# Homework #3 Report

Alejandro Escontrela

## Q1: Average return on pacman

Chart, line chart

Description automatically generated

## Q2: Double Q-Learning (DDQN)

Chart, line chart

Description automatically generated

## Q3: DQN Hyperparameter Sweep

Chart, line chart

Description automatically generated

For this part of the assignment, I performed a sweep over the batchsize. Using a smaller batchsize leads to more stable learning.

## Q4

Chart, line chart

Description automatically generated

While all three configurations achieve the highest score, it is evident that using 100 gradient steps and one critic update per update leads to more stable performance.

## Q5 – Inverted Pendulum

Chart, line chart

Description automatically generated

## Q5 – HalfCheetah

Chart, line chart

Description automatically generated