

# Naive Bayes Classifier using Spark MapReduce

December 19, 2023

```
[ ]: from pyspark.ml.feature import Tokenizer, StopWordsRemover

[ ]: # Load the dataset
input_path = "/FileStore/tables/cancer_data.tsv"
input_df = spark.read.option("header", "true").option("inferSchema", "true").
    ↪option("delimiter", "\t").csv(input_path)
input_df = input_df.withColumnRenamed("RT @JarrodTheLord: Yall creating life_
    ↪and cloning animals but want us to believe there is no cure for Cancer or_
    ↪Aids... alright https://t.c...", "text").withColumnRenamed("no", "label")

[ ]: # Tokenize the text
tokenizer = Tokenizer(inputCol="text", outputCol="words")
words_data = tokenizer.transform(input_df)

[ ]: # Display the tokenized data
display(words_data)

[ ]: from pyspark.sql.types import ArrayType, StringType
from pyspark.sql.functions import col
# Define a UDF to filter tokens
# Define a UDF to filter tokens based on their length
def filter_short_tokens(words):
    return [word for word in words if len(word) >= 3]

# Register the UDF
filter_short_tokens_udf = udf(filter_short_tokens, ArrayType(StringType()))

# Apply the UDF to the DataFrame and add a new column
filtered_data = words_data.withColumn("filtered_words",_
    ↪filter_short_tokens_udf(col("words")))

# Show the result
display(filtered_data)

[ ]: # Remove stop words
stopwords_remover = StopWordsRemover(inputCol="filtered_words",_
    ↪outputCol="filtered")
filtered_data = stopwords_remover.transform(filtered_data)
```

```
[ ]: # Display the data after stop words removal
display(filtered_data)
```

```
[ ]: # Select relevant columns
required_data = filtered_data.select(['filtered', 'label'])
```

```
[ ]: # Display the required data
display(required_data)
```

```
[ ]: # Split the data into training and testing sets
train_data, test_data = required_data.randomSplit([0.80, 0.20], seed=42)
train_data.display()
```

```
[ ]: # Calculate probabilities
total_count = train_data.count()
yes_tweets = train_data.filter(train_data["label"] == 'yes').count()
no_tweets = train_data.filter(train_data["label"] == 'no').count()
p_yes = yes_tweets / total_count
p_no = no_tweets / total_count
```

```
[ ]: print("Total count is", total_count)
print("Total number of yes tweets are", yes_tweets)
print("Total number of no tweets are", no_tweets)
print("The probability of yes tweets", p_yes)
print("The probability of no tweets Texts", p_no)
```

Total count is 1248

Total number of yes tweets are 175

Total number of no tweets are 1072

The probability of yes tweets 0.14022435897435898

The probability of no tweets Texts 0.8589743589743589

```
[ ]: # Prepare RDD for further processing
train_rdd = train_data.rdd.map(lambda x: (x[1], x[0]))
```

```
[ ]: display(train_rdd.collect())
```

```
[ ]: train_rdd_filtered = train_rdd
```

```
[ ]: display(train_rdd_filtered.collect())
```

```
[ ]: # Process data for word counting
train_rdd_word_counts = train_rdd_filtered.flatMap(lambda x: [(x[0], word) for
    ↪ word in x[1]]).map(lambda x: ((x[0], x[1]), 1))

# Laplace smoothing constant
alpha = 0.1
```

```
word_counts = train_rdd_word_counts.map(lambda x: ((x[0][1], x[0][0]), x[1])).
    ↪reduceByKey(lambda x, y: x + y)
```

```
[ ]: word_counts.collect()
```

```
Out[18]: [(( '#armyhivestreamingparty', 'no'), 1),
  (('purest', 'no'), 2),
  (('thing', 'no'), 21),
  (('ever', 'no'), 16),
  (('experienced.', 'no'), 1),
  (('beyonce', 'no'), 1),
  (('bts', 'no'), 3),
  (('together', 'no'), 15),
  (('cure', 'no'), 1142),
  (('cancer', 'no'), 816),
  (('diseases/illnesses', 'no'), 1),
  (('bring', 'no'), 26),
  (('world', 'no'), 106),
  (('peace', 'no'), 34),
  ((' ', 'no'), 1),
  ((' #ayurveda_a_panacea', 'yes'), 3),
  ((' "mother', 'yes'), 1),
  (('healing"', 'yes'), 1),
  ((' @gurmeetramrahim', 'yes'), 2),
  (('tells', 'yes'), 2),
  (('ayurveda', 'yes'), 4),
  (('first', 'yes'), 5),
  (('scientific', 'yes'), 3),
  (('medical', 'yes'), 11),
  (('branch', 'yes'), 2),
  (('originated', 'yes'), 2),
  (('india.', 'yes'), 2),
  (('yoga', 'yes'), 2),
  (('practices,', 'yes'), 2),
  (('becomes', 'yes'), 2),
  (('boon', 'yes'), 2),
  (('cure', 'yes'), 184),
  (('cancer,', 'yes'), 33),
  (('kidney', 'yes'), 7),
  (('disorders,', 'yes'), 2),
  (('cardiac', 'yes'), 2),
  (('arrest,', 'yes'), 2),
  (('etc.', 'yes'), 4),
  (('https://t.co/7jdqksviqf', 'yes'), 1),
  ((' "mother", 'yes'), 1),
  (('healing'.', 'yes'), 1),
  (('https://t.co/vkj1uq8sod', 'yes'), 1),
  ((' #cureforcancer', 'no'), 1),
```

(('100', 'no'), 3),  
 (('gave', 'no'), 15),  
 (('medical', 'no'), 16),  
 (('research', 'no'), 44),  
 (('teams', 'no'), 2),  
 (('back', 'no'), 23),  
 (('2014', 'no'), 2),  
 (('know', 'no'), 46),  
 (('hid', 'no'), 1),  
 (('sick', 'no'), 8),  
 (('needed', 'no'), 5),  
 (('lie', 'no'), 5),  
 (('anyone', 'no'), 19),  
 (('true', 'no'), 9),  
 (('noname', 'no'), 1),  
 (('executiveorder', 'no'), 2),  
 (('sick,', 'no'), 1),  
 (('deranged', 'no'), 1),  
 (('liberals', 'no'), 23),  
 (('happy', 'no'), 18),  
 (('president', 'no'), 52),  
 (('found', 'no'), 65),  
 (('cancer.', 'no'), 137),  
 (('morons', 'no'), 2),  
 (('say,', 'no'), 8),  
 (('no!', 'no'), 1),  
 (('want', 'no'), 74),  
 (('people', 'no'), 153),  
 (('suffer', 'no'), 5),  
 (('die!', 'no'), 1),  
 (('unhinged', 'no'), 2),  
 (('nothing', 'no'), 18),  
 (('please', 'no'), 59),  
 (('them.', 'no'), 12),  
 (('thrive', 'no'), 2),  
 (('something', 'no'), 50),  
 (('bitch', 'no'), 12),  
 (('about.', 'no'), 7),  
 (('garethsouthgatewould', 'no'), 1),  
 (('solve', 'no'), 13),  
 (('hunger', 'no'), 25),  
 (('even', 'no'), 63),  
 (('take', 'no'), 28),  
 (('credit', 'no'), 8),  
 (('say', 'no'), 66),  
 (('the', 'no'), 5),  
 (('lads', 'no'), 1),  
 (('well"', 'no'), 1),

(('worldcup18', 'no'), 1),  
 (('england', 'no'), 1),  
 (('lgbtactionplan', 'no'), 1),  
 (('important', 'no'), 8),  
 (('step', 'no'), 6),  
 (('become', 'no'), 10),  
 (('equally', 'no'), 1),  
 (('diverse', 'no'), 1),  
 (('place', 'no'), 3),  
 (('love.', 'no'), 4),  
 (('gay,', 'no'), 1),  
 (('anything', 'no'), 23),  
 (('straight', 'no'), 1),  
 (('curable', 'no'), 3),  
 (('cure.', 'no'), 12),  
 (('work', 'no'), 25),  
 (('hiv', 'no'), 14),  
 (('otd', 'no'), 1),  
 (('years', 'no'), 35),  
 (('ago,', 'no'), 4),  
 (('terry', 'no'), 2),  
 (('fox', 'no'), 4),  
 (('passed', 'no'), 3),  
 (('away', 'no'), 16),  
 (('new', 'no'), 20),  
 (('westminister,', 'no'), 1),  
 (('bc.', 'no'), 1),  
 (('speaking', 'no'), 3),  
 (('run,', 'no'), 1),  
 (('said', 'no'), 20),  
 (('dreamer,', 'no'), 1),  
 (('saying', 'no'), 17),  
 (('initiate', 'no'), 1),  
 (('kind', 'no'), 8),  
 (('definitive', 'no'), 1),  
 (('answer', 'no'), 4),  
 (('believe', 'no'), 35),  
 (('miracles.', 'no'), 1),  
 (('to."', 'no'), 1),  
 (('watch', 'no'), 9),  
 (('heritageminute', 'no'), 1),  
 (('here.', 'no'), 3),  
 (('https://t.co/oqonptqesm', 'no'), 1),  
 (('predict', 'no'), 2),  
 (('amp;', 'no'), 95),  
 (('cure', 'no'), 2),  
 (('next', 'no'), 13),  
 (('cancer', 'no'), 26),

```

(('https://t.co/gyztoi46yi', 'no'), 1),
(('minutes', 'no'), 3),
(('#cancercure', 'no'), 7),
(('#healed', 'no'), 2),
(('#reiki', 'no'), 2),
(('#rtrrt', 'no'), 4),
(('#bigpharma', 'no'), 4),
(('#loa', 'no'), 2),
(('#fridayreads', 'no'), 2),
(('#nlp', 'no'), 2),
(('#acim', 'no'), 2),
(('https://t.co/uqx20hdk8c', 'no'), 1),
(('https://t.co/oqh0emyqpk', 'no'), 1),
(('home', 'no'), 9),
(('#academiczodiac', 'no'), 2),
(('#o5iris', 'no'), 2),
(('predicta...', 'no'), 2),
(('https://t.co/jvbdy7arqy', 'no'), 1),
(('https://t.co/4r1zgo8opo', 'no'), 1),
(('https://t.co/ojby25xlyk', 'no'), 1),
(('#righttotry', 'no'), 2),
(('1st', 'no'), 7),
(('lyrics', 'no'), 1),
(('trump', 'no'), 203),
(('released?', 'no'), 1),
(('give', 'no'), 49),
(('vote', 'no'), 7),
(('not', 'no'), 1),
(('hillary', 'no'), 3),
(('dr.', 'no'), 5),
(('video=', 'no'), 1),
(('blocked', 'no'), 2),
(('bill', 'no'), 6),
(('clinton', 'no'), 2),
(('since', 'no'), 13),
(('94', 'no'), 1),
(('#drburzynskirhymeswithlewinski', 'no'), 2),
(('#trumptrainjumponit', 'no'), 1),
((' https://t.co/6odiawuwga', 'no'), 1),
(('via', 'no'), 11),
(('@youtube', 'no'), 2),
(('https://t.co/0uta7ccgwk', 'no'), 1),
(('poor', 'no'), 7),
(('clinton!', 'no'), 1),
(('decades', 'no'), 2),
(('suppressing', 'no'), 1),
(('cancer-cures', 'no'), 1),
(('wants', 'no'), 7),

```

(('righttocry', 'no'), 1),  
 (('righttodie', 'no'), 1),  
 (('ask', 'no'), 15),  
 (('patients', 'no'), 24),  
 (('that's', 'no'), 21),  
 (('right!', 'no'), 5),  
 (('rhymes', 'no'), 1),  
 (('mistress', 'no'), 1),  
 (('both', 'no'), 1),  
 (('f\*cked', 'no'), 1),  
 (('https://t.co/xpiketzppcz', 'no'), 1),  
 (('shocking', 'no'), 1),  
 (('government', 'no'), 1),  
 (('petitions', 'no'), 1),  
 (('dogs', 'no'), 1),  
 (('get', 'no'), 85),  
 (('signatures', 'no'), 1),  
 (('life', 'no'), 39),  
 (('saving', 'no'), 5),  
 (('funding', 'no'), 5),  
 (('terminal', 'no'), 5),  
 (('diseases', 'no'), 16),  
 (('known', 'no'), 12),  
 (('cause', 'no'), 20),  
 (('kill', 'no'), 10),  
 (('them!', 'no'), 1),  
 (('ipf', 'no'), 2),  
 (('pulmonaryfibrosis', 'no'), 2),  
 (('needs', 'no'), 17),  
 (('like', 'no'), 140),  
 (('cancer!', 'no'), 2),  
 (('humanity?', 'no'), 1),  
 (('show', 'no'), 11),  
 (('care ', 'no'), 1),  
 (('https://t.co/zzcp7rnld0', 'no'), 1),  
 (('sisterscare:', 'no'), 1),  
 (('cbs2's', 'no'), 2),  
 (('@burrelltv', 'no'), 1),  
 (('@efinchtv', 'no'), 2),  
 (('@danatylertv', 'no'), 2),  
 (('took', 'no'), 12),  
 (('part', 'no'), 10),  
 (('sisters', 'no'), 2),  
 (('event', 'no'), 13),  
 (('today,', 'no'), 7),  
 (('addressing', 'no'), 1),  
 (('disparities', 'no'), 1),  
 (('diagnosis', 'no'), 5),

(('treatment', 'no'), 18),  
 (('breast', 'no'), 32),  
 (('african', 'no'), 2),  
 (('american', 'no'), 12),  
 (('women.', 'no'), 4),  
 (('https://t.co/unjsqhcycd6', 'no'), 1),  
 (('teachersleading', 'no'), 1),  
 (('come', 'no'), 31),  
 (('doctor', 'no'), 9),  
 (('revolution', 'no'), 1),  
 (('teaching', 'no'), 3),  
 (('going', 'no'), 27),  
 (('teachers', 'no'), 2),  
 (('@nickdonohuenmef', 'no'), 1),  
 (('teamhalon', 'yes'), 1),  
 (('supports', 'yes'), 1),  
 (('national', 'yes'), 1),  
 (('randomactsoflightday', 'yes'), 1),  
 (('stands', 'yes'), 1),  
 (('llsusa', 'yes'), 1),  
 (('continued', 'yes'), 1),  
 (('research', 'yes'), 7),  
 (('cancer.', 'yes'), 6),  
 (('learn', 'yes'), 3),  
 (('https://t.co/q2lt7y56km', 'yes'), 1),  
 (('trump', 'no'), 9),  
 (('cancer,', 'no'), 150),  
 (('eradicate', 'no'), 5),  
 (('poverty', 'no'), 10),  
 (('still', 'no'), 157),  
 (('msm', 'no'), 10),  
 (('pull', 'no'), 4),  
 (('put', 'no'), 34),  
 (('negative', 'no'), 10),  
 (('spin', 'no'), 9),  
 (('it.', 'no'), 54),  
 (('war', 'no'), 8),  
 (('conflict', 'no'), 2),  
 (('preferable', 'no'), 1),  
 (('wrong', 'no'), 14),  
 (('consistently', 'no'), 1),  
 (('are.', 'no'), 5),  
 (('trumpkimsummit', 'no'), 15),  
 (('maga', 'no'), 14),  
 (('discover', 'no'), 9),  
 (('trumpderangementsyndrome', 'no'), 5),  
 (('sufferers', 'no'), 2),  
 (('screaming', 'no'), 1),



```

(('injection', 'no'), 1),
(('stinging.#liberalhypocrisy', 'no'), 1),
(('#singaporesummit', 'no'), 2),
(('#trumpkim', 'no'), 1),
(('#deepstate', 'no'), 1),
(('#thestorm', 'no'), 1),
(('self-inflicted', 'no'), 1),
(('metastasized', 'no'), 1),
(('full', 'no'), 15),
(('blown', 'no'), 1),
(('psychological', 'no'), 1),
(('insanity', 'no'), 2),
(('outbreak', 'no'), 1),
(('across', 'no'), 9),
(('democrat', 'no'), 4),
(('party,', 'no'), 1),
(('liberal', 'no'), 12),
(('media', 'no'), 28),
(('nevertrumpers', 'no'), 1),
(('lead', 'no'), 10),
(('ultimate', 'no'), 2),
(('demise.', 'no'), 1),
(('time', 'no'), 41),
(('let', 'no'), 22),
(('go.', 'no'), 2),
(('#walkaway', 'no'), 2),
((' $agen', 'yes'), 4),
(('ceo', 'yes'), 2),
(('@garo_armen', 'yes'), 1),
(('@agenus_bio', 'yes'), 1),
(('#wearwhite', 'yes'), 2),
(('@cancerresearch', 'yes'), 2),
(('show', 'yes'), 2),
(('commitment', 'yes'), 1),
(('fight', 'yes'), 4),
(('#cancer', 'yes'), 6),
(('#immunotherapy', 'yes'), 3),
(('#oncology', 'yes'), 1),
(('#curecancer', 'yes'), 1),
(('https://t.co/f5vigul6ky', 'yes'), 1),
(('proud', 'yes'), 1),
(('work', 'yes'), 4),
(('pioneer', 'yes'), 1),
(('#immunotherapy,', 'yes'), 1),
(('dr.', 'yes'), 6),
(('@garo_armen,', 'yes'), 1),
(('&', 'yes'), 28),
(('stand', 'yes'), 1),

```

(('courageous', 'yes'), 1),  
 (('patients', 'yes'), 4),  
 (('physicians', 'yes'), 1),  
 (('in pursuit of', 'yes'), 3),  
 (('cancer.', 'yes'), 25),  
 (('https://t.co/ve97oxqakw', 'yes'), 1),  
 (('innovation', 'yes'), 4),  
 (('engine', 'yes'), 1),  
 (('produced', 'yes'), 1),  
 (('inds', 'yes'), 2),  
 (('18 mos;', 'yes'), 1),  
 (('planned', 'yes'), 2),  
 (('2018', 'yes'), 2),  
 (('1h 2019.', 'yes'), 1),  
 (('molecules', 'yes'), 1),  
 (('discovery', 'yes'), 2),  
 (('ind', 'yes'), 1),  
 (('yrs.', 'yes'), 2),  
 (('plus,', 'yes'), 1),  
 (('>100', 'yes'), 1),  
 (('pts', 'yes'), 8),  
 (('treated', 'yes'), 2),  
 (('agen1884', 'yes'), 1),  
 (('&/or', 'yes'), 1),  
 (('agen2034', 'yes'), 1),  
 (('asco2018.', 'yes'), 1),  
 (('speed', 'yes'), 3),  
 (('action!', 'yes'), 1),  
 (('https://t.co/vrwjomjmvo', 'yes'), 1),  
 (('driven', 'yes'), 2),  
 (('powered', 'yes'), 2),  
 (('right', 'yes'), 8),  
 (('combination', 'yes'), 2),  
 (('science', 'yes'), 6),  
 (('people-see', 'yes'), 2),  
 (('pursuing', 'yes'), 2),  
 (('passion', 'yes'), 2),  
 (('advancing', 'yes'), 2),  
 (('toward', 'yes'), 2),  
 (('goal', 'yes'), 2),  
 (('story:', 'yes'), 2),  
 (('https://t.co/apxu2dplkm', 'yes'), 2),  
 (('big3biobos', 'yes'), 2),  
 (('celz', 'no'), 1),  
 (('imho,', 'no'), 1),  
 (('beaten', 'no'), 3),  
 (('many', 'no'), 49),  
 (('countries.;', 'no'), 1),

(('rate', 'no'), 18),  
 (('95%+', 'no'), 1),  
 (('fda', 'no'), 1),  
 (('big', 'no'), 17),  
 (('pharma', 'no'), 1),  
 (('keep', 'no'), 27),  
 (('treatments', 'no'), 10),  
 (('illegal', 'no'), 6),  
 (('u.s', 'no'), 2),  
 (('capitalize', 'no'), 1),  
 (('150,000-500,000', 'no'), 1),  
 (('unneeded', 'no'), 1),  
 (('surgeries', 'no'), 5),  
 (('chemo', 'no'), 11),  
 (('celz's', 'no'), 1),  
 (('amniostem', 'no'), 1),  
 (('help', 'no'), 106),  
 (('cachexia', 'no'), 1),  
 (('huge!', 'no'), 1),  
 (('&gt;', 'yes'), 1),  
 (('kind', 'yes'), 4),  
 (('advice', 'yes'), 2),  
 (('isn't', 'yes'), 5),  
 (('helpful.', 'yes'), 1),  
 (('i'm', 'yes'), 4),  
 (('sure', 'yes'), 2),  
 (('it's', 'yes'), 13),  
 (('said', 'yes'), 5),  
 (('wanting', 'yes'), 1),  
 (('helpful,', 'yes'), 1),  
 (('sucks.', 'yes'), 1),  
 (('good', 'yes'), 10),  
 (('deeds', 'yes'), 1),  
 (('faith', 'yes'), 1),  
 (('quinoa', 'yes'), 1),  
 (('hope', 'yes'), 15),  
 (('many', 'yes'), 18),  
 (('impacted', 'yes'), 1),  
 (('it.', 'yes'), 11),  
 (('worse,', 'yes'), 1),  
 (('none', 'yes'), 1),  
 (('around', 'yes'), 3),  
 (('surgery', 'yes'), 1),  
 (('treatment', 'yes'), 9),  
 (('days.', 'yes'), 2),  
 (('&gt;"the', 'no'), 1),  
 (('idea', 'no'), 6),  
 (('tell', 'no'), 17),

(('story', 'no'), 13),  
 (('lesnar', 'no'), 1),  
 (('looking', 'no'), 10),  
 (('leaving', 'no'), 4),  
 (('belt', 'no'), 1),  
 (('ufc"', 'no'), 1),  
 (('garner', 'no'), 1),  
 (('sympathy', 'no'), 2),  
 (('reigns.', 'no'), 1),  
 (('roman', 'no'), 1),  
 (('reigns', 'no'), 1),  
 (('end', 'no'), 54),  
 (('hunger,', 'no'), 12),  
 (('find', 'no'), 204),  
 (('booed', 'no'), 1),  
 (('pushed', 'no'), 1),  
 (('hard.', 'no'), 1),  
 (('https://t.co/uxful5sx6q', 'no'), 1),  
 (('\*breaking', 'no'), 1),  
 (('news\*', 'no'), 1),  
 (('hypothetical', 'no'), 1),  
 (('situation)', 'no'), 1),  
 (('finds', 'no'), 7),  
 (('cancer!', 'no'), 28),  
 (('everyone', 'no'), 26),  
 (('mother:', 'no'), 1),  
 (('what', 'no'), 2),  
 (('piece', 'no'), 5),  
 (('shit!', 'no'), 1),  
 (('probably', 'no'), 15),  
 (('knew', 'no'), 7),  
 (('whole', 'no'), 12),  
 (('time!'", 'no'), 1),  
 (('dont', 'no'), 7),  
 (('agree', 'no'), 2),  
 (('shit', 'no'), 20),  
 (('guy', 'no'), 13),  
 (('does/says,', 'no'), 1),  
 (('ya'll", 'no'), 1),  
 (('never', 'no'), 57),  
 (('satisfied', 'no'), 2),  
 (('\*opens', 'no'), 1),  
 (('cab-booking', 'no'), 1),  
 (('app\*', 'no'), 1),  
 (('cab', 'no'), 2),  
 (('\*books', 'no'), 1),  
 (('cab\*', 'no'), 1),  
 (('years,', 'no'), 5),

(('months', 'no'), 1),  
 (('weeks', 'no'), 1),  
 (('days', 'no'), 8),  
 (('hours.', 'no'), 1),  
 (('kids', 'no'), 22),  
 (('born', 'no'), 6),  
 (('grow', 'no'), 12),  
 (('knowing', 'no'), 4),  
 (('mother', 'no'), 1),  
 (('scientists', 'no'), 18),  
 (('volcano', 'no'), 1),  
 (('erupt', 'no'), 1),  
 (('wipe', 'no'), 4),  
 (('hawaii', 'no'), 1),  
 (('away.', 'no'), 1),  
 (('\*this', 'yes'), 1),  
 (('chimes\*', 'yes'), 1),  
 (('positivity', 'yes'), 2),  
 (('can't', 'yes'), 3),  
 (('neither', 'yes'), 2),  
 (('fighting', 'yes'), 2),  
 (('talk.', 'yes'), 1),  
 (('battle"', 'yes'), 1),  
 (('narrative', 'yes'), 1),  
 (('created', 'yes'), 3),  
 (('people', 'yes'), 20),  
 (('facing', 'yes'), 1),  
 (('disease', 'yes'), 4),  
 (('denied', 'yes'), 2),  
 (('narratives.', 'yes'), 1),  
 (('couldn't', 'yes'), 1),  
 (('agree', 'yes'), 4),  
 (('more\*', 'yes'), 1),  
 (('lifelife', 'yes'), 1),  
 (('https://t.co/akj8zdd3jz', 'yes'), 1),  
 (('abbvie', 'no'), 2),  
 (('alphabet-backed', 'no'), 2),  
 (('calico', 'no'), 2),  
 (('spend', 'no'), 8),  
 (('another', 'no'), 12),  
 (('1b', 'no'), 2),  
 (('https://t.co/a7tqwrwtrv', 'no'), 2),  
 (('erictrump:', 'no'), 1),  
 (('we're', 'no'), 1),  
 (('raising', 'no'), 14),  
 (('money', 'no'), 76),  
 (('we're', 'no'), 11),  
 (('win', 'no'), 16),

(('fight.', 'no'), 2),  
 (('fight', 'no'), 35),  
 (('pediatric', 'no'), 12),  
 (('cancer."', 'no'), 2),  
 (('https://t.co/7pyfcd7xzy', 'no'), 1),  
 (('thank', 'no'), 39),  
 (('@erictrump', 'no'), 2),  
 (('son', 'no'), 5),  
 (('alive', 'no'), 2),  
 (('due', 'no'), 3),  
 (('started', 'no'), 4),  
 (('jude.', 'no'), 1),  
 (('year', 'no'), 29),  
 (('survivor', 'no'), 11),  
 (('week.', 'no'), 4),  
 (('loose', 'no'), 2),  
 (('least', 'no'), 6),  
 (('every', 'no'), 50),  
 (('month', 'no'), 5),  
 (('local', 'no'), 3),  
 (('children's', 'no'), 2),  
 (('hospital', 'no'), 8),  
 (('childhood', 'no'), 24),  
 (('https://t.co/gwqtxflqau', 'no'), 1),  
 (('11:11', 'no'), 4),  
 (('excellent', 'no'), 4),  
 (('grades!!!', 'no'), 2),  
 (('job', 'no'), 14),  
 (('best', 'no'), 17),  
 (('friend', 'no'), 10),  
 (('happiness', 'no'), 6),  
 (('friends', 'no'), 13),  
 (('gun', 'no'), 7),  
 (('control', 'no'), 11),  
 (('@arianagrande', 'no'), 4),  
 (('notice', 'no'), 5),  
 (('@shawnmendes', 'no'), 4),  
 (('follow', 'no'), 9),  
 (('bye', 'no'), 5),  
 (('anxiety', 'no'), 6),  
 (('fun', 'no'), 7),  
 (('summer', 'no'), 4),  
 (('lion', 'no'), 9),  
 (('rid', 'no'), 9),  
 (('depression', 'no'), 16),  
 (('hug', 'no'), 6),  
 (('love', 'no'), 26),  
 (('good', 'no'), 44),

(('grades', 'no'), 1),  
 (('summer!', 'no'), 1),  
 (('harry', 'no'), 3),  
 (('potter', 'no'), 1),  
 (('reunion', 'no'), 3),  
 (('weight!', 'no'), 1),  
 (('16-year-old', 'no'), 4),  
 (('jeremiah', 'no'), 4),  
 (('thomas,', 'no'), 3),  
 (('battling', 'no'), 5),  
 (('used', 'no'), 11),  
 (('make', 'no'), 84),  
 (('wish', 'no'), 12),  
 (('abortion.', 'no'), 5),  
 (('you,', 'no'), 9),  
 (('jeremiah,', 'no'), 3),  
 (('beautiful', 'no'), 8),  
 (('life,', 'no'), 13),  
 (('inspiring', 'no'), 7),  
 (('many.', 'no'), 4),  
 (('https://t.co/bumzmgyay', 'no'), 3),  
 (('sad', 'no'), 13),  
 (('young,', 'no'), 1),  
 (('brainwashed', 'no'), 2),  
 (('maybe', 'no'), 21),  
 (('health', 'no'), 23),  
 (('care', 'no'), 19),  
 (('all,', 'no'), 3),  
 (('imaginary', 'no'), 1),  
 (('exist', 'no'), 4),  
 (('https://t.co/27twcb3oek', 'no'), 1),  
 (('selfless', 'no'), 1),  
 (('way"', 'no'), 1),  
 (('idk', 'no'), 1),  
 (('mann', 'no'), 1),  
 (('it's', 'no'), 59),  
 (('pretty', 'no'), 5),  
 (('selfish', 'no'), 1),  
 (('try', 'no'), 18),  
 (('use', 'no'), 10),  
 (('religion', 'no'), 11),  
 (('controls', 'no'), 1),  
 (('someone', 'no'), 29),  
 (('else's', 'no'), 1),  
 (('body...', 'no'), 1),  
 (('coulda', 'no'), 1),  
 (('wished', 'no'), 4),  
 (('https://t.co/tyx3hynfa8', 'no'), 1),

(('17yr', 'no'), 3),  
 (('stage', 'no'), 18),  
 (('carcinoid', 'no'), 12),  
 (('survivor.', 'no'), 4),  
 (('disease', 'no'), 48),  
 (('longest', 'no'), 2),  
 (('surviving', 'no'), 2),  
 (('patient', 'no'), 7),  
 (('world.', 'no'), 15),  
 (('survival', 'no'), 11),  
 (('5yrs.', 'no'), 6),  
 (('plz', 'no'), 14),  
 (('read', 'no'), 24),  
 (('remarkable', 'no'), 10),  
 (('miracle', 'no'), 12),  
 (('book', 'no'), 13),  
 (('living', 'no'), 14),  
 (('silent', 'no'), 11),  
 (('enemy', 'no'), 11),  
 (('coming', 'no'), 18),  
 (('soon.', 'no'), 11),  
 (('support', 'no'), 31),  
 (('mdacc.', 'no'), 1),  
 (('inoperable', 'no'), 5),  
 (('neuroendocrine', 'no'), 7),  
 (('misdiagnosed', 'no'), 7),  
 (('lets', 'no'), 3),  
 (('cancer#', 'no'), 1),  
 (('17yr.', 'no'), 1),  
 (('endcancer#', 'no'), 1),  
 (('1965:', 'no'), 2),  
 (('bet', 'no'), 10),  
 (('there'll', 'no'), 4),  
 (('flying', 'no'), 8),  
 (('cars', 'no'), 7),  
 (('future"', 'no'), 1),  
 (('1987:', 'no'), 1),  
 (('future', 'no'), 12),  
 (('2018:', 'no'), 2),  
 (('prototype', 'no'), 1),  
 (('potentially', 'no'), 5),  
 (('viable', 'no'), 3),  
 (('cures', 'no'), 12),  
 (('cancers', 'no'), 9),  
 (('given', 'no'), 5),  
 (('1998:', 'no'), 2),  
 (('in', 'no'), 1),  
 (('twenty', 'no'), 2),



```

(('cancer!', 'no'), 1),
(('dennis', 'no'), 1),
(('rodman', 'no'), 1),
(('saved', 'no'), 5),
(('nuclear', 'no'), 4),
(('celebrity', 'no'), 2),
(('apprentice', 'no'), 1),
(('hosted', 'no'), 2),
(('united', 'no'), 3),
(('states', 'no'), 1),
(('dictator', 'no'), 3),
(('north', 'no'), 8),
(('korea!', 'no'), 1),
(('@panmass', 'no'), 1),
(('rider', 'no'), 1),
(('@danafarber.', 'no'), 1),
(('fund', 'no'), 14),
(('retweet!', 'no'), 1),
(('https://t.co/qdwqn3zzre', 'no'), 1),
(('https://t.co/lafbipfydc', 'no'), 1),
(('2018', 'no'), 9),
(('3mt', 'no'), 1),
(('semi-finalist', 'no'), 1),
(('ubc', 'no'), 1),
(('science', 'no'), 12),
(('master', 'no'), 1),
(('student', 'no'), 2),
(('evan', 'no'), 1),
(('warner', 'no'), 1),
(('presented', 'no'), 1),
(('shakespeare's', 'no'), 1),
(('much', 'no'), 48),
(('thing?', 'no'), 2),
(('asked.', 'no'), 1),
(('#tbt', 'no'), 1),
(('talk:', 'no'), 1),
(('https://t.co/vjazjwo5em', 'no'), 1),
(('22,', 'no'), 1),
(('ethan', 'no'), 1),
(('hawes', 'no'), 1),
(('diagnosed', 'no'), 8),
(('multiple', 'no'), 4),
(('myeloma,', 'no'), 1),
(('typically', 'no'), 1),
(('afflicting', 'no'), 2),
(('age', 'no'), 4),
(('65.', 'no'), 1),
(('hospitalizations,', 'no'), 1),

```

(('chemotherapy', 'no'), 1),  
 (('stem', 'no'), 3),  
 (('cell', 'no'), 5),  
 (('replacement', 'no'), 1),  
 (('followed.', 'no'), 1),  
 (('https://t.co/f0lztvodcb', 'no'), 1),  
 (('5,000', 'no'), 2),  
 (('young', 'no'), 12),  
 (('women', 'no'), 16),  
 (('country', 'no'), 15),  
 (('participated', 'no'), 2),  
 (('year's', 'no'), 4),  
 (('lax', 'no'), 1),  
 (('raise', 'no'), 24),  
 (('funds', 'no'), 12),  
 (('awareness', 'no'), 13),  
 (('tournament', 'no'), 2),  
 (('raised', 'no'), 16),  
 (('235,000', 'no'), 1),  
 (('komen', 'no'), 2),  
 (('csnj', 'no'), 1),  
 (('@komennja.', 'no'), 1),  
 (('more:', 'no'), 2),  
 (('https://t.co/4qlw3bei7k', 'no'), 1),  
 (('https://t.co/uwdnketoie', 'no'), 1),  
 (('6th', 'no'), 4),  
 (('40tude', 'no'), 3),  
 (('thames', 'no'), 2),  
 (('sup', 'no'), 1),  
 (('marathon', 'no'), 3),  
 (('saturday', 'no'), 2),  
 (('already', 'no'), 16),  
 (('£63k', 'no'), 1),  
 (('marks', 'no'), 1),  
 (('colon', 'no'), 4),  
 (('add', 'no'), 7),  
 (('£180k', 'no'), 1),  
 (('past', 'no'), 4),  
 (('years.', 'no'), 8),  
 (('miles', 'no'), 2),  
 (('sups', 'no'), 1),  
 (('finishing', 'no'), 1),  
 (('putney.', 'no'), 1),  
 (('@curebowelcancer', 'no'), 1),  
 (('@active360', 'no'), 1),  
 (('https://t.co/dn5cjvm9gz', 'no'), 1),  
 (('87%', 'no'), 2),  
 (('f-ing', 'no'), 2),

(('republicans', 'no'), 6),  
 (('trump.', 'no'), 9),  
 (('omg.', 'no'), 1),  
 (('how?', 'no'), 2),  
 (('children', 'no'), 32),  
 (('ripped', 'no'), 1),  
 (('mother's', 'no'), 2),  
 (('arms.', 'no'), 1),  
 (('truly', 'no'), 8),  
 (('80%', 'no'), 2),  
 (('news', 'no'), 9),  
 (('20%', 'no'), 3),  
 (('don't', 'no'), 50),  
 (('on.', 'no'), 4),  
 (('disgusted', 'no'), 1),  
 (('must', 'no'), 20),  
 (('accept', 'no'), 4),  
 (('sinister', 'no'), 1),  
 (('brainwashing.', 'no'), 1),  
 (('can't', 'no'), 29),  
 (('comprehend', 'no'), 2),  
 (('maga', 'no'), 3),  
 (('grasp', 'no'), 1),  
 (('shooting', 'no'), 2),  
 (('afternoon', 'no'), 1),  
 (('nature.', 'no'), 2),  
 (('one', 'no'), 89),  
 (('voting', 'no'), 2),  
 (('overwhelming', 'no'), 2),  
 (('numbers.', 'no'), 1),  
 (('https://t.co/s7ctd0lyc6', 'no'), 1),  
 (('1405scituateave', 'no'), 1),  
 (('realjameswoods', 'no'), 4),  
 (('funny', 'no'), 7),  
 (('is,', 'no'), 5),  
 (('discovered', 'no'), 17),  
 (('tomorrow,', 'no'), 3),  
 (('donald', 'no'), 12),  
 (('telling', 'no'), 12),  
 (('somehow', 'no'), 4),  
 (('instrumental', 'no'), 1),  
 (('it,', 'no'), 19),  
 (('jackass', 'no'), 1),  
 (('trumpuppets', 'no'), 1),  
 (('word', 'no'), 4),  
 (('it!!!!', 'no'), 1),  
 (('...and', 'no'), 1),  
 (('you're', 'no'), 10),

```

(('saying;', 'no'), 1),
(('maybe', 'no'), 1),
(('be"', 'no'), 1),
(('right', 'no'), 18),
(('now!!!', 'no'), 1),
(('@1776stonewall', 'no'), 2),
(('pathetic', 'no'), 1),
(('wasn't', 'no'), 1),
(('enough!!!', 'no'), 1),
(('talk', 'no'), 24),
(('pitiful', 'no'), 1),
(('souls', 'no'), 2),
(('way', 'no'), 43),
(('trump!!!!', 'no'), 1),
(('@1999mcm11', 'no'), 1),
(('denuclearization', 'no'), 3),
(('achieved.', 'no'), 2),
(('gets', 'no'), 6),
(('taste', 'no'), 1),
(('western', 'no'), 3),
(('culture,', 'no'), 1),
(('mcdonalds,', 'no'), 1),
(('coke,', 'no'), 1),
(('xbox,', 'no'), 1),
(('etc...', 'no'), 1),
(('over.', 'no'), 1),
(('achieve', 'no'), 5),
(('peace...', 'no'), 1),
(('complain', 'no'), 32),
(('him.', 'no'), 31),
(('cuz', 'no'), 4),
(('he's', 'no'), 22),
(('orange', 'no'), 2),
(('tweets', 'no'), 2),
(('stupid', 'no'), 2),
(('stuff.', 'no'), 2),
(('@2mattandgaby2', 'yes'), 1),
(('no.', 'yes'), 2),
(('laetrile', 'yes'), 3),
(('cancer', 'yes'), 122),
(('anything', 'yes'), 4),
(('else.', 'yes'), 1),
(('studied,', 'yes'), 1),
(('work.', 'yes'), 2),
(('please', 'yes'), 3),
(('stop', 'yes'), 10),
(('spreading', 'yes'), 2),
(('dangerous', 'yes'), 3),

```

(('bullshit.', 'yes'), 1),  
 (('https://t.co/7z0jpa3osp', 'yes'), 1),  
 (('@420stonercorner', 'yes'), 2),  
 (('.@420stonercorner', 'yes'), 2),  
 (('fascinating', 'yes'), 2),  
 (('subject', 'yes'), 2),  
 (('matter', 'yes'), 3),  
 (('thought', 'yes'), 3),  
 (('pot', 'yes'), 3),  
 (('heads', 'yes'), 2),  
 (('such,', 'yes'), 2),  
 (('rampant', 'yes'), 2),  
 (('family', 'yes'), 5),  
 (('cbd', 'yes'), 5),  
 (('maybe', 'yes'), 4),  
 (('use', 'yes'), 18),  
 (('thc', 'yes'), 3),  
 (('instead', 'yes'), 3),  
 (('morphine!', 'yes'), 2),  
 (('#cannabis', 'yes'), 2),  
 (('#weeds', 'yes'), 2),  
 (('#pmqs', 'yes'), 1),  
 (('medically', 'yes'), 1),  
 (('unqualified', 'yes'), 1),  
 (('opinion?', 'yes'), 1),  
 (('#theresamay', 'yes'), 1),  
 (('@weedseeds\_uk', 'yes'), 1),  
 (('@ukcannabisclub', 'yes'), 1),  
 (('@ukcannabisclubs', 'yes'), 1),  
 (('@younglabouruk', 'yes'), 1),  
 (('#bbcqt', 'yes'), 1),  
 (('#bbcaq', 'yes'), 1),  
 (('assholes', 'yes'), 1),  
 (('criminalizing', 'yes'), 1),  
 (('@weedfeed', 'yes'), 1),  
 (('@norml', 'yes'), 2),  
 (('#nyc', 'yes'), 1),  
 (('#nypd', 'yes'), 1),  
 (('@ukleap', 'yes'), 1),  
 (('@uklabour', 'yes'), 1),  
 (('@weednucleus', 'yes'), 1),  
 (('cannabis', 'yes'), 15),  
 (('communion', 'yes'), 1),  
 (('brings', 'yes'), 2),  
 (('great', 'yes'), 3),  
 (('peace', 'yes'), 1),  
 (('forever!', 'yes'), 1),  
 (('https://t.co/f583xnt3jy', 'yes'), 1),

```

(('@_cape_', 'no'), 1),
(('@__hunchohoudini', 'no'), 1),
(('provide', 'no'), 6),
(('food', 'no'), 11),
(('shelter', 'no'), 1),
(('homeless', 'no'), 2),
(('can't', 'no'), 2),
(('aids', 'no'), 32),
(('although', 'no'), 2),
(('it)', 'no'), 3),
(('shot', 'no'), 5),
(('tupac.we', 'no'), 1),
(('investigators', 'no'), 1),
(('support.', 'no'), 4),
(('swear', 'no'), 7),
(('better', 'no'), 25),
(('sometimes', 'no'), 1),
(('always', 'no'), 15),
(('dumb', 'no'), 2),
(('ppl', 'no'), 6),
(('fuck', 'no'), 18),
(('@_duckyfresh', 'no'), 1),
(('@babygrotesque', 'no'), 1),
(('underlying', 'no'), 1),
(('economic', 'no'), 4),
(('problems', 'no'), 5),
(('contribute', 'no'), 2),
(('gotham's", 'no'), 1),
(('crime.', 'no'), 3),
(('stop', 'no'), 25),
(('beating', 'no'), 4),
(('mentally', 'no'), 2),
(('ill', 'no'), 5),
(('throwing', 'no'), 1),
(('inhumane', 'no'), 1),
(('caricature', 'no'), 1),
(('mental', 'no'), 7),
(('facility.', 'no'), 1),
(('bruce', 'no'), 1),
(('literally', 'no'), 18),
(('spidey', 'no'), 1),
(('meme', 'no'), 1),
(('tea', 'no'), 2),
(('@_kingofthehill1:', 'no'), 1),
(('sexual', 'no'), 4),
(('https://t.co/qt6n6wrqse', 'no'), 1),
(('@aacrfoundation:', 'yes'), 1),
(('leading', 'yes'), 3),

