Statistical Learning Questions:

1. Why is standardization useful for inference in a linear regression model
2. For homework 3, what columns of data would you standardize and include in your inference model.
3. What is model bias – Linear regression is a high bias model
4. What is model variance – Decision tree is high variance
5. What is the difference between reducible and irreducible error.  Provide a couple of concrete examples of irreducible error.
6. What is maximum likelihood estimation – estimating distribution parameters
7. How does maximum likelihood relate to least squares in linear regression.
8. What is inference (as per the ISLR book). What’s the difference between inference and prediction.  Describe how you might use inference in a data science problem.
9. What’s the difference between parametric and non parametric methods.
10. Describe the flexibility vs. interpretability trade off.
11. What is the theoretical minimum training error which can be achieved with a machine learning model. – Irreducible error

Model should have balance of variance and bias.