

Project

December 18, 2020

1 POGS 9 Final Project Analysis

We want to know if there is a correlation between the increase in popularity of K-pop in America and the frequency of English lyrics in K-pop songs. We are testing this by looking at k-pop songs in the Billboard Top 100.

Hypothesis: We believe there is a positive correlation between the amount of English lyrics in a K-pop song and its ranking in the Billboard Top 100. More specifically, we believe that this correlation is cyclical. As interest in the K-pop genre increases in America, K-pop producers attempt to capitalize on this by incorporating more English lyrics in order to appeal to the Western audience. In turn, the increased frequency of English lyrics bolsters the overall popularity of K-pop songs and this cyclical process may continue.

```
[1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
plt.style.use('fivethirtyeight')
import openpyxl
from scipy.stats import linregress
spotify_data = pd.read_excel(
    'https://query.data.world/s/ey7am5nefbypy4m3l5vw3fnv3yz7wn',
    engine='openpyxl',
)

hot_stuff = pd.read_csv('https://query.data.world/s/
→puic3odpttlalwzemfc5tn7w5fkuml')
```

1.1 Raw Data:

```
[2]: spotify_data
```

```
[2]:
```

	SongID \
0	AdictoTainy, Anuel AA & Ozuna
1	The Ones That Didn't Make It Back HomeJustin M...
2	ShallowLady Gaga & Bradley Cooper
3	EnemiesPost Malone Featuring DaBaby
4	Bacc At It AgainYella Beezy, Gucci Mane & Quavo

```

...
28487      Tonight, TonightThe Smashing Pumpkins
28488  Tonight's The Night (Gonna Be Alright)Rod Stewart
28489      Tonight's The NightBLACKstreet
28490      Tonight's The NightS.S.O.
28491      Tonight's The NightSolomon Burke

```

```

                Performer \
0      Tainy, Anuel AA & Ozuna
1      Justin Moore
2      Lady Gaga & Bradley Cooper
3      Post Malone Featuring DaBaby
4      Yella Beezy, Gucci Mane & Quavo

```

```

...
28487      The Smashing Pumpkins
28488      Rod Stewart
28489      BLACKstreet
28490      S.S.O.
28491      Solomon Burke

```

```

                Song \
0      Adicto
1      The Ones That Didn't Make It Back Home
2      Shallow
3      Enemies
4      Bacc At It Again
...
28487      Tonight, Tonight
28488  Tonight's The Night (Gonna Be Alright)
28489      Tonight's The Night
28490      Tonight's The Night
28491      Tonight's The Night

```

```

                spotify_genre \
0      ['pop reggaeton']
1      ['arkansas country', 'contemporary country', '...
2      ['dance pop', 'pop']
3      ['dfw rap', 'melodic rap', 'rap']
4      ['dfw rap', 'rap', 'southern hip hop', 'trap']
...
28487  ['alternative metal', 'alternative rock', 'gru...
28488  ['adult standards', 'mellow gold', 'soft rock']
28489  ['boy band', 'dance pop', 'hip hop', 'hip pop'...
28490                                     []
28491  ['brill building pop', 'classic soul', 'funk',...

```

```

                spotify_track_id \

```

0	3jbT1Y5MoPwEIpZndDDwVq
1	NaN
2	2VxeLyX666F8uXCJ0dZF8B
3	0Xek5rqai2jcOWCYWJfVCF
4	2biNa12dMbHJrHVFRt8Jy0
...	...
28487	7bu0znpSbTks006I98ij0W
28488	10LRRKPmx8uHcSxLVWK0ct
28489	3scEwKiAkfdPUdZ6k1AEoW
28490	5Gxa3zKb3xoS9G0y4fJnP1
28491	18TM70njJRFH4Fm4ZuwLd8

	spotify_track_preview_url \
0	NaN
1	NaN
2	NaN
3	NaN
4	https://p.scdn.co/mp3-preview/fa6fa6f6f363be29...
...	...
28487	NaN
28488	https://p.scdn.co/mp3-preview/de4772519fbe8134...
28489	NaN
28490	https://p.scdn.co/mp3-preview/fb63477c965b182b...
28491	https://p.scdn.co/mp3-preview/3bc545be6705bc2e...

	spotify_track_album \
0	Adicto (with Anuel AA & Ozuna)
1	NaN
2	A Star Is Born Soundtrack
3	Hollywood's Bleeding
4	Bacc At It Again
...	...
28487	Mellon Collie And The Infinite Sadness (Deluxe...
28488	A Night on the Town (Deluxe Edition)
28489	Blackstreet
28490	Funky Chicken: Belgian Grooves From The 70's
28491	The Very Best Of Solomon Burke

	spotify_track_explicit	spotify_track_duration_ms \
0	0.0	270740.0
1	NaN	NaN
2	0.0	215733.0
3	1.0	196760.0
4	1.0	228185.0
...
28487	0.0	254626.0
28488	0.0	237120.0

28489	0.0	258026.0
28490	0.0	170506.0
28491	0.0	161186.0

	spotify_track_popularity	...	key	loudness	mode	speechiness	\
0	91.0	...	10.0	-4.803	0.0	0.0735	
1	NaN	...	NaN	NaN	NaN	NaN	
2	88.0	...	7.0	-6.362	1.0	0.0308	
3	86.0	...	6.0	-4.169	1.0	0.2100	
4	61.0	...	8.0	-5.725	0.0	0.1680	
...	
28487	69.0	...	6.0	-9.381	1.0	0.0368	
28488	55.0	...	11.0	-12.617	1.0	0.0523	
28489	32.0	...	6.0	-5.561	0.0	0.0759	
28490	13.0	...	NaN	NaN	NaN	NaN	
28491	28.0	...	7.0	-9.534	1.0	0.0290	

	acousticness	instrumentalness	liveness	valence	tempo	\
0	0.017000	0.000016	0.1790	0.623	80.002	
1	NaN	NaN	NaN	NaN	NaN	
2	0.371000	0.000000	0.2310	0.323	95.799	
3	0.058800	0.000000	0.0955	0.667	76.388	
4	0.001240	0.000001	0.0716	0.856	135.979	
...	
28487	0.000063	0.748000	0.2370	0.222	148.928	
28488	0.336000	0.125000	0.3820	0.724	136.542	
28489	0.133000	0.000011	0.1730	0.591	76.060	
28490	NaN	NaN	NaN	NaN	NaN	
28491	0.401000	0.000000	0.6000	0.633	173.305	

	time_signature
0	4.0
1	NaN
2	4.0
3	4.0
4	4.0
...	...
28487	4.0
28488	3.0
28489	4.0
28490	NaN
28491	3.0

[28492 rows x 22 columns]

