***Question 1**: How many parameters (i.e., weights) are present in each of the three models you used? Explain how you know/determined this. Answers given without an explanation or supporting calculations receive zero credit.*

The linear model is 3 weights. There are 2 inputs and 1 output. The function is in the form of $y = w_0 + w_1x_1 + w_2x_2$

The quadratic function has 6 weights. If the linear model is squared the terms are in the form of $1 + x_1 + x_2 + x_{11} + x_{22} + x_{12}$. Which results in 6 potential coefficients.

The cubic function has 10 weights. It has 1, along with x_1, x_{11}, x_{111}, and x_2, x_{22}, x_{222}. It also has x_{12}, x_{112}, and x_{122}. This can also be determined by the statistical function "choose". The form is n+d C d. Where n is the number of measurable variables, and d is the degree, in this case 2 and 3 respectively. So (3+2)! / ((3! * 2!) = 5*4 / 2! which is 10.

***Question 2:** Suppose we fit a 5th-order polynomial to the same data. Based on trends in **training** loss you observed for the linear, quadratic and cubic models, do you expect training loss for the 5th-order polynomial to be higher or lower than the cubic polynomial? Explain the rationale for your answer.*

I expect the training loss to continue downwards as the training loss has decreased with an increase in order of polynomial. While it *could* overfit the data for testing, training loss will always go down (or at least stay the same) as order increases, because the lower order polynomials are subsets of the higher order polynomials.

***Question 3**: If any of these models overfit the training data, what would you expect to see in the results?*

If they overfit the training data then I would expect the testing data MSE to begin to increase.

***Question 4**: When considering both model accuracy and model complexity, which of the three models do you consider to be best? That is, which would you say should be the final model selected for this problem? Explain your reasoning.*

I think the quadratic model should be selected for this problem. The most accurate model for the testing was the cubic model. The cubic model is also the most complex.

The cubic and quadratic model testing errors are relatively close, so due to the accuracy and simplicity the quadratic model is the best.