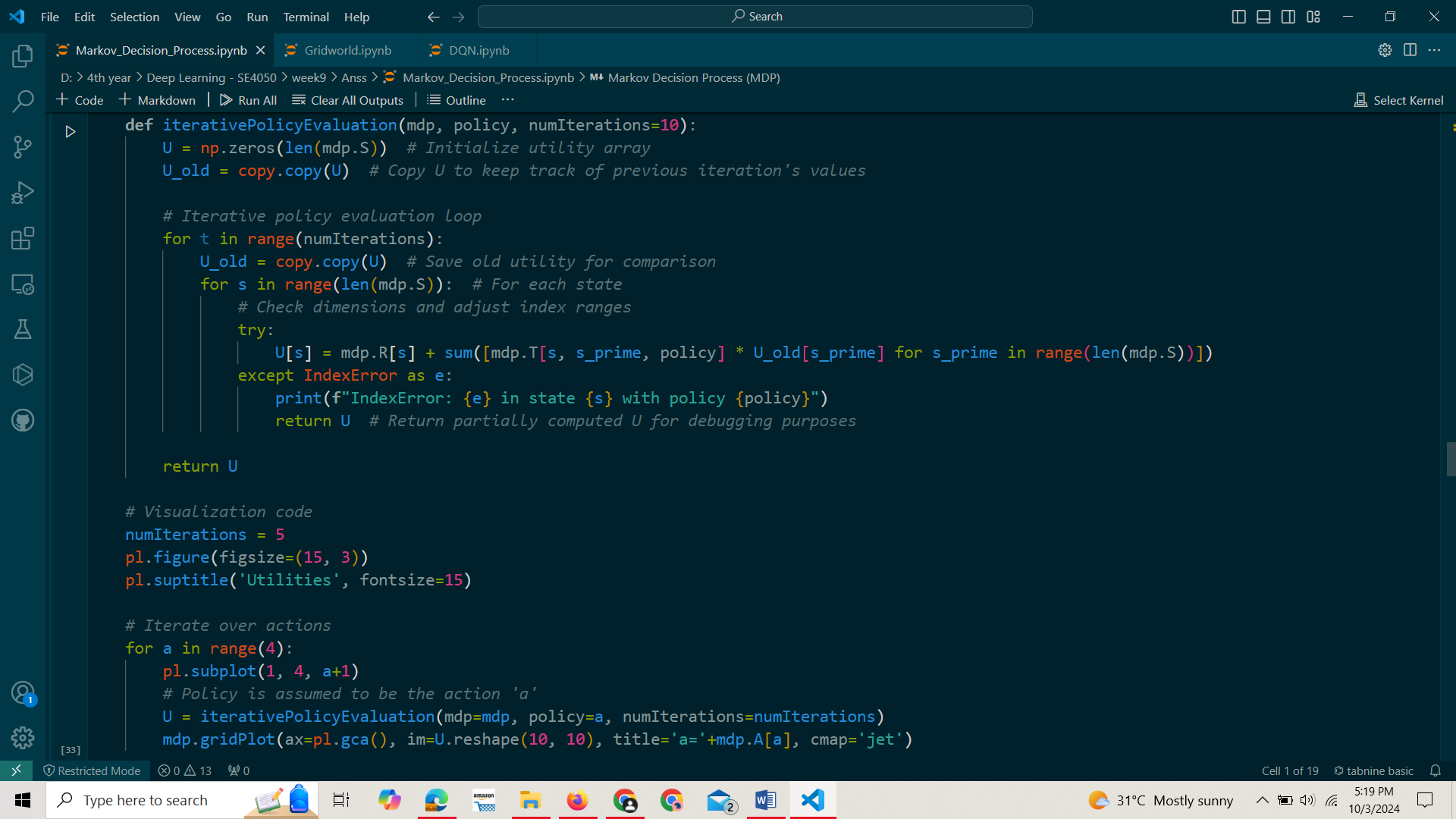
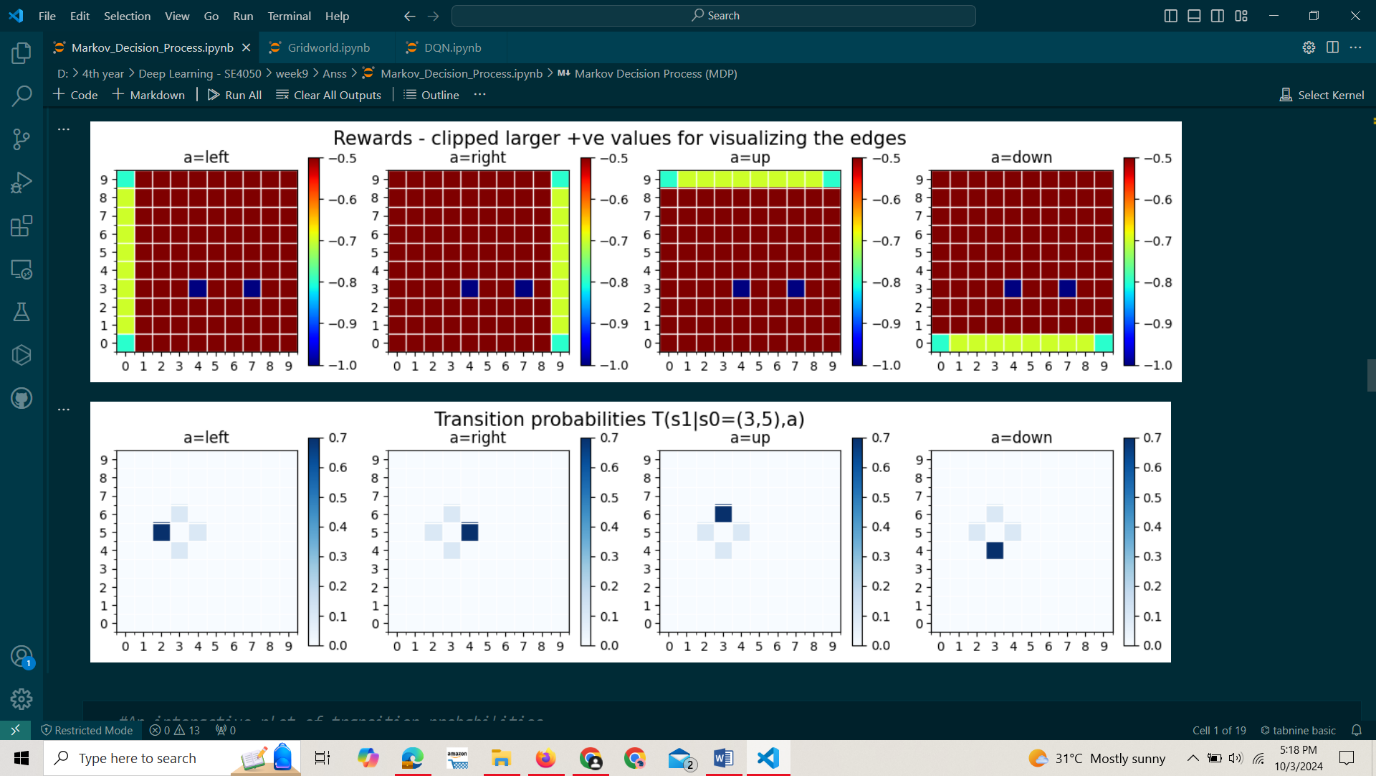
**Reinforcement Learning**

