

The goal of having the Chi-square measure is to calculate how likely a distribution observed is due to chance. It calculates how well the observed data distribution matches with the predicted distribution if the variables are independent. However, PMI represents the discrepancy between the probability of the coincidence between the two words given their joint distribution and their individual distributions. In addition, PMI gave us so many diagrams with equal value. Therefore, Chi-square measure works better with collocation.

```
C:\Users\Lenovo\Desktop\Assignment>python Collocations.py Collocations "chi-square"
of the 4299525524.344475
in the 2967685268.5730453
to the 550239988.0873303
for the 518049079.2873304
the company 437643592.7730447
on the 429357021.80161613
that the 227233932.34447327
and the 208153548.5730447
said the 196130234.94447327
to be 185818439.0873304
in a 177106782.14447322
said it 169898735.77304468
of a 166030188.8016161
at the 155311245.34447324
new york 131501000.77304468
by the 129228880.0873304
from the 124744040.77304468
with the 123635206.37304467
a share 94114354.28733039
will be 92672004.08733039
```

```
C:\Users\Lenovo\Desktop\Assignment>python Collocations.py Collocations "PMI"
stockard channing 18.572478192570014
brideshead revisited 18.572478192570014
schizoid horror 18.572478192570014
ku klux 18.572478192570014
klux klan 18.572478192570014
limply constructed 18.572478192570014
pained expressions 18.572478192570014
coy lazy 18.572478192570014
eloquently explores 18.572478192570014
moldy mildewy 18.572478192570014
smoldering sofa 18.572478192570014
lefty driesell 18.572478192570014
lavender jasmine 18.572478192570014
wafting aromas 18.572478192570014
keyless boga 18.572478192570014
greenwald flocked 18.572478192570014
lada niva 18.572478192570014
bubblelike sneaker 18.572478192570014
leeza spider 18.572478192570014
hijet dumbo 18.572478192570014
```