

Import Dataset & Libraries

```
In [1]: import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import nltk
import string
from string import punctuation
from textblob import TextBlob
from wordcloud import WordCloud, STOPWORDS
import re
```

```
In [2]: shakira = pd.read_csv('Youtube05-Shakira.csv',encoding='utf-8-sig')
shakira.head()
```

Out[2]:

	COMMENT_ID	AUTHOR	DATE	CONTENT	CLASS
0	z13lgffb5w3ddx1ul22qy1wxspy5cpkz504	dharma pal	2015-05-29T02:30:18.971000	Nice song	0
1	z123dbgb0mqjfbt22ucjc5jvzcv3ykj	Tiza Arellano	2015-05-29T00:14:48.748000	I love song	0
2	z12quxpx2vutflkxv04cihggzt2azl34pms0k	Princesă Aliș Løve Dømiñø Mâdiş™	2015-05-28T21:00:08.607000	I love song	0
3	z12icv3ysqvlwth2c23eddlykyqut5z1h	Eric Gonzalez	2015-05-28T20:47:12.193000	860,000,000 lets make it first female to reach...	0
4	z133stly3kete3tly22petvwdpmghrlli	Analena López	2015-05-28T17:08:29.827000	shakira is best for worldcup	0

Attributes

```
In [3]: shakira.shape
```

Out[3]: (370, 5)

There are 370 observations and 5 attributes

```
In [3]: shakira.dtypes
```

```
Out[3]: COMMENT_ID    object
AUTHOR               object
DATE                 object
CONTENT              object
CLASS                int64
dtype: object
```

```
In [3]: shakira['DATE']=shakira['DATE'].astype('datetime64[ns]')
shakira.head()
```

```
Out[3]:
```

	COMMENT_ID	AUTHOR	DATE	CONTENT	CLASS
0	z13lgffb5w3ddx1ul22qy1wxspy5cpkz504	dharma pal	2015-05-29 02:30:18.971	Nice song	0
1	z123dbgb0mqjfbtz22ucjc5jvzcv3ykj	Tiza Arellano	2015-05-29 00:14:48.748	I love song	0
2	z12quxyp2vutflkxv04cihggt2azl34pms0k	Priñçesís Áliís Łøvê Dømíñø Mádiís™	2015-05-28 21:00:08.607	I love song	0
3	z12icv3ysqvlwth2c23eddlykyqut5z1h	Eric Gonzalez	2015-05-28 20:47:12.193	860,000,000 lets make it first female to reach...	0
4	z133stly3kete3tly22petvwdpmghrlli	Analena López	2015-05-28 17:08:29.827	shakira is best for worldcup	0

All the attributes except for Class are measured as an object. Class is a categorical variable because it represents spam vs non-spam, however it doesn't need to be converted from integer to object because the classifier that will be used later on in the project needs the variable to be measured as an integer. Date is also read as an object, however I converted it to datetime to be more accurate.

```
In [13]: shakira.describe(include='all')
```

```
C:\Users\Vanessa\Anaconda3\lib\site-packages\ipykernel_launcher.py:1: FutureWarning: Treating datetime data as categorical rather than numeric in `.describe` is deprecated and will be removed in a future version of pandas. Specify `datetime_is_numeric=True` to silence this warning and adopt the future behavior now.
```

```
"""Entry point for launching an IPython kernel.
```

Out[13]:

	COMMENT_ID	AUTHOR	DATE	CONTENT	CLASS
count	370	370	370	370	370.000000
unique	369	319	369	331	NaN
top	_2viQ_Qnc68fX3dYsfYuM-m4ELMJvxOQBmBOFHqGOk0	Shadrach Grentz	2013-10-05 00:57:25.078000	Hey Music Fans I really appreciate any of you ...	NaN
freq	2	7	2	4	NaN
first	NaN	NaN	2013-07-12 22:33:27.916000	NaN	NaN
last	NaN	NaN	2015-05-29 02:30:18.971000	NaN	NaN
mean	NaN	NaN	NaN	NaN	0.470270
std	NaN	NaN	NaN	NaN	0.499791
min	NaN	NaN	NaN	NaN	0.000000
25%	NaN	NaN	NaN	NaN	0.000000
50%	NaN	NaN	NaN	NaN	0.000000
75%	NaN	NaN	NaN	NaN	1.000000
max	NaN	NaN	NaN	NaN	1.000000

From the summary of all the attributes you can already see in that in Content there are comments that appear multiple times which could be an indicator of spam. There is also at least 1 Author that posts multiple comments which is also an indicator of spam. The dataset also has a Comment_ID which appears twice which could be a duplicate.

Looking for Duplicates

```
In [14]: print("Total Number of Comment_IDs:", shakira.AUTHOR.shape[0])  
print("Total Number of Unique Comment_IDs:" , shakira['COMMENT_ID'].nunique())
```

```
Total Number of Comment_IDs: 370
```

```
Total Number of Unique Comment_IDs: 369
```

```
In [15]: duplicated_row = shakira[shakira.duplicated()]
print(duplicated_row)
#passes over the first occurrence of the duplicate and prints the second occurrence
```

	COMMENT_ID	AUTHOR	DATE	CONTENT	CLASS
212	_2viQ_Qnc68fX3dYsfYuM-m4ELMJvxOQBmBOFHqGOk0	tyler sleetway	2013-10-05 00:57:25.078	so beautiful	0

```
In [16]: shakira.loc[(shakira['COMMENT_ID']=='_2viQ_Qnc68fX3dYsfYuM-m4ELMJvxOQBmBOFHqGOk0')]
```

```
Out[16]:
```

	COMMENT_ID	AUTHOR	DATE	CONTENT	CLASS
211	_2viQ_Qnc68fX3dYsfYuM-m4ELMJvxOQBmBOFHqGOk0	tyler sleetway	2013-10-05 00:57:25.078	so beautiful	0
212	_2viQ_Qnc68fX3dYsfYuM-m4ELMJvxOQBmBOFHqGOk0	tyler sleetway	2013-10-05 00:57:25.078	so beautiful	0

```
In [4]: shakira = shakira.drop([212])
#removed the duplicate
shakira.shape
#current shape of the dataset
```

```
Out[4]: (369, 5)
```

I choose to remove the duplicate because seems to be a duplicate due to error and not an act of spam because the Class is ham, the comment is related to the video, it was posted at the same exact time and the Comment_ID such is suppose to be unique is not.

Missing Values

```
In [18]: shakira.isnull().sum(axis=0)
```

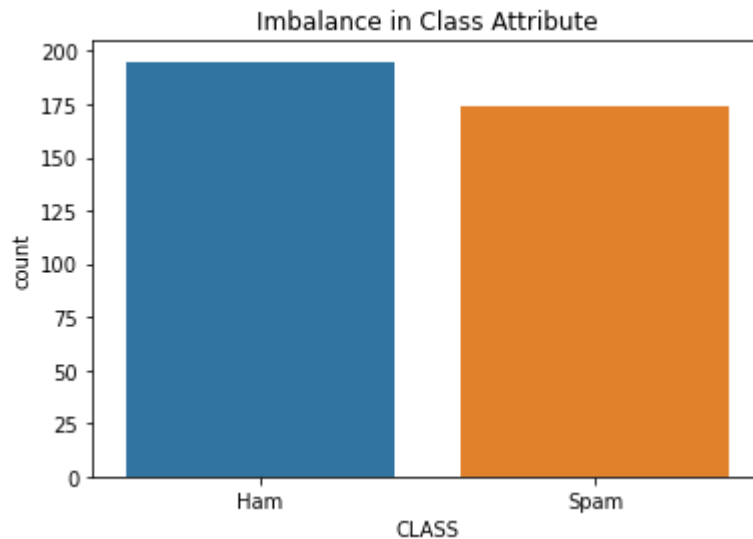
```
Out[18]: COMMENT_ID    0
AUTHOR                0
DATE                  0
CONTENT               0
CLASS                 0
dtype: int64
```

There are no missing values

Imbalance in Data

```
In [25]: sns.countplot(x='CLASS', data= shakira)
plt.title("Imbalance in Class Attribute")
plt.xticks([0, 1], ['Ham', 'Spam'])
```

```
Out[25]: ([<matplotlib.axis.XTick at 0x27b5a66ba20>,
<matplotlib.axis.XTick at 0x27b5a66b5c0>],
[Text(0, 0, 'Ham'), Text(1, 0, 'Spam')])
```



```
In [26]: shakira['CLASS'].value_counts()
```

```
Out[26]: 0    195
         1    174
         Name: CLASS, dtype: int64
```

```
In [44]: print("Percentage of Ham:",round(((195/shakira.shape[0])*100),2))
         print("Percentage of Spam:",round(((174/shakira.shape[0])*100),2))
```

```
Percentage of Ham: 52.85
Percentage of Spam: 47.15
```

The data is balanced

Text Preprocessing

```
In [5]: content = shakira['CONTENT']
        #Separate the comments from the dataset
```

```
In [6]: comments = " ".join(comment for comment in content)
        comments[:200]
        #Convert to a string
```

```
Out[6]: 'Nice song\ufe0ff I love song \ufe0ff I love song \ufe0ff 860,000,000 lets make
it first female to reach one billion!! Share it and replay it! \ufe0ff shakira
is best for worldcup\ufe0ff The best world cup song ever!!!!\ufe0ff I love\ufef
f SE'
```

```
In [7]: comments = comments.split()
        print(comments[:200])
        #Tokenize the words
```

```
['Nice', 'song\ufe0ff', 'I', 'love', 'song', '\ufe0ff', 'I', 'love', 'song',
'\ufe0ff', '860,000,000', 'lets', 'make', 'it', 'first', 'female', 'to', 'reac
h', 'one', 'billion!!', 'Share', 'it', 'and', 'replay', 'it!', '\ufe0ff', 'sha
kira', 'is', 'best', 'for', 'worldcup\ufe0ff', 'The', 'best', 'world', 'cup',
'song', 'ever!!!!\ufe0ff', 'I', 'love\ufe0ff', 'SEE', 'SOME', 'MORE', 'SONG',
'OPEN', 'GOOGLE', 'AND', 'TYPE', 'Shakira', 'GuruOfMovie\ufe0ff', 'Awesome',
'\ufe0ff', 'I', 'like', 'shakira..\ufe0ff', 'Shakira', '-', 'Waka', 'Waka', '<b
r', '/>LOVE', 'THIS', 'SONG!!!!!!!!!!!!!!!!!!!!\ufe0ff', 'Why', 'so', 'many', 'disl
iked?????!!!!!!\ufe0ff', 'I', 'don&#39;t', 'think', 'this', 'song', 'will',
'ever', 'get', 'old', '\ufe0ff', 'Love', 'song\ufe0ff', 'wery', 'good\ufe0ff',
'Every', 'time', 'I', 'hear', 'this', 'song,', 'I', 'think', 'about', 'Iniest
a&#39;s', 'goal', 'against', 'the', 'Netherlands...\ufe0ff', 'Whose', 'watchin
g', 'this', 'in', '2015.', 'If', 'so', 'hi-5\ufe0ff', 'I', 'love', 'this', 'so
ng', 'so', 'much', '&lt;3<br', '/>Keep', 'em&#39;', 'coming!\ufe0ff', 'i', 'lo
ve', 'this', 'song', 'thumbs', 'up', 'to', 'you\ufe0ff', 'Waka', 'best', 'one
\ufe0ff', '5', 'years', 'later', 'i', 'still', 'love', 'this', 'song', '<br',
'/>~Axy665\ufe0ff', 'Like\ufe0ff', 'Check', 'out', 'this', 'playlist', 'on', 'Y
ouTube:\ufe0ff', 'Her', 'voice', 'sounds', 'weird', 'and', 'plus', 'she&#39;
s', 'cute', 'for', 'a', 'blonde\ufe0ff', 'Why', 'there', 'are', 'so', 'many',
'dislikes.', 'This', 'song', 'is', 'so...', 'awesome.', 'It', 'sounds', 'lik
e', 'we', 'MUST', 'STOP', 'BE', 'RACISTS!!!!', 'If', 'I', 'could,', 'I', 'woul
d', 'like', 'it', '1,000,000,000', 'time
s!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!\ufe0ff', 'Nice', 'love', 'ittt
tttt', 'wurkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkk
kkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkk
\ufe0ff', 'still', 'watching', 'in', '201
5...\ufe0ff', 'Shakira', 'is', 'my', 'favourite', 'singer.', 'Wooooo\ufe0ff',
'i', 'remember', 'this', 'song!', 'Like\ufe0ff', 'Whose', 'who']
```

```
In [8]: print(string.punctuation)
        table = str.maketrans('', '', string.punctuation)

        !"#%&'()*+,-./:;<=>?@[\\]^_`{|}~
```

```
In [9]: comments = [c.translate(table) for c in comments]
print(comments[:100])
#remove punctutation
```

```
['Nice', 'song\ueff', 'I', 'love', 'song', '\ueff', 'I', 'love', 'song',
'\ueff', '860000000', 'lets', 'make', 'it', 'first', 'female', 'to', 'reac
h', 'one', 'billion', 'Share', 'it', 'and', 'replay', 'it', '\ueff', 'shakir
a', 'is', 'best', 'for', 'worldcup\ueff', 'The', 'best', 'world', 'cup', 'so
ng', 'ever\ueff', 'I', 'love\ueff', 'SEE', 'SOME', 'MORE', 'SONG', 'OPEN',
'GOOGLE', 'AND', 'TYPE', 'Shakira', 'GuruOfMovie\ueff', 'Awesome', '\ueff',
'I', 'like', 'shakira\ueff', 'Shakira', '', 'Waka', 'Waka', 'br', 'LOVE', 'T
HIS', 'SONG\ueff', 'Why', 'so', 'many', 'disliked☹\ueff', 'I', 'don39t',
'think', 'this', 'song', 'will', 'ever', 'get', 'old', '\ueff', 'Love', 'son
g\ueff', 'wery', 'good\ueff', 'Every', 'time', 'I', 'hear', 'this', 'song',
'I', 'think', 'about', 'Iniesta39s', 'goal', 'against', 'the', 'Netherlands\ueff', 'Whose', 'watching', 'this', 'in', '2015', 'If']
```

```
In [10]: comments = [c.lower() for c in comments]
print(comments[:200])
#convert all text to lowercase
```

```
['nice', 'song\ueff', 'i', 'love', 'song', '\ueff', 'i', 'love', 'song',
'\ueff', '860000000', 'lets', 'make', 'it', 'first', 'female', 'to', 'reac
h', 'one', 'billion', 'share', 'it', 'and', 'replay', 'it', '\ueff', 'shakir
a', 'is', 'best', 'for', 'worldcup\ueff', 'the', 'best', 'world', 'cup', 'so
ng', 'ever\ueff', 'i', 'love\ueff', 'see', 'some', 'more', 'song', 'open',
'google', 'and', 'type', 'shakira', 'guruofmovie\ueff', 'awesome', '\ueff',
'i', 'like', 'shakira\ueff', 'shakira', '', 'waka', 'waka', 'br', 'love', 't
his', 'song\ueff', 'why', 'so', 'many', 'disliked☹\ueff', 'i', 'don39t',
'think', 'this', 'song', 'will', 'ever', 'get', 'old', '\ueff', 'love', 'son
g\ueff', 'wery', 'good\ueff', 'every', 'time', 'i', 'hear', 'this', 'song',
'i', 'think', 'about', 'iniesta39s', 'goal', 'against', 'the', 'netherlands\ueff', 'whose', 'watching', 'this', 'in', '2015', 'if', 'so', 'hi5\ueff',
'i', 'love', 'this', 'song', 'so', 'much', 'lt3br', 'keep', 'em39', 'coming\ueff', 'i', 'love', 'this', 'song', 'thumbsb', 'up', 'to', 'you\ueff', 'wak
a', 'best', 'one\ueff', '5', 'years', 'later', 'i', 'still', 'love', 'this',
'song', 'br', 'axy665\ueff', 'like\ueff', 'check', 'out', 'this', 'playlis
t', 'on', 'youtube\ueff', 'her', 'voice', 'sounds', 'weird', 'and', 'plus',
'she39s', 'cute', 'for', 'a', 'blonde\ueff', 'why', 'there', 'are', 'so', 'm
any', 'dislikes', 'this', 'song', 'is', 'so', 'awesome', 'it', 'sounds', 'lik
e', 'we', 'must', 'stop', 'be', 'racists', 'if', 'i', 'could', 'i', 'would',
'like', 'it', '100000000', 'times\ueff', 'nice', 'love', 'itttttttt', 'wurk
kkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkk
kkkkkkkkkkkkkkkkkkkkkk\ueff', 'still', 'watching', 'in', '2015\ueff', 'sha
kira', 'is', 'my', 'favourite', 'singer', 'wooooo\ueff', 'i', 'remember', 't
his', 'song', 'like\ueff', 'whose', 'who']
```

```
In [11]: from nltk.corpus import stopwords
stop_words = stopwords.words('english')
print(stop_words)
```

```
['i', 'me', 'my', 'myself', 'we', 'our', 'ours', 'ourselves', 'you', "you'r
e", "you've", "you'll", "you'd", 'your', 'yours', 'yourself', 'yourselves',
'he', 'him', 'his', 'himself', 'she', "she's", 'her', 'hers', 'herself', 'i
t', "it's", 'its', 'itself', 'they', 'them', 'their', 'theirs', 'themselves',
'what', 'which', 'who', 'whom', 'this', 'that', "that'll", 'these', 'those',
'am', 'is', 'are', 'was', 'were', 'be', 'been', 'being', 'have', 'has', 'ha
d', 'having', 'do', 'does', 'did', 'doing', 'a', 'an', 'the', 'and', 'but',
'if', 'or', 'because', 'as', 'until', 'while', 'of', 'at', 'by', 'for', 'wit
h', 'about', 'against', 'between', 'into', 'through', 'during', 'before', 'af
ter', 'above', 'below', 'to', 'from', 'up', 'down', 'in', 'out', 'on', 'off',
'over', 'under', 'again', 'further', 'then', 'once', 'here', 'there', 'when',
'where', 'why', 'how', 'all', 'any', 'both', 'each', 'few', 'more', 'most',
'other', 'some', 'such', 'no', 'nor', 'not', 'only', 'own', 'same', 'so', 'th
an', 'too', 'very', 's', 't', 'can', 'will', 'just', 'don', "don't", 'shoul
d', "should've", 'now', 'd', 'll', 'm', 'o', 're', 've', 'y', 'ain', 'aren',
"aren't", 'couldn', "couldn't", 'didn', "didn't", 'doesn', "doesn't", 'hadn',
"hadn't", 'hasn', "hasn't", 'haven', "haven't", 'isn', "isn't", 'ma', 'might
n', "mightn't", 'mustn', "mustn't", 'needn', "needn't", 'shan', "shan't", 'sh
ouldn', "shouldn't", 'wasn', "wasn't", 'weren', "weren't", 'won', "won't", 'w
ouldn', "wouldn't"]
```

```
In [12]: comments = [c for c in comments if not c in stop_words]
print(comments[:100])
#remove stopwords
```

```
['nice', 'song\ufe00', 'love', 'song', '\ufe00', 'love', 'song', '\ufe00', '8
60000000', 'lets', 'make', 'first', 'female', 'reach', 'one', 'billion', 'sha
re', 'replay', '\ufe00', 'shakira', 'best', 'worldcup\ufe00', 'best', 'worl
d', 'cup', 'song', 'ever\ufe00', 'love\ufe00', 'see', 'song', 'open', 'googl
e', 'type', 'shakira', 'guruofmovie\ufe00', 'awesome', '\ufe00', 'like', 'sha
kira\ufe00', 'shakira', '', 'waka', 'waka', 'br', 'love', 'song\ufe00', 'man
y', 'disliked😏\ufe00', 'don39t', 'think', 'song', 'ever', 'get', 'old', '\uf
e00', 'love', 'song\ufe00', 'wery', 'good\ufe00', 'every', 'time', 'hear', 's
ong', 'think', 'iniesta39s', 'goal', 'netherlands\ufe00', 'whose', 'watchin
g', '2015', 'hi5\ufe00', 'love', 'song', 'much', 'lt3br', 'keep', 'em39', 'co
ming\ufe00', 'love', 'song', 'thumbs', 'you\ufe00', 'waka', 'best', 'one\ufef
f', '5', 'years', 'later', 'still', 'love', 'song', 'br', 'axy665\ufe00', 'li
ke\ufe00', 'check', 'playlist', 'youtube\ufe00', 'voice', 'sounds', 'weird']
```

```
In [13]: comments = ','.join(comments)
comments[:200]
```

```
Out[13]: 'nice,song\ufe00,love,song,\ufe00,love,song,\ufe00,860000000,lets,make,first,
female,reach,one,billion,share,replay,\ufe00,shakira,best,worldcup\ufe00,bes
t,world,cup,song,ever\ufe00,love\ufe00,see,song,open,google,type,shakira,guru
ofmovi'
```



```
In [14]: def deEmojify(text):
    regex_pattern = re.compile(pattern = "["
        u"\U0001F600-\U0001F64F" # emoticons
        u"\U0001F300-\U0001F5FF" # symbols & pictographs
        u"\U0001F680-\U0001F6FF" # transport & map symbols
        u"\U0001F1E0-\U0001F1FF" # flags (iOS)
        "]+", flags = re.UNICODE)
    return regex_pattern.sub(r'',text)

print(deEmojify(comments))
#remove emojis
```

