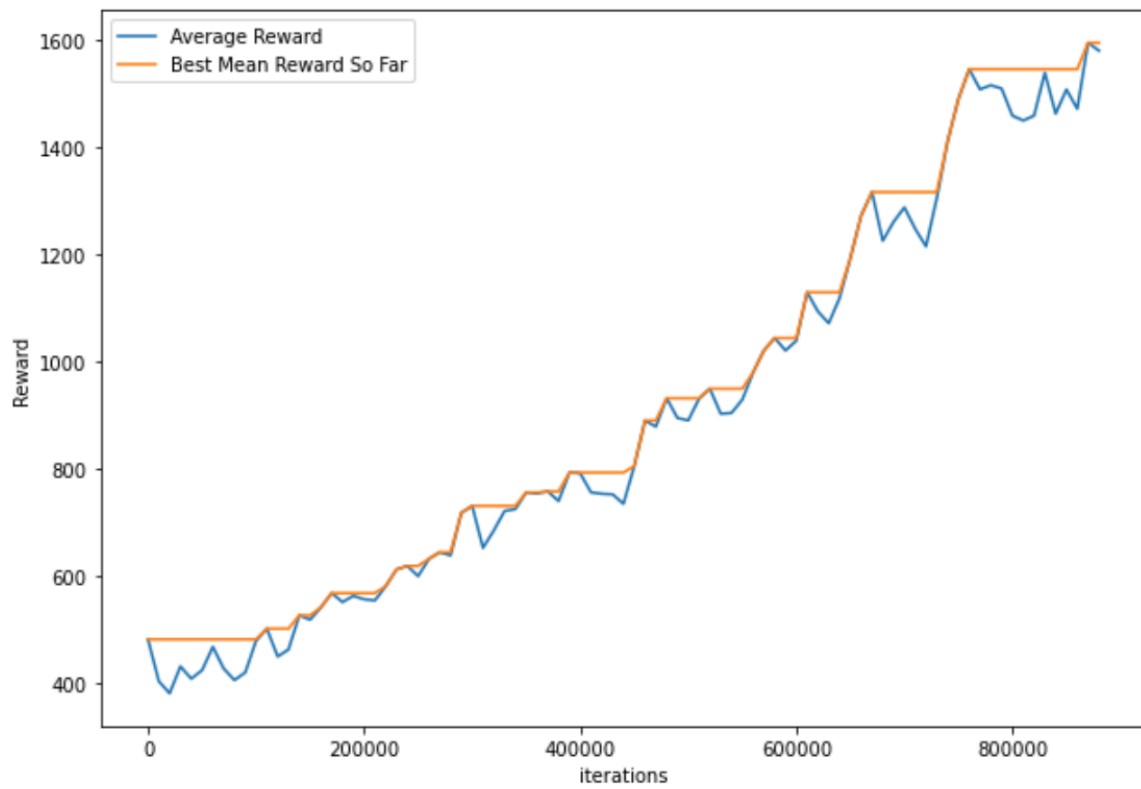
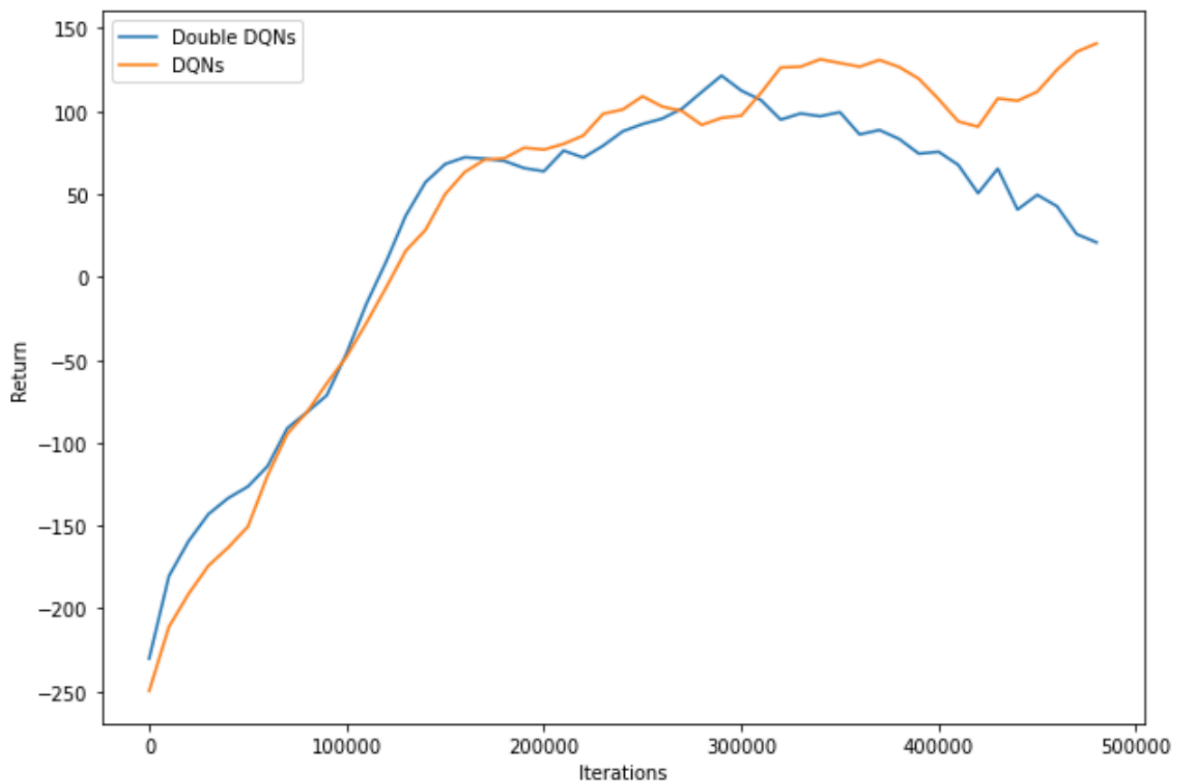


