

Out[66]:

| | 0 | 1 | 2 | 3 | ... | 1840 | 1841 | 1842 | 1843 |
|-----|----------|----------|----------|----------|-----|------|------|------|------|
| 0 | 0.279303 | 0.281644 | 0.279118 | 0.277990 | ... | NaN | NaN | NaN | NaN |
| 1 | 0.277499 | 0.275546 | 0.274464 | 0.277357 | ... | NaN | NaN | NaN | NaN |
| 2 | 0.278139 | 0.277292 | 0.283209 | 0.281009 | ... | NaN | NaN | NaN | NaN |
| 3 | 0.282112 | 0.279055 | 0.278079 | 0.280397 | ... | NaN | NaN | NaN | NaN |
| 4 | 0.280623 | 0.277431 | 0.283111 | 0.282117 | ... | NaN | NaN | NaN | NaN |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 295 | 0.280947 | 0.277148 | 0.278684 | 0.277602 | ... | NaN | NaN | NaN | NaN |
| 296 | 0.280107 | 0.283195 | 0.280672 | 0.277256 | ... | NaN | NaN | NaN | NaN |
| 297 | 0.280817 | 0.280277 | 0.277665 | 0.282029 | ... | NaN | NaN | NaN | NaN |
| 298 | 0.278062 | 0.279478 | 0.279437 | 0.279340 | ... | NaN | NaN | NaN | NaN |
| 299 | 0.282934 | 0.285722 | 0.284294 | 0.279728 | ... | NaN | NaN | NaN | NaN |

[300 rows x 1844 columns]

In [67]: avg_rew_0_df = pd.DataFrame(avg_rew_0)

In [68]: avg_rew_0_df

Out[68]:

| | 0 | 1 | 2 | 3 | ... | 1840 | 1841 | 1842 | 1843 |
|-----|----------|----------|----------|----------|-----|------|------|------|------|
| 0 | 0.279303 | 0.281644 | 0.279118 | 0.277990 | ... | NaN | NaN | NaN | NaN |
| 1 | 0.277499 | 0.275546 | 0.274464 | 0.277357 | ... | NaN | NaN | NaN | NaN |
| 2 | 0.278139 | 0.277292 | 0.283209 | 0.281009 | ... | NaN | NaN | NaN | NaN |
| 3 | 0.282112 | 0.279055 | 0.278079 | 0.280397 | ... | NaN | NaN | NaN | NaN |
| 4 | 0.280623 | 0.277431 | 0.283111 | 0.282117 | ... | NaN | NaN | NaN | NaN |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 295 | 0.280947 | 0.277148 | 0.278684 | 0.277602 | ... | NaN | NaN | NaN | NaN |
| 296 | 0.280107 | 0.283195 | 0.280672 | 0.277256 | ... | NaN | NaN | NaN | NaN |
| 297 | 0.280817 | 0.280277 | 0.277665 | 0.282029 | ... | NaN | NaN | NaN | NaN |
| 298 | 0.278062 | 0.279478 | 0.279437 | 0.279340 | ... | NaN | NaN | NaN | NaN |
| 299 | 0.282934 | 0.285722 | 0.284294 | 0.279728 | ... | NaN | NaN | NaN | NaN |

[300 rows x 1844 columns]

In [69]: avg_rew_0_df.dropna(axis=1)

Out[69]:

| | 0 | 1 | 2 | ... | 707 | 708 | 709 |
|-----|----------|----------|----------|-----|----------|----------|----------|
| 0 | 0.279303 | 0.281644 | 0.279118 | ... | 0.323757 | 0.323465 | 0.323872 |
| 1 | 0.277499 | 0.275546 | 0.274464 | ... | 0.328651 | 0.328350 | 0.327730 |
| 2 | 0.278139 | 0.277292 | 0.283209 | ... | 0.319953 | 0.320075 | 0.319560 |
| 3 | 0.282112 | 0.279055 | 0.278079 | ... | 0.327804 | 0.327856 | 0.327905 |
| 4 | 0.280623 | 0.277431 | 0.283111 | ... | 0.324429 | 0.324322 | 0.324132 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 295 | 0.280947 | 0.277148 | 0.278684 | ... | 0.319455 | 0.319850 | 0.320096 |
| 296 | 0.280107 | 0.283195 | 0.280672 | ... | 0.332006 | 0.332125 | 0.332332 |
| 297 | 0.280817 | 0.280277 | 0.277665 | ... | 0.319468 | 0.319405 | 0.319124 |
| 298 | 0.278062 | 0.279478 | 0.279437 | ... | 0.335819 | 0.335612 | 0.335793 |
| 299 | 0.282934 | 0.285722 | 0.284294 | ... | 0.335501 | 0.335515 | 0.335639 |

[300 rows x 710 columns]

In [70]: avg_rew_0_df

Out[70]:

| | 0 | 1 | 2 | 3 | ... | 1840 | 1841 | 1842 | 1843 |
|-----|----------|----------|----------|----------|-----|------|------|------|------|
| 0 | 0.279303 | 0.281644 | 0.279118 | 0.277990 | ... | NaN | NaN | NaN | NaN |
| 1 | 0.277499 | 0.275546 | 0.274464 | 0.277357 | ... | NaN | NaN | NaN | NaN |
| 2 | 0.278139 | 0.277292 | 0.283209 | 0.281009 | ... | NaN | NaN | NaN | NaN |
| 3 | 0.282112 | 0.279055 | 0.278079 | 0.280397 | ... | NaN | NaN | NaN | NaN |
| 4 | 0.280623 | 0.277431 | 0.283111 | 0.282117 | ... | NaN | NaN | NaN | NaN |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 295 | 0.280947 | 0.277148 | 0.278684 | 0.277602 | ... | NaN | NaN | NaN | NaN |
| 296 | 0.280107 | 0.283195 | 0.280672 | 0.277256 | ... | NaN | NaN | NaN | NaN |

```

297 0.280817 0.280277 0.277665 0.282029 ... NaN NaN NaN NaN
298 0.278062 0.279478 0.279437 0.279340 ... NaN NaN NaN NaN
299 0.282934 0.285722 0.284294 0.279728 ... NaN NaN NaN NaN

```

[300 rows x 1844 columns]

```

In [71]: avg_rew_0_dropNaN_df = avg_rew_0_df.dropna(axis=1)
...: avg_rew_1_dropNaN_df = avg_rew_1_df.dropna(axis=1)

```

```

In [72]: avg_rew_1_dropNaN_df

```

```

Out[72]:
      0      1      2      ...      707      708      709
0  0.279303 0.281644 0.279118 ... 0.323757 0.323465 0.323872
1  0.277499 0.275546 0.274464 ... 0.328651 0.328350 0.327730
2  0.278139 0.277292 0.283209 ... 0.319953 0.320075 0.319560
3  0.282112 0.279055 0.278079 ... 0.327804 0.327856 0.327905
4  0.280623 0.277431 0.283111 ... 0.324429 0.324322 0.324132
..      ...      ...      ...      ...      ...      ...
295 0.280947 0.277148 0.278684 ... 0.319455 0.319850 0.320096
296 0.280107 0.283195 0.280672 ... 0.332006 0.332125 0.332332
297 0.280817 0.280277 0.277665 ... 0.319468 0.319405 0.319124
298 0.278062 0.279478 0.279437 ... 0.335819 0.335612 0.335793
299 0.282934 0.285722 0.284294 ... 0.335501 0.335515 0.335639

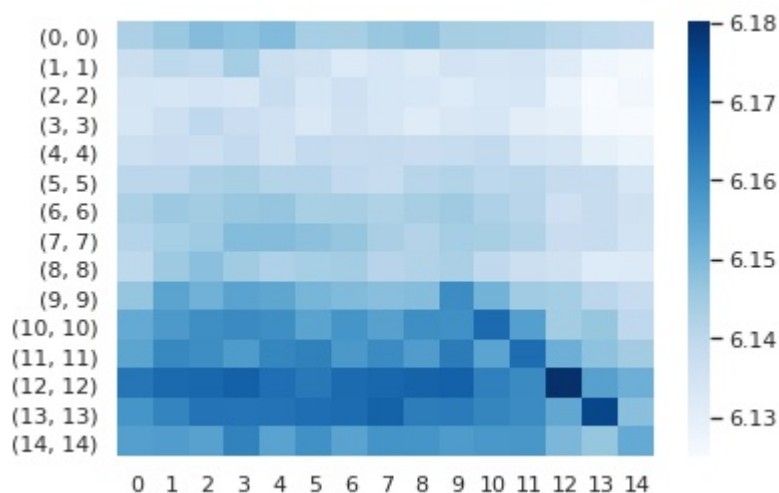
```

