να σημειωθεί ότι πολλοί από τους δείκτες ήταν ελλιπείς (μέχρι το 2003) και έγινε προσπάθεια συμπλήρωσής τους μέχρι το 2007, κάτι που ήταν δύσκολο, αφού η περίοδος 2004 - 2007 θεωρείται σχετικά «πρόσφατο παρελθόν» και για την περίοδο αυτή δεν υπάρχουν αρκετά επίσημα στοιχεία.

Δείκτες πλαισίου των Αγώνων (context indexes) :

- ✓ Ec01 Δομή Απασχόλησης
- ✓ Ec04 Κοινωνικο-οικονομικές τάσεις
- ✓ Ec05 Κατανάλωση ενέργειας κατά κεφαλήν
- ✓ Ec06 Κατανάλωση ενέργειας κατά πρωτογενή πηγή
- ✓ Ec07 Κατανάλωση ενέργειας ανά χρήση
- ✓ Ec08 Ενεργειακή αυτάρκεια
- ✓ Ec12 Ξενοδοχειακή Υποδομή
- ✓ Ec13 Βαθμός Πληρότητας Καταλυμάτων
- ✓ Ec14 Καταγεγραμμένοι Επιβάτες σε Αεροδρόμια

- ✓ Ec15 Αφίξεις Περιηγητών σε Καταλύματα
- ✓ Ec16 Διανυκτερεύσεις Περιηγητών
- ✓ Ec17 Εποχικότητα Ξενοδοχειακής Υποδομής
- ✓ Ec18 Δομή Οικονομικού Προϊόντος
- ✓ Ec19 Κατά Κεφαλήν Ακαθάριστο Εγχώριο Προϊόν
- ✓ Ec20 Αυτοδυναμία οικονομίας
- ✓ Ec21 Αποδοχές
- ✓ Ec22 Ποσοστό Φτώχειας
- ✓ Ec23 Δείκτης Εισοδηματικής Κατανομής (Gini Index)
- ✓ Ec24 Δείκτης Τιμών Καταναλωτή
- ✓ Ec25 Δείκτες Τιμών
- ✓ Ec26 Εισαγωγές-Εξαγωγές Προϊόντων
- ✓ Ec27 Δυναμική παροχής υπηρεσιών
- ✓ Ec28 Ισοτιμίες
- ✓ Ec29 Επενδυτικός Κίνδυνος
- ✓ Ec30 Οικονομικός ρόλος του κράτους
- ✓ Ec31 Δομή Δημοσιοοικονομικών Εξόδων
- ✓ Ec32 Δομή Δημοσιοοικονομικών Εσόδων
- ✓ Ec33 Κατά Κεφαλήν Ακαθάριστο Χρέος
- ✓ Ec34 Βαρύτητα Δημοσίου Χρέους

Δείκτες γεγονότος των Αγώνων (event indexes) :

- ✓ Ec40 Θέσεις εργασίας που δημιουργήθηκαν
- ✓ Ec41 Ολυμπιακή Φιλοξενία
- ✓ Ec42 Ολυμπιακά Οχήματα
- ✓ Ec43 Δομή Εσόδων ΟΕΟΑ
- ✓ Ec44 Δομή Εξόδων ΟΕΟΑ
- ✓ Ec45 Καταλυτική Επίδραση των Αγώνων
- ✓ Ec46 Αναλογίες Κόστους Σχετιζόμενες με Ολυμπιακές Δραστηριότητες
- ✓ Ec47 Λειτουργικά Έξοδα Ολυμπιακών Δραστηριοτήτων
- ✓ Ec48 Επενδυτικά Έξοδα Υποδομών Ολυμπιακών Δραστηριοτήτων
- ✓ Ec49 Λειτουργικά Έξοδα Συναφών με τις Ολυμπιακές Δραστηριοτήτων
- ✓ Ec50 Επενδυτικά Έξοδα Υποδομών Συναφών με τις Ολυμπιακές Δραστηριοτήτων
- ✓ Ec54 Δημόσιες Δαπάνες Ολυμπιακών Δραστηριοτήτων

- ✓ Ec55 Δημόσιες Δαπάνες Συναφών με τις Ολυμπιακές Δραστηριοτήτων
- ✓ Ec56 Έσοδα από Ολυμπιακές Δραστηριότητες

Από τους δείκτες αυτούς επιλέγονται μόνο οι δείκτες πλαισίου³¹ και πάνω σε αυτούς γίνεται το όποιο «φιλτράρισμα». Όλοι αυτοί οι δείκτες παρατίθενται στο Παράρτημα και είναι ενδεικτικοί των σπουδαιότερων οικονομικών και άλλων μεταβολών του χρονικού διαστήματος 1996 - 2007.

6.1.2. Μορφοποίηση Πίνακα Δεδομένων.

Εν συνεχεία, γίνεται η μετατροπή του συγκεντρωτικού data table από αρχείο .xls σε αρχείο .csv, μέσω των επιλογών «Save as →save as type →name.csv». Το υπάρχον πλέον .csv αρχείο ανοίγεται με .txt editor, π.χ. notepad, ώστε να γίνει η τυποποίησή του σε αρχείο .arff και να ανοιχτεί στο WEKA. Στο notepad γράφουμε το όνομα του data set που επεξεργαζόμαστε (@relation), τις πληροφορίες των χαρακτηριστικών (@attribute, μία σειρά για κάθε χαρακτηριστικό) και τη σειρά @data.³² Ας παραθέσουμε στο σημείο αυτό ένα κομμάτι του notepad που δημιουργήσαμε στην παρούσα εργασία (ολόκληρο το notepad αναλυτικά υπάρχει στο Παράρτημα), για καλύτερη κατανόηση :

@relation olympic_games

@attribute year numeric

@attribute A_-_Agriculture_hunting_and_forestry numeric

@attribute B_-_Fishing numeric

@attribute C_-_Mining_and_quarrying numeric

@attribute D_-_Manufacturing numeric

@data

1996	639	26	18	628	44
1997	618	24	17	605	42
1998	712	12	21	583	37
1999	710	13	20	575	41
.

³¹ Βλ. σχετικά το κεφάλαιο 3.4. της παρούσας εργασίας.

³² Βλ. Τζιραλής Γιώργος, «Αλγόριθμοι Εξόρυξης Πληροφορίας, Διάλεξη 2, Συνιστώσες Δεδομένων-Οπτικοποίηση-Εξερεύνηση», <http://gtziralis.googlepages.com/Lecture02_DataComponentsVisualizationExploration.pdf>, 10/6/2009

.
.

Τελικά σώζουμε το αρχείο ονομάζοντάς το `olympic_games.arff` και είναι έτοιμο για άνοιγμά του στο WEKA. Η τυποποίηση `.arff` αποτελεί τη φυσική μέθοδο αποθήκευσης δεδομένων του WEKA.

6.2] Προεπεξεργασία Δεδομένων.

Στο στάδιο αυτό γίνεται ουσιαστικά μία διεργασία εξεύρεσης των σπουδαιότερων - με βάση την ελληνική και παγκόσμια οικονομία και τα δεδομένα αυτών - χαρακτηριστικών (attributes), τα οποία χρησιμοποιούνται στην πορεία, στις διάφορες λειτουργίες του λογισμικού, ως «μεταβλητές - στόχοι» (classes). Αφού καθοριστούν οι «μεταβλητές - στόχοι», γίνεται μία διακριτοποίηση (discretize) των χαρακτηριστικών, μέσω της επιλογής κατάλληλου φίλτρου (filters → unsupervised → attribute → discretize). Με τον όρο «διακριτοποίηση» εννοούμε τη μετατροπή των αριθμητικών (numeric) χαρακτηριστικών σε ονομαστικά (nominal) [18]. Η διακριτοποίηση επιβάλλεται, αφού μερικοί αλγόριθμοι ταξινόμησης και ομαδοποίησης χειρίζονται μόνο ονομαστικά χαρακτηριστικά και δεν μπορούν να χειριστούν όσα βρίσκονται σε αριθμητική μορφή, ενώ ακόμα και οι αλγόριθμοι μάθησης που μπορούν να χειριστούν αριθμητικά χαρακτηριστικά, συχνά λειτουργούν πολύ πιο γρήγορα όταν χρησιμοποιούνται μόνο με διακριτά χαρακτηριστικά.³³ Τελική φάση του σταδίου αυτού είναι ο καθαρισμός των χαρακτηριστικών, μέσω της οπτικοποίησής τους (visualize) [18]. Με την οπτικοποίηση των χαρακτηριστικών αποκτά δηλαδή ο χρήστης μια πρώτη εποπτική ματιά στα χαρακτηριστικά και τις σχέσεις μεταξύ τους, μέσω της γραφικής απεικόνισής τους. Παρακάτω γίνεται απλή αναφορά των χαρακτηριστικών εκείνων που θεωρήθηκαν στην παρούσα εργασία τα «σπουδαιότερα» - σε σχέση με την ελληνική οικονομία - και τέθηκαν ως μεταβλητές - στόχοι στο λογισμικό WEKA :

- i. Gross Domestic Product (G.D.P.) - Ακαθάριστο Εγχώριο Προϊόν (Α.Ε.Π.)
- ii. Gross National Product (G.N.P.) - Ακαθάριστο Εθνικό Προϊόν (Α.ΕΘ.Π.)
- iii. Gini Index - Δείκτης Εισοδηματικής Κατανομής (Δ.Ε.Κ.)
- iv. Consumer Price Index (C.P.I.) - Δείκτης Τιμών Καταναλωτή (Δ.Τ.Κ.)
- v. Gross Fixed Capital Formation - Ακαθάριστος Σχηματισμός Παγίου Κεφαλαίου
- vi. Fiscal Revenue - Φορολογικό Εισόδημα
- vii. Inflation - Πληθωρισμός
- viii. Exchange Rate Euro - U.S.Dollar - Ισοτιμία Συναλλάγματος Ευρώ - Αμερικανικού Δολαρίου
- ix. Low Wage Population - Χαμηλόμισθος Πληθυσμός
- x. Minimum Daily Wage of Workers - Ελάχιστος Ημερήσιος Μισθός των Εργαζόμενων
- xi. Per Capita Gross Dept Current Prices - Κατά Κεφαλή Ακαθάριστο Χρέος

³³ Βλ. Τζιραλής Γιώργος, «Αλγόριθμοι Εξόρυξης Πληροφορίας, Διάλεξη 3, Προεπεξεργασία και Επιλογή Δεδομένων», <http://gtziralis.googlepages.com/Lecture02_DataComponentsVisualization_Exploration.pdf>, 10/6/2009.

- xii. Public Spending - Δημόσιες Δαπάνες
- xiii. Ratio of Services Exports to Imports - Αναλογία των Υπηρεσιών Εξαγωγές προς Εισαγωγές
- xiv. Total Energy Gross Inland Consumption - Συνολική Ακαθάριστη Ενεργειακή Εγχώρια Κατανάλωση
- xv. Total Numbers of Passengers (all airports) – Συνολικός Αριθμός Επιβατών (όλα τα αεροδρόμια)
- xvi. Unemployment Rate - Ποσοστό Ανεργίας

6.3] Επιλογή των πιο σχετικών, μεταξύ τους, χαρακτηριστικών (select attributes).

Ίσως αποτελεί το σπουδαιότερο όλων των σταδίων της μεθοδολογίας που ακολουθήθηκε στην παρούσα εργασία, αλλά και της μεθοδολογίας που ακολουθείται γενικά στις διάφορες τεχνικές data mining. Στις περισσότερες εφαρμογές data mining ο αριθμός των χαρακτηριστικών είναι πολύ μεγάλος και προκαλεί αρκετή σύγχυση, τόσο ως προς τον ορθό χειρισμό τους, όσο και ως προς την εξεύρεση συσχετίσεων μεταξύ τους [19].

Η παρούσα διπλωματική εργασία δεν αποτελεί την εξαίρεση στον κανόνα. Μέσα από το λογισμικό WEKA, στο οποίο αποθηκεύσαμε περίπου 400 χαρακτηριστικά (attributes), θέλουμε να εξάγουμε κανόνες συσχετίσεων μεταξύ των χαρακτηριστικών, ούτως ώστε να καταλήξουμε σε κάποια συμπεράσματα αναφορικά με το πώς επηρεάζει (λίγο, πολύ ή καθόλου) ο κάθε οικονομικός δείκτης OGGI τον άλλο. Για να εξαχθούν όμως οι κανόνες αυτοί, πρέπει να αποκτήσουμε πρώτα μια γενική εικόνα του πόσο σχετίζονται οι «μεταβλητές - στόχοι» με τα υπόλοιπα χαρακτηριστικά και κατά πόσον επηρεάζονται από αυτά. Αυτό μπορεί να μας το προσφέρει το data mining, μέσω κατάλληλων αλγορίθμων της επιλογής «select attribute» (στην περίπτωση μας ο αλγόριθμος που χρησιμοποιείται είναι ο CFsSubsetEval). Αφού γίνει η επιλογή του αλγορίθμου και καθοριστεί το χαρακτηριστικό που θα αποτελεί την εκάστοτε «μεταβλητή - στόχο», ξεκινά η διαδικασία επιλογής των σχετικότερων μεταξύ τους χαρακτηριστικών από το λογισμικό που χρησιμοποιούμε. Το λογισμικό (εδώ WEKA), μέσα στα πλαίσια της Τεχνητής Νοημοσύνης που εφαρμόζει, μας παρουσιάζει εν τέλει τα «συμπεράσματά» του, επιλέγοντας ένα αριθμό χαρακτηριστικών, τα οποία μας «δηλώνει» (αριθμητικά ως σύνολο και ονομαστικά ένα - ένα) ότι σχετίζονται άμεσα με το χαρακτηριστικό που θέσαμε προηγουμένως ως «μεταβλητή - στόχο». Φυσικά στο σημείο αυτό χρειάζεται να γίνει μία επισήμανση. Ποτέ τα αποτελέσματα μιας διαδικασίας, η οποία εφαρμόζει τεχνικές Τεχνητής Νοημοσύνης και πιο συγκεκριμένα Data Mining, δε γίνονται αποδεκτά «τυφλά», αλλά τυγχάνουν επεξεργασίας και διήθησης από την ανθρώπινο παράγοντα και απομένουν αυτά που πραγματικά έχουν νόημα [20]. Τα αποτελέσματα του σταδίου αυτού είναι χρήσιμα πλέον δεδομένα για το επόμενο στάδιο.

6.4] Επιλογή κατάλληλων αλγορίθμων για περαιτέρω επεξεργασία των δεδομένων.

Στο τελικό στάδιο επιλέγουμε κατάλληλους αλγόριθμους για να επεξεργαστούν τα δεδομένα που προκύπτουν από το στάδιο επιλογής των σχετικών μεταξύ τους χαρακτηριστικών. Στην ουσία γίνεται μία ταξινόμηση (classification) των δεδομένων, με σκοπό την τελική εξαγωγή κανόνων για τα διάφορα χαρακτηριστικά.

Η ταξινόμηση αποτελεί, ίσως, την πιο δημοφιλή και αποτελεσματική μέθοδο της τεχνικής data mining [17]. Οι αλγόριθμοι ταξινόμησης εφαρμόζονται σε δεδομένα τα οποία έχουν προταξινομηθεί σε συγκεκριμένες κλάσεις, με στόχο την εξαγωγή κανόνων οι οποίοι μπορεί μετέπειτα να χρησιμοποιηθούν για ταξινόμηση νέων δεδομένων στις ίδιες κλάσεις. Ένα σύνολο εξαγόμενων κανόνων ονομάζεται ταξινομητής (classifier) [21]. Στην παρούσα εργασία χρησιμοποιούνται δύο συγκεκριμένοι αλγόριθμοι, ο JRip και ο PART. Η λειτουργία των αλγορίθμων ταξινόμησης αποτελείται από τα εξής στάδια :

(α) Εισάγουμε ένα σύνολο από δεδομένα (data set) σε έναν αλγόριθμο ταξινόμησης.

(β) Ο αλγόριθμος έπειτα, με τη σειρά του, «κατανοεί» τους κανόνες βάση των οποίων ταξινομήθηκαν τα δεδομένα.

(γ) Ακολούθως, βάση των συγκεκριμένων κανόνων, ο αλγόριθμος έχει την ικανότητα να ταξινομήσει νέα δεδομένα.

Υπάρχουν δύο βασικοί τύποι αλγορίθμων ταξινόμησης, ανάλογα με το είδος του ταξινομητή που παράγουν :

(i) Οι αλγόριθμοι που παράγουν λίστες αποφάσεων.

(ii) Οι αλγόριθμοι που παράγουν δέντρα αποφάσεων.

Οι αλγόριθμοι JRip και PART είναι αλγόριθμοι που παράγουν λίστες αποφάσεων, έχουν τη μορφή λογικών κανόνων που βγάζουν ανάλογα συμπεράσματα και είναι μία σχετικά νέα μορφή αλγορίθμων [19]. Τα αποτελέσματα της ταξινόμησης που γίνεται στο στάδιο αυτό είναι χρήσιμα στην τελική αποτίμηση όλων των χαρακτηριστικών μεταξύ τους και στην εξαγωγή κανόνων γι' αυτά.

6.5] Παράδειγμα επεξεργασίας ενός χαρακτηριστικού μέσα από τα διάφορα στάδια της μεθοδολογίας.

Προκειμένου να γίνουν πιο σαφείς οι διαδικασίες που ακολουθούνται στα διάφορα στάδια της τεχνικής Data Mining, θα παρατεθεί στο σημείο αυτό το παράδειγμα επεξεργασίας ενός χαρακτηριστικού (attribute). Το λογισμικό (software) το οποίο χρησιμοποιήσαμε για να εφαρμόσουμε την τεχνική Data Mining ήταν, όπως προαναφέρθηκε, το WEKA. Μέσα από τις διάφορες λειτουργίες και τους αλγόριθμους που χρησιμοποιήθηκαν στο WEKA και παρουσιάζονται παρακάτω, γίνεται πιο κατανοητή η επεξεργασία που τυγχάνουν τα διάφορα δεδομένα μέσα στα πλαίσια της μεθοδολογίας που εφαρμόζει η τεχνική Data Mining.

Το χαρακτηριστικό που επιλέχθηκε να αναλυθεί παρακάτω ως παράδειγμα είναι ένας από τους σπουδαιότερους οικονομικούς δείκτες, το Ακαθάριστο Εγχώριο Προϊόν (Gross Domestic Product) ή, συντομογραφικά, το Α.Ε.Π. (G.D.P.). Ως Α.Ε.Π. ορίζεται το σύνολο όλων των προϊόντων και αγαθών που παράγει μια οικονομία εκφρασμένο σε χρηματικές μονάδες. Με άλλα λόγια είναι η συνολική αξία όλων των τελικών αγαθών (υλικών και άυλων) που παράχθηκαν εντός μιας χώρας σε διάστημα ενός έτους, ακόμα και αν μέρος αυτού παράχθηκε από παραγωγικές μονάδες που ανήκουν σε κατοίκους του εξωτερικού.³⁴ Παρακάτω γίνεται μια ανάλυση επεξεργασίας του attribute «Α.Ε.Π.» από το στάδιο 3.2. - Προεπεξεργασία Δεδομένων (preprocessing) μέχρι το τελικό στάδιο 3.4. - Επιλογή και «τρέξιμο» κατάλληλων αλγορίθμων για περαιτέρω επεξεργασία των δεδομένων, μέσω του λογισμικού WEKA.

Ανάλυση επεξεργασίας του attribute «Α.Ε.Π.» μέσω του λογισμικού WEKA :

1^ο Στάδιο - Προεπεξεργασία Δεδομένων :

Θέτουμε το Α.Ε.Π. ως «μεταβλητή - στόχο» (class). Ακολούθως, γίνεται μία διακριτοποίηση³⁵ (discretize) των χαρακτηριστικών, μέσω της επιλογής κατάλληλου φίλτρου (filters → unsupervised → attribute → discretize). Η διακριτοποίηση ολοκληρώνεται με την επιλογή «apply», όπως φαίνεται και στο παρακάτω σχήμα.

³⁴ Βλ. «Ακαθάριστο Εγχώριο Προϊόν, Από τη Βικιπαίδεια, την ελεύθερη εγκυκλοπαίδεια», <[http:// el.wikipedia.org/wiki](http://el.wikipedia.org/wiki)>, 11/6/2009.

³⁵ Βλ. σχετικά με τον όρο το στάδιο 6.2. της Μεθοδολογίας.

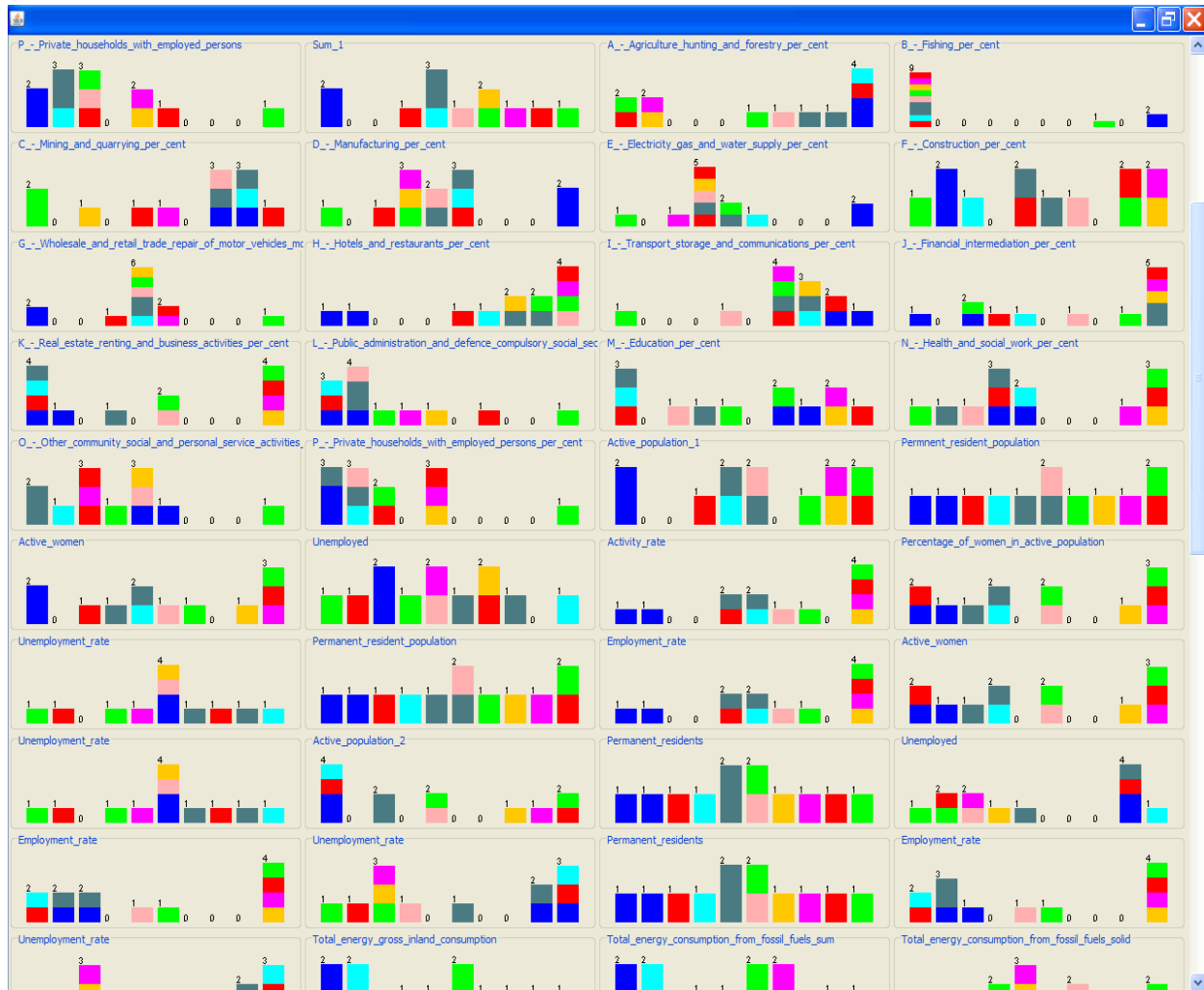
The screenshot shows the Weka Explorer interface. The 'Filter' pane on the left has 'Discretize' selected under the 'attribute' category. The 'Selected attribute' pane shows 'Name: year' with 'Type: Nominal' and 'Unique: 8 (67%)'. Below this is a table of labels and counts:

Label	Count
{-inf,-1997.1]}	2
{1997.1-1998.2]}	1
{1998.2-1999.3]}	1
{1999.3-2000.4]}	1
{2000.4-2001.5]}	1
{2001.5-2002.6]}	1
{2002.6-2003.7]}	1
{2003.7-2004.8]}	1
{2004.8-2005.9]}	1
{2005.9-inf}	2

The 'Class' dropdown is set to 'Gross Domestic Product_GDP (Nom)'. The 'Visualize All' button is highlighted. The main visualization area shows a bar chart with 10 bars of different colors (blue, red, cyan, grey, pink, green, yellow, magenta, red, red). The status bar at the bottom shows 'Log' and 'x0'.

Σχήμα 3.5.α. : Προεπεξεργασία Δεδομένων, Διακριτοποίηση και Επιλογή G.D.P. ως μεταβλητή - στόχος (class) στο WEKA.

Τελική φάση του σταδίου αυτού είναι ο καθαρισμός των χαρακτηριστικών, μέσω της οπτικοποίησής τους (visualize). Με την οπτικοποίηση των χαρακτηριστικών αποκτά δηλαδή ο χρήστης μια πρώτη εποπτική ματιά στα χαρακτηριστικά και τις σχέσεις μεταξύ τους, μέσω της γραφικής απεικόνισής τους.

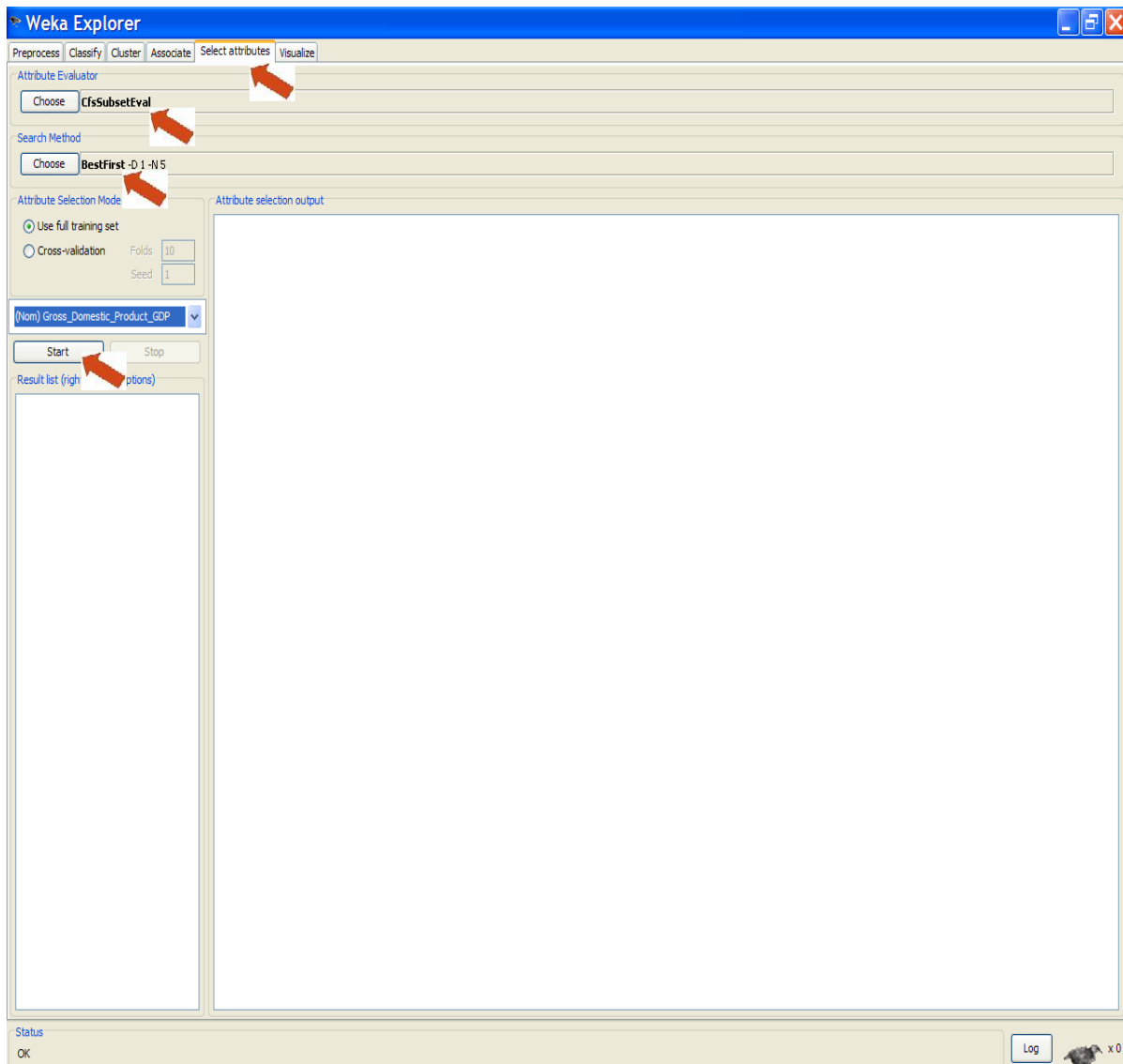


Σχήμα 3.5.β. : Οπτικοποίηση Δεδομένων στο WEKA (επιλογή «visualize all»).

2^ο Στάδιο - Επιλογή των πιο σχετικών με το «G.D.P.» χαρακτηριστικών :

Μετά την προεπεξεργασία των δεδομένων και τη διακριτοποίηση που προηγήθηκε (1^ο στάδιο), γίνεται η επιλογή των πιο σχετικών με το «G.D.P.» χαρακτηριστικών, μέσω της επιλογής του WEKA «select attribute». Επιλέγονται οι κατάλληλοι αλγόριθμοι. Ως εκτιμητής χαρακτηριστικών (attribute evaluator) επιλέγεται ο αλγόριθμος «CFsSubsetEval», ο οποίος αξιολογεί ένα υποσύνολο χαρακτηριστικών με την εξέταση της μεμονωμένης προβλεπτικής δυνατότητας κάθε χαρακτηριστικού γνωρίσματος μαζί με το βαθμό πλεονασμού μεταξύ τους. Ως μέθοδος αναζήτησης (search method) επιλέγεται η μέθοδος «BestFirst», η οποία διερευνά το διάστημα των υποσυνόλων των χαρακτηριστικών ακολουθώντας το σύστημα «σκαρφαλώματος του λόφου» (hillclimbing), δηλαδή από κάτω προς τα πάνω. Αφού γίνει η επιλογή των αλγορίθμων και καθοριστεί το χαρακτηριστικό που θα αποτελεί τη «μεταβλητή - στόχο», το

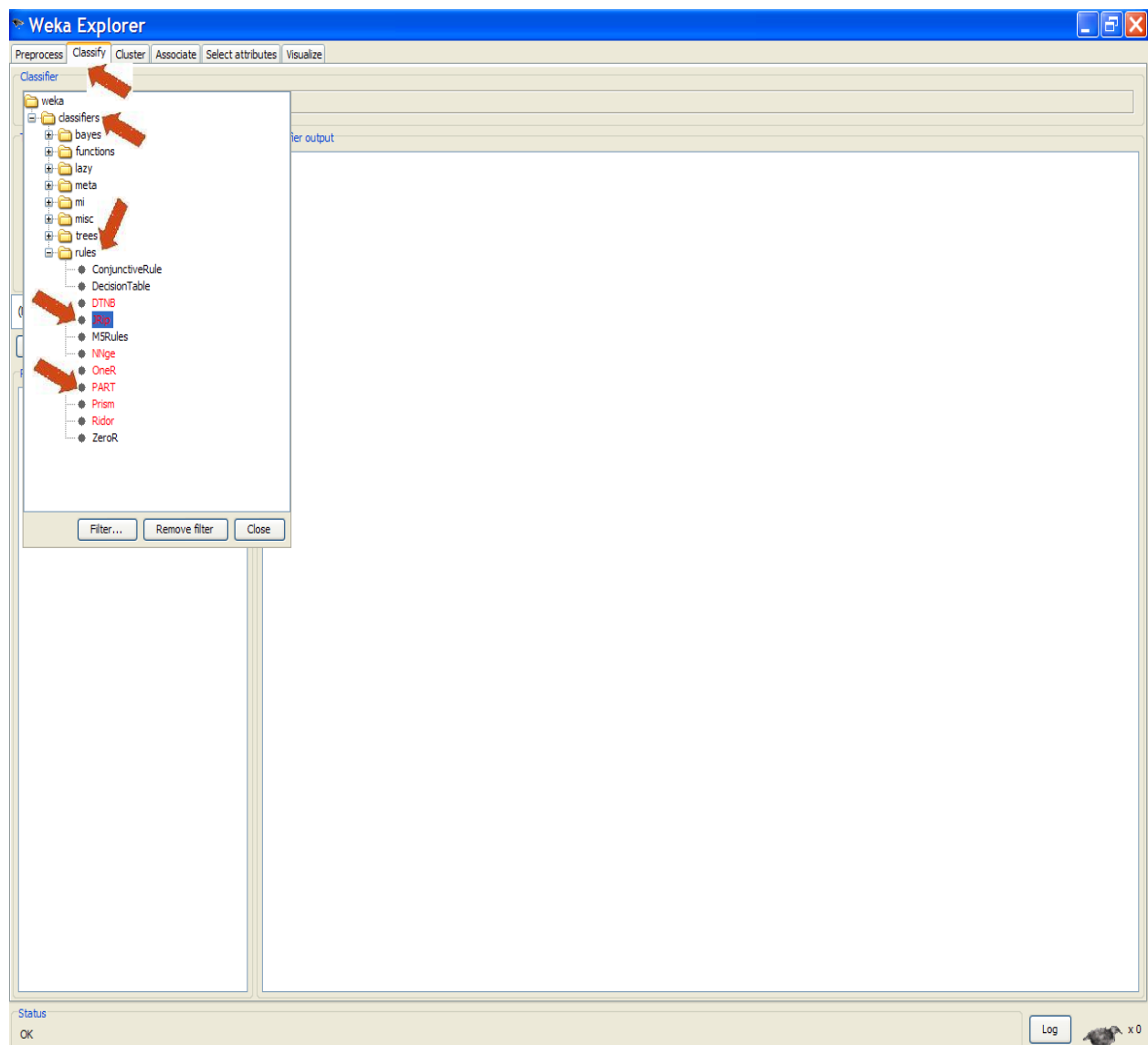
Gross Domestic Product, ξεκινά (start) η διαδικασία επιλογής των σχετικότερων με το G.D.P. χαρακτηριστικών από το λογισμικό.



Σχήμα 3.5.γ. : Η επιλογή «select attributes» και οι κατάλληλοι αλγόριθμοι για επιλογή των πιο σχετικών με το G.D.P. χαρακτηριστικών, όπως εμφανίζονται στο WEKA.

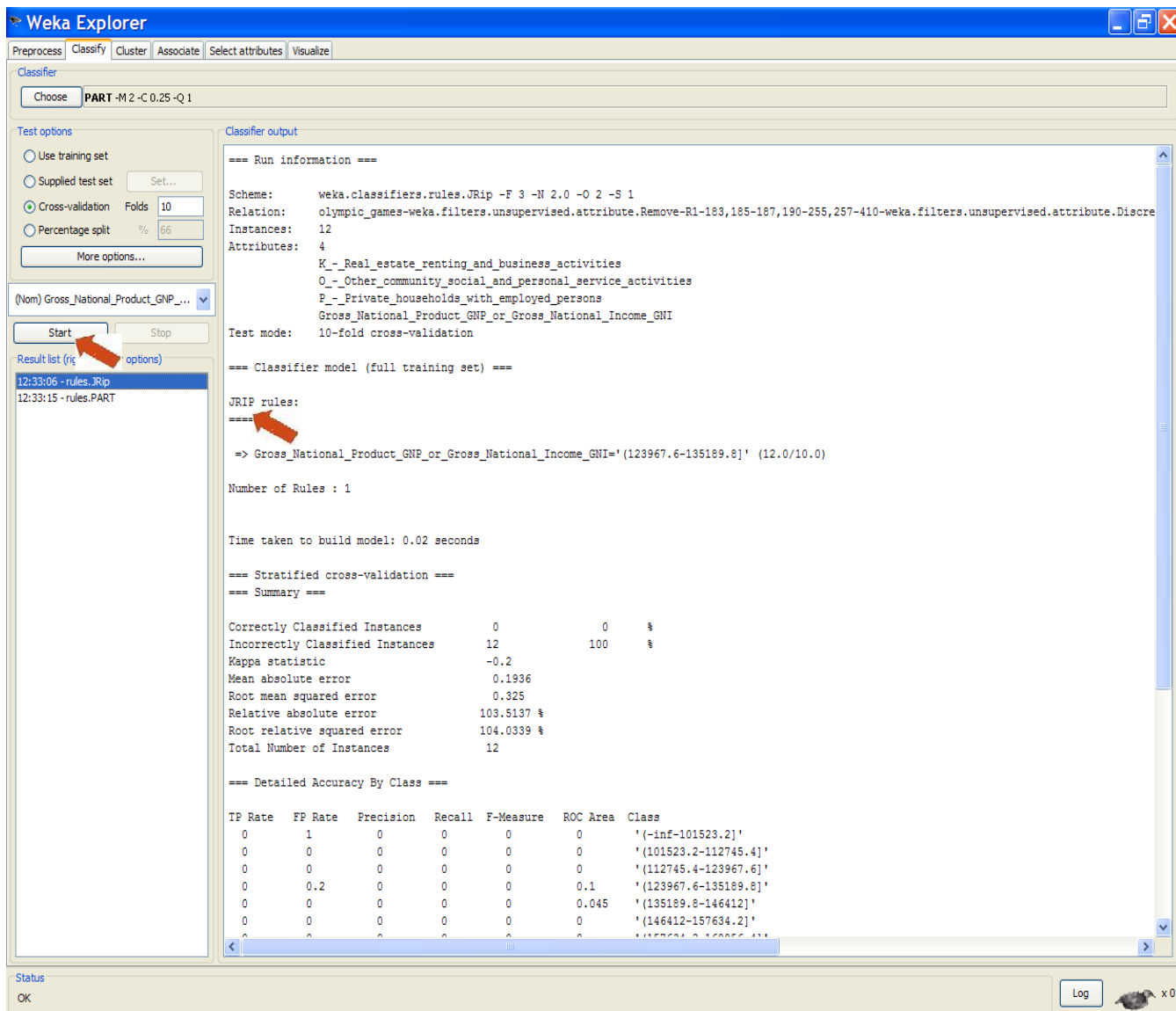
Το WEKA, μέσα στα πλαίσια του Data Mining που εφαρμόζει, μας παρουσιάζει εν τέλει τα «συμπεράσματά» του, επιλέγοντας ένα αριθμό χαρακτηριστικών, τα οποία μας «δηλώνει» (αριθμητικά ως σύνολο και ονομαστικά ένα - ένα) ότι σχετίζονται άμεσα με το χαρακτηριστικό «G.D.P.».

εξαγωγή κανόνων για τα διάφορα χαρακτηριστικά. Πριν επιλεγούν οι αλγόριθμοι αυτοί, επιστρέφουμε στην επιλογή «preprocess» του WEKA και επιλέγουμε ένα πιο μικρό data set (σύνολο δεδομένων), με το οποίο θα καταλήξουμε σε κανόνες ταξινόμησης. Το νέο data set αποτελείται, πλέον, μόνο από τα χαρακτηριστικά (attributes) εκείνα που επιλέγηκαν στο 2^ο στάδιο της επεξεργασίας (selected attributes). Ακολούθως, από την επιλογή «classify» του WEKA διαλέγουμε τον αλγόριθμο - ταξινομητή που επιθυμούμε να επεξεργαστεί τα δεδομένα και να παράξει κανόνες ταξινόμησης (classifiers → rules → JRip ή PART) και στη συνέχεια καθορίζουμε τη «μεταβλητή - στόχο» που θέλουμε, επιλέγοντας κάθε φορά ένα από τα selected attributes για το G.D.P. και θέτοντάς το ως «μεταβλητή - στόχο».



Σχήμα 3.5.ε. : Η επιλογή «classify» (ταξινόμηση) και οι αλγόριθμοι - ταξινομητές στο WEKA.

Αφού γίνει η επιλογή των αλγορίθμων και καθοριστεί το χαρακτηριστικό που θα αποτελεί τη «μεταβλητή - στόχο» κάθε φορά, ξεκινά (start) η διαδικασία ταξινόμησης μεταξύ των χαρακτηριστικών του νέου data set και εμφανίζονται στο πλάι (classifier output) οι διάφοροι κανόνες που προκύπτουν, όπως φαίνεται στο σχήμα παρακάτω.



```
=== Run information ===
Scheme:      weka.classifiers.rules.JRip -F 3 -N 2.0 -O 2 -S 1
Relation:    olympic_games-weka.filters.unsupervised.attribute.Remove-R1-183,185-187,190-255,257-410-weka.filters.unsupervised.attribute.Discretize
Instances:   12
Attributes:  4
             K_ Real_estate_renting_and_business_activities
             O_ Other_community_social_and_personal_service_activities
             P_ Private_households_with_employed_persons
             Gross_National_Product_GNP_or_Gross_National_Income_GNI
Test mode:   10-fold cross-validation

=== Classifier model (full training set) ===

JRIP rules:
=====
=> Gross_National_Product_GNP_or_Gross_National_Income_GNI='(123967.6-135189.8)!' (12.0/10.0)

Number of Rules : 1

Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      0          0 %
Incorrectly Classified Instances    12         100 %
Kappa statistic                     -0.2
Mean absolute error                 0.1936
Root mean squared error             0.325
Relative absolute error             103.5137 %
Root relative squared error        104.0339 %
Total Number of Instances          12

=== Detailed Accuracy By Class ===

TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
0         1         0          0        0          0         '(-inf-101523.2)!'
0         0         0          0        0          0         '(101523.2-112745.4)!'
0         0         0          0        0          0         '(112745.4-123967.6)!'
0         0.2       0          0        0          0.1       '(123967.6-135189.8)!'
0         0         0          0        0          0.045     '(135189.8-146412)!'
0         0         0          0        0          0         '(146412-157634.2)!'
0         0         0          0        0          0         '(157634.2-168856.4)!'

Status
OK
```

Σχήμα 3.5.στ. : Παρουσίαση κανόνων ταξινόμησης (εδώ JRip Rules).

Μετά την όλη επεξεργασία ενός χαρακτηριστικού (attribute) ακολουθεί η παρουσίαση και ανάλυση των αποτελεσμάτων και, βεβαίως, η ερμηνεία τους, ώστε να εξαχθούν κάποια συμπεράσματα και να γίνουν οι οποιεσδήποτε παρατηρήσεις προκύπτουν. Τα αποτελέσματα του data mining πρέπει να «εξεταστούν» πλέον από τον ανθρώπινο παράγοντα.

7] Παρουσίαση και Ανάλυση Αποτελεσμάτων.

Μέσα από την τεχνική data mining και πιο συγκεκριμένα την επεξεργασία των σπουδαιότερων χαρακτηριστικών (attributes) από το λογισμικό WEKA, προκύπτουν κάποια αποτελέσματα «τεχνητής νοημοσύνης», τα οποία τυγχάνουν μεθοδικής ανάλυσης και επεξεργασίας, ώστε να βγουν κάποια συμπεράσματα που να τυγχάνουν φυσικής και χρησιμοποιήσιμης σημασίας. Τα αποτελέσματα αυτά παρουσιάζονται και αναλύονται στο κεφάλαιο αυτό, ενώ παρατίθενται ταυτόχρονα και οι όποιες παρατηρήσεις και τα όποια συμπεράσματα εξάγονται από αυτά. Βέβαια, εδώ πρέπει να τονιστεί ότι κατά την πορεία ανάλυσης των αποτελεσμάτων αυτών, συναντήθηκαν αρκετά προβλήματα επακριβούς κατανόησης αρκετών από αυτά, εξαιτίας, κυρίως, του γεγονότος ότι πολλές φορές έπρεπε να ανιχνεύσουμε πληροφορίες κρυμμένες που προκαλούσαν αβεβαιότητα ως προς τη φυσική τους έννοια, λόγω των ελάχιστων υποδειγμάτων (instances) που είχαμε στο σύνολο των δεδομένων (data set). Αυτό ήταν αναμενόμενο, αφού θα έπρεπε η αναλογία «υποδείγματα / χαρακτηριστικά» (instances / attributes) να είναι περίπου 10 / 1. Το να έχουμε τη σωστή αναλογία «υποδείγματα / χαρακτηριστικά = 10 / 1» καθίσταται ιδιαίτερα δύσκολο, αφού τότε, υπολογίζοντας και τον αριθμό των χαρακτηριστικών (περίπου 375), θα οδηγούμασταν στα 3750 υποδείγματα, κάτι που θα σήμαινε ότι ως υποδείγματα θα έπρεπε να τεθούν οι μήνες ή ακόμα και οι εβδομάδες της περιόδου 1996 - 2007, τα αναλυτικά στοιχεία των οποίων είναι ακόμη πιο δυσεύρετα. Παρόλ' αυτά, σκοπός δεν είναι η οικονομική ανάλυση όλων των δεικτών που χρησιμοποιήθηκαν από το WEKA ως χαρακτηριστικά, αλλά η εξόρυξη κανόνων και προτύπων (patterns) που πηγάζουν από την ανάλυση των δεικτών αυτών μέσω data mining. Δεν παρουσιάζονται στο κεφάλαιο αυτό όλοι οι κανόνες που προέκυψαν, αλλά οι σπουδαιότεροι. Ίσως κάποιοι κανόνες απ' αυτούς αποτελέσουν το εφελήριο λάκτισμα για άλλες μελέτες των δεικτών OGGI, που μπορεί να επακολουθήσουν, στα πλαίσια της εξόρυξης γνώσης μέσα από δεδομένα.

Παρακάτω ορίζονται τα 15 χαρακτηριστικά που θεωρήθηκαν σπουδαιότερα και έτυχαν περαιτέρω επεξεργασίας³⁶, αναλύονται τα αποτελέσματα που προκύπτουν από την επεξεργασία τους (επιλογή των πιο σχετικών μεταξύ τους χαρακτηριστικών, κανόνες ταξινόμησης) και καταγράφονται κάποιοι κανόνες. Εδώ παρουσιάζονται οι σπουδαιότεροι αυτών.

³⁶ Βλ. σχετικά το κεφάλαιο 6.2. της Μεθοδολογίας.

7.1] Gross Domestic Product (G.D.P.) - Ακαθάριστο Εγχώριο Προϊόν (Α.Ε.Π.).

7.1.1. Ορισμός.

Ως Α.Ε.Π. ορίζεται το σύνολο όλων των προϊόντων και αγαθών που παράγει μια οικονομία εκφρασμένο σε χρηματικές μονάδες. Με άλλα λόγια είναι η συνολική αξία όλων των τελικών αγαθών (υλικών και άυλων) που παράχθηκαν εντός μιας χώρας σε διάστημα ενός έτους, ακόμα και αν μέρος αυτού παράχθηκε από παραγωγικές μονάδες που ανήκουν σε κατοίκους του εξωτερικού.³⁷

7.1.2. Επιλογή των πιο σχετικών, με το Α.Ε.Π., χαρακτηριστικών (selected attributes).

Τα αποτελέσματα που παρουσιάζουν τα σχετικότερα μεταξύ τους χαρακτηριστικά ίσως να αποτελούν και τα σπουδαιότερα στοιχεία που προκύπτουν από μια ανάλυση δεδομένων με τεχνικές data mining [19]. Στις περισσότερες εφαρμογές data mining ο αριθμός των χαρακτηριστικών είναι πολύ μεγάλος και προκαλεί αρκετή σύγχυση, τόσο ως προς τον ορθό χειρισμό τους όσο και ως προς την εξεύρεση συσχετίσεων μεταξύ τους. Το λογισμικό (εδώ WEKA), μέσα στα πλαίσια του Data Mining που εφαρμόζει, μας παρουσιάζει εν τέλει τα «συμπεράσματά» του, επιλέγοντας ένα αριθμό χαρακτηριστικών, τα οποία μας «δηλώνει» (αριθμητικά ως σύνολο και ονομαστικά ένα - ένα) ότι σχετίζονται άμεσα με το χαρακτηριστικό που εξετάζουμε. Έτσι, αναγκαζόμαστε να αναλύσουμε και συσχετίσεις μεταξύ χαρακτηριστικών - μη ορατές μέχρι πρότινος -, οι οποίες μάλιστα μπορεί να αποδειχθούν σημαντικές.

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με το Α.Ε.Π., όπως προκύπτουν από την όλη διαδικασία που ακολουθήθηκε στην παρούσα εργασία, είναι τα εξής τέσσερα :

- ✓ Real estate renting and business activities - Διαχείριση ακίνητης περιουσίας, εκμισθώσεις και επιχειρηματικές δραστηριότητες (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στους συγκεκριμένους τομείς)
- ✓ Private households with employed persons - Ιδιωτικά νοικοκυριά που απασχολούν οικιακό προσωπικό (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)

³⁷ Βλ. «Ακαθάριστο Εγχώριο Προϊόν, Από τη Βικιπαίδεια, την ελεύθερη εγκυκλοπαίδεια», <<http://el.wikipedia.org/wiki/>> , 20/6/2009.

- ✓ Other community social and personal service activities - Παραγωγή άλλων υπηρεσιών υπέρ του κοινωνικού συνόλου (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Gross National Product (G.N.P.) or Gross National Income (G.N.I.) - Ακαθάριστο Εθνικό Προϊόν (Α.ΕΘ.Π.) ή Ακαθάριστο Εθνικό Εισόδημα (Α.ΕΘ.Ε.) (ορίζεται στο 7.2)

Η εύρεση του Α.ΕΘ.Π. ως μεγέθους συσχετιζόμενου του Α.Ε.Π. είναι κάτι που μπορεί να χαρακτηριστεί, εν μέρει, ως αναμενόμενο, αφού τα δύο μεγέθη αλληλοσυσχετίζονται μόνο και μόνο από τη φυσική τους σημασία. Ωστόσο, η συσχέτιση του μεγέθους της διαχείρισης της ακίνητης περιουσίας, των εκμισθώσεων και των επιχειρηματικών δραστηριοτήτων με το Α.Ε.Π. δεν έχει ξεκάθαρο νόημα, αλλά παρουσιάζει ιδιαίτερο ενδιαφέρον. Παρόλ' αυτά, προβληματισμό στις συσχετίσεις αυτές προκαλεί ο περιορισμένος αριθμός δεδομένων. Ανάλογα συμπεράσματα εξάγονται από όλα τα χαρακτηριστικά που αναλύονται παρακάτω.

7.1.3. Κανόνες Ταξινόμησης.

Μετά την επιλογή των πιο σχετικών, με το κάθε χαρακτηριστικό που θεωρείται «μεταβλητή - στόχος», χαρακτηριστικών, πραγματοποιείται μία ταξινόμηση (classification) των δεδομένων, με σκοπό την τελική εξαγωγή κανόνων για τα διάφορα χαρακτηριστικά.³⁸ Οι αλγόριθμοι ταξινόμησης εφαρμόζονται σε δεδομένα τα οποία έχουν προταξινομηθεί σε συγκεκριμένες κλάσεις, με στόχο την εξαγωγή κανόνων, οι οποίοι μπορεί μετέπειτα να χρησιμοποιηθούν για ταξινόμηση νέων δεδομένων στις ίδιες κλάσεις. Ένα σύνολο εξαγόμενων κανόνων ονομάζεται ταξινομητής (classifier). Στην παρούσα εργασία χρησιμοποιούνται δυο συγκεκριμένοι αλγόριθμοι, ο JRip και ο PART. Οι συγκεκριμένοι αλγόριθμοι JRip και PART παράγουν λίστες αποφάσεων και έχουν τη μορφή λογικών κανόνων που βγάζουν ανάλογα συμπεράσματα. Όλα τα αποτελέσματα που προκύπτουν από την ταξινόμηση των χαρακτηριστικών, παρατίθενται στο Παράρτημα, στο τέλος της παρούσας εργασίας. Ωστόσο, οι σημαντικότεροι εκ των κανόνων ταξινόμησης, παρουσιάζονται στο παρόν στάδιο. Όλοι αυτοί οι κανόνες που αναφέρονται, περικλείουν ασάφειες και προκαλούν ιδιαίτερες αμφιβολίες, σχετικά με την πραγματική τους ερμηνεία. Σκοπός δεν είναι η εξαγωγή στερεότυπων κανόνων (rules) και προτύπων (patterns), αλλά η συγκέντρωση κάποιων «ευέλικτων» κανόνων, οι οποίοι προέκυψαν μέσω μίας μεθοδευμένης ανάλυσης data mining και οι οποίοι μπορούν να

³⁸ Βλ. σχετικά στην παρούσα εργασία : Κεφάλαιο 6.5, «3^ο Στάδιο - Επιλογή και «τρέξιμο» κατάλληλων αλγορίθμων για περαιτέρω επεξεργασία των δεδομένων, μέσω του λογισμικού WEKA», της Μεθοδολογίας.

επεκταθούν σε ένα νέο σύνολο δεδομένων (data set), ώστε να βοηθήσουν στην εξαγωγή χρήσιμων συμπερασμάτων.

Από το Α.Ε.Π. και τα συσχετιζόμενα με αυτό χαρακτηριστικά δεν προκύπτουν οποιοδήποτε «χρήσιμοι» κανόνες ταξινόμησης (JRip ή PART). Στο σημείο αυτό επιβάλλεται η διευκρίνιση των όρων «χρήσιμος» και λιγότερο σημαντικός» κανόνας που θα χρησιμοποιούνται από εδώ και στο εξής. Όλοι οι κανόνες είναι σημαντικοί στο βαθμό που ορίζει το σφάλμα τους, απλά κάποιοι είναι μεγαλύτερης χρησιμότητας και εφαρμοσιμότητας από τους υπόλοιπους. Παρακάτω παρουσιάζεται ένας «λιγότερο σημαντικός» κανόνας, ούτως ώστε να γίνει κατανοητή η έννοια του «χρήσιμου» κανόνα.

=> Gross National Product GNP or Gross National Income GNI='(123967.6-135189.8)'
(12.0/10.0)

Correctly Classified Instances	0	0	%
Incorrectly Classified Instances	12	100	%
Total Number of Instances	12		

Όπως βλέπουμε έχουμε στο σύνολο 12 υποδείγματα (Total Number of Instances). Από αυτά 0 υποδείγματα έχουν ταξινομηθεί «ορθά» (Correctly Classified Instances), ενώ 12 ταξινομήθηκαν «λανθασμένα» (Incorrectly Classified Instances). Εμείς, λαμβάνοντας υπόψη το σύνολο των κανόνων που προέκυψαν, επιλέγουμε μόνο τους «χρήσιμους» κανόνες, δηλαδή τους κανόνες των οποίων το σύνολο των «ορθά ταξινομημένων υποδειγμάτων» (Correctly Classified Instances) είναι ίσο ή μεγαλύτερο από έξι. Βέβαια, όσο μεγαλύτερος είναι ο αριθμός των «ορθά ταξινομημένων υποδειγμάτων», τόσο πιο «χρήσιμος» είναι ο κανόνας ταξινόμησης που προκύπτει. Κανόνας ταξινόμησης με 6 «ορθά ταξινομημένα υποδείγματα» (minimum) είναι κατά 50 % «χρήσιμος», ενώ κανόνας ταξινόμησης με 12 «ορθά ταξινομημένα υποδείγματα» (maximum) θεωρείται 100 % «χρήσιμος». Στην παρουσίαση των «χρήσιμων» κανόνων ταξινόμησης, που επιχειρείται στο παρόν στάδιο, αναφέρονται οι κανόνες με τουλάχιστον 6 «ορθά ταξινομημένα υποδείγματα».

7.2] Gross National Product (G.N.P.) - Ακαθάριστο Εθνικό Προϊόν (Α.ΕΘ.Π.).

7.2.1. Ορισμός.

Ως Ακαθάριστο Εθνικό Προϊόν (Α.ΕΘ.Π.) ή Ακαθάριστο Εθνικό Εισόδημα (Α.ΕΘ.Ε.) ορίζεται το συνολικό Προϊόν ή Εισόδημα που αποκτούν οι κάτοικοι μιας χώρας ανεξάρτητα του σε ποια χώρα το αποκτούν. Με άλλα λόγια είναι η συνολική αξία όλων των τελικών αγαθών (υλικών και άυλων) που αποκτούν οι κάτοικοι μιας χώρας σε διάστημα ενός έτους. Διαφέρει από το Ακαθάριστο Εγχώριο Προϊόν στο ότι συμπεριλαμβάνει και το εισόδημα που απέκτησαν οι κάτοικοι μιας χώρας στο εξωτερικό και το αποστέλλουν στη χώρα, ενώ δε συμπεριλαμβάνει το τμήμα του εισοδήματος που δημιουργείται μέσα στη χώρα από τους αλλοδαπούς και αποστέλλεται στο εξωτερικό.³⁹

7.2.2. Επιλογή των πιο σχετικών, με το Α.ΕΘ.Π., χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με το Α.ΕΘ.Π. είναι τα εξής επτά :

- ✓ Consumer Price Index (C.P.I.) - Δείκτης Τιμών Καταναλωτή (Δ.Τ.Κ.) (ορίζεται στο 7.4)
- ✓ Gross Domestic Product (G.D.P.) - Ακαθάριστο Εγχώριο Προϊόν (Α.Ε.Π.)
- ✓ Interests - Τόκοι (δημοσιονομικά έξοδα)
- ✓ Personnel expenditure - Δαπάνες Προσωπικού (δημοσιονομικά έξοδα)
- ✓ Ministry of Employment and Social Protection - Υπουργείο Απασχόλησης και Κοινωνικής Προστασίας (δημοσιονομικά έξοδα ανά φορέα, σε εκατομμύρια ευρώ)
- ✓ Non-tax income - Μη φορολογικά έσοδα (δημοσιονομικά έσοδα)
- ✓ Total gross debt of public administration - Συνολικό ακαθάριστο χρέος του ευρύτερου δημόσιου τομέα

7.2.3. Κανόνες Ταξινόμησης.

Από το Α.ΕΘ.Π. και τα συσχετιζόμενα με αυτό χαρακτηριστικά δεν προκύπτουν οποιοδήποτε «χρήσιμοι» κανόνες ταξινόμησης (JRip ή PART).

³⁹ Βλ. «Ακαθάριστο Εθνικό Προϊόν, Από τη Βικιπαίδεια, την ελεύθερη εγκυκλοπαίδεια», <<http://el.wikipedia.org/wiki/>> , 20/6/2009.

7.3] Gini Index - Δείκτης Εισοδηματικής Κατανομής (Δ.Ε.Κ.).

7.3.1. Ορισμός.

Ο Δείκτης Εισοδηματικής Κατανομής Gini αποτελεί στατιστικό δείκτη μέτρησης της διασποράς των εισοδημάτων των εργαζόμενων σε μια χώρα. Χρησιμοποιείται, συνήθως, ως μέτρο της εισοδηματικής ανισότητας που υπάρχει σε μια χώρα. Αποτελεί μια αναλογία και μπορεί να κυμανθεί από 0 έως 1 (0% - 100%). Ένας χαμηλός δείκτης Gini δείχνει μεγαλύτερη ισοκατανομή εισοδήματος ή πλούτου, με το 0 να αντιστοιχεί στην τέλεια ισοκατανομή (όλοι λαμβάνουν το ίδιο ακριβώς εισόδημα), ενώ οι υψηλότεροι δείκτες Gini δείχνουν άνιση κατανομή εισοδήματος, με το 1 να αντιστοιχεί στην τέλεια ανισοκατανομή (μια κατάσταση με περισσότερα από ένα άτομα, όπου ένα άτομο λαμβάνει όλο το εισόδημα).⁴⁰

7.3.2. Επιλογή των πιο σχετικών, με το Δ.Ε.Κ. Gini, χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με το Δείκτη Εισοδηματικής Κατανομής Gini είναι τα εξής πέντε :

- ✓ Fossil fuels - Φυσικά καύσιμα (κατανάλωση ενέργειας στην Ελλάδα, ανά πρωτογενή πηγή)
- ✓ Vacation centers - Παραθεριστικά κέντρα (βαθμός πληρότητας σε ένα χρόνο)
- ✓ Domestic users - Οικιακοί χρήστες (τιμή αγοράς ηλεκτρικής ενέργειας σε € / kWh, μόνο για οικιακή χρήση)
- ✓ Ministry of National Defense - Υπουργείο Εθνικής Άμυνας (δημοσιονομικά έξοδα ανά φορέα, σε εκατομμύρια ευρώ)
- ✓ Ministry of Development - Υπουργείο Ανάπτυξης (δημοσιονομικά έξοδα ανά φορέα, σε εκατομμύρια ευρώ)

7.3.3. Κανόνες Ταξινόμησης.

Από το Δ.Ε.Κ. Gini και τα συσχετιζόμενα με αυτόν χαρακτηριστικά δεν προκύπτουν οποιοδήποτε κανόνες ταξινόμησης από τον αλγόριθμο JRip. (JRip rules). Αντιθέτως προκύπτουν δύο κανόνες από τον αλγόριθμο PART (PART decision list).

PART decision list :

⁴⁰ Βλ. «Gini Coefficient, From Wikipedia, the free encyclopedia», <http://en.wikipedia.org/wiki/Gini_coefficient>, 20/6/2009.

1^{ος} κανόνας :

Vacation centers = '(-inf-33.94]': '(-inf-2734.429]' (3.0)

Έχουμε την προϋπόθεση (Vacation centers = '(-inf-33.94]') και το αποτέλεσμα (: '(-inf-2734.429]') του κανόνα. Στο αποτέλεσμα, μετά τα διαλυτικά (:), όλα τα νούμερα αναφέρονται στο χαρακτηριστικό που θέτουμε ως «μεταβλητή - στόχο» κάθε φορά. Εδώ, ως «μεταβλητή - στόχο», θέσαμε τα δημοσιονομικά έξοδα (σε εκατομμύρια ευρώ) του Υπουργείου Εθνικής Άμυνας. Ως εκ τούτου, ο κανόνας αυτός λέει ότι εάν τα παραθεριστικά κέντρα στην Ελλάδα παρουσιάζουν, στο σύνολό τους (συμπεριλαμβάνονται και Έλληνες και αλλοδαποί παραθεριστές), βαθμό πληρότητας μέχρι 33,94%, για ένα χρόνο, τότε τα δημοσιονομικά έξοδα του Υπουργείου Εθνικής Άμυνας είναι το πολύ 2,734 δισεκατομμύρια ευρώ. Ο αριθμός 3 στην παρένθεση δηλώνει ότι τα «ορθά ταξινομημένα υποδείγματα» που ελέγχονται από τον κανόνα αυτό είναι, στο σύνολό τους, 3. Ο αριθμός 0 δεν εμφανίζεται εδώ, αλλά εξυπακούεται και δηλώνει τον αριθμό των «λανθασμένα ταξινομημένων υποδειγμάτων» που ελέγχονται από αυτό τον κανόνα.

2^{ος} κανόνας :

: '(3037.607-3189.196]' (9.0/6.0)

Κανόνες αυτής της μορφής, στους οποίους δεν υπάρχει προϋπόθεση παρά μόνο το αποτέλεσμα, δηλώνουν τυπικά, ότι η τιμή της «μεταβλητής - στόχου» κυμαίνεται στο διάστημα που αναφέρεται κάθε φορά, ανεξάρτητα του τι συμβαίνει με τα υπόλοιπα χαρακτηριστικά. Εδώ, δηλαδή, ο πιο πάνω κανόνας λέει ότι τα δημοσιονομικά έξοδα του Υπουργείου Εθνικής Άμυνας (που είναι η «μεταβλητή - στόχος») κυμαίνονται μεταξύ 3,04 και 3,19 δισεκατομμυρίων ευρώ, ανεξαρτήτως (μη λαμβάνοντας υπόψη) των υπολοίπων χαρακτηριστικών (attributes).

Number of Rules : 2

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	6	50	%
Incorrectly Classified Instances	6	50	%
Kappa statistic	0.3333		
Mean absolute error	0.1292		
Root mean squared error	0.2695		
Relative absolute error	72.109	%	
Root relative squared error	89.0378	%	
Total Number of Instances	12		

Παραπάνω παρατίθενται και κάποια στοιχεία της μεθόδου που ακολουθείται για ταξινόμηση (cross - validation), τα οποία προσδίδουν κάποιες επιπλέον πληροφορίες για τους κανόνες που προκύπτουν κάθε φορά από τους αλγόριθμους JRip ή PART. Τα «ορθά» (correctly) και «λανθασμένα» (incorrectly) ταξινομημένα υποδείγματα (Classified Instances), καθώς επίσης και ο συνολικός αριθμός των υποδειγμάτων (Total Number of Instances) έχουν αναλυθεί εκτενώς στους κανόνες ταξινόμησης του Α.Ε.Π. Το στατιστικό μέγεθος Κάπα (kappa statistic) καταδεικνύει τη συμφωνία πρόβλεψης ενός κανόνα ή χαρακτηριστικού σε σχέση με την πραγματικότητα. Όταν παίρνει την τιμή 1, σημαίνει απόλυτη συμφωνία μεταξύ πρόβλεψης και πραγματικότητας. Τα υπόλοιπα στοιχεία αποτελούν ενδείξεις διάφορων σφαλμάτων (Mean absolute error - μέσο απόλυτο σφάλμα, Root mean squared error - μέσο σφάλμα τετραγωνικής ρίζας, Relative absolute error - σχετικό απόλυτο σφάλμα, Root relative squared error - σχετικό σφάλμα τετραγωνικής ρίζας), οι οποίες δεν τυγχάνουν συστηματικής ανάλυσης στην παρούσα εργασία.

7.4] Consumer Price Index (C.P.I.) - Δείκτης Τιμών Καταναλωτή (Δ.Τ.Κ.).

7.4.1. Ορισμός.

Ο Δείκτης Τιμών Καταναλωτή αποτελεί δείκτη μέτρησης του σταθμισμένου μέσου όρου των τιμών του «καλαθιού της νοικοκυράς» όσον αφορά τα καταναλωτικά αγαθά και τις υπηρεσίες, όπως η μεταφορά, η διατροφή και η ιατρική φροντίδα. Ο Δ.Τ.Κ. ελέγχει τις μεταβολές των τιμών των αγαθών ή των υπηρεσιών, όπως αυτές προκαθορίζονται από το τρεχούμενο «καλάθι της νοικοκυράς», και υπολογίζει το μέσο όρο των μεταβολών αυτών, λαμβάνοντας ταυτόχρονα υπόψη τη βαρύτητα του κάθε αγαθού ή υπηρεσίας, ανάλογα με τη χρησιμότητά τους. Οι αλλαγές στο Δ.Τ.Κ. χρησιμοποιούνται για να αξιολογήσουν τις μεταβολές των τιμών που συνδέονται με το κόστος ζωής, κατά την εκάστοτε περίοδο υπολογισμού του δείκτη. Πολλές φορές ο Δ.Τ.Κ. εκφράζεται και ως «πρωτεύον πληθωρισμός» («headline inflation»). Ο Δ.Τ.Κ. εφαρμόζει μια από τις πλέον χρησιμοποιημένες στατιστικές μεθόδους για τον προσδιορισμό των περιόδων πληθωρισμού ή αντιπληθωρισμού. Μεγάλη αύξηση του Δ.Τ.Κ., κατά τη διάρκεια μιας μικρής χρονικής περιόδου, φανερώνει περίοδο πληθωρισμού, ενώ, αντίθετα, μεγάλη πτώση του Δ.Τ.Κ. καταδεικνύει περίοδο αντιπληθωρισμού.⁴¹

7.4.2. Επιλογή των πιο σχετικών, με το Δ.Τ.Κ., χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με το Δ.Τ.Κ. είναι τα εξής δεκαέξι :

- ✓ Manufacturing (per cent) - Μεταποίηση (επί τοις εκατό της απασχόλησης των εργαζομένων στην Ελλάδα)
- ✓ Wholesale and retail trade repair of motor vehicles motorcycles and personal and household goods (per cent) - Χονδρικό και λιανικό εμπόριο, επισκευές οχημάτων και ειδών νοικοκυριών (επί τοις εκατό της απασχόλησης των εργαζομένων στην Ελλάδα)
- ✓ Households – Νοικοκυριά (κατανάλωση ενέργειας σε terajoules)
- ✓ Vacation centers - Παραθεριστικά κέντρα (βαθμός πληρότητας σε ένα χρόνο)
- ✓ Arrivals / Foreigners (per cent) - Αφίξεις Αλλοδαπών (επί τοις εκατό των συνολικών αφίξεων σε όλα τα αεροδρόμια της Ελλάδας)
- ✓ Financial intermediation - Ενδιάμεσοι Χρηματοπιστωτικοί Οργανισμοί (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)

⁴¹ Βλ. «INVESTOPEDIA, Consumer Price Index – C.P.I.», <http://www.investopedia.com/terms/c/consumer_price_index.asp>, 20/6/2009.

- ✓ Education – Εκπαίδευση (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Health and social work - Υγεία και κοινωνική εργασία (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ G.D.P. / G.N.P. - Αναλογία Α.Ε.Π. / Α.ΕΘ.Π.
- ✓ Low wage population - Χαμηλόμισθος πληθυσμός (Ποσοστό πληθυσμού υπό του ορίου της φτώχειας σε δείγμα 3000 οικογενειών)
- ✓ Local calls - Τοπικές κλήσεις (αξία σε ευρώ ανά 10 λεπτά κλήσης)
- ✓ Ratio of imports to exports - Αναλογία εισαγωγών - εξαγωγών
- ✓ Ministry of Public Order - Υπουργείο Δημόσιας Τάξης (δημοσιονομικά έξοδα ανά φορέα, σε εκατομμύρια ευρώ)
- ✓ Ministry of Culture - Υπουργείο Πολιτισμού (δημοσιονομικά έξοδα ανά φορέα, σε εκατομμύρια ευρώ)
- ✓ Permanent population - Μόνιμος Πληθυσμός
- ✓ Gross debt as a percentage of G.D.P. - Ακαθάριστο χρέος ως ποσοστό του Α.Ε.Π.

7.4.3. Κανόνες Ταξινόμησης.

Από το Δ.Τ.Κ. και τα συσχετιζόμενα με αυτόν χαρακτηριστικά προκύπτουν, από τους αλγόριθμους JRip και PART, οι εξής κανόνες:

A) JRIP rules:

=>Wholesale and retail trade repair of motor vehicles motorcycles and personal and household goods per cent='(0.16844-0.1755]' (12.0/6.0)

Κανόνες αυτής της μορφής, στους οποίους δεν υπάρχει προϋπόθεση παρά μόνο το αποτέλεσμα, δηλώνουν τυπικά, ότι η τιμή της «μεταβλητής - στόχου» κυμαίνεται στο διάστημα που αναφέρεται κάθε φορά, ανεξάρτητα του τι συμβαίνει με τα υπόλοιπα χαρακτηριστικά. Εδώ, δηλαδή, ο πιο πάνω κανόνας λέει ότι το χονδρικό και λιανικό εμπόριο και οι επισκευές οχημάτων και ειδών νοικοκυριών αποτελούν το (16,84 - 17,55) % της συνολικής απασχόλησης των εργαζομένων στην Ελλάδα, ανεξαρτήτως των υπολοίπων χαρακτηριστικών (attributes).

Number of Rules : 1

B) JRIP rules:

1^{ος} κανόνας :

GDP / GNP = '(-inf-0.9789]') => Local calls='(-inf-0.18]' (3.0/0.0)

Έχουμε την προϋπόθεση (GDP / GNP = '(-inf-0.9789]') και το αποτέλεσμα (=> Local calls='(-inf-0.18]' (3.0/0.0) του κανόνα. Στο αποτέλεσμα, μετά το «συνεπάγεται» (=>), όλα τα νούμερα αναφέρονται στο χαρακτηριστικό που θέτουμε ως «μεταβλητή - στόχο» κάθε φορά. Εδώ, ως «μεταβλητή - στόχο», θέσαμε τις τοπικές κλήσεις (αξία σε ευρώ ανά 10 λεπτά κλήσης). Ως εκ τούτου, ο κανόνας αυτός λέει ότι εάν η αναλογία Α.Ε.Π. / Α.ΕΘ.Π. είναι μέχρι 0,9789, τότε οι τοπικές κλήσεις κοστίζουν το πολύ 0,18 ευρώ / 10' κλήσης. Ο αριθμός 3.0, που υπάρχει στην παρένθεση, δηλώνει ότι, από το σύνολο των υποδειγμάτων μας, ο κανόνας αυτός καλύπτει (covers) 3 υποδείγματα. Ο αριθμός 0.0 δηλώνει τον αριθμό των «λανθασμένων» (misclassified) από αυτά τα «covered instances».

2^{ος} κανόνας :

=> Local calls='(0.3-0.32]' (9.0/3.0)

Οι τοπικές κλήσεις κοστίζουν 0,3 - 0,32 ευρώ ανά 10 λεπτά κλήσης, ανεξάρτητα από το πώς επηρεάζονται από όλα τα άλλα χαρακτηριστικά του συνόλου δεδομένων που έχουμε. Καλύπτονται 9 από τα 12 υποδείγματα και 3 από αυτά ταξινομήθηκαν «λανθασμένα».

Number of Rules : 2

Γ) PART decision list : («μεταβλητή - στόχος» είναι η κατανάλωση ενέργειας από τα νοικοκυριά στην Ελλάδα σε terajoules).

1^{ος}, 2^{ος} και 3^{ος} κανόνας :

Manufacturing per cent = '(0.12837-0.13426]': '(224447.2-231897.1]' (3.0)

Manufacturing per cent = '(0.14015-0.14604]': '(172297.9-179747.8]' (3.0/1.0)

Manufacturing per cent = '(0.13426-0.14015]': '(179747.8-187197.7]' (2.0/1.0)

Εάν η επί τοις εκατό απασχόληση των εργαζομένων στον τομέα της Μεταποίησης κυμαίνεται μεταξύ 12,84 και 13,43 %, τότε τα νοικοκυριά καταναλώνουν από 224447,2 μέχρι 231897,1 terajoules ηλεκτρικής ενέργειας. Αντίστοιχα και οι κανόνες 2 και 3.

4^{ος} κανόνας :

Gross debt as a percentage of GDP ≤ 1.141 : '(231897.1-inf)' (2.0)

Εάν το ακαθάριστο χρέος, ως ποσοστό του Α.Ε.Π., είναι μικρότερο ή ίσο από 114,1 %, τότε τα νοικοκυριά καταναλώνουν τουλάχιστον 231897,1 τεταζουλες ηλεκτρικής ενέργειας.

5^{ος} κανόνας :

: '(-inf-172297.9]' (2.0)

Η κατανάλωση ενέργειας από τα νοικοκυριά στην Ελλάδα είναι το πολύ 172297,9 τεταζουλες, ανεξάρτητα από το πώς επηρεάζεται από όλα τα άλλα χαρακτηριστικά του συνόλου δεδομένων που έχουμε.

Number of Rules : 5

7.5] Gross Fixed Capital Formation - Ακαθάριστος Σχηματισμός Παγίου Κεφαλαίου.

7.5.1. Ορισμός.

Ως Ακαθάριστος Σχηματισμός Παγίου Κεφαλαίου ορίζεται το άθροισμα των δαπανών για τα εξής επενδυτικά αγαθά: αγροτικά προϊόντα, μεταλλικά προϊόντα, μηχανήματα, εξοπλισμός μεταφορών, κατοικίες, άλλες κατασκευές, μεταβολές αποθεμάτων και άλλα προϊόντα.⁴²

7.5.2. Επιλογή των πιο σχετικών, με τον Ακαθάριστο Σχηματισμό Παγίου Κεφαλαίου, χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με τον Ακαθάριστο Σχηματισμό Παγίου Κεφαλαίου είναι τα εξής είκοσι :

- ✓ Mining and quarrying - Ορυχεία και λατομεία (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Hotels and restaurants - Ξενοδοχεία και εστιατόρια (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Other community social and personal service activities - Παραγωγή άλλων υπηρεσιών υπέρ του κοινωνικού συνόλου (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Employment rate - Ποσοστό απασχόλησης
- ✓ Hotels / Motels / Bungalows A - Ξενοδοχεία / Μοτέλ / Bungalows A (κατηγορίας A) (βαθμός πληρότητας σε ένα χρόνο)
- ✓ Hotels / Motels / Bungalows B - Ξενοδοχεία / Μοτέλ / Bungalows B (κατηγορίας B) (βαθμός πληρότητας σε ένα χρόνο)
- ✓ Furnished apartments - Επιπλωμένα διαμερίσματα (βαθμός πληρότητας σε ένα χρόνο)
- ✓ Hotels / Motels / Bungalows D+E - Ξενοδοχεία / Μοτέλ / Bungalows Δ+Ε (κατηγορίας Δ+Ε) (βαθμός πληρότητας σε ένα χρόνο)
- ✓ Arrivals / Greeks (per cent) - Αφίξεις Ελλήνων (επί τοις εκατό των συνολικών αφίξεων σε όλα τα αεροδρόμια της Ελλάδας)

⁴² Βλ. Παπανδρέου Ανδρέας, «Προϋπολογισμός 2007, Μια κριτική αποτίμηση», Δεκέμβριος 2006, σελ.10, <http://www.istame-apapandreou.gr/files/pdf/Budget2007_final.pdf>, 20/6/2009.

- ✓ Private households with employed persons - Ιδιωτικά νοικοκυριά που απασχολούν οικιακό προσωπικό (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Agriculture hunting and forestry - Γεωργία, θήρα και δασοκομία (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στους συγκεκριμένους τομείς)
- ✓ Wholesale and retail trade repair of motor vehicles motorcycles and personal and household goods - Χονδρικό και λιανικό εμπόριο, επισκευές οχημάτων και ειδών νοικοκυριών (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στους συγκεκριμένους τομείς)
- ✓ Manufacturing - Μεταποίηση (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Low wage population - Χαμηλόμισθος πληθυσμός (Ποσοστό πληθυσμού υπό του ορίου της φτώχειας σε δείγμα 3000 οικογενειών)
- ✓ Wage-earning population - Μισθωτός πληθυσμός (Ποσοστό πληθυσμού άνω του ορίου της φτώχειας σε δείγμα 3000 οικογενειών)
- ✓ Ratio of imports to exports - Αναλογία εισαγωγών - εξαγωγών
- ✓ Ratio of services exports to imports - Αναλογία εισαγωγών - εξαγωγών για τις υπηρεσίες
- ✓ Exchange rate euro / drachma - Συναλλαγματική ισοτιμία ευρώ / δραχμή
- ✓ Long-Term Interest Rates 10-year Government Bond Yields Secondary Markets - Μακροπρόθεσμα επιτόκια: Απόδοση δεκαετούς κρατικού ομολόγου (δευτερογενής αγορά)
- ✓ Ministry of Macedonia and Thrace - Υπουργείο Μακεδονίας και Θράκης (δημοσιονομικά έξοδα ανά φορέα, σε εκατομμύρια ευρώ)

7.5.3. Κανόνες Ταξινόμησης.

Από τον Ακαθάριστο Σχηματισμό Παγίου Κεφαλαίου και τα συσχετιζόμενα με αυτόν χαρακτηριστικά προκύπτουν, από τους αλγόριθμους JRip και PART, οι εξής κανόνες:

A) JRIP rules:

=> Manufacturing='(0.118-0.12]' (12.0/2.0)

Η απασχόληση των εργαζομένων στην Ελλάδα στον τομέα της Μεταποίησης καλύπτει το διάστημα μεταξύ 11,8 % και 12 % της συνολικής (σε όλους τους τομείς) πλήρους απασχόλησης των εργαζομένων, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

B) JRIP rules:

=> Exchange rate euro / drachma='(337.23-inf)' (12.0/5.0)

Η συναλλαγματική ισοτιμία ευρώ / δραχμή κυμαίνεται σταθερά πάνω από 1 / 337,23, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

Γ) JRIP rules:

=> Ministry of Macedonia and Thrace='(-inf-13.316]' (12.0/5.0)

Τα δημοσιονομικά έξοδα του Υπουργείου Μακεδονίας και Θράκης είναι το πολύ 13,316 δισεκατομμύρια ευρώ, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

Δ) PART decision list : («μεταβλητή - στόχος» είναι η απασχόληση των εργαζομένων στην Ελλάδα στον τομέα της Μεταποίησης, σαν ποσοστό επί της πλήρους απασχόλησης των εργαζομένων σε όλους τους τομείς).

: '(0.118-0.12]' (12.0/2.0)

Η απασχόληση των εργαζομένων στην Ελλάδα στον τομέα της Μεταποίησης καλύπτει το διάστημα μεταξύ 11,8 % και 12 % της συνολικής (σε όλους τους τομείς) πλήρους απασχόλησης των εργαζομένων, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

Ε) PART decision list : («μεταβλητή - στόχος» είναι τα δημοσιονομικά έξοδα του Υπουργείου Μακεδονίας και Θράκης).

: '(-inf-13.316]' (12.0/5.0)

Τα δημοσιονομικά έξοδα του Υπουργείου Μακεδονίας και Θράκης είναι το πολύ 13,316 δισεκατομμύρια ευρώ, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

7.6] Fiscal Revenue - Φορολογικό Εισόδημα.

7.6.1. Ορισμός.

