Class 6: Lab Supplement

Given Code:

```
# Can you improve this analysis code?
library(bio3d)

s1 <- read.pdb("4AKE")
```

```
  Note: Accessing on-line PDB file
```

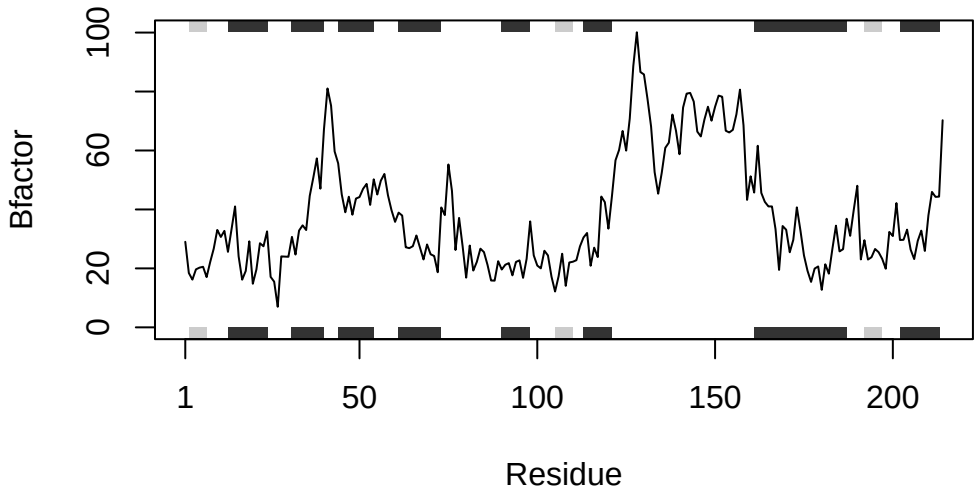
```
# kinase with drug
s2 <- read.pdb("1AKE")
```

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

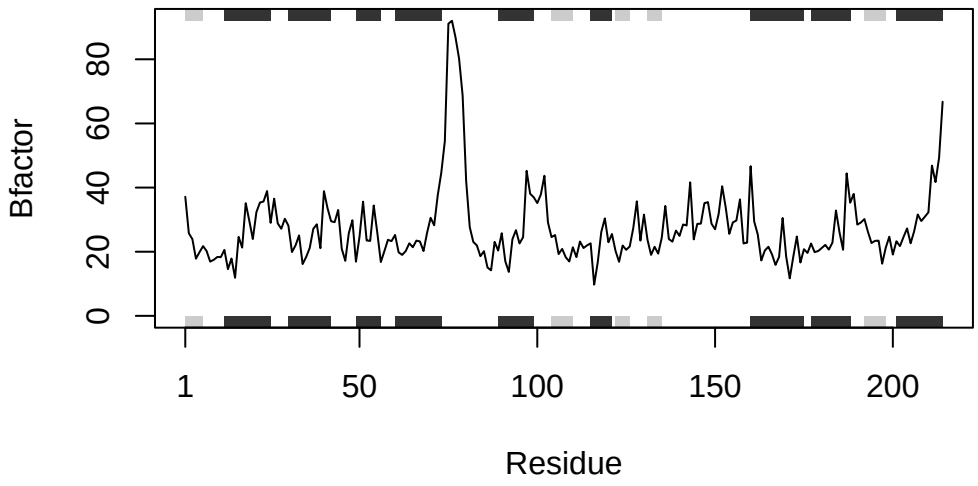
```
# kinase no drug
s3 <- read.pdb("1E4Y")
```

```
  Note: Accessing on-line PDB file
```

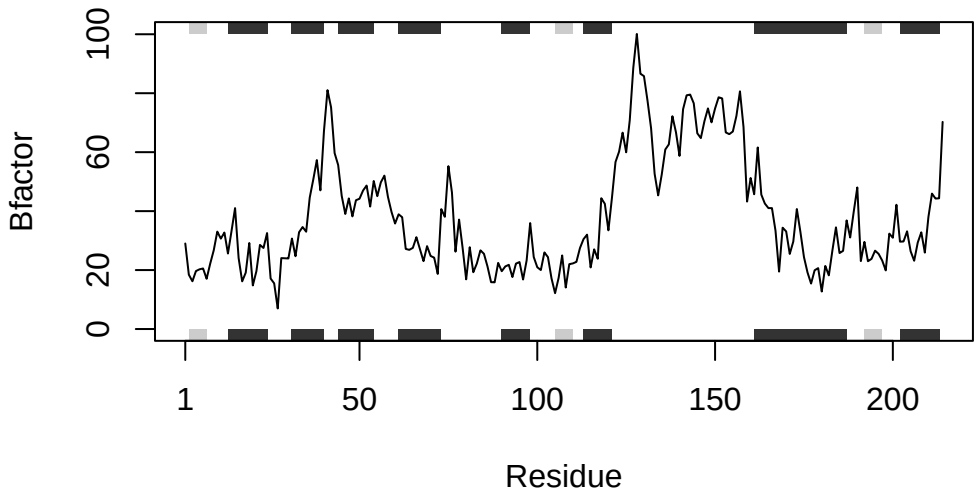
```
# kinase with drug
s1.chainA <- trim.pdb(s1, chain="A", elety="CA")
s2.chainA <- trim.pdb(s2, chain="A", elety="CA")
s3.chainA <- trim.pdb(s1, chain="A", elety="CA")
s1.b <- s1.chainA$atom$b
s2.b <- s2.chainA$atom$b
s3.b <- s3.chainA$atom$b
plotb3(s1.b, sse=s1.chainA, typ="l", ylab="Bfactor")
```




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



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



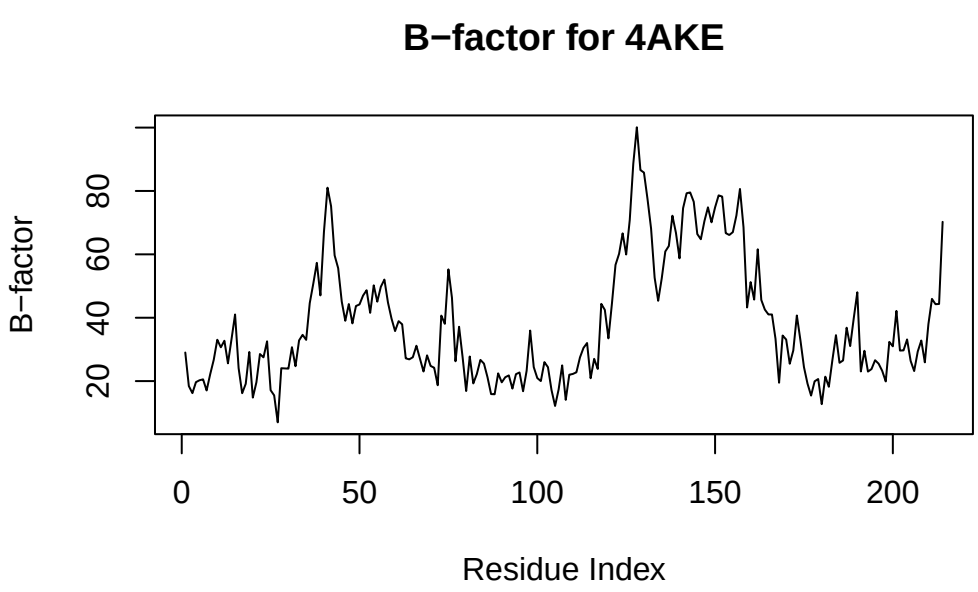
Q6. Improve on and simply the code:

```
# Load bio3d or library for data used
library(bio3d)

# Define a generalized function to analyze protein structures
analyze_proteins <- function(pdb_files, chain="A", elety="CA") {

   # the pdb_files is the vector with the files and their names used in the function
  # the chain in the function is meant to identify the specifc name in the files
  # the elety is the specific element like alpha carbons

  # The function is meant to plot and read for the B-factors of different protein structures

  # Go through each PDB file that is available
  for (pdb_file in pdb_files) {
  # Read the file
    pdb <- read.pdb(pdb_file)

    # In the function for pdb_chain is to trim and extract what is being specified
```

