> # Exercise Download "Auto.csv" from http://www-bcf.usc.edu/~gareth/ISL/data.html

> # and estimate the regression of mpg on weight using the KNN method.

> # Draw the regression line on the scatterplot and compare it with linear regression line.

> rm(list=ls())

>

> setwd("D:/Documents (Louis Booth)/R/Big Data")

> Auto = read.table("Auto.data", header=TRUE, na.strings="?")

>

> y <- Auto$mpg

> x <- Auto$weight

> pdf("Homework 3 Plot.pdf")

> plot(x,y)

> reg <- lm(y~x)

> abline(reg, col='red')

>

> knn = function(x0, X, Y, k) {

+ x0 <- matrix(rep(x0, length(y)), byrow=TRUE)

+ X <- matrix(x)

+ distance <- rowSums((x0-X)^2)

+ rank <- order(distance)

+ y\_k <- Y[rank][1:k]

+ mean(y\_k)

+ }

>

>

> X <- sort(Auto$weight)

> fhat <- matrix(rep(NA, 1588), 397, 4)

> for (j in 1:4) {

+ K = 2\*j-1

+ for (i in 1:397) {

+ fhat[i,j] <- knn(X[i], x, y, K)

+ }

+ }

>

> lines(X, fhat[,1], col='red', lwd=2)

> lines(X, fhat[,2], col='blue', lwd=2)

> lines(X, fhat[,3], col='yellow', lwd=2)

> lines(X, fhat[,4], col='orange', lwd=2)

>

> dev.off()

null device

1