Με τον όρο Φορολογικό Εισόδημα (ετήσιο) εννοείται ο καθαρός μηνιαίος μισθός (είναι το ποσό που προκύπτει, αφού από τις ακαθάριστες αποδοχές αφαιρεθούν οι διάφορες κρατήσεις που γίνονται για τα ασφαλιστικά ταμεία και τον Ο.Γ.Α.) των εργαζομένων και των συνταξιούχων, ο οποίος πολλαπλασιάζεται με το 14. Με βάση το φορολογικό εισόδημα βρίσκουμε το φόρο που αναλογεί για τον κάθε φορολογούμενο.⁴³

7.6.2. Επιλογή των πιο σχετικών, με το Φορολογικό Εισόδημα, χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με το Φορολογικό Εισόδημα είναι τα εξής είκοσι τρία :

- ✓ Wholesale and retail trade repair of motor vehicles motorcycles and personal and household goods - Χονδρικό και λιανικό εμπόριο, επισκευές οχημάτων και ειδών νοικοκυριών (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στους συγκεκριμένους τομείς)
- ✓ Mining and quarrying (per cent) - Ορυχεία και λατομεία (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Transport storage and communications (per cent) - Μεταφορές, αποθηκείσεις και επικοινωνίες (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Public administration and defense compulsory social security (per cent) - Δημόσια διοίκηση και άμυνα, υποχρεωτική κοινωνική ασφάλιση (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Other community social and personal service activities (per cent) - Παραγωγή άλλων υπηρεσιών υπέρ του κοινωνικού συνόλου (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Private households with employed persons (per cent) - Ιδιωτικά νοικοκυριά που απασχολούν οικιακό προσωπικό (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Unemployment rate - Ποσοστό Ανεργίας
- ✓ Permanent residents - Μόνιμοι κάτοικοι

⁴³ Βλ. Υπουργείο Οικονομικών, «Με τη νέα κλίμακα η παρακράτηση φόρου, Δημοσιοποιήθηκε η σχετική εγκύκλιος του Υπουργείου», *Ριζοσπάστης*, 4 / 12 / 2001, σ.24.

- ✓ Fossil fuels (per cent) - Φυσικά καύσιμα (επί τοις εκατό της συνολικής κατανάλωσης ενέργειας)
- ✓ Fossil fuels solid (per cent) - Φυσικά στερεά καύσιμα (επί τοις εκατό της συνολικής κατανάλωσης ενέργειας)
- ✓ Hotels / Motels / Bungalows AA - Ξενοδοχεία / Μοτέλ / Bungalows AA (κατηγορίας AA) (διαθέσιμες κλίνες - και για τους Έλληνες και για τους αλλοδαπούς - αυτής της κατηγορίας καταλυμάτων, κατά τη χρονική περίοδο 1996 - 2007, σαν ποσοστό επί των συνολικών διαθέσιμων κλινών όλων των κατηγοριών καταλυμάτων)
- ✓ Hotels / Motels / Bungalows A - Ξενοδοχεία / Μοτέλ / Bungalows A (κατηγορίας A) (όπως παραπάνω)
- ✓ Hotels / Motels / Bungalows B - Ξενοδοχεία / Μοτέλ / Bungalows B (κατηγορίας B) (όπως παραπάνω)
- ✓ Hotels / Motels / Bungalows C - Ξενοδοχεία / Μοτέλ / Bungalows Γ (κατηγορίας Γ) (όπως παραπάνω)
- ✓ Arrivals / Foreigners (per cent) - Αφίξεις Αλλοδαπών (επί τοις εκατό των συνολικών αφίξεων σε όλα τα αεροδρόμια της Ελλάδας)
- ✓ Agriculture hunting and forestry - Γεωργία, θήρα και δασοκομία (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στους συγκεκριμένους τομείς)
- ✓ G.D.P. / G.N.P. - Αναλογία Α.Ε.Π. / Α.ΕΘ.Π.
- ✓ Mean annual wages ratio of men versus women - Μέση ετήσια αναλογία αμοιβών ανδρών / γυναικών
- ✓ GINI index - Δείκτης Εισοδηματικής Κατανομής
- ✓ Inflation - Πληθωρισμός (ορίζεται στο 7.7)
- ✓ Imports - Εισαγωγές
- ✓ Hellenic Parliament - Ελληνικό Κοινοβούλιο (δημοσιονομικά έξοδα ανά φορέα, σε εκατομμύρια ευρώ)
- ✓ Indirect taxation - Έμμεση φορολογία (δημοσιονομικά έσοδα για ένα έτος, μέσω της έμμεσης φορολόγησης των εργαζομένων)

7.6.3. Κανόνες Ταξινόμησης.

Από το Φορολογικό Εισόδημα και τα συσχετιζόμενα με αυτό χαρακτηριστικά προκύπτουν, από τους αλγόριθμους JRip και PART, οι εξής κανόνες:

A) JRIP rules:

=> Hotels / Motels / Bungalows AA='(-inf-0.08756]' (12.0/4.0)

Οι διαθέσιμες κλίνες (και για τους Έλληνες και για τους αλλοδαπούς) των Ξενοδοχείων / Μοτέλ / Bungalows AA (κατηγορίας AA), κατά τη χρονική περίοδο 1996 - 2007, σαν ποσοστό επί των συνολικών διαθέσιμων κλινών όλων των κατηγοριών καταλυμάτων, δεν ξεπερνούν το 8,756 %, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

B) JRIP rules:

=> Hotels / Motels / Bungalows C='(0.24089-inf)' (12.0/4.0)

Οι διαθέσιμες κλίνες (και για τους Έλληνες και για τους αλλοδαπούς) των Ξενοδοχείων / Μοτέλ / Bungalows C (κατηγορίας C), κατά τη χρονική περίοδο 1996 - 2007, σαν ποσοστό επί των συνολικών διαθέσιμων κλινών όλων των κατηγοριών καταλυμάτων, καλύπτουν κατά το ελάχιστο, το 24,089 % του συνόλου, , ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

Γ) JRIP rules:

1^{ος} κανόνας :

Public administration and defense compulsory social security per cent = '(-inf-0.0728]' => Mean annual wages ratio of men versus women='(-inf-1.022]' (3.0/0.0)

Εάν η απασχόληση των εργαζομένων στην Ελλάδα στον τομέα της Δημόσιας Διοίκησης και Άμυνας, καθώς επίσης και στον τομέα δραστηριοτήτων της υποχρεωτικής κοινωνικής ασφάλισης δεν ξεπερνά το 7,28 % της συνολικής (σε όλους τους τομείς) πλήρους απασχόλησης των εργαζομένων, τότε η μέση ετήσια αναλογία των αμοιβών των αντρών προς τις αντίστοιχες των γυναικών δεν ξεπερνά το 1,022 (102,2 %).

2^{ος} κανόνας :

=> Mean annual wages ratio of men versus women='(1.038-inf)' (9.0/3.0)

Η μέση ετήσια αναλογία των αμοιβών των αντρών προς τις αντίστοιχες των γυναικών είναι, κατά το ελάχιστο, 1,038 ή 103,8 %, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 2

Δ) PART decision list : («μεταβλητή - στόχος» είναι ο αριθμός, σε χιλιάδες, των εργαζομένων πλήρους απασχόλησης στους τομείς του χονδρικού και του λιανικού εμπορίου και πιο συγκεκριμένα, στις επισκευές οχημάτων και στα είδη νοικοκυριών).

1^{ος} κανόνας :

Mean annual wages ratio of men versus women = '(-inf-1.022]' AND
Indirect taxation <= 15625.99: '(-inf-564.4]' (2.0)

Εάν η μέση ετήσια αναλογία των αμοιβών των αντρών προς τις αντίστοιχες των γυναικών είναι, κατά το μέγιστο, 1,022 ή 102,2 % και, ταυτόχρονα, ισχύει ότι η έμμεση φορολογία των εργαζομένων από το κράτος είναι (για ένα έτος) μικρότερη ή ίση από 15,626 δισεκατομμύρια ευρώ, τότε ο αριθμός των εργαζομένων που απασχολούνται πλήρως στους τομείς του χονδρικού και του λιανικού εμπορίου και πιο συγκεκριμένα, στις επισκευές οχημάτων και στα είδη νοικοκυριών, δεν πρέπει να ξεπερνά το όριο των 564,4 χιλιάδων εργαζομένων.

2^{ος} κανόνας :

Mean annual wages ratio of men versus women = '(1.038-inf)' AND
Indirect taxation <= 23722: '(700.6-746]' (4.0)

Εάν η μέση ετήσια αναλογία των αμοιβών των αντρών προς τις αντίστοιχες των γυναικών είναι, κατά το ελάχιστο, 1,038 ή 103,8 % και, ταυτόχρονα, ισχύει ότι η έμμεση φορολογία των εργαζομένων από το κράτος είναι (για ένα έτος) μικρότερη ή ίση από 23,722 δισεκατομμύρια ευρώ, τότε ο αριθμός των εργαζομένων που απασχολούνται πλήρως στους τομείς του χονδρικού και του λιανικού εμπορίου και πιο συγκεκριμένα, στις επισκευές οχημάτων και στα είδη νοικοκυριών, κυμαίνεται μεταξύ 700,6 και 746 χιλιάδων εργαζομένων.

3^{ος} κανόνας :

GINI INDEX = '(-inf-0.332]': '(746-791.4]' (3.0)

Εάν ο Δείκτης Εισοδηματικής Κατανομής GINI δεν ξεπερνά το 0,332 (33,2 %), τότε ο αριθμός των εργαζομένων που απασχολούνται πλήρως στους τομείς του χονδρικού και του λιανικού εμπορίου και πιο συγκεκριμένα, στις επισκευές οχημάτων και στα είδη νοικοκυριών, κυμαίνεται μεταξύ 746 και 791,4 χιλιάδων εργαζομένων.

4^{ος} κανόνας :

:(655.2-700.6]' (3.0/1.0)

Ο αριθμός των εργαζομένων που απασχολούνται πλήρως στους τομείς του χονδρικού και του λιανικού εμπορίου και πιο συγκεκριμένα, στις επισκευές οχημάτων και στα είδη νοικοκυριών, κυμαίνεται μεταξύ 655,2 και 700,6 χιλιάδων εργαζομένων, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 4

E) PART decision list : («μεταβλητή - στόχος» είναι οι διαθέσιμες κλίνες - και για τους Έλληνες και για τους αλλοδαπούς - των Ξενοδοχείων / Μοτέλ / Bungalows AA (κατηγορίας AA), κατά τη χρονική περίοδο 1996 - 2007, σαν ποσοστό επί των συνολικών διαθέσιμων κλινών όλων των κατηγοριών καταλυμάτων της Ελλάδας).

1^{ος} κανόνας :

Indirect taxation <= 21484: '(-inf-0.08756]' (8.0)

Εάν η έμμεση φορολογία των εργαζομένων από το κράτος είναι (για ένα έτος) μικρότερη ή ίση από 21,484 δισεκατομμύρια ευρώ, τότε οι διαθέσιμες κλίνες - και για τους Έλληνες και για τους αλλοδαπούς - των Ξενοδοχείων / Μοτέλ / Bungalows AA (κατηγορίας AA), σαν ποσοστό επί των συνολικών διαθέσιμων κλινών όλων των κατηγοριών καταλυμάτων της Ελλάδας, δεν πρέπει να ξεπερνούν το 8,76 %.

2^{ος} κανόνας :

Fossil fuels solid per cent = '(-inf-0.29624]': '(0.11908-0.13484]' (2.0/1.0)

Εάν η εγχώρια κατανάλωση ενέργειας στην Ελλάδα από τα φυσικά στερεά καύσιμα αποτελεί, κατά το μέγιστο, το 29,62 % της συνολικής εγχώριας κατανάλωσης ενέργειας, τότε οι διαθέσιμες κλίνες - και για τους Έλληνες και για τους αλλοδαπούς - των Ξενοδοχείων / Μοτέλ / Bungalows

AA (κατηγορίας AA), κυμαίνονται μεταξύ 11,91 % και 13,48 % επί των συνολικών διαθέσιμων κλινών όλων των κατηγοριών καταλυμάτων της Ελλάδας.

3^{ος} κανόνας :

: '(0.10332-0.11908]' (2.0/1.0)

Οι διαθέσιμες κλίνες - και για τους Έλληνες και για τους αλλοδαπούς - των Ξενοδοχείων / Μοτέλ / Bungalows AA (κατηγορίας AA), κυμαίνονται μεταξύ 10,33 % και 11,91 % επί των συνολικών διαθέσιμων κλινών όλων των κατηγοριών καταλυμάτων της Ελλάδας, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 3

ΣΤ) PART decision list : («μεταβλητή - στόχος» είναι οι διαθέσιμες κλίνες - και για τους Έλληνες και για τους αλλοδαπούς - των Ξενοδοχείων / Μοτέλ / Bungalows C (κατηγορίας C), κατά τη χρονική περίοδο 1996 - 2007, σαν ποσοστό επί των συνολικών διαθέσιμων κλινών όλων των κατηγοριών καταλυμάτων της Ελλάδας).

1^{ος} κανόνας :

Indirect taxation <= 21484: '(0.24089-inf)' (8.0)

Εάν η έμμεση φορολογία των εργαζομένων από το κράτος είναι (για ένα έτος) μικρότερη ή ίση από 21,484 δισεκατομμύρια ευρώ, τότε οι διαθέσιμες κλίνες - και για τους Έλληνες και για τους αλλοδαπούς - των Ξενοδοχείων / Μοτέλ / Bungalows C (κατηγορίας C), σαν ποσοστό επί των συνολικών διαθέσιμων κλινών όλων των κατηγοριών καταλυμάτων της Ελλάδας, δεν πρέπει να είναι κάτω από 24,09 %.

2^{ος} κανόνας :

Fossil fuels solid per cent = '(-inf-0.29624]': '(-inf-0.16641]' (2.0/1.0)

Εάν η εγχώρια κατανάλωση ενέργειας στην Ελλάδα από τα φυσικά στερεά καύσιμα αποτελεί, κατά το μέγιστο, το 29,62 % της συνολικής εγχώριας κατανάλωσης ενέργειας, τότε οι διαθέσιμες κλίνες - και για τους Έλληνες και για τους αλλοδαπούς - των Ξενοδοχείων / Μοτέλ / Bungalows C (κατηγορίας C), σαν ποσοστό επί των συνολικών διαθέσιμων κλινών όλων των κατηγοριών καταλυμάτων της Ελλάδας, δεν πρέπει να ξεπερνούν το 16,64 %.

3^{ος} κανόνας :

: '(0.18503-0.19434]' (2.0/1.0)

Οι διαθέσιμες κλίνες - και για τους Έλληνες και για τους αλλοδαπούς - των Ξενοδοχείων / Μοτέλ / Bungalows C (κατηγορίας C), κυμαίνονται μεταξύ 18,5 % και 19,43 % επί των συνολικών διαθέσιμων κλινών όλων των κατηγοριών καταλυμάτων της Ελλάδας, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 3

Z) PART decision list : («μεταβλητή - στόχος» είναι η μέση ετήσια αναλογία των αμοιβών των αντρών προς τις αντίστοιχες των γυναικών).

1^{ος} κανόνας :

Indirect taxation > 17969.67: '(1.038-inf)' (8.0/2.0)

Εάν η έμμεση φορολογία των εργαζομένων από το κράτος είναι (για ένα έτος) μεγαλύτερη από 17,97 δισεκατομμύρια ευρώ, τότε η μέση ετήσια αναλογία των αμοιβών των αντρών προς τις αντίστοιχες των γυναικών είναι, κατά το ελάχιστο, 1,038 ή 103,8 %.

2^{ος} κανόνας :

: '(-inf-1.022]' (4.0)

Η μέση ετήσια αναλογία των αμοιβών των αντρών προς τις αντίστοιχες των γυναικών είναι, κατά το μέγιστο, 1,022 ή 102,2 %, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 2

7.7] Inflation - Πληθωρισμός.

7.7.1. Ορισμός.

Πληθωρισμός είναι η ποσοστιαία μεταβολή του γενικού επιπέδου των τιμών μιας οικονομίας μέσα σε μια συγκεκριμένη χρονική περίοδο. Ο πληθωρισμός μπορεί να είναι είτε θετικός, είτε αρνητικός (οπότε μιλάμε για αντιπληθωρισμό). Ας σημειωθεί ότι πληθωρισμός είναι η κίνηση των τιμών. Δεν υφίσταται όταν οι τιμές σταθεροποιηθούν, ανεξαρτήτως αν είναι υψηλές ή όχι. Σε μια οικονομία όταν μετράμε τον πληθωρισμό, στην ουσία μελετάμε την ποσοστιαία μεταβολή του επιπέδου των τιμών, όχι για το σύνολο των αγαθών ή υπηρεσιών που καταναλώνονται, αλλά για κάποια συγκεκριμένα αγαθά ή υπηρεσίες, το σύνολο των οποίων παλαιότερα καλούνταν «καλάθι της νοικοκυράς», ενώ πλέον χρησιμοποιείται το πολιτικά ορθότερο «καλάθι του καταναλωτή».⁴⁴

7.7.2. Επιλογή των πιο σχετικών, με τον Πληθωρισμό, χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με τον Πληθωρισμό είναι τα εξής δεκαέξι :

- ✓ Manufacturing (per cent) - Μεταποίηση (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Electricity gas and water supply (per cent) - Παροχή ηλεκτρικού ρεύματος, φυσικού αερίου και νερού (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Wholesale and retail trade repair of motor vehicles motorcycles and personal and household goods (per cent) - Χονδρικό και λιανικό εμπόριο, επισκευές οχημάτων και ειδών νοικοκυριών (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Transport storage and communications (per cent) - Μεταφορές, αποθηκεύσεις και επικοινωνίες (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Activity rate - Ποσοστό απασχόλησης
- ✓ Permanent resident population - Μόνιμος εδρεύων πληθυσμός
- ✓ Transport (per cent) - Μεταφορά (επί τοις εκατό της συνολικής κατανάλωσης ενέργειας)

⁴⁴ Βλ. «Πληθωρισμός, Από τη Βικιπαίδεια, την ελεύθερη εγκυκλοπαίδεια», <<http://el.wikipedia.org/wiki/>>, 20/6/2009.

- ✓ Hotels / Motels / Bungalows AA - Ξενοδοχεία / Μοτέλ / Bungalows AA (κατηγορίας AA) (πληρότητα των Ξενοδοχείων / Μοτέλ / Bungalows AA (κατηγορίας AA) σαν ποσοστό επί της συνολικής πληρότητας όλων των καταλυμάτων στην Ελλάδα)
- ✓ Education – Εκπαίδευση (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Ratio of services exports to imports - Αναλογία εισαγωγών - εξαγωγών για τις υπηρεσίες
- ✓ Long-Term Interest Rates 10-year Government Bond Yields Secondary Markets - Μακροπρόθεσμα επιτόκια: Απόδοση δεκαετούς κρατικού ομολόγου (δευτερογενής αγορά)
- ✓ Deficit – Έλλειμμα (Έξοδα μείον έσοδα του κράτους)
- ✓ Expenditure on public investments - Δαπάνες για τις δημόσιες επενδύσεις
- ✓ Ministry for the Environment Physical Planning and Public Works - Υπουργείο Περιβάλλοντος, Χωροταξίας και Δημοσίων Έργων (Υ.ΠΕ.ΧΩ.ΔΕ.) (δημοσιονομικά έξοδα ανά φορέα, σε εκατομμύρια ευρώ)
- ✓ Indirect taxation - Έμμεση φορολογία (δημοσιονομικά έσοδα για ένα έτος, μέσω της έμμεσης φορολόγησης των εργαζομένων)
- ✓ Non-tax income - Μη φορολογικά έσοδα (δημοσιονομικά έσοδα)

7.7.3. Κανόνες Ταξινόμησης.

Από τον πληθωρισμό και τα συσχετιζόμενα με αυτόν χαρακτηριστικά προκύπτουν, από τους αλγόριθμους JRip και PART οι εξής κανόνες:

A) JRIP rules:

=>Wholesale and retail trade repair of motor vehicles motorcycles and personal and household goods per cent='(0.16844-0.1755]' (12.0/6.0)

Το ποσοστό του αριθμού των εργαζομένων που απασχολούνται πλήρως στους τομείς του χονδρικού και του λιανικού εμπορίου και πιο συγκεκριμένα, στις επισκευές οχημάτων και στα είδη νοικοκυριών, κυμαίνεται μεταξύ 16,84 % και 17,55 % επί της πλήρους απασχόλησης όλων των εργαζομένων στην Ελλάδα σε όλους του τομείς, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

B) JRIP rules:

1^{ος} κανόνας :

Hotels / Motels / Bungalows AA = '(10.9-11.12]') => Education='(-inf-0.041]' (2.0/0.0)

Εάν η πληρότητα των Ξενοδοχείων / Μοτέλ / Bungalows AA (κατηγορίας AA) αποτελεί το 10,9 % με 11,12 % της συνολικής πληρότητας όλων των καταλυμάτων στην Ελλάδα, τότε οι Έλληνες που απασχολούνται πλήρως στον τομέα της εκπαίδευσης αποτελούν, κατά το μέγιστο, το 4,1 % όλων των εργαζόμενων Ελλήνων που απασχολούνται πλήρως σε όλους ανεξαρτήτως τους τομείς.

2^{ος} κανόνας :

=> Education='(0.049-inf)' (10.0/0.0)

Οι Έλληνες που απασχολούνται πλήρως στον τομέα της εκπαίδευσης αποτελούν, κατά το ελάχιστο, το 4,9 % όλων των εργαζόμενων Ελλήνων που απασχολούνται πλήρως σε όλους τους τομείς, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 2

Γ) JRIP rules:

1^{ος} κανόνας :

Transport per cent = '(-inf-0.38424]') => Ministry for the Environment Physical Planning and Public Works='(-inf-115.914]' (2.0/0.0)

Εάν η κατανάλωση ενέργειας στον τομέα των Μεταφορών αποτελεί, κατά το μέγιστο, το 38,42 % της συνολικής κατανάλωσης ενέργειας, τότε τα δημοσιονομικά έξοδα του Υπουργείου Περιβάλλοντος, Χωροταξίας και Δημοσίων Έργων (Υ.ΠΕ.ΧΩ.ΔΕ.) δεν ξεπερνούν τα 115,91 εκατομμύρια ευρώ.

2^{ος} κανόνας :

=> Ministry for the Environment Physical Planning and Public Works='(115.914-143.478]' (10.0/4.0)

Τα δημοσιονομικά έξοδα του Υπουργείου Περιβάλλοντος, Χωροταξίας και Δημοσίων Έργων (Υ.ΠΕ.ΧΩ.ΔΕ.) κυμαίνονται μεταξύ 115,91 και 143,48 εκατομμύρια ευρώ, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 2

Δ) PART decision list : («μεταβλητή - στόχος» είναι οι Έλληνες που απασχολούνται πλήρως στον τομέα της εκπαίδευσης, σαν ποσοστό επί της συνολικής - σε όλους τους τομείς - πλήρους απασχόλησης όλων των εργαζομένων).

: '(0.049-inf)' (12.0/2.0)

Οι Έλληνες που απασχολούνται πλήρως στον τομέα της εκπαίδευσης αποτελούν, κατά το ελάχιστο, το 4,9 % όλων των εργαζόμενων Ελλήνων που απασχολούνται πλήρως σε όλους τους τομείς, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

7.8] Exchange Rate Euro / U.S.Dollar - Ισοτιμία Συναλλάγματος Ευρώ / Αμερικανικού Δολαρίου.

7.8.1. Ορισμός.

Ισοτιμία συναλλάγματος, γενικά, είναι η αξία του ξένου νομίσματος ανά μονάδα εθνικού νομίσματος.⁴⁵ Π.χ. εάν η ισοτιμία ευρώ / αμερικανικού δολαρίου είναι στις 1,4 μονάδες, αυτό σημαίνει ότι με 1 ευρώ μπορώ να αγοράσω 1,4 δολάρια Αμερικής.

7.8.2. Επιλογή των πιο σχετικών, με την Ισοτιμία Συναλλάγματος Ευρώ / Αμερικανικού Δολαρίου, χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με την Ισοτιμία Συναλλάγματος Ευρώ / Αμερικανικού Δολαρίου είναι τα εξής οκτώ :

- ✓ Fishing (per cent) - Αλιεία (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Financial intermediation (per cent) - Ενδιάμεσοι Χρηματοπιστωτικοί Οργανισμοί (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Health and social work (per cent) - Υγεία και Κοινωνική Μέριμνα (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Fossil fuels (per cent) - Φυσικά καύσιμα (επί τοις εκατό της συνολικής κατανάλωσης ενέργειας)
- ✓ Night Stays / Greeks (per cent) - Διανυκτερεύσεις Ελλήνων (επί τοις εκατό των συνολικών διανυκτερεύσεων περιηγητών σε καταλύματα σε όλη την Ελλάδα)
- ✓ Night Stays / Foreigners (per cent) - Διανυκτερεύσεις αλλοδαπών (επί τοις εκατό των συνολικών διανυκτερεύσεων περιηγητών σε καταλύματα σε όλη την Ελλάδα)
- ✓ Domestic users - Οικιακοί χρήστες (τιμή αγοράς ηλεκτρικής ενέργειας σε € / kWh, μόνο για οικιακή χρήση)
- ✓ Exchange rate euro / Japanese yen - Ισοτιμία συναλλάγματος ευρώ / ιαπωνικού γιεν

⁴⁵ Βλ. Σταμελάτου Μαρία - Ελένη, «Ακάλυπτη Ισοτιμία Επιτοκίων, Οικονομική και Επιχειρησιακή Στρατηγική», Διπλωματική Εργασία, Πανεπιστήμιο Πειραιά, 2006, σ.18, <<http://digilib.lib.unipi.gr/dspace/bitstream/unipi/1539/4/Stamelatou.pdf>>, 20/6/2009.

7.8.3. Κανόνες Ταξινόμησης.

Από τη συναλλαγματική ισοτιμία ευρώ / αμερικανικού δολαρίου και τα συσχετιζόμενα με αυτήν χαρακτηριστικά προκύπτουν, από τους αλγόριθμους JRip και PART οι εξής κανόνες:

A) JRIP rules:

=> Fishing per cent='(-inf-0.00331]' (12.0/3.0)

Οι Έλληνες που απασχολούνται πλήρως στον τομέα της αλιείας δεν ξεπερνούν το 0,33% όλων των εργαζόμενων Ελλήνων που απασχολούνται πλήρως σε όλους τους τομείς, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

B) PART decision list : («μεταβλητή - στόχος» είναι οι Έλληνες απασχολούμενοι στον τομέα της αλιείας, σαν ποσοστό επί της συνολικής - σε όλους τους τομείς - πλήρους απασχόλησης όλων των εργαζομένων).

: '(-inf-0.00331]' (12.0/3.0)

Οι Έλληνες που απασχολούνται πλήρως στον τομέα της αλιείας δεν ξεπερνούν το 0,33% όλων των εργαζόμενων Ελλήνων που απασχολούνται πλήρως σε όλους τους τομείς, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

7.9] Low Wage Population - Χαμηλόμισθος Πληθυσμός.

7.9.1. Ορισμός.

Με τον όρο «χαμηλόμισθος πληθυσμός» εννοούμε τους εργαζόμενους σε μια χώρα, οι οποίοι αμείβονται με μισθό που είναι κάτω του ορίου φτώχειας, όριο το οποίο για κάθε χώρα διαφέρει και καθορίζεται από διάφορους οικονομικούς και άλλους δείκτες. Στην ουσία, το όριο φτώχειας, σε κάθε χώρα, ισούται με το 60 % του μέσου διαθέσιμου εισοδήματος που υπάρχει στη χώρα.

7.9.2. Επιλογή των πιο σχετικών, με το Χαμηλόμισθο Πληθυσμό, χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με το Χαμηλόμισθο Πληθυσμό είναι τα εξής επτά :

- ✓ Mining and quarrying - Ορυχεία και λατομεία (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Transport (per cent) - Μεταφορές (επί τοις εκατό της συνολικής κατανάλωσης ενέργειας)
- ✓ Stock change - Μεταβολή Αποθεμάτων
- ✓ Average of Night Stays / Foreigners - Μέσος όρος διανυκτερεύσεων αλλοδαπών (σε όλη την Ελλάδα)
- ✓ Wage-earning population - Μισθωτός πληθυσμός (εργαζόμενοι οι οποίοι αμείβονται με μισθό που είναι άνω του ορίου φτώχειας - το αντίθετο του «low wage population»)
- ✓ Local calls - Τοπικές κλήσεις (αξία σε ευρώ ανά 10 λεπτά κλήσης)
- ✓ Exchange rate euro / Japanese yen - Ισοτιμία συναλλάγματος ευρώ / ιαπωνικού γιεν

7.9.3. Κανόνες Ταξινόμησης.

Από το χαρακτηριστικό «χαμηλόμισθος πληθυσμός» και τα συσχετιζόμενα με αυτό χαρακτηριστικά δεν προκύπτουν οποιοδήποτε «χρήσιμοι» κανόνες ταξινόμησης (JRip ή PART).

7.10] Minimum Daily Wage of Workers - Ελάχιστος Ημερήσιος Μισθός των Εργαζόμενων.

7.10.1. Ορισμός.

Ο ορισμός εξυπακούεται.

7.10.2. Επιλογή των πιο σχετικών, με τον ελάχιστο ημερήσιο μισθό των Εργαζόμενων, χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με τον Ελάχιστο Ημερήσιο Μισθό των εργαζόμενων είναι τα εξής επτά :

- ✓ Hotels / Motels / Bungalows AA - Ξενοδοχεία / Μοτέλ / Bungalows AA (κατηγορίας AA)
- ✓ Other community social and personal service activities - Παραγωγή άλλων υπηρεσιών υπέρ του κοινωνικού συνόλου (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Mining and quarrying - Ορυχεία και λατομεία (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Education - Εκπαίδευση (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Gross domestic product GDP - Ακαθάριστο εγχώριο προϊόν (Α.Ε.Π.)
- ✓ Mean annual wage of low earning women - Μέσος ετήσιος μισθός χαμηλόμισθων γυναικών
- ✓ Hellenic Parliament - Ελληνικό Κοινοβούλιο (δημοσιονομικά έξοδα ανά φορέα, σε εκατομμύρια ευρώ)

7.10.3. Κανόνες Ταξινόμησης.

Από το χαρακτηριστικό «ελάχιστος ημερήσιος μισθός των εργαζομένων» και τα συσχετιζόμενα με αυτό χαρακτηριστικά προκύπτουν, από τους αλγόριθμους JRip και PART οι εξής κανόνες:

A) JRIP rules:

=> Education='(0.049-inf)' (12.0/2.0)

Οι Έλληνες που απασχολούνται πλήρως στον τομέα της εκπαίδευσης αποτελούν, κατά το ελάχιστο, το 4,9 % όλων των εργαζόμενων Ελλήνων που απασχολούνται πλήρως σε όλους τους τομείς, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

B) PART decision list : («μεταβλητή - στόχος» είναι οι Έλληνες απασχολούμενοι στον τομέα της εκπαίδευσης, σαν ποσοστό επί της συνολικής - σε όλους τους τομείς - πλήρους απασχόλησης όλων των εργαζομένων).

: '(0.049-inf)' (12.0/2.0)

Οι Έλληνες που απασχολούνται πλήρως στον τομέα της εκπαίδευσης αποτελούν, κατά το ελάχιστο, το 4,9 % όλων των εργαζόμενων Ελλήνων που απασχολούνται πλήρως σε όλους τους τομείς, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

7.11] Per Capita Gross Dept - Κατά Κεφαλή Ακαθάριστο Χρέος.

7.11.1. Ορισμός.

Κατά κεφαλή ακαθάριστο χρέος είναι το συνολικό ακαθάριστο χρέος μιας χώρας διαιρεμένο με το μόνιμο πληθυσμό της χώρας.

7.11.2. Επιλογή των πιο σχετικών, με το κατά κεφαλή Ακαθάριστο Χρέος, χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με το κατά κεφαλή Ακαθάριστο Χρέος είναι τα εξής πέντε :

- ✓ Mining and quarrying - Ορυχεία και λατομεία (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Transport storage and communications (per cent) - Μεταφορές, αποθηκεύσεις και επικοινωνίες (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Imports of products (relative change in constant 1996 prices) - Εισαγωγές προϊόντων, (σχετικές αλλαγές με τις σταθερές τιμές του 1996)
- ✓ Ministry of Health and Social Solidarity - Υπουργείο Υγείας και Κοινωνικής Αλληλεγγύης (δημοσιονομικά έξοδα ανά φορέα, σε εκατομμύρια ευρώ)
- ✓ Gross debt of public administration - Ακαθάριστο χρέος της δημόσιας διοίκησης

7.11.3. Κανόνες Ταξινόμησης.

Από το κατά κεφαλή Ακαθάριστο Χρέος και τα συσχετιζόμενα με αυτό χαρακτηριστικά δεν προκύπτουν οποιοδήποτε «χρήσιμοι» κανόνες ταξινόμησης (JRip ή PART).

7.12] Public Spending - Δημόσιες Δαπάνες.

7.12.1. Ορισμός.

Ως δημόσιες δαπάνες ορίζονται οι πληρωμές που διενεργούν οι δημόσιοι φορείς μέσω του προϋπολογισμού, για την επίτευξη ορισμένων στόχων, όπως η άριστη κατανομή των παραγωγικών μέσων, η δίκαιη διανομή του εισοδήματος, η σταθεροποίηση της οικονομίας και η οικονομική μεγέθυνση.⁴⁶

7.12.2. Επιλογή των πιο σχετικών, με τις Δημόσιες Δαπάνες, χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με τις Δημόσιες Δαπάνες είναι τα εξής δέκα :

- ✓ Hotels / Motels / Bungalows A - Ξενοδοχεία / Μοτέλ / Bungalows A (κατηγορίας A)
- ✓ Boarding Houses - Οικοτροφεία
- ✓ Guest houses - Ξενώνες
- ✓ Agriculture hunting and forestry - Γεωργία, θήρα και δασοκομία (αριθμός εργαζομένων πλήρους απασχόλησης στην Ελλάδα, στο συγκεκριμένο τομέα)
- ✓ Minimum daily wage of employees - Ελάχιστος ημερήσιος μισθός υπαλλήλων
- ✓ Low-wage population - Χαμηλόμισθος πληθυσμός
- ✓ Ministry of Economy and Finance - Υπουργείο Οικονομίας και Οικονομικών (δημοσιονομικά έξοδα ανά φορέα, σε εκατομμύρια ευρώ)
- ✓ Ministry of Development - Υπουργείο Ανάπτυξης (δημοσιονομικά έξοδα ανά φορέα, σε εκατομμύρια ευρώ)
- ✓ Public investments revenue - Εισόδημα από δημόσιες επενδύσεις (δημοσιονομικά έσοδα)
- ✓ Non-tax income - Μη φορολογικά έσοδα

7.12.3. Κανόνες Ταξινόμησης.

Από τις Δημόσιες Δαπάνες και τα συσχετιζόμενα με αυτές χαρακτηριστικά δεν προκύπτουν οποιοδήποτε «χρήσιμοι» κανόνες ταξινόμησης (JRip ή PART).

⁴⁶ Βλ. Δριτσάκης Νικόλαος, «Δημόσιες Δαπάνες και Οικονομική Ανάπτυξη, Μια εμπειρική έρευνα για την Ελλάδα με την ανάλυση της συνολοκλήρωσης», περ. *Σπουδαί*, τόμος 53, τεύχος 4^ο, 2003, σ.66.

7.13] Ratio of Services (Exports to Imports) - Αναλογία των Υπηρεσιών (Εξαγωγές προς Εισαγωγές).

7.13.1. Ορισμός.

Ο ορισμός εξυπακούεται.

7.13.2. Επιλογή των πιο σχετικών, με την αναλογία των Υπηρεσιών (Εξαγωγές προς Εισαγωγές), χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με την αναλογία των Υπηρεσιών (Εξαγωγές προς Εισαγωγές) είναι τα εξής είκοσι ένα :

- ✓ Mining and quarrying (per cent) - Ορυχεία και λατομεία (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Construction (per cent) - Κατασκευές (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Employment rate - Ποσοστό εργασίας (αντίθετο του ποσοστού ανεργίας)
- ✓ Households (per cent) - Νοικοκυριά (επί τοις εκατό της συνολικής τελικής κατανάλωσης ενέργειας)
- ✓ Total energy primary Production - Συνολική εγχώρια πρωτογενής παραγωγή ενέργειας
- ✓ Hotels / Motels / Bungalows AA - Ξενοδοχεία / Μοτέλ / Bungalows AA (κατηγορίας AA) (πληρότητα των Ξενοδοχείων / Μοτέλ / Bungalows AA (κατηγορίας AA) σαν ποσοστό επί της συνολικής πληρότητας όλων των καταλυμάτων στην Ελλάδα)
- ✓ Hotels / Motels / Bungalows A - Ξενοδοχεία / Μοτέλ / Bungalows A (κατηγορίας A)
- ✓ Hotels / Motels / Bungalows C - Ξενοδοχεία / Μοτέλ / Bungalows Γ (κατηγορίας Γ)
- ✓ Vacation centers - Παραθεριστικά Κέντρα
- ✓ Arrivals / Greeks (per cent) - Αφίξεις Ελλήνων (επί τοις εκατό των συνολικών αφίξεων σε όλα τα αεροδρόμια της Ελλάδας)
- ✓ Arrivals / Foreigners (per cent) - Αφίξεις Αλλοδαπών (επί τοις εκατό των συνολικών αφίξεων σε όλα τα αεροδρόμια της Ελλάδας)
- ✓ Night Stays / Greeks (per cent) - Διανυκτερεύσεις Ελλήνων (επί τοις εκατό των συνολικών διανυκτερεύσεων περιηγητών σε καταλύματα σε όλη την Ελλάδα)
- ✓ Average of Night Stays / Foreigners - Μέσος όρος διανυκτερεύσεων αλλοδαπών (σε όλη την Ελλάδα)

- ✓ Real estate renting and business activities - Διαχείριση ακίνητης περιουσίας, εκμισθώσεις και επιχειρηματικές δραστηριότητες
- ✓ Education (per cent) – Εκπαίδευση (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Wage-earning population - Μισθωτός πληθυσμός
- ✓ Expenditure on public investments - Δαπάνες για τις δημόσιες επενδύσεις
- ✓ Interests - Τόκοι (δημοσιονομικά έξοδα)
- ✓ Direct taxation - Άμεση Φορολογία
- ✓ Non-tax income - Μη φορολογικά έσοδα
- ✓ Public investments revenue - Εισόδημα από δημόσιες επενδύσεις

7.13.3. Κανόνες Ταξινόμησης.

Από το χαρακτηριστικό «αναλογία των υπηρεσιών - εξαγωγές προς εισαγωγές» και τα συσχετιζόμενα με αυτό χαρακτηριστικά προκύπτουν, από τους αλγόριθμους JRip και PART, οι εξής κανόνες:

A) JRIP rules:

1^{ος} κανόνας :

Hotels / Motels / Bungalows AA = '(10.9-11.12]' => Education='(-inf-0.041]' (2.0/0.0)

Εάν η πληρότητα των Ξενοδοχείων / Μοτέλ / Bungalows AA (κατηγορίας AA) αποτελεί το 10,9 % με 11,12 % της συνολικής πληρότητας όλων των καταλυμάτων στην Ελλάδα, τότε οι Έλληνες που απασχολούνται πλήρως στον τομέα της εκπαίδευσης αποτελούν, κατά το μέγιστο, το 4,1 % όλων των εργαζόμενων Ελλήνων που απασχολούνται πλήρως σε όλους ανεξαρτήτως τους τομείς.

2^{ος} κανόνας :

=> Education='(0.049-inf)' (10.0/0.0)

Οι Έλληνες που απασχολούνται πλήρως στον τομέα της εκπαίδευσης αποτελούν, κατά το ελάχιστο, το 4,9 % όλων των εργαζόμενων Ελλήνων που απασχολούνται πλήρως σε όλους τους τομείς, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 2

B) PART decision list : («μεταβλητή - στόχος» είναι οι Έλληνες απασχολούμενοι στον τομέα της εκπαίδευσης, σαν ποσοστό επί της συνολικής - σε όλους τους τομείς - πλήρους απασχόλησης όλων των εργαζομένων).

: '(0.049-inf)' (12.0/2.0)

Οι Έλληνες που απασχολούνται πλήρως στον τομέα της εκπαίδευσης αποτελούν, κατά το ελάχιστο, το 4,9 % όλων των εργαζόμενων Ελλήνων που απασχολούνται πλήρως σε όλους τους τομείς, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

7.14] Total Energy Gross Inland Consumption - Συνολική Ακαθάριστη Ενεργειακή Εγχώρια Κατανάλωση.

7.14.1. Ορισμός.

Ο ορισμός εξυπακούεται.

7.14.2. Επιλογή των πιο σχετικών, με τη συνολική Ακαθάριστη Ενεργειακή Εγχώρια Κατανάλωση, χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με τη συνολική Ακαθάριστη Ενεργειακή Εγχώρια Κατανάλωση είναι τα εξής δέκα :

- ✓ Total energy consumption from fossil fuels (sum) - Συνολική κατανάλωση ενέργειας από φυσικά καύσιμα (συνολικά όλα τα φυσικά καύσιμα)
- ✓ Total energy consumption from fossil fuels (crude oil and petroleum products) - Συνολική κατανάλωση ενέργειας από φυσικά καύσιμα (αργό πετρέλαιο και προϊόντα αυτού)
- ✓ Total energy gross inland consumption (per capita) - Συνολική εγχώρια ακαθάριστη κατανάλωση ενέργειας (κατά κεφαλή)
- ✓ Other community social and personal service activities - Παραγωγή άλλων υπηρεσιών υπέρ του κοινωνικού συνόλου
- ✓ Agriculture hunting and forestry - Γεωργία, θήρα και δασοκομία
- ✓ Fishing - Αλιεία
- ✓ Real estate renting and business activities - Διαχείριση ακίνητης περιουσίας, εκμισθώσεις και επιχειρηματικές δραστηριότητες
- ✓ Population - Πληθυσμός
- ✓ Imports - Εισαγωγές
- ✓ Non-tax income - Μη φορολογικά έσοδα

7.14.3. Κανόνες Ταξινόμησης.

