# Stat Modeling and Computing Assignment 7

```
library(MASS)
library(faraway)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following object is masked from 'package:MASS':
##
##     select

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(pls)
```

```
##
## Attaching package: 'pls'

## The following object is masked from 'package:stats':
##
##     loadings
```

```
##HW07: Use PCAfatinmeat.R as a guide to perform Principal Component Analysis using all 215 observation
```

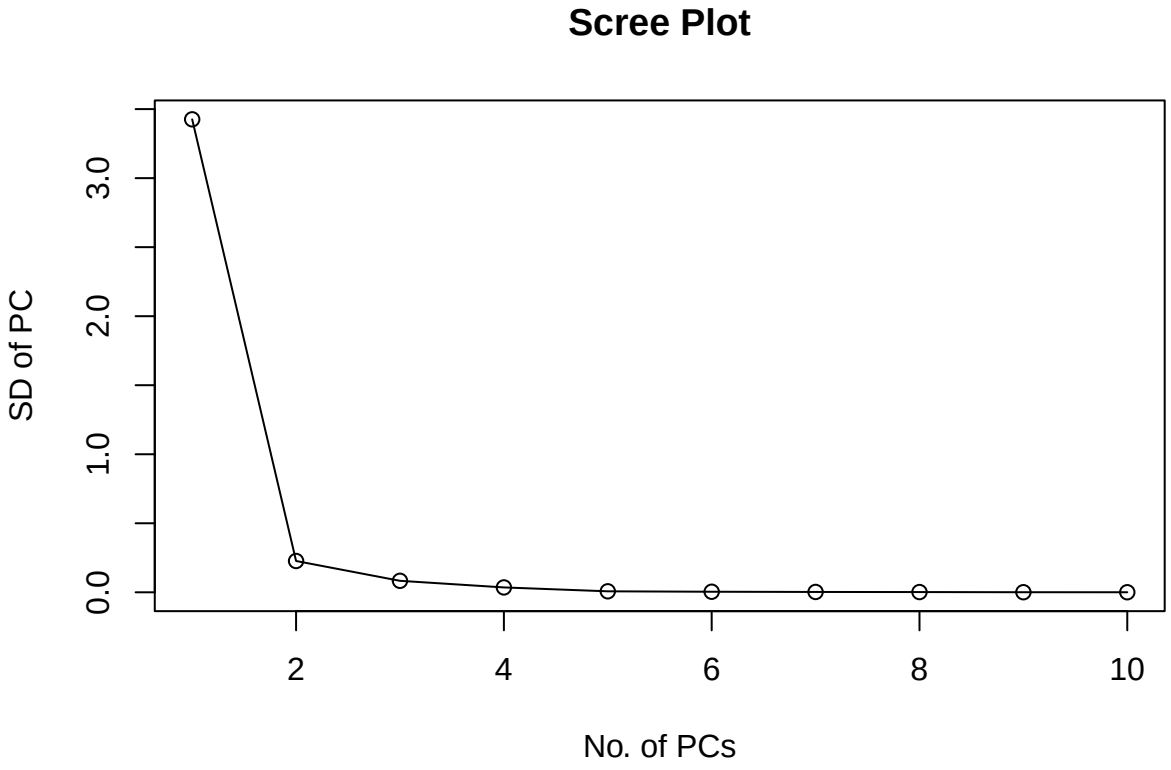
```
data(meatspec)
```

```
meat_data <- meatspec[, 1:50]
```

```
meatpca = prcomp(meat_data)
```

```
#a) Show the SCREE plot. How many components would you extract based on the SCREE plot.

plot(meatpca$sdev[1:10], type = "l", xlab="No. of PCs", ylab = "SD of PC", main = "Scree Plot")
points(1:10, meatpca$sdev[1:10])
```




```
# Based on the scree plot I would use 3 components
```

```
##b) Perform a multiple regression on all 50 predictors and report the rmse from the overall model.

meat_data$fat = meatspec$fat
```

```
lm_model <- lm(fat ~ ., data = meat_data)
```

```
rmse = function(x,y) sqrt(mean((x-y)^2))
rmse(fitted(lm_model), meat_data$fat)
```

```
## [1] 1.528491
```

```
#RMSE = 1.528491
```

```
#c) Perform a Principal Component Regression using all 50 components. How many components are selected
set.seed(10)


pcrmod_1 = pcr(fat ~ ., data=meat_data)
```

```
ypred = predict(pcrmod_1, meat_data)
rmse(ypred, meat_data$fat)
```

```
## [1] 2.709398
```

```r
pcrmod = pcr(fat ~ ., data=meat_data, validation="CV", ncomp=50)
```

```r
pcrCV = RMSEP(pcrmod, estimate="CV")
```

```r
pcrCV
```

```
## (Intercept)      1 comps      2 comps      3 comps      4 comps      5 comps
##      12.770       11.504        5.499        4.173        3.195        3.173
##     6 comps      7 comps      8 comps      9 comps     10 comps     11 comps
##       3.065        2.931        2.938        3.029        2.835        2.772
##    12 comps     13 comps     14 comps     15 comps     16 comps     17 comps
##       2.577        2.514        2.545        2.467        2.465        2.454
##    18 comps     19 comps     20 comps     21 comps     22 comps     23 comps
##       2.473        2.475        2.455        2.463        2.474        2.360
##    24 comps     25 comps     26 comps     27 comps     28 comps     29 comps
##       2.387        2.301        2.443        2.459        2.447        2.460
##    30 comps     31 comps     32 comps     33 comps     34 comps     35 comps
##       2.476        2.481        2.495        2.607        2.611        2.620
##    36 comps     37 comps     38 comps     39 comps     40 comps     41 comps
##       2.640        2.704        2.751        2.809        2.831        2.900
##    42 comps     43 comps     44 comps     45 comps     46 comps     47 comps
##       3.017        3.021        3.049        2.787        2.749        2.743
##    48 comps     49 comps     50 comps
##       2.710        2.630        2.628
```

```r
which.min(pcrCV$val[-1])
```

```
## [1] 25
```

```r
#25 components are selected
```

```r
ypred = predict(pcrmod, meat_data, ncomp=25)
```

```r
rmse(ypred, meat_data$fat)
```

```
## [1] 1.799395
```

```r
#RMSE = 1.799395
```