In your report, mention what you see in the agent's behavior. Does it eventually make it to the target location?

When choosing the action randomly, 47 out of 100 trails reach destination within dealine+100 steps. But most of the agent reach destination beyond deadline limitation.

Justify why you picked these set of states, and how they model the agent and its environment.

There are many ways to choose the state and I use input and next_waypoint as the state. Inputs including the traffic lights and oncoming cars indicates the current traffic conditions. And next_waypoint gives information of the correct direction and helps the agent to learn more effectively.

Q-learning: What changes do you notice in the agent's behavior?
Report what changes you made to your basic implementation of Q-learning to achieve the final version of the agent. How well does it perform?

When using random actions, only 47 episodes arrive at the destination within time limit but after applying Q-learning, all 100 episodes arrive at the destination. What's more, they arrive at destination within deadline.

Does your agent get close to finding an optimal policy, i.e. reach the destination in the minimum possible time and not incur any penalties?
The agent has found the optimal policy.
In order to evaluate the effectiveness of the model, I print out total steps to reach the destination. However, it is unreasonable to use total steps as a bench mark because if the Manhattan Distance is large, total step must be large. Therefore, I use total step/Manhattan Distance as a ratio to assess effectiveness. Firstly, I set epison=0.02, gamma=0.8 and alpha=0.7, the average of 100 trails total step/Manhattan Distance = 2.40. And I change epison = 0.05, gamma = 0.8, alpha = 0.5, the average of 100 trails total step/Manhattan Distance =2.36 and it takes less time to arrive at destination.