

Module 6: lab

BIOSTAT 213: Introduction to Computing in the R Software Environment

Wednesday, February 20, 2019

Save your code in a script file. I suggest adding labels and notes to this script file as comments. You may work in groups of 2–3.

1. Work on (5), from Module 5's lab.
2. Use *nhanes.csv*. Use `tapply()` to find the median systolic blood pressure (`bpxsy1`) in each `bpcat` group. Preferably, first re-order the levels of `bpcat` so they are in scientific order. Note that `bpcat` depends on `bpxsy1` (as well as `bpxdi1`) so you should notice a pattern.
3. Use *nhanes.csv*. Use `tapply()` and a custom function to find the percent of individuals in each `bpcat` group who smoke. Preferably, first re-order the levels of `bpcat` so they are in scientific order. The following is a possible custom function for finding the percent of individuals in a specific group:

```
> percent<-function(variable,group.name){  
+   prop.table(table(variable))[group.name]*100  
+ }
```

For example, to find the percent of individuals in the entire sample with a history of military service:

```
> percent(ii$military,group.name='Military')  
Military  
9.396671
```

4. Write a function that accepts two arguments, representing the two shortest sides of a right triangle. The function should return the length of the longest side of the triangle (the hypotenuse), using [the Pythagorean theorem](#): $c^2 = a^2 + b^2$.
5. Use *unps.csv*. Use `tapply()` to find the number of females (`male`) interviewed per wave (`wave`).
6. Use *kevin_love.txt*. Use `apply()` to find Kevin Love's career highs in field goal percent (`fgp`), 2-point percent (`fg2p`), 3-point percent (`fg3p`), and free-throw percent (`ftp`).