

ADS 509 Assignment 2.1: Tokenization, Normalization, Descriptive Statistics

This notebook holds Assignment 2.1 for Module 2 in ADS 509, Applied Text Mining. Work through this notebook, writing code and answering questions where required.

In the previous assignment you pulled lyrics data on two artists. In this assignment we explore this data set and a pull from the now-defunct Twitter API for the artists Cher and Robyn. 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 Canvas.

This assignment asks you to write a short function to calculate some descriptive statistics on a piece of text. Then you are asked to find some interesting and unique statistics on your corpora.

```
In [80]: import nltk
nltk.download('stopwords')
```

```
[nltk_data] Downloading package stopwords to
[nltk_data]   /Users/kevinbaum/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
```

```
Out[80]: True
```

```
In [81]: import os
import re
import emoji
import pandas as pd
import numpy as np

from collections import Counter, defaultdict
from nltk.corpus import stopwords
from string import punctuation

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

```
In [82]: # Add any additional import statements you need here

# import os to aid in reading in the data
import os
import matplotlib.pyplot as plt
```

KB: I commented out the below file switching code as I decided to simply move the twitter and lyrics files into the current directory I am already in. Thus, when I push this to github and you download it, the files will already be in the same directory together

```
In [83]: # change `data_location` to the location of the folder on your machine.
```

```
# data_location = "/users/chandler/dropbox/teaching/repos/ads-tm-api-scrape/
```

```
In [84]: # These subfolders should still work if you correctly stored the
# data from the Module 1 assignment
twitter_folder = "twitter/"
lyrics_folder = "lyrics/"
```

```
In [85]: def descriptive_stats(tokens, num_tokens = 5, verbose=True) :
        """
        Given a list of tokens, print number of tokens, number of unique tokens,
        number of characters, lexical diversity (https://en.wikipedia.org/wiki/Lexical\_diversity)
        and num_tokens most common tokens. Return a list with the number of
        tokens, number of unique tokens, lexical diversity, and number of characters.

        """

        # Fill in the correct values here.

        # num_tokens is the length of the text that gets passed into the function
        num_tokens = len(tokens)

        # num_unique_tokens contains the "set" data type, which represents a
        # collection of unique elements. It lets us get a unique count of
        # different words in the text.
        num_unique_tokens = len(set(tokens))

        # To get lexical diversity, we divide the number of unique tokens
        # by the total. If the number of unique tokens is identical to the
        # total, then we get a perfect lexical diversity of 1.
        lexical_diversity = num_unique_tokens / num_tokens

        # We use a for loop here to access the letters within the tokens.
        # To do this, we take a length of each word within the text,
        # then sum it.
        num_characters = sum(len(token) for token in tokens)

        if verbose :
            print(f"There are {num_tokens} tokens in the data.")
            print(f"There are {num_unique_tokens} unique tokens in the data.")
            print(f"There are {num_characters} characters in the data.")
            print(f"The lexical diversity is {lexical_diversity:.3f} in the data.")

        # print the five most common tokens
        # "Counter" comes from the "collections" module that we imported above
        # It lets us pass in the number of tokens to make it into a "Counter"
        # object that we can then
        common_tokens = Counter(tokens).most_common(num_tokens)
        print(f"The {num_tokens} most common tokens are: {common_tokens}")

        return([num_tokens, num_unique_tokens,
                lexical_diversity,
                num_characters])
```

```
In [86]: text = """here is some example text with other example text here in this tex
assert(descriptive_stats(text, verbose=True)[0] == 13)
assert(descriptive_stats(text, verbose=False)[1] == 9)
assert(abs(descriptive_stats(text, verbose=False)[2] - 0.69) < 0.02)
assert(descriptive_stats(text, verbose=False)[3] == 55)
```

There are 13 tokens in the data.

There are 9 unique tokens in the data.

There are 55 characters in the data.

The lexical diversity is 0.692 in the data.

The 13 most common tokens are: [('text', 3), ('here', 2), ('example', 2), ('is', 1), ('some', 1), ('with', 1), ('other', 1), ('in', 1), ('this', 1)]

Q: Why is it beneficial to use assertion statements in your code?

A: Assertion statements help the proper functioning of the code. By asserting that a condition is true at a specific point in the code, you can catch errors. Meaning, if an error is produced, that means there is a flaw in the logic and there is less to disentangle later had the assertion statement not been used. In this way, assertion statements validate the code is working as expected and can cut down the time it takes to handle errors and bugs.

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 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.

```
In [87]: # Read in the lyrics data
def read_lyrics_data(folder_path):
    # Set up the dictionary for storing the data
    lyrics_data = defaultdict(dict)

    # Use the imported os dictionary, which lets us access the underlying
    # operating system being run and treats it kind of like an API that
    # it is interacting with.
    # "os.walk()" is a generator that walks a dictionary tree and returns
    # a tuple of 3 values (root, dirs, files). The whole purpose of this is
    # access the files via code within the directories and get the informati
    # out of them to put into the dictionary we set up.
    for root, dirs, files in os.walk(folder_path):
        for file in files:

            # Extract information when dealing with the text files
```

```

        if file.endswith(".txt"):

            # Divides up the artist and song based on the underscore
            artist, song = file[:-4].split('_')

            # Read in the data by opening the files and using f.read()
            with open(os.path.join(root, file), 'r', encoding='utf-8') as f:
                lyrics_data[artist][song] = f.read()

    return lyrics_data

# We use the "lyrics" prefix to access the contents of the lyrics folder
lyrics_folder = "lyrics/"
lyrics_data = read_lyrics_data(lyrics_folder)

```

```

In [88]: def read_twitter_data(folder_path):
    # Setting up a twitter_data dictionary to store the information in
    twitter_data = {}

    # Iterate over each file in the directory for twitter data
    for file in os.listdir(folder_path):

        # Performs the text reading when dealing with the txt files
        if file.endswith(".txt"):

            # Construct the file path
            file_path = os.path.join(folder_path, file)

            # Open the file for reading
            with open(file_path, 'r', encoding='utf-8') as f:

                # Reads all lines from the file and "lines" list
                lines = f.readlines()

                # Extract the "description" field which is the 7th (0-indexed)
                # field in the tab separated twitter data files
                descriptions = [line.split('\t')[6] for line in lines[1:] if

                # Extract the artist name by removing ".txt" from the suffix
                artist = file[:-4]

                # The descriptions are stored in the twitter_data dictionary
                twitter_data[artist] = descriptions

    return twitter_data

twitter_folder = "twitter/"
twitter_data = read_twitter_data(twitter_folder)

```

Data Cleaning

Now clean and tokenize your data. Remove punctuation characters (available in the `punctuation` object in the `string` library), split on whitespace, fold to lowercase, and remove stopwords. Store your cleaned data, which must be accessible as an

iterable for `descriptive_stats`, in new objects or in new columns in your data frame.

```
In [89]: punctuation = set(punctuation) # speeds up comparison

# This function takes in takes a string text, removes punctuation,
# tokenizes it into words, converts to lowercase, and removes stopwords.
# The result is a list of clean tokens.
def clean_and_tokenize(text, keep_hashtags=True):
    """Function to clean and tokenize a given text."""
    # Modify punctuation set to keep '#' if keep_hashtags is True
    if keep_hashtags:
        modified_punctuation = set(punctuation) - set('#')
    else:
        modified_punctuation = set(punctuation)

    # Remove punctuation
    text = ''.join(ch for ch in text if ch not in modified_punctuation)
    # Tokenize and convert to lowercase
    tokens = text.lower().split()
    # Remove stopwords
    tokens = [word for word in tokens if word not in sw]
    return tokens
```

```
In [90]: # create your clean twitter data here
clean_twitter_data = {}
for artist, descriptions in twitter_data.items():

    # Clean the twitter data descriptions using the function
    clean_descriptions = [clean_and_tokenize(desc) for desc in descriptions]
    clean_twitter_data[artist] = clean_descriptions
```

```
In [91]: # create your clean lyrics data here
clean_lyrics_data = {}
for artist, songs in lyrics_data.items():
    clean_songs = {}
    for song, lyrics in songs.items():
        clean_songs[song] = clean_and_tokenize(lyrics)
    clean_lyrics_data[artist] = clean_songs
```

Basic Descriptive Statistics

Call your `descriptive_stats` function on both your lyrics data and your twitter data and for both artists (four total calls).

```
In [92]: # calls to descriptive_stats here

# For Lyrics Data
all_lyrics_artist1 = [token for song in clean_lyrics_data['jonimitchell'].values()]
all_lyrics_artist2 = [token for song in clean_lyrics_data['kanyewest'].values()]

# For Twitter Data
```

```
all_tweets_artist1 = [token for tweet in clean_twitter_data['cher_followers_
all_tweets_artist2 = [token for tweet in clean_twitter_data['robynkonihiwa_

# Descriptive Stats Calls
print("Lyrics Data - Joni Mitchell:")
descriptive_stats(all_lyrics_artist1)

print("\nLyrics Data - Kanye West:")
descriptive_stats(all_lyrics_artist2)

print("\nTwitter Data - Cher:")
descriptive_stats(all_tweets_artist1)