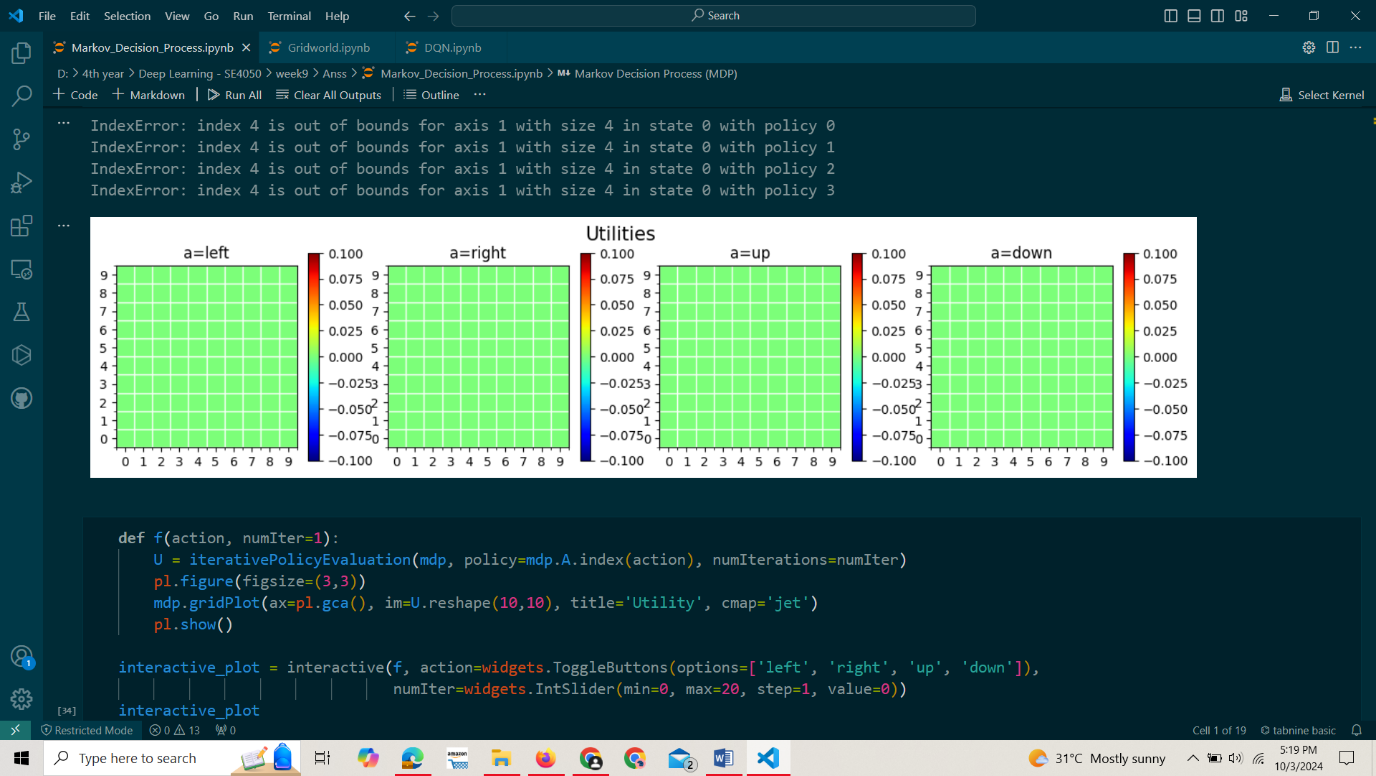
**Question 1: Markov Decision Process and Q-Learning**

