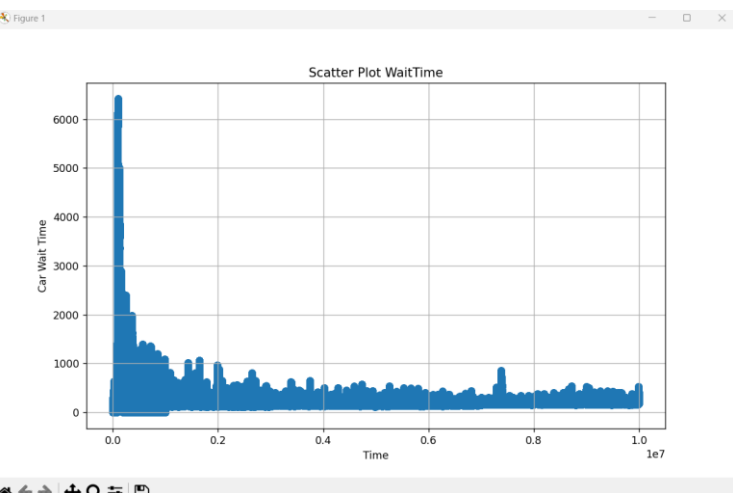
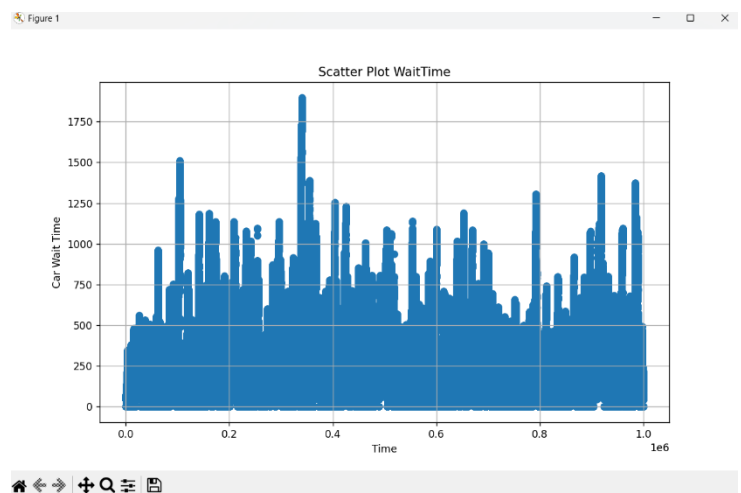


Based on the two figures, a larger step size worked better than a small step size in our case. The faster learning rate performed better. Perhaps if we had let the Agent run longer, it would have surpassed the larger step size.

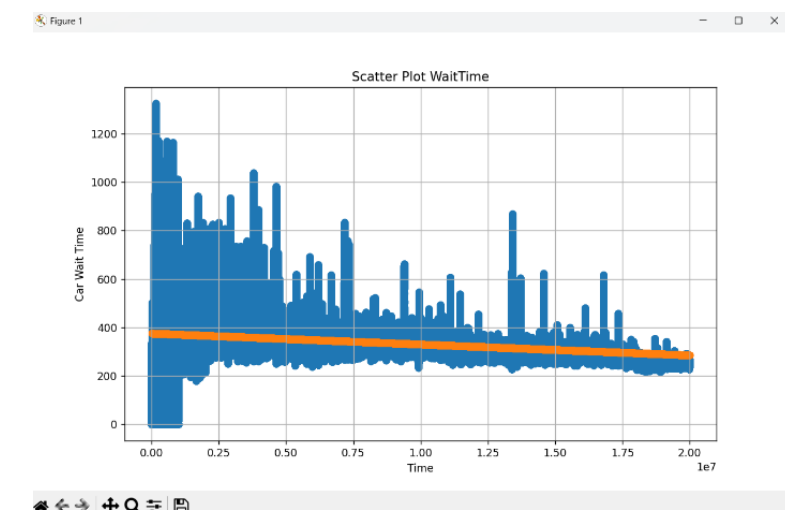


Step_size =0.7

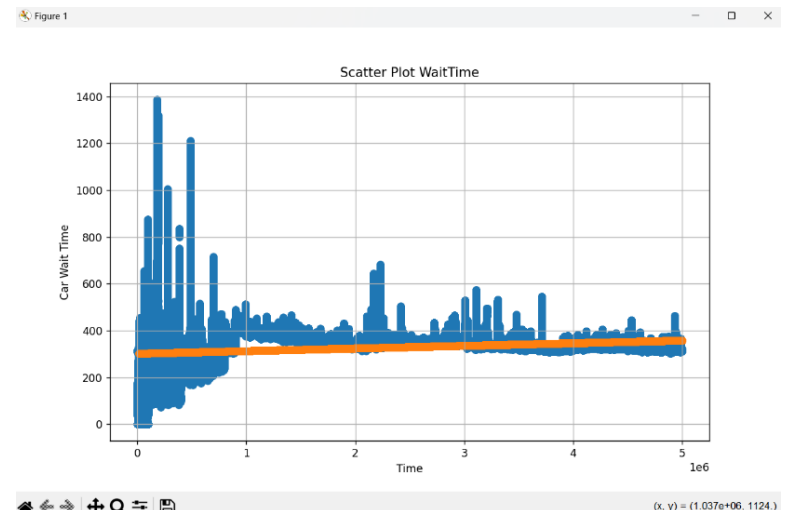


step_size =0.01

We found that a higher positive scaling was better than having positive scaling be equal to or less than negative scaling. Although altogether quite similar, the positive scaling had better relative convergence, if slightly more volatility on average.



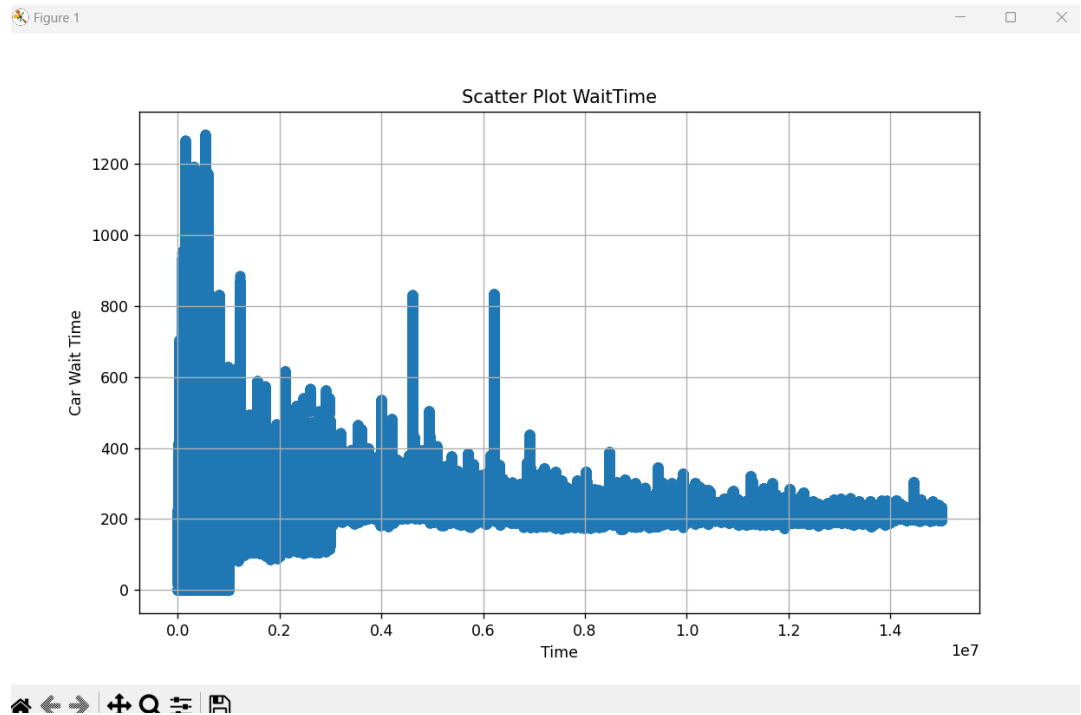
pos_scale = 2, neg_scale = 1.5, buffer_scale = 3



pos_scale = 1.5, neg_scale = 2, buffer_scale = 3

We also tested having a larger free buffer scaling, which we found to be much more impactful on results. We found that average wait time usually converged at a lower value, and with much less variance than with a small buffer scaling.

pos_scale = 1
neg_scale = 0.5
buffer_scale = 9



Given our experiments, we found that the most impactful variables for QLearning were a high buffer reward and a high step size. Positive should also be slightly larger than the negative rewards, but altogether was not overly impactful on results.