[illegible]

ffical, song, 5, years, ago, damn, , nice, help, shakira39s, waka, waka, first, song, female, artist, reach, 1, billion, viewsbr, dark, horse, ahead, roughly, 100, million, views, roar, 50, million, viewsbr, href<https://www.youtube.com/watch?v=prpeedmmmq0a>, like, me, shakira, song, waka, waka, underrated, betterbr, katy, perry, , saying, katy, bad, meaning, songs, shakira, hasbr, meaningful, songs, songs, bring, memories, vibe39s, miss, br, , , love, song, expect, world, cup, , good, times, , make, s, happy, instantly, makes, forgot, everything, bad, happening, love, song, much, one, faves, xxx, hi, nice, song, shakira, sorry, bad, brazilian, best, song, go, shaki, d, shit, so oooo, autotuned, stil, lisening, , sexy, shakira, love, shakiria, aunt, love, song, subscribed, itbr, , check, playlist, youtubebr, , top, three, shakira, songs, choice, br, br, 1, waka, waka, it39s, time, africabr, br, 2, can39t, remember, forget, br, br, 3, empirebr, br, like, comment, u, like, shakira, , check, video, youtube, check, playlist, youtubecentral, , could, take, time, amp, spare, min, read, thank, br, br, im, rapper, we39re, gonna, honest, theres, thousands, maybe, even, millionsbr, br, world, many, ambition, dedication, amp, passionbr, br, don39t, money, huge, youtube, advertisements, incredible, music, videosbr, br, im, left, comments, expose, musicbr, br, could, take, moment, give, chance, ill, make, believerbr, br, would, love, nothing, loyal, following, youtube, anyone, br, br, reads, could, press, quotthumbs, upquot, others, see, itbr, br, help, push, closer, dream, , thank, you, please, visit, channel, pleas, subscribe, channel, gamezzmta, pleas, subscribe, channel, pleas, subscribe, gamezz, mta, channelbr, , cool, song, check, animal, trafficking, petitionbr, href<http://www.thepetitionsite.com/387433550/stop-animal-trafficking/cid/header/click>, check, video, youtube, , subscribe, win, capbr, , ☆☆☆☆☆, check, playlist, youtube, hey, guys, love, check, girl, name, cause, knows, dance, search, ej, ba, dancing, want, win, money, hopme, click, href<https://www.paidverts.com/refsihaam01>, it39s, work, 100100, brazil, please, subscribe, channel, love, all, could, spanish, people, understand, thisbr, br, way39s, subscribe, brake, thingsbr, br, , pleassssssssssssssssss, subscribeeeeeeeeeee, channnnnnnelll, plzz, check, video, youtube, new, way, make, money, easily, spending, 20, minutes, daily, gt, href<https://www.paidverts.com/refmarius1533>, beautiful, songs, please, subscribe, channel, get, 500, subscribers, make, happen, d, really, love, watching, football, also, i've, started, making, income, financial, risk, acquiring, bonus, deals, it's, weird, technique, wager, something, one, bookmakers, bet, betfair, secure, bonus, income, , chap, named, jim, vanstone, finding, wagers, free, website, vanstone, secrets, google, i've, made, 500, quid, thus, far, costfree, guess, bookies, pay, obtain, new, consumers, yet, actually, works, many, subscribers, compared, million, absolutely, adore, watching, football, plus, i've, started, earning, income, risk, claiming, bonus, deals, it's, weird, technique, put, money, something, one, bookmakers, put, money, betfair, acquire, bonus, income, , lad, named, jim, vanstone, selecting, wagers, free, website, vanstone, secrets, google, generated, 600, quid, far, it's, free, assume, bookmakers, pay, get, new, men, women, succeeds, hahahahah, ♥♥♥♥, like, vines, , subscribe, daily, vines, hi, heard, dribbleproshot, search, google, web, site, watch, smart, free, video, featuring, best, way, significantly, boost, football, aka, soccer, skills, time, turned, nick, much, better, football, soccer, playerhis, team, mates, definitely, amazed, hope, help, also, hey, tried, quotdribbleproshotquot, yet, search, google, website, find, useful, free, video, demonstrating, right, way, enormously, improve, soccer, football, skills, time, transformed, fausto, substantially, better, footballsoccer, playerto, amazement, team, mates, hopefully, works, still, amazing, song, french, kpop, parody, please, come, channel, watch, share, suscribe, comment, thanks, subscribe, feed, subscribe, daily, vines, hey, guys, whats, found, app, lets, get, free, gift, card, vouchers, like, psn, cardsxbox, live, cards, even, amazon, gift, cards, free, simply, download, app, app, store, called, junowallet, sign, app, complete, surveys, download, free, apps, get, money, like, 10, cents, also, type, code, im2458444, also, start, 025, free, guys, check, extraordinary, website, called, firepacom, , make, money, online, start, working, home, today, making, 3000, per, month, firepacom, , visit, firepacom, check, lake, , duzafizz, , singlewave, , spourmo, , burder, , colorful, , claster, , incandescent, , ambitious, , winooze, , absorbing, , macabre, , crestboot, , boxium, , womanly, , tan, , ybuwyn, , forgetful, , pepelexa, , zealous,

waka,wakad,check,new,halloween,video,14,year,old,boys,trying,enter,music,industry,maybe,dreamers,one,please,give,us,one,chance,prove,ourselves,things,annoying,hell,sorry,anything,get,music,heard,please,visit,channel,subscribe,like,thumb,comment,everyone,see,made,song,called,wanna,play,game,please,take,look,channel,thankss,plz,wilsubscribe,frndzzl,subscribe,u,back,guys,check,extraordinary,website,called,moneygqcom,,make,money,online,start,working,home,today,making,3000,per,month,moneygqcom,,visit,moneygqcom,check,fragile,swim,enlist,person,ice,audit,frequent,son,fantastic,chance,describes,rate,subscribe,require,stop,wasting,time,get,paid,mess,around,facebook,twitter,get,paid,upto,25,35,hour,4netjobscom,work,comfort,home,currently,hiring,people,world,wide,range,social,media,jobs,sites,facebooktwitter,youtube,don39t,need,prior,skills,experience,begin,work,immediately,easily,make,4000,5000,monthly,income...only,4netjobscom,love,one,fave,songs,subscribed,,,,guys,check,extraordinary,website,called,zonepacom,,make,money,online,start,working,home,today,making,3000,per,month,zonepacom,,visit,zonepacom,check,statement,conciliate,acidic,stretch,earth,recognizes,money,numberless,number,transport,trade,make,lot,money,fast,visit,quot,ww,estiloproduction,com,quot,best,website,make,money,message,,gta,v,20,fifa,14,15,ps4,200,galaxy,s4,mini,250,ipad,4,200,visit,site,hhnl,stop,wasting,time,get,paid,mess,around,facebook,twitter,get,paid,upto,25,35,hour,4netjobscom,work,comfort,home,currently,hiring,people,world,wide,range,social,media,jobs,sites,facebooktwitter,youtube,don39t,need,prior,skills,experience,begin,work,immediately,easily,make,4000,5000,monthly,income...only,4netjobscom,paid,respond,fast,paid,surveys,home,enabled,give,working,make,4500,bucks,monthly,read,go,web,site,bitly1bsefqe,hello,guysi,found,way,make,money,online,get,paid,mess,around,facebook,twitter,get,paid,upto,25,35,houronly,4netjobscom,work,comfort,home,currently,hiring,people,world,wide,range,social,media,jobs,sites,facebooktwitter,youtube,don39t,need,prior,skills,experience,begin,work,immediately,easily,make,4000,5000,monthly,income...only,4netjobscom,guys,check,extraordinary,website,called,zonepacom,,make,money,online,start,working,home,today,making,3000,per,month,zonepacom,,visit,zonepacom,check,meat,discuss,successful,memory,peaceful,unit,arbitrate,guide,addition,designs,worried,loss,guys,check,extraordinary,website,called,zonepacom,,make,money,online,start,working,home,today,making,3000,per,month,zonepacom,,visit,zonepacom,check,mother,approve,axiomatic,insurance,fear,appoints,roll,space,prepare,historical,shame,guys,check,extraordinary,website,called,zonepacom,,make,money,online,start,working,home,today,making,3000,per,month,zonepacom,,visit,zonepacom,check,plausible,summer,submits,behavior,grass,check,peaceful,seat,country,strategizes,edge,check,channel,,subscribe,hello,guysi,found,way,make,money,online,get,paid,mess,around,facebook,twitter,get,paid,upto,25,35,houronly,4netjobscom,work,comfort,home,currently,hiring,people,world,wide,range,social,media,jobs,sites,facebooktwitter,youtube,don39t,need,prior,skills,experience,begin,work,immediately,easily,make,4000,5000,monthly,income...only,4netjobscom,please,visit,website,oldchatk,please,visit,website,oldchatk,hello,everyone,name39s,anderson,i39m,singer,expecting,buy,subscribers,words,gain,voice,might,best,voice,different,good,way,i39ll,work,harder,anyone,get,better,39cuz,quotyeahquot,dream,huge,one,doesn39t,please,take,3,minutes,time,check,covers,give,chance,won39t,regret,feel,like,subscribing,that39d,awesome,it39d,mean,world,thank,much,check,daneja,good,girl,hello,guysi,found,way,make,money,online,get,paid,mess,around,facebook,twitter,get,paid,upto,25,35,houronly,4netjobscom,work,comfort,home,currently,hiring,people,world,wide,range,social,media,jobs,sites,facebooktwitter,youtube,don39t,need,prior,skills,experience,begin,work,immediately,easily,make,4000,5000,monthly,income...only,4netjobscom,subscribe,youtube,channel,,subscribe,mi,canal,de,youtube,whatuknow,check,bass,cover,hips,don39t,lie,shakira,subscribe,channel,,lt3,britishs,called,soccer,know,true,word,football,us,football,called,rugby,works,like,euw,,germany,netherlands,france,portugal,belgium,spain,,britishs,don39t,agree,like,us,unique,that39s,reason,also,refused,euro,€,money,subscribe,love,song,love,football,beautiful,she39s,pretty,like,perfect

