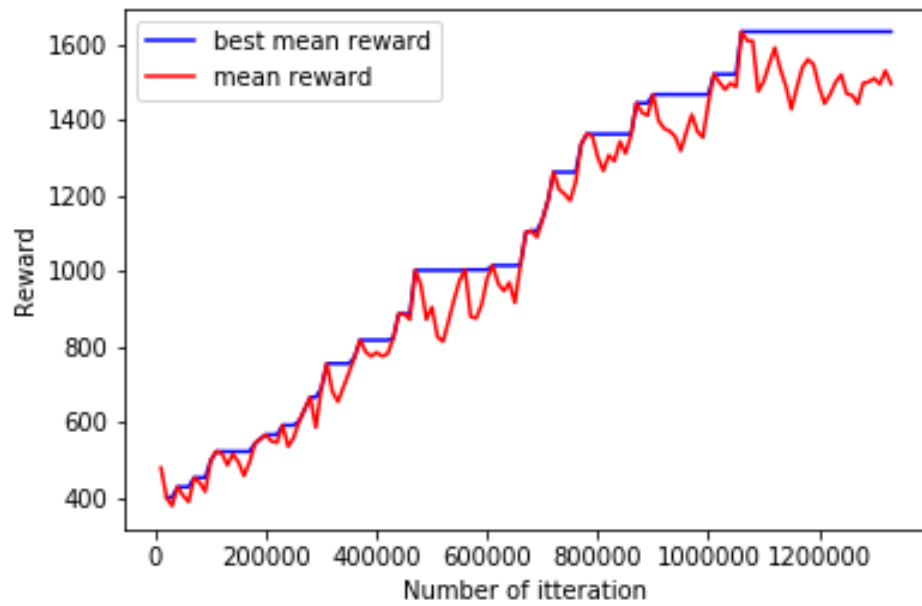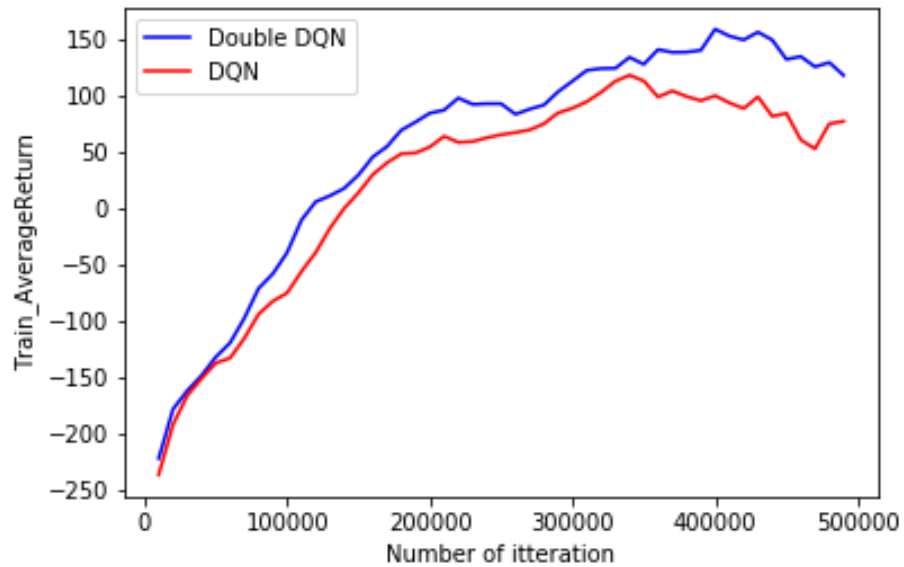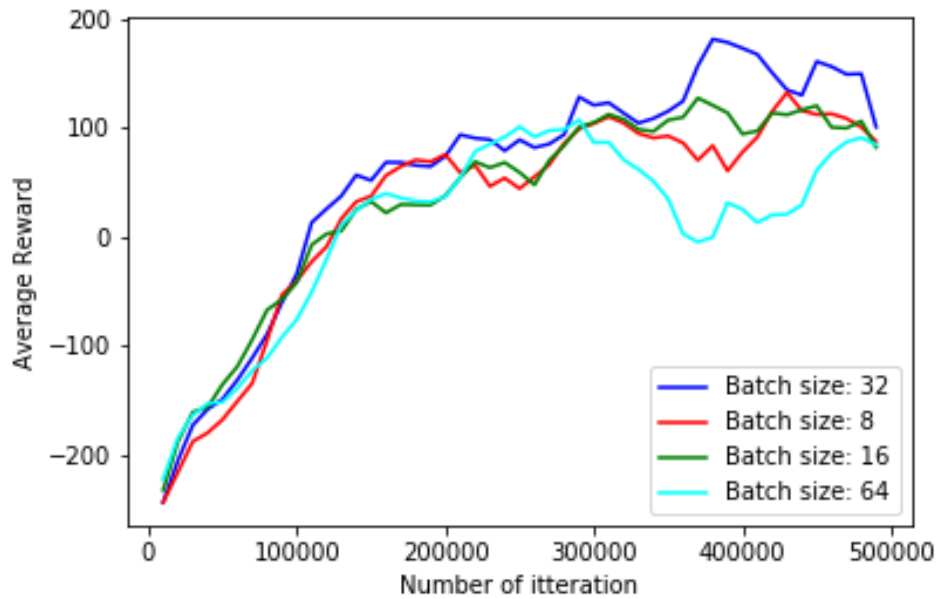# Question 1:

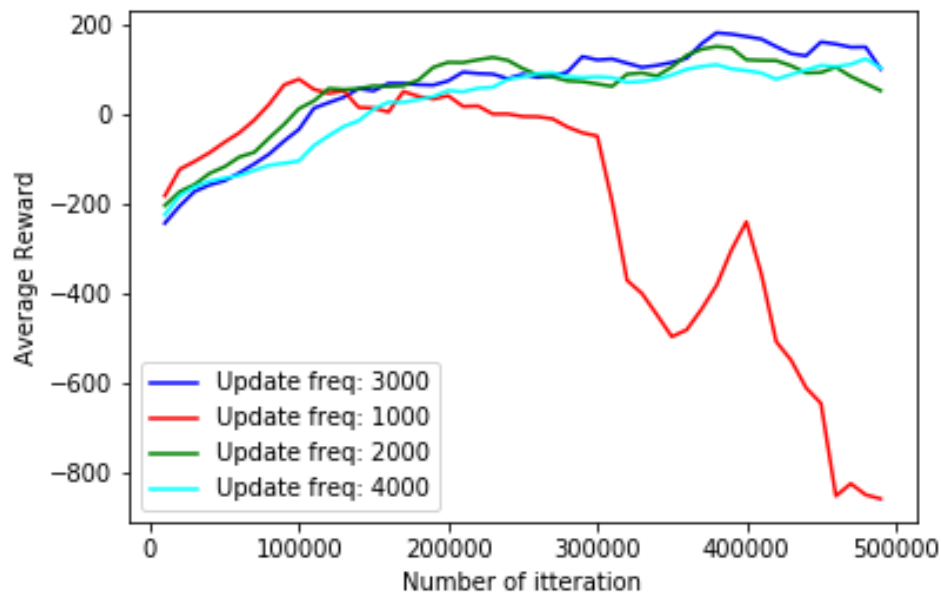Performance of Q-learning in the MsPacman-v0 env.

Question 2:

Performance of Q-learning in the LunarLander-v3 env. With three different seeds
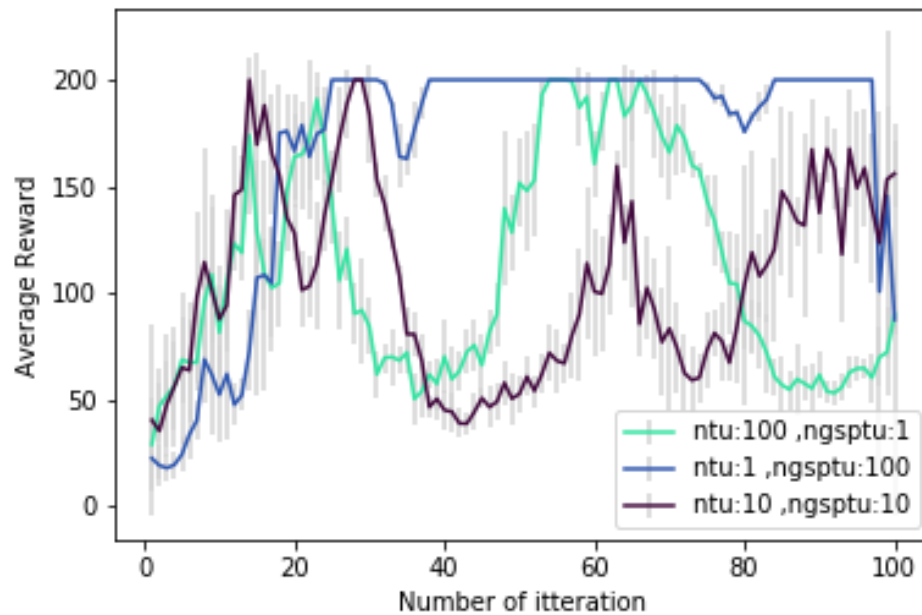
Question 3:

Performance of the Q-learning in LunarLander env with different batch sizes, the batch size equal to 32 has the highest reward
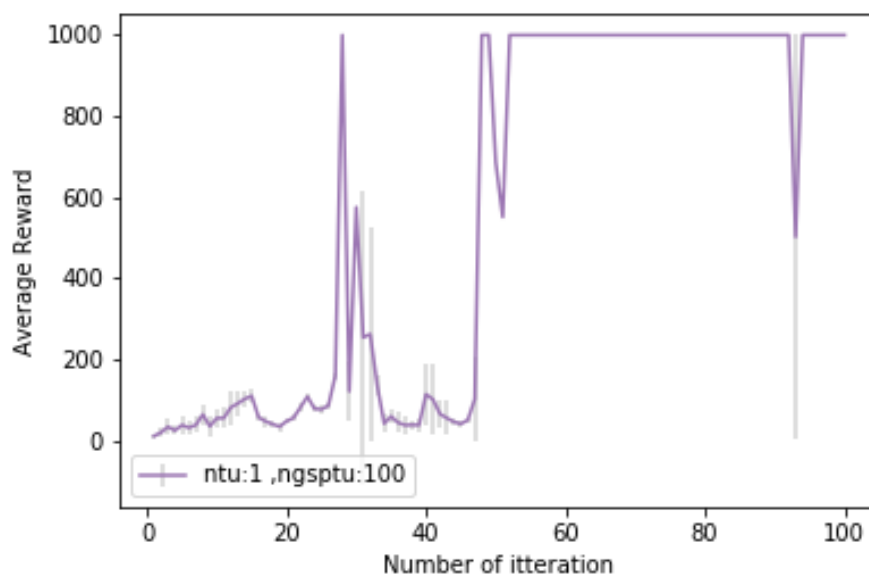

Performance of the Q-learning in LunarLander env with different update frequency of target network, the freq equal to 3000 has the highest reward
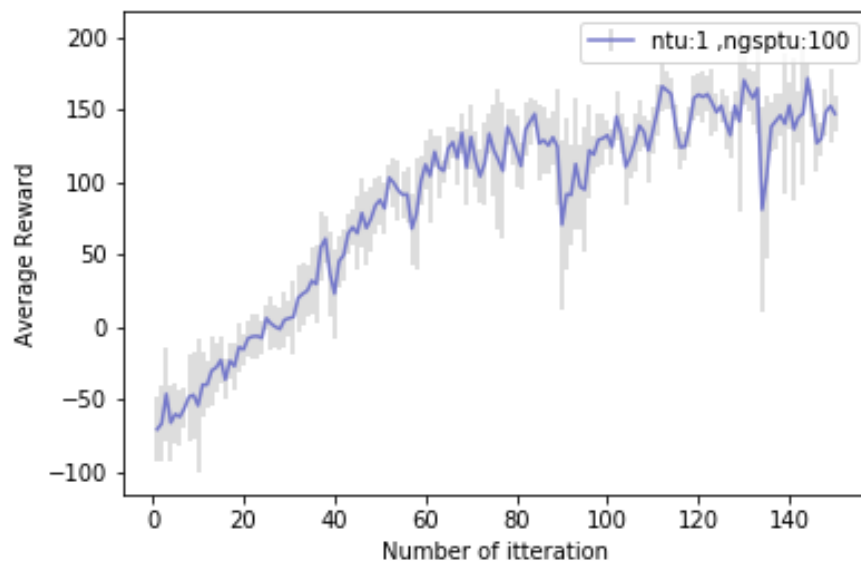
Question 4:

Performance of actor-critic method in cart-pole env. Num of target update= 1 and num of grad steps per target update = 100 is chosen for the following experiments

Question 5:

Actor-critic in InvertedPendulum-v2 env.

Actor-critic in the half chita env.