SandingEnvMiddle

Legend:
- ppo
- ppo_extension

Y-axis: Average Reward (0.0 to 1.4)
X-axis: Timestep (0.0 to 1.0, 1e5)