Από το χαρακτηριστικό «Συνολική ακαθάριστη ενεργειακή εγχώρια κατανάλωση» και τα συσχετιζόμενα με αυτό χαρακτηριστικά δεν προκύπτουν οποιοδήποτε «χρήσιμοι» κανόνες ταξινόμησης (JRip ή PART).

7.15] Unemployment Rate - Ποσοστό Ανεργίας.

7.15.1. Ορισμός.

Ως ποσοστό ανεργίας ορίζεται το ποσοστό των ανέργων μιας περιοχής ως προς το συνολικό εργατικό δυναμικό (το άθροισμα των ανέργων και των απασχολούμενων στη συγκεκριμένη περιοχή) σε αυτήν. Αποτελεί βασικό δείκτη του παραγωγικού συστήματος μιας περιοχής, ο οποίος μετρά και ποσοτικοποιεί την ανεργία σε αυτήν.⁴⁷

7.15.2. Επιλογή των πιο σχετικών, με το ποσοστό ανεργίας, χαρακτηριστικών (selected attributes).

Τα χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με το συνολικό αριθμό επιβατών είναι τα εξής τέσσερα :

- ✓ Hotels / Motels / Bungalows AA - Ξενοδοχεία / Μοτέλ / Bungalows AA (κατηγορίας AA) (βαθμός πληρότητας σε ένα χρόνο)
- ✓ Hotels and restaurants (per cent) - Ξενοδοχεία και εστιατόρια (επί τοις εκατό της απασχόλησης των εργαζομένων)
- ✓ Income (per cent G.D.P.) - Εισόδημα (επί τοις εκατό του Α.Ε.Π.)
- ✓ Non-tax income - Μη φορολογικά έσοδα

7.15.3. Κανόνες Ταξινόμησης.

Από το ποσοστό ανεργίας στην Ελλάδα και τα συσχετιζόμενα με αυτό χαρακτηριστικά προκύπτουν, από τους αλγόριθμους JRip και PART, οι εξής κανόνες:

A) JRIP rules:

1^{ος} κανόνας :

Hotels / Motels / Bungalows AA = '(44.92-46.16]') =>Hotels and restaurants='(-inf-0.071]'
(2.0/0.0)

Εάν ο βαθμός πληρότητας («πόσο γέμισαν») των Ξενοδοχείων / Μοτέλ / Bungalows AA (κατηγορίας AA) στην Ελλάδα, συνυπολογιζομένων των βαθμών πληρότητας των καταλυμάτων αυτής της κατηγορίας και από τους Έλληνες και από τους αλλοδαπούς, κυμαίνεται μεταξύ 44,92% και 46,16 %, τότε οι πλήρως απασχολούμενοι εργαζόμενοι στον τομέα των ξενοδοχείων

⁴⁷ Βλ. Ολοκληρωμένο Πληροφοριακό Σύστημα Περιφερειακής Ανάπτυξης, «Ανεργία», <http://www.infopeloponnisos.gr/innovation/index1_4.html>, 20/6/2009.

και εστιατορίων δεν ξεπερνούν το 7,1 % επί της συνολικής - σε όλους τους τομείς - πλήρους απασχόλησης όλων των εργαζομένων.

2^{ος} κανόνας :

=> Hotels and restaurants='(0.079-inf)' (10.0/2.0)

Οι πλήρως απασχολούμενοι εργαζόμενοι στον τομέα των ξενοδοχείων και εστιατορίων, στην Ελλάδα, αποτελούν, κατά το ελάχιστο, το 7,9 % της συνολικής - σε όλους τους τομείς - πλήρους απασχόλησης όλων των εργαζομένων, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 2

B) PART decision list : («μεταβλητή - στόχος» είναι οι Έλληνες απασχολούμενοι στον τομέα των ξενοδοχείων / εστιατορίων, σαν ποσοστό επί της συνολικής - σε όλους τους τομείς - πλήρους απασχόλησης όλων των εργαζομένων).

: '(0.079-inf)' (12.0/4.0)

Οι πλήρως απασχολούμενοι εργαζόμενοι στον τομέα των ξενοδοχείων και εστιατορίων, στην Ελλάδα, αποτελούν, κατά το ελάχιστο, το 7,9 % της συνολικής - σε όλους τους τομείς - πλήρους απασχόλησης όλων των εργαζομένων, ανεξαρτήτως των υπολοίπων χαρακτηριστικών.

Number of Rules : 1

8] Σύνοψη.

8.1] Απολογισμός.

Η παρούσα διπλωματική εργασία αποτελεί μία καινοτομική εισήγηση στα πλαίσια της διαδικασίας αποτίμησης επιπτώσεων των Ολυμπιακών Αγώνων «Αθήνα 2004». Δεν έχει επιχειρηθεί η συνέχιση παραδοσιακών μεθόδων οικονομικής στατιστικής ανάλυσης, αλλά η εξέταση του ζητήματος των Ολυμπιακών Αγώνων υπό ένα πρωτοποριακό πρίσμα: αυτό της εξόρυξης γνώσης μέσα από δεδομένα (data mining). Πολλές φορές, σε κάποια οικονομικά ή άλλα στοιχεία υπάρχουν, πέραν των αριθμών, πληροφορίες «κρυμμένες». Τις πληροφορίες αυτές είναι δύσκολο - έως αδύνατο πολλές φορές - να τις αντιληφθεί ο ερευνητής, μιας και δεν είναι προφανείς. Έτσι, αποφασίστηκε όπως το παρόν σύγγραμμα επιχειρήσει να αναδείξει αυτές τις «κρυμμένες» πληροφορίες, εξετάζοντας τους οικονομικούς δείκτες OGGI⁴⁸ από μια διαφορετική «οπτική γωνία», αυτήν της Τεχνητής Νοημοσύνης (Artificial Intelligence) και πιο συγκεκριμένα, της Μηχανικής Μάθησης (Machine Learning). Στόχο αποτέλεσε η διαδικασία ανακάλυψης (discovery) προτύπων (patterns) που πριν δεν ήταν γνωστά, αλλά ισχύουν και είναι πιθανόν χρήσιμα, από ένα σύνολο δεδομένων (data set) καθώς και η περαιτέρω ανάλυσή τους, ούτως ώστε να βρούμε μη αναμενόμενες συσχετίσεις μεταξύ των όλων δεδομένων. Αυτά τα πρότυπα που ανακαλύφθηκαν είναι υπό τη μορφή «ευέλικτων» κανόνων, οι οποίοι προέκυψαν μέσω μίας μεθοδευμένης ανάλυσης data mining και οι οποίοι μπορούν, μελλοντικά, να επεκταθούν σε ένα νέο σύνολο δεδομένων (data set), ώστε να βοηθήσουν στην εξαγωγή χρήσιμων συμπερασμάτων.

Αρχικά, συγκεντρώθηκαν και μελετήθηκαν εκτενώς όλοι οι δείκτες του Τομέα της Οικονομίας που υπολογίστηκαν στα πλαίσια του προγράμματος «OGGI-Αθήνα 2004». Αξίζει να σημειωθεί ότι πολλοί από τους δείκτες ήταν ελλιπείς (μέχρι το 2003) και έτσι πραγματοποιήθηκε μια προσπάθεια συμπλήρωσής τους μέχρι το 2007, πράγμα σαφώς δύσκολο, αφού η περίοδος 2004 - 2007 θεωρείται σχετικά «πρόσφατο παρελθόν» και συνεπώς δεν ήταν δυνατόν να επισημανθούν επαρκή επίσημα στοιχεία για αυτήν. Συγκεντρώθηκε τελικά ένα αρκετά μεγάλο σύνολο δεδομένων, το οποίο αποτέλεσε και το «καύσιμο» για την περαιτέρω «αναρρίχηση στα άγνωστα και αβέβαια βουνά» της εξόρυξης γνώσης μέσα από δεδομένα. Ακολούθως, το σύνολο των δεδομένων (data set) που συγκεντρώθηκε τροποποιήθηκε κατάλληλα⁴⁹ και εισήχθηκε μέσα σε ένα λογισμικό, το οποίο εφαρμόζει πιστά όλες τις διαδικασίες που μεθοδεύει

⁴⁸ Βλ. σχετικά στην παρούσα εργασία : κεφάλαιο 6.1.1. της Μεθοδολογίας.

⁴⁹ Βλ. σχετικά στην παρούσα εργασία : κεφάλαιο 6.1.2. της Μεθοδολογίας.

η επιστήμη data mining, το λογισμικό WEKA. Μέσα από το λογισμικό αυτό προέκυψαν αρκετοί κανόνες «τεχνητής νοημοσύνης» (κανόνες συσχέτισης και ταξινόμησης των χαρακτηριστικών που υπάρχουν μέσα στα δεδομένα),⁵⁰ οι οποίοι έτυχαν μεθοδικής ανάλυσης και επεξεργασίας, ώστε να βγουν κάποια συμπεράσματα φυσικής σημασίας και να επεξεργαστούν μελλοντικά σε νέα, βελτιωμένα σύνολα δεδομένων που θα αφορούν δείκτες της οικονομίας σε σχέση με τους Ολυμπιακούς Αγώνες.

Από μόνο του το γεγονός ότι από ένα σύνολο 400 περίπου χαρακτηριστικών (attributes) επιλέχθηκαν τα 15 που θεωρήθηκαν ως τα σπουδαιότερα, είναι ιδιαίτερα σημαντικό. Από εκεί και πέρα, τα 15 αυτά χαρακτηριστικά αναλύθηκαν εκτενέστερα και έτυχαν περαιτέρω επεξεργασίας, από την οποία βγήκαν σημαντικοί κανόνες (rules). Επιχειρήθηκε η εξεύρεση κατάλληλου αλγόριθμου, ώστε να γίνει μία εκ βαθέων αναζήτηση συσχετίσεων των 15 συγκεκριμένων χαρακτηριστικών με τα υπόλοιπα, κάτι που επιτυγχάνεται με την επιλογή «selected attributes» του λογισμικού, κατά την οποία επιλέγονται τα περισσότερα συσχετιζόμενα μεταξύ τους χαρακτηριστικά. Μέσα από μία μεγάλη γκάμα χαρακτηριστικών δηλαδή, καταφέραμε να δημιουργήσουμε μικρότερα υποσύνολα, για περαιτέρω μελέτη. Η εύρεση αυτών όλων των υποσυνόλων είναι σημαντικού ενδιαφέροντος, αφού μέσα από το ευρύ σύνολο των διαθέσιμων μεταβλητών, υποδείχθηκε κάθε φορά το υποσύνολο εκείνων που σχετίζονται με δεδομένο μέγεθος. Παράλληλα, μετά την επιλογή των πιο σχετικών μεταξύ τους χαρακτηριστικών, πραγματοποιήθηκε μία ταξινόμηση (classification) των δεδομένων, με σκοπό την τελική εξαγωγή κανόνων για τα διάφορα χαρακτηριστικά.⁵¹ Συγκεκριμένα, επιλέχθηκαν οι αλγόριθμοι - ταξινομητές JRip και PART του λογισμικού WEKA, οι οποίοι παράγουν λίστες αποφάσεων και έχουν τη μορφή λογικών κανόνων που εξάγουν ανάλογα συμπεράσματα. Οι λίστες αποφάσεων που παράχθηκαν, μολονότι δε δηλώνουν κάτι άμεσα - με μια πρώτη δηλαδή ματιά -, εντούτοις, συσχετίζουν οικονομικούς δείκτες και άλλα δεδομένα, των οποίων ο συνδυασμός - περνώντας ίσως απαρατήρητος από τον ερευνητή - αποτελεί ένα πολυδύναμο σημείο για περαιτέρω μελέτη και προβληματισμό.

Εν τέλει, η αποτίμηση των επιπτώσεων των Ολυμπιακών Αγώνων στους διάφορους δείκτες της Οικονομίας αποτελεί πολύπλευρο ζήτημα. Δεν καθίσταται δυνατόν να δοθεί στην αποτίμηση θετικό ή αρνητικό πρόσημο. Δεν υπάρχει σαφής και τεκμηριωμένη απάντηση σε ένα

⁵⁰ Βλ. σχετικά στην παρούσα εργασία : κεφάλαιο 7 - Παρουσίαση και Ανάλυση Αποτελεσμάτων.

⁵¹ Βλ. σχετικά στην παρούσα εργασία : κεφάλαιο 6.5, «3^ο Στάδιο - Επιλογή και «τρέξιμο» κατάλληλων αλγορίθμων για περαιτέρω επεξεργασία των δεδομένων, μέσω του λογισμικού WEKA», της Μεθοδολογίας.

τέτοιο ζήτημα, παρά μόνο υποκειμενική γνώμη. Η προσωπική μου γνώμη, έπειτα από την εκτενή ενασχόλησή μου με τις οικονομικές επιπτώσεις των Ολυμπιακών Αγώνων με τεχνικές data mining, στα πλαίσια της διπλωματικής εργασίας; Θα έλεγα, χαριτολογώντας, ότι σίγουρα η Αθήνα και κατ' επέκταση όλη η Ελλάδα, αποκόμισε πολλά οφέλη από τους Ολυμπιακούς Αγώνες «Αθήνα 2004», μα ταυτόχρονα ζημιώθηκε από αυτούς. Ωστόσο, το κατά πόσον η ζυγαριά της αποτίμησης των οικονομικών συνεπειών των αθηναϊκών Ολυμπιακών Αγώνων, κλίνει προς την πλευρά του χρέους ή του κέρδους, είναι μια άλλη ιστορία, πολυσυζητημένη μεν, μυθοποιημένη δε, αφού λείπει από τα ελληνικά επιστημονικά δρώμενα η εξονυχιστική εκείνη μελέτη, που θα μπορούσε να προσφέρει μια τεκμηριωμένη και άρτια απάντηση στο πολύπλοκο αυτό ερώτημα.

8.2] Προτεινόμενες Κατευθύνσεις Μελλοντικής Ενασχόλησης.

Οι επιλογές ενασχόλησης που υπήρχαν πριν την εκπόνηση της παρούσας εργασίας, σχετικά με το ζήτημα του πώς επηρέασαν οι Ολυμπιακοί Αγώνες της Αθήνας τους οικονομικούς δείκτες στην Ελλάδα, ήταν πάμπολλες. Θα μπορούσε να γίνει μια στατιστική ανάλυση των δεικτών και να βγουν αριθμητικά αποτελέσματα ή, απλά, να παρασταθούν γραφικά οι διάφορες οικονομικές μεταβολές. Η τελική επιλογή, όμως, της διερεύνησης των επιπτώσεων των Αγώνων για δημιουργία διάφορων κανόνων, μέσω τεχνικών μηχανικής μάθησης - και δη data mining -, αποδείχτηκε πολύ ενδιαφέρουσα και πρωτότυπη, αλλά ταυτόχρονα και λίγο «επισφαλής». Υπήρχαν, παράλληλα, και κάποιες αδυναμίες που οδήγησαν σε σφάλματα, οι οποίες βέβαια δύνανται μελλοντικά να ελαχιστοποιηθούν.

Όπως προαναφέρθηκε, πολλοί από τους δείκτες OGGI ήταν ελλιπείς (μέχρι το 2003). Με το πέρας ωστόσο κάποιων χρόνων, η ανεύρεση στοιχείων για την περίοδο αυτή και όχι μόνο (και μεταγενέστερα) θα είναι περισσότερο εφικτή μα και ορθή, αφού, όσο πιο πλούσιο είναι ένα σύνολο δεδομένων (data set), τόσο πιο έγκυρα και ορθά διεξάγεται μία τεχνική data mining.

Ταυτόχρονα, ένα άλλο πρόβλημα που συναντήθηκε, κατά την παρούσα μελέτη, και το οποίο μπορεί να περιοριστεί σε πιθανή μελλοντική ενασχόληση με το εν λόγω ζήτημα, αποτελεί η αίσθηση αβεβαιότητας που ενυπήρχε καθ' όλη την πορεία διεξαγωγής της εργασίας, γέννημα της ύπαρξης ελάχιστων υποδειγμάτων (instances) στο σύνολο των δεδομένων (data set). Η αναλογία «υποδείγματα / χαρακτηριστικά» (instances / attributes) πρέπει να είναι περίπου «10/1». Το συγκεκριμένο πρόβλημα εναπόκειται στην κρίση και την έρευνα του αυριανού μελετητή, ο οποίος, διαθέτοντας το πλεονέκτημα του επιπρόσθετου, καταγεγραμμένου και συνεπώς πιο έγκυρου υλικού (που αναφέρεται σε ήδη περασμένες χρονιές), θα βρίσκεται σε

θέση ευνοϊκή, ώστε να θέσει, ως υποδείγματα, αντί για έτη, εξάμηνα, μήνες ή ακόμα και εβδομάδες. Συνεπώς, με την αύξηση του αριθμού των υποδειγμάτων, επιτυγχάνεται κατά τρόπο αναλογικό και η άμβλυση του βαθμού αβεβαιότητας.

8.3] Αντί επιλόγου.

Η όλη ενασχόληση με το θέμα αποτίμησης των επιπτώσεων των Ολυμπιακών Αγώνων στους δείκτες OGGI στη σφαίρα της Οικονομίας, μέσω μεθοδολογίας που εφαρμόζει τεχνική Data Mining, αποτελεί, για το μηχανολόγο μηχανικό, τη μύησή του σε ένα άγνωστο ερευνητικό πεδίο. Κατά την καβαφική προσέγγιση της «Οδύσσειας», σημασία δεν έχει η Ιθάκη αυτή καθ' εαυτή, αλλά το ταξίδι προς αυτήν. Συναντήθηκαν αρκετά προβλήματα στην πορεία προς την επίτευξη του στόχου της εργασίας, μα εν τέλει, οι γνώσεις που αποκομίστηκαν απ' όλο αυτό το ενδιαφέρον ταξίδι στο μέχρι πρότινος άγνωστο κόσμο της τεχνητής νοημοσύνης, υπερκάλυψαν τις οποιεσδήποτε αμφιβολίες και δυσφορίες σκέψεις είχαν παρουσιαστεί, σκέψεις που δημιουργούνται εύλογα στον κάθε μελετητή που αναλαμβάνει δειλά δειλά να φέρει εις πέρας ένα σημαντικό έργο, όσον αφορά τη σπουδαιότητα του ζητήματος ενασχόλησης και να προσφέρει με τρόπο καινοτομικό στα επιστημονικά δρώμενα.

Η «τελείωση» του σκοπού, είναι σαφώς, πάντα κατά ένα ποσοστό, ατέλειωτη. Πόσο μάλλον σε μία εργασία μικρής σχετικά έκτασης, όπως και η συγκεκριμένη, που τελείται στα πλαίσια μιας πτυχιακής / διπλωματικής μελέτης. Εκτιμούμε ωστόσο το γεγονός, πως το συγκεκριμένο ζήτημα που τέθηκε με την παρούσα εργασία, αποτέλεσε εν τέλει μια δυναμική αφορμή, ώθηση, για άνοιγμα νέων οριζόντων επιστημονικής έρευνας, για επαφή με νέα δεδομένα και πεδία γνώσεων, άγνωστα ή και ήδη γνωστά, που παρατηρούνται πάντα υπό το πρίσμα μιας νέας οπτικής. Όπως λέει χαρακτηριστικά και ο μεγάλος Έλληνας ποιητής Κ. Π. Καβάφης : «Η Ιθάκη σ' έδωσε τ' ωραίο ταξίδι. Χωρίς αυτήν δεν θα 'βγαίνες στον δρόμο. (...) [και] να εύχεσαι να 'ναι μακρύς ο δρόμος...» [22]. Στη δική μας περίπτωση, το ταξίδι μόλις τώρα αρχίζει...

9] Βιβλιογραφία.

9.1] Συγγράμματα.

- [1] Preuss Holger, *The Economics of Staging the Olympics*, εκδ. MPG Books Ltd, Bodmin, Cornwall, Great Britain, 2004.
- [2] Σάμιτας Αριστείδης, Ζούνης Πέτρος, Καινούργιος Δημήτρης, «Η επίδραση των Ολυμπιακών Αγώνων στο επιχειρηματικό περιβάλλον και οι προσδοκίες των εθνικών χορηγών», *Πρακτικά της ελληνικής διοίκησης επιχειρησιακού αθλητισμού - Διάσκεψη*, τόμος 1, σ. 187, 12/12/2004, Πάτρα.
- [3] Βουτσινάς Βασίλειος, *Θέματα Επιχειρηματικής Νοημοσύνης, Θεωρητική Θεμελίωση και Εφαρμογές*, εκδ. Κωσταράκη Π. Ευρυδίκη, Αθήνα, 2006.
- [4] Maimon Oded, Rokach Lior, *The Data Mining and Knowledge Discovery Handbook*, εκδ. Springer, 2005.
- [5] Τσακαλίδης Αθανάσιος, Χατζηλυγερούδης Ιωάννης, *Τεχνητή Νοημοσύνη και έμπειρα Συστήματα*, Πάτρα, 2002.
- [6] Witten Ian, Eibe Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, εκδ. Morgan Kauffman, 2005.
- [7] Χαλκίδη Μ., Βαρζιγιάννης Μ., *Εξόρυξης γνώσης από βάσεις δεδομένων και τον παγκόσμιο ιστό*, εκδ. Τυπωθήτω / Gutenberg, 2006.
- [8] Cios K. J., Pedrycz W., Swiniarski R. W., Kurgan L. A., *Data Mining : A Knowledge Discovery Approach*, εκδ. Springer, 2007.
- [9] Fayyad M. Usama, Piatesky-Shapiro Gregory, Smyth Padhraic, Uthurusamy Ramasany, *Advances in Knowledge Discovery and Data Mining*, εκδ. American Association for Artificial Intelligence, Menlo Park, California, 1996.
- [10] Dunham H. Margaret, *Data Mining: Introductory and Advanced Topics*, Επιμέλεια ελληνικής έκδοσης : Βερύκιος Β. και Θεοδωρίδης Γ., εκδ. Νέων Τεχνολογιών, 2004.
- [11] Mitchell M. Tom, *Machine Learning*, εκδ. McGraw - Hill International Editions, 1997.

- [12] Tan P., Steinbach M., Kumar V., *Introduction to Data Mining*, εκδ. Addison - Wesley, 2005.
- [13] Guillet Fabrice, Hamilton J. Howard, *Quality Measures in Data Mining*, εκδ. Springer, 2005.
- [14] Pyle Dorian, *Data Preparation for Data Mining, The Morgan Kaufmann Series in Data Management Systems*, εκδ. Morgan Kauffman, 1999.
- [15] Hand D., Mannila H., Smyth P., *Principles of Data Mining*, εκδ. MIT Press, Cambridge, MA, 2001.
- [16] Richter M. Michael, *Advances in Data Mining : Theoretical Aspects and Applications*, εκδ. Springer (Petra Perner), 2007.
- [17] Han Jiawei, Kamber Micheline, *Data Mining : Concepts and Techniques*, εκδ. Morgan Kauffman, 2006.
- [18] Berka Petr, Bruha Ivan, *Discretization and grouping : Preprocessing steps for data mining*, εκδ. Springer Berlin / Heidelberg, 1998.
- [19] Kantardzic M., *Data Mining : Concepts, Models, Methods, and Algorithms*, εκδ. Wiley - IEEE Press, 2002.
- [20] Mohammadian Masoud, *Intelligent Agents : For Data Mining and Information Retrieval*, Canberra, Australia, 2004.
- [21] Δούμπος Μιχάλης, Ζοπουνίδης Κωνσταντίνος, *Πολυκριτήριες Τεχνικές Ταξινόμησης: Θεωρία και Εφαρμογές*, εκδ. Κλειδάριθμος, 2001.
- [22] Καβάφης Κ. Π., *Τα Ποιήματα (1897- 1918)*, τ. Α', επιμέλεια Γ.Π. Σαββίδης, εκδ. Ίκαρος, Αθήνα 2002.

9.2] Εφημερίδες - Περιοδικά.

- Δρισάκης Νικόλαος, «Δημόσιες Δαπάνες και Οικονομική Ανάπτυξη, Μια εμπειρική έρευνα για την Ελλάδα με την ανάλυση της συνολοκλήρωσης», περ. *Σπουδαί*, τόμος 53, τεύχος 4^ο, 2003, σ.66.
- Πασακυριάκος Κώστας, «Πιθανότερες οι αρνητικές επιπτώσεις, Συνέντευξη με το Γρηγόρη Παπανίκο, μέλος του Πανελληνίου Συνδέσμου Αθλητικών Οικονομολόγων και Στελεχών», εφημ. *Ριζοσπάστης*, 1/7/2001, σ.19.
- Υπουργείο Οικονομικών, «Με τη νέα κλίμακα η παρακράτηση φόρου, Δημοσιοποιήθηκε η σχετική εγκύκλιος του Υπουργείου», *Ριζοσπάστης*, 4 / 12 / 2001, σ.24.

9.3] Ηλεκτρονικές Πηγές.

- «What is the Olympic Games Global Impact?», Olympic Review - Official Publication of the Olympic Movement, June 2006,
<http://multimedia.olympic.org/pdf/en_report_1077.pdf>.
- media info 2004, «Ολυμπιακά θέματα : Δημόσια έργα - Μεταφορές»,
< <http://www.mediainfo2004.gr/cgi-bin/hweb?-A=46&-V=olympicissues&-w=>>.
- OGGI Athens 2004, «Athens 2004, Olympic Games Global Impact»,
<http://147.102.46.88/OGGI/front-page/view?set_language=el>.
- Αθηναϊκό Πρακτορείο Ειδήσεων (ΑΠΕ), «Οφέλη από τη διοργάνωση των Ολυμπιακών Αγώνων», 10/10/2004,
< <http://www.forthnet.gr/templates/newsPosting.aspx?p=109413>>.
- Γωγουβίτης Ξένος, «Ολοκληρωμένα άρθρα : Ολυμπιάδα 2004 (η ανάθεση)», *Kybernografoi.gr*, τχ.10, Οκτώβριος 1998,
<<http://www.kybernografoi.gr/modules.php?name=News&file=article&sid=124>>.
- Επίσημα στοιχεία της Γενικής Γραμματείας της Εθνικής Στατιστικής Υπηρεσίας Ελλάδας (Ε.Σ.Υ.Ε.),
<<http://www.statistics.gr/>>.
- Επίσημα στοιχεία της Ευρωπαϊκής Στατιστικής Υπηρεσίας Eurostat,
<<http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home>>.
- Επίσημη ιστοσελίδα του λογισμικού WEKA,
<<http://www.cs.waikato.ac.nz/ml/weka/>>.

- Θεοδωρίδης Γιάννης, Πελέκης Νίκος, «Αποθήκες Δεδομένων και εξόρυξη γνώσης, Εισαγωγή»,
<<http://infolab.cs.unipi.gr/courses/dwdm/slides/1-DWDM-intro.pdf>>.
- Ίδρυμα Μείζονος Ελληνισμού, «Ολυμπιακοί Αγώνες και σύγχρονη ιστορία, οι επικερδείς Αγώνες: από το Λος Άντζελες στο Σίδνεϊ»,
<<http://www.fhw.gr/olympics/modern/gr/history/h302.html>>.
- Μπακογιάννη Ντόρα, «Οι Ολυμπιακοί : Κάτι παραπάνω από απλοί αγώνες, μια παρακαταθήκη για το μέλλον - Πανεπιστήμιο του Πεκίνου»,
<<http://www.dorabak.gr/default.asp?pid=10&riD=1016&la=1>>.
- Ντούτση Ειρήνη, «Ανακάλυψη Γνώσης από Δεδομένα και Εξόρυξη Γνώσης στο εργαλείο WEKA»,
<<http://infolab.cs.unipi.gr/courses/dwdm/tutorials/WEKA-tutorial-2008.pdf>>.
- Ολοκληρωμένο Πληροφοριακό Σύστημα Περιφερειακής Ανάπτυξης, «Ανεργία»,
<http://www.infopeloponnisos.gr/innovation/index1_4.html>.
- Παγουρόπουλος Απόστολος, «Διπλωματική εργασία με θέμα : Data Mining στη Χρηματοοικονομική Ανάλυση», Πάτρα, 17/10/2006,
<http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/867/1/Nimertis_Pagouropoulos.pdf>.
- Παπανδρέου Ανδρέας, «Προϋπολογισμός 2007, Μια κριτική αποτίμηση», Δεκέμβριος 2006,
<http://www.istame-apapandreou.gr/files/pdf/Budget2007_final.pdf>.
- Σταμελάτου Μαρία - Ελένη, *Ακάλυπτη Ισοτιμία Επιτοκίων, Οικονομική και Επιχειρησιακή Στρατηγική*, Διπλωματική Εργασία, Πανεπιστήμιο Πειραιά, 2006,
<<http://digilib.lib.unipi.gr/dspace/bitstream/unipi/1539/4/Stamelatou.pdf>>.
- Τζιραλής Γιώργος, «Αλγόριθμοι Εξόρυξης Πληροφορίας, Διάλεξη 1, Εισαγωγή»,
<http://gtziralis.googlepages.com/Lecture01_Introduction.pdf>.
- Τζιραλής Γιώργος, «Αλγόριθμοι Εξόρυξης Πληροφορίας, Διάλεξη 02 - Συνιστώσες Δεδομένων, Οπτικοποίηση και Εξερεύνηση»,
<http://gtziralis.googlepages.com/Lecture02_DataComponentsVisualizationExploration.pdf>.
- Τζιραλής Γιώργος, «Αλγόριθμοι Εξόρυξης Πληροφορίας, Διάλεξη 3, Προεπεξεργασία και Επιλογή Δεδομένων»,
<http://gtziralis.googlepages.com/Lecture03_DataPreprocessAttributeSelection.pdf>.

- Χριστοδουλάκης Νίκος, «Η ελληνική οικονομία μετά τους Ολυμπιακούς Αγώνες», ομιλία σε συνέδριο της International Herald Tribune, 6/11/2004, Αθήνα, <<http://www.christodoulakis.gr/%289909C513EF7D8FB4372FAAB3986228FDB6283DFD8CFC47F5%29/ecportal.asp?id=4480&nt=18&lang=1>>.
- Ορολογίες διάφορων οικονομικών και άλλων δεικτών, <http://en.wikipedia.org/wiki/Main_Page>, <<http://www.investopedia.com/dictionary/default.asp>>.

9.4] Ηλεκτρονικές Διευθύνσεις για Εικόνες.

- <http://147.102.46.88/OGGI/methodos-analysis>
- http://gtziralis.googlepages.com/Lecture01_Introduction.pdf
- http://gtziralis.googlepages.com/Lecture02_DataComponentsVisualizationExploration.pdf
- <http://infolab.cs.unipi.gr/courses/dwdm/slides/1-DWDM-intro.pdf>
- http://nemertes.lis.upatras.gr/dspace/bitstream/123456789/867/1/Nimertis_Pagouropoulos.pdf
- <http://www.aists.org/en/aists/the-aists-thinking-beyond-the-score.html>

10] Παραρτήματα.

10.1] Παράρτημα Α - Δείκτες ΟGGI πλαισίου των Αγώνων (context indexes).

Ec01 Δομή Απασχόλησης

Δομή Απασχόλησης

Γεωγραφική περιοχή : Ελλάδα

Μονάδες : Ισοδύναμοι εργαζόμενοι πλήρους απασχόλησης (% συνόλου)

Δείκτες	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
A - Γεωργία, θήρα, δασοκομία	17.26%	17.08%	17.73%	17.61%	16.97%	15.78%	15.24%	15.04%	12.31%	12.08%	11.70%	11.16%
B - Αλιεία	0.70%	0.66%	0.29%	0.32%	0.31%	0.32%	0.32%	0.31%	0.30%	0.33%	0.32%	0.59%
C - Ορυχεία , λατομεία	0.49%	0.47%	0.52%	0.49%	0.49%	0.47%	0.47%	0.31%	0.35%	0.42%	0.40%	0.30%
D - Μεταποίηση	16.96%	16.72%	14.51%	14.28%	13.98%	14.12%	13.84%	13.27%	13.06%	12.85%	12.60%	11.07%
E - Παροχή ηλεκτρικού ρεύματος, φυσικού αερίου και νερού	1.19%	1.16%	0.93%	1.01%	0.95%	0.91%	0.89%	0.95%	0.90%	0.88%	0.92%	0.76%
F -Κατασκευές	6.81%	6.88%	7.37%	7.08%	7.34%	7.51%	7.64%	8.09%	8.12%	8.28%	8.10%	6.61%
G - Χονδρικό & λιανικό εμπόριο:επισκευές οχημάτων και ειδών νοικοκυριών	14.02%	14.48%	16.84%	16.98%	17.17%	17.24%	17.26%	17.44%	17.51%	17.79%	17.76%	21.08%
H - Ξενοδοχεία & εστιατόρια	5.24%	5.44%	6.18%	6.30%	6.49%	6.58%	6.75%	6.70%	6.36%	6.82%	6.73%	6.59%
I -Μεταφορές, αποθηκεύσεις, επικοινωνίες	6.59%	6.38%	6.13%	6.23%	6.16%	6.27%	5.94%	6.11%	6.25%	6.20%	6.38%	5.29%
J - Ενδιάμεσοι χρηματοπιστωτικοί οργανισμοί	2.13%	2.27%	2.33%	2.37%	2.63%	2.62%	2.46%	2.58%	2.62%	2.60%	2.61%	2.28%
K - Διαχείριση ακίνητης περιουσίας, εκμισθώσεις και επιχειρηματικές δραστηριότητες	4.81%	5.03%	4.79%	4.87%	4.91%	5.39%	5.84%	5.77%	6.51%	6.55%	6.41%	6.48%
L - Δημόσια διοίκηση και άμυνα, υποχρεωτική κοινωνική ασφάλιση	7.51%	7.10%	7.03%	7.15%	7.46%	7.45%	7.49%	7.56%	8.20%	7.96%	8.55%	9.53%
M - Εκπαίδευση	6.83%	6.99%	6.00%	6.03%	6.09%	6.42%	6.39%	6.62%	7.21%	7.13%	7.38%	6.83%
N - Υγεία και κοινωνική μέριμνα	4.70%	4.61%	4.57%	4.66%	4.57%	4.47%	4.51%	4.33%	5.11%	5.03%	5.07%	5.07%
O - Παραγωγή άλλων υπηρεσιών υπέρ του κοινωνικού συνόλου	3.70%	3.59%	3.47%	3.40%	3.23%	3.25%	3.67%	3.58%	3.62%	3.48%	3.43%	4.12%
P - Ιδιωτικά νοικοκυριά που απασχολούν οικιακό προσωπικό	1.05%	1.13%	1.32%	1.23%	1.24%	1.17%	1.28%	1.33%	1.57%	1.60%	1.63%	2.25%

Πηγή: Εθνική Στατιστική Υπηρεσία Ελλάδος (1996-1997:Eurostat)

Ec04 Κοινωνικο-οικονομικές Τάσεις

Κοινωνικο-οικονομικές Τάσεις

Γεωγραφική περιοχή: Ελλάδα

Μονάδες: Κάτοικοι (χιλιάδες)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Ενεργός πληθυσμός	4318.30	4294.40	4525.75	4586.10	4611.95	4580.25	4656.00	4734.45	4818.85	4846.48	4885.27	4916.80
Μόνιμοι κάτοικοι	10673.70	10744.65	10808.36	10861.40	10903.76	10931.21	10968.71	11006.38	11040.65	11082.75	11125.18	11171.74
Ενεργές γυναίκες	1680.90	1682.40	1765.63	1815.93	1830.75	1805.03	1861.75	1895.10	1959.78	1974.15	1996.97	2005.32
Άνεργοι	446.40	440.40	507.90	554.75	523.48	493.95	480.25	459.95	505.70	477.50	436.27	406.92

Πηγή: ΕΣΥΕ

Σημείωση: Μόνιμοι κάτοικοι στις αρχές κάθε έτους.

