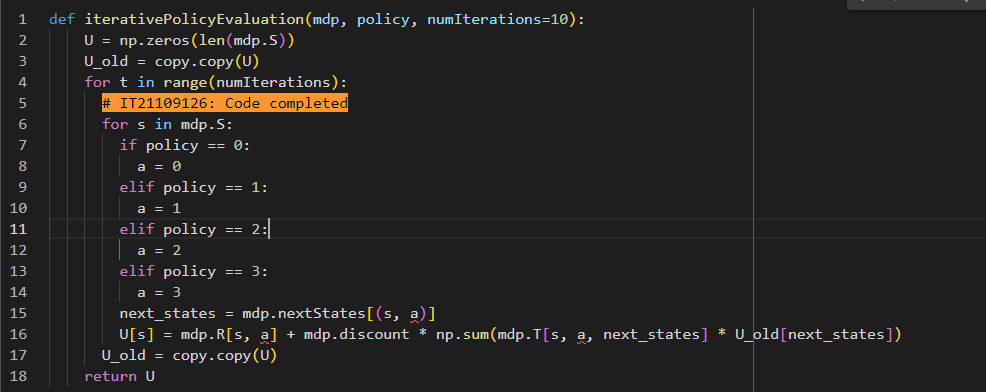
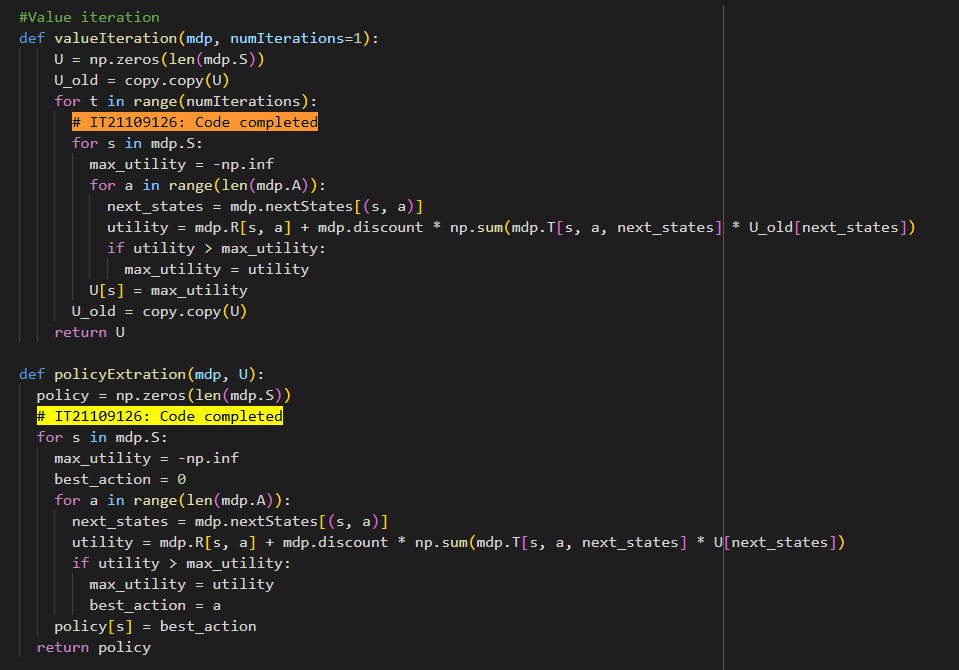
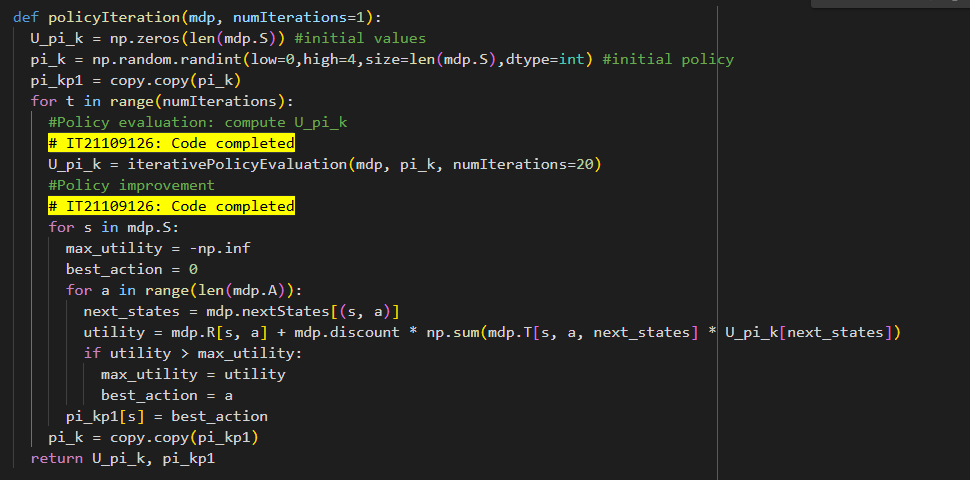
**Question 1: Markov Decision Process and Q-Learning**