print("\nTwitter Data - Robyn:")
descriptive_stats(all_tweets_artist2)
```

Lyrics Data – Joni Mitchell:

There are 2649 tokens in the data.

There are 1642 unique tokens in the data.

There are 16876 characters in the data.

The lexical diversity is 0.620 in the data.

The 2649 most common tokens are: [('know', 31), ('like', 23), ('dont', 23), ('love', 19), ('come', 17), ('think', 17), ('morning', 16), ('one', 15), ('never', 13), ('go', 11), ('ask', 11), ('still', 10), ('comes', 10), ('say', 10), ('heart', 10), ('looked', 9), ('back', 9), ('wont', 9), ('children', 9), ('always', 8), ('another', 8), ('rain', 8), ('saw', 8), ('see', 8), ('find', 8), ('dreams', 8), ('take', 7), ('fit', 7), ('city', 7), ('light', 7), ('got ten', 7), ('man', 7), ('found', 7), ('lady', 6), ('would', 6), ('let', 6), ('day', 6), ('ill', 6), ('heard', 6), ('cant', 6), ('night', 6), ('said', 6), ('youre', 6), ('shes', 6), ('playing', 6), ('sun', 6), ('thats', 6), ('came', 6), ('knows', 6), ('aging', 6), ('may', 5), ('sides', 5), ('clouds', 5), ('life', 5), ('win', 5), ('king', 5), ('im', 5), ('time', 5), ('looks', 5), ('window', 5), ('cannot', 5), ('turn', 5), ('good', 5), ('first', 5), ('chelsea', 5), ('whos', 5), ('hell', 5), ('songs', 4), ('nowfrom', 4), ('so mehowits', 4), ('illusions', 4), ('recalli', 4), ('really', 4), ('laughing', 4), ('shake', 4), ('call', 4), ('sorrow', 4), ('theres', 4), ('hear', 4), ('tell', 4), ('hes', 4), ('till', 4), ('blue', 4), ('left', 4), ('roses', 4), ('sinking', 4), ('brings', 4), ('us', 4), ('knew', 4), ('sometimes', 4), ('dowe', 4), ('thru', 4), ('well', 4), ('free', 4), ('put', 4), ('michael', 4), ('streets', 4), ('gentle', 4), ('please', 4), ('sea', 4), ('promise', 4), ('dream', 4), ('thing', 4), ('someone', 4), ('full', 4), ('stones', 4), ('busy', 4), ('freetheres', 4), ('youve', 4), ('garden', 4), ('seagulls', 4), ('flyout', 4), ('reach', 4), ('ladies', 3), ('canyon', 3), ('leaf', 3), ('welcome', 3), ('empty', 3), ('takes', 3), ('home', 3), ('dear', 3), ('angel', 3), ('waybut', 3), ('things', 3), ('got', 3), ('every', 3), ('leave', 3), ('give', 3), ('loves', 3), ('fears', 3), ('old', 3), ('friends', 3), ('ive', 3), ('lost', 3), ('voice', 3), ('early', 3), ('cani', 3), ('miss', 3), ('wind', 3), ('needs', 3), ('green', 3), ('along', 3), ('show', 3), ('grey', 3), ('went', 3), ('taken', 3), ('curtains', 3), ('anymoreyou', 3), ('keys', 3), ('dooryou', 3), ('thoughts', 3), ('manthey', 3), ('name', 3), ('town', 3), ('pretty', 3), ('run', 3), ('says', 3), ('feel', 3), ('stars', 3), ('away', 3), ('hard', 3), ('place', 3), ('lately', 3), ('black', 3), ('doyes', 3), ('rocking', 3), ('last', 3), ('real', 3), ('play', 3), ('curtain', 3), ('mountainsgo', 3), ('toknow', 3), ('yousomeday', 3), ('dark', 3), ('keep', 3), ('want', 3), ('deadi', 3), ('saint', 3), ('cruelbut', 3), ('youi', 3), ('somewhere', 3), ('seadreams', 3), ('believe', 3), ('song', 3), ('staywell', 3), ('rainbow', 3), ('beads', 3), ('told', 3), ('keeps', 3), ('cactus', 3), ('asked', 3), ('make', 3), ('cityand', 3), ('comeaging', 3), ('fiddle', 3), ('cry', 3), ('cometo', 3), ('trade', 3), ('peace', 3), ('fear', 3), ('pirate', 3), ('penance', 3), ('answerthe', 3), ('danceri', 3), ('dollar', 3), ('nathan', 3), ('ghostly', 3), ('hand', 3), ('understandfear', 3), ('wilderl andstepping', 3), ('midway', 3), ('wearing', 3), ('round', 2), ('circus', 2), ('beveled', 2), ('gathers', 2), ('flowers', 2), ('flowersand', 2), ('sun shine', 2), ('music', 2), ('ice', 2), ('castles', 2), ('canyons', 2), ('block', 2), ('many', 2), ('way', 2), ('care', 2), ('feeling', 2), ('somethings', 2), ('lose', 2), ('lives', 2), ('stand', 2), ('funny', 2), ('looking', 2), ('laughter', 2), ('leaves', 2), ('wrote', 2), ('youand', 2), ('strange', 2), ('guess', 2), ('marcie', 2), ('inside', 2), ('sweet', 2), ('greens', 2), ('letter', 2), ('wash', 2), ('ocean', 2), ('summer', 2), ('yellow', 2), ('stop', 2), ('winter', 2), ('tells', 2), ('thought', 2), ('shopping', 2), ('west', 2), ('painting', 2), ('sings', 2), ('wars', 2), ('get', 2), ('places', 2), ('menight', 2), ('finemusic', 2), ('spilling', 2), ('streetcolors', 2),

('stayathome', 2), ('faces', 2), ('willy', 2), ('child', 2), ('hed', 2), ('live', 2), ('mebut', 2), ('gave', 2), ('stood', 2), ('cars', 2), ('reasons', 2), ('himwilly', 2), ('wants', 2), ('spell', 2), ('breaking', 2), ('tears', 2), ('shinglesi', 2), ('bluei', 2), ('rose', 2), ('begins', 2), ('trembleto', 2), ('togotten', 2), ('cards', 2), ('seasons', 2), ('puts', 2), ('sisotowbell', 2), ('stories', 2), ('windowsometimes', 2), ('water', 2), ('theyre', 2), ('waiting', 2), ('today', 2), ('dirty', 2), ('townand', 2), ('street', 2), ('played', 2), ('mountains', 2), ('colored', 2), ('wellmichael', 2), ('strings', 2), ('hold', 2), ('leads', 2), ('called', 2), ('galleryi', 2), ('liked', 2), ('ones', 2), ('began', 2), ('bedi', 2), ('saidyou', 2), ('flying', 2), ('meand', 2), ('across', 2), ('roll', 2), ('harbor', 2), ('learns', 2), ('themi', 2), ('loving', 2), ('lies', 2), ('seaall', 2), ('seen', 2), ('fly', 2), ('wings', 2), ('brought', 2), ('woke', 2), ('bells', 2), ('dayand', 2), ('talk', 2), ('awayi', 2), ('bring', 2), ('conversationi', 2), ('comfort', 2), ('sometimescomfort', 2), ('consultationhe', 2), ('sees', 2), ('much', 2), ('sharing', 2), ('climbed', 2), ('calls', 2), ('beside', 2), ('forest', 2), ('treewhile', 2), ('since', 2), ('writes', 2), ('tea', 2), ('drum', 2), ('friendand', 2), ('alland', 2), ('whyyou', 2), ('raise', 2), ('sticks', 2), ('falloh', 2), ('friendhow', 2), ('drumyou', 2), ('turnedlike', 2), ('enemies', 2), ('earnedbut', 2), ('rememberall', 2), ('areand', 2), ('pleasecan', 2), ('help', 2), ('staroh', 2), ('port', 2), ('sails', 2), ('met', 2), ('near', 2), ('watch', 2), ('knowpenanceshe', 2), ('morgantown', 2), ('rise', 2), ('everyone', 2), ('morgantownmorning', 2), ('morgantownbuy', 2), ('downmorning', 2), ('namemornings', 2), ('la', 2), ('common', 2), ('daylight', 2), ('paid', 2), ('reached', 2), ('falls', 2), ('voices', 2), ('cryi', 2), ('trina', 1), ('wears', 1), ('wampum', 1), ('beadsshe', 1), ('fills', 1), ('drawing', 1), ('book', 1), ('linessewing', 1), ('lace', 1), ('widows', 1), ('weedsand', 1), ('filigree', 1), ('vinevine', 1), ('filigreeand', 1), ('coats', 1), ('secondhand', 1), ('onetrimmed', 1), ('antique', 1), ('luxuryshe', 1), ('canyonannie', 1), ('sits', 1), ('eatshe', 1), ('makes', 1), ('incats', 1), ('babies', 1), ('feetand', 1), ('fat', 1), ('none', 1), ('thinnone', 1), ('thin', 1), ('fatshe', 1), ('bake', 1), ('brownies', 1), ('todaysaying', 1), ('backshe', 1), ('ladyestrella', 1), ('girlcomes', 1), ('wrapped', 1), ('gypsy', 1), ('shawlssongs', 1), ('tiny', 1), ('hammers', 1), ('hurledat', 1), ('mirrors', 1), ('hallsempy', 1), ('halls', 1), ('mirrorssailing', 1), ('seas', 1), ('climbing', 1), ('banyanscome', 1), ('visit', 1), ('hereto', 1), ('canyontrina', 1), ('paints', 1), ('threadsand', 1), ('weaves', 1), ('pattern', 1), ('ownannie', 1), ('bakes', 1), ('cakes', 1), ('breadsand', 1), ('homefor', 1), ('estrella', 1), ('companioncolors', 1), ('hourspouring', 1), ('canyoncoloring', 1), ('hoursthey', 1), ('rows', 1), ('floes', 1), ('hairand', 1), ('cream', 1), ('airand', 1), ('feather', 1), ('everywhereive', 1), ('sunthey', 1), ('snow', 1), ('everyoneso', 1), ('donebut', 1), ('wayive', 1), ('cloud', 1), ('allmoons', 1), ('junes', 1), ('ferris', 1), ('wheels the', 1), ('dizzy', 1), ('dancing', 1), ('feelas', 1), ('fairy', 1), ('tale', 1), ('realive', 1), ('showyou', 1), ('em', 1), ('goand', 1), ('knowdon t', 1), ('awayive', 1), ('alltears', 1), ('proudto', 1), ('right', 1), ('loud dreams', 1), ('schemes', 1), ('crowdsive', 1), ('acting', 1), ('strangethe y', 1), ('heads', 1), ('changedwell', 1), ('gainedin', 1), ('living', 1), ('dayive', 1), ('allive', 1), ('finding', 1), ('theresunny', 1), ('braiding', 1), ('wild', 1), ('hairpicked', 1), ('pencil', 1), ('finest', 1), ('hand wanted', 1), ('send', 1), ('standtelephone', 1), ('even', 1), ('sound', 1), ('newall', 1), ('alone', 1), ('california', 1), ('talking', 1), ('foolish', 1), ('words', 1), ('plannedi', 1), ('cause', 1), ('standcrickets', 1), ('court ing', 1), ('stardappled', 1), ('greenthickets', 1), ('tall', 1), ('dreamall', 1), ('muted', 1), ('misty', 1), ('drowsy', 1), ('sleep', 1), ('standi', 1), ('coat', 1), ('flowersstops', 1), ('candy', 1), ('storereads', 1), ('sour

still', 1), ('doorso', 1), ('shell', 1), ('flower', 1), ('curtainshang', 1),
 ('drydust', 1), ('tables', 1), ('shirt', 1), ('andwave', 1), ('goodbyemarcie
 s', 1), ('faucet', 1), ('plumbermarcies', 1), ('manred', 1), ('autumn', 1),
 ('summergreens', 1), ('turning', 1), ('sandall', 1), ('beachesstares', 1),
 ('skymarcie', 1), ('buys', 1), ('bag', 1), ('peachesstops', 1), ('postman',
 1), ('passing', 1), ('byand', 1), ('goesfalls', 1), ('sidewalk', 1), ('string',
 1), ('brown', 1), ('paperwinter', 1), ('blowsup', 1), ('river', 1), ('he
 rto', 1), ('seamarcie', 1), ('dresses', 1), ('warm', 1), ('snowingtakes',
 1), ('cab', 1), ('uptownred', 1), ('goingsees', 1), ('rides', 1), ('downdow
 n', 1), ('hudson', 1), ('riverpast', 1), ('shipyards', 1), ('coldstill', 1),
 ('letters', 1), ('deliveredstill', 1), ('days', 1), ('unfoldlike', 1), ('mag
 azinesfading', 1), ('dusty', 1), ('attics', 1), ('cellarsmake', 1), ('dreamd
 ream', 1), ('herwait', 1), ('memarcie', 1), ('doesnt', 1), ('uswhere', 1),
 ('moved', 1), ('awayred', 1), ('angry', 1), ('jealousthat', 1), ('saysomeon
 e', 1), ('sundaywindow', 1), ('rainsomeone', 1), ('bought', 1), ('oneway',
 1), ('ticketand', 1), ('tenement', 1), ('castlrelately', 1), ('pastel', 1),
 ('walls', 1), ('brownhes', 1), ('downhes', 1), ('swept', 1), ('broom', 1),
 ('contemptand', 1), ('rooms', 1), ('ringhes', 1), ('cleaned', 1), ('tearso
 f', 1), ('actor', 1), ('laughters', 1), ('stingi', 1), ('dressed', 1), ('dri
 pdry', 1), ('paisleylately', 1), ('saying', 1), ('crazy', 1), ('blindhe',
 1), ('lives', 1), ('timeladies', 1), ('gingham', 1), ('blushwhile', 1), ('wi
 nebut', 1), ('leather', 1), ('lacei', 1), ('become', 1), ('kindi', 1), ('sal
 trusted', 1), ('carriagewho', 1), ('carried', 1), ('country', 1), ('marriag
 e', 1), ('soonbeware', 1), ('power', 1), ('moonstheres', 1), ('blameno', 1),
 ('traitor', 1), ('herethe', 1), ('kings', 1), ('roadand', 1), ('queens', 1),
 ('grove', 1), ('end', 1), ('yeari', 1), ('cantthey', 1), ('uplight', 1), ('la
 zy', 1), ('eyesmoons', 1), ('nights', 1), ('uptaking', 1), ('surprisenight',
 1), ('timeday', 1), ('hour', 1), ('agocity', 1), ('timemust', 1), ('ready',
 1), ('slowthere', 1), ('gonight', 1), ('flashing', 1), ('timetake', 1), ('of
 ftake', 1), ('shoesbreak', 1), ('offchase', 1), ('bluesstairway', 1), ('stai
 rwaydown', 1), ('crowds', 1), ('streetthey', 1), ('waylooking', 1), ('greetb
 ut', 1), ('meetnight', 1), ('waltzing', 1), ('fatheri', 1), ('lifehe', 1),
 ('ancient', 1), ('injurythat', 1), ('healedhe', 1), ('againlike', 1), ('soon
 he', 1), ('laceat', 1), ('face', 1), ('conquered', 1), ('moonand', 1), ('cou
 nting', 1), ('hilland', 1), ('sillthere', 1), ('whyi', 1), ('joy', 1), ('sor
 rownow', 1), ('hidehe', 1), ('realhe', 1), ('chapels', 1), ('pealing', 1),
 ('silver', 1), ('bellsbut', 1), ('tellwhen', 1), ('wrong', 1), ('realbut',
 1), ('bound', 1), ('loseif', 1), ('blues', 1), ('scared', 1), ('feeland',
 1), ('bornlike', 1), ('shiny', 1), ('stormthere', 1), ('father', 1), ('toshe
 s', 1), ('mysterious', 1), ('devotionsshes', 1), ('zodiac', 1), ('zenshes',
 1), ('tarot', 1), ('potionsshes', 1), ('laying', 1), ('religion', 1), ('frie
 ndson', 1), ('friendsfriends', 1), ('futurefriends', 1), ('friendsbecause',
 1), ('signs', 1), ('suit', 1), ('hershell', 1), ('prophecy', 1), ('death',
 1), ('whenwont', 1), ('whenwhen', 1), ('barterno', 1), ('stacked', 1), ('win
 shell', 1), ('head', 1), ('treat', 1), ('martyrit', 1), ('blackest', 1), ('i
 nputs', 1), ('inin', 1), ('lure', 1), ('youinside', 1), ('selfpity', 1), ('s
 wimin', 1), ('drown', 1), ('haunts', 1), ('wincan', 1), ('winyou', 1), ('las
 ting', 1), ('laurels', 1), ('laughterit', 1), ('reaches', 1), ('arm', 1),
 ('sinkto', 1), ('solitary', 1), ('truth', 1), ('afteryou', 1), ('dare', 1),
 ('priestess', 1), ('thinkhow', 1), ('thinki', 1), ('lane', 1), ('lanenoah',
 1), ('fixing', 1), ('pump', 1), ('rainhe', 1), ('shamewe', 1), ('knewup',
 1), ('hilljovial', 1), ('neighbors', 1), ('willwith', 1), ('tellsometimes',
 1), ('chaireach', 1), ('rocks', 1), ('shareeating', 1), ('muffin', 1), ('bun
 s', 1), ('berriesby', 1), ('steamy', 1), ('kitchen', 1), ('doour', 1), ('ton
 gues', 1), ('bluesisotowbell', 1), ('laneanywhere', 1), ('else', 1), ('see
 m', 1), ('strangethe', 1), ('changingeveryday', 1), ('everywaysometimes',

1), ('springsometimes', 1), ('anything', 1), ('poet', 1), ('singsometimes', 1), ('tryyes', 1), ('trywe', 1), ('chairsomedays', 1), ('rock', 1), ('stareat', 1), ('woodlands', 1), ('grasslands', 1), ('badlands', 1), ('cross', 1), ('riversometimes', 1), ('viewsisotowbell', 1), ('lanego', 1), ('youll', 1), ('againto', 1), ('wade', 1), ('grainyou', 1), ('docome', 1), ('starssweet', 1), ('pickling', 1), ('jarswell', 1), ('lend', 1), ('carwe', 1), ('chairsomeone', 1), ('thererocking', 1), ('rhythms', 1), ('candle', 1), ('wait', 1), ('slept', 1), ('hoteli', 1), ('jewelsthe', 1), ('rushed', 1), ('around', 1), ('schoolsi', 1), ('standing', 1), ('noisy', 1), ('cornerwaiting', 1), ('walking', 1), ('greenacross', 1), ('stoodand', 1), ('goodon', 1), ('clarinet', 1), ('freenow', 1), ('fortunesand', 1), ('velvet', 1), ('callsi', 1), ('limousine', 1), ('two', 1), ('gentlemenescorting', 1), ('hallsand', 1), ('moneyor', 1), ('friend', 1), ('bandby', 1), ('quick', 1), ('lunch', 1), ('standhe', 1), ('freenobody', 1), ('stopped', 1), ('himthough', 1), ('highthey', 1), ('tvso', 1), ('passed', 1), ('byi', 1), ('meant', 1), ('songmaybe', 1), ('harmoni', 1), ('refrainas', 1), ('signal', 1), ('changedhe', 1), ('wakes', 1), ('sweetshe', 1), ('downsidewalk', 1), ('markets', 1), ('locked', 1), ('tightand', 1), ('umbrellas', 1), ('bright', 1), ('backgroundtheres', 1), ('oil', 1), ('puddles', 1), ('taffeta', 1), ('patternsthat', 1), ('drainin', 1), ('arrangementsthat', 1), ('change', 1), ('stick', 1), ('foundmichael', 1), ('parkhe', 1), ('byyellow', 1), ('slickers', 1), ('swingslike', 1), ('puppets', 1), ('hanging', 1), ('skytheyll', 1), ('splash', 1), ('suppers', 1), ('wallpapered', 1), ('kitchenstheir', 1), ('mothers', 1), ('scoldbut', 1), ('youto', 1), ('cold', 1), ('sidewalks', 1), ('drymichael', 1), ('stairshe', 1), ('docats', 1), ('crying', 1), ('keyand', 1), ('dry', 1), ('towel', 1), ('twotheres', 1), ('windowtheres', 1), ('smiles', 1), ('wallyou', 1), ('allbut', 1), ('domichael', 1), ('gallery', 1), ('ladiesthen', 1), ('hang', 1), ('meyou', 1), ('studied', 1), ('portray', 1), ('mein', 1), ('greensand', 1), ('jeansand', 1), ('naked', 1), ('rosethen', 1), ('scenesthat', 1), ('work', 1), ('discloseslady', 1), ('yousomewhere', 1), ('magazinei', 1), ('page', 1), ('josephinewho', 1), ('without', 1), ('house', 1), ('repairi', 1), ('dust', 1), ('portraits', 1), ('dailyyour', 1), ('mail', 1), ('everywherethe', 1), ('writing', 1), ('ladieslady', 1), ('yearsthen', 1), ('weatherand', 1), ('herewhile', 1), ('pleasureand', 1), ('waylike', 1), ('homing', 1), ('pigeontheyve', 1), ('monitored', 1), ('brain', 1), ('sayand', 1), ('changed', 1), ('religionlady', 1), ('bedlady', 1), ('saidwell', 1), ('youwhen', 1), ('ladiesbut', 1), ('follow', 1), ('eyes', 1), ('look', 1), ('shady', 1), ('dawntrader', 1), ('peridots', 1), ('periwinkle', 1), ('medallionsgilded', 1), ('galleons', 1), ('spilled', 1), ('floortreasure', 1), ('whenever', 1), ('mind', 1), ('questions', 1), ('answer', 1), ('forthe', 1), ('wakethe', 1), ('riggering', 1), ('makesthe', 1), ('taste', 1), ('spray', 1), ('takesand', 1), ('givehe', 1), ('aches', 1), ('livehe', 1), ('stakes', 1), ('silveron', 1), ('freemermaids', 1), ('coloniesall', 1), ('mecity', 1), ('satins', 1), ('need', 1), ('mesomething', 1), ('truthful', 1), ('youleave', 1), ('behind', 1), ('mecome', 1), ('neon', 1), ('nightscome', 1), ('tourist', 1), ('sightsrun', 1), ('delights', 1), ('youyou', 1), ('hidesunlight', 1), ('renew', 1), ('prideskin', 1), ('white', 1), ('skin', 1), ('goldenlike', 1), ('freedolphins', 1), ('meseabird', 1), ('pilingsi', 1), ('smiling', 1), ('circles', 1), ('airi', 1), ('sit', 1), ('sleepingfold', 1), ('fleet', 1), ('sharea', 1), ('someonea', 1), ('donea', 1), ('seattheyll', 1), ('crazyand', 1), ('babylike', 1), ('freechildren', 1), ('heardwas', 1), ('outside', 1), ('traffic', 1), ('wordsit', 1), ('areeling', 1), ('christmas', 1), ('rapping', 1), ('pipes', 1), ('drumsoh', 1), ('wear', 1), ('comeswoke', 1), ('sawwas', 1), ('wallblue', 1), ('red', 1), ('gold', 1), ('crimson', 1), ('crystal', 1), ('beckonoh', 1), ('daytheres', 1), ('secondnow', 1), ('opens', 1), ('portrait', 1), ('todayand', 1), ('paved', 1), ('passersbyand', 1), ('pigeons', 1), ('flyand',

