Applied Statistics for Data Scientists with R

Class 03: Data Structures

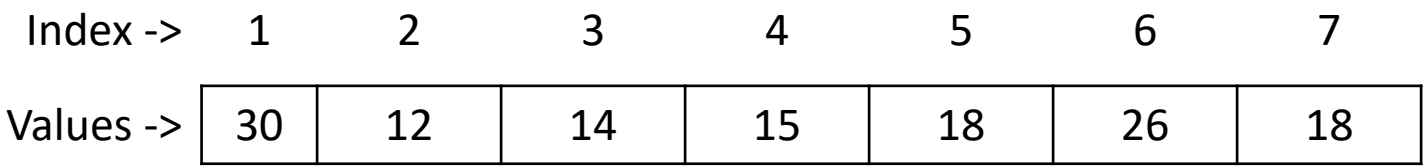
Learning Objective

1. Identifying and creating different data structures
2. Modifying data structures

Common Data Structures in R

Single Data Type	Vector
	Matrix
	Array
Multiple Data Type	Data Frame
	List

Vector



Matrix

2x2 Matrix

	Column 1	Column 2
Row 1	2	3
Row 2	1	5

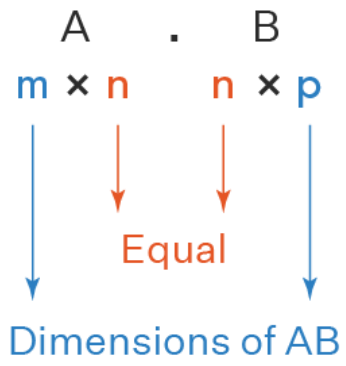
2x3 Matrix

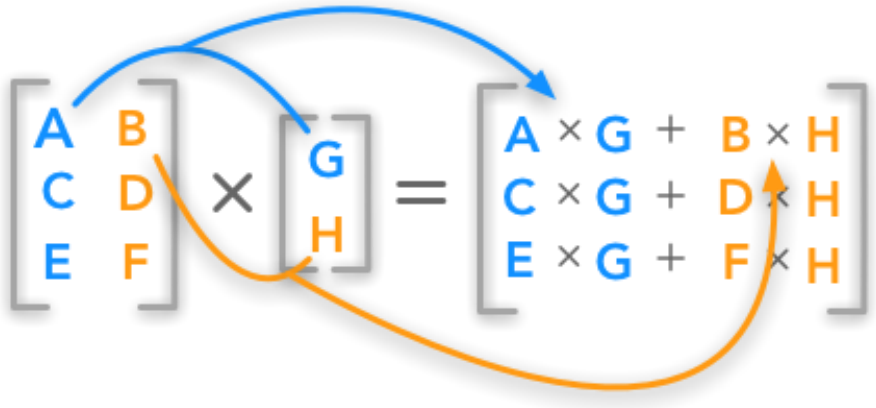
	Column 1	Column 2	Column 3
Row 1	2	3	2
Row 2	1	5	6

