# ASSIGNMENT – 3

## Part2 : Figure 5.1

### for 10000 Episodes

| Usable Ace | No Usable Ace |
|---|---|
|  |  |

Usable Ace, 10000 Episodes

No Usable Ace, 10000 Episodes

### for 500000 Episodes

| Usable Ace | No Usable Ace |
|---|---|
|  |  |

Usable Ace, 500000 Episodes
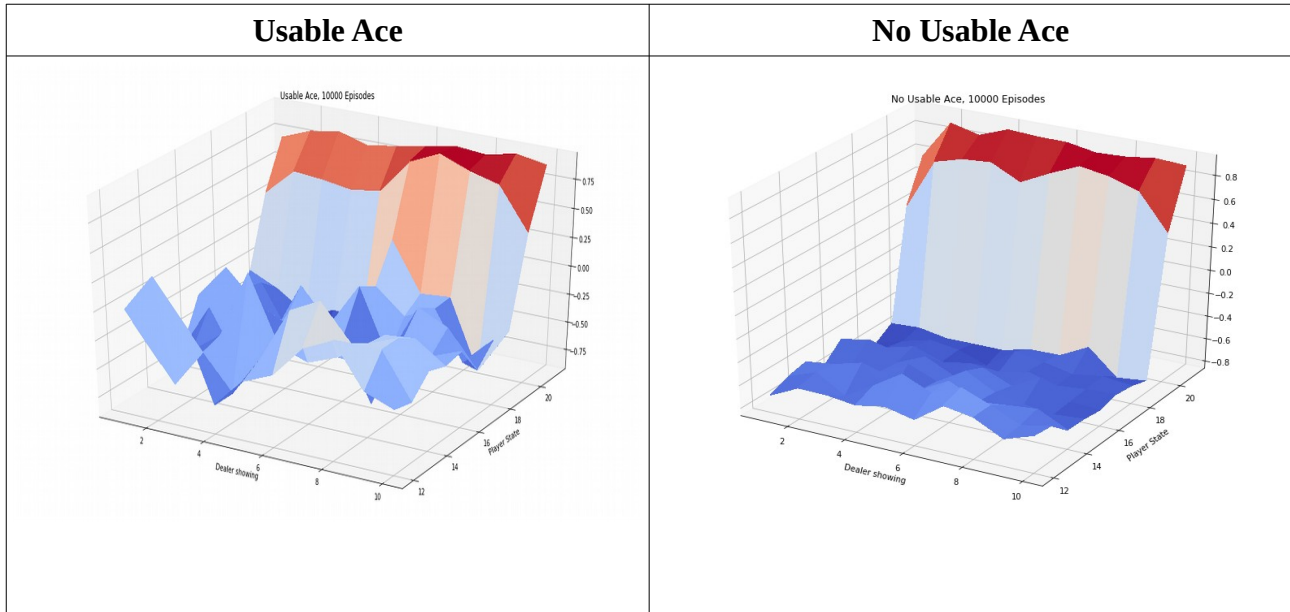
No Usable Ace, 500000 Episodes

**Fig 5.2 : Optimal Values**

| Usable Ace | No Usable Ace |
|:---:|:---:|
| V* with usable Ace | V* without usable Ace |

**Optimal Policy**

| Usable Ace | No Usable Ace |
|:---:|:---:|
| pi* with usable ace | pi* with no usable ace |