1), ('papers', 1), ('liewaiting', 1), ('blow', 1), ('awaywoke', 1), ('knewthere', 1), ('milk', 1), ('toast', 1), ('honey', 1), ('bowl', 1), ('oranges', 1), ('tooand', 1), ('poured', 1), ('butterscotch', 1), ('stuck', 1), ('sense', 1), ('present', 1), ('tenseswhen', 1), ('closes', 1), ('runs', 1), ('incense', 1), ('owls', 1), ('nightby', 1), ('candlelightby', 1), ('jewellight', 1), ('staypretty', 1), ('baby', 1), ('youwake', 1), ('tin', 1), ('varnished', 1), ('weeds', 1), ('jarstarnished', 1), ('tapestrieskept', 1), ('satin', 1), ('boxes', 1), ('arereflections', 1), ('memoriesletters', 1), ('searoses', 1), ('dipped', 1), ('sealing', 1), ('waxvalentines', 1), ('maple', 1), ('leavestucked', 1), ('paperbackguess', 1), ('throw', 1), ('todaydark', 1), ('darker', 1), ('moods', 1), ('henot', 1), ('golden', 1), ('prince', 1), ('comethrough', 1), ('columbines', 1), ('wizardryto', 1), ('sunstill', 1), ('chance', 1), ('seei', 1), ('todaytheres', 1), ('eyeslike', 1), ('made', 1), ('tinwhat', 1), ('happen', 1), ('tryto', 1), ('himin', 1), ('bleecker', 1), ('cafei', 1), ('todayi', 1), ('conversation', 1), ('findi', 1), ('apples', 1), ('cheeseshe', 1), ('playhe', 1), ('pleasesi', 1), ('cafesand', 1), ('helloand', 1), ('knowshow', 1), ('himshe', 1), ('removes', 1), ('ringto', 1), ('handsshe', 1), ('friendsi', 1), ('himsecrets', 1), ('sodathats', 1), ('beganlove', 1), ('story', 1), ('friendits', 1), ('second', 1), ('handbut', 1), ('listen', 1), ('questionsill', 1), ('answers', 1), ('foundhe', 1), ('guessingbut', 1), ('downshe', 1), ('speaks', 1), ('sorry', 1), ('sentencesmiraculous', 1), ('repentancesi', 1), ('hertomorrow', 1), ('speak', 1), ('endlessly', 1), ('whywhy', 1), ('herhe', 1), ('findhe', 1), ('tree', 1), ('sailingin', 1), ('decade', 1), ('dreamsand', 1), ('schoonerand', 1), ('treats', 1), ('queenbearing', 1), ('californiawith', 1), ('amber', 1), ('greenhe', 1), ('harborhe', 1), ('kissed', 1), ('freedomhe', 1), ('starboardin', 1), ('breathingof', 1), ('weedswhile', 1), ('mountainand', 1), ('calling', 1), ('nameand', 1), ('hopes', 1), ('hearthree', 1), ('thousand', 1), ('miles', 1), ('againhe', 1), ('himhe', 1), ('samehe', 1), ('missed', 1), ('forestwhile', 1), ('showed', 1), ('branches', 1), ('sang', 1), ('chorusas', 1), ('scaly', 1), ('towersof', 1), ('sent', 1), ('replyhe', 1), ('travels', 1), ('goodbyehe', 1), ('wish', 1), ('tryhe', 1), ('officewith', 1), ('papersthrough', 1), ('profitshe', 1), ('herfrom', 1), ('memoryand', 1), ('thinks', 1), ('alltheres', 1), ('thinking', 1), ('hertheres', 1), ('callstheres', 1), ('letterswith', 1), ('facts', 1), ('figures', 1), ('scrawlshe', 1), ('sensessthey', 1), ('laughed', 1), ('laughternow', 1), ('rallies', 1), ('defensesfor', 1), ('herfor', 1), ('eternityand', 1), ('sends', 1), ('medalshe', 1), ('bleeding', 1), ('wartheres', 1), ('jouster', 1), ('jesterand', 1), ('owns', 1), ('storetheres', 1), ('drummer', 1), ('dreamerand', 1), ('moreshe', 1), ('themthey', 1), ('followand', 1), ('means', 1), ('themand', 1), ('hollowlike', 1), ('freelike', 1), ('treebeing', 1), ('windless', 1), ('wells', 1), ('wonderby', 1), ('throbbing', 1), ('machinein', 1), ('trance', 1), ('underorders', 1), ('queensongs', 1), ('onepeople', 1), ('hurry', 1), ('quicklydont', 1), ('melodiesin', 1), ('chiming', 1), ('clickingand', 1), ('harmoniesongs', 1), ('onesome', 1), ('dyingcrows', 1), ('ravens', 1), ('whistlinglines', 1), ('weeping', 1), ('cryingso', 1), ('listeningongs', 1), ('onedoes', 1), ('moon', 1), ('silverwhen', 1), ('strums', 1), ('galaxydying', 1), ('theirperfumed', 1), ('rhapsodies', 1), ('mesongs', 1), ('comethis', 1), ('againmy', 1), ('johnny', 1), ('fightin', 1), ('friendwhat', 1), ('thisto', 1), ('handshake', 1), ('fi stand', 1), ('againoh', 1), ('america', 1), ('againyou', 1), ('fighting', 1), ('friendwe', 1), ('beating', 1), ('cranethe', 1), ('anchored', 1), ('wednesdayand', 1), ('wonderto', 1), ('sayshe', 1), ('dances', 1), ('sailorsin', 1), ('smoky', 1), ('cabaret', 1), ('bar', 1), ('undergrounddown', 1), ('cellar', 1), ('towni', 1), ('treasureand', 1), ('reap', 1), ('fullest', 1), ('bo untyhe', 1), ('pleasureso', 1), ('winds', 1), ('morningblew', 1), ('gonei', 1), ('unfurling', 1), ('thursday', 1), ('dawnthe', 1), ('sink', 1), ('kisshe

ll', 1), ('steal', 1), ('sail', 1), ('awaythe', 1), ('dancersaturday', 1), ('cove', 1), ('forestpenancehell', 1), ('drowning', 1), ('flotsam', 1), ('broken', 1), ('baythe', 1), ('dancerdancing', 1), ('wakened', 1), ('dawn', 1), ('birds', 1), ('choruspenancehe', 1), ('heryes', 1), ('saturdaythe', 1), ('dancerask', 1), ('crane', 1), ('uspenancei', 1), ('theni', 1), ('didnt', 1), ('sayingit', 1), ('isnt', 1), ('true', 1), ('hardly', 1), ('himis', 1), ('game', 1), ('playinggo', 1), ('dancer', 1), ('lastthe', 1), ('drew', 1), ('herhe', 1), ('hasnt', 1), ('spring', 1), ('yearthere', 1), ('wouldbring', 1), ('silks', 1), ('sandalwood', 1), ('persian', 1), ('lacethe', 1), ('dancerthe', 1), ('goodbye', 1), ('gopenanceand', 1), ('closeand', 1), ('sailing', 1), ('fireplacethe', 1), ('dancerall', 1), ('dance', 1), ('sorrowpenancei', 1), ('wasnowhere', 1), ('mentioned', 1), ('murder', 1), ('placethe', 1), ('dancersearching', 1), ('comeand', 1), ('knowpenancego', 1), ('dancershe', 1), ('morgantownthe', 1), ('merchants', 1), ('awnings', 1), ('downthe', 1), ('milkltrucks', 1), ('roundsin', 1), ('morgantownwell', 1), ('sunto', 1), ('ride', 1), ('bus', 1), ('yawningand', 1), ('youngin', 1), ('samewell', 1), ('table', 1), ('shadeand', 1), ('sip', 1), ('lemonadeand', 1), ('paradein', 1), ('morgantownladies', 1), ('fashionscolored', 1), ('lights', 1), ('flashingwell', 1), ('wink', 1), ('total', 1), ('strangers', 1), ('passingin', 1), ('saimeid', 1), ('buy', 1), ('everythinga', 1), ('wooden', 1), ('bird', 1), ('painted', 1), ('wingsa', 1), ('ringsin', 1), ('morgantownbut', 1), ('givetto', 1), ('smile', 1), ('withare', 1), ('mornings', 1), ('livein', 1), ('franeer', 1), ('hired', 1), ('coach', 1), ('confusion', 1), ('planeand', 1), ('though', 1), ('shared', 1), ('space', 1), ('meet', 1), ('againthe', 1), ('drive', 1), ('eyebrows', 1), ('furrowed', 1), ('rearview', 1), ('mirrori', 1), ('read', 1), ('plainly', 1), ('written', 1), ('franeeri', 1), ('hurrybut', 1), ('crawled', 1), ('slowlythru', 1), ('buyers', 1), ('sellersthu', 1), ('burglar', 1), ('wishing', 1), ('wellswith', 1), ('gangs', 1), ('girly', 1), ('showsthe', 1), ('growsthe', 1), ('buses', 1), ('bustled', 1), ('bedlam', 1), ('dayi', 1), ('windowglass', 1), ('grumbled', 1), ('greyi', 1), ('cripple', 1), ('selling', 1), ('superman', 1), ('balloonsthe', 1), ('grated', 1), ('chromeplatethe', 1), ('clock', 1), ('struck', 1), ('slowly', 1), ('halfpastnoonthru', 1), ('tunnel', 1), ('tiled', 1), ('turninginto', 1), ('escapingonce', 1), ('goodbyeto', 1), ('symphonies', 1), ('treeswith', 1), ('parks', 1), ('plastic', 1), ('clothesthe', 1), ('growshe', 1), ('morehe', 1), ('cursed', 1), ('facehe', 1), ('hated', 1), ('rideand', 1), ('share', 1), ('spacei', 1), ('picked', 1), ('bags', 1), ('curband', 1), ('stumbled', 1), ('dooranother', 1), ('handanother', 1), ('moreand', 1), ('filled', 1), ('silverand', 1), ('fingers', 1), ('countingand', 1), ('sky', 1), ('goes', 1), ('foreverwithout', 1), ('meter', 1), ('maids', 1), ('paradesyou', 1), ('feed', 1), ('woesthe', 1), ('grows', 1), ('understand', 1), ('upon', 1), ('path', 1), ('behindtoday', 1), ('prey', 1), ('uncertaintythe', 1), ('shadow', 1), ('trembles', 1), ('wrath', 1), ('robbed', 1), ('blackness', 1), ('blindand', 1), ('tasted', 1), ('sunlight', 1), ('clear', 1), ('mei', 1), ('sandnow', 1), ('hills', 1), ('steeple', 1), ('chimebelow', 1), ('sleepy', 1), ('rooftops', 1), ('harborits', 1), ('thirsty', 1), ('fill', 1), ('friendship', 1), ('wineforgetting', 1), ('disregarding', 1), ('heroh', 1), ('sandsometime', 1), ('againback', 1), ('pathway', 1), ('troubled', 1), ('mindwhen', 1), ('forests', 1), ('traveler', 1), ('saneill', 1), ('challenge', 1), ('flashes', 1), ('brighter', 1), ('timeoh', 1), ('sand', 1), ('fair', 1), ('yearand', 1), ('ruby', 1), ('mans', 1), ('earyou', 1), ('horses', 1), ('guitar', 1), ('stringsyou', 1), ('devil', 1), ('wingsyou', 1), ('grand', 1), ('wingsdo', 1), ('tape', 1), ('shoulders', 1), ('singcan', 1), ('flyi', 1), ('flylike', 1), ('eagle', 1), ('doin', 1), ('hunting', 1), ('skyi', 1), ('followed', 1), ('sideshows', 1), ('trailer', 1), ('camping', 1), ('groundsyoun', 1), ('betting', 1), ('lover', 1), ('shaking', 1), ('diceand', 1), ('cheating', 1),

('twice', 1), ('twicei', 1), ('bid', 1), ('twicewere', 1), ('wondering', 1), ('gamble', 1), ('worth', 1), ('pricepack', 1), ('ini', 1), ('pack', 1), ('in was', 1), ('fold', 1), ('could', 1), ('winso', 1), ('hiding', 1), ('newsan d', 1), ('races', 1), ('ticket', 1), ('stubs', 1), ('bluesand', 1), ('number s', 1), ('mentions', 1), ('mineand', 1), ('working', 1), ('overtime', 1), ('overtimeive', 1), ('fire', 1), ('overtimealways', 1), ('playin', 1), ('dim eslowin', 1), ('gettin', 1), ('tiredslowin', 1), ('downand', 1), ('envy', 1), ('valley', 1), ('foundcause', 1), ('midwayslowin', 1), ('seagull', 1), ('silly', 1), ('seabirdno', 1), ('possess', 1), ('youno', 1), ('blame', 1), ('youfor', 1), ('wingsmy', 1), ('relationshave', 1), ('names', 1), ('must', 1), ('mefor', 1), ('freedomof', 1), ('thingsmy', 1), ('lived', 1), ('crusoeo n', 1), ('island', 1), ('noisein', 1), ('cobblestone', 1), ('seaand', 1), ('beaches', 1), ('concreteand', 1), ('billand', 1), ('blossoms', 1), ('hun g', 1), ('falseon', 1), ('store', 1), ('treesmy', 1), ('cryout', 1), ('seasi deto', 1), ('shouldersand', 1), ('hairbut', 1), ('sandcastles', 1), ('crumbl eand', 1), ('hunger', 1), ('humanand', 1), ('humans', 1), ('hungryfor', 1), ('worlds', 1), ('sharemy', 1), ('seagullwho', 1), ('dives', 1), ('watersan d', 1), ('catches', 1), ('silverfinedinner', 1), ('alonecrying', 1), ('footp rintsthat', 1), ('danced', 1), ('beachesand', 1), ('hands', 1), ('cast', 1), ('wishesthat', 1), ('sunk', 1), ('stonemy', 1)]

Lyrics Data – Kanye West:

There are 8496 tokens in the data.

There are 3270 unique tokens in the data.

There are 46238 characters in the data.

The lexical diversity is 0.385 in the data.

The 8496 most common tokens are: [('get', 153), ('like', 135), ('im', 106), ('know', 89), ('aint', 83), ('got', 78), ('dont', 73), ('gonna', 67), ('go', 56), ('well', 49), ('say', 46), ('gon', 41), ('man', 41), ('gotta', 40), ('c ant', 38), ('back', 36), ('one', 35), ('said', 33), ('still', 33), ('shit', 33), ('girl', 32), ('want', 32), ('em', 32), ('take', 31), ('right', 30), ('niggas', 30), ('tell', 29), ('thats', 29), ('sky', 29), ('pull', 28), ('pe ople', 27), ('kanye', 26), ('work', 26), ('wanna', 26), ('ya', 25), ('wont', 25), ('falls', 25), ('nigga', 24), ('money', 23), ('west', 23), ('leave', 23), ('day', 23), ('put', 23), ('oh', 23), ('cause', 23), ('make', 22), ('loo k', 22), ('come', 22), ('need', 21), ('really', 21), ('thing', 20), ('tryn a', 20), ('head', 19), ('told', 19), ('yall', 19), ('keep', 19), ('feel', 19), ('way', 19), ('think', 19), ('ass', 18), ('fuck', 18), ('see', 18), ('ye ah', 18), ('time', 18), ('even', 18), ('never', 18), ('went', 17), ('workou t', 17), ('beat', 17), ('made', 17), ('walk', 17), ('give', 16), ('new', 16), ('ill', 16), ('everybody', 16), ('things', 16), ('high', 16), ('words', 16), ('gold', 15), ('baby', 15), ('kids', 15), ('yo', 15), ('wasnt', 15), ('youre', 15), ('breathe', 15), ('party', 15), ('world', 15), ('hands', 15), ('touch', 14), ('buy', 14), ('could', 14), ('set', 14), ('ever', 14), ('le t', 14), ('id', 14), ('school', 14), ('lookin', 13), ('care', 13), ('love', 13), ('plan', 13), ('call', 13), ('mean', 13), ('til', 13), ('downi', 12), ('talk', 12), ('played', 12), ('play', 12), ('two', 12), ('rap', 12), ('prom ised', 12), ('step', 12), ('hear', 11), ('first', 11), ('stop', 11), ('us', 11), ('throw', 11), ('ask', 11), ('beats', 11), ('uh', 11), ('working', 11), ('show', 11), ('die', 11), ('spaceship', 11), ('nowhere', 11), ('tellin', 11), ('thingsall', 11), ('sayin', 10), ('girls', 10), ('heard', 10), ('littl e', 10), ('god', 10), ('name', 10), ('boy', 10), ('hes', 10), ('wanted', 10), ('nothin', 10), ('downget', 10), ('family', 10), ('smoke', 9), ('least', 9), ('whats', 9), ('good', 9), ('came', 9), ('kanyes', 9), ('every', 9), ('l ets', 9), ('glasses', 9), ('raise', 9), ('jay', 9), ('life', 9), ('spit', 9), ('ima', 9), ('fly', 9), ('fast', 9), ('whatever', 9), ('youll', 9), ('ex

