Question 1:
Training set accuracy: 0.9177
Test set accuracy: 0.9172
The training set accuracy is slightly higher than the test set accuracy. I think this slight different is due to the fact that the training set data is specifically designed for the model to study, but the test set data are not chosen that intentionally. Thus, the training set data might be biased to some extent and the model gives higher accuracy.

Question 2:
After changing the loop time to 10, the accuracy becomes 0.7102
After changing the loop time to 10000, the accuracy becomes 0.9228
In every loop, we randomly choose 100 data points from our training set. Thus, with larger number of loops, the data points used for training becomes more and the model can better identify the digits with more training. Thus, the accuracy increases as the number of loops becomes larger.

Question3:
When W and b are initialized with zeroes, the accuracy is 0.9172
When W and b are initialized with ones, the accuracy is 0.9166
There is no significant difference between two kinds of initializations. This observation is reasonable because the training result should not depend on our initialized values. As the training set becomes larger, the impact of initializations on the result becomes smaller.