

# Class06\_HW

Krysten Jones (A10553682)

First thing is to install the package we'll be using. In this case it's "bio3d" use the following command `install.packages("bio3d")`

Then we'll want to call it from our library

```
library(bio3d)
```

## Section 1: Improving analysis code by writing functions

**Section 1A. Improve this regular R code by abstracting the main activities in your own new function. The main steps should entail running through the code to see if it works, simplifying to a core working code snippet, reducing any calculation duplication, and finally transferring your new streamlined code into a more useful function for you.**

Original

```
# (A. Can you improve this analysis code?)
df <- data.frame(a=1:10, b=seq(200,400,length=10),c=11:20,d=NA)

df$a <- (df$a - min(df$a))/(max(df$a) - min(df$a))
df$b <- (df$b - min(df$a))/(max(df$b) - min(df$b))
df$c <- (df$c - min(df$c))/(max(df$c) - min(df$c))
df$d <- (df$d - min(df$d))/(max(df$a) - min(df$d))
df$a

[1] 0.0000000 0.1111111 0.2222222 0.3333333 0.4444444 0.5555556 0.6666667
[8] 0.7777778 0.8888889 1.0000000
```

Let's start by examining our dataframe

```
df <- data.frame(a=1:10, b=seq(200,400,length=10),c=11:20,d=NA)
df
```

```
      a      b  c  d
1  1 200.0000 11 NA
2  2 222.2222 12 NA
3  3 244.4444 13 NA
4  4 266.6667 14 NA
5  5 288.8889 15 NA
6  6 311.1111 16 NA
7  7 333.3333 17 NA
8  8 355.5556 18 NA
9  9 377.7778 19 NA
10 10 400.0000 20 NA
```

Strange, it just looks like random numbers. That's ok lets see if we can condense. First, we have already been given the body of our function, instead

```
A_ans <- function(x){x - min(x)/(max(x)-min(x))}
apply(df, 2, A_ans)
```

```
      a      b      c  d
[1,] 0.8888889 199.0000  9.777778 NA
[2,] 1.8888889 221.2222 10.777778 NA
[3,] 2.8888889 243.4444 11.777778 NA
[4,] 3.8888889 265.6667 12.777778 NA
[5,] 4.8888889 287.8889 13.777778 NA
[6,] 5.8888889 310.1111 14.777778 NA
[7,] 6.8888889 332.3333 15.777778 NA
[8,] 7.8888889 354.5556 16.777778 NA
[9,] 8.8888889 376.7778 17.777778 NA
[10,] 9.8888889 399.0000 18.777778 NA
```

## Section 1B:

Run through the code to see if it works, fix any copy/paste errors before simplifying to a core working code snippet, reducing any calculation duplication, and finally transferring it into a more useful function for you.

Original Code:

```
# Can you improve this analysis code?  
library(bio3d)  
s1 <- read.pdb("4AKE") # kinase with drug
```

Note: Accessing on-line PDB file

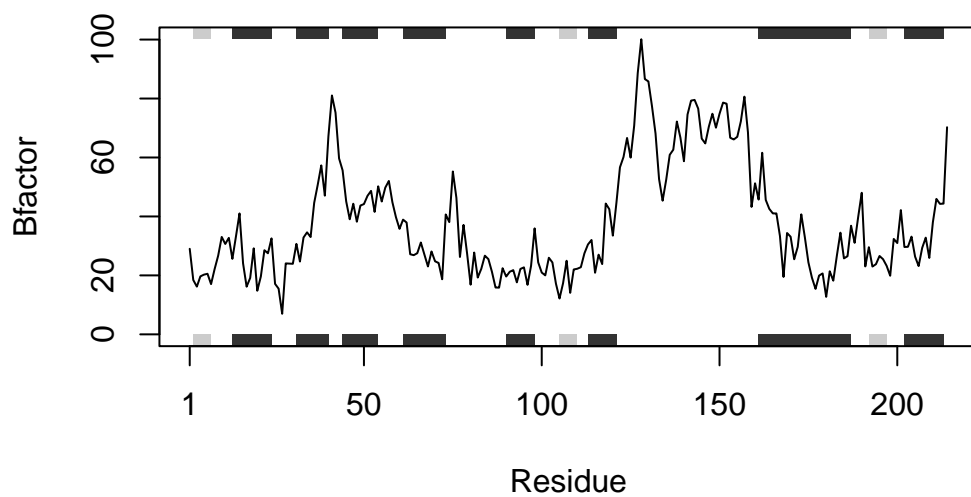
```
s2 <- read.pdb("1AKE") # kinase no drug
```

