Statistical models : Homework 3

Priyanshi Shah and Sourabh Prakash

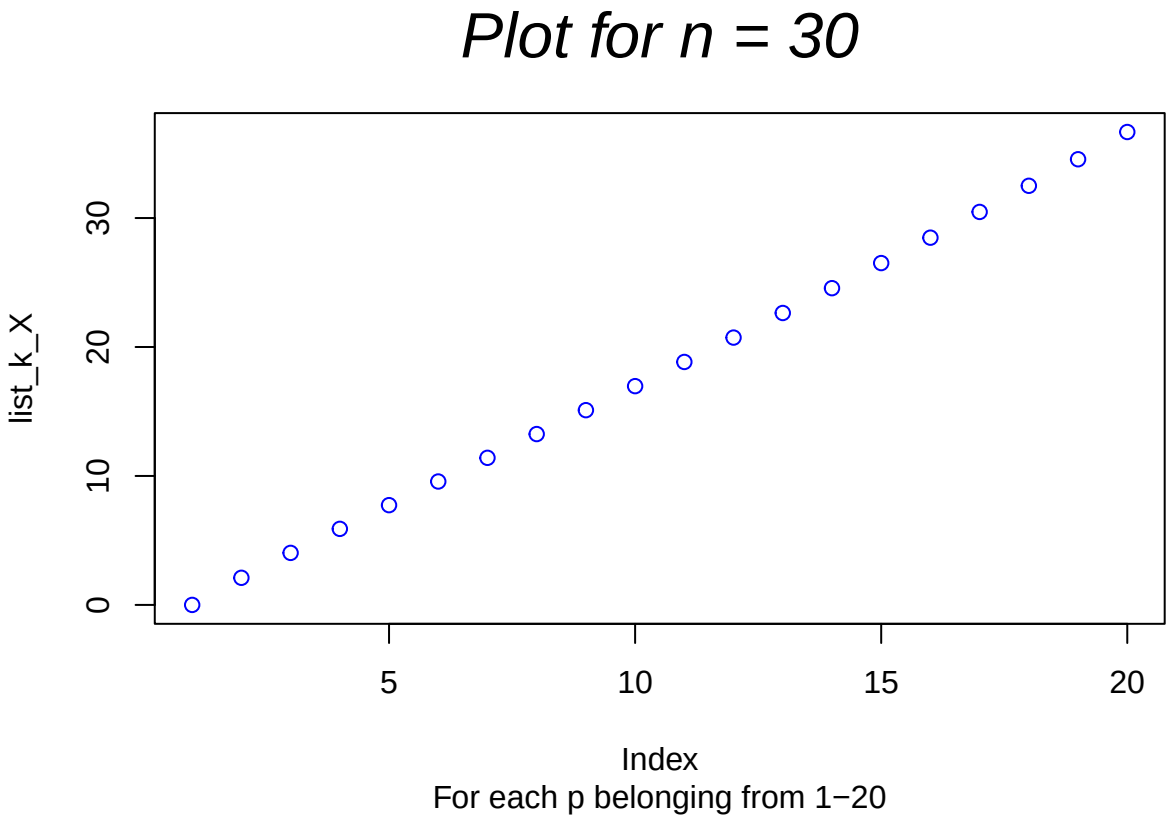
2023-02-03

Question 1

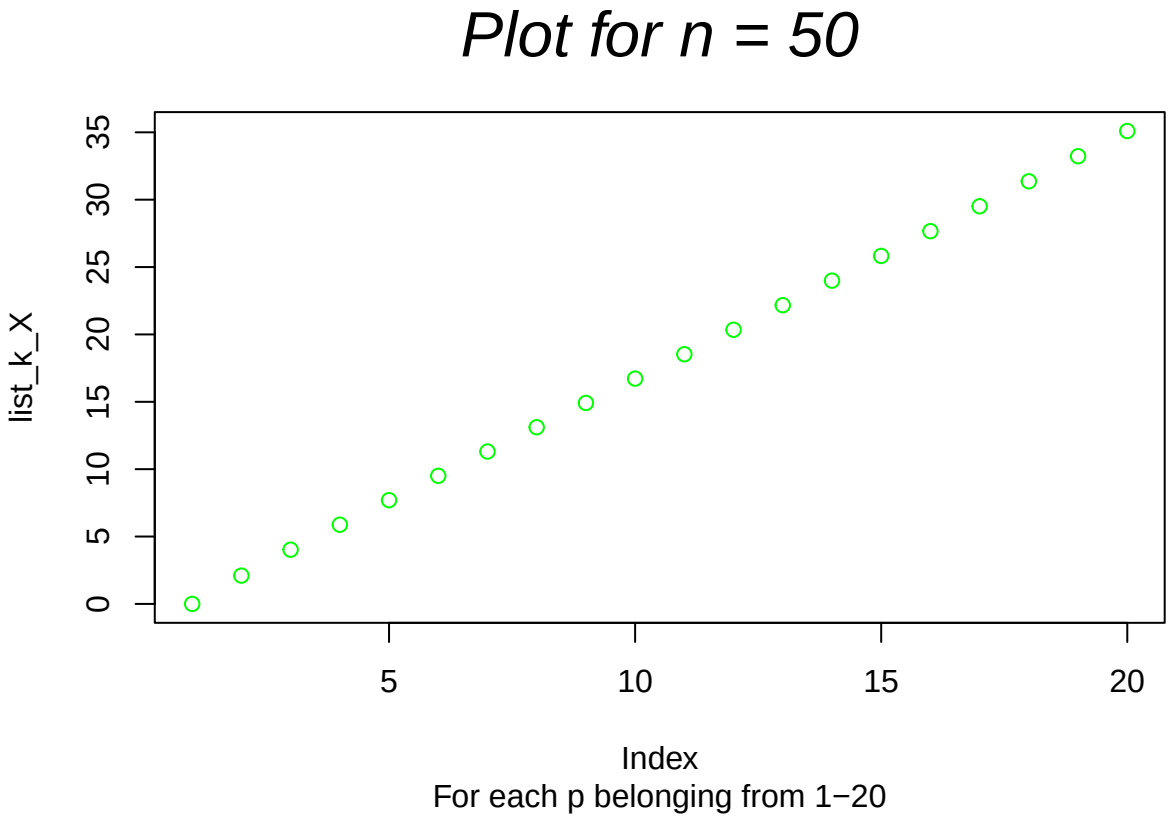
```
fun <- function(row,col) row^(col-1)
final_list = c()
# Creating a function for condition number
condition_number = function(n, color="blue"){
  list_k_X = c()
  i = n
  #looping from 1 to 20
  for (j in 1:20){
    x = seq(from = 0, to = 1, by = (1/(i+1)))[2:i+1]
    rows = x
    cols = 1:j+1
    X = outer(rows,cols,FUN=fun)
    svd_X = svd(X)
    k_X = max(svd_X$d)/min(svd_X$d)
    list_k_X = append(list_k_X, log(k_X))
    #plot(list_k_X, col=color)
  }
  #Visualizing the plot
  plot(list_k_X, col=color)
  title(main = paste('Plot for n =', n), sub = "For each p belonging from 1-20",
      cex.main = 2,   font.main= 3, col.main= "black")
  final_list = append(final_list,list_k_X)
  }

#plot.new()
```

