BTS Word Cloud

```
# Importing Libraries
import pandas as pd
import re
import nltk
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
from nltk.stem import WordNetLemmatizer
from wordcloud import WordCloud
import matplotlib.pyplot as plt
# Importing the file
df = pd.read_csv('lyrics-v6.csv',encoding='latin1')
df.head()
\overline{z}
         id album_title eng_album_title album_rd album_seq track_title kor_track_title eng_track_title
                                                                                                                            lyrics hidden_track rem
                                                                                                                           weâ□□re
                                                                                                                          now going
                                                                       Intro: 2 Cool
                  2 Cool 4
                                               2013-06-
                                                                                                           Intro: 2 Cool 4
                                                                                                                                  to
      0
          0
                                2 Cool 4 Skool
                                                                         4 Skool (ft.
                                                                                                 NaN
                                                                                                                                              False
                                                                                                                                                     Fa
                                                                                                        Skool (ft. DJ Friz)
                     Skool
                                                      12
                                                                                                                            progress
                                                                           DJ Friz)
                                                                                                                            to some
                                                                                                                          steps\nw...
                                                                                                                              (what)
                                                                                                                            give it to
                                                                            We Are
                                               2013-06-
                  2 Cool 4
                                                                                                                 We Are
                                                                                                                               me\n
                                2 Cool 4 Skool
          1
                                                                   2
                                                                         Bulletproof
                                                                                                 NaN
                                                                                                                                              False Fa
                     Skool
                                                      12
                                                                                                          Bulletproof Pt.2
                                                                                                                           (what) be
                                                                              Pt.2
                                                                                                                           nervous\n
                                                                                                                               (wh...
                                                                                                                                 rap
                                                                                                                          monster: it
                                                                         Skit: Circle
                                                                                                        Skit: Circle Room
                  2 Cool 4
                                               2013-06-
                                                                                                                           was a big
                                2 Cool 4 Skool
      2
         2
                                                                   3
                                                                                                 NaN
                                                                                                                                              False Fa
                     Skool
                                                      12
                                                                         Room Talk
                                                                                                                    Talk
                                                                                                                              hit\nv:
                                                                                                                               year
                                                                                                                           2006!\n...
                                                                                                                               hey,
                                                                                                                              what's
                                                                                                                               your
                  2 Cool 4
                                                2013-06-
                                                                          No More
      3
          3
                                2 Cool 4 Skool
                                                                                                 NaN
                                                                                                         No More Dream
                                                                                                                           dream?\n
                                                                                                                                              False
                                                                                                                                                    Fa
                     Skool
                                                      12
                                                                            Dream
                                                                                                                                hey,
                                                                                                                          whatâ□□s
                                                                                                                            your d...
                  2 Cool 4
                                                2013-06-
                                2 Cool 4 Skool
                                                                   5
                                                                          Interlude
                                                                                                 NaN
                                                                                                                Interlude
                                                                                                                               NaN
                                                                                                                                              False
                                                                                                                                                    Fa
                     Skool
                                                      12
               Generate code with df
                                          View recommended plots
                                                                           New interactive sheet
 Next steps:
# Extracting Lyrics column
lyrics = df['lyrics']
lyrics.head()
₹
                                                  lyrics
      0 weâ□□re now going to progress to some steps\nw...
      1
              (what) give it to me\n (what) be nervous\n (wh...
      2
               rap monster: it was a big hit\nv: year 2006!\n...
          hey, what's your dream?\n hey, whatâ□□s your d...
      3
      4
                                                    NaN
```

```
# Removing Null Values
lyrics = lyrics.dropna()
lyrics.head()
→
                                                           lyrics
        0 weâ□□re now going to progress to some steps\nw...
                 (what) give it to me\n (what) be nervous\n (wh...
       1
        2
                  rap monster: it was a big hit\nv: year 2006!\n...
              hey, what's your dream?\n hey, whatâ□□s your d...
        3
        \textbf{5} \quad \text{wanna be loved} \\ \textbf{a} \\ \textbf{a} \\ \textbf{h} \\ \textbf{n} \\ \textbf{d} \\ \textbf{a} \\ \textbf{t} \\ \textbf{wanna be fool wan...}
# Removing special characters except a-zA-Z
lyrics = lyrics.apply(lambda x: re.sub(r'[^a-zA-Z\s]', '', x))
lyrics.head()
<del>_</del>
                                                         lyrics
        0 were now going to progress to some steps\nwhic...
       1
              what give it to me\n what be nervous\n what th...
        2
               rap monster it was a big hit\nv year \nrap mon...
        3 hey whats your dream\n hey whats your dream\n ...
        5 wanna be loved\n\ndont wanna be fool wanna be ...
