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
```

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]:
```
obs = np.array([[50,142,131,70],[435,1523,1356,750]])                #array conversion
obs
```

Out[3]:
```
array([[  50,  142,  131,   70],
       [ 435, 1523, 1356,  750]])
```

In [4]:
```
val = stats.chi2_contingency(obs)                                    #expected values ir
val
```

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]:
```
Expected_Values = val[3]
Expected_Values
```

Out[5]:
```
array([[  42.76531299,  146.81287862,  131.11756787,   72.30424052],
       [ 442.23468701, 1518.18712138, 1355.88243213,  747.69575948]])
```

In [6]:
```
p_value = val[0]
p_value
```

Out[6]:
```
1.595945538661058
```

In [7]:
```
if p_value <0.05:                                    # alpha value is 0.05 or 5%
    print(" we are rejecting null hypothesis")
else:
    print("we are accepting null hypothesis")
```

```
we are accepting null hypothesis
```

Since the p- value is (1.59) > 0.05, there is a significant difference in male-female ratio across regions