Missing Data

Dhruv Aggarwal

12/4/2020

# Missing data plays a crucial role in computing and Statistics  
# R has two types of missing data - NA and NULL  
# while they are similar, but they behave differently and hence needs attention!  
  
# NA - Missing data - Missing Value  
z = c(1,2,NA,8,3,NA,3)  
# z = c(1,2,na,8,3,na,3)  
z

## [1] 1 2 NA 8 3 NA 3

# "is.na" tests each element of a vector for missingness  
is.na(z)

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

#Another example  
z\_char = c("Hockey", NA ,"Cricket")  
z\_char

## [1] "Hockey" NA "Cricket"

is.na(z\_char)

## [1] FALSE TRUE FALSE

# NULL - Absence of anything. It is not exactly missingness, but nothingness  
# Eg: Having Brain but thinking Nothing! - Makes Sense!!!  
# Functions can sometimes return NULL and their arguments can be NULL.  
# Important difference is, NULL is atomical and cannot exist within a vector...   
# ...If used inside a vector, it simply disappears! Let's see...  
z= c(1,NULL,3)  
z

## [1] 1 3

x = c(1,NA,3)  
x

## [1] 1 NA 3

# Notice, here the "NULL" didnot get stored in "z", infact "z" has only length of 2!  
length(z)

## [1] 2

length(x)

## [1] 3

# Assigning NULL and checking!  
d = NULL  
is.null(d)

## [1] TRUE