

R Programming Quizzes

Martín Macías

Diciembre de 2015

1 Quiz 1

1. The R language is a dialect of which of the following programming languages?

R is a dialect of the S language which was developed at Bell Labs.

2. The definition of free software consists of four freedoms (freedoms 0 through 3). Which of the following is NOT one of the freedoms that are part of the definition?

The freedom to sell the software for any price. This is not part of the free software definition. The free software definition does not mention anything about selling software (although it does not disallow it).

3. In R the following are all atomic data types EXCEPT

data frame is not an atomic data type in R.

4. If I execute the expression `x <- 4L` in R, what is the class of the object `x` as determined by the `class()` function?

The L suffix creates an integer vector as opposed to a numeric vector.

```
x <- 4L
class(x)

[1] "integer"
```

5. What is the class of the object defined by `x <- c(4, TRUE)`? The numeric class is the "lowest common denominator" here and so all elements will be coerced into that class.

```
x <- c(4, TRUE)
class(x)

[1] "numeric"
```

Question Explanation: R does automatic coercion of vectors so that all elements of the vector are the same data class.

6. If I have two vectors `x <- c(1,3, 5)` and `y <- c(3, 2, 10)`, what is produced by the expression `cbind(x, y)`?

The `cbind` function treats vectors as if they were columns of a matrix. It then takes those vectors and binds them together column-wise to create a matrix.

```
x <- c(1,3, 5); y <- c(3, 2, 10)
cbind(x, y)

      x  y
[1,] 1  3
[2,] 3  2
[3,] 5 10
```

7. A key property of vectors in R is that *Elements of a vector all must be of the same class*
8. Suppose I have a list defined as `x <- list(2, "a", "b", TRUE)`. What does `x[[1]]` give me? *a numeric vector of length 1.*

```
x <- list(2, "a", "b", TRUE)
x[[1]]

[1] 2
```

9. Suppose I have a vector `x <- 1:4` and `y <- 2:3`. What is produced by the expression `x + y`? *an integer vector with the values 3, 5, 5, 7.*

```
x <- 1:4; y <- 2:3
x + y

[1] 3 5 5 7
```

10. Suppose I have a vector `x <- c(3, 5, 1, 10, 12, 6)` and I want to set all elements of this vector that are less than 6 to be equal to zero. What R code achieves this? *You can create a logical vector with the expression `x < 6` and then use the `[]` operator to subset the original vector `x`.*

```
x <- c(3, 5, 1, 10, 12, 6)
x[x < 6] <- 0
```

11. In the dataset provided for this Quiz, what are the column names of the dataset? *You can get the column names of a data frame with the `names()` function.*

```
setwd("/Users/Martin/Desktop/GitHub/datasciencecoursera")
getwd()

[1] "/Users/Martin/Desktop/GitHub/datasciencecoursera"

data <- read.csv("hw1_data.csv")
names(data)

[1] "Ozone"    "Solar.R" "Wind"     "Temp"     "Month"    "Day"
```

12. Extract the first 2 rows of the data frame and print them to the console. What does the output look like? *You can extract the first two rows using the `[` operator and an integer sequence to index the rows.*

```
data[1:2, ]

  Ozone Solar.R Wind Temp Month Day
1   41     190  7.4   67     5    1
2   36     118  8.0   72     5    2
```

13. How many observations (i.e. rows) are in this data frame? *You can use the `nrows()` function to compute the number of rows in a data frame.*

```
dim(data)

[1] 153  6
```

14. Extract the last 2 rows of the data frame and print them to the console. What does the output look like? *The `tail()` function is an easy way to extract the last few elements of an R object.*

```
data[152:153,]

  Ozone Solar.R Wind Temp Month Day
152   18     131  8.0   76     9   29
153   20     223 11.5   68     9   30
```

15. What is the value of `Ozone` in the 47th row? *The single bracket `[` operator can be used to extract individual rows of a data frame.*

```
data[47, 1]

[1] 21
```

16. How many missing values are in the `Ozone` column of this data frame? The `is.na` function can be used to test for missing values.

```
na <- data[, 1]
sum(is.na(na))

[1] 37
```

17. What is the mean of the `Ozone` column in this dataset? Exclude missing values (coded as `NA`) from this calculation. *The `mean` function can be used to calculate the mean.*

```
bad <- is.na(na)
mean(na[!bad])

[1] 42.12931
```

18. Extract the subset of rows of the data frame where `Ozone` values are above 31 and `Temp` values are above 90. What is the mean of `Solar.R` in this subset? *You need to construct a logical vector in R to match the question's requirements. Then use that logical vector to subset the data frame.*

```
Solar <- subset(data, Ozone > 31 & Temp > 90, select = c(Solar.R))
mean(Solar[, 1])

[1] 212.8
```

19. What is the mean of `Temp` when `Month` is equal to 6?

```
Temp <- subset(data, Month == 6, select = Temp)
mean(Temp[, 1])

[1] 79.1
```

20. What was the maximum ozone value in the month of May (i.e. `Month = 5`)?

```
ozono <- subset(data, Month == 5, select = Ozone)
max(ozono, na.rm = TRUE)

[1] 115
```