

In [6]:

```
# Sentiment Analysis
```

In [7]:

```
### Import Important Libraries
```

In [51]:

```
import pandas as pd
import numpy as np
import pickle
import os
import io
import re
from sys import path
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.naive_bayes import MultinomialNB
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelBinarizer
import matplotlib.pyplot as plt
from string import punctuation,digits
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk.tokenize import RegexpTokenizer
import wordcloud
from PIL import Image
from wordcloud import WordCloud,STOPWORDS
```

In [9]:

```
# Load the data

#Amazon Data
inpu_file="amazon_cells_labelled.txt"
amazon=pd.read_csv(inpu_file,sep='\t')
amazon.columns=['Sentence','Class']
amazon.head()
```

Out[9]:

	Sentence	Class
0	Good case, Excellent value.	1
1	Great for the jawbone.	1
2	Tied to charger for conversations lasting more...	0
3	The mic is great.	1
4	I have to jiggle the plug to get it to line up...	0

In [10]:

```
#Imdb
inpu_file="imdb_labelled (1).txt"
imdb=pd.read_csv(inpu_file,sep='\t')
imdb.columns=['Sentence','Class']
imdb.head()
```

Out[10]:

	Sentence	Class
0	Not sure who was more lost - the flat characte...	0
1	Attempting artiness with black & white and cle...	0
2	Very little music or anything to speak of.	0
3	The best scene in the movie was when Gerardo i...	1
4	The rest of the movie lacks art, charm, meanin...	0

In [11]:

```
#Yelp
inpu_file='yelp_labelled.txt'
yelp=pd.read_csv(inpu_file,sep='\t')
yelp.columns=['Sentence','Class']
yelp.head()
```

Out[11]:

	Sentence	Class
0	Crust is not good.	0
1	Not tasty and the texture was just nasty.	0
2	Stopped by during the late May bank holiday of...	1
3	The selection on the menu was great and so wer...	1
4	Now I am getting angry and I want my damn pho.	0

In [12]:

```
# Create a new Dataframe using all the three input files
data=pd.DataFrame()
data=pd.concat([amazon,imdb,yelp])
data['index']=data.index
data
```

Out[12]:

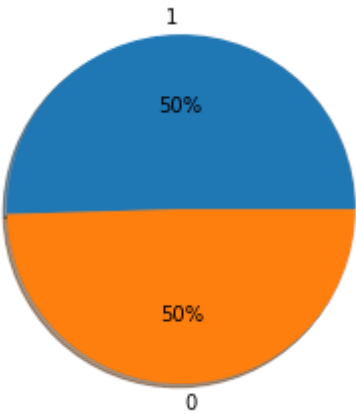
	Sentence	Class	index
0	Good case, Excellent value.	1	0
1	Great for the jawbone.	1	1
2	Tied to charger for conversations lasting more...	0	2
3	The mic is great.	1	3
4	I have to jiggle the plug to get it to line up...	0	4
...	...	...	...
994	I think food should have flavor and texture an...	0	994
995	Appetite instantly gone.	0	995
996	Overall I was not impressed and would not go b...	0	996
997	The whole experience was underwhelming, and I ...	0	997
998	Then, as if I hadn't wasted enough of my life ...	0	998

2745 rows × 3 columns

In [13]:

```
#Checking if the data is well balanced?

plt.pie(data['Class'].value_counts(),autopct='%0f%%',shadow=True,labels=data['Class'].
unique())
plt.show()
```



In [14]:

```
## Text Processing Steps

#Lower String
#Remove Email Address
#Remove IP Address
#Remove Punctuation and Special Characters
# Remove Numbers
#Remove Stop words
```

In [15]:

```
#Text Preprocessing
columns=["index","Class","Sentence"]
df=pd.DataFrame(columns=columns)

#Lower String
data['Sentence']=[term.lower() for term in data['Sentence']]
data['Sentence']
```

Out[15]:

```
0          good case, excellent value.
1          great for the jawbone.
2    tied to charger for conversations lasting more...
3          the mic is great.
4    i have to jiggle the plug to get it to line up...
...
994  i think food should have flavor and texture an...
995          appetite instantly gone.
996  overall i was not impressed and would not go b...
997  the whole experience was underwhelming, and i ...
998  then, as if i hadn't wasted enough of my life ...
Name: Sentence, Length: 2745, dtype: object
```

In [16]:

```
#Remove Email Address
data['Sentence']=data['Sentence'].replace('[a-zA-Z0-9-_.]+@[a-zA-Z0-9-_.]+',' ',regex=True)
data['Sentence']
```

Out[16]:

```
0          good case, excellent value.
1          great for the jawbone.
2    tied to charger for conversations lasting more...
3          the mic is great.
4    i have to jiggle the plug to get it to line up...
...
994  i think food should have flavor and texture an...
995          appetite instantly gone.
996  overall i was not impressed and would not go b...
997  the whole experience was underwhelming, and i ...
998  then, as if i hadn't wasted enough of my life ...
Name: Sentence, Length: 2745, dtype: object
```

In [17]:

```
#Remove IP Address
```

```
data['Sentence']=data['Sentence'].replace('((25[0-5]|2[0-4][0-9]|[01][0-9][0-9]?)(\.|
$)){4}',"",regex=True)
data['Sentence']
```

Out[17]:

```
0          good case, excellent value.
1          great for the jawbone.
2    tied to charger for conversations lasting more...
3          the mic is great.
4    i have to jiggle the plug to get it to line up...
...
994  i think food should have flavor and texture an...
995          appetite instantly gone.
996  overall i was not impressed and would not go b...
997  the whole experience was underwhelming, and i ...
998  then, as if i hadn't wasted enough of my life ...
Name: Sentence, Length: 2745, dtype: object
```

In [18]:

```
#Remove Punctuation and special Characters
```

```
data['Sentence']=data['Sentence'].replace('[^\w\s]','',regex=True)
data['Sentence']
```

Out[18]:

```
0          good case excellent value
1          great for the jawbone
2    tied to charger for conversations lasting more...
3          the mic is great
4    i have to jiggle the plug to get it to line up...
...
994  i think food should have flavor and texture an...
995          appetite instantly gone
996  overall i was not impressed and would not go back
997  the whole experience was underwhelming and i t...
998  then as if i hadnt wasted enough of my life th...
Name: Sentence, Length: 2745, dtype: object
```

In [19]:

```
#Remove Numbers
data['Sentence']=data['Sentence'].replace('\d','',regex=True)
data['Sentence']
```

Out[19]:

```
0                good case excellent value
1                great for the jawbone
2    tied to charger for conversations lasting more...
3                the mic is great
4    i have to jiggle the plug to get it to line up...
...
994    i think food should have flavor and texture an...
995                appetite instantly gone
996    overall i was not impressed and would not go back
997    the whole experience was underwhelming and i t...
998    then as if i hadnt wasted enough of my life th...
Name: Sentence, Length: 2745, dtype: object
```

In [20]:

```
stop_nltk=stopwords.words("English")
```

In [21]:

```
#remove stop words

#We can do this way but it would be a lengthy one(tokenize then remove stop words) so
go with the next cell as its a short cut :P

def del_stop(ind):
    res=[term for term in data['Sentence'] if term not in stop_nltk]
    return res
```

In [22]:

```
for index,row in data.iterrows():
    word_tokens=word_tokenize(row['Sentence'])
    filtered_sentence=[w for w in word_tokens if not w in stop_nltk]
    df=df.append({"index":row['index'], "Class":row['Class'], "Sentence": " ".join(filtere
d_sentence[0:])}),ignore_index=True)
data=df
```

In [23]:

```
data
```

Out[23]:

	index	Class	Sentence
0	0	1	good case excellent value
1	1	1	great jawbone
2	2	0	tied charger conversations lasting minutesmajo...
3	3	1	mic great
4	4	0	jiggle plug get line right get decent volume
...	...	...	...
2740	994	0	think food flavor texture lacking
2741	995	0	appetite instantly gone
2742	996	0	overall impressed would go back
2743	997	0	whole experience underwhelming think well go n...
2744	998	0	hadnt wasted enough life poured salt wound dra...

2745 rows × 3 columns

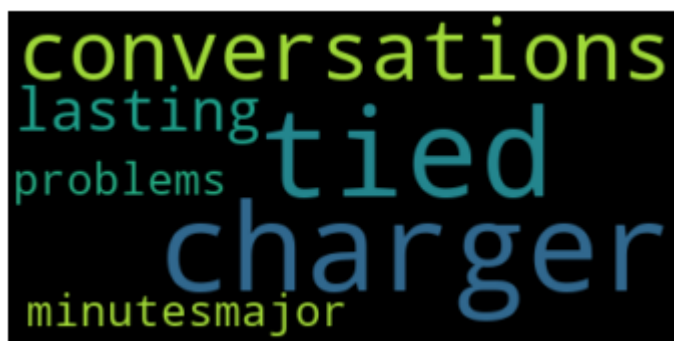
In [44]:

```
#WORDCLOUD
```

In [40]:

```
# Word Cloud for only one Review

text=data['Sentence'][2]
wordcloud=WordCloud().generate(text)
plt.imshow(wordcloud,interpolation='bilinear')
plt.axis('off')
plt.show()
```



In [ ]:

```
# Word Cloud for all the Reviews
```

In [48]:

```
text=" ".join(review for review in data['Sentence'])
print('There are {} words in the combination of all review : '.format(len(text)))

wordcloud=WordCloud(background_color='White',min_font_size=10).generate(text)

plt.imshow(wordcloud,interpolation='bilinear')
plt.axis('off')
plt.show()
```

There are 125910 words in the combination of all review :





In [69]:

```
#NOW if we want to change the squaaed Word Cloud to A wine bottle word cloud

wine_mask = np.array(Image.open("wine_mask.png"))
wine_mask

def transform_format(val):
    if val == 0:
        return 255
    else:
        return val

transformed_wine_mask = np.ndarray((wine_mask.shape[0],wine_mask.shape[1]), np.int32)

for i in range(len(wine_mask)):
    transformed_wine_mask[i] = list(map(transform_format, wine_mask[i]))

transformed_wine_mask

# Create a word cloud image
wc = WordCloud(background_color="white", max_words=1000, mask=transformed_wine_mask,
               contour_width=3, contour_color='firebrick')

# Generate a wordcloud
wc.generate(text)

# store to file
wc.to_file("wine1.png")

# show
plt.figure(figsize=[20,10])
plt.imshow(wc, interpolation='bilinear')
plt.axis("off")
plt.show()
```



In [72]:

```
# Train Test Split
x_train,x_test,y_train,y_test=train_test_split(data['Sentence'].values.astype('U'),data
['Class'].values.astype('int32'),test_size=0.10,random_state=2)
classes=data['Class'].unique()
```

In [73]:

```
#Training the model
# Convert the text into a TFIDF Representation
# Train the ML Model
```

In [156]:

```
from sklearn.metrics import confusion_matrix
from sklearn.model_selection import cross_val_score
from sklearn.metrics import accuracy_score
from sklearn.neural_network import MLPClassifier
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import SGDClassifier

#Grid Search Result
vectorizer=TfidfVectorizer(analyzer='word',ngram_range=(1,3),max_features=50000,max_df=
0.5,use_idf=True,norm='l1')
#max_df only consider those words whose tf-idf value is greater than 0.5
#use_idf =true is not only tf do idf too

counts=vectorizer.fit_transform(x_train)
vocab=vectorizer.vocabulary_

classifier=SGDClassifier(alpha=1e-05,max_iter=50,penalty='l1')
classifier=classifier.fit(counts,y_train)
example_counts=vectorizer.transform(x_test)

predictions=classifier.predict(example_counts)
```

In [157]:

```
from sklearn.metrics import precision_score
from sklearn.metrics import recall_score
from sklearn.metrics import classification_report

#Model Evaluation
acc=accuracy_score(y_test,predictions,normalize=True)
hit=precision_score(y_test,predictions,average=None,labels=classes)
capture=recall_score(y_test,predictions,average=None,labels=classes)

print('Model Accuracy :%2f'%acc)

print(classification_report(y_test,predictions))
```

Model Accuracy :0.821818

	precision	recall	f1-score	support
0	0.84	0.80	0.82	141
1	0.80	0.84	0.82	134
accuracy			0.82	275
macro avg	0.82	0.82	0.82	275
weighted avg	0.82	0.82	0.82	275

In [118]:

```
hit
```

Out[118]:

```
array([0.79577465, 0.84210526])
```

In [119]:

```
capture
```

Out[119]:

```
array([0.84328358, 0.79432624])
```

In [132]:

```
# TO see the details of classification report
```

```
classification_report?
```

In [137]:

```
feature_name=vectorizer.get_feature_names()  
feature_name
```

Out[137]:

```
['aailiyah',  
'aailiyah pretty',  
'aailiyah pretty good',  
'aailiyah pretty good akasha',  
'abandoned',  
'abandoned factory',  
'abandoned factory ready',  
'abandoned factory ready poorly',  
'abhor',  
'ability',  
'ability actually',  
'ability actually know',  
'ability actually know receiving',  
'ability dwight',  
'ability dwight schultz',  
'ability meld',  
'ability meld two',  
'ability meld two highly',  
'ability phone',  
'ability pull',  
'ability pull part',  
'able',  
'able roam',  
'able roam around',  
'able roam around house',  
'able use',  
'able use one',  
'able use one headset',  
'able voice',  
'able voice dialing',  
'able voice dialing car',  
'abovepretty',  
'abovepretty useless',  
'abroad',  
'abroad interacting',  
'abroad interacting nationalities',  
'abroad interacting nationalities although',  
'absolute',  
'absolute must',  
'absolute must visit',  
'absolutel',  
'absolutel junk',  
'absolutely',  
'absolutely abysmal',  
'absolutely amazing',  
'absolutely appalling',  
'absolutely back',  
'absolutely clue',  
'absolutely clue meals',  
'absolutely clue meals contain',  
'absolutely delicious',  
'absolutely delicious time',  
'absolutely flatlined',  
'absolutely flatlined excuse',  
'absolutely flatlined excuse restaurant',  
'absolutely flavor',  
'absolutely flavor meat',  
'absolutely flavor meat totally',  
'absolutely great',
```