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

Μονάδες: Ποσοστό %

Δείκτες	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Ποσοστό απασχόλησης	40.46%	39.97%	41.87%	42.22%	42.30%	41.90%	42.45%	43.02%	43.65%	43.73%	43.91%	44.01%
Ποσοστό γυναικών στον ενεργό πληθυσμό	38.93%	39.18%	39.01%	39.60%	39.70%	39.41%	39.99%	40.03%	40.67%	40.73%	40.88%	40.79%
Ποσοστό ανεργίας	10.34%	10.26%	11.22%	12.10%	11.35%	10.78%	10.31%	9.71%	10.49%	9.85%	8.93%	8.28%

Πηγή: ΕΣΥΕ

Ec5 Κατανάλωση Ενέργειας ανά Κάτοικο

Κατανάλωση Ενέργειας ανά Κάτοικο

Γεωγραφική περιοχή: Ελλάδα

Μονάδα μέτρησης: Terajoules (Καθαρή θερμοδική αξία)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Συνολική ακαθάριστη εγχώρια κατανάλωση ενέργειας	1,063,670	1,071,206	1,125,212	1,120,360	1,175,483	1,211,529	1,244,987	1,262,697	1,282,454	1,302,211	1,361,482	1,400,996
Πληθυσμός	10,709	10,777	10,835	10,883	10,917	10,950	10,988	11,006	11,041	11,104	11,149	11,193
Κατανάλωση ενέργειας ανά κάτοικο	99.32	99.40	103.85	102.95	107.67	110.64	113.31	114.72	116.16	117.27	122.12	125.17

πηγή: ΕΣΥΕ, Eurostat

Ec6 Κατανάλωση ενέργειας κατά πρωτογενή πηγή

Γεωγραφική περιοχή: Ελλάδα

Μονάδα μέτρησης: Ποσοστό / Terajoules (Καθαρή θερμοδική αξία)

Δείκτες	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
φυσικά καύσιμα (%)	94.59%	94.76%	95.05%	94.61%	95.00%	95.44%	95.31%	94.87%	94.91%	95.06%	95.04%	95.46%
φυσικά καύσιμα (στερεά, %)	35.24%	34.46%	34.07%	31.85%	32.20%	32.17%	31.40%	29.53%	29.71%	29.30%	29.86%	29.00%
φυσικά καύσιμα (πετρέλαιο και προϊόντα αυτού)	58.70%	58.86%	57.77%	58.15%	56.74%	56.72%	57.02%	58.02%	57.13%	56.93%	56.15%	56.97%
πυρηνικά καύσιμα (%)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
ανανεώσιμες πηγές (%)	5.41%	5.24%	4.95%	5.39%	5.00%	4.56%	4.69%	5.13%	5.09%	6.63%	8.13%	8.81%
Συνολική ακαθάριστη εγχώρια κατανάλωση ενέργειας/κάτοικο	99.65	99.70	104.11	103.15	107.81	110.83	113.50	114.72	116.16	117.27	122.12	125.17

πηγή: Eurostat

Ec7 Κατανάλωση Ενέργειας ανά Χρήση

Γεωγραφική περιοχή: Ελλάδα

Μονάδες: Terajoules (Net calorific value = NCV)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Βιομηχανία	180,616	181,669	185,081	173,935	185,558	188,243	186,640	180,222	169,346	165,735	182,826	184,992
Νοικοκυριά	164,848	169,436	175,174	176,798	187,147	196,088	205,073	228,936	224,680	224,947	235,614	239,347
Υπηρεσίες	86,168	89,828	94,714	97,390	100,893	107,566	112,251	121,613	120,343	128,366	133,046	139,063
Μεταφορές	274,695	281,587	305,304	312,058	301,301	308,270	312,316	326,578	333,268	348,695	373,374	398,560
Συνολική τελική κατανάλωση	706,327	722,520	760,274	760,181	774,899	800,167	816,280	857,349	847,637	867,743	924,860	961,962

πηγή: Eurostat

Ec8 Ενεργειακή Αυτάρκεια

Γεωγραφική περιοχή: Ελλάδα

Μονάδες: Terajoules (Net calorific value = NCV)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Συνολική εγχώρια ακαθάριστη κατανάλωση ενέργειας	993,356	1,002,651	1,046,134	1,076,327	1,125,834	1,159,439	1,196,292	1,244,580	1,250,564	1,264,225	1,264,413	1,275,058
Συνολική εγχώρια πρωτογενής παραγωγή ενέργειας	350,571	342,228	344,424	352,294	371,562	381,598	402,469	395,019	403,396	388,749	377,779	398,486
Εισαγωγές ενέργειας	973,147	962,265	1,011,004	981,745	1,086,364	1,083,840	1,142,266	1,211,071	1,273,105	1,228,588	1,361,952	1,273,475
Εξαγωγές ενέργειας	194,357	149,475	136,292	159,801	171,772	162,062	178,407	248,524	240,044	245,848	299,529	247,720
Ενεργειακά αποθέματα	134,364	135,201	143,112	133,559	144,445	150,054	134,900	138,040	139,502	123,467	135,764	136,128
Μεταβολή αποθεμάτων	-1,641	-17,166	-29,890	35,648	-15,875	6,117	-35,136	25,054	-46,391	16,203	-40,024	-13,055

πηγή: Eurostat

σημείωση: Οι τιμές για το 2007 αποτελούν εκτίμηση

Ec12 Ξενοδοχειακή Υποδομή

Αριθμός Κλινών

Σύνολο Χώρας	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Ξενοδοχεία / Μοτέλ / Bungalows AA	204,465	221,929	220,750	236,841	254,336	213,190	217,600	236,804	346,257	389,054	611,554	837,829
Ξενοδοχεία / Μοτέλ / Bungalows A	752,434	795,690	789,834	794,691	841,429	788,563	788,531	839,268	874,089	914,559	963,945	1,120,586
Ξενοδοχεία / Μοτέλ / Bungalows B	733,848	789,067	798,648	779,845	799,828	727,281	684,467	729,099	674,206	745,468	718,333	687,301
Ξενοδοχεία / Μοτέλ / Bungalows Γ	696,149	751,467	746,068	755,752	760,554	712,208	681,000	712,489	726,053	699,044	633,404	573,737
Ξενοδοχεία / Μοτέλ / Bungalows Δ+Ε	197,735	201,448	195,809	191,901	188,550	171,960	162,021	158,245	162,032	141,794	103,609	109,901
Ξενώνες	68,885	69,428	69,905	67,156	66,797	60,686	60,830	63,062	6,531	1,633	0	0
Οικοτροφεία	10,502	10,461	9,817	8,883	7,626	6,818	6,779	6,158	612	163	0	0
Επιπλωμένα Διαμερίσματα	153,693	155,021	147,701	157,999	181,596	166,226	167,893	174,614	284,025	298,936	296,697	308,637
Κέντρα Παραθεριστικά	31,314	28,505	28,761	27,784	27,655	20,848	21,822	18,890	18,875	15,322	15,108	13,998
Σύνολο	2,849,025	3,023,016	3,007,293	3,020,852	3,128,371	2,867,780	2,790,943	2,938,629	3,092,680	3,205,973	3,342,650	3,651,989
Ξενοδοχεία / Μοτέλ / Bungalows AA	7.18%	7.34%	7.34%	7.84%	8.13%	7.43%	7.80%	8.06%	11.20%	12.14%	18.30%	22.94%
Ξενοδοχεία / Μοτέλ / Bungalows A	26.41%	26.32%	26.26%	26.31%	26.90%	27.50%	28.25%	28.56%	28.26%	28.53%	28.84%	30.68%
Ξενοδοχεία / Μοτέλ / Bungalows B	25.76%	26.10%	26.56%	25.82%	25.57%	25.36%	24.52%	24.81%	21.80%	23.25%	21.49%	18.82%
Ξενοδοχεία / Μοτέλ / Bungalows Γ	24.43%	24.86%	24.81%	25.02%	24.31%	24.83%	24.40%	24.25%	23.48%	21.80%	18.95%	15.71%
Ξενοδοχεία / Μοτέλ / Bungalows Δ+Ε	6.94%	6.66%	6.51%	6.35%	6.03%	6.00%	5.81%	5.38%	5.24%	4.42%	3.10%	3.01%
Ξενώνες	2.42%	2.30%	2.32%	2.22%	2.14%	2.12%	2.18%	2.15%	0.21%	0.05%	0.00%	0.00%
Οικοτροφεία	0.37%	0.35%	0.33%	0.29%	0.24%	0.24%	0.24%	0.21%	0.02%	0.01%	0.00%	0.00%
Επιπλωμένα Διαμερίσματα	5.39%	5.13%	4.91%	5.23%	5.80%	5.80%	6.02%	5.94%	9.18%	9.32%	8.88%	8.45%
Κέντρα Παραθεριστικά	1.10%	0.94%	0.96%	0.92%	0.88%	0.73%	0.78%	0.64%	0.61%	0.48%	0.45%	0.38%

πηγή: ΕΣΥΕ

Ec13 Βαθμός Πληρότητας Καταλυμάτων

Πληρότητα

Πληρότητα Ημεδαπών %	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Ξενοδοχεία / Μοτέλ / Bungalows AA	11.0	11.7	10.7	10.7	10.9	11.1	10.7	11.5	9.8	11.8	12.0	10.7
Ξενοδοχεία / Μοτέλ / Bungalows A	8.0	8.3	8.6	10.0	9.2	8.7	9.2	9.9	10.7	9.6	10.7	9.2
Ξενοδοχεία / Μοτέλ / Bungalows B	16.1	17.1	17.5	18.0	18.0	17.8	17.7	18.5	17.8	16.5	19.3	16.3
Ξενοδοχεία / Μοτέλ / Bungalows Γ	21.0	21.5	22.4	22.2	22.6	21.8	21.6	21.4	19.1	21.9	21.9	18.9
Ξενοδοχεία / Μοτέλ / Bungalows Δ+Ε	1.1	1.4	1.5	1.3	3.7	18.3	6.1	17.6	16.4	24.9	26.0	12.0
Ξενώνες	17.3	18.4	19.0	20.0	20.2	20.3	22.6	21.1	11.1	13.2	13.1	10.5
Οικοτροφεία	19.8	21.2	21.7	17.8	18.2	18.5	17.0	15.3	0.0	4.3	3.5	4.5
Επιπλωμένα Διαμερίσματα	6.3	7.5	8.8	8.7	8.9	8.4	8.9	9.3	9.2	6.8	10.6	7.1
Κέντρα Παραθεριστικά	3.4	4.1	4.8	4.9	5.6	5.9	8.9	5.8	18.0	10.8	11.9	13.6
Σύνολο	14.3	15.0	15.4	15.9	15.6	15.3	15.2	15.6	14.3	15.5	15.8	12.5

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

Πληρότητα Αλλοδαπών %	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Ξενοδοχεία / Μοτέλ / Bungalows AA	42.5	44.5	46.9	47.4	45.3	45.0	42.7	39.7	37.3	40.5	40.1	35.0
Ξενοδοχεία / Μοτέλ / Bungalows A	60.0	67.1	70.5	75.1	71.3	68.0	65.3	66.9	61.7	64.7	65.2	53.1
Ξενοδοχεία / Μοτέλ / Bungalows B	41.3	44.3	46.5	50.6	49.7	49.4	46.4	44.3	36.2	41.1	42.3	30.4
Ξενοδοχεία / Μοτέλ / Bungalows Γ	26.7	27.4	28.7	32.2	33.0	32.3	30.6	28.3	24.2	28.6	28.7	20.0
Ξενοδοχεία / Μοτέλ / Bungalows Δ+Ε	81.6	74.7	76.1	72.2	69.9	19.3	61.8	16.7	18.6	21.3	17.0	44.4
Ξενώνες	24.3	21.8	24.9	27.5	27.6	32.0	29.8	26.0	18.7	23.1	23.2	15.2
Οικοτροφεία	11.8	8.8	12.4	10.2	15.4	16.4	12.8	17.0	0.0	5.8	5.5	0.0
Επιπλωμένα Διαμερίσματα	55.9	59.3	62.9	67.2	61.1	99.4	56.6	55.2	53.5	58.0	60.6	33.2
Κέντρα Παραθεριστικά	42.1	49.0	53.8	38.0	38.8	40.1	36.6	33.0	13.2	21.4	19.8	21.6
Σύνολο	41.1	44.4	46.7	50.2	49.2	50.5	46.0	45.1	41.3	45.2	45.5	35.5

Πληρότητα (Σύνολο) %	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Ξενοδοχεία / Μοτέλ / Bungalows AA	53.5	56.2	57.7	58.1	56.2	56.1	53.4	51.2	47.2	52.3	52.1	45.7
Ξενοδοχεία / Μοτέλ / Bungalows A	67.9	75.4	79.0	85.1	80.5	76.8	74.5	76.8	72.5	74.4	76.0	62.3
Ξενοδοχεία / Μοτέλ / Bungalows B	57.5	61.5	64.0	68.6	67.8	67.2	64.1	62.8	54.0	57.6	61.6	46.7
Ξενοδοχεία / Μοτέλ / Bungalows Γ	47.6	48.9	51.1	54.4	55.6	54.1	52.1	49.8	43.2	50.5	50.6	38.9
Ξενοδοχεία / Μοτέλ / Bungalows Δ+Ε	82.8	76.1	77.6	73.6	73.6	37.6	67.9	34.3	34.9	46.2	43.0	56.4
Ξενώνες	41.6	40.2	43.9	47.5	47.9	52.3	52.5	47.1	29.7	36.3	36.3	25.7
Οικοτροφεία	31.6	30.0	34.1	28.0	33.6	34.8	29.7	32.3	0.0	10.1	9.0	4.5
Επιπλωμένα Διαμερίσματα	62.2	66.8	71.6	75.9	70.0	107.8	65.5	64.5	62.6	64.8	71.2	40.3
Κέντρα Παραθεριστικά	45.4	53.1	58.6	42.9	44.4	46.0	45.5	38.8	31.2	32.3	31.7	35.2
Σύνολο	55.3	59.4	62.2	66.1	64.8	65.8	61.2	60.7	55.6	60.7	61.3	48.0

πηγή: ΕΣΥΕ

Εκ14 Καταγεγραμμένοι επιβάτες σε αεροδρόμια

Αφίξεις περιηγητών

Σύνολο Επιβατών, Αερολιμένες, Σύνολο Χώρας

	Ιανουάριος	Φεβρουάριος	Μάρτιος	Απρίλιος	Μάιος	Ιούνιος	Ιούλιος	Αύγουστος	Σεπτέμβριος	Οκτώβριος	Νοέμβριος	Δεκέμβριος	Σύνολο
1996	208,801	179,042	301,734	545,048	1,054,677	1,244,330	1,492,465	1,475,013	1,291,128	681,962	190,037	245,459	8,909,696
1997	221,962	186,677	330,873	513,211	1,191,925	1,351,693	1,639,168	1,664,361	1,340,881	740,188	214,553	258,652	9,654,144
1998	221,768	188,481	266,871	589,723	1,233,027	1,409,250	1,751,271	1,780,443	1,389,242	784,564	222,991	259,213	10,096,844
1999	237,168	209,817	331,908	603,890	1,372,011	1,590,320	1,924,108	1,942,240	1,610,636	899,794	242,859	274,751	11,239,502
2000	246,695	220,450	308,237	718,592	1,446,720	1,720,571	2,042,433	1,980,551	1,730,301	936,555	251,226	290,991	11,893,322
2001	233,034	263,161	328,726	725,462	1,516,157	1,839,972	2,178,341	2,104,206	1,797,147	877,748	220,378	280,099	12,364,431
2002	243,834	237,058	388,796	547,112	1,473,855	1,744,472	2,056,392	2,076,850	1,821,270	919,830	263,084	311,675	12,084,228
2003	280,685	249,878	309,784	600,802	1,367,357	1,740,854	2,013,452	2,105,533	1,700,968	928,720	289,558	331,235	11,918,826
2004	298,622	267,267	375,436	650,459	1,408,385	1,612,577	1,989,075	2,058,534	1,712,130	942,290	312,240	343,027	11,970,042
2005	346,129	280,971	414,684	579,007	1,452,936	1,736,629	2,180,093	2,118,079	1,784,946	953,666	303,698	336,548	12,487,386
2006	333,670	270,298	375,502	691,551	1,488,857	1,892,423	2,355,823	2,284,160	1,883,750	980,033	331,794	388,691	13,276,552
2007	318,520	283,116	413,906	750,162	1,664,020	1,983,074	2,397,754	2,394,134	2,002,958	1,069,583	315,199	368,196	13,960,621

πηγή: ΥΠΑ

Ες15 Επισκέψεις Περιηγητών

Εθνική Κλίμακα

Σύνολο Ετών

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Αφιξεις / Ημεδαποί	4,711,928	5,395,258	5,553,448	5,666,787	5,873,621	5,506,155	5,313,148	5,706,989	5,670,521	6,063,890	6,127,573	6,949,608
Αφιξεις / Αλλοδαποί	5,954,992	6,764,649	7,253,471	7,180,793	7,730,636	6,907,166	6,422,457	6,516,993	5,661,062	7,348,706	7,547,667	8,745,901
Αφιξεις / Σύνολο	10,666,920	12,159,907	12,806,919	12,847,580	13,604,257	12,413,321	11,735,605	12,223,982	11,331,583	13,412,596	12,971,148	15,695,509
Διανυκτερεύσεις / Ημεδαποί	12,228,146	13,668,067	13,981,577	14,509,110	14,748,416	13,249,281	12,834,774	13,881,203	13,757,825	14,529,739	14,248,961	16,675,264
Διανυκτερεύσεις / Αλλοδαποί	35,051,483	40,162,916	42,102,959	45,361,949	46,091,509	43,337,173	38,439,563	39,594,500	38,796,250	40,734,354	40,734,354	47,410,260
Διανυκτερεύσεις / Σύνολο	47,279,629	53,830,983	56,084,536	59,871,059	60,839,925	56,586,454	51,274,337	53,475,703	52,554,075	55,264,093	57,293,613	64,085,524
Αφιξεις / Ημεδαποί (%)	44.17%	44.37%	43.36%	44.11%	43.17%	44.36%	45.27%	46.69%	50.04%	45.21%	47.24%	44.28%
Αφιξεις / Αλλοδαποί (%)	55.83%	55.63%	56.64%	55.89%	56.83%	55.64%	54.73%	53.31%	49.96%	54.79%	58.19%	55.72%
Διανυκτερεύσεις / Ημεδαποί (%)	25.86%	25.39%	24.93%	24.23%	24.24%	23.41%	25.03%	25.96%	26.18%	26.29%	24.87%	26.02%
Διανυκτερεύσεις / Αλλοδαποί (%)	74.14%	74.61%	75.07%	75.77%	75.76%	76.59%	74.97%	74.04%	73.82%	73.71%	71.10%	73.98%
Μέσος Αριθμός Διανυκτερεύσεων Ημεδαπών	2.60	2.53	2.52	2.56	2.51	2.41	2.42	2.43	2.43	2.40	2.33	2.40
Μέσος Αριθμός Διανυκτερεύσεων Αλλοδαπών	5.89	5.94	5.80	6.32	5.96	6.27	5.99	6.08	6.85	5.54	5.40	5.42
Μέσος Αριθμός Διανυκτερεύσεων (Σύνολο)	4.43	4.43	4.38	4.66	4.47	4.56	4.37	4.37	4.64	4.12	4.42	4.08

Διανυκτε-ρεύσεις σε καταλύματα

Σύνολο	Ημεδαποί												
	Ιανουάριος	Φεβρουάριος	Μάρτιος	Απρίλιος	Μάιος	Ιούνιος	Ιούλιος	Αύγουστος	Σεπτέμβριος	Οκτώβριος	Νοέμβριος	Δεκέμβριος	Σύνολο
1996	852,688	932,758	1,300,524	2,782,656	5,144,323	6,256,490	7,882,488	9,487,346	6,963,579	4,121,726	1,073,997	481,054	47,279,629
1997	989,151	938,898	1,370,091	2,826,459	5,886,668	7,149,770	9,137,928	11,038,453	7,813,593	4,468,824	1,193,947	1,017,201	53,830,983
1998	1,015,624	1,005,725	1,271,714	3,023,732	6,247,183	7,738,793	9,162,690	11,579,879	8,197,915	4,584,436	1,223,215	1,033,630	56,084,536
1999	1,094,117	1,042,120	1,451,140	2,938,599	6,595,278	8,275,498	10,512,126	11,683,589	8,725,745	5,161,938	1,263,260	1,127,649	59,871,059
2000	1,081,858	1,139,885	1,477,461	3,108,216	6,563,707	8,350,424	10,669,658	11,837,470	8,865,248	5,255,123	1,283,617	1,207,258	60,839,925
2001	924,185	1,086,483	1,426,857	2,820,006	5,887,121	7,734,921	9,662,587	10,341,472	10,231,010	4,446,007	1,114,427	911,378	56,586,454
2002	893,405	980,230	1,465,136	2,389,743	5,491,022	7,232,265	8,173,261	10,328,022	7,834,285	4,338,393	1,112,245	1,036,330	51,274,337
2003	983,542	984,008	1,320,018	2,676,849	5,384,473	7,485,213	9,691,287	10,728,279	7,561,884	4,388,192	1,147,075	1,124,883	53,475,703
2004	828,899	943,205	1,337,085	2,422,046	5,418,574	7,235,729	9,642,797	10,890,700	7,717,705	4,034,601	1,100,611	982,123	52,554,075
2005	877,933	907,821	1,384,753	2,140,027	5,605,008	7,651,903	10,403,935	11,892,227	8,228,185	4,169,619	1,076,467	926,215	55,264,093
2006	960,853	1,026,543	1,449,808	2,600,999	5,927,931	7,967,122	10,249,992	11,623,673	8,656,661	4,556,450	1,177,172	1,096,409	57,293,613
2007	1,022,345	1,086,088	1,520,075	2,648,734	6,509,068	9,403,558	12,222,221	13,349,381	9,312,206	4,320,077	1,410,496	1,281,275	64,085,524

Διανυκτε-ρεύσεις σε καταλύματα

Σύνολο	Αλλοδαποί												
	Ιανουάριος	Φεβρουάριος	Μάρτιος	Απρίλιος	Μάιος	Ιούνιος	Ιούλιος	Αύγουστος	Σεπτέμβριος	Οκτώβριος	Νοέμβριος	Δεκέμβριος	Σύνολο
1996	224,496	263,306	537,339	1,805,448	4,323,131	5,245,616	6,106,176	7,024,724	5,838,956	3,244,530	367,883	69,878	35,051,483
1997	268,175	287,799	599,078	1,697,427	4,896,165	6,049,742	7,216,555	8,358,040	6,543,253	3,562,585	405,452	278,645	40,162,916
1998	253,620	282,124	505,692	1,860,796	5,182,679	6,568,518	7,176,847	8,951,400	6,839,526	3,762,148	447,634	271,975	42,102,959
1999	305,584	321,929	566,127	1,885,951	5,454,018	7,185,775	8,230,052	9,009,191	7,439,293	4,199,993	462,139	301,897	45,361,949
2000	270,107	330,544	539,917	1,974,282	5,545,404	7,096,766	8,521,815	9,133,071	7,483,953	4,327,034	480,347	388,269	46,091,509
2001	235,349	307,696	568,787	1,782,907	4,964,945	6,546,613	7,796,203	7,978,680	8,957,754	3,582,972	379,912	235,355	43,337,173
2002	220,207	281,618	543,324	1,346,191	4,504,665	6,136,813	6,436,449	7,887,747	6,691,333	3,584,958	411,387	394,871	38,439,563
2003	255,799	308,323	417,051	1,514,978	4,409,781	6,222,496	7,740,389	8,084,239	6,307,362	3,527,678	385,590	420,814	39,594,500
2004	222,409	278,915	431,036	1,451,875	4,392,827	5,979,185	7,439,747	8,313,470	6,438,611	3,134,760	389,286	324,129	38,796,250
2005	237,307	297,704	535,348	1,090,366	4,593,278	6,372,737	8,060,735	8,898,653	6,857,731	3,235,578	343,843	211,074	40,734,354
2006	232,971	276,560	454,284	1,330,722	4,663,413	6,777,815	8,564,397	8,850,015	7,192,048	3,460,392	398,475	257,675	40,734,354
2007	264,646	307,953	530,471	1,450,252	5,362,623	7,873,250	9,621,594	10,091,271	7,825,977	3,302,682	467,960	311,581	47,410,260

Ec18 Δομή Οικονομικού Προϊόντος

Δομή Οικονομικού Προϊόντος

Γεωγραφική περιοχή : Ελλάδα

Μονάδες : Προστιθέμενη αξία-Εκατομμύρια Ευρώ (τρέχουσες τιμές)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
A - Γεωργία, θήρα, δασοκομία	7010	7270	7623	7542	7464	7630	8181	8478	8308	8625	7159	7251
B - Αλιεία	303	287	296	505	566	691	804	856	836	980	1024	917
C - Ορυχεία, λατομεία	533	552	596	529	711	748	873	856	945	1027	1048	1074
D - Μεταποίηση	10537	10378	11424	11959	13196	14143	14712	15948	18229	19391	21382	22080
E - Παροχή ηλεκτρικού ρεύματος, φυσικού αερίου και νερού	1725	1782	2000	2051	1989	2144	2296	2479	2581	3141	3254	3509
F - Κατασκευές	5218	5875	6872	7401	8260	9963	10390	12083	12772	13531	14943	15351
G - Χονδρικό & λιανικό εμπόριο: επισκευές οχημάτων και ειδών νοικοκυριών	11661	12811	13154	13480	14017	15747	16201	18054	18650	18455	20528	21643
H - Ξενοδοχεία & εστιατόρια	5632	6904	7437	7186	8181	8845	9599	11123	11813	12827	13724	15072
I - Μεταφορές, αποθηκεύσεις, επικοινωνίες	4984	5689	6487	8683	9716	9975	10936	11830	15828	17064	16618	19371
J - Ενδιάμεσοι χρηματοπιστωτικοί οργανισμοί	3691	4022	4814	5496	6210	6771	6543	7192	8033	8764	8845	9921
K - Διαχείριση ακίνητης περιουσίας, εκμισθώσεις και επιχειρηματικές δραστηριότητες	13991	15475	16715	16299	17467	18577	20032	21597	21878	22825	25116	26650
L - Δημόσια διοίκηση και άμυνα, υποχρεωτική κοινωνική ασφάλιση	5428	6168	6660	7194	7693	8312	9404	9998	11758	12244	12813	13913
M - Εκπαίδευση	3493	4280	4602	4756	5127	5348	6272	6608	6892	7806	7735	8758
N - Υγεία και κοινωνική μέριμνα	4114	4680	5251	5473	5812	6256	6874	7560	8278	9700	9599	10663
O - Παραγωγή άλλων υπηρεσιών υπέρ του κοινωνικού συνόλου	2139	2261	2511	2827	2959	3564	3978	4013	4687	5074	6786	6406
P - Ίδιωτικά νοικοκυριά που απασχολούν οικιακό προσωπικό	213	266	309	441	496	540	587	675	757	833	925	1020
Σύνολο	80672	88700	96751	101822	109864	119254	127682	139350	152245	162287	171499	183599

Πηγή : Εθνική Στατιστική Υπηρεσία

Σημείωση: Τα στοιχεία του έτους 2006 αποτελούν εκτίμηση.

Ηλεκτ. Διεύθυνση : http://www.statistics.gr/table_menu.asp?dt=0&sb=SEL_5&SSnid=Ακαθάριστη%20Προστιθέμενη%20Αξία%20Κατά%20Κλάδο

Μονάδες: Ποσοστό επί του συνόλου

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
A	0.09	0.08	0.08	0.07	0.07	0.06	0.06	0.06	0.05	0.05	0.04	0.04
B	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00
C	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
D	0.13	0.12	0.12	0.12	0.12	0.12	0.12	0.11	0.12	0.12	0.12	0.12
E	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
F	0.06	0.07	0.07	0.07	0.08	0.08	0.08	0.09	0.08	0.08	0.09	0.08
G	0.14	0.14	0.14	0.13	0.13	0.13	0.13	0.13	0.12	0.11	0.12	0.12
H	0.07	0.08	0.08	0.07	0.07	0.07	0.08	0.08	0.08	0.08	0.08	0.08
I	0.06	0.06	0.07	0.09	0.09	0.08	0.09	0.08	0.10	0.11	0.10	0.11
J	0.05	0.05	0.05	0.05	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.05
K	0.17	0.17	0.17	0.16	0.16	0.16	0.16	0.15	0.14	0.14	0.15	0.15
L	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.08	0.08	0.07	0.08
M	0.04	0.05	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.05
N	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.06	0.06
O	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.04	0.03
P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01
Σύνολο	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Ec19 Κατά Κεφαλήν ΑΕΠ

Γεωγραφική περιοχή : Ελλάδα

Μονάδες : Α.Ε.Π. σε Εκατομμύρια Ευρώ (σταθερές τιμές έτους 1996)

Πληθυσμός σε χιλιάδες

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Ακαθάριστο Εγχώριο Προϊόν(Α.Ε.Π.)	87,850.56	97,235.00	105,773.00	112,686.00	124,388.00	133,105.00	143,482.00	155,543.00	168,417.00	181,088.00	195,346.00	209,065.00
Πληθυσμός	10,709.15	10,776.53	10,834.91	10,882.61	10,917.46	10,949.95	10,987.56	11,006.38	11,064.27	11,082.75	11,148.53	11,192.85

Πηγή : Εθνική Στατιστική Υπηρεσία

Κατά Κεφαλήν Α.Ε.Π.

Κατά κεφαλή Α.Ε.Π. (Ελλάδα)	8,203.32	9,022.85	9,762.24	10,354.69	10,321.63	12,155.76	13,058.59	14,132.08	15,221.70	16,339.63	17,522.13	18,678.44
-----------------------------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------

Ec20 Αυτοδυναμία Οικονομίας

Γεωγραφική περιοχή : Ελλάδα

Μονάδες : Εκατομμύρια Ευρώ

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Ακαθάριστο Εγχώριο Προϊόν (Α.Ε.Π.)	87,850.56	97,235.00	105,773.00	112,686.00	124,388.00	133,105.00	143,482.00	155,543.00	168,417.00	181,088.00	195,346.00	209,065.00
Ακαθάριστο Εθνικό Εισόδημα (Α.Ε.Ε.)	90,301.68	99,959.00	108,639.00	114,185.00	125,353.00	132,174.00	142,459.00	153,888.00	167,129.00	177,750.00	190,742.00	202,523.00

(Α.Ε.Π.) / (Α.Ε.Ε.)	0.973	0.973	0.974	0.987	0.992	1.007	1.007	1.011	1.007	1.019	1.007	1.032
---------------------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Πηγή : ΕΣΥΕ

Ec21 Αποδοχές

Γεωγραφική περιοχή : Ελλάδα

Μονάδες : Ευρώ (τρέχουσες τιμές)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Μέσες ετήσιες αποδοχές ανδρών	6552.00	7146.00	7847.00	7690.00	8265.00	8416.00	8904.67	9839.00	10359.00	11477.00	12406.71	13338.26
Μέσες ετήσιες αποδοχές γυναικών	6411.00	6991.00	7714.00	7517.00	7981.00	8118.00	8586.13	9494.00	10019.00	11142.00	11963.84	12848.14
Μέσο εισόδημα χαμηλόμισθων ανδρών	5545.00	5973.00	6529.00	6442.00	7013.00	7227.00	7599.13	8333.00	9000.00	9716.00	10649.38	11470.48
Μέσο εισόδημα χαμηλόμισθων γυναικών	5370.00	5798.00	6299.00	6275.00	6850.00	6974.00	7376.20	8073.00	8650.00	9333.00	10240.06	11018.66

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

αναλογία μέσων ετήσιων αποδοχών ανδρών/ γυναικών	1.02	1.02	1.02	1.02	1.04	1.04	1.04	1.04	1.03	1.03	1.04	1.04
αναλογία μέσου εισοδήματος χαμηλόμισθων ανδρών/ γυναικών	1.03	1.03	1.04	1.03	1.02	1.04	1.03	1.03	1.04	1.04	1.04	1.04

Πηγή : Eurostat, ΕΣΥΕ

Χαμηλόμισθοι: Όσοι έχουν εισόδημα μικρότερο ή ίσο του μέσου εισοδήματος

Κατώτατο ημερομίσθιο εργατών	18.49	19.73	19.45	20.44	20.62	21.05	22.17	23.29	24.5	25.86	27.46	28.93
Κατώτατος μισθός υπαλλήλων	412.92	440.58	434.13	456.24	460.08	469.65	494.77	519.87	547.56	579.04	614.62	647.63
αναλογία	0.940	0.941	0.941	0.941	0.941	0.941	0.941	0.941	0.940	0.938	0.938	0.938

Πηγή : Υπουργείο Εθνικής Οικονομίας και Οικονομικών, Τράπεζα της Ελλάδος

Ec22 Ποσοστό Φτώχειας

Γεωγραφική περιοχή : Ελλάδα

Μονάδες : % ποσοστό σε δείγμα 3000 οικογενειών

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Ποσοστό πληθυσμού υπό του ορίου φτώχειας	21%	22%	22%	21%	20%	20%	21%	20%	20%	21%	22%	21%
Ποσοστό πληθυσμού άνω του ορίου φτώχειας	79%	78%	78%	79%	80%	80%	79%	80%	80%	79%	78%	79%

Πηγή : Εθνική Στατιστική Υπηρεσία (Τομέας κοινωνικών στατιστικών)

(Το όριο φτώχειας είναι το 60% του μέσου διαθέσιμου εισοδήματος)

Ec23 Δείκτης Εισοδηματικής Κατανομής (Gini Index)

Γεωγραφική περιοχή : Ελλάδα

Μονάδες : % ποσοστό

Δείκτης	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
GINI INDEX	0.34	0.35	0.35	0.34	0.33	0.33	0.33	0.35	0.33	0.33	0.33	0.34

Πηγή : Eurostat

Ec24 Δείκτης Τιμών Καταναλωτή

Γεωγραφική περιοχή : Ελλάδα

Μονάδες: Τιμή έτους βάσης = 100,00

Δείκτες	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Δείκτης Τιμών Καταναλωτή	100.00	105.54	110.57	113.48	117.06	121.01	125.40	129.83	133.59	138.32	142.75	147.14
Πληθωρισμός	8.20%	5.54%	4.77%	2.64%	3.15%	3.37%	3.63%	3.53%	2.90%	3.54%	3.20%	3.08%

Πηγή : Εθνική Στατιστική Υπηρεσία

Ec25 Δείκτες τιμών

Τηλεπικοινωνίες

Γεωγραφική περιοχή: Ελλάδα

Μονάδες: Ευρώ ανά 10' κλήσης (τρέχουσες τιμές)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Τοπικές κλήσεις	0.16	0.16	0.18	0.21	0.31	0.36	0.31	0.31	0.31	0.31	0.31	0.30
Υπεραστικές κλήσεις	4.18	3.74	3.15	2.78	1.40	0.98	0.77	0.77	0.73	0.74	0.72	0.70

Ηλεκτρική ενέργεια

Γεωγραφική περιοχή: Ελλάδα

Μονάδες: Ευρώ ανά KWh (τρέχουσες τιμές)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Οικιακοί χρήστες	0.0609	0.0619	0.0627	0.0622	0.0564	0.0564	0.0580	0.0606	0.0621	0.0637	0.0643	0.0650
Βιομηχανικοί χρήστες	0.0571	0.0580	0.0588	0.0583	0.0571	0.0571	0.0590	0.0614	0.0630	0.0645	0.0668	0.0680

Πηγή: Eurostat

Ec26 Εισαγωγές-Εξαγωγές Προϊόντων

Γεωγραφική περιοχή: Ελλάδα

Μονάδες: Εκατομμύρια Ευρώ (τρέχουσες τιμές)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Εισαγωγές προϊόντων	20485.00	22438.00	25423.00	28103.00	35748.00	34142.00	35527.00	38754.00	42641.00	43846.00	49897.00	54837.00
Εξαγωγές προϊόντων	8854.00	9321.00	9800.00	10528.00	12973.00	12836.00	12160.00	12712.00	12749.00	14505.00	16608.00	17522.00
Λόγος εισαγωγών προς εξαγωγές	2.31	2.41	2.59	2.67	2.76	2.66	2.92	3.05	3.34	3.02	3.00	3.13

Πηγή: ΕΣΥΕ

Ec27 Δυναμική παροχής Υπηρεσιών

Γεωγραφική περιοχή: Ελλάδα

Μονάδες: Εκατομμύρια Ευρώ (τρέχουσες τιμές)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Εξαγωγές υπηρεσιών	3999.67	5747.78	9968.30	15575.50	20977.00	22075.90	21131.40	21430.40	26742.50	27560.50	28364.10	31337.30
Εισαγωγές υπηρεσιών	1962.01	2266.27	3895.70	8729.00	12266.00	12925.90	10376.10	9923.90	11275.50	11862.40	13027.00	14745.60

Λόγος εξαγωγών προς εισαγωγές υπηρεσιών	2.04	2.54	2.56	1.78	1.71	1.71	2.04	2.16	2.37	2.32	2.18	2.13
---	------	------	------	------	------	------	------	------	------	------	------	------

Πηγή: Τράπεζα της Ελλάδος

Ec28 Ισοτιμίες ευρώ - κυριότερων νομισμάτων

Μονάδες: Νομισματική ισοτιμία

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Ισοτιμία Ευρώ / Δραχμής	305.55	309.36	330.73	325.76	336.63	340.75	340.75	340.75	340.75	340.75	340.75	340.75
Ισοτιμία Ευρώ / Δολαρίου	1.27	1.13	1.12	1.07	0.92	0.90	0.95	1.13	1.24	1.24	1.19	1.22
Ισοτιμία Ευρώ / Γεν	138.08	137.08	146.42	121.32	99.47	108.68	118.06	130.97	134.44	136.85	139.11	128.64

Πηγή: Eurostat

Ec29 Επενδυτικός Κίνδυνος

Χώρα

δείκτης	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	μονάδες	source
Μακροπρόθεσμα επιτόκια: Απόδοση δεκαετούς κρατικού ομολόγου (δευτερογενής αγορά)	14.62	10.17	8.47	6.30	6.10	5.30	5.12	4.27	4.25	3.58	4.07	4.29	Ετήσια μεταβολή (%)	Eurostat

Το επιτόκιο ορίζεται ως το κόστος αγοράς ή δανεισμού ή το κέρδος από δανεισμό και εκφράζεται συνήθως ως ετήσιο ποσοστό. Η απόδοση των δεκαετών κρατικών ομολόγων χρησιμοποιείται συχνά ως μέτρο των μακροπρόθεσμων επιτοκίων. Οι αποδόσεις κυμαίνονται ανάλογα με την πμή του ομολόγου. Η δευτερογενής αγορά υποδηλώνει ότι ως πμή δεν ορίζεται η σταθερή πμή έκδοσης (πρωτογενής αγορά), αλλά αυτή όπως διαμορφώνεται από την προσφορά και τη ζήτηση της αγοράς.

Ec30 Οικονομικός Ρόλος του Κράτους

Γεωγραφική περιοχή : Ελλάδα

Μονάδες : Εκατομμύρια Ευρώ (σταθερές τιμές έτους 1996)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Έξοδα	35,419.41	35,448.41	33,838.62	36,032.57	36,033.75	35,767.62	35,878.83	38,520.81	51,800.00	55,627.00	50,413.00	53,360.00
Έσοδα	26,023.89	27,711.64	27,390.92	30,166.59	31,097.44	31,267.57	31,624.54	31,568.89	45,540.00	47,110.00	46,400.00	49,170.00
Ακαθάριστο Εγχώριο Προϊόν (Α.Ε.Π.)	87,850.56	91,045.73	94,107.59	97,326.57	101,683.92	106,579.53	110,446.19	115,412.69	164,421.00	176,471.00	195,346.00	209,065.00
Έλλειμμα	9,395.52	7,736.76	6,447.70	5,865.98	4,936.31	4,500.05	4,254.29	6,951.92	6,260.00	8,517.00	4,013.00	4,190.00

Δείκτης	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Έξοδα (%Α.Ε.Π.)	40.32%	38.93%	35.96%	37.02%	35.44%	33.56%	32.49%	33.38%	31.50%	31.52%	25.81%	25.52%
Έσοδα (%Α.Ε.Π.)	29.62%	30.44%	29.11%	31.00%	30.58%	29.34%	28.63%	27.35%	27.70%	26.70%	23.75%	23.52%

Πηγή: Κρατικός Προϋπολογισμός 2005 (σελ. 59)

Ec31 Δομή Δημοσιοοικονομικών Εξόδων

Γεωγραφική περιοχή : Ελλάδα

Μονάδες : Εκατομμύρια Ευρώ (τρέχουσες τιμές)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Δαπάνες προσωπικού	9,933.56	11,355.42	11,610.68	12,624.94	13,275.70	13,866.00	15,353.00	16,526.00	18,425.00	19,325.00	20,549.00	21,909.00
Τόκοι	11,458.35	10,396.64	9,777.46	10,140.59	10,034.76	9,711.00	9,134.00	9,416.00	9,464.00	9,774.00	9,589.00	9,796.00
Λοιπές δαπάνες	10,443.79	10,781.64	10,846.61	11,850.14	12,435.02	13,159.00	14,911.00	17,174.00	20,399.00	22,140.00	22,370.00	26,652.00
Δαπάνες δημόσιων επενδύσεων	3,583.70	5,257.80	5,675.32	6,906.93	7,512.70	7,842.00	7,014.00	8,434.00	9,522.00	7,524.00	8,184.00	8,809.00
Σύνολο	35,419.41	37,791.50	37,910.08	41,522.59	43,258.18	44,578.00	46,412.00	51,550.00	57,810.00	58,763.00	60,692.00	67,166.00

Μονάδες : Ποσοστά %

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Δαπάνες προσωπικού	28.05	30.05	30.63	30.40	30.69	31.11	33.08	32.06	31.87	32.89	33.86	32.62
Τόκοι	32.35	27.51	25.79	24.42	23.20	21.78	19.68	18.27	16.37	16.63	15.80	14.58
Λοιπές δαπάνες	29.49	28.53	28.61	28.54	28.75	29.52	32.13	33.32	35.29	37.68	36.86	39.68
Δαπάνες δημόσιων επενδύσεων	10.12	13.91	14.97	16.63	17.37	17.59	15.11	16.36	16.47	12.80	13.48	13.12

Πηγή : Τράπεζα της Ελλάδος

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

Δομή Δημοσιοοικονομικών Εξόδων / Πιστώσεις ανά Φορέα

Γεωγραφική περιοχή : Ελλάδα

Μονάδες : Εκατομμύρια Ευρώ (τρέχουσες τιμές)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
ΠΡΟΕΔΡΙΑ ΤΗΣ ΔΗΜΟΚΡΑΤΙΑΣ	1.79	2.22	2.42	2.48	2.68	2.62	3.01	2.90	3.00	4.00	5.00	4.84
ΒΟΥΛΗ ΤΩΝ ΕΛΛΗΝΩΝ	63.97	74.30	80.25	102.87	98.39	104.67	111.54	120.07	133.00	160.00	176.00	193.10
ΥΠ. ΕΣΩΤΕΡΙΚΩΝ	220.47	273.26	283.95	295.03	326.37	359.99	436.30	437.54	413.00	431.00	418.00	409.13
ΥΠ. ΕΣΩΤΕΡΙΚΩΝ & ΔΗΜ. ΔΙΟΙΚΗΣΗΣ	917.30	952.57	1,145.35	1,675.26	1,665.65	1,791.60	2,205.13	2,586.54	2,910.00	3,288.00	3,358.00	3,507.32
ΥΠ. ΜΑΚΕΔΟΝΙΑΣ-ΘΡΑΚΗΣ	20.19	25.49	18.42	13.94	12.80	13.87	16.07	19.30	17.00	48.00	18.00	18.19
ΥΠ. ΑΙΓΑΙΟΥ	6.45	13.18	13.18	13.17	13.43	14.82	22.03	24.10	43.00	51.00	30.00	32.98
ΥΠ. ΔΗΜΟΣΙΑΣ ΤΑΞΗΣ	836.37	951.05	996.11	1,114.14	1,293.86	1,395.97	1,484.87	1,579.14	1,792.00	1,872.00	1,915.00	2,103.27
ΥΠ. ΕΘΝΙΚΗΣ ΑΜΥΝΑΣ	2,760.40	3,194.36	3,372.82	3,608.94	3,756.83	4,169.87	3,882.78	4,198.98	3,570.00	3,601.00	3,687.00	6,057.10
ΥΠ. ΠΑΙΔΕΙΑΣ	2,540.07	3,145.24	3,153.47	3,442.14	3,668.65	3,889.70	4,230.48	4,521.95	5,633.00	6,308.00	6,102.00	6,180.38
ΥΠ. ΠΟΛΙΤΙΣΜΟΥ	121.00	158.52	142.84	164.96	177.20	223.62	267.60	276.96	266.00	281.00	316.00	376.54
ΥΠ. ΥΓΕΙΑΣ	1,075.09	846.54	782.82	2,178.18	2,186.00	2,226.91	2,568.86	3,634.66	4,442.00	4,902.00	5,179.00	4,505.16
ΥΠ. ΔΙΚΑΙΟΣΥΝΗΣ	291.74	354.29	341.75	364.93	375.54	395.42	433.79	444.76	534.00	542.00	517.00	526.73
ΥΠ. ΜΕΤΑΦΟΡΩΝ	335.79	327.21	301.77	364.30	332.06	320.89	354.31	361.69	292.00	331.00	340.00	444.80
ΥΠ. ΤΥΠΟΥ	70.54	46.84	53.85	60.97	55.34	56.66	62.66	71.52	78.00	88.00	94.00	111.85
Υ.Π.Ε.Χ.Ω.Δ.Ε	363.99	177.69	147.32	157.80	164.51	175.78	110.79	133.07	170.00	191.00	200.00	155.93
ΥΠ. ΟΙΚΟΝΟΜΙΑΣ & ΟΙΚΟΝΟΜΙΚΩΝ	47,998.09	48,816.45	40,424.54	38,064.45	35,016.32	30,288.51	34,482.37	41,437.77	45,464.00	47,031.00	42,709.00	69,955.39
ΥΠ. ΑΝΑΠΤΥΞΗΣ	128.41	218.11	249.70	274.00	285.23	301.68	355.72	305.17	543.00	500.00	183.00	286.10
ΥΠ. ΕΜΠΟΡΙΚΗΣ ΝΑΥΤΙΛΙΑΣ	402.64	400.95	475.80	595.84	608.39	614.37	664.12	789.48	821.00	1,065.00	1,112.00	1,287.62
ΥΠ. ΓΕΩΡΓΙΑΣ	877.13	768.70	549.58	591.76	618.73	703.66	782.66	710.56	959.00	1,032.00	1,212.00	1,125.98
ΥΠ. ΕΡΓΑΣΙΑΣ	2,976.38	3,264.22	3,117.61	3,276.31	3,680.56	3,965.27	4,196.54	5,137.90	6,565.00	7,261.00	7,476.00	8,272.10
Σύνολο	62,007.79	64,011.19	55,653.57	56,361.47	54,338.56	51,015.9	56,671.6	66,794.1	74,648.0	78,987.0	75,047.0	105,554.5