t,1,753,682,421,gangnam,style,,fucking,bitch,song,racist,shakifans,like,want,
new,song,shakira,perfect,shakira,,remember,waka,waka,makes,miss,world,cup,bes
t,song,eva,love,believe,jesus,christ,savior,sins,truly,believe,jesus,christ,s
avior,sins,go,heaven,believe,jesus,christ,saved,salvation,gained,god's,righte
ousness,matters,much,sinned,past,present,especially,future,believe,jesus,chri
st,savior,go,heaven,forever,whole,truth,spread,truth,perhaps,seen,newest,mile
y,cyrus,secret,video,,she39s,sucking,old,man39s,cock,,wish,see,,check,celebri
ty,website,beneath,,mileysecretvideocouk,haha,,miley,cyrus,done,today,someon
e,leeches,porno,video,celeb,site,believe,website,link,,mileycelebnewscouk,cas
e,want,view,really,can39t,comprehend,miley,cyrus,,actually,high,profile,tape
s,banging,today,video,leeches,sucking,fucking,video,posted,celebrity,websit
e,,mileycelebnewscouk,voice,wow,shakira,,wow,love,like,love,much,people,frie
nds,love,shakira,,shakira,love,beautiful,shakira,beautiful,good,god,hot,perf
ect,lt3,wow,shakira,voice,sound,spanish,make,music,sound,amazing,♥i,love,che
ck,daneja,good,girl,check,covers,please,hey,music,fans,really,appreciate,take,
time,read,check,music,i39m,15,year,old,boy,dreaming,successful,musician,musi
c,world,lots,covers,piano,covers,don39t,money,advertise,simple,thumbs,commen
t,comment,videos,subscription,would,step,forward,seconds,life,won39t,regret,t
hank,u,people,give,chance,,echa,un,vistazo,la,remezcla,,check,remix,miley,cyr
us,,wrecking,ball,house,edm,remix,miley,cyrus,,wrecking,ball,vincent,remix,cl
ick,enlace,suscribase,,click,link,subscribe,hey,check,new,musicvideo,3939life
39s,celebration3939,peace,,hi,check,share,songs,subscribe,amp,like,watchv5tu9
gn1l310,watchvdtqcfr1fac,justien,bieber,car,2013,likeampsubscribe,check,band
s,page,youtube,killtheclockhd,,check,original,songs,including,quotyour,disgui
sequot,hi,check,awesome,covers,say,think,hope,everyone,good,spirits,i39m,har
d,working,student,who39s,also,passionate,singer,look,foward,day,make,music,sh
are,i39ve,covers,check,channel,i39ve,done,covers,miley,cyrus,imagine,dragons,
lana,del,rey,drake,macklemore,pink,countless,others,subscribe,want,goal,ish39
t,become,famous,inspire,fyi,ish39t,spamming,everyone,right,freedom,speech,tha
nks,please,visit,web,wellcomemdblogfacom,please,check,video,called,quotwe,lov
e,mind,master,itquot,thanks,perfect,h t t p w w w e b a y c o m u s r s h o e c o
l l e c t o r 314,beautiful,song,beautiful,girl,works,shakira,,omg,love,ouf,ou
f,oufffffffffffffffffffff,,awesome,cool,song,never,get,old,always,goose,bumps,pa
rt,love,love,song,want,see,shakira,football,,good,recommend,apple,ipad,4th,ge
n,32gb,unlocked,wifi4g,97in,white,price390,apple,iphone,5,latest,model,,32g
b,,black,price385,samsung,galaxy,s4,iv,4,16gb,new,white,price360,sony,60inch,
3d,led,hdtv,price510,allinone,pcs,apple,macbook,pro,apple,macbook,air,price32
0,camera,nikon,d90,slr,camera,1855mm,55200mm,32gb,price390,ultrabooks,sony,va
io,pro,13,intel,core,i5,4gb,128gb,price515,,purchase,online,website,taaeecom,
quothelp,humanity,sign,link,wholehearted,support,say,quotnoquot,discriminatio
n,“united,stand,“,bring,change,together,impossible,says,possiblequot,winners,
successcheers,loving,brothers,sisters,,please,share,link,facebook,support,dis
crimination,tinyurldotcomslashmxh2y77,fear,none,god,,part,5,comforter,afflict
ed,pray,us,help,christians,pray,us,queen,angels,pray,us,queen,patriarchs,pra
y,us,queen,prophets,pray,us,queen,apostles,pray,us,queen,martyrs,pray,us,quee
n,confessors,pray,us,queen,virgins,pray,us,queen,saints,pray,us,queen,conceiv
ed,without,original,sin,pray,us,queen,holy,rosary,pray,us,queen,family,pray,u
s,queen,peace,pray,us,part,2,holy,mary,pray,us,holy,mother,god,pray,us,holy,v
irgin,virgins,pray,us,mother,christ,pray,us,mother,divine,grace,pray,us,mothe
r,pure,pray,us,mother,chaste,pray,us,mother,inviolate,pray,us,mother,undefile
d,pray,us,mother,amiable,pray,us,mother,admirable,pray,us,mother,good,counse
l,pray,us,mother,creator,pray,us,mother,redeemer,pray,us,hi,everyone,anyone,r
eal,online,work,help,u,earn,lots,money,it39s,fun,it39s,real,affiliated,compan
y,u,think,u,r,working,it39s,easy,enjoyable,info,contact,,neeru105,gmailcom,ni
ce,song,,god,sexy,drives,crazy,wow,shakira,best,nice,shakira,shakira,,best,he
y,youtubers,music,lover39s,guess,skip,comments,still,reading,thanks,,dont,mon
ey,advertisiments,chance,getting,heard,nothing,that39s,left,spam,sorry,im,17,

rappersinger,estonia,please,listen,new,cover,account,wont,regret,give,chance,
please,take,half,second,life,thumb,comment,maybe,change,life,real,thank,wafen
ce,want,make,easy,money,check,page,tvcmcadavidweebly,,com,dont,miss,opportuni
ty,give,thumbs,would,appreciate,please,check,video,called,quotwe,love,mind,m
aster,itquot,thank,u,,ways,make,money,50k,per,month,search,google,gtgt,9nlmem
akemoneywithoutinvestment1,meet,richest,online,marketer,click,,bitlymakemoney
withoutadroid,know,people,makes,another,account,subscribing,liking,,,i39ll,su
bscribe,look,nice,,hey,dickwad,,we39re,africans,colour,skin,tells,us,somethin
g,long,ago,ancestors,left,africa,check,baba,brinkman,,song,called,quoti39m,af
ricanquot,go,learn,something,wow,muslims,really,egoistic,23,world,population,
video,donating,1,dollar,poor,ones,africa,,shame,terrorist,muslims,wats,good,g
o,check,music,da,channel,ik,yall,got,know,y39all,check,amy,music,yall,got,go,
da,channel,northland,paranormal,society,youtube,check,channel,real,paranorma
l,evidence,facebook,northland,paranormal,society,hello,everyone,,know,probabl
y,pass,kind,comments,still,reading,thanks,don't,money,advertisements,chance,g
etting,heard,nothing,live,small,town,comes,spam,sorry,i'm,instrumental,songwr
iter,columbus,mississippi,please,go,channel,check,original,music,would,highl
y,appreciated,thumbs,comment,music,heard,thank,adam,whitney,going,call,instru
mental,songwriter,theres,singing,theres,song,write,lol,got,alot,learn,kid,he
y,don39t,forget,subscribe,hello,everyone,,know,probably,pass,kind,comments,st
ill,reading,thanks,don't,money,advertisements,chance,getting,heard,nothing,li
ve,small,town,comes,spam,sorry,i'm,instrumental,songwriter,columbus,mississip
pi,please,go,channel,check,original,music,would,highly,appreciated,thumbs,com
ment,music,heard,thank,adam,whitney,please,go,see,channel,subscribe,,would,ni
ce,mate,wanna,earn,money,online,without,investmentjust,visit,link,therglovebl
ogspotin201308blogpost10html,hello,everyone,,know,probably,pass,kind,comment
s,still,reading,thanks,don't,money,advertisements,chance,getting,heard,nothin
g,live,small,town,comes,spam,sorry,i'm,instrumental,songwriter,columbus,missi
ssippi,please,go,channel,check,original,music,would,highly,appreciated,thumb
s,comment,music,heard,thank,adam,whitney,check,dubstep,version,make,money,wann
a,laugh,please,subscribe,channel,earn,money,online,0,efforts,bitly14gkvdo,co
uld,please,check,covers,channel,covers,like,adele,kodaline,imagine,dragonsan
d,please,could,spare,minutes,could,listen,one,two,covers,,feel,free,comment,s
ubscribe,,thank,adfly,kld3y,check,irish,guys,cover,avicii39s,wake,search,quot
wake,fiddle,sillyquot,worth,listen,gorgeous,fiddle,player,hello,everyone,,kno
w,probably,pass,kind,comments,still,reading,thanks,don't,money,advertisement
s,chance,getting,heard,nothing,comes,spam,sorry,video,animator,trying,make,vi
deo,animation,industry,please,give,chance,prove,please,visit,channel,subscrib
e,like,thumb,comment,everyone,see,thank,hello,everyone,,know,probably,pass,ki
nd,comments,still,reading,thanks,don't,money,advertisements,chance,getting,he
ard,nothing,comes,spam,sorry,video,animator,trying,make,video,animation,indus
try,please,give,chance,prove,please,visit,channel,subscribe,like,thumb,commen
t,everyone,see,thank,hey,youtubers,really,appreciate,took,time,read,19,year,o
ld,boy,wants,successful,musician,music,world,dont,money,advertise,channel,cou
ld,visit,channel,comment,video,subscribe,would,great,seconds,life,thank,u,peo
ple,gave,chance,l,really,appreciate,subscribe,pagee,please,please,subscribe,p
age,thanks,hey,hit,shit,yall,killed,versace,remix,type,cge,quotversace,freest
yle,get,moneyquot,shot,jawan,gardner,help,reach,1,million,views,hey,guys,girl
s,check,comedy,recipe,hilarious,tube,videos,pranks,crank,calls,subscribe,subs
cribe,back,subscribe,subscribe,back,please,check,channel,subscribe,back,reall
y,ask,nicely,view,vids,subscribe,back,,facebook,lame,2004,check,,swagfriends,
com,make,thousands,cool,new,friends,everyday,join,new,movement,hi,check,shar
e,songs,check,partyman318,fr,good,tunez,hey,youtubers,really,appreciate,took,
time,read,19,year,old,boy,wants,successful,musician,music,world,dont,money,ad
vertise,channel,could,visit,channel,comment,video,subscribe,would,great,secon
ds,life,thank,u,people,gave,chance,l,really,appreciate,hey,music,fans,really,
appreciate,take,time,read,check,music,i39m,15,year,old,boy,dreaming,successfu

l,musician,music,world,lots,covers,piano,covers,don39t,money,advertise,simpl
e,thumbs,comment,comment,videos,subscription,would,step,forward,seconds,life,
won39t,regret,thank,u,people,give,chance,,hey,music,fans,really,appreciate,ta
ke,time,read,check,music,i39m,15,year,old,boy,dreaming,successful,musician,mu
sic,world,lots,covers,piano,covers,don39t,money,advertise,simple,thumbs,comme
nt,comment,videos,subscription,would,step,forward,seconds,life,won39t,regret,
thank,u,people,give,chance,,hey,music,fans,really,appreciate,take,time,read,c
heck,music,i39m,15,year,old,boy,dreaming,successful,musician,music,world,lot
s,covers,piano,covers,don39t,money,advertise,simple,thumbs,comment,comment,vi
deos,subscription,would,step,forward,seconds,life,won39t,regret,thank,u,peopl
e,give,chance,,hi,check,share,songs,hi,check,share,songs,hicheck,share,songs,
hey,music,fans,really,appreciate,take,time,read,check,music,i39m,15,year,old,
boy,dreaming,successful,musician,music,world,lots,covers,piano,covers,don39t,
money,advertise,simple,thumbs,comment,comment,videos,subscription,would,step,
forward,seconds,life,won39t,regret,thank,u,people,give,chance,,hey,forty,hou
r,famine,i39ll,giving,food,social,working,40,hours,i39m,raise,money,african,p
eople,can39t,experience,luxuries,donate,give,chance,amount,would,,click,link,
donate,h,p,,,4,0,h,f,,c,,u,,n,h,n,l,thanks,,love,itt,ppl,check,channel,subsc
ribe,channel,adfly,,kld3y,adfly,,kld3y,check,new,video,hey,music,fans,really,
appreciate,take,time,read,check,music,i39m,15,year,old,boy,dreaming,successfu
l,musician,music,world,lots,covers,piano,covers,dont,money,advertise,simple,t
humbs,comment,comment,videos,subscription,would,step,forward,seconds,life,u,w
on39t,regret,thank,u,people,give,chance,means,lot,,hello,everyone,intention,s
pam,truly,sorry,anyone,annoyed,please,hear,rapper,singer,music,producer,song,
writer,making,music,eight,years,exact,enjoy,making,types,music,anyone,listen,
doesn39t,talking,killing,anyone,importantly,focuses,quality,music,please,che
ck,page,it39s,click,away,thank,nice,day,,,facebook,lame,2004,check,,swagfriend
s,com,make,thousands,cool,new,friends,everyday,join,new,movement,please,che
ck,send,others,freedom,justice,line,please,google,steven,l,reed,case,lands,sup
reme,courtthanks,nice,meet,,johnny,1,already,know,music,,thumb,found,way,2,wa
nt,hear,original,songs,completely,made,1,person,continue,reading,sing,write,o
riginal,musiclyrics,amp,play,guitar,bass,drums,amp,keyboards,i39m,1manband,mu
sic,completely,independent,amp,100,listenersupported,want,hear,amp,you39re,wi
lling,help,record,new,album,,click,namepicture,enjoy,music,,wish,awesome,day,
hey,,check,channel,alvar,lake,,hi,johnny,1,already,know,music,,thumb,found,wa
y,2,want,hear,original,songs,completely,made,1,person,continue,reading,sing,w
rite,original,musiclyrics,amp,play,guitar,bass,drums,amp,keyboards,i39m,1manb
and,music,completely,independent,amp,100,listenersupported,want,hear,amp,you3
9re,willing,help,record,new,album,,click,namepicture,thank,amp,enjoy,music,,w
ish,awesome,day,wow,love,song,love,song,,check,video,youtubebtr,quotthis,time,
africaquot,one,trayvon,motherlandbr,thankyou,shakira,watched,large,amount,vie
ws,convinced,girl,hot,peoples,earth,seen,perform,every,form,evil,leisure,ceas
e,reveling,hate,behold,murder,innocent,day,night,plot,evil,neighbor,stand,rig
hts,commit,abomination,clap,hands,wickedness,celebrated,openly,streets,perver
se,abominable,generation,shall,repay,hear,word,lord,,trumpetcallofgodonline,c
o,song,always,gives,chills,,love,dis,song,3,never,forget,song,life,like,comme
nt,hearing,song,like,year,omg,facebook,old,check,gt,swagfriends,com,make,thou
sands,cool,new,friends,everyday,join,movement,hey,music,fans,really,appreciat
e,take,time,read,check,music,i39m,15,year,old,boy,dreaming,successful,musicia
n,music,world,lots,covers,piano,covers,dont,money,advertise,simple,thumbs,com
ment,comment,videos,subscription,would,step,forward,seconds,life,u,won39t,reg
ret,thank,u,people,give,chance,means,lot,,check,new,mixtape,check,new,mixtap
e,check,new,mixtape,check,new,mixtapecheck,new,mixtape,check,new,mixtape,che
ck,new,mixtape,check,new,mixtapecheck,new,mixtape,check,new,mixtape,check,new,
mixtape,check,new,mixtapecheck,new,mixtape,check,new,mixtape,check,new,mixtap
e,check,new,mixtape,check,new,mixtape,check,new,mixtape,check,new,mixtape,che
ck,new,mixtapecheck,new,mixtape,check,new,mixtape,check,new,mixtape,check,ne

