Quiz 8: Direct Q Evaluation A

Supposed to use Direct Evaluation.

$$V^{\pi}(S_t) = E\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1}\right]$$

$$\gamma = 1 \Rightarrow V^{\pi}(S_t) = E\left[\sum_{k=0}^{\infty} R_{t+k+1}\right]$$

So I have to look through each of the episodes and find the average of the sum of rewards:

$$Q((2,2), E) = \frac{(0+0+100) + (0-30)+ (0+0+60)+ (0-10)+(0-30)}{5}$$

$$= \underline{\underline{18}}$$

$Q((1,2), S) = \underline{\underline{0}}$ (Not in the episodes)

$$Q((1,2), E) = ((0+0+0+100)+(0+0-30)+ (0+0+0+60)$$
$$+ (0+0-10)+ (0+0-30))/5 = \underline{\underline{18}}$$