Vectors in R

Minjeong Kim

10/20/2017

```
library(ggplot2)
```

Vectors

Introduction

Welcome to the world of vectors! We talked about vector at the beginning of the semester. It is the basic concept we must master for further learnings in R. But, I thought that it is important to go over the basic steps to learn about more difficult concepts later on. You will explore more about vectors in depth: definition, automicity, coercion, higher-order functions, recycling. Sounds exciting? YEAH!!! There is a quiz created at the end, so read my post carefully!

What is vector?

It is the basic structure in R.

It has three properties: type, length, attributes

A function typeof() is used for the type of vector. It is classified as integer, double, complex, logical, or character.

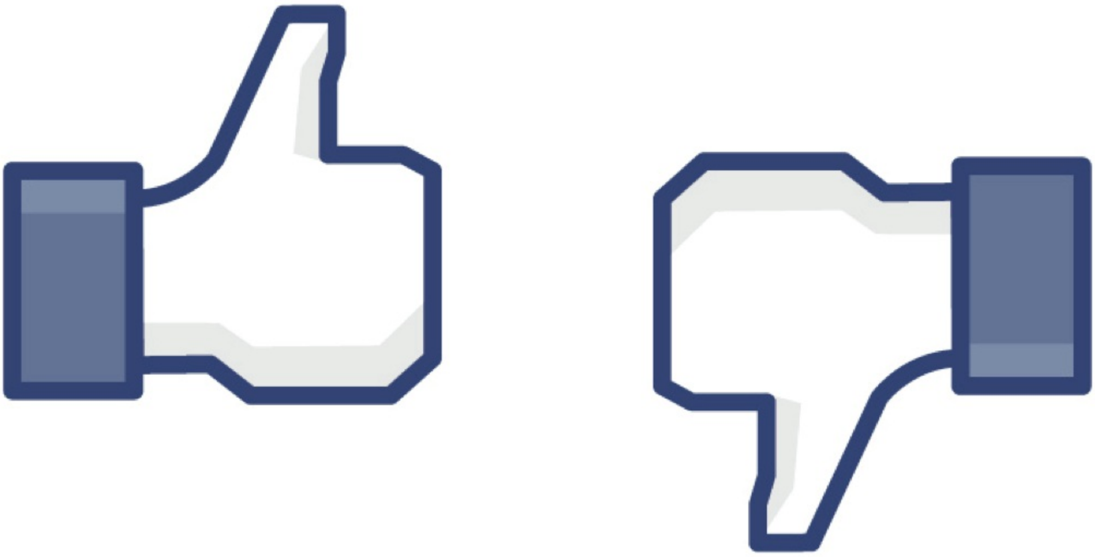
Let’s take a look at how typeof() function works.

```
hair_color <- c("black", "brown", "red", "blonde")
```

```
hair_price <- c(20, 35, 50, 40)
```

```
lhair_price <- c(20L, 35L, 50L, 40L)
```

```
hair_preference <- c(TRUE, TRUE, FALSE, TRUE)
```



```
typeof(hair_color)
```

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

```
typeof(hair_price)
```

```
## [1] "double"
```

```
typeof(lhair_price)
```

```
## [1] "integer"
```

A function length() is used for the length of vector. It indicates that how many elements vector has.

```
color <- c("red", "orange", "yellow", "green", "blue")
price <- c(0, 50, 200)
```

```
length(color)
```

```
## [1] 5
```

```
length(price)
```

```
## [1] 3
```

A function attributes() is used for additional arbitrary matadata. It accesses an object's attributes. The mostattribute assignment takes special for the dim, names, and dimnames attributes. Null objects do not have attributes.

It consists of atomic vectors and lists. I guess you wonder what "atomic" means.

Atomicity

The values in a vector are all of the same type: all integers, reals, complex, characters, or logical. There are two rare types: complex and raw. Atomic vectors are usually created with c() as above, which stands for combine. If you want to determine types of vectors, you can use either typeof() function or "is" function. Is functions, such as is.character(), is.double(), is.integer(), is.logical(), is.atomic(), are used to check specific types of functions.

```
is.character(hair_color)
```

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

```
is.double(hair_price)
```

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

```
is.integer(lhair_price)
```

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

```
is.logical(hair_preference)
```

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

```
is.atomic(c(hair_color, hair_price, lhair_price, hair_preference))
```

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

Note that is.numeric() returns TRUE for both doubles and integers. You see that is.numeric is more general than the others. Huh?

Coercion

All elements of vectors are all of the same type. Combination of a character and the other types of vectors yields a character. Combination of a logical and numbers yields a logical. Combination of a complex and the other numbers yields a complex.

```
c("one", "two", 3, TRUE, FALSE)
```

```
## [1] "one"   "two"   "3"     "TRUE"  "FALSE"
```

```
c(FALSE, 5, 8, TRUE, TRUE)
```

```
## [1] 0 5 8 1 1
```

```
c(3 + 4i, 5L, 6, 7)
```

```
## [1] 3+4i 5+0i 6+0i 7+0i
```

Vectorization

A function takes vectors [x1, x2, x3, …] as inputs and returns vectors [f(x1), f(x2), f(x3), …] as outputs. Unlike looping, vectorization allows you to operate on vectors quickly. It is important that a function does not operate on a scalar vaule but on a vector. Vectorization is applied to not only arithmetic operators but also boolean operators. For more information, you can look at http://alyssafrazee.com/2014/01/29/vectorization.html.

Higher-Order Functions

There is another type of vectorization called fold, also known as reduce. In this situation, a vector is aggregated into a single value. For example, the summation operator sum(x) is commonly used.

Recycling

When you deal with two or more vectors with different lengths, R recycles or repeats the shorter ones until it is long enough to match the longer one.

```
c(3, 6, 3, 8) + 5
```

```
## [1]  8 11  8 13
```

```
c(2, 6, 4, 3, 8) + c(3, 8, 8, 2)
```

```
## Warning in c(2, 6, 4, 3, 8) + c(3, 8, 8, 2): longer object length is not a
## multiple of shorter object length
```

```
## [1]  5 14 12  5 11
```

As you can see the second example above, c(2+3, 6+8, 4+8, 3+2, 8+3) applies.

Subscripting

You use bracket notation for vectors object[index] to extract values from objects. You can use either integer vector, logical vector, or character vector as index.

Integer vector as index

In R, vector index starts from 1, not from 0. You use a vector of integers to specify elements.

```
hair_color[3]
```

```
## [1] "red"
```

```
hair_price[2]
```

```
## [1] 35
```

You use negative integers to specify elements except the specified ones.

```
hair_color[-2]
```

```
## [1] "black"  "red"    "blonde"
```

```
hair_price[-1]
```

```
## [1] 35 50 40
```

```
lhair_price[-1]
```

```
## [1] 35 50 40
```

```
hair_preference[-2]
```

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

Interestingly, real numbers are truncated to integers.

```
hair_color[c(1.2, 4.3)]
```

```
## [1] "black"  "blonde"
```

```
hair_price[c(0.9, 3.3)]
```

```
## [1] 50
```

```
lhair_price[c(1.2, 5.3)]
```

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

```
hair_preference[c(1.7, 2.5)]
```

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

Want to mix positive and negative integers? Sorry, you can't do so. But, 0's can mix with negative integers.

Logical vector as index

If you use logical vector as index, you extract vectors where the logical vector returns TRUE.

```
hair_color[c(TRUE)]
```

```
## [1] "black"  "brown"  "red"    "blonde"
```

```
hair_price[c(FALSE, TRUE, TRUE, FALSE)]
```

```
## [1] 35 50
```

```
lhair_price[c(rep(TRUE,2), rep(FALSE,2))]
```

```
## [1] 20 35
```

```
hair_preference[c(FALSE)]
```

```
## logical(0)
```

Character vector as index

Character vector is used as index when vectors are named.

```
names(hair_price) <- c("a", "b", "c", "d")
hair_price["b"]
```

```
##  b
## 35
```

```
hair_price[c("d", "a")]
```

```
##  d  a
## 40 20
```

Quiz

Since you've mastered vectors, you must deal with the quiz I made easily.
1) What are three properties of vectors?
2) What are two rare types of atomic vectors?
3) What is the length of color?
4) Only natural integers can be used as index. (T/F)
5) One type of vectorizations is fold, also known as reduce. (T/F)

Answers for the Quiz

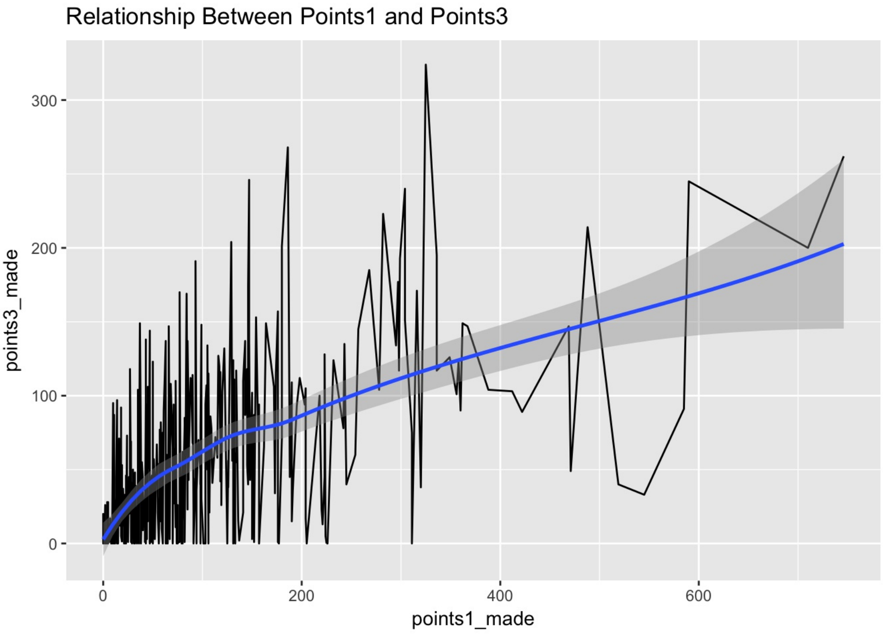
1. type, length, attributes
2. complex, raw
3. 5
4. FALSE
5. TRUE

Congratulations on completing the quiz! On my next post, you will explore more about ggplot. Take a look at the example below.

```
stats <- read.csv('/Users/minjeongkim/stat133/stat133-hws-fall17/hw03/data/nba2017-stats.csv')
```

```
ggplot(stats, aes(points1_made, points3_made)) + geom_line() + geom_smooth() + ggtitle("Relationship Between Point
s1 and Points3")
```

```
## `geom_smooth()` using method = 'loess'
```



```
cor(stats$points1_made, stats$points3_made)
```

```
## [1] 0.5803781
```

The graph shows the relationship between points1 made and points3 made. The smooth line shows that it has a positive, weak relationship between two variables. The correlation between points1 and points3 is around 0.58. Even though it looks positively related, it is fluctuating, so that the relationship is weak.

References

1. http://faculty.nps.edu/sebuttre/home/R/vectors.html
2. https://www.programiz.com/r-programming/vector
3. https://www.safaribooksonline.com/library/view/the-art-of/9781593273842/ch02s03.html
4. http://alyssafrazee.com/2014/01/29/vectorization.html
5. https://www.r-bloggers.com/how-to-use-vectorization-to-streamline-simulations/
6. http://www.dummies.com/programming/r/how-to-vectorize-your-functions-in-r/
7. http://adv-r.had.co.nz/Data-structures.html
8. http://assets.vixendaily.com/wp-content/uploads/2016/06/best-hair-color-for-your-skin-70.jpg
9. https://ei.marketwatch.com//Multimedia/2016/10/17/Photos/ZH/MW-EY056_dollar_20161017211422_ZH.jpg?uuid=33afd4a0-94d0-11e6-8fdf-00137241c023
10. https://gracesalt.files.wordpress.com/2014/04/good-vs-bad-reputation.png

Conclusion

In conclusion, the purpose of this post was for you to master everything about vectors. Once you read the post and completed the quiz, I wish you clearly know all the information about vectors. On the next post, you will explore about ggplot, as introduced above.