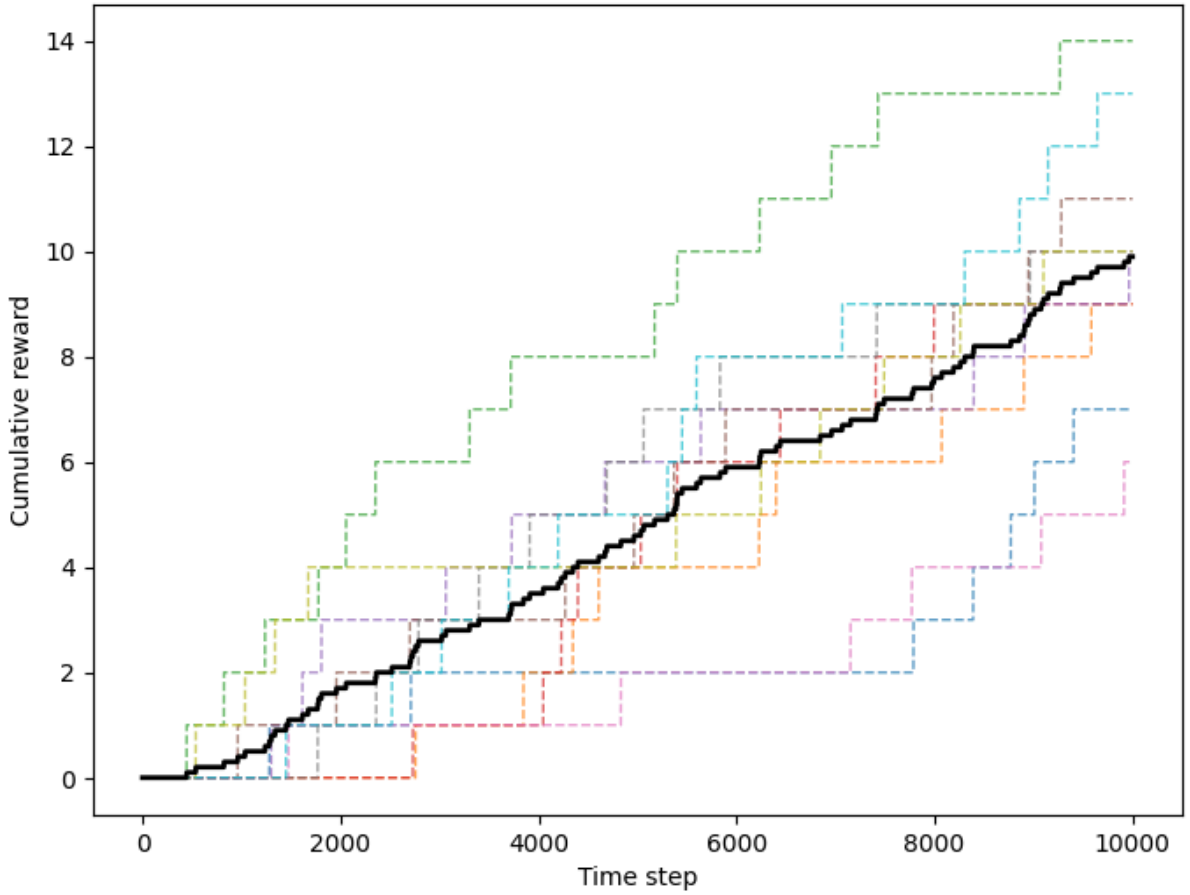
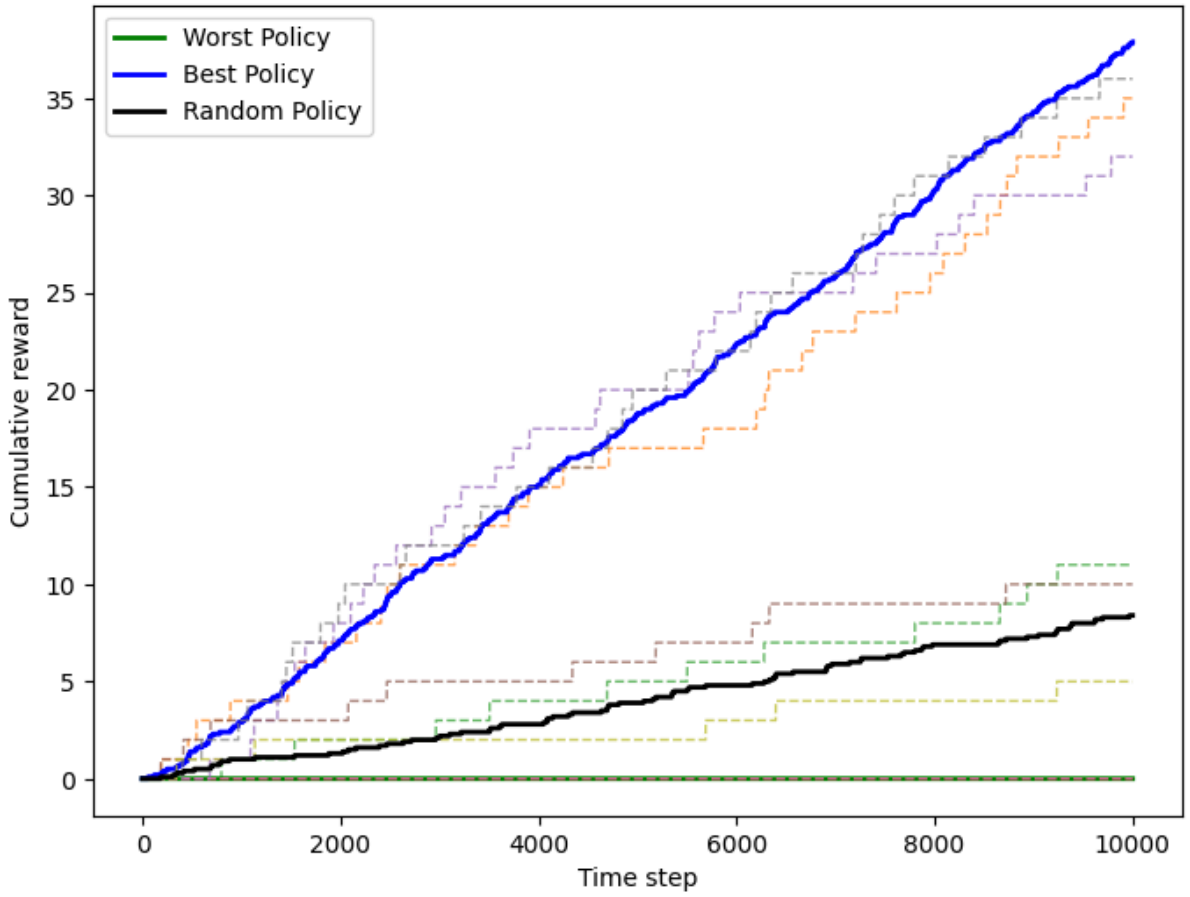
Question3:How do you think this compares with your manual policy? (You do not have to run your manual
policy for 104 steps!) What are some reasons for the difference in performance?

Answer: The Random Policy is not as good as the manual policy. The manual policy is better because user can choose to move right and up to reach the goal since the user is aware of the environment but random policy is not aware of the environment and it can move in any direction thus the random policy often moves in the wrong direction.



Question4:Describe the strategy each policy uses, and why that leads to generally worse/better performance.

Answer: For the worse policy since we know the environment we can move the agent away from the goal by mostly moving left or down because agent will be doing the opposite actions that would reach it to the goal. For the better policy we can move the agent to the goal by mostly moving right or up because agent will be doing the actions that would reach it to the goal since we know the environment the best way to reach the goal is to move right or up.


Worst Policy
Best Policy
Random Policy
Cumulative reward
Time step
0
5
10
15
20
25
30
35
0
2000
4000
6000
8000
10000