Lab 2

Aniket Naik Desai

2019-03-30

# Logical Operators:

1. Use logical operations to get R to agree that “two plus two equals five” is FALSE.

x <- 5  
y <- 2+2  
x==y

## [1] FALSE

1. Use logical operations to test whether 8 ^ 13 is less than 15 ^ 9.

x = 8^13  
y = 15^9  
x>y

## [1] TRUE

# Variables:

1. Create a variable called potato whose value corresponds to the number of potatoes you’ve eaten in the last week. Or something equally ridiculous. Print out the value of potato.

potato = 6   
# I ate a mashed potato tub in the last week  
potato

## [1] 6

1. Calculate the square root of potato using the sqrt() function. Print out the value of potato again to verify that the value of potato hasn’t changed.

sqrt(potato)

## [1] 2.44949

potato

## [1] 6

1. Reassign the value of potato to potato \* 2. Print out the new value of potato to verify that it has changed.

potato = potato \*2  
potato

## [1] 12

1. Try making a character (string) variable and a logical variable . Try creating a variable with a “missing” value NA. You can call these variables whatever you would like. Use class(variablename) to make sure they are the right type of variable.

name = "superman"  
class(name)

## [1] "character"

from\_earth = FALSE  
class(from\_earth)

## [1] "logical"

address =NA  
class(address)

## [1] "logical"

# Vectors:

1. Create a numeric vector with three elements using c().

num\_vec = c(1,2,3)

1. Create a character vector with three elements using c().

char\_vec = c("A","B","C")

1. Create a numeric vector called age whose elements contain the ages of three people you know, where the names of each element correspond to the names of those people.

age = c(28,29,32)  
names(age)= c("Arpit","Aniket","Sharad")  
age

## Arpit Aniket Sharad   
## 28 29 32

1. Use “indexing by number” to get R to print out the first element of one of the vectors you created in the last questions.

age[1]

## Arpit   
## 28

1. Use logical indexing to return all the ages of all people in age greater than 20.

age[age>20]

## Arpit Aniket Sharad   
## 28 29 32

1. Use indexing by name to return the age of one of the people whose ages you’ve stored in age

age["Arpit"]

## Arpit   
## 28

# Dataframes:

1. Load the airquality dataset.
2. Use the $ method to print out the Wind variable in airquality.
3. Print out the third element of the Wind variable.

#Load the dataset  
airquality = airquality  
  
#Wind variable printed  
airquality$Wind

## [1] 7.4 8.0 12.6 11.5 14.3 14.9 8.6 13.8 20.1 8.6 6.9 9.7 9.2 10.9  
## [15] 13.2 11.5 12.0 18.4 11.5 9.7 9.7 16.6 9.7 12.0 16.6 14.9 8.0 12.0  
## [29] 14.9 5.7 7.4 8.6 9.7 16.1 9.2 8.6 14.3 9.7 6.9 13.8 11.5 10.9  
## [43] 9.2 8.0 13.8 11.5 14.9 20.7 9.2 11.5 10.3 6.3 1.7 4.6 6.3 8.0  
## [57] 8.0 10.3 11.5 14.9 8.0 4.1 9.2 9.2 10.9 4.6 10.9 5.1 6.3 5.7  
## [71] 7.4 8.6 14.3 14.9 14.9 14.3 6.9 10.3 6.3 5.1 11.5 6.9 9.7 11.5  
## [85] 8.6 8.0 8.6 12.0 7.4 7.4 7.4 9.2 6.9 13.8 7.4 6.9 7.4 4.6  
## [99] 4.0 10.3 8.0 8.6 11.5 11.5 11.5 9.7 11.5 10.3 6.3 7.4 10.9 10.3  
## [113] 15.5 14.3 12.6 9.7 3.4 8.0 5.7 9.7 2.3 6.3 6.3 6.9 5.1 2.8  
## [127] 4.6 7.4 15.5 10.9 10.3 10.9 9.7 14.9 15.5 6.3 10.9 11.5 6.9 13.8  
## [141] 10.3 10.3 8.0 12.6 9.2 10.3 10.3 16.6 6.9 13.2 14.3 8.0 11.5

#3rd element in Wind  
airquality$Wind[3]