```

```

(('discoveries,', 'yes'), 1),
(('targeting', 'yes'), 3),
(('cures', 'yes'), 7),
(('saving', 'yes'), 2),
(('lives.', 'yes'), 1),
(('follow', 'yes'), 1),
(('see', 'yes'), 7),
(('helping', 'yes'), 1),
(('@abaabood', 'no'), 1),
(('@andreas_krieg', 'no'), 1),
(('me,', 'no'), 1),
(('spared', 'no'), 1),
(('efforts', 'no'), 4),
(('last', 'no'), 30),
(('plus', 'no'), 1),
(('resolve', 'no'), 2),
(('brotherly', 'no'), 1),
(('basis', 'no'), 2),
(('qatar', 'no'), 3),
(('continues', 'no'), 2),
(('path', 'no'), 2),
(('destabilizing', 'no'), 1),
(('saudi', 'no'), 2),
(('arabia', 'no'), 2),
(('region.', 'no'), 1),
(('surgically', 'no'), 1),
(('remove', 'no'), 2),
(('regime.', 'no'), 1),
(('@abby_higgins5:', 'no'), 1),
(('government/pharmacies', 'no'), 1),
(('wanna', 'no'), 6),
(('money...', 'no'), 1),
(('idc', 'no'), 2),
(('https://t.co/ivompefmo0', 'no'), 1),
(('@ac_breezy', 'no'), 1),
(('viruses,', 'no'), 1),
(('spam/trolls', 'no'), 1),
(('exact', 'no'), 1),
(('thing...', 'no'), 1),
(('yet', 'no'), 16),
(('enough', 'no'), 12),
(('realize', 'no'), 6),
(('earned', 'no'), 1),
(('millennium', 'no'), 1),
(('prize', 'no'), 4),
...]
```

```
[ ]: # Prepare data for further analysis
train_rdd_for_analysis = word_counts.map(lambda x: (x[0][1], x[0][0], x[1]))
```

```
[ ]: train_rdd_for_analysis.collect()
```

```
Out[20]: [('no', '#armyhivestreamingparty', 1),
('no', 'purest', 2),
('no', 'thing', 21),
('no', 'ever', 16),
('no', 'experienced.', 1),
('no', 'beyonce', 1),
('no', 'bts', 3),
('no', 'together', 15),
('no', 'cure', 1142),
('no', 'cancer', 816),
('no', 'diseases/illnesses', 1),
('no', 'bring', 26),
('no', 'world', 106),
('no', 'peace', 34),
('no', ' ', 1),
('yes', '#ayurveda_a_panacea', 3),
('yes', '"mother', 1),
('yes', 'healing"', 1),
('yes', '@gurmeetramrahim', 2),
('yes', 'tells', 2),
('yes', 'ayurveda', 4),
('yes', 'first', 5),
('yes', 'scientific', 3),
('yes', 'medical', 11),
('yes', 'branch', 2),
('yes', 'originated', 2),
('yes', 'india.', 2),
('yes', 'yoga', 2),
('yes', 'practices,', 2),
('yes', 'becomes', 2),
('yes', 'boon', 2),
('yes', 'cure', 184),
('yes', 'cancer,', 33),
('yes', 'kidney', 7),
('yes', 'disorders,', 2),
('yes', 'cardiac', 2),
('yes', 'arrest,', 2),
('yes', 'etc.', 4),
('yes', 'https://t.co/7jdqksviqf', 1),
('yes', '"mother"', 1),
('yes', 'healing'.', 1),
('yes', 'https://t.co/vkj1uq8sod', 1),
('no', '#cureforcancer', 1),
```



('no', '100', 3),  
('no', 'gave', 15),  
('no', 'medical', 16),  
('no', 'research', 44),  
('no', 'teams', 2),  
('no', 'back', 23),  
('no', '2014', 2),  
('no', 'know', 46),  
('no', 'hid', 1),  
('no', 'sick', 8),  
('no', 'needed', 5),  
('no', 'lie', 5),  
('no', 'anyone', 19),  
('no', 'true', 9),  
('no', '#noname', 1),  
('no', '#executiveorder', 2),  
('no', 'sick,', 1),  
('no', 'deranged', 1),  
('no', 'liberals', 23),  
('no', 'happy', 18),  
('no', 'president', 52),  
('no', 'found', 65),  
('no', 'cancer.', 137),  
('no', 'morons', 2),  
('no', 'say,', 8),  
('no', '"no!', 1),  
('no', 'want', 74),  
('no', 'people', 153),  
('no', 'suffer', 5),  
('no', 'die!"', 1),  
('no', 'unhinged', 2),  
('no', 'nothing', 18),  
('no', 'please', 59),  
('no', 'them.', 12),  
('no', 'thrive', 2),  
('no', 'something', 50),  
('no', 'bitch', 12),  
('no', 'about.', 7),  
('no', '#garethsouthgatewould', 1),  
('no', 'solve', 13),  
('no', 'hunger', 25),  
('no', 'even', 63),  
('no', 'take', 28),  
('no', 'credit', 8),  
('no', 'say', 66),  
('no', '"the', 5),  
('no', 'lads', 1),  
('no', 'well"', 1),

('no', '#worldcup18', 1),  
('no', '#england', 1),  
('no', '#lgbtactionplan', 1),  
('no', 'important', 8),  
('no', 'step', 6),  
('no', 'become', 10),  
('no', 'equally', 1),  
('no', 'diverse', 1),  
('no', 'place', 3),  
('no', 'love.', 4),  
('no', 'gay,', 1),  
('no', 'anything', 23),  
('no', 'straight', 1),  
('no', 'curable', 3),  
('no', 'cure.', 12),  
('no', 'work', 25),  
('no', 'hiv', 14),  
('no', '#otd', 1),  
('no', 'years', 35),  
('no', 'ago,', 4),  
('no', 'terry', 2),  
('no', 'fox', 4),  
('no', 'passed', 3),  
('no', 'away', 16),  
('no', 'new', 20),  
('no', 'westminister,', 1),  
('no', 'bc.', 1),  
('no', 'speaking', 3),  
('no', 'run,', 1),  
('no', 'said', 20),  
('no', 'dreamer,', 1),  
('no', 'saying', 17),  
('no', 'initiate', 1),  
('no', 'kind', 8),  
('no', 'definitive', 1),  
('no', 'answer', 4),  
('no', 'believe', 35),  
('no', 'miracles.', 1),  
('no', 'to."', 1),  
('no', 'watch', 9),  
('no', '#heritageminute', 1),  
('no', 'here.', 3),  
('no', 'https://t.co/oqonptqcs', 1),  
('no', '#predict', 2),  
('no', '&', 95),  
('no', '#cure', 2),  
('no', 'next', 13),  
('no', '#cancer', 26),

('no', 'https://t.co/gyztoi46yi', 1),  
 ('no', 'minutes', 3),  
 ('no', '#cancercure', 7),  
 ('no', '#healed', 2),  
 ('no', '#reiki', 2),  
 ('no', '#rtrrt', 4),  
 ('no', '#bigpharma', 4),  
 ('no', '#loa', 2),  
 ('no', '#fridayreads', 2),  
 ('no', '#nlp', 2),  
 ('no', '#acim', 2),  
 ('no', 'https://t.co/uqx20hdk8c', 1),  
 ('no', 'https://t.co/oqh0emyqpk', 1),  
 ('no', 'home', 9),  
 ('no', '#academiczodiac', 2),  
 ('no', '#o5iris', 2),  
 ('no', 'predicta...', 2),  
 ('no', 'https://t.co/jvbdy7arqy', 1),  
 ('no', 'https://t.co/4r1zgo8opo', 1),  
 ('no', 'https://t.co/ojby25xlyk', 1),  
 ('no', '#righttotry', 2),  
 ('no', '1st', 7),  
 ('no', 'lyrics', 1),  
 ('no', 'trump', 203),  
 ('no', 'released?', 1),  
 ('no', 'give', 49),  
 ('no', 'vote', 7),  
 ('no', '-not', 1),  
 ('no', 'hillary', 3),  
 ('no', 'dr.', 5),  
 ('no', 'video=', 1),  
 ('no', 'blocked', 2),  
 ('no', 'bill', 6),  
 ('no', 'clinton', 2),  
 ('no', 'since', 13),  
 ('no', '94', 1),  
 ('no', '#drburzynskirhymeswithlewinski', 2),  
 ('no', '#trumptrainjumponit', 1),  
 ('no', 'https://t.co/6odiawuwga', 1),  
 ('no', 'via', 11),  
 ('no', '@youtube', 2),  
 ('no', 'https://t.co/0uta7ccgwk', 1),  
 ('no', 'poor', 7),  
 ('no', 'clinton!', 1),  
 ('no', 'decades', 2),  
 ('no', 'suppressing', 1),  
 ('no', 'cancer-cures', 1),  
 ('no', 'wants', 7),

('no', '#righttocry', 1),  
 ('no', '#righttodie', 1),  
 ('no', 'ask', 15),  
 ('no', 'patients', 24),  
 ('no', 'that's', 21),  
 ('no', 'right!', 5),  
 ('no', 'rhymes', 1),  
 ('no', 'mistress', 1),  
 ('no', '-both', 1),  
 ('no', 'f\*cked', 1),  
 ('no', 'https://t.co/xpikezppcz', 1),  
 ('no', '#shocking', 1),  
 ('no', '#government', 1),  
 ('no', '#petitions', 1),  
 ('no', '#dogs', 1),  
 ('no', 'get', 85),  
 ('no', 'signatures', 1),  
 ('no', 'life', 39),  
 ('no', 'saving', 5),  
 ('no', 'funding', 5),  
 ('no', 'terminal', 5),  
 ('no', 'diseases', 16),  
 ('no', 'known', 12),  
 ('no', 'cause', 20),  
 ('no', 'kill', 10),  
 ('no', 'them!', 1),  
 ('no', '#ipf', 2),  
 ('no', '#pulmonaryfibrosis', 2),  
 ('no', 'needs', 17),  
 ('no', 'like', 140),  
 ('no', '#cancer!', 2),  
 ('no', '#humanity?', 1),  
 ('no', 'show', 11),  
 ('no', 'care ', 1),  
 ('no', 'https://t.co/zzcp7rnld0', 1),  
 ('no', '#sisterscare:', 1),  
 ('no', 'cbs2's", 2),  
 ('no', '@burrelltv,', 1),  
 ('no', '@efinchtv', 2),  
 ('no', '@danatylertv', 2),  
 ('no', 'took', 12),  
 ('no', 'part', 10),  
 ('no', 'sisters', 2),  
 ('no', 'event', 13),  
 ('no', 'today,', 7),  
 ('no', 'addressing', 1),  
 ('no', 'disparities', 1),  
 ('no', 'diagnosis', 5),

('no', 'treatment', 18),  
 ('no', 'breast', 32),  
 ('no', 'african', 2),  
 ('no', 'american', 12),  
 ('no', 'women.', 4),  
 ('no', 'https://t.co/unjsqhcyd6', 1),  
 ('no', '#teachersleading', 1),  
 ('no', 'come', 31),  
 ('no', 'doctor', 9),  
 ('no', 'revolution', 1),  
 ('no', 'teaching', 3),  
 ('no', 'going', 27),  
 ('no', 'teachers', 2),  
 ('no', '@nickdonohuenmef', 1),  
 ('yes', '#teamhalon', 1),  
 ('yes', 'supports', 1),  
 ('yes', 'national', 1),  
 ('yes', '#randomactsoflightday', 1),  
 ('yes', 'stands', 1),  
 ('yes', '@llsusa', 1),  
 ('yes', 'continued', 1),  
 ('yes', 'research', 7),  
 ('yes', '#cancer.', 6),  
 ('yes', 'learn', 3),  
 ('yes', 'https://t.co/q2lt7y56km', 1),  
 ('no', '#trump', 9),  
 ('no', 'cancer,', 150),  
 ('no', 'eradicate', 5),  
 ('no', 'poverty', 10),  
 ('no', 'still', 157),  
 ('no', 'msm', 10),  
 ('no', 'pull', 4),  
 ('no', 'put', 34),  
 ('no', 'negative', 10),  
 ('no', 'spin', 9),  
 ('no', 'it.', 54),  
 ('no', 'war', 8),  
 ('no', 'conflict', 2),  
 ('no', 'preferable', 1),  
 ('no', 'wrong', 14),  
 ('no', 'consistently', 1),  
 ('no', 'are.', 5),  
 ('no', '#trumpkimsummit', 15),  
 ('no', '#maga', 14),  
 ('no', 'discover', 9),  
 ('no', '#trumpderangementsyndrome', 5),  
 ('no', 'sufferers', 2),  
 ('no', 'screaming', 1),

```

('no', 'injection', 1),
('no', 'stinging.#liberalhypocrisy', 1),
('no', '#singaporeshummit', 2),
('no', '#trumpkim', 1),
('no', '#deepstate', 1),
('no', '#thestorm', 1),
('no', 'self-inflicted', 1),
('no', 'metastasized', 1),
('no', 'full', 15),
('no', 'blown', 1),
('no', 'psychological', 1),
('no', 'insanity', 2),
('no', 'outbreak', 1),
('no', 'across', 9),
('no', 'democrat', 4),
('no', 'party,', 1),
('no', 'liberal', 12),
('no', 'media', 28),
('no', 'nevertrumpers', 1),
('no', 'lead', 10),
('no', 'ultimate', 2),
('no', 'demise.', 1),
('no', 'time', 41),
('no', 'let', 22),
('no', 'go.', 2),
('no', '#walkaway', 2),
('yes', '$agen', 4),
('yes', 'ceo', 2),
('yes', '@garo_armen', 1),
('yes', '@agenus_bio', 1),
('yes', '#wearwhite', 2),
('yes', '@cancerresearch', 2),
('yes', 'show', 2),
('yes', 'commitment', 1),
('yes', 'fight', 4),
('yes', '#cancer', 6),
('yes', '#immunotherapy', 3),
('yes', '#oncology', 1),
('yes', '#curecancer', 1),
('yes', 'https://t.co/f5vigul6ky', 1),
('yes', 'proud', 1),
('yes', 'work', 4),
('yes', 'pioneer', 1),
('yes', '#immunotherapy,', 1),
('yes', 'dr.', 6),
('yes', '@garo_armen,', 1),
('yes', '&', 28),
('yes', 'stand', 1),

```

('yes', 'courageous', 1),  
 ('yes', 'patients', 4),  
 ('yes', 'physicians', 1),  
 ('yes', '#inpursuitof', 3),  
 ('yes', 'cancer.', 25),  
 ('yes', 'https://t.co/ve97oxqakw', 1),  
 ('yes', 'innovation', 4),  
 ('yes', 'engine', 1),  
 ('yes', 'produced', 1),  
 ('yes', 'inds', 2),  
 ('yes', '18mos;', 1),  
 ('yes', 'planned', 2),  
 ('yes', '2018', 2),  
 ('yes', '1h2019.', 1),  
 ('yes', 'molecules', 1),  
 ('yes', 'discovery', 2),  
 ('yes', 'ind', 1),  
 ('yes', 'yrs.', 2),  
 ('yes', 'plus,', 1),  
 ('yes', '>100', 1),  
 ('yes', 'pts', 8),  
 ('yes', 'treated', 2),  
 ('yes', 'agen1884', 1),  
 ('yes', '&/or', 1),  
 ('yes', 'agen2034', 1),  
 ('yes', '#asco2018.', 1),  
 ('yes', 'speed', 3),  
 ('yes', 'action!', 1),  
 ('yes', 'https://t.co/vrwjomjmvo', 1),  
 ('yes', 'driven', 2),  
 ('yes', 'powered', 2),  
 ('yes', 'right', 8),  
 ('yes', 'combination', 2),  
 ('yes', 'science', 6),  
 ('yes', 'people-see', 2),  
 ('yes', 'pursuing', 2),  
 ('yes', 'passion', 2),  
 ('yes', 'advancing', 2),  
 ('yes', 'toward', 2),  
 ('yes', 'goal', 2),  
 ('yes', 'story:', 2),  
 ('yes', 'https://t.co/apxu2dplkm', 2),  
 ('yes', '@big3biobos', 2),  
 ('no', '\$celz', 1),  
 ('no', 'imho,', 1),  
 ('no', 'beaten', 3),  
 ('no', 'many', 49),  
 ('no', 'countries.;', 1),

('no', 'rate', 18),  
 ('no', '95%+', 1),  
 ('no', '#fda', 1),  
 ('no', 'big', 17),  
 ('no', '#pharma', 1),  
 ('no', 'keep', 27),  
 ('no', 'treatments', 10),  
 ('no', 'illegal', 6),  
 ('no', 'u.s', 2),  
 ('no', 'capitalize', 1),  
 ('no', '\$150,000-\$500,000', 1),  
 ('no', 'unneeded', 1),  
 ('no', 'surgeries', 5),  
 ('no', 'chemo', 11),  
 ('no', 'celz's', 1),  
 ('no', '#amniostem', 1),  
 ('no', 'help', 106),  
 ('no', '#cachexia', 1),  
 ('no', 'huge!', 1),  
 ('yes', '&gt;', 1),  
 ('yes', 'kind', 4),  
 ('yes', 'advice', 2),  
 ('yes', 'isn't', 5),  
 ('yes', 'helpful.', 1),  
 ('yes', 'i'm', 4),  
 ('yes', 'sure', 2),  
 ('yes', 'it's', 13),  
 ('yes', 'said', 5),  
 ('yes', 'wanting', 1),  
 ('yes', 'helpful,', 1),  
 ('yes', 'sucks.', 1),  
 ('yes', 'good', 10),  
 ('yes', 'deeds', 1),  
 ('yes', 'faith', 1),  
 ('yes', 'quinoa', 1),  
 ('yes', 'hope', 15),  
 ('yes', 'many', 18),  
 ('yes', 'impacted', 1),  
 ('yes', 'it.', 11),  
 ('yes', 'worse,', 1),  
 ('yes', 'none', 1),  
 ('yes', 'around', 3),  
 ('yes', 'surgery', 1),  
 ('yes', 'treatment', 9),  
 ('yes', 'days.', 2),  
 ('no', '&gt;"the', 1),  
 ('no', 'idea', 6),  
 ('no', 'tell', 17),



('no', 'story', 13),  
 ('no', 'lesnar', 1),  
 ('no', 'looking', 10),  
 ('no', 'leaving', 4),  
 ('no', 'belt', 1),  
 ('no', 'ufc"', 1),  
 ('no', 'garner', 1),  
 ('no', 'sympathy', 2),  
 ('no', 'reigns.', 1),  
 ('no', 'roman', 1),  
 ('no', 'reigns', 1),  
 ('no', 'end', 54),  
 ('no', 'hunger,', 12),  
 ('no', 'find', 204),  
 ('no', 'booed', 1),  
 ('no', 'pushed', 1),  
 ('no', 'hard.', 1),  
 ('no', 'https://t.co/uxful5sx6q', 1),  
 ('no', '\*breaking', 1),  
 ('no', 'news\*', 1),  
 ('no', '(hypothetical', 1),  
 ('no', 'situation)', 1),  
 ('no', 'finds', 7),  
 ('no', 'cancer!', 28),  
 ('no', 'everyone', 26),  
 ('no', 'mother:', 1),  
 ('no', '"what', 2),  
 ('no', 'piece', 5),  
 ('no', 'shit!', 1),  
 ('no', 'probably', 15),  
 ('no', 'knew', 7),  
 ('no', 'whole', 12),  
 ('no', 'time!"', 1),  
 ('no', 'dont', 7),  
 ('no', 'agree', 2),  
 ('no', 'shit', 20),  
 ('no', 'guy', 13),  
 ('no', 'does/says,', 1),  
 ('no', "ya'll", 1),  
 ('no', 'never', 57),  
 ('no', 'satisfied', 2),  
 ('no', '\*opens', 1),  
 ('no', 'cab-booking', 1),  
 ('no', 'app\*', 1),  
 ('no', 'cab', 2),  
 ('no', '\*books', 1),  
 ('no', 'cab\*', 1),  
 ('no', 'years,', 5),

('no', 'months,', 1),  
 ('no', 'weeks,', 1),  
 ('no', 'days', 8),  
 ('no', 'hours.', 1),  
 ('no', 'kids', 22),  
 ('no', 'born', 6),  
 ('no', 'grow', 12),  
 ('no', 'knowing', 4),  
 ('no', 'mother,', 1),  
 ('no', 'scientists', 18),  
 ('no', 'volcano', 1),  
 ('no', 'erupt', 1),  
 ('no', 'wipe', 4),  
 ('no', 'hawaii', 1),  
 ('no', 'away.', 1),  
 ('yes', '\*this', 1),  
 ('yes', 'chimes\*', 1),  
 ('yes', 'positivity', 2),  
 ('yes', 'can't', 3),  
 ('yes', 'neither', 2),  
 ('yes', 'fighting', 2),  
 ('yes', 'talk.', 1),  
 ('yes', '"battle"', 1),  
 ('yes', 'narrative', 1),  
 ('yes', 'created', 3),  
 ('yes', 'people', 20),  
 ('yes', 'facing', 1),  
 ('yes', 'disease', 4),  
 ('yes', 'denied', 2),  
 ('yes', 'narratives.', 1),  
 ('yes', 'couldn't', 1),  
 ('yes', 'agree', 4),  
 ('yes', 'more\*', 1),  
 ('yes', '#lifewithcancer', 1),  
 ('yes', 'https://t.co/akj8zdd3jz', 1),  
 ('no', '@abbvie', 2),  
 ('no', 'alphabet-backed', 2),  
 ('no', 'calico', 2),  
 ('no', 'spend', 8),  
 ('no', 'another', 12),  
 ('no', '\$1b', 2),  
 ('no', 'https://t.co/a7tqwrwtrv', 2),  
 ('no', '@erictrump:', 1),  
 ('no', '"we're', 1),  
 ('no', 'raising', 14),  
 ('no', 'money', 76),  
 ('no', 'we're', 11),  
 ('no', 'win', 16),

('no', 'fight.', 2),  
 ('no', 'fight', 35),  
 ('no', 'pediatric', 12),  
 ('no', 'cancer."', 2),  
 ('no', 'https://t.co/7pyfcd7xzy', 1),  
 ('no', 'thank', 39),  
 ('no', '@erictrump', 2),  
 ('no', 'son', 5),  
 ('no', 'alive', 2),  
 ('no', 'due', 3),  
 ('no', 'started', 4),  
 ('no', 'jude.', 1),  
 ('no', 'year', 29),  
 ('no', 'survivor', 11),  
 ('no', 'week.', 4),  
 ('no', 'loose', 2),  
 ('no', 'least', 6),  
 ('no', 'every', 50),  
 ('no', 'month', 5),  
 ('no', 'local', 3),  
 ('no', 'children's', 2),  
 ('no', 'hospital', 8),  
 ('no', 'childhood', 24),  
 ('no', 'https://t.co/gwqtxflqau', 1),  
 ('no', '11:11', 4),  
 ('no', 'excellent', 4),  
 ('no', 'grades!!!', 2),  
 ('no', 'job', 14),  
 ('no', 'best', 17),  
 ('no', 'friend', 10),  
 ('no', 'happiness', 6),  
 ('no', 'friends', 13),  
 ('no', 'gun', 7),  
 ('no', 'control', 11),  
 ('no', '@arianagrande', 4),  
 ('no', 'notice', 5),  
 ('no', '@shawnmendes', 4),  
 ('no', 'follow', 9),  
 ('no', 'bye', 5),  
 ('no', 'anxiety', 6),  
 ('no', 'fun', 7),  
 ('no', 'summer', 4),  
 ('no', 'lion', 9),  
 ('no', 'rid', 9),  
 ('no', 'depression', 16),  
 ('no', 'hug', 6),  
 ('no', 'love', 26),  
 ('no', 'good', 44),

('no', 'grades', 1),  
('no', 'summer!', 1),  
('no', 'harry', 3),  
('no', 'potter', 1),  
('no', 'reunion', 3),  
('no', 'weight!', 1),  
('no', '16-year-old', 4),  
('no', 'jeremiah', 4),  
('no', 'thomas,', 3),  
('no', 'battling', 5),  
('no', 'used', 11),  
('no', 'make', 84),  
('no', 'wish', 12),  
('no', 'abortion.', 5),  
('no', 'you,', 9),  
('no', 'jeremiah,', 3),  
('no', 'beautiful', 8),  
('no', 'life,', 13),  
('no', 'inspiring', 7),  
('no', 'many.', 4),  
('no', 'https://t.co/bumzmggyay', 3),  
('no', 'sad', 13),  
('no', 'young,', 1),  
('no', 'brainwashed', 2),  
('no', 'maybe', 21),  
('no', 'health', 23),  
('no', 'care', 19),  
('no', 'all,', 3),  
('no', 'imaginary', 1),  
('no', 'exist', 4),  
('no', 'https://t.co/27twcb3oek', 1),  
('no', '"selfless', 1),  
('no', 'way"', 1),  
('no', 'idk', 1),  
('no', 'mann', 1),  
('no', 'it's', 59),  
('no', 'pretty', 5),  
('no', 'selfish', 1),  
('no', 'try', 18),  
('no', 'use', 10),  
('no', 'religion', 11),  
('no', 'controls', 1),  
('no', 'someone', 29),  
('no', 'else's', 1),  
('no', 'body...', 1),  
('no', 'coulda', 1),  
('no', 'wished', 4),  
('no', 'https://t.co/tyx3hynfa8', 1),

('no', '17yr', 3),  
 ('no', 'stage', 18),  
 ('no', 'carcinoid', 12),  
 ('no', 'survivor.', 4),  
 ('no', 'disease', 48),  
 ('no', 'longest', 2),  
 ('no', 'surviving', 2),  
 ('no', 'patient', 7),  
 ('no', 'world.', 15),  
 ('no', 'survival', 11),  
 ('no', '5yrs.', 6),  
 ('no', 'plz', 14),  
 ('no', 'read', 24),  
 ('no', 'remarkable', 10),  
 ('no', 'miracle', 12),  
 ('no', 'book', 13),  
 ('no', 'living', 14),  
 ('no', 'silent', 11),  
 ('no', 'enemy', 11),  
 ('no', 'coming', 18),  
 ('no', 'soon.', 11),  
 ('no', 'support', 31),  
 ('no', 'mdacc.', 1),  
 ('no', 'inoperable', 5),  
 ('no', 'neuroendocrine', 7),  
 ('no', 'misdiagnosed', 7),  
 ('no', 'lets', 3),  
 ('no', 'cancer#', 1),  
 ('no', '17yr.', 1),  
 ('no', 'endcancer#', 1),  
 ('no', '1965:', 2),  
 ('no', 'bet', 10),  
 ('no', "there'll", 4),  
 ('no', 'flying', 8),  
 ('no', 'cars', 7),  
 ('no', 'future"', 1),  
 ('no', '1987:', 1),  
 ('no', 'future', 12),  
 ('no', '2018:', 2),  
 ('no', 'prototype', 1),  
 ('no', 'potentially', 5),  
 ('no', 'viable', 3),  
 ('no', 'cures', 12),  
 ('no', 'cancers', 9),  
 ('no', 'given', 5),  
 ('no', '1998:', 2),  
 ('no', '"in', 1),  
 ('no', 'twenty', 2),

('no', 'cancer!', 1),  
 ('no', '"dennis', 1),  
 ('no', 'rodman', 1),  
 ('no', 'saved', 5),  
 ('no', 'nuclear', 4),  
 ('no', 'celebrity', 2),  
 ('no', 'apprentice,', 1),  
 ('no', 'hosted', 2),  
 ('no', 'united', 3),  
 ('no', 'states,', 1),  
 ('no', 'dictator', 3),  
 ('no', 'north', 8),  
 ('no', 'korea!', 1),  
 ('no', '@panmass', 1),  
 ('no', 'rider', 1),  
 ('no', '@danafarber.', 1),  
 ('no', 'fund', 14),  
 ('no', 'retweet!', 1),  
 ('no', 'https://t.co/qdwqn3zzre', 1),  
 ('no', 'https://t.co/lafbipfydc', 1),  
 ('no', '2018', 9),  
 ('no', '3mt', 1),  
 ('no', 'semi-finalist', 1),  
 ('no', 'ubc', 1),  
 ('no', 'science', 12),  
 ('no', 'master', 1),  
 ('no', 'student', 2),  
 ('no', 'evan', 1),  
 ('no', 'warner', 1),  
 ('no', 'presented', 1),  
 ('no', 'shakespeare's', 1),  
 ('no', 'much', 48),  
 ('no', 'thing?', 2),  
 ('no', 'asked.', 1),  
 ('no', '#tbt', 1),  
 ('no', 'talk:', 1),  
 ('no', 'https://t.co/vjazjwo5em', 1),  
 ('no', '22,', 1),  
 ('no', 'ethan', 1),  
 ('no', 'hawes', 1),  
 ('no', 'diagnosed', 8),  
 ('no', 'multiple', 4),  
 ('no', 'myeloma,', 1),  
 ('no', 'typically', 1),  
 ('no', 'afflicting', 2),  
 ('no', 'age', 4),  
 ('no', '65.', 1),  
 ('no', 'hospitalizations,', 1),

('no', 'chemotherapy', 1),  
 ('no', 'stem', 3),  
 ('no', 'cell', 5),  
 ('no', 'replacement', 1),  
 ('no', 'followed.', 1),  
 ('no', 'https://t.co/f0lztvodcb', 1),  
 ('no', '5,000', 2),  
 ('no', 'young', 12),  
 ('no', 'women', 16),  
 ('no', 'country', 15),  
 ('no', 'participated', 2),  
 ('no', 'year's', 4),  
 ('no', 'lax', 1),  
 ('no', 'raise', 24),  
 ('no', 'funds', 12),  
 ('no', 'awareness', 13),  
 ('no', 'tournament', 2),  
 ('no', 'raised', 16),  
 ('no', '\$235,000', 1),  
 ('no', 'komen', 2),  
 ('no', 'csnj', 1),  
 ('no', '@komennja.', 1),  
 ('no', 'more:', 2),  
 ('no', 'https://t.co/4qlw3bei7k', 1),  
 ('no', 'https://t.co/uwdnketoie', 1),  
 ('no', '6th', 4),  
 ('no', '40tude', 3),  
 ('no', 'thames', 2),  
 ('no', 'sup', 1),  
 ('no', 'marathon', 3),  
 ('no', 'saturday', 2),  
 ('no', 'already', 16),  
 ('no', '£63k', 1),  
 ('no', 'marks', 1),  
 ('no', 'colon', 4),  
 ('no', 'add', 7),  
 ('no', '£180k', 1),  
 ('no', 'past', 4),  
 ('no', 'years.', 8),  
 ('no', 'miles', 2),  
 ('no', 'sup', 1),  
 ('no', 'finishing', 1),  
 ('no', 'putney.', 1),  
 ('no', '@curebowelcancer', 1),  
 ('no', '@active360', 1),  
 ('no', 'https://t.co/dn5cjvm9gz', 1),  
 ('no', '87%', 2),  
 ('no', 'f-ing', 2),

('no', 'republicans', 6),  
('no', 'trump.', 9),  
('no', 'omg.', 1),  
('no', 'how?', 2),  
('no', 'children', 32),  
('no', 'ripped', 1),  
('no', 'mother's', 2),  
('no', 'arms.', 1),  
('no', 'truly', 8),  
('no', '80%', 2),  
('no', 'news', 9),  
('no', '20%', 3),  
('no', 'don't', 50),  
('no', 'on.', 4),  
('no', '#disgusted', 1),  
('no', 'must', 20),  
('no', 'accept', 4),  
('no', 'sinister', 1),  
('no', 'brainwashing.', 1),  
('no', 'can't', 29),  
('no', 'comprehend', 2),  
('no', 'maga', 3),  
('no', 'grasp', 1),  
('no', 'shooting', 2),  
('no', 'afternoon', 1),  
('no', 'nature.', 2),  
('no', 'one', 89),  
('no', 'voting', 2),  
('no', 'overwhelming', 2),  
('no', 'numbers.', 1),  
('no', 'https://t.co/s7ctd0lyc6', 1),  
('no', '@1405scituateave', 1),  
('no', '@realjameswoods', 4),  
('no', 'funny', 7),  
('no', 'is,', 5),  
('no', 'discovered', 17),  
('no', 'tomorrow,', 3),  
('no', 'donald', 12),  
('no', 'telling', 12),  
('no', 'somehow', 4),  
('no', 'instrumental', 1),  
('no', 'it,', 19),  
('no', 'jackass', 1),  
('no', 'trumpuppets', 1),  
('no', 'word', 4),  
('no', 'it!!!!', 1),  
('no', '...and', 1),  
('no', 'you're', 10),



('no', 'saying;', 1),  
 ('no', '"maybe', 1),  
 ('no', 'be"', 1),  
 ('no', 'right', 18),  
 ('no', 'now!!!', 1),  
 ('no', '@1776stonewall', 2),  
 ('no', 'pathetic', 1),  
 ('no', 'wasn't', 1),  
 ('no', 'enough!!', 1),  
 ('no', 'talk', 24),  
 ('no', 'pitiful', 1),  
 ('no', 'souls', 2),  
 ('no', 'way', 43),  
 ('no', 'trump!!!!', 1),  
 ('no', '@1999mcm11', 1),  
 ('no', 'denuclearization', 3),  
 ('no', 'achieved.', 2),  
 ('no', 'gets', 6),  
 ('no', 'taste', 1),  
 ('no', 'western', 3),  
 ('no', 'culture,', 1),  
 ('no', 'mcdonalds,', 1),  
 ('no', 'coke,', 1),  
 ('no', 'xbox,', 1),  
 ('no', 'etc...', 1),  
 ('no', 'over.', 1),  
 ('no', 'achieve', 5),  
 ('no', 'peace...', 1),  
 ('no', 'complain', 32),  
 ('no', 'him.', 31),  
 ('no', 'cuz', 4),  
 ('no', 'he's', 22),  
 ('no', 'orange', 2),  
 ('no', 'tweets', 2),  
 ('no', 'stupid', 2),  
 ('no', 'stuff.', 2),  
 ('yes', '@2mattandgaby2', 1),  
 ('yes', 'no.', 2),  
 ('yes', 'laetrile', 3),  
 ('yes', 'cancer', 122),  
 ('yes', 'anything', 4),  
 ('yes', 'else.', 1),  
 ('yes', 'studied,', 1),  
 ('yes', 'work.', 2),  
 ('yes', 'please', 3),  
 ('yes', 'stop', 10),  
 ('yes', 'spreading', 2),  
 ('yes', 'dangerous', 3),

('yes', 'bullshit.', 1),  
('yes', 'https://t.co/7z0jpa3osp', 1),  
('yes', '@420stonercorner', 2),  
('yes', '.@420stonercorner', 2),  
('yes', 'fascinating', 2),  
('yes', 'subject', 2),  
('yes', 'matter', 3),  
('yes', 'thought', 3),  
('yes', 'pot', 3),  
('yes', 'heads', 2),  
('yes', 'such,', 2),  
('yes', 'rampant', 2),  
('yes', 'family', 5),  
('yes', 'cbd', 5),  
('yes', 'maybe', 4),  
('yes', 'use', 18),  
('yes', 'thc', 3),  
('yes', 'instead', 3),  
('yes', 'morphine!', 2),  
('yes', '#cannabis', 2),  
('yes', '#weeds', 2),  
('yes', '#pmqs', 1),  
('yes', 'medically', 1),  
('yes', 'unqualified', 1),  
('yes', 'opinion?', 1),  
('yes', '#theresamay', 1),  
('yes', '@weedseeds\_uk', 1),  
('yes', '@ukcannabisclub', 1),  
('yes', '@ukcannabisclubs', 1),  
('yes', '@younglabouruk', 1),  
('yes', '#bbcqt', 1),  
('yes', '#bbcaq', 1),  
('yes', 'assholes', 1),  
('yes', 'criminalizing', 1),  
('yes', '@weedfeed', 1),  
('yes', '@norml', 2),  
('yes', '#nyc', 1),  
('yes', '#nypd', 1),  
('yes', '@ukleap', 1),  
('yes', '@uklabour', 1),  
('yes', '@weednucleus', 1),  
('yes', 'cannabis', 15),  
('yes', 'communion', 1),  
('yes', 'brings', 2),  
('yes', 'great', 3),  
('yes', 'peace', 1),  
('yes', 'forever!', 1),  
('yes', 'https://t.co/f583xnt3jy', 1),

('no', '@\_cape\_', 1),  
 ('no', '@\_\_hunchohoudini', 1),  
 ('no', 'provide', 6),  
 ('no', 'food', 11),  
 ('no', 'shelter', 1),  
 ('no', 'homeless', 2),  
 ('no', ',can't', 2),  
 ('no', 'aids', 32),  
 ('no', 'although', 2),  
 ('no', 'it)', 3),  
 ('no', 'shot', 5),  
 ('no', 'tupac.we', 1),  
 ('no', 'investigators', 1),  
 ('no', 'support.', 4),  
 ('no', 'swear', 7),  
 ('no', 'better', 25),  
 ('no', 'sometimes', 1),  
 ('no', 'always', 15),  
 ('no', 'dumb', 2),  
 ('no', 'ppl', 6),  
 ('no', 'fuck', 18),  
 ('no', '@\_dukyfresh', 1),  
 ('no', '@babygrotesque', 1),  
 ('no', 'underlying', 1),  
 ('no', 'economic', 4),  
 ('no', 'problems', 5),  
 ('no', 'contribute', 2),  
 ('no', 'gotham's', 1),  
 ('no', 'crime.', 3),  
 ('no', 'stop', 25),  
 ('no', 'beating', 4),  
 ('no', 'mentally', 2),  
 ('no', 'ill', 5),  
 ('no', 'throwing', 1),  
 ('no', 'inhumane', 1),  
 ('no', 'caricature', 1),  
 ('no', 'mental', 7),  
 ('no', 'facility.', 1),  
 ('no', 'bruce', 1),  
 ('no', 'literally', 18),  
 ('no', 'spidey', 1),  
 ('no', 'meme', 1),  
 ('no', 'tea', 2),  
 ('no', '@\_kingofthehill1:', 1),  
 ('no', 'sexual', 4),  
 ('no', 'https://t.co/qt6n6wrqse', 1),  
 ('yes', '@aacrfoundation:', 1),  
 ('yes', 'leading', 3),

```

('yes', 'discoveries,', 1),
('yes', 'targeting', 3),
('yes', 'cures', 7),
('yes', 'saving', 2),
('yes', 'lives.', 1),
('yes', 'follow', 1),
('yes', 'see', 7),
('yes', 'helping', 1),
('no', '@abaabood', 1),
('no', '@andreas_krieg', 1),
('no', 'me,', 1),
('no', 'spared', 1),
('no', 'efforts', 4),
('no', 'last', 30),
('no', 'plus', 1),
('no', 'resolve', 2),
('no', 'brotherly', 1),
('no', 'basis', 2),
('no', 'qatar', 3),
('no', 'continues', 2),
('no', 'path', 2),
('no', 'destabilizing', 1),
('no', 'saudi', 2),
('no', 'arabia', 2),
('no', 'region.', 1),
('no', 'surgically', 1),
('no', 'remove', 2),
('no', 'regime.', 1),
('no', '@abby_higgins5:', 1),
('no', 'government/pharmacies', 1),
('no', 'wanna', 6),
('no', 'money...', 1),
('no', 'idc', 2),
('no', 'https://t.co/ivompefmo0', 1),
('no', '@ac_breezy', 1),
('no', 'viruses,', 1),
('no', 'spam/trolls', 1),
('no', 'exact', 1),
('no', 'thing...', 1),
('no', 'yet', 16),
('no', 'enough', 12),
('no', 'realize', 6),
('no', 'earned', 1),
('no', 'millennium', 1),
('no', 'prize', 4),
...]
```

```
[ ]: # Count the occurrences of each word
word_counts = train_rdd_word_counts.map(lambda x: (x[0][1], 1)).
    ↪reduceByKey(lambda x, y: x + y)
list_of_words = word_counts.keys().collect()
```

```
[ ]: list_of_words
```

```
Out[22]: ['#armyhivestreamingparty',
'purest',
'thing',
'ever',
'experienced.',
'beyonce',
'bts',
'together',
'cure',
'cancer',
'diseases/illnesses',
'bring',
'world',
'peace',
' ',
'#ayurveda_a_panacea',
'"mother',
'healing"',
'@gurmeetramrahim',
'tells',
'ayurveda',
'first',
'scientific',
'medical',
'branch',
'originated',
'india.',
'yoga',
'practices,',
'becomes',
'boon',
'cancer,',
'kidney',
'disorders,',
'cardiac',
'arrest,',
'etc.',
'https://t.co/7jdqksviqf',
'"mother"',
'healing'.',
'https://t.co/vkj1uq8sod',
```

'#cureforcancer',  
'100',  
'gave',  
'research',  
'teams',  
'back',  
'2014',  
'know',  
'hid',  
'sick',  
'needed',  
'lie',  
'anyone',  
'true',  
'#noname',  
'#executiveorder',  
'sick,',  
'deranged',  
'liberals',  
'happy',  
'president',  
'found',  
'cancer.',  
'morons',  
'say,',  
'"no!",  
'want',  
'people',  
'suffer',  
'die!"',  
'unhinged',  
'nothing',  
'please',  
'them.',  
'thrive',  
'something',  
'bitch',  
'about.',  
'#garethsouthgatewould',  
'solve',  
'hunger',  
'even',  
'take',  
'credit',  
'say',  
'"the',  
'lads',  
'well"',

'#worldcup18',  
'#england',  
'#lgbtactionplan',  
'important',  
'step',  
'become',  
'equally',  
'diverse',  
'place',  
'love.',  
'gay,',  
'anything',  
'straight',  
'curable',  
'cure.',  
'work',  
'hiv',  
'#otd',  
'years',  
'ago,',  
'terry',  
'fox',  
'passed',  
'away',  
'new',  
'westminister,',  
'bc.',  
'speaking',  
'run,',  
'said',  
'dreamer,',  
'saying',  
'initiate',  
'kind',  
'definitive',  
'answer',  
'believe',  
'miracles.',  
'to."',  
'watch',  
'#heritageminute',  
'here.',  
'https://t.co/oqonptqcsn',  
'#predict',  
'&',  
'#cure',  
'next',  
'#cancer',

'https://t.co/gyztoi46yi',  
'minutes',  
'#cancercure',  
'#healed',  
'#reiki',  
'#rtrrt',  
'#bigpharma',  
'#loa',  
'#fridayreads',  
'#nlp',  
'#acim',  
'https://t.co/uqx20hdk8c',  
'https://t.co/oqh0emyqpk',  
'home',  
'#academiczodiac',  
'#o5iris',  
'predicta...',  
'https://t.co/jvbdy7arqy',  
'https://t.co/4r1zgo8opo',  
'https://t.co/ojby25xlyk',  
'#righttotry',  
'1st',  
'lyrics',  
'trump',  
'released?',  
'give',  
'vote',  
'-not',  
'hillary',  
'dr.',  
'video=',  
'blocked',  
'bill',  
'clinton',  
'since',  
'94',  
'#drburzynskirhymeswithlewinski',  
'#trumptrainjumponit',  
' https://t.co/6odiawuwga',  
'via',  
'@youtube',  
'https://t.co/0uta7ccgwk',  
'poor',  
'clinton!',  
'decades',  
'suppressing',  
'cancer-cures',  
'wants',