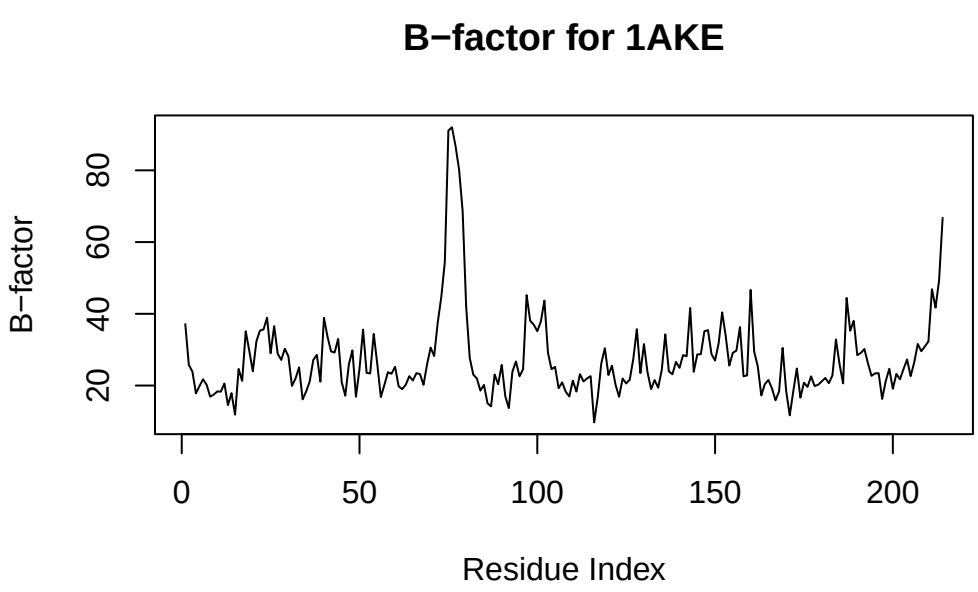
```
    pdb_chain <- trim.pdb(pdb, chain="A", elety="CA")

    # B_factor is used to obtain the atoms that are in the chain
    b_factor <- pdb_chain$atom$b

    # To generate the plot, use the bfactors, what the axis names are, and generate a title
    plot(b_factor, type="l", ylab="B-factor", xlab="Residue Index",
         main=paste("B-factor for", pdb_file))
  }
}

# Example of how the function would work with the pdb files given in the original code:

pdb_files <- c("4AKE", "1AKE", "1E4Y")

# Input general function to produce the graph:
analyze_proteins(pdb_files)
```

```
  Note: Accessing on-line PDB file
```

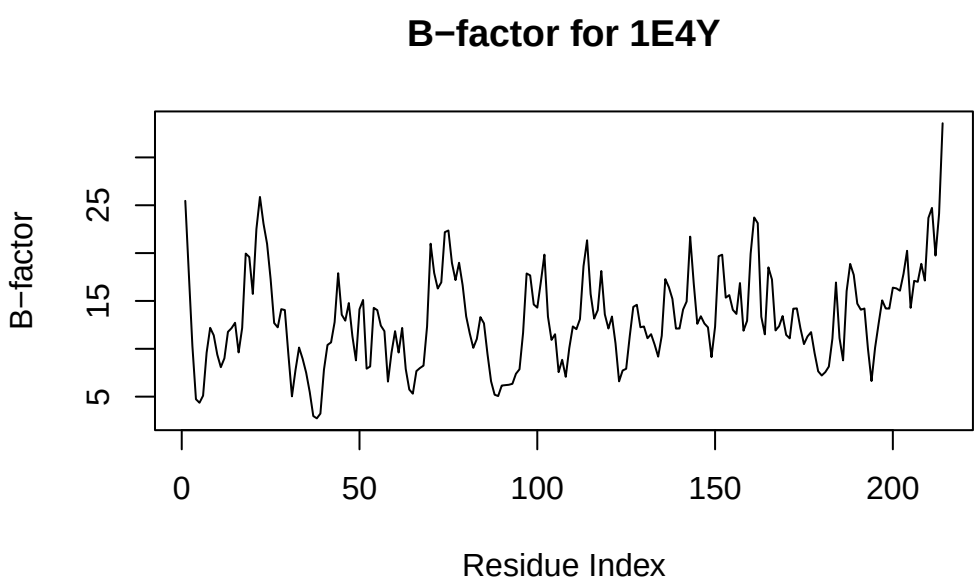
```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
C:\Users\NATALI~1\AppData\Local\Temp\Rtmp84q1dX/4AKE.pdb exists. Skipping
download
```




```
  Note: Accessing on-line PDB file

Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
C:\Users\NATALI~1\AppData\Local\Temp\Rtmp84q1dX/1AKE.pdb exists. Skipping
download

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




```
  Note: Accessing on-line PDB file

Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
C:\Users\NATALI~1\AppData\Local\Temp\Rtmp84q1dX/1E4Y.pdb exists. Skipping
download
```


B−factor for 1E4Y
B−factor
5
15
25
0
50
100
150
200
Residue Index