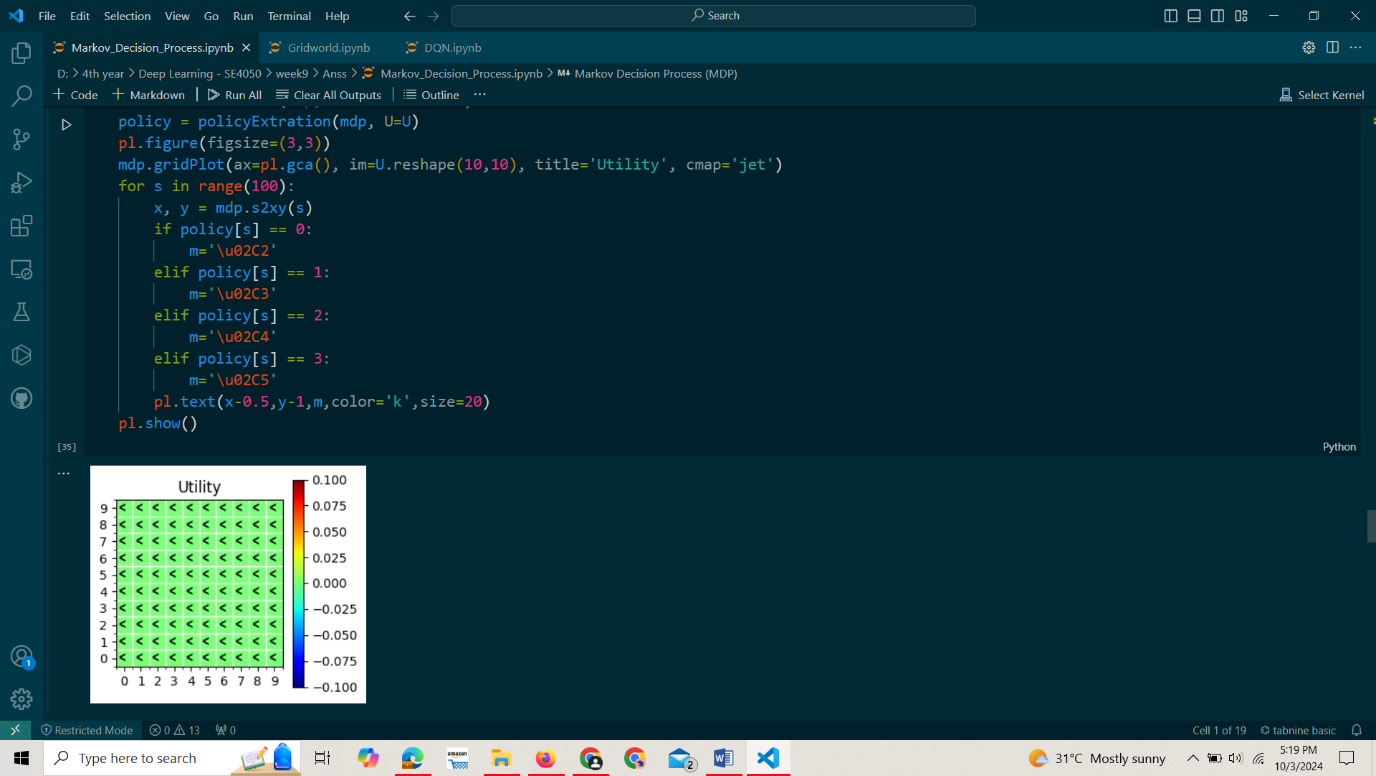
**3.**

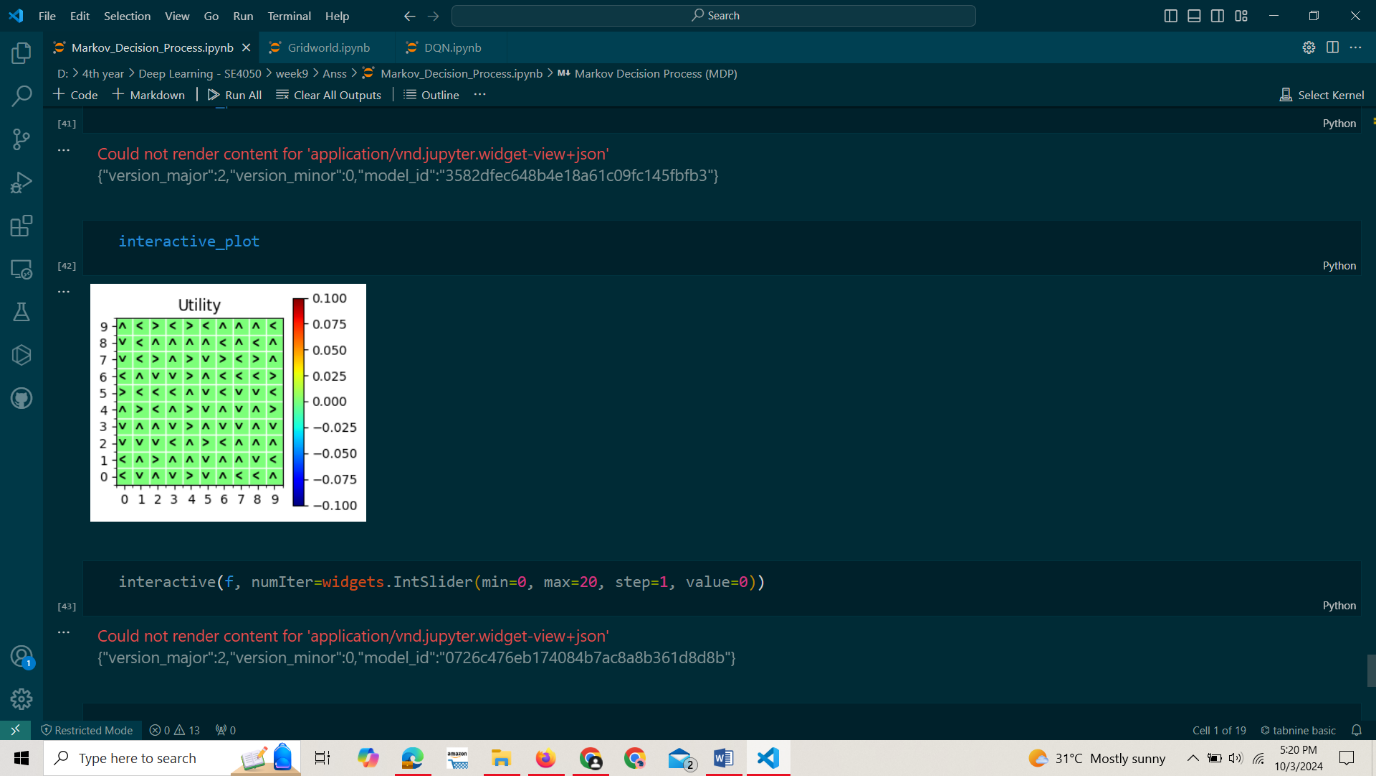
**Markov\_Decision\_Process.ipynb screen shots**

