R code Sheet

```
objects()
load("~/Documents/STAT_201_B/Homework/HW6/berkhousing.RData")
objects()

head(berkhousing)
dim(berkhousing)

berkhousing = berkhousing[-63,]
dim(berkhousing)

## Question 3(a)

install.packages("fields")
k_n_n <- function(x, y, xseq, k){
   require(fields)
   dmat <- rdist(x, xseq)
   indices <- order(dmat)[1:k]
   return(mean(y[indices]))
}

kseq = 1:(dim(berkhousing)[1]-1)

k_n_n.risk = sapply(kseq, FUN=function(k){
                  sum((berkhousing$price-
                           sapply(1:dim(berkhousing)[1],
                                    FUN=function(i){
                                       k_n_n(x=berkhousing$sqft[-i],
                                                y=berkhousing$price[-i],
                                                xseq=berkhousing$sqft[i],
                                                k=k)
                                    }))^2)
   }
)
kseq[which(k_n_n.risk==min(k_n_n.risk))]

## Question 3(b)

n_w_k.risk <- function(h, x, y){
   require(fields)
   dmat <- rdist(x)
```

```
   K <- dnorm(dmat/h)
   rhat <- sapply(1:length(x), function(j){
      sum(K[,j]/sum(K[,j])*y)
   })
   sum((y-rhat)^2 / (1-dnorm(0)/apply(K, 1, sum))^2)
}

h_opt <- optimize(n_w_k.risk, lower=0.00001,
                          upper=diff(range(berkhousing$sqft)), x=berkhousing$sqft,
                          y=berkhousing$price)$min

n_w_k.risk(h=h_opt, x=berkhousing$sqft, y=berkhousing$price)
```

```
> kseq[which(k_n_n.risk==min(k_n_n.risk))]
[1] 29
```

```
> min(k_n_n.risk)
[1] 869219
```

```
> n_w_k.risk(h=h_opt, x=berkhousing$sqft, y=berkhousing$price)
[1] 893382.8
```

```
> h_opt
[1] 124.7581
```