[3]: hot_stuff

[3]:

	url	WeekID \
0	http://www.billboard.com/charts/hot-100/1958-0...	8/2/1958
1	http://www.billboard.com/charts/hot-100/1995-1...	12/2/1995
2	http://www.billboard.com/charts/hot-100/1997-1...	10/11/1997
3	http://www.billboard.com/charts/hot-100/2006-0...	7/1/2006
4	http://www.billboard.com/charts/hot-100/2009-1...	10/24/2009
...
320490	https://www.billboard.com/charts/hot-100/2019-...	1/5/2019
320491	https://www.billboard.com/charts/hot-100/2019-...	12/7/2019
320492	https://www.billboard.com/charts/hot-100/2019-...	12/14/2019
320493	https://www.billboard.com/charts/hot-100/2019-...	12/21/2019
320494	https://www.billboard.com/charts/hot-100/2019-...	12/28/2019

	Week Position	Song \
0	1	Poor Little Fool
1	1	One Sweet Day
2	1	Candle In The Wind 1997/Something About The Wa...
3	1	Do I Make You Proud
4	1	3
...
320490	8	Jingle Bell Rock
320491	47	Jingle Bell Rock
320492	23	Jingle Bell Rock
320493	15	Jingle Bell Rock
320494	9	Jingle Bell Rock

	Performer \
0	Ricky Nelson
1	Mariah Carey & Boyz II Men
2	Elton John
3	Taylor Hicks
4	Britney Spears
...	...
320490	Bobby Helms
320491	Bobby Helms
320492	Bobby Helms
320493	Bobby Helms
320494	Bobby Helms

	SongID	Instance \
0	Poor Little FoolRicky Nelson	1
1	One Sweet DayMariah Carey & Boyz II Men	1
2	Candle In The Wind 1997/Something About The Wa...	1
3	Do I Make You ProudTaylor Hicks	1
4	3Britney Spears	1
...
320490	Jingle Bell RockBobby Helms	7

320491	Jingle Bell Rock	Bobby Helms	8
320492	Jingle Bell Rock	Bobby Helms	8
320493	Jingle Bell Rock	Bobby Helms	8
320494	Jingle Bell Rock	Bobby Helms	8

	Previous Week Position	Peak Position	Weeks on Chart
0	NaN	1	1
1	NaN	1	1
2	NaN	1	1
3	NaN	1	1
4	NaN	1	1
...
320490	13.0	8	25
320491	NaN	47	26
320492	47.0	23	27
320493	23.0	15	28
320494	15.0	9	29

[320495 rows x 10 columns]

1.2 Data Wrangling:

```
[4]: hotter_stuff = hot_stuff.drop(columns = ['url', 'WeekID', 'Previous Week_
↳Position'])
hotter_stuff
```

```
[4]:      Week Position      Song \
0          1      Poor Little Fool
1          1      One Sweet Day
2          1  Candle In The Wind 1997/Something About The Wa...
3          1      Do I Make You Proud
4          1          3
...      ...      ...
320490      8      Jingle Bell Rock
320491     47      Jingle Bell Rock
320492     23      Jingle Bell Rock
320493     15      Jingle Bell Rock
320494      9      Jingle Bell Rock
```

```
      Performer \
0      Ricky Nelson
1  Mariah Carey & Boyz II Men
2      Elton John
3      Taylor Hicks
4      Britney Spears
...      ...
```

320490	Bobby Helms
320491	Bobby Helms
320492	Bobby Helms
320493	Bobby Helms
320494	Bobby Helms

		SongID	Instance	\
0	Poor Little Fool	Ricky Nelson	1	
1	One Sweet Day	Mariah Carey & Boyz II Men	1	
2	Candle In The Wind 1997/Something About The Wa...		1	
3	Do I Make You Proud	Taylor Hicks	1	
4	3	Britney Spears	1	
...	
320490	Jingle Bell Rock	Bobby Helms	7	
320491	Jingle Bell Rock	Bobby Helms	8	
320492	Jingle Bell Rock	Bobby Helms	8	
320493	Jingle Bell Rock	Bobby Helms	8	
320494	Jingle Bell Rock	Bobby Helms	8	

	Peak Position	Weeks on Chart
0	1	1
1	1	1
2	1	1
3	1	1
4	1	1
...
320490	8	25
320491	47	26
320492	23	27
320493	15	28
320494	9	29

[320495 rows x 7 columns]

```
[5]: #we queried for only genres containing the word "k-pop",
#however, we ended up finding that the query was not completely accurate as to
↳ what we want
mostly_kpop_songs = spotify_data[spotify_data.get('spotify_genre').str.
↳ contains('k-pop', na=False)]

#this gets rid of genres that were accidentally in the 'k-pop' query
↳ because 'k-pop' is in "fol'k-pop"
actual_kpop_songs = mostly_kpop_songs[mostly_kpop_songs.get('spotify_genre').
↳ str.contains('folk-pop') == False]

ordered_kpop_songs = actual_kpop_songs.
↳ sort_values(by='spotify_track_popularity', ascending=False)
```

```

cleaned_kpop_songs = pd.DataFrame().assign(SongID = ordered_kpop_songs.
    ↳get('SongID'),
        Performer = ordered_kpop_songs.get('SongID'),
        Song = ordered_kpop_songs.get('Song'),
        spotify_genre = ordered_kpop_songs.get('spotify_genre'),
        spotify_track_album = ordered_kpop_songs.
    ↳get('spotify_track_album'),
        spotify_track_popularity = ordered_kpop_songs.
    ↳get('spotify_track_popularity')
    )
cleaned_kpop_songs

```

```

[5]:
SongID \
431      Boy With LuvBTS Featuring Halsey
17715      Make It RightBTS
15      Make It RightBTS Featuring Lauv
5697      Kill This LoveBLACKPINK
7272      IDOLBTS Featuring Nicki Minaj
337      Chicken Noodle Soupj-hope Featuring Becky G.
7258      Fake LoveBTS
7447      MIC DropBTS Featuring Designer
19691      DNABTS
11647      Gangnam StylePSY
22280      DaddyPSY Featuring CL
17620      LIFTEDCL
7426      HangoverPSY Featuring Snoop Dogg
14196      NobodyWonder Girls
2419      JellyheadCRUSH
4621      DDU-DU DDU-DUBBLACKPINK
7280      GentlemanPSY

Performer      Song \
431      Boy With LuvBTS Featuring Halsey      Boy With Luv
17715      Make It RightBTS      Make It Right
15      Make It RightBTS Featuring Lauv      Make It Right
5697      Kill This LoveBLACKPINK      Kill This Love
7272      IDOLBTS Featuring Nicki Minaj      IDOL
337      Chicken Noodle Soupj-hope Featuring Becky G.      Chicken Noodle Soup
7258      Fake LoveBTS      Fake Love
7447      MIC DropBTS Featuring Designer      MIC Drop
19691      DNABTS      DNA
11647      Gangnam StylePSY      Gangnam Style
22280      DaddyPSY Featuring CL      Daddy
17620      LIFTEDCL      LIFTED
7426      HangoverPSY Featuring Snoop Dogg      Hangover
14196      NobodyWonder Girls      Nobody