**Fig 5.3**

**Part 6: Example 6.2**

| Estimated Value | Empirical Error |
|---|---|
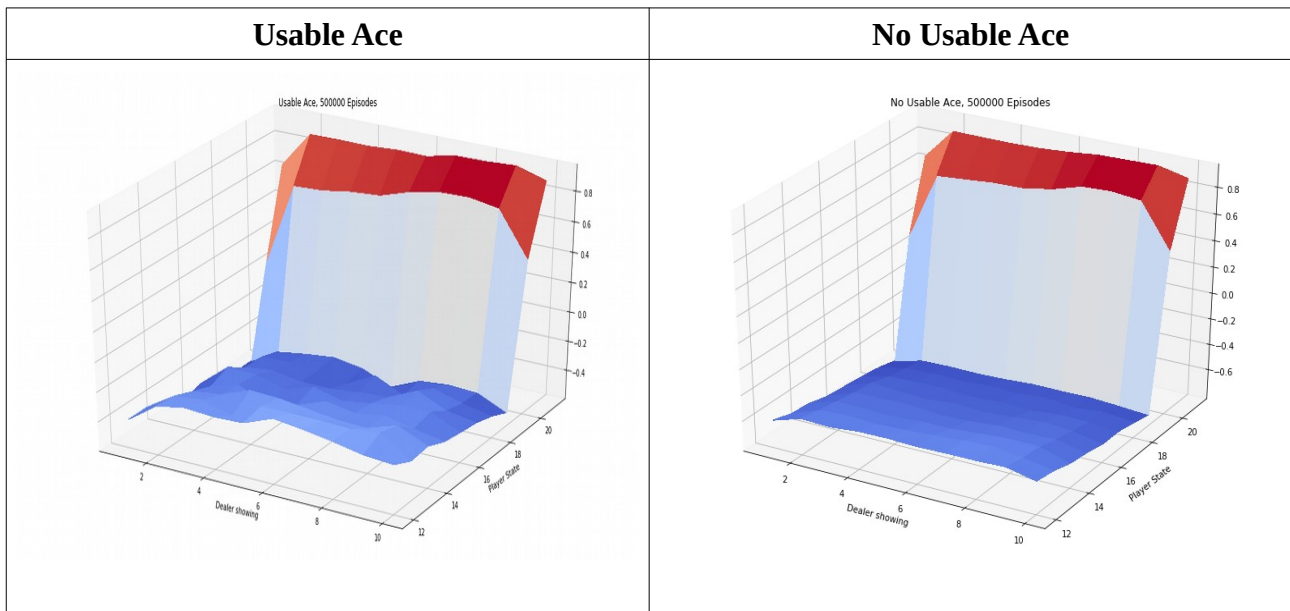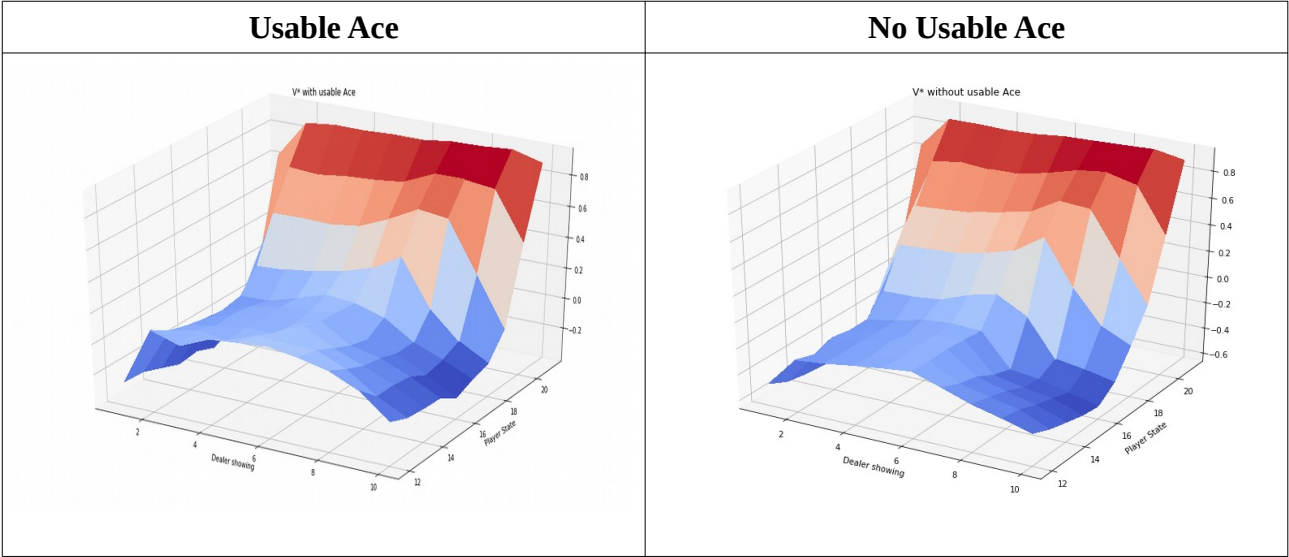