Mini-project 1: Deep Q-learning for Epidemic Mitigation

Amin Asadi Sarijalou
IC, EPFL
amin.asadisarijalou@epfl.ch

Ilker Gül
IC, EPFL
ilker.gul@epfl.ch

Question 1: No Epidemic Mitigation

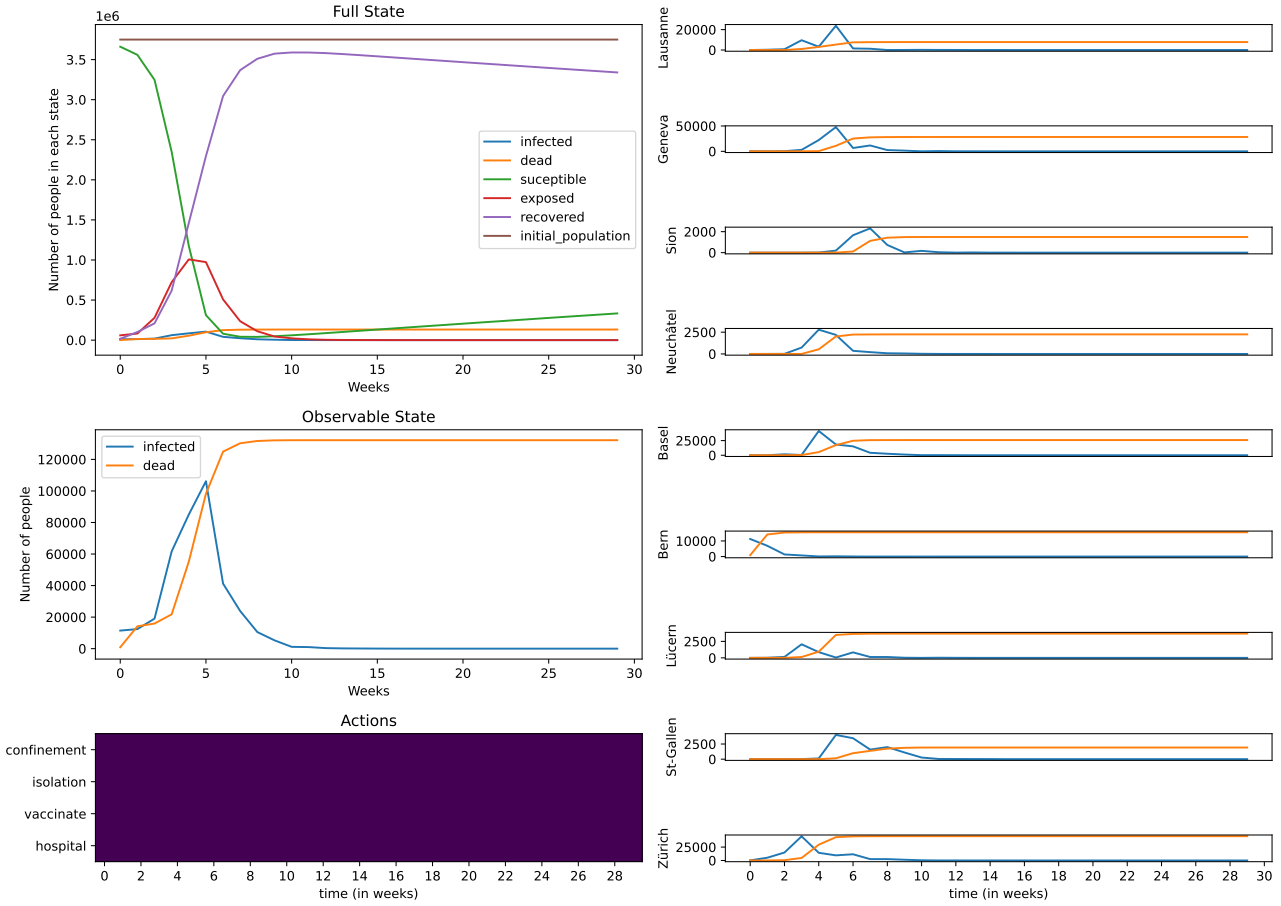

Figure 1: As we can see in this plot, most of the change in the variables occur in the first 10 weeks: The number of susceptible individuals drop significantly because they either get exposed or infected and afterwards they get recovered or they die. In this regard, we can observe a notable increase in the number of recovered individuals. This metric peaks at week 10 where it gets very close to the initial population size. This means, a significant portion of the population gets infected or at least exposed in the first 10 days. The recovered people gain immunity against the virus; thus, for a long time afterwards, no significant change occurs in the variables. We also observe that the number of susceptible and recovered people start to increase and decrease, respectively, after around week 10. This is because some recovered people lose their immunity and become susceptible again. The number of exposed and infected individuals both peak at week 5; nevertheless, the number of exposed is much higher (≈ 1 M) than the number of infected (≈ 0.1 M), which means most of the exposed individuals, get recovered before getting contagious. Finally, the number of dead people grow sharply in the first 7 weeks, but following that, as people get recovered and gain immunity, it almost levels off at around $130,000$, which is relatively high for a population of size $\approx 3.7M$.

Question 2: Prof. Russo's Policy

2.a: Implement Pr. Russo's Policy

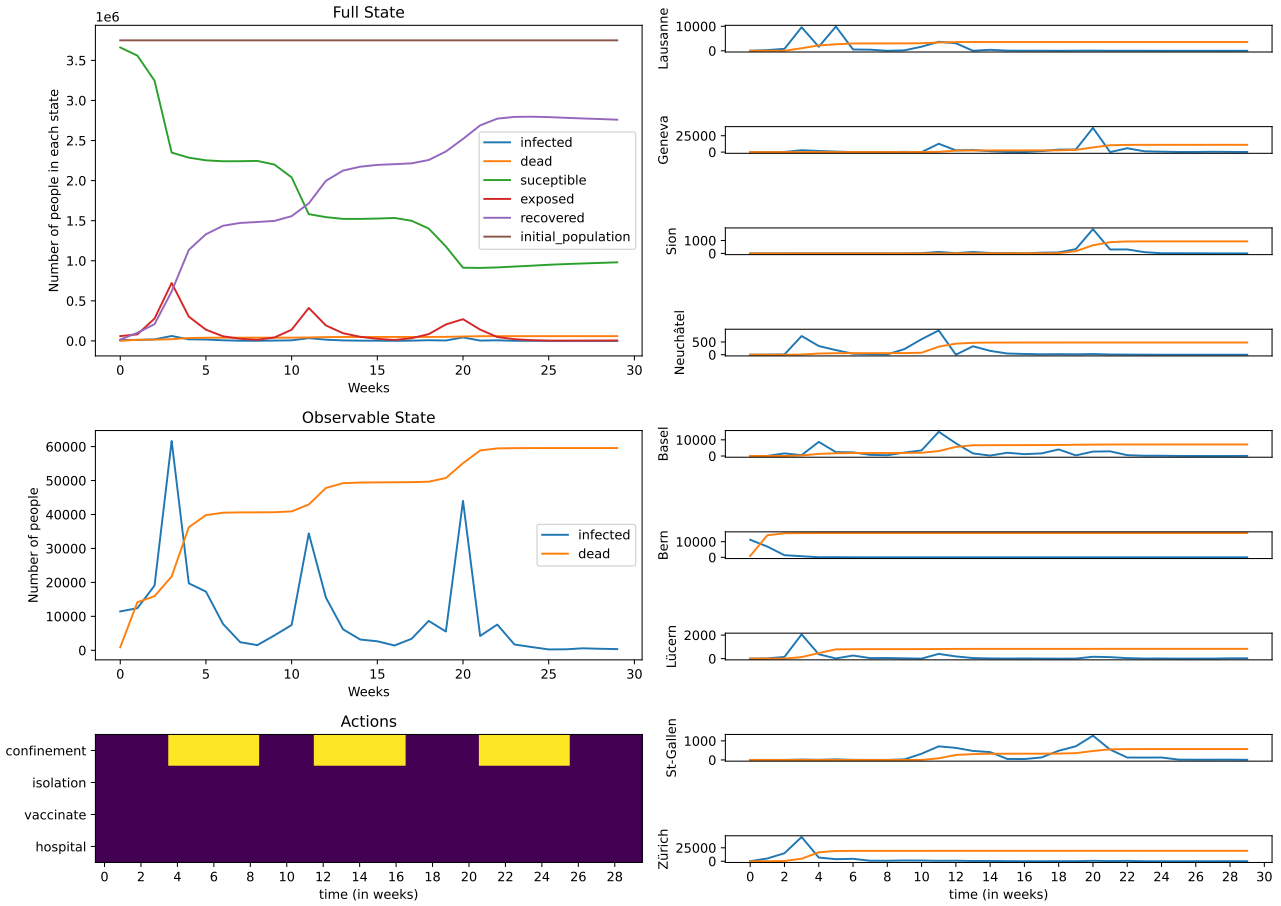

Figure 2: We observe that as expected, each time the number of infected reaches above 20,000 individuals, the confinement action is taken for the next 2 weeks, which results in a sharp decrase in the number of infected and exposed variables. The death toll is almost half of the death toll compared to the unmitigated scenario. Also, the increase in the number of deaths happens slower and more gradually with the new policy. Another important observation is that the number of susceptible people levels off at $\approx 1M$ people (after week 20), which means that $\approx 1M$ of people were never infected by the virus, so they never gained immunity and remained susceptible to the virus.

2.b: Evaluate Pr. Russo's Policy

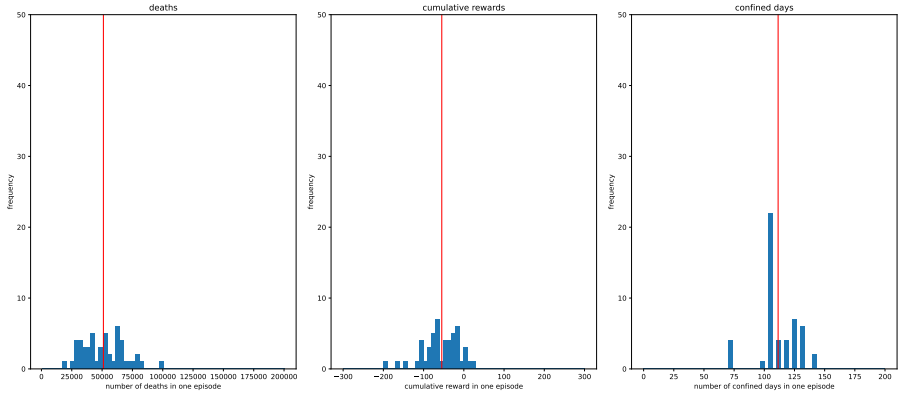

Figure 3: The histograms resulting from the evaluation of the Pr. Russo's Policy for 50 episodes. Average death number: 51079.26, Average cumulative reward: -54.80, Average number of confined days: 111.44.

Question 3: A Deep Q-learning approach

3.a: Implementing Deep Q-Learning

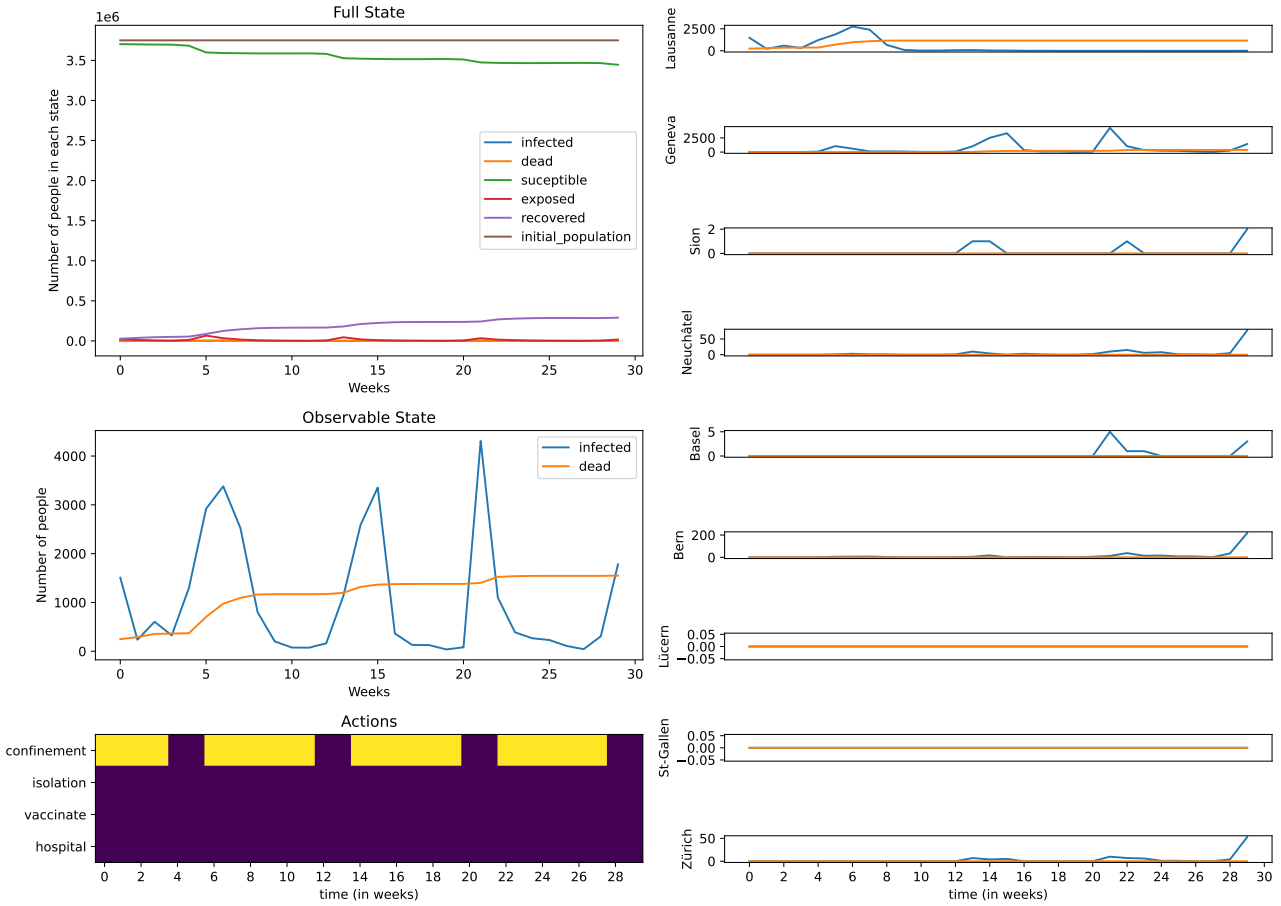

Figure 4: An example episode of DQN policy. (Note that the action space is only *confinement* or *no-confinement*). As we can observe, the policy is meaningful and also effective. As soon as the number of infected individuals reaches $\approx 3,000$, it takes the confinement action.

.

Question 4

Question 5