Average Reward vs. Number of Training Iterations

discount=1
discount=0.999
discount=0.99