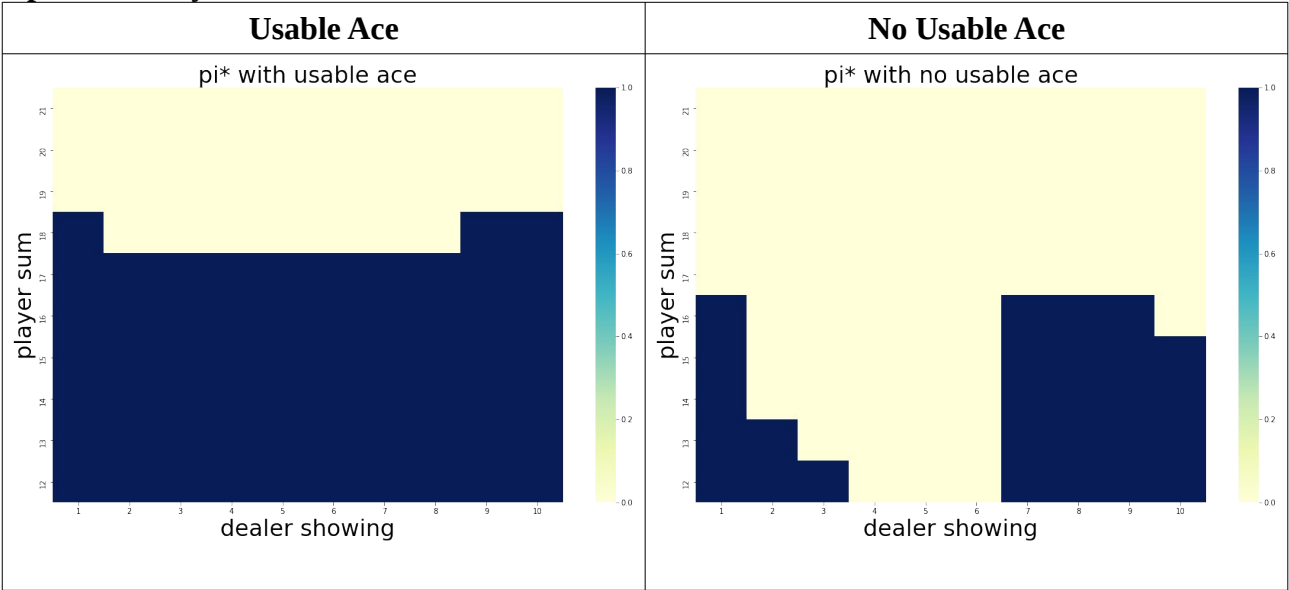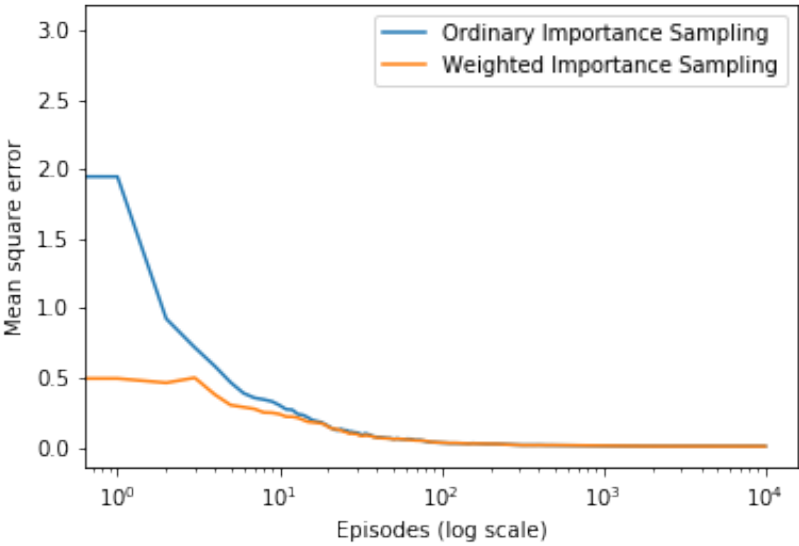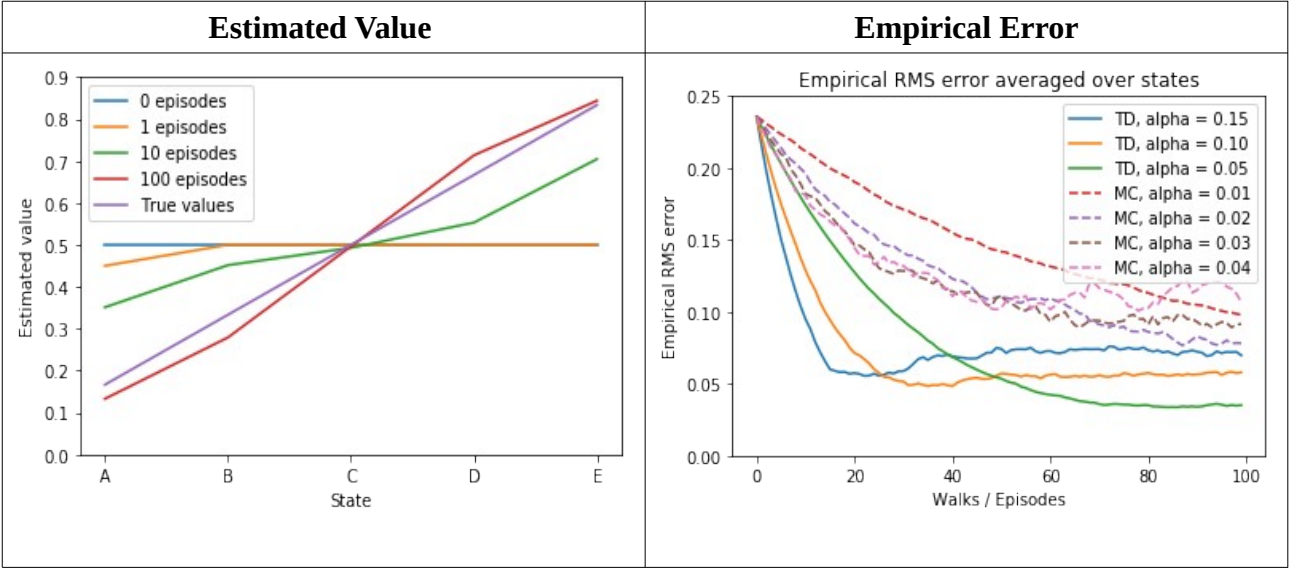


**Part 7**

**Comparison of SARSA and Q-learning Sum of Rewards during Episodes**