'#righttocry',  
'#righttodie',  
'ask',  
'patients',  
'that's',  
'right!',  
'rhymes',  
'mistress',  
'-both',  
'f\*cked',  
'https://t.co/xpikezppcz',  
'#shocking',  
'#government',  
'#petitions',  
'#dogs',  
'get',  
'signatures',  
'life',  
'saving',  
'funding',  
'terminal',  
'diseases',  
'known',  
'cause',  
'kill',  
'them! ',  
'#ipf',  
'#pulmonaryfibrosis',  
'needs',  
'like',  
'#cancer!',  
'#humanity?',  
'show',  
'care ',  
'https://t.co/zzcp7rnld0',  
'#sisterscare:',  
'cbs2's",  
'@burrelltv,',  
'@efinchtv',  
'@danatylertv',  
'took',  
'part',  
'sisters',  
'event',  
'today,',  
'addressing',  
'disparities',  
'diagnosis',

'treatment',  
'breast',  
'african',  
'american',  
'women.',  
'https://t.co/unjsqhcycd6',  
'#teachersleading',  
'come',  
'doctor',  
'revolution',  
'teaching',  
'going',  
'teachers',  
'@nickdonohuenmef',  
'#teamhalon',  
'supports',  
'national',  
'#randomactsoflightday',  
'stands',  
'@llsusa',  
'continued',  
'#cancer.',  
'learn',  
'https://t.co/q2lt7y56km',  
'#trump',  
'eradicate',  
'poverty',  
'still',  
'msm',  
'pull',  
'put',  
'negative',  
'spin',  
'it.',  
'war',  
'conflict',  
'preferable',  
'wrong',  
'consistently',  
'are.',  
'#trumpkimsummit',  
'#maga',  
'discover',  
'#trumpderangementsyndrome',  
'sufferers',  
'screaming',  
'injection',  
'stinging.#liberalhypocrisy',

'#singaporeshummit',  
'#trumpkim',  
'#deepstate',  
'#thetorm',  
'self-inflicted',  
'metastasized',  
'full',  
'blown',  
'psychological',  
'insanity',  
'outbreak',  
'across',  
'democrat',  
'party,',  
'liberal',  
'media',  
'nevertrumpers',  
'lead',  
'ultimate',  
'demise.',  
'time',  
'let',  
'go.',  
'#walkaway',  
'\$agen',  
'ceo',  
'@garo\_armen',  
'@agenus\_bio',  
'#wearwhite',  
'@cancerresearch',  
'commitment',  
'fight',  
'#immunotherapy',  
'#oncology',  
'#curecancer',  
'https://t.co/f5vigul6ky',  
'proud',  
'pioneer',  
'#immunotherapy,',  
'@garo\_armen,',  
'stand',  
'courageous',  
'physicians',  
'#inpursuitof',  
'https://t.co/ve97oxqakw',  
'innovation',  
'engine',  
'produced',

'inds',  
'18mos;',  
'planned',  
'2018',  
'1h2019.',  
'molecules',  
'discovery',  
'ind',  
'yrs.',  
'plus,',  
'&gt;100',  
'pts',  
'treated',  
'agen1884',  
'&/or',  
'agen2034',  
'#asco2018.',  
'speed',  
'action!',  
'https://t.co/vrwjomjmvo',  
'driven',  
'powered',  
'right',  
'combination',  
'science',  
'people-see',  
'pursuing',  
'passion',  
'advancing',  
'toward',  
'goal',  
'story:',  
'https://t.co/apxu2dplkm',  
'@big3biobos',  
'\$celz',  
'imho,',  
'beaten',  
'many',  
'countries.;',  
'rate',  
'95%+.',  
'#fda',  
'big',  
'#pharma',  
'keep',  
'treatments',  
'illegal',  
'u.s',

'capitalize',  
'\$150,000-\$500,000',  
'unneeded',  
'surgeries',  
'chemo',  
'celz's',  
'#amniostem',  
'help',  
'#cachexia',  
'huge!',  
'&gt;',  
'advice',  
'isn't',  
'helpful.',  
'i'm',  
'sure',  
'it's',  
'wanting',  
'helpful,',  
'sucks.',  
'good',  
'deeds',  
'faith',  
'quinoa',  
'hope',  
'impacted',  
'worse,',  
'none',  
'around',  
'surgery',  
'days.',  
'&gt;"the',  
'idea',  
'tell',  
'story',  
'lesnar',  
'looking',  
'leaving',  
'belt',  
'ufc"',  
'garner',  
'sympathy',  
'reigns.',  
'roman',  
'reigns',  
'end',  
'hunger,',  
'find',

'booed',  
'pushed',  
'hard.',  
'https://t.co/uxful5sx6q',  
'\*breaking',  
'news\*',  
'(hypothetical',  
'situation)',  
'finds',  
'cancer!',  
'everyone',  
'mother:',  
'"what',  
'piece',  
'shit!',  
'probably',  
'knew',  
'whole',  
'time!"',  
'dont',  
'agree',  
'shit',  
'guy',  
'does/says,',  
'ya'll",  
'never',  
'satisfied',  
'\*opens',  
'cab-booking',  
'app\*',  
'cab',  
'\*books',  
'cab\*',  
'years,',  
'months,',  
'weeks,',  
'days',  
'hours.',  
'kids',  
'born',  
'grow',  
'knowing',  
'mother,',  
'scientists',  
'volcano',  
'erupt',  
'wipe',  
'hawaii',

'away.',  
'\*this',  
'chimes\*',  
'positivity',  
'can't',  
'neither',  
'fighting',  
'talk.',  
'"battle"',  
'narrative',  
'created',  
'facing',  
'disease',  
'denied',  
'narratives.',  
'couldn't',  
'more\*',  
'#lifewithcancer',  
'https://t.co/akj8zdd3jz',  
'@abbvie',  
'alphabet-backed',  
'calico',  
'spend',  
'another',  
'\$1b',  
'https://t.co/a7tqwrwtrv',  
'@erictrump:',  
'"we're',  
'raising',  
'money',  
'we're',  
'win',  
'fight.',  
'pediatric',  
'cancer."',  
'https://t.co/7pyfcd7xzy',  
'thank',  
'@erictrump',  
'son',  
'alive',  
'due',  
'started',  
'jude.',  
'year',  
'survivor',  
'week.',  
'loose',  
'least',

'every',  
'month',  
'local',  
'children's',  
'hospital',  
'childhood',  
'https://t.co/gwqtxflqau',  
'11:11',  
'excellent',  
'grades!!!',  
'job',  
'best',  
'friend',  
'happiness',  
'friends',  
'gun',  
'control',  
'@arianagrande',  
'notice',  
'@shawnmendes',  
'follow',  
'bye',  
'anxiety',  
'fun',  
'summer',  
'lion',  
'rid',  
'depression',  
'hug',  
'love',  
'grades',  
'summer!',  
'harry',  
'potter',  
'reunion',  
'weight!',  
'16-year-old',  
'jeremiah',  
'thomas,',  
'battling',  
'used',  
'make',  
'wish',  
'abortion.',  
'you,',  
'jeremiah,',  
'beautiful',  
'life,',



'inspiring',  
'many.',  
'https://t.co/bumzmggyay',  
'sad',  
'young',  
'brainwashed',  
'maybe',  
'health',  
'care',  
'all',  
'imaginary',  
'exist',  
'https://t.co/27twcb3oek',  
'"selfless',  
'way"',  
'idk',  
'mann',  
'pretty',  
'selfish',  
'try',  
'use',  
'religion',  
'controls',  
'someone',  
'else's',  
'body...',  
'coulda',  
'wished',  
'https://t.co/tyx3hynfa8',  
'17yr',  
'stage',  
'carcinoid',  
'survivor.',  
'longest',  
'surviving',  
'patient',  
'world.',  
'survival',  
'5yrs.',  
'plz',  
'read',  
'remarkable',  
'miracle',  
'book',  
'living',  
'silent',  
'enemy',  
'coming',

'soon.',  
'support',  
'mdacc.',  
'inoperable',  
'neuroendocrine',  
'misdiagnosed',  
'lets',  
'cancer#',  
'17yr.',  
'endcancer#',  
'1965:',  
'bet',  
'there'll",  
'flying',  
'cars',  
'future"',  
'1987:',  
'future',  
'2018:',  
'prototype',  
'potentially',  
'viable',  
'cures',  
'cancers',  
'given',  
'1998:',  
'"in',  
'twenty',  
'cancer!""',  
'"dennis',  
'rodman',  
'saved',  
'nuclear',  
'celebrity',  
'apprentice,',  
'hosted',  
'united',  
'states,',  
'dictator',  
'north',  
'korea!""',  
'@panmass',  
'rider',  
'@danafarber.',  
'fund',  
'retweet! ',  
'https://t.co/qdwqn3zzre',  
'https://t.co/lafbipfydc',

'3mt',  
'semi-finalist',  
'ubc',  
'master',  
'student',  
'evan',  
'warner',  
'presented',  
"shakespeare's",  
'much',  
'thing?',  
'asked.',  
'#tbt',  
'talk:',  
'https://t.co/vjazjwo5em',  
'22,',  
'ethan',  
'hawes',  
'diagnosed',  
'multiple',  
'myeloma,',  
'typically',  
'afflicting',  
'age',  
'65.',  
'hospitalizations,',  
'chemotherapy,',  
'stem',  
'cell',  
'replacement',  
'followed.',  
'https://t.co/f0lztvodcb',  
'5,000',  
'young',  
'women',  
'country',  
'participated',  
'year's',  
'lax',  
'raise',  
'funds',  
'awareness',  
'tournament',  
'raised',  
'\$235,000',  
'komen',  
'csnj',  
'@komennja.',

'more:',  
'https://t.co/4qlw3bei7k',  
'https://t.co/uwdnketoie',  
'6th',  
'40tude',  
'thames',  
'sup',  
'marathon',  
'saturday',  
'already',  
'£63k',  
'marks',  
'colon',  
'add',  
'£180k',  
'past',  
'years.',  
'miles',  
'sups',  
'finishing',  
'putney.',  
'@curebowelcancer',  
'@active360',  
'https://t.co/dn5cjvm9gz',  
'87%',  
'f-ing',  
'republicans',  
'trump.',  
'omg.',  
'how?',  
'children',  
'ripped',  
'mother's',  
'arms.',  
'truly',  
'80%',  
'news',  
'20%',  
'don't',  
'on.',  
'#disgusted',  
'must',  
'accept',  
'sinister',  
'brainwashing.',  
'comprehend',  
'maga',  
'grasp',

'shooting',  
'afternoon',  
'nature.',  
'one',  
'voting',  
'overwhelming',  
'numbers.',  
'https://t.co/s7ctd0lyc6',  
'@1405scituateave',  
'@realjameswoods',  
'funny',  
'is,',  
'discovered',  
'tomorrow,',  
'donald',  
'telling',  
'somehow',  
'instrumental',  
'it,',  
'jackass',  
'trumpuppets',  
'word',  
'it!!!!',  
'...and',  
'you're',  
'saying;',  
'"maybe',  
'be",',  
'now!!',  
'@1776stonewall',  
'pathetic',  
'wasn't',  
'enough!!',  
'talk',  
'pitiful',  
'souls',  
'way',  
'trump!!!!',  
'@1999mcm11',  
'denuclearization',  
'achieved.',  
'gets',  
'taste',  
'western',  
'culture,',  
'mcdonalds,',  
'coke,',  
'xbox,',

'etc...',  
'over.',  
'achieve',  
'peace...',  
'complain',  
'him.',  
'cuz',  
'he's',  
'orange',  
'tweets',  
'stupid',  
'stuff.',  
'@2mattandgaby2',  
'no.',  
'laetrile',  
'else.',  
'studied,',  
'work.',  
'stop',  
'spreading',  
'dangerous',  
'bullshit.',  
'https://t.co/7z0jpa3osp',  
'@420stonercorner',  
'.@420stonercorner',  
'fascinating',  
'subject',  
'matter',  
'thought',  
'pot',  
'heads',  
'such,',  
'rampant',  
'family',  
'cbd',  
'thc',  
'instead',  
'morphine!',  
'#cannabis',  
'#weeds',  
'#pmqs',  
'medically',  
'unqualified',  
'opinion?',  
'#theresamay',  
'@weedseeds\_uk',  
'@ukcannabisclub',  
'@ukcannabisclubs',

'@younglabouruk',  
'#bbcqt',  
'#bbcaq',  
'assholes',  
'criminalizing',  
'@weedfeed',  
'@norml',  
'#nyc',  
'#nypd',  
'@ukleap',  
'@uklabour',  
'@weednucleus',  
'cannabis',  
'communion',  
'brings',  
'great',  
'forever!',  
'https://t.co/f583xnt3jy',  
'@\_cape\_',  
'@\_\_hunchohoudini',  
'provide',  
'food',  
'shelter',  
'homeless',  
'can't',  
'aids',  
'although',  
'it)',  
'shot',  
'tupac.we',  
'investigators',  
'support.',  
'swear',  
'better',  
'sometimes',  
'always',  
'dumb',  
'ppl',  
'fuck',  
'@\_duckyfresh',  
'@babygrotesque',  
'underlying',  
'economic',  
'problems',  
'contribute',  
'gotham's',  
'crime.',  
'beating',

'mentally',  
'ill',  
'throwing',  
'inhumane',  
'caricature',  
'mental',  
'facility.',  
'bruce',  
'literally',  
'spidey',  
'meme',  
'tea',  
'@\_kingofthehill1:',  
'sexual',  
'https://t.co/qt6n6wrqse',  
'@aacrfoundation:',  
'leading',  
'discoveries,',  
'targeting',  
'lives.',  
'see',  
'helping',  
'@abaabood',  
'@andreas\_krieg',  
'me,',  
'spared',  
'efforts',  
'last',  
'plus',  
'resolve',  
'brotherly',  
'basis',  
'qatar',  
'continues',  
'path',  
'destabilizing',  
'saudi',  
'arabia',  
'region.',  
'surgically',  
'remove',  
'regime.',  
'@abby\_higgins5:',  
'government/pharmacies',  
'wanna',  
'money...',  
'idc',  
'https://t.co/ivompefmo0',



'@ac\_breezy',  
'viruses',  
'spam/trolls',  
'exact',  
'thing...',  
'yet',  
'enough',  
'realize',  
'earned',  
'millennium',  
'prize',  
'couple',  
'hundred',  
'times',  
'think',  
'https://t.co/29f4udvenv',  
'@acosta',  
'real.',  
'left',  
'vitriol',  
'hatred',  
'wrong.',  
'pity',  
'miserable',  
'people.',  
'@adeabdwhb',  
'tweak',  
'question',  
'little',  
'bit',  
'imagine',  
'youre',  
'dying.',  
'rich',  
'man',  
'giving',  
'choices:',  
'huge',  
'amount',  
'money.',  
'basically',  
'cant',  
'god',  
'choose',  
'bcause',  
'choose.',  
'@adevotedyogi:',  
...]

```
[ ]: word_counts.collect()
```

```
Out[23]: [(' #armyhivestreamingparty', 1),
 ('purest', 2),
 ('thing', 24),
 ('ever', 18),
 ('experienced.', 1),
 ('beyonce', 1),
 ('bts', 3),
 ('together', 16),
 ('cure', 1327),
 ('cancer', 938),
 ('diseases/illnesses', 1),
 ('bring', 27),
 ('world', 108),
 ('peace', 35),
 (' ', 1),
 (' #ayurveda_a_panacea', 3),
 ('"mother', 1),
 ('healing"', 1),
 ('@gurmeetramrahim', 2),
 ('tells', 3),
 ('ayurveda', 4),
 ('first', 27),
 ('scientific', 10),
 ('medical', 27),
 ('branch', 2),
 ('originated', 2),
 ('india.', 2),
 ('yoga', 3),
 ('practices,', 2),
 ('becomes', 2),
 ('boon', 2),
 ('cancer,', 183),
 ('kidney', 8),
 ('disorders,', 3),
 ('cardiac', 2),
 ('arrest,', 2),
 ('etc.', 10),
 ('https://t.co/7jdqksviqf', 1),
 ('"mother"', 1),
 ('healing'.', 1),
 ('https://t.co/vkj1uq8sod', 1),
 ('#cureforcancer', 1),
 ('100', 3),
 ('gave', 16),
 ('research', 51),
 ('teams', 2),
```

('back', 28),  
('2014', 2),  
('know', 60),  
('hid', 1),  
('sick', 9),  
('needed', 7),  
('lie', 7),  
('anyone', 20),  
('true', 9),  
('#noname', 1),  
('#executiveorder', 2),  
('sick,', 1),  
('deranged', 1),  
('liberals', 23),  
('happy', 18),  
('president', 52),  
('found', 68),  
('cancer.', 162),  
('morons', 2),  
('say,', 8),  
('"no!', 1),  
('want', 79),  
('people', 173),  
('suffer', 5),  
('die!'", 1),  
('unhinged', 2),  
('nothing', 20),  
('please', 62),  
('them.', 13),  
('thrive', 2),  
('something', 50),  
('bitch', 12),  
('about.', 7),  
('#garethsouthgatewould', 1),  
('solve', 14),  
('hunger', 25),  
('even', 74),  
('take', 32),  
('credit', 8),  
('say', 70),  
('"the', 5),  
('lads', 1),  
('well"', 1),  
('#worldcup18', 1),  
('#england', 1),  
('#lgbtactionplan', 1),  
('important', 10),  
('step', 6),

('become', 11),  
('equally', 1),  
('diverse', 1),  
('place', 4),  
('love.', 5),  
('gay,', 1),  
('anything', 27),  
('straight', 1),  
('curable', 5),  
('cure.', 12),  
('work', 29),  
('hiv', 15),  
('#otd', 1),  
('years', 52),  
('ago,', 4),  
('terry', 2),  
('fox', 4),  
('passed', 3),  
('away', 19),  
('new', 25),  
('westminister,', 1),  
('bc.', 1),  
('speaking', 3),  
('run,', 1),  
('said', 26),  
('dreamer,', 1),  
('saying', 19),  
('initiate', 2),  
('kind', 12),  
('definitive', 1),  
('answer', 6),  
('believe', 41),  
('miracles.', 1),  
('to."', 1),  
('watch', 11),  
('#heritageminute', 1),  
('here.', 3),  
('https://t.co/oqonptqcs', 1),  
('#predict', 2),  
('&', 123),  
('#cure', 2),  
('next', 15),  
('#cancer', 32),  
('https://t.co/gyztoi46yi', 1),  
('minutes', 3),  
('#cancercure', 8),  
('#healed', 2),  
('#reiki', 2),

('#rtrrt', 4),  
 ('#bigpharma', 4),  
 ('#loa', 2),  
 ('#fridayreads', 2),  
 ('#nlp', 2),  
 ('#acim', 2),  
 ('https://t.co/uqx20hdk8c', 1),  
 ('https://t.co/oqh0emyqpk', 1),  
 ('home', 10),  
 ('#academiczodiac', 2),  
 ('#o5iris', 2),  
 ('predicta...', 2),  
 ('https://t.co/jvbdy7arqy', 1),  
 ('https://t.co/4r1zgo8opo', 1),  
 ('https://t.co/ojby25xlyk', 1),  
 ('#righttotry', 2),  
 ('1st', 7),  
 ('lyrics', 1),  
 ('trump', 203),  
 ('released?', 1),  
 ('give', 56),  
 ('vote', 8),  
 ('-not', 1),  
 ('hillary', 3),  
 ('dr.', 11),  
 ('video=', 1),  
 ('blocked', 2),  
 ('bill', 6),  
 ('clinton', 2),  
 ('since', 17),  
 ('94', 1),  
 ('#drburzynskirhymeswithlewinski', 2),  
 ('#trumptrainjumponit', 1),  
 (' https://t.co/6odiawuwga', 1),  
 ('via', 16),  
 ('@youtube', 3),  
 ('https://t.co/0uta7ccgwk', 1),  
 ('poor', 7),  
 ('clinton!', 1),  
 ('decades', 4),  
 ('suppressing', 1),  
 ('cancer-cures', 1),  
 ('wants', 7),  
 ('#righttocry', 1),  
 ('#righttodie', 1),  
 ('ask', 15),  
 ('patients', 28),  
 ('that's', 24),

('right!', 5),  
 ('rhymes', 1),  
 ('mistress', 1),  
 ('-both', 1),  
 ('f\*cked', 1),  
 ('https://t.co/xpikezppcz', 1),  
 ('#shocking', 1),  
 ('#government', 1),  
 ('#petitions', 1),  
 ('#dogs', 1),  
 ('get', 97),  
 ('signatures', 1),  
 ('life', 44),  
 ('saving', 7),  
 ('funding', 5),  
 ('terminal', 7),  
 ('diseases', 25),  
 ('known', 14),  
 ('cause', 23),  
 ('kill', 15),  
 ('them! ', 1),  
 ('#ipf', 2),  
 ('#pulmonaryfibrosis', 2),  
 ('needs', 19),  
 ('like', 155),  
 ('#cancer!', 2),  
 ('#humanity?', 1),  
 ('show', 13),  
 ('care ', 1),  
 ('https://t.co/zzcp7rnld0', 1),  
 ('#sisterscare:', 1),  
 ('cbs2's', 2),  
 ('@burrelltv', 1),  
 ('@efinchtv', 2),  
 ('@danatylertv', 2),  
 ('took', 12),  
 ('part', 13),  
 ('sisters', 2),  
 ('event', 13),  
 ('today,', 7),  
 ('addressing', 1),  
 ('disparities', 1),  
 ('diagnosis', 6),  
 ('treatment', 27),  
 ('breast', 35),  
 ('african', 4),  
 ('american', 12),  
 ('women.', 4),

('https://t.co/unjsqhcyd6', 1),  
 ('#teachersleading', 1),  
 ('come', 32),  
 ('doctor', 11),  
 ('revolution', 2),  
 ('teaching', 3),  
 ('going', 39),  
 ('teachers', 2),  
 ('@nickdonohuenmef', 1),  
 ('#teamhalon', 1),  
 ('supports', 3),  
 ('national', 3),  
 ('#randomactsoflightday', 1),  
 ('stands', 2),  
 ('@llsusa', 1),  
 ('continued', 2),  
 ('#cancer.', 9),  
 ('learn', 12),  
 ('https://t.co/q2lt7y56km', 1),  
 ('#trump', 9),  
 ('eradicate', 5),  
 ('poverty', 10),  
 ('still', 158),  
 ('msm', 10),  
 ('pull', 4),  
 ('put', 35),  
 ('negative', 10),  
 ('spin', 9),  
 ('it.', 65),  
 ('war', 8),  
 ('conflict', 2),  
 ('preferable', 1),  
 ('wrong', 15),  
 ('consistently', 1),  
 ('are.', 5),  
 ('#trumpkimsummit', 15),  
 ('#maga', 15),  
 ('discover', 9),  
 ('#trumpderangementsyndrome', 5),  
 ('sufferers', 3),  
 ('screaming', 1),  
 ('injection', 1),  
 ('stinging.#liberalhypocrisy', 1),  
 ('#singaporesummit', 2),  
 ('#trumpkim', 1),  
 ('#deepstate', 1),  
 ('#thestorm', 1),  
 ('self-inflicted', 1),

('metastasized', 2),  
 ('full', 16),  
 ('blown', 1),  
 ('psychological', 1),  
 ('insanity', 2),  
 ('outbreak', 1),  
 ('across', 11),  
 ('democrat', 4),  
 ('party,', 1),  
 ('liberal', 12),  
 ('media', 28),  
 ('nevertrumpers', 1),  
 ('lead', 13),  
 ('ultimate', 2),  
 ('demise.', 2),  
 ('time', 50),  
 ('let', 22),  
 ('go.', 2),  
 ('#walkaway', 2),  
 ('\$agen', 6),  
 ('ceo', 2),  
 ('@garo\_armen', 1),  
 ('@agenus\_bio', 1),  
 ('#wearwhite', 2),  
 ('@cancerresearch', 2),  
 ('commitment', 1),  
 ('fight', 39),  
 ('#immunotherapy', 4),  
 ('#oncology', 2),  
 ('#curecancer', 1),  
 ('https://t.co/f5vigul6ky', 1),  
 ('proud', 16),  
 ('pioneer', 1),  
 ('#immunotherapy,', 1),  
 ('@garo\_armen,', 1),  
 ('stand', 9),  
 ('courageous', 2),  
 ('physicians', 2),  
 ('#inpursuitof', 3),  
 ('https://t.co/ve97oxqakw', 1),  
 ('innovation', 5),  
 ('engine', 1),  
 ('produced', 1),  
 ('inds', 2),  
 ('18mos;', 1),  
 ('planned', 3),  
 ('2018', 11),  
 ('1h2019.', 1),



('molecules', 1),  
('discovery', 3),  
('ind', 1),  
('yrs.', 3),  
('plus,', 2),  
('>100', 1),  
('pts', 8),  
('treated', 7),  
('agen1884', 1),  
('&/or', 1),  
('agen2034', 1),  
('#asco2018.', 1),  
('speed', 5),  
('action!', 1),  
('https://t.co/vrwjomjmvo', 1),  
('driven', 3),  
('powered', 3),  
('right', 26),  
('combination', 3),  
('science', 18),  
('people-see', 3),  
('pursuing', 3),  
('passion', 6),  
('advancing', 3),  
('toward', 4),  
('goal', 7),  
('story:', 4),  
('https://t.co/apxu2dplkm', 3),  
('@big3biobos', 3),  
('\$celz', 1),  
('imho,', 1),  
('beaten', 4),  
('many', 67),  
('countries.:', 1),  
('rate', 20),  
('95%+', 1),  
('#fda', 1),  
('big', 19),  
('#pharma', 1),  
('keep', 30),  
('treatments', 17),  
('illegal', 8),  
('u.s', 2),  
('capitalize', 1),  
('\$150,000-\$500,000', 1),  
('unneeded', 1),  
('surgeries', 5),  
('chemo', 17),

('celz's', 1),  
 ('#amniostem', 1),  
 ('help', 124),  
 ('#cachexia', 1),  
 ('huge!', 1),  
 ('&gt;', 2),  
 ('advice', 2),  
 ('isn't', 13),  
 ('helpful.', 1),  
 ('i'm', 42),  
 ('sure', 17),  
 ('it's', 73),  
 ('wanting', 1),  
 ('helpful,', 1),  
 ('sucks.', 7),  
 ('good', 54),  
 ('deeds', 1),  
 ('faith', 7),  
 ('quinoa', 2),  
 ('hope', 41),  
 ('impacted', 1),  
 ('worse,', 1),  
 ('none', 5),  
 ('around', 9),  
 ('surgery', 6),  
 ('days.', 4),  
 ('&gt;"the', 1),  
 ('idea', 7),  
 ('tell', 20),  
 ('story', 15),  
 ('lesnar', 1),  
 ('looking', 15),  
 ('leaving', 4),  
 ('belt', 1),  
 ('ufc"', 1),  
 ('garner', 1),  
 ('sympathy', 2),  
 ('reigns.', 1),  
 ('roman', 1),  
 ('reigns', 1),  
 ('end', 57),  
 ('hunger,', 12),  
 ('find', 210),  
 ('booed', 1),  
 ('pushed', 1),  
 ('hard.', 1),  
 ('https://t.co/uxful5sx6q', 1),  
 ('\*breaking', 1),

('news\*', 1),  
('hypothetical', 1),  
('situation', 1),  
('finds', 7),  
('cancer!', 31),  
('everyone', 30),  
('mother:', 1),  
('"what', 2),  
('piece', 6),  
('shit!', 1),  
('probably', 15),  
('knew', 8),  
('whole', 12),  
('time!'", 1),  
('dont', 8),  
('agree', 6),  
('shit', 23),  
('guy', 14),  
('does/says,', 1),  
("ya'll", 1),  
('never', 60),  
('satisfied', 2),  
('\*opens', 1),  
('cab-booking', 1),  
('app\*', 1),  
('cab', 2),  
('\*books', 1),  
('cab\*', 1),  
('years,', 6),  
('months,', 1),  
('weeks,', 1),  
('days', 9),  
('hours.', 1),  
('kids', 25),  
('born', 6),  
('grow', 16),  
('knowing', 11),  
('mother,', 1),  
('scientists', 24),  
('volcano', 1),  
('erupt', 1),  
('wipe', 4),  
('hawaii', 1),  
('away.', 1),  
('\*this', 1),  
('chimes\*', 1),  
('positivity', 3),  
('can't', 32),

('neither', 4),  
 ('fighting', 9),  
 ('talk.', 2),  
 ('"battle"', 1),  
 ('narrative', 3),  
 ('created', 5),  
 ('facing', 1),  
 ('disease', 52),  
 ('denied', 2),  
 ('narratives.', 1),  
 ('couldn't', 5),  
 ('more\*', 1),  
 ('#lifewithcancer', 1),  
 ('https://t.co/akj8zdd3jz', 1),  
 ('.@abbvie', 2),  
 ('alphabet-backed', 2),  
 ('calico', 2),  
 ('spend', 8),  
 ('another', 14),  
 ('\$1b', 2),  
 ('https://t.co/a7tqwrwtrv', 2),  
 ('.@erictrump', 1),  
 ('"we're', 1),  
 ('raising', 14),  
 ('money', 80),  
 ('we're', 11),  
 ('win', 16),  
 ('fight.', 2),  
 ('pediatric', 12),  
 ('cancer."', 3),  
 ('https://t.co/7pyfcd7xzy', 1),  
 ('thank', 41),  
 ('@erictrump', 2),  
 ('son', 7),  
 ('alive', 2),  
 ('due', 4),  
 ('started', 5),  
 ('jude.', 1),  
 ('year', 31),  
 ('survivor', 12),  
 ('week.', 4),  
 ('loose', 2),  
 ('least', 6),  
 ('every', 54),  
 ('month', 6),  
 ('local', 7),  
 ('children's', 2),  
 ('hospital', 8),

('childhood', 25),  
('https://t.co/gwqtxflqau', 1),  
('11:11', 4),  
('excellent', 4),  
('grades!!!', 2),  
('job', 14),  
('best', 23),  
('friend', 12),  
('happiness', 7),  
('friends', 15),  
('gun', 7),  
('control', 12),  
('@arianagrande', 4),  
('notice', 5),  
('@shawnmendes', 4),  
('follow', 10),  
('bye', 5),  
('anxiety', 12),  
('fun', 9),  
('summer', 4),  
('lion', 9),  
('rid', 10),  
('depression', 22),  
('hug', 6),  
('love', 28),  
('grades', 1),  
('summer!', 1),  
('harry', 3),  
('potter', 1),  
('reunion', 3),  
('weight!', 1),  
('16-year-old', 4),  
('jeremiah', 4),  
('thomas,', 3),  
('battling', 6),  
('used', 14),  
('make', 92),  
('wish', 17),  
('abortion.', 5),  
('you,', 10),  
('jeremiah,', 3),  
('beautiful', 9),  
('life,', 14),  
('inspiring', 7),  
('many.', 4),  
('https://t.co/bumzmggyay', 3),  
('sad', 17),  
('young,', 1),

('brainwashed', 2),  
 ('maybe', 25),  
 ('health', 31),  
 ('care', 21),  
 ('all,', 4),  
 ('imaginary', 1),  
 ('exist', 6),  
 ('https://t.co/27twcb3oek', 1),  
 ("selfless", 1),  
 ('way"', 1),  
 ('idk', 1),  
 ('mann', 1),  
 ('pretty', 5),  
 ('selfish', 1),  
 ('try', 20),  
 ('use', 28),  
 ('religion', 11),  
 ('controls', 1),  
 ('someone', 34),  
 ('else's', 1),  
 ('body...', 1),  
 ('coulda', 1),  
 ('wished', 4),  
 ('https://t.co/tyx3hynfa8', 1),  
 ('17yr', 3),  
 ('stage', 23),  
 ('carcinoid', 12),  
 ('survivor.', 4),  
 ('longest', 2),  
 ('surviving', 2),  
 ('patient', 8),  
 ('world.', 15),  
 ('survival', 12),  
 ('5yrs.', 6),  
 ('plz', 14),  
 ('read', 26),  
 ('remarkable', 10),  
 ('miracle', 13),  
 ('book', 13),  
 ('living', 17),  
 ('silent', 11),  
 ('enemy', 11),  
 ('coming', 18),  
 ('soon.', 12),  
 ('support', 32),  
 ('mdacc.', 1),  
 ('inoperable', 5),  
 ('neuroendocrine', 7),

('misdiagnosed', 7),  
 ('lets', 3),  
 ('cancer#', 1),  
 ('17yr.', 1),  
 ('endcancer#', 1),  
 ('1965:', 2),  
 ('bet', 10),  
 ("there'll", 4),  
 ('flying', 8),  
 ('cars', 7),  
 ('future"', 1),  
 ('1987:', 1),  
 ('future', 14),  
 ('2018:', 2),  
 ('prototype', 1),  
 ('potentially', 5),  
 ('viable', 3),  
 ('cures', 19),  
 ('cancers', 17),  
 ('given', 6),  
 ('1998:', 2),  
 ('"in', 1),  
 ('twenty', 2),  
 ('cancer!'", 1),  
 ('"dennis', 1),  
 ('rodman', 1),  
 ('saved', 5),  
 ('nuclear', 4),  
 ('celebrity', 2),  
 ('apprentice,', 1),  
 ('hosted', 2),  
 ('united', 3),  
 ('states,', 1),  
 ('dictator', 3),  
 ('north', 8),  
 ('korea!'", 1),  
 ('@panmass', 1),  
 ('rider', 1),  
 ('@danafarber.', 1),  
 ('fund', 14),  
 ('retweet!', 1),  
 ('https://t.co/qdwqn3zzre', 1),  
 ('https://t.co/lafbipfydc', 1),  
 ('3mt', 1),  
 ('semi-finalist', 1),  
 ('ubc', 1),  
 ('master', 1),  
 ('student', 2),

('evan', 1),  
 ('warner', 1),  
 ('presented', 1),  
 ("shakespeare's", 1),  
 ('much', 51),  
 ('thing?', 2),  
 ('asked.', 1),  
 ('#tbt', 1),  
 ('talk:', 1),  
 ('https://t.co/vjazjwo5em', 1),  
 ('22,', 1),  
 ('ethan', 1),  
 ('hawes', 1),  
 ('diagnosed', 10),  
 ('multiple', 12),  
 ('myeloma,', 1),  
 ('typically', 1),  
 ('afflicting', 2),  
 ('age', 4),  
 ('65.', 1),  
 ('hospitalizations,', 1),  
 ('chemotherapy,', 3),  
 ('stem', 4),  
 ('cell', 7),  
 ('replacement', 1),  
 ('followed.', 1),  
 ('https://t.co/f0lztvodcb', 1),  
 ('5,000', 2),  
 ('young', 13),  
 ('women', 17),  
 ('country', 16),  
 ('participated', 2),  
 ('year's', 4),  
 ('lax', 1),  
 ('raise', 24),  
 ('funds', 12),  
 ('awareness', 13),  
 ('tournament', 2),  
 ('raised', 16),  
 ('\$235,000', 1),  
 ('komen', 2),  
 ('csnj', 1),  
 ('@komennja.', 1),  
 ('more:', 2),  
 ('https://t.co/4qlw3bei7k', 1),  
 ('https://t.co/uwdnketoie', 1),  
 ('6th', 4),  
 ('40tude', 3),



('thames', 2),  
('sup', 1),  
('marathon', 3),  
('saturday', 2),  
('already', 20),  
('£63k', 1),  
('marks', 1),  
('colon', 4),  
('add', 7),  
('£180k', 1),  
('past', 5),  
('years.', 10),  
('miles', 2),  
('sups', 1),  
('finishing', 1),  
('putney.', 1),  
('@curebowelcancer', 1),  
('@active360', 1),  
('https://t.co/dn5cjvm9gz', 1),  
('87%', 2),  
('f-ing', 2),  
('republicans', 6),  
('trump.', 9),  
('omg.', 1),  
('how?', 2),  
('children', 33),  
('ripped', 1),  
('mother's', 2),  
('arms.', 1),  
('truly', 11),  
('80%', 2),  
('news', 11),  
('20%', 3),  
('don't', 58),  
('on.', 5),  
('#disgusted', 1),  
('must', 21),  
('accept', 4),  
('sinister', 1),  
('brainwashing.', 1),  
('comprehend', 2),  
('maga', 3),  
('grasp', 1),  
('shooting', 2),  
('afternoon', 1),  
('nature.', 2),  
('one', 100),  
('voting', 3),

('overwhelming', 3),  
('numbers.', 1),  
('https://t.co/s7ctd0lyc6', 1),  
('@1405scituateave', 1),  
('@realjameswoods', 4),  
('funny', 7),  
('is,', 7),  
('discovered', 17),  
('tomorrow,', 3),  
('donald', 12),  
('telling', 14),  
('somehow', 5),  
('instrumental', 1),  
('it,', 23),  
('jackass', 1),  
('trumpuppets', 1),  
('word', 4),  
('it!!!!', 1),  
('...and', 1),  
('you're', 12),  
('saying;', 1),  
('"maybe', 1),  
('be"', 1),  
('now!!', 1),  
('@1776stonewall', 2),  
('pathetic', 1),  
('wasn't', 1),  
('enough!!', 1),  
('talk', 27),  
('pitiful', 1),  
('souls', 2),  
('way', 47),  
('trump!!!!', 1),  
('@1999mcm11', 1),  
('denuclearization', 3),  
('achieved.', 2),  
('gets', 6),  
('taste', 1),  
('western', 3),  
('culture,', 1),  
('mcdonalds,', 1),  
('coke,', 1),  
('xbox,', 1),  
('etc...', 1),  
('over.', 1),  
('achieve', 5),  
('peace...', 1),  
('complain', 32),

('him.', 31),  
('cuz', 4),  
('he's', 22),  
('orange', 2),  
('tweets', 2),  
('stupid', 2),  
('stuff.', 2),  
('@2mattandgaby2', 1),  
('no.', 5),  
('laetrile', 3),  
('else.', 2),  
('studied,', 1),  
('work.', 8),  
('stop', 35),  
('spreading', 9),  
('dangerous', 3),  
('bullshit.', 1),  
('https://t.co/7z0jpa3osp', 1),  
('@420stonercorner', 2),  
('.@420stonercorner', 2),  
('fascinating', 2),  
('subject', 2),  
('matter', 27),  
('thought', 16),  
('pot', 3),  
('heads', 3),  
('such,', 3),  
('rampant', 2),  
('family', 26),  
('cbd', 6),  
('thc', 3),  
('instead', 16),  
('morphine!', 2),  
('#cannabis', 2),  
('#weeds', 2),  
('#pmqs', 1),  
('medically', 1),  
('unqualified', 1),  
('opinion?', 1),  
('#theresamay', 1),  
('@weedseeds\_uk', 1),  
('@ukcannabisclub', 1),  
('@ukcannabisclubs', 1),  
('@younglabouruk', 1),  
('#bbcqt', 1),  
('#bbcaq', 1),  
('assholes', 2),  
('criminalizing', 1),

('@weedfeed', 1),  
('@norml', 2),  
('#nyc', 1),  
('#nypd', 1),  
('@ukleap', 1),  
('@uklabour', 1),  
('@weednucleus', 1),  
('cannabis', 16),  
('communion', 1),  
('brings', 3),  
('great', 34),  
('forever!', 1),  
('https://t.co/f583xnt3jy', 1),  
('@\_cape\_', 1),  
('@\_\_hunchohoudini', 1),  
('provide', 7),  
('food', 16),  
('shelter', 1),  
('homeless', 2),  
(',can't', 2),  
('aids', 34),  
('although', 3),  
('it)', 4),  
('shot', 5),  
('tupac.we', 1),  
('investigators', 1),  
('support.', 4),  
('swear', 7),  
('better', 30),  
('sometimes', 2),  
('always', 17),  
('dumb', 3),  
('ppl', 11),  
('fuck', 18),  
('@\_duckyfresh', 1),  
('@babygrotesque', 1),  
('underlying', 1),  
('economic', 5),  
('problems', 5),  
('contribute', 2),  
("gotham's", 1),  
('crime.', 3),  
('beating', 4),  
('mentally', 2),  
('ill', 6),  
('throwing', 1),  
('inhumane', 1),  
('caricature', 1),

('mental', 8),  
('facility.', 1),  
('bruce', 1),  
('literally', 20),  
('spidey', 1),  
('meme', 2),  
('tea', 2),  
('@\_kingofthehill1:', 1),  
('sexual', 4),  
('https://t.co/qt6n6wrqse', 1),  
('@aacrfoundation:', 1),  
('leading', 7),  
('discoveries,', 1),  
('targeting', 3),  
('lives.', 4),  
('see', 41),  
('helping', 14),  
('@abaabood', 1),  
('@andreas\_krieg', 1),  
('me,', 2),  
('spared', 1),  
('efforts', 4),  
('last', 39),  
('plus', 4),  
('resolve', 2),  
('brotherly', 1),  
('basis', 3),  
('qatar', 3),  
('continues', 2),  
('path', 2),  
('destabilizing', 1),  
('saudi', 2),  
('arabia', 2),  
('region.', 1),  
('surgically', 1),  
('remove', 2),  
('regime.', 1),  
('@abby\_higgins5:', 1),  
('government/pharmacies', 1),  
('wanna', 6),  
('money...', 1),  
('idc', 2),  
('https://t.co/ivompefmo0', 1),  
('@ac\_breezy', 1),  
('viruses,', 1),  
('spam/trolls', 1),  
('exact', 1),  
('thing...', 1),

```

('yet', 18),
('enough', 16),
('realize', 8),
('earned', 1),
('millennium', 1),
('prize', 5),
('couple', 2),
('hundred', 4),
('times', 5),
('think', 61),
('https://t.co/29f4udvenv', 1),
('@acosta', 4),
('real.', 3),
('left', 56),
('vitriol', 2),
('hatred', 5),
('wrong.', 7),
('pity', 2),
('miserable', 2),
('people.', 10),
('@adeabdwhb', 1),
('tweak', 1),
('question', 7),
('little', 13),
('bit', 7),
('imagine', 7),
('youre', 3),
('dying.', 1),
('rich', 3),
('man', 22),
('giving', 12),
('choices:', 1),
('huge', 5),
('amount', 5),
('money.', 8),
('basically', 4),
('cant', 7),
('god', 30),
('choose', 4),
('bcause', 1),
('choose.', 1),
('@adevotedyogi:', 1),
...]