[300 rows x 710 columns]

```

In [73]: index = []
...: for i in range(15):
...:     for j in range(15):
...:         if i==j:
...:             index.append(s_dict[(i,j)])
...:
...: labels = []
...: for i in index:
...:     labels.append(list(s_dict.keys())[i])
...:
...:
...: sns.heatmap(final_avg_q[index,], yticklabels=labels, cmap="Blues")
...: plt.savefig('fig1.png', dpi=600)
...: plt.show()

```

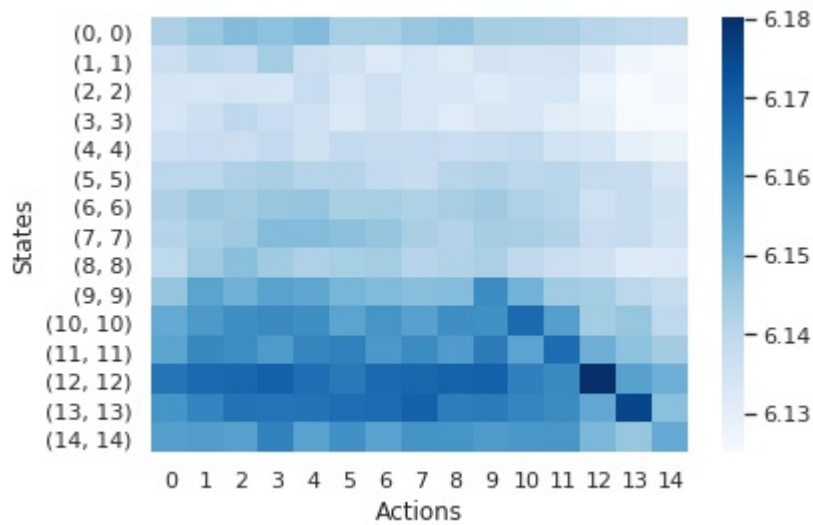


```

In [74]: sns.heatmap(final_avg_q[index,], yticklabels=labels, cmap="Blues")
...: plt.xlabel('Actions')
...: plt.ylabel('States')
...: plt.savefig('fig1.png', dpi=600)

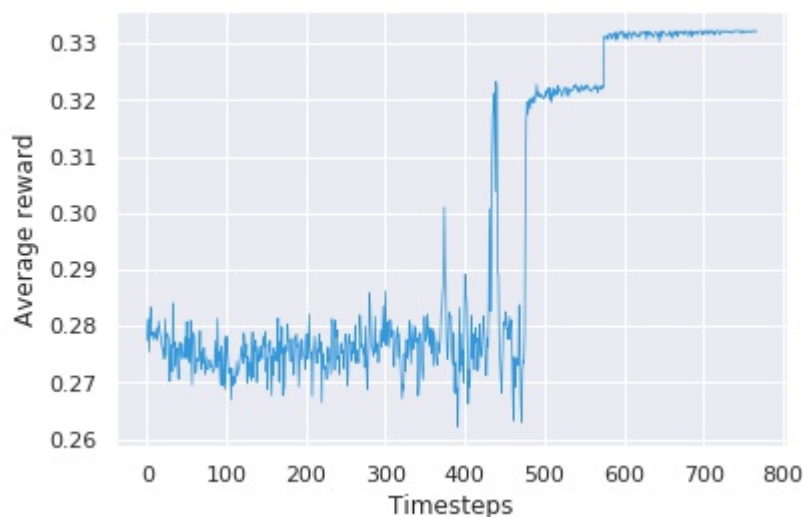
```

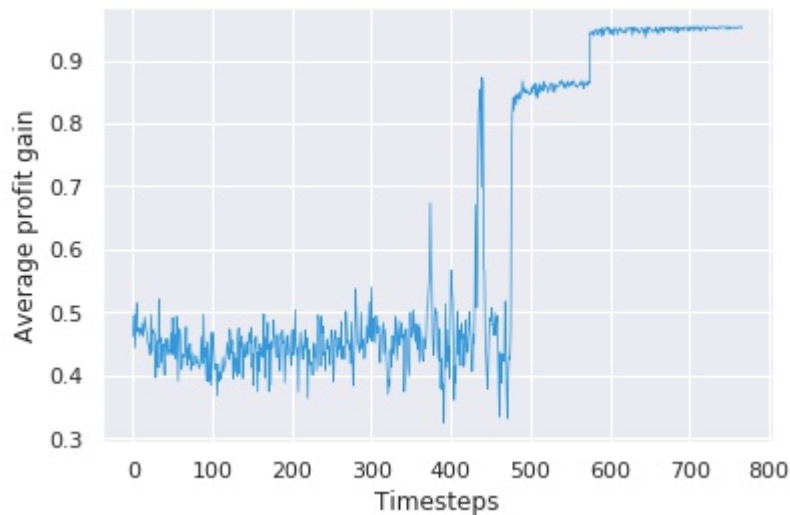
```
...: plt.show()
```



```
In [75]: i = np.random.randint(NUM_EPISODES)
...: avg_delta_0 = (np.array(avg_rew_0[i]) - 0.22589)/(0.337472 - 0.22589)
...: avg_delta_1 = (np.array(avg_rew_1[i]) - 0.22589)/(0.337472 - 0.22589)
...:
...: print(f"Showing training progress for Episode {i} of {NUM_EPISODES}")
...: plt.plot(avg_rew_0[i], label='agent_0', c="#3498db", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average reward')
...: plt.savefig('fig2.png', dpi=600)
...: plt.show()
...:
...: plt.plot(avg_delta_0, label='agent_0', c="#3498db", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average profit gain')
...: plt.savefig('fig3.png', dpi=600)
...: plt.show()
```

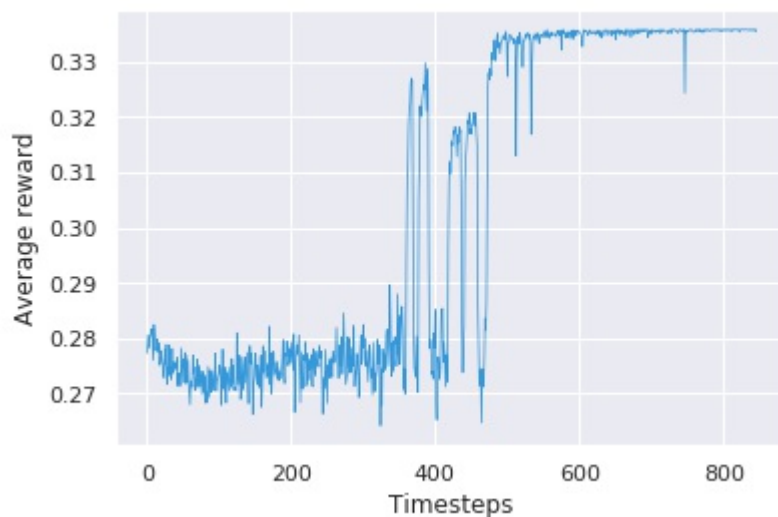
Showing training progress for Episode 274 of 300

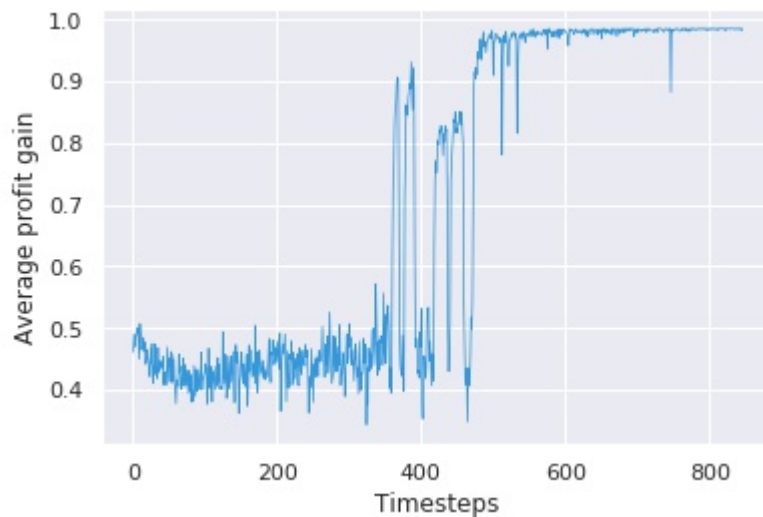




```
In [76]: i = np.random.randint(NUM_EPISODES)
...: avg_delta_0 = (np.array(avg_rew_0[i]) - 0.22589)/(0.337472 - 0.22589)
...: avg_delta_1 = (np.array(avg_rew_1[i]) - 0.22589)/(0.337472 - 0.22589)
...:
...: print(f"Showing training progress for Episode {i} of {NUM_EPISODES}")
...: plt.plot(avg_rew_0[i], label='agent_0', c="#3498db", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average reward')
...: plt.savefig(f'fig2-{i}.png', dpi=600)
...: plt.show()
...:
...: plt.plot(avg_delta_0, label='agent_0', c="#3498db", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average profit gain')
...: plt.savefig(f'fig3-{i}.png', dpi=600)
...: plt.show()
```

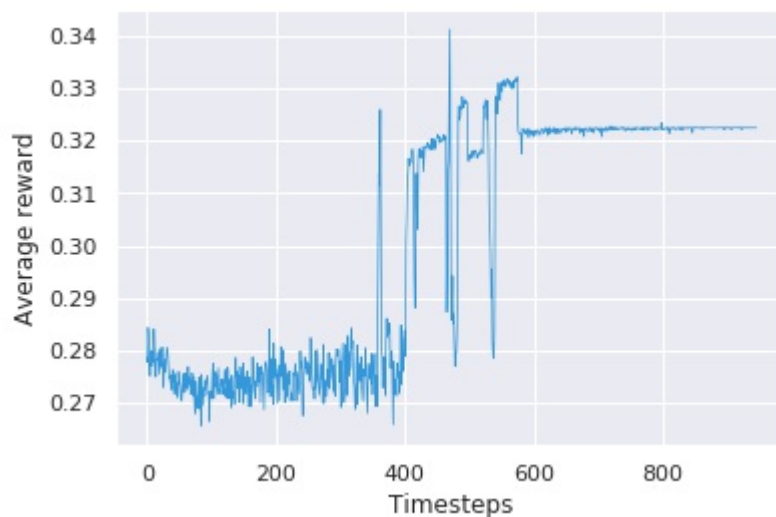
Showing training progress for Episode 78 of 300

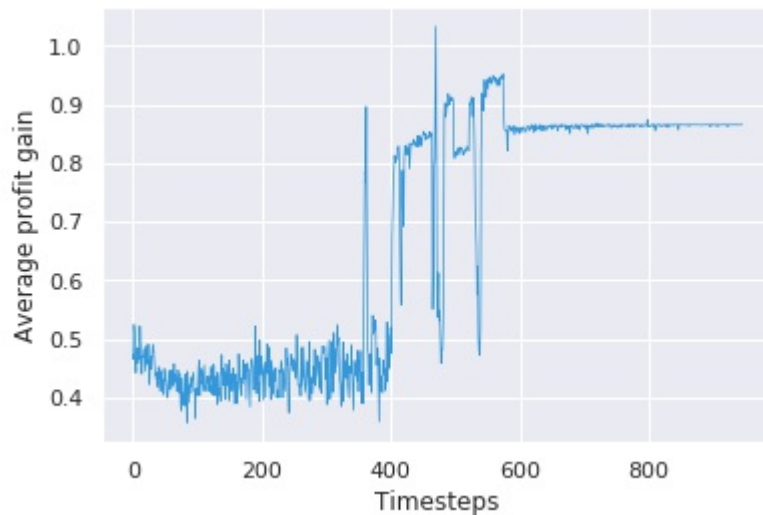




```
In [77]: i = np.random.randint(NUM_EPISODES)
...: avg_delta_0 = (np.array(avg_rew_0[i]) - 0.22589)/(0.337472 - 0.22589)
...: avg_delta_1 = (np.array(avg_rew_1[i]) - 0.22589)/(0.337472 - 0.22589)
...:
...: print(f"Showing training progress for Episode {i} of {NUM_EPISODES}")
...: plt.plot(avg_rew_0[i], label='agent_0', c="#3498db", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average reward')
...: plt.savefig(f'fig2-{i}.png', dpi=600)
...: plt.show()
...:
...: plt.plot(avg_delta_0, label='agent_0', c="#3498db", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average profit gain')
...: plt.savefig(f'fig3-{i}.png', dpi=600)
...: plt.show()
```

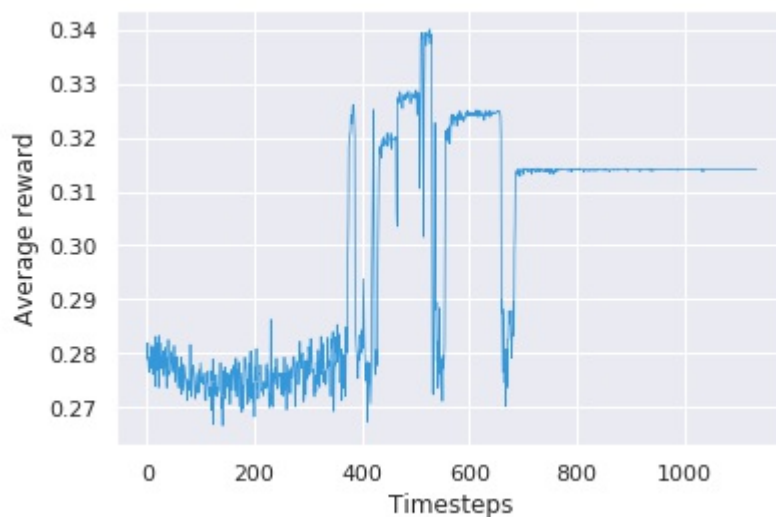
Showing training progress for Episode 33 of 300

