Code

1 box plotting

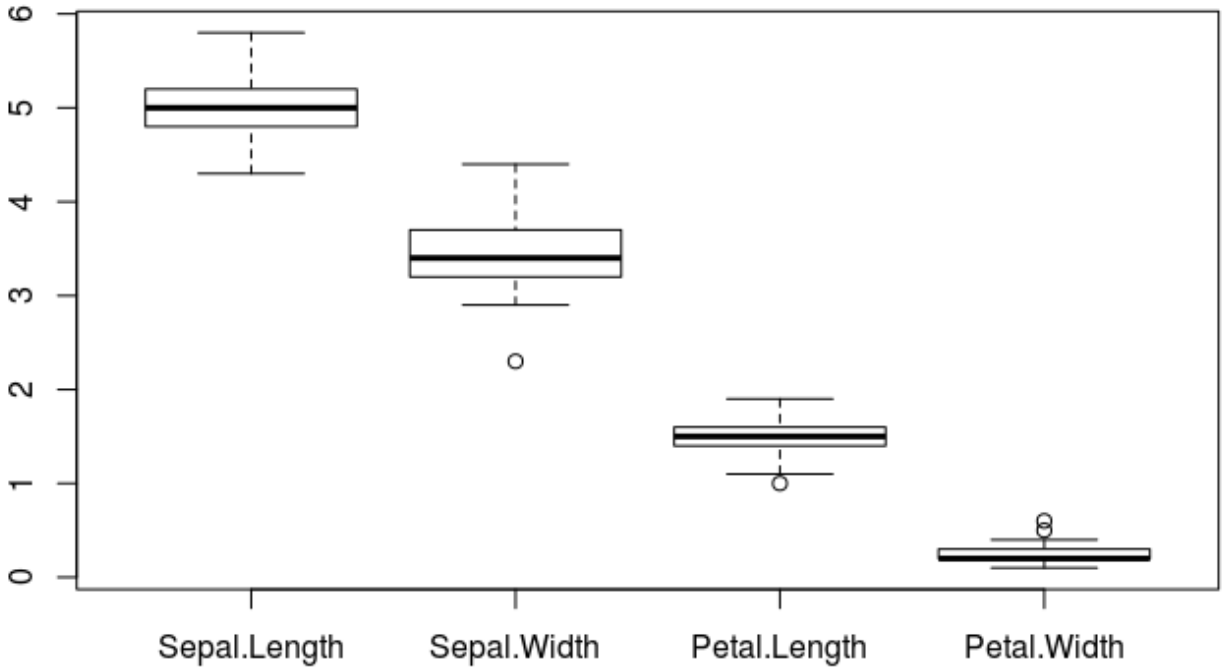
Hide

```
data("iris")
str(iris)
```

```
'data.frame':   150 obs. of  5 variables:
 $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
 $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
 $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
 $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
 $ Species     : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1 1 1 1 1 1 ...
```