## [1] 12.6

1. Create a new data frame called aq that includes only the first 10 cases. Hint: typing c(1,2,3,4,5,6,7,8,9,10) is tedious. R allows you to use 1:10 as a shorthand method!
2. Use logical indexing to print out all days (ie. cases) in aq where the Ozone level was higher than 20.
   1. What did the output do with NA values?
3. Use subset() to do the same thing. Notice the difference in the output.

aq = airquality[1:10,]  
  
aq[aq$Ozone>20,]

## Ozone Solar.R Wind Temp Month Day  
## 1 41 190 7.4 67 5 1  
## 2 36 118 8.0 72 5 2  
## NA NA NA NA NA NA NA  
## 6 28 NA 14.9 66 5 6  
## 7 23 299 8.6 65 5 7  
## NA.1 NA NA NA NA NA NA

1. Create a TooWindy variable inside aq, which is a logical variable that is TRUE if Wind is greater than 10, and FALSE otherwise.

aq$TooWindy = aq$Wind>10

1. Use the length() function to determine the number of observations in the airquality dataframe.

length(airquality$Day)

## [1] 153

#number of observations in Day = number of observations in airquality dataframe

1. Calculate the mean and standard deviation of one of the variables in airquality.

mean(airquality$Wind)

## [1] 9.957516

sd(airquality$Wind)

## [1] 3.523001

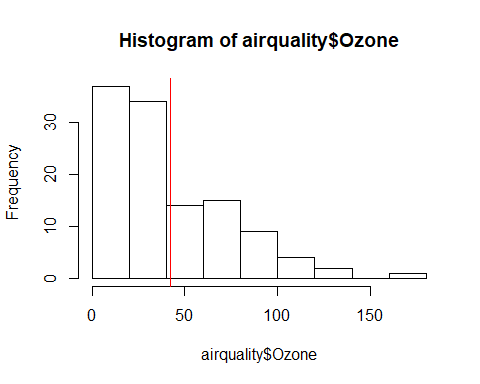
1. Make a table of the Temp values.

table(airquality$Temp)

##   
## 56 57 58 59 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81   
## 1 3 2 2 3 2 1 2 2 3 4 4 3 1 3 3 5 4 4 9 7 6 6 5 11   
## 82 83 84 85 86 87 88 89 90 91 92 93 94 96 97   
## 9 4 5 5 7 5 3 2 3 2 5 3 2 1 1

1. Make a histogram of the Ozone column. Is it a normal distribution? Why or why not?

hist(airquality$Ozone)  
abline(v=mean(airquality$Ozone,na.rm = TRUE), col= "Red")



#mean plotted as a red line on the plot  
#This is not a normal distribution because it does not have a symetry on both sides of its mean

# Functions:

1. Make a simple function to calculate x + 6.

addtox = function(x){  
 x=x+6  
 return(x)}

1. Use that function add 6 to the Temp column in airquality.

airquality$Temp = addtox(airquality$Temp)

# Packages:

1. Install the ggplot2 package.
2. Install the car package.
3. Install the ez package. (no output necessary for these three)
4. Load the car library.

#install.packages("ggplot2")  
#install.packages("car")  
#install.packages("ez")  
  
#I get the error when I try to Knit the output  
# Quitting from lines 198-202 (Lab2.Rmd)   
# Error in contrib.url(repos, "source") :   
# trying to use CRAN without setting a mirror  
# Calls: <Anonymous> ... withVisible -> eval -> eval -> install.packages -> contrib.url  
# Execution halted  
  
  
library(car)

## Warning: package 'car' was built under R version 3.4.4

## Loading required package: carData

## Warning: package 'carData' was built under R version 3.4.4

# Files

1. Import the csv file provided on Moodle

setwd("C:/Users/scyth/OneDrive/Documents/HarrisburgU/AN\_500\_52\_MK/Lab")  
raw = read.csv("Lab2.csv",header = TRUE)  
head(raw)

## variable1 stuff2 thing3  
## 1 3 cheese TRUE  
## 2 4 feta FALSE  
## 3 6 swiss TRUE  
## 4 4 cheese FALSE  
## 5 7 feta FALSE  
## 6 2 swiss TRUE