```


2419		Jellyhead	CRUSH	Jellyhead
4621		DDU-DU	DDU-DU BLACKPINK	DDU-DU
7280		Gentleman	PSY	Gentleman

	spotify_genre	spotify_track_album \
431	['k-pop', 'k-pop boy group']	MAP OF THE SOUL : PERSONA
17715	['k-pop', 'k-pop boy group']	Make It Right (feat. Lauv)
15	['k-pop', 'k-pop boy group']	Make It Right (feat. Lauv)
5697	['k-pop', 'k-pop girl group']	KILL THIS LOVE
7272	['k-pop', 'k-pop boy group']	Love Yourself 'Answer'
337	['k-hop', 'k-pop']	Chicken Noodle Soup (feat. Becky G)
7258	['k-pop', 'k-pop boy group']	Love Yourself 'Tear'
7447	['k-pop', 'k-pop boy group']	Love Yourself 'Answer'
19691	['k-pop', 'k-pop boy group']	Love Yourself 'Her'
11647	['k-hop', 'k-pop']	Gangnam Style ()
22280	['k-hop', 'k-pop']	PSY 7TH ALBUM
17620	['k-pop']	LIFTED
7426	['k-hop', 'k-pop']	Hangover
14196	['k-pop', 'k-pop girl group']	2 Different Tears
2419	['k-pop', 'korean r&b']	Crush
4621	['k-pop', 'k-pop girl group']	NaN
7280	['k-hop', 'k-pop']	NaN

	spotify_track_popularity
431	87.0
17715	86.0
15	84.0
5697	80.0
7272	77.0
337	76.0
7258	75.0
7447	75.0
19691	73.0
11647	72.0
22280	62.0
17620	54.0
7426	50.0
14196	49.0
2419	33.0
4621	NaN
7280	NaN

```
[6]: #merged k-pop songs from the Spotify dataset with relevant Billboard
      ↪ information
merged = cleaned_kpop_songs.merge(hotter_stuff, left_on = 'SongID', right_on =
      ↪ 'SongID')
```

```
[7]: #used groupby to get rid of repeat entries and find the longest period of time
      ↳ on chart
      weeks_on_chart_groupby = merged.groupby('SongID').max()
```

```
[8]: #used groupby to get rid of repeat entries and find the highest rank of the
      ↳ song on chart
      peak_position_groupby = merged.groupby('SongID').min()
```

```
[9]: #combined relevant information from charts above
      combined_group_by = weeks_on_chart_groupby.assign(peak_position =
      ↳ peak_position_groupby.get("Peak Position") )
```

```
[10]: #dropped unnecessary columns
      combined_group_by = combined_group_by.drop(columns=
      ↳ ["Performer_x", "Song_x", "Peak Position", "Instance"])
```

```
[11]: combined_group_by
```

```
[11]:                                     spotify_genre \
```

SongID	
Boy With LuvBTS Featuring Halsey	['k-pop', 'k-pop boy group']
Chicken Noodle Soupj-hope Featuring Becky G.	['k-hop', 'k-pop']
DDU-DU DDU-DUBLACKPINK	['k-pop', 'k-pop girl group']
DNABTS	['k-pop', 'k-pop boy group']
DaddyPSY Featuring CL	['k-hop', 'k-pop']
Fake LoveBTS	['k-pop', 'k-pop boy group']
Gangnam StylePSY	['k-hop', 'k-pop']
GentlemanPSY	['k-hop', 'k-pop']
HangoverPSY Featuring Snoop Dogg	['k-hop', 'k-pop']
IDOLBTS Featuring Nicki Minaj	['k-pop', 'k-pop boy group']
JellyheadCRUSH	['k-pop', 'korean r&b']
Kill This LoveBLACKPINK	['k-pop', 'k-pop girl group']
LIFTEDCL	['k-pop']
MIC DropBTS Featuring Desiigner	['k-pop', 'k-pop boy group']
Make It RightBTS	['k-pop', 'k-pop boy group']
Make It RightBTS Featuring Lauv	['k-pop', 'k-pop boy group']
NobodyWonder Girls	['k-pop', 'k-pop girl group']

spotify_track_album \

SongID

Boy With LuvBTS Featuring Halsey

MAP OF THE SOUL :

PERSONA

Chicken Noodle Soupj-hope Featuring Becky G. Chicken Noodle Soup (feat. Becky G)

DDU-DU DDU-DUBLACKPINK

NaN

DNABTS

Love Yourself

'Her'	
DaddyPSY Featuring CL	PSY 7TH
ALBUM	
Fake LoveBTS	Love Yourself
'Tear'	
Gangnam StylePSY	Gangnam Style
()	
GentlemanPSY	
NaN	
HangoverPSY Featuring Snoop Dogg	
Hangover	
IDOLBTS Featuring Nicki Minaj	Love Yourself
'Answer'	
JellyheadCRUSH	
Crush	
Kill This LoveBLACKPINK	KILL THIS
LOVE	
LIFTEDCL	
LIFTED	
MIC DropBTS Featuring Desiigner	Love Yourself
'Answer'	
Make It RightBTS	Make It Right (feat.
Lauv)	
Make It RightBTS Featuring Lauv	Make It Right (feat.
Lauv)	
NobodyWonder Girls	2 Different
Tears	

SongID	spotify_track_popularity \
Boy With LuvBTS Featuring Halsey	87.0
Chicken Noodle Soupj-hope Featuring Becky G.	76.0
DDU-DU DDU-DUBBLACKPINK	NaN
DNABTS	73.0
DaddyPSY Featuring CL	62.0
Fake LoveBTS	75.0
Gangnam StylePSY	72.0
GentlemanPSY	NaN
HangoverPSY Featuring Snoop Dogg	50.0
IDOLBTS Featuring Nicki Minaj	77.0
JellyheadCRUSH	33.0
Kill This LoveBLACKPINK	80.0
LIFTEDCL	54.0
MIC DropBTS Featuring Desiigner	75.0
Make It RightBTS	86.0
Make It RightBTS Featuring Lauv	84.0
NobodyWonder Girls	49.0

SongID	Week Position \
Boy With LuvBTS Featuring Halsey	90
Chicken Noodle Soupj-hope Featuring Becky G.	81
DDU-DU DDU-DUBLACKPINK	55
DNABTS	98
DaddyPSY Featuring CL	97
Fake LoveBTS	76
Gangnam StylePSY	64
GentlemanPSY	95
HangoverPSY Featuring Snoop Dogg	26
IDOLBTS Featuring Nicki Minaj	92
JellyheadCRUSH	97
Kill This LoveBLACKPINK	100
LIFTEDCL	94
MIC DropBTS Featuring Desiigner	98
Make It RightBTS	95
Make It RightBTS Featuring Lauv	76
NobodyWonder Girls	76

