In [10]:
```
import pandas as pd
from scipy import stats as stats
import numpy as np
```

In [14]:
```
df= pd.read_csv("C:/Users/rjas/Downloads/ds assignments/3a/BuyerRatio.csv")
```

In [15]:
```
df.head()
```

Out[15]:

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

In [16]:
```
df_table=df.iloc[:,1:6]
df_table
```

Out[16]:

	East	West	North	South
0	50	142	131	70
1	435	1523	1356	750

In [17]:
```
df_table.values
```

Out[17]:
```
array([[  50,  142,  131,   70],
       [ 435, 1523, 1356,  750]], dtype=int64)
```

In [18]:
```
val=stats.chi2_contingency(df_table)
```

In [19]:
```
val
```

Out[19]:
```
(1.595945538661058,
 0.6603094907091882,
 3,
 array([[  42.76531299,  146.81287862,  131.11756787,   72.30424052],
        [ 442.23468701, 1518.18712138, 1355.88243213,  747.69575948]]))
```

In [20]:
```
type(val)
```

Out[20]:
```
tuple
```

In [21]:
```
no_of_rows=len(df_table.iloc[0:2,0])
no_of_columns=len(df_table.iloc[0,0:4])
degree_of_f=(no_of_rows-1)*(no_of_columns-1)
print('Degree of Freedom=',degree_of_f)
```

```
Degree of Freedom= 3
```

In [22]:
```
Expected_value=val[3]
```

In [23]:
```
Expected_value
```

Out[23]:
```
array([[  42.76531299,  146.81287862,  131.11756787,   72.30424052],
       [ 442.23468701, 1518.18712138, 1355.88243213,  747.69575948]])
```

In [24]:
```
from scipy.stats import chi2
chi_square=sum([(o-e)**2/e for o,e in zip(df_table.values,Expected_value)])
chi_square_statistic=chi_square[0]+chi_square[1]
chi_square_statistic
```

Out[24]:
```
1.5152956451130446
```

In [25]:
```
critical_value=chi2.ppf(0.95,3)
critical_value
```

Out[25]:
```
7.814727903251179
```

In [30]:
```
if chi_square_statistic >= critical_value:
    print('Dependent (reject H0)')
else:
    print('Independent (fail to reject H0)')
```

```
Independent (fail to reject H0)
```

In [31]:
```
pvalue=1-chi2.cdf(chi_square_statistic,3)
pvalue
```

Out[31]:
```
0.6787446296467897
```

In [32]:
```
if pvalue <= 0.05:
    print('Dependent (reject H0)')
else:
    print('Independent (fail to reject H0)')
```

```
Independent (fail to reject H0)
```

P-value is 0.678 > 0.05=>P,Accept Ho, hence Average are same As per results we can say that there is proportion of male and female buying is similar