tra', 9), ('act', 9), ('mr', 9), ('definitely', 9), ('messin', 8), ('broke', 8), ('seen', 8), ('car', 8), ('without', 8), ('nothing', 8), ('break', 8), ('soul', 8), ('songs', 8), ('black', 8), ('ive', 8), ('hiphop', 8), ('talkin', 8), ('going', 8), ('glass', 8), ('took', 8), ('actually', 8), ('deal', 8), ('remember', 8), ('guess', 8), ('catch', 8), ('ready', 8), ('westi', 8), ('testifycome', 8), ('spot', 8), ('flyfore', 8), ('wish', 8), ('chi', 8), ('nothings', 8), ('tomorrow', 8), ('true', 8), ('rightare', 8), ('outif', 8), ('motherfuckin', 8), ('coming', 8), ('foxxshe', 7), ('better', 7), ('muc h', 7), ('around', 7), ('dude', 7), ('next', 7), ('tryin', 7), ('might', 7), ('ooh', 7), ('meand', 7), ('hit', 7), ('piece', 7), ('move', 7), ('last', 7), ('whole', 7), ('called', 7), ('use', 7), ('dame', 7), ('always', 7), ('t rying', 7), ('try', 7), ('something', 7), ('real', 7), ('song', 7), ('home', 7), ('heart', 7), ('didnt', 7), ('bout', 7), ('cool', 7), ('everything', 7), ('top', 7), ('away', 7), ('graveshiftand', 7), ('shiti', 7), ('flypast', 7), ('feet', 7), ('hot', 7), ('fire', 7), ('jamie', 6), ('friend', 6), ('diggerw hen', 6), ('four', 6), ('bad', 6), ('benz', 6), ('used', 6), ('supposed', 6), ('somethin', 6), ('westnow', 6), ('hell', 6), ('white', 6), ('able', 6), ('rapper', 6), ('maybe', 6), ('ok', 6), ('salad', 6), ('year', 6), ('nobod y', 6), ('please', 6), ('left', 6), ('started', 6), ('youso', 6), ('sign', 6), ('album', 6), ('capitol', 6), ('damn', 6), ('thought', 6), ('would', 6), ('boys', 6), ('enough', 6), ('past', 6), ('knew', 6), ('weed', 6), ('pass', 6), ('three', 6), ('find', 6), ('comes', 6), ('hate', 6), ('chance', 6), ('s tepgangstas', 6), ('pimps', 6), ('talkooh', 6), ('hecky', 6), ('nah', 6), ('pulled', 5), ('somebody', 5), ('mama', 5), ('drive', 5), ('found', 5), ('r oll', 5), ('sing', 5), ('light', 5), ('big', 5), ('phone', 5), ('plane', 5), ('favorite', 5), ('yes', 5), ('couldnt', 5), ('picked', 5), ('kid', 5), ('ga me', 5), ('already', 5), ('city', 5), ('face', 5), ('jesus', 5), ('walks', 5), ('save', 5), ('dawg', 5), ('accident', 5), ('slow', 5), ('dealin', 5), ('hard', 5), ('alive', 5), ('blow', 5), ('couldbuy', 5), ('job', 5), ('man y', 5), ('peace', 5), ('gostill', 5), ('hustlers', 5), ('johnson', 5), ('mus ic', 5), ('spirit', 5), ('downand', 5), ('club', 5), ('marvin', 5), ('luther', 5), ('wire', 5), ('hoses', 5), ('couple', 5), ('diamond', 5), ('shes', 4), ('headkanye', 4), ('met', 4), ('louis', 4), ('vuitton', 4), ('doctor', 4), ('michael', 4), ('eat', 4), ('week', 4), ('theres', 4), ('lil', 4), ('fo llow', 4), ('sit', 4), ('drop', 4), ('friends', 4), ('outwork', 4), ('mout h', 4), ('tracks', 4), ('kill', 4), ('brought', 4), ('listening', 4), ('rod e', 4), ('old', 4), ('berserkeat', 4), ('dessert', 4), ('deserve', 4), ('mot herfucking', 4), ('talking', 4), ('dick', 4), ('almost', 4), ('mos', 4), ('s traight', 4), ('mom', 4), ('push', 4), ('named', 4), ('thinking', 4), ('doi n', 4), ('change', 4), ('company', 4), ('days', 4), ('bed', 4), ('together', 4), ('another', 4), ('streets', 4), ('bring', 4), ('room', 4), ('ones', 4), ('sure', 4), ('young', 4), ('word', 4), ('g', 4), ('must', 4), ('lyin', 4), ('gun', 4), ('highim', 4), ('twentyfivejokes', 4), ('saywe', 4), ('looked', 4), ('cop', 4), ('girlgirl', 4), ('swear', 4), ('mine', 4), ('lot', 4), ('pr ay', 4), ('busters', 4), ('thatstill', 4), ('def', 4), ('pon', 4), ('degree s', 4), ('smart', 4), ('meone', 4), ('foundoh', 4), ('drink', 4), ('gaye', 4), ('anita', 4), ('wellshe', 4), ('minnie', 4), ('riperton', 4), ('loving', 4), ('gladly', 4), ('risk', 4), ('delta', 4), ('believe', 4), ('iced', 4), ('sleeves', 4), ('outpush', 4), ('truck', 4), ('keys', 4), ('outgirls', 4), ('wild', 4), ('ds', 4), ('outbreathe', 4), ('outlet', 4), ('fight', 4), ('we ave', 4), ('desert', 4), ('es', 4), ('outwhen', 4), ('fail', 4), ('across', 4), ('woo', 4), ('woos', 4), ('westoh', 4), ('rain', 4), ('walkgod', 4), ('d evils', 4), ('downjesus', 4), ('niggasshe', 3), ('niggasi', 3), ('rock', 3), ('fuckin', 3), ('best', 3), ('years', 3), ('eighteen', 3), ('tv', 3), ('pa y', 3), ('watch', 3), ('stay', 3), ('start', 3), ('graduation', 3), ('summe r', 3), ('ladies', 3), ('wants', 3), ('tight', 3), ('2', 3), ('3', 3), ('che

ck', 3), ('whos', 3), ('night', 3), ('weak', 3), ('yea', 3), ('bought', 3),
 ('cold', 3), ('mess', 3), ('6', 3), ('tape', 3), ('since', 3), ('paid', 3),
 ('free', 3), ('yeah', 3), ('workmove', 3), ('dessertget', 3), ('hope', 3),
 ('berserk', 3), ('run', 3), ('westand', 3), ('amheres', 3), ('rocand', 3),
 ('themheres', 3), ('rocafellaraise', 3), ('andheres', 3), ('rocthis', 3),
 ('alcohol', 3), ('rockefeller', 3), ('chain', 3), ('end', 3), ('college',
 3), ('air', 3), ('freeway', 3), ('kweli', 3), ('gas', 3), ('suits', 3), ('be
 hind', 3), ('miracle', 3), ('askin', 3), ('send', 3), ('um', 3), ('dynasty',
 3), ('kind', 3), ('track', 3), ('needed', 3), ('jayz', 3), ('verse', 3), ('t
 hisand', 3), ('signed', 3), ('gettin', 3), ('gucci', 3), ('later', 3), ('wat
 ching', 3), ('eyes', 3), ('single', 3), ('ars', 3), ('3h', 3), ('cam', 3),
 ('jacob', 3), ('25', 3), ('hundred', 3), ('plaques', 3), ('kinda', 3), ('loo
 king', 3), ('saying', 3), ('skykanye', 3), ('felt', 3), ('anything', 3),
 ('k', 3), ('sent', 3), ('bystack', 3), ('alivethrow', 3), ('dope', 3), ('cra
 ck', 3), ('highkids', 3), ('singwe', 3), ('second', 3), ('five', 3), ('clai
 m', 3), ('drug', 3), ('schooland', 3), ('cut', 3), ('#1girl', 3), ('six',
 3), ('video', 3), ('bet', 3), ('smoking', 3), ('rhymes', 3), ('kanyeive',
 3), ('ohive', 3), ('skyoh', 3), ('ye', 3), ('late', 3), ('easy', 3), ('snatc
 hed', 3), ('mecause', 3), ('pick', 3), ('todayand', 3), ('forever', 3), ('ho
 nest', 3), ('babeit', 3), ('hurts', 3), ('may', 3), ('aunt', 3), ('cigarette
 s', 3), ('cousin', 3), ('oohh', 3), ('dead', 3), ('boosters', 3), ('movetw
 o', 3), ('jack', 3), ('harlem', 3), ('chitown', 3), ('skit', 3), ('dad', 3),
 ('father', 3), ('getting', 3), ('part', 3), ('highest', 3), ('else', 3), ('v
 andross', 3), ('edition', 3), ('degree', 3), ('guys', 3), ('nearly', 3), ('k
 hanthrough', 3), ('limit', 3), ('wallfor', 3), ('allthrough', 3), ('mayfor',
 3), ('wayright', 3), ('understand', 3), ('safe', 3), ('christmas', 3), ('spe
 nd', 3), ('book', 3), ('outkanye', 3), ('thththrow', 3), ('handsget', 3),
 ('manget', 3), ('handskeep', 3), ('losin', 3), ('againkeep', 3), ('mei', 3),
 ('stage', 3), ('omega', 3), ('stepkappa', 3), ('sigma', 3), ('rawaka', 3),
 ('stepsg', 3), ('rho', 3), ('zeta', 3), ('downcmon', 3), ('glitters', 3),
 ('businessand', 3), ('methe', 3), ('digger', 2), ('bomb', 2), ('herkanye',
 2), ('crib', 2), ('bigger', 2), ('super', 2), ('shorty', 2), ('moneyshe',
 2), ('prenup', 2), ('birthday', 2), ('stick', 2), ('deray', 2), ('probably',
 2), ('nba', 2), ('carso', 2), ('blunt', 2), ('disrespect', 2), ('son', 2),
 ('makes', 2), ('meso', 2), ('meshe', 2), ('butaint', 2), ('aboutunless', 2),
 ('freaking', 2), ('outthen', 2), ('nowmaybe', 2), ('twista', 2), ('breath',
 2), ('homies', 2), ('fit', 2), ('player', 2), ('thank', 2), ('woooo', 2),
 ('outside', 2), ('double', 2), ('manand', 2), ('mall', 2), ('jackson', 2),
 ('clap', 2), ('making', 2), ('themr', 2), ('wallkanye', 2), ('fucking', 2),
 ('goatee', 2), ('paul', 2), ('herethey', 2), ('bougie', 2), ('buzz', 2), ('f
 unny', 2), ('feeling', 2), ('guy', 2), ('gap', 2), ('blame', 2), ('demo',
 2), ('motherfucker', 2), ('dream', 2), ('killers', 2), ('power', 2), ('liste
 n', 2), ('moreyou', 2), ('tour', 2), ('purse', 2), ('hand', 2), ('changed',
 2), ('balled', 2), ('turned', 2), ('american', 2), ('lyrical', 2), ('rocafel
 la', 2), ('truth', 2), ('producers', 2), ('rapi', 2), ('dunno', 2), ('produc
 er', 2), ('ghost', 2), ('production', 2), ('wouldnt', 2), ('poppin', 2), ('s
 tayin', 2), ('chicago', 2), ('apartment', 2), ('acts', 2), ('lights', 2),
 ('pelle', 2), ('layaway', 2), ('sample', 2), ('crazy', 2), ('drums', 2), ('w
 erent', 2), ('rapped', 2), ('sound', 2), ('fresh', 2), ('tracked', 2), ('mee
 t', 2), ('soulful', 2), ('studio', 2), ('uhaul', 2), ('drove', 2), ('jerse
 y', 2), ('beans', 2), ('baseline', 2), ('matter', 2), ('walked', 2), ('bucke
 t', 2), ('classics', 2), ('story', 2), ('john', 2), ('bunch', 2), ('line',
 2), ('wack', 2), ('okay', 2), ('blueprint', 2), ('built', 2), ('relationship
 s', 2), ('relationship', 2), ('happen', 2), ('theyd', 2), ('gimme', 2), ('he
 y', 2), ('fucked', 2), ('though', 2), ('heads', 2), ('thatll', 2), ('naw',
 2), ('mel', 2), ('stupid', 2), ('records', 2), ('wowi', 2), ('chorus', 2),

('front', 2), ('figured', 2), ('b', 2), ('rained', 2), ('clothes', 2), ('sky
 gotta', 2), ('westback', 2), ('hurt', 2), ('team', 2), ('momma', 2), ('stan
 d', 2), ('died', 2), ('hiahhighhighhightop', 2), ('worldalalalalahlalahtop',
 2), ('havin', 2), ('goin', 2), ('became', 2), ('sheets', 2), ('skyyou', 2),
 ('ear', 2), ('body', 2), ('layin', 2), ('westwe', 2), ('manwe', 2), ('hande
 d', 2), ('saythis', 2), ('finna', 2), ('hood', 2), ('dogs', 2), ('shine',
 2), ('bootlegged', 2), ('packs', 2), ('lottery', 2), ('tickets', 2), ('nigga
 s-drug', 2), ('12', 2), ('park', 2), ('hold', 2), ('bus', 2), ('live', 2),
 ('#1ay', 2), ('finally', 2), ('stuff', 2), ('itgirl', 2), ('tired', 2), ('co
 pped', 2), ('manager', 2), ('stole', 2), ('hits', 2), ('help', 2), ('himso',
 2), ('hustleyou', 2), ('blaowkanye', 2), ('glc', 2), ('grind', 2), ('state',
 2), ('mind', 2), ('alone', 2), ('half', 2), ('ball', 2), ('ughgotta', 2),
 ('reach', 2), ('pants', 2), ('police', 2), ('rightand', 2), ('finished', 2),
 ('cars', 2), ('turn', 2), ('flyi', 2), ('upand', 2), ('knows', 2), ('heave
 n', 2), ('wake', 2), ('wayso', 2), ('screen', 2), ('dreams', 2), ('adam',
 2), ('levineuh', 2), ('saynothings', 2), ('today', 2), ('todaybut', 2), ('wa
 yand', 2), ('nowkanye', 2), ('places', 2), ('folks', 2), ('shitcalm', 2),
 ('ghetto', 2), ('thisgame', 2), ('lock', 2), ('choirnow', 2), ('hoesthrow',
 2), ('chains', 2), ('rhyme', 2), ('meyou', 2), ('place', 2), ('heat', 2),
 ('guns', 2), ('backs', 2), ('boss', 2), ('freshman', 2), ('screamin', 2),
 ('jimmy', 2), ('glad', 2), ('moneybut', 2), ('image', 2), ('inside', 2), ('c
 harts', 2), ('hov', 2), ('blood', 2), ('feelin', 2), ('itbut', 2), ('done',
 2), ('land', 2), ('reason', 2), ('side', 2), ('lovethats', 2), ('full', 2),
 ('open', 2), ('getget', 2), ('roc', 2), ('bikes', 2), ('jordan', 2), ('lyric
 s', 2), ('hail', 2), ('morning', 2), ('foxx', 2), ('biggie', 2), ('fine',
 2), ('skinned', 2), ('beare', 2), ('welljamie', 2), ('robinson', 2), ('gree
 n', 2), ('faster', 2), ('sip', 2), ('rims', 2), ('bend', 2), ('finish', 2),
 ('books', 2), ('read', 2), ('mastersthen', 2), ('says', 2), ('jaw', 2), ('wi
 red', 2), ('rappin', 2), ('firei', 2), ('nowits', 2), ('shed', 2), ('apologi
 ze', 2), ('coke', 2), ('wrong', 2), ('death', 2), ('locked', 2), ('plate',
 2), ('classes', 2), ('bag', 2), ('doyou', 2), ('wit', 2), ('youyou', 2), ('h
 undreds', 2), ('bitch', 2), ('fix', 2), ('outand', 2), ('uhim', 2), ('westnn
 ow', 2), ('highall', 2), ('highnow', 2), ('highand', 2), ('highkanye', 2),
 ('usually', 2), ('picture', 2), ('arm', 2), ('filled', 2), ('rappers', 2),
 ('mcs', 2), ('record', 2), ('graduated', 2), ('wait', 2), ('rawi', 2), ('ear
 ly', 2), ('cuzi', 2), ('cuzwoothere', 2), ('wasi', 2), ('cuza', 2), ('cuzwo
 o', 2), ('woothere', 2), ('syleena', 2), ('downsouthside', 2), ('southsidew
 e', 2), ('rightim', 2), ('downwestside', 2), ('westsidewe', 2), ('rightoh',
 2), ('allkanye', 2), ('westman', 2), ('promise', 2), ('daddy', 2), ('long',
 2), ('nowim', 2), ('cmon', 2), ('downoh', 2), ('livin', 2), ('downcome', 2),
 ('onim', 2), ('jail', 2), ('cover', 2), ('house', 2), ('downim', 2), ('busin
 ess', 2), ('gone', 2), ('itand', 2), ('rings', 2), ('ringsthey', 2), ('fanc
 y', 2), ('thingsi', 2), ('weight', 2), ('goldall', 2), ('awaylet', 2), ('su
 n', 2), ('children', 2), ('dogyou', 2), ('storm', 2), ('thingsthey', 2), ('t
 hingall', 2), ('war', 2), ('pieces', 2), ('nowjesus', 2), ('walkand', 2),
 ('wongsjesus', 2), ('afraid', 2), ('spoke', 2), ('longjesus', 2), ('needyea
 h', 1), ('triflin', 1), ('indeedoh', 1), ('diggerway', 1), ('town', 1), ('di
 gs', 1), ('mekanye', 1), ('westcutie', 1), ('beauty', 1), ('salonwith', 1),
 ('underarmshe', 1), ('charmfar', 1), ('flocki', 1), ('charm', 1), ('armbut',
 1), ('hermy', 1), ('psychic', 1), ('shell', 1), ('serenatrina', 1), ('jennif
 er', 1), ('lopez', 1), ('kidsand', 1), ('asses', 1), ('showbizokay', 1), ('f
 riendsi', 1), ('inwe', 1), ('din', 1), ('payif', 1), ('paidyou', 1), ('herfr
 om', 1), ('bustamy', 1), ('usher', 1), ('none', 1), ('westeighteen', 1),
 ('yearsshe', 1), ('years', 1), ('payin', 1), ('child', 1), ('support', 1),
 ('kidshis', 1), ('hisyou', 1), ('given', 1), ('sundaywin', 1), ('bowl', 1),
 ('hyundaishe', 1), ('tyco', 1), ('lipo', 1), ('walkin', 1), ('moneyshouldv

e', 1), ('insured', 1), ('geico', 1), ('moneyif', 1), ('punkholla', 1), ('ye
 ahits', 1), ('havecause', 1), ('halfeighteen', 1), ('yearsand', 1), ('18th',
 1), ('hiskanye', 1), ('needsyou', 1), ('weedyou', 1), ('leavetheres', 1),
 ('dishes', 1), ('sleevesbut', 1), ('washin', 1), ('himhe', 1), ('atsunhe',
 1), ('ambition', 1), ('eyesthis', 1), ('moppin', 1), ('floors', 1), ('friess
 o', 1), ('sidei', 1), ('dudes', 1), ('ballin', 1), ('niceand', 1), ('calli
 n', 1), ('girland', 1), ('girlkanye', 1), ('foxxi', 1), ('westlet', 1), ('in
 tro', 1), ('daviskanye', 1), ('minuteme', 1), ('faculty', 1), ('members',
 1), ('wonderincould', 1), ('somsomethin', 1), ('beautiful', 1), ('itthats',
 1), ('jumpin', 1), ('sharin', 1), ('candy', 1), ('stuffthink', 1), ('poppe
 d', 1), ('westget', 1), ('tapeand', 1), ('instructions', 1), ('exactlyyou',
 1), ('playerman', 1), ('stomachnobody', 1), ('ass1', 1), ('4', 1), ('ups',
 1), ('andtuck', 1), ('tummy', 1), ('crunches', 1), ('thisgive', 1), ('weaved
 ont', 1), ('weedpick', 1), ('seedits', 1), ('tonight', 1), ('excitedtell',
 1), ('invited', 1), ('dickah', 1), ('scary', 1), ('ishenny', 1), ('halle',
 1), ('berry', 1), ('excuse', 1), ('miss', 1), ('forgot', 1), ('namethank',
 1), ('bless', 1), ('cameits', 1), ('outoh', 1), ('silhouette', 1), ('cigaret
 temy', 1), ('jigga', 1), ('wetoo', 1), ('harshcover', 1), ('sarsoff', 1),
 ('carsstill', 1), ('16', 1), ('barswe', 1), ('sweating', 1), ('oldies', 1),
 ('juking', 1), ('beatmaybe', 1), ('boneso', 1), ('brag', 1), ('nowbut', 1),
 ('girl1', 1), ('hop', 1), ('benz4door', 1), ('difference', 1), ('5', 1),
 ('7', 1), ('8all', 1), ('mocha', 1), ('lattes', 1), ('pilatesyou', 1), ('po
 p', 1), ('datinghustlers', 1), ('gangstas', 1), ('ballasits', 1), ('outhi',
 1), ('jill', 1), ('thanks', 1), ('iwas', 1), ('shop', 1), ('rodeodrive', 1),
 ('woooooomy', 1), ('lasandra', 1), ('tokanyes', 1), ('bill', 1), ('soundsand',
 1), ('13s', 1), ('cavalier', 1), ('trip', 1), ('tocancun', 1), ('importantl
 y', 1), ('withrayrays', 1), ('morework', 1), ('morejuke', 1), ('pump', 1),
 ('morechitown', 1), ('moremy', 1), ('ellamay', 1), ('mobile', 1), ('alabam
 a', 1), ('sincelistening', 1), ('date', 1), ('familyi', 1), ('wide', 1), ('p
 lanethanks', 1), ('planim', 1), ('envy', 1), ('friendssee', 1), ('baller',
 1), ('againlemme', 1), ('ofmy', 1), ('ohi', 1), ('outyeah', 1), ('outits',
 1), ('shitthats', 1), ('deserveits', 1), ('plani', 1), ('gentlemenits', 1),
 ('allow', 1), ('introduce', 1), ('myselfhis', 1), ('thisthats', 1), ('stopit
 s', 1), ('breathei', 1), ('tiredits', 1), ('yahis', 1), ('okdouble', 1), ('t
 imethats', 1), ('woomove', 1), ('oweat', 1), ('ughget', 1), ('deservethats',
 1), ('ughthats', 1), ('woook', 1), ('thatd', 1), ('goodi', 1), ('appreciat
 e', 1), ('timei', 1), ('mewhoa', 1), ('whoa', 1), ('yeahdododododododododod
 o', 1), ('jayzaight', 1), ('ityo', 1), ('foremostfor', 1), ('muhfuckerhad',
 1), ('roomcause', 1), ('fuckingkanye', 1), ('westid', 1), ('propose', 1),
 ('toasti', 1), ('toast', 1), ('motherfuckerkanye', 1), ('westthe', 1), ('dig
 ital', 1), ('underground', 1), ('pacthe', 1), ('rudolph', 1), ('rednosed',
 1), ('reindeer', 1), ('roci', 1), ('15', 1), ('seconds', 1), ('fameand', 1),
 ('gameaint', 1), ('expect', 1), ('topthey', 1), ('expected', 1), ('dropout',
 1), ('flopthen', 1), ('savin', 1), ('himselfrocafellas', 1), ('helpedmy',
 1), ('thinner', 1), ('sean', 1), ('pauls', 1), ('hairnow', 1), ('jean', 1),
 ('gaultier', 1), ('cologne', 1), ('fill', 1), ('bigheadedwould', 1), ('isflo
 w', 1), ('infectious', 1), ('10', 1), ('secondsill', 1), ('insects', 1), ('t
 exasits', 1), ('interestedtil', 1), ('killed', 1), ('lexuskanye', 1), ('over
