

Διαδικτυακό εκπαιδευτικό περιβάλλον για εφαρμογές της στατιστικής

Αλέξανδρος Καράκος

Αναπληρωτής Καθηγητής – Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Δημοκρίτειο Πανεπιστήμιο Θράκης
Ξάνθη - 67100
karakos@ee.duth.gr

ΠΕΡΙΛΗΨΗ

Στην εργασία αυτή παρουσιάζεται η δημιουργία ενός Περιβάλλοντος Διαδραστικών Εφαρμογών δηλαδή, μιας διαδικτυακής εφαρμογής η οποία παρέχει το κατάλληλο εκπαιδευτικό περιβάλλον για στοιχειώδεις στατιστικές επεξεργασίες και εφαρμογές πολυδιάστατης ανάλυσης δεδομένων. Μέχρι την προηγούμενη δεκαετία το περιβάλλον εργασίας για την επεξεργασία των δεδομένων ήταν κυρίως το DOS και το Unix. Το περιβάλλον αυτό βελτιώθηκε με νέες εκδόσεις των ιδίων προγραμμάτων για τα γραφικά περιβάλλοντα που παρέχουν πλέον τα λειτουργικά συστήματα Linux και Windows. Το νέο γραφικό περιβάλλον διασύνδεσης χρήστη-εφαρμογής είναι κατά πολύ ανώτερο και εύχρηστο, όμως οι εφαρμογές είναι προσανατολισμένες σε συγκεκριμένες πλατφόρμες. Τους περιορισμούς αυτούς έρχεται να λύσει η προτεινόμενη υλοποίηση, όπου πλέον η παρεχόμενη υπηρεσία είναι ανεξάρτητη από το λειτουργικό σύστημα, δεν χρειάζεται να διανεμηθεί και απλώς απαιτεί σύνδεση στο διαδίκτυο και ένα διαδικτυακό πλοηγό (Browser) νέας γενιάς. Η νέα υπηρεσία έχει στηριχθεί στο περιβάλλον .NET και έχουν χρησιμοποιηθεί οι γλώσσες C# και Javascript.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Διαδίκτυο, στατιστική, εκπαίδευση, .NET

ΕΙΣΑΓΩΓΗ

Η εργασία αυτή παρουσιάζει τη δημιουργία μιας διαδικτυακής εφαρμογής με την οποία παρέχονται δωρεάν υπηρεσίες στατιστικής επεξεργασίας και πολυδιάστατης ανάλυσης δεδομένων (Παπαδημητρίου, 1990). Είναι εξειδικευμένο και προσαρμοσμένο στις ειδικές ανάγκες της στατιστικής επεξεργασίας δεδομένων και απαλλαγμένο από την γενικότητα, πολυπλοκότητα και την δυσχρηστία προγραμμάτων τέτοιου είδους. Τέλος δόθηκε μεγάλη βαρύτητα στην συνεργασία του προγράμματος με άλλα προγράμματα έτσι ώστε να είναι δυνατή η ανταλλαγή δεδομένων σε μια ευρεία γκάμα από μορφές (format) αρχείων.

Το πρόγραμμα υποστηρίζει ένα φύλλο εργασίας όπου μπορούν να γίνουν οι συνήθεις εργασίες όπως αντιγραφή, διαγραφή, μετακίνηση περιεχομένων κελιών ή και ολόκληρων γραμμών ή στηλών και όλα αυτά μέσω εύχρηστων κουμπιών. Επίσης με απλά κουμπιά γίνονται όλες οι άλλες εργασίες όπως άνοιγμα και αποθήκευση, προσθήκη και αφαίρεση γραμμών, μετονομασία κ.ά. Από τα μενού επιλογών, υπάρχει η δυνατότητα για μια σειρά από περαιτέρω στατιστικές επεξεργασίες και αναλύσεις μέσω βοηθητικών παραθύρων που ανοίγουν για την κάθε εργασία και καθοδηγούν το χρήστη στην εισαγωγή κατάλληλων τιμών.

Έχει αναπτυχθεί με τις γλώσσες προγραμματισμού C++ (Harold 2002, Gunnerson 2001), και Javascript (Flannagan 1997, Goodman 2001), στο νέο περιβάλλον εργασίας .NET (Payne 2002, Douglas 2001, Esposito 2002) και βρίσκεται δωρεάν διαθέσιμο στο διαδίκτυο και στη διεύθυνση <http://peramos.ee.duth.gr/dataanalysis/webform1.aspx>.

