

Is the Agent policy better than a Heuristic?

DQN Runs [a-i]

Mean Reward per Step