```

```

[ ]: # Filter words for Yes and No texts
yes_words = train_rdd_for_analysis.filter(lambda x: x[0] == 'yes').map(lambda x:
    ↪ (x[1], x[2]))

```

```
no_words = train_rdd_for_analysis.filter(lambda x: x[0] == 'no').map(lambda x:
↳(x[1], x[2]))
```

```
[ ]: yes_words.collect()
```

```
Out[25]: [('ayurveda_a_panacea', 3),
('mother', 1),
('healing"', 1),
('@gurmeetramrahim', 2),
('tells', 2),
('ayurveda', 4),
('first', 5),
('scientific', 3),
('medical', 11),
('branch', 2),
('originated', 2),
('india.', 2),
('yoga', 2),
('practices,', 2),
('becomes', 2),
('boon', 2),
('cure', 184),
('cancer,', 33),
('kidney', 7),
('disorders,', 2),
('cardiac', 2),
('arrest,', 2),
('etc.', 4),
('https://t.co/7jdqksviqf', 1),
('"mother"', 1),
('healing'.', 1),
('https://t.co/vkj1uq8sod', 1),
('#teamhalon', 1),
('supports', 1),
('national', 1),
('#randomactsoflightday', 1),
('stands', 1),
('@llsusa', 1),
('continued', 1),
('research', 7),
('#cancer.', 6),
('learn', 3),
('https://t.co/q2lt7y56km', 1),
('$agen', 4),
('ceo', 2),
('@garo_armen', 1),
('@agenus_bio', 1),
('#wearwhite', 2),
```

('@cancerresearch', 2),  
 ('show', 2),  
 ('commitment', 1),  
 ('fight', 4),  
 ('#cancer', 6),  
 ('#immunotherapy', 3),  
 ('#oncology', 1),  
 ('#curecancer', 1),  
 ('https://t.co/f5vigul6ky', 1),  
 ('proud', 1),  
 ('work', 4),  
 ('pioneer', 1),  
 ('#immunotherapy,', 1),  
 ('dr.', 6),  
 ('@garo\_armen,', 1),  
 ('&', 28),  
 ('stand', 1),  
 ('courageous', 1),  
 ('patients', 4),  
 ('physicians', 1),  
 ('#inpursuitof', 3),  
 ('cancer.', 25),  
 ('https://t.co/ve97oxqakw', 1),  
 ('innovation', 4),  
 ('engine', 1),  
 ('produced', 1),  
 ('inds', 2),  
 ('18mos;', 1),  
 ('planned', 2),  
 ('2018', 2),  
 ('1h2019.', 1),  
 ('molecules', 1),  
 ('discovery', 2),  
 ('ind', 1),  
 ('yrs.', 2),  
 ('plus,', 1),  
 ('>100', 1),  
 ('pts', 8),  
 ('treated', 2),  
 ('agen1884', 1),  
 ('&/or', 1),  
 ('agen2034', 1),  
 ('#asco2018.', 1),  
 ('speed', 3),  
 ('action!', 1),  
 ('https://t.co/vrwjomjmvo', 1),  
 ('driven', 2),  
 ('powered', 2),



('right', 8),  
('combination', 2),  
('science', 6),  
('people-see', 2),  
('pursuing', 2),  
('passion', 2),  
('advancing', 2),  
('toward', 2),  
('goal', 2),  
('story:', 2),  
('https://t.co/apxu2dplkm', 2),  
('@big3biobos', 2),  
('&gt;', 1),  
('kind', 4),  
('advice', 2),  
('isn't', 5),  
('helpful.', 1),  
('i'm', 4),  
('sure', 2),  
('it's', 13),  
('said', 5),  
('wanting', 1),  
('helpful,', 1),  
('sucks.', 1),  
('good', 10),  
('deeds', 1),  
('faith', 1),  
('quinoa', 1),  
('hope', 15),  
('many', 18),  
('impacted', 1),  
('it.', 11),  
('worse,', 1),  
('none', 1),  
('around', 3),  
('surgery', 1),  
('treatment', 9),  
('days.', 2),  
('\*this', 1),  
('chimes\*', 1),  
('positivity', 2),  
('can't', 3),  
('neither', 2),  
('fighting', 2),  
('talk.', 1),  
('"battle"', 1),  
('narrative', 1),  
('created', 3),

('people', 20),  
('facing', 1),  
('disease', 4),  
('denied', 2),  
('narratives.', 1),  
('couldn't', 1),  
('agree', 4),  
('more\*', 1),  
('#lifewithcancer', 1),  
('https://t.co/akj8zdd3jz', 1),  
('@2mattandgaby2', 1),  
('no.', 2),  
('laetrile', 3),  
('cancer', 122),  
('anything', 4),  
('else.', 1),  
('studied,', 1),  
('work.', 2),  
('please', 3),  
('stop', 10),  
('spreading', 2),  
('dangerous', 3),  
('bullshit.', 1),  
('https://t.co/7z0jpa3osp', 1),  
('@420stonercorner', 2),  
('.@420stonercorner', 2),  
('fascinating', 2),  
('subject', 2),  
('matter', 3),  
('thought', 3),  
('pot', 3),  
('heads', 2),  
('such,', 2),  
('rampant', 2),  
('family', 5),  
('cbd', 5),  
('maybe', 4),  
('use', 18),  
('thc', 3),  
('instead', 3),  
('morphine!', 2),  
('#cannabis', 2),  
('#weeds', 2),  
('#pmqs', 1),  
('medically', 1),  
('unqualified', 1),  
('opinion?', 1),  
('#theresamay', 1),

('@weedseeds\_uk', 1),  
('@ukcannabisclub', 1),  
('@ukcannabisclubs', 1),  
('@younglabouruk', 1),  
('#bbcqt', 1),  
('#bbcaq', 1),  
('assholes', 1),  
('criminalizing', 1),  
('@weedfeed', 1),  
('@norml', 2),  
('#nyc', 1),  
('#nypd', 1),  
('@ukleap', 1),  
('@uklabour', 1),  
('@weednucleus', 1),  
('cannabis', 15),  
('communion', 1),  
('brings', 2),  
('great', 3),  
('peace', 1),  
('forever!', 1),  
('https://t.co/f583xnt3jy', 1),  
('@aacrfoundation:', 1),  
('leading', 3),  
('discoveries,', 1),  
('targeting', 3),  
('cures', 7),  
('saving', 2),  
('lives.', 1),  
('follow', 1),  
('see', 7),  
('helping', 1),  
('@adevotedyogi:', 1),  
('cockroach', 1),  
('sushi?', 1),  
('inside', 1),  
('farming', 2),  
('revolution', 1),  
('compost', 1),  
('waste', 1),  
('shake', 1),  
('menus', 1),  
('https://t.co/twbh...', 1),  
('@adustingofkief', 1),  
('@kkaiserrr', 1),  
('@c\_cannabiscafe', 1),  
('trying', 2),  
('convince', 1),

('power', 5),  
('10yrs', 1),  
('ppl', 5),  
('seem', 2),  
('care', 2),  
('recent', 1),  
('exposure', 1),  
('open', 3),  
('lot', 2),  
('ppls', 1),  
('eyes', 1),  
('going', 12),  
('on.', 1),  
('endocannabinoid', 1),  
('deficiency', 1),  
('https://t.co/r2dzgjavu', 1),  
('@aeidinger', 1),  
('@potus', 1),  
('@bencarsonizer', 1),  
('@npr', 1),  
('@wamu885', 1),  
('@kojoshow', 1),  
('@i\_luv\_thinkers', 1),  
('held', 1),  
('patent', 1),  
('diseases', 9),  
('since', 4),  
('1930's..', 1),  
('patent,', 1),  
('legalize', 2),  
('it?', 1),  
('also', 10),  
('renew', 1),  
('every', 4),  
('15-30', 1),  
('years', 17),  
('so.', 2),  
('@asw2belle', 1),  
('@bornfreefdn', 1),  
('tragically,', 1),  
('sadly,', 1),  
('mommy', 1),  
('baby', 2),  
('slaughtered', 1),  
('flickin', 1),  
('killers', 1),  
('remote', 1),  
('part', 3),

('shamwari', 2),  
 ('reserve.', 1),  
 ('happened', 1),  
 ('6/7', 1),  
 ('time', 9),  
 ('king', 1),  
 ('gone', 1),  
 ('enjoying', 1),  
 ('himself.', 1),  
 ('disgusting!', 1),  
 ('rhino', 7),  
 ('horns', 4),  
 ('grow', 4),  
 ('penises', 2),  
 ('@bparent64', 1),  
 ('@mountsinainyc', 1),  
 ('foolish.', 1),  
 ('holistic', 1),  
 ('approach', 1),  
 ('works', 3),  
 ('best', 6),  
 ('including', 5),  
 ('nutrition,', 2),  
 ('medication,', 1),  
 ('spiritual,', 1),  
 ('alternative', 1),  
 ('approaches.', 1),  
 ('courses', 1),  
 ('chemo', 6),  
 ('particular', 1),  
 ('cancers.', 1),  
 ('btw,', 1),  
 ('case', 4),  
 ('smart', 1),  
 ('enough', 4),  
 ('realize', 2),  
 ('it,', 2),  
 ('life', 5),  
 ('results', 1),  
 ('100%', 8),  
 ('death.', 2),  
 ('@campaignkate', 1),  
 ('told', 4),  
 ('"negativity"', 1),  
 ('towards', 2),  
 ('"positivity"', 1),  
 ('hasten', 1),  
 ('demise.', 1),

('someone', 5),  
 ('undergoing', 1),  
 ('treatment.they', 1),  
 ('statement', 3),  
 ('saying', 2),  
 ('"positive"', 1),  
 ('help', 17),  
 ('charming', 1),  
 ('person.', 1),  
 ('@cannaworldsite:', 1),  
 ('documentary', 2),  
 ('amelia', 2),  
 ('powers', 2),  
 ('extraordinary', 2),  
 ('experience', 2),  
 ('using', 8),  
 ('brain', 11),  
 ('tumor', 2),  
 ('@carlyebaines', 1),  
 ('fungus', 1),  
 ('scientists', 6),  
 ('mining', 1),  
 ('dna-sequence', 1),  
 ('data', 4),  
 ('mushrooms', 1),  
 ('mold', 1),  
 ('find', 6),  
 ('new', 5),  
 ('drugs.', 1),  
 ('https://t.co/93mj3riqly', 1),  
 ('https://t.co/o78q0kgzoa', 1),  
 ('@chaeyoungpulp:', 1),  
 ('pepto', 1),  
 ('bismol:', 1),  
 ('https://t.co/pngwyh689w', 1),  
 ('@clarkemicah', 2),  
 ('peter', 1),  
 ('favour', 2),  
 ('medicinal', 4),  
 ('oil', 13),  
 ('epilepsy', 1),  
 ('etc?', 1),  
 ('evidence', 6),  
 ('seems', 1),  
 ('overwhelming', 1),  
 ('change', 2),  
 ('people's', 2),  
 ('standard', 1),

('better.', 1),  
 ('medicines', 1),  
 ('rigorously', 1),  
 ('tested', 2),  
 ('efficacy', 1),  
 ('shown', 1),  
 ('free', 5),  
 ('side-effects.', 1),  
 ('otherwise', 1),  
 ('not.', 1),  
 ('https://t.co/dcurelgpsr', 1),  
 ('@devarshi\_21', 2),  
 ('@sharma\_jyo', 2),  
 ('@mohitku89856580', 2),  
 ('@swetasamadhiya', 2),  
 ('@india00541401', 2),  
 ('@indianismist', 2),  
 ('@raghu500000', 2),  
 ('@krishna333110', 2),  
 ('@sp\_umang', 2),  
 ('@me\_giezelle', 2),  
 ('@surewrap', 2),  
 ('@jyotsnavarma9', 2),  
 ('@chakdeindia11', 2),  
 ('@ragiing\_bull', 2),  
 ('@rakesha12572734', 2),  
 ('@vera\_anz', 2),  
 ('@purushueme', 2),  
 ('@ind\_rajat92', 2),  
 ('@sureshraoayala1', 2),  
 ('@premlataasopa', 2),  
 ('@rajbir2728', 2),  
 ('@pranjal\_dureja', 2),  
 ('@abdulbarri4444', 2),  
 ('@gouranga1964', 2),  
 ('@sgauni', 2),  
 ('@ar\_jag84', 2),  
 ('@vel\_jaya', 2),  
 ('@berretblack', 2),  
 ('@chittisarma18', 2),  
 ('@pksrivastava6', 2),  
 ('@borana\_borana', 2),  
 ('@wetwokrishna', 2),  
 ('@debdukkolkata', 2),  
 ('@sbhats', 2),  
 ('@muralydoctrack1', 2),  
 ('@subratsmsm', 2),  
 ('@pkban', 2),

('@modiji\_helpus', 2),  
 ('@bujji40397018', 2),  
 ('@shakun1931', 2),  
 ('@sharmaanks08', 2),  
 ('@raibahadur13', 2),  
 ('@arrorasanjeev', 2),  
 ('@\_ankahi', 2),  
 ('@025\_batman', 2),  
 ('@dayonesh', 2),  
 ('@dpka102', 1),  
 ('@vinish\_ind', 1),  
 ('hyderabad-based', 1),  
 ('solve', 1),  
 ('method', 1),  
 ('neem', 6),  
 ('city', 2),  
 ('decoded', 2),  
 ('phytochemical', 2),  
 ('obtained', 2),  
 ('leaves', 4),  
 ('flowers', 2),  
 ('tree,', 1),  
 ('nimbolide', 1),  
 ('https://t.co/nhyoias9ql', 1),  
 ('@drpoojaa', 1),  
 ('kinda', 1),  
 ('feel', 2),  
 ('sad', 4),  
 ('prefix', 1),  
 ('name', 1),  
 ('thinks', 2),  
 ('called', 4),  
 ('conventional', 2),  
 ('treatments', 7),  
 ('finding', 2),  
 ('today.', 1),  
 ('one', 11),  
 ('suggested', 1),  
 ('cow', 11),  
 ('urine,', 2),  
 ('yes,', 1),  
 ('curcumin.', 1),  
 ('@edvinsoriano', 1),  
 ('opposed', 1),  
 ('curing', 2),  
 ('w/marijuana', 1),  
 ('literally', 2),  
 ('everyone', 4),



('think', 8),  
('there's', 6),  
('conspiracy', 2),  
('hide', 1),  
('labs', 1),  
('working', 1),  
('asses', 1),  
('off.', 1),  
('say', 4),  
('fully', 1),  
('support', 1),  
('marijuana', 6),  
('opioids', 2),  
('treat', 3),  
('symptoms', 2),  
('@flimsin:', 1),  
('talk', 3),  
('@thepooluk', 1),  
('#thepool', 1),  
('https://t.co/tqchjq6wi7', 1),  
('@gardeniagal4', 1),  
('@jkellyca', 1),  
('@sarah\_sxlp', 1),  
('@kenjaques', 1),  
('@doritmi', 1),  
('@mmelgar09', 1),  
('@mcfunny', 1),  
('@veritasever', 1),  
('@dansmonkeyshack', 1),  
('@fiski70', 1),  
('@maunablissed', 1),  
('@babynname\_maven', 1),  
('@whats\_the\_harm', 1),  
('@thereal\_truther', 1),  
('@toppediatrician', 1),  
('@rugbyrhino16', 1),  
('@emmagpaley', 1),  
('@rightwingblack2', 1),  
('@just4thecause', 1),  
('@geoffschuler', 1),  
('@proantivaxxer', 1),  
('@chrisvcsefalvay', 1),  
('@somyrerrh', 1),  
('@pharmanemesis', 1),  
('@joshual\_5', 1),  
('@ghoppe', 1),  
('@andre\_\_levy', 1),  
('@vaccineuk', 1),

```

('@janem1276', 1),
('@liamkav', 1),
('@lynnleemavakay', 1),
('@itsmepanda1', 1),
('@quietlynowshhh', 1),
('@kitemanargues', 1),
('@playerking95', 1),
('@lifebiomedguru', 1),
('@skepticalraptor', 1),
('@melissa232220', 1),
('@pileofgoop', 1),
('@werewecrooked', 1),
('@kidsdoc1962', 1),
('@and_kell', 1),
('@viva__lala', 1),
('@milanovnina', 1),
('@steffieschiltz', 1),
('@bluelionblog', 1),
('@dkegel', 1),
('@kevinault', 1),
('@joshmazer2018', 1),
('@badzoot7', 1),
('kids', 3),
('already', 4),
('food', 5),
('make', 8),
('worse.', 2),
('kid', 2),
('getting', 5),
('chemo,', 1),
('keep', 3),
('cake', 1),
('ice', 1),
('cream', 1),
('down,', 1),
('get', 12),
('diagnosis', 1),
('terminal,', 1),
('even', 11),
('high', 4),
('horse.', 1),
('@gatesfoundation', 1),
('challenge', 1),
('gates', 1),
('foundation', 1),
('meet', 3),
('inuo.', 1),
('cured', 6),

```

('rate', 2),  
 ('chemotherapy,', 2),  
 ('radiation.', 1),  
 ('sir', 1),  
 ('surgery.', 1),  
 ('honolulu', 1),  
 ('hawaii.', 1),  
 ('need', 9),  
 ('japanese', 2),  
 ('interpreter.', 1),  
 ('@grpetersen1', 1),  
 ('@vincentrk', 1),  
 ('@myeloma\_doc', 1),  
 ('@pkelpe', 1),  
 ('@tuftsb', 1),  
 ('@myelomateacher', 1),  
 ('@rfonsil', 1),  
 ('@northtxmsg', 1),  
 ('@fransaraco', 1),  
 ('@puentelori', 1),  
 ('@jackmaiello', 1),  
 ('@cure\_talk', 1),  
 ('@mistyacallahan', 1),  
 ('@mtmdphd', 1),  
 ('@imfsupport', 1),  
 ('@gregbrozeit', 1),  
 ('@cgiampalmi', 1),  
 ('@danaholmes911', 1),  
 ('problem', 1),  
 ('heme', 1),  
 ('cancers', 8),  
 ('truly', 3),  
 ('curable', 2),  
 ('decades', 2),  
 ('all,', 1),  
 ('aml,', 1),  
 ('dlbcl,', 1),  
 ('hd,', 1),  
 ('hcl.', 1),  
 ('inherently', 1),  
 ('curable.', 1),  
 ('rxs', 1),  
 ('increased', 1),  
 ('curability.', 1),  
 ('myeloma', 1),  
 ('different.', 1),  
 ('tough.', 1),  
 ('trying.', 1),

('https://t.co/o2rnnr1j8v', 1),  
 ('@iamsonalibendre', 5),  
 ('know', 14),  
 ('effective', 8),  
 ('india,', 6),  
 ('urine', 15),  
 ('therapy', 9),  
 ('natural', 7),  
 ('ingredients', 5),  
 ('cancer...many', 5),  
 ('last', 9),  
 ('stages', 7),  
 ('cured..', 5),  
 ('wish', 5),  
 ('recover', 5),  
 ('soon!', 5),  
 ('don't', 8),  
 ('person', 4),  
 ('joking', 1),  
 ('is,', 2),  
 ('shouldn't', 1),  
 ('mock', 1),  
 ('stricken', 1),  
 ('illness;', 1),  
 ('she's', 2),  
 ('not,', 1),  
 ('plain', 1),  
 ('dumb.', 1),  
 ('encouraging', 1),  
 ('situation', 1),  
 ('either', 1),  
 ('ways.', 1),  
 ('https://t.co/wpq9jcju2y', 1),  
 ('everytime', 1),  
 ('like', 15),  
 ('seen', 4),  
 ('idiotic', 3),  
 ('bhakt', 4),  
 ('ever', 2),  
 ('life,', 1),  
 ('come', 1),  
 ('across', 2),  
 ('super', 1),  
 ('makes', 12),  
 ('old', 4),  
 ('sensible', 1),  
 ('bhakt.', 1),  
 ('https://t.co/60cery8k8y', 1),

('modi', 2),  
 ('recommending', 1),  
 ('cow', 2),  
 ('therapy"', 2),  
 ('sonali', 2),  
 ('bendre', 2),  
 ('claiming', 3),  
 ('many', 1),  
 ('(of', 1),  
 ('cancer)', 2),  
 ('cured"', 1),  
 ('manohar', 1),  
 ('parrikar', 1),  
 ('went', 2),  
 ('usa,', 1),  
 ('there,', 1),  
 ('came', 2),  
 ('back.', 1),  
 ('https://t.co/ujscephajl', 1),  
 ('saw', 2),  
 ('woman', 8),  
 ('advising', 1),  
 ('recently', 1),  
 ('diagnosed', 2),  
 ('metastatic', 1),  
 ('undergo', 1),  
 ('therapy.', 1),  
 ('wtf', 1),  
 ('visit', 1),  
 ('profile', 1),  
 ('cover', 1),  
 ('page.', 1),  
 ('guess', 2),  
 ('earlier', 1),  
 ('urine...', 1),  
 ('seriously', 1),  
 ('https://t.co/hjjmow73it', 1),  
 ('@inzcooler', 1),  
 ('@owenjones84', 1),  
 ('buried', 2),  
 ('friend', 2),  
 ('heroin', 1),  
 ('user', 1),  
 ('friends', 2),  
 ('low', 3),  
 ('depression,', 1),  
 ('chronic', 4),  
 ('pain,', 1),

('cluster', 1),  
('headaches', 2),  
('control', 1),  
('pain', 1),  
('far', 2),  
('lower', 2),  
('side', 1),  
('effects', 4),  
('medication', 1),  
('@itsbravin', 1),  
('cannabis,', 1),  
('avacado', 2),  
('leafs,papaya', 1),  
('leafs,', 1),  
('soursop', 1),  
('leave,tumeric,', 1),  
('dandelion', 1),  
('roots', 1),  
('ginger,moriga', 1),  
('cancer.imagine', 1),  
('africa', 1),  
('infact', 1),  
('free.bravin', 1),  
('online', 2),  
('shop', 2),  
('powerder', 1),  
('papaya', 4),  
('ill', 1),  
('buy.', 1),  
('@jeenasingh:', 1),  
('scientist', 1),  
('single', 2),  
('sentence!', 1),  
('living', 3),  
('interesting', 2),  
('times', 1),  
('indeed!', 1),  
('https://t.co/fdqbbhfd...', 1),  
('@jksmith34:', 1),  
('#nanotechnology', 2),  
('https://t.co/7c98encvph', 1),  
('@luna\_dna', 1),  
('behind', 3),  
('bullshit?', 1),  
('want', 5),  
('eating', 8),  
('sugar', 1),  
('carbs.', 1),

('corporations', 1),  
 ('started', 1),  
 ('feeding', 1),  
 ('live', 2),  
 ('stock.', 1),  
 ('give', 7),  
 ('dna.', 1),  
 ('@mari3nava:', 1),  
 ('https://t.co/jeqaflxse5', 1),  
 ('#tech', 1),  
 ('#ehealth', 1),  
 ('#ai', 2),  
 ('@meghanmccain', 2),  
 ('advancements', 1),  
 ('gbm', 2),  
 ('health', 8),  
 ('happiness', 1),  
 ('year', 2),  
 ('survivor', 1),  
 ('vaccines', 1),  
 ('https://t.co/axwsgdjt6q', 1),  
 ('vitamins', 1),  
 ('relaxation.\xa0', 1),  
 ('https://t.co/bwfoy9gab2', 1),  
 ('cannabinoids', 2),  
 ('-cbd/thc', 1),  
 ('https://t.co/kxdbvs6kw9', 1),  
 ('there!', 1),  
 ('are!', 1),  
 ('plenty!', 1),  
 ('of!', 1),  
 ('natural!', 1),  
 ('cures!', 1),  
 ('for!', 1),  
 ('cancer!', 3),  
 ('"there', 1),  
 ('acceptable', 1),  
 ('radiation."', 1),  
 ('lie', 2),  
 ('damned', 1),  
 ('deathcare', 1),  
 ('industry!', 1),  
 ('non-existent', 1),  
 ('early', 2),  
 ("1900's!", 1),  
 ('bring', 1),  
 ('body', 7),  
 ('die!', 1),

```

('@mirrorweirdnews:', 1),
('swears', 1),
('drinking', 4),
('dog's", 5),
('acne', 5),
('@mussavir', 1),
('@azizakhmad', 1),
('@ali_madeeh', 1),
('trees', 1),
('useful.', 1),
('azadirachta', 1),
('indica', 1),
('(azad', 1),
('drakht', 1),
('india', 2),
('persian).', 1),
('dhrek', 1),
('bakain', 1),
('melia', 1),
('azadarach.', 1),
('useful', 1),
('element', 1),
('azadirachtin', 1),
('think)', 1),
('day', 1),
('alzheimers,', 2),
('pakinsons,', 1),
('miracle', 1),
('tree.', 1),
('@nathanafttanas:', 1),
('breaking', 1),
('news:', 1),
('chips', 1),
('salsa', 1),
('confirmed', 1),
('cancer..', 1),
('check', 3),
('source', 2),
('below ', 1),
('https://t.co/gyle2cmype', 1),
('@northkoreatimes', 1),
('#northkorea', 1),
('#dogmeatt', 1),
('serve', 1),
('"viagra"', 1),
('cruelty', 1),
('!!!!', 1),
('educate', 1),

```



('please!!!', 1),  
 ('', 1),  
 ('https://t.co/nijz7j6tw6', 1),  
 ('@ogatimi', 1),  
 ('@asemota', 1),  
 ('currently', 1),  
 ('mums', 1),  
 ('pelvic', 1),  
 ('past', 1),  
 ('two', 3),  
 ('months', 1),  
 ('what???', 1),  
 ('everything', 5),  
 ('everything...', 1),  
 ('weed', 6),  
 ('hasn't', 2),  
 ('legalized', 3),  
 ('till', 2),  
 ('@sandeepcol', 1),  
 ('@kailashkaushik8', 1),  
 ('@hmoindia', 1),  
 ('@pmoindia', 1),  
 ('behold', 1),  
 ('"american', 1),  
 ('claims', 4),  
 ('helped', 2),  
 ('goes', 2),  
 ('state', 3),  
 ('dog', 1),  
 ('contains', 1),  
 ('vitamin', 6),  
 ('grams', 1),  
 ('calcium', 1),  
 ('potential', 3),  
 ('cancer."', 3),  
 ('https://t.co/nozmkz5l6h', 1),  
 ('@suespurgin', 1),  
 ('@elephas\_maximus', 1),  
 ('@tamsyncarstens', 1),  
 ('@tselsmark', 1),  
 ('@saveourrhinok9', 1),  
 ('@elerhinomarch', 1),  
 ('@sharonhoole', 1),  
 ('@louisal1000', 1),  
 ('@h0t\_p0ppy', 1),  
 ('rhinos', 1),  
 ('murdered!!', 1),  
 ('#china', 2),

('vietnam', 2),  
('@scmpnews', 1),  
('@tuoitrennews', 1),  
('#rhinohorncuresnothing', 1),  
('https://t.co/o1mstdrto5', 1),  
('horrific!', 1),  
('mom', 2),  
('baby...people', 1),  
('sick!', 1),  
('july.', 1),  
('hacked', 1),  
('axe.', 1),  
('attn', 1),  
('people,', 2),  
('yourselves...', 1),  
('penises.', 1),  
('really!!!', 1),  
('leave', 2),  
('animals', 1),  
('alone', 1),  
('https://t.co/7h28bgysz5', 1),  
('@trumpmoments411', 1),  
('@peterbakernyt', 1),  
('sweet', 1),  
('pea,', 1),  
('fail', 1),  
('understand', 8),  
('sometimes', 1),  
('take', 4),  
('dose', 1),  
('knows', 2),  
('that.', 1),  
('long', 4),  
('game.', 1),  
('gop', 1),  
('grows', 2),  
('older', 2),  
('older.', 1),  
('enjoy', 1),  
('gasp.', 1),  
('@upshelp', 1),  
('month', 1),  
('son', 2),  
('terminal', 2),  
('tissue', 1),  
('shipped', 1),  
('usc', 1),  
('try', 2),

```

('develop', 1),
('sitting', 1),
('ontario,', 1),
('facility', 1),
('delivered', 1),
('cant', 1),
('wait', 1),
('til', 1),
('monday', 1),
('ruined.', 1),
('@ups', 1),
('@wordbird67', 3),
('@takethatchem', 2),
('ignoring', 1),
('may', 7),
('differ', 1),
('stance...', 1),
('same...', 1),
('call', 4),
('blocks...', 1),
('fun', 2),
('from...', 1),
('"proved"', 1),
('wrong', 1),
('b-17', 1),
('deleted', 1),
('meme...', 1),
('choose', 1),
('troll...', 1),
('lib?', 1),
('https://t.co/ocm4qvghtl', 1),
('aren't', 1),
('discussing', 1),
('political', 3),
('issue,', 1),
('you're', 2),
('posting', 2),
('misinformation', 1),
('internet', 1),
('out.', 2),
('b17', 3),
('despite', 1),
('name.', 1),
('"cancer', 1),
('experts"', 1),
...]
```

```
[ ]: no_words.collect()
```

```
Out[26]: [('armyhivestreamingparty', 1),
('purest', 2),
('thing', 21),
('ever', 16),
('experienced.', 1),
('beyonce', 1),
('bts', 3),
('together', 15),
('cure', 1142),
('cancer', 816),
('diseases/illnesses', 1),
('bring', 26),
('world', 106),
('peace', 34),
(' ', 1),
('#cureforcancer', 1),
('100', 3),
('gave', 15),
('medical', 16),
('research', 44),
('teams', 2),
('back', 23),
('2014', 2),
('know', 46),
('hid', 1),
('sick', 8),
('needed', 5),
('lie', 5),
('anyone', 19),
('true', 9),
('#noname', 1),
('#executiveorder', 2),
('sick,', 1),
('deranged', 1),
('liberals', 23),
('happy', 18),
('president', 52),
('found', 65),
('cancer.', 137),
('morons', 2),
('say,', 8),
('"no!', 1),
('want', 74),
('people', 153),
('suffer', 5),
('die!"', 1),
```

('unhinged', 2),  
('nothing', 18),  
('please', 59),  
('them.', 12),  
('thrive', 2),  
('something', 50),  
('bitch', 12),  
('about.', 7),  
('#garethsouthgatewould', 1),  
('solve', 13),  
('hunger', 25),  
('even', 63),  
('take', 28),  
('credit', 8),  
('say', 66),  
('"the', 5),  
('lads', 1),  
('well"', 1),  
('#worldcup18', 1),  
('#england', 1),  
('#lgbtactionplan', 1),  
('important', 8),  
('step', 6),  
('become', 10),  
('equally', 1),  
('diverse', 1),  
('place', 3),  
('love.', 4),  
('gay,', 1),  
('anything', 23),  
('straight', 1),  
('curable', 3),  
('cure.', 12),  
('work', 25),  
('hiv', 14),  
('#otd', 1),  
('years', 35),  
('ago,', 4),  
('terry', 2),  
('fox', 4),  
('passed', 3),  
('away', 16),  
('new', 20),  
('westminister,', 1),  
('bc.', 1),  
('speaking', 3),  
('run,', 1),  
('said', 20),

('dreamer', 1),  
 ('saying', 17),  
 ('initiate', 1),  
 ('kind', 8),  
 ('definitive', 1),  
 ('answer', 4),  
 ('believe', 35),  
 ('miracles.', 1),  
 ('to."', 1),  
 ('watch', 9),  
 ('#heritageminute', 1),  
 ('here.', 3),  
 ('https://t.co/oqonptqcs', 1),  
 ('#predict', 2),  
 ('&', 95),  
 ('#cure', 2),  
 ('next', 13),  
 ('#cancer', 26),  
 ('https://t.co/gyztoi46yi', 1),  
 ('minutes', 3),  
 ('#cancercure', 7),  
 ('#healed', 2),  
 ('#reiki', 2),  
 ('#rtrrt', 4),  
 ('#bigpharma', 4),  
 ('#loa', 2),  
 ('#fridayreads', 2),  
 ('#nlp', 2),  
 ('#acim', 2),  
 ('https://t.co/uqx20hdk8c', 1),  
 ('https://t.co/oqh0emyqpk', 1),  
 ('home', 9),  
 ('#academiczodiac', 2),  
 ('#o5iris', 2),  
 ('predicta...', 2),  
 ('https://t.co/jvbdy7arqy', 1),  
 ('https://t.co/4r1zgo8opo', 1),  
 ('https://t.co/ojby25xlyk', 1),  
 ('#righttotry', 2),  
 ('1st', 7),  
 ('lyrics', 1),  
 ('trump', 203),  
 ('released?', 1),  
 ('give', 49),  
 ('vote', 7),  
 ('-not', 1),  
 ('hillary', 3),  
 ('dr.', 5),

('video=', 1),  
 ('blocked', 2),  
 ('bill', 6),  
 ('clinton', 2),  
 ('since', 13),  
 (''94', 1),  
 ('#drburzynskirhymeswithlewinski', 2),  
 ('#trumptrainjumponit', 1),  
 (' https://t.co/6odiawuwga', 1),  
 ('via', 11),  
 ('@youtube', 2),  
 ('https://t.co/0uta7ccgwk', 1),  
 ('poor', 7),  
 ('clinton!', 1),  
 ('decades', 2),  
 ('suppressing', 1),  
 ('cancer-cures', 1),  
 ('wants', 7),  
 ('#righttocry', 1),  
 ('#righttodie', 1),  
 ('ask', 15),  
 ('patients', 24),  
 ('that's', 21),  
 ('right!', 5),  
 ('rhymes', 1),  
 ('mistress', 1),  
 ('-both', 1),  
 ('f\*cked', 1),  
 ('https://t.co/xpikezppcz', 1),  
 ('#shocking', 1),  
 ('#government', 1),  
 ('#petitions', 1),  
 ('#dogs', 1),  
 ('get', 85),  
 ('signatures', 1),  
 ('life', 39),  
 ('saving', 5),  
 ('funding', 5),  
 ('terminal', 5),  
 ('diseases', 16),  
 ('known', 12),  
 ('cause', 20),  
 ('kill', 10),  
 ('them! ', 1),  
 ('#ipf', 2),  
 ('#pulmonaryfibrosis', 2),  
 ('needs', 17),  
 ('like', 140),

('#cancer!', 2),  
 ('#humanity?', 1),  
 ('show', 11),  
 ('care ', 1),  
 ('https://t.co/zzcp7rnld0', 1),  
 ('#sisterscare:', 1),  
 ("cbs2's", 2),  
 ('@burrelltv', 1),  
 ('@efinchtv', 2),  
 ('@danatylertv', 2),  
 ('took', 12),  
 ('part', 10),  
 ('sisters', 2),  
 ('event', 13),  
 ('today,', 7),  
 ('addressing', 1),  
 ('disparities', 1),  
 ('diagnosis', 5),  
 ('treatment', 18),  
 ('breast', 32),  
 ('african', 2),  
 ('american', 12),  
 ('women.', 4),  
 ('https://t.co/unjsqhcycd6', 1),  
 ('#teachersleading', 1),  
 ('come', 31),  
 ('doctor', 9),  
 ('revolution', 1),  
 ('teaching', 3),  
 ('going', 27),  
 ('teachers', 2),  
 ('@nickdonohuenmef', 1),  
 ('#trump', 9),  
 ('cancer,', 150),  
 ('eradicate', 5),  
 ('poverty', 10),  
 ('still', 157),  
 ('msm', 10),  
 ('pull', 4),  
 ('put', 34),  
 ('negative', 10),  
 ('spin', 9),  
 ('it.', 54),  
 ('war', 8),  
 ('conflict', 2),  
 ('preferable', 1),  
 ('wrong', 14),  
 ('consistently', 1),



('are.', 5),  
 ('#trumpkimsummit', 15),  
 ('#maga', 14),  
 ('discover', 9),  
 ('#trumpderangementsyndrome', 5),  
 ('sufferers', 2),  
 ('screaming', 1),  
 ('injection', 1),  
 ('stinging.#liberalhypocrisy', 1),  
 ('#singaporesummit', 2),  
 ('#trumpkim', 1),  
 ('#deepstate', 1),  
 ('#thestorm', 1),  
 ('self-inflicted', 1),  
 ('metastasized', 1),  
 ('full', 15),  
 ('blown', 1),  
 ('psychological', 1),  
 ('insanity', 2),  
 ('outbreak', 1),  
 ('across', 9),  
 ('democrat', 4),  
 ('party,', 1),  
 ('liberal', 12),  
 ('media', 28),  
 ('nevertrumpers', 1),  
 ('lead', 10),  
 ('ultimate', 2),  
 ('demise.', 1),  
 ('time', 41),  
 ('let', 22),  
 ('go.', 2),  
 ('#walkaway', 2),  
 ('\$celz', 1),  
 ('imho,', 1),  
 ('beaten', 3),  
 ('many', 49),  
 ('countries.;', 1),  
 ('rate', 18),  
 ('95%+', 1),  
 ('#fda', 1),  
 ('big', 17),  
 ('#pharma', 1),  
 ('keep', 27),  
 ('treatments', 10),  
 ('illegal', 6),  
 ('u.s', 2),  
 ('capitalize', 1),

('\$150,000-\$500,000', 1),  
 ('unneeded', 1),  
 ('surgeries', 5),  
 ('chemo', 11),  
 ('celz's', 1),  
 ('#amniostem', 1),  
 ('help', 106),  
 ('#cachexia', 1),  
 ('huge!', 1),  
 ('&gt;"the', 1),  
 ('idea', 6),  
 ('tell', 17),  
 ('story', 13),  
 ('lesnar', 1),  
 ('looking', 10),  
 ('leaving', 4),  
 ('belt', 1),  
 ('ufc"', 1),  
 ('garner', 1),  
 ('sympathy', 2),  
 ('reigns.', 1),  
 ('roman', 1),  
 ('reigns', 1),  
 ('end', 54),  
 ('hunger,', 12),  
 ('find', 204),  
 ('booed', 1),  
 ('pushed', 1),  
 ('hard.', 1),  
 ('https://t.co/uxful5sx6q', 1),  
 ('\*breaking', 1),  
 ('news\*', 1),  
 ('(hypothetical', 1),  
 ('situation)', 1),  
 ('finds', 7),  
 ('cancer!', 28),  
 ('everyone', 26),  
 ('mother:', 1),  
 ('"what', 2),  
 ('piece', 5),  
 ('shit!', 1),  
 ('probably', 15),  
 ('knew', 7),  
 ('whole', 12),  
 ('time!'", 1),  
 ('dont', 7),  
 ('agree', 2),  
 ('shit', 20),

('guy', 13),  
('does/says,', 1),  
('ya'll', 1),  
('never', 57),  
('satisfied', 2),  
('\*opens', 1),  
('cab-booking', 1),  
('app\*', 1),  
('cab', 2),  
('\*books', 1),  
('cab\*', 1),  
('years,', 5),  
('months,', 1),  
('weeks,', 1),  
('days', 8),  
('hours.', 1),  
('kids', 22),  
('born', 6),  
('grow', 12),  
('knowing', 4),  
('mother,', 1),  
('scientists', 18),  
('volcano', 1),  
('erupt', 1),  
('wipe', 4),  
('hawaii', 1),  
('away.', 1),  
('.@abbvie', 2),  
('alphabet-backed', 2),  
('calico', 2),  
('spend', 8),  
('another', 12),  
('\$1b', 2),  
('https://t.co/a7tqwrwtrv', 2),  
('.@erictrump:', 1),  
('"we're', 1),  
('raising', 14),  
('money', 76),  
('we're', 11),  
('win', 16),  
('fight.', 2),  
('fight', 35),  
('pediatric', 12),  
('cancer."', 2),  
('https://t.co/7pyfcd7xzy', 1),  
('thank', 39),  
('@erictrump', 2),  
('son', 5),

('alive', 2),  
('due', 3),  
('started', 4),  
('jude.', 1),  
('year', 29),  
('survivor', 11),  
('week.', 4),  
('loose', 2),  
('least', 6),  
('every', 50),  
('month', 5),  
('local', 3),  
('children's', 2),  
('hospital', 8),  
('childhood', 24),  
('https://t.co/gwqtxflqau', 1),  
('11:11', 4),  
('excellent', 4),  
('grades!!!', 2),  
('job', 14),  
('best', 17),  
('friend', 10),  
('happiness', 6),  
('friends', 13),  
('gun', 7),  
('control', 11),  
('@arianagrande', 4),  
('notice', 5),  
('@shawnmendes', 4),  
('follow', 9),  
('bye', 5),  
('anxiety', 6),  
('fun', 7),  
('summer', 4),  
('lion', 9),  
('rid', 9),  
('depression', 16),  
('hug', 6),  
('love', 26),  
('good', 44),  
('grades', 1),  
('summer!', 1),  
('harry', 3),  
('potter', 1),  
('reunion', 3),  
('weight!', 1),  
('16-year-old', 4),  
('jeremiah', 4),

('thomas,', 3),  
 ('battling', 5),  
 ('used', 11),  
 ('make', 84),  
 ('wish', 12),  
 ('abortion.', 5),  
 ('you,', 9),  
 ('jeremiah,', 3),  
 ('beautiful', 8),  
 ('life,', 13),  
 ('inspiring', 7),  
 ('many.', 4),  
 ('https://t.co/bumzmggyay', 3),  
 ('sad', 13),  
 ('young,', 1),  
 ('brainwashed', 2),  
 ('maybe', 21),  
 ('health', 23),  
 ('care', 19),  
 ('all,', 3),  
 ('imaginary', 1),  
 ('exist', 4),  
 ('https://t.co/27twcb3oek', 1),  
 ('"selfless', 1),  
 ('way"', 1),  
 ('idk', 1),  
 ('mann', 1),  
 ('it's', 59),  
 ('pretty', 5),  
 ('selfish', 1),  
 ('try', 18),  
 ('use', 10),  
 ('religion', 11),  
 ('controls', 1),  
 ('someone', 29),  
 ('else's', 1),  
 ('body...', 1),  
 ('coulda', 1),  
 ('wished', 4),  
 ('https://t.co/tyx3hynfa8', 1),  
 ('17yr', 3),  
 ('stage', 18),  
 ('carcinoid', 12),  
 ('survivor.', 4),  
 ('disease', 48),  
 ('longest', 2),  
 ('surviving', 2),  
 ('patient', 7),

('world.', 15),  
('survival', 11),  
('5yrs.', 6),  
('plz', 14),  
('read', 24),  
('remarkable', 10),  
('miracle', 12),  
('book', 13),  
('living', 14),  
('silent', 11),  
('enemy', 11),  
('coming', 18),  
('soon.', 11),  
('support', 31),  
('mdacc.', 1),  
('inoperable', 5),  
('neuroendocrine', 7),  
('misdiagnosed', 7),  
('lets', 3),  
('cancer#', 1),  
('17yr.', 1),  
('endcancer#', 1),  
('1965:', 2),  
('bet', 10),  
("there'll", 4),  
('flying', 8),  
('cars', 7),  
('future"', 1),  
('1987:', 1),  
('future', 12),  
('2018:', 2),  
('prototype', 1),  
('potentially', 5),  
('viable', 3),  
('cures', 12),  
('cancers', 9),  
('given', 5),  
('1998:', 2),  
('"in', 1),  
('twenty', 2),  
('cancer!'", 1),  
('"dennis', 1),  
('rodman', 1),  
('saved', 5),  
('nuclear', 4),  
('celebrity', 2),  
('apprentice,', 1),  
('hosted', 2),

('united', 3),  
 ('states,', 1),  
 ('dictator', 3),  
 ('north', 8),  
 ('korea!'", 1),  
 ('@panmass', 1),  
 ('rider', 1),  
 ('@danafarber.', 1),  
 ('fund', 14),  
 ('retweet!', 1),  
 ('https://t.co/qdwqn3zzre', 1),  
 ('https://t.co/lafbipfydc', 1),  
 ('2018', 9),  
 ('3mt', 1),  
 ('semi-finalist', 1),  
 ('ubc', 1),  
 ('science', 12),  
 ('master', 1),  
 ('student', 2),  
 ('evan', 1),  
 ('warner', 1),  
 ('presented', 1),  
 ("shakespeare's", 1),  
 ('much', 48),  
 ('thing?', 2),  
 ('asked.', 1),  
 ('#tbt', 1),  
 ('talk:', 1),  
 ('https://t.co/vjazjwo5em', 1),  
 ('22,', 1),  
 ('ethan', 1),  
 ('hawes', 1),  
 ('diagnosed', 8),  
 ('multiple', 4),  
 ('myeloma,', 1),  
 ('typically', 1),  
 ('afflicting', 2),  
 ('age', 4),  
 ('65.', 1),  
 ('hospitalizations,', 1),  
 ('chemotherapy,', 1),  
 ('stem', 3),  
 ('cell', 5),  
 ('replacement', 1),  
 ('followed.', 1),  
 ('https://t.co/f0lztvodcb', 1),  
 ('5,000', 2),  
 ('young', 12),

