## Module 1: Exercises

### Question 1:

I am trying to use R as a calculator, and am receiving an error. Can you spot my error and update the code?

(42-5/(88+6)

### Answer 1:

My brackets are incorrectly specified as I open two sets of brackets but only close one. The correct solution is probably either (42-5)/(88+6) or 42-(5/(88+6)), depending on what I actually wanted to do The former is probably more likely!

(42-5)/(88+6)

## [1] 0.393617

### Question 2

This question will use the same earthquakes dataset from the quiz, showing the magnitude of earthquakes occuring in the ocean around Fiji since 1964, as well as the number of different stations reporting the earthquake. This has been loaded into the R sessions as a data frame called quakes

### Question 2a

Write a command to determine the largest magnitude (mag) earthquake recorded?

### Answer 2a

I would need to use the max command, and then specify that i want to use the mag column within the quakes dataset by using the data frame name quakes followed by a $ follwed by the column name mag

max(quakes$mag)

## [1] 6.4

### Question 2b

Write a command to determine the smallest depth (depth) below surface that an earthquake was recorded?

### Answer 2b

I would need to use the min command, and then specify that i want to use the depth column within the quakes dataset by using the data frame name quakes followed by a $ follwed by the column name depth

min(quakes$mag)

## [1] 4

### Question 2c

I would like to obtain the standard deviation of the earthquake magnitude column from the quakes dataset. See if you can find the function for standard deviation in R (we have not mentioned it in the course workbook so far) and then apply it to the relevant column

### Answer 2c

From a bit of searching, hopefully you found the sd function. This works in the same way as we have seen with max and min

sd(quakes$mag)

## [1] 0.402773

### Question 3

I am again using the airquality data and I am now interested in looking at the Solar.R variable. I know that there can sometimes be very extreme outlying values for this variable, so rather than looking at the minimum and maximum I would instead like to look at the 5th percentile and the 95th percentile. Find these two values using the quantile function.

### Answer 3

There are a few tricks to notice here - firstly when you look at the Solar.R variable, make sure you notice that there are missing values

airquality$Solar.R

## [1] 190 118 149 313 NA NA 299 99 19 194 NA 256 290 274 65 334 307 78  
## [19] 322 44 8 320 25 92 66 266 NA 13 252 223 279 286 287 242 186 220  
## [37] 264 127 273 291 323 259 250 148 332 322 191 284 37 120 137 150 59 91  
## [55] 250 135 127 47 98 31 138 269 248 236 101 175 314 276 267 272 175 139  
## [73] 264 175 291 48 260 274 285 187 220 7 258 295 294 223 81 82 213 275  
## [91] 253 254 83 24 77 NA NA NA 255 229 207 222 137 192 273 157 64 71  
## [109] 51 115 244 190 259 36 255 212 238 215 153 203 225 237 188 167 197 183  
## [127] 189 95 92 252 220 230 259 236 259 238 24 112 237 224 27 238 201 238  
## [145] 14 139 49 20 193 145 191 131 223

So when we use the quantile function we obtain an error

quantile(airquality$Solar.R)

## Error in quantile.default(airquality$Solar.R): missing values and NaN's not allowed if 'na.rm' is FALSE

So we need to use the option na.rm=TRUE

quantile(airquality$Solar.R,na.rm=T)

## 0% 25% 50% 75% 100%   
## 7.00 115.75 205.00 258.75 334.00

The default output is also not what we want - the question is asking for the 5th percentile and the 95th percentile. we need to use the probs option to set the quantile. we could write this once for each percentile like this. Note that we need to specifiy the percentage as a decimal - i.e. 5%=0.05:

quantile(airquality$Solar.R,na.rm=T,probs=0.05)

## 5%   
## 24.25

quantile(airquality$Solar.R,na.rm=T,probs=0.95)

## 95%   
## 311.5

But a better way would be to use the c() function to combine 0.05 and 0.95 and ask for the two percentiles within one line

quantile(airquality$Solar.R,na.rm=T,probs=c(0.05,0.95))

## 5% 95%   
## 24.25 311.50

### Question 4:

A task I am sure you remember from school is to solve a quadratic equation using the formula . Write some R code to find the two values of x when . As a reminder, in the quadratic equation formula from this particular example a would be 1; b would be -9 and c would be 19

### Answer 4

I am going to write this in general terms so that if i wanted to change my code for a different formula later, then i could. However, I didn’t explicitly ask you to do that, so if you directly plugged in 1, -9 and 19 into a formula this would still be fine.

Brackets are incredibly easy to get wrong on this one. Be extremely careful working out where to place them! My solution below has the minimum number of brackets necessary (due to BODMAS rules), but there is no harm in including extra brackets, just to be safe, if you want to ensure that the order of operations acts as you expect.

Make sure to write 4\*a\*c to multiply these together; R would not be able to interpret 4ac. R can interpet -b providing b is a number. But writing -1\*b here is maybe a little more explicit and clearer.

a<-1  
b<--9  
c<-19  
  
((-1\*b)+sqrt(b^2-4\*a\*c))/(2\*a)

## [1] 5.618034

((-1\*b)-sqrt(b^2-4\*a\*c))/(2\*a)

## [1] 3.381966

An extra trick would be to use the c() function to replace the operator as this is equivalent to saying “plus one times” or “minus one times”, so we can provide a vector of -1 and 1 to the code we write so that we only have to write one line for the main part of the solution.

a<-1  
b<--9  
c<-19  
  
x<-((-1\*b)+c(1,-1)\*sqrt(b^2-4\*a\*c))/(2\*a)  
  
x

## [1] 5.618034 3.381966

As i’m sure you learnt at school - it’s always good practice to plug these numbers back into the equation to see if it makes sense! Saving my object as x in the previous step makes this easy.

x^2 -(9\*x) +19

## [1] 3.552714e-15 0.000000e+00

You will see here that you don’t actually get a zero for the first number. Unfortunately R will sometimes come up with rounding errors. Remember in scientific notation that “3.552714e-15” is equal to “0.00000000000000355”.So i think i am happy to conclude that I got my formula correct!