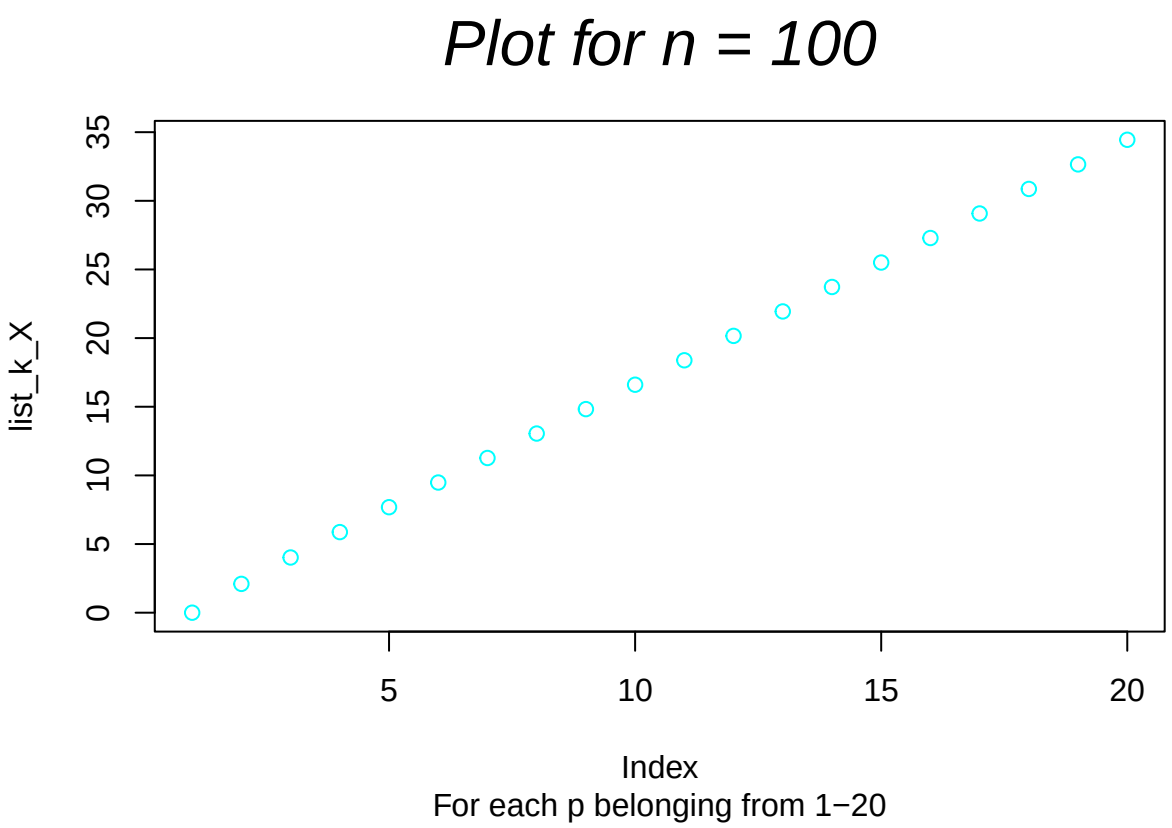
```
condition_number(30, "blue")
```



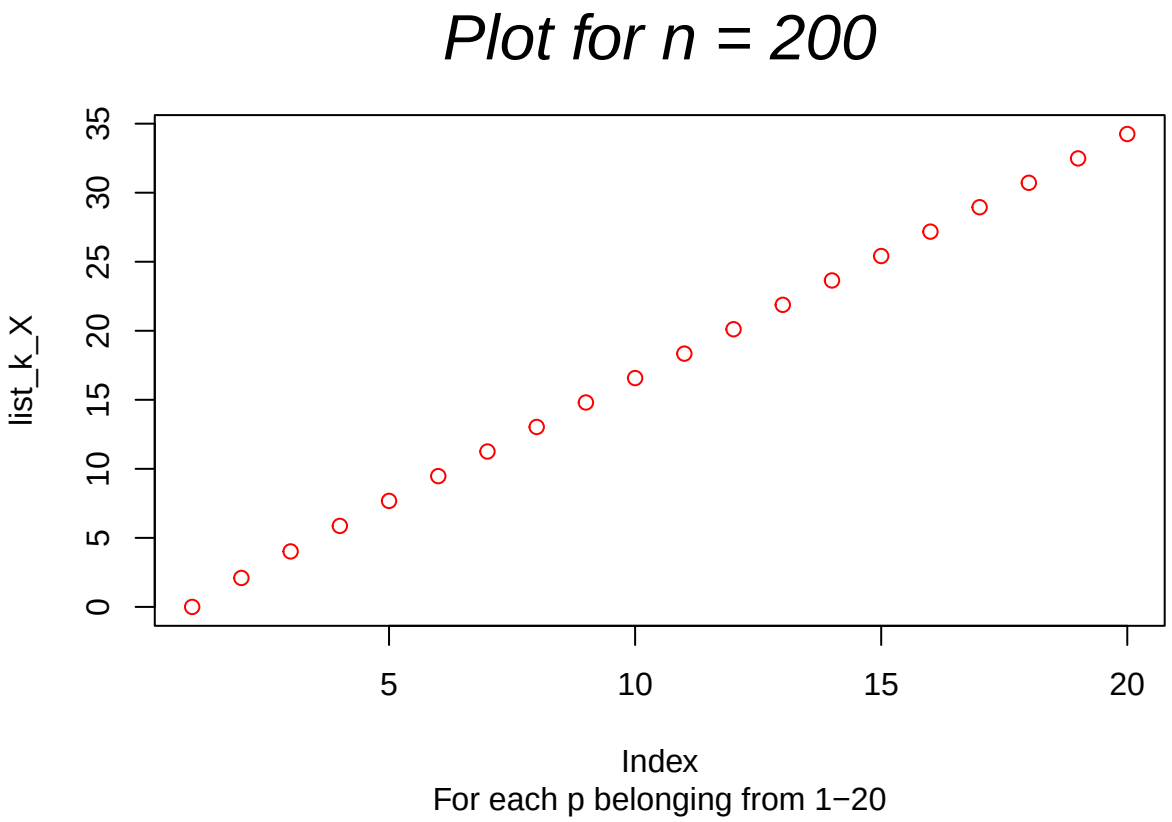

```
condition_number(50, "green")
```



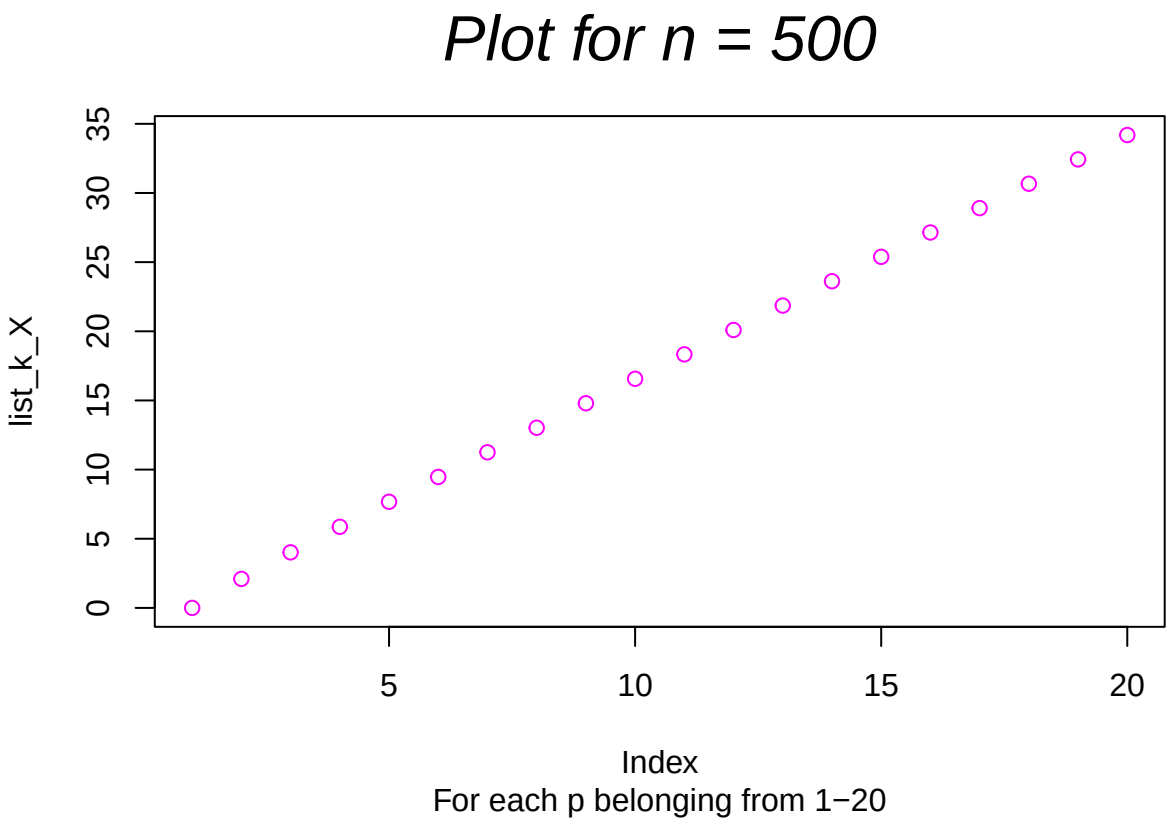
```
condition_number(100, "cyan")
```