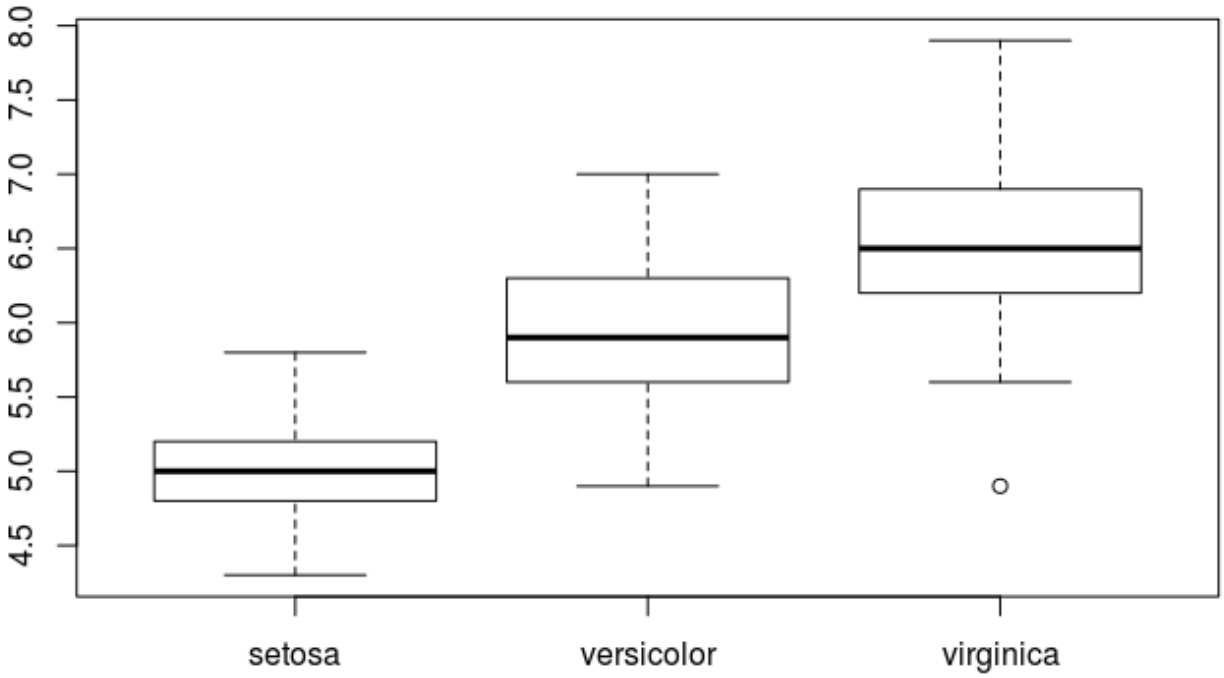
Hide

```
boxplot(as.list(iris[1:50, 1:4]))
```



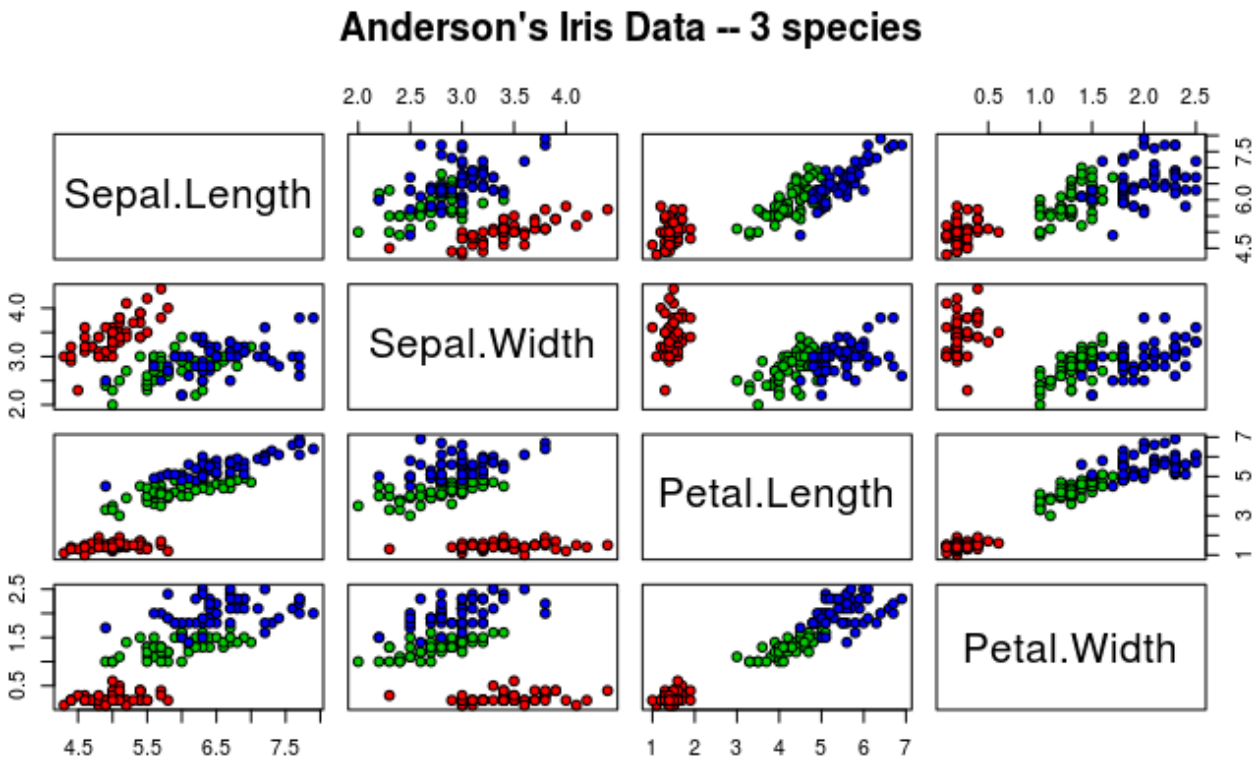
Hide

```
boxplot(Sepal.Length ~ Species, data = iris)
```



