1. **What is the gradient descent method and how is it computed? What were we using GD for in class?**
   The Gradient descent method finds the gradient of the loss function (the direction of greatest change), and goes in the opposite direction by some learning rate to minimize the loss function.

2. **What is the learning rate?**
   The learning rate is the speed in which this method converges (ie. the step size we use in the direction opposite of the gradient)

3. **Why normalize the data?**
   Sometimes it is hard to find a starting point to converge, and normalizing can help with that regard.

4. **What does it mean when the gradient decent method converges?**
   When you the gradient approaches 0, meaning your function has essentially stopped changing and is (locally) minimized or maximized.

5. **State your findings from the logistic regression you conducted with the BreastCancer data set with respect to the above.**
   (Uploaded results on github)