Question 1 Result for MC

Optimal Policy (MC First-Visit): [0 1 2 3 4]

Q-values (MC First-Visit):

```
 [[ 0.          -6.20929839 -4.23003228 -5.6897062
-7.10641275]

[-6.20929839  0.          -4.22915407 -0.65450284
-5.49900662]

[-4.23003228 -4.22915407  0.          -3.58076173
-2.92299256]

[-5.6897062  -0.65450284 -3.58076173  0.
-5.00644545]

[-7.10641275 -5.49900662 -2.92299256 -5.00644545  0.    ]]
```

Q-values (MC Every-Visit):

```
[[ 0.      -6.20929839 -4.23003228 -5.6897062  -7.10641275]

[-6.20929839  0.          -4.22915407 -0.65450284
-5.49900662]

[-4.23003228 -4.22915407  0.          -3.58076173
-2.92299256]

[-5.6897062  -0.65450284 -3.58076173  0.
-5.00644545]

[-7.10641275 -5.49900662 -2.92299256 -5.00644545  0.    ]]
```

Question 1 Result for DP

```
Optimal Values (DP):

[-746.1999451  -745.68954299 -746.19999421
-745.43432589 -745.74351698]

Optimal Policy (DP): [2 0 0 2 0]
```

Question 2 result

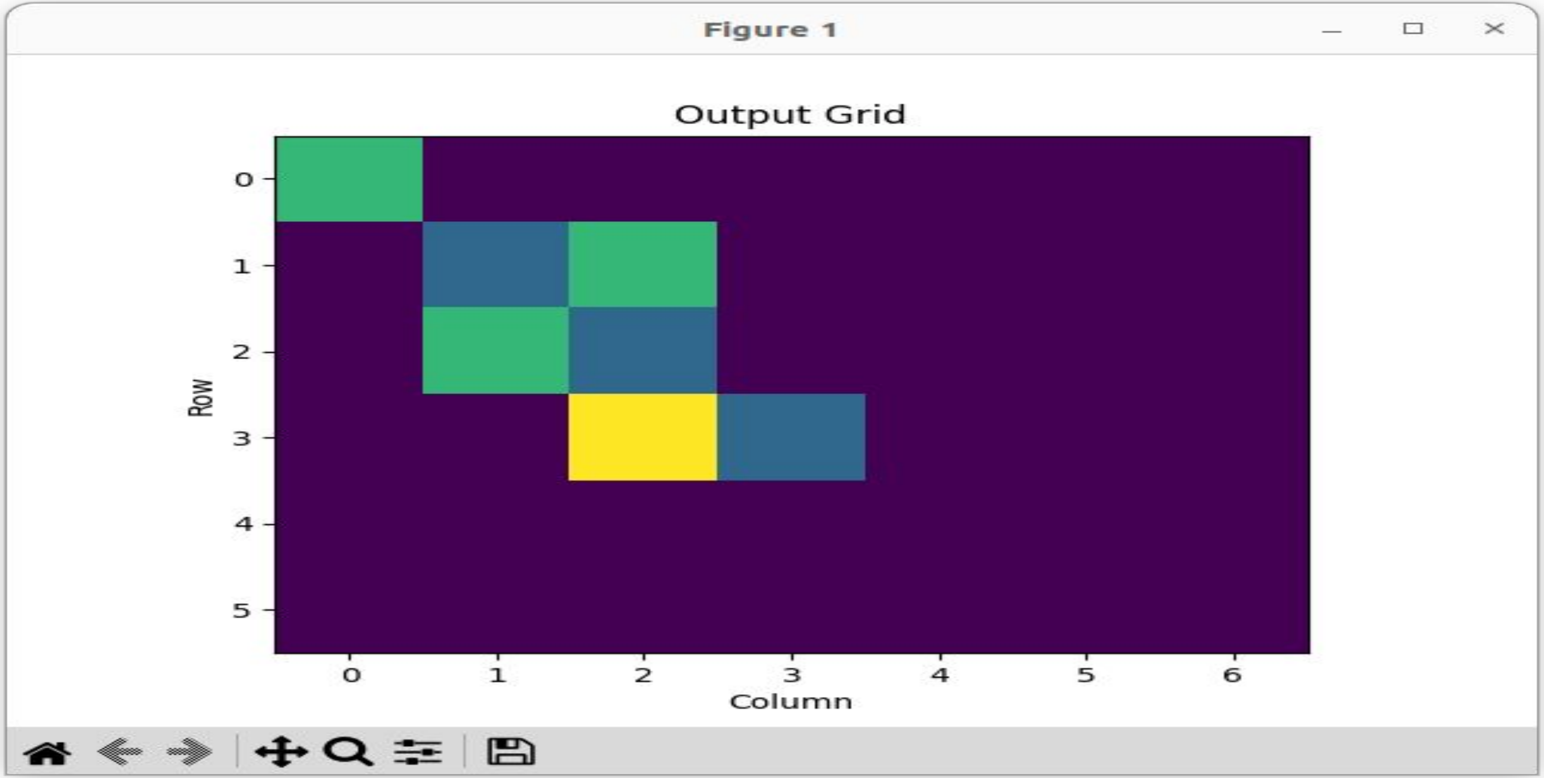
Figure 1
Output Grid
Row
0
1
2
3
4
5
Column
0
1
2
3
4
5
6