Hide

```
pairs(iris[1:4], main = "Anderson's Iris Data -- 3 species",   pch = 21,
      bg = c("red", "green3", "blue")[unclass(iris$Species)])
```



2 decision tree

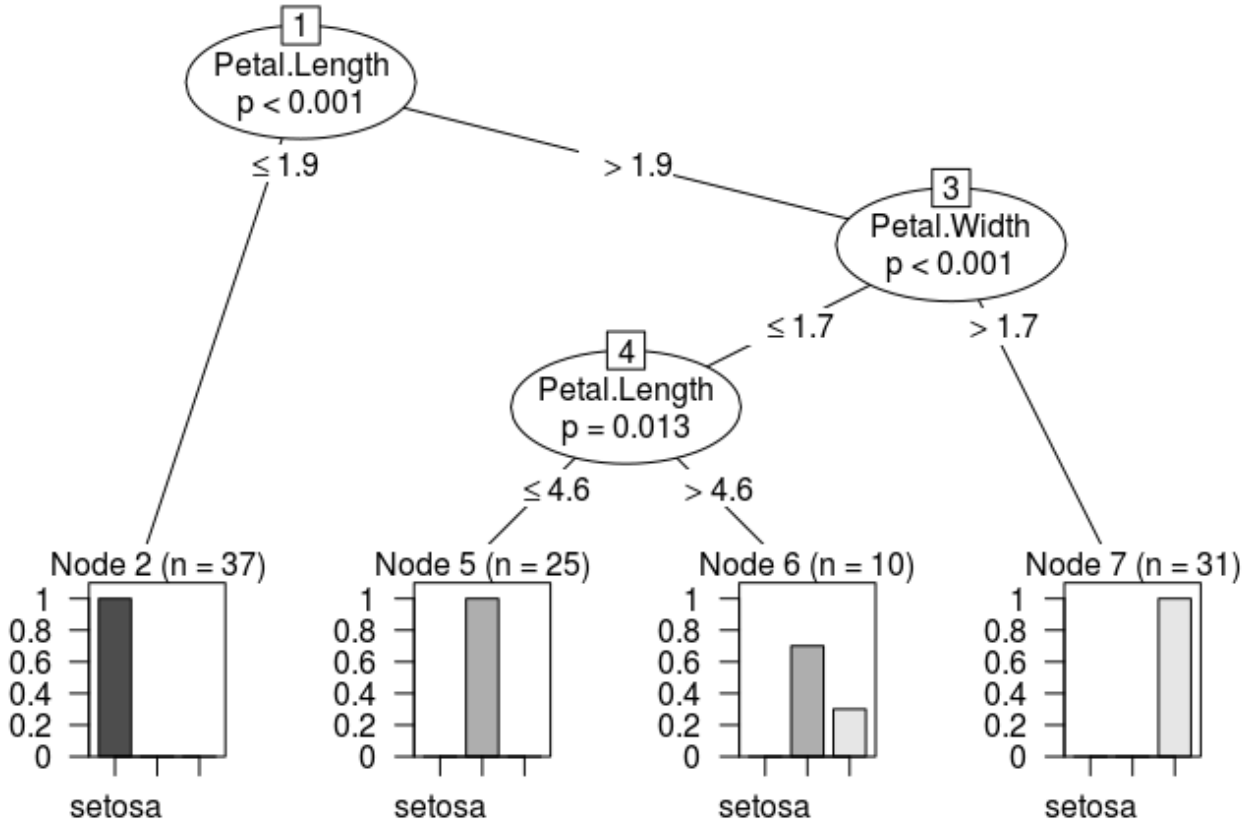
Hide

```
idx <- sample(2, nrow(iris), replace=TRUE, prob=c(0.7, 0.3))
trainData <- iris[idx==1,]
testData <- iris[idx==2,]
#install.packages("party")
require(party)
formula <- Species ~ Sepal.Length + Sepal.Width + Petal.Length + Petal.Width
iris_ctree <- ctree(formula, data=trainData)
pred <- predict(iris_ctree, testData)
conf.mat <- table(pred, testData$Species)
(accuracy <- sum(diag(conf.mat))/sum(conf.mat) * 100)
```

```
[1] 93.61702
```

