

Ανάλυση Δεδομένων με χρήση της Γλώσσας R

Δομή Εκπαιδευτικού Προγράμματος ανά θεματική ενότητα

Ενότητα 1η

Τίτλος: Εισαγωγή στην R

Σκοπός- Στόχοι: Στην ενότητα “Εισαγωγή στην R”, θα μάθετε τα βασικά εργαλεία αυτής της ευρέως χρησιμοποιούμενης γλώσσας ανοιχτού κώδικα. Θα μάθετε να διαχειρίζεστε αντικείμενα δεδομένων όπως διανύσματα, παράγοντες (factors), πίνακες, πλαίσια δεδομένων και λίστες. Τέλος, θα δείτε πως μπορείτε να φορτώσετε και να διαχειριστείτε πακέτα της R.

Υποενότητες:

1. Εισαγωγή στις βασικές έννοιες της R
2. Διανύσματα
3. Πίνακες
4. Παράγοντες
5. Πλαίσια Δεδομένων
6. Λίστες
7. Πακέτα R
8. Εργαστηριακές Ασκήσεις

Ώρες: 10

Τρόπος αξιολόγησης των αποκτηθεισών γνώσεων και δεξιοτήτων: Εργαστηριακές ασκήσεις / Quiz & τελική εργασία

Η διδασκαλία της θεματικής ενότητας απαιτεί αίθουσα πληροφορικής;: ΝΑΙ

Ενότητα 2η

Τίτλος: Εισαγωγή Δεδομένων στην R

Σκοπός- Στόχοι: Η εισαγωγή δεδομένων στο R θα έπρεπε να είναι το ευκολότερο βήμα στην διαδικασία της ανάλυσης. Δυστυχώς, όμως τα δεδομένα μπορεί να διατίθενται σε διάφορες μορφές όπως για παράδειγμα σε αρχεία .csv, .txt, spss, sas, dta, κτλ, με αποτέλεσμα να μην είναι πάντα εύκολη η εισαγωγή τους στην ανάλυση. Σε αυτή την ενότητα, θα ξεκινήσετε μαθαίνοντας πώς να εισάγετε αρχεία .csv, txt με εντολές της R και στη συνέχεια, θα διερευνήσετε τα πακέτα readr και data.table τα οποία αποτελούν μια εναλλακτική και πιο αποτελεσματική εισαγωγή δεδομένων τέτοιου είδους αρχείων. Ακολούθως, θα μάθετε τον τρόπο με τον οποίο γίνεται η εισαγωγή αρχείων .xls στην R χρησιμοποιώντας το πακέτο readxl. Τέλος θα μάθετε πώς να εισάγετε δεδομένα από τον ιστό καθώς και από πακέτα στατιστικών λογισμικών SPSS, SAS και STATA .

Υποενότητες:

1. Εισαγωγή δεδομένων με το πακέτο `utils`
2. `readr` και `data.table`
3. Εισαγωγή δεδομένων από Excel
4. Εισαγωγή δεδομένων από τον ιστό
5. Εισαγωγή δεδομένων από βάσεις δεδομένων
6. Εισαγωγή δεδομένων από πακέτα στατιστικών λογισμικών
7. Εισαγωγή δεδομένων – R Studio
8. Εργαστηριακές Ασκήσεις

Ώρες: 16

Τρόπος αξιολόγησης των αποκτηθεισών γνώσεων και δεξιοτήτων: Εργαστηριακές ασκήσεις / Quiz & τελική εργασία

Η διδασκαλία της θεματικής ενότητας απαιτεί αίθουσα πληροφορικής;: ΝΑΙ

Ενότητα 3η

Τίτλος: Διαχείριση Δεδομένων `dplyr`

Σκοπός- Στόχοι: Στην ενότητα αυτή θα δείτε πως μπορείτε να χρησιμοποιήσετε εντολές του πακέτου `dplyr` για να διαχειριστείτε τα σύνολα δεδομένων σας. Θα δουλέψετε για πρώτη φορά με τον τελεστή `pipe` (`%>%`) καθώς και με τη δομή δεδομένων `tibble`. Αρχικά θα μάθετε πως μπορείτε να επιλέγετε συγκεκριμένες μεταβλητές, να φιλτράρετε τις παρατηρήσεις σας και να τις αναδιατάσετε. Επίσης, θα μάθετε πως να δημιουργείτε νέες μεταβλητές ως συνάρτηση των ήδη υπάρχουσών μεταβλητών, καθώς και πως να μετονομάσετε τις μεταβλητές σας. Ακόμη, θα μάθετε πως να ομαδοποιείται τα δεδομένα σας αλλά και πως να συνοψίζετε τις αρχικές παρατηρήσεις βάσει μιας συνάρτησης. Τέλος, θα δείτε πως μπορείτε να ενώσετε σε έναν πίνακα δεδομένα από δύο διαφορετικά σύνολα δεδομένων.