('women', 16),  
('country', 15),  
('participated', 2),  
('year's', 4),  
('lax', 1),  
('raise', 24),  
('funds', 12),  
('awareness', 13),  
('tournament', 2),  
('raised', 16),  
('\$235,000', 1),  
('komen', 2),  
('csnj', 1),  
('@komennja.', 1),  
('more:', 2),  
('https://t.co/4qlw3bei7k', 1),  
('https://t.co/uwdnketoie', 1),  
('6th', 4),  
('40tude', 3),  
('thames', 2),  
('sup', 1),  
('marathon', 3),  
('saturday', 2),  
('already', 16),  
('£63k', 1),  
('marks', 1),  
('colon', 4),  
('add', 7),  
('£180k', 1),  
('past', 4),  
('years.', 8),  
('miles', 2),  
('sups', 1),  
('finishing', 1),  
('putney.', 1),  
('@curebowelcancer', 1),  
('@active360', 1),  
('https://t.co/dn5cjvm9gz', 1),  
('87%', 2),  
('f-ing', 2),  
('republicans', 6),  
('trump.', 9),  
('omg.', 1),  
('how?', 2),  
('children', 32),  
('ripped', 1),  
('mother's', 2),  
('arms.', 1),



('truly', 8),  
('80%', 2),  
('news', 9),  
('20%', 3),  
('don't', 50),  
('on.', 4),  
('#disgusted', 1),  
('must', 20),  
('accept', 4),  
('sinister', 1),  
('brainwashing.', 1),  
('can't', 29),  
('comprehend', 2),  
('maga', 3),  
('grasp', 1),  
('shooting', 2),  
('afternoon', 1),  
('nature.', 2),  
('one', 89),  
('voting', 2),  
('overwhelming', 2),  
('numbers.', 1),  
('https://t.co/s7ctd0lyc6', 1),  
('@1405scituateave', 1),  
('@realjameswoods', 4),  
('funny', 7),  
('is,', 5),  
('discovered', 17),  
('tomorrow,', 3),  
('donald', 12),  
('telling', 12),  
('somehow', 4),  
('instrumental', 1),  
('it,', 19),  
('jackass', 1),  
('trumpuppets', 1),  
('word', 4),  
('it!!!!', 1),  
('...and', 1),  
('you're', 10),  
('saying;', 1),  
('"maybe', 1),  
('be"', 1),  
('right', 18),  
('now!!', 1),  
('@1776stonewall', 2),  
('pathetic', 1),  
('wasn't', 1),

('enough!!!', 1),  
 ('talk', 24),  
 ('pitiful', 1),  
 ('souls', 2),  
 ('way', 43),  
 ('trump!!!!', 1),  
 ('@1999mcm11', 1),  
 ('denuclearization', 3),  
 ('achieved.', 2),  
 ('gets', 6),  
 ('taste', 1),  
 ('western', 3),  
 ('culture,', 1),  
 ('mcdonalds,', 1),  
 ('coke,', 1),  
 ('xbox,', 1),  
 ('etc...', 1),  
 ('over.', 1),  
 ('achieve', 5),  
 ('peace...', 1),  
 ('complain', 32),  
 ('him.', 31),  
 ('cuz', 4),  
 ('he's', 22),  
 ('orange', 2),  
 ('tweets', 2),  
 ('stupid', 2),  
 ('stuff.', 2),  
 ('@\_cape\_', 1),  
 ('@\_\_hunchohoudini', 1),  
 ('provide', 6),  
 ('food', 11),  
 ('shelter', 1),  
 ('homeless', 2),  
 ('can't', 2),  
 ('aids', 32),  
 ('although', 2),  
 ('it)', 3),  
 ('shot', 5),  
 ('tupac.we', 1),  
 ('investigators', 1),  
 ('support.', 4),  
 ('swear', 7),  
 ('better', 25),  
 ('sometimes', 1),  
 ('always', 15),  
 ('dumb', 2),  
 ('ppl', 6),

('fuck', 18),  
 ('@\_duckyfresh', 1),  
 ('@babygrotesque', 1),  
 ('underlying', 1),  
 ('economic', 4),  
 ('problems', 5),  
 ('contribute', 2),  
 ("gotham's", 1),  
 ('crime.', 3),  
 ('stop', 25),  
 ('beating', 4),  
 ('mentally', 2),  
 ('ill', 5),  
 ('throwing', 1),  
 ('inhumane', 1),  
 ('caricature', 1),  
 ('mental', 7),  
 ('facility.', 1),  
 ('bruce', 1),  
 ('literally', 18),  
 ('spidey', 1),  
 ('meme', 1),  
 ('tea', 2),  
 ('@\_kingofthehill1:', 1),  
 ('sexual', 4),  
 ('https://t.co/qt6n6wrqse', 1),  
 ('@abaabood', 1),  
 ('@andreas\_krieg', 1),  
 ('me,', 1),  
 ('spared', 1),  
 ('efforts', 4),  
 ('last', 30),  
 ('plus', 1),  
 ('resolve', 2),  
 ('brotherly', 1),  
 ('basis', 2),  
 ('qatar', 3),  
 ('continues', 2),  
 ('path', 2),  
 ('destabilizing', 1),  
 ('saudi', 2),  
 ('arabia', 2),  
 ('region.', 1),  
 ('surgically', 1),  
 ('remove', 2),  
 ('regime.', 1),  
 ('@abby\_higgins5:', 1),  
 ('government/pharmacies', 1),

('wanna', 6),  
('money...', 1),  
('idc', 2),  
('https://t.co/ivompefmo0', 1),  
('@ac\_breezy', 1),  
('viruses,', 1),  
('spam/trolls', 1),  
('exact', 1),  
('thing...', 1),  
('yet', 16),  
('enough', 12),  
('realize', 6),  
('earned', 1),  
('millennium', 1),  
('prize', 4),  
('couple', 1),  
('hundred', 2),  
('times', 4),  
('think', 53),  
('https://t.co/29f4udvenv', 1),  
('@acosta', 4),  
('real.', 3),  
('left', 54),  
('vitriol', 2),  
('hatred', 5),  
('wrong.', 6),  
('pity', 2),  
('miserable', 2),  
('people.', 9),  
('@adeabdwbb', 1),  
('tweak', 1),  
('question', 6),  
('little', 12),  
('bit', 7),  
('imagine', 6),  
('youre', 3),  
('dying.', 1),  
('rich', 3),  
('man', 20),  
('giving', 11),  
('choices:', 1),  
('huge', 5),  
('amount', 4),  
('money.', 7),  
('basically', 3),  
('cant', 6),  
('god', 28),  
('choose', 3),

('bcause', 1),  
('choose.', 1),  
('@adryzzle', 1),  
('repeal', 2),  
('thought', 13),  
('obama', 19),  
('made', 21),  
('happen,', 1),  
('cult', 2),  
('followers', 4),  
('cheer', 2),  
('tumors', 2),  
('killed', 9),  
('@adustingofkief', 1),  
('@kkaiserrr', 1),  
('@c\_cannabiscafe', 1),  
('may', 29),  
('noticed', 3),  
('adverts', 2),  
('pre', 3),  
('2015', 1),  
('"help', 1),  
('sooner"', 1),  
('"making', 1),  
('chemotherapy', 6),  
('safer', 1),  
('making', 21),  
('radiotherapy', 1),  
('safer"', 1),  
('finding', 40),  
('obvious', 3),  
('@cr\_uk', 2),  
('frauds', 1),  
('protecting', 1),  
('pharmaceutical', 4),  
('industry.', 3),  
('https://t.co/mb6kdnslja', 1),  
('@ajpapa\_', 1),  
('@roxyreynolds', 1),  
('pinky', 1),  
('stepped', 1),  
('scene', 1),  
('problem', 6),  
('suck', 1),  
('dick', 1),  
('nut', 2),  
('special', 8),  
('shout', 2),

('outs', 1),  
('venessa', 1),  
('blue', 5),  
('kapri', 1),  
('styles', 1),  
('skin', 7),  
('diamond', 2),  
('lacey', 1),  
('duvalle', 1),  
('julia', 1),  
('chanel', 1),  
('@ajplus', 1),  
('intentions', 1),  
('good,', 1),  
('reduce', 4),  
('heel', 2),  
('men', 5),  
('praised', 2),  
('indirectly', 1),  
('trying', 13),  
('instead', 13),  
('directly', 3),  
('helping', 13),  
('themselves.', 1),  
('male', 2),  
('equivalent', 1),  
('white', 5),  
('girl', 2),  
('feminism', 1),  
('impressive.', 1),  
('guys', 10),  
('stuff', 4),  
('wearing', 5),  
('pink', 8),  
('awareness.', 1),  
('https://t.co/kzcbgzcdsp', 1),  
('@akinwole\_', 1),  
('@damben1', 1),  
('similar', 3),  
('saying:', 1),  
('asked', 3),  
('components?', 1),  
('response', 4),  
('"go', 3),  
('lab', 5),  
('yours".', 1),  
('nutshell,', 1),  
('dat', 1),

('means', 9),  
('true,', 2),  
('works,', 1),  
('real,', 1),  
('proud', 15),  
('release', 8),  
('simple.', 2),  
('@alyssa\_milano', 2),  
('seriously', 5),  
('right.', 2),  
('jobs', 11),  
('lost', 33),  
('healthcare', 11),  
('and/or', 1),  
('pharmaceuticals.', 1),  
('@amaroqwolf', 1),  
('ha.', 1),  
('referencing', 1),  
('line', 2),  
('elliott', 1),  
('benefits', 1),  
('time-travel.', 1),  
('told', 7),  
('katerina', 1),  
('wanted', 8),  
('"give', 1),  
('tomorrow."', 1),  
('@americafirst181', 1),  
('@potus', 25),  
('america', 14),  
('million', 11),  
('dollars', 11),  
('sure', 15),  
('house', 3),  
('live', 22),  
('in.', 2),  
('private', 1),  
('plane', 2),  
('rino's', 1),  
('throw', 2),  
('fit', 3),  
('fuel', 2),  
('plane.', 1),  
('https://t.co/pmkmejsj9o', 1),  
('@amike4761', 2),  
('applying', 3),  
('east', 2),  
('german', 3),

```
( 'communist', 1),
( 'autocracy', 1),
( 'policies,', 2),
( 'irretrievably', 1),
( 'spoiling', 1),
( 'both.', 1),
( 'undermined', 1),
( 'culture', 2),
( 'social', 6),
( 'chemo.', 2),
( '@amy_siskind', 1),
( 'create', 16),
( 'empty', 1),
( 'trojan', 1),
( 'horse', 2),
( 'agreement', 1),
( 'russia', 2),
( 'korea', 5),
( 'helps', 5),
( 'knuckleheaded', 1),
( 'base', 9),
( 'everything', 15),
( 'says...', 1),
( 'negotiated', 2),
( 'safe"', 1),
( '"all', 1),
( 'hail', 1),
( 'zod"', 1),
( 'https://t.co/ntarclkepa', 1),
( '@ancientaliens', 1),
( 'real', 18),
( 'illuminati,', 1),
( 'alien', 3),
( 'tech,', 1),
( 'origins,', 1),
( 'identity', 2),
( 'species', 1),
( 'creator,', 1),
( 'amongst', 1),
( 'things.', 1),
( '@andersonala45', 1),
( 'tries', 3),
...]
```

```
[ ]: # Join word counts for yes and no
joined_data = word_counts.leftOuterJoin(yes_words)
word_table = joined_data.leftOuterJoin(no_words)
```



```
[ ]: joined_data.collect()
```

```
Out[28]: [('purest', (2, None)),
('ever', (18, 2)),
('beyonce', (1, None)),
('bts', (3, None)),
('cure', (1327, 184)),
('diseases/illnesses', (1, None)),
('bring', (27, 1)),
('world', (108, 2)),
('peace', (35, 1)),
(' ', (1, None)),
('"mother', (1, 1)),
('healing"', (1, 1)),
('tells', (3, 2)),
('scientific', (10, 3)),
('medical', (27, 11)),
('yoga', (3, 2)),
('boon', (2, 2)),
('cancer,', (183, 33)),
('disorders,', (3, 2)),
('cardiac', (2, 2)),
('etc.', (10, 4)),
('#cureforcancer', (1, None)),
('100', (3, None)),
('gave', (16, 1)),
('research', (51, 7)),
('2014', (2, None)),
('know', (60, 14)),
('needed', (7, 2)),
('lie', (7, 2)),
('anyone', (20, 1)),
('true', (9, None)),
('#noname', (1, None)),
('cancer.', (162, 25)),
('morons', (2, None)),
('"no!', (1, None)),
('suffer', (5, None)),
('die!"', (1, None)),
('them.', (13, 1)),
('something', (50, None)),
('bitch', (12, None)),
('#garethsouthgatewould', (1, None)),
('solve', (14, 1)),
('even', (74, 11)),
('take', (32, 4)),
('say', (70, 4)),
('"the', (5, None)),
```

```

('well"', (1, None)),
('#worldcup18', (1, None)),
('step', (6, None)),
('become', (11, 1)),
('equally', (1, None)),
('diverse', (1, None)),
('love.', (5, 1)),
('gay,', (1, None)),
('anything', (27, 4)),
('work', (29, 4)),
('hiv', (15, 1)),
('years', (52, 17)),
('ago,', (4, None)),
('passed', (3, None)),
('away', (19, 3)),
('new', (25, 5)),
('run,', (1, None)),
('said', (26, 5)),
('saying', (19, 2)),
('kind', (12, 4)),
('definitive', (1, None)),
('believe', (41, 6)),
('miracles.', (1, None)),
('watch', (11, 2)),
('here.', (3, None)),
('#predict', (2, None)),
('&', (123, 28)),
('#cure', (2, None)),
('#cancer', (32, 6)),
('minutes', (3, None)),
('#reiki', (2, None)),
('#rtrrt', (4, None)),
('#bigpharma', (4, None)),
('#loa', (2, None)),
('#acim', (2, None)),
('home', (10, 1)),
('predicta...', (2, None)),
('https://t.co/4r1zgo8opo', (1, None)),
('https://t.co/ojby25xlyk', (1, None)),
('1st', (7, None)),
('give', (56, 7)),
('vote', (8, 1)),
('-not', (1, None)),
('hillary', (3, None)),
('dr.', (11, 6)),
('video=', (1, None)),
('clinton', (2, None)),
('94', (1, None)),

```

```

('#trumptrainjumponit', (1, None)),
('via', (16, 5)),
('poor', (7, None)),
('clinton!', (1, None)),
('decades', (4, 2)),
('suppressing', (1, None)),
('cancer-cures', (1, None)),
('wants', (7, None)),
('#righttocry', (1, None)),
('#righttodie', (1, None)),
('ask', (15, None)),
('patients', (28, 4)),
('right!', (5, None)),
('-both', (1, None)),
('f*cked', (1, None)),
('https://t.co/xpikezppcz', (1, None)),
('#petitions', (1, None)),
('#dogs', (1, None)),
('signatures', (1, None)),
('terminal', (7, 2)),
('diseases', (25, 9)),
('known', (14, 2)),
('cause', (23, 3)),
('kill', (15, 5)),
('them! ', (1, None)),
('needs', (19, 2)),
('like', (155, 15)),
('care ', (1, None)),
('https://t.co/zzcp7rnld0', (1, None)),
('#sisterscare:', (1, None)),
('@burrelltv', (1, None)),
('@efinchtv', (2, None)),
('@danatylertv', (2, None)),
('took', (12, None)),
('event', (13, None)),
('today,', (7, None)),
('disparities', (1, None)),
('diagnosis', (6, 1)),
('african', (4, 2)),
('https://t.co/unjsqhcycd6', (1, None)),
('doctor', (11, 2)),
('revolution', (2, 1)),
('teaching', (3, None)),
('#teamhalon', (1, 1)),
('supports', (3, 1)),
('#randomactsoflightday', (1, 1)),
('continued', (2, 1)),
('https://t.co/q2lt7y56km', (1, 1)),

```

```

('eradicate', (5, None)),
('put', (35, 1)),
('negative', (10, None)),
('it.', (65, 11)),
('consistently', (1, None)),
('are.', (5, None)),
('discover', (9, None)),
('injection', (1, None)),
('stinging.#liberalhypocrisy', (1, None)),
('#deepstate', (1, None)),
('self-inflicted', (1, None)),
('metastasized', (2, 1)),
('full', (16, 1)),
('outbreak', (1, None)),
('nevertrumpers', (1, None)),
('demise.', (2, 1)),
('let', (22, None)),
('go.', (2, None)),
('ceo', (2, 2)),
('@garo_armen', (1, 1)),
('@agenus_bio', (1, 1)),
('@cancerresearch', (2, 2)),
('fight', (39, 4)),
('#oncology', (2, 1)),
('#curecancer', (1, 1)),
('https://t.co/f5vigul6ky', (1, 1)),
('stand', (9, 1)),
('courageous', (2, 1)),
('physicians', (2, 1)),
('#inpursuitof', (3, 3)),
('https://t.co/ve97oxqakw', (1, 1)),
('engine', (1, 1)),
('produced', (1, 1)),
('inds', (2, 2)),
('18mos;', (1, 1)),
('planned', (3, 2)),
('1h2019.', (1, 1)),
('molecules', (1, 1)),
('ind', (1, 1)),
('plus,', (2, 1)),
('treated', (7, 2)),
('agen1884', (1, 1)),
('agen2034', (1, 1)),
('#asco2018.', (1, 1)),
('speed', (5, 3)),
('https://t.co/vrwjomjmvo', (1, 1)),
('right', (26, 8)),
('combination', (3, 2)),

```

('science', (18, 6)),  
 ('people-see', (3, 2)),  
 ('passion', (6, 2)),  
 ('advancing', (3, 2)),  
 ('toward', (4, 2)),  
 ('goal', (7, 2)),  
 ('https://t.co/apxu2dplkm', (3, 2)),  
 ('@big3biobos', (3, 2)),  
 ('\$celz', (1, None)),  
 ('imho,', (1, None)),  
 ('beaten', (4, 1)),  
 ('rate', (20, 2)),  
 ('#pharma', (1, None)),  
 ('treatments', (17, 7)),  
 ('illegal', (8, 2)),  
 ('u.s', (2, None)),  
 ('celz's', (1, None)),  
 ('help', (124, 17)),  
 ('huge!', (1, None)),  
 ('helpful.', (1, 1)),  
 ('i'm', (42, 4)),  
 ('sure', (17, 2)),  
 ('it's', (73, 13)),  
 ('wanting', (1, 1)),  
 ('helpful,', (1, 1)),  
 ('good', (54, 10)),  
 ('deeds', (1, 1)),  
 ('faith', (7, 1)),  
 ('impacted', (1, 1)),  
 ('none', (5, 1)),  
 ('around', (9, 3)),  
 ('idea', (7, 1)),  
 ('tell', (20, 3)),  
 ('story', (15, 2)),  
 ('lesnar', (1, None)),  
 ('looking', (15, 5)),  
 ('belt', (1, None)),  
 ('roman', (1, None)),  
 ('reigns', (1, None)),  
 ('end', (57, 3)),  
 ('hunger,', (12, None)),  
 ('pushed', (1, None)),  
 ('hard.', (1, None)),  
 ('news\*', (1, None)),  
 ('finds', (7, None)),  
 ('mother:', (1, None)),  
 ('"what', (2, None)),  
 ('shit!', (1, None)),

```

('probably', (15, None)),
('time!""', (1, None)),
('shit', (23, 3)),
('does/says,', (1, None)),
("ya'll", (1, None)),
('never', (60, 3)),
('*opens', (1, None)),
('cab-booking', (1, None)),
('*books', (1, None)),
('cab*', (1, None)),
('years,', (6, 1)),
('months,', (1, None)),
('weeks,', (1, None)),
('days', (9, 1)),
('hours.', (1, None)),
('scientists', (24, 6)),
('volcano', (1, None)),
('erupt', (1, None)),
('*this', (1, 1)),
('chimes*', (1, 1)),
('can't', (32, 3)),
('neither', (4, 2)),
('fighting', (9, 2)),
('narrative', (3, 1)),
('created', (5, 3)),
('narratives.', (1, 1)),
('more*', (1, 1)),
('#lifewithcancer', (1, 1)),
('@abbvie', (2, None)),
('alphabet-backed', (2, None)),
('@erictrump:', (1, None)),
('money', (80, 4)),
('we're', (11, None)),
('win', (16, None)),
('fight.', (2, None)),
('pediatric', (12, None)),
('https://t.co/7pyfcd7xzy', (1, None)),
('son', (7, 2)),
('due', (4, 1)),
('started', (5, 1)),
('jude.', (1, None)),
('year', (31, 2)),
('survivor', (12, 1)),
('loose', (2, None)),
('month', (6, 1)),
('children's', (2, None)),
('hospital', (8, None)),
('childhood', (25, 1)),

```

('https://t.co/gwqtxflqau', (1, None)),  
 ('11:11', (4, None)),  
 ('excellent', (4, None)),  
 ('grades!!!', (2, None)),  
 ('job', (14, None)),  
 ('best', (23, 6)),  
 ('happiness', (7, 1)),  
 ('gun', (7, None)),  
 ('@shawnmendes', (4, None)),  
 ('bye', (5, None)),  
 ('summer', (4, None)),  
 ('lion', (9, None)),  
 ('rid', (10, 1)),  
 ('love', (28, 2)),  
 ('grades', (1, None)),  
 ('summer!', (1, None)),  
 ('harry', (3, None)),  
 ('potter', (1, None)),  
 ('reunion', (3, None)),  
 ('weight!', (1, None)),  
 ('16-year-old', (4, None)),  
 ('thomas,', (3, None)),  
 ('battling', (6, 1)),  
 ('used', (14, 3)),  
 ('make', (92, 8)),  
 ('abortion.', (5, None)),  
 ('jeremiah,', (3, None)),  
 ('life,', (14, 1)),  
 ('many.', (4, None)),  
 ('https://t.co/bumzmggyay', (3, None)),  
 ('sad', (17, 4)),  
 ('brainwashed', (2, None)),  
 ('maybe', (25, 4)),  
 ('idk', (1, None)),  
 ('mann', (1, None)),  
 ('pretty', (5, None)),  
 ('try', (20, 2)),  
 ('use', (28, 18)),  
 ('religion', (11, None)),  
 ('controls', (1, None)),  
 ('else's', (1, None)),  
 ('body...', (1, None)),  
 ('coulda', (1, None)),  
 ('https://t.co/tyx3hynfa8', (1, None)),  
 ('stage', (23, 5)),  
 ('carcinoid', (12, None)),  
 ('longest', (2, None)),  
 ('survival', (12, 1)),

('5yrs.', (6, None)),  
 ('read', (26, 2)),  
 ('miracle', (13, 1)),  
 ('book', (13, None)),  
 ('living', (17, 3)),  
 ('enemy', (11, None)),  
 ('support', (32, 1)),  
 ('misdiagnosed', (7, None)),  
 ('endcancer#', (1, None)),  
 ('flying', (8, None)),  
 ('future"', (1, None)),  
 ('1987:', (1, None)),  
 ('2018:', (2, None)),  
 ('potentially', (5, None)),  
 ('cancers', (17, 8)),  
 ('twenty', (2, None)),  
 ('cancer!'", (1, None)),  
 ('"dennis', (1, None)),  
 ('rodman', (1, None)),  
 ('nuclear', (4, None)),  
 ('united', (3, None)),  
 ('states,', (1, None)),  
 ('korea!'", (1, None)),  
 ('@panmass', (1, None)),  
 ('rider', (1, None)),  
 ('https://t.co/qdwqn3zzre', (1, None)),  
 ('semi-finalist', (1, None)),  
 ('ubc', (1, None)),  
 ('master', (1, None)),  
 ('student', (2, None)),  
 ('warner', (1, None)),  
 ('presented', (1, None)),  
 ("shakespeare's", (1, None)),  
 ('thing?', (2, None)),  
 ('asked.', (1, None)),  
 ('#tbt', (1, None)),  
 ('https://t.co/vjazjwo5em', (1, None)),  
 ('22,', (1, None)),  
 ('diagnosed', (10, 2)),  
 ('multiple', (12, 8)),  
 ('typically', (1, None)),  
 ('afflicting', (2, None)),  
 ('hospitalizations,', (1, None)),  
 ('chemotherapy,', (3, 2)),  
 ('followed.', (1, None)),  
 ('https://t.co/f0lztvodcb', (1, None)),  
 ('5,000', (2, None)),  
 ('young', (13, 1)),



('women', (17, 1)),  
 ('country', (16, 1)),  
 ('participated', (2, None)),  
 ('year's', (4, None)),  
 ('tournament', (2, None)),  
 ('raised', (16, None)),  
 ('\$235,000', (1, None)),  
 ('csnj', (1, None)),  
 ('more:', (2, None)),  
 ('https://t.co/4qlw3bei7k', (1, None)),  
 ('https://t.co/uwdnketoie', (1, None)),  
 ('6th', (4, None)),  
 ('already', (20, 4)),  
 ('£63k', (1, None)),  
 ('colon', (4, None)),  
 ('past', (5, 1)),  
 ('years.', (10, 2)),  
 ('miles', (2, None)),  
 ('finishing', (1, None)),  
 ('@curebowelcancer', (1, None)),  
 ('https://t.co/dn5cjvm9gz', (1, None)),  
 ('republicans', (6, None)),  
 ('trump.', (9, None)),  
 ('omg.', (1, None)),  
 ('how?', (2, None)),  
 ('ripped', (1, None)),  
 ('mother's', (2, None)),  
 ('truly', (11, 3)),  
 ('80%', (2, None)),  
 ('news', (11, 2)),  
 ('#disgusted', (1, None)),  
 ('must', (21, 1)),  
 ('grasp', (1, None)),  
 ('afternoon', (1, None)),  
 ('nature.', (2, None)),  
 ('numbers.', (1, None)),  
 ('https://t.co/s7ctd0lyc6', (1, None)),  
 ('@realjameswoods', (4, None)),  
 ('is,', (7, 2)),  
 ('somehow', (5, 1)),  
 ('instrumental', (1, None)),  
 ('it,', (23, 2)),  
 ('jackass', (1, None)),  
 ('it!!!!', (1, None)),  
 ('you're', (12, 2)),  
 ('be"', (1, None)),  
 ('pathetic', (1, None)),  
 ('enough!!', (1, None)),

```

('souls', (2, None)),
('way', (47, 4)),
('trump!!!!', (1, None)),
('@1999mcm11', (1, None)),
('denuclearization', (3, None)),
('achieved.', (2, None)),
('xbox,', (1, None)),
('peace...', (1, None)),
('complain', (32, None)),
('orange', (2, None)),
('stupid', (2, None)),
('stuff.', (2, None)),
('@2mattandgaby2', (1, 1)),
('no.', (5, 2)),
('laetrile', (3, 3)),
('else.', (2, 1)),
('studied,', (1, 1)),
('work.', (8, 2)),
('stop', (35, 10)),
('https://t.co/7z0jpa3osp', (1, 1)),
('@420stonercorner', (2, 2)),
('subject', (2, 2)),
('matter', (27, 3)),
('thought', (16, 3)),
('such,', (3, 2)),
('rampant', (2, 2)),
('family', (26, 5)),
('cbd', (6, 5)),
('instead', (16, 3)),
('#cannabis', (2, 2)),
('medically', (1, 1)),
('unqualified', (1, 1)),
('opinion?', (1, 1)),
('#bbcaq', (1, 1)),
('assholes', (2, 1)),
('criminalizing', (1, 1)),
('@norml', (2, 2)),
('#nypd', (1, 1)),
('brings', (3, 2)),
('forever!', (1, 1)),
('https://t.co/f583xnt3jy', (1, 1)),
('provide', (7, 1)),
('shelter', (1, None)),
(',can't', (2, None)),
('aids', (34, 2)),
('although', (3, 1)),
('sometimes', (2, 1)),
('always', (17, 2)),

```

('ppl', (11, 5)),  
 ('fuck', (18, None)),  
 ('underlying', (1, None)),  
 ('crime.', (3, None)),  
 ('beating', (4, None)),  
 ('mentally', (2, None)),  
 ('ill', (6, 1)),  
 ('caricature', (1, None)),  
 ('bruce', (1, None)),  
 ('literally', (20, 2)),  
 ('meme', (2, 1)),  
 ('sexual', (4, None)),  
 ('leading', (7, 3)),  
 ('lives.', (4, 1)),  
 ('helping', (14, 1)),  
 ('spared', (1, None)),  
 ('efforts', (4, None)),  
 ('last', (39, 9)),  
 ('resolve', (2, None)),  
 ('brotherly', (1, None)),  
 ('basis', (3, 1)),  
 ('continues', (2, None)),  
 ('path', (2, None)),  
 ('saudi', (2, None)),  
 ('region.', (1, None)),  
 ('remove', (2, None)),  
 ('@abby\_higgins5:', (1, None)),  
 ('government/pharmacies', (1, None)),  
 ('wanna', (6, None)),  
 ('money...', (1, None)),  
 ('https://t.co/ivompefmo0', (1, None)),  
 ('thing...', (1, None)),  
 ('yet', (18, 2)),  
 ('earned', (1, None)),  
 ('millennium', (1, None)),  
 ('prize', (5, 1)),  
 ('think', (61, 8)),  
 ('@acosta', (4, None)),  
 ('hatred', (5, None)),  
 ('wrong.', (7, 1)),  
 ('pity', (2, None)),  
 ('people.', (10, 1)),  
 ('@adeabdwbb', (1, None)),  
 ('tweak', (1, None)),  
 ('question', (7, 1)),  
 ('youre', (3, None)),  
 ('dying.', (1, None)),  
 ('choices:', (1, None)),

```

('huge', (5, None)),
('amount', (5, 1)),
('money.', (8, 1)),
('basically', (4, 1)),
('bcause', (1, None)),
('choose.', (1, None)),
('@adevotedyogi:', (1, 1)),
('cockroach', (1, 1)),
('sushi?', (1, 1)),
('farming', (2, 2)),
('shake', (1, 1)),
('menus', (1, 1)),
('@adryzzle', (1, None)),
('repeal', (2, None)),
('killed', (11, 2)),
('may', (36, 7)),
('adverts', (2, None)),
('pre', (3, None)),
('2015', (1, None)),
('"help', (1, None)),
('"making', (1, None)),
('safer', (1, None)),
('making', (22, 1)),
('safer"', (1, None)),
('@cr_uk', (2, None)),
('frauds', (1, None)),
('protecting', (1, None)),
('pharmaceutical', (5, 1)),
('https://t.co/mb6kdnslja', (1, None)),
('trying', (15, 2)),
('power', (10, 5)),
('10yrs', (1, 1)),
('exposure', (1, 1)),
('open', (8, 3)),
('ppls', (1, 1)),
('eyes', (2, 1)),
('endocannabinoid', (1, 1)),
('https://t.co/r2dzgjavu', (1, 1)),
('@aeidinger', (1, 1)),
('@potus', (26, 1)),
('@bencarsonizer', (1, 1)),
('1930's..', (1, 1)),
('renew', (1, 1)),
('15-30', (1, 1)),
('so.', (5, 2)),
('stepped', (1, None)),
('scene', (1, None)),
('nut', (2, None)),

```

('shout', (2, None)),  
 ('skin', (7, None)),  
 ('lacey', (1, None)),  
 ('duvalle', (1, None)),  
 ('intentions', (1, None)),  
 ('good,', (1, None)),  
 ('reduce', (6, 2)),  
 ('heel', (2, None)),  
 ('praised', (2, None)),  
 ('indirectly', (1, None)),  
 ('feminism', (1, None)),  
 ('wearing', (5, None)),  
 ('awareness.', (1, None)),  
 ('https://t.co/kzcbgzcdsp', (1, None)),  
 ('similar', (3, None)),  
 ('response', (5, 1)),  
 ('"go', (3, None)),  
 ('lab', (5, None)),  
 ('yours".', (1, None)),  
 ('nutshell,', (1, None)),  
 ('means', (9, None)),  
 ('works,', (1, None)),  
 ('simple.', (2, None)),  
 ('seriously', (6, 1)),  
 ('right.', (2, None)),  
 ('lost', (38, 5)),  
 ('and/or', (1, None)),  
 ('pharmaceuticals.', (1, None)),  
 ('@amaroqwolf', (1, None)),  
 ('line', (2, None)),  
 ('benefits', (2, 1)),  
 ('time-travel.', (1, None)),  
 ('told', (11, 4)),  
 ('"give', (1, None)),  
 ('america', (14, None)),  
 ('house', (3, None)),  
 ('live', (24, 2)),  
 ('private', (1, None)),  
 ('rino's", (1, None)),  
 ('throw', (2, None)),  
 ('fuel', (3, 1)),  
 ('plane.', (1, None)),  
 ('applying', (3, None)),  
 ('east', (2, None)),  
 ('german', (3, None)),  
 ('communist', (1, None)),  
 ('autocracy', (1, None)),  
 ('policies,', (2, None)),

('spoiling', (1, None)),  
 ('both.', (2, 1)),  
 ('undermined', (1, None)),  
 ('social', (6, None)),  
 ('agreement', (1, None)),  
 ('knuckleheaded', (1, None)),  
 ('everything', (20, 5)),  
 ('negotiated', (2, None)),  
 ('"all', (1, None)),  
 ('@ancientaliens', (1, None)),  
 ('alien', (3, None)),  
 ('identity', (2, None)),  
 ('creator,', (2, 1)),  
 ('amongst', (1, None)),  
 ('moves', (3, None)),  
 ('tube', (2, None)),  
 ('cream', (2, 1)),  
 ('low', (5, 3)),  
 ('nominated', (1, None)),  
 ('dictatorial', (1, None)),  
 ('overreach.', (1, None)),  
 ('headlines', (4, None)),  
 ('read:', (2, None)),  
 ('"millions', (1, None)),  
 ('die', (19, 6)),  
 ('acts."', (1, None)),  
 ('@angelicjane03', (1, None)),  
 ('@abctdiagnosis', (1, None)),  
 ('@nursedave34', (1, None)),  
 ('primarily', (1, None)),  
 ('totally', (5, 1)),  
 ('ignores', (2, 1)),  
 ('cards', (1, None)),  
 ('nixon', (2, None)),  
 ('legal', (1, None)),  
 ('babies', (7, None)),  
 ('could've', (5, None)),  
 ('female', (6, None)),  
 ('#overturnroevswade', (1, None)),  
 ('@angielovesusa', (1, None)),  
 ('true.', (4, None)),  
 ('focused', (4, None)),  
 ('rather', (3, None)),  
 ('happen', (4, None)),  
 ('sooner.', (2, None)),  
 ('generation', (3, 1)),  
 ('two', (14, 3)),  
 ('training', (3, None)),

('witness', (2, 1)),  
 ('rare', (5, 1)),  
 ('@aplauriekellman', (1, None)),  
 ('hacks', (1, None)),  
 ('unbelievable.', (1, None)),  
 ('typing', (1, None)),  
 ('yellow', (1, None)),  
 ('emirati', (1, None)),  
 ('solution', (6, None)),  
 ('global', (9, 3)),  
 ('deserve', (6, None)),  
 ('awards', (1, None)),  
 ('streets', (2, None)),  
 ('idiots', (2, None)),  
 ('nation', (1, None)),  
 ('@artobsessed4', (1, None)),  
 ('@foxnews', (5, None)),  
 ('does,', (6, None)),  
 ('haters', (10, None)),  
 ('hate', (36, 4)),  
 ('protest', (4, None)),  
 ('@ashtonbirdie:', (1, None)),  
 ('sexist', (2, None)),  
 ('didn't', (13, None)),  
 ('@bornfreefdn', (1, 1)),  
 ('tragically,', (1, 1)),  
 ('mommy', (1, 1)),  
 ('slaughtered', (1, 1)),  
 ('remote', (2, 1)),  
 ('reserve.', (1, 1)),  
 ('happened', (2, 1)),  
 ('king', (1, 1)),  
 ('enjoying', (1, 1)),  
 ('disgusting!', (1, 1)),  
 ('rhino', (8, 7)),  
 ('penises', (2, 2)),  
 ('got', (28, 4)),  
 ('they'll', (3, None)),  
 ('public', (8, None)),  
 ('there's', (25, 6)),  
 ('treating', (4, 2)),  
 ('curing', (9, 2)),  
 ('guess.', (1, None)),  
 ('single', (10, 2)),  
 ('swoop', (1, None)),  
 ('@aynrandpaulryan', (1, None)),  
 ('antidote.', (1, None)),  
 ('waiting', (3, None)),

```

('penalty', (1, None)),
('shirt', (2, None)),
('admin', (1, None)),
('@bangronknandez', (1, None)),
('@chilomane303', (1, None)),
('fucked', (4, None)),
('@bankdeny', (1, None)),
('@geertwilderspvv', (1, None)),
('speech', (3, None)),
('progress..', (1, None)),
('ultimately', (2, None)),
('possible', (6, None)),
('truth.', (2, None)),
('worse', (7, None)),
('colors', (2, None)),
('show.', (1, None)),
('@bertholftre:', (1, None)),
('really', (44, 15)),
('https://t.co/rcsshyy4k5', (1, None)),
('cured', (16, 6)),
('@sheffnorr', (1, None)),
('socialism', (9, None)),
('capitalism.', (4, None)),
('springboard', (4, None)),
('nazism,', (4, None)),
('rotten', (4, None)),
('sickness', (8, None)),
('killing', (11, 2)),
('processes', (4, None)),
('spirit', (4, None)),
('party.', (5, None)),
('eczema?', (1, None)),
('@bleakfacts:', (1, None)),
('results', (4, 1)),
('@bobo_697', (1, None)),
('@cnn', (12, None)),
('potential', (9, 3)),
('shares', (5, 1)),
('kinda', (4, 1)),
('wya', (1, None)),
('@donlemon', (2, None)),
('#takemehome', (1, None)),
('@bocavista2016', (1, None)),
('free', (24, 5)),
('terrible', (8, None)),
('putting', (16, None)),
('business.', (7, 1)),
('walk', (12, None)),

```



('water', (9, None)),  
 ('establish', (2, None)),  
 ('wait!', (1, None)),  
 ('@boegershausen\_', (1, None)),  
 ('@feysperson', (1, None)),  
 ('worked', (5, None)),  
 ('incredible', (10, None)),  
 ('hours', (7, None)),  
 ('send', (5, 1)),  
 ('compassionate', (1, None)),  
 ('wonder', (10, None)),  
 ('cancer.we', (1, None)),  
 ('immigrants', (5, None)),  
 ('placed', (1, None)),  
 ('non-curable', (1, None)),  
 ('@mountsinainyc', (1, 1)),  
 ('approach', (1, 1)),  
 ('medication,', (1, 1)),  
 ('alternative', (2, 1)),  
 ('particular', (1, 1)),  
 ('cancers.', (4, 1)),  
 ('smart', (3, 1)),  
 ('death.', (4, 2)),  
 ('@realdonaldtrump', (46, None)),  
 ('dead...', (1, None)),  
 ('.25', (1, None)),  
 ('cents', (2, None)),  
 ('gallon', (1, None)),  
 ('conservatives', (1, None)),  
 ('said...', (1, None)),  
 ('quite', (3, None)),  
 ('grounded', (1, None)),  
 ('solving', (2, None)),  
 ('not.', (6, 1)),  
 ('let's', (9, 1)),  
 ('hook', (2, None)),  
 ('capacity,', (1, None)),  
 ('ambition', (1, None)),  
 ('@nicolledwallace', (2, None)),  
 ('auto', (2, None)),  
 ('staving', (1, None)),  
 ('depression?', (3, None)),  
 ('millions', (12, None)),  
 ('expense', (1, None)),  
 ('professional', (2, None)),  
 ('feel', (8, 2)),  
 ('obama.', (1, None)),  
 ('@bulldog2026', (1, None)),

```

('@georgehwbush', (1, None)),
('claim', (13, 1)),
('evidence,', (1, None)),
('lauded', (1, None)),
('genius', (2, None)),
('hero.', (2, None)),
('gullibility.', (1, None)),
('@bushrahussein1:', (1, None)),
('£15000', (1, None)),
("mother's", (3, None)),
('@justgiving', (1, None)),
('https://t.co/...', (1, None)),
('@busywriting', (1, None)),
('@krystal337', (1, None)),
('@brithume', (1, None)),
('talking', (10, 3)),
('points', (2, None)),
('permeating', (1, None)),
('subconscious.', (1, None)),
('dead,', (3, None)),
('eliminate', (2, None)),
('libs', (8, None)),
('points.', (1, None)),
(' ', (1, None)),
('@c_3c_3', (1, None)),
('acknowledge', (3, None)),
('attempted', (1, None)),
('coup', (1, None)),
('fire', (1, None)),
('deeper', (1, None)),
('underground', (2, None)),
('org', (1, None)),
('run', (14, None)),
('jarret', (1, None)),
('co.', (1, None)),
('@cabeyoslatina:', (1, None)),
('https://t.co/vlvvmg460o', (1, None)),
('@calvincardwell:', (1, None)),
('operate', (1, None)),
('ca...', (1, None)),
('@campaignkate', (1, 1)),
('towards', (11, 2)),
('"positivity"', (1, 1)),
('statement', (4, 3)),
('"positive"', (1, 1)),
('charming', (1, 1)),
('@cannaworldsite:', (1, 1)),
('amelia', (2, 2)),

```

```

('using', (15, 8)),
('tumor', (5, 2)),
('fungus', (1, 1)),
('mining', (2, 1)),
('dna-sequence', (1, 1)),
('mushrooms', (1, 1)),
('https://t.co/o78q0kgzoa', (1, 1)),
('@cass_alexandra_', (1, None)),
('neurotypical', (1, None)),
('"drink', (1, None)),
('things', (15, 1)),
('different', (18, 2)),
('"how', (1, None)),
('reminded', (3, 1)),
('@crackedscie...', (1, None)),
('@cfourkids44:', (1, None)),
('@chaeyoungpulp:', (1, 1)),
('pepto', (1, 1)),
('bismol:', (1, 1)),
('https://t.co/pngwyh689w', (1, 1)),
('hour.', (1, None)),
('@caramariamtv', (1, None)),
('@charlestillman', (1, None)),
('no,', (5, 2)),
('journalists', (2, None)),
('investigate', (1, None)),
('match', (3, None)),
('agenda.', (2, None)),
('stations', (1, None)),
('thing.', (6, None)),
('@chercooper2', (1, None)),
('squirming', (1, None)),
('corrupt', (4, None)),
('demand', (3, 1)),
('@gop', (4, None)),
('@housegop', (1, None)),
('@caitrionambalfe', (1, None)),
('@hillaryclinton', (1, None)),
('way.', (5, 1)),
('@chicagoinno:', (1, None)),
('point', (10, 1)),
('korea,', (2, None)),
('middle', (4, 1)),
('east,', (1, None)),
('redeem', (1, None)),
('awfulness', (1, None)),
('@chrystalfigaro:', (1, None)),
('etc.?', (1, None)),

```

('https://t.co/fcv3picioe', (1, None)),  
('@amjoyshow', (1, None)),  
('60,000,000', (1, None)),  
('murderous', (2, None)),  
('chance', (12, None)),  
('minds', (6, None)),  
('developed', (2, None)),  
('universal', (4, None)),  
('"clean"', (1, None)),  
('replace', (2, 1)),  
('about,', (5, None)),  
('snuffed', (1, None)),  
('of.', (2, 1)),  
('@cifertherhyme', (1, None)),  
('@dccomics', (1, None)),  
('@nytimes', (1, None)),  
('married', (1, None)),  
('did,', (2, None)),  
('promote', (4, None)),  
('intensely', (1, None)),  
('months', (6, 1)),  
('"i\'ve', (1, None)),  
('cancer"', (4, 1)),  
('bored', (1, None)),  
('https://t.co/szdcpokayd', (1, None)),  
('there\'ll', (2, None)),  
('type', (8, None)),  
('major', (5, 1)),  
('long', (21, 4)),  
('resulting', (1, None)),  
('infertility', (1, None)),  
('rallying', (1, None)),  
('prevention', (4, 1)),  
('wonderful', (5, 2)),  
('treatment.', (4, 1)),  
('epilepsy', (1, 1)),  
('evidence', (11, 6)),  
('change', (11, 2)),  
('people's', (2, 2)),  
('rigorously', (1, 1)),  
('tested', (2, 2)),  
('side-effects.', (1, 1)),  
('https://t.co/dcurelgpsr', (1, 1)),  
('@cnamulindwa', (1, None)),  
('@ugandacancerins', (1, None)),  
('due!', (1, None)),  
('deal', (5, None)),  
('types...', (1, None)),

```

('notable', (1, None)),
('democrats.', (4, None)),
('would.say', (1, None)),
('individuals.', (1, None)),
('immigrant', (5, None)),
('https://t.co/4g5sgjykfi', (1, None)),
('@stcollinson', (1, None)),
('families', (13, None)),
('bad.', (8, None)),
('summit', (8, None)),
('kim:', (4, None)),
('listen,', (1, None)),
('cnn,', (1, None)),
('@thedailybeast', (1, None)),
('call', (10, 4)),
('up!!!!', (1, None)),
('crazy', (5, None)),
('sun', (3, None)),
('glow,', (2, None)),
('two-faced...', (1, None)),
('@ellimayradrep', (1, None)),
('hate!', (2, None)),
('blame', (9, None)),
('#realtalk', (1, None)),
(' ', (1, None)),
('jacket?!', (1, None)),
('wasted', (2, None)),
('#teamtrump', (1, None)),
('cleaning', (1, None)),
('https://t.co/ga9bsyuk6k', (1, None)),
('@daneyellnicole', (1, None)),
('usa', (3, None)),
('countries', (6, None)),
('fix', (8, 1)),
('oxygen', (2, None)),
('alkaline', (4, 2)),
('@sheisfondofbks', (1, None)),
('@rizzlebetta', (1, None)),
('@sashadevlin', (1, None)),
('@ashleyn1cole', (1, None)),
('depends', (1, None)),
('"evil".', (1, None)),
...]