Note: Accessing on-line PDB file  
PDB has ALT records, taking A only, rm.alt=TRUE

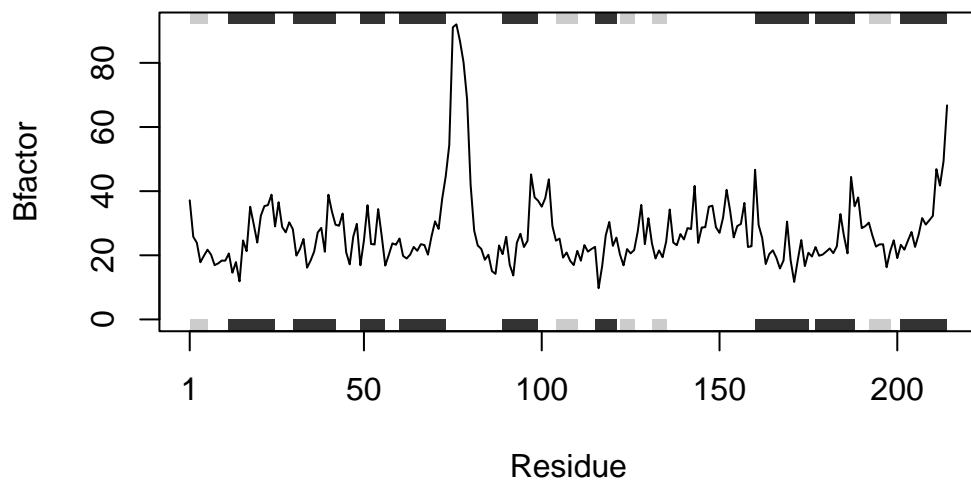
```
s3 <- read.pdb("1E4Y") # kinase with drug
```

Note: Accessing on-line PDB file

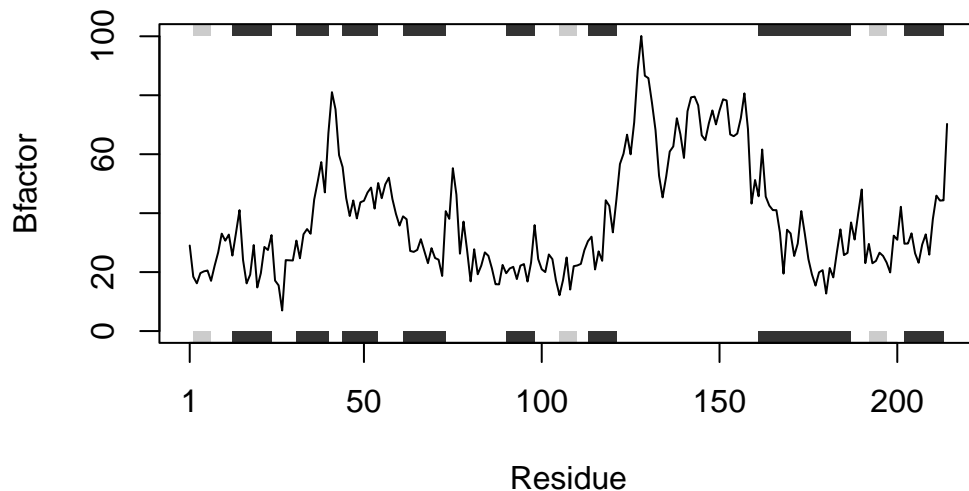
```
s1.chainA <- trim.pdb(s1, chain="A", elety="CA")  
s2.chainA <- trim.pdb(s2, chain="A", elety="CA")  
s3.chainA <- trim.pdb(s1, chain="A", elety="CA")  
  
s1.b <- s1.chainA$atom$b  
s2.b <- s2.chainA$atom$b  
s3.b <- s3.chainA$atom$b  
  
plotb3(s1.b, sse=s1.chainA, typ="l", ylab="Bfactor")
```



```
plotb3(s2.b, sse=s2.chainA, typ="l", ylab="Bfactor")
```



```
plotb3(s3.b, sse=s3.chainA, typ="l", ylab="Bfactor")
```



**Q6. How would you generalize the original code above to work with any set of input**

protein structures?