'absolutely horrible',  
'absolutely horrible reception',  
'absolutely horrible reception apartment',  
'absolutely loved',  
'absolutely nothing',  
'absolutely nothing whatsoever',  
'absolutely problem',  
'absolutely problem headset',  
'absolutely problem headset linking',  
'absolutely ray',  
'absolutely ray charles',  
'absolutely ray charles performance',  
'absolutely recommend',  
'absolutely recommend movie',  
'absolutely recommend movie anyone',  
'absolutely stars',  
'absolutely suspense',  
'absolutely suspense buildup',  
'absolutely suspense buildup slightest',  
'absolutely warmth',  
'absolutely warmth charm',  
'absolutely warmth charm scenes',  
'absolutley',  
'absolutley fantastic',  
'abstruse',  
'abstruse culture',  
'abstruse culture flat',  
'abstruse culture flat reenactments',  
'abysmal',  
'abysmal everything',  
'abysmal everything stinks',  
'abysmal everything stinks trouble',  
'ac',  
'ac charger',  
'ac charger included',  
'ac charger included make',  
'accents',  
'accents absolutely',  
'accents absolutely abysmal',  
'accept',  
'accept anything',  
'accept anything except',  
'accept anything except cbr',  
'accept cassette',  
'acceptable',  
'access',  
'access fine',  
'access fine rare',  
'access fine rare instance',  
'access phone',  
'access phone handy',  
'access phone handy detachable',  
'accessible',  
'accessing',  
'accessing internet',  
'accessing internet downloading',  
'accessing internet downloading ringtones',  
'accessory',  
'accessory good',  
'accessory good manufacturer',  
'accessoryone',

'accident',  
'accident could',  
'accident could happier',  
'accidentally',  
'accidentally activate',  
'accidentally activate gentletouch',  
'accidentally activate gentletouch buttons',  
'accidentally touch',  
'accidentally touch phone',  
'accidentally touch phone face',  
'acclaimed',  
'acclaimed novella',  
'acclaimed novella dead',  
'accommodations',  
'accommodations vegetarian',  
'accommodations vegetarian daughter',  
'accomodate',  
'accomodate veganveggie',  
'accomodate veganveggie options',  
'accompanied',  
'accompanied software',  
'accompanied software almost',  
'accompanied software almost brilliant',  
'according',  
'according people',  
'according people called',  
'according people called applifies',  
'accordingly',  
'accountant',  
'accountant know',  
'accountant know im',  
'accountant know im getting',  
'accurate',  
'accurate portrayal',  
'accurate portrayal night',  
'accurate portrayal night raver',  
'accurately',  
'accurately defined',  
'accurately defined pretentious',  
'accurately defined pretentious crap',  
'accused',  
'accused murder',  
'accused murder pure',  
'accused murder pure horror',  
'accused murdering',  
'accused murdering white',  
'accused murdering white man',  
'ache',  
'ache felt',  
'ache felt sick',  
'ache felt sick rest',  
'achievement',  
'achievement made',  
'achievement made timely',  
'achievement made timely perhaps',  
'achievement sobaditsgood',  
'achievement sobaditsgood sobaditsmemorable',  
'achievement sobaditsgood sobaditsmemorable movie',  
'achille',  
'achille philippa',  
'achille philippa beautifully',



'achille philippa beautifully sing',  
'ackerman',  
'ackerman totally',  
'ackerman totally believable',  
'ackerman totally believable great',  
'acknowledged',  
'acknowledged another',  
'acknowledged another minutes',  
'acknowledged another minutes get',  
'across',  
'act',  
'act film',  
'act film glad',  
'act film glad youre',  
'act one',  
'act one least',  
'act one least scary',  
'acted',  
'acted done',  
'acted done tv',  
'acted done tv movie',  
'acted played',  
'acted played man',  
'acted played man ray',  
'acting',  
'acting ability',  
'acting ability dwight',  
'acting ability dwight schultz',  
'acting absolutely',  
'acting absolutely appalling',  
'acting action',  
'acting action location',  
'acting action location work',  
'acting alongside',  
'acting alongside olivia',  
'acting alongside olivia de',  
'acting bad',  
'acting bad dialogs',  
'acting bad dialogs extremely',  
'acting bad kind',  
'acting bad kind word',  
'acting bad really',  
'acting bad really bad',  
'acting bad script',  
'acting better',  
'acting beyond',  
'acting beyond abysmal',  
'acting beyond abysmal everything',  
'acting blah',  
'acting cinematography',  
'acting cinematography solid',  
'acting coach',  
'acting coach fascinating',  
'acting decent',  
'acting decent budget',  
'acting decidely',  
'acting decidely wooden',  
'acting decidely wooden though',  
'acting effective',  
'acting effective music',  
'acting effective music riz',