Hide

```
plot(iris_ctree)
```



3 data type

Hide

```
sex <- factor("m", c("m", "f"))
sex <- factor(c("m", "f", "m"))
sex
```

```
[1] m f m
Levels: f m
```

Hide

```
nlevels(sex)
```

```
[1] 2
```

Hide

```
levels(sex)
```

```
[1] "f" "m"
```

Hide

```
levels(sex) <- c("male", "female") # level 값 전체를 바꾸면, 해당 값을 가진 변수도 변한다.
sex
```

```
[1] female male   female
Levels: male female
```

Hide

```
ordered(c("a", "b", "c"))
```

```
[1] a b c
Levels: a < b < c
```

Hide

```
factor(c("a", "b", "c"), ordered=TRUE)
```

```
[1] a b c
Levels: a < b < c
```

4 vector

Hide

```
x <- c(1, 3, 4)
names(x) <- c("kim", "seo", "park")
x
```

```
 kim  seo park
   1    3    4
```

Hide

```
x[c("kim")]
```

```
kim
  1
```

Hide

```
x["kim"]
```

```
kim
  1
```

Hide

```
nrow(x)
```

```
NULL
```

Hide

```
NROW(x)
```

```
[1] 3
```

Hide

```
"a" %in% c("a", "b", "c")
```

```
[1] TRUE
```

Hide

```
"d" %in% c("a", "b", "c")
```

```
[1] FALSE
```

Hide

```
setdiff(c("a", "b", "c"), c("a", "d"))
```

```
[1] "b" "c"
```

Hide

```
union(c("a", "b", "c"), c("a", "d"))
```

```
[1] "a" "b" "c" "d"
```

Hide

```
intersect(c("a", "b", "c"), c("a", "d"))
```

```
[1] "a"
```

Hide

```
seq(1, 5)
```

```
[1] 1 2 3 4 5
```

Hide

```
seq(1, 5, 2)
```

```
[1] 1 3 5
```

Hide

```
1:5   # vector를 생성함에 주의
```

```
[1] 1 2 3 4 5
```

5 list

Hide

```
x <- list(namw="foo", height=70)
x
```

```
$namw
[1] "foo"

$height
[1] 70
```

Hide

```
x <- list(name="foo", height=c(1,3,5))
x$name
```

```
[1] "foo"
```

Hide

```
x$height
```

```
[1] 1 3 5
```

Hide

```
x[1]
```

```
$name
[1] "foo"
```

Hide

```
x[[1]]
```

```
[1] "foo"
```

Hide

```
x[2]
```

```
$height
[1] 1 3 5
```

Hide

```
x[[2]]
```

```
[1] 1 3 5
```

6 matrix

Hide

```
matrix(c(1,2,3,4,5,6,7,8,9),  nrow=3) # column major
```

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

Hide

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

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

Hide

```
matrix(c(1,2,3,4,5,6,7,8,9),  nrow=3, byrow=T) # row major
```

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

Hide

```
matrix(c(1,2,3,4,5,6,7,8,9),  nrow=3,
                dimnames=list(c("item1", "item2", "item3"),
                              c("feature1", "feature2", "feature3")))
```

```
      feature1 feature2 feature3
item1        1        4        7
item2        2        5        8
item3        3        6        9
```

Hide

```
x <- matrix(c(1,2,3,4,5,6,7,8,9),  nrow=3)
rownames(x) <- c("r1", "r2", "r3")
colnames(x) <- c("c1", "c2", "c3")
x
```

```
   c1 c2 c3
r1  1  4  7
r2  2  5  8
r3  3  6  9
```