```

```
[ ]: word_table.collect()
```

```

Out[29]: [('ever', ((18, 2), 16)),
          ('scientific', ((10, 3), 7)),

```

('medical', ((27, 11), 16)),  
('yoga', ((3, 2), 1)),  
('cancer,', ((183, 33), 150)),  
('cardiac', ((2, 2), None)),  
('#cureforcancer', ((1, None), 1)),  
('100', ((3, None), 3)),  
('2014', ((2, None), 2)),  
('know', ((60, 14), 46)),  
('needed', ((7, 2), 5)),  
('cancer.', ((162, 25), 137)),  
('bitch', ((12, None), 12)),  
('even', ((74, 11), 63)),  
('become', ((11, 1), 10)),  
('diverse', ((1, None), 1)),  
('gay,', ((1, None), 1)),  
('away', ((19, 3), 16)),  
('new', ((25, 5), 20)),  
('run,', ((1, None), 1)),  
('saying', ((19, 2), 17)),  
('kind', ((12, 4), 8)),  
('&', ((123, 28), 95)),  
('#reiki', ((2, None), 2)),  
('clinton', ((2, None), 2)),  
('#trumptrainjumponit', ((1, None), 1)),  
('clinton!', ((1, None), 1)),  
('suppressing', ((1, None), 1)),  
('cancer-cures', ((1, None), 1)),  
('wants', ((7, None), 7)),  
('#righttodie', ((1, None), 1)),  
('ask', ((15, None), 15)),  
('-both', ((1, None), 1)),  
('f\*cked', ((1, None), 1)),  
('https://t.co/xpikezppcz', ((1, None), 1)),  
('#petitions', ((1, None), 1)),  
('terminal', ((7, 2), 5)),  
('diseases', ((25, 9), 16)),  
('them! ', ((1, None), 1)),  
('needs', ((19, 2), 17)),  
('like', ((155, 15), 140)),  
('@danatylertv', ((2, None), 2)),  
('took', ((12, None), 12)),  
('disparities', ((1, None), 1)),  
('diagnosis', ((6, 1), 5)),  
('african', ((4, 2), 2)),  
('teaching', ((3, None), 3)),  
('#teamhalon', ((1, 1), None)),  
('https://t.co/q2lt7y56km', ((1, 1), None)),  
('negative', ((10, None), 10)),

```

('are.', ((5, None), 5)),
('injection', ((1, None), 1)),
('stinging.#liberalhypocrisy', ((1, None), 1)),
('#deepstate', ((1, None), 1)),
('metastasized', ((2, 1), 1)),
('full', ((16, 1), 15)),
('outbreak', ((1, None), 1)),
('nevertrumpers', ((1, None), 1)),
('demise.', ((2, 1), 1)),
('let', ((22, None), 22)),
('#inpursuitof', ((3, 3), None)),
('1h2019.', ((1, 1), None)),
('agen2034', ((1, 1), None)),
('#asco2018.', ((1, 1), None)),
('https://t.co/vrwjomjmvo', ((1, 1), None)),
('right', ((26, 8), 18)),
('combination', ((3, 2), 1)),
('people-see', ((3, 2), 1)),
('toward', ((4, 2), 2)),
('@big3biobos', ((3, 2), 1)),
('sure', ((17, 2), 15)),
('deeds', ((1, 1), None)),
('impacted', ((1, 1), None)),
('lesnar', ((1, None), 1)),
('looking', ((15, 5), 10)),
('pushed', ((1, None), 1)),
('finds', ((7, None), 7)),
('shit!', ((1, None), 1)),
('probably', ((15, None), 15)),
('years,', ((6, 1), 5)),
('weeks,', ((1, None), 1)),
('scientists', ((24, 6), 18)),
('volcano', ((1, None), 1)),
('chimes*', ((1, 1), None)),
('can't', ((32, 3), 29)),
('narrative', ((3, 1), 2)),
('#lifewithcancer', ((1, 1), None)),
('@abbvie', ((2, None), 2)),
('@erictrump:', ((1, None), 1)),
('we're', ((11, None), 11)),
('win', ((16, None), 16)),
('https://t.co/7pyfcd7xzy', ((1, None), 1)),
('son', ((7, 2), 5)),
('https://t.co/gwqtxflqau', ((1, None), 1)),
('excellent', ((4, None), 4)),
('job', ((14, None), 14)),
('gun', ((7, None), 7)),
('bye', ((5, None), 5)),

```

('love', ((28, 2), 26)),  
 ('reunion', ((3, None), 3)),  
 ('used', ((14, 3), 11)),  
 ('make', ((92, 8), 84)),  
 ('many.', ((4, None), 4)),  
 ('maybe', ((25, 4), 21)),  
 ('pretty', ((5, None), 5)),  
 ('religion', ((11, None), 11)),  
 ('else's', ((1, None), 1)),  
 ('body...', ((1, None), 1)),  
 ('stage', ((23, 5), 18)),  
 ('survival', ((12, 1), 11)),  
 ('book', ((13, None), 13)),  
 ('living', ((17, 3), 14)),  
 ('misdiagnosed', ((7, None), 7)),  
 ('flying', ((8, None), 8)),  
 ('future"', ((1, None), 1)),  
 ('1987:', ((1, None), 1)),  
 ('potentially', ((5, None), 5)),  
 ('twenty', ((2, None), 2)),  
 ('rider', ((1, None), 1)),  
 ('semi-finalist', ((1, None), 1)),  
 ('master', ((1, None), 1)),  
 ('thing?', ((2, None), 2)),  
 ('asked.', ((1, None), 1)),  
 ('https://t.co/vjazjwo5em', ((1, None), 1)),  
 ('22,', ((1, None), 1)),  
 ('diagnosed', ((10, 2), 8)),  
 ('typically', ((1, None), 1)),  
 ('5,000', ((2, None), 2)),  
 ('women', ((17, 1), 16)),  
 ('tournament', ((2, None), 2)),  
 ('more:', ((2, None), 2)),  
 ('£63k', ((1, None), 1)),  
 ('years.', ((10, 2), 8)),  
 ('finishing', ((1, None), 1)),  
 ('trump.', ((9, None), 9)),  
 ('80%', ((2, None), 2)),  
 ('#disgusted', ((1, None), 1)),  
 ('must', ((21, 1), 20)),  
 ('grasp', ((1, None), 1)),  
 ('nature.', ((2, None), 2)),  
 ('@realjameswoods', ((4, None), 4)),  
 ('you're', ((12, 2), 10)),  
 ('denuclearization', ((3, None), 3)),  
 ('peace...', ((1, None), 1)),  
 ('complain', ((32, None), 32)),  
 ('orange', ((2, None), 2)),



```

('studied,', ((1, 1), None)),
('stop', ((35, 10), 25)),
('such,', ((3, 2), 1)),
('rampant', ((2, 2), None)),
('family', ((26, 5), 21)),
('cbd', ((6, 5), 1)),
('instead', ((16, 3), 13)),
('unqualified', ((1, 1), None)),
('#bbcaq', ((1, 1), None)),
('criminalizing', ((1, 1), None)),
('@norml', ((2, 2), None)),
('forever!', ((1, 1), None)),
(',', 'can't', ((2, None), 2)),
('aids', ((34, 2), 32)),
('although', ((3, 1), 2)),
('always', ((17, 2), 15)),
('ppl', ((11, 5), 6)),
('underlying', ((1, None), 1)),
('beating', ((4, None), 4)),
('mentally', ((2, None), 2)),
('ill', ((6, 1), 5)),
('caricature', ((1, None), 1)),
('bruce', ((1, None), 1)),
('literally', ((20, 2), 18)),
('leading', ((7, 3), 4)),
('path', ((2, None), 2)),
('saudi', ((2, None), 2)),
('@abby_higgins5:', ((1, None), 1)),
('government/pharmacies', ((1, None), 1)),
('https://t.co/ivompefmo0', ((1, None), 1)),
('prize', ((5, 1), 4)),
('youre', ((3, None), 3)),
('dying.', ((1, None), 1)),
('choices:', ((1, None), 1)),
('cockroach', ((1, 1), None)),
('shake', ((1, 1), None)),
('menus', ((1, 1), None)),
('pre', ((3, None), 3)),
('safer', ((1, None), 1)),
('frauds', ((1, None), 1)),
('protecting', ((1, None), 1)),
('open', ((8, 3), 5)),
('@aeidinger', ((1, 1), None)),
("1930's..", ((1, 1), None)),
('15-30', ((1, 1), None)),
('nut', ((2, None), 2)),
('shout', ((2, None), 2)),
('indirectly', ((1, None), 1)),

```

('awareness.', ((1, None), 1)),  
 ('similar', ((3, None), 3)),  
 ('lab', ((5, None), 5)),  
 ('yours".', ((1, None), 1)),  
 ('means', ((9, None), 9)),  
 ('time-travel.', ((1, None), 1)),  
 ('told', ((11, 4), 7)),  
 ('america', ((14, None), 14)),  
 ('live', ((24, 2), 22)),  
 ('fuel', ((3, 1), 2)),  
 ('plane.', ((1, None), 1)),  
 ('east', ((2, None), 2)),  
 ('social', ((6, None), 6)),  
 ('knuckleheaded', ((1, None), 1)),  
 ('everything', ((20, 5), 15)),  
 ('@ancientaliens', ((1, None), 1)),  
 ('alien', ((3, None), 3)),  
 ('amongst', ((1, None), 1)),  
 ('read:', ((2, None), 2)),  
 ('@angelicjane03', ((1, None), 1)),  
 ('@nursedave34', ((1, None), 1)),  
 ('primarily', ((1, None), 1)),  
 ('cards', ((1, None), 1)),  
 ('legal', ((1, None), 1)),  
 ('#overturnroevswade', ((1, None), 1)),  
 ('@angielovesusa', ((1, None), 1)),  
 ('rather', ((3, None), 3)),  
 ('sooner.', ((2, None), 2)),  
 ('two', ((14, 3), 11)),  
 ('solution', ((6, None), 6)),  
 ('nation', ((1, None), 1)),  
 ('@artobsessed4', ((1, None), 1)),  
 ('@foxnews', ((5, None), 5)),  
 ('does,', ((6, None), 6)),  
 ('hate', ((36, 4), 32)),  
 ('protest', ((4, None), 4)),  
 ('sexist', ((2, None), 2)),  
 ('didn't', ((13, None), 13)),  
 ('tragically,', ((1, 1), None)),  
 ('slaughtered', ((1, 1), None)),  
 ('reserve.', ((1, 1), None)),  
 ('king', ((1, 1), None)),  
 ('disgusting!', ((1, 1), None)),  
 ('rhino', ((8, 7), 1)),  
 ('penises', ((2, 2), None)),  
 ('there's', ((25, 6), 19)),  
 ('curing', ((9, 2), 7)),  
 ('antidote.', ((1, None), 1)),

('@bankdeny', ((1, None), 1)),  
 ('@geertwilderspvv', ((1, None), 1)),  
 ('progress..', ((1, None), 1)),  
 ('possible', ((6, None), 6)),  
 ('https://t.co/rcsshYu4k5', ((1, None), 1)),  
 ('cured', ((16, 6), 10)),  
 ('springboard', ((4, None), 4)),  
 ('nazism,', ((4, None), 4)),  
 ('sickness', ((8, None), 8)),  
 ('spirit', ((4, None), 4)),  
 ('@bleakfacts:', ((1, None), 1)),  
 ('results', ((4, 1), 3)),  
 ('kinda', ((4, 1), 3)),  
 ('#takemehome', ((1, None), 1)),  
 ('free', ((24, 5), 19)),  
 ('walk', ((12, None), 12)),  
 ('water', ((9, None), 9)),  
 ('worked', ((5, None), 5)),  
 ('incredible', ((10, None), 10)),  
 ('hours', ((7, None), 7)),  
 ('send', ((5, 1), 4)),  
 ('compassionate', ((1, None), 1)),  
 ('placed', ((1, None), 1)),  
 ('@mountsinainyc', ((1, 1), None)),  
 ('approach', ((1, 1), None)),  
 ('medication,', ((1, 1), None)),  
 ('smart', ((3, 1), 2)),  
 ('death.', ((4, 2), 2)),  
 ('dead...', ((1, None), 1)),  
 ('solving', ((2, None), 2)),  
 ('hook', ((2, None), 2)),  
 ('capacity,', ((1, None), 1)),  
 ('ambition', ((1, None), 1)),  
 ('auto', ((2, None), 2)),  
 ('depression?', ((3, None), 3)),  
 ('millions', ((12, None), 12)),  
 ('expense', ((1, None), 1)),  
 ('professional', ((2, None), 2)),  
 ('obama.', ((1, None), 1)),  
 ('@bulldog2026', ((1, None), 1)),  
 ('@georgehwbush', ((1, None), 1)),  
 ('evidence,', ((1, None), 1)),  
 ('genius', ((2, None), 2)),  
 ('£15000', ((1, None), 1)),  
 ('@justgiving', ((1, None), 1)),  
 ('https://t.co/...', ((1, None), 1)),  
 ('permeating', ((1, None), 1)),  
 ('points.', ((1, None), 1)),

```

('@c_3c_3', ((1, None), 1)),
('run', ((14, None), 14)),
('co.', ((1, None), 1)),
('https://t.co/vlvimg460o', ((1, None), 1)),
('@calvincardwell:', ((1, None), 1)),
('operate', ((1, None), 1)),
('ca...', ((1, None), 1)),
('@campaignkate', ((1, 1), None)),
('tumor', ((5, 2), 3)),
('bismol:', ((1, 1), None)),
('https://t.co/pngwyh689w', ((1, 1), None)),
('hour.', ((1, None), 1)),
('no,', ((5, 2), 3)),
('investigate', ((1, None), 1)),
('@chercooper2', ((1, None), 1)),
('@caitrionambalfe', ((1, None), 1)),
('@hillaryclinton', ((1, None), 1)),
('way.', ((5, 1), 4)),
('@chicagoinno:', ((1, None), 1)),
('korea,', ((2, None), 2)),
('60,000,000', ((1, None), 1)),
('developed', ((2, None), 2)),
('replace', ((2, 1), 1)),
('@cifertherhyme', ((1, None), 1)),
('bored', ((1, None), 1)),
('type', ((8, None), 8)),
('resulting', ((1, None), 1)),
('prevention', ((4, 1), 3)),
('wonderful', ((5, 2), 3)),
('change', ((11, 2), 9)),
('rigorously', ((1, 1), None)),
('due!', ((1, None), 1)),
('types...', ((1, None), 1)),
('notable', ((1, None), 1)),
('individuals.', ((1, None), 1)),
('summit', ((8, None), 8)),
('kim:', ((4, None), 4)),
('cnn,', ((1, None), 1)),
('up!!!!', ((1, None), 1)),
('sun', ((3, None), 3)),
('two-faced...', ((1, None), 1)),
('@ellimayradrep', ((1, None), 1)),
('hate!', ((2, None), 2)),
('blame', ((9, None), 9)),
('#realtalk', ((1, None), 1)),
(' ', ((1, None), 1)),
('#teamtrump', ((1, None), 1)),
('alkaline', ((4, 2), 2)),

```

```

('evil.', ((1, None), 1)),
('man.', ((3, None), 3)),
('you've', ((4, 1), 3)),
('estimating', ((1, None), 1)),
('mind', ((18, 2), 16)),
('fair.', ((1, None), 1)),
('@dannijean1640', ((1, None), 1)),
('@klingensmith_j', ((1, None), 1)),
('fatal', ((1, None), 1)),
('companies', ((9, None), 9)),
('work!', ((2, 1), 1)),
('@dearillegirl', ((1, None), 1)),
('group', ((5, None), 5)),
('share', ((21, 2), 19)),
('entire', ((6, 1), 5)),
('she's', ((10, 2), 8)),
('queensland', ((1, None), 1)),
('plot.', ((1, None), 1)),
('"other"', ((1, None), 1)),
('for!', ((2, 1), 1)),
('@devarshi_21', ((2, 2), None)),
('@mohitku89856580', ((2, 2), None)),
('@swetasamadhiya', ((2, 2), None)),
('@sp_umang', ((2, 2), None)),
('@rakesha12572734', ((2, 2), None)),
('@vera_anz', ((2, 2), None)),
('@ar_jag84', ((2, 2), None)),
('@debduttakolkata', ((2, 2), None)),
('@sbhats', ((2, 2), None)),
('@raibahadur13', ((2, 2), None)),
('@arrorasanjeev', ((2, 2), None)),
('leaves', ((8, 4), 4)),
('flowers', ((2, 2), None)),
('tree,', ((1, 1), None)),
('lose', ((12, None), 12)),
('weight,', ((2, None), 2)),
('cash', ((3, None), 3)),
('ago', ((7, 1), 6)),
('follows', ((2, None), 2)),
('out.', ((7, 2), 5)),
('@realcandaceo', ((2, None), 2)),
('whine,', ((1, None), 1)),
('diseases?trump', ((1, None), 1)),
('discriminatory!', ((1, None), 1)),
('aunts', ((1, None), 1)),
('treatment?', ((1, None), 1)),
('convinced', ((5, None), 5)),
('ready', ((4, None), 4)),

```

```

('@docbastard', ((1, None), 1)),
('cahoots', ((2, None), 2)),
('@tatamemorial', ((1, None), 1)),
('latest', ((8, 3), 5)),
('#cancon2018', ((1, None), 1)),
('@cytecare', ((1, None), 1)),
('improve', ((5, None), 5)),
('@freedomfromfear', ((1, None), 1)),
('@secpompeo', ((1, None), 1)),
('save', ((23, None), 23)),
('@drjonesaa', ((1, None), 1)),
('@carequalitycomm', ((1, None), 1)),
('@zeromilly', ((1, None), 1)),
('@compassnincare', ((1, None), 1)),
('@jasonginnelly1', ((1, None), 1)),
('@kenzeroharm', ((1, None), 1)),
('@shaunlintern', ((1, None), 1)),
('@bbcnews', ((1, None), 1)),
('honest', ((1, None), 1)),
('system', ((4, 1), 3)),
('chibok', ((1, None), 1)),
('girls', ((1, None), 1)),
('neck', ((2, None), 2)),
('akwa', ((1, None), 1)),
('ways...trump', ((1, None), 1)),
('https://t.co/b0jnv1vwhh', ((1, None), 1)),
('@drmendelsohn', ((1, None), 1)),
('@jerome_corsi', ((1, None), 1)),
('thinks', ((3, 2), 1)),
('conventional', ((2, 2), None)),
('cow', ((13, 11), 2)),
('meltdown', ((1, None), 1)),
('https://t.co/60aq9x33et', ((1, None), 1)),
('defend', ((3, None), 3)),
('cancer..', ((1, None), 1)),
('@education4libs', ((2, None), 2)),
('showing', ((4, None), 4)),
('backlash', ((1, None), 1)),
('https://t.co/3xtsyqurrx', ((1, None), 1)),
('americans', ((8, 1), 7)),
('pri#k!', ((1, None), 1)),
('@exandshadow:', ((1, None), 1)),
('crossing', ((1, None), 1)),
('riots', ((2, None), 2)),
('jaw', ((1, None), 1)),
('@fandincollingo:', ((1, None), 1)),
('@williewumpkins', ((1, None), 1)),
('shoes', ((3, None), 3)),

```

```

('hiding', ((10, None), 10)),
('gallops', ((1, None), 1)),
('fart', ((1, None), 1)),
('came', ((8, 2), 6)),
('@thepooluk', ((2, 1), 1)),
('right?', ((3, None), 3)),
('head', ((6, None), 6)),
('youtube', ((1, None), 1)),
('newspaper', ((2, None), 2)),
('@fourfour_stu:', ((1, None), 1)),
('@frankcottrell_b:', ((1, None), 1)),
('@frankie_k97:', ((1, None), 1)),
('answered', ((1, None), 1)),
('exist.', ((1, None), 1)),
('https://t.co/cjrirlguxbg', ((1, None), 1)),
('medicine', ((10, 5), 5)),
('#funnythingsdadssay', ((1, None), 1)),
('@somemyrrh', ((1, 1), None)),
('@joshua1_5', ((1, 1), None)),
('@lynnleemavakay', ((1, 1), None)),
('@melissa232220', ((1, 1), None)),
('@pileofgoop', ((1, 1), None)),
('@kidsdoc1962', ((1, 1), None)),
('@and_kell', ((1, 1), None)),
('@bluelionblog', ((1, 1), None)),
('high', ((13, 4), 9)),
('challenge', ((3, 1), 2)),
('inuo.', ((1, 1), None)),
('radiation.', ((2, 1), 1)),
('hawaii.', ((1, 1), None)),
('person', ((18, 4), 14)),
('pump', ((1, None), 1)),
('planning', ((1, None), 1)),
('@gorskon:', ((1, None), 1)),
('see!', ((1, None), 1)),
('promoting', ((2, 1), 1)),
('@pkelpe', ((1, 1), None)),
('@northtxmsg', ((1, 1), None)),
('@jackmaiello', ((1, 1), None)),
('@imfsupport', ((1, 1), None)),
('@wwwillstand', ((2, None), 2)),
('@integrumwm', ((1, None), 1)),
('everything!', ((1, None), 1)),
('act', ((4, None), 4)),
('low-level', ((1, None), 1)),
('creature', ((1, None), 1)),
('sampling', ((1, None), 1)),
('music', ((2, None), 2)),

```

```

('poorly', ((2, None), 2)),
('benefit', ((7, 2), 5)),
('you?', ((6, 1), 5)),
('taxes,', ((1, None), 1)),
('pipeline', ((1, None), 1)),
('hostages', ((1, None), 1)),
('https://t.co/oz1b4nrpux', ((1, None), 1)),
('@erwindean96', ((1, None), 1)),
('@hbronnenberg', ((1, None), 1)),
('@bronsonsrevenge', ((1, None), 1)),
('@3millfam', ((1, None), 1)),
('@patriot_m1a', ((1, None), 1)),
('@sand_manx', ((1, None), 1)),
('@loveshoulas', ((1, None), 1)),
('@chighbe', ((1, None), 1)),
('@pistolgrip6spd1', ((1, None), 1)),
('@seeingred02', ((1, None), 1)),
('@natty_bumppo45', ((1, None), 1)),
('business', ((10, None), 10)),
('join', ((16, None), 16)),
('@henkelmanrb:', ((1, None), 1)),
('suppresses', ((2, 1), 1)),
('this,', ((6, None), 6)),
('https://t.co/x8pl7ybuz7', ((1, None), 1)),
('@hubie0', ((1, None), 1)),
('@halcyondon', ((1, None), 1)),
('@stablegeniusrn', ((1, None), 1)),
('@captainair727', ((1, None), 1)),
('@5th_ape', ((1, None), 1)),
('some.', ((1, None), 1)),
('chemo!"', ((1, None), 1)),
('@wordpress...', ((1, None), 1)),
('@iamamazinglucas:', ((1, None), 1)),
('india,', ((6, 6), None)),
('therapy', ((9, 9), None)),
('cancer...many', ((5, 5), None)),
('cured..', ((5, 5), None)),
('mock', ((1, 1), None)),
('stricken', ((1, 1), None)),
('situation', ((2, 1), 1)),
('everytime', ((1, 1), None)),
('seen', ((9, 4), 5)),
('https://t.co/60cery8k8y', ((1, 1), None)),
('https://t.co/ujscephajl', ((1, 1), None)),
('they'd', ((8, None), 8)),
('pay', ((7, None), 7)),
('generate', ((2, None), 2)),
('billions', ((8, None), 8)),

```



```

('too.', ((7, 2), 5)),
('wouldn't', ((12, None), 12)),
('feed', ((3, None), 3)),
('hungry', ((2, None), 2)),
('scotland', ((1, None), 1)),
('@inzcooler', ((1, 1), None)),
('depression,', ((2, 1), 1)),
('researchers,', ((8, None), 8)),
('diseases?', ((2, 1), 1)),
('cancer.imagine', ((1, 1), None)),
('free.bravin', ((1, 1), None)),
('holding', ((3, None), 3)),
('can,', ((3, None), 3)),
('genetic', ((4, 1), 3)),
('hats', ((1, None), 1)),
('hype', ((1, None), 1)),
('crap', ((5, None), 5)),
('unknown', ((1, None), 1)),
('mean', ((19, 9), 10)),
('"just"', ((1, None), 1)),
('@jdugudichi', ((1, None), 1)),
('@daniscotchirish', ((3, None), 3)),
('@kellyannegoebb', ((1, None), 1)),
('@gregthedoughboy', ((1, None), 1)),
('@joanne48640679', ((1, None), 1)),
('@leopoldjody', ((1, None), 1)),
('@james_peluso', ((1, None), 1)),
('@pendingvictory', ((1, None), 1)),
('@sgriffin_87to98', ((1, None), 1)),
('@dcarver1018', ((1, None), 1)),
('@wendidavies', ((1, None), 1)),
('@rynlds56', ((1, None), 1)),
('fool', ((1, None), 1)),
('playground', ((1, None), 1)),
('support"', ((1, None), 1)),
('tomorrow', ((7, 1), 6)),
('putin', ((1, None), 1)),
('@jeremypoxon', ((1, None), 1)),
('murdered', ((7, None), 7)),
('kills', ((5, 1), 4)),
('virginity', ((1, None), 1)),
('caused', ((3, None), 3)),
('drinking', ((7, 4), 3)),
('sebi', ((7, 1), 6)),
('amonth', ((1, None), 1)),
('thought,', ((1, None), 1)),
('archon', ((1, None), 1)),
('their/our', ((1, None), 1)),

```

```

('played', ((3, None), 3)),
('"donald', ((2, None), 2)),
('*nukes', ((1, None), 1)),
('@lplzydeco', ((1, None), 1)),
('undercut', ((1, None), 1)),
('expertise', ((1, None), 1)),
('cancer?', ((27, 2), 25)),
('president,', ((5, None), 5)),
('@jowalker2960:', ((1, None), 1)),
('democrats', ((20, None), 20)),
('@foxandfriends', ((2, None), 2)),
('home.', ((3, None), 3)),
('ass', ((10, None), 10)),
('"the', ((8, 1), 7)),
('legacy', ((2, None), 2)),
('yes', ((7, None), 7)),
('squirm', ((1, None), 1)),
('we'll', ((2, None), 2)),
('@ctfletcherisymf', ((1, None), 1)),
('!!!!', ((2, None), 2)),
('@botagain', ((1, None), 1)),
('@jrmtm50', ((1, None), 1)),
('beautifully', ((1, None), 1)),
('helpful', ((2, 1), 1)),
('proactive', ((3, None), 3)),
('precipitous', ((1, None), 1)),
('fall', ((4, None), 4)),
('6wks.', ((1, None), 1)),
('@ksnazzylawler', ((1, None), 1)),
('mindset', ((2, None), 2)),
('draft,', ((1, None), 1)),
('especially', ((3, 1), 2)),
('@maga', ((1, None), 1)),
('state', ((8, 3), 5)),
('$ncp', ((2, None), 2)),
('members', ((2, None), 2)),
('recently,', ((1, None), 1)),
('publicity', ((1, None), 1)),
('@dave_attewell', ((1, None), 1)),
('@bjaninemorison', ((1, None), 1)),
('@laura2qt:', ((1, None), 1)),
('dinosaurs', ((5, None), 5)),
('dumb???', ((1, None), 1)),
('sixers', ((1, None), 1)),
('nice', ((4, None), 4)),
('@armandodkos', ((1, None), 1)),
('beyond', ((2, None), 2)),
('belief', ((1, None), 1)),

```

```

('@lifezette:', ((1, None), 1)),
('host', ((3, None), 3)),
('@lilsweetstina', ((1, None), 1)),
('snakes,', ((1, None), 1)),
('norris', ((1, None), 1)),
('yawns,', ((1, None), 1)),
('#emptysouls', ((1, None), 1)),
('filmed', ((1, None), 1)),
('@lp62o', ((1, None), 1)),
('easily', ((2, None), 2)),
('mexican', ((2, None), 2)),
('fence', ((2, None), 2)),
('hair', ((5, None), 5)),
('form', ((7, 2), 5)),
('@lunchboxfr', ((1, None), 1)),
('turn!', ((1, None), 1)),
('ignorant!', ((1, None), 1)),
('causes', ((2, None), 2)),
('kick', ((3, None), 3)),
('rhetoric', ((6, 3), 3)),
('forgive', ((2, None), 2)),
('@thom_not_a_bot', ((1, None), 1)),
('bro', ((1, None), 1)),
('...confusion', ((1, None), 1)),
('exercise', ((1, None), 1)),
('losses', ((1, None), 1)),
('mr.', ((4, None), 4)),
('@uberfacts', ((1, None), 1)),
('@stevejazz', ((1, None), 1)),
('@willkinginri', ((1, None), 1)),
('boring', ((1, None), 1)),
('politics', ((3, None), 3)),
('holds', ((2, None), 2)),
('regarded', ((1, None), 1)),
('vitamins', ((2, 1), 1)),
('relaxation.\xa0', ((2, 1), 1)),
('bigoted', ((1, None), 1)),
('(just', ((1, None), 1)),
('tissue)', ((1, None), 1)),
('@ellebeaublog', ((1, None), 1)),
('@facebook', ((1, None), 1)),
('encouraged', ((1, None), 1)),
('#juiceplus', ((1, None), 1)),
('removed', ((5, 1), 4)),
('behaviours', ((1, None), 1)),
('@mikeggibbs', ((1, None), 1)),
('conservatives.', ((1, None), 1)),
('eliminates', ((1, None), 1)),

```

```

('@mirrorweirdnews:', ((1, 1), None)),
('proved', ((1, None), 1)),
('2020!', ((1, None), 1)),
('@dccra', ((1, None), 1)),
('@marilyncapps', ((1, None), 1)),
('@downsouth505', ((1, None), 1)),
('@girl_daisey', ((1, None), 1)),
('@jahrtt', ((1, None), 1)),
('@callisti2024', ((1, None), 1)),
('@assemblyroomusa', ((1, None), 1)),
('@msgoddessrises', ((1, None), 1)),
('@libertyscott63', ((1, None), 1)),
('@nude4peace', ((1, None), 1)),
('2018/', ((1, None), 1)),
('pissed', ((3, None), 3)),
('did.', ((2, None), 2)),
('@msscha2', ((1, None), 1)),
('idiot', ((2, None), 2)),
('clinton.', ((1, None), 1)),
('shows.', ((1, None), 1)),
('views!!!', ((1, None), 1)),
('@mussavir', ((1, 1), None)),
('useful.', ((1, 1), None)),
('indica', ((1, 1), None)),
('dhrek', ((1, 1), None)),
('useful', ((3, 1), 2)),
('tree.', ((1, 1), None)),
('@mkstalin', ((1, None), 1)),
('@fouress', ((1, None), 1)),
('@lillymarypinto', ((1, None), 1)),
('@suryahsg', ((1, None), 1)),
('heard', ((6, None), 6)),
('published', ((6, 4), 2)),
('@mrdanecw', ((1, None), 1)),
('begin?!', ((1, None), 1)),
('algorithms', ((1, None), 1)),
('source', ((2, 2), None)),
('@nc4everblessed', ((1, None), 1)),
('@tuckercarlson', ((1, None), 1)),
('rug', ((1, None), 1)),
('niagara', ((1, None), 1)),
('falls!', ((1, None), 1)),
('remission', ((3, 1), 2)),
('diseases.', ((4, 1), 3)),
('@nintendoamerica', ((1, None), 1)),
('incorrect.', ((1, None), 1)),
('alrighty', ((1, None), 1)),
('pressure.', ((1, None), 1)),

```

```

('numpty', ((1, None), 1)),
('bureaucrats.', ((1, None), 1)),
('@northkoreatimes', ((1, 1), None)),
('#northkorea', ((2, 1), 1)),
('please!!!', ((1, 1), None)),
('aids,', ((10, 2), 8)),
('"they', ((2, None), 2)),
('weed', ((6, 6), None)),
('@telegraphnews', ((1, None), 1)),
('pabgirl', ((1, None), 1)),
('"didn\'t', ((1, None), 1)),
('scam-artists:', ((1, None), 1)),
('@pastormarkburns', ((1, None), 1)),
('pastor,', ((1, None), 1)),
('cancer...', ((2, None), 2)),
('@peacerangel', ((1, None), 1)),
('extreme', ((3, 1), 2)),
('honest.', ((3, None), 3)),
('liked', ((1, None), 1)),
('clinton's", ((1, None), 1)),
('common', ((6, None), 6)),
('cold.', ((3, None), 3)),
('city.', ((1, None), 1)),
('@ranjha001', ((1, None), 1)),
('abortion', ((3, None), 3)),
('burning', ((1, None), 1)),
('building...i', ((1, None), 1)),
('who"will', ((1, None), 1)),
('seed', ((5, 2), 3)),
('#qanon', ((4, 1), 3)),
('#noko', ((1, None), 1)),
('nobel', ((4, None), 4)),
('peace.', ((4, None), 4)),
('allcharity', ((2, None), 2)),
('@just2bizi', ((1, None), 1)),
('afraid.', ((1, None), 1)),
('https://t.co/td5dddjce0', ((1, None), 1)),
('homelessness,', ((1, None), 1)),
('change"', ((1, None), 1)),
('hospitals', ((2, 1), 1)),
('rain', ((1, None), 1)),
('credit!', ((1, None), 1)),
('@hrdministry', ((1, None), 1)),
('@akashvaniair', ((1, None), 1)),
('jio', ((2, None), 2)),
('ended', ((3, None), 3)),
('reversed', ((1, None), 1)),
('climate', ((6, None), 6)),

```

```

('ask', ((1, None), 1)),
('covenant', ((1, None), 1)),
('+2349057353987,', ((1, None), 1)),
('bollinger', ((1, None), 1)),
('cancer.and', ((1, None), 1)),
('ann', ((2, None), 2)),
('https://t.co/nozmkz5l6h', ((1, 1), None)),
('virtue', ((1, None), 1)),
('#dems', ((1, None), 1)),
('akhi', ((1, None), 1)),
('applauding', ((1, None), 1)),
('strive', ((1, None), 1)),
('#bestpresidentever', ((1, None), 1)),
('https://t.co/zywsqvp0bj', ((1, None), 1)),
('unemployment".', ((1, None), 1)),
('headline', ((3, None), 3)),
('#nksummit', ((1, None), 1)),
('universe', ((3, 1), 2)),
('daughters', ((1, None), 1)),
('neurological', ((1, None), 1)),
('sad.', ((1, None), 1)),
('tragic.', ((1, None), 1)),
('record', ((1, None), 1)),
('anymore', ((2, None), 2)),
('hdmi', ((1, None), 1)),
('@hazyhula', ((1, None), 1)),
('@_iamcannabis', ((1, None), 1)),
('@cannilive', ((1, None), 1)),
('cute', ((1, None), 1)),
('selling', ((8, 4), 4)),
('google,', ((1, None), 1)),
('researching!', ((1, None), 1)),
('guns...', ((1, None), 1)),
('@deeescalante', ((1, None), 1)),
('@alexandcohen', ((2, None), 2)),
('@blogliberally', ((2, None), 2)),
('@emergvictorious', ((2, None), 2)),
('@carolyncolson', ((2, None), 2)),
('@stopbaddocs', ((2, None), 2)),
('@thomasklinemd', ((2, None), 2)),
('@michelleziemba', ((2, None), 2)),
('@chrisellington5', ((2, None), 2)),
('space', ((3, None), 3)),
('cousins', ((1, None), 1)),
('dhs,', ((2, None), 2)),
('lying,', ((1, None), 1)),
('father', ((5, 1), 4)),
('ensure', ((4, None), 4)),

```

('outside,', ((1, None), 1)),  
 ('lenin,', ((1, None), 1)),  
 ('@realdeancain', ((1, None), 1)),  
 ('@stclairashley', ((1, None), 1)),  
 ('technology', ((4, None), 4)),  
 ('turn', ((6, None), 6)),  
 ('child', ((13, 5), 8)),  
 ('operations', ((1, None), 1)),  
 ('https://t.co/vkftlvxyuf', ((1, None), 1)),  
 ('@saylesjon', ((1, None), 1)),  
 ('says', ((12, 2), 10)),  
 ('(nay', ((1, None), 1)),  
 ('...', ((1, None), 1)),  
 ('guitars', ((1, None), 1)),  
 ('', ((1, None), 1)),  
 ('https://t.co/je3gryeiw7', ((1, None), 1)),  
 ('he'd', ((3, None), 3)),  
 ('soldiers', ((1, None), 1)),  
 ('drs', ((1, None), 1)),  
 ('demokrats', ((1, None), 1)),  
 ('foreign', ((1, None), 1)),  
 ('schools', ((1, None), 1)),  
 ('race', ((17, 1), 16)),  
 ('hustler', ((2, None), 2)),  
 ('pro-choice', ((1, None), 1)),  
 ('born.', ((1, None), 1)),  
 ('racecar', ((2, None), 2)),  
 ('email.', ((3, None), 3)),  
 ('https://t.co/cosx4qbck3', ((1, None), 1)),  
 ('@rileyjameswarr', ((1, None), 1)),  
 ('scores', ((1, None), 1)),  
 ('one)', ((1, None), 1)),  
 ('touch', ((1, None), 1)),  
 ('politicians...', ((1, None), 1)),  
 ('dummy', ((1, None), 1)),  
 ('america,', ((1, None), 1)),  
 ('fricking', ((1, None), 1)),  
 ('crying', ((6, 1), 5)),  
 ('supposedly', ((4, 2), 2)),  
 ('rinos', ((1, None), 1)),  
 ('https://t.co/8bmt73v7hr', ((1, None), 1)),  
 ('fail,', ((1, None), 1)),  
 ('payments,', ((1, None), 1)),  
 ('allow', ((4, None), 4)),  
 ('plays', ((2, None), 2)),  
 ('vision', ((7, 1), 6)),  
 ('libs,', ((1, None), 1)),  
 ('trumpers:', ((1, None), 1)),

```

('@2b7c89526', ((1, None), 1)),
('@cmccitizenmedia', ((1, None), 1)),
('donate.', ((2, None), 2)),
('@elerhinomarch', ((1, 1), None)),
('@louisal1000', ((1, 1), None)),
('change,', ((1, None), 1)),
('perhaps', ((3, None), 3)),
('okkkk', ((1, None), 1)),
('@supremesaxx:', ((1, None), 1)),
('genuis', ((1, None), 1)),
('jerk', ((2, None), 2)),
('@srxrabit:', ((1, None), 1)),
('@grxxvec', ((1, None), 1)),
('recognize', ((1, None), 1)),
('@teapainusa', ((1, None), 1)),
('population!"', ((1, None), 1)),
('.no', ((2, 1), 1)),
('#sendsarahskiddies2disney', ((1, None), 1)),
('@thehill', ((4, None), 4)),
('disappear,', ((1, None), 1)),
('determined', ((1, None), 1)),
('details.', ((1, None), 1)),
('qualities,', ((1, None), 1)),
('starting', ((2, None), 2)),
('reverse', ((1, None), 1)),
('#impeach', ((1, None), 1)),
('jackasses', ((1, None), 1)),
('happening', ((2, None), 2)),
('@deji14_official', ((1, None), 1)),
('@manesti336', ((1, None), 1)),
('nightmare', ((1, None), 1)),
('brough', ((1, None), 1)),
('terminally', ((1, None), 1)),
('https://t.co/r4eo6zjya1', ((1, None), 1)),
('opinions?', ((1, None), 1)),
('@elcontador2000', ((1, None), 1)),
('@willfranken', ((1, None), 1)),
('@poshfruit', ((1, None), 1)),
('spelling', ((1, None), 1)),
('already.', ((1, None), 1)),
('wasting', ((3, None), 3)),
('senate,', ((1, None), 1)),
('elites', ((1, None), 1)),
('industry,', ((1, None), 1)),
('fbi', ((1, None), 1)),
('$3.89', ((1, None), 1)),
('higher', ((4, 1), 3)),
('indie-band', ((1, None), 1)),