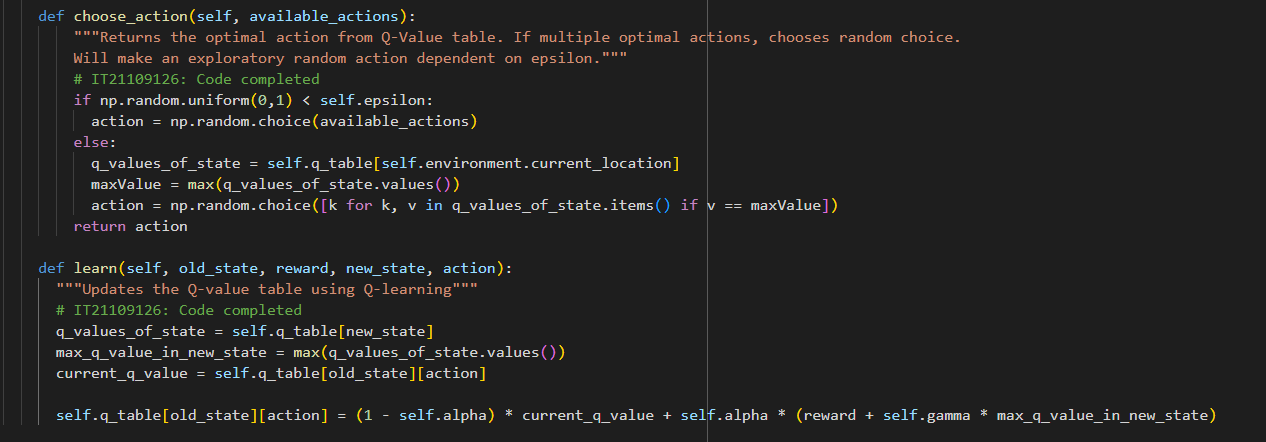
Markov\_Decision\_Process Notebook Screenshots

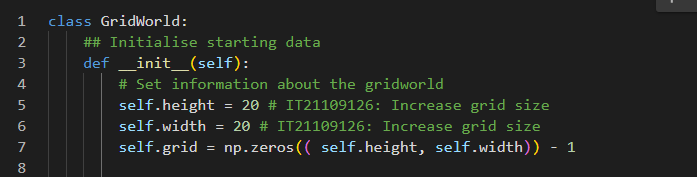


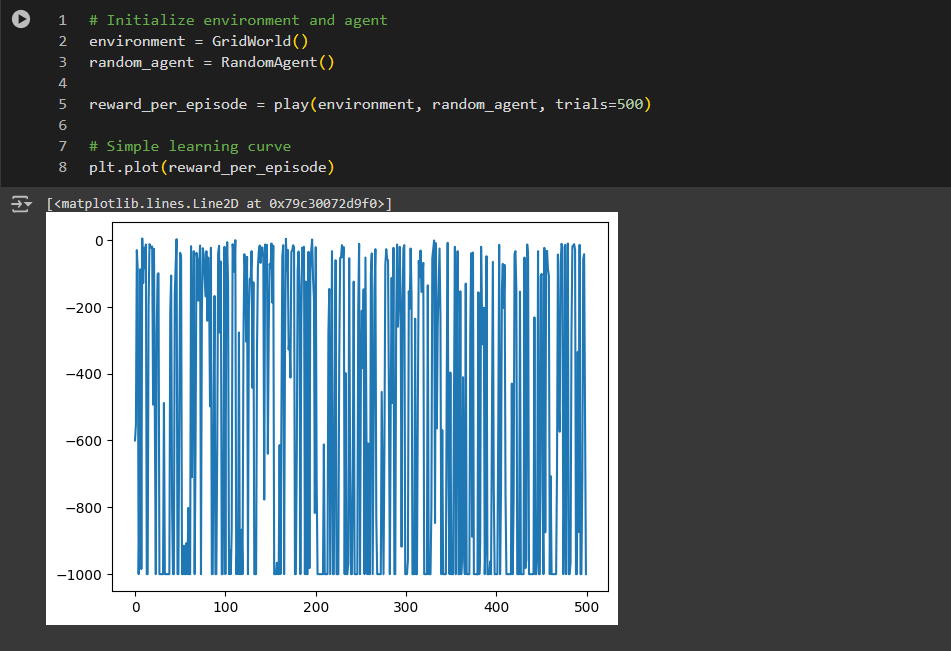




GridWorld Notebook Screenshots







A screenshot of a computer

Description automatically generated

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

Model-Based vs Model-Free

Model-free algorithms like Q-Learning, SARSA, and PPO learn directly from interactions with the environment, focusing on policies or value functions without building a model. While they are simpler to implement, they typically require more real-world interactions to learn effectively.

In contrast, model-based algorithms such as Dyna-Q and Model-Based Value Iteration build a model of the environment’s dynamics. This allows for planning and simulating experiences, making them more sample-efficient but also more complex.

The key differences are summarized below.

**Learning Process**

* *Model-Free:* Learns policies or value functions directly from observed transitions and rewards.
* *Model-Based:* Learns a model of the environment’s dynamics first and then uses this model to plan and simulate future actions.

**Efficiency**

* *Model-Free:* Often requires more real-world interactions.
* *Model-Based:* Can be more sample-efficient due to the ability to simulate interactions.

**Complexity:**

* *Model-Free:* Generally simpler to implement.
* *Model-Based:* Involves additional complexity due to model learning and planning.

Screenshots

A screen shot of a computer program

Description automatically generated

A screenshot of a computer program

Description automatically generated

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

Screenshots

A screen shot of a computer program

Description automatically generated

A screen shot of a computer program

Description automatically generated

A screenshot of a computer screen

Description automatically generated

The plot illustrates a general trend of increasing rewards as the agent learns and explores the environment.