



# R Programming

INT232



# Variables and Data Types in R

# Variable

Variables are nothing but reserved memory locations to store values. This means that when you create a variable you reserve some space in memory.

# Data Types

A data type, in programming, is a classification that specifies which type of value a variable has and what type of mathematical, relational or logical operations can be applied to it without causing an error.

# Types of Data Types

- Vectors
- Matrix
- Array
- List
- Data Frame

# Vector

- Vector is a sequence of data elements of the same basic type
- There are 5 atomic vectors, also termed as 5 classes of vectors.

# Vector Types

- |             |                 |
|-------------|-----------------|
| ○ Logical   | True or False   |
| ○ Integer   | 15L, 30L, 1699L |
| ○ Numeric   | 5, 3.14, 9452   |
| ○ Complex   | $4+3i$          |
| ○ Character | 'A', "Hey"      |

## Matrices

A matrix is a two-dimensional rectangular data set. It can be created using a vector input to the matrix function.

```
# Create a matrix.
```

```
M = matrix( c('a','a','b','c','b','a'), nrow = 2, ncol = 3,  
byrow = TRUE)
```

```
print(M)
```

## Output

```
[,1] [,2] [,3] [1,] "a" "a" "b" [2,] "c" "b" "a"
```



## **Lists**

- A list is an R-object which can contain many different types of elements inside it like vectors, functions and even another list inside it.

### **# Create a list.**

```
list1 <- list(c(2,5,3),21.3,sin)
```

### **# Print the list.**

```
print(list1)
```

## Arrays

While matrices are confined to two dimensions, arrays can be of any number of dimensions. The array function takes a dim attribute which creates the required number of dimension.

In the below example we create an array with two elements which are 3x3 matrices each.

**# Create an array.**

```
a <- array(c('green','yellow'),dim = c(3,3,2))  
print(a)
```

## Factors

- Factors are the r-objects which are created using a vector. It stores the vector along with the distinct values of the elements in the vector as labels.
- The labels are always character irrespective of whether it is numeric or character or Boolean etc. in the input vector. They are useful in statistical modeling.
- Factors are created using the **factor()** function. The **nlevels** functions gives the count of levels.

**# Create a vector.**

```
apple_colors <-  
c('green','green','yellow','red','red','red','green')
```

**# Create a factor object.**

```
factor_apple <- factor(apple_colors)
```

**# Print the factor.**

```
print(factor_apple)  
print(nlevels(factor_apple))
```

## Data Frames

- Data frames are tabular data objects. Unlike a matrix in data frame each column can contain different modes of data.
- The first column can be numeric while the second column can be character and third column can be logical. It is a list of vectors of equal length.
- Data Frames are created using the **data.frame()** function.

- **# Create the data frame.**

```
BMI <- data.frame(  
  gender = c("Male", "Male", "Female"),  
  height = c(152, 171.5, 165),  
  weight = c(81, 93, 78),  
  Age = c(42, 38, 26) )  
print(BMI)
```

# Output

**gender height weight Age**

Male 152.0 81 42

Male 171.5 93 38

Female 165.0 78 26



- Numeric, Integer ,Character ,Logical and Date are the basic data types in R.
- **Class** function returns the data type of an object.
- **is.data\_type** tests whether an object is of specified data type
  - is.numeric
  - is.integer
  - is.character
  - is.logical
- **as.data\_type** will coerce objects to the specified data type
  - as.numeric
  - as.integer
  - as.character
  - as.logical