```



```

('walk.', ((1, None), 1)),
('timeline', ((1, None), 1)),
('jetpacks', ((1, None), 1)),
('pea,', ((1, 1), None)),
('gop', ((1, 1), None)),
('@webbgolf:', ((1, None), 1)),
('peanut', ((1, None), 1)),
('jelly', ((1, None), 1)),
('usc', ((1, 1), None)),
('ontario,', ((1, 1), None)),
('disgraceful.', ((1, None), 1)),
('decided', ((3, None), 3)),
('@vinayprasadmd', ((1, None), 1)),
('palliation', ((1, None), 1)),
('@vinesauce', ((1, None), 1)),
('charity', ((4, None), 4)),
('pleased', ((2, None), 2)),
('announce', ((3, None), 3)),
('sunday,', ((3, None), 3)),
('"lol"', ((1, None), 1)),
('highschool,', ((1, None), 1)),
('blaming', ((2, None), 2)),
('rising', ((1, None), 1)),
('(weird) ', ((1, None), 1)),
('tweet!', ((1, None), 1)),
('canada.', ((1, None), 1)),
('stance...', ((1, 1), None)),
('from...', ((1, 1), None)),
('meme...', ((1, 1), None)),
('lib?', ((1, 1), None)),
('discussing', ((1, 1), None)),
('internet', ((2, 1), 1)),
('"cancer', ((1, 1), None)),
('experts"', ((1, 1), None)),
('start', ((13, 5), 8)),
('exaggerating', ((1, None), 1)),
('shelves.', ((1, None), 1)),
('high,', ((1, None), 1)),
('shkreli', ((1, None), 1)),
('bit,', ((1, None), 1)),
('hydrogen', ((1, None), 1)),
('prevents', ((1, None), 1)),
('contributions,', ((1, None), 1)),
('actively', ((3, None), 3)),
('#scotuspick', ((1, None), 1)),
('score', ((8, None), 8)),
('certain', ((1, 1), None)),
('offers', ((1, 1), None)),

```

```
(('capable', ((1, 1), None)),
('esophageal', ((1, 1), None)),
('advocate', ((1, None), 1)),
('daydream', ((1, None), 1)),
('fantasies', ((2, 1), 1)),
('loved', ((6, 1), 5)),
('note,', ((1, None), 1)),
('advancement', ((1, None), 1)),
('issues', ((2, 1), 1)),
('situations', ((1, 1), None)),
('existence:)', ((1, 1), None)),
('https://t.co/ymefygzjgx', ((1, 1), None)),
('factory', ((1, 1), None)),
('poisoning', ((2, 2), None)),
('fooled', ((1, 1), None)),
('rape', ((1, 1), None)),
('times.', ((2, None), 2)),
('https://t.co/ya9tolltgr', ((1, None), 1)),
('ahh', ((1, None), 1)),
('organize', ((1, None), 1)),
('#hflpride', ((1, None), 1)),
('afterlife', ((1, None), 1)),
('theres', ((1, 1), None)),
('prone', ((3, None), 3)),
('mars', ((1, None), 1)),
('https://t.co/bo5rvk0worm', ((1, None), 1)),
('signing', ((1, None), 1)),
('store', ((3, 2), 1)),
('raided', ((1, 1), None)),
('apricot', ((6, 6), None)),
('threatened', ((1, 1), None)),
('ladies', ((3, None), 3)),
('measles,', ((1, None), 1)),
('importance.', ((1, None), 1)),
('ian', ((1, 1), None)),
('wright', ((1, 1), None)),
('world!', ((1, 1), None)),
('gratitude', ((1, None), 1)),
...]
```

```
[ ]: # Convert data to the desired format
word_table = word_table.map(lambda x: (x[0], x[1][0][0], x[1][1], x[1][0][1])).
    ↪map(lambda x: tuple(0 if val is None else val for val in x))
```

```
[ ]: word_table.collect()
```

```
Out[31]: [('ever', 18, 16, 2),
('scientific', 10, 7, 3),
```

('medical', 27, 16, 11),  
 ('yoga', 3, 1, 2),  
 ('cancer,', 183, 150, 33),  
 ('cardiac', 2, 0, 2),  
 ('#cureforcancer', 1, 1, 0),  
 ('100', 3, 3, 0),  
 ('2014', 2, 2, 0),  
 ('know', 60, 46, 14),  
 ('needed', 7, 5, 2),  
 ('cancer.', 162, 137, 25),  
 ('bitch', 12, 12, 0),  
 ('even', 74, 63, 11),  
 ('become', 11, 10, 1),  
 ('diverse', 1, 1, 0),  
 ('gay,', 1, 1, 0),  
 ('away', 19, 16, 3),  
 ('new', 25, 20, 5),  
 ('run,', 1, 1, 0),  
 ('saying', 19, 17, 2),  
 ('kind', 12, 8, 4),  
 ('&', 123, 95, 28),  
 ('#reiki', 2, 2, 0),  
 ('clinton', 2, 2, 0),  
 ('#trumptrainjumponit', 1, 1, 0),  
 ('clinton!', 1, 1, 0),  
 ('suppressing', 1, 1, 0),  
 ('cancer-cures', 1, 1, 0),  
 ('wants', 7, 7, 0),  
 ('#righttodie', 1, 1, 0),  
 ('ask', 15, 15, 0),  
 ('-both', 1, 1, 0),  
 ('f\*cked', 1, 1, 0),  
 ('https://t.co/xpikezppcz', 1, 1, 0),  
 ('#petitions', 1, 1, 0),  
 ('terminal', 7, 5, 2),  
 ('diseases', 25, 16, 9),  
 ('them! ', 1, 1, 0),  
 ('needs', 19, 17, 2),  
 ('like', 155, 140, 15),  
 ('@danatylertv', 2, 2, 0),  
 ('took', 12, 12, 0),  
 ('disparities', 1, 1, 0),  
 ('diagnosis', 6, 5, 1),  
 ('african', 4, 2, 2),  
 ('teaching', 3, 3, 0),  
 ('#teamhalon', 1, 0, 1),  
 ('https://t.co/q2lt7y56km', 1, 0, 1),  
 ('negative', 10, 10, 0),

('are.', 5, 5, 0),  
 ('injection', 1, 1, 0),  
 ('stinging.#liberalhypocrisy', 1, 1, 0),  
 ('#deepstate', 1, 1, 0),  
 ('metastasized', 2, 1, 1),  
 ('full', 16, 15, 1),  
 ('outbreak', 1, 1, 0),  
 ('nevertrumpers', 1, 1, 0),  
 ('demise.', 2, 1, 1),  
 ('let', 22, 22, 0),  
 ('#inpursuitof', 3, 0, 3),  
 ('1h2019.', 1, 0, 1),  
 ('agen2034', 1, 0, 1),  
 ('#asco2018.', 1, 0, 1),  
 ('https://t.co/vrwjomjmvo', 1, 0, 1),  
 ('right', 26, 18, 8),  
 ('combination', 3, 1, 2),  
 ('people-see', 3, 1, 2),  
 ('toward', 4, 2, 2),  
 ('@big3biobos', 3, 1, 2),  
 ('sure', 17, 15, 2),  
 ('deeds', 1, 0, 1),  
 ('impacted', 1, 0, 1),  
 ('lesnar', 1, 1, 0),  
 ('looking', 15, 10, 5),  
 ('pushed', 1, 1, 0),  
 ('finds', 7, 7, 0),  
 ('shit!', 1, 1, 0),  
 ('probably', 15, 15, 0),  
 ('years,', 6, 5, 1),  
 ('weeks,', 1, 1, 0),  
 ('scientists', 24, 18, 6),  
 ('volcano', 1, 1, 0),  
 ('chimes\*', 1, 0, 1),  
 ('can't', 32, 29, 3),  
 ('narrative', 3, 2, 1),  
 ('#lifewithcancer', 1, 0, 1),  
 ('.@abbvie', 2, 2, 0),  
 ('.@erictrump:', 1, 1, 0),  
 ('we're', 11, 11, 0),  
 ('win', 16, 16, 0),  
 ('https://t.co/7pyfcd7xzy', 1, 1, 0),  
 ('son', 7, 5, 2),  
 ('https://t.co/gwqtxflqau', 1, 1, 0),  
 ('excellent', 4, 4, 0),  
 ('job', 14, 14, 0),  
 ('gun', 7, 7, 0),  
 ('bye', 5, 5, 0),

('love', 28, 26, 2),  
 ('reunion', 3, 3, 0),  
 ('used', 14, 11, 3),  
 ('make', 92, 84, 8),  
 ('many.', 4, 4, 0),  
 ('maybe', 25, 21, 4),  
 ('pretty', 5, 5, 0),  
 ('religion', 11, 11, 0),  
 ('else's', 1, 1, 0),  
 ('body...', 1, 1, 0),  
 ('stage', 23, 18, 5),  
 ('survival', 12, 11, 1),  
 ('book', 13, 13, 0),  
 ('living', 17, 14, 3),  
 ('misdiagnosed', 7, 7, 0),  
 ('flying', 8, 8, 0),  
 ('future"', 1, 1, 0),  
 ('1987:', 1, 1, 0),  
 ('potentially', 5, 5, 0),  
 ('twenty', 2, 2, 0),  
 ('rider', 1, 1, 0),  
 ('semi-finalist', 1, 1, 0),  
 ('master', 1, 1, 0),  
 ('thing?', 2, 2, 0),  
 ('asked.', 1, 1, 0),  
 ('https://t.co/vjazjwo5em', 1, 1, 0),  
 ('22,', 1, 1, 0),  
 ('diagnosed', 10, 8, 2),  
 ('typically', 1, 1, 0),  
 ('5,000', 2, 2, 0),  
 ('women', 17, 16, 1),  
 ('tournament', 2, 2, 0),  
 ('more:', 2, 2, 0),  
 ('£63k', 1, 1, 0),  
 ('years.', 10, 8, 2),  
 ('finishing', 1, 1, 0),  
 ('trump.', 9, 9, 0),  
 ('80%', 2, 2, 0),  
 ('#disgusted', 1, 1, 0),  
 ('must', 21, 20, 1),  
 ('grasp', 1, 1, 0),  
 ('nature.', 2, 2, 0),  
 ('@realjameswoods', 4, 4, 0),  
 ('you're', 12, 10, 2),  
 ('denuclearization', 3, 3, 0),  
 ('peace...', 1, 1, 0),  
 ('complain', 32, 32, 0),  
 ('orange', 2, 2, 0),

('studied', 1, 0, 1),  
 ('stop', 35, 25, 10),  
 ('such', 3, 1, 2),  
 ('rampant', 2, 0, 2),  
 ('family', 26, 21, 5),  
 ('cbd', 6, 1, 5),  
 ('instead', 16, 13, 3),  
 ('unqualified', 1, 0, 1),  
 ('#bbcaq', 1, 0, 1),  
 ('criminalizing', 1, 0, 1),  
 ('@norml', 2, 0, 2),  
 ('forever!', 1, 0, 1),  
 ('can't', 2, 2, 0),  
 ('aids', 34, 32, 2),  
 ('although', 3, 2, 1),  
 ('always', 17, 15, 2),  
 ('ppl', 11, 6, 5),  
 ('underlying', 1, 1, 0),  
 ('beating', 4, 4, 0),  
 ('mentally', 2, 2, 0),  
 ('ill', 6, 5, 1),  
 ('caricature', 1, 1, 0),  
 ('bruce', 1, 1, 0),  
 ('literally', 20, 18, 2),  
 ('leading', 7, 4, 3),  
 ('path', 2, 2, 0),  
 ('saudi', 2, 2, 0),  
 ('@abby\_higgins5:', 1, 1, 0),  
 ('government/pharmacies', 1, 1, 0),  
 ('https://t.co/ivompefmo0', 1, 1, 0),  
 ('prize', 5, 4, 1),  
 ('youre', 3, 3, 0),  
 ('dying.', 1, 1, 0),  
 ('choices:', 1, 1, 0),  
 ('cockroach', 1, 0, 1),  
 ('shake', 1, 0, 1),  
 ('menus', 1, 0, 1),  
 ('pre', 3, 3, 0),  
 ('safer', 1, 1, 0),  
 ('frauds', 1, 1, 0),  
 ('protecting', 1, 1, 0),  
 ('open', 8, 5, 3),  
 ('@aeidinger', 1, 0, 1),  
 ('1930's..', 1, 0, 1),  
 ('15-30', 1, 0, 1),  
 ('nut', 2, 2, 0),  
 ('shout', 2, 2, 0),  
 ('indirectly', 1, 1, 0),

('awareness.', 1, 1, 0),  
 ('similar', 3, 3, 0),  
 ('lab', 5, 5, 0),  
 ('yours".', 1, 1, 0),  
 ('means', 9, 9, 0),  
 ('time-travel.', 1, 1, 0),  
 ('told', 11, 7, 4),  
 ('america', 14, 14, 0),  
 ('live', 24, 22, 2),  
 ('fuel', 3, 2, 1),  
 ('plane.', 1, 1, 0),  
 ('east', 2, 2, 0),  
 ('social', 6, 6, 0),  
 ('knuckleheaded', 1, 1, 0),  
 ('everything', 20, 15, 5),  
 ('@ancientaliens', 1, 1, 0),  
 ('alien', 3, 3, 0),  
 ('amongst', 1, 1, 0),  
 ('read:', 2, 2, 0),  
 ('@angelicjane03', 1, 1, 0),  
 ('@nursedave34', 1, 1, 0),  
 ('primarily', 1, 1, 0),  
 ('cards', 1, 1, 0),  
 ('legal', 1, 1, 0),  
 ('#overturnroevswade', 1, 1, 0),  
 ('@angielovesusa', 1, 1, 0),  
 ('rather', 3, 3, 0),  
 ('sooner.', 2, 2, 0),  
 ('two', 14, 11, 3),  
 ('solution', 6, 6, 0),  
 ('nation', 1, 1, 0),  
 ('@artobsessed4', 1, 1, 0),  
 ('@foxnews', 5, 5, 0),  
 ('does,', 6, 6, 0),  
 ('hate', 36, 32, 4),  
 ('protest', 4, 4, 0),  
 ('sexist', 2, 2, 0),  
 ('didn't', 13, 13, 0),  
 ('tragically,', 1, 0, 1),  
 ('slaughtered', 1, 0, 1),  
 ('reserve.', 1, 0, 1),  
 ('king', 1, 0, 1),  
 ('disgusting!', 1, 0, 1),  
 ('rhino', 8, 1, 7),  
 ('penises', 2, 0, 2),  
 ('there's', 25, 19, 6),  
 ('curing', 9, 7, 2),  
 ('antidote.', 1, 1, 0),

('@bankdeny', 1, 1, 0),  
 ('@geertwilderspvv', 1, 1, 0),  
 ('progress..', 1, 1, 0),  
 ('possible', 6, 6, 0),  
 ('https://t.co/rcsshYu4k5', 1, 1, 0),  
 ('cured', 16, 10, 6),  
 ('springboard', 4, 4, 0),  
 ('nazism,', 4, 4, 0),  
 ('sickness', 8, 8, 0),  
 ('spirit', 4, 4, 0),  
 ('@bleakfacts:', 1, 1, 0),  
 ('results', 4, 3, 1),  
 ('kinda', 4, 3, 1),  
 ('#takemehome', 1, 1, 0),  
 ('free', 24, 19, 5),  
 ('walk', 12, 12, 0),  
 ('water', 9, 9, 0),  
 ('worked', 5, 5, 0),  
 ('incredible', 10, 10, 0),  
 ('hours', 7, 7, 0),  
 ('send', 5, 4, 1),  
 ('compassionate', 1, 1, 0),  
 ('placed', 1, 1, 0),  
 ('@mountsinainyc', 1, 0, 1),  
 ('approach', 1, 0, 1),  
 ('medication,', 1, 0, 1),  
 ('smart', 3, 2, 1),  
 ('death.', 4, 2, 2),  
 ('dead...', 1, 1, 0),  
 ('solving', 2, 2, 0),  
 ('hook', 2, 2, 0),  
 ('capacity,', 1, 1, 0),  
 ('ambition', 1, 1, 0),  
 ('auto', 2, 2, 0),  
 ('depression?', 3, 3, 0),  
 ('millions', 12, 12, 0),  
 ('expense', 1, 1, 0),  
 ('professional', 2, 2, 0),  
 ('obama.', 1, 1, 0),  
 ('@bulldog2026', 1, 1, 0),  
 ('@georgehwbush', 1, 1, 0),  
 ('evidence,', 1, 1, 0),  
 ('genius', 2, 2, 0),  
 ('£15000', 1, 1, 0),  
 ('@justgiving', 1, 1, 0),  
 ('https://t.co/...', 1, 1, 0),  
 ('permeating', 1, 1, 0),  
 ('points.', 1, 1, 0),



('@c\_3c\_3', 1, 1, 0),  
 ('run', 14, 14, 0),  
 ('co.', 1, 1, 0),  
 ('https://t.co/vlvlmg460o', 1, 1, 0),  
 ('@calvincardwell:', 1, 1, 0),  
 ('operate', 1, 1, 0),  
 ('ca...', 1, 1, 0),  
 ('@campaignkate', 1, 0, 1),  
 ('tumor', 5, 3, 2),  
 ('bismol:', 1, 0, 1),  
 ('https://t.co/pngwyh689w', 1, 0, 1),  
 ('hour.', 1, 1, 0),  
 ('no,', 5, 3, 2),  
 ('investigate', 1, 1, 0),  
 ('@chercooper2', 1, 1, 0),  
 ('@caitrionambalfe', 1, 1, 0),  
 ('@hillaryclinton', 1, 1, 0),  
 ('way.', 5, 4, 1),  
 ('@chicagoinno:', 1, 1, 0),  
 ('korea,', 2, 2, 0),  
 ('60,000,000', 1, 1, 0),  
 ('developed', 2, 2, 0),  
 ('replace', 2, 1, 1),  
 ('@cifertherhyme', 1, 1, 0),  
 ('bored', 1, 1, 0),  
 ('type', 8, 8, 0),  
 ('resulting', 1, 1, 0),  
 ('prevention', 4, 3, 1),  
 ('wonderful', 5, 3, 2),  
 ('change', 11, 9, 2),  
 ('rigorously', 1, 0, 1),  
 ('due!', 1, 1, 0),  
 ('types...', 1, 1, 0),  
 ('notable', 1, 1, 0),  
 ('individuals.', 1, 1, 0),  
 ('summit', 8, 8, 0),  
 ('kim:', 4, 4, 0),  
 ('cnn,', 1, 1, 0),  
 ('up!!!!', 1, 1, 0),  
 ('sun', 3, 3, 0),  
 ('two-faced...', 1, 1, 0),  
 ('@ellimayradrep', 1, 1, 0),  
 ('hate!', 2, 2, 0),  
 ('blame', 9, 9, 0),  
 ('#realtalk', 1, 1, 0),  
 ('', 1, 1, 0),  
 ('#teamtrump', 1, 1, 0),  
 ('alkaline', 4, 2, 2),

```

('evil.', 1, 1, 0),
('man.', 3, 3, 0),
('you've', 4, 3, 1),
('estimating', 1, 1, 0),
('mind', 18, 16, 2),
('fair.', 1, 1, 0),
('@dannijean1640', 1, 1, 0),
('@klingensmith_j', 1, 1, 0),
('fatal', 1, 1, 0),
('companies', 9, 9, 0),
('work!', 2, 1, 1),
('@dearillegirl', 1, 1, 0),
('group', 5, 5, 0),
('share', 21, 19, 2),
('entire', 6, 5, 1),
('she's', 10, 8, 2),
('queensland', 1, 1, 0),
('plot.', 1, 1, 0),
('"other"', 1, 1, 0),
('for!', 2, 1, 1),
('@devarshi_21', 2, 0, 2),
('@mohitku89856580', 2, 0, 2),
('@swetasamadhiya', 2, 0, 2),
('@sp_umang', 2, 0, 2),
('@rakesha12572734', 2, 0, 2),
('@vera_anz', 2, 0, 2),
('@ar_jag84', 2, 0, 2),
('@debduttakolkata', 2, 0, 2),
('@sbhats', 2, 0, 2),
('@raibahadur13', 2, 0, 2),
('@arrorasanjeev', 2, 0, 2),
('leaves', 8, 4, 4),
('flowers', 2, 0, 2),
('tree,', 1, 0, 1),
('lose', 12, 12, 0),
('weight,', 2, 2, 0),
('cash', 3, 3, 0),
('ago', 7, 6, 1),
('follows', 2, 2, 0),
('out.', 7, 5, 2),
('@realcandaceo', 2, 2, 0),
('whine,', 1, 1, 0),
('diseases?trump', 1, 1, 0),
('discriminatory!"', 1, 1, 0),
('aunts', 1, 1, 0),
('treatment?', 1, 1, 0),
('convinced', 5, 5, 0),
('ready', 4, 4, 0),

```

('@docbastard', 1, 1, 0),  
 ('cahoots', 2, 2, 0),  
 ('@tatamemorial', 1, 1, 0),  
 ('latest', 8, 5, 3),  
 ('#cancon2018', 1, 1, 0),  
 ('@cytecare', 1, 1, 0),  
 ('improve', 5, 5, 0),  
 ('@freedomfromfear', 1, 1, 0),  
 ('@secpompeo', 1, 1, 0),  
 ('save', 23, 23, 0),  
 ('@drjonesaa', 1, 1, 0),  
 ('@carequalitycomm', 1, 1, 0),  
 ('@zeromilly', 1, 1, 0),  
 ('@compassnincare', 1, 1, 0),  
 ('@jasonginnelly1', 1, 1, 0),  
 ('@kenzeroharm', 1, 1, 0),  
 ('@shaunlintern', 1, 1, 0),  
 ('@bbcnews', 1, 1, 0),  
 ('honest', 1, 1, 0),  
 ('system', 4, 3, 1),  
 ('chibok', 1, 1, 0),  
 ('girls', 1, 1, 0),  
 ('neck', 2, 2, 0),  
 ('akwa', 1, 1, 0),  
 ('ways...trump', 1, 1, 0),  
 ('https://t.co/b0jnv1vwhh', 1, 1, 0),  
 ('@drmendelsohn', 1, 1, 0),  
 ('@jerome\_corsi', 1, 1, 0),  
 ('thinks', 3, 1, 2),  
 ('conventional', 2, 0, 2),  
 ('cow', 13, 2, 11),  
 ('meltdown', 1, 1, 0),  
 ('https://t.co/60aq9x33et', 1, 1, 0),  
 ('defend', 3, 3, 0),  
 ('cancer..', 1, 1, 0),  
 ('@education4libs', 2, 2, 0),  
 ('showing', 4, 4, 0),  
 ('backlash', 1, 1, 0),  
 ('https://t.co/3xtsyqurrx', 1, 1, 0),  
 ('americans', 8, 7, 1),  
 ('pri#k!', 1, 1, 0),  
 ('@exandshadow:', 1, 1, 0),  
 ('crossing', 1, 1, 0),  
 ('riots', 2, 2, 0),  
 ('jaw', 1, 1, 0),  
 ('@fandincollingo:', 1, 1, 0),  
 ('@williewumpkins', 1, 1, 0),  
 ('shoes', 3, 3, 0),

```

('hiding', 10, 10, 0),
('gallops', 1, 1, 0),
('fart', 1, 1, 0),
('came', 8, 6, 2),
('@thepooluk', 2, 1, 1),
('right?', 3, 3, 0),
('head', 6, 6, 0),
('youtube', 1, 1, 0),
('newspaper', 2, 2, 0),
('@fourfour_stu:', 1, 1, 0),
('@frankcottrell_b:', 1, 1, 0),
('@frankie_k97:', 1, 1, 0),
('answered', 1, 1, 0),
('exist.', 1, 1, 0),
('https://t.co/cjrrlguxbg', 1, 1, 0),
('medicine', 10, 5, 5),
('#funnythingsdadssay', 1, 1, 0),
('@somemyrrh', 1, 0, 1),
('@joshua1_5', 1, 0, 1),
('@lynnleemavakay', 1, 0, 1),
('@melissa232220', 1, 0, 1),
('@pileofgoop', 1, 0, 1),
('@kidsdoc1962', 1, 0, 1),
('@and_kell', 1, 0, 1),
('@bluelionblog', 1, 0, 1),
('high', 13, 9, 4),
('challenge', 3, 2, 1),
('inuo.', 1, 0, 1),
('radiation.', 2, 1, 1),
('hawaii.', 1, 0, 1),
('person', 18, 14, 4),
('pump', 1, 1, 0),
('planning', 1, 1, 0),
('@gorskon:', 1, 1, 0),
('see!', 1, 1, 0),
('promoting', 2, 1, 1),
('@pkelpe', 1, 0, 1),
('@northtxmsg', 1, 0, 1),
('@jackmaiello', 1, 0, 1),
('@imfsupport', 1, 0, 1),
('@wwwwillstand', 2, 2, 0),
('@integrumwm', 1, 1, 0),
('everything!', 1, 1, 0),
('act', 4, 4, 0),
('low-level', 1, 1, 0),
('creature', 1, 1, 0),
('sampling', 1, 1, 0),
('music', 2, 2, 0),

```

('poorly', 2, 2, 0),  
 ('benefit', 7, 5, 2),  
 ('you?', 6, 5, 1),  
 ('taxes,', 1, 1, 0),  
 ('pipeline', 1, 1, 0),  
 ('hostages', 1, 1, 0),  
 ('https://t.co/oz1b4nrpux', 1, 1, 0),  
 ('@erwindean96', 1, 1, 0),  
 ('@hbronnenberg', 1, 1, 0),  
 ('@bronsonsrevenge', 1, 1, 0),  
 ('@3millfam', 1, 1, 0),  
 ('@patriot\_m1a', 1, 1, 0),  
 ('@sand\_manx', 1, 1, 0),  
 ('@loveshoulas', 1, 1, 0),  
 ('@chighbe', 1, 1, 0),  
 ('@pistolgrip6spd1', 1, 1, 0),  
 ('@seeingred02', 1, 1, 0),  
 ('@natty\_bumppo45', 1, 1, 0),  
 ('business', 10, 10, 0),  
 ('join', 16, 16, 0),  
 ('@henkelmanrb:', 1, 1, 0),  
 ('suppresses', 2, 1, 1),  
 ('this,', 6, 6, 0),  
 ('https://t.co/x8pl7ybuz7', 1, 1, 0),  
 ('@hubie0', 1, 1, 0),  
 ('@halcyondon', 1, 1, 0),  
 ('@stablegeniusrn', 1, 1, 0),  
 ('@captainair727', 1, 1, 0),  
 ('@5th\_ape', 1, 1, 0),  
 ('some.', 1, 1, 0),  
 ('chemo!"', 1, 1, 0),  
 ('@wordpress...', 1, 1, 0),  
 ('@iamamazinglucas:', 1, 1, 0),  
 ('india,', 6, 0, 6),  
 ('therapy', 9, 0, 9),  
 ('cancer...many', 5, 0, 5),  
 ('cured..', 5, 0, 5),  
 ('mock', 1, 0, 1),  
 ('stricken', 1, 0, 1),  
 ('situation', 2, 1, 1),  
 ('everytime', 1, 0, 1),  
 ('seen', 9, 5, 4),  
 ('https://t.co/60cery8k8y', 1, 0, 1),  
 ('https://t.co/ujscephajl', 1, 0, 1),  
 ('they'd', 8, 8, 0),  
 ('pay', 7, 7, 0),  
 ('generate', 2, 2, 0),  
 ('billions', 8, 8, 0),

('too.', 7, 5, 2),  
 ('wouldn't', 12, 12, 0),  
 ('feed', 3, 3, 0),  
 ('hungry', 2, 2, 0),  
 ('scotland', 1, 1, 0),  
 ('@inzcooler', 1, 0, 1),  
 ('depression,', 2, 1, 1),  
 ('researchers,', 8, 8, 0),  
 ('diseases?', 2, 1, 1),  
 ('cancer.imagine', 1, 0, 1),  
 ('free.bravin', 1, 0, 1),  
 ('holding', 3, 3, 0),  
 ('can,', 3, 3, 0),  
 ('genetic', 4, 3, 1),  
 ('hats', 1, 1, 0),  
 ('hype', 1, 1, 0),  
 ('crap', 5, 5, 0),  
 ('unknown', 1, 1, 0),  
 ('mean', 19, 10, 9),  
 ('"just"', 1, 1, 0),  
 ('@jdugudichi', 1, 1, 0),  
 ('@daniscotchirish', 3, 3, 0),  
 ('@kellyannegoebb', 1, 1, 0),  
 ('@gregthedoughboy', 1, 1, 0),  
 ('@joanne48640679', 1, 1, 0),  
 ('@leopoldjody', 1, 1, 0),  
 ('@james\_peluso', 1, 1, 0),  
 ('@pendingvictory', 1, 1, 0),  
 ('@sgriffin\_87to98', 1, 1, 0),  
 ('@dcarver1018', 1, 1, 0),  
 ('@wendidavies', 1, 1, 0),  
 ('@rynlds56', 1, 1, 0),  
 ('fool', 1, 1, 0),  
 ('playground', 1, 1, 0),  
 ('support"', 1, 1, 0),  
 ('tomorrow', 7, 6, 1),  
 ('putin', 1, 1, 0),  
 ('@jeremypoxon', 1, 1, 0),  
 ('murdered', 7, 7, 0),  
 ('kills', 5, 4, 1),  
 ('virginity', 1, 1, 0),  
 ('caused', 3, 3, 0),  
 ('drinking', 7, 3, 4),  
 ('sebi', 7, 6, 1),  
 ('amonth', 1, 1, 0),  
 ('thought,', 1, 1, 0),  
 ('archon', 1, 1, 0),  
 ('their/our', 1, 1, 0),

('played', 3, 3, 0),  
 ('donald', 2, 2, 0),  
 ('\*nukes', 1, 1, 0),  
 ('@lplzydeco', 1, 1, 0),  
 ('undercut', 1, 1, 0),  
 ('expertise', 1, 1, 0),  
 ('cancer?', 27, 25, 2),  
 ('president,', 5, 5, 0),  
 ('@jowalker2960:', 1, 1, 0),  
 ('democrats', 20, 20, 0),  
 ('@foxandfriends', 2, 2, 0),  
 ('home.', 3, 3, 0),  
 ('ass', 10, 10, 0),  
 ('the', 8, 7, 1),  
 ('legacy', 2, 2, 0),  
 ('yes', 7, 7, 0),  
 ('squirm', 1, 1, 0),  
 ('we'll', 2, 2, 0),  
 ('@ctfletcherisymf', 1, 1, 0),  
 ('!!!', 2, 2, 0),  
 ('@botagain', 1, 1, 0),  
 ('@jrmtm50', 1, 1, 0),  
 ('beautifully', 1, 1, 0),  
 ('helpful', 2, 1, 1),  
 ('proactive', 3, 3, 0),  
 ('precipitous', 1, 1, 0),  
 ('fall', 4, 4, 0),  
 ('6wks.', 1, 1, 0),  
 ('@ksnazzylawler', 1, 1, 0),  
 ('mindset', 2, 2, 0),  
 ('draft,', 1, 1, 0),  
 ('especially', 3, 2, 1),  
 ('@maga', 1, 1, 0),  
 ('state', 8, 5, 3),  
 ('\$ncp', 2, 2, 0),  
 ('members', 2, 2, 0),  
 ('recently,', 1, 1, 0),  
 ('publicity', 1, 1, 0),  
 ('@dave\_attewell', 1, 1, 0),  
 ('@bjaninemorison', 1, 1, 0),  
 ('@laura2qt:', 1, 1, 0),  
 ('dinosaurs', 5, 5, 0),  
 ('dumb???', 1, 1, 0),  
 ('sixers', 1, 1, 0),  
 ('nice', 4, 4, 0),  
 ('@armandodkos', 1, 1, 0),  
 ('beyond', 2, 2, 0),  
 ('belief', 1, 1, 0),

('@lifezette:', 1, 1, 0),  
 ('host', 3, 3, 0),  
 ('@lilsweetstina', 1, 1, 0),  
 ('snakes,', 1, 1, 0),  
 ('norris', 1, 1, 0),  
 ('yawns,', 1, 1, 0),  
 ('#emptysouls', 1, 1, 0),  
 ('filmed', 1, 1, 0),  
 ('@lp62o', 1, 1, 0),  
 ('easily', 2, 2, 0),  
 ('mexican', 2, 2, 0),  
 ('fence', 2, 2, 0),  
 ('hair', 5, 5, 0),  
 ('form', 7, 5, 2),  
 ('@lunchboxfr', 1, 1, 0),  
 ('turn!', 1, 1, 0),  
 ('ignorant!', 1, 1, 0),  
 ('causes', 2, 2, 0),  
 ('kick', 3, 3, 0),  
 ('rhetoric', 6, 3, 3),  
 ('forgive', 2, 2, 0),  
 ('@thom\_not\_a\_bot', 1, 1, 0),  
 ('bro', 1, 1, 0),  
 ('...confusion', 1, 1, 0),  
 ('exercise', 1, 1, 0),  
 ('losses', 1, 1, 0),  
 ('mr.', 4, 4, 0),  
 ('@uberfacts', 1, 1, 0),  
 ('@stevejazz', 1, 1, 0),  
 ('@willkinginri', 1, 1, 0),  
 ('boring', 1, 1, 0),  
 ('politics', 3, 3, 0),  
 ('holds', 2, 2, 0),  
 ('regarded', 1, 1, 0),  
 ('vitamins', 2, 1, 1),  
 ('relaxation.\xa0', 2, 1, 1),  
 ('bigoted', 1, 1, 0),  
 ('(just', 1, 1, 0),  
 ('tissue)', 1, 1, 0),  
 ('@ellebeaublog', 1, 1, 0),  
 ('@facebook', 1, 1, 0),  
 ('encouraged', 1, 1, 0),  
 ('#juiceplus', 1, 1, 0),  
 ('removed', 5, 4, 1),  
 ('behaviours', 1, 1, 0),  
 ('@mikeggibbs', 1, 1, 0),  
 ('conservatives.', 1, 1, 0),  
 ('eliminates', 1, 1, 0),



('@mirrorweirdnews:', 1, 0, 1),  
 ('proved', 1, 1, 0),  
 ('2020!', 1, 1, 0),  
 ('@dccra', 1, 1, 0),  
 ('@marilyncapps', 1, 1, 0),  
 ('@downsouth505', 1, 1, 0),  
 ('@girl\_daisey', 1, 1, 0),  
 ('@jahrtt', 1, 1, 0),  
 ('@callisti2024', 1, 1, 0),  
 ('@assemblyroomusa', 1, 1, 0),  
 ('@msgoddessrises', 1, 1, 0),  
 ('@libertyscott63', 1, 1, 0),  
 ('@nude4peace', 1, 1, 0),  
 ('2018/', 1, 1, 0),  
 ('pissed', 3, 3, 0),  
 ('did.', 2, 2, 0),  
 ('@msscha2', 1, 1, 0),  
 ('idiot', 2, 2, 0),  
 ('clinton.', 1, 1, 0),  
 ('shows.', 1, 1, 0),  
 ('views!!!', 1, 1, 0),  
 ('@mussavir', 1, 0, 1),  
 ('useful.', 1, 0, 1),  
 ('indica', 1, 0, 1),  
 ('dhrek', 1, 0, 1),  
 ('useful', 3, 2, 1),  
 ('tree.', 1, 0, 1),  
 ('@mkstalin', 1, 1, 0),  
 ('@fouress', 1, 1, 0),  
 ('@lillymarypinto', 1, 1, 0),  
 ('@suryahsg', 1, 1, 0),  
 ('heard', 6, 6, 0),  
 ('published', 6, 2, 4),  
 ('@mrdanecw', 1, 1, 0),  
 ('begin?!', 1, 1, 0),  
 ('algorithms', 1, 1, 0),  
 ('source', 2, 0, 2),  
 ('@nc4everblessed', 1, 1, 0),  
 ('@tuckercarlson', 1, 1, 0),  
 ('rug', 1, 1, 0),  
 ('niagara', 1, 1, 0),  
 ('falls!', 1, 1, 0),  
 ('remission', 3, 2, 1),  
 ('diseases.', 4, 3, 1),  
 ('@nintendoamerica', 1, 1, 0),  
 ('incorrect.', 1, 1, 0),  
 ('alrighty', 1, 1, 0),  
 ('pressure.', 1, 1, 0),

```

('numpty', 1, 1, 0),
('bureaucrats.', 1, 1, 0),
('@northkoreatimes', 1, 0, 1),
('#northkorea', 2, 1, 1),
('please!!!', 1, 0, 1),
('aids,', 10, 8, 2),
('"they', 2, 2, 0),
('weed', 6, 0, 6),
('@telegraphnews', 1, 1, 0),
('pabgirl', 1, 1, 0),
('"didn\'t', 1, 1, 0),
('scam-artists:', 1, 1, 0),
('@pastormarkburns', 1, 1, 0),
('pastor,', 1, 1, 0),
('cancer...', 2, 2, 0),
('@peacerangel', 1, 1, 0),
('extreme', 3, 2, 1),
('honest.', 3, 3, 0),
('liked', 1, 1, 0),
('clinton\'s"', 1, 1, 0),
('common', 6, 6, 0),
('cold.', 3, 3, 0),
('city.', 1, 1, 0),
('@ranjha001', 1, 1, 0),
('abortion', 3, 3, 0),
('burning', 1, 1, 0),
('building...i', 1, 1, 0),
('who"will', 1, 1, 0),
('seed', 5, 3, 2),
('#qanon', 4, 3, 1),
('#noko', 1, 1, 0),
('nobel', 4, 4, 0),
('peace.', 4, 4, 0),
('allcharity', 2, 2, 0),
('@just2bizi', 1, 1, 0),
('afraid.', 1, 1, 0),
('https://t.co/td5dddjce0', 1, 1, 0),
('homelessness,', 1, 1, 0),
('change"', 1, 1, 0),
('hospitals', 2, 1, 1),
('rain', 1, 1, 0),
('credit!', 1, 1, 0),
('@hrdministry', 1, 1, 0),
('@akashvaniair', 1, 1, 0),
('jio', 2, 2, 0),
('ended', 3, 3, 0),
('reversed', 1, 1, 0),
('climate', 6, 6, 0),

```

```

('ask', 1, 1, 0),
('covenant', 1, 1, 0),
('+2349057353987,', 1, 1, 0),
('bollinger', 1, 1, 0),
('cancer.and', 1, 1, 0),
('ann', 2, 2, 0),
('https://t.co/nozmkz5l6h', 1, 0, 1),
('virtue', 1, 1, 0),
('#dems', 1, 1, 0),
('akhi', 1, 1, 0),
('applauding', 1, 1, 0),
('strive', 1, 1, 0),
('#bestpresidentever', 1, 1, 0),
('https://t.co/zywsqvp0bj', 1, 1, 0),
('unemployment".', 1, 1, 0),
('headline', 3, 3, 0),
('#nksummit', 1, 1, 0),
('universe', 3, 2, 1),
('daughters', 1, 1, 0),
('neurological', 1, 1, 0),
('sad.', 1, 1, 0),
('tragic.', 1, 1, 0),
('record', 1, 1, 0),
('anymore', 2, 2, 0),
('hdmi', 1, 1, 0),
('@hazyhula', 1, 1, 0),
('@_iamcannabis', 1, 1, 0),
('@cannilive', 1, 1, 0),
('cute', 1, 1, 0),
('selling', 8, 4, 4),
('google,', 1, 1, 0),
('researching!', 1, 1, 0),
('guns...', 1, 1, 0),
('@deeescalante', 1, 1, 0),
('@alexandcohen', 2, 2, 0),
('@blogliberally', 2, 2, 0),
('@emergvictorious', 2, 2, 0),
('@carolyncolson', 2, 2, 0),
('@stopbaddocs', 2, 2, 0),
('@thomasklinemd', 2, 2, 0),
('@michelleziemba', 2, 2, 0),
('@chrisellington5', 2, 2, 0),
('space', 3, 3, 0),
('cousins', 1, 1, 0),
('dhs,', 2, 2, 0),
('lying,', 1, 1, 0),
('father', 5, 4, 1),
('ensure', 4, 4, 0),

```

('outside,', 1, 1, 0),  
 ('lenin,', 1, 1, 0),  
 ('@realdeancain', 1, 1, 0),  
 ('@stclairashley', 1, 1, 0),  
 ('technology', 4, 4, 0),  
 ('turn', 6, 6, 0),  
 ('child', 13, 8, 5),  
 ('operations', 1, 1, 0),  
 ('https://t.co/vkftlvxyuf', 1, 1, 0),  
 ('@saylesjon', 1, 1, 0),  
 ('says', 12, 10, 2),  
 ('(nay', 1, 1, 0),  
 ('...', 1, 1, 0),  
 ('guitars', 1, 1, 0),  
 ('', 1, 1, 0),  
 ('https://t.co/je3gryeiw7', 1, 1, 0),  
 ('he'd', 3, 3, 0),  
 ('soldiers', 1, 1, 0),  
 ('drs', 1, 1, 0),  
 ('demokrats', 1, 1, 0),  
 ('foreign', 1, 1, 0),  
 ('schools', 1, 1, 0),  
 ('race', 17, 16, 1),  
 ('hustler', 2, 2, 0),  
 ('pro-choice', 1, 1, 0),  
 ('born.', 1, 1, 0),  
 ('racecar', 2, 2, 0),  
 ('email.', 3, 3, 0),  
 ('https://t.co/cosx4qbck3', 1, 1, 0),  
 ('@rileyjameswarr', 1, 1, 0),  
 ('scores', 1, 1, 0),  
 ('one)', 1, 1, 0),  
 ('touch', 1, 1, 0),  
 ('politicians...', 1, 1, 0),  
 ('dummy', 1, 1, 0),  
 ('america,', 1, 1, 0),  
 ('fricking', 1, 1, 0),  
 ('crying', 6, 5, 1),  
 ('supposedly', 4, 2, 2),  
 ('rinos', 1, 1, 0),  
 ('https://t.co/8bmt73v7hr', 1, 1, 0),  
 ('fail,', 1, 1, 0),  
 ('payments,', 1, 1, 0),  
 ('allow', 4, 4, 0),  
 ('plays', 2, 2, 0),  
 ('vision', 7, 6, 1),  
 ('libs,', 1, 1, 0),  
 ('trumpers:', 1, 1, 0),

