In [1]:
```
import pandas as pd
from matplotlib import pyplot as plt
%matplotlib inline
import numpy as np
import seaborn as sns
from scipy import stats
```

Q1 Answer

H0 : There is no significant difference in diameter

H1 : There is a significant difference in diameter

In [2]:
```
cutlets_data = pd.read_csv('Cutlets.csv')
cutlets_data
```

Out[2]:

	Unit A	Unit B
0	6.8090	6.7703
1	6.4376	7.5093
2	6.9157	6.7300
3	7.3012	6.7878
4	7.4488	7.1522
5	7.3871	6.8110
6	6.8755	7.2212
7	7.0621	6.6606
8	6.6840	7.2402
9	6.8236	7.0503
10	7.3930	6.8810
11	7.5169	7.4059
12	6.9246	6.7652
13	6.9256	6.0380
14	6.5797	7.1581
15	6.8394	7.0240
16	6.5970	6.6672
17	7.2705	7.4314
18	7.2828	7.3070
19	7.3495	6.7478
20	6.9438	6.8889
21	7.1560	7.4220

	Unit A	Unit B
22	6.5341	6.5217
23	7.2854	7.1688
24	6.9952	6.7594
25	6.8568	6.9399
26	7.2163	7.0133
27	6.6801	6.9182
28	6.9431	6.3346
29	7.0852	7.5459
30	6.7794	7.0992
31	7.2783	7.1180
32	7.1561	6.6965
33	7.3943	6.5780
34	6.9405	7.3875

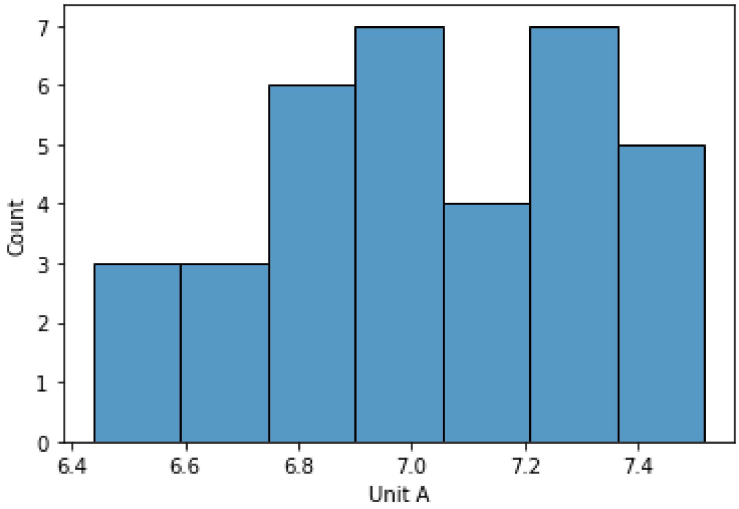
In [5]:
```
cutlets_data['Unit A'].mean()
```

Out[5]:
```
7.01909142857143
```

In [6]:
```
cutlets_data['Unit A'].std()
```

Out[6]:
```
0.2884084841815496
```

In [9]:
```
sns.histplot(cutlets_data['Unit A'])
plt.show()
```



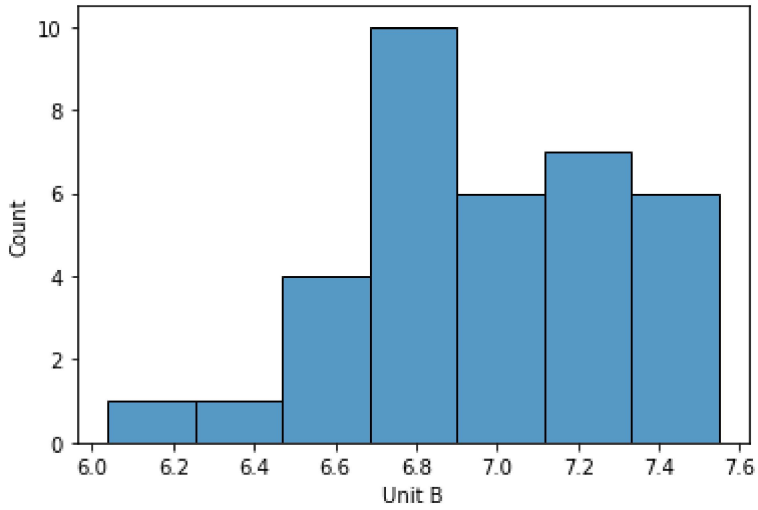
In [10]:
```
cutlets_data['Unit B'].mean()
```

Out[10]: 6.964297142857142

In [11]:
```
cutlets_data['Unit B'].std()
```

Out[11]: 0.343400647063108

In [12]:
```
sns.histplot(cutlets_data['Unit B'])
plt.show()
```



In [26]:
```
t_value,p_value=stats.ttest_ind(cutlets_data['Unit A'],cutlets_data['Unit B'])
```

In [27]:
```
t_value
```

Out[27]: 0.7228688704678063

In [28]:
```
p_value
```

Out[28]: 0.4722394724599501

In [34]:
```
    if p_value<0.05:
        print('we reject the null hypothesis')
    else:
        print('we would not reject the null hypothesis')
```

```
we would not reject the null hypothesis
```

ie, there is no significant difference in diameter

Q2 Answer

H0 : There is no difference in the average TAT

H1: There is an difference in the average TAT

In [35]:
```
lab_data = pd.read_csv('LabTAT.csv')
lab_data
```

Out[35]:

	Laboratory 1	Laboratory 2	Laboratory 3	Laboratory 4
0	185.35	165.53	176.70	166.13
1	170.49	185.91	198.45	160.79
2	192.77	194.92	201.23	185.18
3	177.33	183.00	199.61	176.42
4	193.41	169.57	204.63	152.60
...	...	...	...	...
115	178.49	170.66	193.80	172.68
116	176.08	183.98	215.25	177.64
117	202.48	174.54	203.99	170.27
118	182.40	197.18	194.52	150.87
119	182.09	215.17	221.49	162.21

120 rows × 4 columns

In [40]:
```
lab_data['Laboratory 1'].mean() ,lab_data['Laboratory 1'].std()
```

Out[40]: (178.36158333333339, 13.173593589458577)

In [41]:
```
lab_data['Laboratory 2'].mean() ,lab_data['Laboratory 2'].std()
```

Out[41]: (178.9029166666668, 14.957113728438193)

In [42]:
```
lab_data['Laboratory 3'].mean() ,lab_data['Laboratory 3'].std()
```

Out[42]: (199.91325000000003, 16.539032831329596)

In [45]:
```
lab_data['Laboratory 4'].mean() ,lab_data['Laboratory 4'].std()
```

Out[45]: (163.6827499999999, 15.08507977671142)

In [53]:
```
statistic,p_value = stats.f_oneway(lab_data['Laboratory 1'],lab_data['Laboratory 2'],la
```

In [54]:
```
p_value
```

Out[54]: 2.1156708949992414e-57

In [55]:
```
if p_value < 0.05 :
```

```
    print('We reject null hypothesis')
else:
    print('We do not reject null hypothesis')
```

```
We reject null hypothesis
```

ie, There is a difference in the average TAT of the labortaries

Q3 Answer

H0 : All proportions are equal

H1 : Not all proportions are equal

In [56]:
```
buyer_data = pd.read_csv('BuyerRatio.csv')
buyer_data
```

Out[56]:

	Observed Values	East	West	North	South
0	Males	50	142	131	70
1	Females	435	1523	1356	750

In [57]:
```
del buyer_data['Observed Values']
```

In [58]:
```
chi2,p_value,dof,expected = stats.chi2_contingency(buyer_data)
```

In [59]:
```
p_value
```

Out[59]:
```
0.6603094907091882
```

In [60]:
```
if p_value < 0.05 :
    print('We reject null hypothesis')
else:
    print('We do not reject null hypothesis')
```

```
We do not reject null hypothesis
```

ie , All proportions are same

Q4 Answer

H0 : There is no significant change

H1 : There is an significant change

In [78]:
```
customer_data = pd.read_csv('Costomer+OrderForm.csv')
customer_data
```

Out[78]:

	Phillippines	Indonesia	Malta	India
0	Error Free	Error Free	Defective	Error Free
1	Error Free	Error Free	Error Free	Defective
2	Error Free	Defective	Defective	Error Free
3	Error Free	Error Free	Error Free	Error Free
4	Error Free	Error Free	Defective	Error Free
...	...	...	...	...
295	Error Free	Error Free	Error Free	Error Free
296	Error Free	Error Free	Error Free	Error Free
297	Error Free	Error Free	Defective	Error Free
298	Error Free	Error Free	Error Free	Error Free
299	Error Free	Defective	Defective	Error Free

300 rows × 4 columns

In [79]:
```
customer_data.replace({'Error Free': 0 , 'Defective' :1},inplace=True)
```

In [80]:
```
customer_data
```

Out[80]:

	Phillippines	Indonesia	Malta	India
0	0	0	1	0
1	0	0	0	1
2	0	1	1	0
3	0	0	0	0
4	0	0	1	0
...	...	...	...	...
295	0	0	0	0
296	0	0	0	0
297	0	0	1	0
298	0	0	0	0
299	0	1	1	0

300 rows × 4 columns

In [81]:
```
customer_data['Phillippines'].value_counts()
```

Out[81]:
```
0    271
1     29
```

```
Name: Phillippines, dtype: int64
```

In [82]:
```
customer_data['Indonesia'].value_counts()
```

Out[82]:
```
0    267
1     33
Name: Indonesia, dtype: int64
```

In [83]:
```
customer_data['Malta'].value_counts()
```

Out[83]:
```
0    269
1     31
Name: Malta, dtype: int64
```

In [84]:
```
customer_data['India'].value_counts()
```

Out[84]:
```
0    280
1     20
Name: India, dtype: int64
```

In [85]:
```
customer_counts = np.array([[271,267,269,280],[29,33,31,20]])
```

In [86]:
```
customer_counts
```

Out[86]:
```
array([[271, 267, 269, 280],
       [ 29,  33,  31,  20]])
```

In [89]:
```
chi2,p_value,dof,expected =stats.chi2_contingency(customer_counts)
```

In [90]:
```
p_value
```

Out[90]:
```
0.2771020991233135
```

In [91]:
```
if p_value < 0.05 :
    print('We reject the null hypothesis')
else:
    print('We do not rejct the null hypothesis')
```

```
We do not rejct the null hypothesis
```

ie, there is no significant change

In [ ]:
```

```