



ARISTOTLE
UNIVERSITY OF
THESSALONIKI

FACULTY OF HEALTH SCIENCES - SCHOOL OF MEDICINE
MSc Health Statistics and Data Analytics

data structures

KONSTANTINOS I. BOUGIOUKAS

PHYSICIST, BIOSTATISTICIAN AND RESEARCH METHODOLOGIST

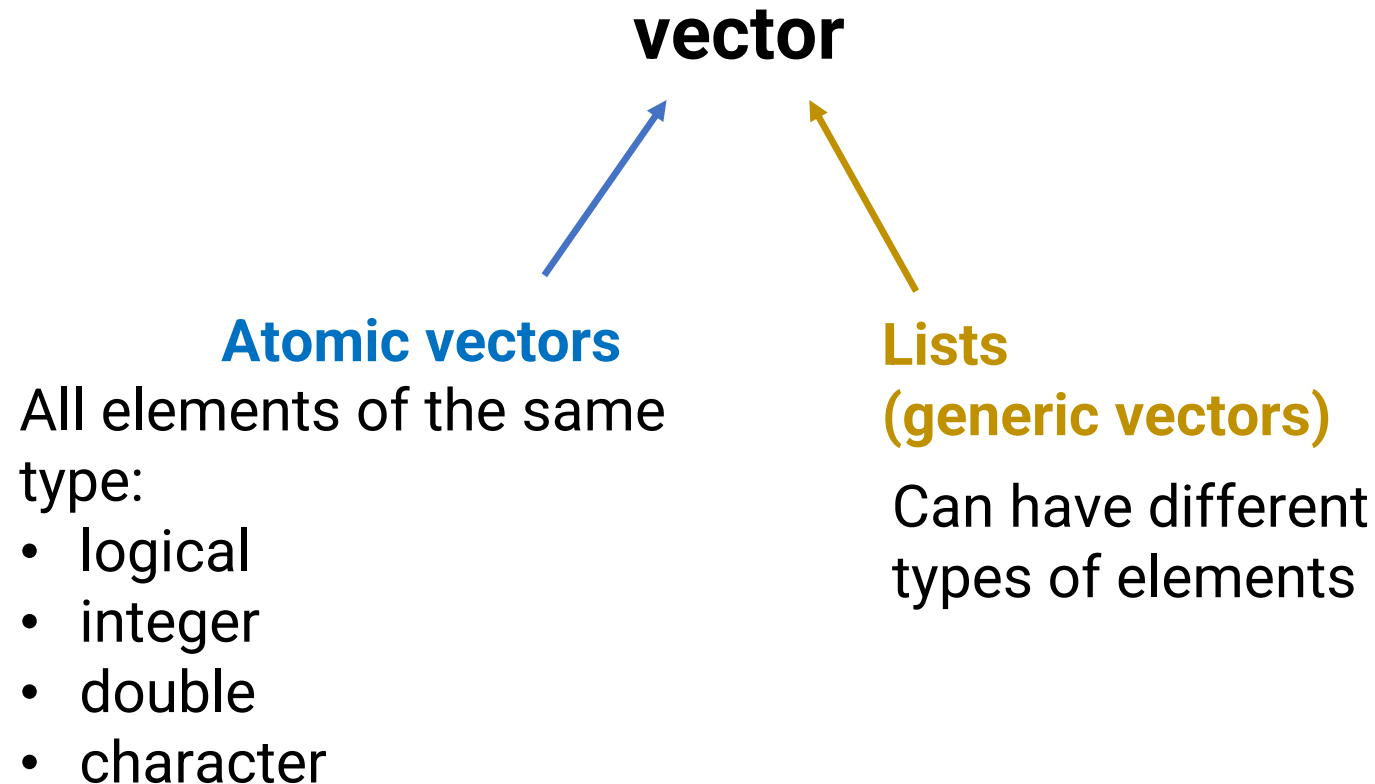


THESSALONIKI 2021-22



Basic data structures

The most fundamental concept in base R are the **vectors**.