# Converting all the texts to lower case
lyrics = lyrics.apply(lambda x: x.lower())
# Tokenization
nltk.download('punkt')
df['lyrics_tokenized'] = lyrics.apply(lambda x: word_tokenize(x))
[nltk_data] Downloading package punkt to /root/nltk_data... [nltk_data] Package punkt is already up-to-date!
df.head()
```

```
\overline{2}
          id album_title eng_album_title album_rd album_seq track_title kor_track_title eng_track_title
                                                                                                                                      lyrics hidden_track rem
                                                                                                                                    weâ□□re
                                                                                                                                   now going
                                                                            Intro: 2 Cool
                    2 Cool 4
                                                   2013-06-
                                                                                                                   Intro: 2 Cool 4
       n
           0
                                  2 Cool 4 Skool
                                                                              4 Skool (ft.
                                                                                                        NaN
                                                                                                                                                        False
                                                                                                                                                               Fa
                                                                                                                Skool (ft. DJ Friz)
                      Skool
                                                                                                                                     progress
                                                                                DJ Friz)
                                                                                                                                     to some
                                                                                                                                   steps\nw...
                                                                                                                                       (what)
                                                                                                                                     give it to
                                                                                 We Are
                    2 Cool 4
                                                   2013-06-
                                                                                                                         We Are
                                                                                                                                        me\n
           1
                                  2 Cool 4 Skool
                                                                        2
                                                                              Bulletproof
                                                                                                        NaN
                                                                                                                                                        False
                                                                                                                                                               Fa
                                                                                                                  Bulletproof Pt.2
                      Skool
                                                          12
                                                                                                                                    (what) be
                                                                                    Pt.2
                                                                                                                                    nervous\n
                                                                                                                                        (wh...
                                                                                                                                          rap
                                                                                                                                   monster: it
                    2 Cool 4
                                                   2013-06-
                                                                              Skit: Circle
                                                                                                                Skit: Circle Room
                                                                                                                                    was a big
          2
                                  2 Cool 4 Skool
       2
                                                                        3
                                                                                                        NaN
                                                                                                                                                        False Fa
                      Skool
                                                          12
                                                                              Room Talk
                                                                                                                             Talk
                                                                                                                                       hit\nv:
                                                                                                                                         vear
                                                                                                                                    2006!\n...
                                                                                                                                         hey,
                                                                                                                                       what's
                                                                                                                                         your
                    2 Cool 4
                                                   2013-06-
                                                                                No More
           3
                                  2 Cool 4 Skool
       3
                                                                                                        NaN
                                                                                                                 No More Dream
                                                                                                                                    dream?\n
                                                                                                                                                                Fa
                                                                        4
                                                                                                                                                        False
                      Skool
                                                          12
                                                                                  Dream
                                                                                                                                         hey,
                                                                                                                                   whatâ□□s
                                                                                                                                     your d...
                                                   2013-06-
                    2 Cool 4
                                  2 Cool 4 Skool
           4
                                                                        5
                                                                                Interlude
                                                                                                        NaN
                                                                                                                        Interlude
                                                                                                                                         NaN
                                                                                                                                                        False
                                                                                                                                                                Fa
                      Skool
                                                          12
 Next steps:
                Generate code with df
                                             View recommended plots
                                                                                 New interactive sheet
# Drop all other columns from the dataframe except lyrics and lyrics tokenized
df = df.drop(['id', 'album_title', 'eng_album_title', 'album_rd', 'album_seq','track_title','kor_track_title','eng_track_title','hidden_t
df.head()
→
                                                     lvrics
                                                                                         lyrics_tokenized
                                                                                                                扁
       \textbf{0} \quad \text{we} \textbf{\^{a}} \ \Box \ \text{re now going to progress to some steps} \\ \text{nw...} \quad \text{[were, now, going, to, progress, to, some, ste...]}
       1
               (what) give it to me\n (what) be nervous\n (wh...
                                                               [what, give, it, to, me, what, be, nervous, wh...
       2
                rap monster: it was a big hit\nv: year 2006!\n...
                                                                    [rap, monster, it, was, a, big, hit, v, year, ...
       3
           hey, what's your dream?\n hey, whatâ □ □s your d... [hey, whats, your, dream, hey, whats, your, dr...
       4
                                                        NaN
                                                                                                        NaN
 Next steps:
                Generate code with df
                                             View recommended plots
                                                                                 New interactive sheet
df=df,dropna()
df.head()
\overline{2}
                                                       lyrics
                                                                                          lyrics_tokenized
                                                                                                                 0
           weâ□□re now going to progress to some steps\nw... [were, now, going, to, progress, to, some, ste...
                                                                                                                 d.
       1
                (what) give it to me\n (what) be nervous\n (wh...
                                                                 [what, give, it, to, me, what, be, nervous, wh...
       2
                 rap monster: it was a big hit\nv: year 2006!\n...
                                                                     [rap, monster, it, was, a, big, hit, v, year, ...
       3
            hey, what's your dream?\n hey, whatâ □ □s your d... [hey, whats, your, dream, hey, whats, your, dr...
          wanna be lovedâ□!\n\ndonâ□□t wanna be fool wan.
