

Challenge-3

Chen Zi Xin

2023-08-28

I. Questions

Question 1: Emoji Expressions Imagine you're analyzing social media posts for sentiment analysis. If you were to create a variable named "postSentiment" to store the sentiment of a post using emojis (for positive, for neutral, for negative), what data type would you assign to this variable? Why? (*narrative type question, no code required*)

Solution: I would use "String" as the emojis are manipulated by words such as "positive".

Question 2: Hashtag Havoc In a study on trending hashtags, you want to store the list of hashtags associated with a post. What data type would you choose for the variable "postHashtags"? How might this data type help you analyze and categorize the hashtags later? (*narrative type question, no code required*)

Solution: List of strings as it creates a structured data pool for Hashtags of different lengths which can be analysed further.

Question 3: Time Traveler's Log You're examining the timing of user interactions on a website. Would you use a numeric or non-numeric data type to represent the timestamp of each interaction? Explain your choice (*narrative type question, no code required*)

Solution: Numeric as timestamp can include up to several decimal places and even running numbers thus numeric is most suitable to have a more accurate data.

Question 4: Event Elegance You're managing an event database that includes the date and time of each session. What data type(s) would you use to represent the session date and time? (*narrative type question, no code required*)

Solution: DateTime as it stores both data in a single element.

Question 5: Nominee Nominations You're analyzing nominations for an online award. Each participant can nominate multiple candidates. What data type would be suitable for storing the list of nominated candidates for each participant? (*narrative type question, no code required*)

Solution: I would create a list to store and organise the data.

Question 6: Communication Channels In a survey about preferred communication channels, respondents choose from options like "email," "phone," or "social media." What data type would you assign to the variable "preferredChannel"? (*narrative type question, no code required*)

Solution: Categorical nominal as there is no natural ordering to the options.

Question 7: Colorful Commentary In a design feedback survey, participants are asked to describe their feelings about a website using color names (e.g., “warm red,” “cool blue”). What data type would you choose for the variable “feedbackColor”? (*narrative type question, no code required*)

Solution: Character, Non numeric, Nominal

Question 8: Variable Exploration Imagine you’re conducting a study on social media usage. Identify three variables related to this study, and specify their data types in R. Classify each variable as either numeric or non-numeric.

Solution: 1. Character (Non-numeric) 2.Double (Numeric) 3.Logical (Non-numeric)

Question 9: Vector Variety Create a numeric vector named “ages” containing the ages of five people: 25, 30, 22, 28, and 33. Print the vector.

Solution:

```
x<-c(25, 30, 22, 28,33)
ages<-x
print(ages)
```

```
## [1] 25 30 22 28 33
```

Question 10: List Logic Construct a list named “student_info” that contains the following elements:

- A character vector of student names: “Alice,” “Bob,” “Catherine”
- A numeric vector of their respective scores: 85, 92, 78
- A logical vector indicating if they passed the exam: TRUE, TRUE, FALSE

Print the list.

Solution:

```
student_info<-list(names=c("Alice", "Bob", "Catherine"),scores=c(85, 92, 78),passed=c(TRUE, TRUE, FALSE))
print(student_info)
```

```
## $names
## [1] "Alice"      "Bob"        "Catherine"
##
## $scores
## [1] 85 92 78
##
## $passed
## [1] TRUE TRUE FALSE
```

Question 11: Type Tracking You have a vector “data” containing the values 10, 15.5, “20”, and TRUE. Determine the data types of each element using the typeof() function.

Solution:

```
x<-c(10, 15.5, "20",TRUE)
typeof(x)
```

```
## [1] "character"
```

Question 12: Coercion Chronicles You have a numeric vector “prices” with values 20.5, 15, and “25”. Use explicit coercion to convert the last element to a numeric data type. Print the updated vector.

Solution:

```
prices<-c(20.5, 15,"25")
as.numeric("25")
```

```
## [1] 25
```

```
print(prices)
```

```
## [1] "20.5" "15"  "25"
```

Question 13: Implicit Intuition Combine the numeric vector `c(5, 10, 15)` with the character vector `c("apple", "banana", "cherry")`. What happens to the data types of the combined vector? Explain the concept of implicit coercion.

Solution:

```
x<-c(5,10,15)
y<-c("apple", "banana", "cherry")
z<-c(x,y)
typeof(z)
```

```
## [1] "character"
```

```
#It refers to the automatic conversion of data.
```

Question 14: Coercion Challenges You have a vector “numbers” with values 7, 12.5, and “15.7”. Calculate the sum of these numbers. Will R automatically handle the data type conversion? If not, how would you handle it?

