# Worksheet-2 in R

**Worksheet for R Programming**

**Instructions:**

- Use RStudio or the RStudio Cloud to accomplish this worksheet.
- Save the R script as Worksheet_lastname #2.R.
- Commit and push the R script and your Rmarkdown file in html to your own repo. Do not forget to comment your Git repo
  Accomplish this worksheet by answering the questions being asked and writing the code manually

.

## Using Vectors

1. Create a vector using : operator

a. Sequence from -5 to 5. Write the R code and its output.
Describe its output.

```
vector_1 <- -5:5
vector_1
```

```
[1] -5 -4 -3 -2 -1  0  1  2  3  4  5
```

The code uses the {:} operator to generate a vector from {-5} to {5}, resulting in a sequence of numbers in that range. The numbers from {-5} to {5}, increased by 1, are included in the output.

b. x <- 1:7. What will be the value of x?

```
[1] 1 2 3 4 5 6 7
```

2. * Create a vector using seq() function

a. seq(1, 3, by=0.2) # specify step size
Write the R script and its output. Describe the output.

```
vector_2 <- seq(1, 3, by=0.2)
vector_2
```

```
[1] 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0
```

The seq() function creates a sequence starting at 1, ending at 3, with a step size of 0.2. The output is a numeric vector with values incremented by 0.2 from 1 to 3.

3. A factory has a census of its workers. There are 50 workers in total. The following list shows their ages: 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?

```
[1] 22
```

b. Access 2nd and 4th element, what are the values?

```
[1] 28 36
```

c. Access all but the 4th and 12th element is not included. Write the R script and its output.

```
[1] 34 28 22 27 18 52 39 42 29 35 27 22 37 34 19 20 57 49 50 37 46 2
5 17 37
[25] 43 53 41 51 35 24 33 41 53 40 18 44 38 41 48 27 39 19 30 61 54
58 26 18
```

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

a. Print the results. Then access x [c("first", "third")].
Describe the output.

The vector `x` is created with named elements: "first" (3), "second" (0), and "third" (9). Accessing `x[c("first", "third")]` retrieves the values `3` and `9`, corresponding to the names "first" and "third". The output shows only these two values, excluding "second".

b.Write the code and its output.

x[c("first", "third")]

```
first third
    3     9
```

5. Create a sequence x from -3:2.

a. Modify 2nd element and change it to 0;

```
x [2] <- 0
x
```

Describe the output.

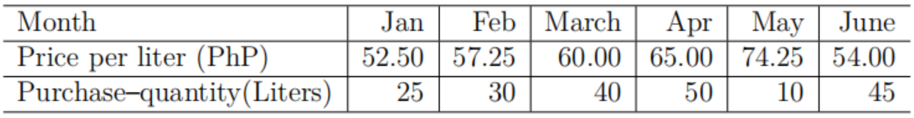
The sequence x is created with values from -3 to 2. After modifying the 2nd element to 0, the output updates the sequence to -3 0 -1 0 1 2, where the original -2 is replaced by 0

b. Write the code and its output.

```
x[2] <- 0
x
```

```
[1] -3  0 -1  0  1  2
```

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.

```
cruz_data <- data.frame(
  Month = c("Jan", "Feb", "March", "Apr", "May", "June"),
  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)
)
cruz_data
```

```
Month Price_per_liter Purchase_quantity
1   Jan           52.50                25
2   Feb           57.25                30
3 March           60.00                40
4   Apr           65.00                50
5   May           74.25                10
6  June           54.00                45
```

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.

```
cruz_data$Total_expenditure <- cruz_data$Price_per_liter * cruz_data$Purchase_quantity
average_expenditure <- weighted.mean(cruz_data$Total_expenditure, cruz_data$Purchase_q
uantity)
average_expenditure
```

```
[1] 2298.062
```

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))

b.What are the results?