'acting especially',  
'acting especially lead',  
'acting especially lead professor',  
'acting fantastic',  
'acting fantastic stories',  
'acting fantastic stories seamlessly',  
'acting helps',  
'acting helps writing',  
'acting helps writing along',  
'acting involved',  
'acting involved includes',  
'acting involved includes like',  
'acting least',  
'acting least type',  
'acting least type film',  
'acting like',  
'acting like watching',  
'acting like watching wooden',  
'acting look',  
'acting look film',  
'acting look film added',  
'acting main',  
'acting main person',  
'acting main person low',  
'acting make',  
'acting make one',  
'acting make one plain',  
'acting mess',  
'acting mess script',  
'acting mess script sinking',  
'acting mostly',  
'acting mostly lesserknown',  
'acting mostly lesserknown cast',  
'acting poor',  
'acting poor plot',  
'acting poor plot lines',  
'acting postproduction',  
'acting postproduction editing',  
'acting postproduction editing directing',  
'acting skilled',  
'acting stanwycks',  
'acting stanwycks singing',  
'acting story',  
'acting story characters',  
'acting story characters supposedly',  
'acting sucked',  
'acting sucked concert',  
'acting sucked concert sequences',  
'acting sucks',  
'acting sucks music',  
'acting sucks music sucks',  
'acting superb',  
'acting superb tom',  
'acting superb tom wilkinson',  
'acting terrible',  
'acting terrible writing',  
'acting terrible writing worse',  
'acting utterly',  
'acting utterly predictable',  
'acting utterly predictable bad',  
'acting weak',

'acting weak around',  
'acting weak around pm',  
'acting whole',  
'acting whole cast',  
'acting whole cast could',  
'acting youd',  
'acting youd expect',  
'acting youd expect cast',  
'actingeven',  
'actingeven professionals',  
'actingeven professionals drago',  
'actingeven professionals drago debbie',  
'actingwise',  
'actingwise either',  
'actingwise either surprisingly',  
'actingwise either surprisingly solid',  
'action',  
'action check',  
'action check directors',  
'action check directors filmography',  
'action location',  
'action location work',  
'action location work balance',  
'action movie',  
'action movie screenwriter',  
'action movie screenwriter air',  
'action movies',  
'action movies ive',  
'action movies ive ever',  
'action scenes',  
'action scenes little',  
'action scenes little comedy',  
'action suspense',  
'action suspense less',  
'action suspense less unneeded',  
'actions',  
'actions minutes',  
'activate',  
'activate gentletouch',  
'activate gentletouch buttons',  
'activate gentletouch buttons accidentally',  
'activated',  
'activated suddenly',  
'activated suddenly died',  
'actor',  
'actor dads',  
'actor dads favourite',  
'actor dads favourite actor',  
'actor enjoyed',  
'actor enjoyed reading',  
'actor enjoyed reading book',  
'actor grew',  
'actor grew watching',  
'actor grew watching films',  
'actor played',  
'actor played supporting',  
'actor played supporting roles',  
'actor playing',  
'actor playing villain',  
'actor playing villain low',  
'actor screen',

'actor scripting',  
'actor scripting subtle',  
'actor scripting subtle comedy',  
'actors',  
'actors actresses',  
'actors actresses bonus',  
'actors actresses bonus showthese',  
'actors agreed',  
'actors agreed film',  
'actors around',  
'actors awkwardly',  
'actors awkwardly babbling',  
'actors awkwardly babbling overwrought',  
'actors cant',  
'actors cant even',  
'actors cant even tell',  
'actors deliver',  
'actors deliver sharply',  
'actors deliver sharply scripted',  
'actors give',  
'actors give wonderful',  
'actors give wonderful performance',  
'actors havent',  
'actors havent seen',  
'actors like',  
'actors like thomerson',  
'actors like thomerson james',  
'actors master',  
'actors master director',  
'actors master director significant',  
'actors nostalgia',  
'actors nostalgia trip',  
'actors period',  
'actors screen',  
'actors screen think',  
'actors screen think must',  
'actors telly',  
'actors telly savalas',  
'actors telly savalas peter',  
'actors truly',  
'actors truly understand',  
'actors truly understand become',  
'actors twice',  
'actors twice much',  
'actors twice much worry',  
'actorsan',  
'actorsan ugly',  
'actorsan ugly cartoon',  
'actorsan ugly cartoon crafted',  
'actress',  
'actress repeating',  
'actress repeating robotic',  
'actress repeating robotic face',  
'actress worse',  
'actress worse used',  
'actress worse used june',  
'actresses',  
'actresses bonus',  
'actresses bonus showthese',  
'actresses bonus showthese characters',  
'actresses think',

