

RWorkshee__Sobusa#2.rmd

Nexon Sobusa

2024-10-02

1.a Creates a sequence from -5 to 5

```
operator <- -5:5 operator
```

1.b Creates a sequence from 1 to 7

```
x <- 1:7 x
```

2. Creates a sequence from 1 to 3 with increments of 0.2

```
seq(1, 3, by = 0.2)
```

3. Manipulating the ‘ages’ vector

3.a Accesses the 3rd element (22)

```
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) age_3 <- ages[3] age_3
```

3.b Accesses the 2nd and 4th elements (28 and 36)

```
secondfourth <- ages[c(2, 4)] secondfourth
```

3.c Removes the 4th and 12th elements

```
allbut4and12 <- ages[-c(4, 12)] allbut4and12
```

4. Named vector operations

4.a Creates a named vector and accesses elements by name

```
x <- c("first" = 3, "second" = 0, "third" = 9) x x[c("first", "third")]
```

5 Modifying vectors

5.a Replaces the 2nd element of x with 0

```
x <- -3:2 x[2] <- 0 x
```

6. Data frame operations

6.a Creates a data frame with months, fuel prices, and quantities

```
month <- c("Jan", "Feb", "March", "Apr", "May", "June") ppl <- c(52.50, 57.25, 60.00, 65.00, 74.25, 54.00)
pq <- c(25, 30, 40, 50, 10, 45) df <- data.frame(month, ppl, pq) df
```

6.b Calculates the weighted mean (average fuel expenditure)

```
afe <- weighted.mean(ppl, pq) afe
```

7. Summary statistics for the rivers dataset

7.a Calculates various statistics for the ‘rivers’ dataset

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

8. Celebrity rankings and manipulations

8.a Creates vectors for rankings, celebrities, and their pay

```
pr <- 1:25 celebrity <- c("Tom Cruise", "Rolling Stones", "Oprah Winfrey", "U2", "Tiger Woods", "Steven
Spielberg", "Howard Stern", "50 Cent", "Cast of the Sopranos", "Dan Brown", "Bruce Springsteen", "Donal
Trump", "Muhammad Ali", "Paul McCartney", "George Lucas", "Elton John", "David Letterman", "Phil
Mickelson", "J.K. Rowling", "Brad Pitt", "Peter Jackson", "Dr. Phil McGraw", "Jay Leno", "Celine Dion",
"Kobe Bryant") pay <- c(67, 90, 225, 110, 90, 332, 302, 41, 52, 88, 55, 44, 55, 40, 233, 60, 40, 47, 75, 25, 39,
45, 32, 40, 31) pr celebrity pay
```

8.b Updates ranking and pay for J.K. Rowling

```
Change <- which(celebrity == "J.K. Rowling") pr[Change] <- 15 pay[Change] <- 90 pr pay
```

8.c Create a CSV file from the table, save it as PowerRanking.csv, and then import it into RStudio.

```
ranking_data <- data.frame(Rank = pr, Celebrity = celebrity, Pay = pay) write.csv(ranking_data, file =
"C:/Users/kurts/Desktop/R-Code/RWorksheet_2/PowerRanking.csv", row.names = FALSE)
```

Now try to read it

```
if (file.exists("C:/Users/kurts/Desktop/R-Code/RWorksheet_2/PowerRanking.csv")) { power_ranking_data
<- read.csv("C:/Users/kurts/Desktop/R-Code/RWorksheet_2/PowerRanking.csv") print(power_ranking_data)
} else { stop("File 'PowerRanking.csv' does not exist.") }
```

8.d Access rows 10 to 20, save it as Ranks.RData, and provide the R code for the process.

```
subset_ranks <- power_ranking_data[10:20, ] save(subset_ranks, file = "Ranks.RData") load("Ranks.RData")
subset_ranks
```

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

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

```
hotels_vienna_data <- read_excel("C:/Users/kurts/Desktop/R-Code/RWorksheet_2/hotels-vienna.xlsx")
hotels_vienna_data
```

9.b Checks the dimensions of the hotelsvienna dataset

```
dimensions <- dim(hotels_vienna_data)
```

9.c Select columns country, neighbourhood, price, stars, accommodation_type, and ratings

```
selected_columns <- hotels_vienna_data[, c("country", "neighbourhood", "price", "stars", "accommodation_type", "rating")] selected_columns
```

9.d Save the data as new.RData

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

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

```
load("new.RData") first_six_rows <- head(selected_columns) first_six_rows last_six_rows <-
tail(selected_columns) last_six_rows
```

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

10.a Creates a list of vegetables

```
vegetables <- c("Carrot", "Broccoli", "Spinach", "Tomato", "Potato", "Cucumber", "Zucchini", "Bell Pepper",
"Cauliflower", "Onion") vegetables
```

10.b Adds more vegetables to the list

```
vegetables <- c(vegetables, "Eggplant", "Asparagus") vegetables
```

10.c Add 4 additional vegetables after index 5. How many datapoints does your vegetable list have?

```
vegetables <- append(vegetables, c("Lettuce", "Radish", "Beetroot", "Pumpkin"), after = 5) total_vegetables  
<- length(vegetables) vegetables total_vegetables
```

10.d Removes 3 specific vegetables from the list

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