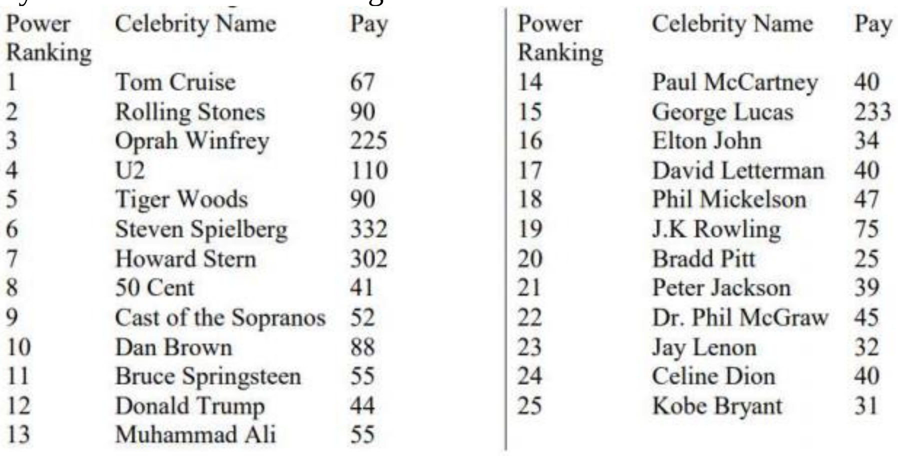
```
[1]    141.0000  83357.0000    591.1844    425.0000 243908.4086    4
93.8708
[7]    135.0000   3710.0000
```

c.  Write the R scripts and its outputs.

```
data("rivers")
data <- c(length(rivers), sum(rivers), mean(rivers), median(rivers),
         var(rivers), sd(rivers), min(rivers), max(rivers))
data
```

```
[1]    141.0000  83357.0000    591.1844    425.0000 243908.4086    4
93.8708
[7]    135.0000   3710.0000
```

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.



| Power Ranking | Celebrity Name | Pay | Power Ranking | Celebrity Name | Pay |
|---|---|---|---|---|---|
| 1 | Tom Cruise | 67 | 14 | Paul McCartney | 40 |
| 2 | Rolling Stones | 90 | 15 | George Lucas | 233 |
| 3 | Oprah Winfrey | 225 | 16 | Elton John | 34 |
| 4 | U2 | 110 | 17 | David Letterman | 40 |
| 5 | Tiger Woods | 90 | 18 | Phil Mickelson | 47 |
| 6 | Steven Spielberg | 332 | 19 | J.K Rowling | 75 |
| 7 | Howard Stern | 302 | 20 | Bradd Pitt | 25 |
| 8 | 50 Cent | 41 | 21 | Peter Jackson | 39 |
| 9 | Cast of the Sopranos | 52 | 22 | Dr. Phil McGraw | 45 |
| 10 | Dan Brown | 88 | 23 | Jay Lenon | 32 |
| 11 | Bruce Springsteen | 55 | 24 | Celine Dion | 40 |
| 12 | Donald Trump | 44 | 25 | Kobe Bryant | 31 |
| 13 | Muhammad Ali | 55 | | | |

Figure 1: Forbes Ranking

a. Create vectors according to the above table.
Write the R scripts and its output.

```
power_ranking <- c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13,
                   14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25)
celebrity_name <- 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")
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)

print(power_ranking)
print(celebrity_name)
print(pay)
```

```
[1]  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 2
2 23 24
[25] 25
```

```
[1] "Tom Cruise"
 [2] "Rolling Stones"
 [3] "Oprah Winfrey"
 [4] "U2"
 [5] "Tiger Woods"
 [6] "Steven Spielberg"
 [7] "Howard Stern"
 [8] "50 Cent"
 [9] "Cast of the Sopranos"
[10] "Dan Brown"
[11] "Bruce Springsteen"
[12] "Donald Trump"
[13] "Muhammad Ali"
[14] "Paul McCartney"
[15] "George Lucas"
[16] "Elton John"
[17] "David Letterman"
[18] "Phil Mickelson"
[19] "J.K Rowling"
[20] "Bradd Pitt"
[21] "Peter Jackson"
[22] "Dr. \n                    Phil McGraw"
[23] "Jay Lenon"
[24] "Celine Dion"
[25] "Kobe Bryant"
```

```
 [1]  67  90 225 110  90 332 302  41  52  88  55  44  55  40 233  34  40  47
 [19]  75  25  39  45  32  40  31
```

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.

```
power_ranking[19] <- 15
pay[19] <- 90

print(power_ranking)
print(pay)
```

```
[1]  67  90 225 110  90 332 302  41  52  88  55  44  55  40 233  34
 40  47
[19]  90  25  39  45  32  40  31
```

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?

```
power_ranking_data <- data.frame(power_ranking, celebrity_name, pay)
write.csv(power_ranking_data, "PowerRanking.csv", row.names = FALSE)
power_ranking_data <- read.csv("PowerRanking.csv")
print(power_ranking_data)
```

d. Access the rows 10 to 20 and save it as Ranks.RData
Write the R script and its output.

```
power_ranking_data <- read.csv("PowerRanking.csv")
ranks_data <- power_ranking_data[10:20, ]
print(ranks_data)
save(ranks_data, file = "Ranks.RData")
```

```
power_ranking    celebrity_name pay
10            10         Dan Brown  88
11            11 Bruce Springsteen  55
12            12      Donald Trump  44
13            13      Muhammad Ali  55
14            14    Paul McCartney  40
15            15      George Lucas 233
16            16        Elton John  34
17            17   David Letterman  40
18            18    Phil Mickelson  47
19            15       J.K Rowling  90
20            20        Bradd Pitt  25
```

d. Describe its output.

