RWorksheet_calvario#2.R.

Jolien

2024-09-25

{r setup, include=FALSE} knitr::opts_chunk$set(echo = TRUE)

1.

a. Sequence from -5 to 5. {r} seq_vec <- -5:5 seq_vec
b. x <- 1:7. What will be the value of x? {r} x <- 1:7 x

2.

a. seq(1, 3, by=0.2) # specify step size {r} vector <- seq(1, 3, by=0.2) print(vector)

3. "'{r} ages <- c(34, 28, 22, 36, 27, 18, 52, 39, 42, 29, 35, 31, 27, 22, 37, 34, 19, 20, 57, 49, 50, 37, 46, 25, 17, 37, 43, 53, 41, 51, 35, 24, 33, 41, 53, 40, 18, 44, 38, 41, 48, 27, 39, 19, 30, 61, 54, 58, 26, 18)

a. Access 3rd element, what is the value?'
```{r}
ages[3]
```

b. Access 2nd and 4th element, what are the values? {r} ages[c(2, 4)]
c. Access all but the 4th and 12th element is not included. Write the R script and its output. {r} ages[-c(4, 12)]

4. *Create a vector x <- c("first"=3, "second"=0, "third"=9). Then named the vector, names(x). {r} x <- c("first" = 3, "second" = 0, "third" = 9) print(names(x))

a. Print the results. Then access x[c("first", "third")]. Describe the output. {r} result <- x[c("first", "third")] print(result)

5. Create a sequence x from -3:2. {r} x <- -3:2 print(x)

a. Modify 2nd element and change it to 0; x[2] <- 0 x Describe the output. {r} x[2] <- 0 print(x)

6. *The following data shows the diesel fuel purchased by Mr. Cruz.

Month Jan Feb March Apr May June Price per liter (PhP) 52.50 57.25 60.00 65.00 74.25 54.00 Purchase–quantity(Liters) 25 30 40 50 10 45

a. Create a data frame for month, price per liter (php) and purchase-quantity (liter). Write the R scripts and its output. "'{r} # Creating vectors for each column month <- c("Jan", "Feb", "Mar", "Apr", "May", "Jun") price_per_liter <- c(52.50, 57.25, 60.00, 65.00, 74.25, 54.00) purchase_quantity <- c(25, 30, 40, 50, 10, 45)

Creating the data frame

fuel_data <- data.frame(month, price_per_liter, purchase_quantity)

Display the data frame

print(fuel_data)

```
b. What is the average fuel expenditure of Mr. Cruz from
Jan to June? Note: Use ‘weighted.mean(liter,
purchase)‘. Write the R scripts and its output.
```{r}
# Calculating the total expenditure for each month
expenditure <- price_per_liter * purchase_quantity

# Calculating the average fuel expenditure using weighted.mean
average_expenditure <- weighted.mean(price_per_liter, purchase_quantity)

# Display the average expenditure
print(average_expenditure)
```

7. R has actually lots of built-in datasets. For example, the rivers data “gives the lengths

(in miles) of 141 “major” rivers in North America, as compiled by the US Geological Survey”.

a. Type “rivers” in your R console.

Create a vector data with 7 elements, containing the number of elements (length) in rivers, their sum (sum), mean (mean), median(median), variance(var), standard deviation(sd), minimum (min) and maximum (max). data <- c(length(rivers), sum(rivers), mean(rivers), median(rivers), var(rivers), sd(rivers), min(rivers), max(rivers))

“‘{r} data <- c(length(rivers), sum(rivers), mean(rivers), median(rivers), var(rivers), sd(rivers), min(rivers), max(rivers))

print(data)

```
# 8. The table below gives the 25 most powerful celebrities and their annual pay as ranked
by the editions of Forbes magazine and as listed on the Forbes.com website.

# a. Create vectors according to the above table.
Write the R scripts and its output.
```{r}
# Creating vectors
names <- c("Tom Cruise", "Rolling Stones", "Oprah Winfrey", "U2", "Tiger Woods",
           "Steven Spielberg", "Howard Stern", "50 Cent", "Cast of the Sopranos", "Dan Brown",
           "Bruce Springsteen", "Donald Trump", "Muhammad Ali", "Paul McCartney", "George Lucas",
           "Elton John", "David Letterman", "Phil Mickelson", "J.K Rowling", "Bradd Pitt",
           "Peter Jackson", "Dr. Phil McGraw", "Jay Lenon", "Celine Dion", "Kobe Bryant")
power_ranking <- 1:25
annual_pay <- c(67, 90, 225, 110, 90, 332, 302, 41, 52, 88,
                55, 44, 55, 40, 233, 34, 40, 47, 75, 25,
                39, 45, 32, 40, 31)

celebrity_data <- data.frame(names, power_ranking, annual_pay)

print(celebrity_data)
```

b. Modify the power ranking and pay of J.K. Rowling.

Change power ranking to 15 and pay to 90. Write the R scripts and its output. “‘{r} celebrity_data$power_ranking[celebrity_data$ar
== “J.K. Rowling”] <- 19 celebrity_data$annual_pay[celebrity_data$names == “J.K. Rowling”] <- 75

print(celebrity_data)

```
# c. Create an excel file from the table above and save it
as csv file(PowerRanking). Import the csv file into
the RStudio. What is the R script?
```{r}
write.csv(celebrity_data, "PowerRanking.csv", row.names = FALSE)

imported_data <- read.csv("PowerRanking.csv")

print(imported_data)
```

d. Access the rows 10 to 20 and save it as Ranks.RData.

Write the R script and its output. “‘{r} selected_rows <- celebrity_data[10:20,]

save(selected_rows, file = “Ranks.RData”)

```
# e. Describe its output.
# The output from the print(selected_rows) command will display a subset of the original data frame con


# 9a. Import the Excel file
library(readxl)
data <- read_excel("hotels-vienna.xlsx")

# 9b. Check dimensions
dimensions <- dim(data)
print(dimensions)

# 9c. Select specific columns
library(dplyr)
selected_data <- data %>%
  select(country, neighbourhood, price, stars, accommodation_type, rating)

# 9d. Save the data
save(selected_data, file = "new.RData")

# 9e. Load the saved data and display rows
load("new.RData")
head(selected_data)
tail(selected_data)

# 10. Create a list of ten (10) vegetables you ate during your lifetime. If none, just list down.

# a. Write the R scripts and its output.
```{r}
```