SongID	Song_y \
Boy With LuvBTS Featuring Halsey	Boy With Luv
Chicken Noodle Soupj-hope Featuring Becky G.	Chicken Noodle Soup
DDU-DU DDU-DUBLACKPINK	DDU-DU DDU-DU
DNABTS	DNA
DaddyPSY Featuring CL	Daddy
Fake LoveBTS	Fake Love
Gangnam StylePSY	Gangnam Style
GentlemanPSY	Gentleman
HangoverPSY Featuring Snoop Dogg	Hangover
IDOLBTS Featuring Nicki Minaj	IDOL
JellyheadCRUSH	Jellyhead
Kill This LoveBLACKPINK	Kill This Love
LIFTEDCL	LIFTED
MIC DropBTS Featuring Desiigner	MIC Drop
Make It RightBTS	Make It Right
Make It RightBTS Featuring Lauv	Make It Right
NobodyWonder Girls	Nobody

SongID	Performer_y \
Boy With LuvBTS Featuring Halsey	BTS Featuring Halsey
Chicken Noodle Soupj-hope Featuring Becky G.	j-hope Featuring Becky G.
DDU-DU DDU-DUBLACKPINK	BLACKPINK
DNABTS	BTS

DaddyPSY Featuring CL	PSY Featuring CL
Fake LoveBTS	BTS
Gangnam StylePSY	PSY
GentlemanPSY	PSY
HangoverPSY Featuring Snoop Dogg	PSY Featuring Snoop Dogg
IDOLBTS Featuring Nicki Minaj	BTS Featuring Nicki Minaj
JellyheadCRUSH	CRUSH
Kill This LoveBLACKPINK	BLACKPINK
LIFTEDCL	CL
MIC DropBTS Featuring Desiigner	BTS Featuring Desiigner
Make It RightBTS	BTS
Make It RightBTS Featuring Lauv	BTS Featuring Lauv
NobodyWonder Girls	Wonder Girls

SongID	Weeks on Chart	peak_position
Boy With LuvBTS Featuring Halsey	8	8
Chicken Noodle Soupj-hope Featuring Becky G.	1	81
DDU-DU DDU-DUBBLACKPINK	1	55
DNABTS	4	67
DaddyPSY Featuring CL	1	97
Fake LoveBTS	6	10
Gangnam StylePSY	31	2
GentlemanPSY	15	5
HangoverPSY Featuring Snoop Dogg	1	26
IDOLBTS Featuring Nicki Minaj	3	11
JellyheadCRUSH	21	72
Kill This LoveBLACKPINK	4	41
LIFTEDCL	1	94
MIC DropBTS Featuring Desiigner	9	28
Make It RightBTS	1	95
Make It RightBTS Featuring Lauv	1	76
NobodyWonder Girls	1	76

```
[12]: #removed JellyheadCRUSH because it was not actually a k-pop song
filtered = combined_group_by.drop("JellyheadCRUSH")
```

```
[13]: #one of the songs had Spanish lyrics, so we decided to exclude those, hence the
      ↪name
      #Korean and English words rather than total words
total_korean_n_eng_words = np.array([347, 581, 265, 200, 466, 406, 262, 263,
      ↪838, 413, 308, 394, 387, 331, 376, 324])
total_eng_words = np.array([192, 473, 117, 35, 381, 194, 21, 188, 802, 303,
      ↪185, 394, 138, 131, 226, 176])
total_unique_eng_words = np.array([52, 91, 43, 14, 53, 34, 8, 27, 158, 132, 85,
      ↪121, 55, 24, 67, 32])
```

```
[14]: #added hand-inputted data into dataset
word_info = filtered.assign(total_korean_n_eng_words =
    ↳total_korean_n_eng_words, total_eng_words = total_eng_words,
    ↳total_unique_eng_words = total_unique_eng_words)

word_info_proportion = word_info.assign(Proportion_English = word_info.
    ↳get('total_eng_words') / word_info.get('total_korean_n_eng_words'))
word_info_proportion
```

```
[14]:                                     spotify_genre \
SongID
Boy With LuvBTS Featuring Halsey          ['k-pop', 'k-pop boy group']
Chicken Noodle Soupj-hope Featuring Becky G. ['k-hop', 'k-pop']
DDU-DU DDU-DUBLACKPINK                    ['k-pop', 'k-pop girl group']
DNABTS                                     ['k-pop', 'k-pop boy group']
DaddyPSY Featuring CL                     ['k-hop', 'k-pop']
Fake LoveBTS                             ['k-pop', 'k-pop boy group']
Gangnam StylePSY                         ['k-hop', 'k-pop']
GentlemanPSY                             ['k-hop', 'k-pop']
HangoverPSY Featuring Snoop Dogg          ['k-hop', 'k-pop']
IDOLBTS Featuring Nicki Minaj             ['k-pop', 'k-pop boy group']
Kill This LoveBLACKPINK                  ['k-pop', 'k-pop girl group']
LIFTEDCL                                  ['k-pop']
MIC DropBTS Featuring Desiigner           ['k-pop', 'k-pop boy group']
Make It RightBTS                         ['k-pop', 'k-pop boy group']
Make It RightBTS Featuring Lauv           ['k-pop', 'k-pop boy group']
NobodyWonder Girls                       ['k-pop', 'k-pop girl group']

spotify_track_album \
SongID
Boy With LuvBTS Featuring Halsey          MAP OF THE SOUL :
PERSONA
Chicken Noodle Soupj-hope Featuring Becky G. Chicken Noodle Soup (feat. Becky
G)
DDU-DU DDU-DUBLACKPINK
NaN
DNABTS                                     Love Yourself
'Her'
DaddyPSY Featuring CL                     PSY 7TH
ALBUM
Fake LoveBTS                             Love Yourself
'Tear'
Gangnam StylePSY                         Gangnam Style
( )
GentlemanPSY
NaN
HangoverPSY Featuring Snoop Dogg
```

Hangover	
IDOLBTS Featuring Nicki Minaj	Love Yourself
'Answer'	
Kill This LoveBLACKPINK	KILL THIS
LOVE	
LIFTEDCL	
LIFTED	
MIC DropBTS Featuring Desiigner	Love Yourself
'Answer'	
Make It RightBTS	Make It Right (feat.
Lauv)	
Make It RightBTS Featuring Lauv	Make It Right (feat.
Lauv)	
NobodyWonder Girls	2 Different
Tears	

SongID	spotify_track_popularity \
Boy With LuvBTS Featuring Halsey	87.0
Chicken Noodle Soupj-hope Featuring Becky G.	76.0
DDU-DU DDU-DUBBLACKPINK	NaN
DNABTS	73.0
DaddyPSY Featuring CL	62.0
Fake LoveBTS	75.0
Gangnam StylePSY	72.0
GentlemanPSY	NaN
HangoverPSY Featuring Snoop Dogg	50.0
IDOLBTS Featuring Nicki Minaj	77.0
Kill This LoveBLACKPINK	80.0
LIFTEDCL	54.0
MIC DropBTS Featuring Desiigner	75.0
Make It RightBTS	86.0
Make It RightBTS Featuring Lauv	84.0
NobodyWonder Girls	49.0

