Ex 7.6

K-nearest-neighbors is defined as $\hat{y}(x) = \frac{1}{K} \sum_{x_i \in N_K(x)} y_i \quad (*)$

Where $N_K(x)$ is a set of the K nearest elements to the given Point x.

We cand define H as being an indicaktor matrix where the $(i,j)^{th}$ element takes the Value 1 if the j^{th} Observation is in the neighborhood of the i^{th} observation.

More concretely,

$$H_{ij} = K_K(x_i, x_j) = I(\|x_j - x_i\| \leq \|x_{(K)} - x_i\|)$$

given this, we can rewrite $(*)$ as:

$$\tfrac{1}{K} H y = \hat{y}$$

and Since $\frac{1}{K}H$ depends only on the data X and the neighborhood size K we can think of it as being equivalent to the smoother matrix S.

Then, for K-nearest-neighbors, we can write:

$$\hat{y} = \tfrac{1}{K} H y = S y$$

Now, effective number of Parameters or effective degrees of freedom is defined as:

$$df(S) = \text{trace}(S) = \text{trace}(\tfrac{1}{K}H) = \tfrac{1}{K}\,\text{trace}(H)$$
$$= \tfrac{1}{K}(n) = n/K$$

(Since every observation must be a member of its own neighborhood)