7. The table below provides a training data set containing six observations, three predictors, and one qualitative response variable.

Obs.	X_1	X_2	X_3	Y
1	0	3	0	Red
2	2	0	0	Red
3	0	1	3	Red
4	0	1	2	Green
5	−1	0	1	Green
6	1	1	1	Red

Suppose we wish to use this data set to make a prediction for Y when $X_1 = X_2 = X_3 = 0$ using K-nearest neighbors.

(a) Compute the Euclidean distance between each observation and the test point, $X_1 = X_2 = X_3 = 0$.

(b) What is our prediction with $K = 1$? Why?

(c) What is our prediction with $K = 3$? Why?

(d) If the Bayes decision boundary in this problem is highly nonlinear, then would we expect the *best* value for K to be large or small? Why?

Answer/

K-Nearest Neighbors Classification

(a) Euclidean Distances

$$d = \sqrt{(x_1 - 0)^2} + (x_2 - 0)^2 + (x_3 - 0)^2$$

$$d = \sqrt{(0 - 0)^2} + (3 - 0)^2 + (0 - 0)^2$$

Obs 1: d= 3

$$d = \sqrt{2^2} + (0-0)^2 + (0-0)^2$$

Obs 2: d= 2

Obs 3: d= 3.16

Obs 4: d= 2.24

Obs 5: d= 1.41

Obs 6: d= 1.73

(b) Prediction for K = 1

Nearest neighbor = **Obs 5** (distance ≈ 1.41)
Its class = **Green**

Prediction: Green, because KNN assigns the class of the closest point.

(c) Prediction for K = 3

Three closest points:

Votes:

- Red = 2
- Green = 1

Prediction: Red, because the majority of the 3 nearest neighbors are Red.

d) Best K if Bayes decision boundary is highly non-linear?

Small K is better.

because

A highly non-linear boundary requires a flexible model. Small K allows KNN to adapt closely to local patterns, while large K smooths too much and misses complex boundaries.