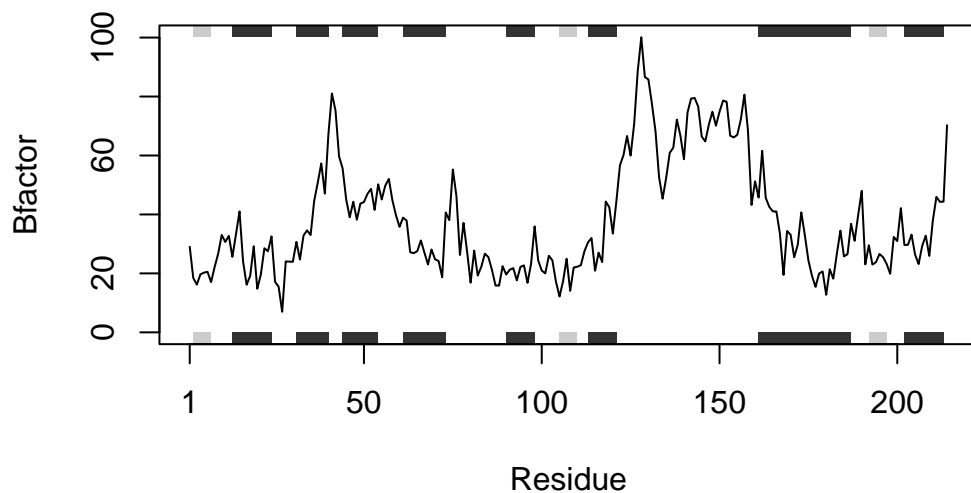
```
library(bio3d)

Bfunc <- function(x) {
  ChainA <- trim.pdb(read.pdb(x), chain="A", elety="CA")
  plotb3(ChainA$atom$b, sse=ChainA, typ="l", ylab="Bfactor")
}

Bfunc("4AKE")
```

Note: Accessing on-line PDB file

Warning in get.pdb(file, path = tempdir(), verbose = FALSE):  
C:\Users\kryst\AppData\Local\Temp\Rtmp0W0lXM\4AKE.pdb exists. Skipping download

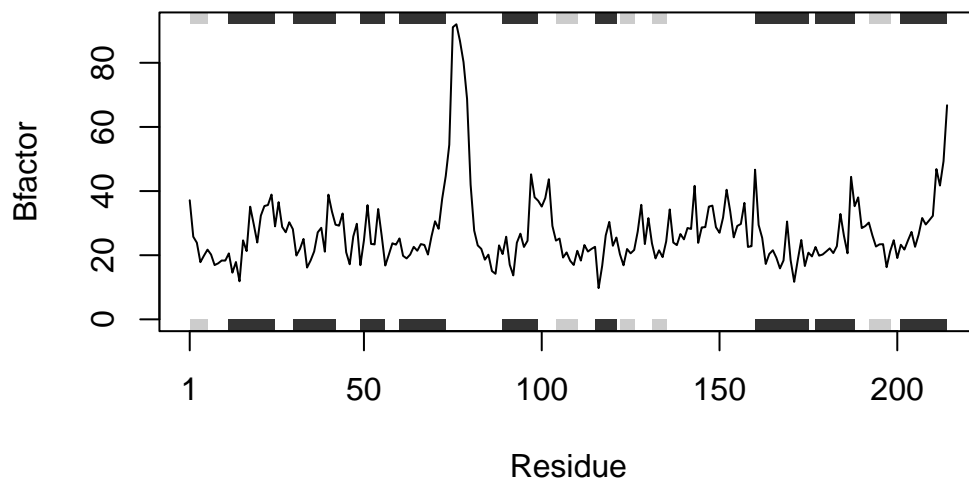


```
Bfunc("1AKE")
```

Note: Accessing on-line PDB file

Warning in get.pdb(file, path = tempdir(), verbose = FALSE):  
C:\Users\kryst\AppData\Local\Temp\Rtmp0W0lXM\1AKE.pdb exists. Skipping download

PDB has ALT records, taking A only, rm.alt=TRUE



```
Bfunc("1E4Y")
```

Note: Accessing on-line PDB file

Warning in get.pdb(file, path = tempdir(), verbose = FALSE):  
C:\Users\kryst\AppData\Local\Temp\Rtmp0W01XM\1E4Y.pdb exists. Skipping download

