3.) ISLR 2.4.1

a) Better - a more flexible approach would be obtained will fit the data closer and with the large sample size a better fit than an inflexible approach.

b) Worse - a flexible method would overfit the small no. of observations

c) Better - with more degrees of freedom, a flexible model would obtain a better fit.

d) Worse - Flexible method would fit to the noise in the error terms and increase variance.

4.) ISLR 2.4.7

a)

Observations	X1	X2	X3	Distance (0,0,0)	Y
1	0	3	0	3	Red
2	2	0	0	2	Red
3	0	1	3	$\sqrt{10} \approx 3.2$	Red
4	0	1	2	$\sqrt{5} \approx 2.2$	Green
5	-1	0	1	$\sqrt{2} \approx 1.4$	Green
6	1	1	1	$\sqrt{3} \approx 1.7$	Red

b) If $K=1$, then $x_5 \in \mathcal{N}_0$

$$P(Y=\text{Red} \mid X=x_0) = \frac{1}{1} \sum_{i \in \mathcal{N}_0} I(y_i = \text{Red}) = I(y_5 = \text{Red}) = 0$$

$$P(Y=\text{Green} \mid X=x_0) = \frac{1}{1} \sum_{i \in \mathcal{N}_0} I(y_i = \text{Green}) = I(y_5 = \text{Green}) = 1$$

$\Rightarrow$ Observation #5 is the closest neighbour for $K=1$.

c.) If $K=3$, then $x_2, x_5, x_6 \in \mathcal{N}_0$

$$P(Y=\text{Red} \mid X=x_0) = \frac{1}{3} \sum_{i \in \mathcal{N}_0} I(y_i = \text{Red}) = \frac{1}{3}(1+0+1) = \frac{2}{3}$$

$$P(Y=\text{Green} \mid X=x_0) = \frac{1}{3} \sum_{i \in \mathcal{N}_0} I(y_i = \text{Green}) = \frac{1}{3}(0+1+0) = \frac{1}{3}$$

$\Rightarrow$ Observation #2, 5, 6 are the closest neighbors for $K=3$. 2 is Red, 5 is Green, and 6 is Red.

d.) Small. A small K would be flexible for a non-linear decision boundary, whereas a large K would try to fit a more linear boundary because it takes more points into consideration.