SongID	Week Position \
Boy With LuvBTS Featuring Halsey	90
Chicken Noodle Soupj-hope Featuring Becky G.	81
DDU-DU DDU-DUBBLACKPINK	55
DNABTS	98
DaddyPSY Featuring CL	97
Fake LoveBTS	76
Gangnam StylePSY	64
GentlemanPSY	95
HangoverPSY Featuring Snoop Dogg	26
IDOLBTS Featuring Nicki Minaj	92

Kill This Love	BLACKPINK	100
LIFTED	DCL	94
MIC Drop	BTS Featuring Desiigner	98
Make It Right	BTS	95
Make It Right	BTS Featuring Lauv	76
Nobody	Wonder Girls	76

Song_y \

SongID	
Boy With Luv	BTS Featuring Halsey
Chicken Noodle Soup	j-hope Featuring Becky G.
DDU-DU DDU-DU	BLACKPINK
DNABTS	
Daddy	PSY Featuring CL
Fake Love	BTS
Gangnam Style	PSY
Gentleman	PSY
Hangover	PSY Featuring Snoop Dogg
IDOL	BTS Featuring Nicki Minaj
Kill This Love	BLACKPINK
LIFTED	DCL
MIC Drop	BTS Featuring Desiigner
Make It Right	BTS
Make It Right	BTS Featuring Lauv
Nobody	Wonder Girls

Performer_y \

SongID	
Boy With Luv	BTS Featuring Halsey
Chicken Noodle Soup	j-hope Featuring Becky G.
DDU-DU DDU-DU	BLACKPINK
DNABTS	BTS
Daddy	PSY Featuring CL
Fake Love	BTS
Gangnam Style	PSY
Gentleman	PSY
Hangover	PSY Featuring Snoop Dogg
IDOL	BTS Featuring Nicki Minaj
Kill This Love	BLACKPINK
LIFTED	DCL
MIC Drop	BTS Featuring Desiigner
Make It Right	BTS
Make It Right	BTS Featuring Lauv
Nobody	Wonder Girls

Weeks on Chart peak_position \

SongID

Boy With LuvBTS Featuring Halsey	8	8
Chicken Noodle Soupj-hope Featuring Becky G.	1	81
DDU-DU DDU-DUBLACKPINK	1	55
DNABTS	4	67
DaddyPSY Featuring CL	1	97
Fake LoveBTS	6	10
Gangnam StylePSY	31	2
GentlemanPSY	15	5
HangoverPSY Featuring Snoop Dogg	1	26
IDOLBTS Featuring Nicki Minaj	3	11
Kill This LoveBLACKPINK	4	41
LIFTEDCL	1	94
MIC DropBTS Featuring Desiigner	9	28
Make It RightBTS	1	95
Make It RightBTS Featuring Lauv	1	76
NobodyWonder Girls	1	76

total_korean_n_eng_words \

SongID	
Boy With LuvBTS Featuring Halsey	347
Chicken Noodle Soupj-hope Featuring Becky G.	581
DDU-DU DDU-DUBLACKPINK	265
DNABTS	200
DaddyPSY Featuring CL	466
Fake LoveBTS	406
Gangnam StylePSY	262
GentlemanPSY	263
HangoverPSY Featuring Snoop Dogg	838
IDOLBTS Featuring Nicki Minaj	413
Kill This LoveBLACKPINK	308
LIFTEDCL	394
MIC DropBTS Featuring Desiigner	387
Make It RightBTS	331
Make It RightBTS Featuring Lauv	376
NobodyWonder Girls	324

total_eng_words \

SongID	
Boy With LuvBTS Featuring Halsey	192
Chicken Noodle Soupj-hope Featuring Becky G.	473
DDU-DU DDU-DUBLACKPINK	117
DNABTS	35
DaddyPSY Featuring CL	381
Fake LoveBTS	194
Gangnam StylePSY	21
GentlemanPSY	188
HangoverPSY Featuring Snoop Dogg	802

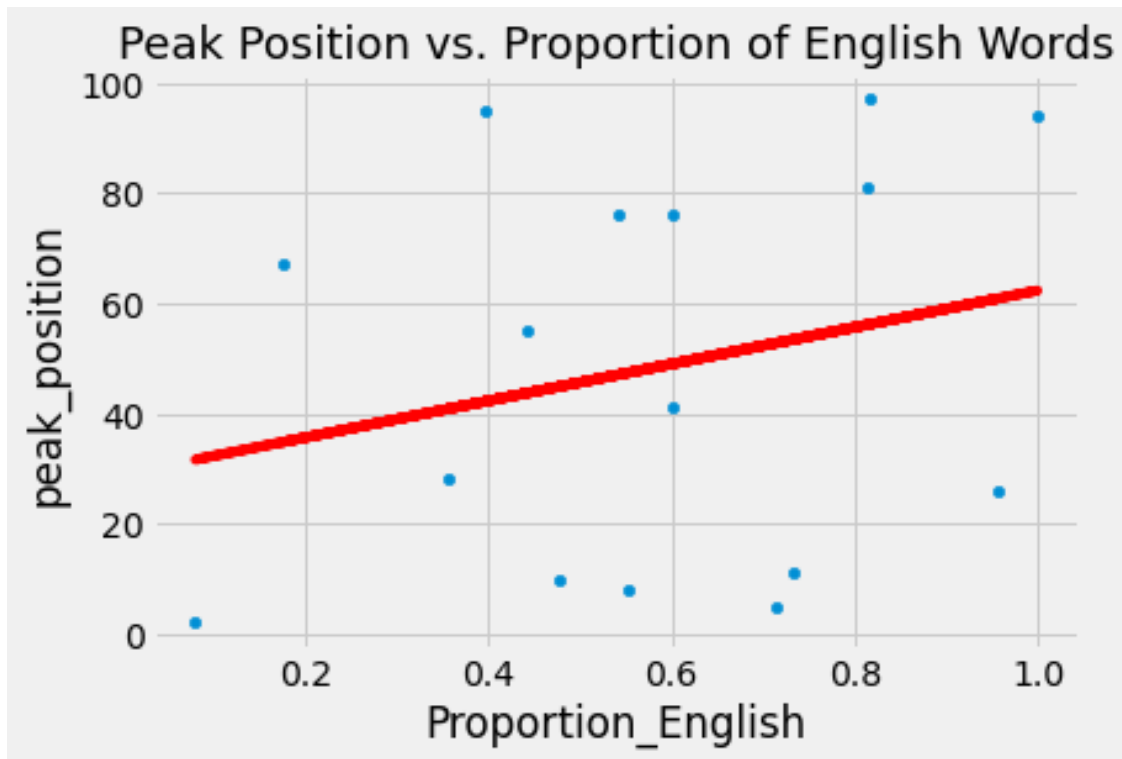
IDOLBTS Featuring Nicki Minaj	303
Kill This LoveBLACKPINK	185
LIFTEDCL	394
MIC DropBTS Featuring Desiigner	138
Make It RightBTS	131
Make It RightBTS Featuring Lauv	226
NobodyWonder Girls	176

SongID	total_unique_eng_words \
Boy With LuvBTS Featuring Halsey	52
Chicken Noodle Soupj-hope Featuring Becky G.	91
DDU-DU DDU-DUBBLACKPINK	43
DNABTS	14
DaddyPSY Featuring CL	53
Fake LoveBTS	34
Gangnam StylePSY	8
GentlemanPSY	27
HangoverPSY Featuring Snoop Dogg	158
IDOLBTS Featuring Nicki Minaj	132
Kill This LoveBLACKPINK	85
LIFTEDCL	121
MIC DropBTS Featuring Desiigner	55
Make It RightBTS	24
Make It RightBTS Featuring Lauv	67
NobodyWonder Girls	32

SongID	Proportion_English
Boy With LuvBTS Featuring Halsey	0.553314
Chicken Noodle Soupj-hope Featuring Becky G.	0.814114
DDU-DU DDU-DUBBLACKPINK	0.441509
DNABTS	0.175000
DaddyPSY Featuring CL	0.817597
Fake LoveBTS	0.477833
Gangnam StylePSY	0.080153
GentlemanPSY	0.714829
HangoverPSY Featuring Snoop Dogg	0.957041
IDOLBTS Featuring Nicki Minaj	0.733656
Kill This LoveBLACKPINK	0.600649
LIFTEDCL	1.000000
MIC DropBTS Featuring Desiigner	0.356589
Make It RightBTS	0.395770
Make It RightBTS Featuring Lauv	0.601064
NobodyWonder Girls	0.543210

1.3 Graphing:

```
[15]: word_info_proportion.plot(kind="scatter", x="Proportion_English", y =  
      ↪ "peak_position")  
plt.title('Peak Position vs. Proportion of English Words', fontsize= 18)  
r = np.corrcoef(word_info_proportion['Proportion_English'],  
      ↪ word_info_proportion['peak_position'])[0][1]  
slope, intercept, r_value, p_value, std_err = linregress(word_info_proportion.  
      ↪ get('Proportion_English'), word_info_proportion.get('peak_position'))  
plt.plot(word_info_proportion.get('Proportion_English'), intercept +  
      ↪ slope*word_info_proportion.get('Proportion_English'), 'r', label='fitted_  
      ↪ line')  
plt.show()  
print('The correlation coefficent for this graph is: ' + str(r))
```



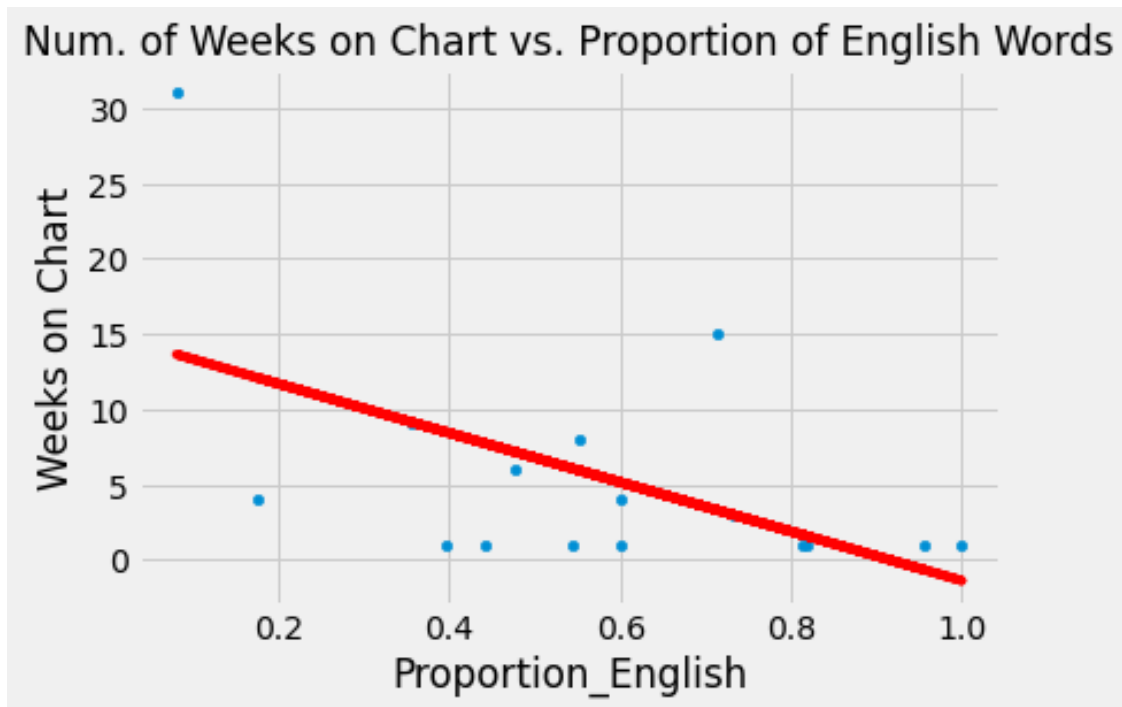
The correlation coefficent for this graph is: 0.241555687033949

```
[16]: word_info_proportion.plot(kind="scatter", x="Proportion_English", y = "Weeks on_  
      ↪ Chart")  
plt.title('Num. of Weeks on Chart vs. Proportion of English Words', fontsize=17)  
r = np.corrcoef(word_info_proportion['Proportion_English'],  
      ↪ word_info_proportion['Weeks on Chart'])[0][1]
```

```

slope, intercept, r_value, p_value, std_err = linregress(word_info_proportion.
    ↳get('Proportion_English'), word_info_proportion.get('Weeks on Chart'))
plt.plot(word_info_proportion.get('Proportion_English'), intercept +
    ↳slope*word_info_proportion.get('Proportion_English'), 'r', label='fitted_
    ↳line')
plt.show()
print('The correlation coefficient for this graph is: ' + str(r))

