Your Tasks

1. **(TURN THIS IN, 5 points)** First, read the assignment specification and *__estimate how long you think it will take you__* and write it down.
2. *Task 1: Understand Policy Iteration:*
   A. Investigate PolicyIterationAgent.py's implementation to understand how policy is stored and updated during the policy iteration. Note: do not spend too long on this, understanding the high level process is enough.
   B. Run PolicyIterationLabtest1 and PolicyIterationLabtest2 to check the updates of Stationary POLICY (The policy stores the action the agent will take in each state, stationary means that it does not vary with time)
   C. **[TURN THIS IN, 10 points]** Compare and contrast the final policy and initial random policy for both MDPs, and provide a written interpretation
3. *Task 2: Measuring Policy Iteration:*
   A. Investigate the measureAgentPerformance() function in this file to understand how to evaluate a policy
   B. Run PolicyIterationLabtest3 to see how well the PolicyIterationAgent performs on a simple MDP
   C. **[TURN THIS IN, 15 points]** Interpret the chart: How does the initial random policy perform relative to the final policy
4. *Task 3: Test Policy Iteration on different MDPs*
   A. Run test4 to see how policy iteration performs on a simple parking MDP
   B. Run test5 to see how policy iteration performs on a hard parking MDP
   C. **[TURN THIS IN, 15 points]** Interpret the two charts. How big is the performance gap between the two MDPs? Why do you think it exists?
5. *Task 4: Compare 3x Parking Agents: Probabilistic Policies, QLearningAgent and PolicyIterationAgent:*
   A. Run test6 to see results for each agent on the parking MDP
   B. **[TURN THIS IN, 15 points]** Provide a written interpretation of what you see in the chart.
6. Task 5: Compare three agent types on a RANDOM MDP
   A. Run test7 to see results for each agent on the parking MDP
   B. **[TURN THIS IN, 6 points]** Provide a written interpretation of what you see in the chart.
   C. **[TURN THIS IN, 4 points]** While showing your work, compute the size of the Q-value table required for this MDP.
   D. **[TURN THIS IN, 4 points]** Next, compute the number of updates this training loop will perform on the Q-value table.

E. **[TURN THIS IN, 6 points]** Explain why you think this isn't enough training.
F. **[TURN THIS IN, 4 points]** Describe how you would like to adjust the parameters WILL be enough and rerun the test.
G. **[TURN THIS IN, 5 points]** Create a chart illustrating how your adjustment impacted the results.
H. **[TURN THIS IN, 6 points]** Explain what you see in your chart, with a focus on whether or not your adjustments were sufficient.
I. Repeat steps G+H+I until you are satisfied that values have converged and achieve a good performance.

7. **(TURN THIS IN, 5 points)** Upon completing the lab, determine how long you actually spent on the lab, and report that timeframe in addition to your estimate beforehand.

Submit

A file that is readable (pdf, docx, etc) containing your charts, explanations, and mathematics.