Η εφαρμογή είναι προσανατολισμένη κυρίως στην εφαρμογή μεθόδων πολυδιάστατης ανάλυσης δεδομένων (Benzecri J.P. 1973 και 1979) και προς το παρόν, έχουν ενσωματωθεί η Παραγοντική Ανάλυση των Αντιστοιχιών - ΠΑΑ (Analyse Factorielle des Correspondences - AFC), η προς τα Πάνω Ιεραρχική Ταξινόμηση - ΠΙΤ (Classification Ascendante Hierarchique - CAH) και η προς τα Κάτω Ιεραρχική Ταξινόμηση - ΚΙΤ (Classification Descendante Hierarchique - CDH).

Η διαδικτυακή έκδοσή του μπορεί να βοηθήσει αποτελεσματικά τους ειδικούς αλλά και τους μη εξειδικευμένους στη στατιστική, να διαλέξουν τη σωστή διαδικασία και να φτάσουν γρήγορα και χωρίς λάθη στην ανάλυση των δεδομένων τους και στην εμφάνιση ή εκτύπωση των αποτελεσμάτων της επεξεργασίας. Έτσι, μπορεί να επεξεργαστεί τόσο ποιοτικές όσο και ποσοτικές μεταβλητές. Στις ποιοτικές μεταβλητές δίνει τη δυνατότητα στον χρήστη να ομογενοποιήσει τα δεδομένα του με πολλές επιλογές έτσι ώστε το τελικό σύνολο των δεδομένων που θα επεξεργαστούν να αποδίδει όσο γίνεται πιο πιστά την πραγματική εικόνα του φαινομένου που προσπαθεί να ερμηνεύσει (Tenenbaums, 1994). Στην παρούσα του έκδοση με τον αναλυτικό τρόπο βοήθειας και απλότητας στην επεξεργασία θα μπορούσε να αποτελέσει βασικό εργαλείο διδασκαλίας τόσο για εισαγωγικές έννοιες όσο και προχωρημένες εφαρμογές της ανάλυσης δεδομένων (Karakos, 1991).

ΠΑΡΟΥΣΙΑΣΗ ΤΟΥ ΠΡΟΓΡΑΜΜΑΤΟΣ

Όταν ο χρήστης πληκτρολογήσει τη διαδικτυακή διεύθυνση της υπηρεσίας, συνδέεται με τον εξυπηρετητή που την παρέχει και δημιουργείται ένα μοναδικό στιγμιότυπο της υπηρεσίας. Αφού επιτευχθεί η σύνδεση με τον εξυπηρετητή το βασικό παράθυρο της υπηρεσίας παρουσιάζεται στον πλοηγό του χρήστη.

Το αρχικό αυτό παράθυρο προτρέπει το χρήστη να δημιουργήσει ένα νέο πίνακα από την επιλογή «Αρχείο» του κεντρικού μενού. Στην περίπτωση που έχει αποθηκευμένα τα δεδομένα του τότε μπορεί από την ίδια επιλογή να ανοίξει το αρχείο στο οποίο βρίσκονται οι πληροφορίες. Ο εξ' ορισμού πίνακας δεδομένων αποτελείται από δέκα γραμμές και πέντε στήλες.

Θεωρούμε τη γενικότερη περίπτωση στην οποία ο χρήστης δημιουργεί ένα νέο πίνακα. Αφού λοιπόν δώσει τον αριθμό των γραμμών και στηλών του πίνακα που θέλει να δημιουργήσει τότε θα εμφανιστεί το βασικό παράθυρό της υπηρεσίας (εικόνα 1). Στην συγκεκριμένη εικόνα ο αριθμός των στηλών τέθηκε σε τέσσερις και ο αριθμός των γραμμών σε πέντε. Οι τιμές των κελιών θέτονται με την αρχική τιμή μηδέν.

Στην πρώτη γραμμή της ιστοσελίδας εμφανίζεται το κεντρικό μενού της υπηρεσίας με τις κύριες επιλογές : α) *Αρχείο*, β) *Επεξεργασία*, γ) *Δεδομένα*, δ) *Αναλύσεις*, ε) *Βοήθεια* στις οποίες και θα αναφερθούμε αργότερα. Επίσης, στην επόμενη γραμμή εμφανίζεται ένας μικρός πίνακας πληροφοριών που μας δηλώνει τον αριθμό των γραμμών και στηλών του κεντρικού πίνακα δεδομένων. Δίπλα από τον πίνακα πληροφοριών εμφανίζεται μία αναδιπλούμενη λίστα (Drop Down List), η λειτουργία της οποίας εξηγείται αργότερα.

Στην επόμενη γραμμή βρίσκεται η γραμμή εργαλείων, της οποίας κάθε εικονίδιο αντιστοιχεί στην αντίστοιχη λειτουργία κάποιας επιλογής από το κεντρικό μενού.