****

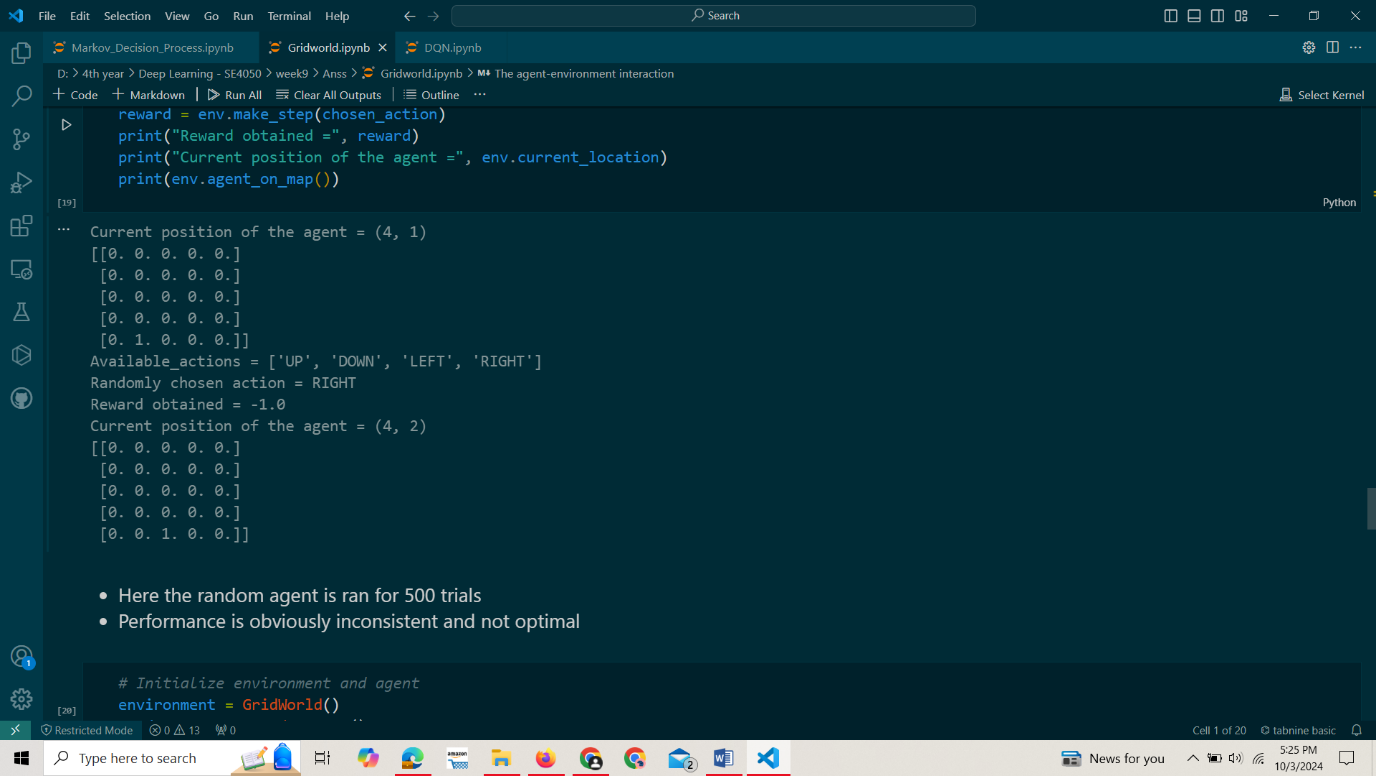
****

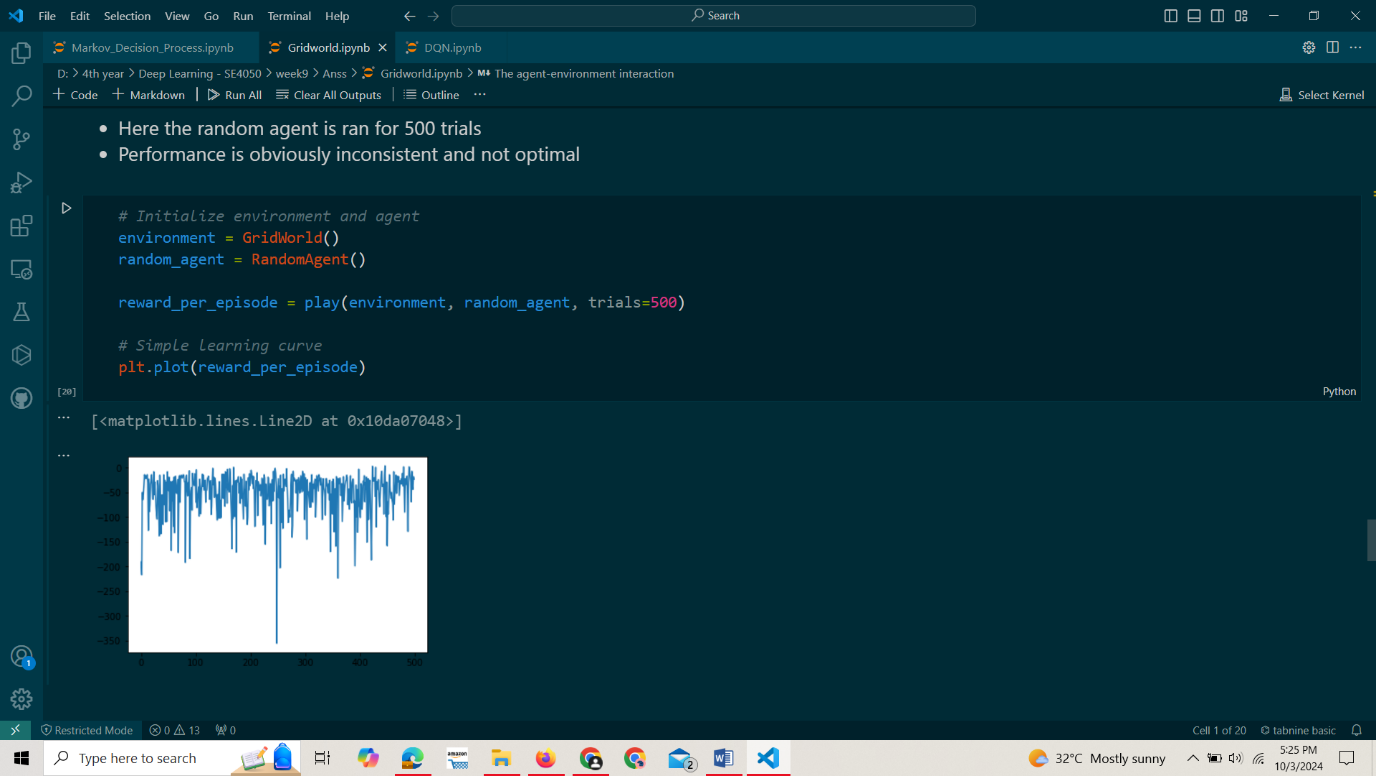
****

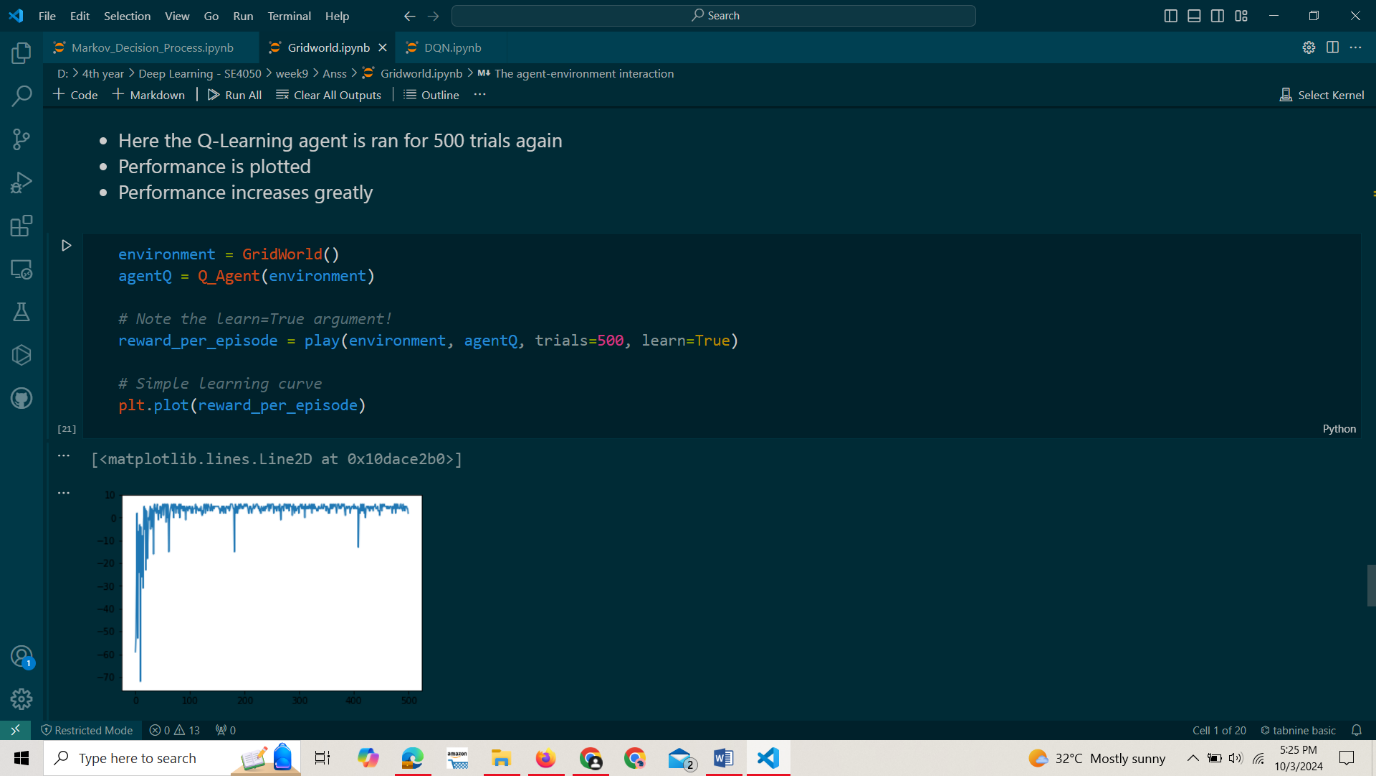
****

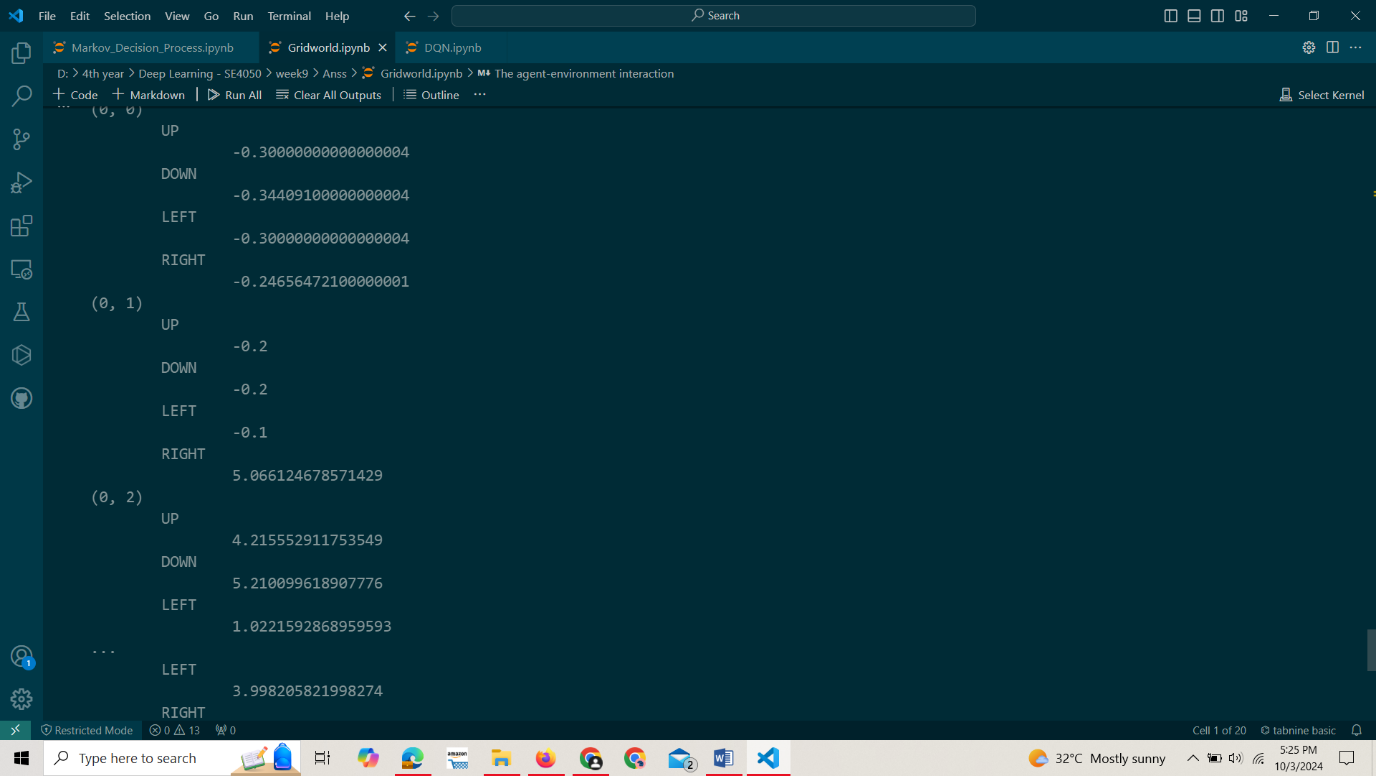
****

**Gridworld.ipynb**

****

****

****

****

**Question 2: Model-Based vs Model-Free Reinforcement Learning**

**2.**

**Model-Based Algorithms**

**Learns a model of the environment; uses planning to optimize decisions.**

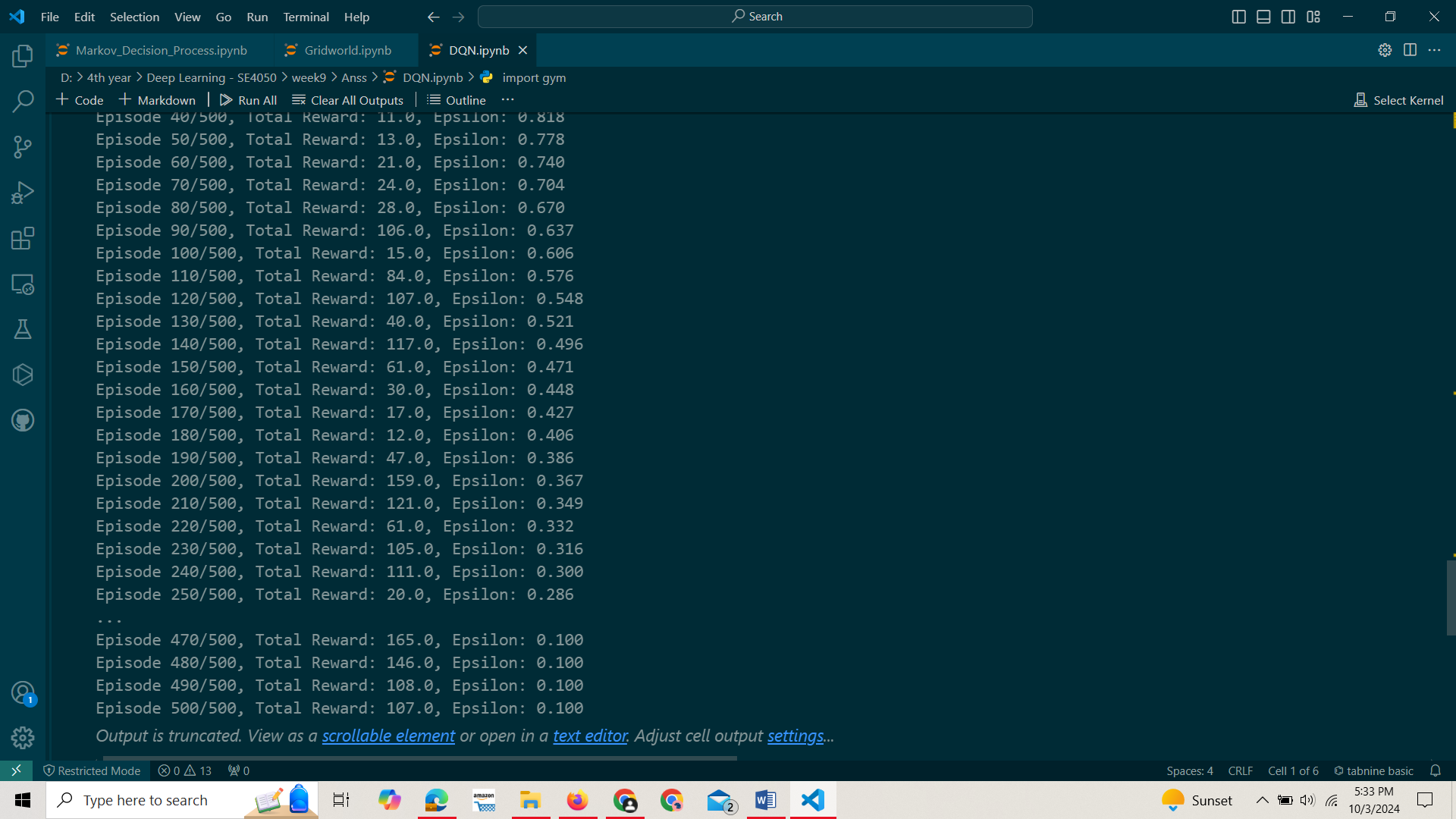
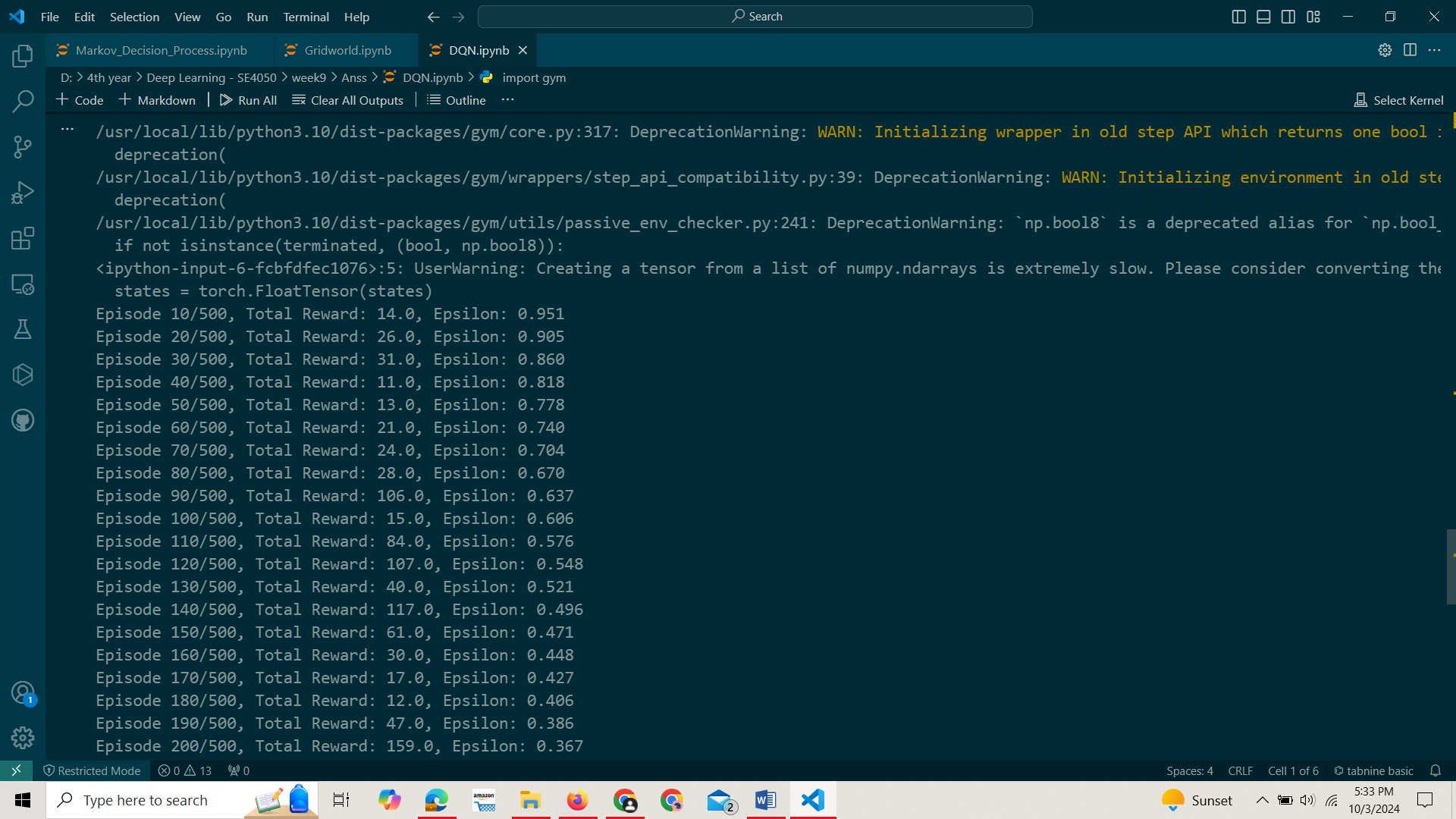
* **Definition**: These algorithms explicitly build a model of the environment. This model captures the dynamics of the environment, including state transitions and rewards.
* **How They Work**: They use the model to simulate future states and evaluate the consequences of actions. This allows the agent to plan its actions based on predictions about the outcomes.
* **Examples**: Value Iteration, Policy Iteration, Dyna-Q.
* **Advantages**:
  + Can be more sample-efficient since they leverage the model to simulate experiences.
  + Suitable for environments where the dynamics are known or can be learned effectively.
* **Disadvantages**:
  + Building an accurate model can be computationally expensive and complex, especially in large or continuous spaces.

