**DSI -06 Homework 5:**

Introduction with Statistical Learning with Applications to Python (ISLP) pg. 286-287

9. In this exercise, we will predict the number of applications received using the other variables in the College data set.

1. Split the data set into a training set and a test set.
2. Fit a linear model using least squares on the training set, and report the test error obtained.
3. Fit a ridge regression model on the training set, with λ chosen by cross-validation. Report the test error obtained.
4. Fit a lasso model on the training set, with λ chosen by cross validation. Report the test error obtained, along with the number of non-zero coefficient estimates.

Additional Practice Questions:

Explain the importance of a training and test set to a non-technical audience.

How would you describe this exercise in an interview to both a technical and non-technical interviewer? What are the key insights you would want to show?

Can you think of a business context where this exercise would have applications?