Implementation of Q-Learning for Dots and Boxes

# Implementation of Q-Learning Technique – 2 x 2 grid

Learning rate : 0.6

Discount factor: 0.7

Epsilon : 0.6 – for epsilon greedy policy

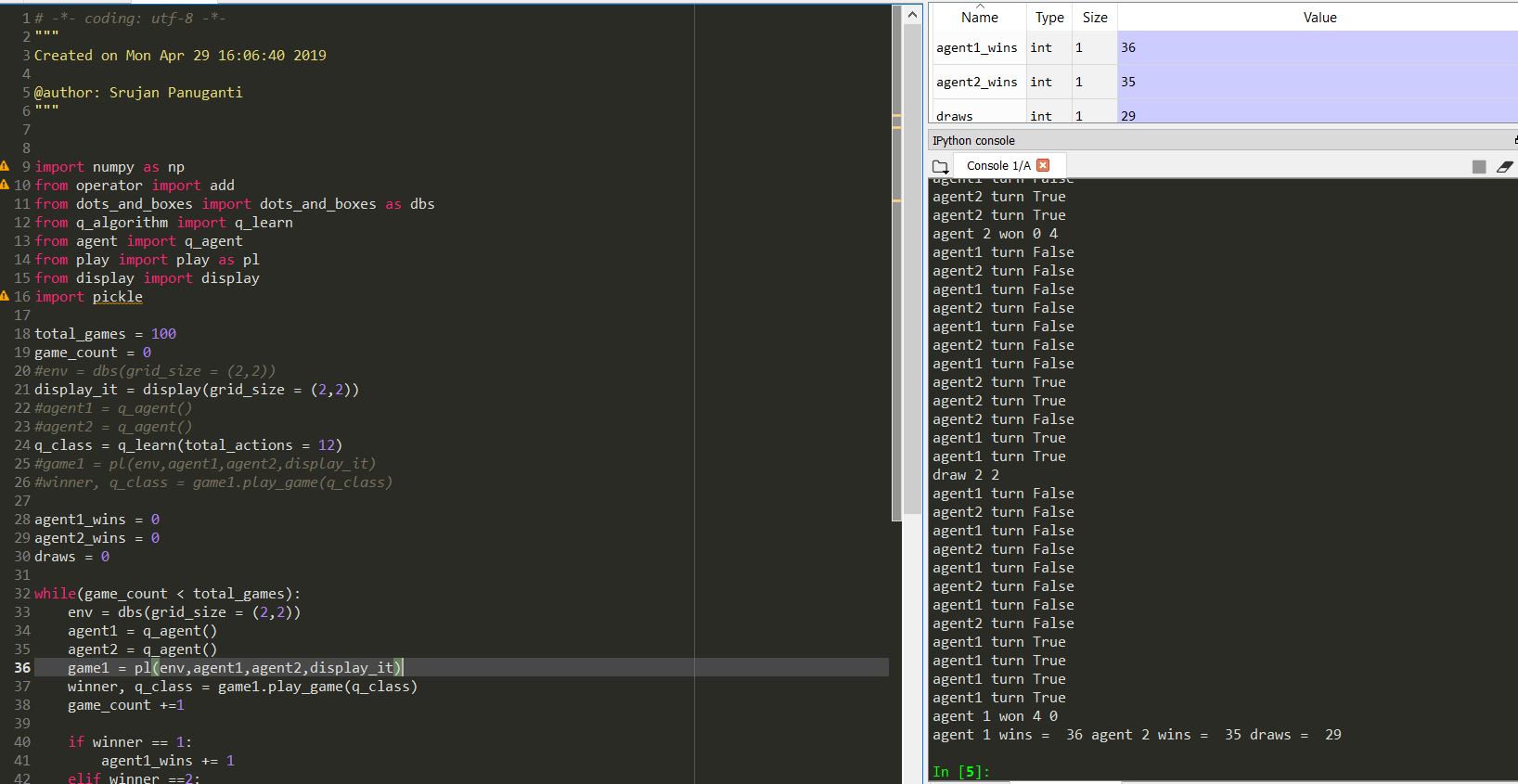
**For 100 games:**

Time consumed : 1 second

Agent 1 wins : 36

Agent 2 wins : 35

Draws : 29



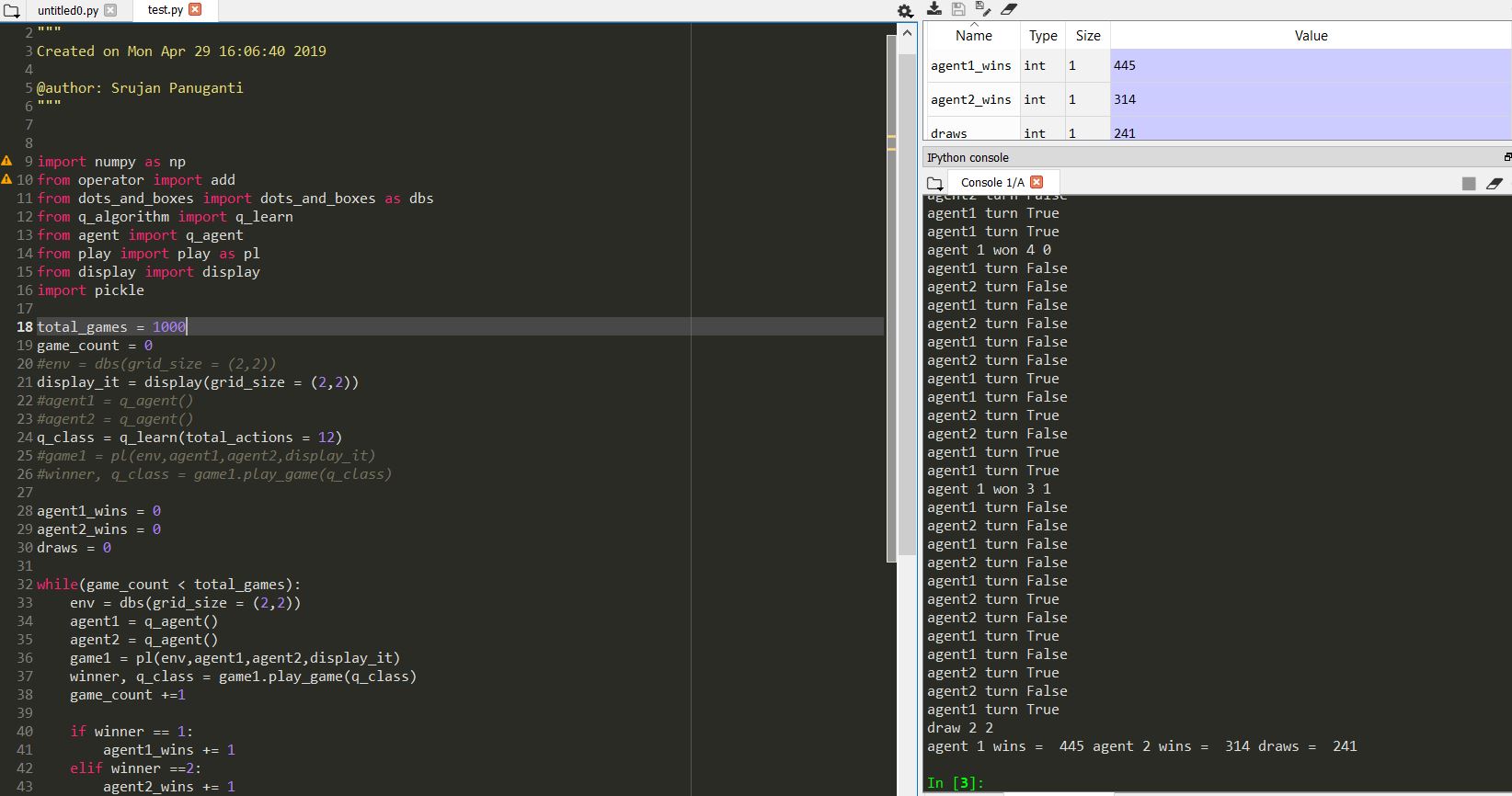
**For 1000 games:**

Time consumed: 13 seconds