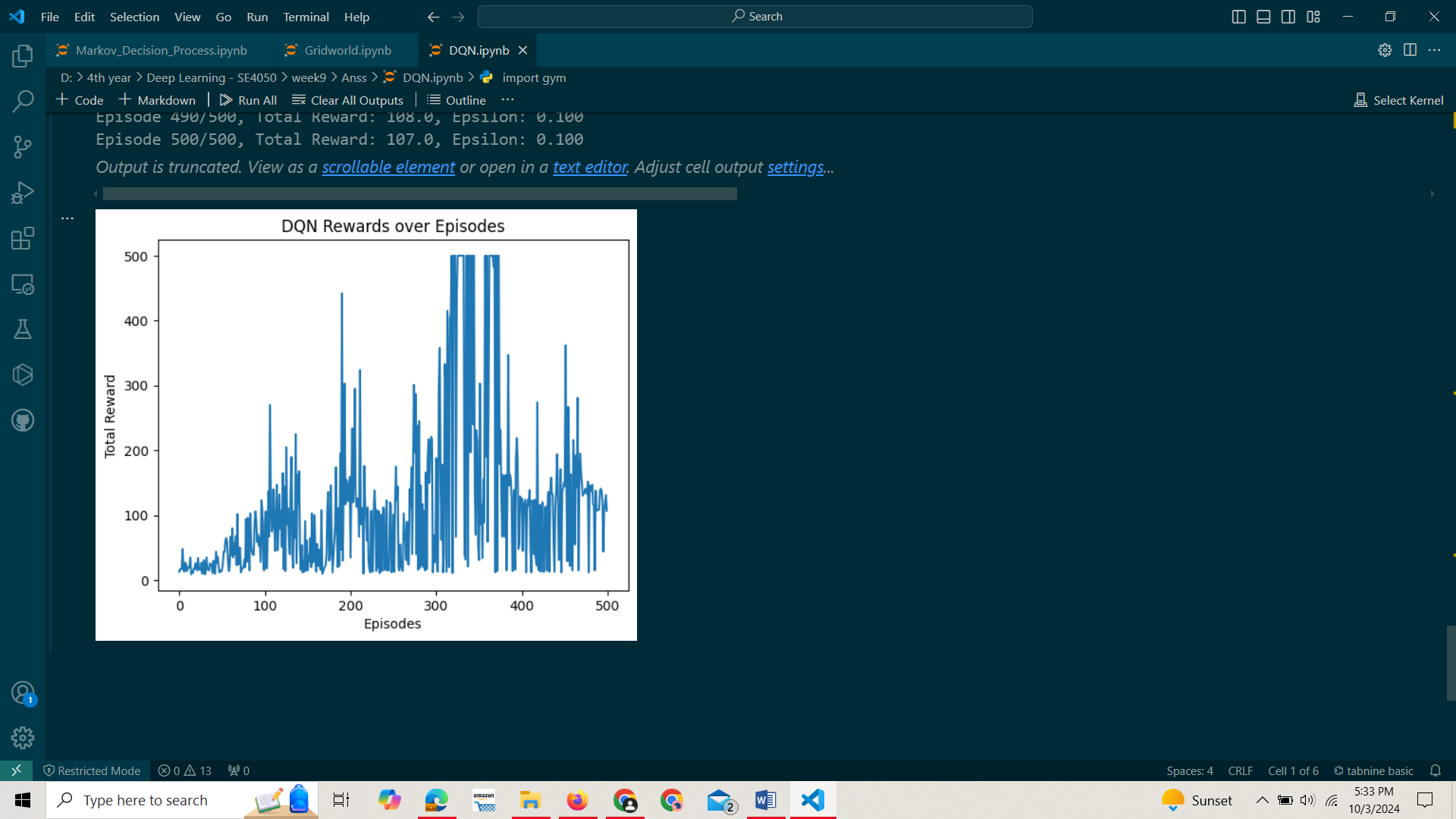
**Model-Free Algorithms**

**Learns policies directly from experience; relies on trial and error without a model.**

* **Definition**: These algorithms do not build or utilize a model of the environment. Instead, they learn directly from the interactions with the environment.
* **How They Work**: They rely on exploration and exploitation to update the value of states or actions based on received rewards, without simulating future states.
* **Examples**: Q-Learning, SARSA, REINFORCE.
* **Advantages**:
  + Simpler and often easier to implement since they don’t require a model.
  + Effective in environments where the dynamics are unknown or too complex to model.
* **Disadvantages**:
  + Typically less sample-efficient, as they may require many more interactions with the environment to learn optimal policies.

**Question 3: Introduction to Deep Q-Learning (DQN)**

**4.**

****