More Data Structures

- Vectors
- Factors
- Lists
- Matrices
- Arrays
- Data Frames

Array

Arrays are the R data objects which can store data in more than two dimensions.

```
result <- array((1:9),dim = c(3,3,2))
print(result)
```

Naming Columns and Rows

```
>r <- c("ROW1","ROW2","ROW3")
>c <- c("COL1","COL2","COL3")
>m <- c("Matrix1","Matrix2")

>array(1:9,c(3,3,2),list(r, c,m))
```

Accessing array elements.

- # Print the third row of the second matrix of the array.
  print(result[3,,2])

- # Print the element in the 1st row and 3rd column of the 1st matrix.
  print(result[1,3,1])

- # Print the 2nd Matrix. print(result[,,2])

Data Frame

- A **data frame** is used for storing data tables. It is a list of vectors of equal length. The following variable df is a data frame containing three vectors n, s, b.
- > n = c(2, 3, 5)
  > s = c(“Football", “Cricket", “Tennis")
  > b = c(TRUE, FALSE, TRUE)
  > df = data.frame(n, s, b) # df is a data frame

Data Frame-Characteristics

- The column names should be non-empty.
- The row names should be unique.
- The data stored in a data frame can be any basic data type.
- Each column should contain same number of data items.

Access the items from Data Frame

- result <- data.frame(dfn,dfs)

Extract the first two rows and then all columns

- result <- df[1:2,]

Bind the two data frames.

Use rbind and cbind for row bind and column bind operations

Row Names

row.names(df)=c(“one”,”two”,”three”)

row.names(df)=c(1,2,3)

R Environment

- ls() to list the variables declared
- getwd()
- setwd()
- list.files() to see the list of files in the workspace

Reading Data into R

```
read.table("marks.txt")
read.table(file.choose())

p<-data.frame(read.csv(“Pizza.csv"))
print(is.data.frame(p))
head(p)
max(p$Rating)
```

Saving the workspace

save.image()

load(".RData")

- save.image("new.Rdata")
- load("new.Rdata")

- Creating an Rscript file
- Loading an RScript