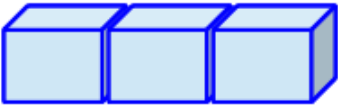
NULL

closely related to vectors and often serves the role of a generic **zero** length vector

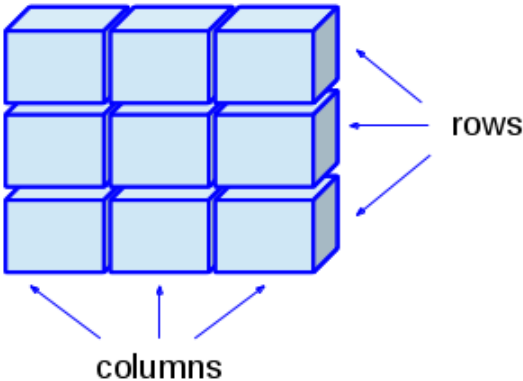
Basic data structures

Atomic vectors

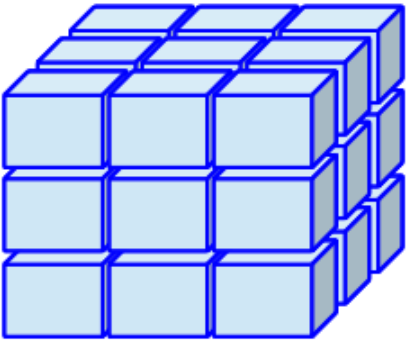
Vector



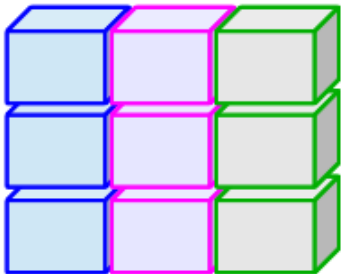
Matrix (2D vector)



Array (3D vector)

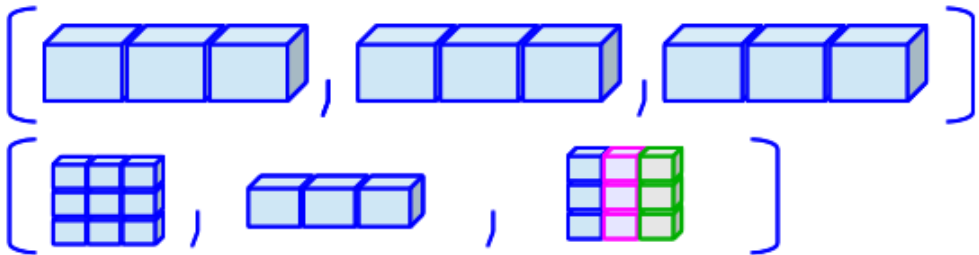


Data Frame
(Table)



Lists (generic vectors)

Lists



Rectangular table

Basic data structures: vector

A collection of one or more objects of the same type. The combine function `c()` is used to form a vector. For example:

```
a <- c(4, 6, 20, 10) # numeric vector
```

```
b <- c("Male", "Female", "Male") # character vector
```

```
d <- c(TRUE, TRUE, TRUE, FALSE, TRUE) # logical vector
```

- **Factor** is a special vector that represents categorical data (eg. vector b). It can be ordered or unordered.

Basic data structures: matrix

A matrix is collection of data elements arranged in a two-dimensional rectangular layout. The following is an example of a matrix with 2 rows and 3 columns (2x3=6 elements):

```
a <- c(2, 1, 4, 5, 3, 7) # numeric vector with six elements  
  
# Matrices are constructed column-wise starting in the “upper  
# Left” corner and running down the columns  
A <- matrix(a, nrow=2, ncol=3)
```

	1 st column	2 nd column	3 rd column
	[,1]	[,2]	[,3]
1 st row	[1,] 2	4	3
2 nd row	[2,] 1	5	7

Basic data structures: data frame

In a data frame the different columns can have different types of data (numeric, character, date etc.).

observations	age_yrs	sex	patient_is_male	adm_date	weight_kg	pain
	52	Male	TRUE	2009-10-15	78.3	mild
	57	Female	FALSE	2009-11-01	64.7	severe
	63	Female	FALSE	2009-10-21	67.2	moderate
	60	Male	TRUE	2009-10-28	NA	mild

```
df <- tibble(age_yrs, sex, patient_is_male,  
             adm_date, weight_kg, pain)
```