Agent 1 wins :445

Agent 2 wins: 314

Draws : 241



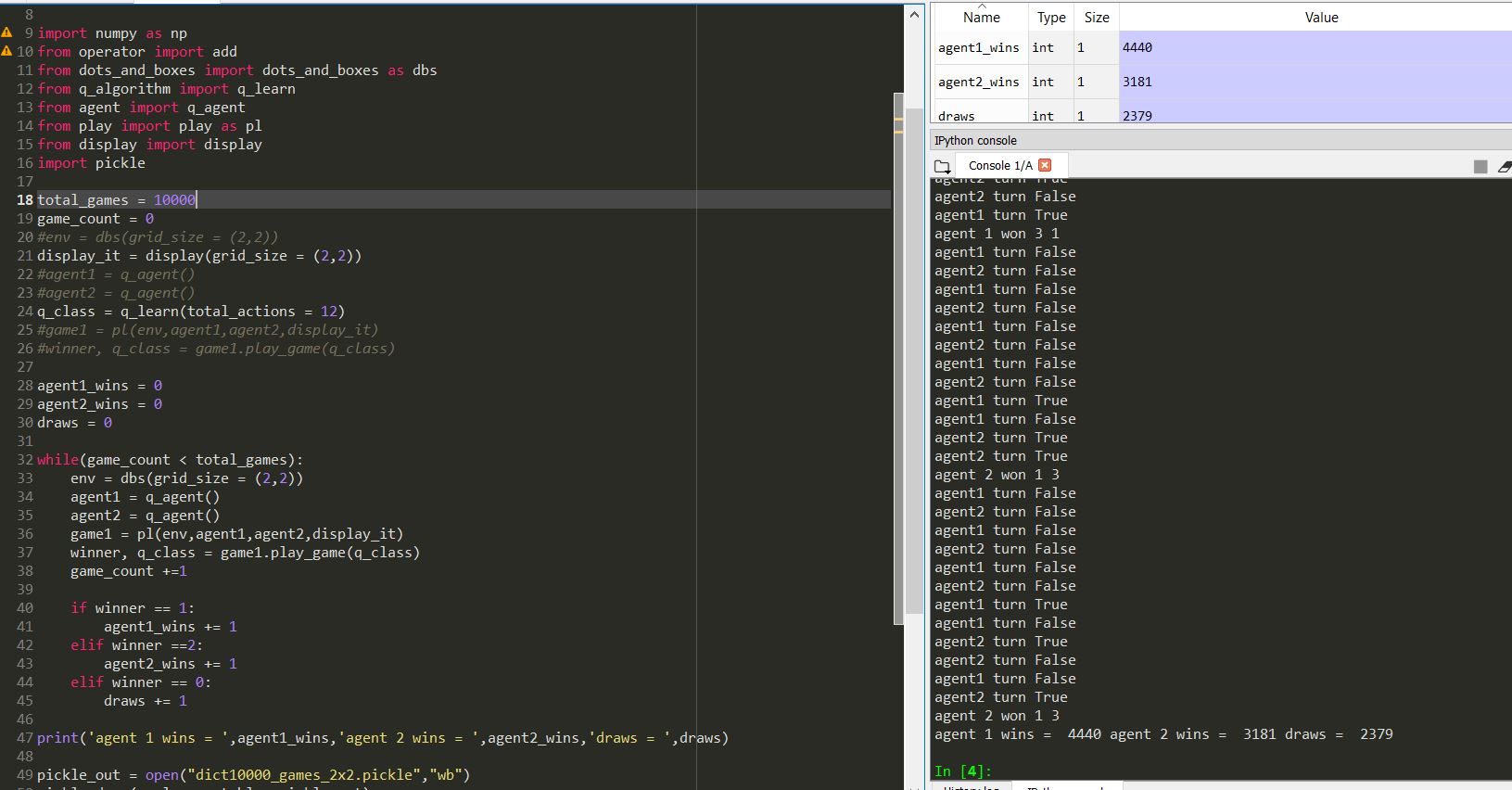
**For 10000 games:**

Time consumed 1 minute 45 seconds

Agent 1 wins : 4440

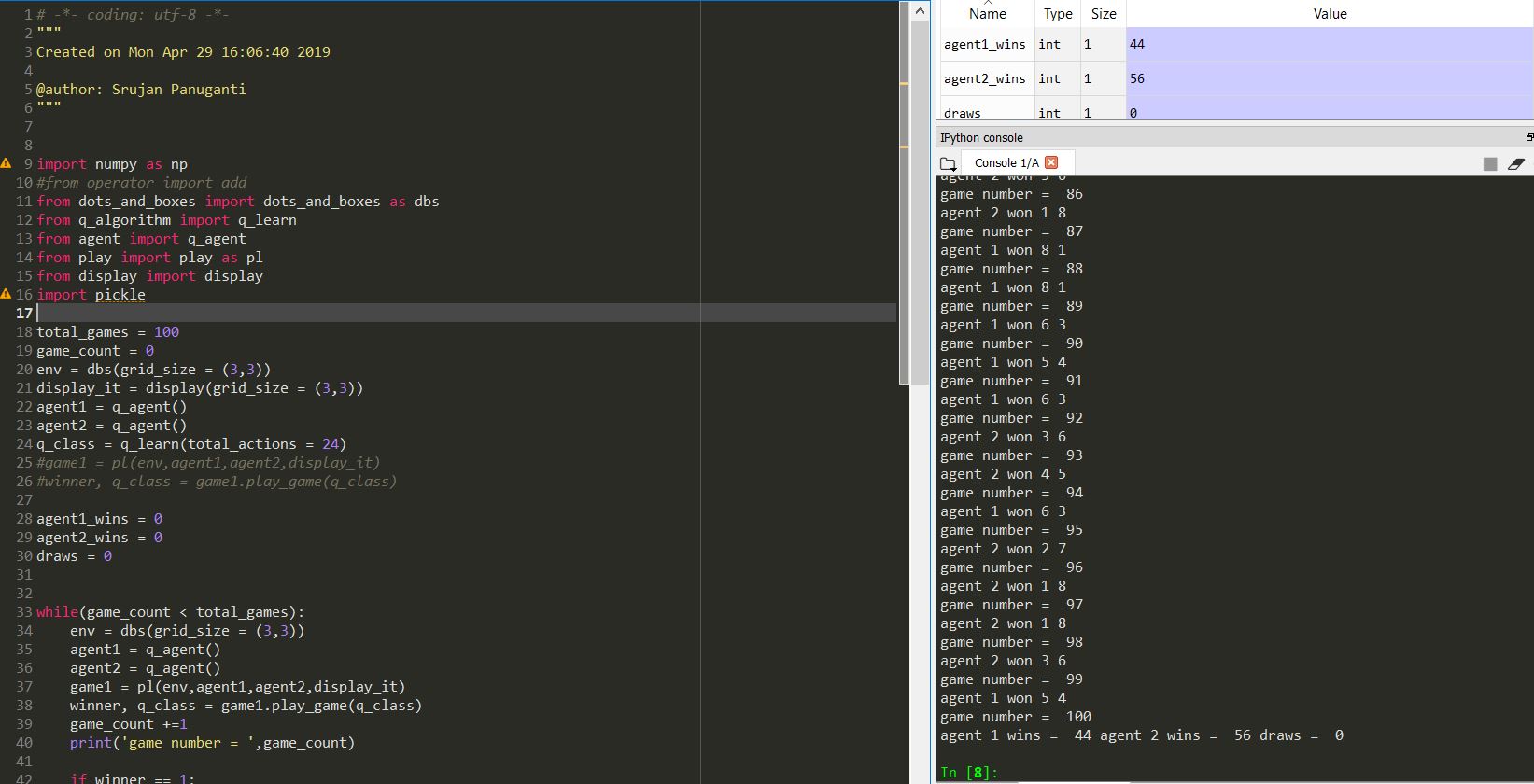
Agent 2 wins : 3181

Draws : 2379



# Implementation of Q-Learning Technique – 3 x 3 grid

Time consumed: 1 minute 30 seconds



For 1000 games:

Time consumed : 14 minutes 10 seconds