Πηγή : Γενικό Λογιστήριο του Κράτους

Ec32 Δομή Δημοσιοοικονομικών Εσόδων

Γεωγραφική περιοχή: Ελλάδα

Μονάδες: Εκατομμύρια Ευρώ (τρέχουσες τιμές)

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Άμεσοι φόροι	7580.10	8944.88	10859.01	12412.82	13849.92	13584.69	14813.00	15397.00	16484.00	18371.00	18704.00	19832.00
Έμμεσοι φόροι	14072.66	15625.99	15869.14	17969.67	18848.29	19499.00	20989.00	21484.00	23000.00	23722.00	26287.00	28573.00
Μη φορολογικά έσοδα	2513.83	2800.71	1991.65	2135.31	1978.14	3490.00	3246.00	3000.00	2571.00	2667.00	3694.00	3372.00
Έσοδα δημοσίων επενδύσεων	1857.31	2321.00	2763.58	3100.44	3303.33	2795.00	2003.00	1823.00	2894.00	2686.00	3775.00	4875.00
Σύνολο	26023.89	29692.58	31483.39	35618.25	37979.68	39368.69	41051.00	41704.00	44949.00	47446.00	52460.00	56652.00

Μονάδες: Ποσοστό %

Δείκτες	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Άμεσοι φόροι	29.13%	30.12%	34.49%	34.85%	36.47%	34.51%	36.08%	36.92%	36.67%	38.72%	35.65%	35.01%
Έμμεσοι φόροι	54.08%	52.63%	50.40%	50.45%	49.63%	49.53%	51.13%	51.52%	51.17%	50.00%	50.11%	50.44%
Μη φορολογικά έσοδα	9.66%	9.43%	6.33%	6.00%	5.21%	8.86%	7.91%	7.19%	5.72%	5.62%	7.04%	5.95%
Έσοδα δημοσίων επενδύσεων	7.14%	7.82%	8.78%	8.70%	8.70%	7.10%	4.88%	4.37%	6.44%	5.66%	7.20%	8.61%

Πηγή : Γενικό Λογιστήριο του Κράτους

Ec33 Κατά Κεφαλήν Ακαθάριστο Χρέος

Γεωγραφική περιοχή : Ελλάδα

Μονάδες: Εκατομμύρια ευρώ

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Συνολικό ακαθάριστο χρέος του ευρύτερου δημόσιου τομέα, τρέχουσες	113544.72	122128.21	129952.85	137687.53	163679.26	151869.00	158887.00	167723.00	182702.00	194666.00	202770.00	216381.00
Μόνιμοι κάτοικοι	10673.70	10744.65	10808.36	10861.40	10903.76	10931.21	10968.71	11006.38	11040.65	11082.75	11125.18	11185.48
Κατά κεφαλήν ακαθάριστο χρέος του ευρύτερου δημόσιου τομέα, τρέχουσες	10637.81	11366.42	12023.37	12676.77	15011.27	13893.16	14485.48	15238.71	16548.12	17564.77	18226.22	19344.81

Πηγή : Τράπεζα της Ελλάδος & Εθνική Στατιστική Υπηρεσία

Ec34 Βαρύτητα Δημόσιου Χρέους

Γεωγραφική περιοχή : Ελλάδα

Μονάδες: Εκατομμύρια ευρώ, τρέχουσες τιμές

Δεδομένα	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Συνολικό ακαθάριστο χρέος του ευρύτερου δημόσιου τομέα, τρέχουσες τιμές	113545	122128	129953	137688	163679	151869	158887	168025	183157	195338	204394	216381
Ακαθάριστο Εγχώριο Προϊόν (Α.Ε.Π.)	87851	97235	105773	112686	124388	133105	144892	156302	169224	181855	194318	208593
Ακαθάριστο χρέος ως ποσοστό του Α.Ε.Π (%)	129.2%	125.6%	122.9%	122.2%	131.6%	114.1%	109.7%	107.5%	108.2%	107.4%	105.2%	103.7%

Πηγή : Τράπεζα της Ελλάδος

10.2] Παράρτημα Β - Τυποποίηση σε αρχείο .arff, για εισαγωγή των δεδομένων στο λογισμικό WEKA.

```
@relation olympic_games

@attribute year numeric
@attribute A_-_Agriculture_hunting_and_forestry numeric
@attribute B_-_Fishing numeric
@attribute C_-_Mining_and_quarrying numeric
@attribute D_-_Manufacturing numeric
@attribute E_-_Electricity_gas_and_water_supply numeric
@attribute F_-_Construction numeric
@attribute G_-_
_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycles_and_personal
_and_household_goods numeric
@attribute H_-_Hotels_and_restaurants numeric
@attribute I_-_Transport_storage_and_communications numeric
@attribute J_-_Financial_intermediation numeric
@attribute K_-_Real_estate_renting_and_business_activities numeric
@attribute L_-_Public_administration_and_defence_compulsory_social_security
numeric
@attribute M_-_Education numeric
@attribute N_-_Health_and_social_work numeric
@attribute O_-_Other_community_social_and_personal_service_activities numeric
@attribute P_-_Private_households_with_employed_persons numeric
@attribute Sum_1 numeric
@attribute A_-_Agriculture_hunting_and_forestry_per_cent numeric
@attribute B_-_Fishing_per_cent numeric
@attribute C_-_Mining_and_quarrying_per_cent numeric
@attribute D_-_Manufacturing_per_cent numeric
@attribute E_-_Electricity_gas_and_water_supply_per_cent numeric
@attribute F_-_Construction_per_cent numeric
@attribute G_-_
_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycles_and_personal
_and_household_goods_per_cent numeric
@attribute H_-_Hotels_and_restaurants_per_cent numeric
@attribute I_-_Transport_storage_and_communications_per_cent numeric
@attribute J_-_Financial_intermediation_per_cent numeric
@attribute K_-_Real_estate_renting_and_business_activities_per_cent numeric
@attribute L_-_
_Public_administration_and_defence_compulsory_social_security_per_cent
numeric
@attribute M_-_Education_per_cent numeric
@attribute N_-_Health_and_social_work_per_cent numeric
@attribute O_-_
_Other_community_social_and_personal_service_activities_per_cent numeric
@attribute P_-_Private_households_with_employed_persons_per_cent numeric
@attribute Active_population_1 numeric
@attribute Permment_resident_population numeric
@attribute Active_women numeric
@attribute Unemployed numeric
@attribute Activity_rate numeric
@attribute Percentage_of_women_in_active_population numeric
@attribute Unemployment_rate numeric
@attribute Permanent_resident_population numeric
```

@attribute Employment_rate numeric
@attribute Active_women numeric
@attribute Unemployment_rate numeric
@attribute Active_population_2 numeric
@attribute Permanent_residents numeric
@attribute Unemployed numeric
@attribute Employment_rate numeric
@attribute Unemployment_rate numeric
@attribute Permanent_residents numeric
@attribute Employment_rate numeric
@attribute Unemployment_rate numeric
@attribute Total_energy_gross_inland_consumption numeric
@attribute Total_energy_consumption_from_fossil_fuels_sum numeric
@attribute Total_energy_consumption_from_fossil_fuels_solid numeric
@attribute
Total_energy_consumption_from_fossil_fuels_crude_oil_and_petroleum_products
numeric
@attribute Total_energy_consumption_from_nuclear_sources numeric
@attribute Total_energy_consumption_from_renewable_sources numeric
@attribute Population_in_thousands numeric
@attribute Fossil_fuels_per_cent numeric
@attribute fossil_fuels_solid_per_cent numeric
@attribute fossil_fuels_(crude_oil_and_petroleum_products_per_cent numeric
@attribute nuclear_per_cent numeric
@attribute renewable_per_cent numeric
@attribute Total_energy_gross_inland_consumption_per_capita numeric
@attribute Industry numeric
@attribute Households numeric
@attribute Services numeric
@attribute Transport numeric
@attribute Total_final_consumption numeric
@attribute Industry_per_cent numeric
@attribute Households_per_cent numeric
@attribute Services_per_cent numeric
@attribute Transport_per_cent numeric
@attribute Total_energy_gross_inland_consumption numeric
@attribute Total_energy_primary_Production numeric
@attribute Energy_imports numeric
@attribute Energy_exports numeric
@attribute Bunkers numeric
@attribute Stock_change numeric
@attribute Production_as_per_cent_of_consumption numeric
@attribute Imports_as_per_cent_of_consumption numeric
@attribute Imports_as_per_cent_of_exports numeric
@attribute Bunkers_as_per_cent_of_consumption numeric
@attribute Hotels_/Motels_/Bungalows_AA numeric
@attribute Hotels_/Motels_/Bungalows_A numeric
@attribute Hotels_/Motels_/Bungalows_B numeric
@attribute Hotels_/Motels_/Bungalows_C numeric
@attribute Hotels_/Motels_/Bungalows_D+E numeric
@attribute Guest_houses numeric
@attribute Boarding_Houses numeric
@attribute Furnished_apartments numeric
@attribute Vacation_centers numeric
@attribute Total numeric
@attribute Hotels_/Motels_/Bungalows_AA numeric

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

@attribute Hotels_/_Motels_/_Bungalows_A numeric
@attribute Hotels_/_Motels_/_Bungalows_B numeric
@attribute Hotels_/_Motels_/_Bungalows_C numeric
@attribute Hotels_/_Motels_/_Bungalows_D+E numeric
@attribute Guest_houses numeric
@attribute Boarding_Houses numeric
@attribute Furnished_apartments numeric
@attribute Vacation_centers numeric
@attribute Hotels_/_Motels_/_Bungalows_AA numeric
@attribute Hotels_/_Motels_/_Bungalows_A numeric
@attribute Hotels_/_Motels_/_Bungalows_B numeric
@attribute Hotels_/_Motels_/_Bungalows_C numeric
@attribute Hotels_/_Motels_/_Bungalows_D+E numeric
@attribute Guest_houses numeric
@attribute Boarding_Houses numeric
@attribute Furnished_apartments numeric
@attribute Vacation_centers numeric
@attribute Total numeric
@attribute Hotels_/_Motels_/_Bungalows_AA numeric
@attribute Hotels_/_Motels_/_Bungalows_A numeric
@attribute Hotels_/_Motels_/_Bungalows_B numeric
@attribute Hotels_/_Motels_/_Bungalows_C numeric
@attribute Hotels_/_Motels_/_Bungalows_D+E numeric
@attribute Guest_houses numeric
@attribute Boarding_Houses numeric
@attribute Furnished_apartments numeric
@attribute Vacation_centers numeric
@attribute Total numeric
@attribute Hotels_/_Motels_/_Bungalows_AA numeric
@attribute Hotels_/_Motels_/_Bungalows_A numeric
@attribute Hotels_/_Motels_/_Bungalows_B numeric
@attribute Hotels_/_Motels_/_Bungalows_C numeric
@attribute Hotels_/_Motels_/_Bungalows_D+E numeric
@attribute Guest_houses numeric
@attribute Boarding_Houses numeric
@attribute Furnished_apartments numeric
@attribute Vacation_centers numeric
@attribute Total numeric
@attribute Total numeric
@attribute Total numeric
@attribute Arrivals_/_Greeks numeric
@attribute Arrivals_/_Foreigners numeric
@attribute Arrivals_/_Total numeric
@attribute Night_Stays_/_Greeks numeric
@attribute Night_Stays_/_Foreigners numeric
@attribute Night_Stays_/_Total numeric
@attribute Arrivals_/_Greeks_per_cent numeric
@attribute Arrivals_/_Foreigners_per_cent numeric
@attribute Night_Stays_/_Greeks_per_cent numeric
@attribute Night_Stays_/_Foreigners_per_cent numeric
@attribute Average_of_Night_Stays_/_Greeks numeric
@attribute Average_of_Night_Stays_/_Foreigners numeric
@attribute Average_of_Night_Stays_/_Total numeric
@attribute A_-_Agriculture_hunting_and_forestry numeric
@attribute B_-_Fishing numeric
@attribute C_-_Mining_and_quarrying numeric

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

@attribute D_-_Manufacturing numeric
@attribute E_-_Electricity_gas_and_water_supply numeric
@attribute F_-_Construction numeric
@attribute G_-_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycles_and_personal_and_household_goods numeric
@attribute H_-_Hotels_and_restaurants numeric
@attribute I_-_Transport_storage_and_communications numeric
@attribute J_-_Financial_intermediation numeric
@attribute K_-_Real_estate_renting_and_business_activities numeric
@attribute L_-_Public_administration_and_defence_compulsory_social_security numeric
@attribute M_-_Education numeric
@attribute N_-_Health_and_social_work numeric
@attribute O_-_Other_community_social_and_personal_service_activities numeric
@attribute P_-_Private_households_with_employed_persons numeric
@attribute Total numeric
@attribute consumer_price_index numeric
@attribute A_-_Agriculture_hunting_and_forestry numeric
@attribute B_-_Fishing numeric
@attribute C_-_Mining_and_quarrying numeric
@attribute D_-_Manufacturing numeric
@attribute E_-_Electricity_gas_and_water_supply numeric
@attribute F_-_Construction numeric
@attribute G_-_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycles_and_personal_and_household_goods numeric
@attribute H_-_Hotels_and_restaurants numeric
@attribute I_-_Transport_storage_and_communications numeric
@attribute J_-_Financial_intermediation numeric
@attribute K_-_Real_estate_renting_and_business_activities numeric
@attribute L_-_Public_administration_and_defence_compulsory_social_security numeric
@attribute M_-_Education numeric
@attribute N_-_Health_and_social_work numeric
@attribute O_-_Other_community_social_and_personal_service_activities numeric
@attribute P_-_Private_households_with_employed_persons numeric
@attribute Total numeric
@attribute A_-_Agriculture_hunting_and_forestry numeric
@attribute B_-_Fishing numeric
@attribute C_-_Mining_and_quarrying numeric
@attribute D_-_Manufacturing numeric
@attribute E_-_Electricity_gas_and_water_supply numeric
@attribute F_-_Construction numeric
@attribute G_-_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycles_and_personal_and_household_goods numeric
@attribute H_-_Hotels_and_restaurants numeric
@attribute I_-_Transport_storage_and_communications numeric
@attribute J_-_Financial_intermediation numeric
@attribute K_-_Real_estate_renting_and_business_activities numeric
@attribute L_-_Public_administration_and_defence_compulsory_social_security numeric
@attribute M_-_Education numeric
@attribute N_-_Health_and_social_work numeric
@attribute O_-_Other_community_social_and_personal_service_activities numeric

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

@attribute P_-_Private_households_with_employed_persons numeric
@attribute Total numeric
@attribute A_-_Agriculture_hunting_and_forestry numeric
@attribute B_-_Fishing numeric
@attribute C_-_Mining_and_quarrying numeric
@attribute D_-_Manufacturing numeric
@attribute E_-_Electricity_gas_and_water_supply numeric
@attribute F_-_Construction numeric
@attribute G_-_
_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycles_and_personal
_and_household_goods numeric
@attribute H_-_Hotels_and_restaurants numeric
@attribute I_-_Transport_storage_and_communications numeric
@attribute J_-_Financial_intermediation numeric
@attribute K_-_Real_estate_renting_and_business_activities numeric
@attribute L_-_Public_administration_and_defence_compulsory_social_security
numeric
@attribute M_-_Education numeric
@attribute N_-_Health_and_social_work numeric
@attribute O_-_Other_community_social_and_personal_service_activities numeric
@attribute P_-_Private_households_with_employed_persons numeric
@attribute Total numeric
@attribute Gross_domestic_product_GDP numeric
@attribute Population numeric
@attribute Per_capita_gross_domestic_product numeric
@attribute consumer_price_index numeric
@attribute constant_prices numeric
@attribute Gross_domestic_product_GDP numeric
@attribute Population numeric
@attribute Per_capita_gross_domestic_product_Attica_region numeric
@attribute consumer_price_index numeric
@attribute constant_prices numeric
@attribute Per_capita_GDP_Greece numeric
@attribute Per_capita_GDP_Attica numeric
@attribute Gross_Domestic_Product_GDP numeric
@attribute Gross_National_Product_GNP_or_Gross_National_Income_GNI numeric
@attribute GDP/_GNP numeric
@attribute Mens_mean_annual_wages numeric
@attribute Womens_mean_annual_wages numeric
@attribute Mean_annual_wage_of_low_earning_men numeric
@attribute Mean_annual_wage_of_low_earning_women numeric
@attribute Mean_annual_wages_ratio_of_men_versus_women numeric
@attribute Mean_annual_wage_ration_of_low_earning_men_versus_women numeric
@attribute Minimum_daily_wage_of_workers numeric
@attribute Minimum_daily_wage_of_employees numeric
@attribute ratio numeric
@attribute constant_prices_1996 numeric
@attribute Mens_mean_annual_wages numeric
@attribute Womens_mean_annual_wages numeric
@attribute Mean_annual_wage_of_low_earning_men numeric
@attribute Mean_annual_wage_of_low_earning_women numeric
@attribute Minimum_daily_wage_of_workers numeric
@attribute Minimum_daily_wage_of_employees numeric
@attribute Low-wage_population numeric
@attribute Wage-earning_population numeric
@attribute Low-wage_population numeric

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

@attribute Wage-earning_population numeric
@attribute GINI_INDEX numeric
@attribute Annual_mean numeric
@attribute Inflation numeric
@attribute Local_calls numeric
@attribute International_calls numeric
@attribute Domestic_users numeric
@attribute Industrial_users numeric
@attribute Imports numeric
@attribute Exports numeric
@attribute Imports numeric
@attribute Exports numeric
@attribute Ratio_of_imports_to_exports numeric
@attribute Imports_of_products_relative_change_in_constant_1996_prices numeric
@attribute Exports_of_products_relative_change_in_constant_1996_prices numeric
@attribute Consumer_Price_Index numeric
@attribute Services_exported numeric
@attribute Services_imported numeric
@attribute Services_exported numeric
@attribute Services_imported numeric
@attribute Ratio_of_services_exports_to_imports numeric
@attribute Services_exported_relative_change_in_constant_1996_prices numeric
@attribute Services_imported_relative_change_in_constant_1996_prices numeric
@attribute Consumer_Price_Index numeric
@attribute Gross_fixed_capital_formation numeric
@attribute Exchange_rate_euro/_drachma numeric
@attribute Exchange_rate_euro/_US_dollar numeric
@attribute Exchange_rate_euro/_japanese_yen numeric
@attribute Exchange_rate_euro/_drachma numeric
@attribute Exchange_rate_euro/_US_dollar numeric
@attribute Exchange_rate_euro/_japanese_yen numeric
@attribute Long-Term_Interest_Rates_10-year_Government_Bond_Yields_Secondary_Markets numeric
@attribute Expenses numeric
@attribute Income numeric
@attribute Gross_Domestic_Product_G.D.P. numeric
@attribute Deficit numeric
@attribute Expenses_per_cent_G.D.P. numeric
@attribute Income_per_cent_G.D.P. numeric
@attribute Expenses_per_cent_G.D.P. numeric
@attribute Income_per_cent_G.D.P. numeric
@attribute Personnel_expenditure numeric
@attribute Interests numeric
@attribute Other_expenditure numeric
@attribute Expenditure_on_public_investments numeric
@attribute Total numeric
@attribute Personnel_expenditure numeric
@attribute Interests numeric
@attribute Other_expenditure numeric
@attribute Expenditure_on_public_investments numeric
@attribute Total numeric
@attribute Consumer_price_index numeric
@attribute Personnel_expenditure numeric
@attribute Interests numeric
@attribute Other_expenditure numeric
@attribute Expenditure_on_public_investments numeric

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

@attribute Total numeric
@attribute Presidency_of_the_Hellenic_republic numeric
@attribute Hellenic_Parliament numeric
@attribute Ministry_of_Foreign_Affairs numeric
@attribute Ministry_of_Interior_Public_Administration_and_Decentralisation numeric
@attribute Ministry_of_Macedonia_and_Thrace numeric
@attribute Ministry_of_Aegean numeric
@attribute Ministry_of_Public_Order numeric
@attribute Ministry_of_National_Defense numeric
@attribute Ministry_of_National_Education_and_religious_Affairs numeric
@attribute Ministry_of_Culture numeric
@attribute Ministry_of_Health_and_Social_Solidarity numeric
@attribute Ministry_of_Justice numeric
@attribute Hellenic_Ministry_of_transport_&_Communications numeric
@attribute Secretariat_General_of_Communication_&_Information numeric
@attribute Ministry_for_the_Environment_Physical_Planning_and_Public_Works numeric
@attribute Ministry_of_Economy_and_Finance numeric
@attribute Ministry_of_Development numeric
@attribute Ministry_of_Merchantile_Marine numeric
@attribute Ministry_of_Rural_Development_and_Food numeric
@attribute Ministry_of_Employment_and_Social_Protection numeric
@attribute Total numeric
@attribute Presidency_of_the_Hellenic_republic numeric
@attribute Hellenic_Parliament numeric
@attribute Ministry_of_Foreign_Affairs numeric
@attribute Ministry_of_Interior_Public_Administration_and_Decentralisation numeric
@attribute Ministry_of_Macedonia_and_Thrace numeric
@attribute Ministry_of_Aegean numeric
@attribute Ministry_of_Public_Order numeric
@attribute Ministry_of_National_Defense numeric
@attribute Ministry_of_National_Education_and_religious_Affairs numeric
@attribute Ministry_of_Culture numeric
@attribute Ministry_of_Health_and_Social_Solidarity numeric
@attribute Ministry_of_Justice numeric
@attribute Hellenic_Ministry_of_transport_&_Communications numeric
@attribute Secretariat_General_of_Communication_&_Information numeric
@attribute Ministry_for_the_Environment_Physical_Planning_and_Public_Works numeric
@attribute Ministry_of_Economy_and_Finance numeric
@attribute Ministry_of_Development numeric
@attribute Ministry_of_Merchantile_Marine numeric
@attribute Ministry_of_Rural_Development_and_Food numeric
@attribute Ministry_of_Employment_and_Social_Protection numeric
@attribute Total numeric
@attribute Direct_taxation numeric
@attribute Indirect_taxation numeric
@attribute Non-tax_income numeric
@attribute Public_investments_revenue numeric
@attribute Sum numeric
@attribute Direct_taxation numeric
@attribute Indirect_taxation numeric
@attribute Non-tax_income numeric
@attribute Public_investments_revenue numeric

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

@attribute Direct_taxation numeric
@attribute Indirect_taxation numeric
@attribute Non-tax_income numeric
@attribute Public_investments_revenue numeric
@attribute Sum numeric
@attribute Consumer_price_index numeric
@attribute Gross_debt_of_public_administration_current_prices numeric
@attribute Gross_debt_of_public_administration_in_1996_prices numeric
@attribute Permanent_population numeric
@attribute Per_capita_gross_debt_current_prices numeric
@attribute Per_capita_gross_debt_in_1996_prices numeric
@attribute Consumer_price_index numeric
@attribute Total_gross_debt_of_public_administration numeric
@attribute Gross_Domestic_Product_GDP numeric
@attribute Gross_debt_as_a_percentage_of_GDP numeric

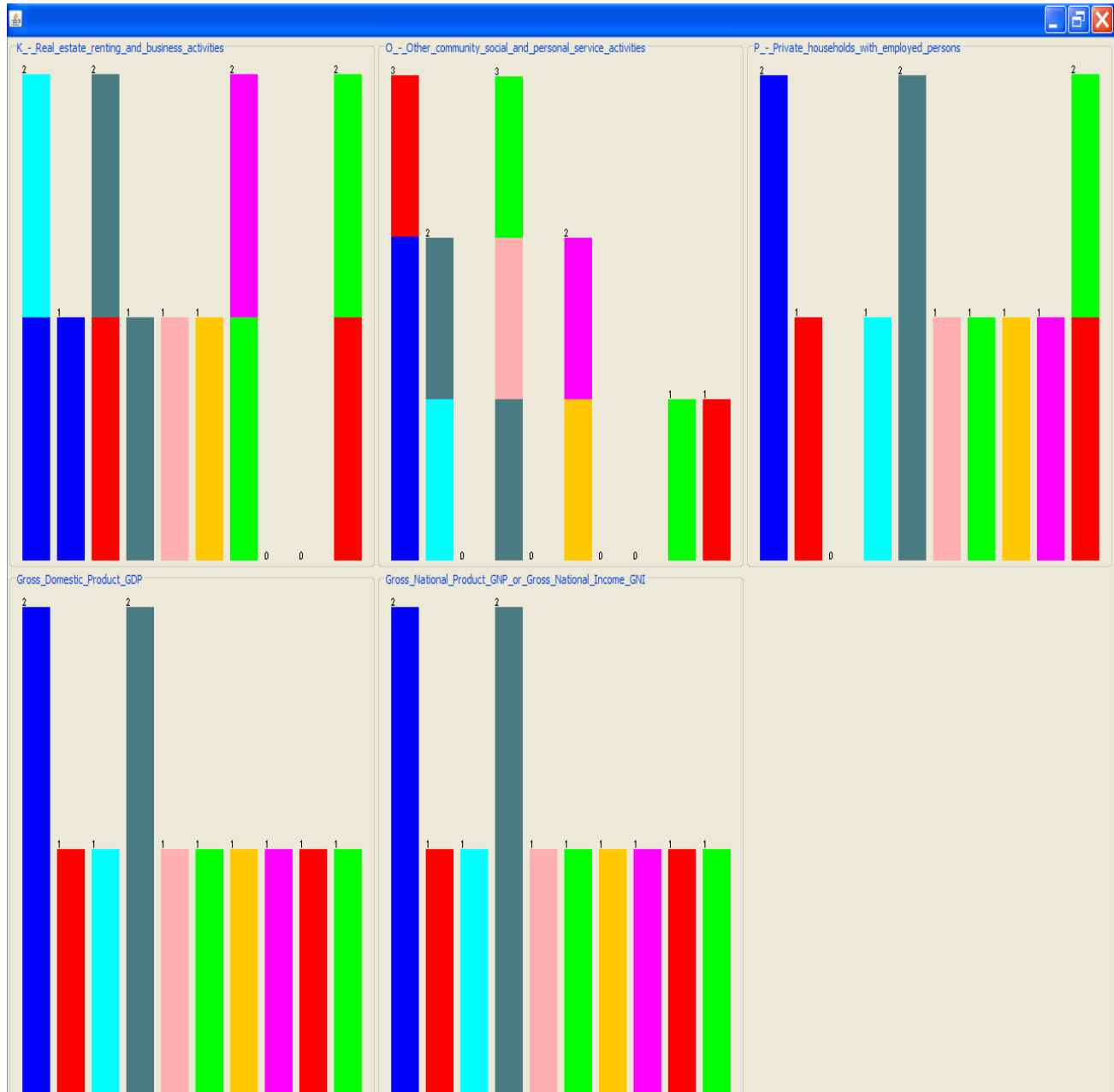
@data

1996	639	26	18	628	44	252	519	194	244	...
1997	618	24	17	605	42	249	524	197	231	...
1998	712	12	21	583	37	296	676	248	246	...
1999	710	13	20	575	41	285	684	254	251	...
2000	694	13	20	572	39	300	702	265	252	...
2001	645	13	19	577	37	307	705	269	256	...
2002	636	13	20	578	37	319	720	282	248	...
2003	643	13	13	567	41	346	745	287	261	...
2004	531	13	15	563	39	350	755	274	269	...
2005	528	14	19	561	38	362	777	298	271	...
2006	521	14	18	561	41	360	790	299	284	...
2007	515	27	14	511	35	305	973	304	244	...

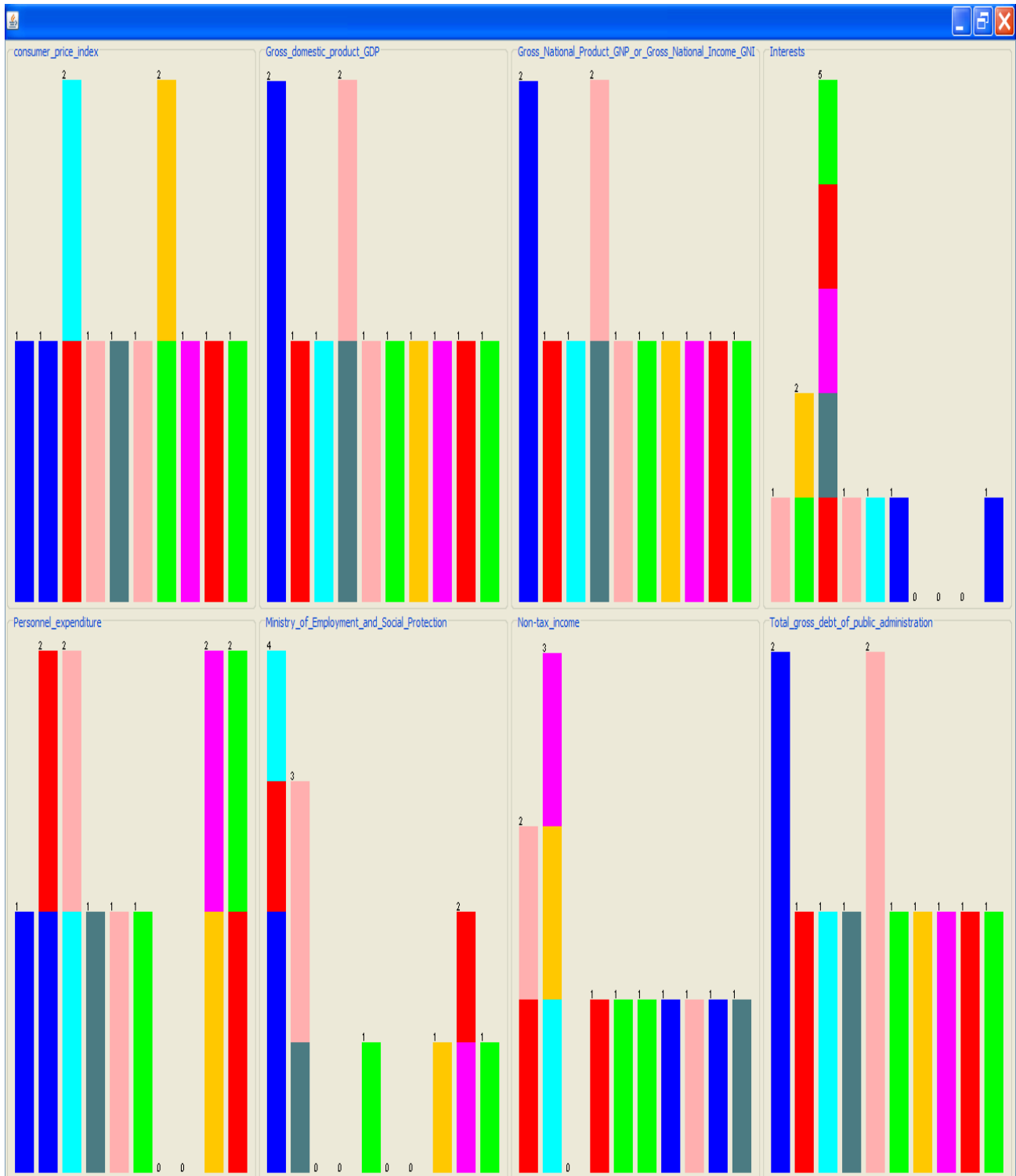
10.3] Οπτικοποίηση χαρακτηριστικών.

Στο παρόν παράρτημα παρουσιάζεται η οπτικοποίηση των χαρακτηριστικών, με την οποία ο ερευνητής αποκτά μια πρώτη εποπτική ματιά στα χαρακτηριστικά και τις σχέσεις μεταξύ τους, μέσω της γραφικής απεικόνισής τους.

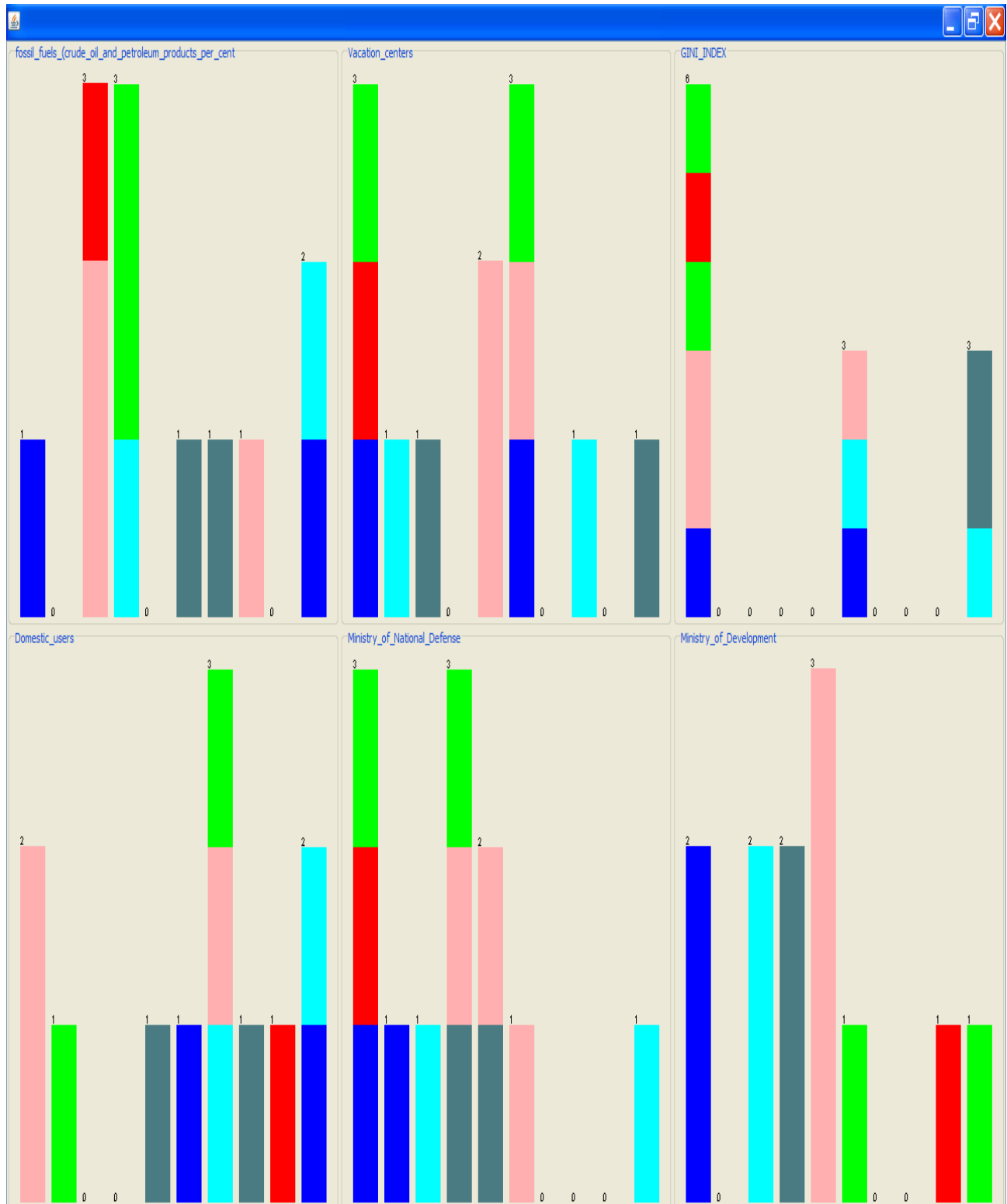
- Για Α.Ε.Π.:



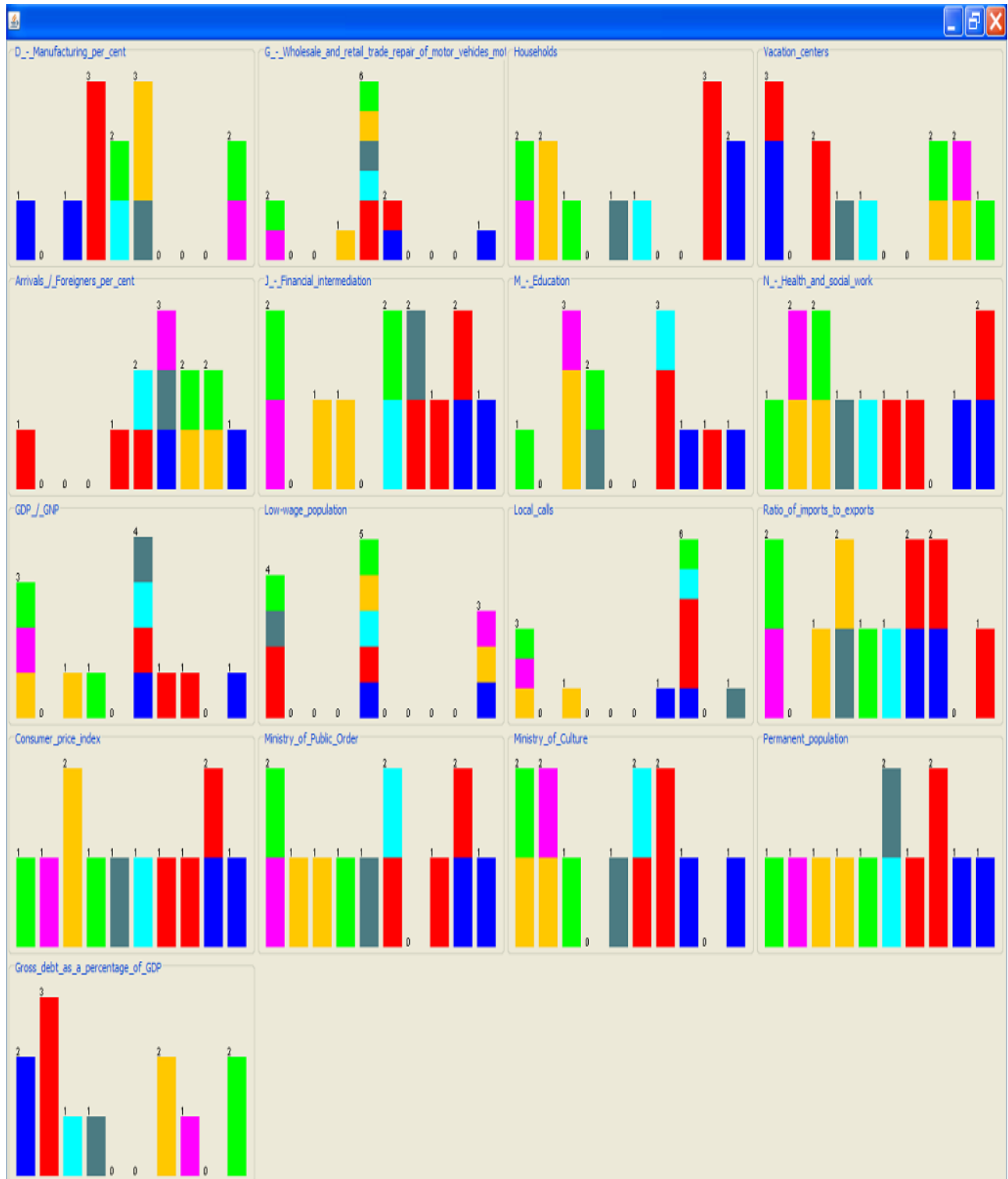
- Για Α.ΕΘ.Π.:



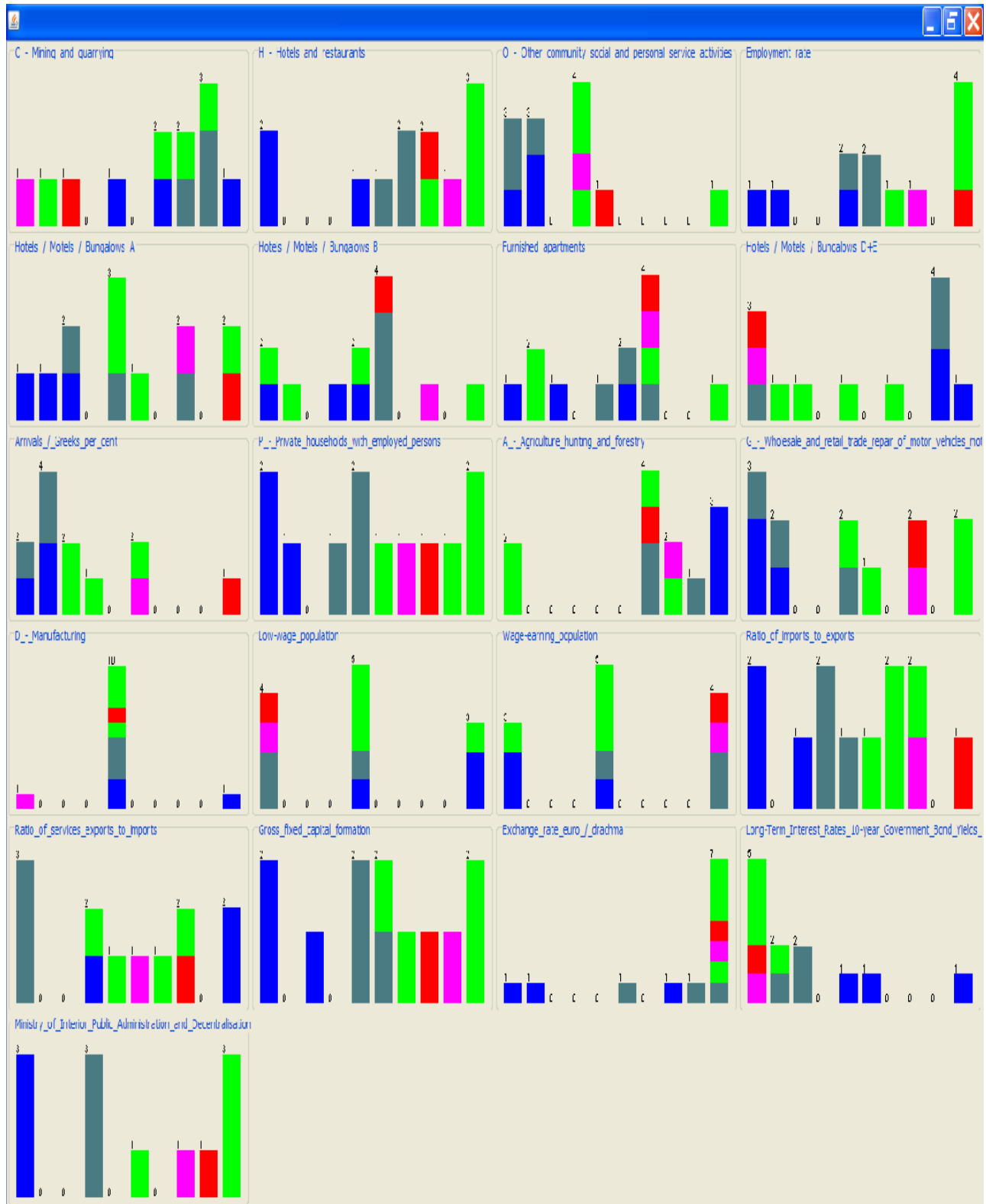
- Για Δ.Ε.Κ.:



- Για Δ.Τ.Κ.:

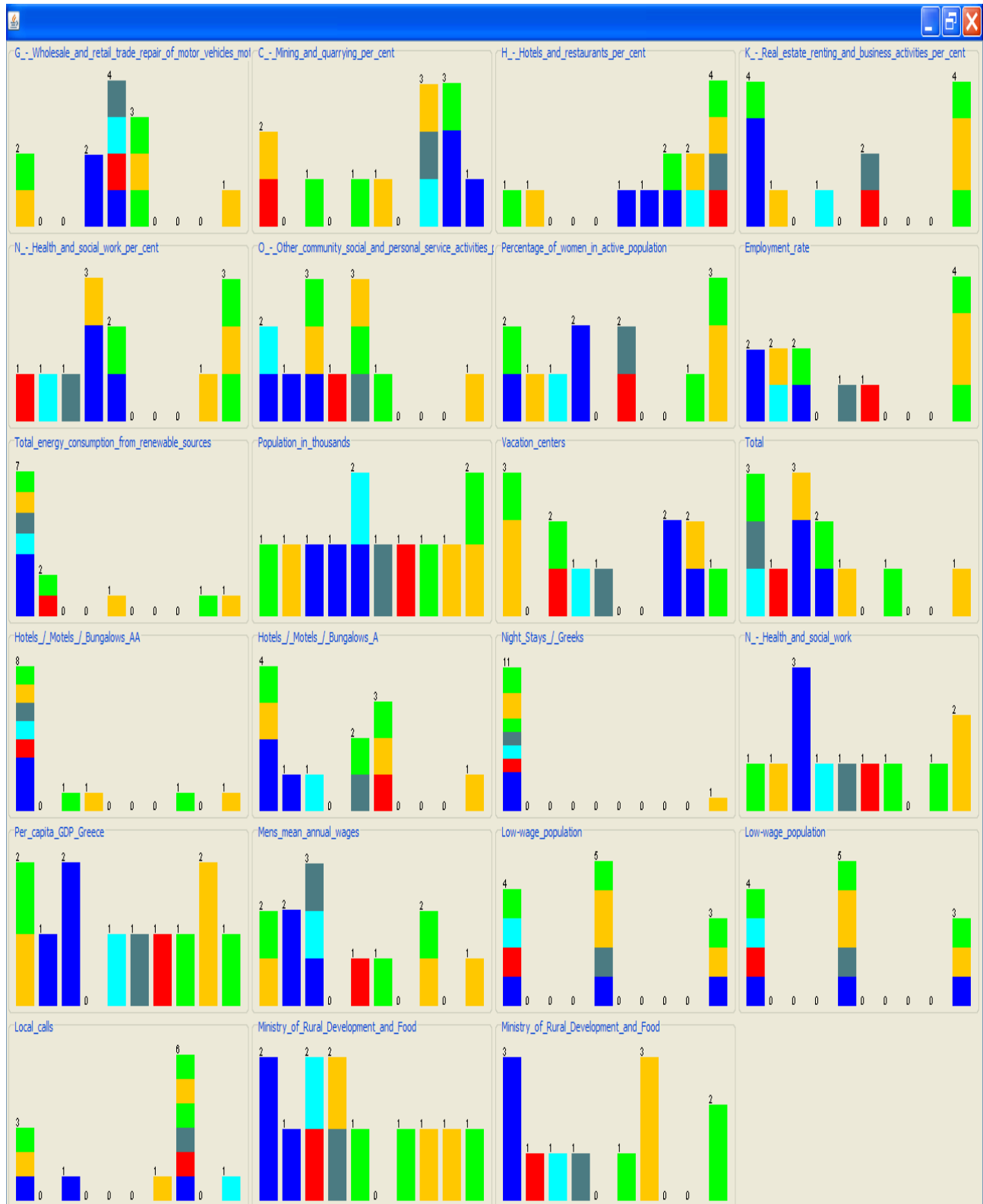


- Για Ακαθάριστο Σχηματισμό Παγίου Κεφαλαίου:

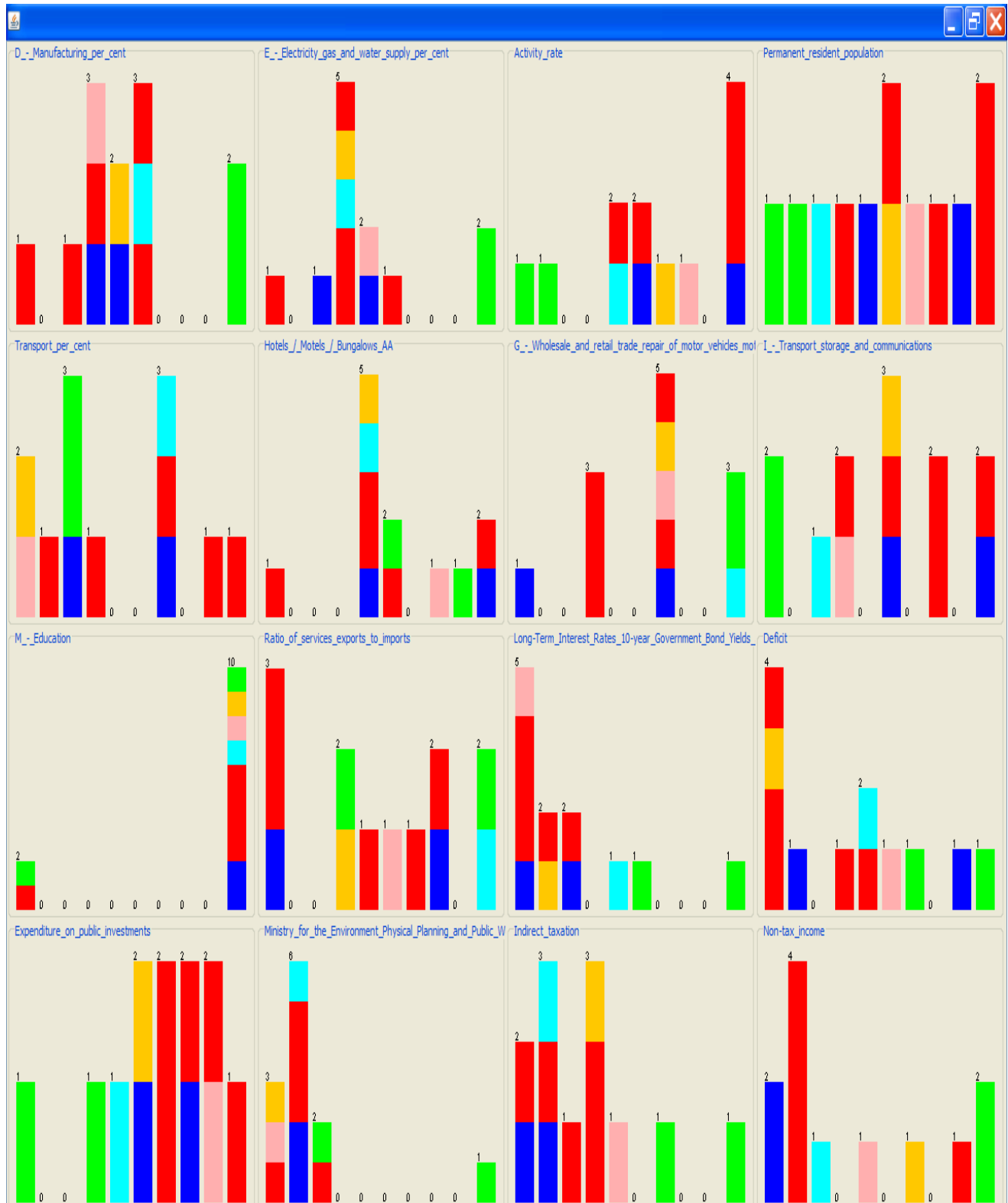


Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

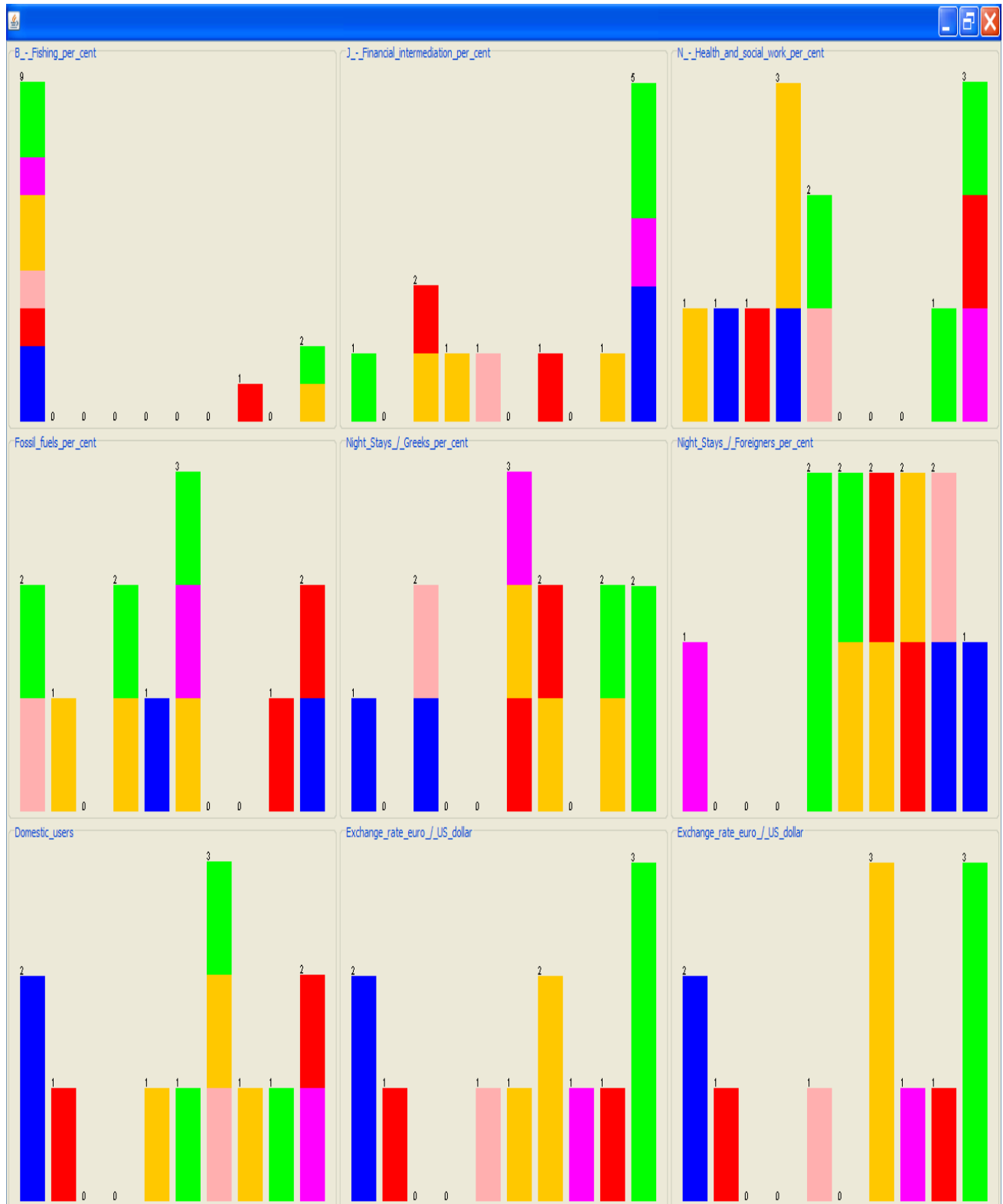
- Για Φορολογικό Εισόδημα:



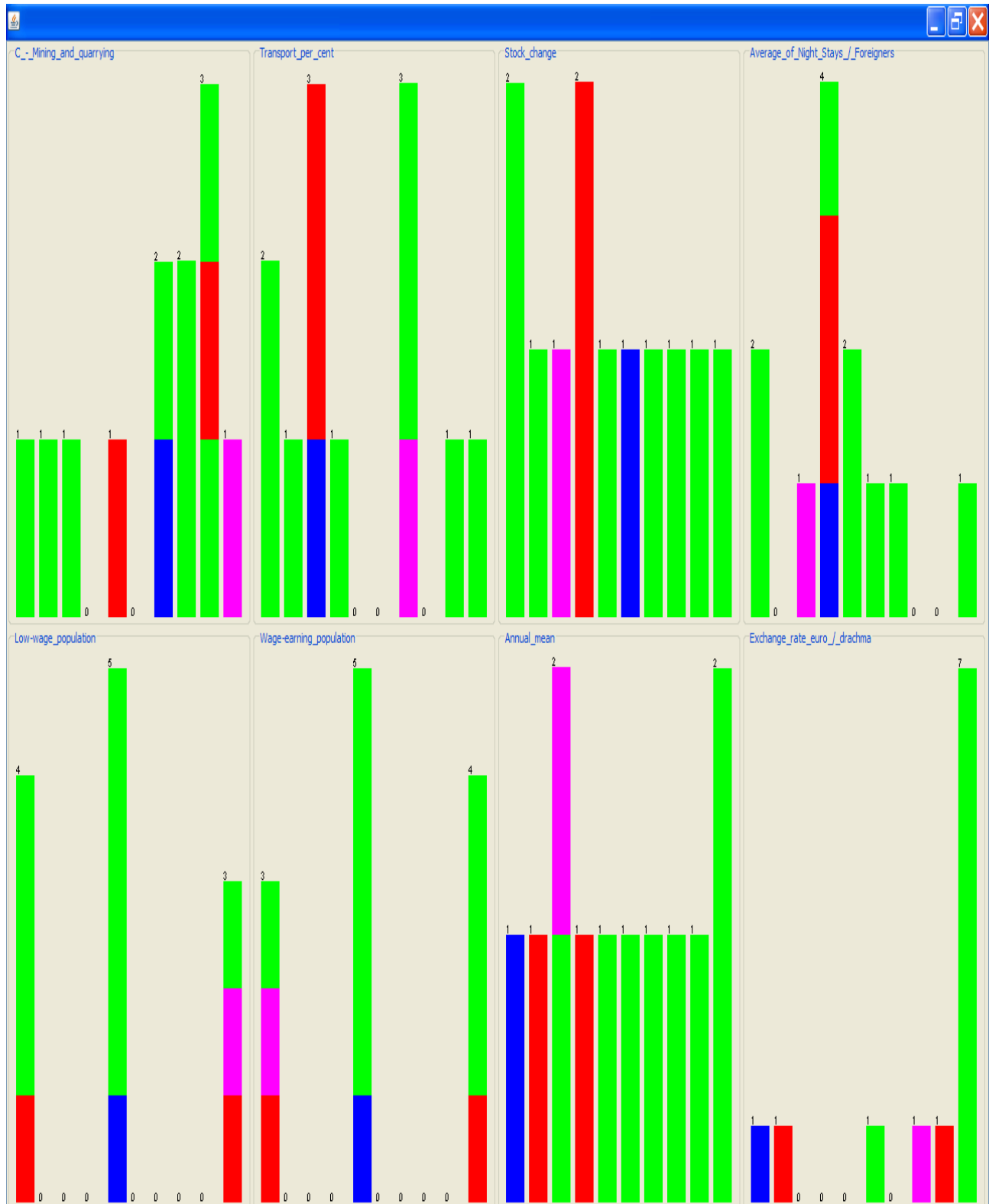
- Για Πληθωρισμό:



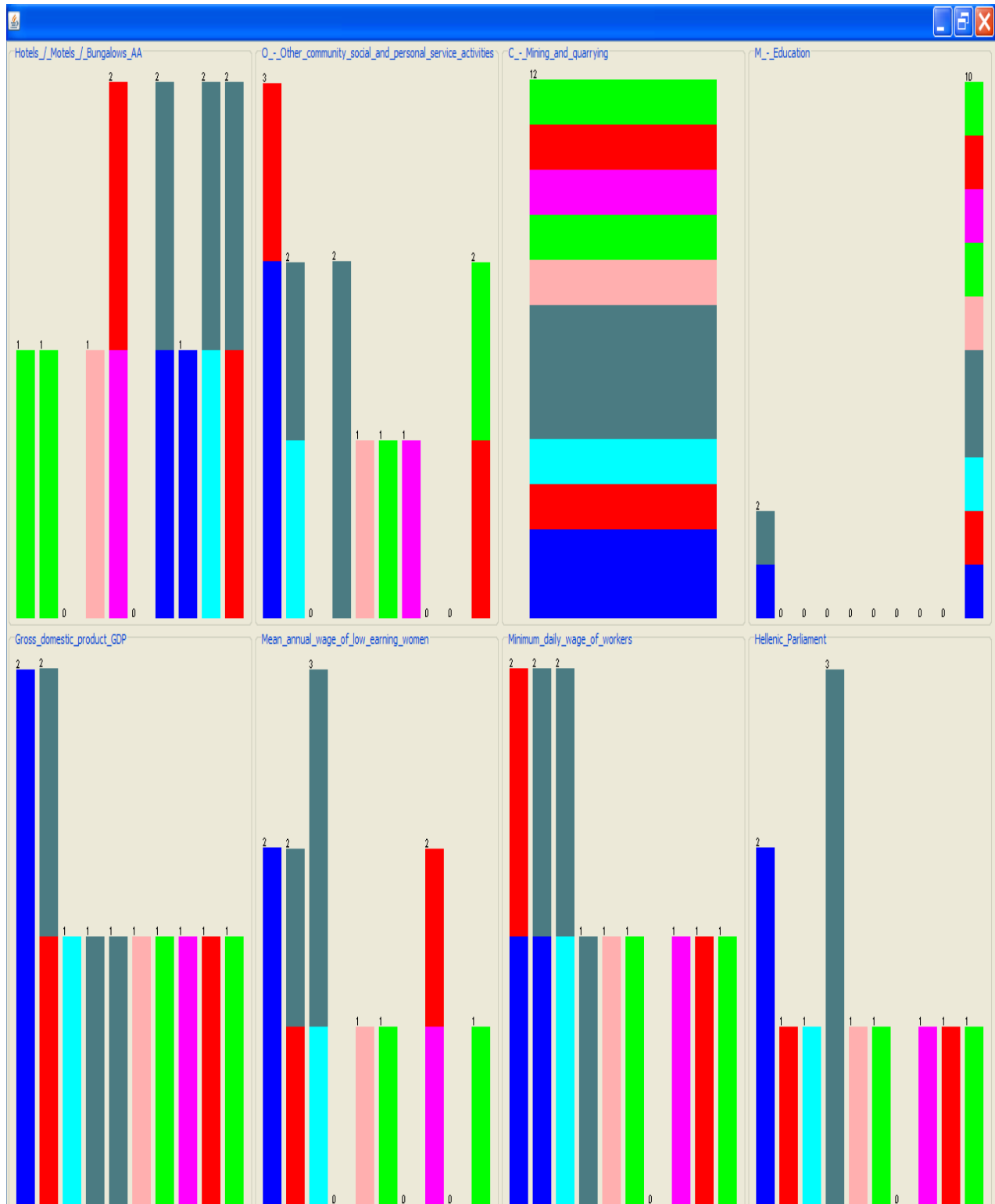
- Για ισοτιμία συναλλάγματος ευρώ / αμερικανικού δολαρίου:



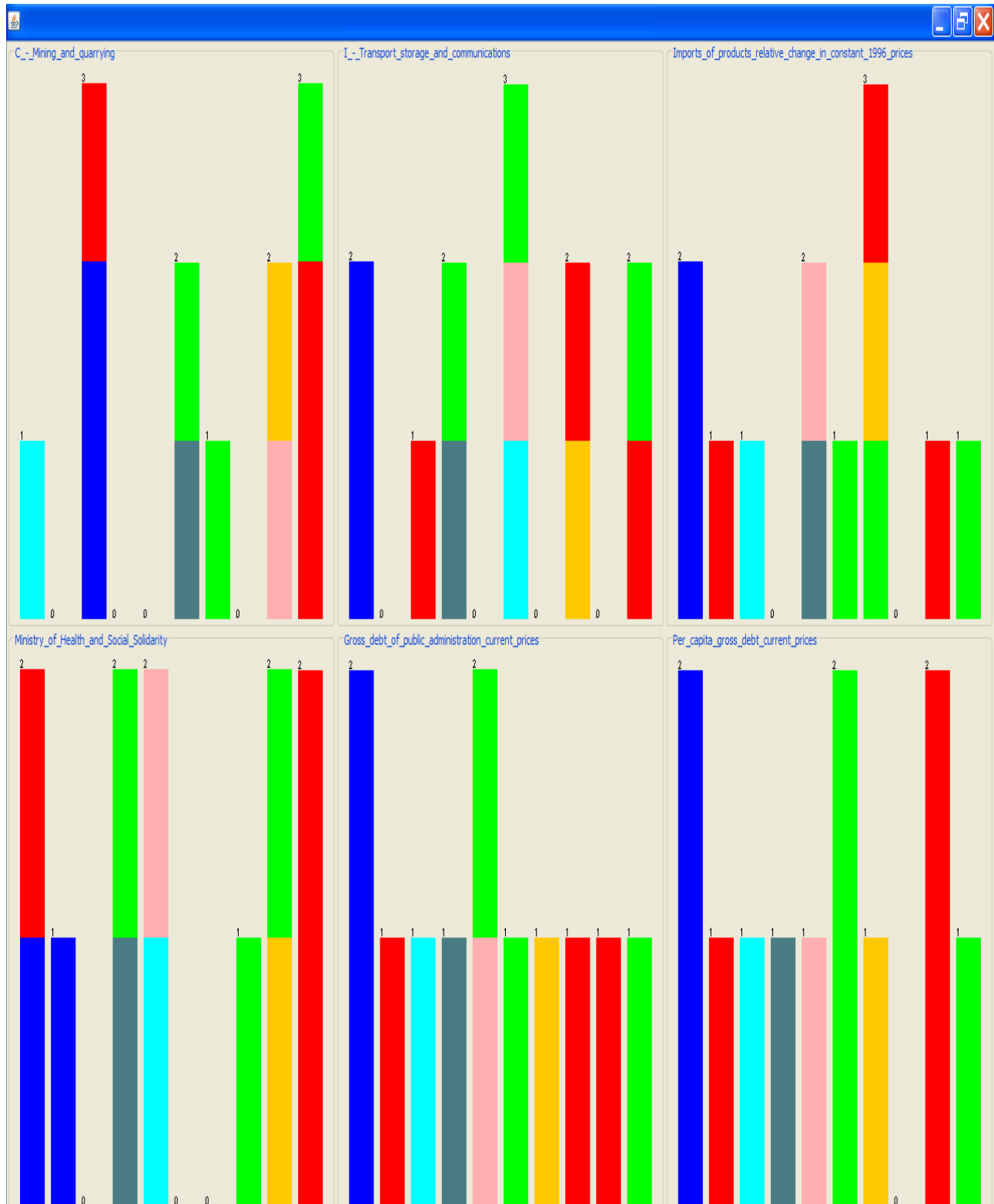
- Για Χαμηλόμισθο Πληθυσμό:



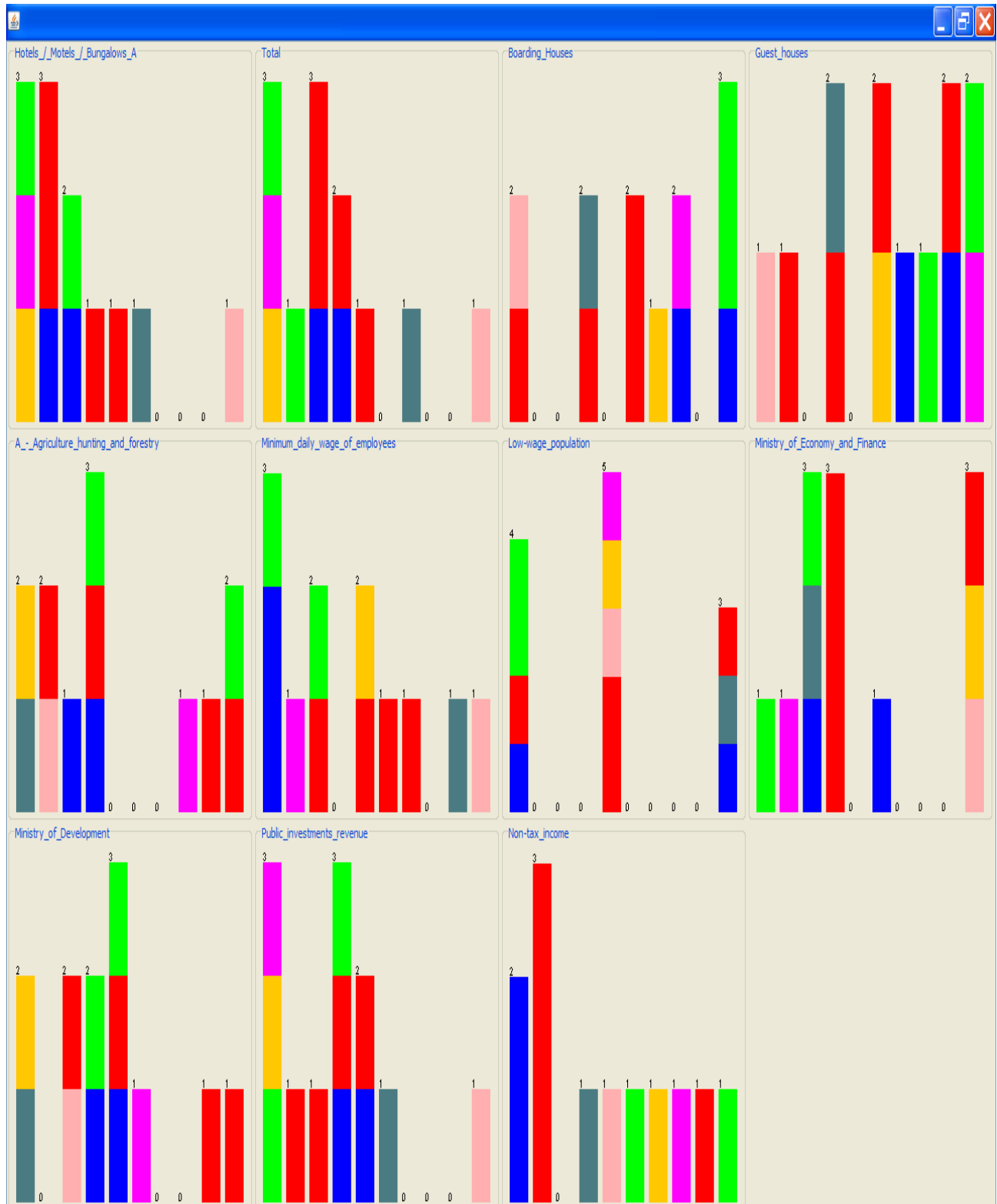
- Για ελάχιστο ημερήσιο μισθό των εργαζομένων:



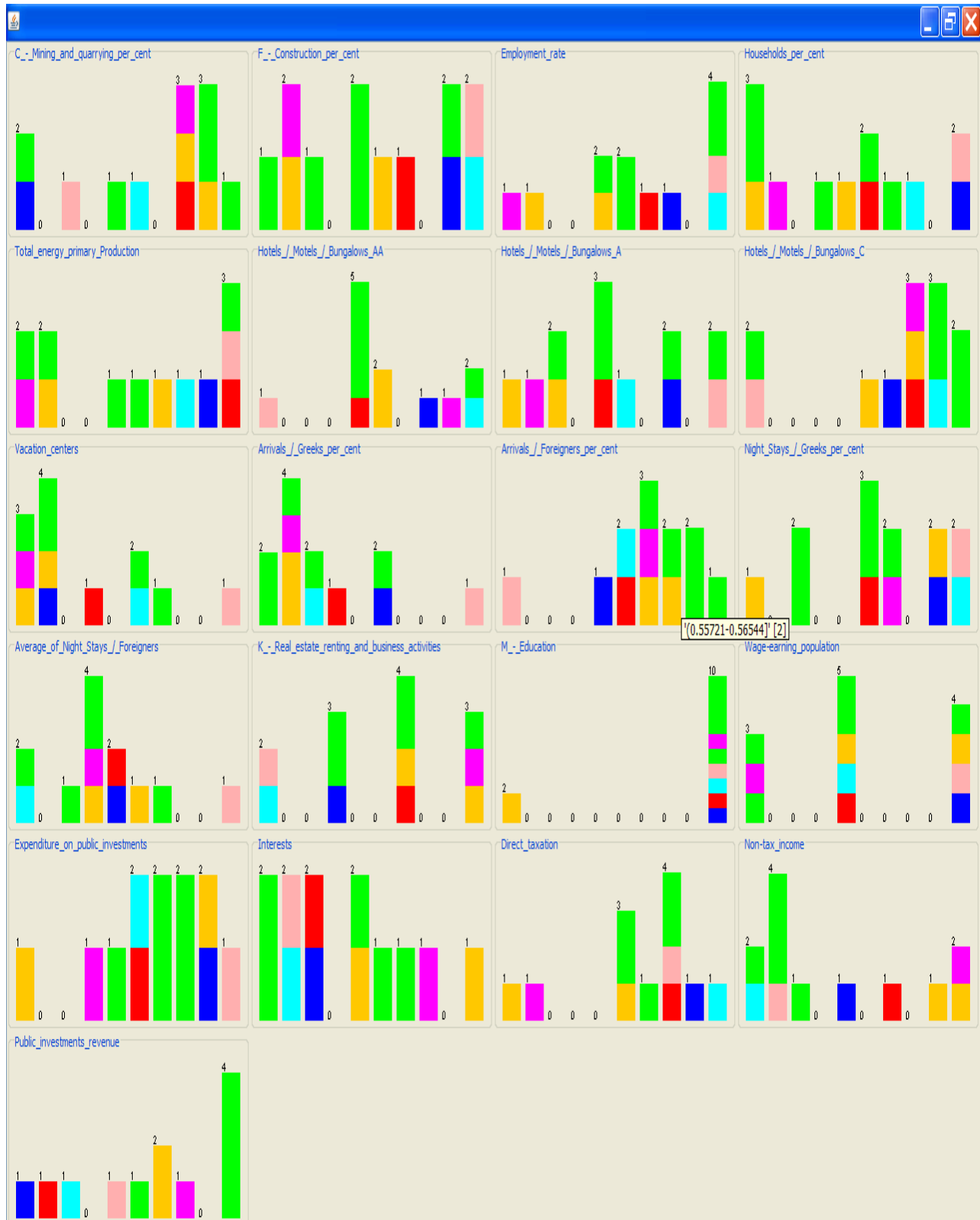
- Για κατά κεφαλή ακαθάριστο χρέος:



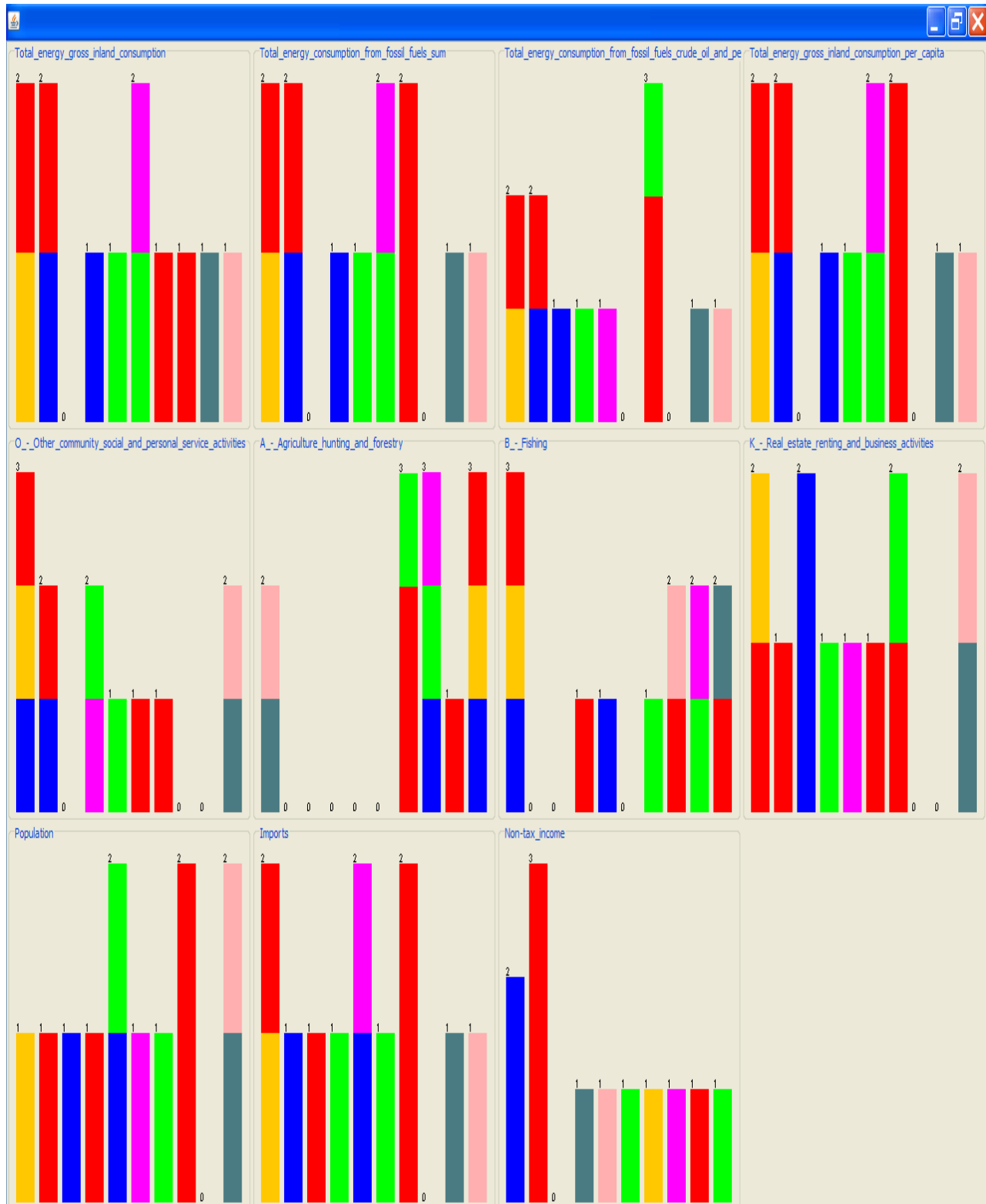
- Για Δημόσιες Δαπάνες:



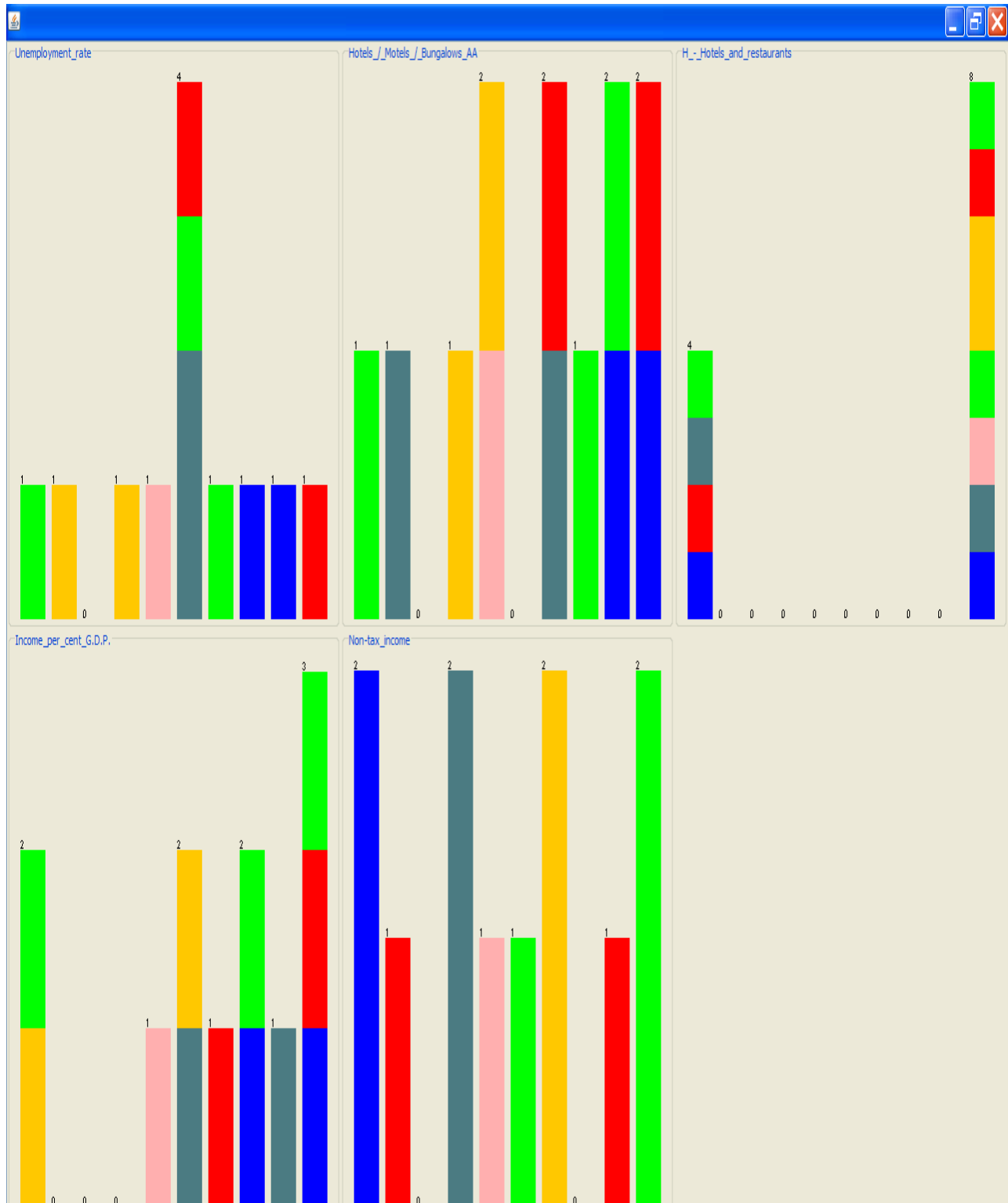
- Για αναλογία των υπηρεσιών (εξαγωγές προς εισαγωγές):



- Για συνολική ακαθάριστη ενεργειακή εγχώρια κατανάλωση:



- Για Ποσοστό Ανεργίας:



10.4] Παράρτημα Δ - Πινακοποίηση των πιο σχετικών, μεταξύ τους, χαρακτηριστικών (select attributes).

Στο παράρτημα αυτό γίνεται μια πινακοποίηση των πιο σχετικών, μεταξύ τους, χαρακτηριστικών, όπως αυτά περιγράφονται στο κεφάλαιο 4 - «Παρουσίαση και Ανάλυση Αποτελεσμάτων». Τα χαρακτηριστικά εμφανίζονται όπως ακριβώς τα παρουσιάζει το WEKA, ενώ οι διάφοροι αριθμοί είναι οι αύξοντες αριθμοί των attributes στο σύνολο των δεδομένων που εισήχθησαν στο WEKA.

