In [2]:
```
import pandas as pd
import numpy as np
from scipy import stats
import statsmodels.api as sm
import warnings
warnings.filterwarnings("ignore")
from PIL import ImageGrab
import matplotlib.pyplot as plt
import seaborn as sns
```

In [3]:
```
data=pd.read_csv('Cutlets.CSV')
data.head()
```

Out[3]:

	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

In [4]:
```
data.describe()
```

Out[4]:

	Unit A	Unit B
count	35.000000	35.000000
mean	7.019091	6.964297
std	0.288408	0.343401
min	6.437600	6.038000
25%	6.831500	6.753600
50%	6.943800	6.939900
75%	7.280550	7.195000
max	7.516900	7.545900

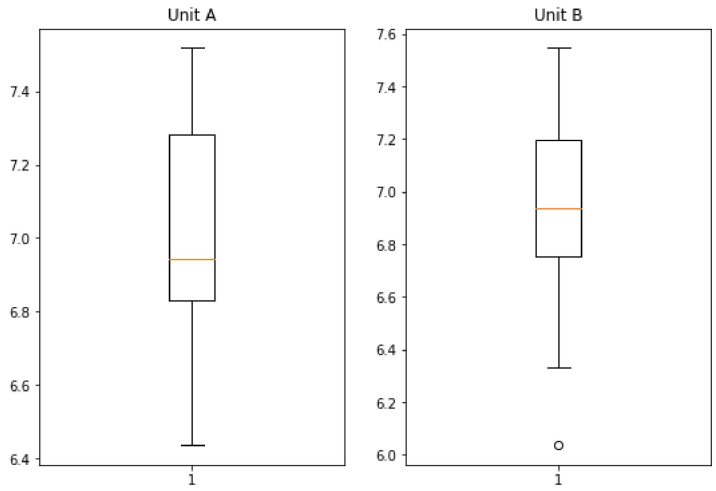
In [5]:
```
data.isnull().sum()
```

Out[5]:
```
Unit A    0
Unit B    0
dtype: int64
```

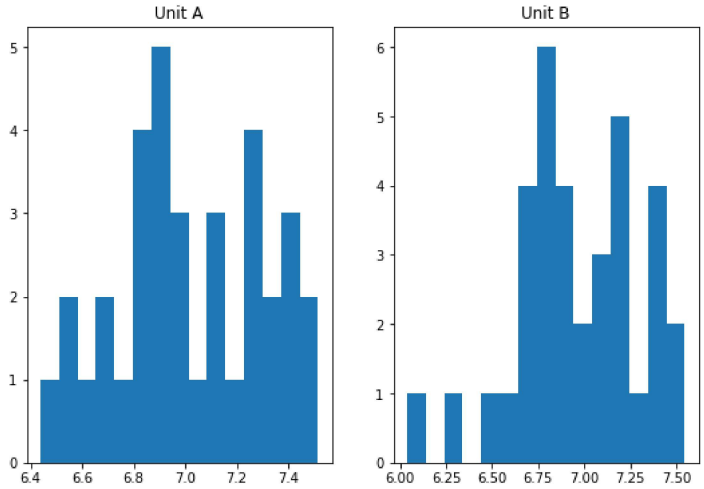
In [6]:
```
data[data.duplicated()].shape
```

Out[6]:
```
(0, 2)
```

In [7]:
```
plt.subplots(figsize = (9,6))
plt.subplot(121)
plt.boxplot(data['Unit A'])
plt.title('Unit A')
plt.subplot(122)
plt.boxplot(data['Unit B'])
plt.title('Unit B')
plt.show()
```

