Question 1

3.

We have different from eighth, ninth and tenth iteration (U8) on box(row = 1,col = 2)

Reason:

We have difference since in R(s) = -0.05 in eighth iteration, the value in box(row = 1,col = 1) is greater than the value in box(row = 1,col = 3). However, in R(s) = -0.1 in eighth iteration, the value in box(row = 1,col = 1) is smaller than the value in box(row = 1, col=3). Thus, we would choose left as optimal policy in R(s) = -0.05 in contract with R(s) = -0.1 which choose right as optimal policy.

As for following states, since we change the policy in eighth iteration, the ninth and tenth iteration will be different in R(s) = -0.05 and R(s) = -0.1