
Introduction to R for Social Sciences

Μεταφόρτωση και Εγκατάσταση της R και του Rstudio
Ανακαλύπτοντας την R

Αναστάσιος Εμβλωτής & Αικατερίνη Σαργιώτη

Εγκατάσταση R και R Studio

- Εγκατάσταση R

<https://cran.r-project.org/> → [Download R for Windows](#)

- Εγκατάσταση R Studio

- Χρειάζεται αρχικά εγκατάσταση της R και έπειτα εγκατάσταση του R Studio

<https://www.rstudio.com/products/rstudio/download/> → επιλογή του installer για windows

Installers	Size	Date	MD5
RStudio 1.2.5001 - Windows 10/8/7 (64-bit)	149.83 MB	2019-09-19	c54d8779f363ec9636c7831e577521bd

Εγκατάσταση πακέτων και βιβλιοθηκών

Εγκατάσταση πακέτων

```
install.packages("test")
```

Φόρτωση βιβλιοθήκης

```
library(test)
```


Σύνταξη εντολών

- Εντολές σε μορφή **εκφράσεων** (expressions) ή **εκχωρήσεων** (assignments)
 - Εκφράσεις: Άμεσος υπολογισμός και εμφάνιση στην οθόνη
 - Εκχωρήσεις: Υπολογισμός έκφρασης καταχωρώντας το αποτέλεσμα σε μια **μεταβλητή** (αντικείμενο)
- Εμφάνιση αποτελέσματος μιας εκχώρησης με την εντολή `print()`
- Εισαγωγή σχολίων με τη χρήση του συμβόλου `#`

Βασικές πράξεις στην R

Τελεστής	Ερμηνεία
+	Πρόσθεση
-	Αφαίρεση
*	Πολλαπλασιασμός
/	Διαίρεση
==	Ισότητα
>	Μεγαλύτερο από
<	Μικρότερο από
>=	Μεγαλύτερο ή ίσο από
<=	Μικρότερο ή ίσο από
!=	Ανισότητα
^	Ύψωση σε δύναμη
= ή <-	Ανάθεση και <u>όχι</u> ισότητα

Κλάσεις αντικειμένων

- Αντικείμενο (object) = Μεταβλητή της R
- Κλάση = Τύπος αντικειμένων – Καθορίζεται από την τιμή του αντικειμένου
- Η κλάση εμφανίζεται με την εντολή `class()` ή `mode()`
- Βασικές κλάσεις της R:
 - Χαρακτήρας (character): $x = \text{"test"}$
 - Αριθμητικός (numeric): $x = 3.5$
 - Ακέραιος (integer): $x = 5L$
 - Σύνθετος (complex): $x = 8 + 5L$
 - Λογικός (logical): $x = \text{TRUE}$ ή $x = \text{F}$

Βασικές δομές αντικειμένων (1)

- Διανύσματα (vectors)
 - Δημιουργία με τη συνάρτηση `c` [π.χ. `x = c(a, b, c, ...)`]
 - Αριθμητικά
 - Χαρακτήρων
 - Λογικά
 - Κατηγοριών
- Πίνακες (matrices)
 - Δισδιάστατα διανύσματα
 - Γραμμές και στήλες πρέπει να έχουν αντικείμενα ίδιας κλάσης
 - Εισαγωγή πίνακα με την εντολή `matrix(c(), nrow=, ncol=)`
 - Η κατασκευή του πίνακα ξεκινάει συμπληρώνοντας με τη σειρά τα αντικείμενα που ορίζουν το διάνυσμα στην κάθε στήλη
 - Το όρισμα `byrow=T` δηλώνει ότι η συμπλήρωση του πίνακα θα γίνει συμπληρώνοντας τις γραμμές του
 - Κατασκευή πίνακα με τη συνάρτηση `dim()`
 - Κατασκευή πίνακα με ένωση διανυσμάτων [κατά γραμμές χρήση της εντολής `rbind()` και κατά στήλες της `cbind()`]

Βασικές δομές αντικειμένων (2)

- Πολλαπλοί – πολυδιάστατοι πίνακες (arrays)
 - Πίνακες με περισσότερες από δύο διαστάσεις
 - Κατασκευή πολυδιάστατων πινάκων με την εντολή `array()`
 - Δήλωση μεγέθους της κάθε διάστασης με την εντολή `dim()`
- Λίστες (lists)
 - Διανύσματα που μπορούν να περιέχουν διαφορετικούς τύπους αντικειμένων
 - Αποτελούνται από στήλες, αλλά δεν είναι υποχρεωτικό να έχουν το ίδιο μήκος
 - Δημιουργία λίστας με την εντολή `list()`
 - Δύο ή περισσότερες λίστες μπορούν να ενωθούν με τη συνάρτηση `c` [π.χ. `list = c(list1, list2)`]
 - Το μήκος της λίστας δίνεται με την εντολή `length()`

Βασικές δομές αντικειμένων (3)

- Πλαίσια δεδομένων (data frames)
 - Χρησιμοποιούνται για την αποθήκευση δεδομένων
 - Έχουν παρόμοια μορφή με τους πίνακες και είναι δισδιάστατα
 - Διαφορετικός τύπος δεδομένων σε κάθε στήλη, αλλά κάθε στήλη έχει αντικείμενα ίδιας κλάσης
 - Εισαγωγή πλαισίων δεδομένων με την εντολή `data.frame()`
 - Προσθήκη στηλών και γραμμών με τις εντολές `cbind()` και `rbind()`, αντίστοιχα
 - Διάταξη πλαισίων δεδομένων ως προς στήλη με την εντολή `order()`
 - Ένωση δύο πλαισίων δεδομένων με την εντολή `merge()`

Βασικές δομές αντικειμένων (4)

- Παράγοντες (factors)
 - Διευκολύνουν την αναπαράσταση και διαχείριση κατηγορικών μεταβλητών
 - Διαθέτουν επίπεδα (levels) που αναπαριστούν τις τιμές της μεταβλητής
 - Εισάγονται με την εντολή `factor()`

Ελλιπείς τιμές

- Συμβολίζονται ως **NA** ή **NaN**
 - NA → Αριθμητική κλάση
 - NaN → Λογική κλάση
- Έλεγχος ελλιπών τιμών με τις εντολές *is.na()* και *is.nan()*