



# RESULTS FROM EXPERIMENTS

SKANDA VAIDYANATH

INSTITUTE OF CREATIVE TECHNOLOGIES

6/24/2019

# ENVIRONMENT REWARD FUNCTION

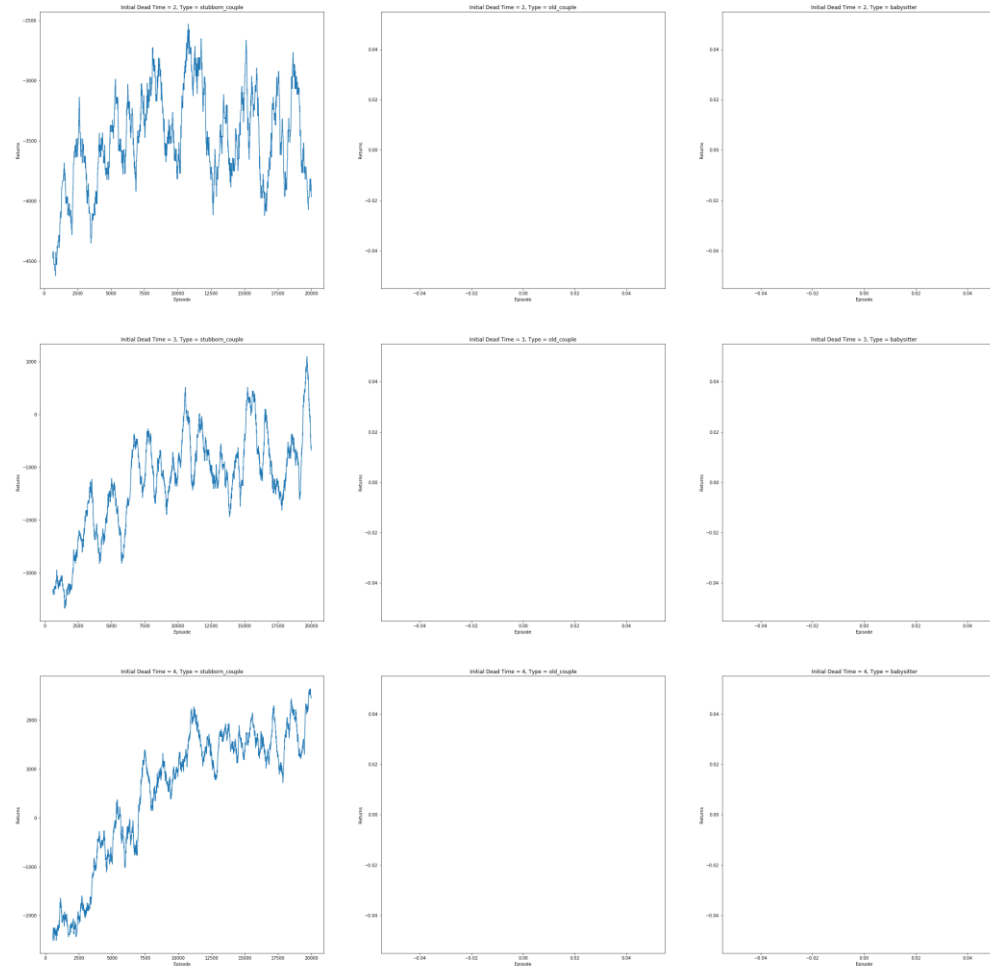
- Group saved: + 5000
- Group killed: -5000
- Interrupt operator:  $op\_busy * -100$
- Interrupt operator when already status ON: -50

# POLICIES ATTEMPTED

1. Only stubborn couple
2. Only babysitter
3. Only old couple
4. Only dead time initialization to 2
5. Only dead time initialization to 3
6. Only dead time initialization to 4

