Reinforcement Learning

I did pt 1 and pt 2 in separate files diverging from some point along pt1 because I was using Google Colab and had limited resources. So the “important” code is towards the end of both notebooks. Sorry.

Pt. 1.

**Learning Rate (alpha):**

* **Plots:** When alpha = 0.1, the cumulative reward plot shows slower growth and might not reach its full potential. When alpha = 0.9, the reward plot may increase rapidly at first but could become unstable. alpha = 0.5 offers a balance between these, demonstrating consistent improvement. The average reward plots also support these observations, showing differences in learning speed and stability. The average steps plots can indicate how efficiently the agent learns to reach the goal.

**Discount Factor (gamma):**

* **Plots:** Higher gamma values (say .99) generally lead to higher cumulative rewards in the FrozenLake scenario, as the agent prioritizes long-term goals. The "Learned Q-values" heatmaps reveal that with higher gamma, Q-values are higher for states closer to the goal, showing the agent values these more.

**Exploration Rate (epsilon):**

* **Plots:** The "States" and "Actions" distribution plots illustrate that higher epsilon values like .5 result in more exploration of the environment. The average reward and steps plots are key indicators of performance here. If epsilon is decreased over time, the average reward tends to improve, and average steps decrease, reflecting better performance.