Τέλος παρουσιάζεται ο κεντρικός πίνακας δεδομένων, με τα κουμπάκια ελέγχου του, τις αναδιπλούμενες λίστες, τις ιδιότητες και τα ονόματα τόσο των μεταβλητών όσο και των αντικειμένων. Στο κάτω μέρος του παραθύρου βρίσκεται η γραμμή κατάστασης του πλοηγού και χρησιμοποιείται από την υπηρεσία για να μας δίνει ανά πάσα στιγμή, σύμφωνα με τη θέση που βρίσκεται ο δείκτης, σε ποια γραμμή και στήλη ανήκει το συγκεκριμένο κελί. Η λειτουργία αυτή είναι ιδιαίτερα χρήσιμη στις περιπτώσεις που έχουμε πλήθος γραμμών και στηλών.

	Ιδιότητες	Υ1	Υ2	Υ3	Υ4
	Ενεργή	Ενεργή	Ενεργή	Ενεργή	Ενεργή
X1	Ενεργή	0	0	0	0
X2	Ενεργή	0	0	0	0
X3	Ενεργή	0	0	0	0
X4	Ενεργή	0	0	0	0
X5	Ενεργή	0	0	0	0

Εικόνα 1: Το βασικό παράθυρο της υπηρεσίας

Ο κεντρικός πίνακας δεδομένων

Ο κεντρικός πίνακας δεδομένων είναι ένα πλέγμα στο οποίο έχουν προστεθεί όσο το δυνατόν περισσότερες πληροφορίες και στοιχεία ελέγχου ώστε ο χρήστης να μπορεί με εύκολο τρόπο να εξετάσει και να τροποποιήσει τα δεδομένα που βρίσκονται στα κελιά, την ιδιότητα και την ονομασία κάθε μεταβλητής και κάθε αντικειμένου.

Όταν μια γραμμή βρίσκεται σε κατάσταση «Edit Mode» δηλ μπορούμε να αλλάξουμε το περιεχόμενο των κελιών της τότε πρέπει να δοθεί προσοχή στα δεδομένα που εισαγάγουμε στα κελιά ώστε εάν είναι πραγματικοί αριθμοί τότε ο διαχωρισμός τους γίνεται με το ελληνικό κόμμα « , ». Σε περίπτωση που ο χρήστης πληκτρολογήσει τελεία « . » τότε δεν θα ληφθεί υπόψη ως το διαχωριστικό του ακεραίου και του δεκαδικού μέρους του αριθμού. Αν γίνει κάποιο λάθος και εισάγει ο χρήστης αλφαριθμητικούς χαρακτήρες τότε μετά την επικύρωση των αλλαγών θα

εμφανιστεί μήνυμα λάθους που θα ενημερώνει το χρήστη ότι βρέθηκε ακατάλληλη τιμή και αυτή τέθηκε από την υπηρεσία σε -999. Τότε ο χρήστης μπορεί να χρησιμοποιήσει είτε την εύρεση του πλοηγού είτε με τις μπάρες ολίσθησης για να εντοπίσει το συγκεκριμένο κελί και αν θέλει μπορεί να προβεί σε διορθώσεις. Η τιμή -999 επιλέχθηκε επειδή σπάνια υπάρχουν τέτοιες τιμές στα δεδομένα των αναλύσεων τα οποία επεξεργάζεται η υπηρεσία.

Το κεντρικό μενού

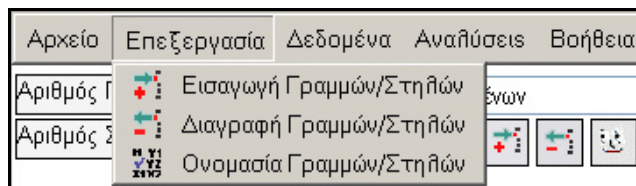
Το κεντρικό μενού της υπηρεσίας αποτελείται από τις κύριες επιλογές : α) *Αρχείο*, β) *Επεξεργασία*, γ) *Δεδομένα*, δ) *Αναλύσεις*, ε) *Βοήθεια* οι οποίες μέσω των υποκατηγοριών τους παρέχουν στον χρήστη τις δυνατότητες της υπηρεσίας. Στις περισσότερες από τις υποκατηγορίες υπάρχουν εικονίδια, παρόμοια με τις εφαρμογές του Microsoft Office, τα οποία αντιστοιχούν σε κουμπάκια που βρίσκονται στη γραμμή εργαλείων.

