In [1]:
```
import scipy.stats as stats
import statsmodels.api as sm
import numpy as np
import pandas as pd
import warnings
warnings.filterwarnings("ignore")
from PIL import ImageGrab
import matplotlib.pyplot as plt
import seaborn as sns
```

In [6]:
```
centers = pd.read_csv('Costomer+OrderForm.csv')
centers.head(10)
```

Out[6]:

	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
5	Error Free	Error Free	Error Free	Error Free
6	Error Free	Defective	Error Free	Error Free
7	Error Free	Error Free	Error Free	Error Free
8	Error Free	Error Free	Error Free	Error Free
9	Error Free	Error Free	Error Free	Error Free

In [7]:
```
centers.describe()
```

Out[7]:

	Phillippines	Indonesia	Malta	India
count	300	300	300	300
unique	2	2	2	2
top	Error Free	Error Free	Error Free	Error Free
freq	271	267	269	280

In [8]:
```
centers.isnull().sum()
```

Out[8]:
```
Phillippines    0
Indonesia       0
Malta           0
India           0
dtype: int64
```

In [9]:
```
centers[centers.isnull().any(axis=1)]
```

Out[9]:

	Phillippines	Indonesia	Malta	India

In [10]:
```
centers.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 300 entries, 0 to 299
Data columns (total 4 columns):
 #   Column        Non-Null Count  Dtype 
---  ------        --------------  ----- 
 0   Phillippines  300 non-null    object
 1   Indonesia     300 non-null    object
 2   Malta         300 non-null    object
 3   India         300 non-null    object
dtypes: object(4)
memory usage: 9.5+ KB
```

In [11]:
```
print(centers['Phillippines'].value_counts(),'\n',centers['Indonesia'].value_count
```

```
Error Free    271
Defective      29
Name: Phillippines, dtype: int64
 Error Free    267
Defective      33
Name: Indonesia, dtype: int64
 Error Free    269
Defective      31
Name: Malta, dtype: int64
 Error Free    280
Defective      20
Name: India, dtype: int64
```

In [12]:
```
contingency_table = [[271,267,269,280],
                     [29,33,31,20]]
print(contingency_table)
```

```
[[271, 267, 269, 280], [29, 33, 31, 20]]
```

In [13]:
```
stat, p, df, exp = stats.chi2_contingency(contingency_table)
print("Statistics = ",stat,"\n",'P_Value = ', p,'\n', 'degree of freedom =', df,'\
```

```
Statistics =  3.858960685820355
 P_Value =  0.2771020991233135
 degree of freedom = 3
 Expected Values =  [[271.75 271.75 271.75 271.75]
 [ 28.25  28.25  28.25  28.25]]
```

In [14]:
```
observed = np.array([271, 267, 269, 280, 29, 33, 31, 20])
expected = np.array([271.75, 271.75, 271.75, 271.75, 28.25, 28.25, 28.25, 28.25])
```

In [15]:
```
test_statistic , p_value = stats.chisquare(observed, expected, ddof = df)
print("Test Statistic = ",test_statistic,'\n', 'p_value =',p_value)
```

```
Test Statistic =  3.858960685820355
 p_value = 0.4254298144535761
```

In [16]:
```
alpha = 0.05
print('Significnace=%.3f, p=%.3f' % (alpha, p_value))
if p_value <= alpha:
    print('We reject Null Hypothesis there is a significance difference between TA
else:
    print('We fail to reject Null hypothesis')
```

```
Significnace=0.050, p=0.425
We fail to reject Null hypothesis
```

In [ ]:
```
```