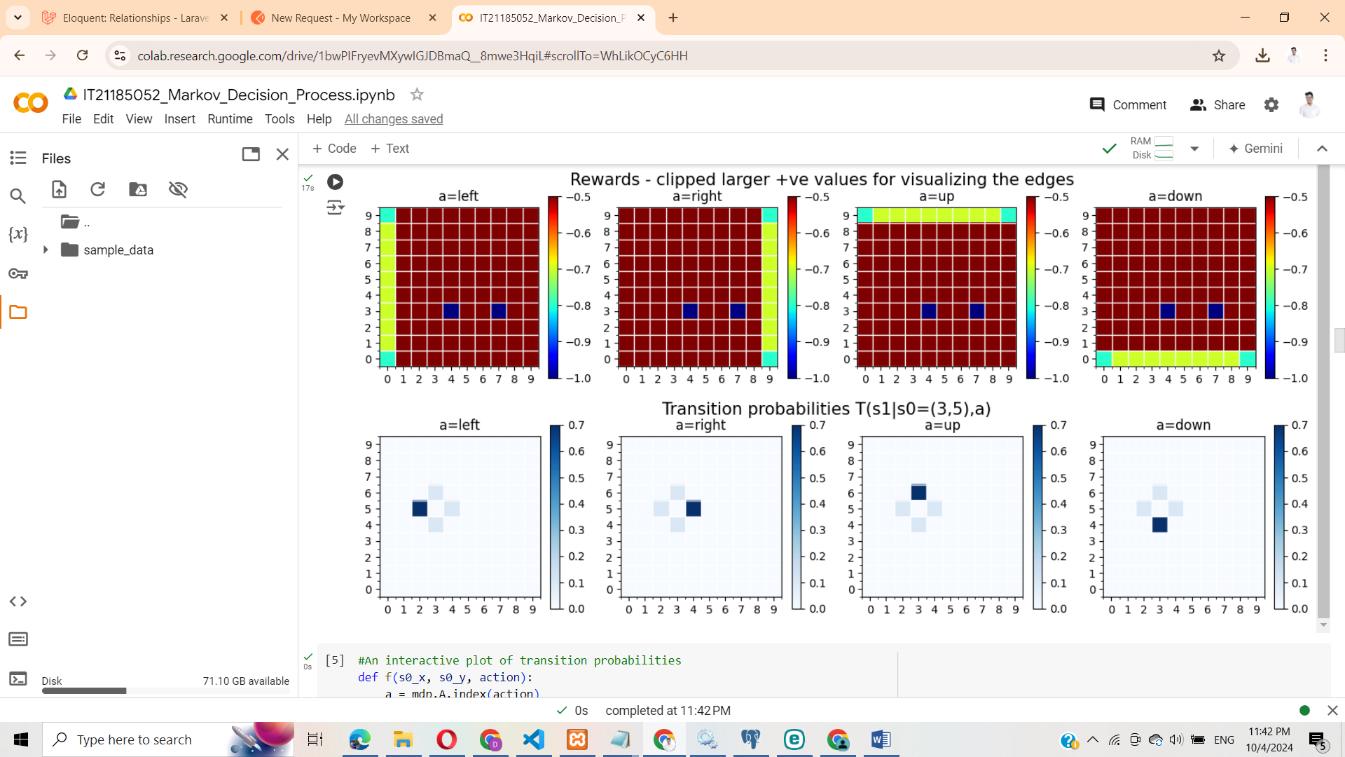
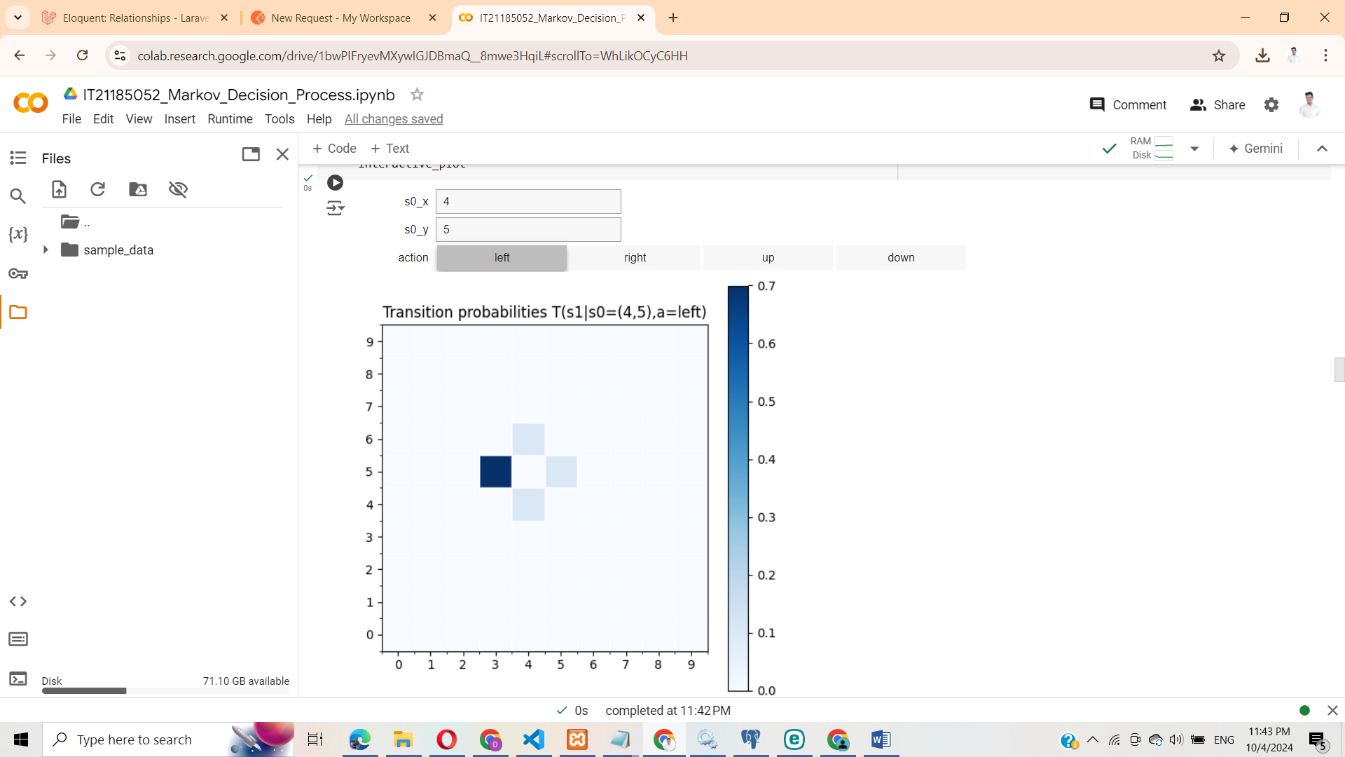
Q1

