import pandas as pd import numpy as np from scipy import stats from scipy.stats import norm,chi2 from scipy.stats import chi2_contingency

-----read dataset-----

data = pd.read_csv('Downloads/Costomer+OrderForm.csv')
data

	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 Freε
4	Error Free	Error Free	Defective	Error Free
295	Error Free	Error Free	Error Free	Error Free
296	Error Free	Error Free	Error Free	Error Freε
297	Error Free	Error Free	Defective	Error Free
298	Error Free	Error Free	Error Free	Error Freε
299	Error Free	Defective	Defective	Error Free

-----print columns

print(data['Phillippines'].value_counts(),data['Indonesia'].value_counts(),data['Malta'].value_counts())

```
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

observed = ([[271,267,269,280],[29,33,31,20]])

observed

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

```
stats,p,dof,expected = chi2_contingency ([[271,267,269,280],[29,33,31,20]])
 stats
  3.858960685820355
 0.2771020991233135
print('dof=%d' %dof)
print(expected)
  dof=3
  [[271.75 271.75 271.75 271.75]
   [ 28.25 28.25 28.25 28.25]]
alpha = 0.05
prob = 1-alpha
critical = chi2.ppf(prob,dof)
print('probability=%.3f','critical=%.3f','stats=%.3f' % (prob,critical,stats))
if abs(stats) >= critical:
  print('Dependent(reject H0), variables are related')
else:
  print('Independent(reject H0), variables are not related')
print ('significance=%.3f,p=%.3f' %(alpha,p))
if p <= alpha:
  print('Dependent (reject H0)')
else:
  print('Independent (fail to reject H0)')
   significance=0.050,p=0.950
   Independent (fail to reject H0)
```