4x3 Matrix

	Column 1	Column 2	Column 3
Row 1	2	3	2
Row 2	1	5	6
Row 3	9	6	4
Row 4	1	2	6

Matrix Multiplication



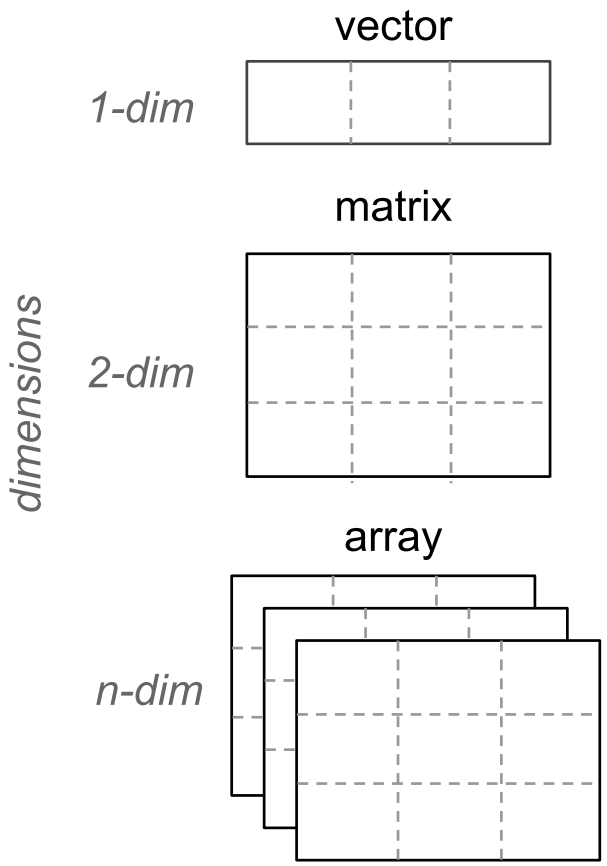


Matrix multiplication

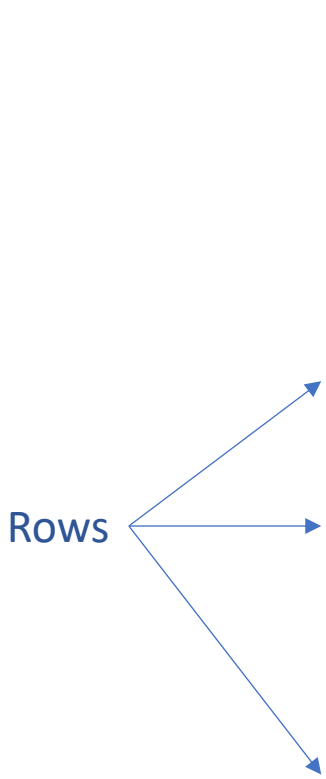
$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \otimes \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} = \begin{bmatrix} 1\begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} & 2\begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} \\ 3\begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} & 4\begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} \end{bmatrix} = \left[\begin{array}{cc|cc} 1\times0 & 1\times5 & 2\times0 & 2\times5 \\ 1\times6 & 1\times7 & 2\times6 & 2\times7 \\ \hline 3\times0 & 3\times5 & 4\times0 & 4\times5 \\ 3\times6 & 3\times7 & 4\times6 & 4\times7 \end{array}\right] = \left[\begin{array}{cc|cc} 0 & 5 & 0 & 10 \\ 6 & 7 & 12 & 14 \\ \hline 0 & 15 & 0 & 20 \\ 18 & 21 & 24 & 28 \end{array}\right].$$

Kronecker product

Array



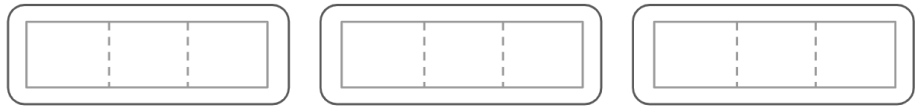
Data Frame



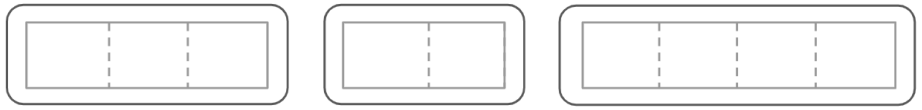
	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa
7	4.6	3.4	1.4	0.3	setosa
8	5	3.4	1.5	0.2	setosa
9	4.4	2.9	1.4	0.2	setosa
10	4.9	3.1	1.5	0.1	setosa
11	5.4	3.7	1.5	0.2	setosa
12	4.8	3.4	1.6	0.2	setosa

List

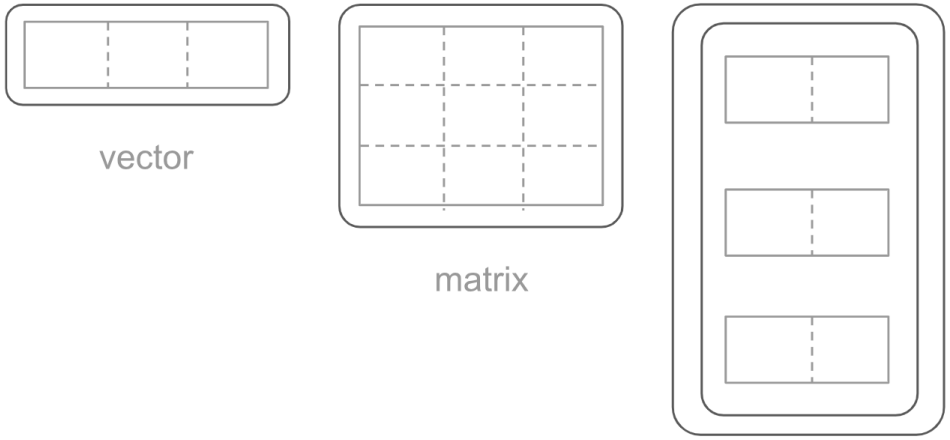
List of Vectors (of equal length)



List of Vectors (of different length)



List of various objects



vector

matrix

Other lists