HW3 Report

Question 1



Question 2



HW3 Report

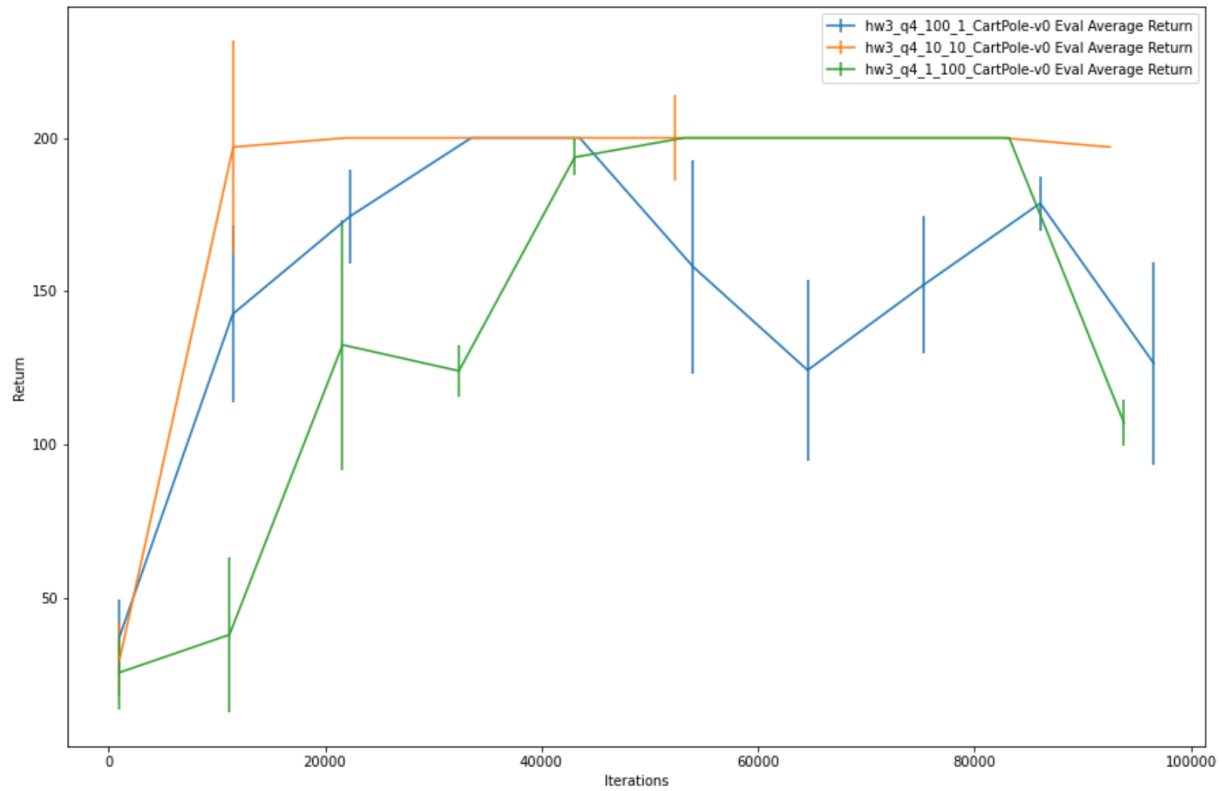
Question 3



The batch size increase resulted in less volatile reward graphs. I believe this is because if we have higher batch size, there is less variance in the calculated gradient, resulting in a more accurate gradient (because the estimator is unbiased) and thus a less volatile return graph.

HW3 Report

Question 4



The one that worked best was q4_10_10

HW3 Report

Question 5

