



Πανεπιστήμιο Αιγαίου

# Εισαγωγή στον Προγραμματισμό Η/Υ

## Ενότητα 6: Δομές Δεδομένων

Μιχάλης Βαΐτης  
Τμήμα Γεωγραφίας



Ευρωπαϊκή Ένωση  
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ  
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

# Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



# Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αιγαίου**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση  
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ  
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



# Σκοποί ενότητας

- Να γνωρίζετε την έννοια των δομών δεδομένων (data structures) σε μια γλώσσα προγραμματισμού.
- Να γνωρίζετε τον ορισμό πινάκων (matrix) στη γλώσσα προγραμματισμού R.
- Να γνωρίζετε τον ορισμό λίστας (list) στη γλώσσα προγραμματισμού R.
- Να γνωρίζετε τον ορισμό πλαισίων δεδομένων (data frame) στη γλώσσα προγραμματισμού R.
- Να μπορείτε να χρησιμοποιείτε αποτελεσματικά τις δομές δεδομένων της γλώσσας R για τη διαχείριση δεδομένων και την επίλυση υπολογιστικών προβλημάτων.

# **Δομές δεδομένων**

# Δομές δεδομένων στην R

- Vector (διάνυσμα μίας διάστασης)
- Matrix (πίνακας δύο διαστάσεων)
- Factor
- List (λίστα – διατεταγμένη συλλογή)
- Data frame (πλαίσιο δεδομένων)
- Array (πίνακας πολλών διαστάσεων)

# Vector

Αποτελεί διατεταγμένο σύνολο τιμών μίας διάστασης και ενός τύπου δεδομένων.

Κατασκευάζεται με τη συνάρτηση `c()`

(από τη λέξη `constructor`).

Παραδείγματα

```
st = c("aa", "bb", "cc", "dd", "ee")
```

```
x = c(74, 122, 235, 111, 292)
```

# Vector

Για να αναφερθούμε σε μια τιμή του διανύσματος, χρησιμοποιούμε το όνομα της μεταβλητής, τις τετράγωνες αγκύλες και τον δείκτη της τιμής – δηλ. τον αριθμό που εκφράζει τη θέση της τιμής στο διάνυσμα.

Παράδειγμα

```
st = c("aa", "bb", "cc", "dd", "ee")  
print(st[2])  
> "bb"
```



# Matrix

Πίνακας δύο διαστάσεων τιμών του ίδιου τύπου δεδομένων.

Κατασκευάζεται με τη συνάρτηση `matrix()`.

Παράμετροι της συνάρτησης είναι: ένα διάνυσμα με τις τιμές, το πλήθος των γραμμών και των στηλών του πίνακα, καθώς και ένδειξη αν οι τιμές θα σχηματίσουν πρώτα στήλες ή πρώτα γραμμές.

# Matrix

Παράδειγμα

```
A = matrix(c(1,4,7,2,5,8,3,6,9),nrow=3,ncol=3)
```

Θα κατασκευαστεί πίνακας 3Χ3 με τιμές αυτές του διανύσματος, σχηματίζοντας πρώτα τις στήλες.

```
> print(A)
```

	[,1]	[,2]	[,3]
[1,]	1	2	3
[2,]	4	5	6
[3,]	7	8	9

# Matrix

Παράδειγμα

```
B = matrix(c(1,2,3,4,5,6,7,8,9),nrow=3,ncol=3,  
byrow=TRUE)
```

Θα κατασκευαστεί πίνακας 3Χ3 με τιμές αυτές του διανύσματος, σχηματίζοντας πρώτα τις γραμμές.

```
> print(B)
```

	[, 1]	[, 2]	[, 3]
[1, ]	1	2	3
[2, ]	4	5	6
[3, ]	7	8	9

# Matrix

Για να αναφερθούμε σε μια τιμή του πίνακα, χρησιμοποιούμε το όνομα της μεταβλητής, τις τετράγωνες αγκύλες και τους δύο δείκτες της τιμής – δηλ. τον αριθμό της γραμμής και τον αριθμό της στήλης που εκφράζει τη θέση της τιμής στο πίνακα.

Παράδειγμα

```
print(A[2, 3])
```

```
> 6
```

# Data frame

Διατεταγμένη λίστα (list) από διανύσματα (vectors) που όλα έχουν το ίδιο μήκος – δηλ. το ίδιο πλήθος τιμών. Ουσιαστικά έχει την έννοια δισδιάστατου πίνακα, που κάθε στήλη περιέχει τιμές ίδιου τύπου, αλλά όχι όλος το πίνακας (σε αντίθεση με το matrix). Κατασκευάζεται με τη συνάρτηση `dataframe()`.

# Data frame

Παράδειγμα

```
a = c(10, 20, 30, 40, 50)
```

```
b = c(2, 2, 4, 4, 8)
```

```
st = c("aa", "bb", "cc", "dd", "ee")
```

```
mydata = data.frame(st, a, b,  
row.names=c("A", "B", "C", "D", "E"),  
stringsAsFactors = F )
```

Η παράμετρος `row.names` καθορίζει τα ονόματα των γραμμών. Τα ονόματα των στηλών είναι τα ονόματα των διανυσμάτων που σχηματίζουν το data frame.

# Data frame

mydata

	row.names	st	a	b
1	A	aa	10	2
2	B	bb	20	2
3	C	cc	30	4
4	D	dd	40	4
5	E	ee	50	8

# Data frame

Για να αναφερθούμε σε μια τιμή του πλαισίου δεδομένων, χρησιμοποιούμε το όνομα της μεταβλητής, τις τετράγωνες αγκύλες και τους δύο δείκτες της τιμής – δηλ. τον αριθμό ή το όνομα της γραμμής και τον αριθμό ή το όνομα της στήλης που εκφράζει τη θέση της τιμής στο πλαίσιο δεδομένων.

Παράδειγμα

```
print (mydata [1 , 2] )
```

```
> 10
```

```
print (mydata [ "A" , "a" ] )
```

```
> 10
```



# Τέλος Ενότητας



Ευρωπαϊκή Ένωση  
Ευρωπαϊκό Κοινωνικό Ταμείο



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

