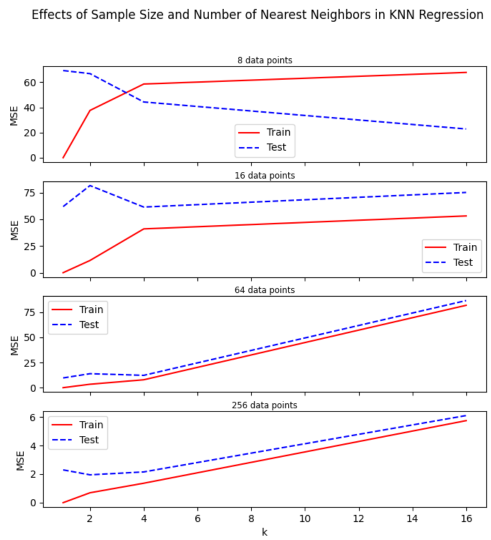
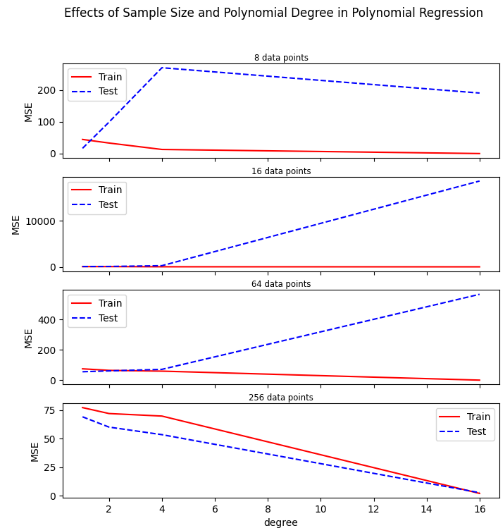
a.

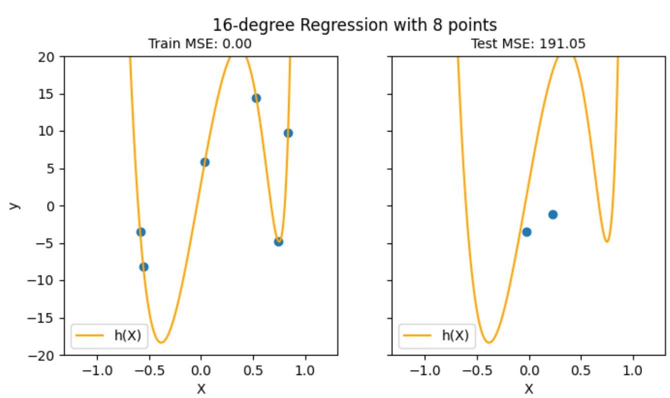




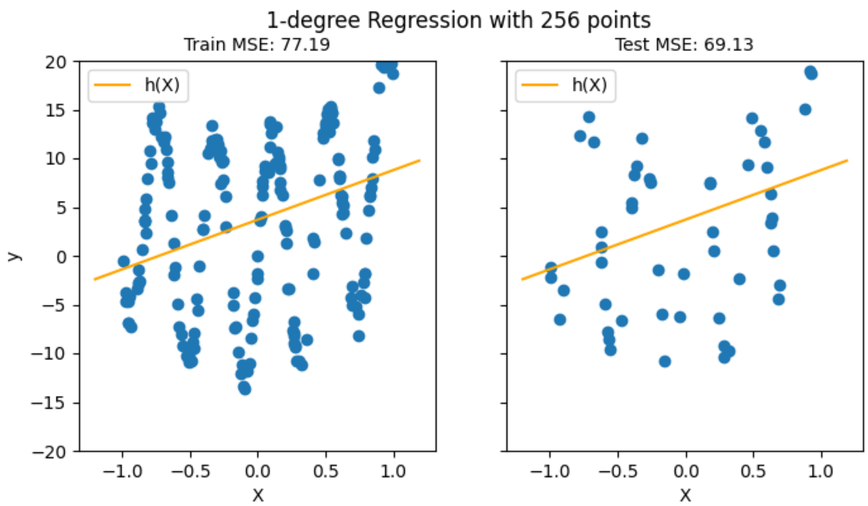
In the case of KNN Regression, we see that MSE decreases significantly as more data is added, but error also tends to increase as k is increased. In the case of Polynomial Regression, we see that the error also decreases significantly as the number of data points is increased, and that while the training error tends to decrease quickly with polynomial degree, the test error increases for large polynomial degrees (except for the 256-point case).

In general, we see that as the learned functions get more complicated, our polynomial models fit the training data better. For smaller amounts of data, they also tend to fit the testing data worse, implying that we are overfitting. In the case of our nearest neighbor regressions, increasing k tends to lead to a worse fit of the training data (this makes sense as k=1 leads to memorization of the training data). Performance on the testing data tends to at first improve, then increase, with increasing k. These findings are consistent with what we have seen in class.

b.



The above plot is clearly overfitting: we see that the model has memorized the training data (it gets it perfectly) and is wildly incorrect on the testing data.



The above plot is an example of underfitting: the first degree polynomial cannot learn a function complex enough to appropriately fit the training data, so both the training and testing data have high MSE values.

c.

Overfitting and underfitting depend in a complex way on H, f(X), and n. For large N, a more complex f(X) (a higher degree polynomial for example) will fit the data better than a less complex one: the less complex f(X) will tend to underfit. For small N, a more complex f(X) will tend to overfit. Overfitting also can be caused by the relation between H and f(X). If H is much simpler than f(X), one will tend to overfit; if H is much more complicated than f(X), one will tend to underfit. This overfitting will tend to be mitigated by more data (larger n), while the underfitting will not be.