1a. Training set is used to train the model (form the original regression). Validation set is used to validate the model (during the validation process to avoid overfitting). The test set is used to evaluate the accuracy of the model (and produce a real-world accuracy prediction).
1b. Split data into train, validation, and testing sets. For each complexity (of your choice), train the model on the training set, and test on the validation set. Keep track of which model produces the lowest validation error. Then test on the test data, and find the model with both the least test and validation errors.

2a. Good.
2b. Bad. Model is symptomatic of overfitting, adding more complexity will not solve the issue.
2c. Bad. A model with 0% test error voids the point of using a test set -- it no longer reflects the model's real-world performance.
2d. Good. A k of 10 is within reasonable range of modern practices.

3a. Basic Model: 248812 Train, 282446 Validation. Advanced Model: 192134 Train, 235226 Validation. Since the advanced model has lower train and validation error, it should perform better in the real world.

4. Zip Codes aren't a metric like square foot or price, and calculations with Zip Codes are meaningless. In a linear regression model, it may learn something like a Zip Code from a certain range of values might be worth more or less than Zip Codes from another range of values, which isn't how value is assigned in the real world.