Q-Learning CartPole — Stability Across Seeds (b=7)