'actresses think superb',  
'actresses think superb job',  
'actual',  
'actual experience',  
'actual experience bit',  
'actual experience bit disappointing',  
'actually',  
'actually gets',  
'actually gets watching',  
'actually gets watching people',  
'actually graphics',  
'actually graphics good',  
'actually graphics good time',  
'actually know',  
'actually know receiving',  
'actually know receiving call',  
'actually lot',  
'actually lot art',  
'actually outperform',  
'actually outperform original',  
'actually outperform original battery',  
'actually pretty',  
'actually pretty cool',  
'actually rare',  
'actually rare give',  
'actually rare give star',  
'actually smart',  
'actually smart movie',  
'actually try',  
'actually try use',  
'actually try use phonemy',  
'actually turned',  
'actually turned pretty',  
'actually turned pretty decent',  
'actually turns',  
'actually turns pda',  
'actually turns pda realworld',  
'actually worth',  
'actually worth seeing',  
'actually worth seeing reason',  
'ad',  
'adams',  
'adams unfortunate',  
'adams unfortunate life',  
'adams unfortunate life subject',  
'adaptation',  
'adaptation james',  
'adaptation james joyces',  
'adaptation james joyces acclaimed',  
'adapter',  
'adapter provide',  
'adapter provide enough',  
'adapter provide enough charging',  
'adapters',  
'adapters time',  
'add',  
'add betty',  
'add betty white',  
'add betty white jean',  
'add kind',  
'add kind boost',

'add kind boost reception',  
'added',  
'added bonuses',  
'added gratuity',  
'added gratuity bill',  
'added gratuity bill since',  
'addition',  
'addition feels',  
'addition feels amp',  
'addition feels amp looks',  
'addition giallo',  
'addition giallo subgenre',  
'addition new',  
'addition new scenes',  
'addition new scenes usual',  
'addition one',  
'addition one lovely',  
'addition one lovely songs',  
'additional',  
'additional costs',  
'additional costs owning',  
'additional costs owning official',  
'additional ear',  
'additional ear gels',  
'additional ear gels provided',  
'address',  
'address book',  
'address book treo',  
'address book treo rebootsoverall',  
'adhesive',  
'admins',  
'admiration',  
'admiration swords',  
'admiration swords etc',  
'admiration swords etc bore',  
'admitted',  
'admitted elsewhere',  
'adorable',  
'adorable eye',  
'adorable eye catching',  
'adorablethe',  
'adorablethe junkyard',  
'adorablethe junkyard scenes',  
'adorablethe junkyard scenes funnyall',  
'adrift',  
'adrift stagy',  
'adrift stagy rest',  
'adrift stagy rest sits',  
'adventure',  
'adventure best',  
'advertised',  
'advise',  
'advise anyone',  
'advise anyone go',  
'advise anyone go see',  
'advise look',  
'advise purchase',  
'advise purchase item',  
'advise purchase item never',  
'aerial',  
'aerial bombardments',

'aerial bombardments london',  
'aerial bombardments london outward',  
'aerial scenes',  
'aerial scenes ought',  
'aerial scenes ought thrilled',  
'aesthetically',  
'aesthetically like',  
'aesthetically like sculpture',  
'affected',  
'affected second',  
'affected second world',  
'affected second world war',  
'affleck',  
'affleck sandra',  
'affleck sandra bullock',  
'affleck sandra bullock film',  
'affordable',  
'affordable prices',  
'affordable prices good',  
'affordable prices good food',  
'afraid',  
'afraid go',  
'afraid go car',  
'afraid go car end',  
'afraid subtitles',  
'afraid subtitles worth',  
'afraid subtitles worth little',  
'africa',  
'africa understand',  
'africa understand past',  
'africa understand past attempted',  
'afternoon',  
'afternoon punish',  
'afternoon punish unemployed',  
'afternoon punish unemployed jobs',  
'age',  
'age john',  
'age john wayne',  
'age john wayne incredible',  
'ages',  
'ages although',  
'ages although younger',  
'ages although younger set',  
'aggravating',  
'ago',  
'ago impressed',  
'ago left',  
'ago left phone',  
'ago left phone trunk',  
'ago march',  
'ago march th',  
'ago march th began',  
'ago nightmare',  
'ago really',  
'ago really liked',  
'ago really liked service',  
'agree',  
'agree jessica',  
'agree jessica movie',  
'agree jessica movie pretty',  
'agreed',