Hide

```
x <- matrix(c(1,2,3,4,5,6,7,8,9),  nrow=3,
       dimnames=list(c("item1", "item2", "item3"),
                     c("feature1", "feature2", "feature3")))
x
```

```
      feature1 feature2 feature3
item1        1        4        7
item2        2        5        8
item3        3        6        9
```

Hide

```
x[ , "feature2"] >= 5
```

```
item1 item2 item3
FALSE  TRUE  TRUE
```

Hide

```
x[ x[, "feature2"] >= 5 ,]
```

```
      feature1 feature2 feature3
item2        2        5        8
item3        3        6        9
```

Hide

```
x
```

```
      feature1 feature2 feature3
item1        1        4        7
item2        2        5        8
item3        3        6        9
```

Hide

```
x[ , 2]>=5
```

```
item1 item2 item3
FALSE  TRUE  TRUE
```

Hide

```
x[x[ , 2]>=5, ]
```

```
      feature1 feature2 feature3
item2        2        5        8
item3        3        6        9
```

7 matrix transpose

Hide

```
x <- matrix(c(1,2,3,4,5,6,7,8,9),  nrow=3)
x
```

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

Hide

```
t(x)
```

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

Hide

```
ncol(x)
```

```
[1] 3
```

Hide

```
nrow(x)
```

```
[1] 3
```

Hide

```
x %*% t(x)
```

```
     [,1] [,2] [,3]
[1,]   66   78   90
[2,]   78   93  108
[3,]   90  108  126
```

Hide

```
x * t(x)
```

```
     [,1] [,2] [,3]
[1,]    1    8   21
[2,]    8   25   48
[3,]   21   48   81
```

8 data.frame

Hide

```
d <- data.frame(x=c(1,2,3,4,5), y=c(2,4,6,8,10))
d
```

x <dbl>	y <dbl>
1	2
2	4
3	6
4	8
5	10

5 rows

Hide

```
x <- c(1,2,3,4,5)
y <- c(2,4,6,8,10)
z <- c("M", "F", "M", "F", "M")
d <- data.frame(x, y, z)
d
```

x <dbl>	y <dbl>	z <fctr>
1	2	M
2	4	F
3	6	M
4	8	F
5	10	M

5 rows

Hide

```
str(d)
```

```
'data.frame':   5 obs. of  3 variables:
 $ x: num  1 2 3 4 5
 $ y: num  2 4 6 8 10
 $ z: Factor w/ 2 levels "F","M": 2 1 2 1 2
```

Hide

```
d <- data.frame(x, y, z, stringsAsFactors = F)
str(d) # 확인
```

```
'data.frame':   5 obs. of  3 variables:
 $ x: num  1 2 3 4 5
 $ y: num  2 4 6 8 10
 $ z: chr  "M" "F" "M" "F" ...
```

Hide

```
d$v <- c(3, 6, 9, 12, 15)
d
```

x <dbl>	y <dbl>	z <chr>	v <dbl>
1	2	M	3
2	4	F	6
3	6	M	9
4	8	F	12
5	10	M	15

5 rows

Hide

```
d$v2 <- c(10, 20, 30, 40, 50)
d
```

x <dbl>	y <dbl>	z <chr>	v <dbl>	v2 <dbl>
1	2	M	3	10
2	4	F	6	20
3	6	M	9	30
4	8	F	12	40
5	10	M	15	50