A data frame containing rows 10 through 20 from the `power_ranking_data} is displayed in the output of `print(ranks_data)`. J.K. Rowling, Dan Brown, and other celebrities are featured in the `pay}, `celebrity_name}, and `power_ranking} columns. rowing in tandem with their incomes and ranks. The rankings and compensation for these celebrities are clearly seen in this selection.

9. Download the Hotels-Vienna https://tinyurl.com/ Hotels- Vienna

a. Import the excel file into your RStudio.
What is the R. script?

```
install.packages("readxl")
library(readxl)
file_path <- "C:/Users/ASITSD/Downloads/hotels-vienna.xlsx"
hotels_data <- read_excel(file_path)
print(hotels_data)
```

b.How many dimensions does the dataset have?
What is the R script? What is its output?

```
dim(hotels_data)
```

```
[1] 428  24
```

c. Select columns country, neighbourhood,
price, stars, accomodation_type, and
ratings. Write the R script.

```
library(dplyr)
selected_data <- select(hotels_data, country, neighbourhood,
                        price, stars, accommodation_type, rating)
head(selected_data)
```

d. Save the data as **new.RData to your RStudio. Write the
R. script.

```
save(selected_data, file = "new.RData")
```

d. Display the first six rows and last six rows of the new.RData. What is the R script?

```
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.

```
vegetable_list <- c("Lettuce", "Potato", "Cabbage", "Eggplant", "Spinach",
"Bittergour",  "Cucumbers","Bell Peppers", "Lady Finger", "Mushrooms")
print(vegetable_list)
```

```
[1]  "Lettuce"      "Potato"       "Cabbage"      "Eggplant"
 [5] "Spinach"      "Bittergour"   "Cucumbers"    "Bell Peppers"
 [9] "Lady Finger"  "Mushrooms"    "Spinach"      "Brocolli"
```

```
vegetable_list <- c(vegetable_list, "Spinach", "Brocolli")

print(vegetable_list)
```

b. Add 2 additional vegetables after the last vegetables in the list. What is the Rscript and its output.

```
vegetable_list <- c(vegetable_list, "Spinach", "Brocolli")
print(vegetable_list)
```

```
[1]  "Lettuce"      "Potato"       "Cabbage"      "Eggplant"
 [5] "Spinach"      "Tomato"       "Ginger"       "Carrot"
 [9] "Corn"         "Bittergour"   "Cucumbers"    "Bell Peppers"
[13] "Lady Finger"  "Mushrooms"    "Spinach"      "Brocolli"
[17] "Spinach"      "Brocolli"
```

c. Add 4 additional vegetables after index 5. How many datapoints does your vegetable list have? What is the R script and its output?

```
egetable_list <- c(vegetable_list[1:5], "Tomato", "Ginger",
          "Carrot", "Corn", vegetable_list[6:12])
print(vegetable_list)
```

```
[1]  "Lettuce"      "Potato"       "Cabbage"      "Eggplant"
 [5] "Spinach"      "Tomato"       "Ginger"       "Carrot"
 [9] "Corn"         "Bittergour"   "Cucumbers"    "Bell Peppers"
[13] "Lady Finger"  "Mushrooms"    "Spinach"      "Brocolli"
[17] "Spinach"      "Brocolli"
```

d. Remove the vegetables in index 5. 10. and 15. How many vegetables were left? Write the codes and its output.

```
vegetable_list <- vegetable_list[-c(5, 10, 15)]

print(vegetable_list)
```

```
[1]  "Lettuce"      "Potato"       "Cabbage"      "Eggplant"
 [5] "Tomato"       "Ginger"       "Carrot"       "Corn"
 [9] "Cucumbers"    "Bell Peppers" "Lady Finger"  "Mushrooms"
[13] "Brocolli"     "Spinach"      "Brocolli"
```

Note: Do not forget to push into your GitHub repo.

Prepared by:
JOYCE F. JAMILE
Instructor
Without ethical considerations, Al becomes a tool of chaos and harm.