Στην επιλογή «Αρχείο» βρίσκονται όλες οι βασικές εντολές διαχείρισης των αρχείων όπως **Νέο αρχείο δεδομένων, Άνοιγμα αρχείου δεδομένων, Αποθήκευση αρχείου δεδομένων κτλ.**

Οι τύποι αρχείων που υποστηρίζονται είναι οι πιο διαδεδομένοι στην ανταλλαγή του είδους των δεδομένων που επεξεργάζεται η υπηρεσία. Υποστηρίζονται αρχεία του τύπου:

CSV (Comma Separated Values – Τιμές διαχωρισμένες με κόμμα), CSV (Τιμές διαχωρισμένες με ελληνικό ερωτηματικό), TXT – Κείμενο με χαρακτήρα διαχωρισμού το [Tab], TXT – Κείμενο με χαρακτήρα διαχωρισμού οριζόμενο από τον χρήστη

Η επιλογή «Επεξεργασία»

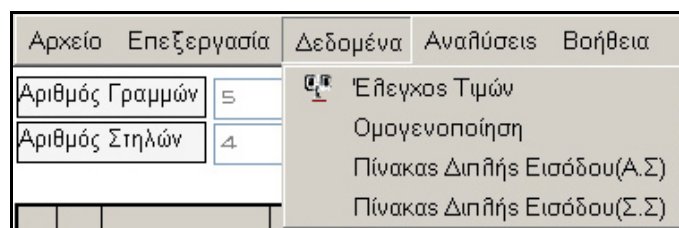


Εικόνα 2 : Οι δυνατότητες της επιλογής επεξεργασία

Στην επιλογή «Επεξεργασία» βρίσκονται οι βασικές εντολές διαχείρισης των στηλών και των γραμμών του κεντρικού πίνακα δεδομένων όπως φαίνεται και στην εικόνα 2.

Η επιλογή «Δεδομένα»

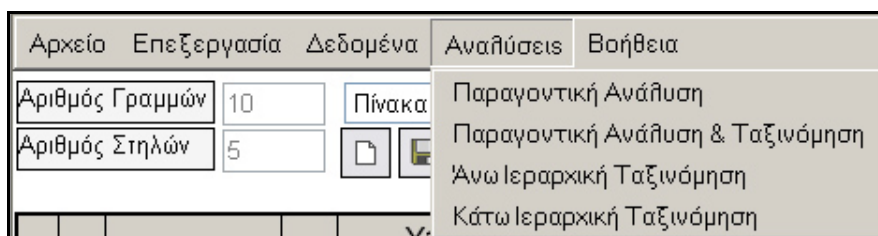
Στην επιλογή «Δεδομένα» (εικόνα 3) βρίσκονται οι βασικές εντολές για τον έλεγχο, την επεξεργασία του κεντρικού πίνακα δεδομένων και για τη δημιουργία των πινάκων διπλής εισόδου, αφού πρώτα προβούμε σε ομογενοποίηση των τιμών του κεντρικού πίνακα δεδομένων.



Εικόνα 3 : Οι δυνατότητες της επιλογής δεδομένα

Η επιλογή «Αναλύσεις»

Στην επιλογή «Αναλύσεις» (εικόνα 4) βρίσκονται οι βασικές εντολές για την εφαρμογή των τριών αναλύσεων που μπορεί να εφαρμοστούν σε οποιοδήποτε από τα τέσσερα είδη πινάκων (Κεντρικό Πίνακα Δεδομένων, Ομογενοποιημένο Πίνακα, Πίνακα Διπλής Εισόδου Α. Σ και Πίνακα Διπλής Εισόδου Σ. Σ) που υποστηρίζονται από την υπηρεσία. Οι τρεις αυτές αναλύσεις, Παραγοντική Ανάλυση των Αντιστοιχιών, Άνω Ιεραρχική Ταξινόμηση και Κάτω Ιεραρχική Ταξινόμηση δεν θα αναπτυχθούν εδώ διότι, η κάθε μια αποτελεί ένα ολόκληρο κεφάλαιο το οποίο είναι πέραν από τις ανάγκες αυτής της παρουσίασης.



Εικόνα 4: Οι βασικές αναλύσεις που επιτρέπει το πρόγραμμα

Η επιλογή «Βοήθεια»

Η επιλογή «Βοήθεια» έχει δύο βασικές υπέρ-συνδέσεις με σκοπό την πληροφόρηση του χρήστη για την υπηρεσία, την προσπάθεια επίλυσης προβλημάτων που αντιμετωπίζει σε σχέση με τη χρήση της, όπως και τρόπους επικοινωνίας με τους υπεύθυνους της υπηρεσίας.