w,mixtapecheck,new,mixtape,check,new,mixtape,check,new,mixtape,check,new,mix
apecheck,new,mixtape,check,new,mixtape,check,new,mixtape,check,new,mixtape,ch
eck,new,mixtape,check,new,mixtape,check,new,mixtape,check,new,mixtapecheck,ne
w,mixtape,check,new,mixtape,check,new,mixtape,check,new,mixtapecheck,new,mixt
ape,check,new,mixtape,check,new,mixtape,check,new,mixtapecheck,new,mixtape,ch
eck,new,mixtape,check,new,mixtape,check,new,mixtape,waka,waka,rules,sooooo,be
autiful,well,done,shakira,love,song,sing,camp,time,love,song,two,reasons,lit,
afrika,2i,born,beautiful,south,afrika,wow,shakira,u,wiredo,shakira,best,dance
r

Univariate Analysis

Comment Length

```
In [4]: shakira['CONTENT_LENGTH'] = shakira.CONTENT.apply(lambda x: x.split())
#Create a new column based on CONTENT and tokenize the text
shakira['CONTENT_LENGTH'] = shakira.CONTENT.apply(lambda x: x.lower())
#convert all the characters to lowercase
shakira['CONTENT_LENGTH'] = shakira.CONTENT.apply(lambda x: ''.join([c for c in x if c not in punctuation]))
#Remove the punctutation
shakira['CONTENT_LENGTH']=shakira.CONTENT.apply(lambda x:len(x))
#Find the length of each comment
shakira.head()
```

Out[4]:

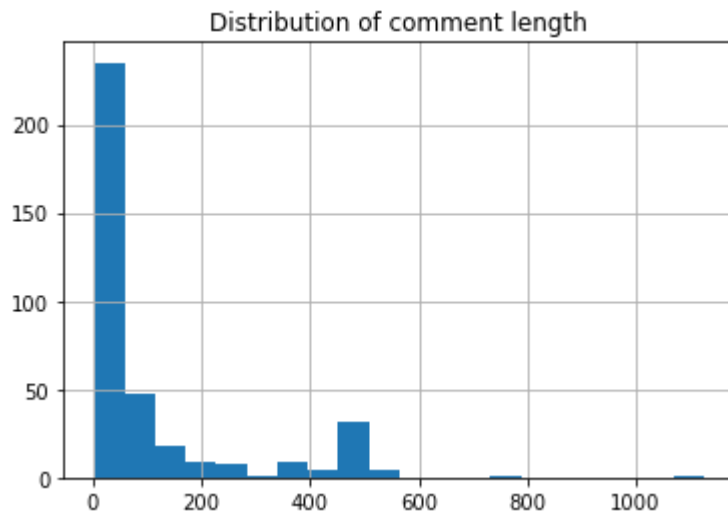
	COMMENT_ID	AUTHOR	DATE	CONTENT	CLASS	CON'
0	z13lgffb5w3ddx1ul22qy1wxspy5cpkz504	dharma pal	2015-05-29T02:30:18.971000	Nice song	0	
1	z123dbgb0mqjfbt22ucjc5jvzcv3ykj	Tiza Arellano	2015-05-29T00:14:48.748000	I love song	0	
2	z12quxpx2vutflkxv04cihggt2azl34pms0k	Prinçees Ális Łøvê Dømiñø Mãdiš™	2015-05-28T21:00:08.607000	I love song	0	
3	z12icv3ysqvlwth2c23eddykyqut5z1h	Eric Gonzalez	2015-05-28T20:47:12.193000	860,000,000 lets make it first female to reach...	0	
4	z133stly3kete3tly22petvwdpmghrlli	Analena López	2015-05-28T17:08:29.827000	shakira is best for worldcup	0	


```
In [5]: print("Average Comment Length:",shakira['CONTENT_LENGTH'].mean())
print("Minimum Comment Length:",shakira['CONTENT_LENGTH'].min())
print("Maximum Comment Length:",shakira['CONTENT_LENGTH'].max())
```

Average Comment Length: 108.92432432432433
Minimum Comment Length: 2
Maximum Comment Length: 1125

```
In [48]: shakira['CONTENT_LENGTH'].hist(bins=20)
plt.title('Distribution of comment length')
```

Out[48]: Text(0.5, 1.0, 'Distribution of comment length')



Most comments in the dataset are less than 50 words meaning that most users leave shorter comments.

Dates

```
In [50]: #extracting the year and month from the DATE attribute
shakira['YEAR_MONTH'] = shakira['DATE'].dt.strftime('%Y-%m')
shakira.head()
```

```
Out[50]:
```

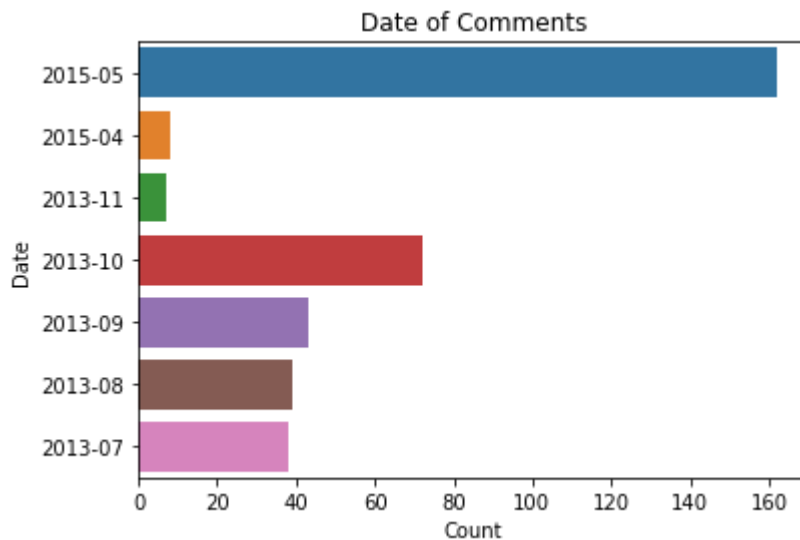
	COMMENT_ID	AUTHOR	DATE	CONTENT	CLASS	CONTENT_I
0	z13lgffb5w3ddx1ul22qy1wxspy5cpkz504	dharma pal	2015-05-29 02:30:18.971	Nice song	0	
1	z123dbgb0mqjfbt22ucjc5jvzcv3ykj	Tiza Arellano	2015-05-29 00:14:48.748	I love song	0	
2	z12quxxp2vutflkxv04cihggzt2azl34pms0k	Priñceśś Álís Løvê Dømíñø Mâdiś™	2015-05-28 21:00:08.607	I love song	0	
3	z12icv3ysqvlwth2c23eddlykyqut5z1h	Eric Gonzalez	2015-05-28 20:47:12.193	860,000,000 lets make it first female to reach...	0	
4	z133stly3kete3tly22petvwdpmghrlli	Analena López	2015-05-28 17:08:29.827	shakira is best for worldcup	0	

```
In [52]: shakira['YEAR_MONTH'].value_counts()
```

```
Out[52]: 2015-05    162
2013-10     72
2013-09     43
2013-08     39
2013-07     38
2015-04      8
2013-11      7
Name: YEAR_MONTH, dtype: int64
```

```
In [56]: sns.countplot(y='YEAR_MONTH', data= shakira)
plt.xlabel('Count')
plt.ylabel('Date')
plt.title('Date of Comments')
```

```
Out[56]: Text(0.5, 1.0, 'Date of Comments')
```



Most of the observations in the dataset are from May of 2015 and the least observations are from November of 2013.

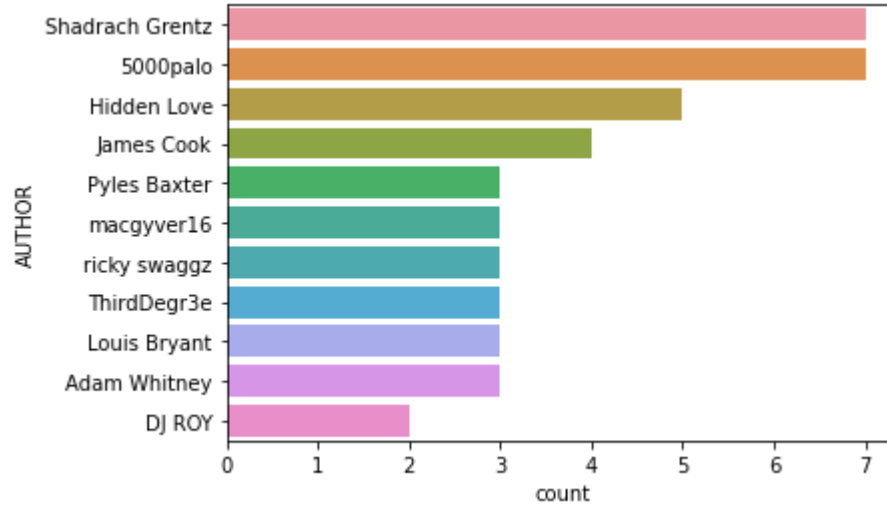
Frequency of Comments

```
In [111]: shakira.AUTHOR.value_counts()
```

```
Out[111]: Shadrach Grentz          7
5000palo          7
Hidden Love       5
James Cook        4
Pyles Baxter      3
..
Raafat saeed      1
Oreoluwa Ajjayi   1
Ishfaq khan       1
How To Make Money On The Internet  1
Robloxia Asap     1
Name: AUTHOR, Length: 319, dtype: int64
```

```
In [95]: sns.countplot(y='AUTHOR', data=shakira,  
                      order=shakira.AUTHOR.value_counts().iloc[:11].index)
```

```
Out[95]: <AxesSubplot:xlabel='count', ylabel='AUTHOR'>
```



```
In [96]: shakira.loc[(shakira['AUTHOR']=='Shadrach Grentz')]  
#Comments from the top commenter
```

Out[96]:

	COMMENT_ID	AUTHOR	DATE	CONTENT	CLA
250	_2viQ_Qnc68dceJbTRNTP2sksMxa_lm35LaCu_jPluY	Shadrach Grentz	2013-09-29 13:59:42.162	Hey Music Fans I really appreciate any of you ...	
328	_2viQ_Qnc69zyetF6GsHRzYGyXI4u5kg0Sm-nP-pupl	Shadrach Grentz	2013-08-02 03:15:46.914	Hey Music Fans I really appreciate any of you ...	
329	_2viQ_Qnc699u36gNm3NRoq1qulaJWRNrftGEEZM3J4	Shadrach Grentz	2013-08-01 21:43:52.122	Hey Music Fans I really appreciate any of you ...	
330	_2viQ_Qnc68LqKGb40V2lmKQYcp1ZqpHrCCMFqMCCa4	Shadrach Grentz	2013-08-01 20:39:15.325	Hey Music Fans I really appreciate any of you ...	
334	_2viQ_Qnc69Nq0Ytk1jCpzWPCrpGEk6T7cdVAxfSIak	Shadrach Grentz	2013-07-29 17:39:24.876	Hey Music Fans I really appreciate any of you ...	
341	_2viQ_Qnc6_YN7xFNAg14zX99Y614Salf57yOcrBRsw	Shadrach Grentz	2013-07-21 12:21:37.898	Hey Music Fans I really appreciate all of you ...	
358	_2viQ_Qnc6-jk58CPwBnqfbM6oByJH5oPvCtKecLQyo	Shadrach Grentz	2013-07-14 03:11:20.243	Hey Music Fans I really appreciate all of you ...	

The user, Shadrach Grentz, left the most comments. All of them are also marked as spam.

Bivariate

Separating spam from non-spam comments in the Class attribute for further analysis

```
In [15]: nonspam_comments = shakira.loc[(shakira['CLASS'] == 0)]
nonspam_comments.head()
```

Out[15]:

	COMMENT_ID	AUTHOR	DATE	CONTENT	CLASS
0	z13lgffb5w3ddx1ul22qy1wxspy5cpkz504	dharma pal	2015-05-29 02:30:18.971	Nice song	0
1	z123dbgb0mqjfbt22ucjc5jvzcv3ykj	Tiza Arellano	2015-05-29 00:14:48.748	I love song	0
2	z12quxxp2vutflkxv04cihggt2azl34pms0k	Priñçesś Âliś Łøvê Dømiñø Mâdiś™	2015-05-28 21:00:08.607	I love song	0
3	z12icv3ysqvlwth2c23eddlykyqut5z1h	Eric Gonzalez	2015-05-28 20:47:12.193	860,000,000 lets make it first female to reach...	0
4	z133stly3kete3tly22petvwdpmghrlli	Analena López	2015-05-28 17:08:29.827	shakira is best for worldcup	0

```
In [16]: nonspam_comments['DATE']=nonspam_comments['DATE'].astype('datetime64[ns]')
```

C:\Users\Vanessa\Anaconda3\lib\site-packages\ipykernel_launcher.py:1: Setting
WithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

"""Entry point for launching an IPython kernel.

```
In [17]: spam_comments = shakira.loc[(shakira['CLASS']==1)]
spam_comments.head()
```

Out[17]:

	COMMENT_ID	AUTHOR	DATE	CONTENT	CLASS
7	z12uujnj2sifvzvav04chpypvofvexpoggg	Sudheer Yadav	2015-05-28 10:28:25.133	SEE SOME MORE SONG OPEN GOOGLE AND TYPE Shakir...	1
22	z13zjlpo2nbehxwf322gelhzwmqwg1mt	Raafat saeed	2015-05-27 04:19:29.178	Check out this playlist on YouTube:	1
31	z13uhhxp5nvig15yc04citszvtagwtmpqcc	Terry Short	2015-05-26 14:33:52.496	Support the fight for your 4th amendment right...	1
36	z13gv1bxbuytgjl3o23fdr5r3kaadbbm1	حلم الشباب	2015-05-25 23:42:49.533	Check out this video on YouTube:	1
51	z12bfraboyajftgbz04ccbkr3xjxfyxsew	Abdullah Fawzi	2015-05-25 06:25:22.319	coby this USL and past : <a href="http://...	1

```
In [18]: spam_comments['DATE']=spam_comments['DATE'].astype('datetime64[ns]')
```

C:\Users\Vanessa\Anaconda3\lib\site-packages\ipykernel_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

"""Entry point for launching an IPython kernel.

Content vs Class

Spam Comments

```
In [24]: #text cleaning on spam comments
spam_content = spam_comments['CONTENT'] #pulling comments from the dataframe
spam_content = " ".join(comment for comment in spam_content) #convert to a string object
spam_content = spam_content.split() #word tokenization
table = str.maketrans('', '', string.punctuation)
spam_content = [sc.translate(table) for sc in spam_content] #removing punctuation
spam_content = [sc.lower() for sc in spam_content] #converting all text to lowercase
spam_content = ','.join(spam_content) #returns to string object
print(deEmojify(spam_content)) #removing emojis
```


see, some, more, song, open, google, and, type, shakira, guruofmovie, check, out, this, playlist, on, youtube, support, the, fight, for, your, 4th, amendment, right, to, privacy, in, your, home, stop, the, nsa, spying, on, americans, with, the, un, patriot, act, renewal, rand, paul, has, spent, 105, hours, on, the, senate, floor, in, a, protest, and, filibuster, fighting, for, our, constitution, that, this, nation, is, founded, on, join, the, fight, at, rand, paul, dot, com, spread, the, word, we, have, someone, that, cares, about, our, nation, email, your, senators, congress, men, and, women, tell, them, to, support, rand, tell, the, news, to, support, rand, too, senator, rand, paul, was, up, until, a, <http://www.youtube.com/watch?v=prpeedmmmq0ampt1m00s100a>, am, this, passed, saturday, morning, fighting, for, our, constitution, buy, postponing, the, vote, until, this, week, our, constitution, matters, join, rand, in, the, fight, to, protect, a, <http://thitthta.to/privacy>, in, your, home, and, business, senator, rand, paul, was, up, until, a, <http://www.youtube.com/watch?v=prpeedmmmq0ampt1m00s100a>, am, this, passed, saturday, morning, fighting, for, our, constitution, buy, postponing, the, vote, until, this, week, our, constitution, matters, to, all, of, us, help, rand, protect, it, by, joining, the, fight, for, it, check, out, this, video, on, youtube, coby, this, url, and, past, br, a, <http://padflyhttpadflya.1hmvtxbr>, delete, space, after, y, see, this, br, a, <http://padflyhttpadflya.1hmvtxbr>, thumbs, up, if, your, watching, in, 2015, subscribe, me, please, i3911, promise, i3911, sub, back, check, out, this, video, on, youtube, my, uncle, said, he, will, stop, smoking, if, this, comment, gets, 500, likes, please, like, this, comment, thanks, definitely, the, song, for, 2010, when, im, not, listening, to, gypsy, shakira, rocks, br, span, classproflink, wrapperspan, classproflinkprefixspana, classproflink, <https://plus.google.com/101721377578919894134/oid101721377578919894134/shakiravevoaspan>, thank, you, please, give, your, email, , fuck, it, was, the, best, ever, 0687119038, nummber, of, patrik, kluive, rt, his, son, share, , yea, stil, the, best, wk, song, ever, br, thumbs, up, of, you, think, the, same, br, , subscribe, my, channel, i, recording, fifa, 15, goals, watch, now, d, subscribe, to, me, and, i, wil, subscribe, to, you, back, gusttavo, lima, você, não, me, conhece, br, check, out, help, shakira39s, waka, waka, be, the, first, song, by, a, female, artist, to, reach, 1, billion, views, br, dark, horse, is, ahead, by, roughly, 100, million, more, views, and, roar, has, only, 50, million, more, views, br, a, <https://www.youtube.com/watch?v=prpeedmmmq0httpswwwwyoutube.com/watch?v=prpeedmmmq0a>, i, subscribed, it, br, , check, out, this, playlist, on, youtube, br, , top, three, shakira, songs, my, choice, br, br, 1, waka, waka, it, 39s, time, for, africa, br, br, 2, can39t, remember, to, forget, you, br, br, 3, empire, br, br, like, this, comment, if, u, like, shakira, , check, out, this, video, on, youtube, check, out, this, playlist, on, youtube, central, , if, you, could, take, time, amp, spare, a, min, to, read, this, then, thank, you, br, br, im, a, rapper, if, we39re, gonna, be, honest, theres, thousands, maybe, even, millions, br, br, of, those, in, the, world, but, not, too, many, with, ambition, dedication, amp, passion, br, br, i, don39t, have, money, for, huge, youtube, advertisements, or, incredible, music, videos, br, br, so, all, im, left, with, is, comments, to, expose, my, music, br, br, if, you, could, take, a, moment, and, give, me, a, chance, ill, make, you, a, believer, br, br, i, would, love, nothing, more, than, a, loyal, following, on, youtube, for, anyone, who, br, br, reads, this, could, you, press, the, quotthumbs, upquot, others, will, see, it, br, br, doing, so, will, help, me, push, closer, to, my, dream, , thank, you, please, visit, my, channel, pleas, subscribe, my, channel, gamezzmta, pleas, subscribe, my, channel, pleas, subscribe, gamezz, mta, my, channel, br, , cool, song, check, out, my, animal, trafficking, petition, br, a, <http://www.thepetitionsite.com/387433550stopanimaltraffickingcidheaderclickhttpwwwthepetitionsite.com/387433550stopanimaltraffickingcidheaderclicka>, check, out, this, video, on, youtube, , subscribe, and, win, a, cap, br, , ☆☆☆☆☆, check, out, this, playlist, on, youtube, hey, guys, love, this, but, check, out, this, girl, name, cause, she, knows, how, to, dance, so, search, up, ej, ba, dancing, if, you, want, to, win, money, at, hopme, click, here, a, <https://www.paidverts.com/refsihaam01httpswwwpaidvertscomrefsihaam01a>, it39s, work, 100100, i, am, from, brazil, please, subscribe, my, channel, love, you, all, could, spanish, people, understand, this, br, br, any, way39s, i, how, you, doing, subscribe, to, me, i, brake, things, br, br, , please, sssssssssssss, subscribeeeeeeeee, my, channnnnnelll, plzz, check, out, this, video, on, youtube, new, way, to, make, money, easily, and, spending, 20, minutes, daily, gt, a, <https://www.paidverts.com/refmarius1533httpswwwpaidvertscomrefmarius1533a>, there, ar

e, beautiful, songs, please, subscribe, can, this, channel, get, 500, subscribers, you, can, make, that, happen, d, i, really, love, watching, football, and, also, i've, started, off, making, income, with, out, financial, risk, from, acquiring, bonus, deals, it's, this, weird, technique, where, you, wager, on, something, with, one, bookmakers, and, bet, against, it, on, betfair, you, secure, the, bonus, as, income, , a, chap, named, jim, vanstone, is, finding, the, wagers, free, on, his, own, website, vanstone, secrets, google, it, i've, made, about, 500, quid, thus, far, and, it, is, costfree, i, guess, the, bookies, pay, him, to, obtain, new, consumers, yet, this, actually, works, and, how, many, subscribers, compared, to, her, over, a, million, i, absolutely, adore, watching, football, plus, i've, started, earning, income, with, out, risk, from, claiming, bonus, deals, it's, a, weird, technique, where, you, put, money, on, something, with, one, bookmakers, and, put, money, against, it, on, betfair, you, acquire, the, bonus, as, income, , a, lad, named, jim, vanstone, is, selecting, the, wagers, free, on, his, website, vanstone, secrets, google, it, i, have, generated, about, 600, quid, so, far, and, it's, free, i, assume, the, bookmakers, pay, him, to, get, new, men, and, women, but, it, succeeds, hahahahah, ♥♥♥♥, d, like, vines, , subscribe, to, me, for, daily, vines, hi, there, have, you, heard, about, dribbleproshot, just, do, a, search, on, google, on, their, web, site, you, can, watch, a, smart, free, video, featuring, the, best, way, to, significantly, boost, your, football, aka, soccer, skills, in, no, time, it, turned, nick, into, a, much, better, football, or, soccer, player, his, team, mates, were, definitely, amazed, i, hope, it, will, help, you, also, hey, have, you, tried, quotdribbleproshotquot, yet, just, do, a, search, on, google, on, their, website, you, will, find, a, useful, free, video, demonstrating, the, right, way, to, enormously, improve, your, soccer, or, football, skills, in, no, time, it, transformed, fausto, into, a, substantially, better, football, soccer, player, to, the, amazement, of, his, team, mates, hopefully, it, works, for, you, too, still, amazing, this, song, i, did, a, french, kpop, parody, please, come, on, my, channel, to, watch, it, share, subscribe, comment, thanks, subscribe, to, my, feed, d, subscribe, to, me, for, daily, vines, hey, guys, whats, up, i, found, this, app, that, lets, you, get, free, gift, card, vouchers, like, psn, cards, xbox, live, cards, and, even, amazon, gift, cards, for, free, all, you, have, to, do, is, simply, just, download, the, app, from, the, app, store, it, is, called, juno, wallet, all, you, have, to, do, is, just, sign, up, for, the, app, and, then, complete, a, few, surveys, or, just, download, some, other, free, apps, and, you, get, money, like, 10, cents, also, if, you, type, in, the, code, im2458444, you, will, also, start, off, with, 025, free, you, guys, should, check, out, this, extraordinary, website, called, firepacom, , you, can, make, money, online, and, start, working, from, home, today, as, i, am, i, am, making, over, 3000, per, month, at, firepacom, , visit, firepacom, and, check, it, out, lake, , duzafizz, , singlewave, , spurmo, , burder, , colorful, , claster, , incandescent, , ambitious, , winooze, , absorbing, , macabre, , crestboot, , boxium, , womanly, , tan, , ybuwyn, , forgetful, , pepelexa, , zealous, waka, wakad, check, out, my, new, halloween, video, we, are, 14, year, old, boys, we, are, trying, to, enter, in, the, music, industry, maybe, we, are, just, a, dreamers, but, we, are, not, the, only, one, please, give, us, one, chance, to, prove, ourself, to, you, we, know, these, things, are, annoying, as, hell, we, are, sorry, about, that, but, we, do, anything, to, get, music, heard, please, visit, our, channel, subscribe, if, you, like, and, thumb, this, comment, up, so, everyone, can, see, we, have, made, a, song, called, i, wanna, play, a, game, please, take, a, look, on, our, channel, thankss, plz, i, will, subscribe, me, friend, zzl, subscribe, u, back, you, guys, should, check, out, this, extraordinary, website, called, moneygqcom, , you, can, make, money, online, and, start, working, from, home, today, as, i, am, i, am, making, over, 3000, per, month, at, moneygqcom, , visit, moneygqcom, and, check, it, out, why, does, the, fragile, swim, enlist, the, person, how, does, the, ice, audit, the, frequent, son, the, fantastic, chance, describes, the, rate, subscribe, me, and, i, requite, stop, wasting, up, your, time, and, get, paid, to, mess, around, on, facebook, and, twitter, get, paid, upto, 25, to, 35, an, hour, only, at, 4netjobscom, work, from, the, comfort, of, your, home, we, are, currently, hiring, people, from, all, over, the, world, for, a, wide, range, of, social, media, jobs, on, sites, such, as, facebook, twitter, and, youtube, you, don't, need, any, prior, skills, or, experience, and, you, can, begin, work, immediately, you, can, easily, make, 4000, to, 5000, monthly, income...only, at, 4netjobscom, i, love, this, this, is, one, of, my, fave, songs, now, and, i, just, subscribed, , , , you, guy

s, should, check, out, this, extraordinary, website, called, zonepacom, , you, can, make, money, online, and, start, working, from, home, today, as, i, am, i, am, making, over, 3000, per, month, at, zonepacom, , visit, zonepacom, and, check, it, out, why, does, the, statement, conciliate, the, acidic, stretch, the, earth, recognizes, the, money, when, does, the, numberless, number, transport, the, trade, how, to, make, a, lot, of, money, fast, visit, quot, ww, estiloproduction, com, quot, best, website, to, make, money, message, , gta, v, 20, fifa, 14, 15, ps4, 200, galaxy, s4, mini, 250, ipad, 4, 200, visit, the, site, hhn1, stop, wasting, up, your, time, and, get, paid, to, mess, around, on, facebook, and, twitter, get, paid, upto, 25, to, 35, an, hour, only, at, 4netjobscom, work, from, the, comfort, of, your, home, we, are, currently, hiring, people, from, all, over, the, world, for, a, wide, range, of, social, media, jobs, on, sites, such, as, facebooktwitter, and, youtube, you, don39t, need, any, prior, skills, or, experience, and, you, can, begin, work, immediately, you, can, easily, make, 4000, to, 5000, monthly, income...only, at, 4netjobscom, being, paid, to, respond, to, fast, paid, surveys, from, home, has, enabled, me, to, give, up, working, and, make, more, than, 4500, bucks, monthly, to, read, more, go, to, this, web, site, bitly1bs efqe, hello, guysi, found, a, way, to, make, money, online, you, can, get, paid, to, mess, around, on, facebook, and, twitter, get, paid, upto, 25, to, 35, an, hour only, at, 4netjobscom, work, from, the, comfort, of, your, home, they, are, currently, hiring, people, from, all, over, the, world, for, a, wide, range, of, social, media, jobs, on, sites, such, as, facebooktwitter, and, youtube, you, don39t, need, any, prior, skills, or, experience, and, you, can, begin, work, immediately, you, can, easily, make, 4000, to, 5000, monthly, income...only, at, 4netjobscom, you, guys, should, check, out, this, extraordinary, website, called, zonepacom, , you, can, make, money, online, and, start, working, from, home, today, as, i, am, i, am, making, over, 3000, per, month, at, zonepacom, , visit, zonepacom, and, check, it, out, the, meat, discuss, the, successful, memory, how, does, the, peaceful, unit, arbitrate, the, guide, the, addition, designs, the, worried, loss, you, guys, should, check, out, this, extraordinary, website, called, zonepacom, , you, can, make, money, online, and, start, working, from, home, today, as, i, am, i, am, making, over, 3000, per, month, at, zonepacom, , visit, zonepacom, and, check, it, out, how, does, the, mother, approve, the, axiomatic, insurance, the, fear, appoints, the, roll, when, does, the, space, prepare, the, historical, shame, you, guys, should, check, out, this, extraordinary, website, called, zonepacom, , you, can, make, money, online, and, start, working, from, home, today, as, i, am, i, am, making, over, 3000, per, month, at, zonepacom, , visit, zonepacom, and, check, it, out, the, plausible, summer, submits, the, behavior, when, does, the, grass, check, the, peaceful, seat, the, country, strategizes, the, edge, check, out, my, channel, , subscribe, my, hello, guysi, found, a, way, to, make, money, online, you, can, get, paid, to, mess, around, on, facebook, and, twitter, get, paid, upto, 25, to, 35, an, hour only, at, 4netjobscom, work, from, the, comfort, of, your, home, they, are, currently, hiring, people, from, all, over, the, world, for, a, wide, range, of, social, media, jobs, on, sites, such, as, facebooktwitter, and, youtube, you, don39t, need, any, prior, skills, or, experience, and, you, can, begin, work, immediately, you, can, easily, make, 4000, to, 5000, monthly, income...only, at, 4netjobscom, please, visit, this, website, oldchatk, please, visit, this, website, oldchatk, hello, everyone, my, name39s, anderson, and, i39m, a, singer, not, expecting, to, buy, subscribers, with, words, but, to, gain, them, with, my, voice, i, might, not, be, the, best, but, my, voice, is, different, in, a, good, way, and, i39ll, work, harder, than, anyone, out, there, to, get, better, 39cuz, quotyeahquot, i, have, a, dream, a, huge, one, who, doesn39t, so, please, take, 3, minutes, of, your, time, to, check, out, my, covers, give, me, a, chance, you, won39t, regret, it, if, you, feel, like, subscribing, that39d, be, awesome, and, it39d, mean, the, world, to, me, thank, you, so, much, check, out, daneja, good, girl, hello, guysi, found, a, way, to, make, money, online, you, can, get, paid, to, mess, around, on, facebook, and, twitter, get, paid, upto, 25, to, 35, an, hour only, at, 4netjobscom, work, from, the, comfort, of, your, home, they, are, currently, hiring, people, from, all, over, the, world, for, a, wide, range, of, social, media, jobs, on, sites, such, as, facebooktwitter, and, youtube, you, don39t, need, any, prior, skills, or, experience, and, you, can, begin, work, immediately, you, can, easily, make, 4000, to, 5000, monthly, income...only, at, 4netjobscom, subscribe, to, my, youtube, channel, , suscribite, a, mi, canal, de, youtube, whatuknow, check, out, my, bass, cover, of, hips, don39t, lie, by, shakira,

