````
---
title: "Homework 6 Week 5"
author: "Lisa Chen A17082974"
date: "2024-02-03"
output: html_document
---

> Q6. How would you generalize the original code above to work with any set of input
protein structures?

```{r}
# installing packages
install.packages("bio3d")

```

```{r}
 # The function below takes the B data and plots their subsequent factors for each
individual protein and reads the trends from the previous questions
protein_analysis <- function(files) {
  # load library and reading data files of the protein
  library(bio3d)
  proteins_data <- lapply(files, read.pdb)
  # B factor information being extracted
  b_factors_list <- lapply(proteins_data, function(protein) trim.pdb(protein, chain="A",
elety="CA")$atom$b)
  # Plot B-factors by using a for loop to iterate the protein structure
  for (index in seq_along(b_factors_list)) {
    plotb3(b_factors_list[[index]], sse=proteins_data[[index]]$atom, typ="l",
ylab="Bfactor", main=paste("Protein ", index))
  }
  # the plot above is the output as it shows the B factors for protein similarity. THe for
loop allows the b factors and protein data to loop and make it more "readable"
  # dendogram
  hierarchical_clustering <- hclust(dist(do.call(rbind, b_factors_list)))
  plot(hierarchical_clustering, main="Cluster Dendrogram")
}# the plot above is the output as it shows the B factors for protein similarity.
protein_files <- c("4AKE", "1AKE", "1E4Y")
# protein files is the input
protein_analysis(protein_files)
```

}
````