 looked', 1), ('sirnow', 1), ('overbooked', 1), ('sirthrough', 1), ('fans',
 1), ('tribe', 1), ('questbut', 1), ('westthatll', 1), ('defcall', 1), ('kwal
 i', 1), ('jayzim', 1), ('banana', 1), ('republic', 1), ('navy', 1), ('ooohi
 t', 1), ('sweeter', 1), ('sadienice', 1), ('bunb', 1), ('source', 1), ('awar
 dsgirl', 1), ('coulda', 1), ('horse', 1), ('awardsand', 1), ('famous', 1),
 ('kanyeim', 1), ('raymondsome', 1), ('arrogant', 1), ('himit', 1), ('embarra
 ssing', 1), ('himlast', 1), ('shoppin', 1), ('shineevery', 1), ('rhymenow',
 1), ('selfesteemor', 1), ('arrogance', 1), ('steam', 1), ('dreamsi', 1), ('g

assedbut', 1), ('ought', 1), ('laughso', 1), ('desk', 1), ('wear', 1), ('dress', 1), ('morei', 1), ('rocked', 1), ('20000', 1), ('niggaim', 1), ('kon', 1), ('donbought', 1), ('momi', 1), ('dealt', 1), ('cardsi', 1), ('prayed', 1), ('skies', 1), ('starsi', 1), ('mall', 1), ('hardoh', 1), ('cardi', 1), ('replied', 1), ('yesbut', 1), ('prefer', 1), ('term', 1), ('african', 1), ('expressbrains', 1), ('muscle', 1), ('puffy', 1), ('russellyour', 1), ('hustle', 1), ('tokillin', 1), ('shitmayonnaise', 1), ('whipskanye', 1), ('niggasmr', 1), ('wallso', 1), ('ar', 1), ('beanie', 1), ('session', 1), ('spittin', 1), ('okayuhh', 1), ('forthjust', 1), ('managing', 1), ('blaze', 1), ('withso', 1), ('youand', 1), ('mentor', 1), ('rapno', 1), ('idyo', 1), ('gassin', 1), ('manage', 1), ('producerhiphopill', 1), ('rapperim', 1), ('dot', 1), ('also', 1), ('figuring', 1), ('local', 1), ('jordans', 1), ('technomarine', 1), ('wore', 1), ('theni', 1), ('sped', 1), ('harold', 1), ('melvin', 1), ('hip', 1), ('compilation', 1), ('soundin', 1), ('dre', 1), ('chronic', 1), ('2001', 1), ('bit', 1), ('xxplosive', 1), ('spedup', 1), ('style', 1), ('lifeand', 1), ('resurgence', 1), ('management', 1), ('appreciated', 1), ('mans', 1), ('level', 1), ('dudejayzoh', 1), ('saidjayzcheck', 1), ('he retell', 1), ('simple', 1), ('type', 1), ('introspective', 1), ('complicated', 1), ('rhy', 1), ('personal', 1), ('opinion', 1), ('asked', 1), ('itjayzo', 1), ('train', 1), ('groups', 1), ('artists', 1), ('ended', 1), ('holdin', 1), ('news', 1), ('evicted', 1), ('packed', 1), ('ten', 1), ('landlord', 1), ('jerk', 1), ('mother', 1), ('tomothercome', 1), ('gonewark', 1), ('hadnt', 1), ('upmotherkanye', 1), ('herei', 1), ('unpacked', 1), ('ikea', 1), ('loaded', 1), ('equipment', 1), ('fact', 1), ('seven', 1), ('zone', 1), ('hat', 1), ('yesterday', 1), ('dmxhiphopno', 1), ('hereand', 1), ('faces', 1), ('ooooooooooooh', 1), ('kanyedameyo', 1), ('kanyeyou', 1), ('gave', 1), ('beat sdamejay', 1), ('gyou', 1), ('starstruck', 1), ('picturing', 1), ('superstars', 1), ('lenny', 1), ('sneakers', 1), ('tyty', 1), ('meneilly', 1), ('killin', 1), ('mayonnaise', 1), ('colored', 1), ('whips', 1), ('saw', 1), ('rest', 1), ('response', 1), ('tightjayzthat', 1), ('hotthat', 1), ('hehe', 1), ('forward', 1), ('h', 1), ('izzo', 1), ('proudly', 1), ('career', 1), ('course', 1), ('allowed', 1), ('appre', 1), ('riders', 1), ('postblueprint', 1), ('happened', 1), ('itd', 1), ('dave', 1), ('lighty', 1), ('labels', 1), ('je ssica', 1), ('rivera', 1), ('manjessicaman', 1), ('realim', 1), ('promotion', 1), ('upyall', 1), ('stupidlets', 1), ('person', 1), ('kicked', 1), ('joe', 1), ('badjoewe', 1), ('buddydame', 1), ('wackdameyou', 1), ('wackthen', 1), ('ahead', 1), ('guru', 1), ('thou', 1), ('11', 1), ('walls', 1), ('nowyou', 1), ('wowlike', 1), ('bite', 1), ('shitdameoh', 1), ('wacki', 1), ('hotdameits', 1), ('hotlike', 1), ('raps', 1), ('office', 1), ('brick', 1), ('brickdameyou', 1), ('brickyou', 1), ('umbrella', 1), ('table', 1), ('names', 1), ('producerrapper', 1), ('planned', 1), ('booking', 1), ('sessions', 1), ('arranging', 1), ('marketing', 1), ('schemes', 1), ('dealmelyo', 1), ('deal and', 1), ('interested', 1), ('price', 1), ('rocafellaso', 1), ('pink', 1), ('polos', 1), ('rocbefore', 1), ('popthe', 1), ('doors', 1), ('closedi', 1), ('street', 1), ('locksnow', 1), ('planme', 1), ('hopped', 1), ('vanany', 1), ('pessimists', 1), ('themplus', 1), ('apartmentlets', 1), ('clubleast', 1), ('hour', 1), ('linei', 1), ('dance', 1), ('houerafter', 1), ('advance', 1), ('shinejay', 1), ('due', 1), ('timenow', 1), ('ama', 1), ('legendi', 1), ('heavenkanye', 1), ('worldnow', 1), ('worldkanye', 1), ('rockback', 1), ('slic k', 1), ('rick', 1), ('popid', 1), ('itdamn', 1), ('loafers', 1), ('pocketbefore', 1), ('anybody', 1), ('beatsme', 1), ('split', 1), ('buffet', 1), ('kfc dog', 1), ('nervous', 1), ('breakdownslike', 1), ('mebaby', 1), ('aeroplane and', 1), ('againsure', 1), ('ticketsbut', 1), ('kick', 1), ('differentany', 1), ('cheated', 1), ('skeeted', 1), ('oncouldnt', 1), ('nia', 1), ('longim', 1), ('wrongsbut', 1), ('wrongs', 1), ('helped', 1), ('write', 1), ('songkanye', 1), ('skylupe', 1), ('fiascoyes', 1), ('thirdlupe', 1), ('steal', 1),

('lupin', 1), ('3rdhere', 1), ('beer', 1), ('curbpeachfuzz', 1), ('beard', 1), ('vergelets', 1), ('syrupbottleshaped', 1), ('mrs', 1), ('butterworthbut', 1), ('wordim', 1), ('block', 1), ('streetim', 1), ('mummrabut', 1), ('en', 1), ('gardeor', 1), ('touché', 1), ('lupe', 1), ('unthawedbut', 1), ('possessed', 1), ('chargecome', 1), ('correct', 1), ('porn', 1), ('starin', 1), ('pair', 1), ('steps', 1), ('foreign', 1), ('represent', 1), ('firstnow', 1), ('horns', 1), ('uhkanye', 1), ('babysky', 1), ('highsky', 1), ('highyea h', 1), ('rollinyeah', 1), ('feels', 1), ('babyfeels', 1), ('perfect', 1), ('singand', 1), ('thats-drug', 1), ('highwe', 1), ('sayif', 1), ('hearing', 1), ('thisyou', 1), ('experience', 1), ('grantedtook', 1), ('manbut', 1), ('manonly', 1), ('adult', 1), ('manflicking', 1), ('starter', 1), ('coats', 1), ('man-man', 1), ('wintercause', 1), ('finnasitting', 1), ('community', 1), ('collegesthis', 1), ('treys', 1), ('scholarshipcause', 1), ('tuition', 1), ('ambitionand', 1), ('loans', 1), ('sitting', 1), ('homesos', 1), ('forced', 1), ('sell', 1), ('jobyou', 1), ('growndrug', 1), ('nine', 1), ('hustlec ause', 1), ('655and', 1), ('selling', 1), ('makeup', 1), ('jacobsand', 1), ('tapes', 1), ('cake', 1), ('upwe', 1), ('backwe', 1), ('income', 1), ('taxw e', 1), ('paidand', 1), ('saymama', 1), ('southscratching', 1), ('houseroun d', 1), ('doe', 1), ('ran', 1), ('housecouldnt', 1), ('jobso', 1), ('hed', 1), ('workthe', 1), ('bulimic', 1), ('weightso', 1), ('homo', 1), ('straight but', 1), ('bakin', 1), ('cakeand', 1), ('yeahstack', 1), ('uhhuhthrow', 1), ('ughyou', 1), ('foolwhen', 1), ('programs', 1), ('dcfs', 1), ('dyslexicthe y', 1), ('50', 1), ('cent', 1), ('questionswe', 1), ('scream', 1), ('rocks', 1), ('smartwe', 1), ('retards', 1), ('teachers', 1), ('thoughthold', 1), ('c ashnow', 1), ('belong', 1), ('classits', 1), ('welfarethey', 1), ('space', 1), ('wheelchairim', 1), ('chrome', 1), ('wheels', 1), ('herelook', 1), ('f athers', 1), ('abandoned', 1), ('uswhen', 1), ('hammers', 1), ('ambulancesom etimes', 1), ('understands', 1), ('usbut', 1), ('#2hey', 1), ('#3hey', 1), ('doingirl', 1), ('#2well', 1), ('agirl', 1), ('workoutgirl', 1), ('#2got', 1), ('pack', 1), ('shakur', 1), ('stuffgirl', 1), ('#1dont', 1), ('#3my', 1), ('putting', 1), ('187', 1), ('pager', 1), ('workin', 1), ('moregirl', 1), ('hofesional', 1), ('shitis', 1), ('rollin', 1), ('lexus', 1), ('acura s', 1), ('#3girl', 1), ('bootleg', 1), ('queen', 1), ('toyou', 1), ('free99b oth', 1), ('#2', 1), ('#3free99girl', 1), ('#3thats', 1), ('usgirl', 1), ('# 2my', 1), ('godgirl', 1), ('#3oh', 1), ('glcive', 1), ('toobuy', 1), ('dopas t', 1), ('ohglc', 1), ('baby-i', 1), ('manive', 1), ('manoh', 1), ('ohkany e', 1), ('westif', 1), ('insults', 1), ('againi', 1), ('assaulting', 1), ('h imafter', 1), ('upthen', 1), ('shorten', 1), ('register', 1), ('uplets', 1), ('gaplook', 1), ('check-wasnt', 1), ('scratchso', 1), ('faultyeah', 1), ('st ole-never', 1), ('caughtthey', 1), ('pat', 1), ('measking', 1), ('khakisbu t', 1), ('ini', 1), ('token', 1), ('blackieoh', 1), ('kanyelets', 1), ('stor eso', 1), ('breaknext', 1), ('malltaking', 1), ('writing', 1), ('hitswritin g', 1), ('playing', 1), ('mindthis', 1), ('quit', 1), ('welcomeyall', 1), ('struggleyall', 1), ('match', 1), ('fathom', 1), ('dudelock', 1), ('roomdoi ng', 1), ('summersthats', 1), ('different', 1), ('cree', 1), ('summersi', 1), ('numbersthe', 1), ('deserves', 1), ('maybachso', 1), ('basementim', 1), ('waitin', 1), ('skytryna', 1), ('ughglcand', 1), ('jobrepresent', 1), ('mo b', 1), ('woothirsty', 1), ('mindlost', 1), ('lost', 1), ('awwlife', 1), ('m inewhy', 1), ('timeleave', 1), ('yallhalf', 1), ('mineonly', 1), ('fall', 1), ('ughgot', 1), ('tec9', 1), ('prime', 1), ('haters', 1), ('respect', 1), ('yeahin', 1), ('schedule', 1), ('9putting', 1), ('shelveswaiting', 1), ('pa tiently', 1), ('myselfwhere', 1), ('belife', 1), ('running', 1), ('streetsho ller', 1), ('beatput', 1), ('sounds', 1), ('sweetyes', 1), ('ol', 1), ('gsam e', 1), ('lowkey', 1), ('nopeholler', 1), ('whyd', 1), ('folkshope', 1), ('f reddie', 1), ('yusef', 1), ('grolly', 1), ('mesmoke', 1), ('count', 1), ('gs got', 1), ('counting', 1), ('rightshouldve', 1), ('niecethen', 1), ('gravesh

iftaww', 1), ('pressureand', 1), ('ohconsequencei', 1), ('member', 1), ('dollar', 1), ('cabcoming', 1), ('nightstanding', 1), ('daytryna', 1), ('coworkers', 1), ('likethis', 1), ('busta', 1), ('nightwell', 1), ('goesno', 1), ('broad', 1), ('service', 1), ('showsthat', 1), ('facultiesall', 1), ('naturally', 1), ('factuallyhad', 1), ('catastrophe', 1), ('empty', 1), ('fridge', 1), ('staring', 1), ('fairi', 1), ('thereso', 1), ('taking', 1), ('takes', 1), ('bowkanye', 1), ('ohtony', 1), ('williamsi', 1), ('chariot', 1), ('brother', 1), ('rideheaven', 1), ('westuh', 1), ('yeakanye', 1), ('todayfrom', 1), ('tim', 1), ('harder', 1), ('robert', 1), ('saybefore', 1), ('todaycan', 1), ('minimum', 1), ('wageand', 1), ('government', 1), ('administer', 1), ('aidsso', 1), ('minister', 1), ('sayallahu', 1), ('akbar', 1), ('carsthings', 1), ('oursbut', 1), ('farwhere', 1), ('starsbut', 1), ('seeing', 1), ('carsi', 1), ('bars', 1), ('-', 1), ('itkanye', 1), ('waykanye', 1), ('westthey', 1), ('seasonsand', 1), ('reasonand', 1), ('clapping', 1), ('squeezingand', 1), ('gran', 1), ('praying', 1), ('believingin', 1), ('himtil', 1), ('footsteps', 1), ('himthe', 1), ('devil', 1), ('breathingclaiming', 1), ('key', 1), ('dreamingand', 1), ('tease', 1), ('usmy', 1), ('pam', 1), ('downso', 1), ('smokin', 1), ('nowhis', 1), ('niggerish', 1), ('nowis', 1), ('skin', 1), ('blacker', 1), ('licorice', 1), ('nowi', 1), ('figure', 1), ('sick', 1), ('wayadam', 1), ('levinewith', 1), ('worthless', 1), ('far', 1), ('awayand', 1), ('lasts', 1), ('wayoohh', 1), ('playa', 1), ('mailits', 1), ('either', 1), ('jailstill', 1), ('gonow', 1), ('hoseeverybody', 1), ('gomos', 1), ('deftwo', 1), ('united', 1), ('states', 1), ('breakslow', 1), ('brow', 1), ('stakes', 1), ('folksbig', 1), ('macs', 1), ('fat', 1), ('ecstasy', 1), ('capsulespresidential', 1), ('scandals', 1), ('cock', 1), ('shitwe', 1), ('bk', 1), ('ny', 1), ('stuytoo', 1), ('harsh', 1), ('hungry', 1), ('whythese', 1), ('playheavy', 1), ('traffic', 1), ('lane', 1), ('cockwe', 1), ('movekanye', 1), ('goand', 1), ('boostershoes', 1), ('gokanye', 1), ('westtwo', 1), ('southside', 1), ('worldwidecause', 1), ('rep', 1), ('dieone', 1), ('neck', 1), ('waist', 1), ('gatsone', 1), ('wall', 1), ('twenty', 1), ('dues', 1), ('thati', 1), ('limelight', 1), ('micsgogetters', 1), ('shouldve', 1), ('twicemost', 1), ('imitated', 1), ('gramminominatedhotel', 1), ('accommodated', 1), ('cheerleader', 1), ('promdatedbarbershop', 1), ('playahated', 1), ('momandpop', 1), ('itfelt', 1), ('roof', 1), ('caved', 1), ('intwo', 1), ('raised', 1), ('crazyso', 1), ('mescreaming', 1), ('judgment', 1), ('melook', 1), ('basically', 1), ('racially', 1), ('profiledcuffed', 1), ('hosed', 1), ('pimped', 1), ('hod', 1), ('downplus', 1), ('downfrom', 1), ('bottom', 1), ('tops', 1), ('highoh', 1), ('gofreewaytwo', 1), ('letters', 1), ('arturn', 1), ('runnerslike', 1), ('jackie', 1), ('joyner', 1), ('sleep', 1), ('burnerthe', 1), ('skeet', 1), ('reef', 1), ('godtwo', 1), ('armsbreak', 1), ('necks', 1), ('steven', 1), ('seagalfree', 1), ('rocwith', 1), ('beef', 1), ('pot', 1), ('brawlforget', 1), ('squad', 1), ('fend', 1), ('yourselfhave', 1), ('law', 1), ('ughttwo', 1), ('slightly', 1), ('retardedfuck', 1), ('clip', 1), ('artist', 1), ('broadmos', 1), ('choirred', 1), ('blue', 1), ('upcalm', 1), ('backmotherfuckers', 1), ('asking', 1), ('thatyou', 1), ('almighty', 1), ('jackmos', 1), ('westghetto', 1), ('chestnorth', 1), ('south', 1), ('east', 1), ('westblack', 1), ('contesthigh', 1), ('manhey', 1), ('goingjimmyi', 1), ('goingmy', 1), ('degreesmy', 1), ('workbut', 1), ('kept', 1), ('learringall', 1), ('parents', 1), ('saving', 1), ('schoolhe', 1), ('greedy', 1), ('degreeand', 1), ('thesebecause', 1), ('regular', 1), ('homeless', 1), ('newspaper', 1), ('havethese', 1), ('documentedmy', 1), ('kidsim', 1), ('learn', 1), ('toomim', 1), ('smartest', 1), ('guywho', 1), ('downjayzyo', 1), ('chartsfirst', 1), ('heartrappers', 1), ('starti', 1), ('watched', 1), ('apartseen', 1), ('rovers', 1), ('ribbonsi', 1), ('repod', 1), ('resold', 1), ('redrivenso', 1), ('reload', 1), ('holds', 1), ('number', 1), ('positionwhen', 1), ('hotand', 1), ('mines', 1), ('sizzlingits', 1), ('plain', 1), ('lif