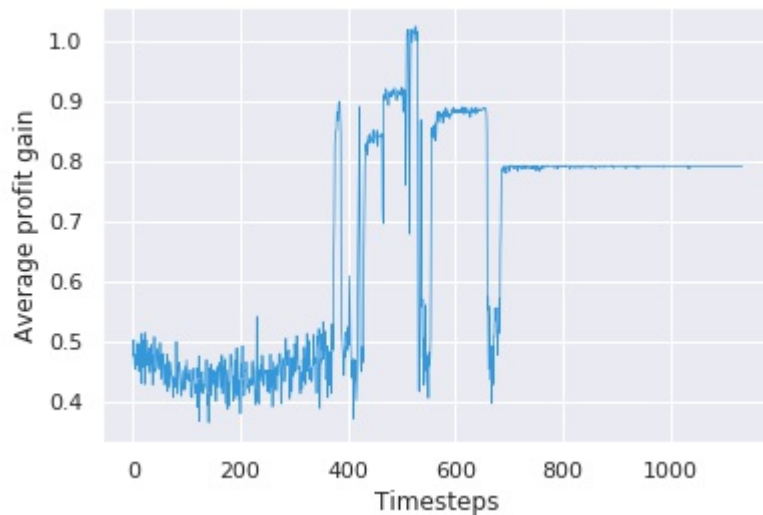




```
In [78]: i = np.random.randint(NUM_EPISODES)
...: avg_delta_0 = (np.array(avg_rew_0[i]) - 0.22589)/(0.337472 - 0.22589)
...: avg_delta_1 = (np.array(avg_rew_1[i]) - 0.22589)/(0.337472 - 0.22589)
...:
...: print(f"Showing training progress for Episode {i} of {NUM_EPISODES}")
...: plt.plot(avg_rew_0[i], label='agent_0', c="#3498db", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average reward')
...: plt.savefig(f'fig2-{i}.png', dpi=600)
...: plt.show()
...:
...: plt.plot(avg_delta_0, label='agent_0', c="#3498db", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average profit gain')
...: plt.savefig(f'fig3-{i}.png', dpi=600)
...: plt.show()
```

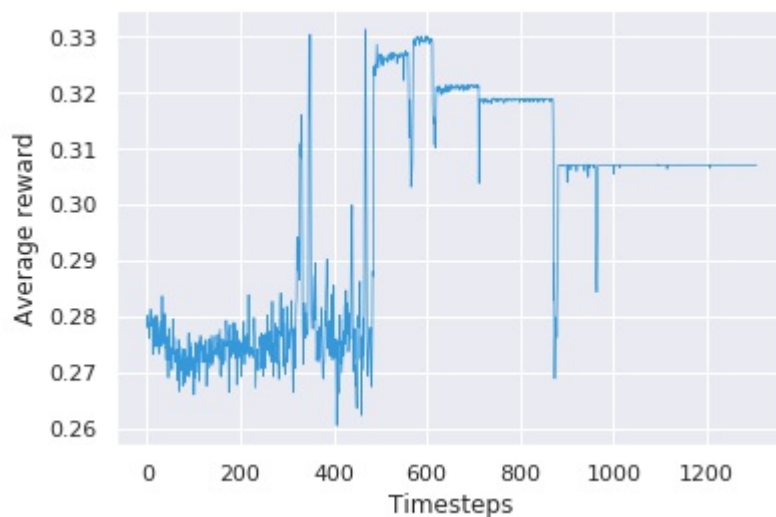
Showing training progress for Episode 82 of 300

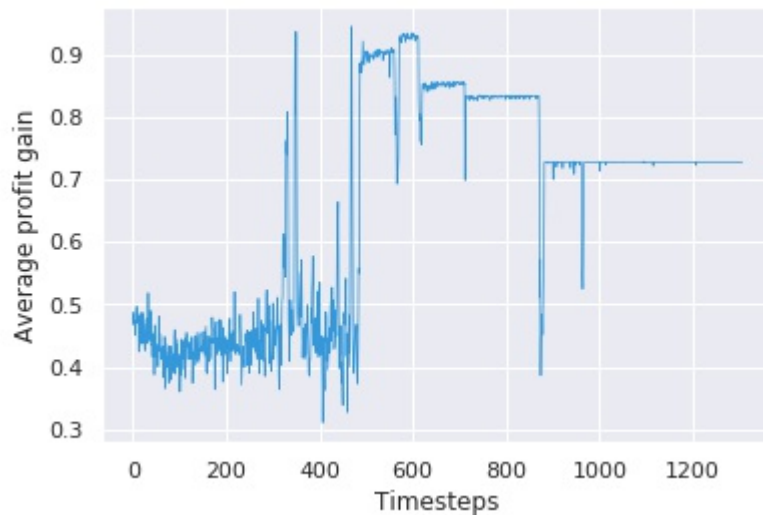




```
In [79]: i = np.random.randint(NUM_EPISODES)
...: avg_delta_0 = (np.array(avg_rew_0[i]) - 0.22589)/(0.337472 - 0.22589)
...: avg_delta_1 = (np.array(avg_rew_1[i]) - 0.22589)/(0.337472 - 0.22589)
...:
...: print(f"Showing training progress for Episode {i} of {NUM_EPISODES}")
...: plt.plot(avg_rew_0[i], label='agent_0', c="#3498db", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average reward')
...: plt.savefig(f'fig2-{i}.png', dpi=600)
...: plt.show()
...:
...: plt.plot(avg_delta_0, label='agent_0', c="#3498db", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average profit gain')
...: plt.savefig(f'fig3-{i}.png', dpi=600)
...: plt.show()
```

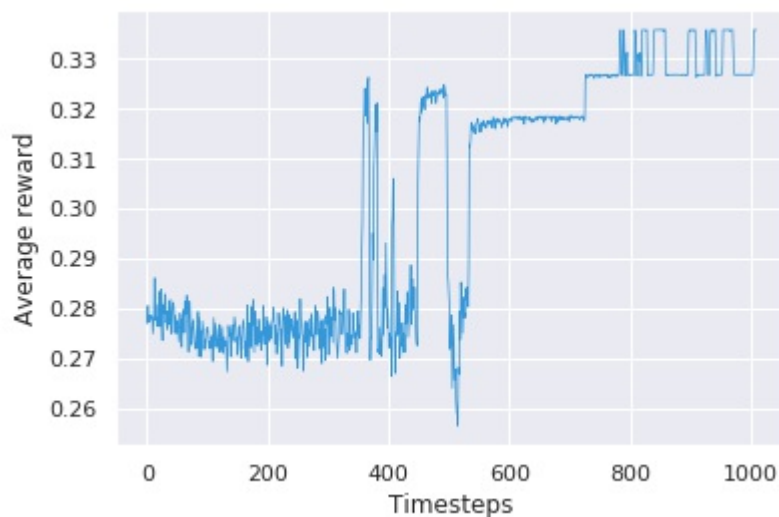
Showing training progress for Episode 195 of 300