'agreed film',  
'air',  
'air enough',  
'air enough anyone',  
'air enough anyone knows',  
'air force',  
'air force one',  
'air force one obliged',  
'aired',  
'aired dribble',  
'aired dribble watched',  
'aired dribble watched putting',  
'airline',  
'airline food',  
'airline food seriously',  
'airport',  
'airport experience',  
'airport experience tasty',  
'airport experience tasty food',  
'akasha',  
'akasha places',  
'akasha places compelling',  
'akasha places compelling first',  
'akin',  
'akin torture',  
'ala',  
'ala cart',  
'alarm',  
'alarm clock',  
'alarm clock docking',  
'alarm clock docking station',  
'albondigas',  
'albondigas soup',  
'albondigas soup warm',  
'albondigas soup warm tasted',  
'alert',  
'alert try',  
'alert try attempt',  
'alert try attempt decipher',  
'alexander',  
'alexander nevsky',  
'alexander nevsky great',  
'alexander nevsky great film',  
'alexander yardley',  
'alexander yardley character',  
'alike',  
'allergy',  
'allergy warnings',  
'allergy warnings menu',  
'allergy warnings menu waitress',  
'allison',  
'allison movie',  
'allison movie yet',  
'allison movie yet enjoy',  
'allot',  
'allot clearer',  
'allot clearer ear',  
'allot clearer ear piece',  
'allow',  
'allow cell',  
'allow cell phone',



'allow cell phone usage',  
'allowing',  
'allowing screen',  
'allowing screen get',  
'allowing screen get scratched',  
'allows',  
'allows connect',  
'allows connect miniusb',  
'allows connect miniusb device',  
'allows possibility',  
'allows possibility double',  
'allows possibility double booking',  
'allstar',  
'allstar cast',  
'allstar cast loved',  
'allstar cast loved opened',  
'almonds',  
'almonds bacon',  
'almonds bacon big',  
'almonds bacon big winner',  
'almost',  
'almost action',  
'almost action scenes',  
'almost action scenes little',  
'almost always',  
'almost always olde',  
'almost always olde english',  
'almost brilliant',  
'almost empty',  
'almost empty excuse',  
'almost everyone',  
'almost everyone involved',  
'almost everyone involved must',  
'almost hear',  
'almost hear nothing',  
'almost melt',  
'almost melt mouth',  
'almost nonexistent',  
'almost right',  
'almost right face',  
'almost songs',  
'almost songs cover',  
'almost songs cover girl',  
'almost totally',  
'almost totally unintelligible',  
'almost totally unintelligible couldnt',  
'almost unbearable',  
'almost unbearable watch',  
'almost unbearable watch screen',  
'almost unrecognizable',  
'almost unrecognizable fans',  
'almost unrecognizable fans recent',  
'alone',  
'alone makes',  
'alone makes say',  
'alone makes say wow',  
'alone vegas',  
'along',  
'along headset',  
'along headset starts',  
'along headset starts ringing',

'along touching',  
'along touching drama',  
'along touching drama hilarious',  
'along verizon',  
'along verizon customer',  
'along verizon customer service',  
'along well',  
'along well maybe',  
'along well maybe idiotsavant',  
'alongside',  
'alongside olivia',  
'alongside olivia de',  
'alongside olivia de havilland',  
'alot',  
'alot money',  
'already',  
'already know',  
'already know racism',  
'already know racism prejudice',  
'already memorized',  
'already probably',  
'already seen',  
'already seen one',  
'already seen one miyazakis',  
'also',  
'also annoying',  
'also area',  
'also area unit',  
'also area unit broke',  
'also bad',  
'also bad barely',  
'also bad barely touched',  
'also bad well',  
'also belowpar',  
'also belowpar script',  
'also belowpar script borrowed',  
'also best',  
'also best cheese',  
'also best cheese crisp',  
'also boasts',  
'also boasts one',  
'also boasts one cutest',  
'also building',  
'also building freezing',  
'also building freezing cold',  
'also came',  
'also came back',  
'also came back check',  
'also combos',  
'also combos like',  
'also combos like burger',  
'also cool',  
'also cool styles',  
'also cool styles like',  
'also decided',  
'also decided send',  
'also decided send back',  
'also didnt',  
'also didnt like',  
'also didnt like button',  
'also dirty',