5 rows

Hide

```
d$x
```

```
[1] 1 2 3 4 5
```

Hide

```
d[1, ]
```

	x <dbl>	y <dbl>	z <chr>	v <dbl>	v2 <dbl>
1	1	2	M	3	10

1 row

Hide

```
d[ , "v"]
```

```
[1]  3  6  9 12 15
```

Hide

```
d[ , c("x", "v")]
```

x <dbl>	v <dbl>
1	3
2	6
3	9
4	12
5	15

5 rows

Hide

```
d[d$x>3, ]
```

	x <dbl>	y <dbl>	z <chr>	v <dbl>	v2 <dbl>
4	4	8	F	12	40
5	5	10	M	15	50

2 rows

Hide

```
d[ , "v", drop=F]
```

v <dbl>
3
6
9
12
15

5 rows

Hide

```
x <- data.frame(1:3, 11:13)
x
```

X1.3 <int>	X11.13 <int>
1	11
2	12
3	13

3 rows

Hide

```
colnames(x) <- c("col1", "col2")
x
```

col1 <int>	col2 <int>
1	11
2	12
3	13

3 rows

Hide

```
rownames(x) <- c("row1", "row2", "row3")
x
```

	col1 <int>	col2 <int>
row1	1	11
row2	2	12

	col1 <int>	col2 <int>
row3	3	13

3 rows

Hide

```
x <- data.frame(matrix(c(1,2,3,4), ncol=2))
x
```

	X1 <dbl>	X2 <dbl>
	1	3
	2	4

2 rows

Hide

```
colnames(x) <- c("X", "Y")
x <- data.frame(list(x=c(1,2), y=c(3,4)))
str(x)
```

```
'data.frame':   2 obs. of  2 variables:
 $ x: num  1 2
 $ y: num  3 4
```

9 if statements

Hide

```
x <- c(1,2,3,4,5)
result <- ifelse( x %% 2 == 0, "even", "odd")
result
```

```
[1] "odd"  "even" "odd"  "even" "odd"
```

10 for statements

Hide

```
for (i in 1:10) { print(i) }
```

```
[1] 1
[1] 2
[1] 3
[1] 4
[1] 5
[1] 6
[1] 7
[1] 8
[1] 9
[1] 10
```

Hide

```
i <- 1
repeat {
  print(i)
  if (i >= 10) { break }
  i <- i+1
}
```

```
[1] 1
[1] 2
[1] 3
[1] 4
[1] 5
[1] 6
[1] 7
[1] 8
[1] 9
[1] 10
```

Hide

```
i <- 0
while (i < 10){
  print(i)
  i <- i+1
}
```

```
[1] 0
[1] 1
[1] 2
[1] 3
[1] 4
[1] 5
[1] 6
[1] 7
[1] 8
[1] 9
```

11 vector operation

Hide

```
d <- data.frame(x=c(1,2,3,4,5), y=c("a", "b", "c", "d", "e"))
d
```

	x <dbl>	y <fctr>
	1	a
	2	b
	3	c
	4	d
	5	e

5 rows

Hide

```
d[c(TRUE, FALSE, TRUE, FALSE, TRUE), ]
```

	x <dbl>	y <fctr>
1	1	a
3	3	c
5	5	e

3 rows

Hide

```
d[d$x%%2==0, ]
```

	x <dbl>	y <fctr>
2	2	b
4	4	d

2 rows

Hide

```
sum(c(1,2,3,4,5))
```

```
[1] 15
```

Hide

```
sum(c(1,2,3,4,NA))
```

```
[1] NA
```

Hide

```
sum(c(1,2,3,4,NA), na.rm = TRUE)
```

```
[1] 10
```

Hide

```
mean(c(1,2,3,4,NA))
```

```
[1] NA
```

Hide

```
mean(c(1,2,3,4,NA), na.rm = TRUE)
```

```
[1] 2.5
```

Hide

```
x <- data.frame(a=c(1,2,3), b=c("a", NA, "c"), c=c("a", "b", NA))
x
```

	a <dbl>	b <fctr>	c <fctr>
	1	a	a
	2	*NA*	b
	3	c	*NA*

3 rows

Hide

```
na.omit(x) # Remove rows with missing values on columns specified
```

	a <dbl>	b <fctr>	c <fctr>
1	1	a	a

1 row

12 functions

Hide

```
fibo <- function(n){
  if (n==1 || n==2) { return(1)}
  else { return(fibo(n-1)+fibo(n-2))}
}
fibo(2)
```

```
[1] 1
```

Hide

```
fibo(10)
```

```
[1] 55
```

Hide

```
f <- function(x, y) { print(x); print(y)}
f(1,2)
```

```
[1] 1
[1] 2
```

Hide

```
x <- c(1,2,3)
y <- c(10, 20, 30)
f(x, y)
```

```
[1] 1 2 3
[1] 10 20 30
```

Hide

```
g <- function(z, ...){
  print(z)
  f(...)
}
g(1,2,3)
```

```
[1] 1
[1] 2
[1] 3
```

13 call-by-value

Hide

```
f <- function(x,i) {x[i] <- 4}
w<-c(10,11,12,13)
print(f(w, 1))
```

```
[1] 4
```

Hide

```
w
```

```
[1] 10 11 12 13
```