Problem 1 (30 points)

a. Part a (OLS fit and data plot) - 5pts
    i. 1pt deducted for missing plot in any part of question
b. Part b (correct MSE calculation) - 5pts
    i. 1pt deducted for missing MSE in any part of question
c. Part c
    i. Rolling mean - 8pts
        1. 2pts for each window size calculation (5,7,15)
    ii. First order differencing - 8pts
        1. 1pt deducted for incorrectly handling 0th point
    iii. Transformation of your own - 4pts

Note: Partial credit was deducted for partially incorrect implementations of any data transformation, depending on the student's implementation.

Problem 2 (20 pts)

a. Plot of data - 3pts
b. Correct decision boundary **with explanation** - 10pts
    i. 3pts deducted if no explanation
    ii. 2pts deducted if decision boundary is off due to some small computational error
c. Classification of test points - 5pts (1pt for each test point)
d. Reasoning for classification decisions on decision boundary (part c) - 2pts

Problem 3 (20 pts)

a. Correct construction of KNN/decision tree for both datasets (iris, wine), with reasonable accuracies for each - 12pts (3pts for each classifier, for each dataset)
b. Pick out best k's - 4pts (2pts for each dataset)
c. Reasoning for difference in results of KNN/decision tree - 4pts
    i. 1-2pts deducted if reasoning is incomplete or incorrect

Problem 4 (30 pts)

a. Plots for n=5, n=50, n=500 - 3pts
b. Sampling random data across p dimensions - 5pts
c. Calculating max norm correctly - 6pts
    i. 5pts deducted if incorrect due to calculation error
d. Correctly obtaining nearest neighbors - 10pts
    i. 5pts deducted if incorrect due to calculation error
    ii. 10pts deducted if no nearest neighbor calculation at all
e. Plotting against median curve for each n - 3pts
f. Interpretation of plots - 3pts
    i. 1pt deducted if interpretation is incomplete