```

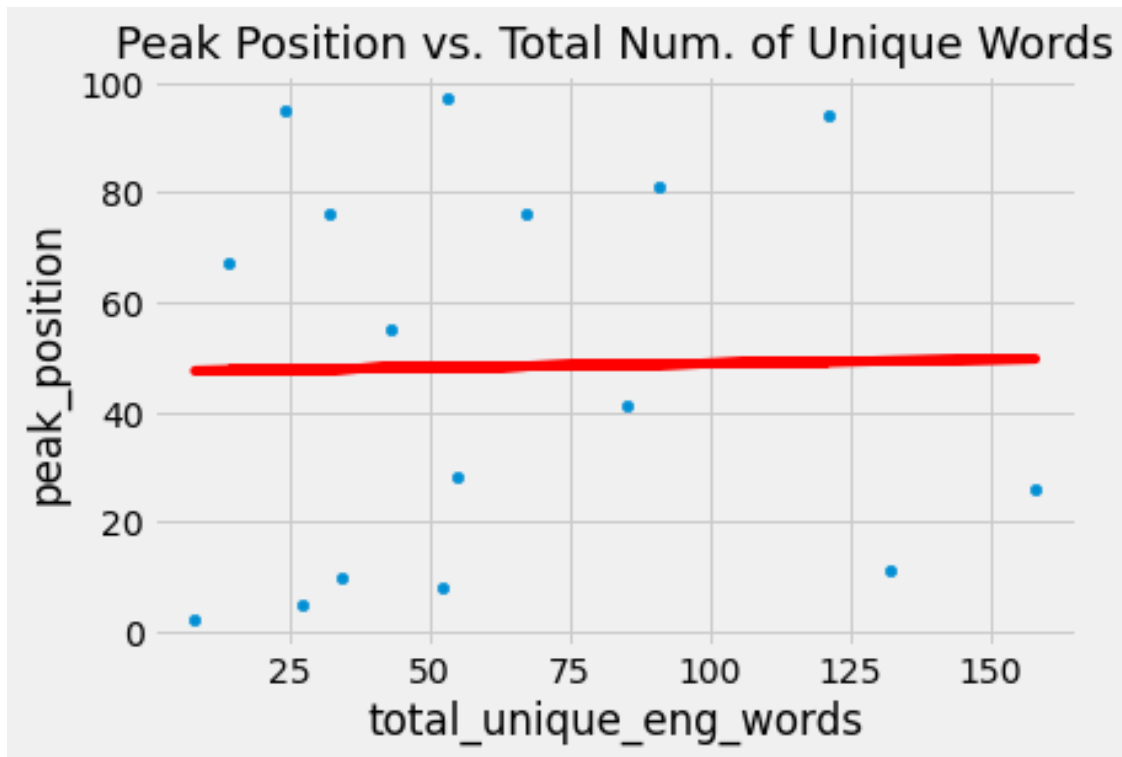


The correlation coefficient for this graph is: -0.5353551677748455

```

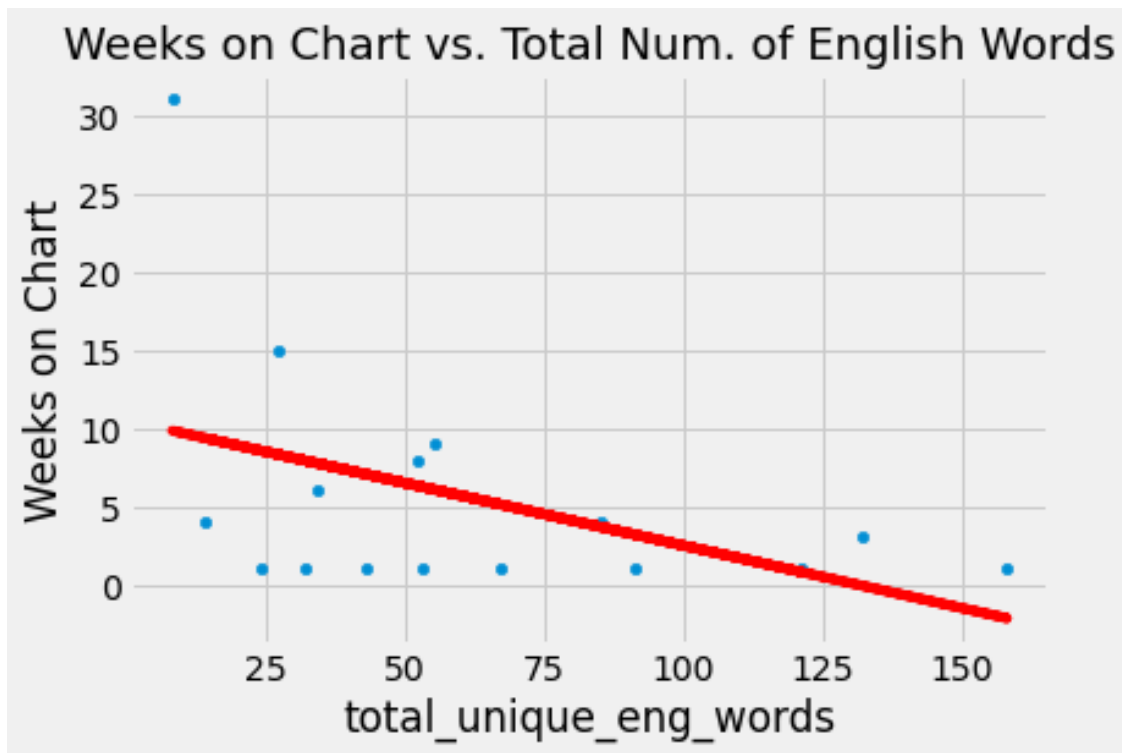
[17]: word_info_proportion.plot(kind="scatter", x="total_unique_eng_words",
    ↳y="peak_position")
plt.title('Peak Position vs. Total Num. of Unique Words', fontsize=18)
r = np.corrcoef(word_info_proportion['total_unique_eng_words'],
    ↳word_info_proportion['peak_position'])[0][1]
slope, intercept, r_value, p_value, std_err = linregress(word_info_proportion.
    ↳get('total_unique_eng_words'), word_info_proportion.get('peak_position'))
plt.plot(word_info_proportion.get('total_unique_eng_words'), intercept +
    ↳slope*word_info_proportion.get('total_unique_eng_words'), 'r', label='fitted_
    ↳line')
plt.show()
print('The correlation coefficient for this graph is: ' + str(r))