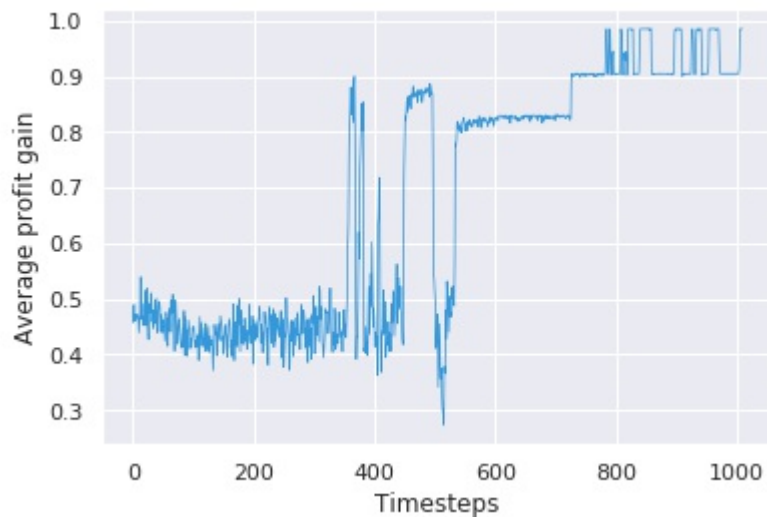




```
In [80]: i = np.random.randint(NUM_EPISODES)
...: avg_delta_0 = (np.array(avg_rew_0[i]) - 0.22589)/(0.337472 - 0.22589)
...: avg_delta_1 = (np.array(avg_rew_1[i]) - 0.22589)/(0.337472 - 0.22589)
...:
...: print(f"Showing training progress for Episode {i} of {NUM_EPISODES}")
...: plt.plot(avg_rew_0[i], label='agent_0', c="#3498db", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average reward')
...: plt.savefig(f'fig2-{i}.png', dpi=600)
...: plt.show()
...:
...: plt.plot(avg_delta_0, label='agent_0', c="#3498db", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average profit gain')
...: plt.savefig(f'fig3-{i}.png', dpi=600)
...: plt.show()
```

Showing training progress for Episode 28 of 300





```
In [81]: steps_at_conv_df
```

```
Out[81]:
```

```
steps_at_conv
0      1240266
1      865005
2      793681
3      992148
4      953446
..      ...
295     814383
296     811064
297     1051875
298     816400
299     916347
```

```
[300 rows x 1 columns]
```

```
In [82]: steps_at_conv_df.mean
```

```
Out[82]:
```

```
<bound method DataFrame.mean of      steps_at_conv
0      1240266
1      865005
2      793681
3      992148
4      953446
..      ...
295     814383
296     811064
297     1051875
298     816400
299     916347
```

```
[300 rows x 1 columns]>
```

```
In [83]: steps_at_conv_df.mean()
```

```
Out[83]:
```

```
steps_at_conv    1.050035e+06
dtype: float64
```

```
In [84]: steps_at_conv_df.describe()
```

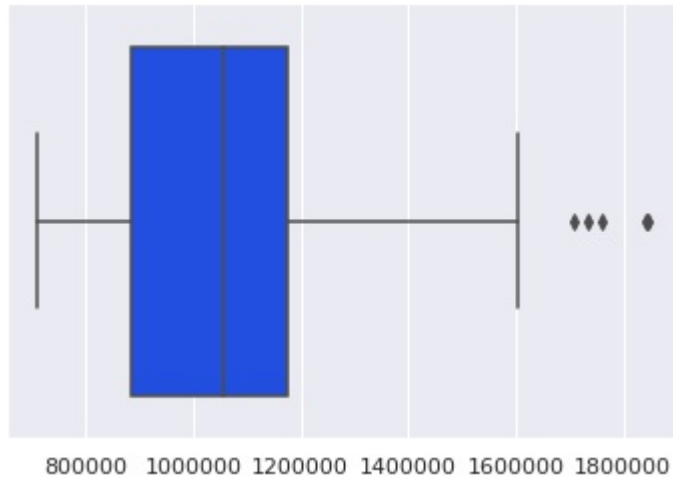
```
Out[84]:
```

```
steps_at_conv
count    3.000000e+02
mean     1.050035e+06
```

```
std      2.072801e+05
min      7.108360e+05
25%      8.842405e+05
50%      1.056762e+06
75%      1.173472e+06
max      1.844669e+06
```

```
In [85]: sns.boxplot(steps_at_conv_df)
```

```
Out[85]: <matplotlib.axes._subplots.AxesSubplot at 0x7f4eec777358>
```



```
In [86]: sns.boxplot(steps_at_conv_df, c="#3498db")
```

```
Traceback (most recent call last):
```

```
File "<ipython-input-86-aa04602aa7c1>", line 1, in <module>
    sns.boxplot(steps_at_conv_df, c="#3498db")
```

```
File "/home/lorenzo/anaconda3/envs/tf-rllib-2/lib/python3.6/site-packages/seaborn/
categorical.py", line 2237, in boxplot
    plotter.plot(ax, kwargs)
```

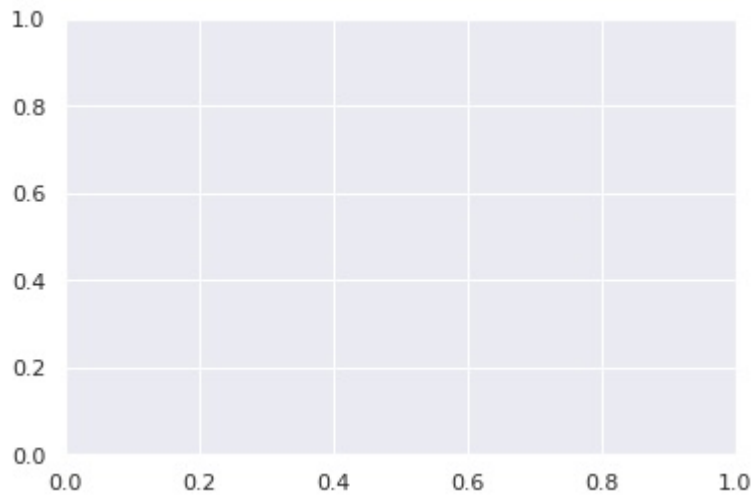
```
File "/home/lorenzo/anaconda3/envs/tf-rllib-2/lib/python3.6/site-packages/seaborn/
categorical.py", line 549, in plot
    self.draw_boxplot(ax, boxplot_kws)
```

```
File "/home/lorenzo/anaconda3/envs/tf-rllib-2/lib/python3.6/site-packages/seaborn/
categorical.py", line 486, in draw_boxplot
    **kws)
```

```
File "/home/lorenzo/anaconda3/envs/tf-rllib-2/lib/python3.6/site-packages/
matplotlib/cbook/deprecation.py", line 307, in wrapper
    return func(*args, **kwargs)
```

```
File "/home/lorenzo/anaconda3/envs/tf-rllib-2/lib/python3.6/site-packages/
matplotlib/_init_.py", line 1589, in inner
    return func(ax, *map(sanitize_sequence, args), **kwargs)
```

```
TypeError: boxplot() got an unexpected keyword argument 'c'
```



In [87]:

```
In [87]: steps_at_conv_df.describe()
...: sns.boxplot(steps_at_conv_df, cmap="#3498db")
Traceback (most recent call last):
```

```
File "<ipython-input-87-575c51862508>", line 2, in <module>
    sns.boxplot(steps_at_conv_df, cmap="#3498db")
```

```
File "/home/lorenzo/anaconda3/envs/tf-rllib-2/lib/python3.6/site-packages/seaborn/
categorical.py", line 2237, in boxplot
    plotter.plot(ax, kwargs)
```