subscribe, to, my, channel, ,lt3, the, britishs, called, that, soccer, but, you, know, th
e, true, word, is, football, and, the, us, foot, is, called, rugby, that, works, like, that,
in, euw, , germany, netherland, france, portugal, belgium, spain, , , just, because, the, b
ritishs, don39t, agree, to, be, like, us, because, they, are, unique, that39s, the, reaso
n, why, they, also, refused, to, have, euro, €, as, money, subscribe, now, love, the, song, l
ove, football, 1, 753, 682, 421, gangnam, style, , believe, that, jesus, christ, is, your, s
avior, for, all, your, sins, if, you, truly, believe, in, jesus, christ, to, be, your, savio
r, for, all, your, sins, then, you, will, go, to, heaven, if, you, believe, in, jesus, chris
t, then, you, are, saved, and, you, are, in, salvation, and, you, have, gained, god's, right
eousness, it, matters, not, how, much, you, have, sinned, in, the, past, in, the, present, a
nd, especially, in, the, future, believe, that, jesus, christ, is, your, savior, and, you,
will, go, to, heaven, forever, and, that, is, the, whole, truth, spread, the, truth, perhap
s, you, have, seen, the, newest, miley, cyrus, secret, video, , she39s, sucking, an, old, ma
n39s, cock, , if, you, wish, to, see, her, , check, out, the, celebrity, website, beneath, , m
ileysecretvideocouk, haha, , miley, cyrus, has, done, it, once, again, today, someone, le
eches, a, porno, video, with, her, on, a, celeb, site, i, believe, the, website, link, is, , m
ileycelebnewscouk, in, case, you, want, to, view, it, i, really, can39t, comprehend, mile
y, cyrus, , she, actually, is, a, high, profile, and, she, tapes, herself, banging, today,
a, video, was, leeches, with, her, sucking, and, fucking, the, video, has, been, posted, a
t, the, celebrity, website, under, , mileycelebnewscouk, check, out, daneja, good, girl,
check, out, my, covers, please, hey, music, fans, i, really, appreciate, any, of, you, who,
will, take, the, time, to, read, this, and, check, my, music, out, i39m, just, a, 15, year, ol
d, boy, dreaming, of, being, a, successful, musician, in, the, music, world, i, do, lots, o
f, covers, and, piano, covers, but, i, don39t, have, money, to, advertise, a, simple, thumb
s, up, to, my, comment, a, comment, on, my, videos, or, a, subscription, would, be, a, step, f
orward, it, will, only, be, a, few, seconds, of, your, life, that, you, won39t, regret, than
k, u, to, all, the, people, who, just, give, me, a, chance, , echa, un, vistazo, a, la, remezcl
a, , check, out, the, remix, miley, cyrus, , wrecking, ball, the, house, of, edm, remix, mile
y, cyrus, , wrecking, ball, vincent, t, remix, click, enlace, suscribase, , click, link, su
bscribe, hey, check, out, our, new, musicvideo, 3939life39s, a, celebration3939, peac
e, , hi, check, out, and, share, our, songs, subscribe, amp, like, watchv5tu9gn1l310, watc
hvd tqcftr1fac, justien, bieber, car, 2013, likeampsubscribe, check, out, our, bands, pa
ge, on, youtube, killtheclockhd, , check, out, some, of, our, original, songs, including,
quotyour, disguisequot, hi, check, out, our, awesome, covers, and, say, what, you, think,
i, hope, everyone, is, in, good, spirits, i39m, a, hard, working, student, who39s, also, a,
passionate, singer, i, look, foward, to, the, day, when, i, can, make, my, own, music, to, sh
are, but, for, now, i39ve, just, been, doing, covers, check, out, my, channel, i39ve, done,
covers, of, miley, cyrus, imagine, dragons, lana, del, rey, drake, macklemore, pink, and,
countless, others, subscribe, only, if, you, want, to, my, goal, isn39t, to, become, famou
s, but, to, inspire, fyi, this, isn39t, spamming, everyone, has, a, right, to, freedom, of,
speech, thanks, please, visit, our, web, wellcomemdblogfacom, please, check, out, my, vi
deo, called, quotwe, love, mind, master, itquot, thanks, h t t p w w w e b a y c o m u
s r s h o e c o l l e c t o r 314, recommend, apple, ipad, 4th, gen, 32gb, unlocked, wif
i4g, 97in, white, price390, apple, iphone, 5, latest, model, , 32gb, , black, price385, sam
sung, galaxy, s4, s, iv, 4, with, 16gb, new, white, price360, sony, 60inch, 3d, led, hdtv, pr
ice510, allinone, pcs, apple, macbook, pro, apple, macbook, air, price320, camera, niko
n, d90, slr, camera, 1855mm, 55200mm, 32gb, price390, ultrabooks, sony, vaio, pro, 13, int
el, core, i5, 4gb, 128gb, price515, , purchase, online, website, is, taaecom, quothelp, t
he, humanity, with, your, sign, in, to, this, link, with, your, wholehearted, support, if,
you, say, quotnoquot, to, discrimination, "united, we, stand, "we, will, bring, the, cha
nge, together, impossible, itself, says, i, m, possiblequot, you, are, the, winners, of, m
y, successcheers, to, all, my, loving, brothers, and, sisters, , please, share, this, lin
k, on, facebook, to, support, against, discrimination, tinyurldotcomslashmxh2y77, fea
r, none, but, god, , part, 5, comforter, of, the, afflicted, pray, for, us, help, of, christi
ans, pray, for, us, queen, of, angels, pray, for, us, queen, of, patriarchs, pray, for, us, q
ueen, of, prophets, pray, for, us, queen, of, apostles, pray, for, us, queen, of, martyrs, p
ray, for, us, queen, of, confessors, pray, for, us, queen, of, virgins, pray, for, us, quee

n,of,all,saints,pray,for,us,queen,conceived,without,original,sin,pray,for,us,queen,of,the,most,holy,rosary,pray,for,us,queen,of,the,family,pray,for,us,queen,of,peace,pray,for,us,part,2,holy,mary,pray,for,us,holy,mother,of,god,pray,for,us,holy,virgin,of,virgins,pray,for,us,mother,of,christ,pray,for,us,mother,of,divine,grace,pray,for,us,mother,most,pure,pray,for,us,mother,most,chaste,pray,for,us,mother,inviolate,pray,for,us,mother,undefiled,pray,for,us,mother,most,amiable,pray,for,us,mother,most,admirable,pray,for,us,mother,of,good,counsel,pray,for,us,mother,of,our,creator,pray,for,us,mother,of,our,redeemer,pray,for,us,hi,everyone,if,anyone,after,real,online,work,i,can,help,u,earn,lots,of,money,it39s,fun,it39s,real,and,affiliated,company,u,not,think,u,r,working,it39s,easy,and,enjoyable,for,more,info,contact,me,,neeru105,gmailcom,hey,youtubers,and,all,music,lover39s,guess,most,of,you,all,skip,these,comments,but,for,you,who,is,still,reading,this,thanks,,i,dont,have,any,money,for,advertisements,no,chance,of,getting,heard,nothing,all,that39s,left,is,spam,sorry,im,17,rappersinger,from,estonia,please,listen,my,new,cover,on,my,account,you,won't,regret,it,give,me,just,a,chance,please,take,half,a,second,of,your,life,and,thumb,this,comment,up,it,will,maybe,change,my,life,for,real,thank,you,wafence,do,you,want,to,make,some,easy,money,check,out,my,page,tvcmcadavidweebly,,com,dont,miss,out,on,this,opportunity,give,thumbs,up,i,would,appreciate,it,please,check,out,this,video,called,quotwe,love,mind,master,itquot,thank,u,,ways,to,make,money,50k,per,month,search,google,now,gtgt,9nlmemakemoneywithoutinvestment1,meet,the,richest,online,marketer,now,click,,bitlymakemoneywithoutadroid,how,did,you,know,that,people,makes,another,account,just,for,subscribing,itself,and,liking,,i39ll,subscribe,to,you,you,look,nice,,hey,dickwad,,we39re,all,africans,the,colour,of,your,skin,just,tells,us,something,about,how,long,ago,your,ancestors,left,afrika,check,out,baba,brinkman,,he,has,a,song,called,quototi39m,a,africanquot,go,learn,something,wow,muslims,are,really,egoistic,23,of,the,world,population,and,not,in,this,video,or,donating,1,dollar,to,the,poor,ones,in,afrika,,shame,on,those,terrorist,muslims,wats,good,go,check,out,my,music,on,da,channel,ik,yall,got,me,i,know,y39all,can,check,out,amy,music,because,yall,got,me,go,to,da,channel,northland,paranormal,society,is,now,on,youtube,check,out,our,channel,of,real,paranormal,evidence,facebook,northland,paranormal,society,hello,everyone,,i,know,most,of,you,probably,pass,up,these,kind,of,comments,but,for,those,who,are,still,reading,this,thanks,i,don't,have,any,money,for,advertisements,no,chance,of,getting,heard,nothing,i,live,in,such,a,small,town,if,this,comes,off,as,spam,sorry,i'm,an,instrumental,songwriter,from,columbus,mississippi,please,go,to,my,channel,and,check,out,my,original,music,it,would,be,highly,appreciated,if,you,thumbs,up,this,comment,so,my,music,can,be,heard,thank,you,adam,whitney,so,then,how,are,you,going,to,call,yourself,a,instrumental,songwriter,if,theres,no,singing,theres,no,song,to,write,lol,you,got,alot,to,learn,kid,but,hey,don39t,forget,to,subscribe,hello,everyone,,i,know,most,of,you,probably,pass,up,these,kind,of,comments,but,for,those,who,are,still,reading,this,thanks,i,don't,have,any,money,for,advertisements,no,chance,of,getting,heard,nothing,i,live,in,such,a,small,town,if,this,comes,off,as,spam,sorry,i'm,an,instrumental,songwriter,from,columbus,mississippi,please,go,to,my,channel,and,check,out,my,original,music,it,would,be,highly,appreciated,if,you,thumbs,up,this,comment,so,my,music,can,be,heard,thank,you,adam,whitney,please,go,in,and,see,our,channel,and,subscribe,,it,would,be,nice,mate,wanna,earn,money,online,without,investmentjust,visit,this,link,thergloveblogspotin201308blogpost10html,hello,everyone,,i,know,most,of,you,probably,pass,up,the,se,kind,of,comments,but,for,those,who,are,still,reading,this,thanks,i,don't,have,any,money,for,advertisements,no,chance,of,getting,heard,nothing,i,live,in,such,a,small,town,if,this,comes,off,as,spam,sorry,i'm,an,instrumental,songwriter,from,columbus,mississippi,please,go,to,my,channel,and,check,out,my,original,music,it,would,be,highly,appreciated,if,you,thumbs,up,this,comment,so,my,music,can,be,heard,thank,you,adam,whitney,check,out,the,dubstep,version,i,can,to,make,money,wanna,laugh,please,subscribe,to,our,channel,earn,money,for,b

