Class 4 Clicker Question 1

$K = 5$

- $R^2 = 0.9242$ and CV RMSE $= 0.0803$

$K = 10$

- $R^2 = 0.9247$ and CV RMSE $= 0.0826$

Which model appears to predict better?

A It's impossible to say

B $K = 5$ because it has the smaller R^2 value

C $K = 10$ because it has the larger R^2 value

D $K = 5$ because it has the smaller CV RMSE value.

E $K = 10$ because it has the larger CV RMSE value

Class 4 Clicker Question 2

10-fold cross-validation is being used here, i.e., $v = 10$.
Consider the linear-splines model with $K = 5$ knots. How many different model fits are there (via `lm` calls)?

A 9

B 10

C 1

D 2

E n.

Class 4 Clicker Question 3

10-fold cross-validation is being used here, i.e., $v = 10$. Consider the linear-splines model with $K = 5$ knots. How many observations are being used in each model fit (via `lm`)?

A n

B 10

C 90

D Approximately $0.1n$

E Approximately $0.9n$.

Class 4 Clicker Question 4

10-fold cross-validation is being used here, i.e., $v = 10$.
Consider the linear-splines model with $K = 5$ knots and observation 1 in the data set.
How many times is observation 1 predicted as a test case in computing CV RMSE?

A 1

B 9

C 10

D Approximately $0.1n$

E Approximately $0.9n$.

Class 4 Clicker Question 5

$$E[Y|x] = \beta_0 + \beta_1 x + \beta_2 x^2 + \sum_{k=1}^{K} \beta_{k+2} f_k(x)$$

In this still a linear model that can be fit via `lm`?

A No, because there is an x^2 term

B No, because the $f_k(x)$ functions are nonlinear

C Yes, because it is a sum of terms

D Yes, because it is linear in the β_j parameters

E I have no idea.

Class 4 Clicker Question 6

For the lidar data, we have considered different numbers of knots in quadratic-splines models.

Based on 10-fold CV, what number of knots would you choose if you want accurate prediction?

A 1

B 2

C 4

D 6

E 10.