ethis', 1), ('godgiventhis', 1), ('hardlivin', 1), ('mixed', 1), ('cristal', 1), ('sippinits', 1), ('consistent', 1), ('hovgive', 1), ('disc', 1), ('andn igga', 1), ('visitinits', 1), ('tradition', 1), ('jeff', 1), ('gordon', 1), ('rapim', 1), ('pole', 1), ('position', 1), ('holla', 1), ('boywhen', 1), ('upkanye', 1), ('grandfather', 1), ('mamamade', 1), ('seat', 1), ('eatat', 1), ('tender', 1), ('age', 1), ('arrested', 1), ('sitinsand', 1), ('born', 1), ('differentnow', 1), ('ballots', 1), ('choose', 1), ('leadershipbut', 1), ('jacobs', 1), ('dealershipthats', 1), ('itracisms', 1), ('concealin', 1), ('clubthey', 1), ('sams', 1), ('clubi', 1), ('dirt', 1), ('church', 1), ('scrubbedswear', 1), ('baptized', 1), ('timesbut', 1), ('praise', 1), ('yuk ons', 1), ('paidits', 1), ('coupons', 1), ('savedlike', 1), ('doo', 1), ('ra gs', 1), ('wavesnothing', 1), ('sad', 1), ('passed', 1), ('awayso', 1), ('ra iney', 1), ('marry', 1), ('daughterand', 1), ('smiling', 1), ('heryou', 1), ('tears', 1), ('upbut', 1), ('complain', 1), ('eyecause', 1), ('eyefirst', 1), ('aalayah', 1), ('romeo', 1), ('diei', 1), ('angels', 1), ('sidewhen', 1), ('downj', 1), ('ivywe', 1), ('particular', 1), ('pathyou', 1), ('curricu lum', 1), ('mathcats', 1), ('delirious', 1), ('seriousthats', 1), ('expose', 1), ('globe', 1), ('worldim', 1), ('girlsim', 1), ('individualmy', 1), ('spi ritualbut', 1), ('hymns', 1), ('lyricalim', 1), ('heavensent', 1), ('instrum entmy', 1), ('rhythmatic', 1), ('regiment', 1), ('navigates', 1), ('melodi c', 1), ('notesfor', 1), ('mentalthats', 1), ('instrumental', 1), ('vibratio ns', 1), ('intoyeah', 1), ('loot', 1), ('rent', 1), ('daybut', 1), ('gives', 1), ('kunta', 1), ('kinteim', 1), ('cinquei', 1), ('hotdetermination', 1), ('dedication', 1), ('motivationim', 1), ('inspirationswhen', 1), ('selfdowni f', 1), ('cliff', 1), ('riffand', 1), ('slipped', 1), ('clinched', 1), ('life in', 1), ('grip', 1), ('foundlet', 1), ('known', 1), ('gods', 1), ('penmanshi p', 1), ('signedwith', 1), ('language', 1), ('deafand', 1), ('confined', 1), ('ears', 1), ('blindi', 1), ('color', 1), ('rhymeso', 1), ('housecause', 1), ('whenever', 1), ('moutha', 1), ('reigns', 1), ('outtake', 1), ('churchwhen', 1), ('downjayz', 1), ('westwho', 1), ('longoh', 1), ('finishedstarted', 1), ('strongthe', 1), ('buildingbuilt', 1), ('pebble', 1), ('peddled', 1), ('youp edaled', 1), ('nephews', 1), ('pedal', 1), ('theyre', 1), ('special', 1), ('w ooooolet', 1), ('fallingwell', 1), ('caught', 1), ('himcause', 1), ('fourth', 1), ('quarter', 1), ('mike', 1), ('themnumber', 1), ('albums', 1), ('themmor e', 1), ('eighth', 1), ('wonder', 1), ('wayclear', 1), ('stayyall', 1), ('chi tter', 1), ('chat', 1), ('thatthis', 1), ('dissing', 1), ('macedwhen', 1), ('spitting', 1), ('religioussix', 1), ('marys', 1), ('forgive', 1), ('usyoun g', 1), ('archbishop', 1), ('pope', 1), ('niggasthe', 1), ('jiggahovs', 1), ('living', 1), ('legend', 1), ('whyeverybody', 1), ('hovs', 1), ('overill', 1), ('awayto', 1), ('joy', 1), ('shall', 1), ('endill', 1), ('awayill', 1), ('gloryill', 1), ('awaywhen', 1), ('hallelujah', 1), ('bye', 1), ('byeill', 1), ('jamz', 1), ('foxxkanye', 1), ('manim', 1), ('knowand', 1), ('emphasis', 1), ('sometimeyou', 1), ('girlshe', 1), ('knowshe', 1), ('water', 1), ('thirs tyshe', 1), ('danced', 1), ('92', 1), ('backwith', 1), ('feelinlike', 1), ('m eare', 1), ('saiidsay', 1), ('oohhluther', 1), ('vandrossare', 1), ('besay', 1), ('bejamie', 1), ('foxxoh', 1), ('ohjamie', 1), ('wellkanye', 1), ('westi t', 1), ('willyou', 1), ('whipbring', 1), ('within', 1), ('whipthen', 1), ('s tripsee', 1), ('chickso', 1), ('onshe', 1), ('grabbing', 1), ('calling', 1), ('shyne', 1), ('homeman', 1), ('homeswhy', 1), ('lying', 1), ('thoughtellin g', 1), ('diamonds', 1), ('rhinestonesshe', 1), ('jacksongot', 1), ('dark', 1), ('jacksoni', 1), ('actionmy', 1), ('dog', 1), ('freak', 1), ('prove', 1), ('wrongim', 1), ('vandrossyoure', 1), ('offim', 1), ('gladys', 1), ('knightm e', 1), ('rightshe', 1), ('vandrossa', 1), ('editionsome', 1), ('foxxsmokey', 1), ('smokey', 1), ('robinsonfreddie', 1), ('wellsmokey', 1), ('ashford', 1), ('simpson', 1), ('beset', 1), ('rightyes', 1), ('wellal', 1), ('al', 1), ('oh h', 1), ('yeahooohwhatcha', 1), ('isley', 1), ('brothersaisha', 1), ('tyler',

1), ('babyits', 1), ('beautifulbut', 1), ('babycan', 1), ('fasterdamn', 1),
 ('twistatwistagot', 1), ('glist', 1), ('wristwhile', 1), ('laying', 1), ('can
 nabiswhen', 1), ('rocking', 1), ('rhythm', 1), ('lutheri', 1), ('messaging',
 1), ('kanman', 1), ('twistfrom', 1), ('besipping', 1), ('hennesseyplay', 1),
 ('rbwhile', 1), ('byou', 1), ('preferably', 1), ('gand', 1), ('wellcome', 1),
 ('evelyn', 1), ('champagneyou', 1), ('whispershit', 1), ('isleysthe', 1), ('m
 oving', 1), ('bump', 1), ('spinnerswhile', 1), ('bdipping', 1), ('streetsbump
 ing', 1), ('rband', 1), ('23sand', 1), ('wellwith', 1), ('earth', 1), ('win
 d', 1), ('firelet', 1), ('wet', 1), ('keith', 1), ('sweatput', 1), ('daze',
 1), ('mazefulfilling', 1), ('temptation', 1), ('jamming', 1), ('deep', 1),
 ('sexyou', 1), ('girlcome', 1), ('moment', 1), ('controllin', 1), ('holding',
 1), ('jodeciand', 1), ('assyou', 1), ('bumping', 1), ('teddy', 1), ('pendergr
 assid', 1), ('melody', 1), ('slownow', 1), ('lastno', 1), ('thug', 1), ('see
 i', 1), ('rb', 1), ('mechi', 1), ('ridejamie', 1), ('tracie', 1), ('spencerja
 mie', 1), ('rollingyou', 1), ('many', 1), ('associates', 1), ('okaythen',
 1), ('bachelors', 1), ('masters', 1), ('doctrineyou', 1), ('quityou', 1), ('m
 anwhen', 1), ('everyone', 1), ('workingyoure', 1), ('moneyyou', 1), ('lifeye
 a', 1), ('52', 1), ('wantbut', 1), ('makingmoney', 1), ('ways', 1), ('spendin
 g', 1), ('smartyou', 1), ('whycause', 1), ('buddy', 1), ('knowwhats', 1), ('w
 arm', 1), ('introlast', 1), ('october', 1), ('grammy', 1), ('fatal',
 1), ('fractured', 1), ('weeks', 1), ('recorded', 1), ('shut', 1), ('painyo',
 1), ('gee', 1), ('theycan', 1), ('hopchaka', 1), ('mantheres', 1), ('manid',
 1), ('lifeordeath', 1), ('situation', 1), ('manyall', 1), ('manits', 1), ('th
 e...chitown', 1), ('oni', 1), ('boost', 1), ('breakfast', 1), ('ensure', 1),
 ('dessertsombody', 1), ('ordered', 1), ('pancakes', 1), ('sizzurpthat', 1),
 ('sane', 1), ('berserknot', 1), ('worry', 1), ('htotheizzos', 1), ('wizzerkho
 w', 1), ('console', 1), ('supporttelling', 1), ('sons', 1), ('supportand', 1),
 ('imagine', 1), ('feelon', 1), ('scared', 1), ('emmett', 1), ('tillshe', 1),
 ('mineshe', 1), ('throwin', 1), ('signno', 1), ('signedtryin', 1), ('milliona
 ire', 1), ('lifelinesin', 1), ('hospital', 1), ('small', 1), ('diedthe', 1),
 ('clots', 1), ('jamaican', 1), ('manstory', 1), ('mtv', 1), ('bandi', 1), ('h
 istory', 1), ('makin', 1), ('manchaka', 1), ('nowif', 1), ('unclear', 1), ('m
 anthey', 1), ('shutfor', 1), ('like...', 1), ('weeksyouknow', 1), ('reconstructi
 ve', 1), ('surgery', 1), ('jawi', 1), ('mirrorand', 1), ('mani', 1), ('manth
 is', 1), ('dawgyeah', 1), ('uhwhat', 1), ('dealon', 1), ('hottest', 1), ('la
 bel', 1), ('aroundbut', 1), ('birdsit', 1), ('spoken', 1), ('cept', 1), ('pu
 ttin', 1), ('explained', 1), ('blacks', 1), ('gloryand', 1), ('gamegood',
 1), ('timein', 1), ('blink', 1), ('eye', 1), ('changedif', 1), ('feltyou',
 1), ('mase', 1), ('mason', 1), ('bethathank', 1), ('belti', 1), ('driver',
 1), ('suei', 1), ('lawyer', 1), ('case', 1), ('safemy', 1), ('dawks', 1),
 ('i...i', 1), ('tom', 1), ('cruise', 1), ('vanilla', 1), ('televisedtheres',
 1), ('geicothey', 1), ('burnt', 1), ('pepsi', 1), ('michaeli', 1), ('angel',
 1), ('missed', 1), ('assunbreakable', 1), ('glasslook', 1), ('pasttoys', 1),
 ('r', 1), ('cashand', 1), ('grow', 1), ('grownass', 1), ('kidswear', 1), ('d
 idbut', 1), ('champion', 1), ('tragedy', 1), ('triumphmake', 1), ('wirechak
 a', 1), ('fireyknow', 1), ('sayinwhen', 1), ('umthat', 1), ('chini', 1), ('r
 ealize', 1), ('jewelry', 1), ('onyou', 1), ('serious', 1), ('1', 1), ('davis
 now', 1), ('thatand', 1), ('mothers', 1), ('colleges', 1), ('wonderfullyou',
 1), ('goodand', 1), ('partying', 1), ('crazinessand', 1), ('forget', 1), ('h
 abit', 1), ('peersheyknow', 1), ('25thous', 1), ('yearand', 1), ('cocainebu
 t', 1), ('itll', 1), ('moneyno', 1), ('borrowing', 1), ('highso', 1), ('tatt
 ood', 1), ('excited', 1), ('itif', 1), ('continue', 1), ('several', 1), ('i
 nterviews', 1), ('godyoull', 1), ('entrylevel', 1), ('positionand', 1), ('th
 atif', 1), ('kiss', 1), ('assyoull', 1), ('levelwhich', 1), ('secretarys',
 1), ('secretaryand', 1), ('greatyou', 1), ('messages', 1), ('secretarywho',
 1), ('collegeshes', 1), ('bosss', 1), ('nieceso', 1), ('familyyou', 1), ('yo

uit', 1), ('manall', 1), ('classnot', 1), ('listenedoki', 1), ('hall', 1),
 ('monitorthis', 1), ('meant', 1), ('beyou', 1), ('tookextra', 1), ('classesn
 o', 1), ('sexbut', 1), ('whatmy', 1), ('keeps', 1), ('satisfiedwhen', 1),
 ('lady', 1), ('sayshey', 1), ('sexyi', 1), ('isbut', 1), ('add', 1), ('hohoh
 oi', 1), ('againwhere', 1), ('goddamn', 1), ('atalways', 1), ('carryin', 1),
 ('thatgoddamn', 1), ('4th', 1), ('grader', 1), ('downhe', 1), ('notin', 1),
 ('lifeyou', 1), ('isyou', 1), ('promisedwell', 1), ('youlook', 1), ('facedoe
 s', 1), ('anythingkanye', 1), ('niggais', 1), ('snorin', 1), ('classwake',
 1), ('westwake', 1), ('ludacrisyeah', 1), ('westgolly', 1), ('bullshit', 1),
 ('ice', 1), ('pologize', 1), ('probablybut', 1), ('goldif', 1), ('ghana',
 1), ('mali', 1), ('malifirst', 1), ('backpackice', 1), ('carti', 1), ('len
 s', 1), ('knapsackalways', 1), ('significantbut', 1), ('againand', 1), ('ben
 jaminsbig', 1), ('faced', 1), ('synonymsstrippers', 1), ('cinnamon', 1), ('c
 hips', 1), ('pentiumwhatcha', 1), ('trend', 1), ('isim', 1), ('stacks', 1),
 ('brokei', 1), ('cracki', 1), ('twelve', 1), ('platinum', 1), ('thatwhat',
 1), ('dogsing', 1), ('along', 1), ('yallludacrisyeah', 1), ('collegeand',
 1), ('dropped', 1), ('quick', 1), ('phd', 1), ('pretty', 1), ('huge', 1),
 ('dickladies', 1), ('ripped', 1), ('givin', 1), ('whale', 1), ('usin', 1),
 ('toothpickwell', 1), ('limited', 1), ('timeact', 1), ('action', 1), ('free9
 9later', 1), ('charge', 1), ('ménageheard', 1), ('flossbut', 1), ('toss',
 1), ('drawersand', 1), ('cost', 1), ('caddycuz', 1), ('grabbed', 1), ('gol
 f', 1), ('balls', 1), ('cluband', 1), ('actin', 1), ('calm', 1), ('denamuh
 h', 1), ('askedcan', 1), ('altima', 1), ('waswhile', 1), ('problems', 1),
 ('manbaby', 1), ('fully', 1), ('planwhile', 1), ('trickin', 1), ('rich', 1),
 ('niggagive', 1), ('piss', 1), ('offludacrisyeah', 1), ('lowclass', 1), ('ca
 shin', 1), ('dash', 1), ('askso', 1), ('popo', 1), ('laughright', 1), ('cras
 hwhatever', 1), ('prepared', 1), ('worstwhatever', 1), ('weaponpullin', 1),
 ('lexuses', 1), ('handso', 1), ('gses', 1), ('ambidextrouscoulda', 1), ('swo
 rn', 1), ('breasteses', 1), ('sendin', 1), ('messagesk', 1), ('mammogrami',
 1), ('handicamall', 1), ('legal', 1), ('amsterdamso', 1), ('candymanand',
 1), ('handymanbut', 1), ('leaveyou', 1), ('breatheludacrisyeah', 1), ('westu
 huh', 1), ('beatim', 1), ('beatkanye', 1), ('westnnnnnow', 1), ('flowis', 1),
 ('pocket', 1), ('wallets', 1), ('bounce', 1), ('hydraulicssi', 1), ('swerve',
 1), ('alcoholicsmy', 1), ('hella', 1), ('problemstil', 1), ('nerve', 1), ('c
 ollegemy', 1), ('teacher', 1), ('ise', 1), ('loser', 1), ('heartand', 1),
 ('stacksyou', 1), ('seeim', 1), ('bashfulbut', 1), ('bastards', 1), ('flow',
 1), ('bash', 1), ('skulland', 1), ('pastor', 1), ('troand', 1), ('droand',
 1), ('forwhy', 1), ('assholesthats', 1), ('forwardcause', 1), ('mokanye',
 1), ('sumeke', 1), ('raineyyouve', 1), ('mail', 1), ('nnnnnnnow', 1), ('thise
 mailing', 1), ('1126telling', 1), ('3626', 1), ('plus', 1), ('dyou', 1), ('p
 lanet', 1), ('bubblyat', 1), ('nyu', 1), ('kansaright', 1), ('lampin', 1),
 ('chillin', 1), ('campussent', 1), ('candicewho', 1), ('great', 1), ('franci
 swhit', 1), ('mamiyour', 1), ('saver', 1), ('tweet', 1), ('kweliyou', 1),
 ('talib', 1), ('ribi', 1), ('cd', 1), ('cribi', 1), ('lyinyo', 1), ('kwe',
 1), ('lineshe', 1), ('linesthen', 1), ('timeand', 1), ('owwwtalib', 1), ('kw
 eliyeah', 1), ('pickin', 1), ('dimesbut', 1), ('outwell', 1), ('twisted',
 1), ('assist', 1), ('charmoyo', 1), ('chick', 1), ('conference', 1), ('momsa
 nd', 1), ('sisters', 1), ('behavioralways', 1), ('hateranyways', 1), ('inter
 netor', 1), ('chicks', 1), ('birth', 1), ('control', 1), ('stuck', 1), ('nic
 oretteyou', 1), ('highif', 1), ('yetnow', 1), ('inconsideratei', 1), ('bubb
 a', 1), ('kush', 1), ('sister', 1), ('yeahcommonyo', 1), ('yoget', 1), ('noo
 n', 1), ('moonor', 1), ('hypefilled', 1), ('dopeyall', 1), ('assumed', 1),
 ('doomed', 1), ('tunebut', 1), ('notes', 1), ('quotesreal', 1), ('remotecont
 rol', 1), ('ofused', 1), ('abuse', 1), ('thugsrock', 1), ('clubs', 1), ('tig
 er', 1), ('woods', 1), ('hoodshould', 1), ('reality', 1), ('survivori', 1),
 ('liver', 1), ('youyous', 1), ('thicker', 1), ('youhow', 1), ('affect', 1),

('methey', 1), ('resurrect', 1), ('memarshas', 1), ('sexy', 1), ('thesethat
 s', 1), ('raw', 1), ('alicia', 1), ('keystoo', 1), ('featured', 1), ('popula
 rtwelve', 1), ('thousand', 1), ('spins', 1), ('coppin', 1), ('aalbum', 1),
 ('garbage', 1), ('ofthe', 1), ('clear', 1), ('looped', 1), ('uplabel', 1),
 ('souped', 1), ('sickwhen', 1), ('loose', 1), ('nutvideo', 1), ('medusaeve
 n', 1), ('booster', 1), ('chimped', 1), ('upwith', 1), ('pimp', 1), ('cup',
 1), ('illiterate', 1), ('infrared', 1), ('bred', 1), ('king', 1), ('simbabol
 der', 1), ('denver', 1), ('madd', 1), ('rapperjust', 1), ('mc', 1), ('temper
 you', 1), ('dancin', 1), ('honey', 1), ('industry', 1), ('crash', 1), ('surv
 ive', 1), ('kanyespittin', 1), ('wires', 1), ('fires', 1), ('retiringot',
 1), ('motherfuckers', 1), ('higher', 1), ('thenkanye', 1), ('commonnow',
 1), ('yeahall', 1), ('mmm', 1), ('uhuhnow', 1), ('uhuhand', 1), ('deraywha
 t', 1), ('ticket', 1), ('robe', 1), ('senioritis', 1), ('slide', 1), ('uplif
 ting', 1), ('yous', 1), ('nice', 1), ('shit-now', 1), ('jokes', 1), ('wave',
 1), ('goodbye', 1), ('everymotherfuckingbody', 1), ('campus', 1), ('kanye joh
 n', 1), ('legendim', 1), ('longer', 1), ('confused', 1), ('anybodyim', 1),
 ('rules', 1), ('anybodyi', 1), ('anybodymy', 1), ('anybodyshe', 1), ('everyb
 odyshe', 1), ('shoes', 1), ('motherfuckeralpha', 1), ('rawima', 1), ('mama
 aim', 1), ('downima', 1), ('lightskinned', 1), ('niggasnever', 1), ('styletol
 d', 1), ('businessthey', 1), ('decided', 1), ('finishedchasing', 1), ('youv
 e', 1), ('plannednow', 1), ('tannedback', 1), ('thereeverything', 1), ('year
 this', 1), ('classi', 1), ('cheesecake', 1), ('waiter', 1), ('therealpha',
 1), ('jones', 1), ('norah', 1), ('sororbring', 1), ('auroratammy', 1), ('bec
 ky', 1), ('laura', 1), ('shirleyim', 1), ('hurrysee', 1), ('pootie', 1), ('t
 ang', 1), ('tippy', 1), ('towrocafella', 1), ('stylerosary', 1), ('catholi
 c', 1), ('stylered', 1), ('kappa', 1), ('styleand', 1), ('pledgecrack', 1),
 ('steering', 1), ('wheel', 1), ('deadif', 1), ('breathingbitch', 1), ('reaso
 nalalpha', 1), ('wasa', 1), ('downyeah', 1), ('babyim', 1), ('downuh', 1), ('u
 poh', 1), ('selfconsciousshe', 1), ('idea', 1), ('collegethat', 1), ('majo
 r', 1), ('majored', 1), ('parentsl', 1), ('funnynow', 1), ('insecurethe',
 1), ('concept', 1), ('seems', 1), ('securesophomore', 1), ('carriershe', 1),
 ('haircause', 1), ('pairsof', 1), ('airs', 1), ('careshes', 1), ('precious',
 1), ('peer', 1), ('pressurecouldnt', 1), ('afford', 1), ('daughter', 1), ('a
 lexishe', 1), ('hair', 1), ('weavethen', 1), ('eveand', 1), ('issues', 1),
 ('believesingle', 1), ('female', 1), ('addicted', 1), ('retail', 1), ('wells
 yleena', 1), ('selfconsciousthats', 1), ('watchesrollies', 1), ('pashas',
 1), ('crazyi', 1), ('pronounce', 1), ('versayseethen', 1), ('spent', 1), ('b
 ucks', 1), ('thisjust', 1), ('grocery', 1), ('storewithout', 1), ('clean',
 1), ('shirt', 1), ('teamit', 1), ('seem', 1), ('dreambut', 1), ('lowest',
 1), ('selfesteemthe', 1), ('prettiest', 1), ('ugliest', 1), ('thingsfor',
 1), ('road', 1), ('riches', 1), ('ringswe', 1), ('floss', 1), ('degrade',
 1), ('uswe', 1), ('40', 1), ('acresand', 1), ('paper', 1), ('low', 1), ('sto
 opeven', 1), ('coupesyleena', 1), ('nowoh', 1), ('downkanye', 1), ('treat',
 1), ('emwe', 1), ('freedomwell', 1), ('emthings', 1), ('insidecause', 1),
 ('ourself', 1), ('wealththats', 1), ('shortys', 1), ('hollerin', 1), ('balle
 rs', 1), ('atdrug', 1), ('dealer', 1), ('crackhead', 1), ('crackand', 1),
 ('thatbut', 1), ('holier', 1), ('thoucause', 1), ('thoubefore', 1), ('againc
 ause', 1), ('106', 1), ('pushin', 1), ('benzi', 1), ('ballerific', 1), ('ter
 rifici', 1), ('pastdue', 1), ('bills', 1), ('specifici', 1), ('problem', 1),
 ('spendin', 1), ('itwe', 1), ('selfconscious', 1), ('admit', 1), ('itsyleen
 a', 1), ('cmonand', 1), ('downchitown', 1), ('chitownwe', 1), ('downnow',
 1), ('belt', 1), ('saved', 1), ('lifecmonoh', 1), ('hows', 1), ('thingshe',
 1), ('thingsaight', 1), ('tall', 1), ('enoughall', 1), ('goldand', 1), ('rea
 lityreal', 1), ('lays', 1), ('mema', 1), ('cook', 1), ('sunday', 1), ('oktha
 ts', 1), ('girlfriendthis', 1), ('usand', 1), ('answers', 1), ('usthis', 1),
 ('sweet', 1), ('photoof', 1), ('grannys', 1), ('ussuper', 1), ('thanksgivin