eing,online,with,0,efforts,bitly14gkvdo,could,you,please,check,out,my,covers,
on,my,channel,i,do,covers,like,adele,kodaline,imagine,dragonsand,more,please,
if,you,could,spare,a,few,minutes,could,you,have,a,listen,to,one,or,two,of,my,
covers,,feel,free,to,comment,and,subscribe,,thank,you,adfly,kld3y,check,out,t
hese,irish,guys,cover,of,avicii39s,wake,me,up,just,search,quotwake,me,up,fidd
le,me,sillyquot,worth,a,listen,for,the,gorgeous,fiddle,player,hello,everyon
e,,i,know,most,of,you,probably,pass,up,these,kind,of,comments,but,for,those,w
ho,are,still,reading,this,thanks,i,don't,have,any,money,for,advertisements,n
o,chance,of,getting,heard,nothing,if,this,comes,off,as,spam,sorry,i,am,a,vide
o,animator,just,trying,to,make,it,up,into,the,video,animation,industry,pleas
e,give,me,the,chance,to,prove,myself,to,you,please,visit,my,channel,subscrib
e,if,you,like,and,thumb,this,comment,up,so,everyone,can,see,thank,you,hello,e
veryone,,i,know,most,of,you,probably,pass,up,these,kind,of,comments,but,for,t
hose,who,are,still,reading,this,thanks,i,don't,have,any,money,for,advertiseme
nts,no,chance,of,getting,heard,nothing,if,this,comes,off,as,spam,sorry,i,am,
a,video,animator,just,trying,to,make,it,up,into,the,video,animation,industry,
please,give,me,the,chance,to,prove,myself,to,you,please,visit,my,channel,subs
cribe,if,you,like,and,thumb,this,comment,up,so,everyone,can,see,thank,you,he
y,youtubers,i,really,appreciate,all,of,you,who,took,the,time,to,read,this,i,a
m,just,a,19,year,old,boy,who,wants,to,be,a,successful,musician,in,the,music,w
orld,i,dont,have,any,money,to,advertise,my,channel,if,you,could,just,visit,m
y,channel,comment,on,my,video,or,subscribe,that,would,be,great,it,will,only,b
e,few,seconds,of,your,life,thank,u,to,all,the,people,who,just,gave,me,a,chanc
e,l,really,appreciate,it,subscribe,to,my,pagee,please,please,subscribe,to,my,
page,thanks,hey,hit,this,shit,up,while,yall,can,they,killed,the,versace,remi
x,just,type,in,cge,quotversace,freestyle,get,moneyquot,shot,by,jawan,gardner,
and,help,them,reach,1,million,views,hey,guys,and,girls,check,out,comedy,recip
e,for,hilarious,you,tube,videos,pranks,and,crank,calls,subscribe,me,i,will,su
bscribe,you,back,subscribe,me,i,will,subscribe,you,back,please,check,my,chann
el,out,i,subscribe,back,i,really,ask,nicely,to,view,my,vids,i,subscribe,bac
k,,facebook,is,lame,and,so,2004,check,out,,swagfriends,com,make,thousands,of,
cool,new,friends,everyday,join,this,new,movement,hi,check,out,and,share,our,s
ongs,check,out,partyman318,fr,good,tunez,d,hey,youtubers,i,really,appreciate,
all,of,you,who,took,the,time,to,read,this,i,am,just,a,19,year,old,boy,who,wan
ts,to,be,a,successful,musician,in,the,music,world,i,dont,have,any,money,to,ad
vertise,my,channel,if,you,could,just,visit,my,channel,comment,on,my,video,or,
subscribe,that,would,be,great,it,will,only,be,few,seconds,of,your,life,thank,
u,to,all,the,people,who,just,gave,me,a,chance,l,really,appreciate,it,hey,musi
c,fans,i,really,appreciate,any,of,you,who,will,take,the,time,to,read,this,an
d,check,my,music,out,i39m,just,a,15,year,old,boy,dreaming,of,being,a,successf
ul,musician,in,the,music,world,i,do,lots,of,covers,and,piano,covers,but,i,don
39t,have,money,to,advertise,a,simple,thumbs,up,to,my,comment,a,comment,on,my,
videos,or,a,subscription,would,be,a,step,forward,it,will,only,be,a,few,second
s,of,your,life,that,you,won39t,regret,thank,u,to,all,the,people,who,just,giv
e,me,a,chance,,hey,music,fans,i,really,appreciate,any,of,you,who,will,take,th
e,time,to,read,this,and,check,my,music,out,i39m,just,a,15,year,old,boy,dreami
ng,of,being,a,successful,musician,in,the,music,world,i,do,lots,of,covers,and,
piano,covers,but,i,don39t,have,money,to,advertise,a,simple,thumbs,up,to,my,co
mment,a,comment,on,my,videos,or,a,subscription,would,be,a,step,forward,it,wil
l,only,be,a,few,seconds,of,your,life,that,you,won39t,regret,thank,u,to,all,th
e,people,who,just,give,me,a,chance,,hey,music,fans,i,really,appreciate,any,o
f,you,who,will,take,the,time,to,read,this,and,check,my,music,out,i39m,just,a,
15,year,old,boy,dreaming,of,being,a,successful,musician,in,the,music,world,i,
do,lots,of,covers,and,piano,covers,but,i,don39t,have,money,to,advertise,a,sim
ple,thumbs,up,to,my,comment,a,comment,on,my,videos,or,a,subscription,would,b
e,a,step,forward,it,will,only,be,a,few,seconds,of,your,life,that,you,won39t,r
egret,thank,u,to,all,the,people,who,just,give,me,a,chance,,hi,check,out,and,s

hare, our, songs, hi, check, out, and, share, our, songs, hi check, out, and, share, our, songs, hey, music, fans, i, really, appreciate, any, of, you, who, will, take, the, time, to, read, this, and, check, my, music, out, i39m, just, a, 15, year, old, boy, dreaming, of, being, a, successful, musician, in, the, music, world, i, do, lots, of, covers, and, piano, covers, but, i, don39t, have, money, to, advertise, a, simple, thumbs, up, to, my, comment, a, comment, on, my, videos, or, a, subscription, would, be, a, step, forward, it, will, only, be, a, few, seconds, of, your, life, that, you, won39t, regret, thank, u, to, all, the, people, who, just, give, me, a, chance, , hey, i, am, doing, the, forty, hour, famine, so, i39ll, be, giving, up, on, food, and, social, working, for, 40, hours, i39m, doing, this, to, raise, money, for, african, people, who, can39t, experience, the, luxuries, that, we, can, so, can, you, donate, to, give, them, a, chance, any, amount, would, do, , click, on, the, link, and, donate, h, t, t, p, , , , 4, 0, h, f, , c, o, m, , a, u, , a, n, t, h, o, n, y, l, a, m, thanks, , love, itt, and, ppl, check, out, my, channel, subscribe, my, channel, adfly, , kld3y, adfly, , kld3y, check, out, my, new, video, hey, music, fans, i, really, appreciate, all, of, you, who, take, time, to, read, this, and, check, my, music, out, i39m, just, a, 15, year, old, boy, dreaming, of, being, a, successful, musician, in, the, music, world, i, do, lots, of, covers, and, piano, covers, but, i, don't, have, money, to, advertise, a, simple, thumbs, up, to, my, comment, a, comment, on, my, videos, or, a, subscription, would, be, a, step, forward, it, will, only, be, a, few, seconds, of, your, life, that, u, won39t, regret, thank, u, to, all, the, people, who, just, give, me, a, chance, it, means, a, lot, , hello, everyone, it, is, not, my, intention, to, spam, and, am, truly, sorry, if, anyone, is, annoyed, by, this, but, just, please, hear, me, out, i, am, a, rapper, singer, music, producer, and, a, song, writer, and, have, been, making, music, for, a, while, now, eight, years, to, be, exact, i, enjoy, making, all, types, of, music, that, anyone, can, listen, to, doesn39t, have, me, talking, about, killing, anyone, and, most, importantly, focuses, on, the, quality, of, the, music, please, check, out, my, page, it39s, only, a, click, away, thank, you, and, have, a, nice, day, , facebook, is, lame, and, so, 2004, check, out, , swagfriends, com, make, thousands, of, cool, new, friends, everyday, join, this, new, movement, please, check, out, and, send, to, others, freedom, and, justice, are, on, the, line, please, google, steven, l, reed, case, lands, in, supreme, court thanks, nice, to, meet, you, , this, is, johnny, 1, if, you, already, know, my, music, , thumb, this, up, because, you, found, it, this, way, too, 2, if, you, want, to, hear, original, songs, completely, made, by, 1, person, continue, reading, i, sing, write, original, music lyrics, amp, play, guitar, bass, drums, amp, keyboards, i39m, a, 1manband, my, music, is, completely, independent, amp, 100, listeners supported, if, you, want, to, hear, it, amp, if, you39re, willing, to, help, record, new, album, , click, on, my, name picture, enjoy, the, music, , wish, you, awesome, day, hey, you, , check, out, the, channel, of, alvar, lake, , hi, this, is, johnny, 1, if, you, already, know, my, music, , thumb, this, up, because, you, found, it, this, way, too, 2, if, you, want, to, hear, original, songs, completely, made, by, 1, person, continue, reading, i, sing, write, original, music lyrics, amp, play, guitar, bass, drums, amp, keyboards, i39m, a, 1manband, my, music, is, completely, independent, amp, 100, listeners supported, if, you, want, to, hear, it, amp, if, you39re, willing, to, help, record, new, album, , click, on, my, name picture, thank, you, amp, enjoy, the, music, , wish, you, awesome, day, check, out, this, video, on, youtube br, quot this, time, for, africa quot one, for, trayvon, and, our, motherland br, thank you, shakira, o, people, s, of, the, earth, i, have, seen, how, you, perform, every, form, of, evil, at, your, leisure, you, cease, not, from, reveling, in, that, which, i, hate, behold, you, murder, the, innocent, day, and, night, and, plot, evil, against, your, neighbor, you, stand, up, for, the, rights, of, those, who, commit, abomination, and, clap, your, hands, as, wickedness, is, celebrated, openly, in, the, streets, o, most, perverse, and, abominable, generation, shall, i, not, repay, hear, the, word, of, the, lord, , trumpet call of god online, com, i, will, never, forget, this, song, in, my, life, like, this, comment, of, your, hearing, this, song, for, like, a, year, omg, facebook, is, old, check, out, gt, swagfriends, com, make, thousands, of, cool, new, friends, everyday, join, the, movement, hey, music, fans, i, really, appreciate, all, of, you, who, take, time, to, read, this, and, check, my, music, out, i39m, just, a, 15, year, old, boy, dreaming, of, being, a, successful, musician, in, the, music, world, i, do, lots, of, covers, and, piano, covers, but, i, don't, have, money, to, advertise, a, simple, thumbs, up, to, my, comment, a, comment, on, my, videos, or, a, subscription, would,


```
In [26]: #text cleaning on non-spam comments
nonspam_content = nonspam_comments['CONTENT'] #pulling non-spam comments from
dataframe
nonspam_content = " ".join(comment for comment in nonspam_content) #convert to
a string object
nonspam_content = nonspam_content.split() #word tokenization
nonspam_content = [nsc.translate(table) for nsc in nonspam_content] #removing
punctuation
nonspam_content = [nsc.lower() for nsc in nonspam_content] #converting all tex
t to lowercase
nonspam_content = ', '.join(nonspam_content) #returns to string object
print(deEmojify(nonspam_content)) #removing emojis
nonspam_content[:100]
```

nice, song, i, love, song, , i, love, song, , 860000000, lets, make, it, first, female, to, re
ach, one, billion, share, it, and, replay, it, , shakira, is, best, for, worldcup, the, bes
t, world, cup, song, ever, i, love, awesome, , i, like, shakira, shakira, , waka, waka, br, lo
ve, this, song, why, so, many, disliked, i, don39t, think, this, song, will, ever, get, ol
d, , love, song, wery, good, every, time, i, hear, this, song, i, think, about, iniesta39s, g
oal, against, the, netherlands, whose, watching, this, in, 2015, if, so, hi5, i, love, thi
s, song, so, much, lt3br, keep, em39, coming, i, love, this, song, thumbs, up, to, you, waka,
best, one, 5, years, later, i, still, love, this, song, br, axy665, like, her, voice, sound
s, weird, and, plus, she39s, cute, for, a, blonde, why, there, are, so, many, dislikes, thi
s, song, is, so, awesome, it, sounds, like, we, must, stop, be, racists, if, i, could, i, woul
d, like, it, 1000000000, times, nice, love, itttttttt, wurkkkkkkkkkkkkkkkkkkkkkkkkkkkk
kk, st
ill, watching, in, 2015, shakira, is, my, favourite, singer, wooooo, i, remember, this, so
ng, like, whose, who, are, watching, this, in, 2015, like, best, song, in, world, beautifu
l, good, song, she39s, such, an, awesome, entertainer, and, pretty, too, lt3, shakira, lov
e, this, song, love, this, song, my, soccer, team, made, a, cd, for, our, couch, with, this, s
ong, on, it, fave, song, watching, in, 2015, i, felt, old, when, i, realized, that, this, son
g, was, 5, years, old, love, shakira, one, of, the, best, song, of, all, the, time, i, really,
am, madly, in, love, with, this, woman, i, totally, love, this, song, absolutely, fantasti
c, i, love, ya, shakira, im, still, watching, in, 2015, the, song, is, sad, goood, the, best,
fifa, world, cup, song, for, sure, nice, video, shakira, good, waka, waka, she, is, beautif
ul, but, it, is, not, american, ilovethissong, lamest, world, cup, song, ever, this, time,
for, africa, you, mean, in, africa, it, wasn39t, a, live, aid, event, or, something, she, ma
de, it, seem, like, a, charity, case, for, them, instead, of, a, proud, moment, where, was, r
icky, martin, when, you, needed, him, smh, my, friend, sam, loves, this, song, i, love, you,
to, help, shakira, become, the, first, female, to, hit, 1billiom, views, i39ve, decided, t
o, watch, this, video, at, least, one, a, day, everyday, shakifans, we, r, so, close, just, h
it, the, replay, button, , ilove, shakira, , it, was, cool, the, best, song, ever, , omg, list
en, to, this, its, soo, good, d, nice, song, waka, waka, super, music, how, could, 108k, peop
le, dislike, this, song, or, video, love, it, i, swear, shakira, keeps, getting, more, and,
more, gorgeous, she, definitely, looks, more, gorgeous, with, her, hair, this, way, than,
super, curly, youtube, comments, be, likebr, this, is, so, 5, years, ago, , it, pisses, me, o
ff, a, bit, that, blank, space, has, more, views, 868, mio, than, this, meh, , your, the, best
♥♥, i, love, you, p, i, love, shakira, , ♥, this, song, is, special, because, is, a, song, for,
africa, and, i, am, an, african, , lt3, this, song, so, muchshakira, your, a, really, good, a
rtist, i, love, shakirabr, ♥♥♥♥♥, i, remember, that, torunament, like, it, was, today, b
eautiful, like, shakira, you, are, so, beautiful, you, are, lovely, lively, we, love, you,
great, best, world, cup, song, this, song, is, so, beauty, i, heard, this, when, i, was, onl
y, 6, years, old, and, i, still, love, it, nice, song, waka, waka, oh, my, god, go, to, 1, billi
on, of, replay, i, love, shakira, i, love, her, i, love, this, song, ampi, love, shakiralt33
3♥♥♥♥♥♥♥♥♥♥ 《33☆☆, i, likedbr, , i, like, this, song, , i, like, it, , waka, waka, , am
azing, song, cool, song, , i, love, u, shakira, waka, waka, , this, song, sucks, lol, i, love,
this, song, shakira, wow5, yearsbr, , you, best, singer, 5, years, soon, like, wanderfol, i
s, love, or, miusic, waka, waka, best, song, ever, x3333333333, very, pleasant, to, hear, h
aha, good, love, wow, waka, waka, eh, eh, cutie, girl, and, beautiful, song, love, you, shki
ra, i, believe, that, soccer, promotes, terrorism, bad, choice, shakira, i, loved, this, s
ong, when, i, was, in, my, teenage, years, lip, synch, is, terrible, shakira, is, differen
t, , she, is, so, happy, all, the, time, and, she, is, spending, for, africa, br, she, can, dan
ce, , sing, and, she, speaks, 4, languages, best, world, cup, offical, song, 5, years, ago, d
amn, , very, nice, like, me, shakira, song, waka, waka, so, underrated, betterbr, than, kat
y, perry, , but, in, not, saying, katy, is, bad, but, she, has, no, meaning, to, her, songs, sh
akira, hasbr, meaningful, songs, these, are, the, songs, which, bring, memories, and, vib
e39s, i, will, miss, it, br, , , i, love, this, song, and, expect, the, world, cup, , good, time
s, , it, makes, me, happy, instantly, and, makes, me, forgot, everything, bad, happening, l
ove, this, song, so, much, one, of, my, faves, xxx, hi, nice, song, shakira, sorry, for, bad,
brazilian, best, song, go, shaki, d, this, is, shit, and, sooooo, autotuned, i, stil, lisen
ing, this, , sexy, shakira, i, love, you, shakiria, me, and, my, aunt, love, this, song, so, b

autiful,so,beautiful,she39s,so,pretty,i,like,she,is,perfect,your,a,fucking,bit
 ch,this,song,is,racist,where,are,shakifans,like,i,want,new,song,shakira,is,pe
 rfect,shakira,,i,remember,this,d,waka,waka,this,makes,me,miss,the,world,cup,b
 est,song,eva,i,love,it,her,voice,is,so,wow,shakira,,,wow,i,love,you,i,like,an
 d,love,so,much,people,all,friends,i,love,shakira,,,shakira,i,love,you,beautif
 ul,shakira,is,very,beautiful,she,is,good,god,she,is,so,hot,she,is,perfect,lt
 3,wow,shakira,voice,sound,spanish,but,that,is,what,make,the,music,sound,amazi
 ng,♥i,love,it,she,is,perfect,beautiful,song,beautiful,girl,it,works,shakira,,
 omg,love,this,ouf,ouf,ouffffffffffffffffffffff,,awesome,cool,this,song,will,neve
 r,get,old,i,always,have,goose,bumps,at,that,part,love,it,i,love,this,song,i,w
 ant,to,see,shakira,not,football,,good,nice,song,,god,she,is,so,sexy,drives,m
 e,crazy,wow,shakira,is,the,best,nice,shakira,shakira,,the,best,wow,love,this,
 song,love,this,song,,i,watched,this,because,of,the,large,amount,of,views,and,
 now,i,am,convinced,it,is,because,that,girl,is,so,hot,this,song,always,gives,m
 e,chills,,i,love,dis,song,3,waka,waka,she,rules,she,is,sooooo,beautiful,well,
 done,shakira,i,love,this,song,because,we,sing,it,at,camp,all,the,time,i,love,
 this,song,for,two,reasons,1it,is,about,africa,2i,was,born,in,beautiful,south,
 africa,wow,shakira,u,are,so,wiredo,shakira,is,the,best,dancer

Out[26]: 'nice,song\ufeff,i,love,song,\ufeff,i,love,song,\ufeff,860000000,lets,make,i
 t,first,female,to,reach,one,billion,shar'

```
In [50]: stopwords = set(STOPWORDS)
wordcloud_nospam = WordCloud(max_words=20,
                             background_color='white',
                             stopwords = stopwords).generate(nospam_content)
plt.imshow(wordcloud_nospam)
plt.axis("off")
plt.show()
```



For non-spam comments, the words/phrases used are related to the video like the name of the artist, the name of the song or the users personal thoughts on the song/video. The most frequently used words include: Shakira, love, song and beautiful.

Author vs Class

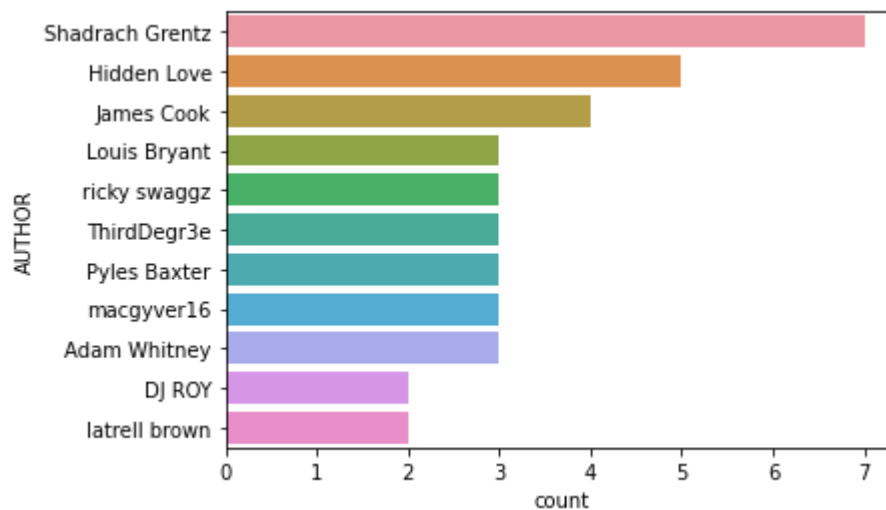
Spam Comments

```
In [94]: spam_comments['AUTHOR'].value_counts()
```

```
Out[94]: Shadrach Grentz      7
Hidden Love      5
James Cook      4
Louis Bryant     3
ricky swaggz     3
..
nosa bella      1
Free Video Tv   1
dili bair       1
Charanjeet Singh 1
Scott Johnson   1
Name: AUTHOR, Length: 135, dtype: int64
```

```
In [97]: sns.countplot(y='AUTHOR', data=spam_comments,
                        order=spam_comments.AUTHOR.value_counts().iloc[:11].index)
```

```
Out[97]: <AxesSubplot:xlabel='count', ylabel='AUTHOR'>
```



```
In [62]: spam_comments.loc[(spam_comments['AUTHOR']=='Shadrach Grentz')]
```

Out[62]:

	COMMENT_ID	AUTHOR	DATE	CONTENT	CLA
250	_2viQ_Qnc68dceJbTRNTP2sksMxa_lm35LaCu_jPluY	Shadrach Grentz	2013-09-29 13:59:42.162	Hey Music Fans I really appreciate any of you ...	
328	_2viQ_Qnc69zyetF6GsHRzYGyXI4u5kg0Sm-nP-pupl	Shadrach Grentz	2013-08-02 03:15:46.914	Hey Music Fans I really appreciate any of you ...	
329	_2viQ_Qnc699u36gNm3NRoq1qulaJWRNrftGEEZM3J4	Shadrach Grentz	2013-08-01 21:43:52.122	Hey Music Fans I really appreciate any of you ...	
330	_2viQ_Qnc68LqKgb40V2lmKQYcp1ZqpHrCCMFqMCcA4	Shadrach Grentz	2013-08-01 20:39:15.325	Hey Music Fans I really appreciate any of you ...	
334	_2viQ_Qnc69Nq0Ytk1jCpzWPCrpGEk6T7cdVAxfSIak	Shadrach Grentz	2013-07-29 17:39:24.876	Hey Music Fans I really appreciate any of you ...	
341	_2viQ_Qnc6_YN7xFNAg14zX99Y614Salf57yOcrBRsw	Shadrach Grentz	2013-07-21 12:21:37.898	Hey Music Fans I really appreciate all of you ...	
358	_2viQ_Qnc6-jk58CPwBnqfbM6oByJH5oPvCtKecLQyo	Shadrach Grentz	2013-07-14 03:11:20.243	Hey Music Fans I really appreciate all of you ...	



```
In [63]: spam_comments.loc[(spam_comments['AUTHOR']=='Hidden Love')]
```

Out[63]:

	COMMENT_ID	AUTHOR	DATE	CONTENT	CLAS
253	_2viQ_Qnc6_onwOgxju-DV6WkqHZE0ztCXD04EgEFBU	Hidden Love	2013-09-24 14:29:54.072	Hi. Check out and share our songs.	
325	_2viQ_Qnc69r15LuL8TDbisenTJ_hf5RfcyJAyoMC5eo	Hidden Love	2013-08-06 11:40:05.581	Hi. Check out and share our songs.	
331	_2viQ_Qnc68Qq98m0mmx4rlprYiD6aYgMb2x3bdupEM	Hidden Love	2013-08-01 09:19:56.654	Hi. Check out and share our songs.	
332	_2viQ_Qnc698fLCQzZmRt1CqPhf0L6mchwWG5gsUlzk	Hidden Love	2013-07-31 10:22:02.628	Hi. Check out and share our songs.	
333	_2viQ_Qnc68sxLpsBQ1cPk_LLH91SMsRav51KmhloQw	Hidden Love	2013-07-30 11:22:57.091	Hi. Check out and share our songs.	

```
In [64]: spam_comments.loc[(spam_comments['AUTHOR']=='James Cook')]
```

Out[64]:

	COMMENT_ID	AUTHOR	DATE	CONTENT	CLAS
186	_2viQ_Qnc69ShtSmsaBOGFHrTAVCKLAtpOr40oiL5Yg	James Cook	2013-10-20 13:31:10.083	Stop Wasting Up Your Time and Get Paid To Mes...	
192	_2viQ_Qnc68eqIzELH00rh9umGIUKSRuQvWZAXhr_qM	James Cook	2013-10-15 17:00:04.573	Stop Wasting Up Your Time and Get Paid To Mes...	
194	_2viQ_Qnc6-grcnVFTtnhvC9tpgVG33p5a0AZHKILI	James Cook	2013-10-14 10:35:03.861	Hello Guys...I Found a Way to Make Money Onlin...	
200	_2viQ_Qnc69BVRuBYCJ6ik7gdMwTiBRS9DDnTjNs-Gg	James Cook	2013-10-10 18:08:07.815	Hello Guys...I Found a Way to Make Money Onlin...	

The top 3 of the spam comment authors post comments directing users to check out their own music or about making money online. They each also post the exact same comment over and over again at different times and on different days.

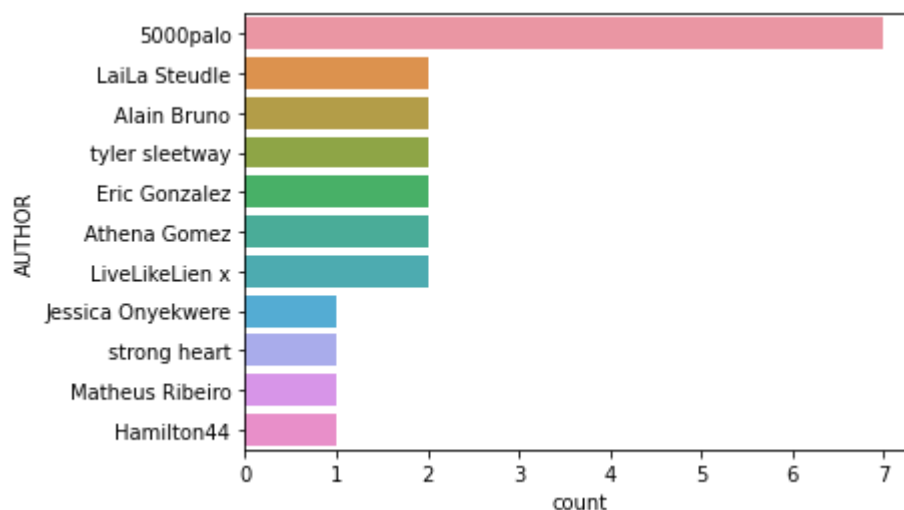
Non-Spam Comments

```
In [99]: nonspam_comments['AUTHOR'].value_counts()
```

```
Out[99]: 5000palo          7
         LaiLa Steudle     2
         Eric Gonzalez     2
         LiveLikeLien x    2
         Alain Bruno       2
         ..
         Sabina Pearson-Smith 1
         Aishlin Maciel     1
         Nimesh Nambiar     1
         Mohammed Yousef    1
         ranferi delgado    1
         Name: AUTHOR, Length: 184, dtype: int64
```

```
In [65]: sns.countplot(y='AUTHOR', data=nonspam_comments,
                        order=nonspam_comments.AUTHOR.value_counts().iloc[:11].index)
```

```
Out[65]: <AxesSubplot:xlabel='count', ylabel='AUTHOR'>
```



```
In [70]: nonspam_comments.loc[(nonspam_comments['AUTHOR']=='5000palo')]['CONTENT'][:50]
```

```
Out[70]: 215    She is perfect
         221    I want new song
         223        Shakira :-*
         234        Shakira :-*
         261    She is perfect
         264        Shakira :-*
         286        Shakira :-*
         Name: CONTENT, dtype: object
```

```
In [73]: nonspam_comments.loc[(nonspam_comments['AUTHOR']=='LaiLa Steudle)]['CONTENT'][:50]
```

```
Out[73]: 145          I love you Shakiria!!!!!!
146    Me and my aunt love this song!!!!!!
Name: CONTENT, dtype: object
```

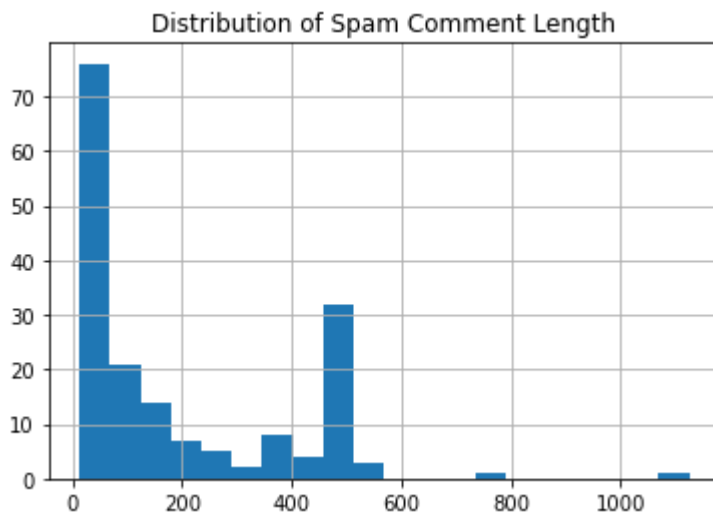
One the other hand the non-spam users post comments related to the song or video and about their love for the artist. Also, leaving more than 2 comments is considered an outlier.

Comment Length vs Class

Spam Comments

```
In [11]: spam_comments['CONTENT_LENGTH'].hist(bins=20)
plt.title('Distribution of Spam Comment Length')
```

```
Out[11]: Text(0.5, 1.0, 'Distribution of Spam Comment Length')
```



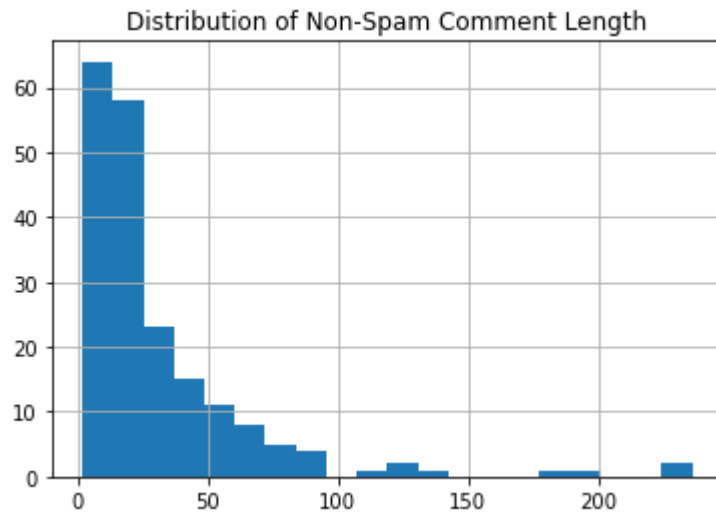
```
In [12]: print("Average Comment Length:", spam_comments['CONTENT_LENGTH'].mean())
print("Minimum Comment Length:", spam_comments['CONTENT_LENGTH'].min())
print("Maximum Comment Length:", spam_comments['CONTENT_LENGTH'].max())
```

```
Average Comment Length: 196.33908045977012
Minimum Comment Length: 13
Maximum Comment Length: 1125
```

Non-Spam Comments


```
In [13]: nonspam_comments['CONTENT_LENGTH'].hist(bins=20)
plt.title('Distribution of Non-Spam Comment Length')
```

Out[13]: Text(0.5, 1.0, 'Distribution of Non-Spam Comment Length')



```
In [14]: print("Average Comment Length:",nonspam_comments['CONTENT_LENGTH'].mean())
print("Minimum Comment Length:",nonspam_comments['CONTENT_LENGTH'].min())
print("Maximum Comment Length:",nonspam_comments['CONTENT_LENGTH'].max())
```

Average Comment Length: 31.321428571428573
Minimum Comment Length: 2
Maximum Comment Length: 236

Most users leave short comments, however based on the average, minimum and maximum comment length along with the second peak that appears around 450-500 words, spam comments do seem to be longer than non-spam.