knn regression

The k nearest-neighbor estimator of $m(t) = E(Y|X = t)$ is defined as

$$\hat{m}(t) = \frac{1}{k} \sum_{i \in N_k(t)} y_i,$$

where $N_k(t)$ is the neighborhood of t defined by the k closest points x_i in the training sample.

Read the Boston housing data from library `MASS`:

```
library(MASS)
data(Boston)
help(Boston)
```

```
## starting httpd help server ...
```

```
##  done
```

Define x and y as follows:

```
x <- Boston$lstat
y <- Boston$medv
# plot(x,y)
```

1. Write a function for computing the k-nn estimator of $m(t)$ for a given value of $t \in R$.

2. Then, define t as a sequence form 1 to 40: `t <- 1:40`.
3. Estimate $m($`t[i]`$)$ for $i = 1, \ldots, 40$ using $k = 50$.
4. Plot `y` against `x`. Then represent the estimated regression function.
5. Repeat the same exercise using different values of k.