1. Βοήθεια για την Υπηρεσία

Με την επιλογή αυτή ο πλοηγός του χρήστη θα ανοίξει μια νέα ιστοσελίδα που περιέχει βοήθεια και χρήσιμες οδηγίες που αφορούν τη χρήση και τις δυνατότητες της υπηρεσίας.

2. Επικοινωνία

Με την επιλογή αυτή ο πλοηγός του χρήστη θα ανοίξει μια νέα ιστοσελίδα (εικόνα 10) που περιέχει χρήσιμες ηλεκτρονικές διευθύνσεις στις οποίες μπορεί να αποταθεί ο χρήστης για προβλήματα, βοήθεια, υποστήριξη και ότι άλλο έχει σχέση με την υπηρεσία.

ΣΥΜΠΕΡΑΣΜΑΤΑ

Η υπηρεσία δεν έχει δοκιμαστεί σε πραγματικές συνθήκες εξυπηρέτησης πολλών χρηστών στις οποίες στιγμιαία ο εξυπηρετητής θα πρέπει να ανταποκριθεί στις απαιτήσεις τους στο ελάχιστο χρονικό διάστημα. Η σταθερότητα, η κατανάλωση των υπολογιστικών πόρων και η εξυπηρέτηση των χρηστών στο συγκεκριμένο μηχανήμα που έχει εγκατασταθεί η υπηρεσία δεν έχουν διερευνηθεί προς το παρόν δηλαδή, αν κατά την ταυτόχρονη εξυπηρέτηση πολλών χρηστών, ο κατακερματισμός και η κατανάλωση της μνήμης του μηχανήματος οδηγεί σε μειωμένη απόδοση.

Πιστεύω ότι η εφαρμογή αυτή παρέχει το κατάλληλο εκπαιδευτικό περιβάλλον για στοιχειώδεις στατιστικές επεξεργασίες και εφαρμογές πολυδιάστατης ανάλυσης δεδομένων, και η ιδιαίτερη εκπαιδευτική αξία του για αξιοποίησή στην εκπαίδευση αφορά την απλότητα στη χρήση του και τη δυνατότητα χρήσης του μέσω του Διαδικτύου.

Ελπίζω τέλος, η προσπάθεια αυτή θα βρει μιμητές για την υλοποίηση παρόμοιων εφαρμογών που θα δώσουν τη δυνατότητα να μπορούν οι χρήστες του διαδικτύου να εκτελούν εφαρμογές με ασφάλεια και οι οποίες θα είναι ανεξάρτητες λειτουργικού συστήματος, θα μπορούν να αναβαθμίζονται χωρίς αναδιανομή του κώδικα και θα παρέχουν τη μέγιστη δυνατή βοήθεια άμεσα.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Benzecri J.P. (1973), *L'analyse des Données-2. Analyse des correspondances*. Dunod.
- Benzecri J.P. (1973), *L'analyse des Données-1. La taxinomie*. Dunod.
- Benzecri J.P., Lebeau M., Jambu M. (1979), Aide a l'interprétation en classification Automatique , *Les cahiers de l'analyse des données* Vol V, no 1 pp 101-123
- Douglas J. Reilly (2001), *Designing Microsoft ASP.NET Applications*, Microsoft Press, ISBN: 0735613486
- Esposito Dino (2002), *Building Web Solutions with ASP.NET and ADO.NET*, Microsoft Press, ISBN: 0-7356-1578-0
- Flannagan David (1997), *JavaScript The Definite Guide*, O'Reilly, ISBN: 1-56592-235-2
- Goodman Danny (2001), *JavaScript Bible Gold Edition*, Hungry Minds, ISBN: 0-7645-4718-6
- Gunnerson Eric (2001), *A Programmer's Introduction to C#*, Apress, ISBN: 1893115860
- Harold Davis (2002), *Visual C#.NET Programming*, SYBEX Inc, ISBN: 0-7821-4046-7
- Karakos Alex. (1991), Logiciel Conversationnel PRAXITELIS . *Symbolic – Numeric Data Analysis and Learning*, Paris-France, pp 18-20
- Payne Chris (2002) *ASP.NET Web Developer 's Guide*, Syngress Publishing Inc, ISBN: 1-928994-51-2
- Tenenbaus M. (1994), *Méthodes Statistiques en Gestion*. Dunod
- Multivariate Statistics: An Introduction*,
<http://trochim.human.cornell.edu/tutorial/flynn/mutlivar.htm>
- StatSoft's Correspondence Analysis*, <http://www.statsoftinc.com/textbook/stcoran.html>
- Παπαδημητρίου Ιωάννης. (1990), *Μέθοδοι ανάλυσης δεδομένων*. Θεσ/νικη