'also dirty lot',  
'also dirty lot time',  
'also ears',  
'also ears hurt',  
'also ears hurt try',  
'also enough',  
'also enough hypocrisy',  
'also enough hypocrisy film',  
'also found',  
'also found watching',  
'also funny',  
'also funny poignant',  
'also funny poignant times',  
'also good',  
'also great',  
'also great see',  
'also great see renowned',  
'also happy',  
'also happy price',  
'also highlight',  
'also highlight place',  
'also horrible',  
'also horrible cause',  
'also horrible cause see',  
'also incredible',  
'also lacked',  
'also lacked depth',  
'also lacked depth woa',  
'also loves',  
'also make',  
'also make ringtones',  
'also make ringtones form',  
'also makes',  
'also makes easier',  
'also makes easier hold',  
'also makes monica',  
'also makes monica bellucci',  
'also managed',  
'also managed give',  
'also managed give away',  
'also marred',  
'also marred constant',  
'also marred constant use',  
'also mention',  
'also mention weird',  
'also mention weird color',  
'also music',  
'also music mark',  
'also music mark snow',  
'also new',  
'also new problem',  
'also nice',  
'also ordered',  
'also ordered spinach',  
'also ordered spinach avocado',  
'also perfectly',  
'also perfectly cooked',  
'also phone',  
'also phone doesnt',  
'also phone doesnt seem',  
'also phone dropped',

'also phone dropped case',  
'also play',  
'also play well',  
'also potentially',  
'also potentially fry',  
'also pretty',  
'also pretty rare',  
'also pretty rare vegas',  
'also problems',  
'also problems phone',  
'also problems phone reading',  
'also quick',  
'also real',  
'also real treat',  
'also real treat see',  
'also really',  
'also really good',  
'also really good pretty',  
'also revealing',  
'also right',  
'also right balance',  
'also right balance war',  
'also served',  
'also served hot',  
'also served hot bread',  
'also slim',  
'also slim enough',  
'also slim enough fit',  
'also story',  
'also story acting',  
'also story acting weak',  
'also take',  
'also take great',  
'also take great pictures',  
'also take pictures',  
'also take pictures case',  
'also taste',  
'also taste homemade',  
'also taste homemade extra',  
'also taste moms',  
'also taste moms multigrain',  
'also tasty',  
'also thrilled',  
'also thrilled made',  
'also thrilled made amazing',  
'also took',  
'also took forever',  
'also took forever bring',  
'also turns',  
'also turns good',  
'also turns good performance',  
'also uninspired',  
'also works',  
'also works well',  
'also works well according',  
'also wrotedirected',  
'also wrotedirected weak',  
'also wrotedirected weak monster',  
'although',  
'although blame',  
'although blame needs',

'although blame needs placed',  
'although circumstances',  
'although circumstances slightly',  
'although circumstances slightly different',  
'although much',  
'although much liked',  
'although much liked look',  
'although never',  
'although never murdered',  
'although never murdered anyone',  
'although rated',  
'although rated impressive',  
'although rated impressive megapixels',  
'although vacant',  
'although vacant tables',  
'although vacant tables folks',  
'although younger',  
'although younger set',  
'although younger set probably',  
'aluminum',  
'aluminum case',  
'aluminum case new',  
'aluminum case new palm',  
'always',  
'always changing',  
'always changing food',  
'always changing food quality',  
'always clean',  
'always cuts',  
'always cuts makes',  
'always cuts makes beep',  
'always easy',  
'always easy hear',  
'always entertaining',  
'always entertaining usual',  
'always evening',  
'always evening wonderful',  
'always evening wonderful food',  
'always friendly',  
'always friendly helpful',  
'always friendly helpful menu',  
'always full',  
'always full never',  
'always full never wait',  
'always good',  
'always good dangerous',  
'always good dangerous nutbag',  
'always great',  
'always great time',  
'always great time dos',  
'always hit',  
'always hit group',  
'always hit lots',  
'always hit lots great',  
'always honor',  
'always honor pizza',  
'always honor pizza hut',  
'always known',  
'always known errol',  
'always known errol flynn',  
'always make',

```
'always make us',
'always make us feel',
'always olde',
'always olde english',
'always olde english sometimes',
'always order',
'always order vegetarian',
'always order vegetarian menu',
'always pleasure',
'always pleasure dealing',
'always super',
'always super friendly',
'always super friendly helpful',
'always turn',
'always used',
'always used corded',
'always used corded headsets',
'amateurish',
'amateurish films',
'amateurish films first',
'amateurish films first watched',
'amaze',
'amaze one',
'amaze one favourite',
'amaze one favourite directors',
'amazed',
...]
```

In [146]:

```
feature_matrix=counts.todense()
feature_matrix[:2]
```

Out[146]:

```
matrix([[0., 0., 0., ..., 0., 0., 0.],
        [0., 0., 0., ..., 0., 0., 0.]])
```

In [138]:

```
counts
```

Out[138]:

```
<2470x39399 sparse matrix of type '<class 'numpy.float64'>'
  with 51750 stored elements in Compressed Sparse Row format>
```

In [147]:

```
feature_matrix.shape
```

Out[147]:

```
(2470, 29643)
```

In [ ]: