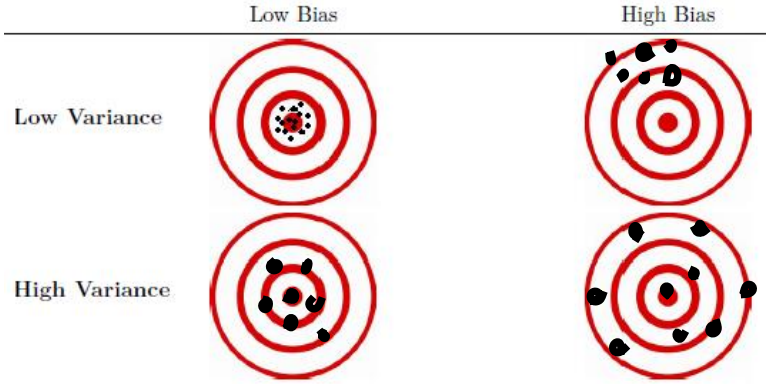
Quiz 1. Overview of Statistical Machine learning

Name: Lisa Liubovich Attempt (circle one): 1 2

1. (3 pts) Draw points on each image where the cloud of points might be relative to the true value point (the red dot in the center) for the specified case. The case is the combination of the row and column heading, e.g., Low Variance and Low Bias, or Low Variance and High Bias. (The Low Variance and Low Bias case has been provided as example.)



2. (7 pts) An appraiser needs to come up with the value of a particular house whose owner applied for a mortgage. She found five (5) similar properties recently sold in the neighborhood, and she plans to use their characteristics (price, square footage, number of bedrooms, backyard area, etc.) to estimate the value of a given house.
   (a) Is this an inference or a prediction problem? Briefly justify your answer.
   (b) Would regression or classification methods be more appropriate? Briefly justify your answer.
   (c) Would you recommend a flexible method?
   (d) (Stat 627) Briefly justify your answer in (c)

2.a. This is a prediction problem because the appraiser isn't trying to find a deep understanding of the relationship between each individual characteristic and the value of the house. She is only interested in getting an accurate prediction of the value of the house given the house characteristics.
2.b. Regression methods would be more appropriate because the response value is quantitatve/ numerical.
2.c and d. I would not recommend a flexible methods, as even though we don't really need the model to be super interpretable, less flexible methods often obtain more accurate predictions because more flexible methods have more potential for overfitting (which makes predictions less accurate).