

Διαχείριση, ανάλυση, και οπτικοποίηση περιβαλλοντικών δεδομένων με R.

1. Γενικά για την R

Θεωρία

- Σύντομο ιστορικό της R
- Βασικά χαρακτηριστικά της R (πλεονεκτήματα, ιδιαιτερότητες, περιορισμοί)

Εργαστήριο

- Εισαγωγή στο περιβάλλον της R
- Εγκατάσταση R και RSTUDIO
- Γνωριμία με το περιβάλλον RSTUDIO

2. Αντικείμενα στην R

Θεωρία

- Κατηγορίες αντικειμένων
- Χρήση και αξιοποίηση των αντικειμένων
- Τύπος και κλάσεις των αντικειμένων

Εργαστήριο

- Διερεύνηση αντικειμένων
- Μετατροπές αντικειμένων
- Αξιοποίηση των αντικειμένων σε πραγματικά δεδομένα

3. Διανύσματα και τελεστές στην R

Θεωρία

- Τελεστές
- Αριθμητικές συναρτήσεις
- Διανύσματα

Εργαστήριο

- Πράξεις με τελεστές και διανύσματα
- Έκφραση αριθμητικών συναρτήσεων σε R
- Δόμηση script επίλυσης προβλήματος με αριθμητική συνάρτηση

4. Πίνακες, πλαίσια δεδομένων και λίστες

Θεωρία

- Δισδιάστατοι/ πολυδιάστατοι πίνακες
- Τύποι διανυσμάτων δεδομένων
- Πλαίσιο δεδομένων και λίστες

Εργαστήριο

- Εισαγωγή δεδομένων
- Μετατροπές μεταξύ διανυσμάτων/ πλαισίων και λιστών
- Εξαγωγή δεδομένων

5. Γνωριμία με την έννοια του project σε R

Θεωρία

- Ανάλυση του project και του οφέλους του να εργάζεται κανείς σε τέτοιο πλαίσιο
- Τα μέρη και οι περιορισμοί του project

Εργαστήριο

- Δημιουργία και εργασία πραγματικού project

6. Λήψη και διαχείριση δεδομένων από απομακρυσμένες βάσεις με R

Θεωρία

- Γνωριμία με μεγάλες βάσεις δεδομένων (big data)
- Διαδικασία και περιορισμοί λήψης δεδομένων από απομακρυσμένες βάσεις
- Διαδικασία διερεύνησης των δεδομένων

Εργαστήριο

- Παράδειγμα λήψης δεδομένων από αποκρουσμένη βάση
- Διερεύνηση των δεδομένων λήψης
- Επίλυση προβλημάτων διαχείρισης των δεδομένων

7. Εξειδικευμένα πακέτα, σχόλια και ονοματοδοσία στην R

Θεωρία

- Τι είναι τα πακέτα της R
- Τα σχόλια και η σημασία του στην συγγραφή του script
- Καλές πρακτικές σχετικά με τα σχόλια και τα ονόματα των αντικειμένων στην R

Εργαστήριο

- Λήψη και εγκατάσταση πακέτων
- Συγγραφή κώδικα με σχόλια και αξιοποίηση τους
- Ολοκληρωμένη συγγραφή script με έμφαση στην καλή οργάνωσή του για διεπιστημονική επικοινωνία

8. Επεξεργασία και ανάλυση δεδομένων με χρήση dplyr

Θεωρία

- Οι βασικές νέες εντολές και συμβολισμούς του dplyr
- Διαδικασία εργασίας με dplyr
- Σύγκριση με άλλες βασικές προσεγγίσεις

Εργαστήριο

- Εκτέλεση ανάλυσης πραγματικών δεδομένων με dplyr
- Εξαγωγή αποτελεσμάτων σε μορφές πίνακα
- Δημιουργία γραφημάτων με τις βασικές λειτουργίες της R

9. Οπτικοποίηση ανάλυσης δεδομένων με χρήση ggplot2

Θεωρία

- Βασική δομή των εντολών του ggplot2
- Γραμματική του ggplot2

Εργαστήριο

- Δημιουργία γραφικών από δεδομένα πραγματικής ανάλυσης
- Δημιουργία γραφικών με πολλαπλά πάνελς
- Εξαγωγή γραφημάτων σε διάφορες διαστάσεις, αναλύσεις και μορφότυπους

10. Βασική στατιστική ανάλυση δεδομένων με R

Θεωρία

- Οι βασικές εντολές εκτέλεσης στατιστικής ανάλυσης
- Εντολές εξαγωγής αποτελεσμάτων περιγραφικής στατιστικής
- Διαδικασία εξαγωγής δεδομένων σε ποικιλία μορφών

Εργαστήριο

- Εκτέλεση ανάλυσης περιγραφικής στατιστικής
- Εκτέλεση ανάλυσης παλινδρόμησης
- Μορφοποίηση και εξαγωγή αποτελεσμάτων ανάλυσης

11. Δημιουργία συναρτήσεων σε R

Θεωρία

- Εισαγωγή στις συναρτήσεις της R
- Βασικά χαρακτηριστικά και δομή συναρτήσεων

Εργαστήριο

- Επίλυση προβλήματος με χρήση συνάρτησης που δημιουργεί ο χρήστης

12. Έλεγχος εκδόσεων κώδικα

Θεωρία

- Εισαγωγή στον έλεγχο εκδόσεων κώδικα
- Γνωριμία με τα βασικά εργαλεία ελέγχου κώδικα

Εργαστήριο

- Εγκατάσταση git
- Λειτουργία git

13. Καλές πρακτικές κατά τη συγγραφή και διόρθωση κώδικα σε R

Θεωρία

- Βασικές δομές κώδικα
- Διαδικασίες αναζήτησης και επιδιόρθωσης σφαλμάτων στον κώδικα
- Γνωριμία με τα βασικά αποθετήρια γνώσης για επίλυση σφαλμάτων κώδικα

Εργαστήριο

- Εκτέλεση κώδικα και αναζήτηση σφαλμάτων
- Αναζήτηση και επιδιόρθωση σφαλμάτων
- Πλοήγηση και αναζήτηση λύσεων στα αποθετήρια γνώσης

Μαθησιακοί στόχοι

Με την επιτυχή ολοκλήρωση του μαθήματος, ο φοιτητής/τρια θα μπορεί να :

- Έχει γνώση του περιβάλλοντος εργασίας της γλώσσας R
- Γνωρίζει να εγκαθιστά και να λειτουργεί την R στον προσωπικό υπολογιστή
- Εισάγει δικά του δεδομένα τοπικά ή από απομακρυσμένες βάσεις
- Τροποποιεί τα δεδομένα ώστε να έχουν την κατάλληλη μορφή για περαιτέρω ανάλυση
- Κάνει βασική ανάλυση των δεδομένων (περιγραφική στατιστική, παλινδρόμηση κ.α.)
- Ολοκληρώνει οπτικοποίηση των δεδομένων και των αποτελεσμάτων της ανάλυσης
- Κάνει εξαγωγή και διαμοιρασμός των δεδομένων και των οπτικοποιημένων αποτελεσμάτων
- Λειτουργεί αποτελεσματικά το πλαίσιο ελέγχου εκδόσεων