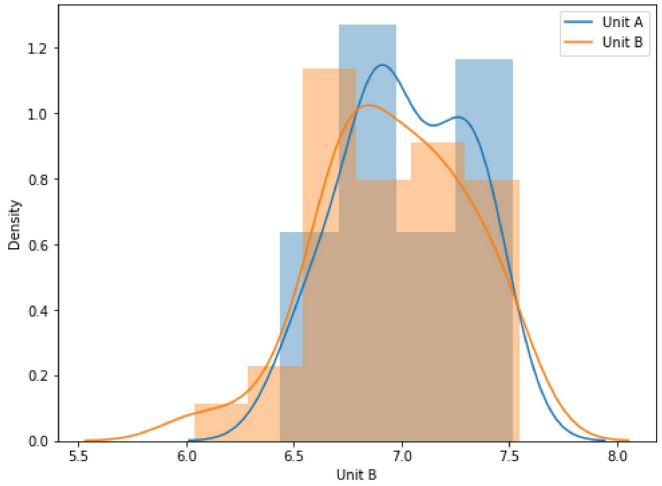


In [8]:
```
plt.subplots(figsize = (9,6))
plt.subplot(121)
plt.hist(data['Unit A'], bins = 15)
plt.title('Unit A')
plt.subplot(122)
plt.hist(data['Unit B'], bins = 15)
plt.title('Unit B')
plt.show()
```

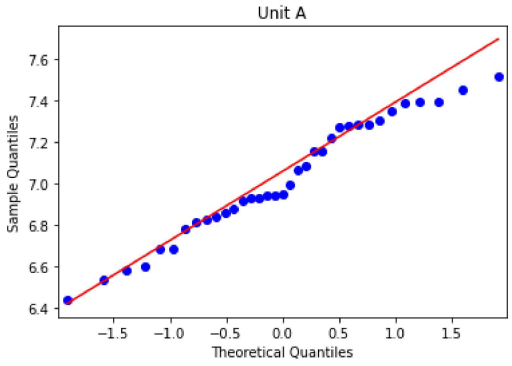


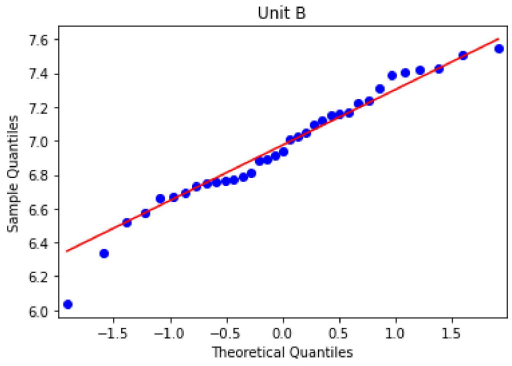
In [9]:
```
plt.figure(figsize = (8,6))
labels = ['Unit A', 'Unit B']
sns.distplot(data['Unit A'], kde = True)
sns.distplot(data['Unit B'],hist = True)
plt.legend(labels)
```

Out[9]: <matplotlib.legend.Legend at 0x1a4b4a091f0>



In [10]:
```
sm.qqplot(data["Unit A"], line = 'q')
plt.title('Unit A')
sm.qqplot(data["Unit B"], line = 'q')
plt.title('Unit B')
plt.show()
```





In [11]:
```
statistic , p_value = stats.ttest_ind(data['Unit A'],data['Unit B'], alternative = 'two-sided')
print('p_value=',p_value)
```

p_value= 0.4722394724599501

In [41]:
```
# compare p value with 'α '(Significane Level)
#If p_value is ≠ 'α ' we failed to reject Null Hypothesis because of lack of evidence
#If p_value is = 'α ' we reject Null Hypothesis
```

In [17]:
```
alpha = 0.025
print('Significnace=%.3f, p=%.3f' % (alpha, p_value))
if p_value <= alpha:
     print('We reject Null Hypothesis there is a significance difference between two Units A and B')
else:
    print('We fail to reject Null hypothesis')
```

```
Significnace=0.025, p=0.472
We fail to reject Null hypothesis
```

In [43]:
```
# we fail to reject hypothesis because of lack of evidence, there is no significant difference between the samples
```

In [ ]:
```

```