```
g', 1), ('rightyeah', 1), ('lifesomebody', 1), ('grace', 1), ('faceand', 1),
('facei', 1), ('food', 1), ('granny', 1), ('itmonkey', 1), ('bread', 1), ('i
twhen', 1), ('guard', 1), ('itas', 1), ('laughwho', 1), ('fastwho', 1), ('gl
assand', 1), ('boutall', 1), ('realitynow', 1), ('thingsreal', 1), ('lease',
1), ('methis', 1), ('standing', 1), ('uscome', 1), ('grammy', 1), ('picturea
bi', 1), ('menow', 1), ('tvnow', 1), ('favoritei', 1), ('deny', 1), ('riderb
ut', 1), ('electric', 1), ('slidinggrandma', 1), ('shook', 1), ('upaw', 1),
('photo', 1), ('upi', 1), ('ruth', 1), ('namebut', 1), ('polaroidsll', 1),
('memory', 1), ('laneyou', 1), ('auntie', 1), ('rudebut', 1), ('holiday',
1), ('eating', 1), ('foodand', 1), ('worst', 1), ('cousinsgot', 1), ('roache
s', 1), ('cousinsact', 1), ('bath', 1), ('cousinsfit', 1), ('yallim', 1),
('legbut', 1), ('pee', 1), ('bedrain', 1), ('sayrain', 1), ('sayi', 1), ('wo
ke', 1), ('minda', 1), ('creative', 1), ('using', 1), ('knives', 1), ('gunsk
eep', 1), ('nose', 1), ('godand', 1), ('risin', 1), ('sunall', 1), ('dogan
d', 1), ('dogi', 1), ('manly', 1), ('heated', 1), ('warminsittin', 1), ('gri
llin', 1), ('george', 1), ('foremanwhy', 1), ('uncle', 1), ('ray', 1), ('shi
ela', 1), ('performinthe', 1), ('inyall', 1), ('reunionand', 1), ('wine',
1), ('communionand', 1), ('isntwe', 1), ('letting', 1), ('uhall', 1), ('real
itythey', 1), ('thingi', 1), ('thingsclk', 1), ('mercedes', 1), ('benzall',
1), ('thingsa', 1), ('lotta', 1), ('moneyall', 1), ('thingsmommy', 1), ('fig
htingall', 1), ('thingslets', 1), ('stevie', 1), ('outta', 1), ('warwe', 1),
('terrorism', 1), ('racismbut', 1), ('ourselvesjesus', 1), ('midwest', 1),
('restlesswhere', 1), ('restless', 1), ('snatch', 1), ('necklaceand', 1),
('lexussomebody', 1), ('isi', 1), ('valley', 1), ('istop', 1), ('floor', 1),
('view', 1), ('breathlesstry', 1), ('hardgetting', 1), ('choked', 1), ('dete
ctives', 1), ('methodthey', 1), ('questions', 1), ('harass', 1), ('arrest',
1), ('ussayin', 1), ('breakfasthuh', 1), ('basiswe', 1), ('casesa', 1), ('tr
unk', 1), ('rental', 1), ('avismy', 1), ('uswell', 1), ('foolbut', 1), ('nov
ember', 1), ('movei', 1), ('hopejesus', 1), ('longso', 1), ('meto', 1), ('mu
rderers', 1), ('dealers', 1), ('scrippersjesus', 1), ('themto', 1), ('victim
s', 1), ('welfare', 1), ('yeahjesus', 1), ('themnow', 1), ('thee', 1), ('cle
arlyi', 1), ('wearycause', 1), ('extinctwe', 1), ('role', 1), ('models', 1),
('thinki', 1), ('argue', 1), ('facial', 1), ('featuresor', 1), ('convert',
1), ('atheists', 1), ('believersim', 1), ('teachersthe', 1), ('kathie', 1),
('lee', 1), ('regis', 1), ('jesusso', 1), ('radio', 1), ('needs', 1), ('this
they', 1), ('except', 1), ('jesusthat', 1), ('means', 1), ('sex', 1), ('lie
s', 1), ('videotapebut', 1), ('huhwell', 1), ('spinswhichll', 1), ('endsthe
n', 1), ('sinsand', 1), ('dreamin', 1), ('aboutnext', 1), ('outjesus', 1)]
```

Twitter Data – Cher:

There are 15674013 tokens in the data.

There are 1555373 unique tokens in the data.

There are 93389697 characters in the data.

The lexical diversity is 0.099 in the data.

IOPub data rate exceeded.

The Jupyter server will temporarily stop sending output to the client in order to avoid crashing it.

To change this limit, set the config variable
`--ServerApp.iopub_data_rate_limit`.

Current values:

ServerApp.iopub_data_rate_limit=1000000.0 (bytes/sec)

ServerApp.rate_limit_window=3.0 (secs)

Out[92]: [1495780, 258391, 0.17274666060516922, 9158569]

Q: How do you think the "top 5 words" would be different if we left stopwords in the data?

A: Without removing the stopwords, the top 5 words would be much more likely to be words like "the," "is," "at," or "which." If we decided not to get rid of them, they would appear much more frequently than other, more meaningful words. Having said that, in both Kanye and Joni we see some similarities.

For Joni, the top 5 words are: ('know', 31), ('like', 23), ('dont', 23), ('love', 19), ('come', 17)

For Kanye, the top 5 words are: ('get', 153), ('like', 135), ('im', 106), ('know', 89), ('aint', 83)

As we can see, both of them contain the words "know," "like," and a variation of "do not" with "dont" and "aint." Thus, even after removing the stop words, there are still some that don't have a lot of importance.

Q: What were your prior beliefs about the lexical diversity between the artists? Does the difference (or lack thereof) in lexical diversity between the artists conform to your prior beliefs?

A: Joni has a much higher lexical diversity at 0.620 than Kanye does at 0.385. I would not have previously expected the lexical diversity to be much different between the two artists. I believe both are very poetic in how they write their songs. What is interesting is that Kanye's 25 songs contain much more content per song than Joni's 25 songs. Joni's songs contain 2649 words, while Kanye's songs contain 8496 words. That's more than 3 times as many words! However, he obviously repeats a lot of words given the relative lack of lexical diversity.

Specialty Statistics

The descriptive statistics we have calculated are quite generic. You will now calculate a handful of statistics tailored to these data.

1. Ten most common emojis by artist in the twitter descriptions.
2. Ten most common hashtags by artist in the twitter descriptions.
3. Five most common words in song titles by artist.
4. For each artist, a histogram of song lengths (in terms of number of tokens)

We can use the `emoji` library to help us identify emojis and you have been given a function to help you.