```



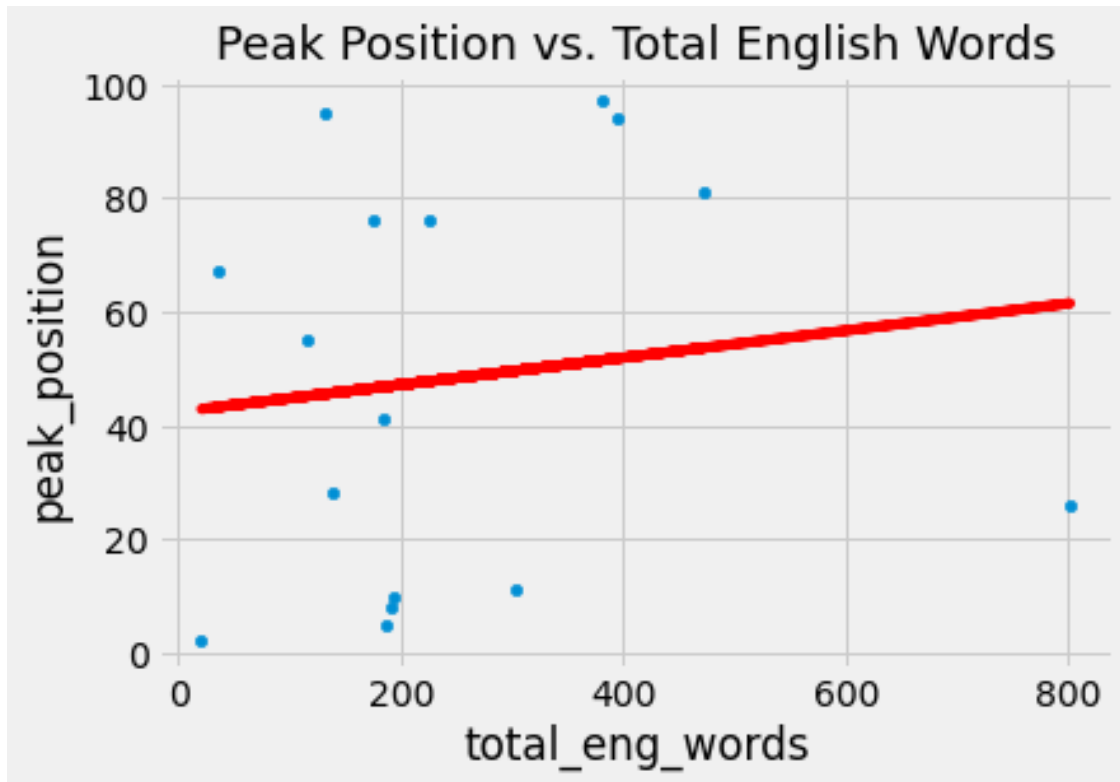
The correlation coefficient for this graph is: 0.018135789163855286

```
[18]: word_info_proportion.plot(kind="scatter", x="total_unique_eng_words", y="Weeks on Chart")
plt.title('Weeks on Chart vs. Total Num. of English Words', fontsize=18)
r = np.corrcoef(word_info_proportion['total_unique_eng_words'],
    word_info_proportion['Weeks on Chart'])[0][1]
slope, intercept, r_value, p_value, std_err = linregress(word_info_proportion.get('total_unique_eng_words'),
    word_info_proportion.get('Weeks on Chart'))
plt.plot(word_info_proportion.get('total_unique_eng_words'), intercept +
    slope*word_info_proportion.get('total_unique_eng_words'), 'r', label='fitted line')
plt.show()
print('The correlation coefficient for this graph is: ' + str(r))
```



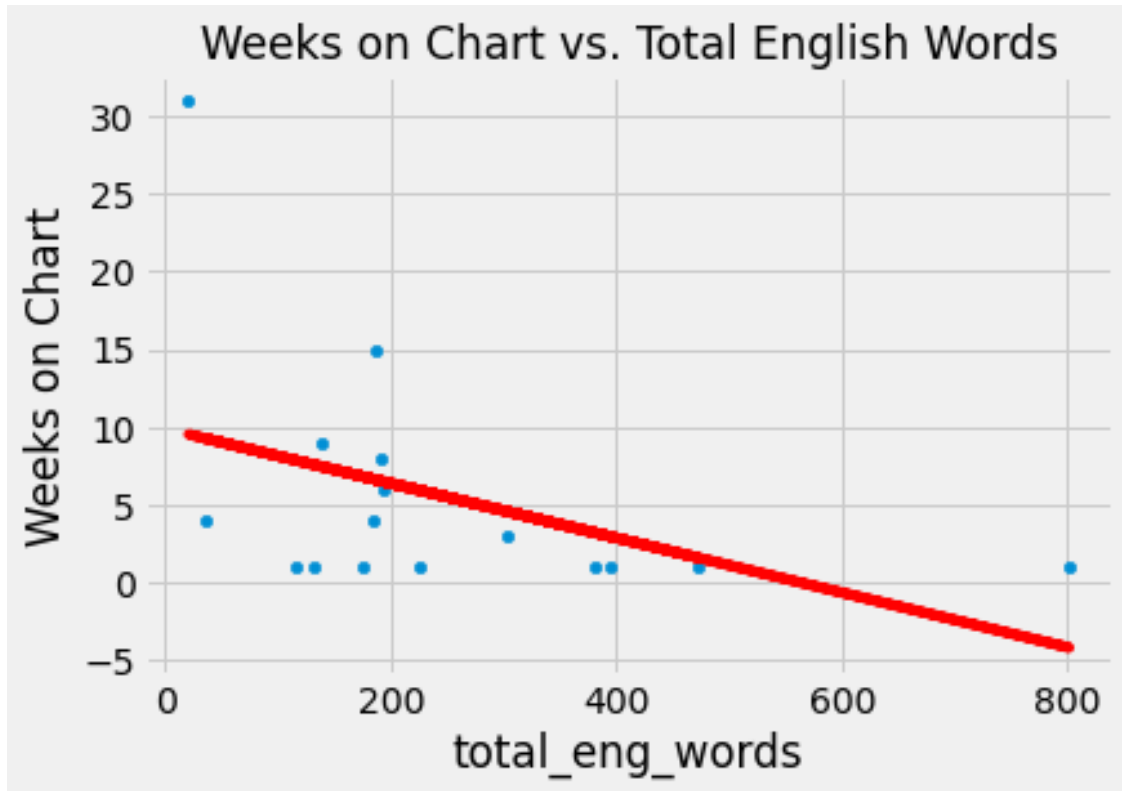
The correlation coefficient for this graph is: -0.4467916102919963

```
[19]: word_info_proportion.plot(kind="scatter", x="total_eng_words",
    ↪y="peak_position")
plt.title('Peak Position vs. Total English Words', fontsize=18)
r = np.corrcoef(word_info_proportion['total_eng_words'],
    ↪word_info_proportion['peak_position'])[0][1]
slope, intercept, r_value, p_value, std_err = linregress(word_info_proportion.
    ↪get('total_eng_words'), word_info_proportion.get('peak_position'))
plt.plot(word_info_proportion.get('total_eng_words'), intercept +
    ↪slope*word_info_proportion.get('total_eng_words'), 'r', label='fitted line')
plt.show()
print('The correlation coefficient for this graph is: ' + str(r))
```



The correlation coefficient for this graph is: 0.12872016676850873

```
[20]: word_info_proportion.plot(kind="scatter", x="total_eng_words", y="Weeks on Chart")
plt.title('Weeks on Chart vs. Total English Words', fontsize=17)
r = np.corrcoef(word_info_proportion['total_eng_words'],
    ↳word_info_proportion['Weeks on Chart'])[0][1]
slope, intercept, r_value, p_value, std_err = linregress(word_info_proportion.
    ↳get('total_eng_words'), word_info_proportion.get('Weeks on Chart'))
plt.plot(word_info_proportion.get('total_eng_words'), intercept +
    ↳slope*word_info_proportion.get('total_eng_words'), 'r', label='fitted line')
plt.show()
print('The correlation coefficient for this graph is: ' + str(r))
```



The correlation coefficient for this graph is: -0.4319204745570109

1.4 Conclusion:

From the possible correlation coefficients we could calculate, most of the correlation coefficients were very low (below an absolute value of 0.5). Furthermore, the highest correlation coefficient was only -0.535 . This leads us to believe that there is not possible correlation between the amount of English lyrics in a K-pop song and its ranking in the Billboard Top 100