In [1]:
```
import pandas as pd
import scipy.stats as st
import numpy as np
```

In [2]:
```
df = pd.read_csv('experiment_result.csv')
df.sample(5)
```

Out[2]:

	n	num_step_taken(h_misplaced)	num_step_taken(h_manhattan)	num_step_taken(h_misplaced-h_manhattan)
82	4	7	7	0
41	4	8	8	0
78	4	11	11	0
56	4	8	8	0
49	4	7	7	0

In [3]:
```
df.mean()
```

Out[3]:
```
n                                       4.00
num_step_taken(h_misplaced)             7.56
num_step_taken(h_manhattan)             6.78
num_step_taken(h_misplaced-h_manhattan)    0.78
dtype: float64
```

In [4]:
```
manhattan = np.array(df['num_step_taken(h_manhattan)'])
misplaced = np.array(df['num_step_taken(h_misplaced)'])
```

In [5]:
```
t_test = st.ttest_ind(a=misplaced,b = manhattan)
```

In [6]:
```
t_test.statistic
```

Out[6]:
```
1.0086922470833157
```

In [7]:
```
t_test.pvalue
```

Out[7]:
```
0.3143533348184606
```

Result: Total Manhattan Distance is better than Number of Misplaced Tiles