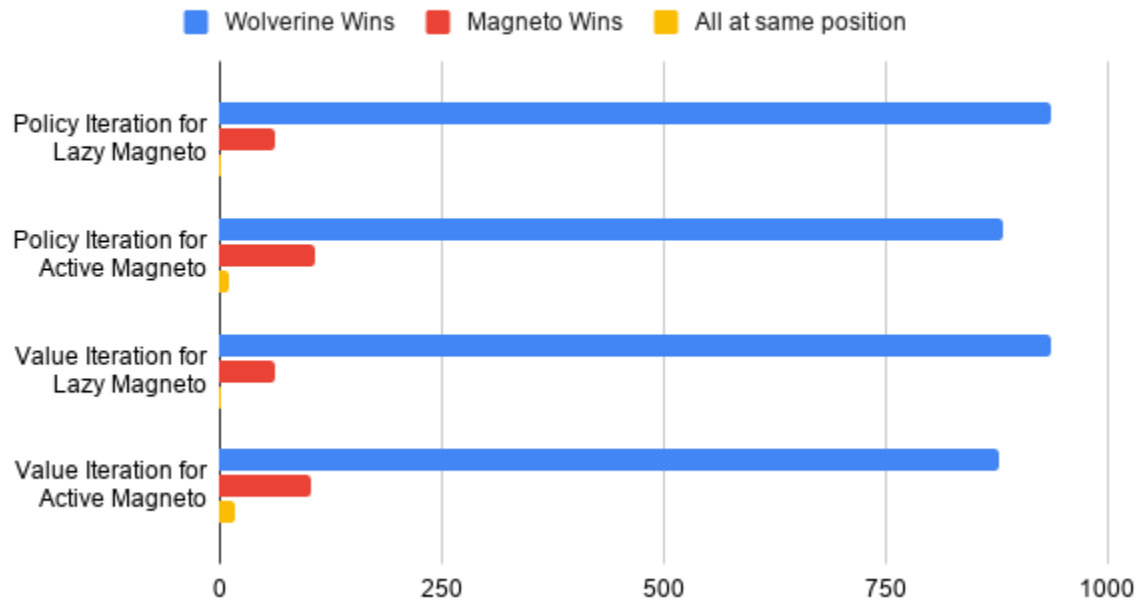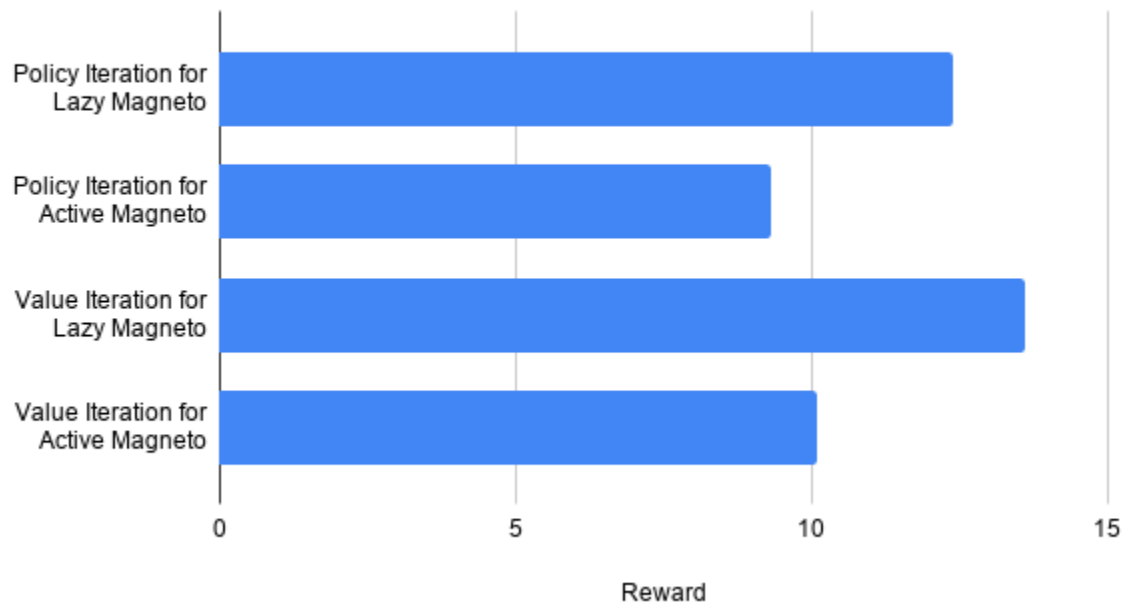Q2. In this question, We got the following results from the output of the program
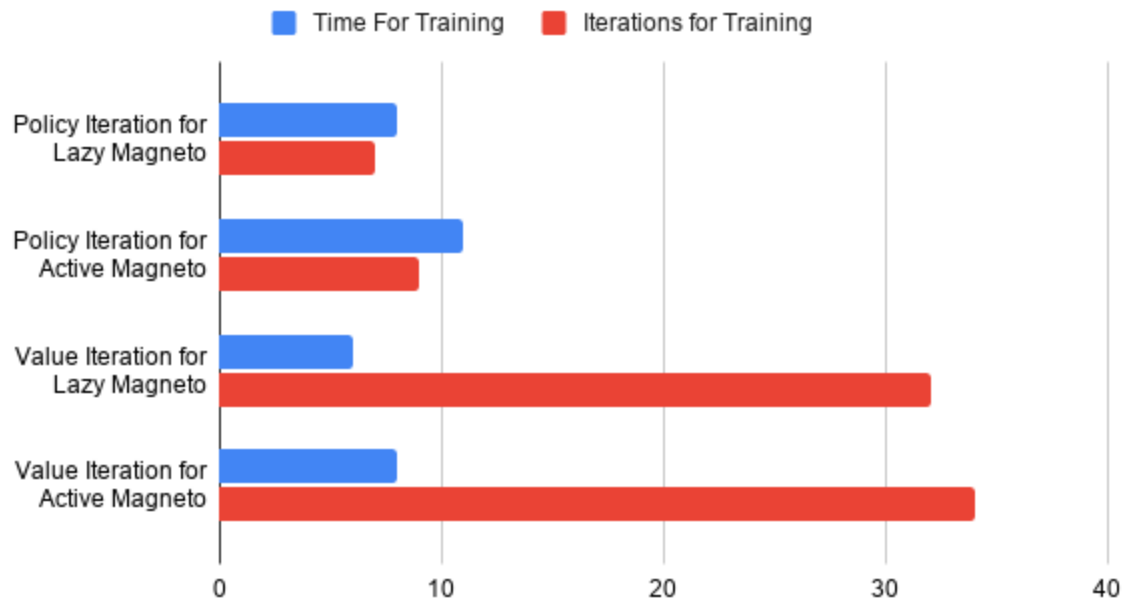
Total Wins in 1000 games

Legend: Wolverine Wins, Magneto Wins, All at same position

Categories (top to bottom):
- Policy Iteration for Lazy Magneto
- Policy Iteration for Active Magneto
- Value Iteration for Lazy Magneto
- Value Iteration for Active Magneto

X-axis: 0, 250, 500, 750, 1000

Reward vs Different Configurations

Categories (top to bottom):
- Policy Iteration for Lazy Magneto
- Policy Iteration for Active Magneto
- Value Iteration for Lazy Magneto
- Value Iteration for Active Magneto

X-axis: 0, 5, 10, 15 — Reward

Time and Iterations for Training

These are the results

Observation:

As we can see the difference between wolverine wins and magneto wins is more in the case of the lazy magneto.

Time for training in the Active magneto is more than Lazy magneto