INSTITUTION: KCA University

NAME: Raphael Kang’eri

UNIT NAME: Programming for Data Science

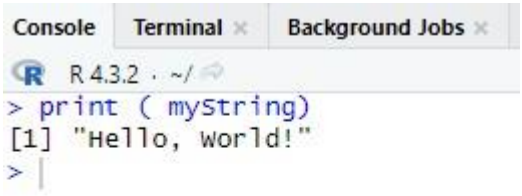
UNIT CODE: BSD 320

INTRODUCTION TO R FOR DATA SCIENCE

1. # My first program in R Programming

   myString <- "Hello, World!"

   print (myString)



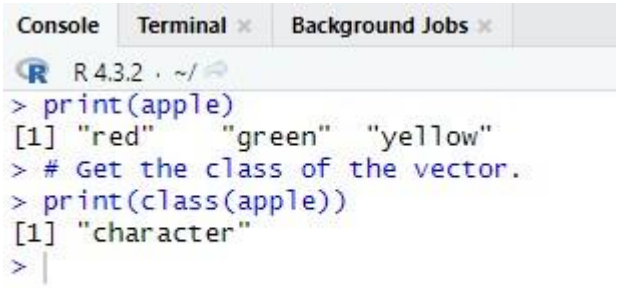
2. # Create a vector.

   apple <- c('red','green',"yellow")

   print(apple)

   # Get the class of the vector.

   print(class(apple))

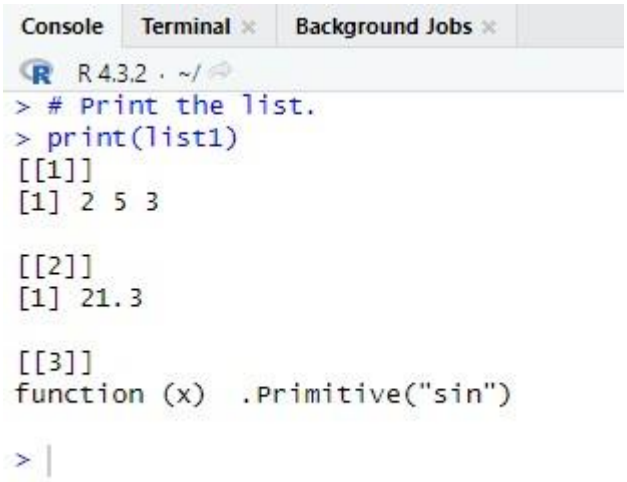


3. # Create a list.

   list1 <- list(c(2,5,3),21.3,sin)

   # Print the list.

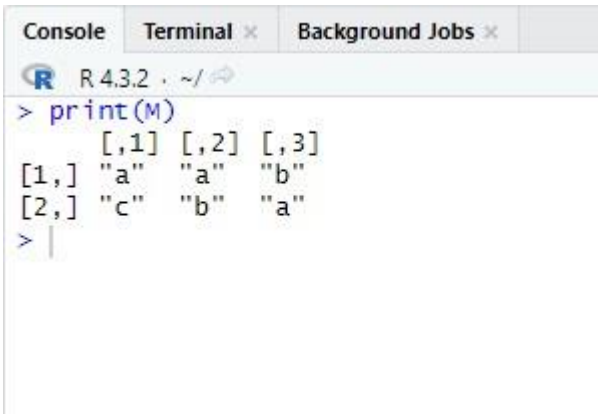
   print(list1)



4. # Create a matrix.

   M = matrix(c('a','a','b','c','b','a'), nrow = 2, ncol = 3, byrow = TRUE)