```
In [93]: assert(emoji.is_emoji("❤️"))
```

```
assert(not emoji.is_emoji(":-"))
```

Emojis 🤗

What are the ten most common emojis by artist in the twitter descriptions?

```
In [94]: # Your code here

# Define the function to gather the most common emojis, with twitter_data and
# number of emojis = 10 passed in as the arguments
def most_common_emojis(twitter_data, num_emojis=10):

    # Initialize and empty dictionary to store counts in
    emoji_counts = {}

    # Use a for loop to search for artist and descriptions within the twitter_data
    for artist, descriptions in twitter_data.items():

        # Outer loop: "for tweet in descriptions" iterates over each tweet in
        # Inner loop: "for token in tweet" takes the tweet obtained from the
        # over each token in the tweet.
        # Conditional statement: returns "True" if the token is an emoji
        # Finally, use a list comprehension (")") to construct a list of emojis
        # in descriptions.
        all_emojis = [token for tweet in descriptions for token in tweet if
                       emoji.is_emoji(token)]

        # Gets the top 10 most common emojis
        emoji_counts[artist] = Counter(all_emojis).most_common(num_emojis)

    return emoji_counts

# Calculate the most common emojis for each artist
common_emojis_by_artist = most_common_emojis(clean_twitter_data)

# Print the results
for artist, emojis in common_emojis_by_artist.items():
    print(f"{artist}: {emojis}")
```

```
cher_followers_data: [('❤️', 14715), ('🌈', 14122), ('❤️', 10154),
('❤️', 9651), ('🌟', 8343), ('🌈', 5474), ('🇺🇸', 3696), ('💙', 3684), ('💜',
3499), ('👤', 3278)]
robynkonihiwa_followers_data: [('🌈', 1701), ('❤️', 1167), ('❤️', 98
7), ('🌟', 751), ('❤️', 653), ('🌈', 570), ('🎵', 272), ('👤', 213), ('❤️', 2
12), ('💜', 205)]
cher_followers: []
robynkonihiwa_followers: []
```

Hashtags

What are the ten most common hashtags by artist in the twitter descriptions?

```
In [95]: # Your code here

# Set up the most common hashtags
def most_common_hashtags(twitter_data, num_hashtags=10):
```

```

# Initialize the hashtag_counts dictionary
hashtag_counts = {}
for artist, descriptions in twitter_data.items():
    all_hashtags = [token for tweet in descriptions for token in tweet if token.startswith('#')]
    hashtag_counts[artist] = Counter(all_hashtags).most_common(num_hashtags)
return hashtag_counts

# Calculate the most common hashtags for each artist
common_hashtags_by_artist = most_common_hashtags(clean_twitter_data)

# Print the results
for artist, hashtags in common_hashtags_by_artist.items():
    print(f"{artist}: {hashtags}")

```

```

cher_followers_data: [('resist', 10540), ('blm', 9582), ('blacklivesmatter', 7775), ('theresistance', 3211), ('fbr', 3131), ('resistance', 2788), ('1', 2423), ('', 2128), ('voteblue', 2062), ('lgbtq', 1791)]
robynkonihiwa_followers_data: [('blacklivesmatter', 584), ('blm', 340), ('music', 289), ('1', 194), ('', 168), ('teamfollowback', 127), ('edm', 108), ('lgbtq', 81), ('resist', 80), ('art', 69)]
cher_followers: []
robynkonihiwa_followers: []

```

Song Titles

What are the five most common words in song titles by artist? The song titles should be on the first line of the lyrics pages, so if you have kept the raw file contents around, you will not need to re-read the data.

In [96]: *# Your code here*

```

# Function to find the most common words in song titles
def most_common_words_in_titles(lyrics_data, num_words=5):
    title_word_counts = {}
    for artist, songs in lyrics_data.items():
        all_title_words = []
        for song, lyrics in songs.items():

            # Split the lyrics by the return lines
            title = lyrics.split('\n')[0]

            title_words = clean_and_tokenize(title)

            # ".extend" is used to add the elements of the list
            # title_words to the end of the list all_title_words
            all_title_words.extend(title_words)
        title_word_counts[artist] = Counter(all_title_words).most_common(num_words)
    return title_word_counts

# Calculate the most common words in song titles for each artist
common_words_by_artist = most_common_words_in_titles(lyrics_data)

# Print the results

```

```
for artist, common_words in common_words_by_artist.items():
    print(f"{artist}: {common_words}")
```

```
jonimitchell: [('morning', 2), ('song', 2), ('ladies', 1), ('canyon', 1),
('sides', 1)]
```

```
kanyewest: [('skit', 3), ('school', 3), ('spirit', 3), ('workout', 2), ('plan', 2)]
```

Song Lengths

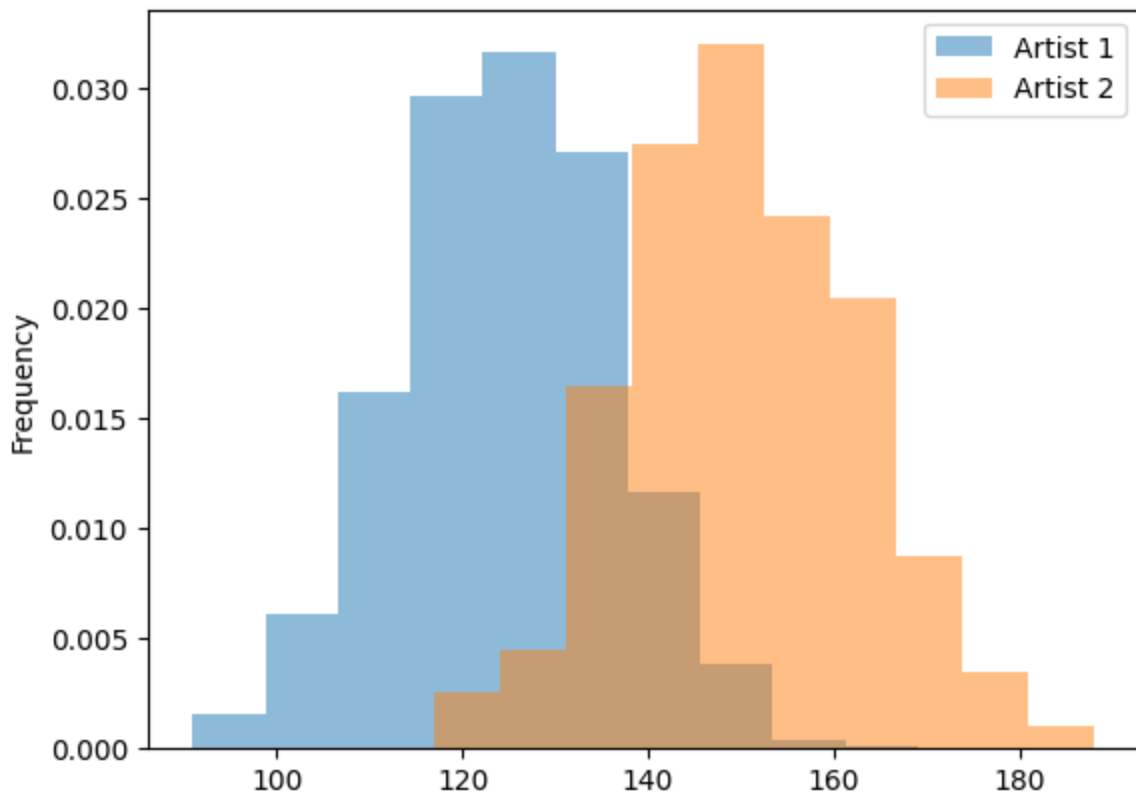
For each artist, a histogram of song lengths (in terms of number of tokens). If you put the song lengths in a data frame with an artist column, matplotlib will make the plotting quite easy. An example is given to help you out.

```
In [97]: num_replicates = 1000

df = pd.DataFrame({
    "artist" : ['Artist 1'] * num_replicates + ['Artist 2'] * num_replicates,
    "length" : np.concatenate((np.random.poisson(125, num_replicates), np.random.poisson(150, num_replicates)))

df.groupby('artist')['length'].plot(kind="hist", density=True, alpha=0.5, legend=True)
```

```
Out[97]: artist
Artist 1    Axes(0.125,0.11;0.775x0.77)
Artist 2    Axes(0.125,0.11;0.775x0.77)
Name: length, dtype: object
```



Since the lyrics may be stored with carriage returns or tabs, it may be useful to have a function that can collapse whitespace, using regular expressions, and be used for

splitting.

Q: What does the regular expression `'\s+'` match on?

A: The regular expression `'\s+'` matches one or more whitespace characters in a string. This includes spaces (' '), tabs ('\t'), newlines ('\n'), and carriage returns ('\r'). The `+` is a quantifier in regular expressions. It matches one or more occurrences of the preceding element. In this case, it applies to the `\s`, meaning the expression will match sequences of one or more consecutive whitespace characters.

```
In [100... collapse_whitespace = re.compile(r'\s+')

def tokenize_lyrics(lyric) :
    """strip and split on whitespace"""
    return([item.lower() for item in collapse_whitespace.split(lyric)])
```

```
In [102... # Your lyric length comparison chart here.

# Calculate song lengths

# Initialize list
song_lengths = []

for artist, songs in lyrics_data.items():
    for song, lyrics in songs.items():
        tokens = tokenize_lyrics(lyrics)
        length = len(tokens)
        song_lengths.append({'artist': artist, 'length': length})

# Create DataFrame
df = pd.DataFrame(song_lengths)

# Plotting
df.groupby('artist')['length'].plot(kind="hist", density=True, alpha=0.5, legend=True)
plt.xlabel("Song Length (number of tokens)")
plt.ylabel("Density")
plt.title("Histogram of Song Lengths by Artist")
plt.show()
```