Χαρακτηριστικό που αποτελεί τη «μεταβλητή - στόχο»	Επιλογή των πιο σχετικών, με τη «μεταβλητή - στόχο», χαρακτηριστικών (select attributes).
231 Gross_Domestic_Product_GDP	Selected attributes: 160,165,199,232 : 4 K_-Real_estate_renting_and_business_activities P_-Private_households_with_employed_persons O_-Other_community_social_and_personal_service_activities Gross_National_Product_GNP_or_Gross_National_Income_GNI
232 Gross_National_Product_GNP_or_Gross_National_Income_GNI	Selected attributes: 167,219,294,298,349,362,372 : 7 consumer_price_index Gross_domestic_product_GDP Interests Personnel_expenditure Ministry_of_Employment_and_Social_Protection Non-tax_income Total_gross_debt_of_public_administration
254 GINI_INDEX	Selected attributes: 63,133,259,337,346 : 5 fossil_fuels_(crude_oil_and_petroleum_products_per_cent Vacation_centers Domestic_users Ministry_of_National_Defense Ministry_of_Development

<p>303 Consumer_price_index</p>	<p>Selected attributes: 22,25,68,94,144,194,197,198,233,250,257,265,315,318,368,374 : 16</p> <p>D_-_Manufacturing_per_cent G_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycles_and_personal_and_household_goods_per_cent Households Vacation_centers Arrivals_/Foreigners_per_cent J_-_Financial_intermediation M_-_Education N_-_Health_and_social_work GDP_/GNP Low-wage_population Local_calls Ratio_of_imports_to_exports Ministry_of_Public_Order Ministry_of_Culture Permanent_population Gross_debt_as_a_percentage_of_GDP</p>
<p>277 Gross_fixed_capital_formation</p>	<p>Selected attributes: 4,9,16,43,106,107,112,129,143,183,185,191,205,252,253,265,273,278,284,333 : 20</p> <p>C_-_Mining_and_quarrying H_-_Hotels_and_restaurants O_-_Other_community_social_and_personal_service_activities Employment_rate Hotels_/Motels_/Bungalows_A Hotels_/Motels_/Bungalows_B Furnished_apartments Hotels_/Motels_/Bungalows_D+E Arrivals_/Greeks_per_cent P_-_Private_households_with_employed_persons A_-_Agriculture_hunting_and_forestry G_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycles_and_personal_and_household_goods D_-_Manufacturing Low-wage_population Wage-earning_population Ratio_of_imports_to_exports Ratio_of_services_exports_to_imports Exchange_rate_euro_/drachma LongTerm_Interest_Rates_10year_Government_Bond_Yields_Secondary_Markets Ministry_of_Macedonia_and_Thrace</p>
	<p>Selected attributes: 8,21,26,29,32,33,40,49,59,60,94,95,96,97,140,181,229,234,250,252,257,327,348 : 23</p> <p>G_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycles_and_</p>

<p>351 Fiscal_revenue</p>	<p>personal_and_household_goods C_-_Mining_and_quarrying_per_cent I_-_Transport_storage_and_communications_per_cent L_-_Public_administration_and_defence_compulsory_social_security_per_cent O_-_Other_community_social_and_personal_service_activities_per_cent P_-_Private_households_with_employed_persons_per_cent Unemployment_rate Permanent_residents Fossil_fuels_per_cent fossil_fuels_solid_per_cent Hotels_/Motels_/Bungalows_AA Hotels_/Motels_/Bungalows_A Hotels_/Motels_/Bungalows_B Hotels_/Motels_/Bungalows_C Arrivals_/Foreigners_per_cent A_-_Agriculture_hunting_and_forestry GDP_/GNP Mean_annual_wages_ratio_of_men_versus_women GINI_INDEX Inflation Imports Hellenic_Parliament Indirect_taxation</p>
<p>256 Inflation</p>	<p>Selected attributes: 22,23,25,27,39,42,75,105,214,273,284,288,301,344,357,358 : 16</p> <p>D_-_Manufacturing_per_cent E_-_Electricity_gas_and_water_supply_per_cent G_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycles_and_personal_and_household_goods_per_cent I_-_Transport_storage_and_communications_per_cent Activity_rate Permanent_resident_population Transport_per_cent Hotels_/Motels_/Bungalows_AA M_-_Education Ratio_of_services_exports_to_imports LongTerm_Interest_Rates_10year_Government_Bond_Yields_Secondary_Markets Deficit Expenditure_on_public_investments Ministry_for_the_Environment_Physical_Planning_and_Public_Works Indirect_taxation Non-tax_income</p>

<p>279 Exchange_rate_euro /_US_dollar</p>	<p>Selected attributes: 20,28,32,61,145,146,259,282: 8</p> <p>B_-_Fishing_per_cent J_-_Financial_intermediation_per_cent N_-_Health_and_social_work_per_cent Fossil_fuels_per_cent Night_Stays /_Greeks_per_cent Night_Stays /_Foreigners_per_cent Domestic_users Exchange_rate_euro /_japanese_yen</p>
<p>250 Low-wage_population</p>	<p>Selected attributes: 4,75,81,148,251,255,278 : 7</p> <p>C_-_Mining_and_quarrying Transport_per_cent Stock_change Average_of_Night_Stays /_Foreigners Wage-earning_population Local_calls Exchange_rate_euro /_japanese_yen</p>
<p>238 Minimum_daily_wage_of_workers</p>	<p>Selected attributes: 114,162,202,212,222,235,308 : 7</p> <p>Hotels /_Motels /_Bungalows_AA O_-_Other_community_social_and_personal_service_activities C_-_Mining_and_quarrying M_-_Education Gross_domestic_product_GDP Mean_annual_wage_of_low_earning_women Hellenic_Parliament</p>
<p>369 Per_capita_gross_debt_current_prices</p>	<p>Selected attributes: 170,210,266,340,366: 5</p> <p>C_-_Mining_and_quarrying I_-_Transport_storage_and_communications Imports_of_products_relative_change_in_constant_1996_prices Ministry_of_Health_and_Social_Solidarity Gross_debt_of_public_administration_current_prices</p>
<p>95 Public_Spending_Total</p>	<p>Selected attributes: 87,121,130,150,249,252,345,346,354,362 : 10</p> <p>Hotels /_Motels /_Bungalows_A Boarding_Houses Guest_houses A_-_Agriculture_hunting_and_forestry Minimum_daily_wage_of_employees Low-wage_population Ministry_of_Economy_and_Finance Ministry_of_Development Public_investments_revenue Non-tax_income</p>

<p>264 Ratio_of_imports _to_exports</p>	<p>Selected attributes: 21,24,42,72,76,104,105,107,112,142,143,144,147,211,213,252,300,304,355,357 ,358 : 21</p> <p>C_- Mining_and_quarrying_per_cent F_- Construction_per_cent Employment_rate Households_per_cent Total_energy_primary_Production Hotels_/Motels_/Bungalows_AA Hotels_/Motels_/Bungalows_A Hotels_/Motels_/Bungalows_C Vacation_centers Arrivals_/Greeks_per_cent Arrivals_/Foreigners_per_cent Night_Stays_/Greeks_per_cent Average_of_Night_Stays_/Foreigners K_- Real_estate_renting_and_business_activities M_- Education Wage-earning_population Expenditure_on_public_investments Interests Direct_taxation Non-tax_income Public_investments_revenue</p>
<p>54 Total_energy_gross_inland_ consumption</p>	<p>Selected attributes: 55,57,66,164,168,169,178,220,261,362 : 10</p> <p>Total_energy_consumption_from_fossil_fuels_sum Total_energy_consumption_from_fossil_fuels_crude_oil_and_petroleum_ products Total_energy_gross_inland_consumption_per_capita O_- Other_community_social_and_personal_service_activities A_- Agriculture_hunting_and_forestry B_- Fishing K_- Real_estate_renting_and_business_activities Population Imports Non-tax_income</p>
<p>41 Unemployment_ rate</p>	<p>Selected attributes: 115,209,290,353 : 4</p> <p>Hotels_/Motels_/Bungalows_AA H_- Hotels_and_restaurants Income_per_cent_G.D.P. Non-tax_income</p>

10.5] Σημαντικότεροι κανόνες ταξινόμησης.

Στο παρόν παράρτημα παρατίθενται οι πιο αντιπροσωπευτικοί, για το κάθε ένα από τα 15 χαρακτηριστικά που τέθηκαν ως «μεταβλητές - στόχοι», «χρήσιμοι» κανόνες ταξινόμησης στη «φυσική» τους μορφή, δηλαδή όπως ακριβώς προέκυψαν κατά την επεξεργασία των δεδομένων στο WEKA. Οι λεπτομέρειες που υπάρχουν δεν αναλύονται στο παρόν σύγγραμμα - κάτι που δεν ήταν σκοπός μας εξαρχής -, αλλά μπορεί να φανούν χρήσιμες σε μελλοντικές μελέτες. Οι πιο αντιπροσωπευτικοί, για το κάθε ένα από τα 15 αυτά χαρακτηριστικά, «χρήσιμοι» κανόνες ταξινόμησης είναι οι εξής :

- Από το Α.Ε.Π. και τα συσχετιζόμενα με αυτό χαρακτηριστικά δεν προκύπτουν οποιοδήποτε «χρήσιμοι» κανόνες ταξινόμησης (JRip ή PART).
- Από το Α.ΕΘ.Π. και τα συσχετιζόμενα με αυτό χαρακτηριστικά δεν προκύπτουν οποιοδήποτε «χρήσιμοι» κανόνες ταξινόμησης (JRip ή PART).
- Από το Δ.Ε.Κ. Gini και τα συσχετιζόμενα με αυτόν χαρακτηριστικά :

=== Run information ===

```
Scheme:          weka.classifiers.rules.PART -M 2 -C 0.25 -Q 1
Relation:        datatable-weka.filters.unsupervised.attribute.Remove-R135-
146,148-159,279-290-weka.filters.unsupervised.attribute.Remove-R1-62,64-
132,134-258,260-336,338-345,347-374-
weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last-
weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last
Instances:       12
Attributes:       5
                  fossil_fuels_(crude_oil_and_petroleum_products_per_cent
                  Vacation_centers
                  Domestic_users
                  Ministry_of_National_Defense
                  Ministry_of_Development
Test mode:       10-fold cross-validation
```

=== Classifier model (full training set) ===

PART decision list

```
Vacation_centers = '(-inf-33.94]': '(-inf-2734.429]' (3.0)
```

```
: '(3037.607-3189.196]' (9.0/6.0)
```

```
Number of Rules :      2
```

```
Time taken to build model: 0 seconds
```

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

```
=== Stratified cross-validation ===  
=== Summary ===
```

Correctly Classified Instances	6	50	%
Incorrectly Classified Instances	6	50	%
Kappa statistic	0.3333		
Mean absolute error	0.1292		
Root mean squared error	0.2695		
Relative absolute error	72.109	%	
Root relative squared error	89.0378	%	
Total Number of Instances	12		

```
=== Detailed Accuracy By Class ===
```

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0	1	1	1	1	'(-inf-2734.429]'
0	0	0	0	0	0.182	'(2734.429-2886.018]'
0	0	0	0	0	0.182	'(2886.018-3037.607]'
1	0.667	0.333	1	0.5	0.333	'(3037.607-3189.196]'
0	0	0	0	0	0.3	'(3189.196-3340.785]'
0	0	0	0	0	0.273	'(3340.785-3492.374]'
0	0	0	0	0	?	'(3492.374-3643.963]'
0	0	0	0	0	?	'(3643.963-3795.552]'
0	0	0	0	0	?	'(3795.552-3947.141]'
0	0	0	0	0	0.136	'(3947.141-inf)'

```
=== Confusion Matrix ===
```

```
a b c d e f g h i j <-- classified as  
3 0 0 0 0 0 0 0 0 0 | a = '(-inf-2734.429]'  
0 0 0 1 0 0 0 0 0 0 | b = '(2734.429-2886.018]'  
0 0 0 1 0 0 0 0 0 0 | c = '(2886.018-3037.607]'  
0 0 0 3 0 0 0 0 0 0 | d = '(3037.607-3189.196]'  
0 0 0 2 0 0 0 0 0 0 | e = '(3189.196-3340.785]'  
0 0 0 1 0 0 0 0 0 0 | f = '(3340.785-3492.374]'  
0 0 0 0 0 0 0 0 0 0 | g = '(3492.374-3643.963]'  
0 0 0 0 0 0 0 0 0 0 | h = '(3643.963-3795.552]'  
0 0 0 0 0 0 0 0 0 0 | i = '(3795.552-3947.141]'  
0 0 0 1 0 0 0 0 0 0 | j = '(3947.141-inf)'
```

- Από το Δ.Τ.Κ. και τα συσχετιζόμενα με αυτόν χαρακτηριστικά :

```
=== Run information ===
```

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

```
Scheme:          weka.classifiers.rules.JRip -F 3 -N 2.0 -O 2 -S 1
Relation:        datatable-weka.filters.unsupervised.attribute.Remove-R135-
146,148-159,279-290-weka.filters.unsupervised.attribute.Remove-R1-21,23-
24,26-67,69-93,95-143,145-193,195-196,199-232,234-249,251-256,258-264,266-
314,316-317,319-367,369-373-weka.filters.unsupervised.attribute.Discretize-
B10-M-1.0-Rfirst-last
Instances:       12
Attributes:      16
                 D_-_Manufacturing_per_cent
                 G_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycl
                 es_and_personal_and_household_goods_per_cent
                 Households
                 Vacation_centers
                 Arrivals_/Foreigners_per_cent
                 J_-_Financial_intermediation
                 M_-_Education
                 N_-_Health_and_social_work
                 GDP_/GNP
                 Low-wage_population
                 Local_calls
                 Ratio_of_imports_to_exports
                 Ministry_of_Public_Order
                 Ministry_of_Culture
                 Permanent_population
                 Gross_debt_as_a_percentage_of_GDP
Test mode:       10-fold cross-validation
```

=== Classifier model (full training set) ===

JRIP rules:

=====

=>

G_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycles_and_person
al_and_household_goods_per_cent='(0.16844-0.1755]' (12.0/6.0)

Number of Rules : 1

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	6	50	%
Incorrectly Classified Instances	6	50	%
Kappa statistic	0.0526		
Mean absolute error	0.1421		
Root mean squared error	0.2929		
Relative absolute error	87.2301	%	
Root relative squared error	104.1954	%	
Total Number of Instances	12		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
---------	---------	-----------	--------	-----------	----------	-------

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

```

0          0          0          0          0          0.5      '(-inf-
0.14726]'
0          0          0          0          0          ?        '(0.14726-
0.15432]'
0          0          0          0          0          ?        '(0.15432-
0.16138]'
0          0          0          0          0          0.091    '(0.16138-
0.16844]'
1          0.833      0.545      1          0.706      0.306    '(0.16844-
0.1755]'
0          0.1        0          0          0          0.1      '(0.1755-
0.18256]'
0          0          0          0          0          ?        '(0.18256-
0.18962]'
0          0          0          0          0          ?        '(0.18962-
0.19668]'
0          0          0          0          0          ?        '(0.19668-
0.20374]'
0          0          0          0          0          0.045    '(0.20374-
inf) '

```

=== Confusion Matrix ===

```

a b c d e f g h i j  <-- classified as
0 0 0 0 2 0 0 0 0 0 0 | a = '(-inf-0.14726]'
0 0 0 0 0 0 0 0 0 0 0 | b = '(0.14726-0.15432]'
0 0 0 0 0 0 0 0 0 0 0 | c = '(0.15432-0.16138]'
0 0 0 0 1 0 0 0 0 0 0 | d = '(0.16138-0.16844]'
0 0 0 0 6 0 0 0 0 0 0 | e = '(0.16844-0.1755]'
0 0 0 0 2 0 0 0 0 0 0 | f = '(0.1755-0.18256]'
0 0 0 0 0 0 0 0 0 0 0 | g = '(0.18256-0.18962]'
0 0 0 0 0 0 0 0 0 0 0 | h = '(0.18962-0.19668]'
0 0 0 0 0 0 0 0 0 0 0 | i = '(0.19668-0.20374]'
0 0 0 0 0 0 1 0 0 0 0 | j = '(0.20374-inf) '

```

- Από τον Ακαθάριστο Σχηματισμό Παγίου Κεφαλαίου και τα συσχετιζόμενα με αυτόν χαρακτηριστικά :

=== Run information ===

```

Scheme:          weka.classifiers.rules.JRip -F 3 -N 2.0 -O 2 -S 1
Relation:        datatable-weka.filters.unsupervised.attribute.Remove-R135-
146,148-159,279-290-weka.filters.unsupervised.attribute.Remove-R307-
weka.filters.unsupervised.attribute.Remove-R1-3,5-8,10-15,17-42,44-105,108-
111,113-128,130-142,144-182,184,186-190,192-204,206-251,254-264,266-272,274-
277,279-283,285-332,334-373-weka.filters.unsupervised.attribute.Discretize-
B10-M-1.0-Rfirst-last-weka.filters.unsupervised.attribute.Discretize-B10-M-
1.0-Rfirst-last
Instances:       12
Attributes:      20
                  C_-_Mining_and_quarrying
                  H_-_Hotels_and_restaurants
                  O_-_Other_community_social_and_personal_service_activities
                  Employment_rate
                  Hotels_/Motels_/Bungalows_A
                  Hotels_/Motels_/Bungalows_B

```

```

Furnished_apartments
Hotels_/Motels_/Bungalows_D+E
Arrivals_/Greeks_per_cent
P_-_Private_households_with_employed_persons
A_-_Agriculture_hunting_and_forestry
G_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycl
es_and_personal_and_household_goods
D_-_Manufacturing
Low-wage_population
Wage-earning_population
Ratio_of_imports_to_exports
Ratio_of_services_exports_to_imports
Exchange_rate_euro_/drachma
LongTerm_Interest_Rates_10year_Government_Bond_Yields_Secondary
_Markets
Ministry_of_Macedonia_and_Thrace
Test mode: 10-fold cross-validation

```

=== Classifier model (full training set) ===

JRIP rules:

=====

=> D_-_Manufacturing='(0.118-0.12]' (12.0/2.0)

Number of Rules : 1

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	10	83.3333 %
Incorrectly Classified Instances	2	16.6667 %
Kappa statistic	0	
Mean absolute error	0.0609	
Root mean squared error	0.1843	
Relative absolute error	51.5416 %	
Root relative squared error	86.9913 %	
Total Number of Instances	12	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0	0	0	0	0	0.045	'(-inf-0.112]'
0	0	0	0	0	?	'(0.112-
0.114]'						
0	0	0	0	0	?	'(0.114-
0.116]'						
0	0	0	0	0	?	'(0.116-
0.118]'						
1	1	0.833	1	0.909	0.1	'(0.118-0.12]'
0	0	0	0	0	?	'(0.12-0.122]'
0	0	0	0	0	?	'(0.122-
0.124]'						

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

```

0          0          0          0          0          ?          '(0.124-
0.126] '
0          0          0          0          0          ?          '(0.126-
0.128] '
0          0          0          0          0          0.045      '(0.128-inf) '

```

=== Confusion Matrix ===

```

a  b  c  d  e  f  g  h  i  j  <-- classified as
0  0  0  0  1  0  0  0  0  0  | a = '(-inf-0.112] '
0  0  0  0  0  0  0  0  0  0  | b = '(0.112-0.114] '
0  0  0  0  0  0  0  0  0  0  | c = '(0.114-0.116] '
0  0  0  0  0  0  0  0  0  0  | d = '(0.116-0.118] '
0  0  0  0 10  0  0  0  0  0  | e = '(0.118-0.12] '
0  0  0  0  0  0  0  0  0  0  | f = '(0.12-0.122] '
0  0  0  0  0  0  0  0  0  0  | g = '(0.122-0.124] '
0  0  0  0  0  0  0  0  0  0  | h = '(0.124-0.126] '
0  0  0  0  0  0  0  0  0  0  | i = '(0.126-0.128] '
0  0  0  0  1  0  0  0  0  0  | j = '(0.128-inf) '

```

- Από το Φορολογικό Εισόδημα και τα συσχετιζόμενα με αυτό χαρακτηριστικά :

=== Run information ===

```

Scheme:          weka.classifiers.rules.PART -M 2 -C 0.25 -Q 1
Relation:        datatable-weka.filters.unsupervised.attribute.Remove-R135-
146,148-159,279-290-weka.filters.unsupervised.attribute.Remove-
R47,114,134,364-weka.filters.unsupervised.attribute.Remove-R25-
weka.filters.unsupervised.attribute.Remove-R1-7,9-20,22-25,27-28,30-31,34-
39,41-48,50-58,61-93,98-130,132-139,141-180,182-228,230-233,235-249,251,253-
256,258-326,328-347,349-369-weka.filters.unsupervised.attribute.Discretize-
B10-M-1.0-Rfirst-last
Instances:       12
Attributes:      24
G_Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcycl
es_and_personal_and_household_goods
C_-_Mining_and_quarrying_per_cent
I_-_Transport_storage_and_communications_per_cent
L_Public_administration_and_defence_compulsory_social_security
_per_cent
O_Other_community_social_and_personal_service_activities_per_c
ent
P_-_Private_households_with_employed_persons_per_cent
Unemployment_rate
Permanent_residents
Fossil_fuels_per_cent
fossil_fuels_solid_per_cent
Hotels_/Motels_/Bungalows_AA
Hotels_/Motels_/Bungalows_A
Hotels_/Motels_/Bungalows_B
Hotels_/Motels_/Bungalows_C
Total
Arrivals_/Foreigners_per_cent
A_-_Agriculture_hunting_and_forestry
GDP_/GNP
Mean_annual_wages_ratio_of_men_versus_women

```


Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

```
GINI_INDEX
Inflation
Imports
Hellenic_Parliament
Indirect_taxation
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

PART decision list
-----

Indirect_taxation <= 21484: '(-inf-0.08756]' (8.0)

fossil_fuels_solid_per_cent = '(-inf-0.29624]': '(0.11908-0.13484]' (2.0/1.0)
: '(0.10332-0.11908]' (2.0/1.0)

Number of Rules :      3

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      7           58.3333 %
Incorrectly Classified Instances    5           41.6667 %
Kappa statistic                    0.2771
Mean absolute error                 0.0833
Root mean squared error            0.2386
Relative absolute error            57.3457 %
Root relative squared error        93.2196 %
Total Number of Instances         12

=== Detailed Accuracy By Class ===

TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
0.875    0         1          0.875  0.933     0.938    '(-inf-
0.08756]'
0         0         0          0      0         ?        '(0.08756-
0.10332]'
0         0.273    0          0      0         0.364    '(0.10332-
0.11908]'
0         0.182    0          0      0         0.318    '(0.11908-
0.13484]'
0         0         0          0      0         ?        '(0.13484-
0.1506]'
0         0         0          0      0         ?        '(0.1506-
0.16636]'
0         0         0          0      0         ?        '(0.16636-
0.18212]'
0         0         0          0      0         0.364    '(0.18212-
0.19788]'
0         0         0          0      0         ?        '(0.19788-
0.21364]'
```

```
0 0 0 0 0 0.318 '(0.21364-  
inf)'
```

```
=== Confusion Matrix ===
```

```
a b c d e f g h i j <-- classified as  
7 0 0 1 0 0 0 0 0 0 0 | a = '(-inf-0.08756]'  
0 0 0 0 0 0 0 0 0 0 0 | b = '(0.08756-0.10332]'  
0 0 0 1 0 0 0 0 0 0 0 | c = '(0.10332-0.11908]'  
0 0 1 0 0 0 0 0 0 0 0 | d = '(0.11908-0.13484]'  
0 0 0 0 0 0 0 0 0 0 0 | e = '(0.13484-0.1506]'  
0 0 0 0 0 0 0 0 0 0 0 | f = '(0.1506-0.16636]'  
0 0 0 0 0 0 0 0 0 0 0 | g = '(0.16636-0.18212]'  
0 0 1 0 0 0 0 0 0 0 0 | h = '(0.18212-0.19788]'  
0 0 0 0 0 0 0 0 0 0 0 | i = '(0.19788-0.21364]'  
0 0 1 0 0 0 0 0 0 0 0 | j = '(0.21364-inf)'
```

- Από τον πληθωρισμό και τα συσχετιζόμενα με αυτόν χαρακτηριστικά :

```
=== Run information ===
```

```
Scheme: weka.classifiers.rules.JRip -F 3 -N 2.0 -O 2 -S 1  
Relation: datatable-weka.filters.unsupervised.attribute.Remove-R135-  
146,148-159,279-290-weka.filters.unsupervised.attribute.Remove-R1-  
21,24,26,28-38,40-41,43-74,76-104,106-213,215-272,274-283,285-287,289-  
300,302-343,345-356,359-374-weka.filters.unsupervised.attribute.Discretize-  
B10-M-1.0-Rlast-weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-  
Rfirst-last-weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-  
last  
Instances: 12  
Attributes: 16  
D - Manufacturing_per_cent  
E - Electricity_gas_and_water_supply_per_cent  
G - Wholesale_and_retail_trade_repair_of_motor_vehicles_motorcyc  
les_and_personal_and_household_goods_per_cent  
I - Transport_storage_and_communications_per_cent  
Activity_rate  
Permanent_resident_population  
Transport_per_cent  
Hotels/_Motels/_Bungalows_AA  
M - Education  
Ratio_of_services_exports_to_imports  
LongTerm_Interest_Rates_10year_Government_Bond_Yields_Secondary  
_Markets  
Deficit  
Expenditure_on_public_investments  
Ministry_for_the_Environment_Physical_Planning_and_Public_Works  
Indirect_taxation  
Non-tax_income  
Test mode: 10-fold cross-validation
```

```
=== Classifier model (full training set) ===
```

```
JRIP rules:  
=====
```

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

```
(Transport_per_cent = '(-inf-0.38424]') =>
Ministry_for_the_Environment_Physical_Planning_and_Public_Works='(-inf-
115.914]' (2.0/0.0)
=>
Ministry_for_the_Environment_Physical_Planning_and_Public_Works='(115.914-
143.478]' (10.0/4.0)
```

Number of Rules : 2

Time taken to build model: 0 seconds

```
=== Stratified cross-validation ===
=== Summary ===
```

Correctly Classified Instances	6	50	%
Incorrectly Classified Instances	6	50	%
Kappa statistic	0		
Mean absolute error	0.1377		
Root mean squared error	0.2829		
Relative absolute error	86.0611	%	
Root relative squared error	101.8712	%	
Total Number of Instances	12		

```
=== Detailed Accuracy By Class ===
```

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0	0	0	0	0	0.463	'(-inf-115.914]'
1	1	0.5	1	0.667	0.111	'(115.914-143.478]'
0	0	0	0	0	0.1	'(143.478-171.042]'
0	0	0	0	0	?	'(171.042-198.606]'
0	0	0	0	0	?	'(198.606-226.17]'
0	0	0	0	0	?	'(226.17-253.734]'
0	0	0	0	0	?	'(253.734-281.298]'
0	0	0	0	0	?	'(281.298-308.862]'
0	0	0	0	0	?	'(308.862-336.426]'
0	0	0	0	0	0.045	'(336.426-inf)'

```
=== Confusion Matrix ===
```

```
a b c d e f g h i j  <-- classified as
0 3 0 0 0 0 0 0 0 0 | a = '(-inf-115.914] '
0 6 0 0 0 0 0 0 0 0 | b = '(115.914-143.478] '
0 2 0 0 0 0 0 0 0 0 | c = '(143.478-171.042] '
0 0 0 0 0 0 0 0 0 0 | d = '(171.042-198.606] '
```

```
0 0 0 0 0 0 0 0 0 0 | e = '(198.606-226.17]'  
0 0 0 0 0 0 0 0 0 0 | f = '(226.17-253.734]'  
0 0 0 0 0 0 0 0 0 0 | g = '(253.734-281.298]'  
0 0 0 0 0 0 0 0 0 0 | h = '(281.298-308.862]'  
0 0 0 0 0 0 0 0 0 0 | i = '(308.862-336.426]'  
0 1 0 0 0 0 0 0 0 0 | j = '(336.426-inf)'
```

- Από τη συναλλαγματική ισοτιμία ευρώ / αμερικανικού δολαρίου και τα συσχετιζόμενα με αυτήν χαρακτηριστικά :

=== Run information ===

```
Scheme:          weka.classifiers.rules.PART -M 2 -C 0.25 -Q 1  
Relation:        datatable-weka.filters.unsupervised.attribute.Remove-R135-  
146,148-159,279-290-weka.filters.unsupervised.attribute.Remove-R282-  
weka.filters.unsupervised.attribute.Remove-R1-19,21-27,29-31,33-60,62-  
144,147-241,243-258,260-281,283-348,350-373-  
weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last-  
weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last  
Instances:       12  
Attributes:      10  
                 B_-_Fishing_per_cent  
                 J_-_Financial_intermediation_per_cent  
                 N_-_Health_and_social_work_per_cent  
                 Fossil_fuels_per_cent  
                 Night_Stays_/_Greeks_per_cent  
                 Night_Stays_/_Foreigners_per_cent  
                 ratio  
                 Domestic_users  
                 Exchange_rate_euro_/_japanese_yen  
                 Total  
Test mode:       10-fold cross-validation
```

=== Classifier model (full training set) ===

PART decision list

: '(0.9407-inf)' (12.0/5.0)

Number of Rules : 1

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	7	58.3333 %
Incorrectly Classified Instances	5	41.6667 %
Kappa statistic	0	
Mean absolute error	0.1221	
Root mean squared error	0.2571	
Relative absolute error	81.3711 %	
Root relative squared error	97.8007 %	
Total Number of Instances	12	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.9383]	0	0	0	0	0	'(-inf-
0.9386]	0	0	0	0	?	'(0.9383-
0.9389]	0	0	0	0	?	'(0.9386-
0.9392]	0	0	0	0	?	'(0.9389-
0.9395]	0	0	0	0	?	'(0.9392-
0.9398]	0	0	0	0	?	'(0.9395-
0.9401]	0	0	0	0	0.1	'(0.9398-
0.9404]	0	0	0	0	?	'(0.9401-
0.9407]	0	0	0	0	?	'(0.9404-
1	1	0.583	1	0.737	0.057	'(0.9407-inf)'

=== Confusion Matrix ===

```

a b c d e f g h i j  <-- classified as
0 0 0 0 0 0 0 0 0 3 | a = '(-inf-0.9383]'
0 0 0 0 0 0 0 0 0 0 | b = '(0.9383-0.9386]'
0 0 0 0 0 0 0 0 0 0 | c = '(0.9386-0.9389]'
0 0 0 0 0 0 0 0 0 0 | d = '(0.9389-0.9392]'
0 0 0 0 0 0 0 0 0 0 | e = '(0.9392-0.9395]'
0 0 0 0 0 0 0 0 0 0 | f = '(0.9395-0.9398]'
0 0 0 0 0 0 0 0 0 2 | g = '(0.9398-0.9401]'
0 0 0 0 0 0 0 0 0 0 | h = '(0.9401-0.9404]'
0 0 0 0 0 0 0 0 0 0 | i = '(0.9404-0.9407]'
0 0 0 0 0 0 0 0 0 7 | j = '(0.9407-inf)'
```

- Από το χαρακτηριστικό «χαμηλόμισθος πληθυσμός» και τα συσχετιζόμενα με αυτό χαρακτηριστικά δεν προκύπτουν οποιοδήποτε «χρήσιμοι» κανόνες ταξινόμησης (JRip ή PART).
- Από το χαρακτηριστικό «ελάχιστος ημερήσιος μισθός των εργαζομένων» και τα συσχετιζόμενα με αυτό χαρακτηριστικά :

=== Run information ===

```

Scheme:          weka.classifiers.rules.JRip -F 3 -N 2.0 -O 2 -S 1
Relation:        datatable-weka.filters.unsupervised.attribute.Remove-R135-
146,148-159,279-290-weka.filters.unsupervised.attribute.Remove-R136-
weka.filters.unsupervised.attribute.Remove-R16-
weka.filters.unsupervised.attribute.Remove-R1-113,115-161,163-201,203-
211,213-221,223-234,236-238,240-307,309-372-
weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last-
weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rlast-
```

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

```
weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last-
weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rlast
Instances:      12
Attributes:     8
                Hotels_/_Motels_/_Bungalows_AA
                O_-_Other_community_social_and_personal_service_activities
                C_-_Mining_and_quarrying
                M_-_Education
                Gross_domestic_product_GDP
                Mean_annual_wage_of_low_earning_women
                Minimum_daily_wage_of_employees
                Hellenic_Parliament
Test mode:     10-fold cross-validation
```

=== Classifier model (full training set) ===

JRIP rules:

=====

=> M_-_Education='(0.049-inf)' (12.0/2.0)

Number of Rules : 1

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	10	83.3333 %
Incorrectly Classified Instances	2	16.6667 %
Kappa statistic	0	
Mean absolute error	0.0576	
Root mean squared error	0.1782	
Relative absolute error	49.4179 %	
Root relative squared error	85.3934 %	
Total Number of Instances	12	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0	0	0	0	0	0.1	'(-inf-0.041]'
0	0	0	0	0	?	'(0.041-
0.042]'						
0	0	0	0	0	?	'(0.042-
0.043]'						
0	0	0	0	0	?	'(0.043-
0.044]'						
0	0	0	0	0	?	'(0.044-
0.045]'						
0	0	0	0	0	?	'(0.045-
0.046]'						
0	0	0	0	0	?	'(0.046-
0.047]'						
0	0	0	0	0	?	'(0.047-
0.048]'						

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

```

0      0      0      0      0      ?      '(0.048-
0.049]'
1      1      0.833  1      0.909  0.1    '(0.049-inf)'

```

=== Confusion Matrix ===

```

a b c d e f g h i j <-- classified as
0 0 0 0 0 0 0 0 0 2 | a = '(-inf-0.041]'
0 0 0 0 0 0 0 0 0 0 | b = '(0.041-0.042]'
0 0 0 0 0 0 0 0 0 0 | c = '(0.042-0.043]'
0 0 0 0 0 0 0 0 0 0 | d = '(0.043-0.044]'
0 0 0 0 0 0 0 0 0 0 | e = '(0.044-0.045]'
0 0 0 0 0 0 0 0 0 0 | f = '(0.045-0.046]'
0 0 0 0 0 0 0 0 0 0 | g = '(0.046-0.047]'
0 0 0 0 0 0 0 0 0 0 | h = '(0.047-0.048]'
0 0 0 0 0 0 0 0 0 0 | i = '(0.048-0.049]'
0 0 0 0 0 0 0 0 0 10 | j = '(0.049-inf)'

```

- Από το κατά κεφαλή Ακαθάριστο Χρέος και τα συσχετιζόμενα με αυτό χαρακτηριστικά δεν προκύπτουν οποιοδήποτε «χρήσιμοι» κανόνες ταξινόμησης (JRip ή PART).
- Από τις Δημόσιες Δαπάνες και τα συσχετιζόμενα με αυτές χαρακτηριστικά δεν προκύπτουν οποιοδήποτε «χρήσιμοι» κανόνες ταξινόμησης (JRip ή PART).
- Από το χαρακτηριστικό «αναλογία των υπηρεσιών - εξαγωγές προς εισαγωγές» και τα συσχετιζόμενα με αυτό χαρακτηριστικά :

=== Run information ===

```

Scheme:      weka.classifiers.rules.JRip -F 3 -N 2.0 -O 2 -S 1
Relation:    datatable-weka.filters.unsupervised.attribute.Remove-R135-
146,148-159,279-290-weka.filters.unsupervised.attribute.Remove-R29-
weka.filters.unsupervised.attribute.Remove-R1-20,22-23,25-41,43-71,73-75,77-
103,106,108-111,113-141,145-146,148-210,212,214-251,253-299,301-303,305-
354,356,359-373-weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-
Rfirst-last-weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-
last
Instances:   12
Attributes:  21
C_-_Mining_and_quarrying_per_cent
F_-_Construction_per_cent
Employment_rate
Households_per_cent
Total_energy_primary_Production
Hotels_/Motels_/Bungalows_AA
Hotels_/Motels_/Bungalows_A
Hotels_/Motels_/Bungalows_C
Vacation_centers
Arrivals_/Greeks_per_cent
Arrivals_/Foreigners_per_cent
Night_Stays_/Greeks_per_cent
Average_of_Night_Stays_/Foreigners
K_-_Real_estate_renting_and_business_activities
M_-_Education
Wage-earning_population

```

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

```
Expenditure_on_public_investments
Interests
Direct_taxation
Non-tax_income
Public_investments_revenue
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

JRIP rules:
=====

(Hotels_/Motels/_Bungalows_AA = '(10.9-11.12]') => M_-_Education='(-inf-
0.041] '(2.0/0.0)
=> M_-_Education='(0.049-inf)' (10.0/0.0)

Number of Rules : 2

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          10           83.3333 %
Incorrectly Classified Instances         2           16.6667 %
Kappa statistic                          0
Mean absolute error                      0.0333
Root mean squared error                   0.1653
Relative absolute error                   28.6104 %
Root relative squared error               79.2429 %
Total Number of Instances               12

=== Detailed Accuracy By Class ===

TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
0         0         0          0        0          0.9      '(-inf-0.041] '
0         0         0          0        0          ?        '(0.041-
0.042] '
0         0         0          0        0          ?        '(0.042-
0.043] '
0         0         0          0        0          ?        '(0.043-
0.044] '
0         0         0          0        0          ?        '(0.044-
0.045] '
0         0         0          0        0          ?        '(0.045-
0.046] '
0         0         0          0        0          ?        '(0.046-
0.047] '
0         0         0          0        0          ?        '(0.047-
0.048] '
0         0         0          0        0          ?        '(0.048-
0.049] '
1         1         0.833     1        0.909     0.9      '(0.049-inf) '

=== Confusion Matrix ===
```



```
a b c d e f g h i j <-- classified as
0 0 0 0 0 0 0 0 0 2 | a = '(-inf-0.041]'
0 0 0 0 0 0 0 0 0 0 | b = '(0.041-0.042]'
0 0 0 0 0 0 0 0 0 0 | c = '(0.042-0.043]'
0 0 0 0 0 0 0 0 0 0 | d = '(0.043-0.044]'
0 0 0 0 0 0 0 0 0 0 | e = '(0.044-0.045]'
0 0 0 0 0 0 0 0 0 0 | f = '(0.045-0.046]'
0 0 0 0 0 0 0 0 0 0 | g = '(0.046-0.047]'
0 0 0 0 0 0 0 0 0 0 | h = '(0.047-0.048]'
0 0 0 0 0 0 0 0 0 0 | i = '(0.048-0.049]'
0 0 0 0 0 0 0 0 0 10 | j = '(0.049-inf)'
```

- Από το χαρακτηριστικό «Συνολική ακαθάριστη ενεργειακή εγχώρια κατανάλωση» και τα συσχετιζόμενα με αυτό χαρακτηριστικά δεν προκύπτουν οποιοιδήποτε «χρήσιμοι» κανόνες ταξινόμησης (JRip ή PART).
- Από το ποσοστό ανεργίας στην Ελλάδα και τα συσχετιζόμενα με αυτό χαρακτηριστικά :

=== Run information ===

```
Scheme:          weka.classifiers.rules.JRip -F 3 -N 2.0 -O 2 -S 1
Relation:        datatable-weka.filters.unsupervised.attribute.Remove-R135-
146,148-159,279-290-weka.filters.unsupervised.attribute.Remove-R1-44,46-
114,116-208,210-289,291-352,354-374-
weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last-
weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last
Instances:       12
Attributes:      5
                 Unemployment_rate
                 Hotels_/Motels_/Bungalows_AA
                 H_-_Hotels_and_restaurants
                 Income_per_cent_G.D.P.
                 Non-tax_income
Test mode:       10-fold cross-validation
```

=== Classifier model (full training set) ===

JRIP rules:
=====

```
(Hotels_/Motels_/Bungalows_AA = '(44.92-46.16]') => H_-_
Hotels_and_restaurants='(-inf-0.071]' (2.0/0.0)
=> H_-_Hotels_and_restaurants='(0.079-inf)' (10.0/2.0)
```

Number of Rules : 2

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	8	66.6667 %
Incorrectly Classified Instances	4	33.3333 %

Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα

```
Kappa statistic          0
Mean absolute error      0.0825
Root mean squared error  0.2151
Relative absolute error  60.8544 %
Root relative squared error 89.4155 %
Total Number of Instances 12
```

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0	0	0	0	0	0.625	'(-inf-0.071]'
0	0	0	0	0	?	'(0.071-0.072]'
0	0	0	0	0	?	'(0.072-0.073]'
0	0	0	0	0	?	'(0.073-0.074]'
0	0	0	0	0	?	'(0.074-0.075]'
0	0	0	0	0	?	'(0.075-0.076]'
0	0	0	0	0	?	'(0.076-0.077]'
0	0	0	0	0	?	'(0.077-0.078]'
0	0	0	0	0	?	'(0.078-0.079]'
1	1	0.667	1	0.8	0.625	'(0.079-inf)'

=== Confusion Matrix ===

```
a b c d e f g h i j  <-- classified as
0 0 0 0 0 0 0 0 0 4 | a = '(-inf-0.071] '
0 0 0 0 0 0 0 0 0 0 | b = '(0.071-0.072] '
0 0 0 0 0 0 0 0 0 0 | c = '(0.072-0.073] '
0 0 0 0 0 0 0 0 0 0 | d = '(0.073-0.074] '
0 0 0 0 0 0 0 0 0 0 | e = '(0.074-0.075] '
0 0 0 0 0 0 0 0 0 0 | f = '(0.075-0.076] '
0 0 0 0 0 0 0 0 0 0 | g = '(0.076-0.077] '
0 0 0 0 0 0 0 0 0 0 | h = '(0.077-0.078] '
0 0 0 0 0 0 0 0 0 0 | i = '(0.078-0.079] '
0 0 0 0 0 0 0 0 0 8 | j = '(0.079-inf) '
```