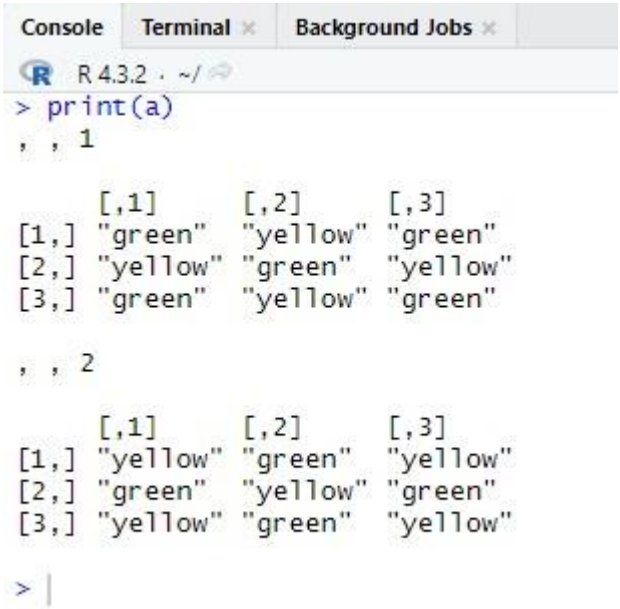
   print(M)



5. # Create an array.

   a <- array(c('green','yellow'),dim = c(3,3,2))

   print(a)



6. # Create a vector.

   apple_colors <- c('green','green','yellow','red','red','red','green')

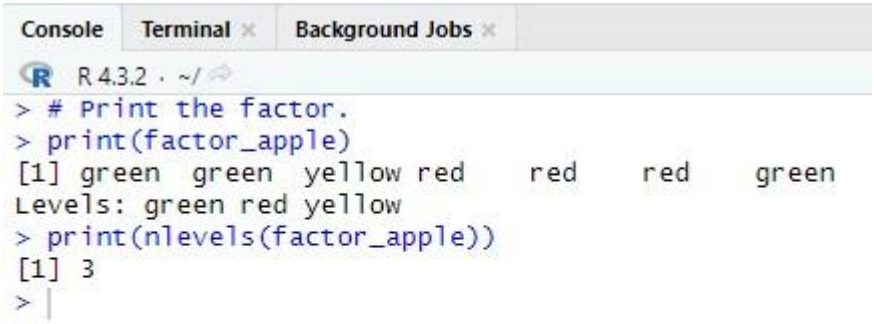
   # Create a factor object.

   factor_apple <- factor(apple_colors)

   # Print the factor.

   print(factor_apple)

   print(nlevels(factor_apple))



7. # Create the data frame.

   BMI <- data.frame(

   gender = c("Male", "Male","Female"),

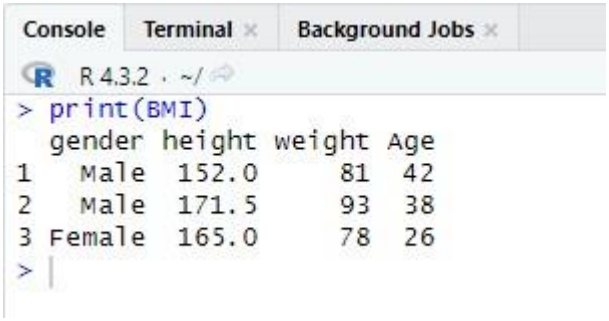
   height = c(152, 171.5, 165),

   weight = c(81,93, 78),

   Age = c(42,38,26)

)

   print(BMI)



8. # Assignment using equal operator.

   var.1 = c(0,1,2,3)

   # Assignment using leftward operator.

   var.2 <- c("learn","R")

   # Assignment using rightward operator.

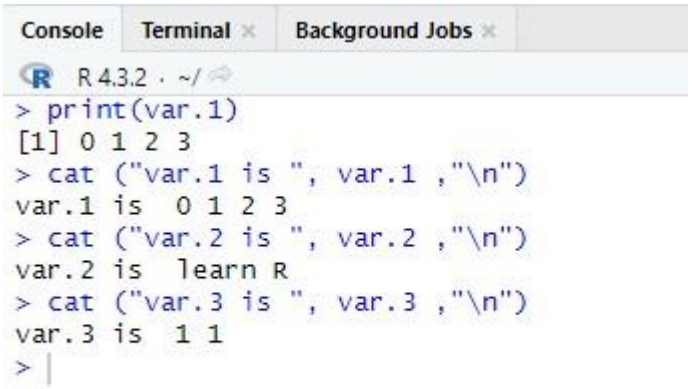
   c(TRUE,1) -> var.3

   print(var.1)

   cat ("var.1 is ", var.1 ,"\n")

   cat ("var.2 is ", var.2 ,"\n")

   cat ("var.3 is ", var.3 ,"\n")



