In [1]:
```python
# Importing libraries
import warnings
warnings.filterwarnings("ignore")

import sqlite3
import pandas as pd
import numpy as np
import nltk
import string
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
from sklearn.feature_extraction.text import TfidfTransformer
from sklearn.feature_extraction.text import TfidfVectorizer

from sklearn.feature_extraction.text import CountVectorizer
from nltk.stem.porter import PorterStemmer

import re

import string
from nltk.corpus import stopwords
from nltk.stem import PorterStemmer
from nltk.stem.wordnet import WordNetLemmatizer

from gensim.models import Word2Vec
from gensim.models import KeyedVectors
import pickle
```

```
C:\Users\Sai charan\Anaconda3\lib\site-packages\ipykernel\parentpoller.py:116: U
serWarning: Parent poll failed.  If the frontend dies,
                the kernel may be left running.  Please let us know
                about your system (bitness, Python, etc.) at
                ipython-dev@scipy.org
  ipython-dev@scipy.org""")
C:\Users\Sai charan\Anaconda3\lib\site-packages\gensim\utils.py:1197: UserWarnin
g: detected Windows; aliasing chunkize to chunkize_serial
  warnings.warn("detected Windows; aliasing chunkize to chunkize_serial")
```

```
In [2]:  con1 = sqlite3.connect('final.sqlite')
         filtered_data = pd.read_sql_query(" SELECT * FROM Reviews ", con1)


         print(filtered_data.shape)
         filtered_data.head()
```

(364171, 12)

Out[2]:

	index	Id	ProductId	UserId	ProfileName	HelpfulnessNumerator	HelpfulnessDenominato
0	138706	150524	0006641040	ACITT7DI6IDDL	shari zychinski	0	
1	138688	150506	0006641040	A2IW4PEEKO2R0U	Tracy	1	
2	138689	150507	0006641040	A1S4A3IQ2MU7V4	sally sue "sally sue"	1	
3	138690	150508	0006641040	AZGXZ2UUK6X	Catherine Hallberg "(Kate)"	1	
4	138691	150509	0006641040	A3CMRKGE0P909G	Teresa	3	

Data Cleaning: Deduplication

In [3]:
```python
sorted_data=filtered_data.sort_values('ProductId', axis=0, ascending=True, inplace=
False, kind='quicksort', na_position='last')

#Deduplication of entries
final=sorted_data.drop_duplicates(subset={"UserId","ProfileName","Time","Text"}, ke
ep='first', inplace=False)
print(final.shape)

#Checking to see how much % of data still remains
((final.shape[0]*1.0)/(filtered_data.shape[0]*1.0)*100)
```

(364171, 12)

Out[3]: 100.0

```
In [4]: final = final[final.HelpfulnessNumerator <= final.HelpfulnessDenominator]

print(final.shape)
final[30:50]
```

```
(364171, 12)
```

Out[4]:

	index	Id	ProductId	UserId	ProfileName	HelpfulnessNumerator	HelpfulnessDenomin
12	138700	150518	0006641040	AK1L4EJBA23JF	L. M. Kraus	0	
13	138701	150519	0006641040	A12HY5OZ2QNK4N	Elizabeth H. Roessner	0	
14	138702	150520	0006641040	ADBFSA9KTQANE	James L. Hammock "Pucks Buddy"	0	
8	138696	150514	0006641040	A2ONB6ZA292PA	Rosalind Matzner	0	
34	476617	515426	141278509X	AB1A5EGHHVA9M	CHelmic	1	
36	22620	24750	2734888454	A13ISQV0U9GZIC	Sandikaye	1	
35	22621	24751	2734888454	A1C298ITT645B6	Hugh G. Pritchard	0	
37	284375	308077	2841233731	A3QD68O22M2XHQ	LABRNTH	0	
142	157910	171225	7310172001	A314APAWYQFKBJ	Diana Hersholt "dog lover"	1	
143	157909	171224	7310172001	AK0CENM3LUM28	Ana Mardoll	1	
144	157924	171240	7310172001	A2JCG7KT8HRSUJ	B. Le	0	
145	157925	171241	7310172001	ACBK6OXAMOHTE	Andrew Fox	0	
146	157926	171242	7310172001	A12L1NY994GXSF	Dawn Rene	0	
147	157907	171222	7310172001	A1CV6GLAPUIP80	Reader Rabbit	1	

Stemming, stop-word removal and Lemmatization

```python
In [5]: from nltk.corpus import stopwords
        stop = set(stopwords.words('english'))
        words_to_keep = set(('not'))
        stop -= words_to_keep
        #initialising the snowball stemmer
        sno = nltk.stem.SnowballStemmer('english')

         #function to clean the word of any html-tags
        def cleanhtml(sentence):
            cleanr = re.compile('<.*?>')
            cleantext = re.sub(cleanr, ' ', sentence)
            return cleantext

        #function to clean the word of any punctuation or special characters
        def cleanpunc(sentence):
            cleaned = re.sub(r'[?|!|\'|"|#]',r'',sentence)
            cleaned = re.sub(r'[.|,|)|(|\|/]',r' ',cleaned)
            return  cleaned
```

```python
In [6]: i=0
        str1=' '
        final_string=[]
        all_positive_words=[] # store words from +ve reviews here
        all_negative_words=[] # store words from -ve reviews here.
        s=''
        for sent in final['Text'].values:
            filtered_sentence=[]
            #print(sent);
            sent=cleanhtml(sent) # remove HTMl tags
            for w in sent.split():
                for cleaned_words in cleanpunc(w).split():
                    if((cleaned_words.isalpha()) & (len(cleaned_words)>2)):
                        if(cleaned_words.lower() not in stop):
                            s=(sno.stem(cleaned_words.lower())).encode('utf8')
                            filtered_sentence.append(s)
                            if (final['Score'].values)[i] == 'positive':
                                all_positive_words.append(s) #list of all words used to des
        cribe positive reviews
                            if(final['Score'].values)[i] == 'negative':
                                all_negative_words.append(s) #list of all words used to des
        cribe negative reviews reviews
                        else:
                            continue
                    else:
                        continue

            str1 = b" ".join(filtered_sentence) #final string of cleaned words


            final_string.append(str1)
            i+=1
```

```
In [7]: final['CleanedText']=final_string
        final['CleanedText']=final['CleanedText'].str.decode("utf-8")
        #below the processed review can be seen in the CleanedText Column
        print('Shape of final',final.shape)
        final.head()
```

Shape of final (364171, 12)

Out[7]:

	index	Id	ProductId	UserId	ProfileName	HelpfulnessNumerator	HelpfulnessDenominato
0	138706	150524	0006641040	ACITT7DI6IDDL	shari zychinski	0	
19	138708	150526	0006641040	A3E9QZFE9KXH8J	R. Mitchell	11	1
20	138709	150529	0006641040	A25ACLV5KPB4W	Matt Hetling "Matt"	0	
21	138699	150517	0006641040	ABW4IC5G5G8B5	kevin clark	0	
22	138686	150504	0006641040	AQEYF1AXARWJZ	Les Sinclair "book maven"	1	

```
In [8]: ##Sorting data according to Time in ascending order for Time Based Splitting
        time_sorted_data = final.sort_values('Time', axis=0, ascending=True, inplace=False,
        kind='quicksort', na_position='last')

        # We will collect different 5K rows without repetition from time_sorted_data datafr
        ame
        my_final = time_sorted_data.take(np.random.permutation(len(final))[:5000])

        x = my_final['CleanedText'].values
```

Bag of Words (BoW)

```
In [9]: count_vect = CountVectorizer(min_df = 10)
        data = count_vect.fit_transform(x)
        print("the type of count vectorizer :",type(data))
        print("the shape of out text BOW vectorizer : ",data.get_shape())
        print("the number of unique words :", data.get_shape()[1])
```

```
the type of count vectorizer : <class 'scipy.sparse.csr.csr_matrix'>
the shape of out text BOW vectorizer :  (5000, 1931)
the number of unique words : 1931
```

Hierarchical Clustering with 2 clusters

```python
from sklearn.cluster import AgglomerativeClustering

model = AgglomerativeClustering(n_clusters=2).fit(data.toarray())

reviews = my_final['Text'].values
# Getting all the reviews in different clusters
cluster1 = []
cluster2 = []

for i in range(model.labels_.shape[0]):
    if model.labels_[i] == 0:
        cluster1.append(reviews[i])
    else :
        cluster2.append(reviews[i])


# Number of reviews in different clusters
print("No. of reviews in Cluster-1 : ",len(cluster1))
print("\nNo. of reviews in Cluster-2 : ",len(cluster2))
```

```
No. of reviews in Cluster-1 :  244

No. of reviews in Cluster-2 :  4756
```

READING REVIEWS MANUALLY

In [11]:
```python
# Three Reviews of cluster 1
count=1
for i in range(3):
    if i < len(cluster1):
        print('Review-%d : \n %s\n'%(count,cluster1[i]))
        count +=1
```

Review-1 :
 This English Breakfast tea from Higgins & Burke, is good: it's basic, good tast
ing black tea (English Breakfast blend) and is tasty. I wasn't blown away by th
e flavor, but it was good: I enjoyed the tea, it brewed up quickly in these nice
tea bags (I recommend about one minute of brew time), and then added a splash of
cream: it was quite nice!

Review-2 :
 I have never had black tea this good before. I love lychee fruit and wanted so
me black tea for the caffeine content. Problem is I do not like the taste of bl
ack tea. This tea is different it is not bitter. It is smooth in taste. I do
not add sugar to my teas so they have to stand on their own. This one does. Th
at is why I purchase teas with flavors in them. This may seem expensive, but re
alize this is 100 bags of tea and comes in a huge box. If you look a regular te
a it is a much smaller box. This looks over three times the size. Great tea tr
y it for your caffenated tea. It is wonderful tea. This beats out earl gray or
english breakfast tea in taste. This tea is a pleasure to drink. I just wanted
to get some caffeine in a tolerable way with no artificial sweeteners or regular
sugars. This is it. In the highly sweetened american culture there is a way ou
t drink this tea from china. It is calorie free with no artifical sweeteners. I
am not asian, but I love asian products.

Review-3 :
 I have been drinking tea for over thirty years, and tea in America has improved
! Vine offered me a chance to try another basic tea, and its a good entry level,
AM drinking tea.
I found the vanilla just a little strong for my taste, so
I mixed it with black tea and it makes a large pot, perfect with stevia and a to
uch of milk.
I do not think this would taste well without milk, but then I
like the english school of tea, and do not understand why Americans try to drink
black tea black, like its coffee?
Not impressed my the less caffeine argume
nt on the side of the box, the only way to get no /less caffeine in teas is to d
rink steeped herbs, this has tea in it, so it has cafffeine, duh.
So enjoy,
its good flavorful choice when tired of the basic teas.

```
In [12]:  # Three Reviews of cluster 2
          count=1
          for i in range(3):
              if i < len(cluster2):
                  print('Review-%d : \n %s\n'%(count,cluster2[i]))
                  count +=1
```

Review-1 :
 I like the mint, but the only reason I have bought these mint are because I lik
e the tin. The tins are very nice. There are a lot of project I can build using
these candy tins. The mint is not bad either, very freshing.

They sel
l these mint for $2.50/tin here in San Francisco, so 12 tins will cost me $30, a
nd I saved about 10 dollars to get it here.

Review-2 :
 This is my second review request for this mustard. Again, I will say that it i
s my favorite mustard and hard to find in select stores. I was so happy to see
that I could purchase it on Amazon. This is sure a great place to shop from my
mustard to my hard to find shampoos. This mustard is wonderful on all kind of m
eats. My husband is not a mustard lover like me, but he does like this brand an
d selection.

Review-3 :
 Great product. Great price. I use it as a substitute for white sugar in everyth
ing from my coffee to cooking and baking. I use Amazon's Subscribe and Save prog
ram for this product. Enjoy.

Hierarchical Clustering with 5 clusters

In [13]:
```python
model = AgglomerativeClustering(n_clusters=5).fit(data.toarray())

# Getting all the reviews in different clusters
cluster1 = []
cluster2 = []
cluster3 = []
cluster4 = []
cluster5 = []

for i in range(model.labels_.shape[0]):
    if model.labels_[i] == 0:
        cluster1.append(reviews[i])
    elif model.labels_[i] == 1:
        cluster2.append(reviews[i])
    elif model.labels_[i] == 2:
        cluster3.append(reviews[i])
    elif model.labels_[i] == 3:
        cluster4.append(reviews[i])
    else :
        cluster5.append(reviews[i])

# Number of reviews in different clusters
print("No. of reviews in Cluster-1 : ",len(cluster1))
print("\nNo. of reviews in Cluster-2 : ",len(cluster2))
print("\nNo. of reviews in Cluster-3 : ",len(cluster3))
print("\nNo. of reviews in Cluster-4 : ",len(cluster4))
print("\nNo. of reviews in Cluster-5 : ",len(cluster5))
```

```
No. of reviews in Cluster-1 :  543

No. of reviews in Cluster-2 :  243

No. of reviews in Cluster-3 :  1

No. of reviews in Cluster-4 :  3943

No. of reviews in Cluster-5 :  270
```

In [14]:
```python
#Three Reviews of cluster 1
count=1
for i in range(3):
    if i < len(cluster1):
        print('Review-%d : \n %s\n'%(count,cluster1[i]))
        count +=1
```

Review-1 :
 It's very difficult to tell if this product is working. I applied it around my deck to repel rabbits. We have them in the yard, and sometimes they hover around the hostas around the deck. I haven't noticed any rabbits near the house since a pplication, but we had rain nearly every day for a week, so I'm not sure if the product is effective anymore or not. I also sprinkled it around one of my garden s that doesn't have a fence to deter rabbits, groundhogs and moles. Since I don' t have a camera set up, and I'm not monitoring it round the clock, I just can't tell if it's keeping all or some of them away.

When I sprinkle it on, the product has a very strong smell of black pepper. It's easy to apply, but doe sn't go very far. It's a little expensive for me to purchase for additional appl ications.

For those wondering what the other 99% of inactive ingredie nts might be, I'm 99% sure it's sand. So, this is black pepper, Piperine, and Ca psaicin, and a ton of sand. If they made it in a liquid that customers could mix with sand for spreadability, I wonder if they could make it at a better price.

Review-2 :
 The dog food got good ratings at consumers, so I bought it and the dogs just lo ve it.. It seems to be a very healtful and complete food.. I just add a little h ot water and its gobbled up in no time

Review-3 :
 We bought this with a coupon at our warehouse grocery store wondering if we'd e ver use it all, but here we are on our second canister now! Let me tell you, thi s stuff is very versatile! I know Nestle just advertises it as a chocolate milk drink mix, but we've found many uses for it.

I use it with milk to ma ke hot cocoas (and it's become my preference over Swiss Miss packets and the lik e), we've used it to add chocolaty flavors to baking projects, we obviously make our toddler the occasional chocolate milk for a treat, but my favorite thing to do with Nesquik powder is to make myself an iced mocha! (Although you could cert ainly make hot mochas as well). My husband has recently been on my case for spen ding almost $5 every day (sometimes twice) on Starbuck's iced mochas. I have a n ewborn, you see, so I'm sleep deprived and my coffee consumption has gone up. I' m addicted! But we realized the Starbuck's charges were adding up quite a bit, s o as an alternative, most days I turn to Nesquik instead. It doesn't taste like Starbuck's mochas, but it does make for a tasty coffee drink if you like them ch ocolaty. I use one shot of espresso (or strong coffee if I'm away from our espre sso machine), two spoonfuls of Nesquik, and fill the glass up the rest of the wa y with about 1 cup of milk. It's yummy! You can add ice, or keep it hot if you l ike. In the long run, we've saved money for sure by using the Nesquik.

Finally, I have to say that as a mom, the ingredient list is satisfactory. Thi s obviously isn't what you'd call health food, and certainly not advertised as o rganic or anything - but as far as the ingredient list goes it's pretty mild. No scary sounding ingredients I can't pronounce on the label - mostly minerals, vit amins, and sugar. There are no artificial sweeteners, as advertised. I absolutel y expect us to remain regular buyers of Nesquik, it fits in nicely with my famil y!

In [15]:
```python
count=1
for i in range(3):
    if i < len(cluster2):
        print('Review-%d : \n %s\n'%(count,cluster2[i]))
        count +=1
```

Review-1 :
 This English Breakfast tea from Higgins & Burke, is good: it's basic, good tast
ing black tea (English Breakfast blend) and is tasty. I wasn't blown away by th
e flavor, but it was good: I enjoyed the tea, it brewed up quickly in these nice
tea bags (I recommend about one minute of brew time), and then added a splash of
cream: it was quite nice!

Review-2 :
 I have never had black tea this good before. I love lychee fruit and wanted so
me black tea for the caffeine content. Problem is I do not like the taste of bl
ack tea. This tea is different it is not bitter. It is smooth in taste. I do
not add sugar to my teas so they have to stand on their own. This one does. Th
at is why I purchase teas with flavors in them. This may seem expensive, but re
alize this is 100 bags of tea and comes in a huge box. If you look a regular te
a it is a much smaller box. This looks over three times the size. Great tea tr
y it for your caffenated tea. It is wonderful tea. This beats out earl gray or
english breakfast tea in taste. This tea is a pleasure to drink. I just wanted
to get some caffeine in a tolerable way with no artificial sweeteners or regular
sugars. This is it. In the highly sweetened american culture there is a way ou
t drink this tea from china. It is calorie free with no artifical sweeteners. I
am not asian, but I love asian products.

Review-3 :
 I have been drinking tea for over thirty years, and tea in America has improved
! Vine offered me a chance to try another basic tea, and its a good entry level,
AM drinking tea.
I found the vanilla just a little strong for my taste, so
I mixed it with black tea and it makes a large pot, perfect with stevia and a to
uch of milk.
I do not think this would taste well without milk, but then I
like the english school of tea, and do not understand why Americans try to drink
black tea black, like its coffee?
Not impressed my the less caffeine argume
nt on the side of the box, the only way to get no /less caffeine in teas is to d
rink steeped herbs, this has tea in it, so it has cafffeine, duh.
So enjoy,
its good flavorful choice when tired of the basic teas.

```
In [16]: count=1
         for i in range(3):
             if i < len(cluster3):
                 print('Review-%d : \n %s\n'%(count,cluster3[i]))
                 count +=1
```

```
count=1
for i in range(3):
    if i < len(cluster3):
        print('Review-%d : \n %s\n'%(count,cluster3[i]))
        count +=1
```

Review-1 :
 QUALITY
In my opinion, having used a dozen or so brands of olive oil, Cola
vita is simply the best reasonably priced olive oil which is widely available.
I've been using Colavita for several years now, and have never been dissappointe
d. Colavita is excellent quality, dependable, and reasonably priced---the perfe
ct combination for a daily general purpose olive oil.

FLAVOR
Alt
hough not an "olive oil gourmand", I'd describe Colavita as "rich, smooth, and n
eutral". Gourmands probably have other silly terms for the flavor, such as "frui
ty" (but a "civilian" would never call it fruity). It is not harsh or bitter lik
e most common olive oil, not spicy like Whole Foods 360, and not notably "olivey
". Colavita is an evoo ("extra virgin olive oil"---to us non-olive oil snobs),
which means "first pressing". Most olive-oil nuts consider the words "extra vir
gin" to be sacred. Practically speaking though, the mere label "extra virgin" i
s not a guarantee of quality---most grocery store brands (evoo or not) are harsh
and bitter, but some are good cooking oils.

PREMIUM OILS
Botique
premium oils are are as variable year to year as wines, comparably expensive, an
d have to be bought the way you buy fine wine (that is, ideally at tastings, whe
re you confirm the quality and flavor before purchase). Some are genuinely (as
in NOT-gourmet-speak) buttery and nutty---great for dipping. Some taste like th
e fragrance of flowers---great for salads. Never cook premium olive oil---use i
t for dipping or drizzling, or on delicate salads. Colavita is a good general p
urpose oil, better than most grocery store olive oil, but is not a premium oil.<
br />
COOKING OIL
Non-extra virgin olive oils (second or later pressin
gs) are usually darker and stronger flavored---which adds complexity to soups, s
auses, stews, bean dishes etc. If you are using a bitter or strong-tasting oli
ve oil, cook it WITH the dish. I add 1/4 to 1/2 cup of olive oil to every sause
, soup, stew, chili, cassrole, or bean dish, etc. However, olive oil can "go ba
d"---so unless you use alot, it doesn't sense to use a different cooking oil tha
n you use on salads, for dipping etc. One usually purchases good cooking oil in
multi-gallon tins. Colavita is a very good general purpose oil, and therefore a
good cooking oil---albeit a bit expensive for a cooking oil.

DIPPING
OIL
Traditionally (and in the best restaurants), Italian bread, oil and bal
samic vinegar is served as an appetizer. The diner usually mixes his own oil an
d vinegar at the table. If all three are top quality, the combination can be de
licious, and can even be a light meal. Add figs or slices of apple or pear and
some olives, and you do have a meal. <a href="http://www.amazon.com/gp/product/B
000FAKUP0">Byzantine Fresh Olive Antipasto, 5-Pound Bag Albeit probably her
esy, in fact you can substitute fine wine vinegar or fresh lemon juice. If I do
n't have olives, I usually add soy sause or Braggs for saltiness to the mixture.
Bragg Liquid Aminos, Natur
al Soy Sauce Alternative, 32-Ounce Bottle , (Pack of 3) Colavita is a very g
ood "everyday" dipping oil.

SALAD OIL
The requirements for an "e
veryday" salad oil are a little less stringent than that for a dipping oil, beca
use the heavy flavor of the vinegar, and often pamessan, bacon bits, etc. usuall
y dominate. However, I rarely prepare use oil and vinegar the traditional way.
For everyday salads (or as a dip for raw vegies) I prefer olive-oil based mayo w
ith a little freshly ground pepper, and some citrus zest. For special salads, I
use a fruity premium olive oil (and just a hint of lemon and/or orange juice and
zest). Colavita is a very good "everyday" salad oil.

DRIZZING OIL
A drizzling oil is an oil you pour over a dish immediately before serving, or
which you serve in a carafe on the table. If you REALLY like the flavor of "ord
inary good" olive oil (I don't)---then you will find Colavita to be a good drizz
ling oil. My hispanic wife frequently drizzles Colavita on the dishes I've cook
ed (with Colavita), because she likes the olive-oil flavor. By the way, the sam
e principle applies to the use of butter, pepper, garlic, cheese, even soy sause
---if you want the flavor to dominate, add it just before serving; if you want t
he flavor to be substle, cook the seasoning with the dish. For most dishes I pr
efer a rich blend of subtle flavors.

HEALTHY COOKING
I go to a g
reat deal of effort to minimize animal fat in my diet, by very strict trimming,
thorough cooking, and by chilling broth to remove the solidified fat. However,
the body needs fat and craves fat. The ideal solution is to replace animal fat
with vegetable oils which are high in omega-3s. Olive oil is ideal for the purp
ose.

My Southern grandmother could not cook any vegetables without a
thick slice of fatback (fatty bacon)---and indeed her vegetables were very tasty

```
In [17]: count=1
         for i in range(3):
             if i < len(cluster4):
                 print('Review-%d : \n %s\n'%(count,cluster4[i]))
                 count +=1
```

```
Review-1 :
 I like the mint, but the only reason I have bought these mint are because I lik
e the tin. The tins are very nice. There are a lot of project I can build using
these candy tins. The mint is not bad either, very freshing.<br /><br />They sel
l these mint for $2.50/tin here in San Francisco, so 12 tins will cost me $30, a
nd I saved about 10 dollars to get it here.

Review-2 :
 This is my second review request for this mustard.  Again, I will say that it i
s my favorite mustard and hard to find in select stores.  I was so happy to see
that I could purchase it on Amazon.  This is sure a great place to shop from my
mustard to my hard to find shampoos.  This mustard is wonderful on all kind of m
eats.  My husband is not a mustard lover like me, but he does like this brand an
d selection.

Review-3 :
 Great product. Great price. I use it as a substitute for white sugar in everyth
ing from my coffee to cooking and baking. I use Amazon's Subscribe and Save prog
ram for this product. Enjoy.
```

```
In [ ]: count=1
        for i in range(3):
            if i < len(cluster5):
                print('Review-%d : \n %s\n'%(count,cluster5[i]))
                count +=1
```

Hierarchical Clustering with 10 clusters

```python
In [18]: model = AgglomerativeClustering(n_clusters=10).fit(data.toarray())
         # Getting all the reviews in different clusters
         cluster1 = []
         cluster2 = []
         cluster3 = []
         cluster4 = []
         cluster5 = []
         cluster6 = []
         cluster7 = []
         cluster8 = []
         cluster9 = []
         cluster10 = []

         for i in range(model.labels_.shape[0]):
             if model.labels_[i] == 0:
                 cluster1.append(reviews[i])
             elif model.labels_[i] == 1:
                 cluster2.append(reviews[i])
             elif model.labels_[i] == 2:
                 cluster3.append(reviews[i])
             elif model.labels_[i] == 3:
                 cluster4.append(reviews[i])
             elif model.labels_[i] == 4:
                 cluster5.append(reviews[i])
             elif model.labels_[i] == 5:
                 cluster6.append(reviews[i])
             elif model.labels_[i] == 6:
                 cluster7.append(reviews[i])
             elif model.labels_[i] == 7:
                 cluster8.append(reviews[i])
             elif model.labels_[i] == 8:
                 cluster9.append(reviews[i])
             else :
                 cluster10.append(reviews[i])
```

In [19]:
```python
# Number of reviews in different clusters
print("No. of reviews in Cluster-1 : ",len(cluster1))
print("\nNo. of reviews in Cluster-2 : ",len(cluster2))
print("\nNo. of reviews in Cluster-3 : ",len(cluster3))
print("\nNo. of reviews in Cluster-4 : ",len(cluster4))
print("\nNo. of reviews in Cluster-5 : ",len(cluster5))
print("\nNo. of reviews in Cluster-6 : ",len(cluster6))
print("\nNo. of reviews in Cluster-7 : ",len(cluster7))
print("\nNo. of reviews in Cluster-8 : ",len(cluster8))
print("\nNo. of reviews in Cluster-9 : ",len(cluster9))
print("\nNo. of reviews in Cluster-10 : ",len(cluster10))
```

No. of reviews in Cluster-1 : 316

No. of reviews in Cluster-2 : 2

No. of reviews in Cluster-3 : 224

No. of reviews in Cluster-4 : 1117

No. of reviews in Cluster-5 : 270

No. of reviews in Cluster-6 : 1

No. of reviews in Cluster-7 : 1

No. of reviews in Cluster-8 : 199

No. of reviews in Cluster-9 : 44

No. of reviews in Cluster-10 : 2826

```
In [20]:  # Three Reviews of cluster 1
          count=1
          for i in range(3):
              if i < len(cluster1):
                  print('Review-%d : \n %s\n'%(count,cluster1[i]))
                  count +=1
```

Review-1 :
 It's very difficult to tell if this product is working. I applied it around my
deck to repel rabbits. We have them in the yard, and sometimes they hover around
the hostas around the deck. I haven't noticed any rabbits near the house since a
pplication, but we had rain nearly every day for a week, so I'm not sure if the
product is effective anymore or not. I also sprinkled it around one of my garden
s that doesn't have a fence to deter rabbits, groundhogs and moles. Since I don'
t have a camera set up, and I'm not monitoring it round the clock, I just can't
tell if it's keeping all or some of them away.

When I sprinkle it on,
the product has a very strong smell of black pepper. It's easy to apply, but doe
sn't go very far. It's a little expensive for me to purchase for additional appl
ications.

For those wondering what the other 99% of inactive ingredie
nts might be, I'm 99% sure it's sand. So, this is black pepper, Piperine, and Ca
psaicin, and a ton of sand. If they made it in a liquid that customers could mix
with sand for spreadability, I wonder if they could make it at a better price.

Review-2 :
 We bought this with a coupon at our warehouse grocery store wondering if we'd e
ver use it all, but here we are on our second canister now! Let me tell you, thi
s stuff is very versatile! I know Nestle just advertises it as a chocolate milk
drink mix, but we've found many uses for it.

I use it with milk to ma
ke hot cocoas (and it's become my preference over Swiss Miss packets and the lik
e), we've used it to add chocolaty flavors to baking projects, we obviously make
our toddler the occasional chocolate milk for a treat, but my favorite thing to
do with Nesquik powder is to make myself an iced mocha! (Although you could cert
ainly make hot mochas as well). My husband has recently been on my case for spen
ding almost $5 every day (sometimes twice) on Starbuck's iced mochas. I have a n
ewborn, you see, so I'm sleep deprived and my coffee consumption has gone up. I'
m addicted! But we realized the Starbuck's charges were adding up quite a bit, s
o as an alternative, most days I turn to Nesquik instead. It doesn't taste like
Starbuck's mochas, but it does make for a tasty coffee drink if you like them ch
ocolaty. I use one shot of espresso (or strong coffee if I'm away from our espre
sso machine), two spoonfuls of Nesquik, and fill the glass up the rest of the wa
y with about 1 cup of milk. It's yummy! You can add ice, or keep it hot if you l
ike. In the long run, we've saved money for sure by using the Nesquik.

Finally, I have to say that as a mom, the ingredient list is satisfactory. Thi
s obviously isn't what you'd call health food, and certainly not advertised as o
rganic or anything - but as far as the ingredient list goes it's pretty mild. No
scary sounding ingredients I can't pronounce on the label - mostly minerals, vit
amins, and sugar. There are no artificial sweeteners, as advertised. I absolutel
y expect us to remain regular buyers of Nesquik, it fits in nicely with my famil
y!

Review-3 :
 White Gold Raw Unfiltered Honey is wonderful.

I've always loved hone
y...believed that the honey I purchased was just that...HONEY. Ignorance is bli
ss...eh? I've been really dismayed to discover that some of the honey (the proc
essed/filtered/clear kind) out there is adulterated with HIGH FRUCTOSE CORN SYRU
P (HFCS). I HATE this stuff. How would I or you know? We can't...it's an impo
ssibility...and that's exactly how HFCS manufacturers want to keep it. If HFCS
listed as an ingredient on any product I pick up at the grocery store I place it
back on the shelf and walk away. I'm sure that's EXACTLY why corn producers and
HFCS manufacturers do not want regulations that require it being listed in the i
ngredients on products that contain HFCS...if we only knew. Sadly, some bee far
mers use HFCS to supplement their bees diets...could this be part of the cause f
or bee colony collapse? I wonder.

I'm stepping off my soapbox now an
d want you to know that this White Gold Raw/Unfiltered Honey is absolutely the b
est honey I've ever tasted. It's mild, buttery, and aromatic all at the same ti
me. It just melts in your mouth and has a clean, yummy aftertaste.

In [21]:
```python
# Three Reviews of cluster 2
count=1
for i in range(3):
    if i < len(cluster2):
        print('Review-%d : \n %s\n'%(count,cluster2[i]))
        count +=1
```

Review-1 :
 "We use and believe in stone milling because no machinery has yet been develope
d that grinds grains into flours, cereals, and meals quite as well as the flint-
hard quartz stones used by master millers since early Roman times." ~Bob's Red M
ill Site

The 100 % Stone Ground Whole Wheat Flour is my favorite flou
r for use in a variety of whole wheat bread recipes. It is all natural and doesn
't contain any preservatives. You can use this flour for bread machine baking to
o and it produces superior results. I've used the recipe on the back of the bag
and adapted it by using a mixture of whole wheat and regular bread flour because
I didn't have any "gluten" as called for in the recipe. If you substitute 1 ½ cu
ps of regular bread flour for part of the whole wheat flour you don't seem to ne
ed the gluten.

Bob's Red Mill products remind me of my childhood when
my father took me to the Graue Mill in Oakbrook, IL where we watched corn being
milled. It is now the only operating waterwheel gristmill in Illinois. While Bob
's Red Mill products don't arrive in a cloth bag, the paper bags do have an old
fashioned feel and have a string-sewn closure. I love the crinkly sound of the h
igh-quality paper bags.

Bob's Red Mill uses slow turning, 100-year-ol
d millstones to naturally grind together the bran, endosperm and germ that conta
ins the nutritious oil. This is a "cool stone grinding process" which preserves
nutrients normally lost in high-speed, high-heat milling.

This compan
y produces more than 400 products. Look for everything from hot cereals and mixe
s to flour milled from a wide variety of grains. If you love making natural brea
ds in your bread machine, the flours are superior and produce deliciously nutrit
ious breads.

Whole Wheat Oat Rolls

1 c. water, warmed to 1
10°F
5 T. fat-free milk powder
2 T. clover honey
2 T. salted butt
er, melted
2 1/2 tsp. active dry yeast
1 lg. egg
2 c. bread flour

1 c. Bob's Red Mill whole-wheat flour
1/2 c. Bob's Red Mill thick-rol
led oats
2 T. wheat germ
1 1/2 tsp. salt
Melted butter
Oats<
br />
The dough may be made in a bread machine or by hand. Bread machine me
thod: Place the water, milk powder, honey, melted butter, yeast, egg, bread<br /
>flour, whole-wheat flour, oats, wheat germ and salt into a bread machine pan. S
et bread machine for whole wheat, dough and large loaf cycle. Form rolls.
H
and method: In a large bowl, whisk the water, milk powder, honey, melted butter,
yeast and egg. In an extra large bowl, mix the flour, whole wheat flour,
oa
ts, wheat germ and salt. Stir in the yeast mixture.

Place dough on a
well-floured breadboard. Knead until smooth and elastic. Place in an oiled bowl,
turn once, then cover with a damp cloth. Turn oven on to 200°F, when you can fee
l the heat on your hand, turn the oven off. Place bowl in the warm oven and let
the dough rise until double.

Punch the dough down and cut into 16 pie
ces. Form rolls then brush with melted butter. Sprinkle the tops of the rolls wi
th the oats. Let rise for 45 minutes or until double then bake in a preheated 40
0°F oven 10-15 minutes or until lightly browned.

Makes 16 rolls

~The Rebecca Review

Review-2 :
 My apologies in advance for what is a lengthy review. However, if I can help a
nyone save a little time compared to what I spent trying to choose a good puppy
food, that's my goal. So I wrote this not only to review the Castor & Pollux pr
oduct, but also to provide some information on other types of dry dog food.<br /
>
It is no secret that you can buy some very cheap and some very expensive
dog/puppy foods. Common sense tells you that you can spend very little and get
lots of filler and by-products, or spend a good deal more and get wholesome, hea
lthy food. My personal opinion is that long term, both your dog and you are bet
ter off if s/he is fed good food. The little extra money spent now may offset p
otential medical bills and heartache later.
That being the case, I set out
to find a really good dog food for my Golden Retriever puppy, while, of course,
trying to keep cost down. I spent a good deal of time researching foods. I dis
regarded low end foods that are largely grains, fillers and meat by-products. I
looked at average to premium foods.

I narrowed my choices to several
including Castor & Pollux Natural Ultramix. The others were Blue, Eukanuba, Inn
ova EVO, Merrick, Natural Balance, and Wellness.

After that I used a
few review websites and ratings tables. These can be helpful. I used one that
starts each perspective food off with 100 points and then subtracts for poor ing
redients (example: -10 points of containing by-products) and then adds points fo

```
In [22]:  # Three Reviews of cluster 3
          count=1
          for i in range(3):
              if i < len(cluster3):
                  print('Review-%d : \n %s\n'%(count,cluster3[i]))
                  count +=1
```

Review-1 :
 The dog food got good ratings at consumers, so I bought it and the dogs just lo
ve it.. It seems to be a very healtful and complete food.. I just add a little h
ot water and its gobbled up in no time

Review-2 :
 In terms of convenience and taste, these are great. Overall, I have only one co
mplaint. Given the design of the pouch it's pretty much impossible to get all th
e food out, which seems like a waste every time. It's not much, but more so than
a regular glass container of baby food.

Breakdown:

PROS:<b
r />1. Baby loves it, and savors it.
2. Organic ingredients
3. Super e
asy to open (simple twist)
4. Tastes great (yeah, I try her food to make su
re we're giving her food that has some flavor to it)
5. 4.5 ounces of food
(larger than the normal/standard of 4 ounces.

Neutral
1. In gene
ral I wish that baby food emphasized more the vegetables than the fruit. (these
are plenty sweet, they could have easily used half as much pear and the rest cou
ld've been some healthy greens.
2. Feel very liquid, more so than most baby
food that I've bought.

CONS:
1. Due to puch design, baby food is
wasted.

Review-3 :
 When you love Indian food, but don't have the time to cook it, Kitchens of Indi
a is a really great alternative. I have yet to be disappointed in one of their d
ishes and this one is no exception. The great thing about Indian food is that th
ey are mostly sauce (or "gravy" as they say), so I will cook up some meat or ste
am up some veggies and add one of these packets to make a fabulous meal. Their r
eally aren't much in the way of veggies in these packets, but that is typical in
Indian food, no matter where it comes from. American's tend to use gravy or sauc
e as the secondary component while Indian's use it as the primary component and
sop it up with lots of bread or rice. By steaming or sauteing some additional vi
ttles, you can really stretch one of these packets.

In [23]:
```python
# Three Reviews of cluster 4
count=1
for i in range(3):
    if i < len(cluster4):
        print('Review-%d : \n %s\n'%(count,cluster4[i]))
        count +=1
```

Review-1 :
 I like the mint, but the only reason I have bought these mint are because I like the tin. The tins are very nice. There are a lot of project I can build using these candy tins. The mint is not bad either, very freshing.

They sell these mint for $2.50/tin here in San Francisco, so 12 tins will cost me $30, and I saved about 10 dollars to get it here.

Review-2 :
 Great product. Great price. I use it as a substitute for white sugar in everything from my coffee to cooking and baking. I use Amazon's Subscribe and Save program for this product. Enjoy.

Review-3 :
 Kitchens of India is the closest to a homemade curry we have ever had. We love it. The sauce is great and it is fast, so curry a few times a week is possible. We often make a curry from scratch but it is time consuming and now we can have a good curry whenever we want. We have also tried the Rogan Josh as well as the Korma. Add some onions and raisins, or a bit of honey. The ingredients have little or no preservatives. The Best!!

In [24]:
```python
# Three Reviews of cluster 5
count=1
for i in range(3):
    if i < len(cluster5):
        print('Review-%d : \n %s\n'%(count,cluster5[i]))
        count +=1
```

Review-1 :
 This is a good deal. Pack of 2, one for me and one to give as a gift. Very easy to use. I use it at night to make a cup of decaf. Makes good coffee and you can make coffee your desired strength. I use 1 scoop of coffee for every 5oz of water.

Review-2 :
 I bought this coffee on a whim as it was a little cheaper than green mountain. I really like the flavor, it is a great medium flavor coffee. I will purchase again.

Review-3 :
 Marley Coffee & Tea One Love Organic Coffee makes a tasty cup of coffee. However, I have to say that nowhere on the box, or the individual packets, are instructions for actually making the coffee. Also it doesn't tell you what type of coffee maker to use or even how much water to use for each packet. I use K-cups and my daughter uses T-cups. I had to make this product in my old automatic drip coffee maker, and I had to guess on how much water to use. Once I was able to make a cup, it was very good.

In [25]:
```python
# Three Reviews of cluster 6
count=1
for i in range(3):
    if i < len(cluster6):
        print('Review-%d : \n %s\n'%(count,cluster6[i]))
        count +=1
```

Review-1 :

QUALITY
In my opinion, having used a dozen or so brands of olive oil, Colavita is simply the best reasonably priced olive oil which is widely available. I've been using Colavita for several years now, and have never been dissappointed. Colavita is excellent quality, dependable, and reasonably priced---the perfect combination for a daily general purpose olive oil.

FLAVOR
Although not an "olive oil gourmand", I'd describe Colavita as "rich, smooth, and neutral". Gourmands probably have other silly terms for the flavor, such as "fruity" (but a "civilian" would never call it fruity). It is not harsh or bitter like most common olive oil, not spicy like Whole Foods 360, and not notably "olivey". Colavita is an evoo ("extra virgin olive oil"---to us non-olive oil snobs), which means "first pressing". Most olive-oil nuts consider the words "extra virgin" to be sacred. Practically speaking though, the mere label "extra virgin" is not a guarantee of quality---most grocery store brands (evoo or not) are harsh and bitter, but some are good cooking oils.

PREMIUM OILS
Botique premium oils are are as variable year to year as wines, comparably expensive, and have to be bought the way you buy fine wine (that is, ideally at tastings, where you confirm the quality and flavor before purchase). Some are genuinely (as in NOT-gourmet-speak) buttery and nutty---great for dipping. Some taste like the fragrance of flowers---great for salads. Never cook premium olive oil---use it for dipping or drizzling, or on delicate salads. Colavita is a good general purpose oil, better than most grocery store olive oil, but is not a premium oil.

COOKING OIL
Non-extra virgin olive oils (second or later pressings) are usually darker and stronger flavored---which adds complexity to soups, sauces, stews, bean dishes etc. If you are using a bitter or strong-tasting olive oil, cook it WITH the dish. I add 1/4 to 1/2 cup of olive oil to every sause, soup, stew, chili, cassrole, or bean dish, etc. However, olive oil can "go bad"---so unless you use alot, it doesn't sense to use a different cooking oil than you use on salads, for dipping etc. One usually purchases good cooking oil in multi-gallon tins. Colavita is a very good general purpose oil, and therefore a good cooking oil---albeit a bit expensive for a cooking oil.

DIPPING OIL
Traditionally (and in the best restaurants), Italian bread, oil and balsamic vinegar is served as an appetizer. The diner usually mixes his own oil and vinegar at the table. If all three are top quality, the combination can be delicious, and can even be a light meal. Add figs or slices of apple or pear and some olives, and you do have a meal. Byzantine Fresh Olive Antipasto, 5-Pound Bag Albeit probably heresy, in fact you can substitute fine wine vinegar or fresh lemon juice. If I don't have olives, I usually add soy sause or Braggs for saltiness to the mixture. Bragg Liquid Aminos, Natural Soy Sauce Alternative, 32-Ounce Bottle , (Pack of 3) Colavita is a very good "everyday" dipping oil.

SALAD OIL
The requirements for an "everyday" salad oil are a little less stringent than that for a dipping oil, because the heavy flavor of the vinegar, and often pamessan, bacon bits, etc. usually dominate. However, I rarely prepare use oil and vinegar the traditional way. For everyday salads (or as a dip for raw vegies) I prefer olive-oil based mayo with a little freshly ground pepper, and some citrus zest. For special salads, I use a fruity premium olive oil (and just a hint of lemon and/or orange juice and zest). Colavita is a very good "everyday" salad oil.

DRIZZING OIL
/>A drizzling oil is an oil you pour over a dish immediately before serving, or which you serve in a carafe on the table. If you REALLY like the flavor of "ordinary good" olive oil (I don't)---then you will find Colavita to be a good drizzling oil. My hispanic wife frequently drizzles Colavita on the dishes I've cooked (with Colavita), because she likes the olive-oil flavor. By the way, the same principle applies to the use of butter, pepper, garlic, cheese, even soy sause---if you want the flavor to dominate, add it just before serving; if you want the flavor to be substle, cook the seasoning with the dish. For most dishes I prefer a rich blend of subtle flavors.

HEALTHY COOKING
I go to a great deal of effort to minimize animal fat in my diet, by very strict trimming, thorough cooking, and by chilling broth to remove the solidified fat. However, the body needs fat and craves fat. The ideal solution is to replace animal fat with vegetable oils which are high in omega-3s. Olive oil is ideal for the purpose.

My Southern grandmother could not cook any vegetables without a thick slice of fatback (fatty bacon)---and indeed her vegetables were very tasty

In [26]:
```python
# Three Reviews of cluster 7
count=1
for i in range(3):
    if i < len(cluster7):
        print('Review-%d : \n %s\n'%(count,cluster7[i]))
        count +=1
```

Review-1 :
 Review:
I've eaten ramen for a long time (30+ years) and have eaten numerous varieties and the Shin ramen is probably my #1 tasting. The reasons are as follows:
* Best flavor broth - it's spicy (but adjustable), leaves a good aftertaste (yes, even when you burp), tint of mushroom flavor coupled with a good balance of red pepper, onions and spices.
* Noodles have good texture - depending how you cook them of course but with bad ones, no matter how you cook them they come out soggy/soft. These come out chewy if you want or softer if you over cook them.
* It's not oily - I've noticed with some noodles, you can see a ring of oil around the pot after you cook the noodles. So with these I sometimes had to boil an extra pot of water and then transfer the cooked noodles to the fresh pot of water to get rid of the oil. Not with Shin Ramen.
* Economical - for less than a dollar a pack, they beat any dollar menu item
* Great for any weather - particularly winter to keep you warm. In the summer, they help to break out a sweat and opens your pores
* Can be a hearty meal - you can add all types of things to this ramen as others have posted. An egg is recommended but so are vegetables and meat.

Tips:

Amount of water - add about 2 mug size cup fulls of water. If not sure, add a little less. You can always add more after it's fully cooked. Adding cold water at the end cools down the boiling hot soup so it's easier to eat.

How to cook - wait til the water boils. As soon as it does, drop in everything: the soup base, noodles, dried vegie packet. If you want to add vegies, let it cook for about a minute before adding e.g. mushrooms, bean sprouts, tomatoes, green onions. Don't put more than a quarter couple of any vegies or meat cuz you'll kill the flavor. For eggs, if you like them well cooked, add them at the beginning. You can add at the end of cooking too but this is a bit too raw for my taste

Cooking time - boil the noodles 2 - 4 minutes. This depends on how hard or soft you like your noodles. Just pull a strand out and taste it every 30 seconds after the first minute. There's no right or wrong here.

Soup base - add all of it. Don't worry about it being too spicy. Per above you can always water it down later. The soup will be plenty hot. Remember it just finished boiling and it's difficult to put something in your mouth when it was just boiling.

How to serve - there are 3 ways I'd recommend
1) The traditional Korean way - folks like to eat straight from the pot. There's actually an 'art' to this. After removing it from the stove, place the pot on something that wont burn. Grab the lid of the pot, a spoon (yes a spoon) and chop sticks (fork if you can't wield them).

2) In a bowl. Don't use a plastic one unless it can stand the heat. I don't like to use plastic period for health reasons but up to you. Remove some of the excess soup from the pot first and then pour everything directly into the bowl. Grab a spoon and chopsticks/fork and place the bowl down and go to work.

3) combination of 1 and 2 above. Place the pot of ramen and get a bowl along with a spoon and chopsticks/fork.

How to eat:
First, forget what your parents have told you on slurping. To eat ramen, you must slurp. Why? Because 1) if you don't you'll burn your lips because no slurping means the noodles must pass through your lips slower and enclosed more around the noodles. 2) because your lips are more enclosed, most of the soup around the noodles are sucked away from noodles, leaving you with less flavorful noodles. So forget the decorum. Eat it asian style and slurp away.

2nd, I'd recommend the Korean way (#1 above). You grab the pot lid then take about 2 mouthful amounts of noodle and dump it into the lid. You should be holding the lid with your left hand (switch if you're right handed) and quickly grab a large spoon and add several spoonfuls of soup broth into the lid. You're doing 3 things here 1) you're saturating the noodles with the flavorful broth. 2) by adding the noodles to the lid, you're cooling it down for consumption while keeping the rest of the noodles hot. 3) When you eat straight from the lid, you first eat the noodles and immediately follow up with a slurp from the lid itself to drink the soup base. To do this, you put the lid close to your mouth and then push the noodles in with your utensil, drink the soup right after or together as you push the noodles in. I know this sounds crude but the flavor is maximized this way. You can add any vegetables or meat to eat bite this way too.

If you want to eat using method #2, well, it needs no explanation, just make sure you slurp and drink the broth as you eat the noodles.

For method #3, you're basically substituting the bowl for the lid. Works as well but assuming your lid is metal, it keeps the noodles hotter. Also, people tend to put more noodles into a bowl then a lid and

In [27]:
```python
# Three Reviews of cluster 8
count=1
for i in range(3):
    if i < len(cluster8):
        print('Review-%d : \n %s\n'%(count,cluster8[i]))
        count +=1
```

Review-1 :
 This English Breakfast tea from Higgins & Burke, is good: it's basic, good tast
ing black tea (English Breakfast blend) and is tasty. I wasn't blown away by th
e flavor, but it was good: I enjoyed the tea, it brewed up quickly in these nice
tea bags (I recommend about one minute of brew time), and then added a splash of
cream: it was quite nice!

Review-2 :
 I purchased this tea because of all the good reviews. I like it very much and
will order again. Maybe Davidson's has change the ingredients because I found a
lot of rosehips and had to really search for blueberries and found only 1. I als
o have the chamomile flowers and mixed that with the spring fruit and flowers in
my tea pot, the combo made an excellent tea. I do love rosehips and the flavor
of the tea but was disappointed that I couldn't taste blueberry. For this reaso
n alone I gave this tea 4 stars. I just got my second order of this tea and it
was full of blueberries, no rosehips and has a blueberry taste to it. I will ke
ep ordering this tea and I'm giving it 5 stars now.

Review-3 :
 I was looking all over for an organic Earl Grey tea for a reasonable price and
found that it was either ridiculous expensive or not Earl Gray which I prefer. I
tried 2 other brands but didn't like the taste. I'm very pleased with this one.
Smooth and pleasant flavor and even better tasting than the one I used to buy he
re in San Francisco in a local store (I got tired of always feeling that I paid
far too much for that tea). One review complained about bitterness and I'm sensi
tive to that so that was a worry... but that's not the case at all for me even w
hen making it strong... this is an excellent quality and I have had different ki
nd of teas every morning for more than 40 years both here in the US and in Europ
e. Highly recommended.

```
In [28]:  # Three Reviews of cluster 8
          count=1
          for i in range(3):
              if i < len(cluster9):
                  print('Review-%d : \n %s\n'%(count,cluster9[i]))
                  count +=1
```

Review-1 :
 I have never had black tea this good before. I love lychee fruit and wanted so
me black tea for the caffeine content. Problem is I do not like the taste of bl
ack tea. This tea is different it is not bitter. It is smooth in taste. I do
not add sugar to my teas so they have to stand on their own. This one does. Th
at is why I purchase teas with flavors in them. This may seem expensive, but re
alize this is 100 bags of tea and comes in a huge box. If you look a regular te
a it is a much smaller box. This looks over three times the size. Great tea tr
y it for your caffenated tea. It is wonderful tea. This beats out earl gray or
english breakfast tea in taste. This tea is a pleasure to drink. I just wanted
to get some caffeine in a tolerable way with no artificial sweeteners or regular
sugars. This is it. In the highly sweetened american culture there is a way ou
t drink this tea from china. It is calorie free with no artifical sweeteners. I
am not asian, but I love asian products.

Review-2 :
 I have been drinking tea for over thirty years, and tea in America has improved
! Vine offered me a chance to try another basic tea, and its a good entry level,
AM drinking tea.
I found the vanilla just a little strong for my taste, so
I mixed it with black tea and it makes a large pot, perfect with stevia and a to
uch of milk.
I do not think this would taste well without milk, but then I
like the english school of tea, and do not understand why Americans try to drink
black tea black, like its coffee?
Not impressed my the less caffeine argume
nt on the side of the box, the only way to get no /less caffeine in teas is to d
rink steeped herbs, this has tea in it, so it has cafffeine, duh.
So enjoy,
its good flavorful choice when tired of the basic teas.

Review-3 :
 I have used Triple Leaf brand tea in the past and tried several variations. Ea
ch time I end up using two bags, it may just be that I prefer a stronger tea, bu
t even the bags seem a bit smaller than normal, maybe it's intentional and I am
overdoing it. This product in particular is a good mix, ginseng, ginko, decaf g
reen tea. I tried to find a product with this blend for awhile with no luck. T
he tea tastes great with some honey, it has a bit of an earthy taste, as to be e
xpected, but it is not overpowering and the honey easily smooths it out. Ginsen
g and ginko have a long history and many reported benefits, I do know buy into a
ll of it, but some of it must be true. Clinical studies have shown ginko to hel
p people who have occasional claudication, but had no effect on individuals with
noraml operating memories. Of course green tea is antioxidant rich, and the lon
gest living people drink it daily. These things seem to change everyday though,
I remember when butter was bad for us... Basicaly, a good product, priced right
. Get's one star off for the weakness of the brew, but as I said, maybe it shou
ld get full stars and it was intentional.

```
In [ ]:  # Three Reviews of cluster 8
         count=1
         for i in range(3):
             if i < len(cluster10):
                 print('Review-%d : \n %s\n'%(count,cluster10[i]))
                 count +=1
```

TFIDF

```
In [29]: tf_idf_vect = TfidfVectorizer(min_df=10)
         data = tf_idf_vect.fit_transform(x)
         print("the type of count vectorizer :",type(data))
         print("the shape of out text TFIDF vectorizer : ",data.get_shape())
         print("the number of unique words :", data.get_shape()[1])
```

```
the type of count vectorizer : <class 'scipy.sparse.csr.csr_matrix'>
the shape of out text TFIDF vectorizer :  (5000, 1931)
the number of unique words : 1931
```

Hierarchical Clustering with 2 clusters

```
In [30]: model = AgglomerativeClustering(n_clusters=2).fit(data.toarray())

         reviews = my_final['Text'].values
         # Getting all the reviews in different clusters
         cluster1 = []
         cluster2 = []

         for i in range(model.labels_.shape[0]):
             if model.labels_[i] == 0:
                 cluster1.append(reviews[i])
             else :
                 cluster2.append(reviews[i])


         # Number of reviews in different clusters
         print("No. of reviews in Cluster-1 : ",len(cluster1))
         print("\nNo. of reviews in Cluster-2 : ",len(cluster2))
```

```
No. of reviews in Cluster-1 :  4743

No. of reviews in Cluster-2 :  257
```

READING REVIEWS MANUALLY:

In [31]:
```python
# Three Reviews of cluster 1
count=1
for i in range(3):
    print('Review-%d : \n %s\n'%(count,cluster1[i]))
    count +=1
```

Review-1 :
 I like the mint, but the only reason I have bought these mint are because I like the tin. The tins are very nice. There are a lot of project I can build using these candy tins. The mint is not bad either, very freshing.

They sell these mint for $2.50/tin here in San Francisco, so 12 tins will cost me $30, and I saved about 10 dollars to get it here.

Review-2 :
 This is my second review request for this mustard. Again, I will say that it is my favorite mustard and hard to find in select stores. I was so happy to see that I could purchase it on Amazon. This is sure a great place to shop from my mustard to my hard to find shampoos. This mustard is wonderful on all kind of meats. My husband is not a mustard lover like me, but he does like this brand and selection.

Review-3 :
 Great product. Great price. I use it as a substitute for white sugar in everything from my coffee to cooking and baking. I use Amazon's Subscribe and Save program for this product. Enjoy.

In [32]:
```python
# Three Reviews of cluster 1
count=1
for i in range(3):
    print('Review-%d : \n %s\n'%(count,cluster2[i]))
    count +=1
```

Review-1 :
 Do not know a lot about this type of tea, so couldn't compare it to other brands. I can say the product is rather strong tasting, but not unpleasantly so. Since a portion is pretty small, that is not much of a problem anyway. I purchased it after reading about its many health benefits. It does open my eyes in the morning, so the caffeine in it is really nice. As recommended by others, I keep it in the refrigerator inside the tin, in its original package, and it's kept fresh-tasting for a couple of weeks now. The color is very bright green, and it does taste healthy--hopefully, it is.

Review-2 :
 I have to say that this tea is very helpful during those stressful times. I don't pay attention to the taste, but next thing I know, I'm like, "WOW - I'm so relaxed." All that tension slowly melts away. Great to have around during the time of the month.

Review-3 :
 This English Breakfast tea from Higgins & Burke, is good: it's basic, good tasting black tea (English Breakfast blend) and is tasty. I wasn't blown away by the flavor, but it was good: I enjoyed the tea, it brewed up quickly in these nice tea bags (I recommend about one minute of brew time), and then added a splash of cream: it was quite nice!

Hierarchical Clustering with 5 clusters

In [33]:
```python
model = AgglomerativeClustering(n_clusters=5).fit(data.toarray())

# Getting all the reviews in different clusters
cluster1 = []
cluster2 = []
cluster3 = []
cluster4 = []
cluster5 = []

for i in range(model.labels_.shape[0]):
    if model.labels_[i] == 0:
        cluster1.append(reviews[i])
    elif model.labels_[i] == 1:
        cluster2.append(reviews[i])
    elif model.labels_[i] == 2:
        cluster3.append(reviews[i])
    elif model.labels_[i] == 3:
        cluster4.append(reviews[i])
    else :
        cluster5.append(reviews[i])

# Number of reviews in different clusters
print("No. of reviews in Cluster-1 : ",len(cluster1))
print("\nNo. of reviews in Cluster-2 : ",len(cluster2))
print("\nNo. of reviews in Cluster-3 : ",len(cluster3))
print("\nNo. of reviews in Cluster-4 : ",len(cluster4))
print("\nNo. of reviews in Cluster-5 : ",len(cluster5))
```

```
No. of reviews in Cluster-1 :  4019

No. of reviews in Cluster-2 :  262

No. of reviews in Cluster-3 :  358

No. of reviews in Cluster-4 :  257

No. of reviews in Cluster-5 :  104
```

READING REVIEWS MANUALLY:

In [34]:
```python
# Three Reviews of cluster 1
count=1
for i in range(3):
    if i < len(cluster1):
        print('Review-%d : \n %s\n'%(count,cluster1[i]))
        count +=1
```

Review-1 :
 I like the mint, but the only reason I have bought these mint are because I like the tin. The tins are very nice. There are a lot of project I can build using these candy tins. The mint is not bad either, very freshing.

They sell these mint for $2.50/tin here in San Francisco, so 12 tins will cost me $30, and I saved about 10 dollars to get it here.

Review-2 :
 This is my second review request for this mustard. Again, I will say that it is my favorite mustard and hard to find in select stores. I was so happy to see that I could purchase it on Amazon. This is sure a great place to shop from my mustard to my hard to find shampoos. This mustard is wonderful on all kind of meats. My husband is not a mustard lover like me, but he does like this brand and selection.

Review-3 :
 Great product. Great price. I use it as a substitute for white sugar in everything from my coffee to cooking and baking. I use Amazon's Subscribe and Save program for this product. Enjoy.

In [35]:
```python
# Three Reviews of cluster 2
count=1
for i in range(3):
    if i < len(cluster2):
        print('Review-%d : \n %s\n'%(count,cluster2[i]))
        count +=1
```

Review-1 :
 I have a serious addiction to this gum but cannot find it anywhere except online. I go through at least a pack a day. I used to chew Peppermint or Spearmint or Winterfrost but I can't chew anything but Exotic Mint any more. The flavor is just right and it lasts forever.

Seriously, if you see this gum in a store, buy all of it. It's like little candy gold bullion. I have opted to pay twice as much online for this gum versus gum I can buy on the store shelves just so I can get this particular flavor. Who would do that if the gum weren't the best gum ever? Nobody. That's who.

I also wrote Wrigley to see if they'll send me a year-supply for free. Cross your fingers and wish me luck.

Review-2 :
 We've had this for about a week now. I have a 1 year old Beagle that destroys ever furry "super strong" soft toy, but he isn't interested in a lot of hard toys like tires, etc. We got this hoping that his food motivation would keep him occupied with a toy he can't tear apart.

The Good:
  * It is VERY durable. He hasn't even managed to put a dent it.
  * We stuck a treat inside it, which is impossible for him to remove. He will happily spend several hours occupied trying to get it out. Yaaay!

The Ugly:
  * It took him about 20 mins to pop the regular "everlasting" treat out. Without the ball to protect it, it's a 10 minute treat. :-(He can do this every time now.

Review-3 :
 The dog food got good ratings at consumers, so I bought it and the dogs just love it.. It seems to be a very healtful and complete food.. I just add a little hot water and its gobbled up in no time

In [36]:
```python
# Three Reviews of cluster 3
count=1
for i in range(3):
    if i < len(cluster3):
        print('Review-%d : \n %s\n'%(count,cluster3[i]))
        count +=1
```

Review-1 :
 This is a good deal. Pack of 2, one for me and one to give as a gift. Very ea
sy to use. I use it at night to make a cup of decaf. Makes good coffee and you
can make coffee your desired strength. I use 1 scoop of coffee for every 5oz of
water.

Review-2 :
 Pilao is my favorite coffe. I have tried several brands from around the globe b
ut so far this is the best tasting. Not bitter, nice and smooth.

Review-3 :
 I bought this coffee on a whim as it was a little cheaper than green mountain.
I really like the flavor, it is a great medium flavor coffee. I will purchase a
gain.

In [37]:
```python
# Three Reviews of cluster 4
count=1
for i in range(3):
    if i < len(cluster4):
        print('Review-%d : \n %s\n'%(count,cluster4[i]))
        count +=1
```

Review-1 :
 Do not know a lot about this type of tea, so couldn't compare it to other brand
s. I can say the product is rather strong tasting, but not unpleasantly so. Sinc
e a portion is pretty small, that is not much of a problem anyway. I purchased i
t after reading about its many health benefits. It does open my eyes in the morn
ing, so the caffeine in it is really nice. As recommended by others, I keep it i
n the refrigerator inside the tin, in its original package, and it's kept fresh-
tasting for a couple of weeks now. The color is very bright green, and it does t
aste healthy--hopefully, it is.

Review-2 :
 I have to say that this tea is very helpful during those stressful times. I do
n't pay attention to the taste, but next thing I know, I'm like, "WOW - I'm so r
elaxed." All that tension slowly melts away. Great to have around during the t
ime of the month.

Review-3 :
 This English Breakfast tea from Higgins & Burke, is good: it's basic, good tast
ing black tea (English Breakfast blend) and is tasty. I wasn't blown away by th
e flavor, but it was good: I enjoyed the tea, it brewed up quickly in these nice
tea bags (I recommend about one minute of brew time), and then added a splash of
cream: it was quite nice!

```
In [38]:  # Three Reviews of cluster 5
          count=1
          for i in range(3):
              if i < len(cluster5):
                  print('Review-%d : \n %s\n'%(count,cluster5[i]))
                  count +=1
```

Review-1 :
 We really like Feline Pine Cat Litter, it is a very low maintenance type of pro
duct and friendly to the environment.

Review-2 :
 This is great if you work long shifts and have a small dog. I have a 7 lb bolo
gnese dog and the breeder litter box trains the dogs from the puppy stage on. W
e got her home and she just jumped in the litter box and went. She has used it
now for a year with very little accidents. It's great and I only replace the li
tter every 7 - 10 days!

Review-3 :
 One of my two cats LOVES this and gobbles it down. The other only likes fancy
feast unfortunately. I love the ingredients and the smell. It looks just like
a can of chicken that any human could eat! The label says it all--feeding your
cat weruva can increase your cat's lives from nine to ten. Thank you weruva and
amazon for providing this fabulous cat food at affordable prices. Puuurrs!

Hierarchical Clustering with 10 clusters

```python
In [39]: model = AgglomerativeClustering(n_clusters=10).fit(data.toarray())
         # Getting all the reviews in different clusters
         cluster1 = []
         cluster2 = []
         cluster3 = []
         cluster4 = []
         cluster5 = []
         cluster6 = []
         cluster7 = []
         cluster8 = []
         cluster9 = []
         cluster10 = []

         for i in range(model.labels_.shape[0]):
             if model.labels_[i] == 0:
                 cluster1.append(reviews[i])
             elif model.labels_[i] == 1:
                 cluster2.append(reviews[i])
             elif model.labels_[i] == 2:
                 cluster3.append(reviews[i])
             elif model.labels_[i] == 3:
                 cluster4.append(reviews[i])
             elif model.labels_[i] == 4:
                 cluster5.append(reviews[i])
             elif model.labels_[i] == 5:
                 cluster6.append(reviews[i])
             elif model.labels_[i] == 6:
                 cluster7.append(reviews[i])
             elif model.labels_[i] == 7:
                 cluster8.append(reviews[i])
             elif model.labels_[i] == 8:
                 cluster9.append(reviews[i])
             else :
                 cluster10.append(reviews[i])
```

```
In [40]:  # Number of reviews in different clusters
          print("No. of reviews in Cluster-1 : ",len(cluster1))
          print("\nNo. of reviews in Cluster-2 : ",len(cluster2))
          print("\nNo. of reviews in Cluster-3 : ",len(cluster3))
          print("\nNo. of reviews in Cluster-4 : ",len(cluster4))
          print("\nNo. of reviews in Cluster-5 : ",len(cluster5))
          print("\nNo. of reviews in Cluster-6 : ",len(cluster6))
          print("\nNo. of reviews in Cluster-7 : ",len(cluster7))
          print("\nNo. of reviews in Cluster-8 : ",len(cluster8))
          print("\nNo. of reviews in Cluster-9 : ",len(cluster9))
          print("\nNo. of reviews in Cluster-10 : ",len(cluster10))
```

```
No. of reviews in Cluster-1 :  3089

No. of reviews in Cluster-2 :  262

No. of reviews in Cluster-3 :  596

No. of reviews in Cluster-4 :  146

No. of reviews in Cluster-5 :  104

No. of reviews in Cluster-6 :  65

No. of reviews in Cluster-7 :  358

No. of reviews in Cluster-8 :  257

No. of reviews in Cluster-9 :  81

No. of reviews in Cluster-10 :   42
```

READING REVIEWS MANUALLY

In [41]:
```python
# Three Reviews of cluster 1
count=1
for i in range(3):
    if i < len(cluster1):
        print('Review-%d : \n %s\n'%(count,cluster1[i]))
        count +=1
```

Review-1 :
 This is my second review request for this mustard. Again, I will say that it i
s my favorite mustard and hard to find in select stores. I was so happy to see
that I could purchase it on Amazon. This is sure a great place to shop from my
mustard to my hard to find shampoos. This mustard is wonderful on all kind of m
eats. My husband is not a mustard lover like me, but he does like this brand an
d selection.

Review-2 :
 Kitchens of India is the closest to a homemade curry we have ever had. We love
it. The sauce is great and it is fast, so curry a few times a week is possible.
We often make a curry from scratch but it is time consuming and now we can have
a good curry whenever we want. We have also tried the Rogan Josh as well as the
Korma. Add some onions and raisins, or a bit of honey. The ingredients have litt
le or no preservatives. The Best!!

Review-3 :
 This product is light, not too sweet and very satisfying. Just what it says it
is and shows on the box. There is no sacrifice of delectable quality in the pur
suit of a gluten free product at all. Needs no frosting at all! Guests told me
how wonderful the cake was even when they didn't know it was good for everybody,
including us gluten free folk.

In [42]:
```python
# Three Reviews of cluster 2
count=1
for i in range(3):
    if i < len(cluster2):
        print('Review-%d : \n %s\n'%(count,cluster2[i]))
        count +=1
```

Review-1 :
 I have a serious addiction to this gum but cannot find it anywhere except online. I go through at least a pack a day. I used to chew Peppermint or Spearmint or Winterfrost but I can't chew anything but Exotic Mint any more. The flavor is just right and it lasts forever.

Seriously, if you see this gum in a store, buy all of it. It's like little candy gold bullion. I have opted to pay twice as much online for this gum versus gum I can buy on the store shelves just so I can get this particular flavor. Who would do that if the gum weren't the best gum ever? Nobody. That's who.

I also wrote Wrigley to see if they'll send me a year-supply for free. Cross your fingers and wish me luck.

Review-2 :
 We've had this for about a week now. I have a 1 year old Beagle that destroys ever furry "super strong" soft toy, but he isn't interested in a lot of hard toys like tires, etc. We got this hoping that his food motivation would keep him occupied with a toy he can't tear apart.

The Good:
  * It is VERY durable. He hasn't even managed to put a dent it.
  * We stuck a treat inside it, which is impossible for him to remove. He will happily spend several hours occupied trying to get it out. Yaaay!

The Ugly:
  * It took him about 20 mins to pop the regular "everlasting" treat out. Without the ball to protect it, it's a 10 minute treat. :-(He can do this every time now.

Review-3 :
 The dog food got good ratings at consumers, so I bought it and the dogs just love it.. It seems to be a very healtful and complete food.. I just add a little hot water and its gobbled up in no time

In [43]:
```python
# Three Reviews of cluster 3
count=1
for i in range(3):
    if i < len(cluster3):
        print('Review-%d : \n %s\n'%(count,cluster3[i]))
        count +=1
```

Review-1 :
 I like the mint, but the only reason I have bought these mint are because I like the tin. The tins are very nice. There are a lot of project I can build using these candy tins. The mint is not bad either, very freshing.

They sell these mint for $2.50/tin here in San Francisco, so 12 tins will cost me $30, and I saved about 10 dollars to get it here.

Review-2 :
 Great product. Great price. I use it as a substitute for white sugar in everything from my coffee to cooking and baking. I use Amazon's Subscribe and Save program for this product. Enjoy.

Review-3 :
 For all you celiacs out there, this is the best cereal ever. You can eat with milk in the am or dry in the afternoon as a snack. Great price point when you order without shipping charge.

In [44]:
```python
# Three Reviews of cluster 4
count=1
for i in range(3):
    if i < len(cluster4):
        print('Review-%d : \n %s\n'%(count,cluster4[i]))
        count +=1
```

Review-1 :
 I started eating almonds regularly after going on the South Beach Diet. I find
them quite delicious . . . so delicious that I have to remind myself that I'm on
ly supposed to eat 15 or so a day. Oh, well.

I could easily eat a ha
ndful . . . and still wouldn't feel like I had gotten enough flavor.

I had no idea what a cocoa roast almond would taste like, but I was pleasantly s
urprised that the taste was superb. It reminded me a little of eating a high qua
lity fudge with lots of nuts in it.

But the nice surprise was that th
e cocoa taste lingered in my mouth so that I continued to enjoy the flavor long
after the almonds were eaten. Even though the cocoa roasting adds some calories
and some fat, I would be satisfied after eating far fewer of these almonds. I pl
an to buy and eat these on a regular basis.

Try them. I think you'll
like them!

Review-2 :
 I absolutely love Justin's Nut Butter Maple Almond! It is so good I eat it out
of the jar (like an earlier reviewer stated). But don't take my word for it...i
t's Justin's favorite too!!! As an FYI, it is not as smooth as peanut butter.
There is an almost grainy consistency. This does not deter from it's flavor and
it is still easy to spread.

Review-3 :
 Purchased one out one day and really enjoyed it. Found them on Amazon and orde
red this flavor, which is my favorite and Coconut. These bars taste great & are
not like other protein or natural bars I have tried in the past. I would recomm
end these bars for a quick snack, after work out or even breakfast.

In [45]:
```python
# Three Reviews of cluster 5
count=1
for i in range(3):
    if i < len(cluster5):
        print('Review-%d : \n %s\n'%(count,cluster5[i]))
        count +=1
```

Review-1 :
 We really like Feline Pine Cat Litter, it is a very low maintenance type of pro
duct and friendly to the environment.

Review-2 :
 This is great if you work long shifts and have a small dog. I have a 7 lb bolo
gnese dog and the breeder litter box trains the dogs from the puppy stage on. W
e got her home and she just jumped in the litter box and went. She has used it
now for a year with very little accidents. It's great and I only replace the li
tter every 7 - 10 days!

Review-3 :
 One of my two cats LOVES this and gobbles it down. The other only likes fancy
feast unfortunately. I love the ingredients and the smell. It looks just like
a can of chicken that any human could eat! The label says it all--feeding your
cat weruva can increase your cat's lives from nine to ten. Thank you weruva and
amazon for providing this fabulous cat food at affordable prices. Puuurrs!

In [46]:
```python
# Three Reviews of cluster
count=1
for i in range(3):
    if i < len(cluster6):
        print('Review-%d : \n %s\n'%(count,cluster6[i]))
        count +=1
```

Review-1 :
 For someone who is vegan and gluten-free, finding snacks out and about can be a challenge. Lucy's cookies are the perfect answer! Not many places carry them, but, luckily, they can be ordered by the case from Amazon for lunchbox and purse stocking!

All the flavors are good. Sometimes cookies are broken, but that is to be expected when something like this is shipped, and that just makes the package last longer!

Review-2 :
 The Back to Nature Fudge Mint cookies and these Fudge Striped cookies are my favorites.

The Fudge Striped cookies have a light crunch and a great chocolaty taste. For some reason when I eat them upside-down they taste even better. These are a bit healthier than Keebler's which have trans fat in their fudge striped cookies.

Review-3 :
 I really enjoyed these cookies. They were soft and yummy. Almost like hommade. I would repurchase them again. Quaker is a brand I love and trust for quality.

In [47]:
```python
# Three Reviews of cluster 7
count=1
for i in range(3):
    if i < len(cluster7):
        print('Review-%d : \n %s\n'%(count,cluster7[i]))
        count +=1
```

Review-1 :
 This is a good deal. Pack of 2, one for me and one to give as a gift. Very easy to use. I use it at night to make a cup of decaf. Makes good coffee and you can make coffee your desired strength. I use 1 scoop of coffee for every 5oz of water.

Review-2 :
 Pilao is my favorite coffe. I have tried several brands from around the globe but so far this is the best tasting. Not bitter, nice and smooth.

Review-3 :
 I bought this coffee on a whim as it was a little cheaper than green mountain. I really like the flavor, it is a great medium flavor coffee. I will purchase again.

In [48]:
```python
# Three Reviews of cluster 8
count=1
for i in range(3):
    if i < len(cluster8):
        print('Review-%d : \n %s\n'%(count,cluster8[i]))
        count +=1
```

Review-1 :
 Do not know a lot about this type of tea, so couldn't compare it to other brand
s. I can say the product is rather strong tasting, but not unpleasantly so. Sinc
e a portion is pretty small, that is not much of a problem anyway. I purchased i
t after reading about its many health benefits. It does open my eyes in the morn
ing, so the caffeine in it is really nice. As recommended by others, I keep it i
n the refrigerator inside the tin, in its original package, and it's kept fresh-
tasting for a couple of weeks now. The color is very bright green, and it does t
aste healthy--hopefully, it is.

Review-2 :
 I have to say that this tea is very helpful during those stressful times. I do
n't pay attention to the taste, but next thing I know, I'm like, "WOW - I'm so r
elaxed." All that tension slowly melts away. Great to have around during the t
ime of the month.

Review-3 :
 This English Breakfast tea from Higgins & Burke, is good: it's basic, good tast
ing black tea (English Breakfast blend) and is tasty. I wasn't blown away by th
e flavor, but it was good: I enjoyed the tea, it brewed up quickly in these nice
tea bags (I recommend about one minute of brew time), and then added a splash of
cream: it was quite nice!

```
In [49]:  # Three Reviews of cluster 9
          count=1
          for i in range(3):
              if i < len(cluster9):
                  print('Review-%d : \n %s\n'%(count,cluster9[i]))
                  count +=1
```

Review-1 :
 I am helping a 71 yr.old sufferer from Parkinson's to completely recover from t
he syndrome, partly by replacing animal-product fish oil with pure virgin coconu
t oil, and with very satisfying results!

Review-2 :
 After doing research for some time on coconut oils, I settled on Tropical Trad
itions. I've been using it for years and usually put some in my oatmeal. I've al
so used it on my skin after extended sun exposure. I've been very happy with the
product. In my research I found that Coconut Oil has very strong anti-viral prop
erties. A friend of mine's daughter came down with a bad case of Mono. She was t
old she would be out of Softball for several months. I gave him a jar and instru
cted him to start her out with one spoonful in the a.m. and then work up to 3 a
day. Within several days she was showing signs of being better and within two we
eks was back in the Texas Summer heat playing softball and never had any issues.
Was it the Coconut Oil? I think yes but regardless, sometimes, you have nothing
to lose and everything to gain by trying what Native Islanders have known for ma
ny centuries. I can't speak to the flavor of other VGOs but I've have good resu
lts with this one so I'm sticking with it.

Review-3 :
 I gave my wife the new "CLEAR Scalp and Hair Therapy shampoo and conditioner fo
r damage and color repair" for her birthday. She is picky about her hair product
s and has trained me to purchase only those for color treated hair, though she r
arely admits to having color added. I've learned my lesson, about her pickiness,
when buying shampoo and conditioner, but I dared to try something new.

Whew! Did I ever score with CLEAR! She told me "the bottle shapes are so cute,
and the open/close mechanism was very secure," something she doesn't always find
on her regular brands. She also said she doesn't need to use as big a portion to
clean as she has before, making it very economical. She said the suds felt thick
, smooth, and it had a nice mild scent.

After just a couple days of u
se, and especially after a full week of daily use, she thought her hair was inde
ed shinier thatn before, had much better texture, and seemed smooth, manageable,
and extremely soft. Usually in the humidity of warm weather she has the "frizzie
s." When the CLEAR runs out, she asked me to get her more. Boy, do I now have li
ttle gifts for future occasions--the jackpot of suprise gift-giving!

```
In [50]: # Three Reviews of cluster 10
         count=1
         for i in range(3):
             if i < len(cluster10):
                 print('Review-%d : \n %s\n'%(count,cluster10[i]))
                 count +=1
```

```
Review-1 :
 Bought the popper about two years ago and have been enjoying the delicious fres
h buttery salty (as I want) best popcorn ever.  love it and it's a staple snack
in our house. Would never micorwave corn again.

Review-2 :
 I have eaten all kinds of popcorn for many years.  Many years ago, we made popc
orn the old fashioned way; in a pot on the stove.  I remember it being crisp and
crunchy.  Then, along came microwave popcorn.  I tried many different brands and
none of them had the same crisp, and sometimes they were just plain mushy. I tho
ught this was just a common attribute among microwavable popcorn products until
several years ago when I tried making it the old fashioned way with the same res
ults.<br /><br />What I have come to realize is, it's not how you prepare the po
pcorn, it's whether or not it's "organic."  The first time I tried "Newman's Own
," it had that crisp and crunchy texture that I remembered from so many years ag
o.  Perfect!

Review-3 :
 Pricey and hard to get Kettle Corn popcorn seasoning from Kernel Seasons. Snap
it up if its here because it may not be tomorrow. Not sure why its 50 cents a bo
ttle higher than all the other flavors.
```

Word2Vec

```
In [51]: sent_x = []
         for sent in x :
             sent_x.append(sent.split())


         # Train your own Word2Vec model using your own train text corpus
         # min_count = 5 considers only words that occured atleast 5 times
         w2v_model=Word2Vec(sent_x,min_count=5,size=50, workers=4)

         w2v_words = list(w2v_model.wv.vocab)
         print("number of words that occured minimum 5 times ",len(w2v_words))
```

```
number of words that occured minimum 5 times  3177
```

```
In [52]: train_vectors = [];
         for sent in sent_x:
             sent_vec = np.zeros(50)
             cnt_words =0;
             for word in sent: #
                 if word in w2v_words:
                     vec = w2v_model.wv[word]
                     sent_vec += vec
                     cnt_words += 1
             if cnt_words != 0:
                 sent_vec /= cnt_words
             train_vectors.append(sent_vec)

         data = train_vectors
```

Hierarchical Clustering with 2 clusters

```
In [53]:  model = AgglomerativeClustering(n_clusters=2).fit(data)

          reviews = my_final['Text'].values
          # Getting all the reviews in different clusters
          cluster1 = []
          cluster2 = []

          for i in range(model.labels_.shape[0]):
              if model.labels_[i] == 0:
                  cluster1.append(reviews[i])
              else :
                  cluster2.append(reviews[i])


          # Number of reviews in different clusters
          print("No. of reviews in Cluster-1 : ",len(cluster1))
          print("\nNo. of reviews in Cluster-2 : ",len(cluster2))
```

```
No. of reviews in Cluster-1 :  2814

No. of reviews in Cluster-2 :  2186
```

```
In [54]:  # Three Reviews of cluster 1
          count=1
          for i in range(3):
              print('Review-%d : \n %s\n'%(count,cluster1[i]))
              count +=1
```

```
Review-1 :
 Great product. Great price. I use it as a substitute for white sugar in everyth
ing from my coffee to cooking and baking. I use Amazon's Subscribe and Save prog
ram for this product. Enjoy.

Review-2 :
 Kitchens of India is the closest to a homemade curry we have ever had. We love
it. The sauce is great and it is fast, so curry a few times a week is possible.
We often make a curry from scratch but it is time consuming and now we can have
a good curry whenever we want.  We have also tried the Rogan Josh as well as the
Korma. Add some onions and raisins, or a bit of honey. The ingredients have litt
le or no preservatives. The Best!!

Review-3 :
 Teff in general is a delicious healthy grain, and this Teff is no different.<br
/>Cook according to pkg. directions and add to morning oatmeal for a hearty heal
thy breakfast.<br />If I had one wish it would be for it to be "Organic Brown Te
ff."
```

In [55]:
```python
# Three Reviews of cluster 2
count=1
for i in range(3):
    if i < len(cluster2):
        print('Review-%d : \n %s\n'%(count,cluster2[i]))
        count +=1
```

Review-1 :
 I like the mint, but the only reason I have bought these mint are because I lik
e the tin. The tins are very nice. There are a lot of project I can build using
these candy tins. The mint is not bad either, very freshing.

They sel
l these mint for $2.50/tin here in San Francisco, so 12 tins will cost me $30, a
nd I saved about 10 dollars to get it here.

Review-2 :
 This is my second review request for this mustard. Again, I will say that it i
s my favorite mustard and hard to find in select stores. I was so happy to see
that I could purchase it on Amazon. This is sure a great place to shop from my
mustard to my hard to find shampoos. This mustard is wonderful on all kind of m
eats. My husband is not a mustard lover like me, but he does like this brand an
d selection.

Review-3 :
 For all you celiacs out there, this is the best cereal ever. You can eat with
milk in the am or dry in the afternoon as a snack. Great price point when you o
rder without shipping charge.

Hierarchical Clustering with 5 clusters

In [56]:
```python
model = AgglomerativeClustering(n_clusters=5).fit(data)

# Getting all the reviews in different clusters
cluster1 = []
cluster2 = []
cluster3 = []
cluster4 = []
cluster5 = []

for i in range(model.labels_.shape[0]):
    if model.labels_[i] == 0:
        cluster1.append(reviews[i])
    elif model.labels_[i] == 1:
        cluster2.append(reviews[i])
    elif model.labels_[i] == 2:
        cluster3.append(reviews[i])
    elif model.labels_[i] == 3:
        cluster4.append(reviews[i])
    else :
        cluster5.append(reviews[i])

# Number of reviews in different clusters
print("No. of reviews in Cluster-1 : ",len(cluster1))
print("\nNo. of reviews in Cluster-2 : ",len(cluster2))
print("\nNo. of reviews in Cluster-3 : ",len(cluster3))
print("\nNo. of reviews in Cluster-4 : ",len(cluster4))
print("\nNo. of reviews in Cluster-5 : ",len(cluster5))
```

```
No. of reviews in Cluster-1 :  2221

No. of reviews in Cluster-2 :  751

No. of reviews in Cluster-3 :  593

No. of reviews in Cluster-4 :  473

No. of reviews in Cluster-5 :  962
```

In [ ]:
```python
# Three Reviews of cluster 1
count=1
for i in range(3):
    if i < len(cluster1):
        print('Review-%d : \n %s\n'%(count,cluster1[i]))
        count +=1
```

In [57]:
```python
# Three Reviews of cluster 2
count=1
for i in range(3):
    if i < len(cluster2):
        print('Review-%d : \n %s\n'%(count,cluster2[i]))
        count +=1
```

Review-1 :
 This is my second review request for this mustard. Again, I will say that it i
s my favorite mustard and hard to find in select stores. I was so happy to see
that I could purchase it on Amazon. This is sure a great place to shop from my
mustard to my hard to find shampoos. This mustard is wonderful on all kind of m
eats. My husband is not a mustard lover like me, but he does like this brand an
d selection.

Review-2 :
 For all you celiacs out there, this is the best cereal ever. You can eat with
milk in the am or dry in the afternoon as a snack. Great price point when you o
rder without shipping charge.

Review-3 :
 These arrived quickly but the box was really banged up. had to open a couple of
bags just to be sure the chips weren't too broken up since I'm selling these in
a vending machine, they were pretty crunched, I can't sell these.

In [58]:
```python
# Three Reviews of cluster 3
count=1
for i in range(3):
    if i < len(cluster3):
        print('Review-%d : \n %s\n'%(count,cluster3[i]))
        count +=1
```

Review-1 :
 This is a good deal. Pack of 2, one for me and one to give as a gift. Very ea
sy to use. I use it at night to make a cup of decaf. Makes good coffee and you
can make coffee your desired strength. I use 1 scoop of coffee for every 5oz of
water.

Review-2 :
 I Have been ordering this k cup cappucino drink for awhile now and have never h
ad a problem! Love it tastes good!

Review-3 :
 This water taste great. We love the bottle. Perfect water for everyone :) I lov
e that it is 1.5 liters because I am trying to drink more water. You will love t
his water.

```
In [59]:  # Three Reviews of cluster 4
          count=1
          for i in range(3):
              if i < len(cluster4):
                  print('Review-%d : \n %s\n'%(count,cluster4[i]))
                  count +=1
```

```
Review-1 :
 Had this in cart and ready to buy when after a bit of research (info not readil
y available) realized the excessive cost of shipping.  Such cost make this produ
ct less than feasible to try?

Review-2 :
 I love this product, well, my cat loves this product. They no longer sell it in
the pet stores, I was so grateful to find it on Amazon, my true shopping place.<
br /><br />The service was excellent, fast and on time. Would buy again

Review-3 :
 I do not live in the United States, and I am used of Splenda, and where I live
nowadays, splenda is very expensive and not easy to find, and I found out that i
f I order it online from Amazon.com, I will get it much cheaper than where I liv
e.
```

```
In [60]:  # Three Reviews of cluster 5
          count=1
          for i in range(3):
              if i < len(cluster5):
                  print('Review-%d : \n %s\n'%(count,cluster5[i]))
                  count +=1
```

```
Review-1 :
 I like the mint, but the only reason I have bought these mint are because I lik
e the tin. The tins are very nice. There are a lot of project I can build using
these candy tins. The mint is not bad either, very freshing.<br /><br />They sel
l these mint for $2.50/tin here in San Francisco, so 12 tins will cost me $30, a
nd I saved about 10 dollars to get it here.

Review-2 :
 This product is light, not too sweet and very satisfying.  Just what it says it
is and shows on the box.  There is no sacrifice of delectable quality in the pur
suit of a gluten free product at all.  Needs no frosting at all! Guests told me
how wonderful the cake was even when they didn't know it was good for everybody,
including us gluten free folk.

Review-3 :
 My 8-month baby always screws up his face like he's being poisoned when you giv
e him something new, so I thought he'd need some time to get used to this.  Not
true!  He loved this from the first bite.  And while he tends not to like foods
with a lot of texture yet, the rice-in-sauce texture of this food doesn't bother
him at all.
```

Hierarchical Clustering with 10 clusters

```
In [61]:  model = AgglomerativeClustering(n_clusters=10).fit(data)
          # Getting all the reviews in different clusters
          cluster1 = []
          cluster2 = []
          cluster3 = []
          cluster4 = []
          cluster5 = []
          cluster6 = []
          cluster7 = []
          cluster8 = []
          cluster9 = []
          cluster10 = []

          for i in range(model.labels_.shape[0]):
              if model.labels_[i] == 0:
                  cluster1.append(reviews[i])
              elif model.labels_[i] == 1:
                  cluster2.append(reviews[i])
              elif model.labels_[i] == 2:
                  cluster3.append(reviews[i])
              elif model.labels_[i] == 3:
                  cluster4.append(reviews[i])
              elif model.labels_[i] == 4:
                  cluster5.append(reviews[i])
              elif model.labels_[i] == 5:
                  cluster6.append(reviews[i])
              elif model.labels_[i] == 6:
                  cluster7.append(reviews[i])
              elif model.labels_[i] == 7:
                  cluster8.append(reviews[i])
              elif model.labels_[i] == 8:
                  cluster9.append(reviews[i])
              else :
                  cluster10.append(reviews[i])
```

In [62]:
```python
# Number of reviews in different clusters
print("No. of reviews in Cluster-1 : ",len(cluster1))
print("\nNo. of reviews in Cluster-2 : ",len(cluster2))
print("\nNo. of reviews in Cluster-3 : ",len(cluster3))
print("\nNo. of reviews in Cluster-4 : ",len(cluster4))
print("\nNo. of reviews in Cluster-5 : ",len(cluster5))
print("\nNo. of reviews in Cluster-6 : ",len(cluster6))
print("\nNo. of reviews in Cluster-7 : ",len(cluster7))
print("\nNo. of reviews in Cluster-8 : ",len(cluster8))
print("\nNo. of reviews in Cluster-9 : ",len(cluster9))
print("\nNo. of reviews in Cluster-10 : ",len(cluster10))
```

No. of reviews in Cluster-1 : 593

No. of reviews in Cluster-2 : 962

No. of reviews in Cluster-3 : 893

No. of reviews in Cluster-4 : 732

No. of reviews in Cluster-5 : 447

No. of reviews in Cluster-6 : 392

No. of reviews in Cluster-7 : 431

No. of reviews in Cluster-8 : 320

No. of reviews in Cluster-9 : 26

No. of reviews in Cluster-10 :   204

In [63]:
```python
# Three Reviews of cluster 1
count=1
for i in range(3):
    if i < len(cluster1):
        print('Review-%d : \n %s\n'%(count,cluster1[i]))
        count +=1
```

Review-1 :
 This is a good deal. Pack of 2, one for me and one to give as a gift. Very ea
sy to use. I use it at night to make a cup of decaf. Makes good coffee and you
can make coffee your desired strength. I use 1 scoop of coffee for every 5oz of
water.

Review-2 :
 I Have been ordering this k cup cappucino drink for awhile now and have never h
ad a problem! Love it tastes good!

Review-3 :
 This water taste great. We love the bottle. Perfect water for everyone :) I lov
e that it is 1.5 liters because I am trying to drink more water. You will love t
his water.

```
In [64]: # Three Reviews of cluster 2
         count=1
         for i in range(3):
             if i < len(cluster2):
                 print('Review-%d : \n %s\n'%(count,cluster2[i]))
                 count +=1
```

Review-1 :
 I like the mint, but the only reason I have bought these mint are because I lik
e the tin. The tins are very nice. There are a lot of project I can build using
these candy tins. The mint is not bad either, very freshing.

They sel
l these mint for $2.50/tin here in San Francisco, so 12 tins will cost me $30, a
nd I saved about 10 dollars to get it here.

Review-2 :
 This product is light, not too sweet and very satisfying. Just what it says it
is and shows on the box. There is no sacrifice of delectable quality in the pur
suit of a gluten free product at all. Needs no frosting at all! Guests told me
how wonderful the cake was even when they didn't know it was good for everybody,
including us gluten free folk.

Review-3 :
 My 8-month baby always screws up his face like he's being poisoned when you giv
e him something new, so I thought he'd need some time to get used to this. Not
true! He loved this from the first bite. And while he tends not to like foods
with a lot of texture yet, the rice-in-sauce texture of this food doesn't bother
him at all.

```
In [65]: # Three Reviews of cluster 3
         count=1
         for i in range(3):
             if i < len(cluster3):
                 print('Review-%d : \n %s\n'%(count,cluster3[i]))
                 count +=1
```

Review-1 :
 Teff in general is a delicious healthy grain, and this Teff is no different.
Cook according to pkg. directions and add to morning oatmeal for a hearty heal
thy breakfast.
If I had one wish it would be for it to be "Organic Brown Te
ff."

Review-2 :
 I made this cake for Christmas 2008. My family thought this cake was dry and s
aid that didn't taste good. It actually left a bad aftertaste in the mouth. I
frosted it with Cherrybrook Kitchen Chocolate Frosting Gluten Free and they love
d that. They did not want me to ever repeat this cake but they said they would
be happy to have the frosting again.

Review-3 :
 We love Sunbutter! I don't have to worry about using it, as it and the company
is completely PEANUT FREE!. Sunbutter taste great, unlike Soy Butter which is gr
itty and awful tasting. I use Sunbutter for anything that I would use peanut but
ter for cookies, frostings and candies, dips, etc. Sunbutter has the same consis
tency and texture as peanut butter and doesn't need refrigeration. When using Su
nbutter for baking you do have to cut back to about 1/2 of the baking soda and/o
r baking powder or you can add a little bit of lemon juice to prevent your baked
goods from turning green from the sunflower seeds. If your cookies do turn green
it doesn't affect the taste just the looks (guess how I know that little bit of
information :). Sunbutter also has a great website at sunbutter dot com with rec
ipes and information about their products. Check out their site and defiantly ge
t some Sunbutter.

In [66]:
```python
# Three Reviews of cluster 4
count=1
for i in range(3):
    if i < len(cluster4):
        print('Review-%d : \n %s\n'%(count,cluster4[i]))
        count +=1
```

Review-1 :
 Kitchens of India is the closest to a homemade curry we have ever had. We love it. The sauce is great and it is fast, so curry a few times a week is possible. We often make a curry from scratch but it is time consuming and now we can have a good curry whenever we want. We have also tried the Rogan Josh as well as the Korma. Add some onions and raisins, or a bit of honey. The ingredients have litt le or no preservatives. The Best!!

Review-2 :
 I was looking for yeast with no added ingredients, and this is not it...
I ngredients: Yeast, sorbitan monostearate, ascorbic acid
It also came hot of f the UPS truck, and I hesitate to open it for fear it is activated and dead. W ill return instead.
Not what I was looking for, won't purchase again.

Review-3 :
 My grown kids always dieting love the flavors and that it is carbonated, especi ally the Pomegranate Berry, my grandkids, whose parents are always watching what they are drinking, love them just as much. I had one accident with the grandkid s who had carried it into my bedroom and spilled on my carpet, I didnt think I w as going to be able to get out the pink stain and was very surprised when a rag w/a little cold water got it right out. My frig is usually the first place the whole family looks to see what flavors I have on hand. I think Sparling Ice Moun tain Spring Water is far better that most on the market.

In [67]:
```python
# Three Reviews of cluster 5
count=1
for i in range(3):
    if i < len(cluster5):
        print('Review-%d : \n %s\n'%(count,cluster5[i]))
        count +=1
```

Review-1 :
 Had this in cart and ready to buy when after a bit of research (info not readil y available) realized the excessive cost of shipping. Such cost make this produ ct less than feasible to try?

Review-2 :
 I love this product, well, my cat loves this product. They no longer sell it in the pet stores, I was so grateful to find it on Amazon, my true shopping place.< br />
The service was excellent, fast and on time. Would buy again

Review-3 :
 I do not live in the United States, and I am used of Splenda, and where I live nowadays, splenda is very expensive and not easy to find, and I found out that i f I order it online from Amazon.com, I will get it much cheaper than where I liv e.

In [68]:
```python
# Three Reviews of cluster 6
count=1
for i in range(3):
    if i < len(cluster6):
        print('Review-%d : \n %s\n'%(count,cluster6[i]))
        count +=1
```

Review-1 :
 Great product. Great price. I use it as a substitute for white sugar in everyth
ing from my coffee to cooking and baking. I use Amazon's Subscribe and Save prog
ram for this product. Enjoy.

Review-2 :
 I have this item to ship monthly. I think this is the best honey that I have f
ound. I tried almost all local honeys, I wanted uncook & unfiltered honey for t
he best health benefit. I found this is by far the best flavor and consistency.
My honey arrives on time and in good condition.

Review-3 :
 Do not know a lot about this type of tea, so couldn't compare it to other brand
s. I can say the product is rather strong tasting, but not unpleasantly so. Sinc
e a portion is pretty small, that is not much of a problem anyway. I purchased i
t after reading about its many health benefits. It does open my eyes in the morn
ing, so the caffeine in it is really nice. As recommended by others, I keep it i
n the refrigerator inside the tin, in its original package, and it's kept fresh-
tasting for a couple of weeks now. The color is very bright green, and it does t
aste healthy--hopefully, it is.

In [69]:
```python
# Three Reviews of cluster 7
count=1
for i in range(3):
    if i < len(cluster7):
        print('Review-%d : \n %s\n'%(count,cluster7[i]))
        count +=1
```

Review-1 :
 This is my second review request for this mustard. Again, I will say that it i
s my favorite mustard and hard to find in select stores. I was so happy to see
that I could purchase it on Amazon. This is sure a great place to shop from my
mustard to my hard to find shampoos. This mustard is wonderful on all kind of m
eats. My husband is not a mustard lover like me, but he does like this brand an
d selection.

Review-2 :
 For all you celiacs out there, this is the best cereal ever. You can eat with
milk in the am or dry in the afternoon as a snack. Great price point when you o
rder without shipping charge.

Review-3 :
 These arrived quickly but the box was really banged up. had to open a couple of
bags just to be sure the chips weren't too broken up since I'm selling these in
a vending machine, they were pretty crunched, I can't sell these.

In [70]:
```python
# Three Reviews of cluster 8
count=1
for i in range(3):
    if i < len(cluster8):
        print('Review-%d : \n %s\n'%(count,cluster8[i]))
        count +=1
```

Review-1 :
 Like others, I cannot order this product w/o an ingredients list. Amazon, peop
le with food allergies, intolerances, or other dietary restrictions need complet
e information! Thank you and please resend the G-F for Betty Crocker notice whe
n ingredients lists are included.

Review-2 :
 My son is 7 months old and this has always been his favorite baby food - it is
nice a thick, which he loves!

Review-3 :
 The dog food got good ratings at consumers, so I bought it and the dogs just lo
ve it.. It seems to be a very healtful and complete food.. I just add a little h
ot water and its gobbled up in no time

In [71]:
```python
# Three Reviews of cluster 9
count=1
for i in range(3):
    if i < len(cluster9):
        print('Review-%d : \n %s\n'%(count,cluster10[i]))
        count +=1
```

Review-1 :
 Smells tempting when its being brewed but the flavor just isnt there. I will n
ot purchase this again. I gave the other 2 bags away as gifts (because my roomm
ates weren't impressed with the flavor either).

Review-2 :
 Great drink for working out. I was paying $5 at 24 hour for each drink now got
a case of them for cheap.

Review-3 :
 Wanted to try the best coffee in the world. Recieved as promised and enjoyed fr
om the first cup. Would order again.

```
In [72]:  # Three Reviews of cluster 10
          count=1
          for i in range(3):
              if i < len(cluster10):
                  print('Review-%d : \n %s\n'%(count,cluster10[i]))
                  count +=1
```

Review-1 :
 Smells tempting when its being brewed but the flavor just isnt there. I will n
ot purchase this again. I gave the other 2 bags away as gifts (because my roomm
ates weren't impressed with the flavor either).

Review-2 :
 Great drink for working out. I was paying $5 at 24 hour for each drink now got
a case of them for cheap.

Review-3 :
 Wanted to try the best coffee in the world. Recieved as promised and enjoyed fr
om the first cup. Would order again.

TFIDF-Word2Vec

```
In [73]:  tf_idf_vect = TfidfVectorizer()

          # final_tf_idf1 is the sparse matrix with row= sentence, col=word and cell_val = tf
          idf
          final_tf_idf1 = tf_idf_vect.fit_transform(x)

          # tfidf words/col-names
          tfidf_feat = tf_idf_vect.get_feature_names()

          # compute TFIDF Weighted Word2Vec for each review for sent_x .
          tfidf_vectors = [];
          row=0;
          for sent in sent_x:
              sent_vec = np.zeros(50)
              weight_sum =0;
              for word in sent:
                  if word in w2v_words:
                      vec = w2v_model.wv[word]
                      # obtain the tf_idfidf of a word in a sentence/review
                      tf_idf = final_tf_idf1[row, tfidf_feat.index(word)]
                      sent_vec += (vec * tf_idf)
                      weight_sum += tf_idf
              if weight_sum != 0:
                  sent_vec /= weight_sum
              tfidf_vectors.append(sent_vec)
              row += 1

          data = tfidf_vectors
```

Hierarchical Clustering with 2 clusters

In [74]:
```python
model = AgglomerativeClustering(n_clusters=2).fit(data)

reviews = my_final['Text'].values
# Getting all the reviews in different clusters
cluster1 = []
cluster2 = []

for i in range(model.labels_.shape[0]):
    if model.labels_[i] == 0:
        cluster1.append(reviews[i])
    else :
        cluster2.append(reviews[i])


# Number of reviews in different clusters
print("No. of reviews in Cluster-1 : ",len(cluster1))
print("\nNo. of reviews in Cluster-2 : ",len(cluster2))
```

```
No. of reviews in Cluster-1 :  3830

No. of reviews in Cluster-2 :  1170
```

In [75]:
```python
# Three Reviews of cluster 1
count=1
for i in range(3):
    if i < len(cluster1):
        print('Review-%d : \n %s\n'%(count,cluster1[i]))
        count +=1
```

```
Review-1 :
 I like the mint, but the only reason I have bought these mint are because I lik
e the tin. The tins are very nice. There are a lot of project I can build using
these candy tins. The mint is not bad either, very freshing.<br /><br />They sel
l these mint for $2.50/tin here in San Francisco, so 12 tins will cost me $30, a
nd I saved about 10 dollars to get it here.

Review-2 :
 This is my second review request for this mustard.  Again, I will say that it i
s my favorite mustard and hard to find in select stores.  I was so happy to see
that I could purchase it on Amazon.  This is sure a great place to shop from my
mustard to my hard to find shampoos.  This mustard is wonderful on all kind of m
eats.  My husband is not a mustard lover like me, but he does like this brand an
d selection.

Review-3 :
 Great product. Great price. I use it as a substitute for white sugar in everyth
ing from my coffee to cooking and baking. I use Amazon's Subscribe and Save prog
ram for this product. Enjoy.
```

In [76]:
```python
# Three Reviews of cluster 2
count=1
for i in range(3):
    if i < len(cluster2):
        print('Review-%d : \n %s\n'%(count,cluster2[i]))
        count +=1
```

```
Review-1 :
 Smells tempting when its being brewed but the flavor just isnt there.  I will n
ot purchase this again.  I gave the other 2 bags away as gifts (because my roomm
ates weren't impressed with the flavor either).

Review-2 :
 This is a good deal.  Pack of 2, one for me and one to give as a gift.  Very ea
sy to use.  I use it at night to make a cup of decaf.  Makes good coffee and you
can make coffee your desired strength.  I use 1 scoop of coffee for every 5oz of
water.

Review-3 :
 I Have been ordering this k cup cappucino drink for awhile now and have never h
ad a problem! Love it tastes good!
```

Hierarchical Clustering with 5 clusters

```
In [77]: model = AgglomerativeClustering(n_clusters=5).fit(data)

         # Getting all the reviews in different clusters
         cluster1 = []
         cluster2 = []
         cluster3 = []
         cluster4 = []
         cluster5 = []

         for i in range(model.labels_.shape[0]):
             if model.labels_[i] == 0:
                 cluster1.append(reviews[i])
             elif model.labels_[i] == 1:
                 cluster2.append(reviews[i])
             elif model.labels_[i] == 2:
                 cluster3.append(reviews[i])
             elif model.labels_[i] == 3:
                 cluster4.append(reviews[i])
             else :
                 cluster5.append(reviews[i])

         # Number of reviews in different clusters
         print("No. of reviews in Cluster-1 : ",len(cluster1))
         print("\nNo. of reviews in Cluster-2 : ",len(cluster2))
         print("\nNo. of reviews in Cluster-3 : ",len(cluster3))
         print("\nNo. of reviews in Cluster-4 : ",len(cluster4))
         print("\nNo. of reviews in Cluster-5 : ",len(cluster5))
```

```
No. of reviews in Cluster-1 :  1170

No. of reviews in Cluster-2 :  812

No. of reviews in Cluster-3 :  2173

No. of reviews in Cluster-4 :  490

No. of reviews in Cluster-5 :  355
```

```
In [78]: # Three Reviews of cluster 1
         count=1
         for i in range(3):
             if i < len(cluster1):
                 print('Review-%d : \n %s\n'%(count,cluster1[i]))
                 count +=1
```

```
Review-1 :
 Smells tempting when its being brewed but the flavor just isnt there.  I will n
ot purchase this again.  I gave the other 2 bags away as gifts (because my roomm
ates weren't impressed with the flavor either).

Review-2 :
 This is a good deal.  Pack of 2, one for me and one to give as a gift.  Very ea
sy to use.  I use it at night to make a cup of decaf.  Makes good coffee and you
can make coffee your desired strength.  I use 1 scoop of coffee for every 5oz of
water.

Review-3 :
 I Have been ordering this k cup cappucino drink for awhile now and have never h
ad a problem! Love it tastes good!
```

In [79]:
```python
# Three Reviews of cluster 2
count=1
for i in range(3):
    if i < len(cluster2):
        print('Review-%d : \n %s\n'%(count,cluster2[i]))
        count +=1
```

Review-1 :
 I like the mint, but the only reason I have bought these mint are because I like the tin. The tins are very nice. There are a lot of project I can build using these candy tins. The mint is not bad either, very freshing.

They sell these mint for $2.50/tin here in San Francisco, so 12 tins will cost me $30, and I saved about 10 dollars to get it here.

Review-2 :
 These are just great. Even my dog liked them, stealing a pack off a coffee table and eating darned near the whole thing.

Review-3 :
 I was looking for yeast with no added ingredients, and this is not it...
Ingredients: Yeast, sorbitan monostearate, ascorbic acid
It also came hot off the UPS truck, and I hesitate to open it for fear it is activated and dead. Will return instead.
Not what I was looking for, won't purchase again.

In [80]:
```python
# Three Reviews of cluster 3
count=1
for i in range(3):
    if i < len(cluster3):
        print('Review-%d : \n %s\n'%(count,cluster3[i]))
        count +=1
```

Review-1 :
 This is my second review request for this mustard. Again, I will say that it is my favorite mustard and hard to find in select stores. I was so happy to see that I could purchase it on Amazon. This is sure a great place to shop from my mustard to my hard to find shampoos. This mustard is wonderful on all kind of meats. My husband is not a mustard lover like me, but he does like this brand and selection.

Review-2 :
 Great product. Great price. I use it as a substitute for white sugar in everything from my coffee to cooking and baking. I use Amazon's Subscribe and Save program for this product. Enjoy.

Review-3 :
 For all you celiacs out there, this is the best cereal ever. You can eat with milk in the am or dry in the afternoon as a snack. Great price point when you order without shipping charge.

In [81]:
```python
# Three Reviews of cluster 4
count=1
for i in range(3):
    if i < len(cluster4):
        print('Review-%d : \n %s\n'%(count,cluster4[i]))
        count +=1
```

Review-1 :
 These arrived quickly but the box was really banged up. had to open a couple of bags just to be sure the chips weren't too broken up since I'm selling these in a vending machine, they were pretty crunched, I can't sell these.

Review-2 :
 I have a serious addiction to this gum but cannot find it anywhere except onlin e. I go through at least a pack a day. I used to chew Peppermint or Spearmint or Winterfrost but I can't chew anything but Exotic Mint any more. The flavor is ju st right and it lasts forever.

Seriously, if you see this gum in a st ore, buy all of it. It's like little candy gold bullion. I have opted to pay twi ce as much online for this gum versus gum I can buy on the store shelves just so I can get this particular flavor. Who would do that if the gum weren't the best gum ever? Nobody. That's who.

I also wrote Wrigley to see if they'll send me a year-supply for free. Cross your fingers and wish me luck.

Review-3 :
 These cashews are of decent quality and at a good price. They are not always w hole as pointed out by another reviewer, but for the price they are a nice deal and taste fine.

In [82]:
```python
# Three Reviews of cluster 5
count=1
for i in range(3):
    if i < len(cluster5):
        print('Review-%d : \n %s\n'%(count,cluster5[i]))
        count +=1
```

Review-1 :
 My 8-month baby always screws up his face like he's being poisoned when you giv e him something new, so I thought he'd need some time to get used to this. Not true! He loved this from the first bite. And while he tends not to like foods with a lot of texture yet, the rice-in-sauce texture of this food doesn't bother him at all.

Review-2 :
 My son is 7 months old and this has always been his favorite baby food - it is nice a thick, which he loves!

Review-3 :
 The dog food got good ratings at consumers, so I bought it and the dogs just lo ve it.. It seems to be a very healtful and complete food.. I just add a little h ot water and its gobbled up in no time

Hierarchical Clustering with 10 clusters

```
In [83]:  model = AgglomerativeClustering(n_clusters=10).fit(data)
          # Getting all the reviews in different clusters
          cluster1 = []
          cluster2 = []
          cluster3 = []
          cluster4 = []
          cluster5 = []
          cluster6 = []
          cluster7 = []
          cluster8 = []
          cluster9 = []
          cluster10 = []

          for i in range(model.labels_.shape[0]):
              if model.labels_[i] == 0:
                  cluster1.append(reviews[i])
              elif model.labels_[i] == 1:
                  cluster2.append(reviews[i])
              elif model.labels_[i] == 2:
                  cluster3.append(reviews[i])
              elif model.labels_[i] == 3:
                  cluster4.append(reviews[i])
              elif model.labels_[i] == 4:
                  cluster5.append(reviews[i])
              elif model.labels_[i] == 5:
                  cluster6.append(reviews[i])
              elif model.labels_[i] == 6:
                  cluster7.append(reviews[i])
              elif model.labels_[i] == 7:
                  cluster8.append(reviews[i])
              elif model.labels_[i] == 8:
                  cluster9.append(reviews[i])
              else :
                  cluster10.append(reviews[i])
```

```
In [84]: print("No. of reviews in Cluster-1 : ",len(cluster1))
         print("\nNo. of reviews in Cluster-2 : ",len(cluster2))
         print("\nNo. of reviews in Cluster-3 : ",len(cluster3))
         print("\nNo. of reviews in Cluster-4 : ",len(cluster4))
         print("\nNo. of reviews in Cluster-5 : ",len(cluster5))
         print("\nNo. of reviews in Cluster-6 : ",len(cluster6))
         print("\nNo. of reviews in Cluster-7 : ",len(cluster7))
         print("\nNo. of reviews in Cluster-8 : ",len(cluster8))
         print("\nNo. of reviews in Cluster-9 : ",len(cluster9))
         print("\nNo. of reviews in Cluster-10 : ",len(cluster10))
```

```
No. of reviews in Cluster-1 :  330

No. of reviews in Cluster-2 :  1409

No. of reviews in Cluster-3 :  649

No. of reviews in Cluster-4 :  446

No. of reviews in Cluster-5 :  355

No. of reviews in Cluster-6 :  419

No. of reviews in Cluster-7 :  764

No. of reviews in Cluster-8 :  366

No. of reviews in Cluster-9 :  71

No. of reviews in Cluster-10 :   191
```

```
In [85]: # Three Reviews of cluster 1
         count=1
         for i in range(3):
             if i < len(cluster1):
                 print('Review-%d : \n %s\n'%(count,cluster1[i]))
                 count +=1
```

```
Review-1 :
 This is a good deal.  Pack of 2, one for me and one to give as a gift.  Very ea
sy to use.  I use it at night to make a cup of decaf.  Makes good coffee and you
can make coffee your desired strength.  I use 1 scoop of coffee for every 5oz of
water.

Review-2 :
 This water taste great. We love the bottle. Perfect water for everyone :) I lov
e that it is 1.5 liters because I am trying to drink more water. You will love t
his water.

Review-3 :
 This English Breakfast tea from Higgins & Burke, is good: it's basic, good tast
ing black tea (English Breakfast blend) and is tasty.  I wasn't blown away by th
e flavor, but it was good: I enjoyed the tea, it brewed up quickly in these nice
tea bags (I recommend about one minute of brew time), and then added a splash of
cream: it was quite nice!
```

In [86]:
```python
# Three Reviews of cluster 2
count=1
for i in range(3):
    if i < len(cluster2):
        print('Review-%d : \n %s\n'%(count,cluster2[i]))
        count +=1
```

Review-1 :
 Kitchens of India is the closest to a homemade curry we have ever had. We love
it. The sauce is great and it is fast, so curry a few times a week is possible.
We often make a curry from scratch but it is time consuming and now we can have
a good curry whenever we want. We have also tried the Rogan Josh as well as the
Korma. Add some onions and raisins, or a bit of honey. The ingredients have litt
le or no preservatives. The Best!!

Review-2 :
 Teff in general is a delicious healthy grain, and this Teff is no different.
Cook according to pkg. directions and add to morning oatmeal for a hearty heal
thy breakfast.
If I had one wish it would be for it to be "Organic Brown Te
ff."

Review-3 :
 I have this item to ship monthly. I think this is the best honey that I have f
ound. I tried almost all local honeys, I wanted uncook & unfiltered honey for t
he best health benefit. I found this is by far the best flavor and consistency.
My honey arrives on time and in good condition.

In [87]:
```python
# Three Reviews of cluster 3
count=1
for i in range(3):
    if i < len(cluster3):
        print('Review-%d : \n %s\n'%(count,cluster3[i]))
        count +=1
```

Review-1 :
 Smells tempting when its being brewed but the flavor just isnt there. I will n
ot purchase this again. I gave the other 2 bags away as gifts (because my roomm
ates weren't impressed with the flavor either).

Review-2 :
 I Have been ordering this k cup cappucino drink for awhile now and have never h
ad a problem! Love it tastes good!

Review-3 :
 Preparation could not be easier. The flavor was wonderful. My son who almost
never salts anything found it a bit salty for him. My Mom who should not eat sa
lt said she thought it was a bit salty too.

```python
In [88]:   # Three Reviews of cluster 4
           count=1
           for i in range(3):
               if i < len(cluster4):
                   print('Review-%d : \n %s\n'%(count,cluster4[i]))
                   count +=1
```

```
Review-1 :
 I like the mint, but the only reason I have bought these mint are because I lik
e the tin. The tins are very nice. There are a lot of project I can build using
these candy tins. The mint is not bad either, very freshing.<br /><br />They sel
l these mint for $2.50/tin here in San Francisco, so 12 tins will cost me $30, a
nd I saved about 10 dollars to get it here.

Review-2 :
 These are always going to be more expensive than buying the single big pack - r
emember when oreos were packed into three plastic tubes in a plastic box.<br /><
br />These oreos always come out crunchy with a very pliable filling that can be
easily nudged-rolled off if you don't like the filling.  If Nabisco sold individ
ually wrapped cookies, I'd pay the extra money as I usually only want one cookie
.<br /><br />I will always choose to buy oreos packaged in as small a serving as
possible despite the expense unless I get some kind of vacuum sealer.

Review-3 :
 When I purchaed this item, I thought it will be normal size lollipop... But whe
n I opened, I was really disappointed.. It's so tiny(I'm talking about finger si
ze lollipop)  it's not worth spending That much money.. I should of just went to
candy store... I could of found better lollipop! And With that money I could of
found more and left with some change!,
```

```python
In [89]:   # Three Reviews of cluster 5
           count=1
           for i in range(3):
               if i < len(cluster5):
                   print('Review-%d : \n %s\n'%(count,cluster5[i]))
                   count +=1
```

```
Review-1 :
 My 8-month baby always screws up his face like he's being poisoned when you giv
e him something new, so I thought he'd need some time to get used to this.  Not
true!  He loved this from the first bite.  And while he tends not to like foods
with a lot of texture yet, the rice-in-sauce texture of this food doesn't bother
him at all.

Review-2 :
 My son is 7 months old and this has always been his favorite baby food - it is
nice a thick, which he loves!

Review-3 :
 The dog food got good ratings at consumers, so I bought it and the dogs just lo
ve it.. It seems to be a very healtful and complete food.. I just add a little h
ot water and its gobbled up in no time
```

In [90]:
```python
# Three Reviews of cluster 6
count=1
for i in range(3):
    if i < len(cluster6):
        print('Review-%d : \n %s\n'%(count,cluster6[i]))
        count +=1
```

Review-1 :
 These arrived quickly but the box was really banged up. had to open a couple of
bags just to be sure the chips weren't too broken up since I'm selling these in
a vending machine, they were pretty crunched, I can't sell these.

Review-2 :
 I have a serious addiction to this gum but cannot find it anywhere except onlin
e. I go through at least a pack a day. I used to chew Peppermint or Spearmint or
Winterfrost but I can't chew anything but Exotic Mint any more. The flavor is ju
st right and it lasts forever.

Seriously, if you see this gum in a st
ore, buy all of it. It's like little candy gold bullion. I have opted to pay twi
ce as much online for this gum versus gum I can buy on the store shelves just so
I can get this particular flavor. Who would do that if the gum weren't the best
gum ever? Nobody. That's who.

I also wrote Wrigley to see if they'll
send me a year-supply for free. Cross your fingers and wish me luck.

Review-3 :
 These cashews are of decent quality and at a good price. They are not always w
hole as pointed out by another reviewer, but for the price they are a nice deal
and taste fine.

In [91]:
```python
# Three Reviews of cluster 7
count=1
for i in range(3):
    if i < len(cluster7):
        print('Review-%d : \n %s\n'%(count,cluster7[i]))
        count +=1
```

Review-1 :
 This is my second review request for this mustard. Again, I will say that it i
s my favorite mustard and hard to find in select stores. I was so happy to see
that I could purchase it on Amazon. This is sure a great place to shop from my
mustard to my hard to find shampoos. This mustard is wonderful on all kind of m
eats. My husband is not a mustard lover like me, but he does like this brand an
d selection.

Review-2 :
 Great product. Great price. I use it as a substitute for white sugar in everyth
ing from my coffee to cooking and baking. I use Amazon's Subscribe and Save prog
ram for this product. Enjoy.

Review-3 :
 For all you celiacs out there, this is the best cereal ever. You can eat with
milk in the am or dry in the afternoon as a snack. Great price point when you o
rder without shipping charge.

In [92]:
```python
# Three Reviews of cluster 8
count=1
for i in range(3):
    if i < len(cluster8):
        print('Review-%d : \n %s\n'%(count,cluster8[i]))
        count +=1
```

Review-1 :
 These are just great. Even my dog liked them, stealing a pack off a coffee tab
le and eating darned near the whole thing.

Review-2 :
 I was looking for yeast with no added ingredients, and this is not it...
I
ngredients: Yeast, sorbitan monostearate, ascorbic acid
It also came hot of
f the UPS truck, and I hesitate to open it for fear it is activated and dead. W
ill return instead.
Not what I was looking for, won't purchase again.

Review-3 :
 We love Sunbutter! I don't have to worry about using it, as it and the company
is completely PEANUT FREE!. Sunbutter taste great, unlike Soy Butter which is gr
itty and awful tasting. I use Sunbutter for anything that I would use peanut but
ter for cookies, frostings and candies, dips, etc. Sunbutter has the same consis
tency and texture as peanut butter and doesn't need refrigeration. When using Su
nbutter for baking you do have to cut back to about 1/2 of the baking soda and/o
r baking powder or you can add a little bit of lemon juice to prevent your baked
goods from turning green from the sunflower seeds. If your cookies do turn green
it doesn't affect the taste just the looks (guess how I know that little bit of
information :). Sunbutter also has a great website at sunbutter dot com with rec
ipes and information about their products. Check out their site and defiantly ge
t some Sunbutter.

In [93]:
```python
# Three Reviews of cluster 9
count=1
for i in range(3):
    if i < len(cluster9):
        print('Review-%d : \n %s\n'%(count,cluster9[i]))
        count +=1
```

Review-1 :
 My daughter eats these bars on a daily basis. I usually purchase these at the
local grocery store and thought is would be cost effective to order through Amaz
on. My grocery store charges $.99 per bar. Amazon's price is $1.44 per bar.

Review-2 :
 These sell in organic shops and deli's for twice what you will pay on Amazon.
They are always a hit at parties or family get togethers so you will need a case
which is what Amazon sells.

Review-3 :
 I ordered the 10-bag product and only received 6 bags. Do not order from this
company as it does not fulfill it's obligations to their customers. I could hav
e tolerated the 10-day wait for the order to come in if it had been exactly what
I ordered, but it wasn't.

In [94]:
```python
# Three Reviews of cluster 10
count=1
for i in range(3):
    if i < len(cluster10):
        print('Review-%d : \n %s\n'%(count,cluster10[i]))
        count +=1
```

Review-1 :
 Very strong Ginger flavor, mild carbonation, I like that. You must like Ginger
Ale to like this. The small bottles are about the right size, as their is more f
lavor in 7 Ounce bottle then anyone else has in 12.

Review-2 :
 Even the "plain" flavor was much better than I expected. I had my heart set on
getting the Vanilla Bean, but it was often out of stock. I opted for to try othe
r flavors to test others and was surprised by how good tasting they all were. I
actually enjoy matching my workout moods with the different flavors! Now I will
just get this mixed flavors box every time.

Review-3 :
 This is definitely my favorite new drink mix in the Mocktail line by Crystal Li
ght. It's a great tasting lime flavored drink. The artificial sweetener gives
this a very slight aftertaste, but just slight. And I like the fact that it is
only 5 calories per serving. Each pouch makes enough to fill a large pitcher.
There isn't any type of alcohol flavor, which is fine with me. It's a very refr
eshing drink. I can drink this stuff all day!

In [ ]: