

ADS 509 Week 6 Sentiment Assignment - Minsu Kim

June 19, 2023

1 ADS 509 Sentiment Assignment

This notebook holds the Sentiment Assignment for Module 6 in ADS 509, Applied Text Mining. Work through this notebook, writing code and answering questions where required.

In a previous assignment you put together Twitter data and lyrics data on two artists. In this assignment we apply sentiment analysis to those data sets. If, for some reason, you did not complete that previous assignment, data to use for this assignment can be found in the assignment materials section of Blackboard.

1.1 General Assignment Instructions

These instructions are included in every assignment, to remind you of the coding standards for the class. Feel free to delete this cell after reading it.

One sign of mature code is conforming to a style guide. We recommend the [Google Python Style Guide](#). If you use a different style guide, please include a cell with a link.

Your code should be relatively easy-to-read, sensibly commented, and clean. Writing code is a messy process, so please be sure to edit your final submission. Remove any cells that are not needed or parts of cells that contain unnecessary code. Remove inessential `import` statements and make sure that all such statements are moved into the designated cell.

Make use of non-code cells for written commentary. These cells should be grammatical and clearly written. In some of these cells you will have questions to answer. The questions will be marked by a “Q:” and will have a corresponding “A:” spot for you. *Make sure to answer every question marked with a Q: for full credit.*

```
[1]: import os
import re
import emoji
import pandas as pd
import numpy as np

from collections import Counter, defaultdict
from string import punctuation

from nltk.corpus import stopwords

sw = stopwords.words("english")
```

```
[29]: # Add any additional import statements you need here
from matplotlib import pyplot as plt
import seaborn as sns
```

```
[2]: # change `data_location` to the location of the folder on your machine.
# These subfolders should still work if you correctly stored the
# data from the Module 1 assignment
data_location = "C:/Users/suemi/Downloads/M1 Results/"
twitter_folder = "twitter/"
lyrics_folder = "lyrics/"

positive_words_file = "positive-words.txt"
negative_words_file = "negative-words.txt"
tidy_text_file = "tidytext_sentiments.txt"
```

```
[3]: artist_files = {'cher': 'cher_followers_data.txt',
                    'robyn': 'robynkonihiwa_followers_data.txt'}
```

1.2 Data Input

Now read in each of the corpora. For the lyrics data, it may be convenient to store the entire contents of the file to make it easier to inspect the titles individually, as you'll do in the last part of the assignment. In the solution, I stored the lyrics data in a dictionary with two dimensions of keys: artist and song. The value was the file contents. A Pandas data frame would work equally well.

For the Twitter data, we only need the description field for this assignment. Feel free all the descriptions read it into a data structure. In the solution, I stored the descriptions as a dictionary of lists, with the key being the artist.

```
[4]: artists = []
songs = []
lyrics = []

for item in os.listdir(data_location + lyrics_folder) :
    if os.path.isdir(data_location + lyrics_folder + item) :
        for lyric_page in os.listdir(data_location + lyrics_folder + item) :
            artist,song = lyric_page.split("_")

            song = song.replace(".txt","")

            artists.append(artist)
            songs.append(song)

            with open(data_location + lyrics_folder + item + "/" + lyric_page) as infile :
                next(infile) # skip title
                next(infile) # skip blank
```

```

next(infile) # skip blank
next(infile) # skip final blank

lyrics.append(infile.read())

```

```

lyrics_data = pd.DataFrame()
lyrics_data['artist'] = artists
lyrics_data['song'] = songs
lyrics_data['lyrics'] = lyrics

```

```

[5]: twitter_data = pd.read_csv(data_location + twitter_folder +
    ↳artist_files['cher'],
                                sep="\t",
                                quoting=3)

twitter_data['artist'] = "cher"

```

```

[6]: twitter_data_2 = pd.read_csv(data_location + twitter_folder +
    ↳artist_files['robyn'],
                                sep="\t",
                                quoting=3)

twitter_data_2['artist'] = "robyn"

twitter_data = pd.concat([
    twitter_data, twitter_data_2])

del(twitter_data_2)

```

```

[20]: # this is how I initially read the positive and negative data, into dataframes

positive_words = pd.read_csv('positive-words.txt', comment=';', header=None,
    ↳names=['word'])
negative_words = pd.read_csv('negative-words.txt', comment=';', header=None,
    ↳names=['word'])
tidytext = pd.read_csv('tidytext_sentiments.txt', delimiter='\t')

```

```

[21]: tidytext_df = tidytext
tidytext_df['sentiment'] = tidytext_df['sentiment'].replace({'positive': 1,
    ↳'negative': -1})
tidytext_df = pd.concat([positive_words.assign(sentiment=1), tidytext_df])
tidytext_df = pd.concat([negative_words.assign(sentiment=-1), tidytext_df])
tidytext_df = tidytext_df.drop_duplicates(subset='word')

tidytext_df.head(5)

```

```
[21]:
```

	word	sentiment	lexicon
0	2-faced	-1	NaN
1	2-faces	-1	NaN
2	abnormal	-1	NaN
3	abolish	-1	NaN
4	abominable	-1	NaN

```
[22]: tidytext_df = tidytext_df.drop('lexicon', axis=1)
tidytext_df.head(5)
```

```
[22]:
```

	word	sentiment
0	2-faced	-1
1	2-faces	-1
2	abnormal	-1
3	abolish	-1
4	abominable	-1

```
[ ]: # then it was difficult to proceed with dataframes
# office hours helped with getting the data into a dictionary format
```

```
[9]: # Read in the positive and negative words and the
# tidytext sentiment. Store these so that the positive
# words are associated with a score of +1 and negative words
# are associated with a score of -1. You can use a dataframe or a
# dictionary for this.
```

```
sentiment = dict()

with open(positive_words_file, 'r') as infile :
    for idx, line in enumerate(infile.readlines()) :
        if line[0] == ";" :
            continue

        line = line.strip()

        if line and line.isalpha() :
            sentiment[line.strip()] = 1

with open(negative_words_file, 'r') as infile :
    for idx, line in enumerate(infile.readlines()) :
        if line[0] == ";" :
            continue

        line = line.strip()

        if line and line.isalpha() :
            sentiment[line.strip()] = -1
```

```

with open(tidy_text_file, 'r') as infile :
    next(infile)
    for line in infile.readlines() :
        word, sent, lexicon = line.strip().split("\t")

        if sent == "negative" :
            sentiment[word] = -1
        elif sent == "positive" :
            sentiment[word] = 1
        else :
            print(f"Got {word} with sentiment {sentiment}.")

```

1.3 Sentiment Analysis on Songs

In this section, score the sentiment for all the songs for both artists in your data set. Score the sentiment by manually calculating the sentiment using the combined lexicons provided in this repository.

After you have calculated these sentiments, answer the questions at the end of this section.

```

[10]: def get_sentiment(text,sent=sentiment) :
        text = [w.lower() for w in text.split()]

        sentiment = sum([sent[word] for word in text if word in sent])

        return(sentiment)

```

```

[11]: lyrics_data['sentiment'] = lyrics_data['lyrics'].apply(get_sentiment)

```

```

[12]: average_sentiment = lyrics_data.groupby('artist')['sentiment'].mean()

average_sentiment

```

```

[12]: artist
cher      5.955696
robyn     7.644231
Name: sentiment, dtype: float64

```

```

[13]: first_artist = lyrics_data[lyrics_data['artist'] == 'cher']

sorted_by_sentiment = first_artist.sort_values(by='sentiment')

top_3_songs = sorted_by_sentiment.tail(3)
bottom_3_songs = sorted_by_sentiment.head(3)

```

```

[14]: top_3_songs

```

```
[14]:      artist          song \
198  cher          perfection
108  cher          ifoundyoulove
155  cher  loveandunderstanding

                                lyrics  sentiment
198  Hush little Baby, gotta be strong\n'Cause in t...      43
108  Well I was looking for a new love, a different...      51
155  Here, here in this world\nWhere do we go? Wher...      56
```

```
[15]: bottom_3_songs
```

```
[15]:      artist          song \
16   cher          bangbang
196  cher          outrageous
133  cher  iwalkonguildedsplinters

                                lyrics  sentiment
16   Bang bang you shot me down\nBang bang I hit th...      -70
196  Outrageous, outrageous\n(They say) I'm outrage...      -26
133  Some people think they jive me, but I know the...      -24
```

```
[23]: # print lyrics
for i, row in top_3_songs.iterrows():
    print(row['lyrics'])
    print("-"*50)

print("-"*50)
print("-"*50)
print("-"*50)

for i, row in bottom_3_songs.iterrows():
    print(row['lyrics'])
    print("-"*50)
```

Hush little Baby, gotta be strong
 'Cause in this world we are born to fight
 Be the best, prove them wrong
 A winner's work is never done, reach the top, number one

Oh, perfection
 You drive me crazy with perfection
 I've worn my pride as my protection
 Perfection, ohh

I was taught to be tough
 That the best that you can be ain't enough
 Crack the whip, sacrifice

But I found out paradise had a price

I didn't know it then, but oh I know it now
You gotta work as hard as love to make the love work out
You know this heart of mine has made some big mistakes
I guess that when it comes to love, I just don't know what I takes

All my life I've been driven by perfection
Pushed it to the limit every day and night
Ohh, I've been driven by perfection
But nothing's perfect when love ain't right
Nothing's perfect when the love ain't right

Gave it all, played my part
I gave everything I had but my heart
Worked so hard, made a name
But the loneliness inside stays the same

When love is here today, and then it's gone today
You got a list of lovers lookin' like a resume
You gotta take some time to make it something real
I guess that when it comes to love, I just don't know how to feel

All my life I've been driven by perfection
Pushed it to the limit every day and night
Ohh, I've been driven by perfection
But nothing's perfect when love ain't right
Nothing's perfect when the love ain't right

Oh, perfection
You drive me crazy with perfection
I've worn my pride as my protection
I'm goin' crazy with perfection

Ohh, I didn't know it then, but oh I know it now
You gotta work as hard as love to make the love work out
You know this heart of mine has made some big mistakes
I guess that when it comes to love, I just don't know what I takes

I've been driven by perfection
Pushed it to the limit every day and night
Ohh, I've been driven by perfection
But nothing's perfect when love ain't right

All my life I've been driven by perfection
Pushed it to the limit every day and night
Ohh, I've been driven by perfection
But nothing's perfect when love ain't right

All my life I've been driven by perfection
Pushed it to the limit every day and night
Ohh, I've been driven by perfection
But nothing's perfect when love ain't right

All my life I've been driven by perfection

Well I was looking for a new love, a different kind of true love
Who's gonna treat me right, all day and night
Hey baby I've been looking too
And I have found there's
No other love from me but you
Well I was looking for a new love
A different kind of true love
Who's gonna treat me right
Day and night
Well I found what I was after
Now my life is filled with laughter
I found you love
I was lost with no direction
Then my life was one big question
I was down and out
Filled with doubt
Found what I was after
Now my life is filled with laughter
I found you love
I found you love
I found a new love
He's wonderful and true
He's gonna spent his money
He's gonna call me honey
I gonna tease her
Oh Lord, I gonna squeeze her
Gonna love her plenty
She's gonna make me manly
We're gonna hold love while we can
I've been looking for a new love
A different kind of true love
Who's gonna treat me right
Every day and every night
Now I found what I was after
Now my life is filled with laughter
I found you love
I found you love
I found a new love
He's wonderful and true

He's gonna spent his money
He's gonna call me honey
I gonna please her
Oh Lord, I gonna squeeze her
Gonna love me madly
I'm gonna love her gladly
We're gonna hold love while we can
We're gonna hold love while we can

Here, here in this world
Where do we go? Where can we turn?
When we need some love
It seems that love just can't be found
Where, where do we stand?
When love's supply don't meet love's demand

We got enough stars to light the sky at night
Enough sun to make to make the whole world bright
We got more than enough
But there's one thing there's just not enough of

Not enough love and understanding
We could use some love to ease these troubled times
Not enough love and understanding
Why, oh why?

Spend all of our time
Building buildings up to the sky
Reaching everywhere
But where we need to reach the most
Hearts never can win
Oh, in this race, this race that we're in

We've got enough cars to drive around the world
Enough planes to take us anywhere
We got more than enough
But there's one thing there's just not enough of

Not enough love and understanding
We could use some love to ease these troubled times
Not enough love and understanding
Why, oh why?

Not enough love and understanding
We could use some love to ease these troubled times
Not enough love and understanding
Why, oh why?

We need some understandin'
We need a little more love
Some love and understandin'

Enough stars to light the sky at night
Enough sun to make the whole world bright
Enough hearts to find some love inside
We got more than enough
But there's one thing there's just not enough of

Not enough love and understanding
We could use some love to ease these troubled times
Not enough love and understanding
Why, oh why?

...

Bang bang you shot me down
Bang bang I hit the ground
Bang bang that awful sound
Bang bang my baby shot me down

I was five and you were six
We rode on horses made of sticks
I wore black you wore white
You would always win the fight

Bang bang you shot me down
Bang bang I hit the ground
Bang bang that awful sound
Bang bang my baby shot me down

Seasons came and changed the time
I grew up I called you mine
You would always laugh and say
Remember when we used to play

Bang bang you shot me down
Bang bang and I hit the ground
Bang bang that awful sound
Bang bang my baby shot me down

Music played and people sang
Just for me the church bells rang

After echoes from a gun
We both vowed that we'd be one

Now you're gone I don't know why
Sometimes I cry
You didn't say goodbye
You didn't take the time to lie

Bang bang you shot me down
Bang bang I hit the ground
Bang bang that awful sound
Bang bang my baby shot me down

Bang bang you shot me right between my eyes
Bang bang you can't go paralyzed
Bang bang bang bang bang bang
Bang bang oh baby I'm laying on the ground
Bang bang I'll never come around

Bang bang
Bang bang
Oh baby
Bang bang
Oh baby come and wrap me
Bang bang
You see how sweet it's gonna be

Bang bang
Bang bang
Bang bang
Oh my baby my baby shot me down
Bang bang
I'm up on the ground now

Outrageous, outrageous
(They say) I'm outrageous
It's the rage

I'm gonna wear what I will and spend some
And I will be dress to kill don'tcha know
And when the lights come up
I'm ready I'm ready
To put on a show with class
And if I clash it's cause I want to
What a show and I want everyone to know
They're gonna fly up, get an eyeful
Everything that's craved from me

I'm gonna be, I'm gonna be outrageous

Outrageous

(They say) I'm outrageous

It's the rage it's the rage

With my long black hair hanging way down to my
Ask me no questions and I'll tell you no lies
Don't tell me what to do don't tell me what to be
See I don't trust anybody else's traits about make-up and me

Well in my show I let everything go
Is what you want is whatcha wanna see from me
But when the curtain comes down
And you're on your way back home
I change into my jeans that are split at the seam
I grab my funky black jacket and make quite a racket
You drive like you're an outlaw
Cause everything that's craved from me
I'm gonna be, I'm gonna be outrageous

So outrageous

I'm outrageous honey yes a rage

It's the rage

Outrageous, outrageous

I'm outrageous

It's the rage it's a rage

Outrageous, outrageous

They say I'm outrageous

Some people think they jive me, but I know they must be crazy
They can't see their misfortune, or else they're just too lazy
Je suie le grand zombie
With my yellow chaffien of choisen
Ain't afraid of no tomcat and gonna fill my guts with poison
I walk through the fire
And I'll fly through the smoke
I wanna see my enemies
On the end of my rope
Walk on pins and needles
And I see what they can do
Walk on gilded splinters
With the King of the Zulu

Come to me, get it, come, come
Walk on gilded splinters
Come to me, get it, come, come

Walk on gilded splinters
Till I burn up
Till I burn up
Till I burn up
Till I burn up

When I roll out in my coffin
Drink poison in my chalice
Pride begins to fade
And you all feel my malice
I put gris-gris on your doorstep
So soon you be in the gutter
I'll make your heart melt like butter
I say I can make you stutter

Come to me, get it, come, come
Walk on gilded splinters
Come to me, get it, come, come
Walk on gilded splinters
Till I burn up
Till I burn up
Till I burn up
Till I burn up

Come to me, get it, come, come
Walk on gilded splinters
Come to me, get it, come, come
Walk on gilded splinters
Till I burn up
Till I burn up
Till I burn up
Till I burn up

Come to me, get it, come, come
Walk on gilded splinters
Come to me, get it, come, come
Walk on gilded splinters
Till I burn up
Till I burn up
Till I burn up
Till I burn up

```
[24]: second_artist = lyrics_data[lyrics_data['artist'] == 'robyn']  
  
sorted_by_sentiment2 = second_artist.sort_values(by='sentiment')
```

```
top_3_songs1 = sorted_by_sentiment2.tail(3)
bottom_3_songs1 = sorted_by_sentiment2.head(3)

top_3_songs1
```

```
[24]:      artist          song \
414  robyn      wedancetothebeat
415  robyn  wedancetothebeat114528
380  robyn          loveisfree

                                lyrics  sentiment
414  We dance to the beat\nWe dance to the beat\nWe...      65
415  We dance to the beat\nWe dance to the beat\nWe...      65
380  Free\nLove is free, baby\nFree\nLove is free, ...    116
```

```
[25]: bottom_3_songs1
```

```
[25]:      artist          song \
343  robyn  dontfuckingtellt mewhattodo114520
342  robyn      dontfuckingtellt mewhattodo
334  robyn      criminalintent

                                lyrics  sentiment
343  My drinking is killing me\nMy drinking is kill...    -92
342  My drinking is killing me\nMy drinking is kill...    -92
334  Somebody alert the authorities, I got criminal...    -53
```

1.3.1 Questions

Q: Overall, which artist has the higher average sentiment per song?

A: Robyn

Q: For your first artist, what are the three songs that have the highest and lowest sentiments? Print the lyrics of those songs to the screen. What do you think is driving the sentiment score?

A: The highest and lowest sentiment songs are shown above; obviously, some of the extreme words are seen in the song title and lyrics which maybe driving the scores.

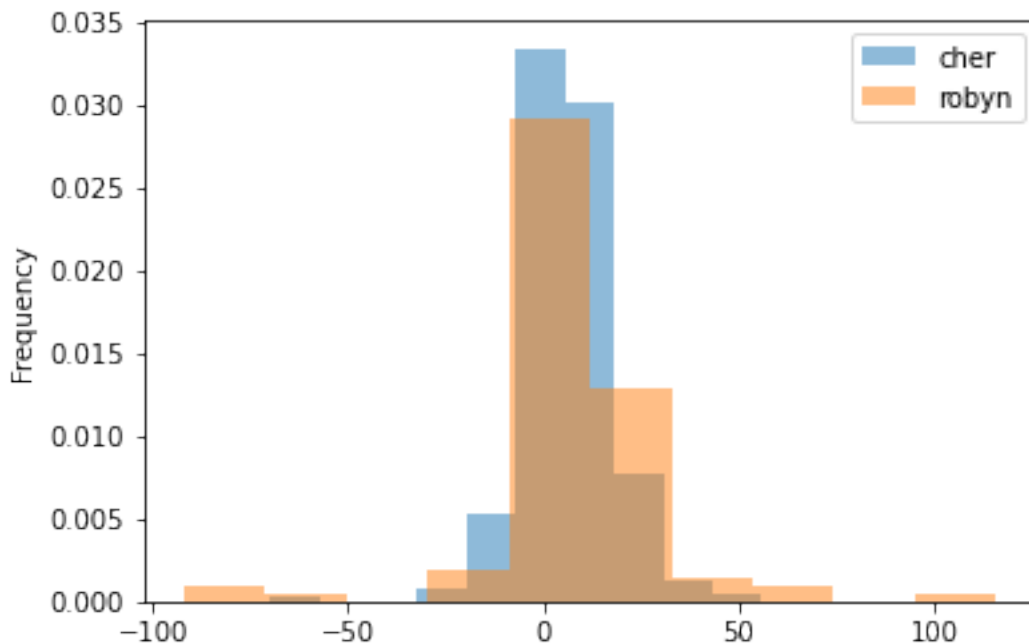
Q: For your second artist, what are the three songs that have the highest and lowest sentiments? Print the lyrics of those songs to the screen. What do you think is driving the sentiment score?

A: The highest and lowest sentiment songs are shown above, and I think the it's for the same reason mentioned in the previous question.

Q: Plot the distributions of the sentiment scores for both artists. You can use `seaborn` to plot densities or plot histograms in `matplotlib`.

```
[34]: lyrics_data.groupby('artist')['sentiment'].
      ↪plot(kind="hist",density=True,alpha=0.5,legend=True)
```

```
[34]: artist
cher      AxesSubplot(0.125,0.125;0.775x0.755)
robyn     AxesSubplot(0.125,0.125;0.775x0.755)
Name: sentiment, dtype: object
```



1.4 Sentiment Analysis on Twitter Descriptions

In this section, define two sets of emojis you designate as positive and negative. Make sure to have at least 10 emojis per set. You can learn about the most popular emojis on Twitter at [the emoji tracker](#).

Associate your positive emojis with a score of +1, negative with -1. Score the average sentiment of your two artists based on the Twitter descriptions of their followers. The average sentiment can just be the total score divided by number of followers. You do not need to calculate sentiment on non-emoji content for this section.

```
[35]: positive_emojis = ['😊','😄','😁','😂','😃','😅','😆','😇','😈','😉','😊','😋','😌','😍','😎','😏','😐','😑','😒','😓','😔','😕','😖','😗','😘','😙','😚','😛','😜','😝','😞','😟','😠','😡','😢','😣','😤','😥','😦','😧','😨','😩','😪','😫','😬','😭','😮','😯','😰','😱','😲','😳','😴','😵','😶','😷','😸','😹','😺','😻','😼','😽','😾','😿','🐼','🦊','🐈','🐓','🐔','🐕','🐶','🐷','🐽','🐾','🐿','🦉','🦋','🦄','🦅','🦆','🦇','🦈','🦉','🦊','🦋','🦌','🦍','🦏','🦐','🦑','🦒','🦓','🦔','🦕','🦖','🦗','🦘','🦙','🦚','🦛','🦜','🦝','🦞','🦟','🦠','🐛','🐜','🐝','🐞','🐟','🐠','🐡','🐢','🐣','🐤','🐥','🐦','🐧','🐨','🐩','🐪','🐫','🐬','🐭','🐮','🐯','🐰','🐱','🐲','🐳','🐴','🐵','🐶','🐷','🐸','🐹','🐺','🐻','🐼','🐽','🐾','🐿']
negative_emojis = ['😞','😟','😠','😡','😢','😣','😤','😥','😦','😧','😨','😩','😪','😫','😬','😭','😮','😯','😰','😱','😲','😳','😴','😵','😶','😷','😸','😹','😺','😻','😼','😽','😾','😿','🐼','🦊','🐈','🐓','🐔','🐕','🐶','🐷','🐽','🐾','🐿']

emoji_sentiment = dict(zip(positive_emojis,[1]*15))
```

```
emoji_sentiment.update(dict(zip(negative_emojis,[-1]*15)))

emoji_sentiment
```

```
[35]: {' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': 1,
      ' ': -1,
      ' ': -1,
      ' ': -1,
      ' ': -1,
      ' ': -1,
      ' ': -1,
      ' ': -1,
      ' ': -1,
      ' ': -1,
      ' ': -1,
      ' ': -1,
      ' ': -1,
      ' ': -1,
      ' ': -1,
      ' ': -1,
      ' ': -1}
```

```
[45]: def get_desc_sentiment(desc,sent=emoji_sentiment) :

      sentiment = sum([sent[ch] for ch in str(desc) if ch in sent])

      return(sentiment)

def extract_emojis(desc,emoji_list) :
    emojis = [ch for ch in str(desc) if ch in emoji_list]

    return(emojis)
```



```
[46]: twitter_data['sentiment'] = twitter_data.description.apply(get_desc_sentiment)
twitter_data['pos_emoji'] = twitter_data.description.
      ↪ apply(extract_emojis, args=(get_desc_sentiment == 1))
twitter_data['neg_emoji'] = twitter_data.description.
      ↪ apply(extract_emojis, args=(get_desc_sentiment == -1))
```

```
-----
TypeError                                Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_38252\2486498128.py in <module>
      1 twitter_data['sentiment'] = twitter_data.description.
      ↪ apply(get_desc_sentiment)
----> 2 twitter_data['pos_emoji'] = twitter_data.description.
      ↪ apply(extract_emojis, args=(get_desc_sentiment == 1))
      3 twitter_data['neg_emoji'] = twitter_data.description.
      ↪ apply(extract_emojis, args=(get_desc_sentiment == -1))

~\miniconda3\lib\site-packages\pandas\core\series.py in apply(self, func,
      ↪ convert_dtype, args, **kwargs)
    4628         dtype: float64
    4629         """
-> 4630         return SeriesApply(self, func, convert_dtype, args, kwargs).
      ↪ apply()
    4631
    4632     def _reduce(

~\miniconda3\lib\site-packages\pandas\core\apply.py in apply(self)
    1023
    1024         # self.f is Callable
-> 1025         return self.apply_standard()
    1026
    1027     def agg(self):

~\miniconda3\lib\site-packages\pandas\core\apply.py in apply_standard(self)
    1074         else:
    1075             values = obj.astype(object)._values
-> 1076             mapped = lib.map_infer(
    1077                 values,
    1078                 f,

~\miniconda3\lib\site-packages\pandas\_libs\lib.pyx in pandas._libs.lib.
      ↪ map_infer()

TypeError: extract_emojis() missing 1 required positional argument: 'emoji_list'
```

```
[48]: twitter_data.query('sentiment > 0').head()
```

```
[48]:
```

	screen_name		name	id	\
110	KieraOsiecki		keezy	559655482	
182	Mimijnm		Memi3	845613602	
320	Good_Person_111	I stand with Ukraine		1040097022307651589	
360	Sarah25Boussi		Sarah25	878738488061042689	
471	VAAF07		Valentina	126409647	

	location	followers_count	friends_count	\
110	NaN	1229	623	
182	Subn Chicago, USA	1383	3100	
320	Illinois, USA	38	111	
360	NaN	0	262	
471	NaN	204	578	

	description	artist	sentiment
110	welcome back	cher	1
182	Love my country, grandkids and dogs! (husband ...	cher	1
320	Stand with Ukraine Dan's wife Mollie's mo...	cher	1
360		cher	1
471	Berlina	cher	1

```
[47]: twitter_data.query('sentiment < 0').head()
```

```
[47]:
```

	screen_name		name	\
168	Mermarietho		Spoopy Jim	
1214	cyndeecexxx		cyndee cexxx	
1543	RussellDeer1		Russell Deer	
1750	portraitxmirage	ari mirage	i can't be your only 1	
1935	SABARBCHILE		Nicki	

	id	location	followers_count	\
168	722620867507724288	Crack House	54	
1214	1502083138813575168	NaN	57	
1543	1399742498923233281	NaN	2	
1750	1331255726468173827	the roof	1400	
1935	1390926049610018816	Cape Town, South Africa	789	

	friends_count	description	artist	\
168	266	I just love my dog, man	cher	
1214	87	communist cum dumpster. here to be shared by t...	cher	
1543	63	,	cher	
1750	1356	just a small town girl tryna make it big m...	cher	
1935	1087	It's them weak bars thinkin' that she dissin' ...	cher	

	sentiment
168	-1
1214	-1

```
1543      -1
1750      -1
1935      -1
```

```
[49]: twitter_data.groupby('artist').sentiment.mean()
```

```
[49]: artist
cher      0.008288
robyn     0.003187
Name: sentiment, dtype: float64
```

```
[50]: def get_top_n_emojis(emoji_list, n=3):
        emoji_count={}
        for emoji in emoji_list:
            if emoji in emoji_count:
                emoji_count[emoji] += 1
            else:
                emoji_count[emoji] = 1
        sorted_emojis = sorted(emoji_count.items(), key=lambda x: x[1],
        ↪reverse=True)
        top_n_emojis = [emoji for emoji, count in sorted_emojis[:n]]
        return top_n_emojis
```

```
[ ]: # was unable to run this code due to not being able to extract 'pos_emoji' above
## emojis_by_artist = twitter_data.groupby('artist')['pos_emoji'].apply(lambda
    ↪x: groupby(twitter_data['artist']))
## emojis_by_artist
```

Q: What is the average sentiment of your two artists?

A: 0.008 and 0.003

Q: Which positive emoji is the most popular for each artist? Which negative emoji?

A: <