ANSWER KEY

1. B – O(n)
2. C – POLYNOMIAL REGRESSION
3. B – GRADIENT DESCENT
4. C – LASSO
5. C – BATCH GRADIENT DESCENT
6. B – FALSE
7. C – IT DOES NOT MATTER WHETHER HALF IS THERE OR NOT
8. B – CORRELATION
9. A & B
10. A & C
11. C & D

12. We can use Gradient Descent because of the large number of features. Gradient Descent is an optimisation algorithm for finding a local/global minimum on a differentiable function. We may use Stochastic or Mini-Batch versions of the Gradient Descent as they do not need to load the entire dataset to make a step of gradient descent.

13. Gradient Descent Algorithm suffers if features in the training set have very different scales. This is because the model will need a longer time to reach the global minimum. This can be mitigated by feature scaling which will help in quicker descent.
    Normal Equations method does not require normalising of the features hence it does not suffer from features in training set having very different scales.