ISLR 2.4 Conceptual Problem 1 (page 52) [63 in pdf]

For each of parts (a) through (d), indicate whether we would generally expect the performance of a flexible statistical learning method to be better or worse than an inflexible method. Justify your answer.

1. The sample size n is extremely large, and the number of predictors p is small.

**A more flexible learning method would have better performance because there is a large amount of sample data for it to produce a tighter and more accurate model fit than more generalized inflexible learning methods.**

1. The number of predictors p is extremely large, and the number of observations n is small.

**A more flexible learning method would have worse performance since one of its drawbacks is the tendency to overfit (or find trends that don’t exist in) data. A small number of observations would make it easier for a more flexible learning method to overfit because there are not enough data points for it to make accurate deductions about the data.**

1. The relationship between the predictors and response is highly non-linear.

**A more flexible learning method would have better performance because an inflexible method would not be able to accurately model a complex non-linear relationship.**

1. The variance of the error terms, i.e. σ2 = Var(ϵ) is extremely high.

**A more flexible learning method would be preferred since it tends to produce a tighter model fit with lower variance to counteract the high irreducible variance.**