                                                                  Iwan. na. be. loved. dont. wan. na. be. fool
                Generate code with df
                                             View recommended plots
 Next steps:
# Lemmatization
nltk.download('wordnet')
nltk.download('omw-1.4')
lemmatizer = WordNetLemmatizer()
\label{eq:df['lyrics_lemmatized'] = df['lyrics_tokenized'].apply(lambda x: [lemmatizer.lemmatize(word) for word in x])} \\
```

```
→ [nltk_data] Downloading package wordnet to /root/nltk_data...
      [nltk_data]
                      Package wordnet is already up-to-date!
      [nltk_data] Downloading package omw-1.4 to /root/nltk_data...
      [nltk data]
                      Package omw-1.4 is already up-to-date!
df.head()
\rightarrow
                                                      lvrics
                                                                                        lyrics_tokenized
                                                                                                                                      lyrics_lemmatized
                                                                    [were, now, going, to, progress, to, some,
                                                                                                                  [were, now, going, to, progress, to, some,
       0
          weâ□□re now going to progress to some steps\nw...
       1
               (what) give it to me\n (what) be nervous\n (wh...
                                                              [what, give, it, to, me, what, be, nervous, wh...
                                                                                                            [what, give, it, to, me, what, be, nervous, wh...
       2
                rap monster: it was a big hit\nv: year 2006!\n...
                                                                   [rap, monster, it, was, a, big, hit, v, year, ...
                                                                                                                  [rap, monster, it, wa, a, big, hit, v, year, r...
                                                                   [hey, whats, your, dream, hey, whats, your,
                                                                                                                 [hey, whats, your, dream, hey, whats, your,
       3
            hey, what's your dream?\n hey, whatâ□□s your d...
 Next steps:
                Generate code with df
                                            View recommended plots
                                                                                New interactive sheet
# Removing stopwords
nltk.download('stopwords')
stop words = set(stopwords.words('english'))
\label{eq:df['lyrics_lemmatized_cleaned'] = df['lyrics_lemmatized'].apply(lambda x: [word for word in x if word not in stop_words])} \\
     [nltk_data] Downloading package stopwords to /root/nltk_data...
      [nltk_data] Package stopwords is already up-to-date!
df.head()
\overline{\Rightarrow}
                                      lvrics
                                                            lyrics_tokenized
                                                                                             lyrics_lemmatized
                                                                                                                            lyrics_lemmatized_cleaned
                                                                                                                                                              weâ\square\,\square re now going to progress to
                                                 [were, now, going, to, progress,
                                                                                   [were, now, going, to, progress,
                                                                                                                         [going, progress, step, bit, difficult,
                                                                                                                                                              ıl.
       0
                             some steps\nw...
                                                                 to, some, ste...
                                                                                                   to, some, ste...
                                                                                                                                                  ready,...
                                                                                   [what, give, it, to, me, what, be,
                (what) give it to me\n (what) be
                                                 [what, give, it, to, me, what, be,
                                                                                                                       faive, nervous, one, end, bulletproof,
                              nervous\n (wh...
                                                                  nervous, wh...
                                                                                                    nervous, wh...
                                                                                                                                                  bulletp...
           rap monster: it was a big hit\nv: year
                                                [rap, monster, it, was, a, big, hit,
                                                                                  [rap, monster, it, wa, a, big, hit, v,
                                                                                                                     [rap, monster, wa, big, hit, v, year, rap,
                                    2006!\n...
                                                                                                          vear. r...
                                                                                                                                                    mon
                                                                      v. vear. ..
                hey, what's your dream?\n hey,
                                                   [hey, whats, your, dream, hey,
                                                                                     [hey, whats, your, dream, hey,
                                                                                                                     [hey, whats, dream, hey, whats, dream,
       3
                           whatâ□□s your d...
                                                                whats, your, dr...
                                                                                                  whats, your, dr...
                                                                                                                                                 hey, wh...
               wanna be lovedâ□!\n\ndonâ□□t
                                                  [wan, na, be, loved, dont, wan,
                                                                                    [wan, na, be, loved, dont, wan,
                                                                                                                        [wan, na, loved, dont, wan, na, fool,
       5
                          wanna be fool wan.
                                                                                                                                                wan. na....
                                                                  na. be. fool. .
                                                                                                    na. be. fool. ...
 Next steps:
                Generate code with df
                                            View recommended plots
                                                                                New interactive sheet
# Creating word cloud
text = ' '.join(df['lyrics_lemmatized_cleaned'].apply(lambda x: ' '.join(x)))
wordcloud = WordCloud(width=800, height=800, background_color='white').generate(text)
plt.figure(figsize=(8, 8), facecolor=None)
plt.imshow(wordcloud)
plt.axis('off')
plt.tight_layout(pad=0)
plt.show()
```