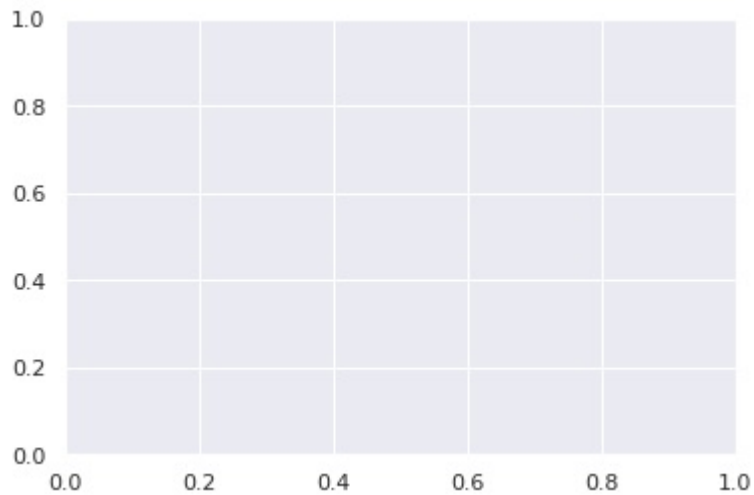
```
File "/home/lorenzo/anaconda3/envs/tf-rllib-2/lib/python3.6/site-packages/seaborn/
categorical.py", line 549, in plot
    self.draw_boxplot(ax, boxplot_kws)
```

```
File "/home/lorenzo/anaconda3/envs/tf-rllib-2/lib/python3.6/site-packages/seaborn/
categorical.py", line 486, in draw_boxplot
    **kws)
```

```
File "/home/lorenzo/anaconda3/envs/tf-rllib-2/lib/python3.6/site-packages/
matplotlib/cbook/deprecation.py", line 307, in wrapper
    return func(*args, **kwargs)
```

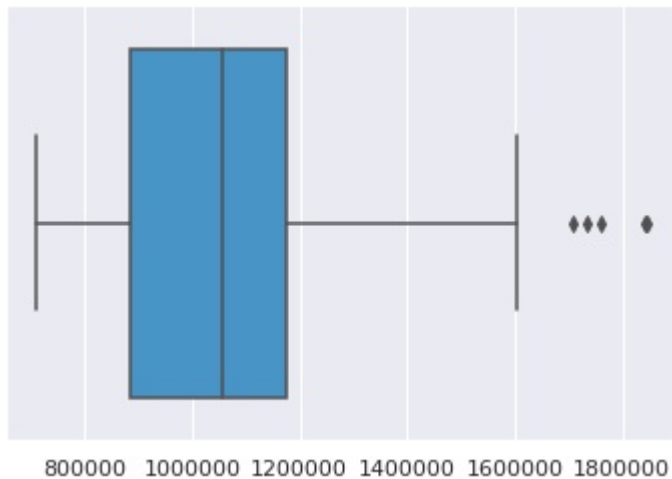
```
File "/home/lorenzo/anaconda3/envs/tf-rllib-2/lib/python3.6/site-packages/
matplotlib/_init_.py", line 1589, in inner
    return func(ax, *map(sanitize_sequence, args), **kwargs)
```

```
TypeError: boxplot() got an unexpected keyword argument 'cmap'
```



In [88]:

```
In [88]: steps_at_conv_df.describe()
...: sns.boxplot(steps_at_conv_df, color="#3498db")
Out[88]: <matplotlib.axes._subplots.AxesSubplot at 0x7f4eec575358>
```

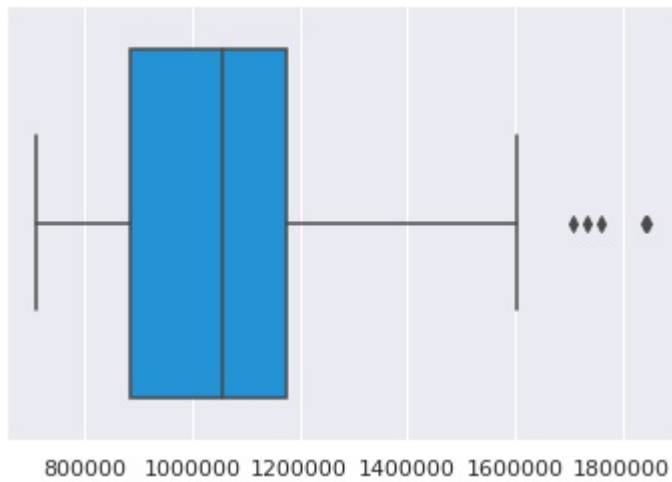


```
In [89]: sns.boxplot(steps_at_conv_df, color="#069af3"
File "<ipython-input-89-df92309f58ae>", line 1
    sns.boxplot(steps_at_conv_df, color="#069af3"
    ^
```

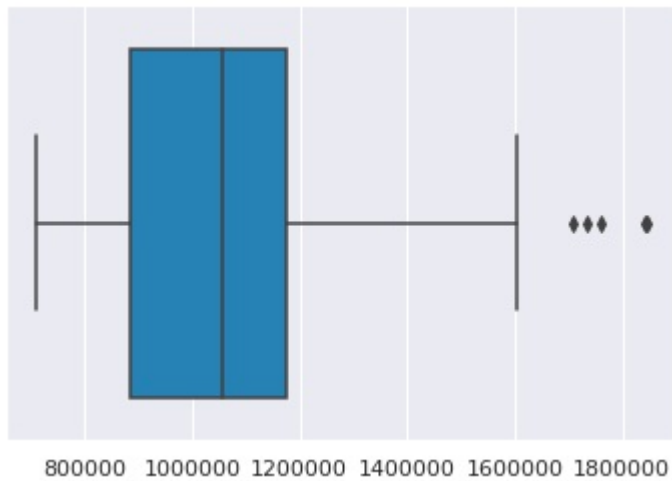
SyntaxError: unexpected EOF while parsing

In [90]:

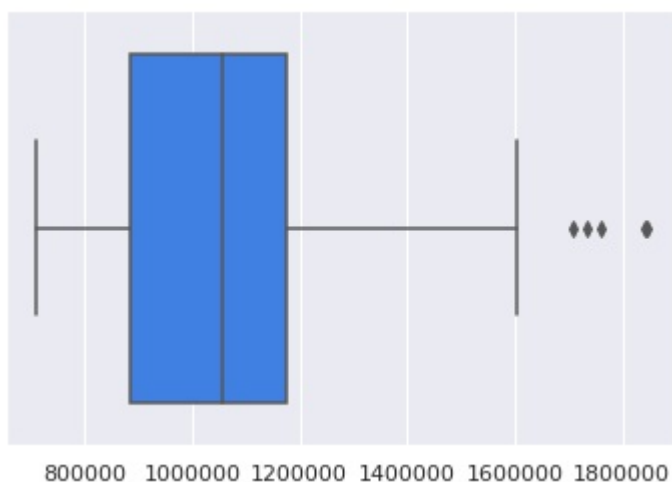
```
In [90]: sns.boxplot(steps_at_conv_df, color="#069af3")
Out[90]: <matplotlib.axes._subplots.AxesSubplot at 0x7f4eec4e4cc0>
```



```
In [91]: sns.boxplot(steps_at_conv_df, color="#0e87cc")
Out[91]: <matplotlib.axes._subplots.AxesSubplot at 0x7f4eec84fb70>
```



```
In [92]: sns.boxplot(steps_at_conv_df, color="#247afd")
Out[92]: <matplotlib.axes._subplots.AxesSubplot at 0x7f4eec424a20>
```



