

MDA102 – Statistical methods

Lecture 10
13 July 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

How to load webpage

- `url()` - To open a webpage in R
- `readLines()` - To read webpage line by line and store it in vector

```
page<-url("https://www.christuniversity.in")  
readLines(page)
```

file.choose()

To select files anywhere in your system using select file v

```
data<-read.csv(file.choose()) #a csv file ca  
window and stored in variable data.  
View(data)
```

loading multiple R scripts

- `source()` - To load multiple R scripts into R console and execute

```
source("w1.R")  
source("w2.R")  
## [1] 11
```

some file functions

file.copy() - To create a copy of a file from R environment

```
source("w1.R")  
file.copy("w1.R", "w1_copy.R")  
## [1] FALSE  
source("w1_copy.R")
```

Missing values

Two types of missing values

- NA , not available
- NaN, not a number
- NaN are NA values but not vice versa
- `is.na()` #returns logic vector based NA values

```
x<-c(1, 8, 5, NA, 8)
```

```
x
```

```
## [1] 1 8 5 NA 8
```

```
is.na(x) #check NA in x and returns a logic
```

```
## [1] FALSE FALSE FALSE TRUE FALSE
```

`complete.cases()`

```
data<-read.csv("data.csv")
```

```
data_1<-
```

```
data.frame(data$Year, data$Area..in.1000hect
```

```
nes., stringsAsFactors = FALSE)
```

```
#View(data_1)
```

```
complete.cases(data_1) #to return true or false if  
is having missing values or not
```

Excellence and Service

Missing values – contd...

```
x<-c(1, 8, 4, NaN, 5, 6)
x
## [1] 1 8 4 NaN 5 6
is.nan(x) #to check NaN
## [1] FALSE FALSE FALSE TRUE F
x<-c(1, 8, 4, NaN, 5, NA)
x
## [1] 1 8 4 NaN 5 NA
is.na(x) #NaN is a subset of NA.
NaN also.
## [1] FALSE FALSE FALSE TRUE F
is.nan(x) #NA is not NaN always
## [1] FALSE FALSE FALSE TRUE F
```

Missing values – contd...

`complete.cases()`

```
data<-read.csv("data.csv")
```

```
data_1<-
```

```
data.frame(data$Year,data$Area..
```

```
roduction..1000.tones.,stringsAs
```

```
#View(data_1)
```

```
complete.cases(data_1) #to return
```

```
depending on wheter row is havin
```

```
not
```


Missing values – contd...

```
## [1] TRUE TRUE TRUE TRUE
TRUE TRUE TRUE TRUE TRUE
## [13] TRUE TRUE TRUE TRUE
TRUE TRUE TRUE TRUE TRUE
## [25] TRUE TRUE TRUE FALSE
TRUE TRUE TRUE FALSE TRUE
## [37] TRUE TRUE FALSE TRUE
TRUE TRUE TRUE TRUE TRUE
## [49] FALSE FALSE
```

`Complete.cases()` is used to subset dataframe values.

```
complete_data<-data_1[complete.cases(data_1),]
View(complete_data)
```