In [1]:

```
import pandas as pd
import numpy as np
from scipy import stats
from scipy.stats import norm
from scipy.stats import chi2_contingency
```

Assume Null Hypothesis as Ho: Independence of categorical variables (male-female buyer rations are similar across regions (does not vary and are not related) Thus Alternate Hypothesis as Ha: Dependence of categorical variables (male-female buyer rations are NOT similar across regions (does vary and somewhat/significantly related

In [2]:

```
data = pd.read_csv('BuyerRatio.csv')
data
```

Out[2]:

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

In [3]:

```
#make dimensional array
array=np.array([[50,142,131,70],[435,1523,1356,750]])
array
```

Out[3]:

```
array([[  50,  142,  131,   70],
       [ 435, 1523, 1356,  750]])
```

In [4]:

```
# Chi2 contengency independence test
chi2_contingency(array) #o/p is (chi2 stats value, p_value,df, expected observations)
```

Out[4]:

```
(1.595945538661058,
 0.6603094907091882,
 3,
 array([[  42.76531299,  146.81287862,  131.11756787,   72.30424052],
        [ 442.23468701, 1518.18712138, 1355.88243213,  747.69575948]]))
```

In [5]:

```
#compare p_value with α = 0.05
```

Inference: As (p-value = 0.6603) > (α = 0.05); Accept the Null Hypothesis i.e. Independence of categorical variables Thus, male-female buyer rations are similar across regions and are not related

In [ ]:

```
```