9. var_x <- "Hello"

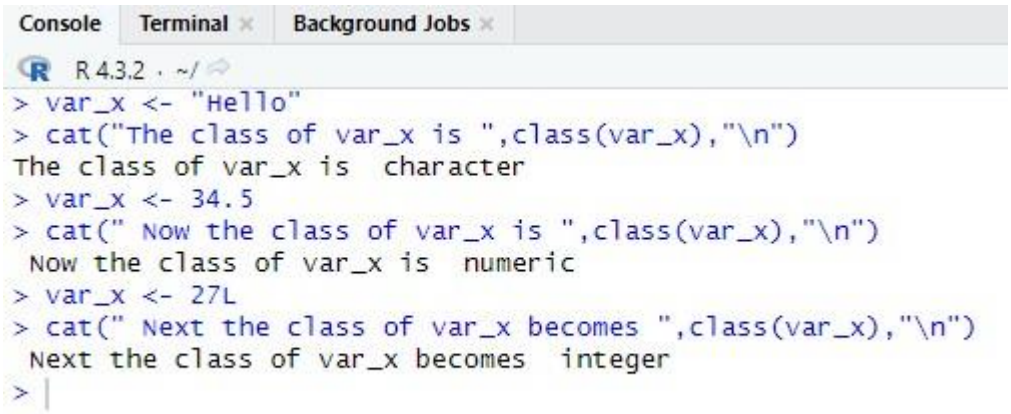
   cat("The class of var_x is ",class(var_x),"\n")

   var_x <- 34.5

   cat(" Now the class of var_x is ",class(var_x),"\n")

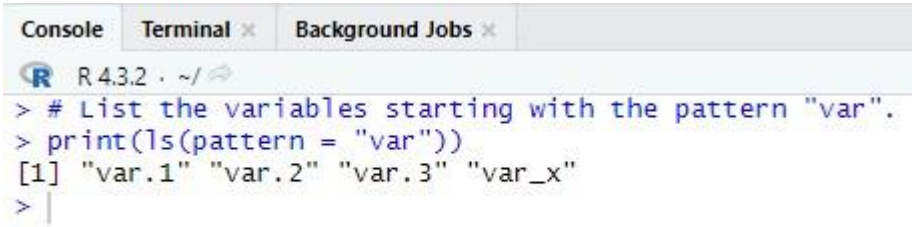
   var_x <- 27L

   cat(" Next the class of var_x becomes ",class(var_x),"\n")

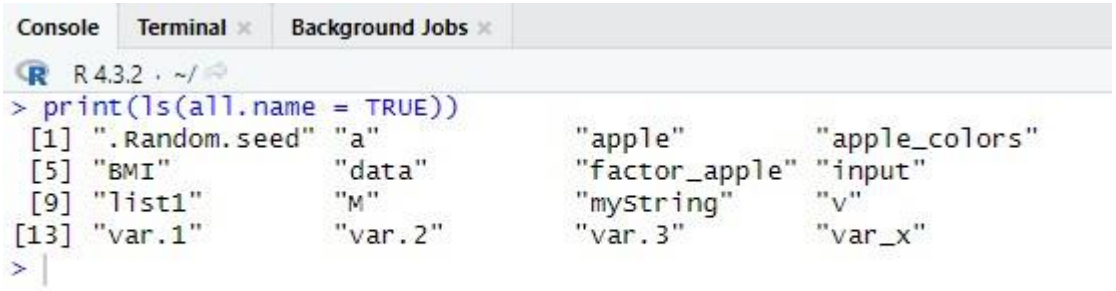


10. # List the variables starting with the pattern "var".

    print(ls(pattern = "var"))



11. print(ls(all.name = TRUE))



1.

INSTITUTION: KCA University

NAME: Raphael Kang’eri

UNIT NAME: Programming for Data Science

UNIT CODE: BSD 320

R-CSV

1. # Get and print current working directory.

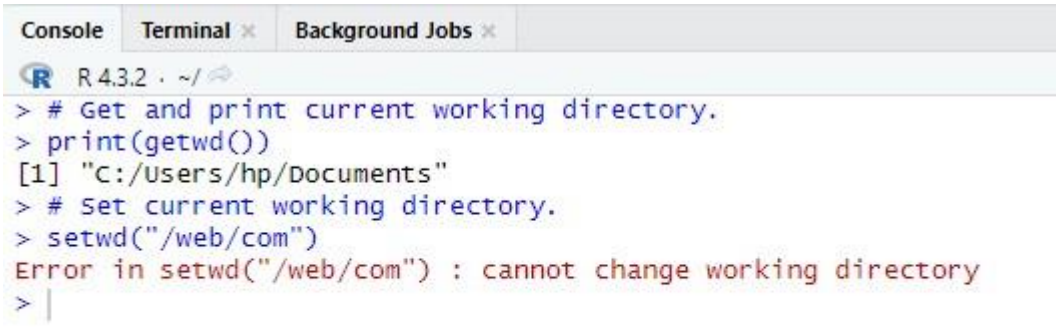
   print(getwd())

   # Set current working directory.

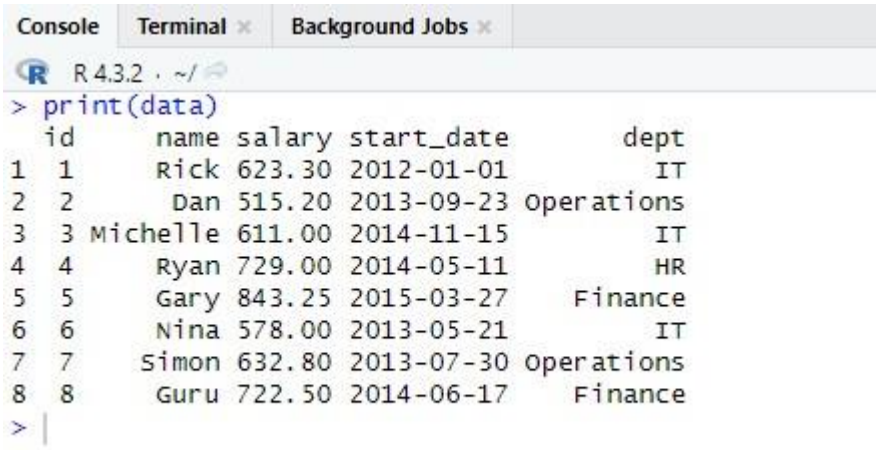
   setwd("/web/com")

   # Get and print current working directory.

   print(getwd())



2. data <- read.csv("input.csv") print(data)

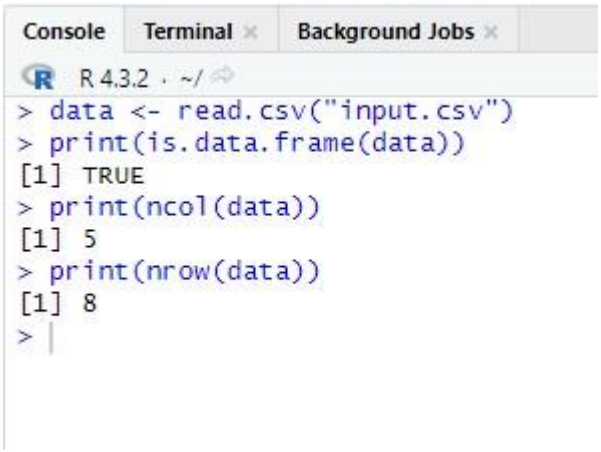


3. data <- read.csv("input.csv")

   print(is.data.frame(data))

   print(ncol(data))

   print(nrow(data))



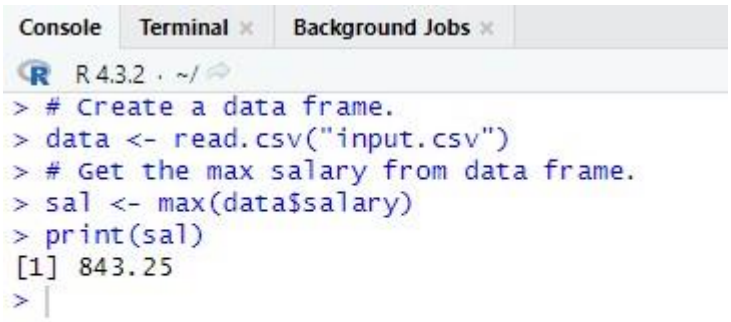
4. # Create a data frame.

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

   # Get the max salary from data frame.

   sal <- max(data$salary)

   print(sal)



5. # Create a data frame.

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

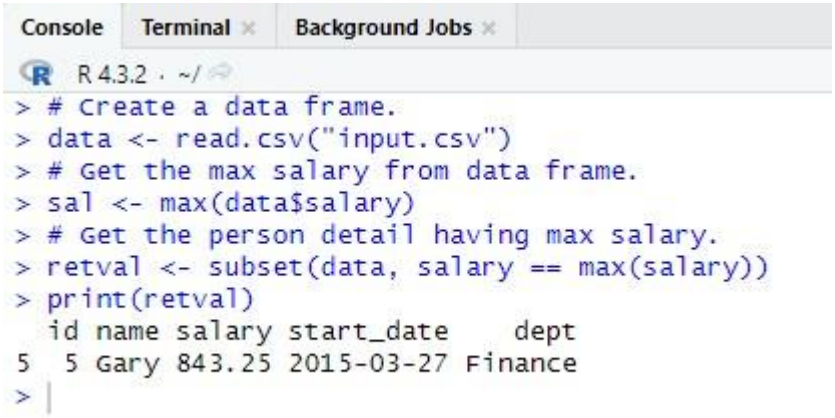
   # Get the max salary from data frame.

   sal <- max(data$salary)

   # Get the person detail having max salary.

   retval <- subset(data, salary == max(salary))

   print(retval)

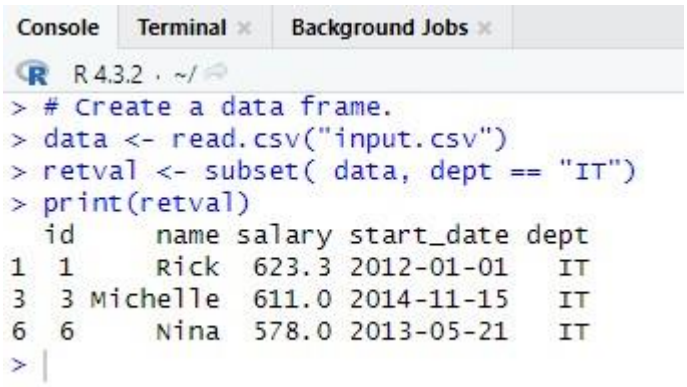


6. # Create a data frame.

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

   retval <- subset(data, dept == "IT")

   print(retval)

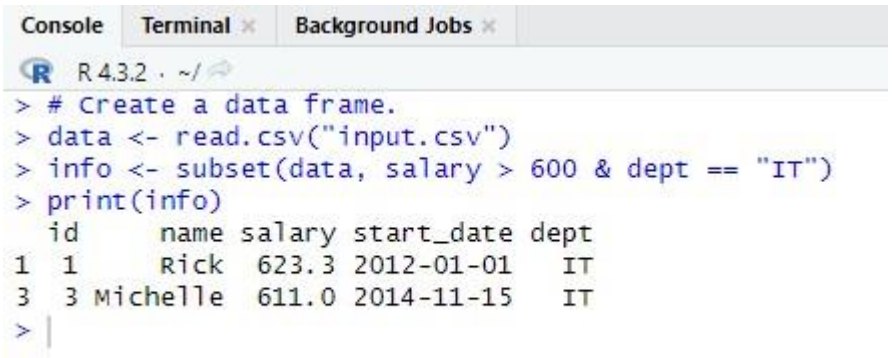


7. # Create a data frame.

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

   info <- subset(data, salary > 600 & dept == "IT")

   print(info)

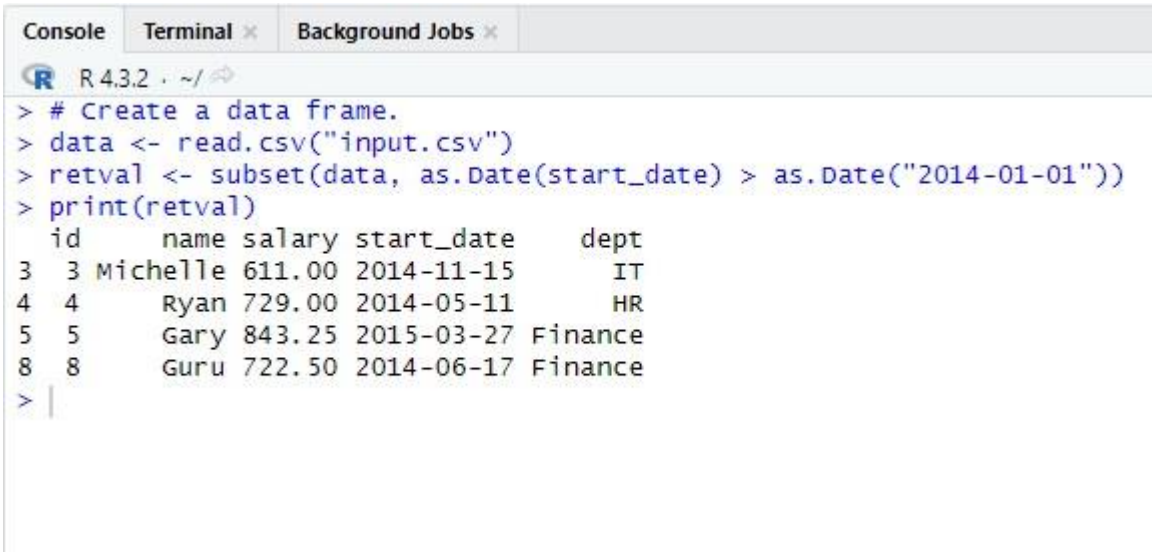


8. # Create a data frame.

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

   retval <- subset(data, as.Date(start_date) > as.Date("2014-01-01"))

   print(retval)



9. # Create a data frame.

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

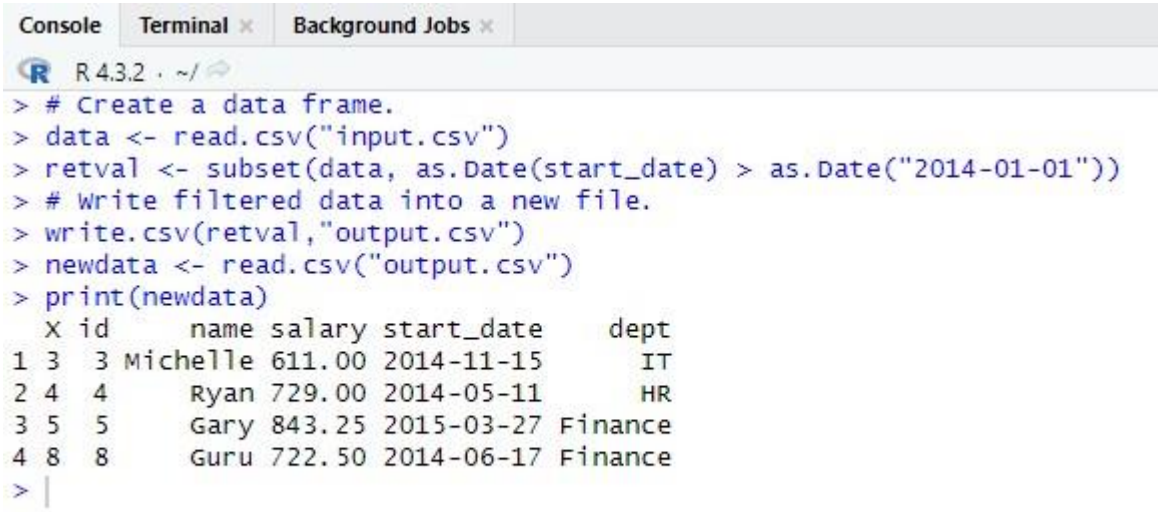
   retval <- subset(data, as.Date(start_date) > as.Date("2014-01-01"))

   # Write filtered data into a new file.

   write.csv(retval,"output.csv")

   newdata <- read.csv("output.csv")

   print(newdata)



10.# Create a data frame.

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

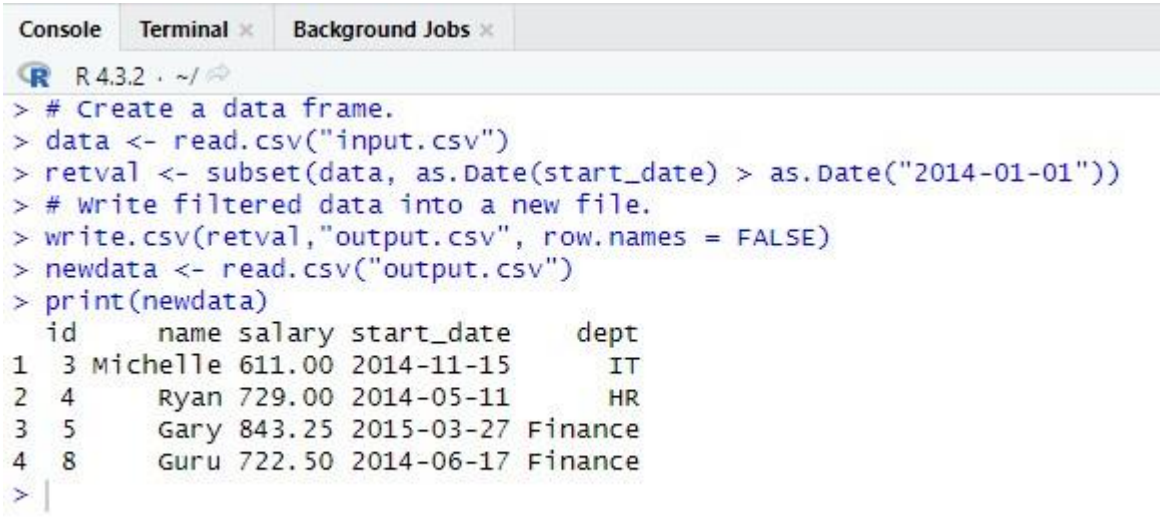
retval <- subset(data, as.Date(start_date) > as.Date("2014-01-01"))

Write filtered data into a new file.

write.csv(retval,"output.csv", row.names = FALSE)

newdata <- read.csv("output.csv")

print(newdata)



INSTITUTION: KCA University

NAME: Raphael Kang’eri

UNIT NAME: Programming for Data Science

UNIT CODE: BSD 320

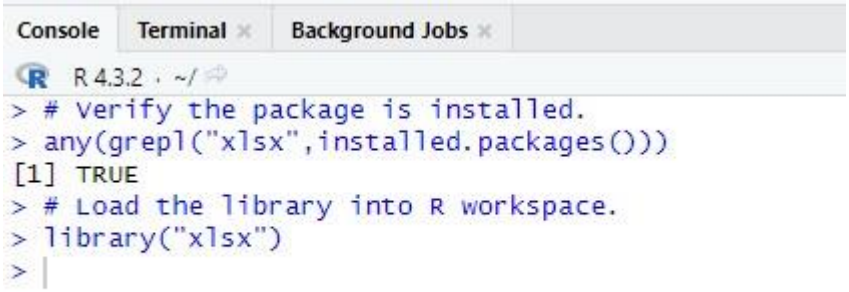
R-EXCEL

1. # Verify the package is installed.

   any(grepl("xlsx",installed.packages()))

   # Load the library into R workspace.

   library("xlsx")



2. # Read the first worksheet in the file input.xlsx.

   data <- read.xlsx("input.xlsx", sheetIndex = 1)

   print(data)