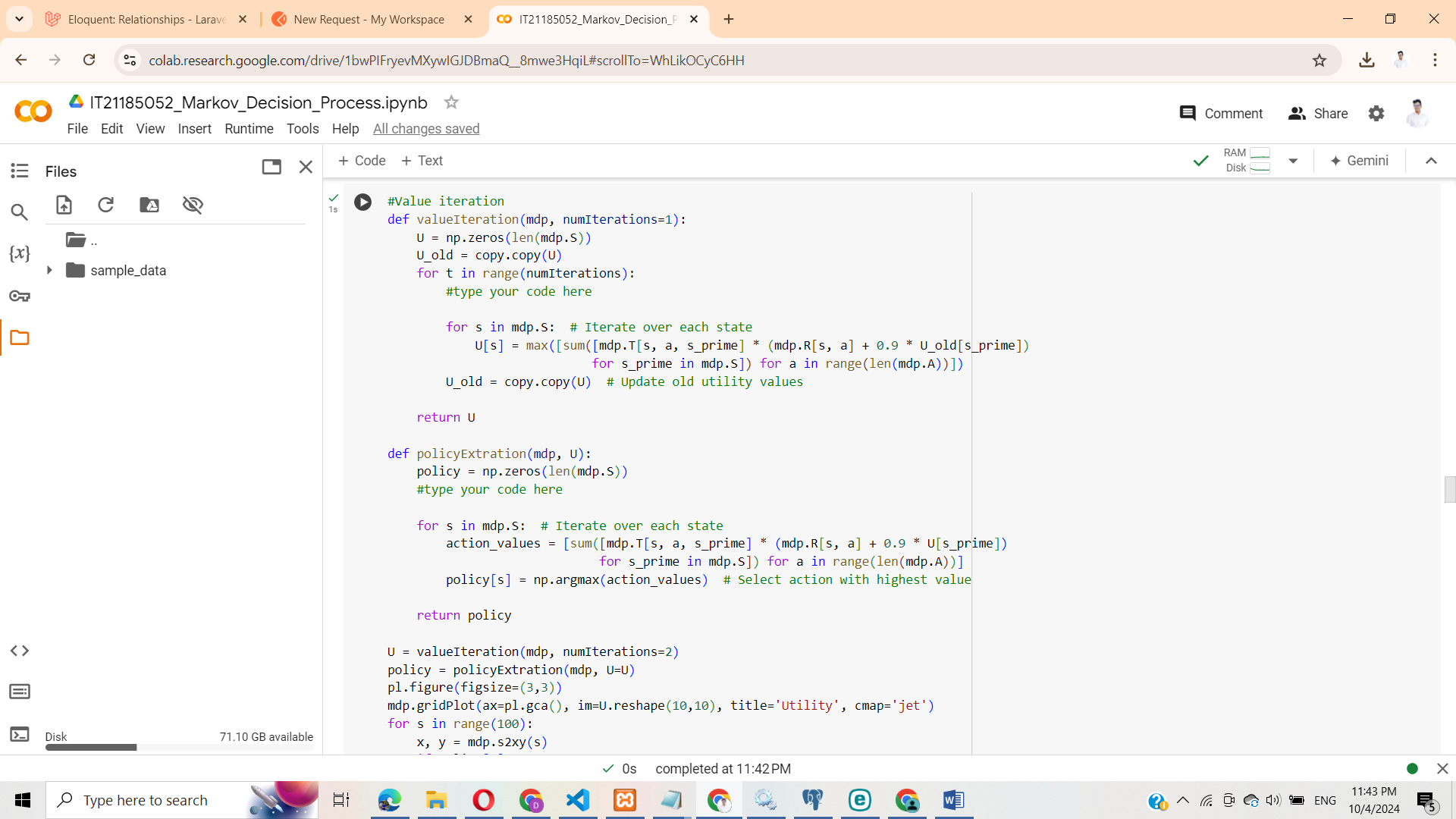
3)