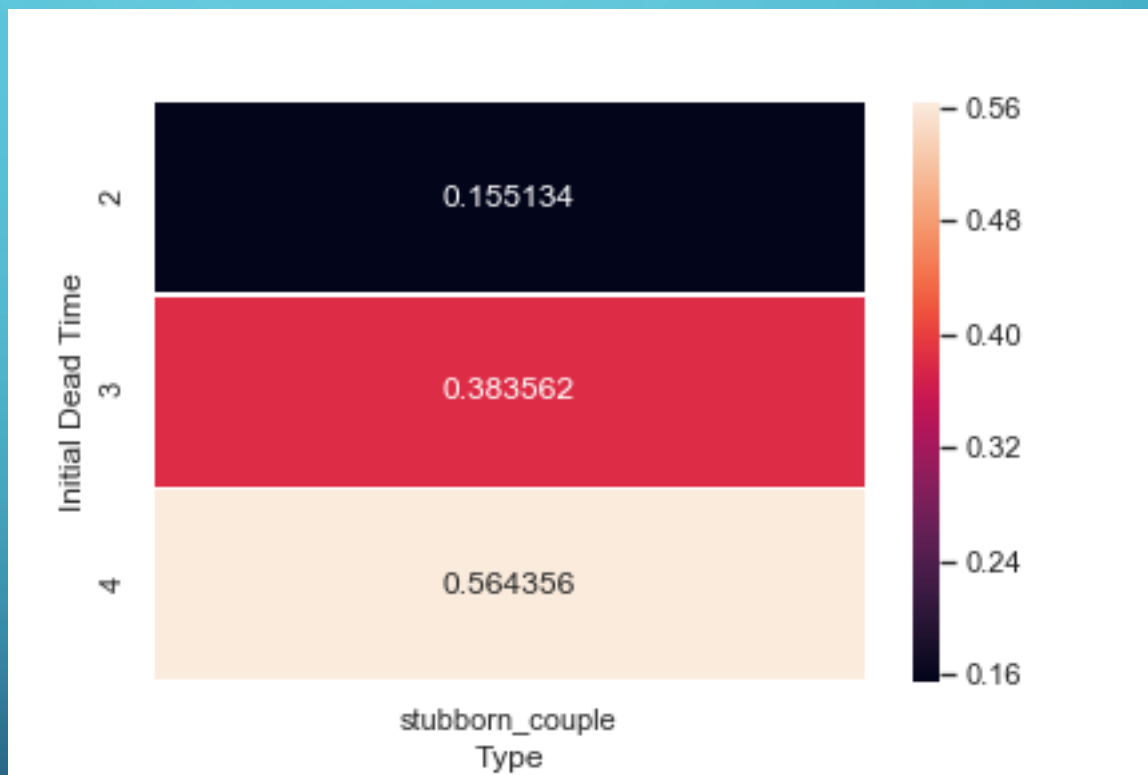
# POLICY 1

- Training curves



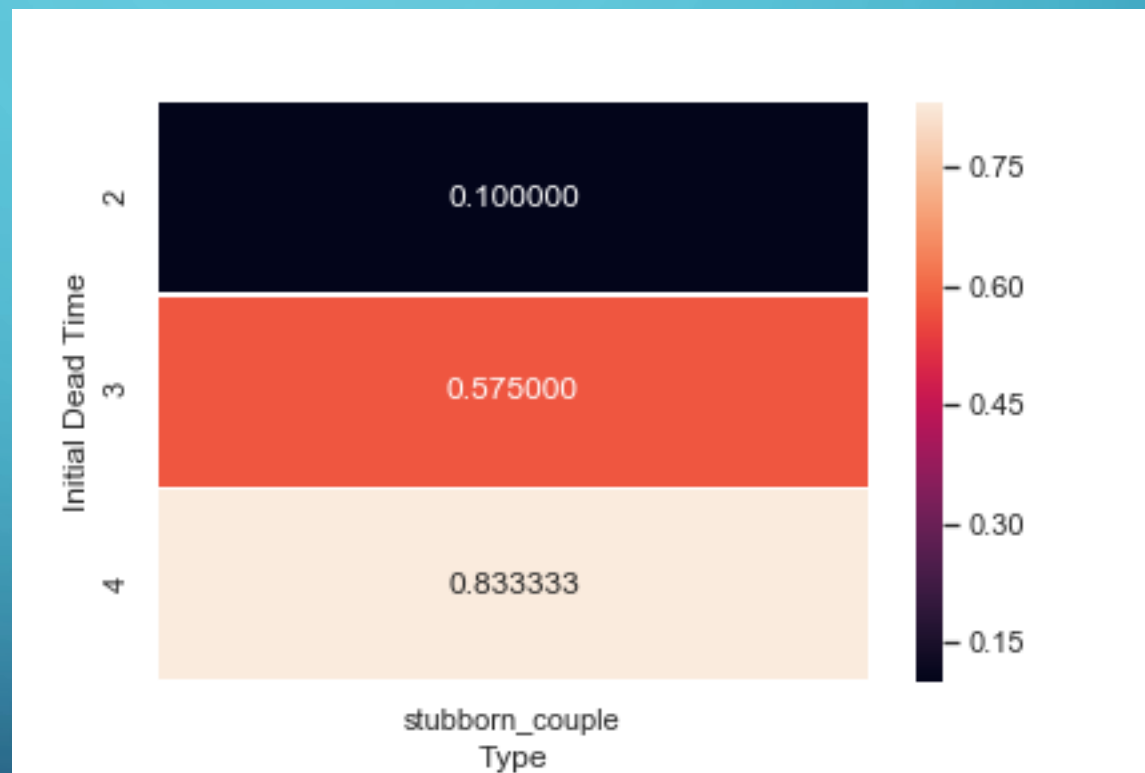
# POLICY 1

- Training statistics



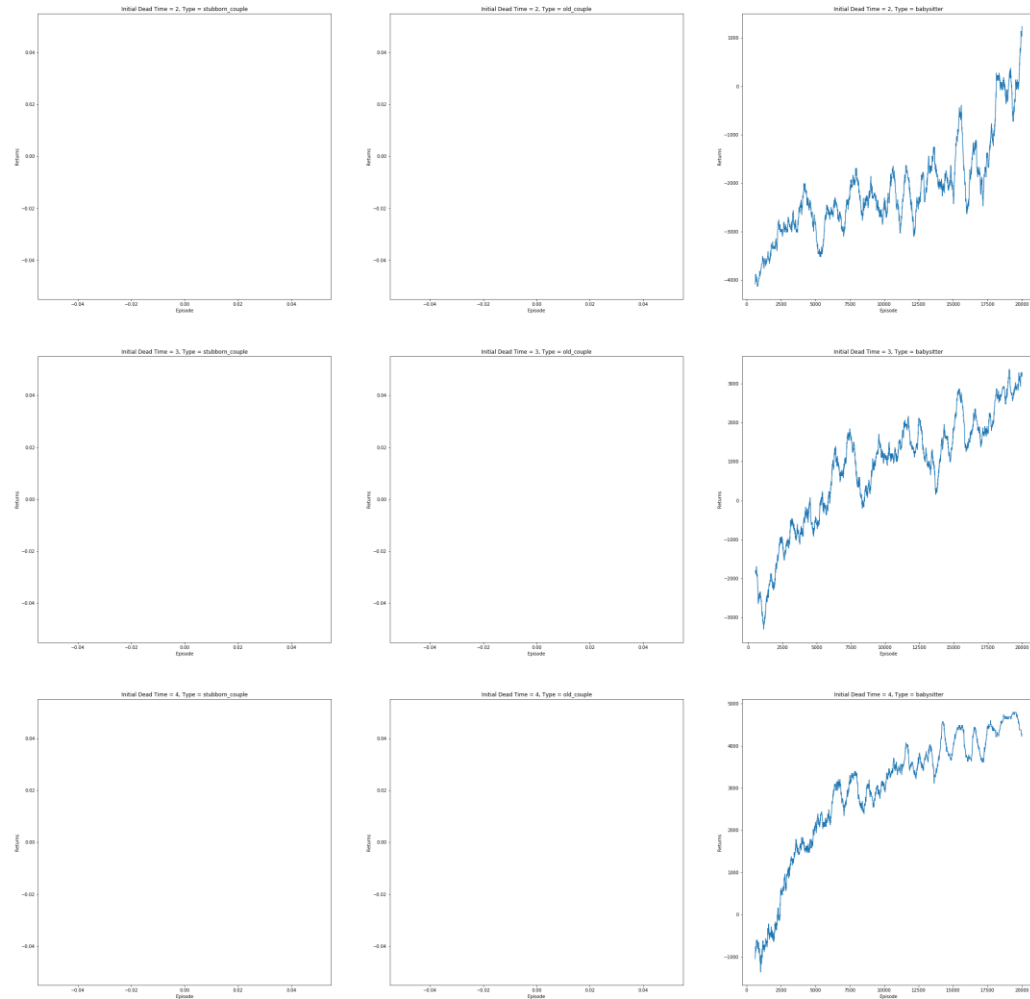
# POLICY 1

- Testing statistics



# POLICY 2

- Training curves



# POLICY 2

- Training statistics





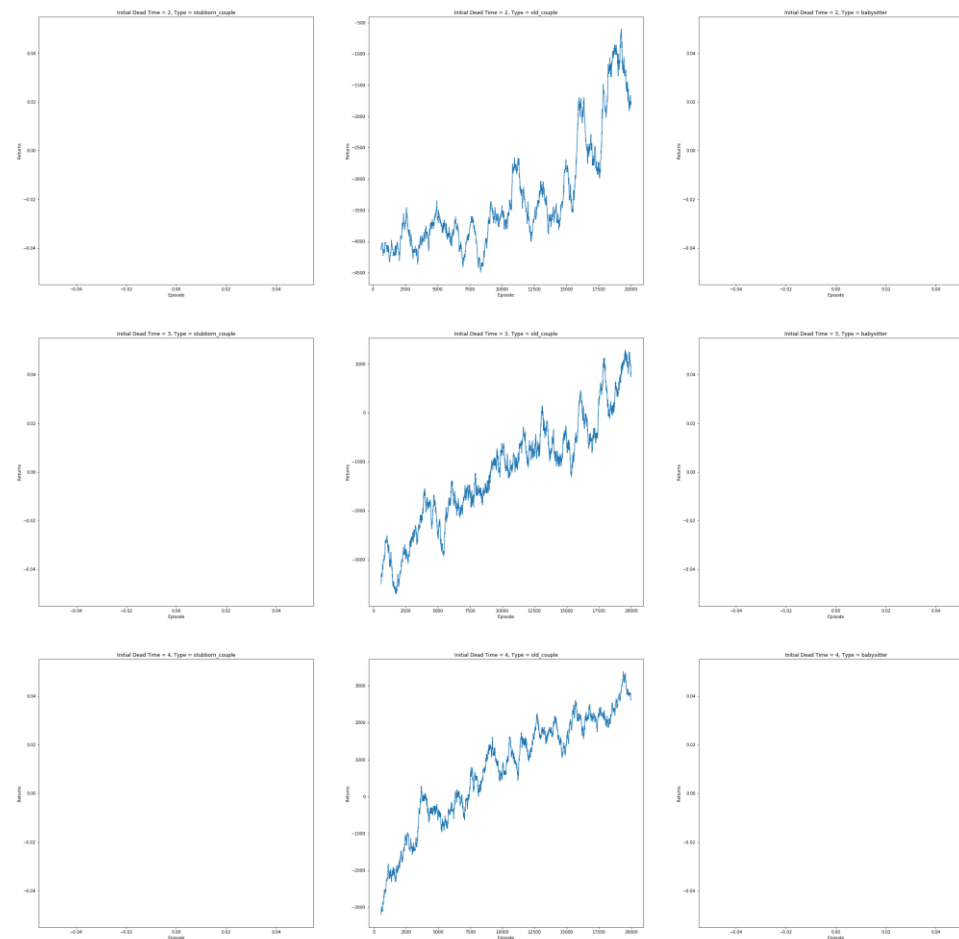
# POLICY 2

- Testing statistics



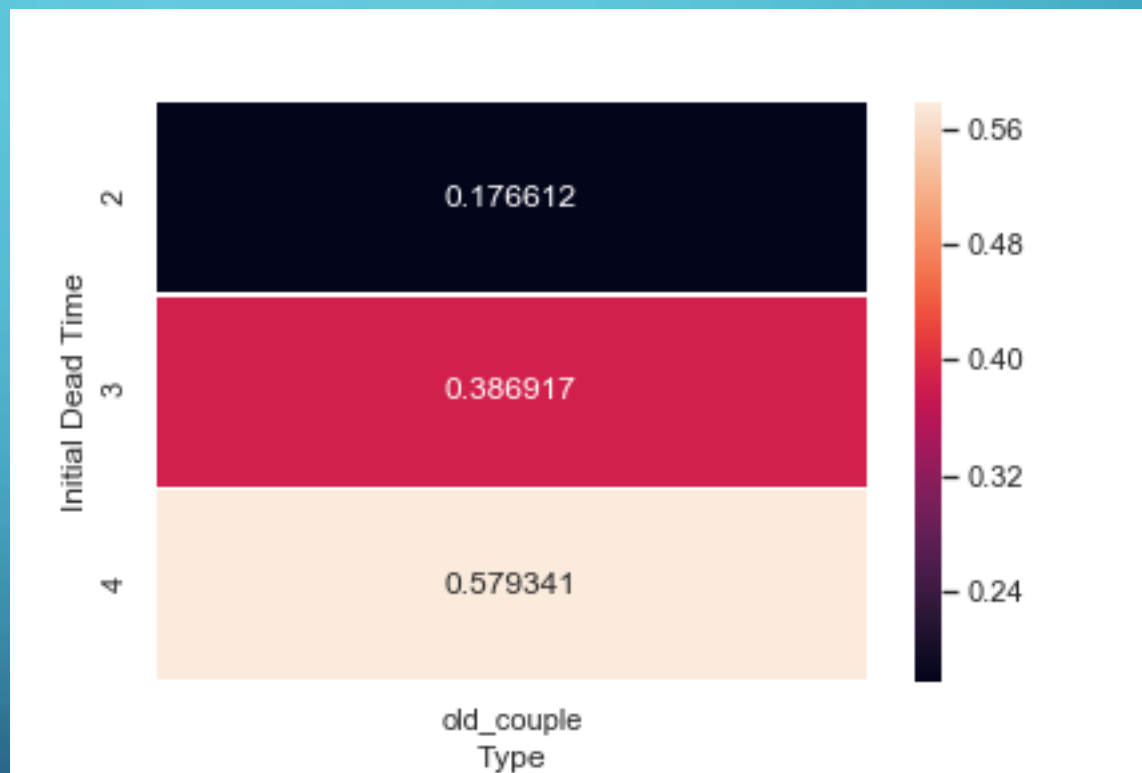
# POLICY 3

- Training curves



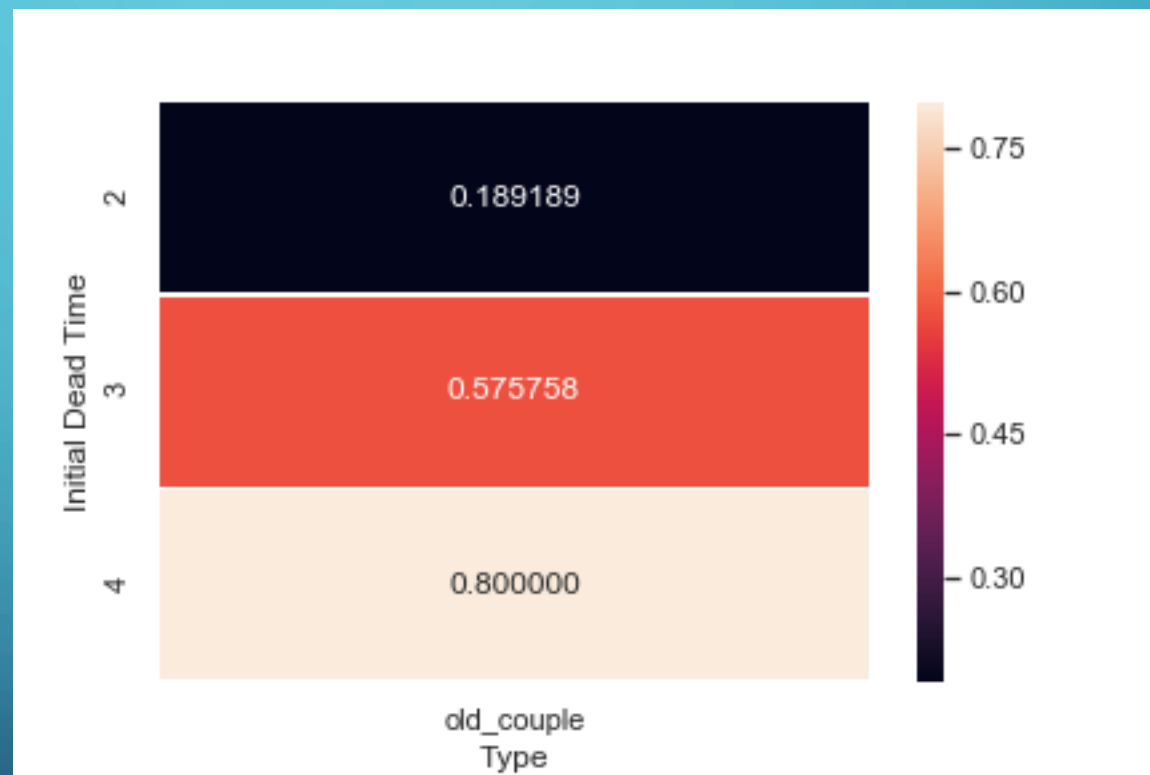
# POLICY 3

- Training statistics



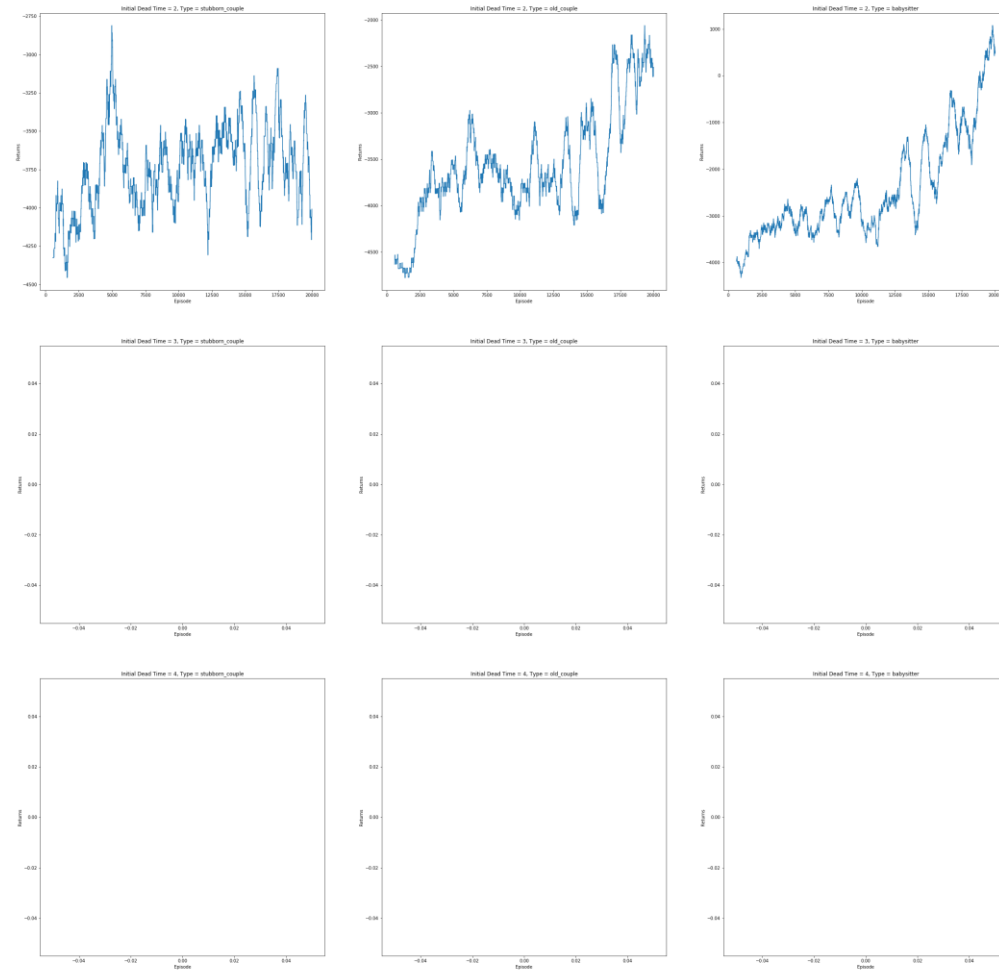
# POLICY 3

- Testing statistics



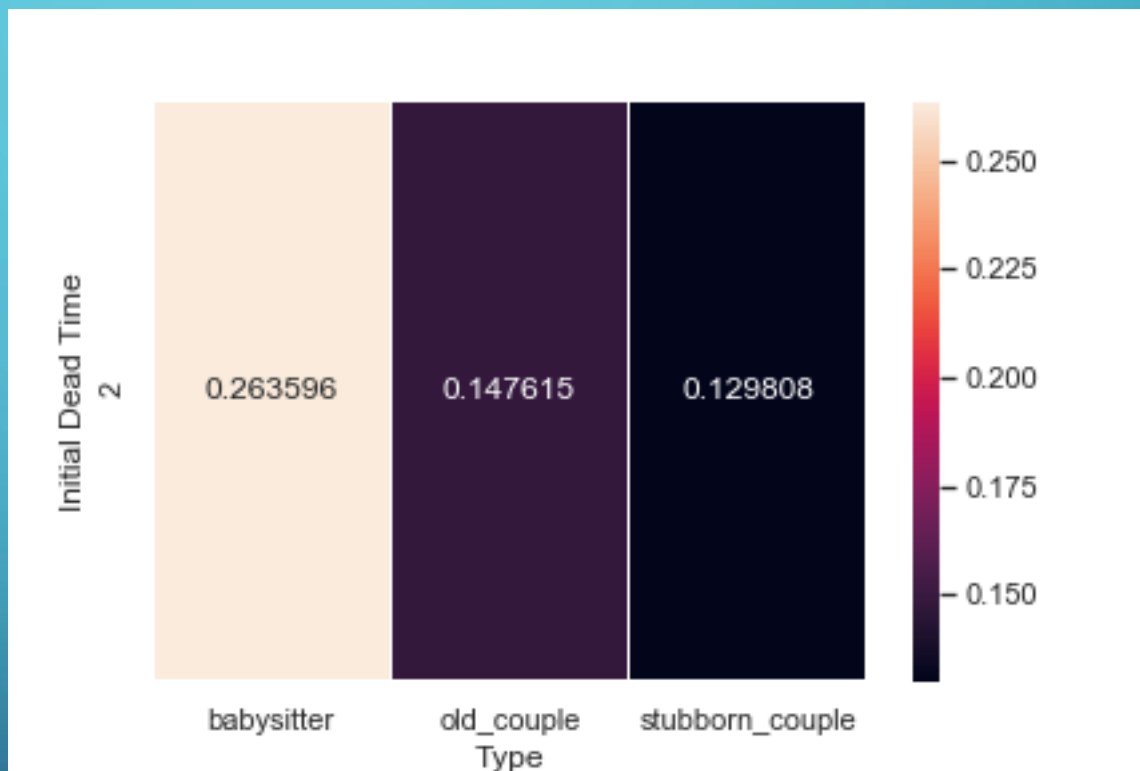
# POLICY 4

- Training curves



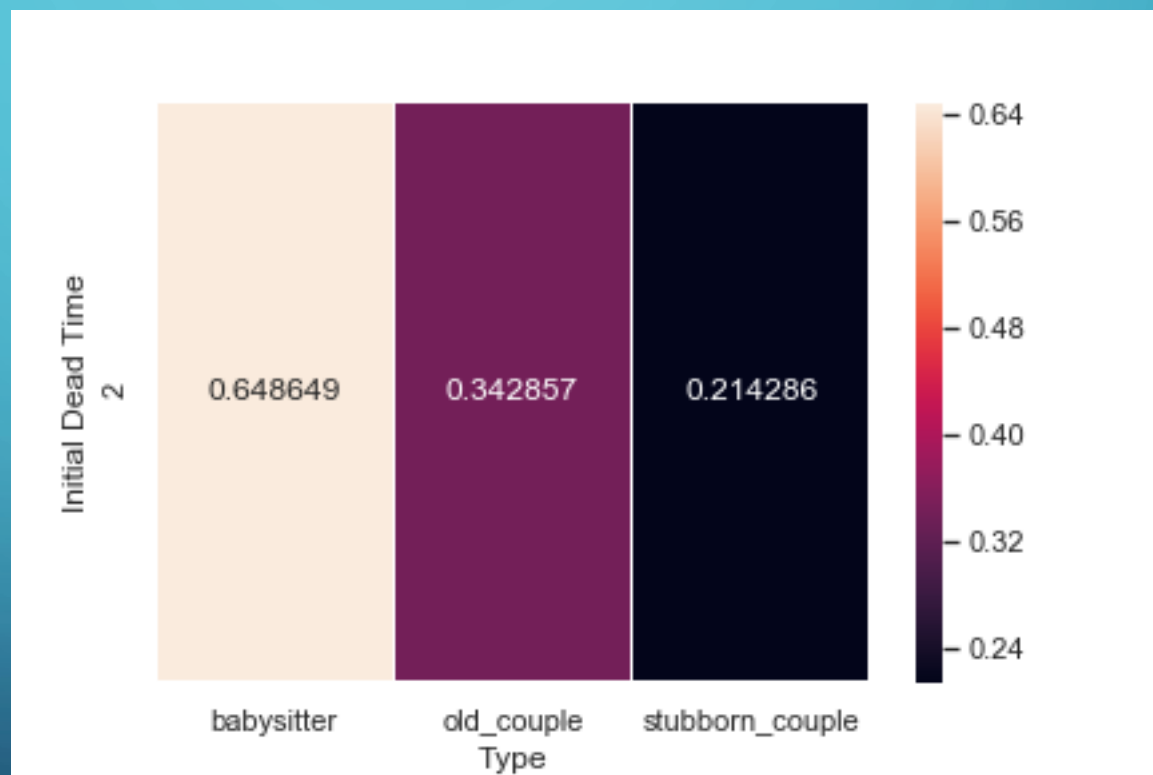
# POLICY 4

- Training statistics



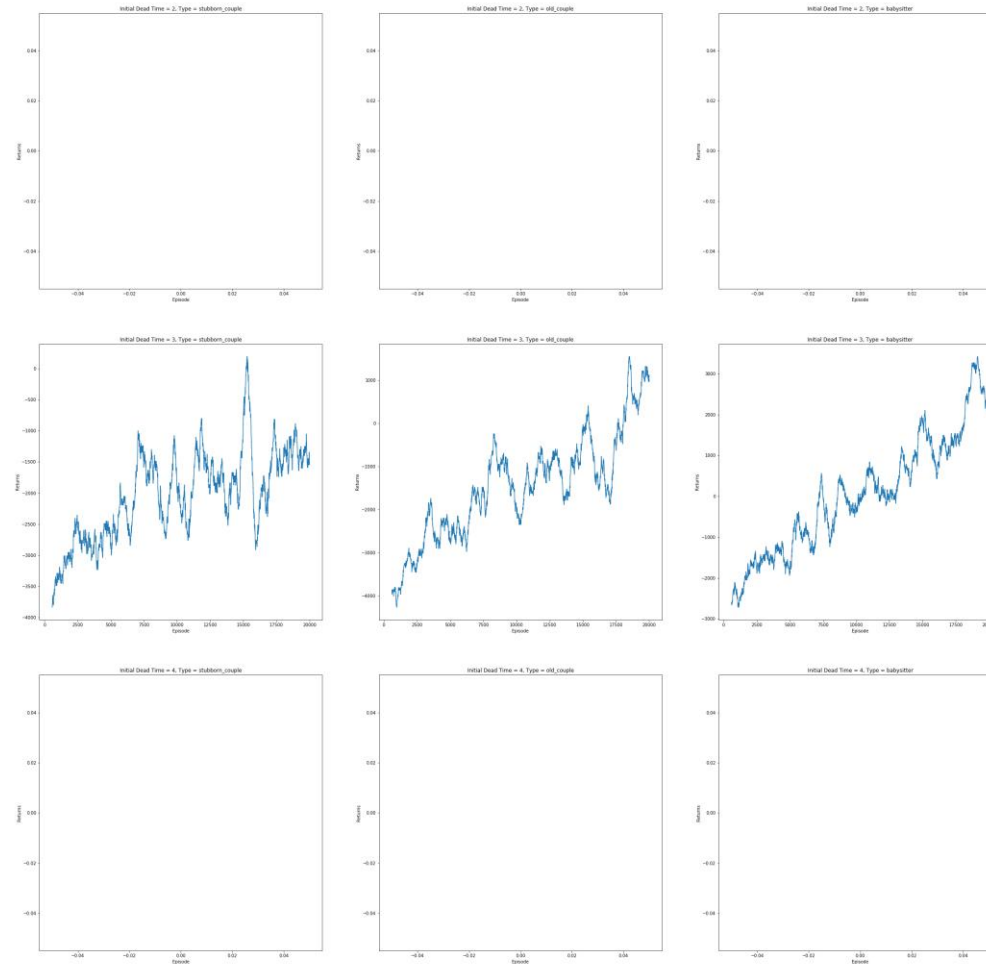
# POLICY 4

- Testing statistics



# POLICY 5

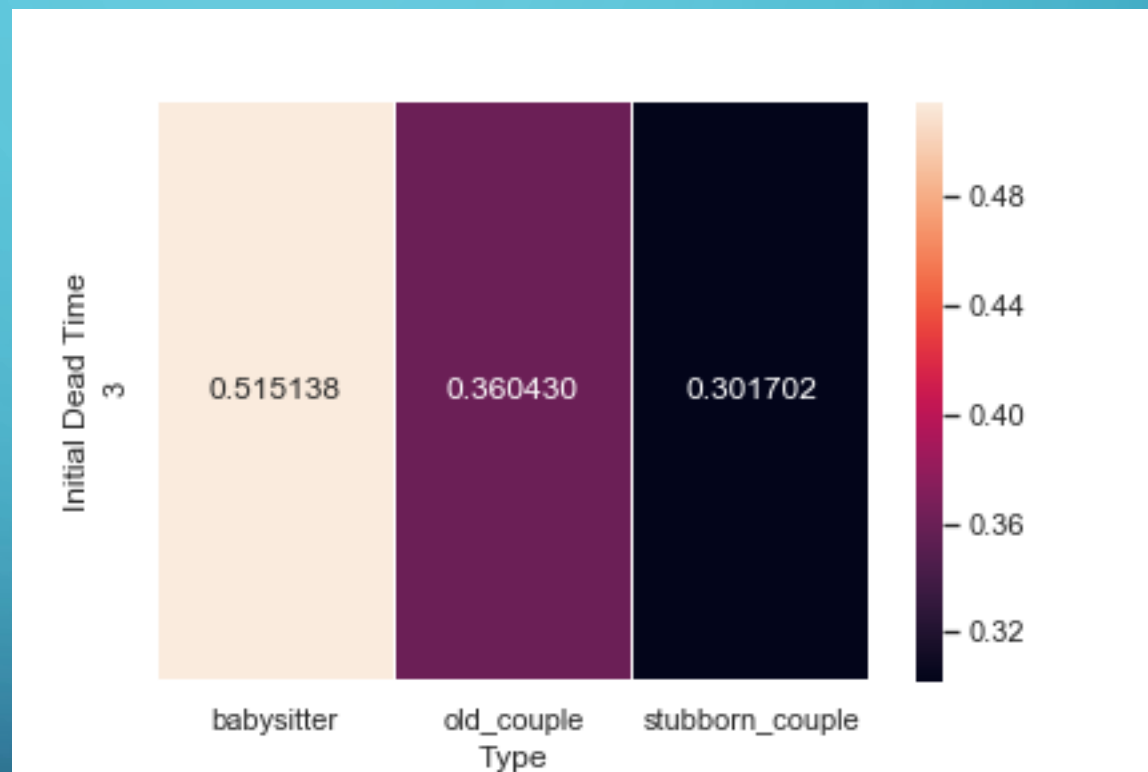
- Training curves





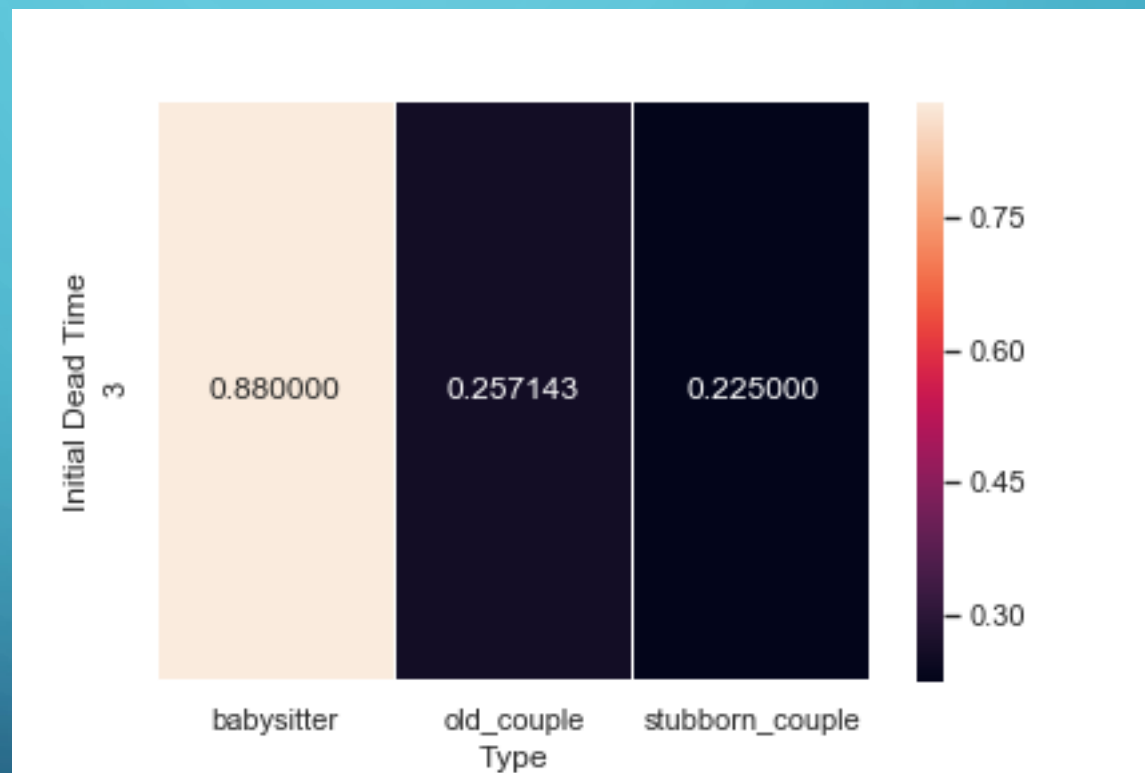
# POLICY 5

- Training statistics



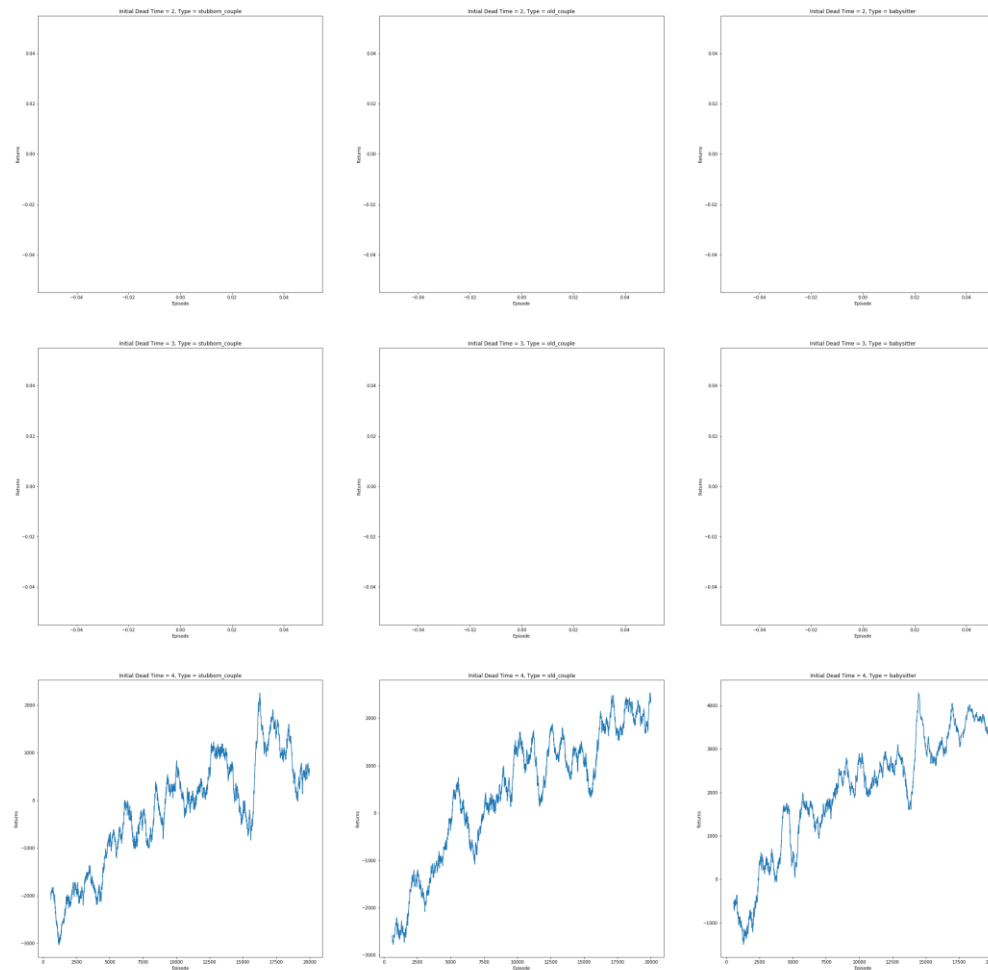
# POLICY 5

- Testing statistics



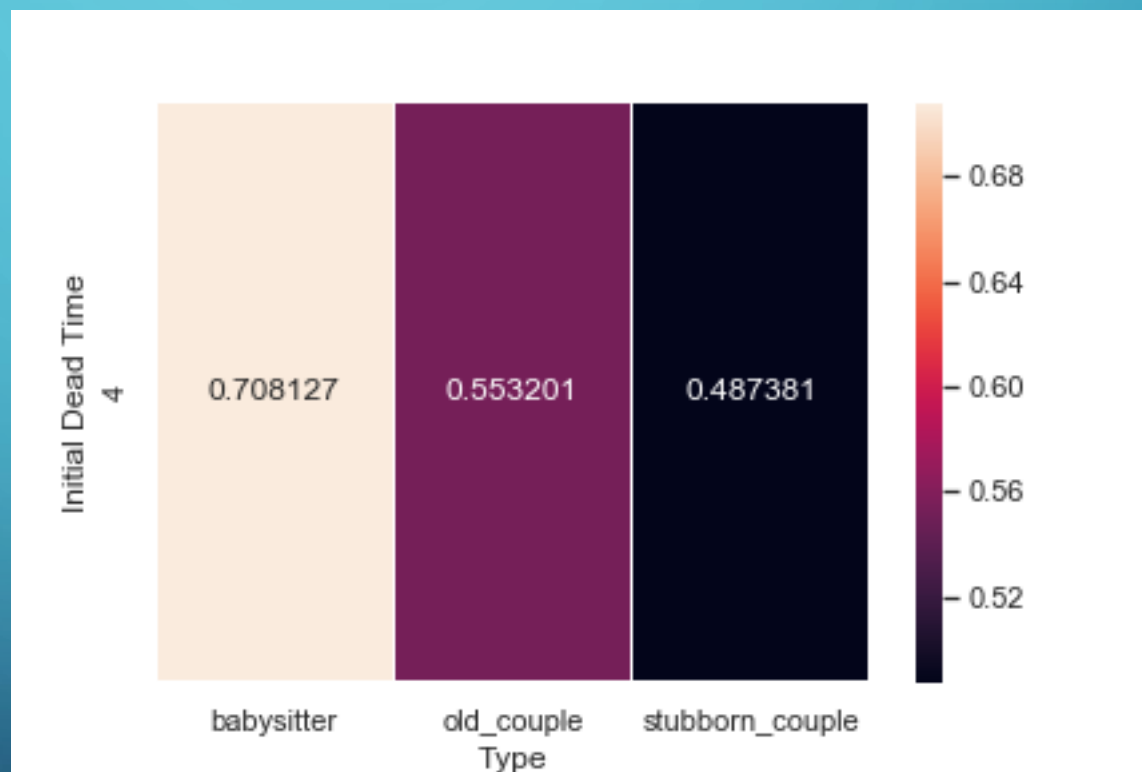