Solution:

```
x<-c(7, 12.5,"15.7")
#sum(x)
#Error in sum(x) : invalid 'type' (character) of argument

#No, R does not automatically convert.

x<-as.double(c(7, 12.5,"15.7"))
sum(x)
```

```
## [1] 35.2
```

Question 15: Coercion Consequences Suppose you want to calculate the average of a vector “grades” with values 85, 90.5, and “75.2”. If you directly calculate the mean using the mean() function, what result do you expect? How might you ensure accurate calculation?

Solution:

```
mean(85, 90.5, "75.2")
```

```
## [1] 85
```

```
#Ensure that all the data are the same type before finding the mean.
```

```
x<-as.double(c(85, 90.5, "75.2"))
mean(x)
```

```
## [1] 83.56667
```

Question 16: Data Diversity in Lists Create a list named “mixed_data” with the following components:

- A numeric vector: 10, 20, 30
- A character vector: “red”, “green”, “blue”
- A logical vector: TRUE, FALSE, TRUE

Calculate the mean of the numeric vector within the list.

Solution:

```
mixed_data<-list("numeric"=c(10,20,30),"character"=c("red", "green", "blue"),"logical"=c(TRUE, FALSE, TRUE))
mean(mixed_data$numeric)
```

```
## [1] 20
```

Question 17: List Logic Follow-up Using the “student_info” list from Question 10, extract and print the score of the student named “Bob.”

Solution:

```
student_info
```

```
## $names
## [1] "Alice"      "Bob"        "Catherine"
##
## $scores
## [1] 85 92 78
##
## $passed
## [1] TRUE TRUE FALSE
```

```
bob_score<-student_info$scores[student_info$names=="Bob"]
bob_score
```

```
## [1] 92
```

Question 18: Dynamic Access Create a numeric vector values with random values. Write R code to dynamically access and print the last element of the vector, regardless of its length.

Solution:

```
vec1 <-c(12,TRUE,24.5)
tail(vec1,1)
```

```
## [1] 24.5
```

Question 19: Multiple Matches You have a character vector words <- c("apple", "banana", "cherry", "apple"). Write R code to find and print the indices of all occurrences of the word "apple."

Solution:

```
x<- c("apple", "banana", "cherry", "apple")
which(x=="apple")
```

```
## [1] 1 4
```

Question 20: Conditional Capture Assume you have a vector ages containing the ages of individuals. Write R code to extract and print the ages of individuals who are older than 30.

Solution:

```
age<-c(23,45,23,65)
older<-age[age>30]
older
```

```
## [1] 45 65
```

Question 21: Extract Every Nth Given a numeric vector sequence <- 1:20, write R code to extract and print every third element of the vector.

Solution:

```
x<- 1:20
nth<-x[seq(from=1,to=20,by=3)]
nth
```

```
## [1] 1 4 7 10 13 16 19
```

Question 22: Range Retrieval Create a numeric vector numbers with values from 1 to 10. Write R code to extract and print the values between the fourth and eighth elements.

Solution:

```
x<-c(1:10)
print(x[4:8])
```

```
## [1] 4 5 6 7 8
```

Question 23: Missing Matters Suppose you have a numeric vector `data <- c(10, NA, 15, 20)`. Write R code to check if the second element of the vector is missing (NA).

Solution:

```
X <- c(10, NA, 15, 20)
print(is.na(X[2]))
```

```
## [1] TRUE
```

Question 24: Temperature Extremes Assume you have a numeric vector `temperatures` with daily temperatures. Create a logical vector `hot_days` that flags days with temperatures above 90 degrees Fahrenheit. Print the total number of hot days.

Solution:

```
daily_temperature<-c(345,2,23,95,34,30)
hot_days<-daily_temperature[daily_temperature>90]
num_hot_days<-length(hot_days)
num_hot_days
```

```
## [1] 2
```

Question 25: String Selection Given a character vector `fruits` containing fruit names, create a logical vector `long_names` that identifies fruits with names longer than 6 characters. Print the long fruit names.

Solution:

```
vect3<-c("Apple", "Pear", "Mangosteen")
long_name<-vect3[nchar(vect3)>6]
long_name
```

```
## [1] "Mangosteen"
```

Question 26: Data Divisibility Given a numeric vector `numbers`, create a logical vector `divisible_by_5` to indicate numbers that are divisible by 5. Print the numbers that satisfy this condition.

Solution:

```
vect1<-c(15,20,24,123)
divisible_by_5<-vect1%%5==0
print(vect1[divisible_by_5])
```

```
## [1] 15 20
```

Question 27: Bigger or Smaller? You have two numeric vectors `vector1` and `vector2`. Create a logical vector comparison to indicate whether each element in `vector1` is greater than the corresponding element in `vector2`. Print the comparison results.

Solution:

```
vect1<-c(10,2,4,32)
vect2<-c(12,43,12,2)
vect1>vect2
```

```
## [1] FALSE FALSE FALSE  TRUE
```