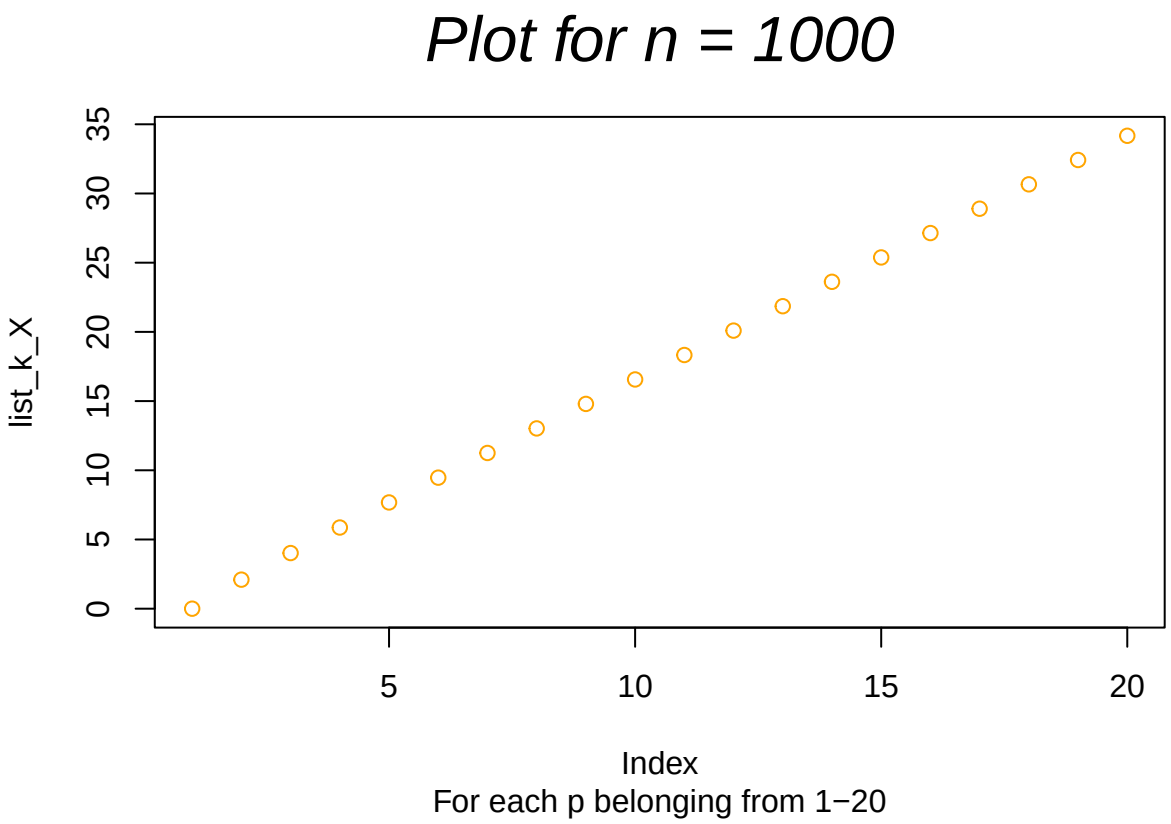

```
condition_number(200, "red")
```




```
condition_number(500, "magenta")
```



```
condition_number(1000, "orange")
```



Inference of question 1

As we can see from the graphs plotted above, for every p belong from 1 to 20 and for all xi = i/(n + 1) for i = 1,. . . ,n where n is ranging {30,50,100,200,500,1000}, we can see that the pattern is the same for all. This means that no matter what the value of n is, the values will be conditioned.

Question 2

```
#Creating Piecewise constant funcyion
piecewiseConstant = function(x,y, L, plot=TRUE){
  n = 2^L
  #K = quantile(x, seq(0, 1, len = n+1), type=1)              # for quantile
  quotient = (range(x)[2]-range(x)[1])/n
  K = seq(from = range(x)[1], to = range(x)[2], by =quotient ) # for equal split
  pts = rep(0,2*n)
  val = rep(0,2*n)
  # looping from 1 to n
  for (j in 1:n){
    I = (K[j] < x)&(x <= K[j+1])
    if (length(I[I==TRUE]) !=0){
      fit = lm(y[I] ~ 1)
      pts[2*j-1] = K[j]
```

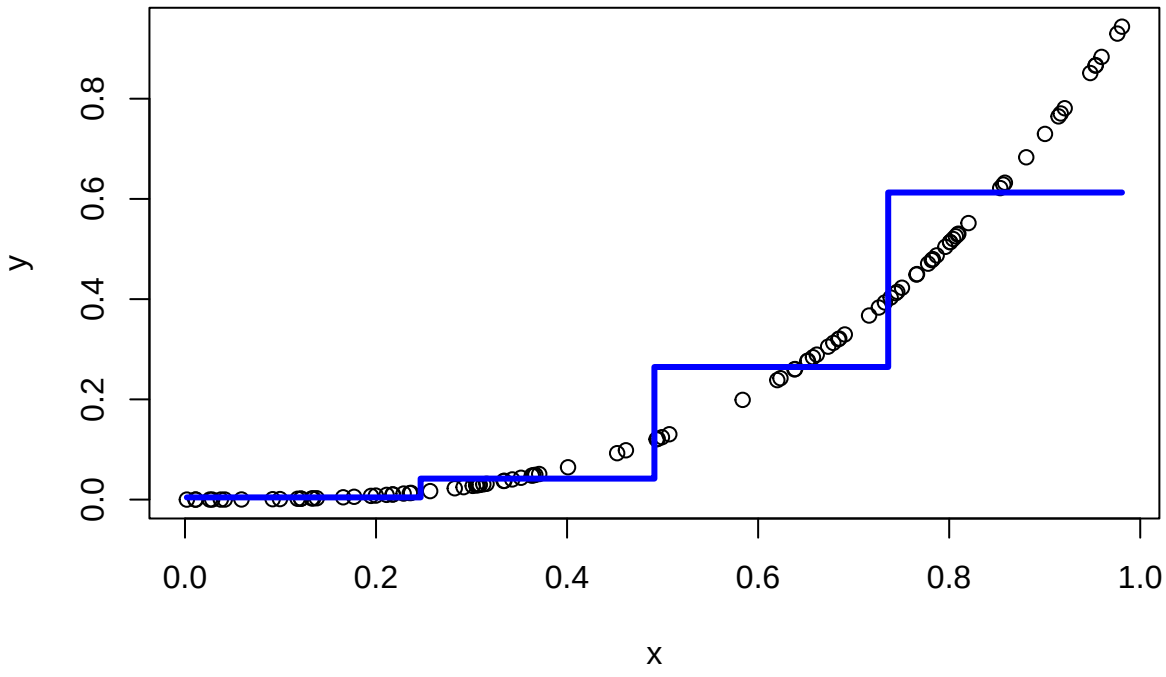
```
      pts[2*j] = K[j+1]
      val[2*j-1] = coef(fit)
      val[2*j] = coef(fit)
    }
    else{
      pts[2*j-1] = K[j]
      pts[2*j] = K[j+1]
      val[2*j-1] = val[2*j-3]
      val[2*j] = val[2*j-2]
    }

  }
  if (plot){
    if (L==2){
      lines(pts, val, col="blue", lwd = 3)
    }
    else if (L==3){
      lines(pts, val, col="green", lwd = 3)
    }
    if (L==4){
      lines(pts, val, col="red", lwd = 3)
    }
  }
}
```

a

Result for 2a:

```
# For part 2a
x = runif(100, min=0, max=1)
y = x^3
plot(y~x)
piecewiseConstant(x,y,2, TRUE)
```

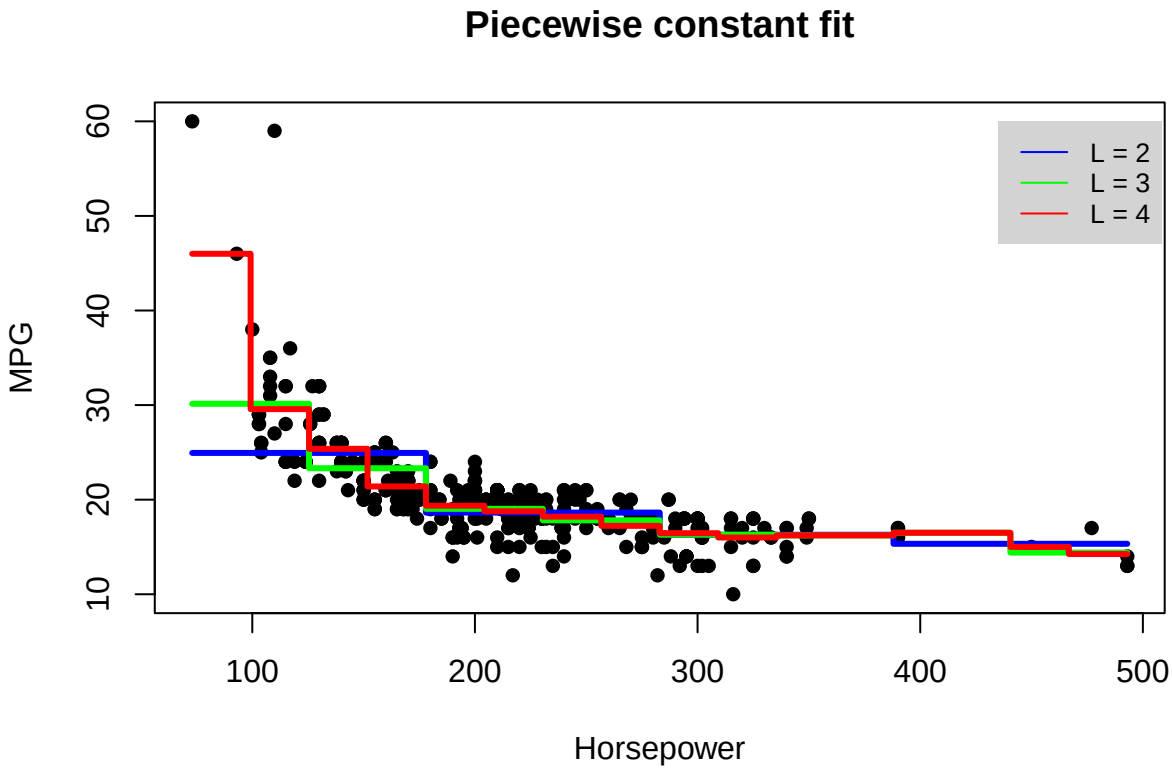


Inference

b

```
#Loading the cars dataset
#setwd("D:/projects/Quarter-2/Stats_model")#
load("../cars/04cars.rda")

data_ = dat
dat = data_[complete.cases(data_),]
#plotting for different values of L
plot(dat$Horsepower,dat$City_MPG, pch = 16, main="Piecewise constant fit", cex = 1, xlab="Horsepower", y
piecewiseConstant(dat$Horsepower,dat$City_MPG,2, TRUE)
piecewiseConstant(dat$Horsepower,dat$City_MPG,3, TRUE)
piecewiseConstant(dat$Horsepower,dat$City_MPG,4, TRUE)
legend(435, 60, legend=c("L = 2", "L = 3", "L = 4"),col=c("blue", "green", "red"), lty=1, cex=0.8, box.l
```



Inference of Question 2

Team Contributions

Both the team members Sourabh Prakash and Priyanshi Shah have contributed equally to the homework by discussing the key points and logic together and doing pair programming. For the implementation part question 1 was contributed by Priyanshi Shah and question 2 by Sourabh Prakash. The infereneces were drawn together by both the team members.