# POLICY 6

- Training curves



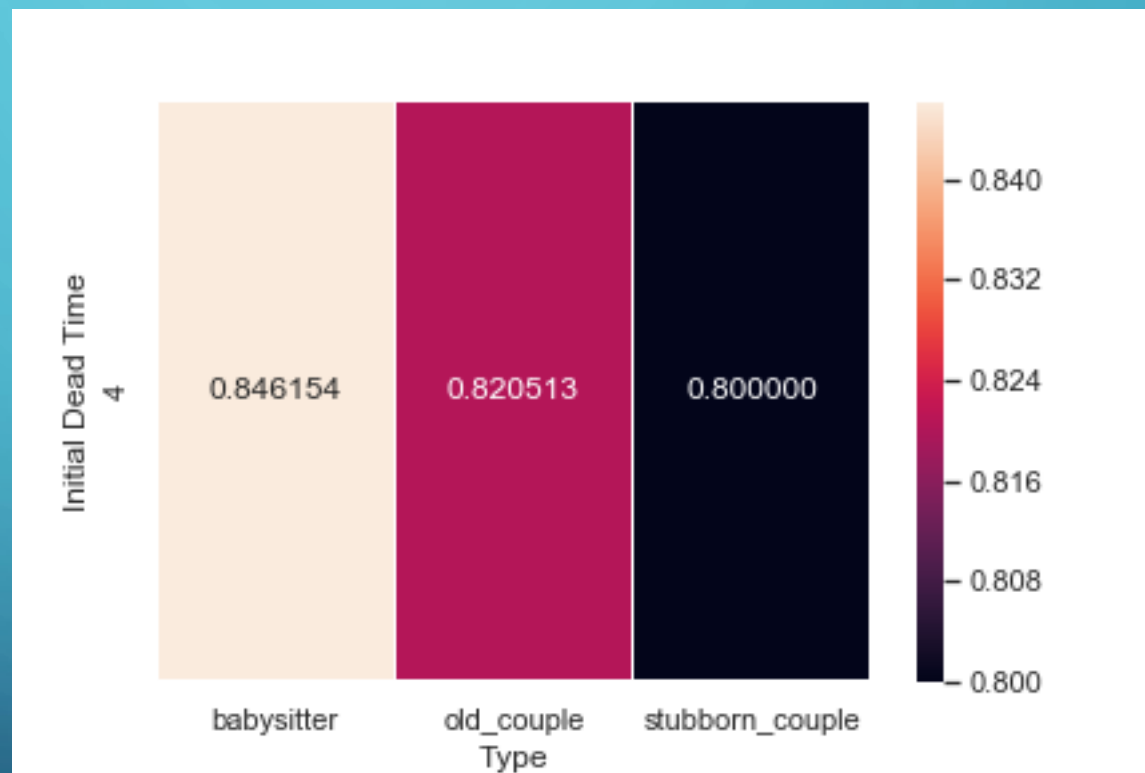
# POLICY 6

- Training statistics



# POLICY 6

- Testing statistics



# OBSERVATIONS – FROM TEST STATISTICS

	Stubborn Couple	Babysitter	Old Couple
Dead time init 2	0.10/0.21	0.59/0.65	0.19/0.34
Dead time init 3	0.58/0.23	0.87/0.88	0.58/0.26
Dead time init 4	0.83/0.80	1.00/0.85	0.80/0.82

Value from individual group policy/ value from individual time policy