

MDA102-Statistical methods

Lecture 5 – 30 June 2020

6.30AM – 8.30AM

Dr Sharon Varghese A

MISSION

CHRIST is a nurturing ground for an individual's holistic development to make effective contribution to the society in a dynamic environment

VISION

Excellence and Service

R objects - Data structure in R - Contd...

Data frames

- is like matrix with columns of different data type creating data frame

```
employees<-c("jack", "Jill")
```

```
salary<-c(89000, 75000)
```


```
d<-data.frame(employees, salary, stringsAsFactors=TRUE)
```

string vector will be converted into factor

d#matrix like view

	employees	salary
1	jack	89000
2	Jill	75000

All matrix like indexing works with data frames

```
d[,1]   
[1] "jack" "Jill"  
d[[1]]  
[1] "jack" "Jill"  
d$employees  
[1] "jack" "Jill"
```

Adding row or column

```
d1<-rbind(d,list("Alan",75000)) #  
data frame
```

d1

	employees	salary
1	jack	89000
2	Jill	75000
3	Alan	75000

```
d2<-cbind(d,list(place='NY'))
```

d2

	employees	salary	place
1	jack	89000	NY
2	Jill	75000	NY

```
d$place<-c('NY','LA')
```

```
d
```

	employees	salary	place
1	jack	89000	NY
2	Jill	75000	LA

Factors -structure for categorical and nominal data from s

```
x<-
```

```
c(rep('Republican',14),rep('Democrat',10),rep('No response',5))
```

```
xf<-factor(x)
```

```
xf #will show vector of x and display levels i.e., categories
```

[1]	Republican	Republican	Republi
	Republican		
[6]	Republican	Republican	Republi
	Republican		
[11]	Republican	Republican	Republi
[16]	Democrat	Democrat	Democra
[21]	Democrat	Democrat	Democra
[26]	Democrat	Democrat	Democra
[31]	Democrat	Democrat	Unaffil
	Unaffiliated		
[36]	Unaffiliated	Unaffiliated	Unaffil
	Unaffiliated		
[41]	Unaffiliated	Unaffiliated	No resp
	response		
[46]	No response	No response	
Levels:	Democrat	No response	Republica

```

length(xf) #number of observations
distinct levels
[1] 47
unclass(xf) #a vector of coded data
[1] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
1 1 1 1 1 1 1 4 4 4 4 4 4
[39] 4 4 4 4 2 2 2 2 2
attr(,"levels")
[1] "Democrat" "No response"
"Unaffiliated"

```


- Arrays - are the R data objects which can store data in more

```
x<-array(c(1, 2, 3, 4, 5, 6, 7, 8), c(2,
```

```
x
```

```
, , 1
```

	[, 1]	[, 2]
[1,]	1	3
[2,]	2	4

```
, , 2
```

	[, 1]	[, 2]
[1,]	5	7
[2,]	6	8

```
dim(x)
```

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

```
x[, , 1] #select 1st matrix
```

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

```
[1,]      1      3
```

```
[2,]      2      4
```

```
x[1, , ] #select 1st row of every matrix
```

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

```
[1,]      1      5
```

```
[2,]      3      7
```

Learn R in R - 'swirl' package

(<https://swirlstats.com/>)

```
install.packages("swirl")  
library(swirl)  
swirl()
```