Comparison of our methods and the baselines

Legend:
- π_{mm} = RL, $\pi_{explore}$ = RL, π_{qa} = Hand-crafted
- π_{mm} = RL, $\pi_{explore}$ = Hand-crafted, π_{qa} = Hand-crafted
- π_{mm} = Random, $\pi_{explore}$ = Hand-crafted, π_{qa} = Hand-crafted
- Baseline (History of observations)

X-axis: History Room Sizes — 1, 6, 12, 24, 48, 100

Y-axis: Average Total Test Episode Rewards