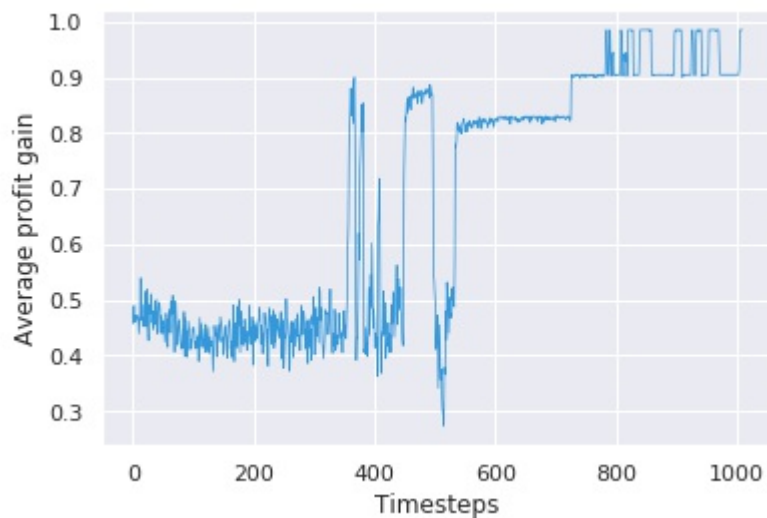
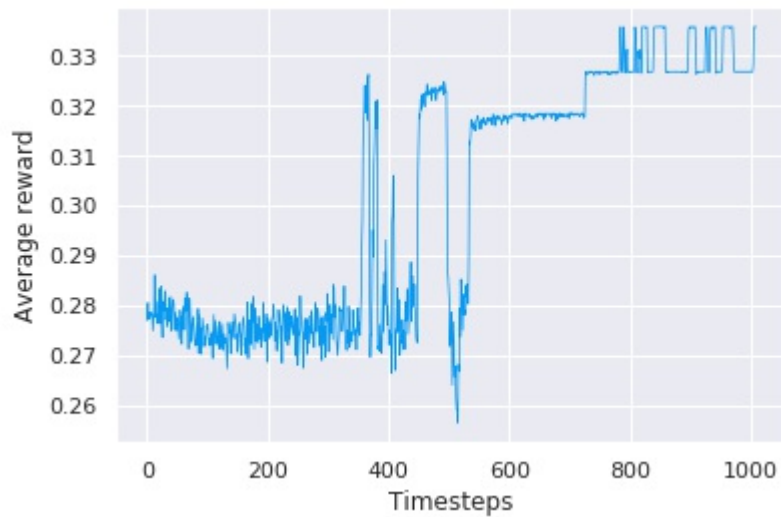
```
In [93]: print(f"Showing training progress for Episode {i} of {NUM_EPISODES}")
...: plt.plot(avg_rew_0[i], label='agent_0', c="#069af3", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average reward')
...: plt.savefig(f'fig2-{i}.png', dpi=600)
```

```

...: plt.show()
...:
...: plt.plot(avg_delta_0, label='agent_0', c="#3498db", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average profit gain')
...: plt.savefig(f'fig3-{i}.png', dpi=600)
...: plt.show()

```

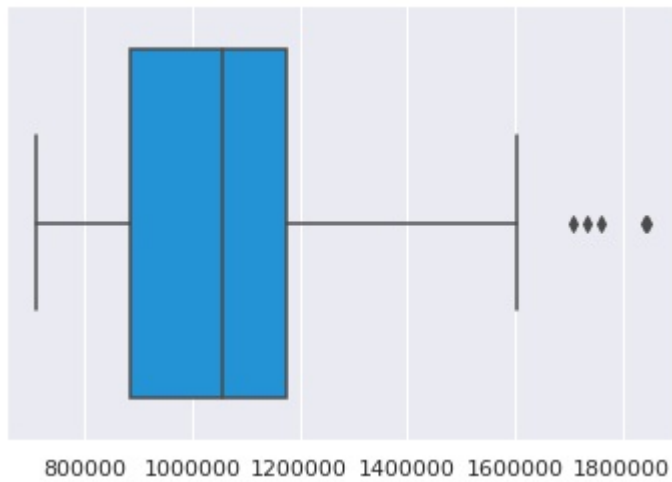
Showing training progress for Episode 28 of 300



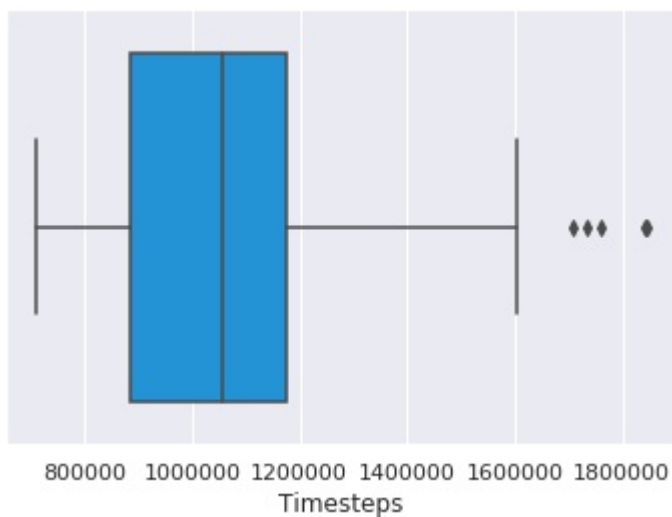
```

In [94]: steps_at_conv_df.describe()
...: sns.boxplot(steps_at_conv_df, color="#069af3")
Out[94]: <matplotlib.axes._subplots.AxesSubplot at 0x7f4eec34ef28>

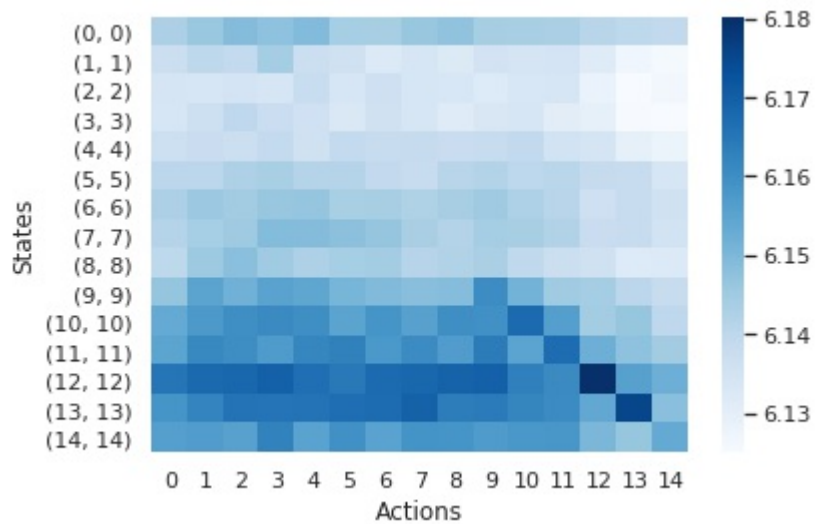
```



```
In [95]: steps_at_conv_df.describe()
...: sns.boxplot(steps_at_conv_df, color="#069af3")
...: plt.xlabel('Timesteps')
...: plt.savefig('ts_boxplot.png', dpi=600)
...: plt.show()
```

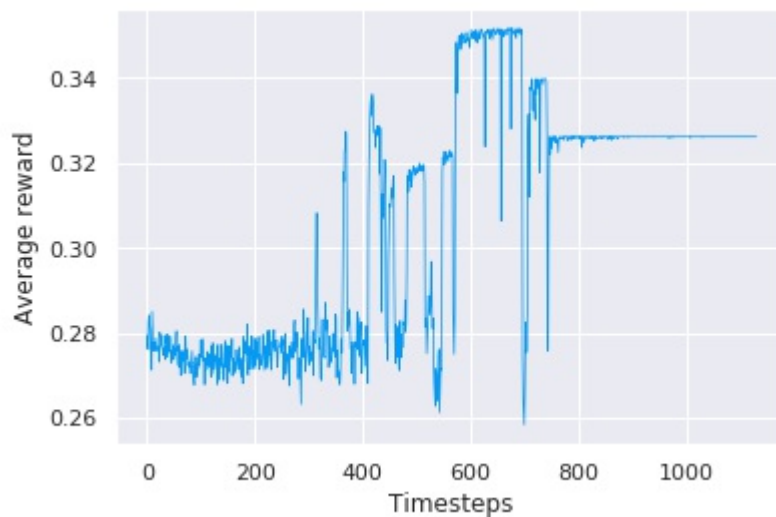


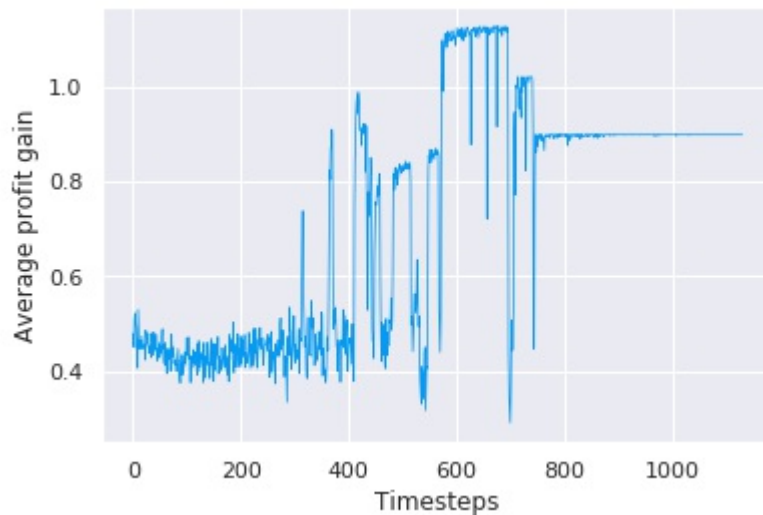
```
In [96]: index = []
...: for i in range(15):
...:     for j in range(15):
...:         if i==j:
...:             index.append(s_dict[(i,j)])
...:
...: labels = []
...: for i in index:
...:     labels.append(list(s_dict.keys())[i])
...:
...:
...: sns.heatmap(final_avg_q[index,], yticklabels=labels, cmap="Blues")
...: plt.xlabel('Actions')
...: plt.ylabel('States')
...: plt.savefig('fig1.png', dpi=600)
...: plt.show()
```



```
In [97]: i = np.random.randint(NUM_EPISODES)
...: avg_delta_0 = (np.array(avg_rew_0[i]) - 0.22589)/(0.337472 - 0.22589)
...: avg_delta_1 = (np.array(avg_rew_1[i]) - 0.22589)/(0.337472 - 0.22589)
...:
...: print(f"Showing training progress for Episode {i} of {NUM_EPISODES}")
...: plt.plot(avg_rew_0[i], label='agent_0', c="#069af3", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average reward')
...: plt.savefig(f'fig2-{i}.png', dpi=600)
...: plt.show()
...:
...: plt.plot(avg_delta_0, label='agent_0', c="#069af3", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average profit gain')
...: plt.savefig(f'fig3-{i}.png', dpi=600)
...: plt.show()
```

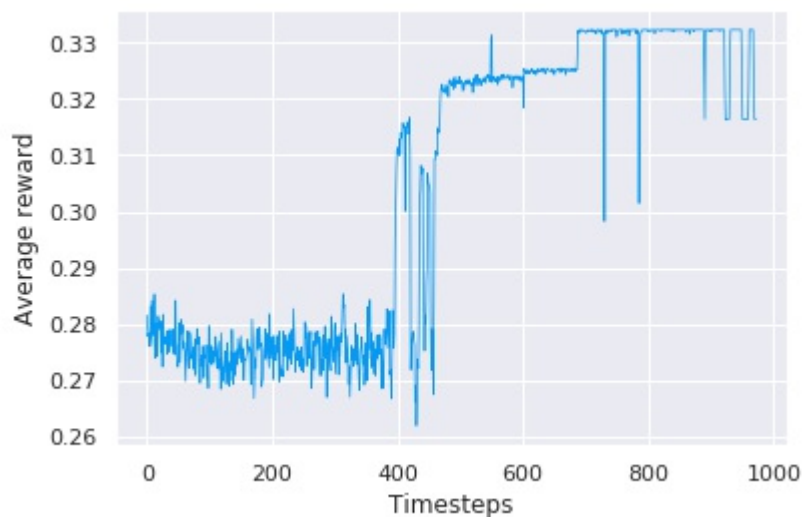
Showing training progress for Episode 146 of 300

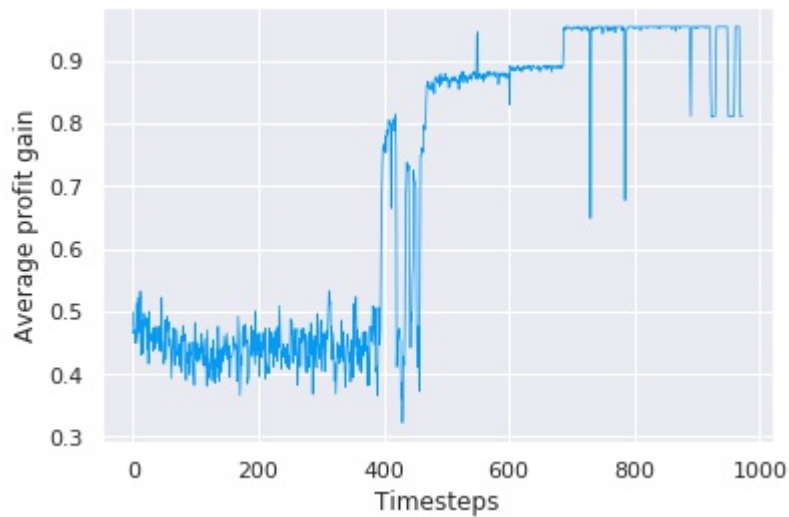




```
In [98]: i = np.random.randint(NUM_EPISODES)
...: avg_delta_0 = (np.array(avg_rew_0[i]) - 0.22589)/(0.337472 - 0.22589)
...: avg_delta_1 = (np.array(avg_rew_1[i]) - 0.22589)/(0.337472 - 0.22589)
...:
...: print(f"Showing training progress for Episode {i} of {NUM_EPISODES}")
...: plt.plot(avg_rew_0[i], label='agent_0', c="#069af3", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average reward')
...: plt.savefig(f'fig2-{i}.png', dpi=600)
...: plt.show()
...:
...: plt.plot(avg_delta_0, label='agent_0', c="#069af3", lw=0.7)
...: plt.xlabel('Timesteps')
...: plt.ylabel('Average profit gain')
...: plt.savefig(f'fig3-{i}.png', dpi=600)
...: plt.show()
```

Showing training progress for Episode 20 of 300





```
In [99]: steps_at_conv_df.describe(
File "<ipython-input-99-65ca54ba1238>", line 1
      steps_at_conv_df.describe(
                                ^
```

SyntaxError: unexpected EOF while parsing

In [100]:

```
In [100]: steps_at_conv_df.describe()
```

```
Out[100]:
      steps_at_conv
count  3.000000e+02
mean   1.050035e+06
std    2.072801e+05
min    7.108360e+05
25%    8.842405e+05
50%    1.056762e+06
75%    1.173472e+06
max    1.844669e+06
```

In [101]: