Lecture 1.8 – Lists and Data Frames

Specific Learning Objectives:

1.1.9 – Create vectors, arrays, matrices, lists, and data frames.

1.1.10 – Understand vectors and vectorized calculations.

1.1.11 – Understand the data classes of R.

1.1.12 – Learn how to index vectors, arrays, matrices, lists, and data frames.

Logical Record Subsets

- Information can also be pulled based on a logical test for most data structures (lists and data frames included!)

    - This is really useful if you don't care what the element position is, you care more about the value of the elements.

You run the following line to create a matrix:

```
my.matrix2 <- matrix(seq(1,21), nrow=7, byrow=TRUE)
```

Which line of code will subset only two-digit numbers (those greater than 9)?

- one option: **`my.matrix2[4:7, ]`**
- another option: **`my.matrix2[my.matrix>9]`**

```
> my.matrix2>9
      [,1]  [,2]  [,3]
[1,] FALSE FALSE FALSE
[2,] FALSE FALSE FALSE
[3,] FALSE FALSE FALSE
[4,]  TRUE  TRUE  TRUE
[5,]  TRUE  TRUE  TRUE
[6,]  TRUE  TRUE  TRUE
[7,]  TRUE  TRUE  TRUE
```

```
> my.matrix2[my.matrix2>9]
 [1] 10 13 16 19 11 14 17 20 12 15 18 21
```

Logical Record Subsets

- In data frames, be a little careful to specify whether you are looking for the record in a column or a row.

    - The easiest way to do this is to come up with the test first and then where it is searching second.

Example: In the **`ToothGrowth`** data set, find all the data associated with animals given orange juice as their supplement.

- We want to first create the test. How can we restrict data to only orange juice (OJ)?

```
ToothGrowth$supp == "OJ"
```

- Now use this logical vector to retrieve the records:

```
ToothGrowth[ToothGrowth$supp == "OJ", ]
```

Get all the rows where this statement is true

And get ALL columns for those rows!

Check Your Understanding

Which class of object would you use if you needed:

a) **Members of different sizes**

b) **Members of different classes**

c) **Both a and b**

	List	Data frame
a)	✓	✗
b)	✓	✓
c)	✓	✗

Check Your Understanding

Create a list in which each member contains one of each data types you've learned so far in the course!

Check Your Understanding

In the `ToothGrowth` data set, how can you print out all the measured tooth lengths from their study?

How can you find the mean and standard deviation of these lengths?

In-class Exercises

1. Catchup with assignments. Any questions on these?

2. Exercise 5.1 a (except ii)

3. Exercise 5.2 a and b

4. Assignments 1.10 and 1.11

Action Items

1. **Complete Assignments 1.10 and 1.11.**

2. **Read Davies Ch. 6 for next time.**