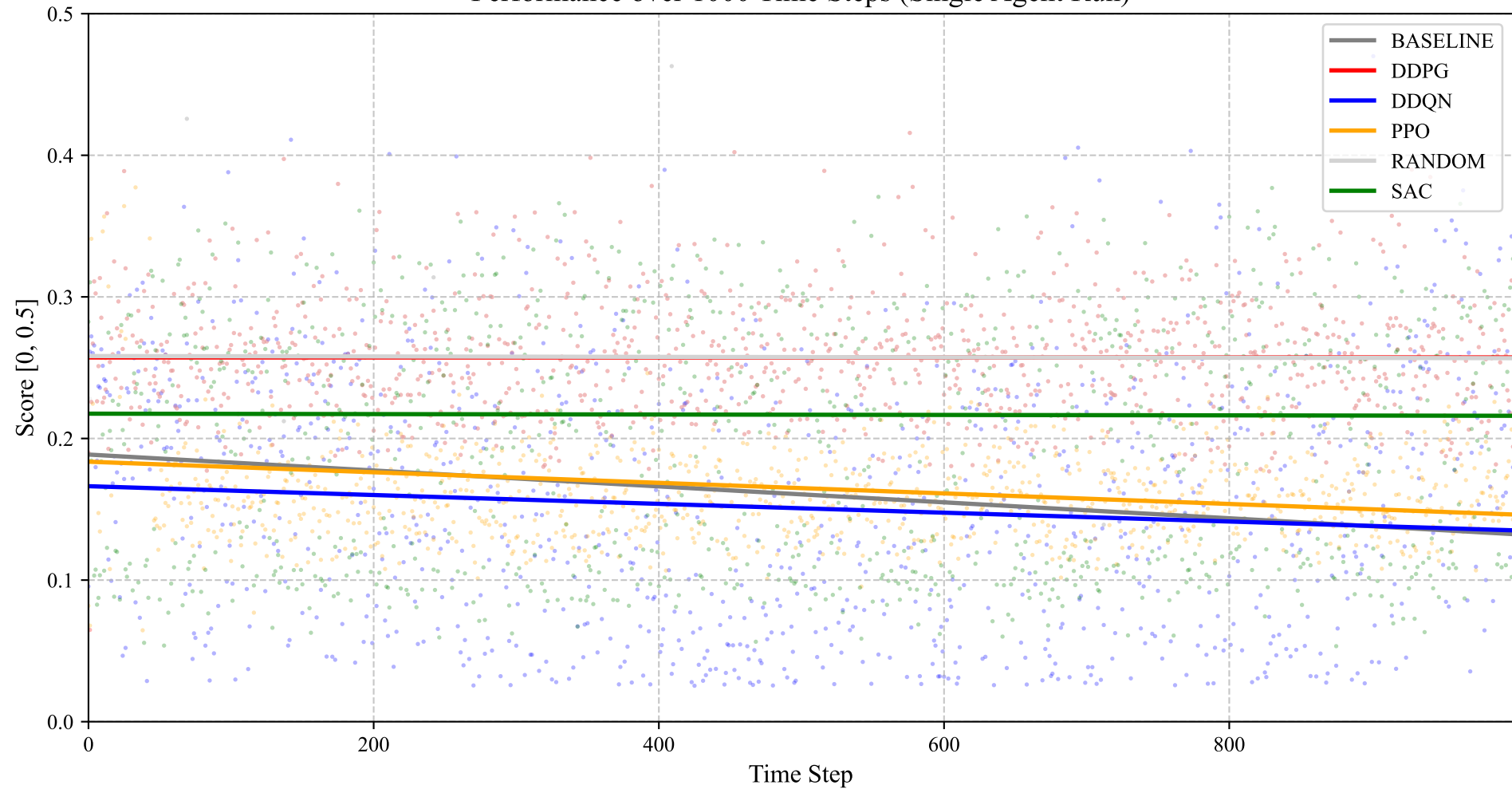


Performance over 1000 Time Steps (Single Agent Run)



Linear Regression (Gradient $\times 10^3$, MSE): BASELINE: (-0.06, 0.90), DDPG: (0.00, 0.00), DDQN: (-0.03, 0.03), PPO: (-0.04, 0.02), RANDOM: (-0.00, 0.00), SAC: (-0.00, 0.03)