Υποενότητες:

1. Εισαγωγή
2. Επιλογή και Μετατροπή Δεδομένων
3. Συγκέντρωση Δεδομένων
4. Ένωση Συνόλων Δεδομένων
5. Εργαστηριακές Ασκήσεις

Ώρες: 12

Τρόπος αξιολόγησης των αποκτηθεισών γνώσεων και δεξιοτήτων: Εργαστηριακές ασκήσεις / Quiz & τελική εργασία

Η διδασκαλία της θεματικής ενότητας απαιτεί αίθουσα πληροφορικής;: ΝΑΙ

Ενότητα 4η

Τίτλος: Οπτικοποίηση Δεδομένων

Σκοπός- Στόχοι: Στην ενότητα αυτή θα μάθετε πως να παράγετε γραφήματα με το πακέτο της R, ggplot2. Αρχικά, θα σας παρουσιαστεί η έννοια της γραμματικής των γραφικών, το πλαίσιο σχεδιασμού στο οποίο χτίστηκε η ggplot2 και στηρίζεται στην παραγωγή γραφημάτων με την προσθήκη επιπέδων. Ακολούθως θα σας γίνει παρουσίαση όλων των επιπέδων. Πρώτα θα δείτε για τα δεδομένα και την αισθητική, ενώ μετά για τις γεωμετρίες, τα στατιστικά, το σύστημα συντεταγμένων, τα γραφήματα σε πάνελ (facets) και τέλος για τα θέματα (themes). Μέσα από την ενότητα αυτή θα δείτε πως να σχεδιάζεται διαγράμματα διασποράς, ραβδογράμματα, ιστογράμματα, διαγράμματα γραμμής καθώς και θηκογράμματα.

Υποενότητες:

1. Εισαγωγή
2. Δεδομένα
3. Aesthetics, Positions and Scales
4. Geometries
5. Statistics
6. Coordinates and Facets
7. Themes
8. Εργαστηριακές Ασκήσεις

Ώρες: 18

Τρόπος αξιολόγησης των αποκτηθεισών γνώσεων και δεξιοτήτων: Εργαστηριακές ασκήσεις / Quiz

&

τελική εργασία

Η διδασκαλία της θεματικής ενότητας απαιτεί αίθουσα πληροφορικής;: ΝΑΙ

Ενότητα 5η

Τίτλος: Διερευνητική Ανάλυση Δεδομένων

Σκοπός- Στόχοι: Η διερευνητική επισκόπηση των δεδομένων είναι το πρώτο βήμα στην ανάλυση δεδομένων. Εκτός από το μέγεθος και το πλήθος των μεταβλητών των δεδομένων, σε αυτή την ενότητα, θα μάθετε τεχνικές οι οποίες αποκαλύπτουν τη δομή των δεδομένων σας, αναζητούν τις μεταβλητές που ενδεχομένως να συσχετίζονται με κάποιες άλλες και εντοπίζουν αν υπάρχουν, τις παρατηρήσεις που είναι «ασυνήθιστες» (παράτυπα σημεία).

Υποενότητες:

1. Εξερεύνηση κατηγορικών δεδομένων
2. Εξερευνώντας Αριθμητικά Δεδομένα
3. Στατιστικά Μέτρα

4. Παράδειγμα
5. Εργαστηριακές Ασκήσεις

Ώρες: 12

Τρόπος αξιολόγησης των αποκτηθεισών γνώσεων και δεξιοτήτων: Εργαστηριακές ασκήσεις / Quiz & τελική εργασία

Η διδασκαλία της θεματικής ενότητας απαιτεί αίθουσα πληροφορικής;: ΝΑΙ

Ενότητα 6η

Τίτλος: Στατιστική Συμπερασματολογία

Σκοπός- Στόχοι: Σε αυτή την ενότητα θα μάθετε τους βασικούς ελέγχους της στατιστικής συμπερασματολογίας. Συγκεκριμένα θα μάθετε να χρησιμοποιείτε παραμετρικούς και μη παραμετρικούς στατιστικούς ελέγχους για ένα δείγμα, για δύο ανεξάρτητα ή εξαρτημένα δείγματα, τους ελέγχους κανονικότητας Kolmogorov-Smirnov, Shapiro-Wilk και τους οπτικούς ελέγχους qqplots καθώς επίσης και τον έλεγχο ανεξαρτησίας χ^2 .

Υποενότητες:

1. Συμπερασματολογία για ένα δείγμα
2. Συμπερασματολογία για δύο δείγματα
3. Έλεγχοι χ^2
4. Εργαστηριακές Ασκήσεις

Ώρες: 12

Τρόπος αξιολόγησης των αποκτηθεισών γνώσεων και δεξιοτήτων: Εργαστηριακές ασκήσεις / Quiz & τελική εργασία

Η διδασκαλία της θεματικής ενότητας απαιτεί αίθουσα πληροφορικής;: ΝΑΙ