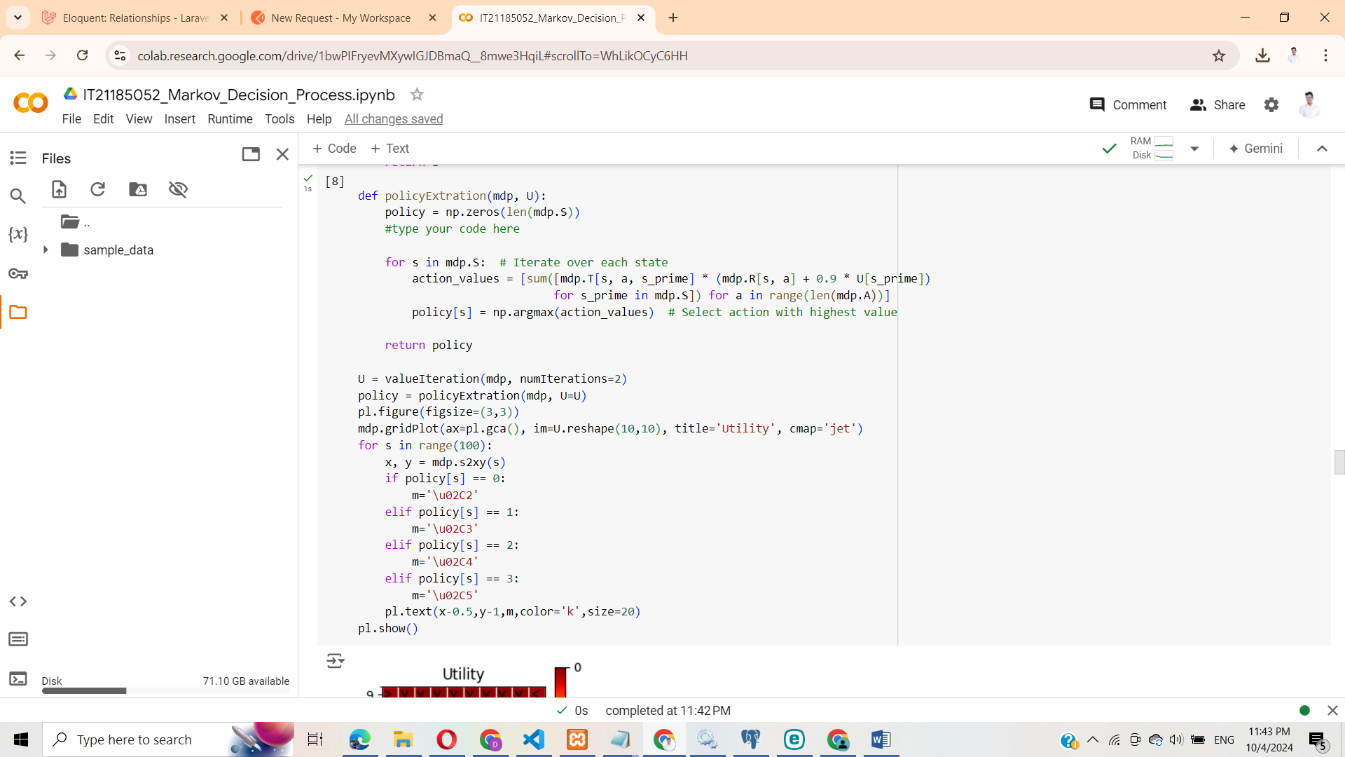


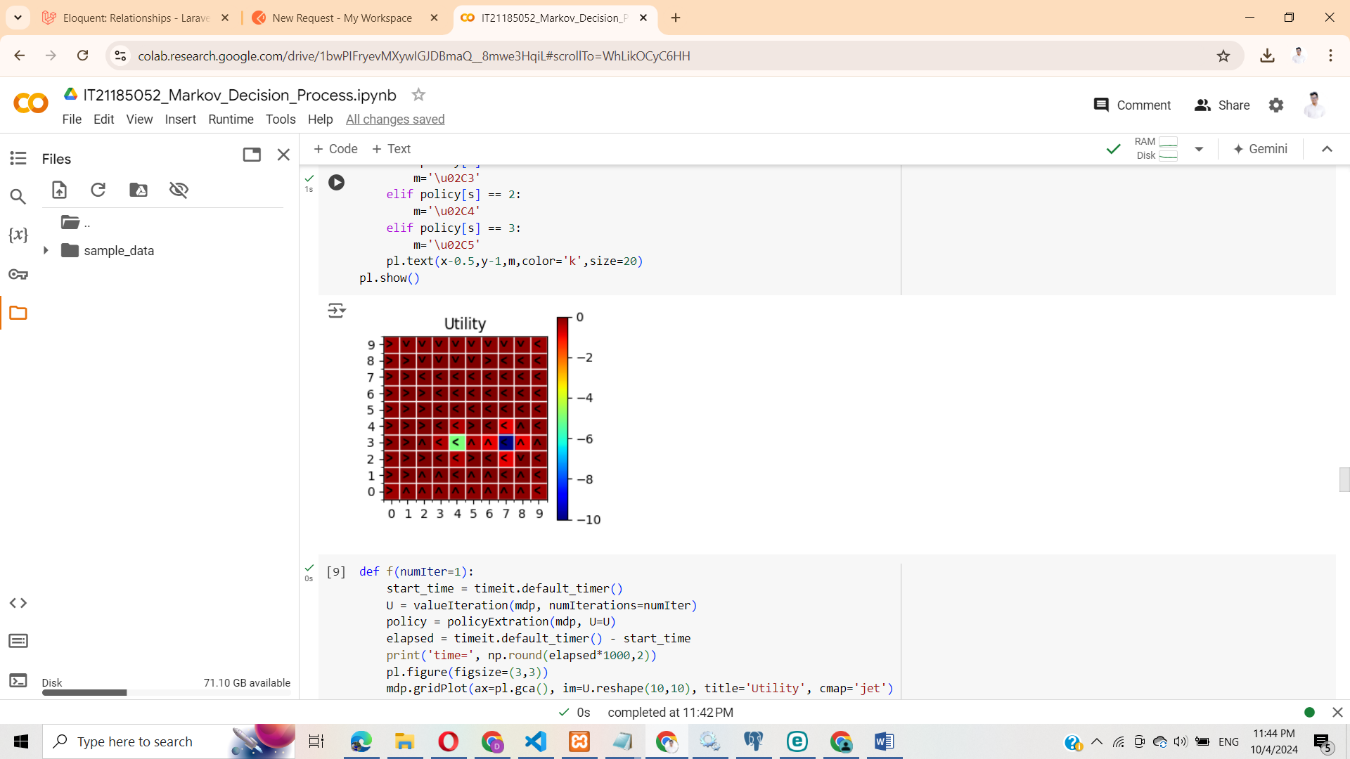


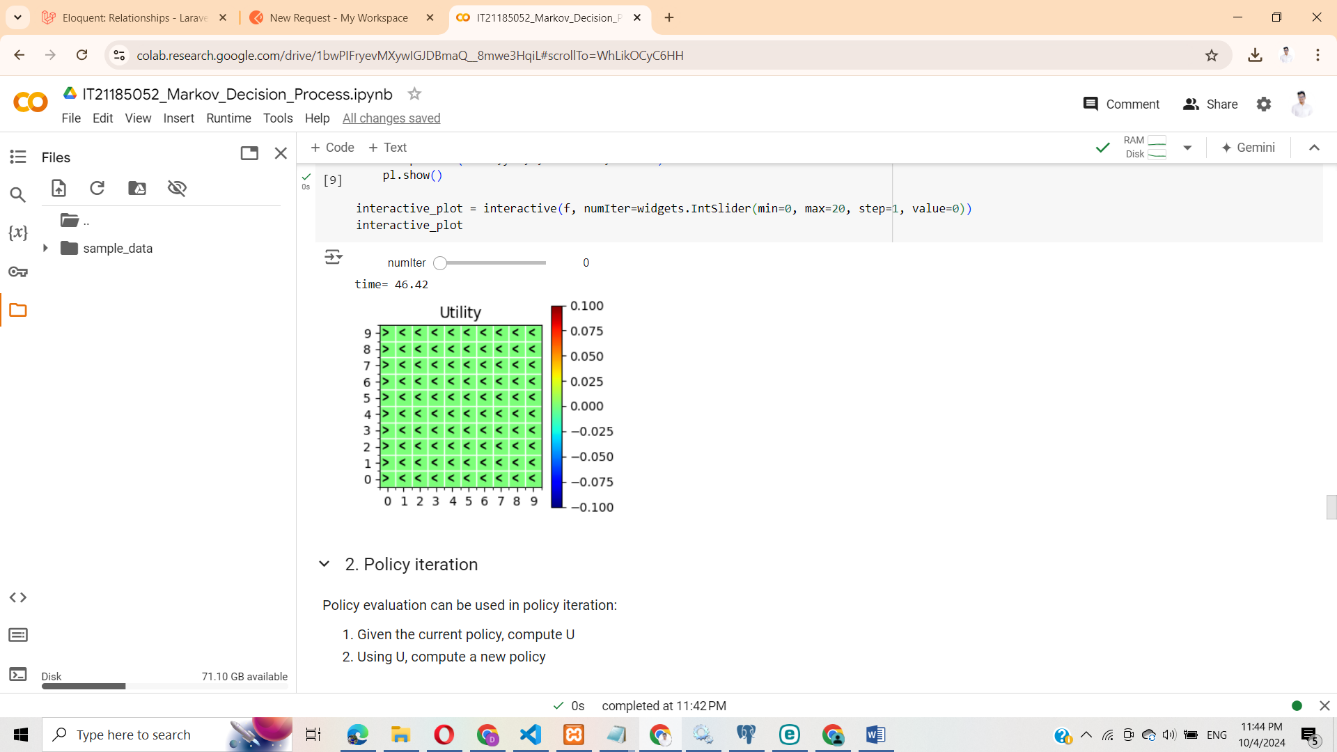


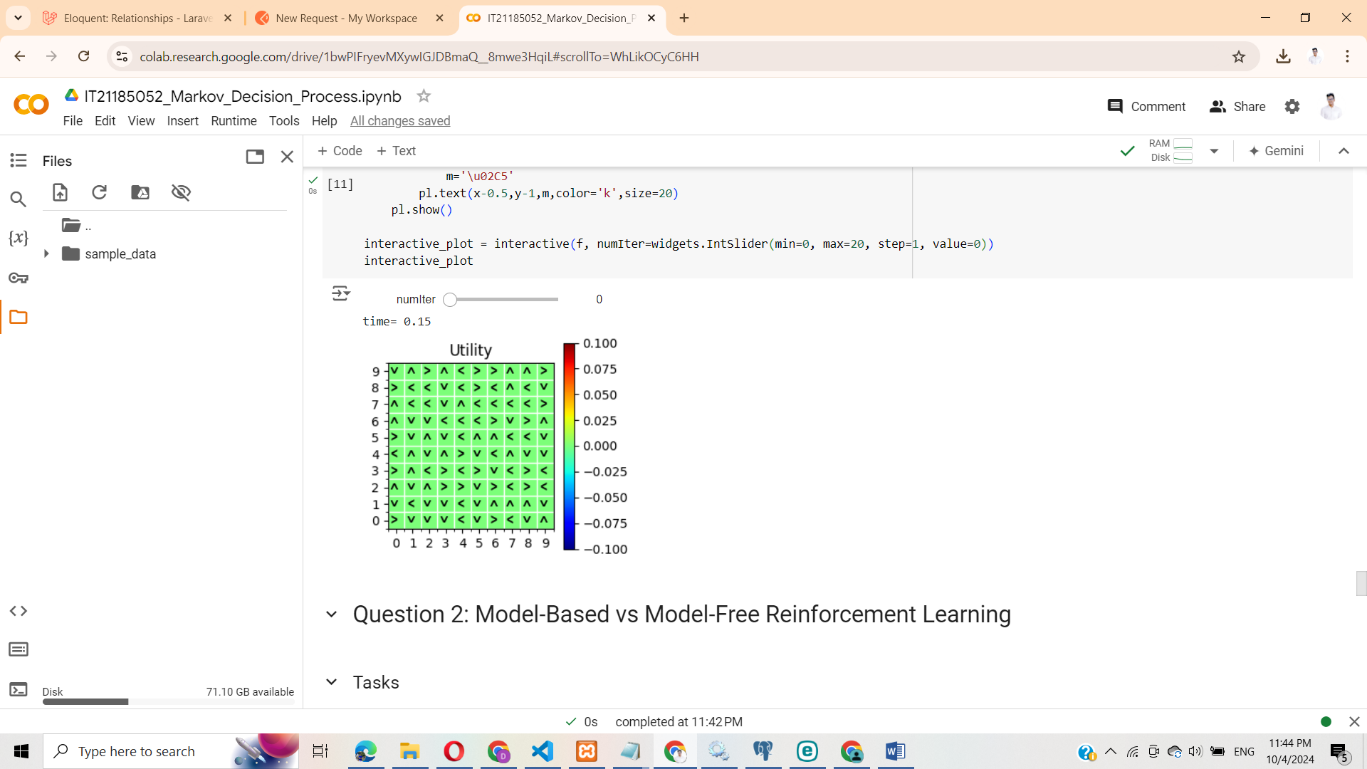


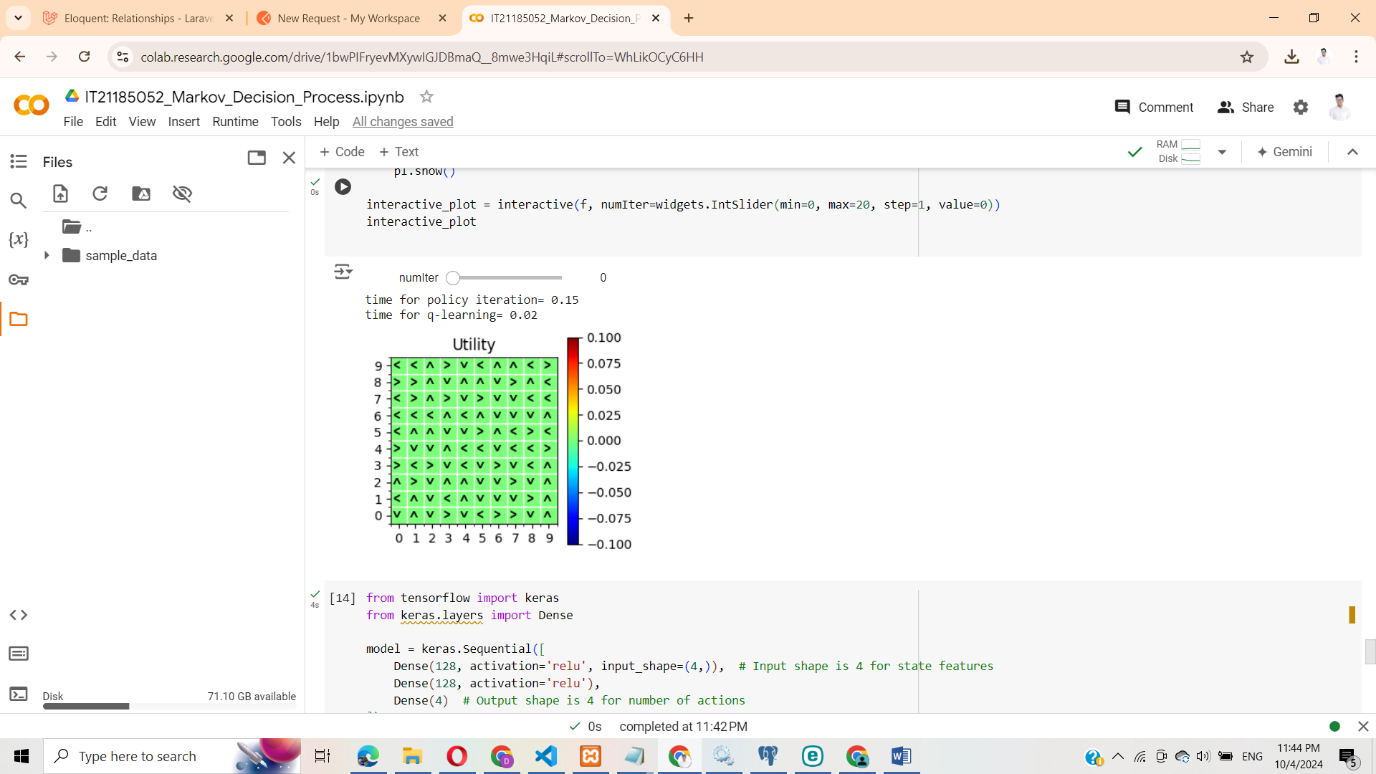




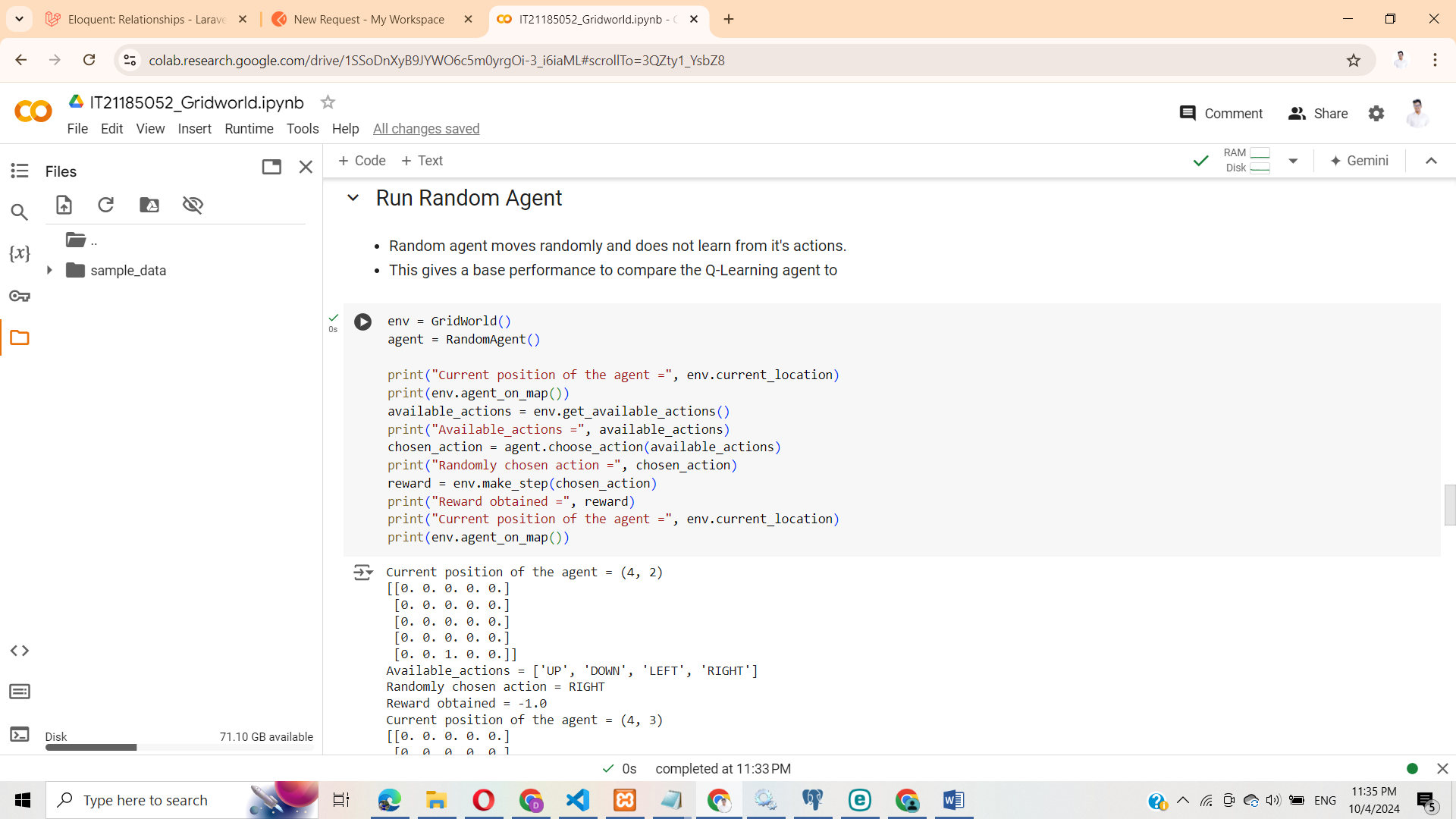


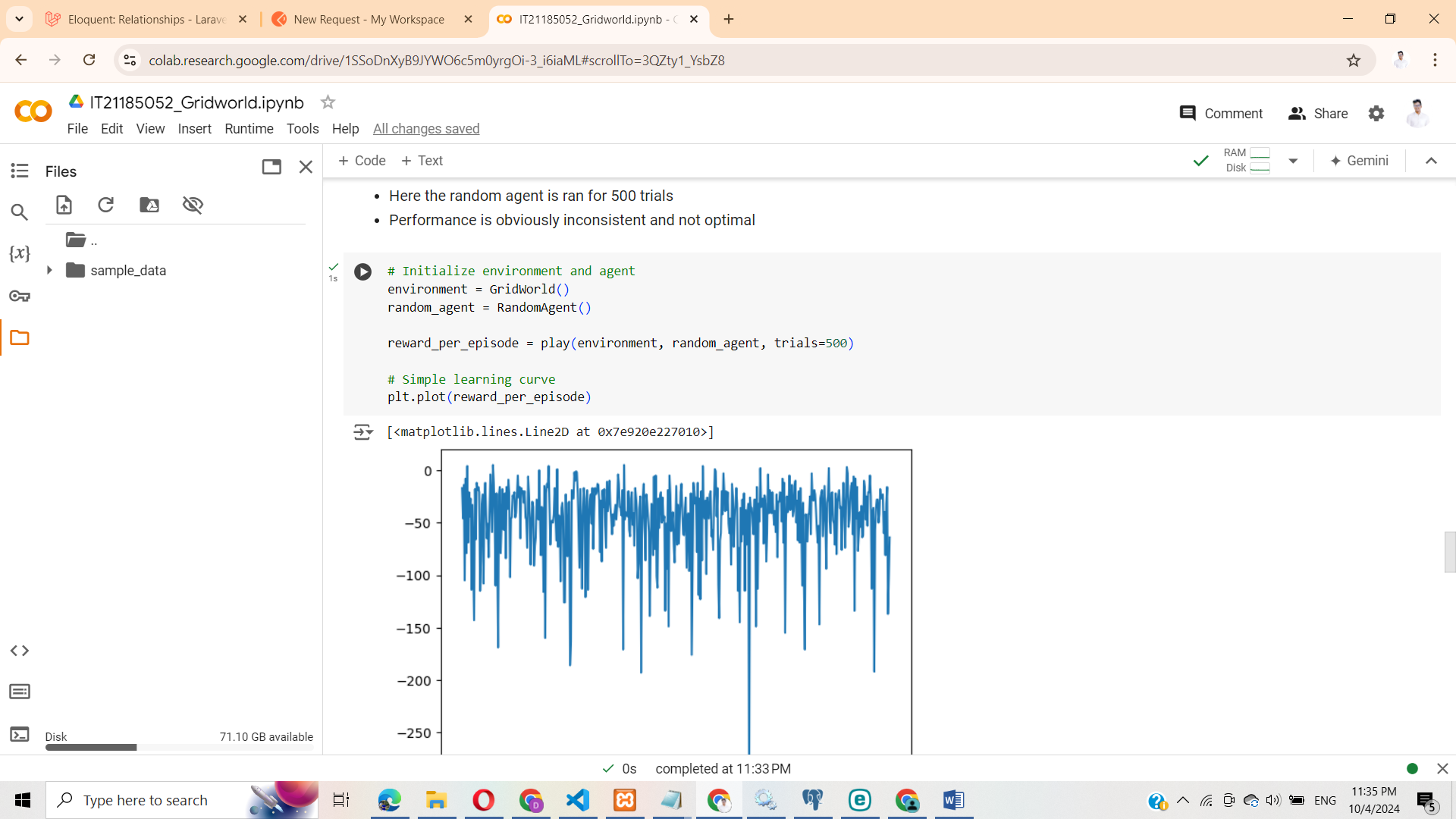




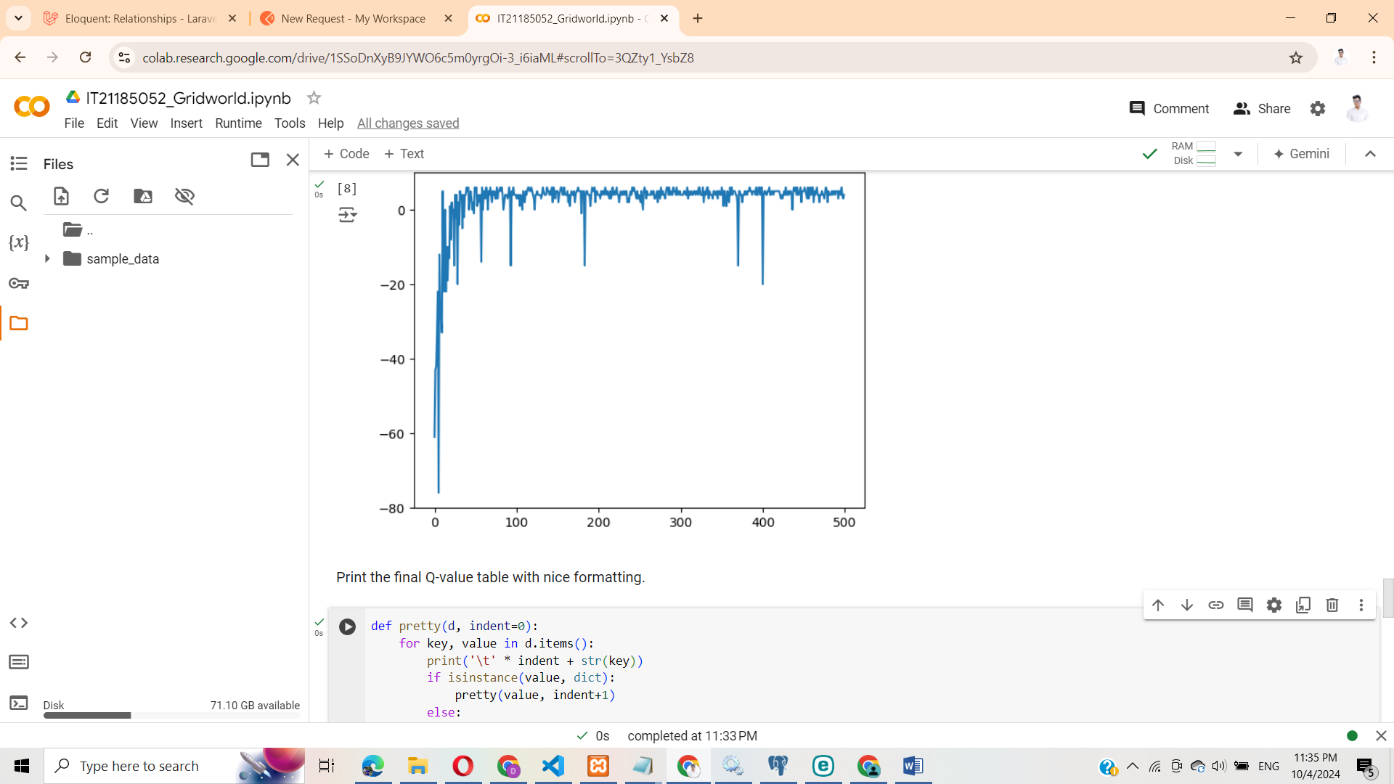


Gridworld.ipynb









Q2

2)

Model-Based Algorithms

These algorithms build a representation of the environment and use it for planning to improve decision-making.

Definition: They create a model that explains how the environment behaves, including how states change and what rewards are given.

How They Work: By using this model, these algorithms can simulate possible future states and evaluate the outcomes of actions. This allows the agent to plan and choose actions based on predicted results.

Examples: Value Iteration, Policy Iteration, Dyna-Q.

Advantages:

They are more efficient in using data since they can simulate experiences using the model.

They work well in environments where the dynamics are known or can be learned effectively.

Disadvantages:

Building an accurate model can be complex and computationally costly, especially in large or continuous environments.

Model-Free Algorithms

These algorithms learn directly from interacting with the environment, without using any model.

Definition: Instead of building a model, model-free algorithms rely on trial and error to learn through experience.

How They Work: They explore the environment, updating the value of actions or states based on the rewards they receive, without simulating future states.

Examples: Q-Learning, SARSA, REINFORCE.

Advantages:

These algorithms are simpler and easier to implement since they don’t need a model.

They are useful in environments where the dynamics are unknown or too complex to model.

Disadvantages:

They are less efficient with data, often needing many more interactions to learn the best strategies.

Q3