('@2b7c89526', 1, 1, 0),  
 ('@cmcitizenmedia', 1, 1, 0),  
 ('donate.', 2, 2, 0),  
 ('@elerhinomarch', 1, 0, 1),  
 ('@louisal1000', 1, 0, 1),  
 ('change,', 1, 1, 0),  
 ('perhaps', 3, 3, 0),  
 ('okkkk', 1, 1, 0),  
 ('@supremesaxx:', 1, 1, 0),  
 ('genuis', 1, 1, 0),  
 ('jerk', 2, 2, 0),  
 ('@srxrabit:', 1, 1, 0),  
 ('@grxxvec', 1, 1, 0),  
 ('recognize', 1, 1, 0),  
 ('@teapainusa', 1, 1, 0),  
 ('population!"', 1, 1, 0),  
 ('.no', 2, 1, 1),  
 ('#sendsarahskiddies2disney', 1, 1, 0),  
 ('@thehill', 4, 4, 0),  
 ('disappear,', 1, 1, 0),  
 ('determined', 1, 1, 0),  
 ('details.', 1, 1, 0),  
 ('qualities,', 1, 1, 0),  
 ('starting', 2, 2, 0),  
 ('reverse', 1, 1, 0),  
 ('#impeach', 1, 1, 0),  
 ('jackasses', 1, 1, 0),  
 ('happening', 2, 2, 0),  
 ('@deji14\_official', 1, 1, 0),  
 ('@manesti336', 1, 1, 0),  
 ('nightmare', 1, 1, 0),  
 ('brough', 1, 1, 0),  
 ('terminally', 1, 1, 0),  
 ('https://t.co/r4eo6zjya1', 1, 1, 0),  
 ('opinions?', 1, 1, 0),  
 ('@elcontador2000', 1, 1, 0),  
 ('@willfranken', 1, 1, 0),  
 ('@poshfruit', 1, 1, 0),  
 ('spelling', 1, 1, 0),  
 ('already.', 1, 1, 0),  
 ('wasting', 3, 3, 0),  
 ('senate,', 1, 1, 0),  
 ('elites', 1, 1, 0),  
 ('industry,', 1, 1, 0),  
 ('fbi', 1, 1, 0),  
 ('\$3.89', 1, 1, 0),  
 ('higher', 4, 3, 1),  
 ('indie-band', 1, 1, 0),

('walk.', 1, 1, 0),  
 ('timeline', 1, 1, 0),  
 ('jetpacks', 1, 1, 0),  
 ('pea,', 1, 0, 1),  
 ('gop', 1, 0, 1),  
 ('@webbgolf:', 1, 1, 0),  
 ('peanut', 1, 1, 0),  
 ('jelly', 1, 1, 0),  
 ('usc', 1, 0, 1),  
 ('ontario,', 1, 0, 1),  
 ('disgraceful.', 1, 1, 0),  
 ('decided', 3, 3, 0),  
 ('@vinayprasadmd', 1, 1, 0),  
 ('palliation', 1, 1, 0),  
 ('@vinesauce', 1, 1, 0),  
 ('charity', 4, 4, 0),  
 ('pleased', 2, 2, 0),  
 ('announce', 3, 3, 0),  
 ('sunday,', 3, 3, 0),  
 ('"lol"', 1, 1, 0),  
 ('highschool,', 1, 1, 0),  
 ('blaming', 2, 2, 0),  
 ('rising', 1, 1, 0),  
 ('(weird) ', 1, 1, 0),  
 ('tweet!', 1, 1, 0),  
 ('canada.', 1, 1, 0),  
 ('stance...', 1, 0, 1),  
 ('from...', 1, 0, 1),  
 ('meme...', 1, 0, 1),  
 ('lib?', 1, 0, 1),  
 ('discussing', 1, 0, 1),  
 ('internet', 2, 1, 1),  
 ('"cancer', 1, 0, 1),  
 ('experts"', 1, 0, 1),  
 ('start', 13, 8, 5),  
 ('exaggerating', 1, 1, 0),  
 ('shelves.', 1, 1, 0),  
 ('high,', 1, 1, 0),  
 ('shkreli', 1, 1, 0),  
 ('bit,', 1, 1, 0),  
 ('hydrogen', 1, 1, 0),  
 ('prevents', 1, 1, 0),  
 ('contributions,', 1, 1, 0),  
 ('actively', 3, 3, 0),  
 ('#scotuspick', 1, 1, 0),  
 ('score', 8, 8, 0),  
 ('certain', 1, 0, 1),  
 ('offers', 1, 0, 1),

```

('capable', 1, 0, 1),
('esophageal', 1, 0, 1),
('advocate', 1, 1, 0),
('daydream', 1, 1, 0),
('fantasies', 2, 1, 1),
('loved', 6, 5, 1),
('note,', 1, 1, 0),
('advancement', 1, 1, 0),
('issues', 2, 1, 1),
('situations', 1, 0, 1),
('existence:)', 1, 0, 1),
('https://t.co/ymefygzjgx', 1, 0, 1),
('factory', 1, 0, 1),
('poisoning', 2, 0, 2),
('fooled', 1, 0, 1),
('rape', 1, 0, 1),
('times.', 2, 2, 0),
('https://t.co/ya9tolltgr', 1, 1, 0),
('ahh', 1, 1, 0),
('organize', 1, 1, 0),
('#hflpride', 1, 1, 0),
('afterlife', 1, 1, 0),
('theres', 1, 0, 1),
('prone', 3, 3, 0),
('mars', 1, 1, 0),
('https://t.co/bo5rvk0worm', 1, 1, 0),
('signing', 1, 1, 0),
('store', 3, 1, 2),
('raided', 1, 0, 1),
('apricot', 6, 0, 6),
('threatened', 1, 0, 1),
('ladies', 3, 3, 0),
('measles,', 1, 1, 0),
('importance.', 1, 1, 0),
('ian', 1, 0, 1),
('wright', 1, 0, 1),
('world!', 1, 0, 1),
('gratitude', 1, 1, 0),
...]
```

```

[ ]: # Calculate probabilities for each word
# alpha = 0.1 # Laplace smoothing constant
word_probabilities = word_table.map(lambda x: (x[0],x[2]/x[1],x[3]/x[1]))

# Small constant for Laplace smoothing
# alpha = 0.1
```

```
word_probabilities = word_table.map(lambda x: (x[0], (x[2] + alpha) / (x[1] +
↪alpha), (x[3] + alpha) / (x[1] + alpha)))
```

```
[ ]: word_probabilities.collect()
```

```
Out[33]: [('ever', 0.8895027624309393, 0.11602209944751381),
('scientific', 0.7029702970297029, 0.3069306930693069),
('medical', 0.5940959409594097, 0.40959409594095936),
('yoga', 0.3548387096774194, 0.6774193548387097),
('cancer,', 0.8197706171490988, 0.1807755324959039),
('cardiac', 0.047619047619047616, 1.0),
('#cureforcancer', 1.0, 0.09090909090909091),
('100', 1.0, 0.03225806451612903),
('2014', 1.0, 0.047619047619047616),
('know', 0.7670549084858569, 0.2346089850249584),
('needed', 0.7183098591549295, 0.295774647887324),
('cancer.', 0.8457742134484886, 0.15484268969771747),
('bitch', 1.0, 0.008264462809917356),
('even', 0.8515519568151148, 0.14979757085020243),
('become', 0.9099099099099099, 0.09909909909909911),
('diverse', 1.0, 0.09090909090909091),
('gay,', 1.0, 0.09090909090909091),
('away', 0.8429319371727749, 0.16230366492146597),
('new', 0.800796812749004, 0.20318725099601592),
('run,', 1.0, 0.09090909090909091),
('saying', 0.8952879581151832, 0.10994764397905758),
('kind', 0.6694214876033058, 0.33884297520661155),
('&', 0.7725426482534524, 0.22826969943135664),
('#reiki', 1.0, 0.047619047619047616),
('clinton', 1.0, 0.047619047619047616),
('#trumptrainjumponit', 1.0, 0.09090909090909091),
('clinton!', 1.0, 0.09090909090909091),
('suppressing', 1.0, 0.09090909090909091),
('cancer-cures', 1.0, 0.09090909090909091),
('wants', 1.0, 0.014084507042253523),
('#righttodie', 1.0, 0.09090909090909091),
('ask', 1.0, 0.006622516556291392),
('-both', 1.0, 0.09090909090909091),
('f*cked', 1.0, 0.09090909090909091),
('https://t.co/xpikezppcz', 1.0, 0.09090909090909091),
('#petitions', 1.0, 0.09090909090909091),
('terminal', 0.7183098591549295, 0.295774647887324),
('diseases', 0.6414342629482072, 0.3625498007968127),
('them! ', 1.0, 0.09090909090909091),
('needs', 0.8952879581151832, 0.10994764397905758),
('like', 0.9032882011605415, 0.0973565441650548),
('@danatylertv', 1.0, 0.047619047619047616),
('took', 1.0, 0.008264462809917356),
```



('disparities', 1.0, 0.09090909090909091),  
 ('diagnosis', 0.8360655737704918, 0.18032786885245905),  
 ('african', 0.5121951219512195, 0.5121951219512195),  
 ('teaching', 1.0, 0.03225806451612903),  
 ('#teamhalon', 0.09090909090909091, 1.0),  
 ('https://t.co/q2lt7y56km', 0.09090909090909091, 1.0),  
 ('negative', 1.0, 0.009900990099009901),  
 ('are.', 1.0, 0.019607843137254905),  
 ('injection', 1.0, 0.09090909090909091),  
 ('stinging.#liberalhypocrisy', 1.0, 0.09090909090909091),  
 ('#deepstate', 1.0, 0.09090909090909091),  
 ('metastasized', 0.5238095238095238, 0.5238095238095238),  
 ('full', 0.9378881987577639, 0.06832298136645963),  
 ('outbreak', 1.0, 0.09090909090909091),  
 ('nevertrumpers', 1.0, 0.09090909090909091),  
 ('demise.', 0.5238095238095238, 0.5238095238095238),  
 ('let', 1.0, 0.004524886877828054),  
 ('#inpursuitof', 0.03225806451612903, 1.0),  
 ('1h2019.', 0.09090909090909091, 1.0),  
 ('agen2034', 0.09090909090909091, 1.0),  
 ('#asco2018.', 0.09090909090909091, 1.0),  
 ('https://t.co/vrwjomjmvo', 0.09090909090909091, 1.0),  
 ('right', 0.6934865900383141, 0.31034482758620685),  
 ('combination', 0.3548387096774194, 0.6774193548387097),  
 ('people-see', 0.3548387096774194, 0.6774193548387097),  
 ('toward', 0.5121951219512195, 0.5121951219512195),  
 ('@big3biobos', 0.3548387096774194, 0.6774193548387097),  
 ('sure', 0.8830409356725145, 0.12280701754385964),  
 ('deeds', 0.09090909090909091, 1.0),  
 ('impacted', 0.09090909090909091, 1.0),  
 ('lesnar', 1.0, 0.09090909090909091),  
 ('looking', 0.6688741721854304, 0.33774834437086093),  
 ('pushed', 1.0, 0.09090909090909091),  
 ('finds', 1.0, 0.014084507042253523),  
 ('shit!', 1.0, 0.09090909090909091),  
 ('probably', 1.0, 0.006622516556291392),  
 ('years,', 0.8360655737704918, 0.18032786885245905),  
 ('weeks,', 1.0, 0.09090909090909091),  
 ('scientists', 0.7510373443983402, 0.25311203319502074),  
 ('volcano', 1.0, 0.09090909090909091),  
 ('chimes\*', 0.09090909090909091, 1.0),  
 ('can't', 0.9065420560747663, 0.09657320872274143),  
 ('narrative', 0.6774193548387097, 0.3548387096774194),  
 ('#lifewithcancer', 0.09090909090909091, 1.0),  
 ('.@abbvie', 1.0, 0.047619047619047616),  
 ('.@erictrump', 1.0, 0.09090909090909091),  
 ('we're', 1.0, 0.009009009009009009),  
 ('win', 1.0, 0.006211180124223602),

('https://t.co/7pyfcd7xzy', 1.0, 0.09090909090909091),  
('son', 0.7183098591549295, 0.295774647887324),  
('https://t.co/gwqtxflqau', 1.0, 0.09090909090909091),  
('excellent', 1.0, 0.02439024390243903),  
('job', 1.0, 0.007092198581560284),  
('gun', 1.0, 0.014084507042253523),  
('bye', 1.0, 0.019607843137254905),  
('love', 0.9288256227758007, 0.07473309608540925),  
('reunion', 1.0, 0.03225806451612903),  
('used', 0.7872340425531915, 0.2198581560283688),  
('make', 0.9131378935939196, 0.08794788273615635),  
('many.', 1.0, 0.02439024390243903),  
('maybe', 0.8406374501992032, 0.1633466135458167),  
('pretty', 1.0, 0.019607843137254905),  
('religion', 1.0, 0.009009009009009009),  
('else's', 1.0, 0.09090909090909091),  
('body...', 1.0, 0.09090909090909091),  
('stage', 0.7835497835497836, 0.22077922077922074),  
('survival', 0.9173553719008264, 0.09090909090909093),  
('book', 1.0, 0.0076335877862595426),  
('living', 0.8245614035087718, 0.18128654970760233),  
('misdiagnosed', 1.0, 0.014084507042253523),  
('flying', 1.0, 0.01234567901234568),  
('future"', 1.0, 0.09090909090909091),  
('1987:', 1.0, 0.09090909090909091),  
('potentially', 1.0, 0.019607843137254905),  
('twenty', 1.0, 0.047619047619047616),  
('rider', 1.0, 0.09090909090909091),  
('semi-finalist', 1.0, 0.09090909090909091),  
('master', 1.0, 0.09090909090909091),  
('thing?', 1.0, 0.047619047619047616),  
('asked.', 1.0, 0.09090909090909091),  
('https://t.co/vjazjwo5em', 1.0, 0.09090909090909091),  
('22,', 1.0, 0.09090909090909091),  
('diagnosed', 0.801980198019802, 0.20792079207920794),  
('typically', 1.0, 0.09090909090909091),  
('5,000', 1.0, 0.047619047619047616),  
('women', 0.9415204678362573, 0.06432748538011696),  
('tournament', 1.0, 0.047619047619047616),  
('more:', 1.0, 0.047619047619047616),  
('£63k', 1.0, 0.09090909090909091),  
('years.', 0.801980198019802, 0.20792079207920794),  
('finishing', 1.0, 0.09090909090909091),  
('trump.', 1.0, 0.01098901098901099),  
('80%', 1.0, 0.047619047619047616),  
('#disgusted', 1.0, 0.09090909090909091),  
('must', 0.95260663507109, 0.05213270142180095),  
('grasp', 1.0, 0.09090909090909091),

('nature.', 1.0, 0.047619047619047616),  
 ('@realjameswoods', 1.0, 0.02439024390243903),  
 ('you're', 0.8347107438016529, 0.17355371900826447),  
 ('denuclearization', 1.0, 0.03225806451612903),  
 ('peace...', 1.0, 0.09090909090909091),  
 ('complain', 1.0, 0.003115264797507788),  
 ('orange', 1.0, 0.047619047619047616),  
 ('studied,', 0.09090909090909091, 1.0),  
 ('stop', 0.7150997150997151, 0.2877492877492877),  
 ('such,', 0.3548387096774194, 0.6774193548387097),  
 ('rampant', 0.047619047619047616, 1.0),  
 ('family', 0.8084291187739464, 0.1954022988505747),  
 ('cbd', 0.18032786885245905, 0.8360655737704918),  
 ('instead', 0.8136645962732918, 0.19254658385093165),  
 ('unqualified', 0.09090909090909091, 1.0),  
 ('#bbcaq', 0.09090909090909091, 1.0),  
 ('criminalizing', 0.09090909090909091, 1.0),  
 ('@norml', 0.047619047619047616, 1.0),  
 ('forever!', 0.09090909090909091, 1.0),  
 ('can't', 1.0, 0.047619047619047616),  
 ('aids', 0.9413489736070382, 0.06158357771260997),  
 ('although', 0.6774193548387097, 0.3548387096774194),  
 ('always', 0.8830409356725145, 0.12280701754385964),  
 ('ppl', 0.5495495495495495, 0.45945945945945943),  
 ('underlying', 1.0, 0.09090909090909091),  
 ('beating', 1.0, 0.02439024390243903),  
 ('mentally', 1.0, 0.047619047619047616),  
 ('ill', 0.8360655737704918, 0.18032786885245905),  
 ('caricature', 1.0, 0.09090909090909091),  
 ('bruce', 1.0, 0.09090909090909091),  
 ('literally', 0.900497512437811, 0.1044776119402985),  
 ('leading', 0.5774647887323944, 0.4366197183098592),  
 ('path', 1.0, 0.047619047619047616),  
 ('saudi', 1.0, 0.047619047619047616),  
 ('@abby\_higgins5:', 1.0, 0.09090909090909091),  
 ('government/pharmacies', 1.0, 0.09090909090909091),  
 ('https://t.co/ivompefmo0', 1.0, 0.09090909090909091),  
 ('prize', 0.803921568627451, 0.21568627450980396),  
 ('youre', 1.0, 0.03225806451612903),  
 ('dying.', 1.0, 0.09090909090909091),  
 ('choices:', 1.0, 0.09090909090909091),  
 ('cockroach', 0.09090909090909091, 1.0),  
 ('shake', 0.09090909090909091, 1.0),  
 ('menus', 0.09090909090909091, 1.0),  
 ('pre', 1.0, 0.03225806451612903),  
 ('safer', 1.0, 0.09090909090909091),  
 ('frauds', 1.0, 0.09090909090909091),  
 ('protecting', 1.0, 0.09090909090909091),

('open', 0.6296296296296297, 0.3827160493827161),  
 ('@aeidinger', 0.09090909090909091, 1.0),  
 ("1930's..", 0.09090909090909091, 1.0),  
 ('15-30', 0.09090909090909091, 1.0),  
 ('nut', 1.0, 0.047619047619047616),  
 ('shout', 1.0, 0.047619047619047616),  
 ('indirectly', 1.0, 0.09090909090909091),  
 ('awareness.', 1.0, 0.09090909090909091),  
 ('similar', 1.0, 0.03225806451612903),  
 ('lab', 1.0, 0.019607843137254905),  
 ('yours".', 1.0, 0.09090909090909091),  
 ('means', 1.0, 0.01098901098901099),  
 ('time-travel.', 1.0, 0.09090909090909091),  
 ('told', 0.6396396396396397, 0.36936936936936937),  
 ('america', 1.0, 0.007092198581560284),  
 ('live', 0.91701244813278, 0.08713692946058091),  
 ('fuel', 0.6774193548387097, 0.3548387096774194),  
 ('plane.', 1.0, 0.09090909090909091),  
 ('east', 1.0, 0.047619047619047616),  
 ('social', 1.0, 0.01639344262295082),  
 ('knuckleheaded', 1.0, 0.09090909090909091),  
 ('everything', 0.7512437810945273, 0.25373134328358204),  
 ('@ancientaliens', 1.0, 0.09090909090909091),  
 ('alien', 1.0, 0.03225806451612903),  
 ('amongst', 1.0, 0.09090909090909091),  
 ('read:', 1.0, 0.047619047619047616),  
 ('@angelicjane03', 1.0, 0.09090909090909091),  
 ('@nursedave34', 1.0, 0.09090909090909091),  
 ('primarily', 1.0, 0.09090909090909091),  
 ('cards', 1.0, 0.09090909090909091),  
 ('legal', 1.0, 0.09090909090909091),  
 ('#overtturnroevswade', 1.0, 0.09090909090909091),  
 ('@angielovesusa', 1.0, 0.09090909090909091),  
 ('rather', 1.0, 0.03225806451612903),  
 ('sooner.', 1.0, 0.047619047619047616),  
 ('two', 0.7872340425531915, 0.2198581560283688),  
 ('solution', 1.0, 0.01639344262295082),  
 ('nation', 1.0, 0.09090909090909091),  
 ('@artobsessed4', 1.0, 0.09090909090909091),  
 ('@foxnews', 1.0, 0.019607843137254905),  
 ('does,', 1.0, 0.01639344262295082),  
 ('hate', 0.8891966759002771, 0.11357340720221605),  
 ('protest', 1.0, 0.02439024390243903),  
 ('sexist', 1.0, 0.047619047619047616),  
 ('didn't', 1.0, 0.0076335877862595426),  
 ('tragically,', 0.09090909090909091, 1.0),  
 ('slaughtered', 0.09090909090909091, 1.0),  
 ('reserve.', 0.09090909090909091, 1.0),

('king', 0.09090909090909091, 1.0),  
('disgusting!', 0.09090909090909091, 1.0),  
('rhino', 0.1358024691358025, 0.8765432098765432),  
('penises', 0.047619047619047616, 1.0),  
('there's', 0.7609561752988048, 0.24302788844621512),  
('curing', 0.7802197802197802, 0.23076923076923078),  
('antidote.', 1.0, 0.09090909090909091),  
('@bankdeny', 1.0, 0.09090909090909091),  
('@geertwilderspvv', 1.0, 0.09090909090909091),  
('progress..', 1.0, 0.09090909090909091),  
('possible', 1.0, 0.01639344262295082),  
('https://t.co/rcsshyu4k5', 1.0, 0.09090909090909091),  
('cured', 0.6273291925465838, 0.3788819875776397),  
('springboard', 1.0, 0.02439024390243903),  
('nazism,', 1.0, 0.02439024390243903),  
('sickness', 1.0, 0.01234567901234568),  
('spirit', 1.0, 0.02439024390243903),  
('@bleakfacts:', 1.0, 0.09090909090909091),  
('results', 0.7560975609756099, 0.26829268292682934),  
('kinda', 0.7560975609756099, 0.26829268292682934),  
('#takemehome', 1.0, 0.09090909090909091),  
('free', 0.7925311203319502, 0.21161825726141076),  
('walk', 1.0, 0.008264462809917356),  
('water', 1.0, 0.01098901098901099),  
('worked', 1.0, 0.019607843137254905),  
('incredible', 1.0, 0.009900990099009901),  
('hours', 1.0, 0.014084507042253523),  
('send', 0.803921568627451, 0.21568627450980396),  
('compassionate', 1.0, 0.09090909090909091),  
('placed', 1.0, 0.09090909090909091),  
('@mountsinainyc', 0.09090909090909091, 1.0),  
('approach', 0.09090909090909091, 1.0),  
('medication,', 0.09090909090909091, 1.0),  
('smart', 0.6774193548387097, 0.3548387096774194),  
('death.', 0.5121951219512195, 0.5121951219512195),  
('dead...', 1.0, 0.09090909090909091),  
('solving', 1.0, 0.047619047619047616),  
('hook', 1.0, 0.047619047619047616),  
('capacity,', 1.0, 0.09090909090909091),  
('ambition', 1.0, 0.09090909090909091),  
('auto', 1.0, 0.047619047619047616),  
('depression?', 1.0, 0.03225806451612903),  
('millions', 1.0, 0.008264462809917356),  
('expense', 1.0, 0.09090909090909091),  
('professional', 1.0, 0.047619047619047616),  
('obama.', 1.0, 0.09090909090909091),  
('@bulldog2026', 1.0, 0.09090909090909091),  
('@georgehwobush', 1.0, 0.09090909090909091),

('evidence', 1.0, 0.09090909090909091),  
('genius', 1.0, 0.047619047619047616),  
('£15000', 1.0, 0.09090909090909091),  
('@justgiving', 1.0, 0.09090909090909091),  
('https://t.co/...', 1.0, 0.09090909090909091),  
('permeating', 1.0, 0.09090909090909091),  
('points.', 1.0, 0.09090909090909091),  
('@c\_3c\_3', 1.0, 0.09090909090909091),  
('run', 1.0, 0.007092198581560284),  
('co.', 1.0, 0.09090909090909091),  
('https://t.co/vlvlmg460o', 1.0, 0.09090909090909091),  
('@calvincardwell:', 1.0, 0.09090909090909091),  
('operate', 1.0, 0.09090909090909091),  
('ca...', 1.0, 0.09090909090909091),  
('@campaignkate', 0.09090909090909091, 1.0),  
('tumor', 0.607843137254902, 0.411764705882353),  
('bismol:', 0.09090909090909091, 1.0),  
('https://t.co/pngwyh689w', 0.09090909090909091, 1.0),  
('hour.', 1.0, 0.09090909090909091),  
('no,', 0.607843137254902, 0.411764705882353),  
('investigate', 1.0, 0.09090909090909091),  
('@chercooper2', 1.0, 0.09090909090909091),  
('@caitrionambalfe', 1.0, 0.09090909090909091),  
('@hillaryclinton', 1.0, 0.09090909090909091),  
('way.', 0.803921568627451, 0.21568627450980396),  
('@chicagoinno:', 1.0, 0.09090909090909091),  
('korea,', 1.0, 0.047619047619047616),  
('60,000,000', 1.0, 0.09090909090909091),  
('developed', 1.0, 0.047619047619047616),  
('replace', 0.5238095238095238, 0.5238095238095238),  
('@cifertherhyme', 1.0, 0.09090909090909091),  
('bored', 1.0, 0.09090909090909091),  
('type', 1.0, 0.01234567901234568),  
('resulting', 1.0, 0.09090909090909091),  
('prevention', 0.7560975609756099, 0.26829268292682934),  
('wonderful', 0.607843137254902, 0.411764705882353),  
('change', 0.8198198198198198, 0.1891891891891892),  
('rigorously', 0.09090909090909091, 1.0),  
('due!', 1.0, 0.09090909090909091),  
('types...', 1.0, 0.09090909090909091),  
('notable', 1.0, 0.09090909090909091),  
('individuals.', 1.0, 0.09090909090909091),  
('summit', 1.0, 0.01234567901234568),  
('kim:', 1.0, 0.02439024390243903),  
('cnn,', 1.0, 0.09090909090909091),  
('up!!!!', 1.0, 0.09090909090909091),  
('sun', 1.0, 0.03225806451612903),  
('two-faced...', 1.0, 0.09090909090909091),

('@ellimayradrep', 1.0, 0.09090909090909091),  
 ('hate!', 1.0, 0.047619047619047616),  
 ('blame', 1.0, 0.01098901098901099),  
 ('#realtalk', 1.0, 0.09090909090909091),  
 ('', 1.0, 0.09090909090909091),  
 ('#teamtrump', 1.0, 0.09090909090909091),  
 ('alkaline', 0.5121951219512195, 0.5121951219512195),  
 ('"evil".', 1.0, 0.09090909090909091),  
 ('man.', 1.0, 0.03225806451612903),  
 ('you've', 0.7560975609756099, 0.26829268292682934),  
 ('estimating', 1.0, 0.09090909090909091),  
 ('mind', 0.8895027624309393, 0.11602209944751381),  
 ('fair.', 1.0, 0.09090909090909091),  
 ('@dannijean1640', 1.0, 0.09090909090909091),  
 ('@klingensmith\_j', 1.0, 0.09090909090909091),  
 ('fatal', 1.0, 0.09090909090909091),  
 ('companies', 1.0, 0.01098901098901099),  
 ('work!', 0.5238095238095238, 0.5238095238095238),  
 ('@dearillegirl', 1.0, 0.09090909090909091),  
 ('group', 1.0, 0.019607843137254905),  
 ('share', 0.9052132701421801, 0.0995260663507109),  
 ('entire', 0.8360655737704918, 0.18032786885245905),  
 ('she's', 0.801980198019802, 0.20792079207920794),  
 ('queensland', 1.0, 0.09090909090909091),  
 ('plot.', 1.0, 0.09090909090909091),  
 ('"other"', 1.0, 0.09090909090909091),  
 ('for!', 0.5238095238095238, 0.5238095238095238),  
 ('@devarshi\_21', 0.047619047619047616, 1.0),  
 ('@mohitku89856580', 0.047619047619047616, 1.0),  
 ('@swetasamadhiya', 0.047619047619047616, 1.0),  
 ('@sp\_umang', 0.047619047619047616, 1.0),  
 ('@rakesha12572734', 0.047619047619047616, 1.0),  
 ('@vera\_anz', 0.047619047619047616, 1.0),  
 ('@ar\_jag84', 0.047619047619047616, 1.0),  
 ('@debduttakolkata', 0.047619047619047616, 1.0),  
 ('@sbhats', 0.047619047619047616, 1.0),  
 ('@raibahadur13', 0.047619047619047616, 1.0),  
 ('@arrorasanjeev', 0.047619047619047616, 1.0),  
 ('leaves', 0.5061728395061729, 0.5061728395061729),  
 ('flowers', 0.047619047619047616, 1.0),  
 ('tree,', 0.09090909090909091, 1.0),  
 ('lose', 1.0, 0.008264462809917356),  
 ('weight,', 1.0, 0.047619047619047616),  
 ('cash', 1.0, 0.03225806451612903),  
 ('ago', 0.8591549295774648, 0.15492957746478875),  
 ('follows', 1.0, 0.047619047619047616),  
 ('out.', 0.7183098591549295, 0.295774647887324),  
 ('@realcandaceo', 1.0, 0.047619047619047616),

('whine,', 1.0, 0.09090909090909091),  
 ('diseases?trump', 1.0, 0.09090909090909091),  
 ('discriminatory!'", 1.0, 0.09090909090909091),  
 ('aunts', 1.0, 0.09090909090909091),  
 ('treatment?', 1.0, 0.09090909090909091),  
 ('convinced', 1.0, 0.019607843137254905),  
 ('ready', 1.0, 0.02439024390243903),  
 ('@docbastard,', 1.0, 0.09090909090909091),  
 ('cahoots', 1.0, 0.047619047619047616),  
 ('@tatamemorial', 1.0, 0.09090909090909091),  
 ('latest', 0.6296296296296297, 0.3827160493827161),  
 ('#cancon2018', 1.0, 0.09090909090909091),  
 ('@cytecare', 1.0, 0.09090909090909091),  
 ('improve', 1.0, 0.019607843137254905),  
 ('@freedomfromfear', 1.0, 0.09090909090909091),  
 ('@secpompeo', 1.0, 0.09090909090909091),  
 ('save', 1.0, 0.004329004329004329),  
 ('@drjonesaa', 1.0, 0.09090909090909091),  
 ('@carequalitycomm', 1.0, 0.09090909090909091),  
 ('@zeromilly', 1.0, 0.09090909090909091),  
 ('@compassnincare', 1.0, 0.09090909090909091),  
 ('@jasonginnelly1', 1.0, 0.09090909090909091),  
 ('@kenzeroharm', 1.0, 0.09090909090909091),  
 ('@shaunlintern', 1.0, 0.09090909090909091),  
 ('@bbcnews', 1.0, 0.09090909090909091),  
 ('honest,', 1.0, 0.09090909090909091),  
 ('system,', 0.7560975609756099, 0.26829268292682934),  
 ('chibok', 1.0, 0.09090909090909091),  
 ('girls,', 1.0, 0.09090909090909091),  
 ('neck', 1.0, 0.047619047619047616),  
 ('akwa', 1.0, 0.09090909090909091),  
 ('ways...trump', 1.0, 0.09090909090909091),  
 ('https://t.co/b0jnv1vwhh', 1.0, 0.09090909090909091),  
 ('@drmendelsohn', 1.0, 0.09090909090909091),  
 ('@jerome\_corsi', 1.0, 0.09090909090909091),  
 ('thinks', 0.3548387096774194, 0.6774193548387097),  
 ('conventional', 0.047619047619047616, 1.0),  
 ('cow', 0.1603053435114504, 0.8473282442748091),  
 ('meltdown,', 1.0, 0.09090909090909091),  
 ('https://t.co/60aq9x33et', 1.0, 0.09090909090909091),  
 ('defend', 1.0, 0.03225806451612903),  
 ('cancer..', 1.0, 0.09090909090909091),  
 ('@education4libs', 1.0, 0.047619047619047616),  
 ('showing', 1.0, 0.02439024390243903),  
 ('backlash', 1.0, 0.09090909090909091),  
 ('https://t.co/3xtsyqurrx', 1.0, 0.09090909090909091),  
 ('americans', 0.8765432098765432, 0.1358024691358025),  
 ('pri#k!', 1.0, 0.09090909090909091),



('@exandshadow:', 1.0, 0.09090909090909091),  
 ('crossing', 1.0, 0.09090909090909091),  
 ('riots', 1.0, 0.047619047619047616),  
 ('jaw', 1.0, 0.09090909090909091),  
 ('@fandincollingo:', 1.0, 0.09090909090909091),  
 ('@williewumpkins', 1.0, 0.09090909090909091),  
 ('shoes', 1.0, 0.03225806451612903),  
 ('hiding', 1.0, 0.009900990099009901),  
 ('gallops', 1.0, 0.09090909090909091),  
 ('fart', 1.0, 0.09090909090909091),  
 ('came', 0.7530864197530864, 0.2592592592592593),  
 ('@thepooluk', 0.5238095238095238, 0.5238095238095238),  
 ('right?', 1.0, 0.03225806451612903),  
 ('head', 1.0, 0.01639344262295082),  
 ('youtube', 1.0, 0.09090909090909091),  
 ('newspaper', 1.0, 0.047619047619047616),  
 ('@fourfour\_stu:', 1.0, 0.09090909090909091),  
 ('@frankcottrell\_b:', 1.0, 0.09090909090909091),  
 ('@frankie\_k97:', 1.0, 0.09090909090909091),  
 ('answered', 1.0, 0.09090909090909091),  
 ('exist.', 1.0, 0.09090909090909091),  
 ('https://t.co/cjrrlguxbg', 1.0, 0.09090909090909091),  
 ('medicine', 0.504950495049505, 0.504950495049505),  
 ('#funnythingsdadssay', 1.0, 0.09090909090909091),  
 ('@somyrrrh', 0.09090909090909091, 1.0),  
 ('@joshua1\_5', 0.09090909090909091, 1.0),  
 ('@lynnleemavakay', 0.09090909090909091, 1.0),  
 ('@melissa232220', 0.09090909090909091, 1.0),  
 ('@pileofgoop', 0.09090909090909091, 1.0),  
 ('@kidsdoc1962', 0.09090909090909091, 1.0),  
 ('@and\_kell', 0.09090909090909091, 1.0),  
 ('@bluelionblog', 0.09090909090909091, 1.0),  
 ('high', 0.6946564885496184, 0.3129770992366412),  
 ('challenge', 0.6774193548387097, 0.3548387096774194),  
 ('inuo.', 0.09090909090909091, 1.0),  
 ('radiation.', 0.5238095238095238, 0.5238095238095238),  
 ('hawaii.', 0.09090909090909091, 1.0),  
 ('person', 0.7790055248618784, 0.22651933701657453),  
 ('pump', 1.0, 0.09090909090909091),  
 ('planning', 1.0, 0.09090909090909091),  
 ('@gorskonskon:', 1.0, 0.09090909090909091),  
 ('see!', 1.0, 0.09090909090909091),  
 ('promoting', 0.5238095238095238, 0.5238095238095238),  
 ('@pkelpe', 0.09090909090909091, 1.0),  
 ('@northtxmsg', 0.09090909090909091, 1.0),  
 ('@jackmaiello', 0.09090909090909091, 1.0),  
 ('@imfsupport', 0.09090909090909091, 1.0),  
 ('@wwwillstand', 1.0, 0.047619047619047616),

('@integrumwm', 1.0, 0.09090909090909091),  
 ('everything!', 1.0, 0.09090909090909091),  
 ('act', 1.0, 0.02439024390243903),  
 ('low-level', 1.0, 0.09090909090909091),  
 ('creature', 1.0, 0.09090909090909091),  
 ('sampling', 1.0, 0.09090909090909091),  
 ('music', 1.0, 0.047619047619047616),  
 ('poorly', 1.0, 0.047619047619047616),  
 ('benefit', 0.7183098591549295, 0.295774647887324),  
 ('you?', 0.8360655737704918, 0.18032786885245905),  
 ('taxes,', 1.0, 0.09090909090909091),  
 ('pipeline', 1.0, 0.09090909090909091),  
 ('hostages', 1.0, 0.09090909090909091),  
 ('https://t.co/ozlb4nrpux', 1.0, 0.09090909090909091),  
 ('@erwindean96', 1.0, 0.09090909090909091),  
 ('@hbronnenberg', 1.0, 0.09090909090909091),  
 ('@bronsonsrevenge', 1.0, 0.09090909090909091),  
 ('@3millfam', 1.0, 0.09090909090909091),  
 ('@patriot\_m1a', 1.0, 0.09090909090909091),  
 ('@sand\_manx', 1.0, 0.09090909090909091),  
 ('@loveshoulas', 1.0, 0.09090909090909091),  
 ('@chighbe', 1.0, 0.09090909090909091),  
 ('@pistolgrip6spd1', 1.0, 0.09090909090909091),  
 ('@seeingred02', 1.0, 0.09090909090909091),  
 ('@natty\_bumppo45', 1.0, 0.09090909090909091),  
 ('business', 1.0, 0.009900990099009901),  
 ('join', 1.0, 0.006211180124223602),  
 ('@henkelmanrb:', 1.0, 0.09090909090909091),  
 ('suppresses', 0.5238095238095238, 0.5238095238095238),  
 ('this,', 1.0, 0.01639344262295082),  
 ('https://t.co/x8pl7ybuz7', 1.0, 0.09090909090909091),  
 ('@hubie0', 1.0, 0.09090909090909091),  
 ('@halcyondon', 1.0, 0.09090909090909091),  
 ('@stablegeniusrn', 1.0, 0.09090909090909091),  
 ('@captainair727', 1.0, 0.09090909090909091),  
 ('@5th\_ape', 1.0, 0.09090909090909091),  
 ('some.', 1.0, 0.09090909090909091),  
 ('chemo!'", 1.0, 0.09090909090909091),  
 ('@wordpress...', 1.0, 0.09090909090909091),  
 ('@iamamazinglucas:', 1.0, 0.09090909090909091),  
 ('india,', 0.01639344262295082, 1.0),  
 ('therapy', 0.01098901098901099, 1.0),  
 ('cancer...many', 0.019607843137254905, 1.0),  
 ('cured..', 0.019607843137254905, 1.0),  
 ('mock', 0.09090909090909091, 1.0),  
 ('stricken', 0.09090909090909091, 1.0),  
 ('situation', 0.5238095238095238, 0.5238095238095238),  
 ('everytime', 0.09090909090909091, 1.0),

('seen', 0.5604395604395604, 0.4505494505494505),  
 ('https://t.co/60cery8k8y', 0.09090909090909091, 1.0),  
 ('https://t.co/ujscephajl', 0.09090909090909091, 1.0),  
 ('they'd', 1.0, 0.01234567901234568),  
 ('pay', 1.0, 0.014084507042253523),  
 ('generate', 1.0, 0.047619047619047616),  
 ('billions', 1.0, 0.01234567901234568),  
 ('too.', 0.7183098591549295, 0.295774647887324),  
 ('wouldn't', 1.0, 0.008264462809917356),  
 ('feed', 1.0, 0.03225806451612903),  
 ('hungry', 1.0, 0.047619047619047616),  
 ('scotland', 1.0, 0.09090909090909091),  
 ('@inzcooler', 0.09090909090909091, 1.0),  
 ('depression,', 0.5238095238095238, 0.5238095238095238),  
 ('researchers,', 1.0, 0.01234567901234568),  
 ('diseases?', 0.5238095238095238, 0.5238095238095238),  
 ('cancer.imagine', 0.09090909090909091, 1.0),  
 ('free.bravin', 0.09090909090909091, 1.0),  
 ('holding', 1.0, 0.03225806451612903),  
 ('can,', 1.0, 0.03225806451612903),  
 ('genetic', 0.7560975609756099, 0.26829268292682934),  
 ('hats', 1.0, 0.09090909090909091),  
 ('hype', 1.0, 0.09090909090909091),  
 ('crap', 1.0, 0.019607843137254905),  
 ('unknown', 1.0, 0.09090909090909091),  
 ('mean', 0.5287958115183246, 0.4764397905759162),  
 ('"just"', 1.0, 0.09090909090909091),  
 ('@jdugudichi', 1.0, 0.09090909090909091),  
 ('@daniscotchirish', 1.0, 0.03225806451612903),  
 ('@kellyannegoebb', 1.0, 0.09090909090909091),  
 ('@gregthedoughboy', 1.0, 0.09090909090909091),  
 ('@joanne48640679', 1.0, 0.09090909090909091),  
 ('@leopoldjody', 1.0, 0.09090909090909091),  
 ('@james\_peluso', 1.0, 0.09090909090909091),  
 ('@pendingvictory', 1.0, 0.09090909090909091),  
 ('@sgriffin\_87to98', 1.0, 0.09090909090909091),  
 ('@dcarver1018', 1.0, 0.09090909090909091),  
 ('@wendidavies', 1.0, 0.09090909090909091),  
 ('@rynlds56', 1.0, 0.09090909090909091),  
 ('fool', 1.0, 0.09090909090909091),  
 ('playground', 1.0, 0.09090909090909091),  
 ('support"', 1.0, 0.09090909090909091),  
 ('tomorrow', 0.8591549295774648, 0.15492957746478875),  
 ('putin', 1.0, 0.09090909090909091),  
 ('@jeremypoxon', 1.0, 0.09090909090909091),  
 ('murdered', 1.0, 0.014084507042253523),  
 ('kills', 0.803921568627451, 0.21568627450980396),  
 ('virginity', 1.0, 0.09090909090909091),

('caused', 1.0, 0.03225806451612903),  
 ('drinking', 0.4366197183098592, 0.5774647887323944),  
 ('sebi', 0.8591549295774648, 0.15492957746478875),  
 ('amonth', 1.0, 0.09090909090909091),  
 ('thought,', 1.0, 0.09090909090909091),  
 ('archon', 1.0, 0.09090909090909091),  
 ('their/our', 1.0, 0.09090909090909091),  
 ('played', 1.0, 0.03225806451612903),  
 ('"donald', 1.0, 0.047619047619047616),  
 ('\*nukes', 1.0, 0.09090909090909091),  
 ('@lplzydeco', 1.0, 0.09090909090909091),  
 ('undercut', 1.0, 0.09090909090909091),  
 ('expertise', 1.0, 0.09090909090909091),  
 ('cancer?', 0.9261992619926199, 0.07749077490774908),  
 ('president,', 1.0, 0.019607843137254905),  
 ('@jowalker2960:', 1.0, 0.09090909090909091),  
 ('democrats', 1.0, 0.004975124378109453),  
 ('@foxandfriends', 1.0, 0.047619047619047616),  
 ('home.', 1.0, 0.03225806451612903),  
 ('ass', 1.0, 0.009900990099009901),  
 ('"the', 0.8765432098765432, 0.1358024691358025),  
 ('legacy', 1.0, 0.047619047619047616),  
 ('yes', 1.0, 0.014084507042253523),  
 ('squirm', 1.0, 0.09090909090909091),  
 ('we'll', 1.0, 0.047619047619047616),  
 ('@ctfletcherisymf', 1.0, 0.09090909090909091),  
 ('!!!!', 1.0, 0.047619047619047616),  
 ('@botagain', 1.0, 0.09090909090909091),  
 ('@jrmtm50', 1.0, 0.09090909090909091),  
 ('beautifully', 1.0, 0.09090909090909091),  
 ('helpful', 0.5238095238095238, 0.5238095238095238),  
 ('proactive', 1.0, 0.03225806451612903),  
 ('precipitous', 1.0, 0.09090909090909091),  
 ('fall', 1.0, 0.02439024390243903),  
 ('6wks.', 1.0, 0.09090909090909091),  
 ('@ksnazzylawler', 1.0, 0.09090909090909091),  
 ('mindset', 1.0, 0.047619047619047616),  
 ('draft,', 1.0, 0.09090909090909091),  
 ('especially', 0.6774193548387097, 0.3548387096774194),  
 ('@maga', 1.0, 0.09090909090909091),  
 ('state', 0.6296296296296297, 0.3827160493827161),  
 ('\$ncp', 1.0, 0.047619047619047616),  
 ('members', 1.0, 0.047619047619047616),  
 ('recently,', 1.0, 0.09090909090909091),  
 ('publicity', 1.0, 0.09090909090909091),  
 ('@dave\_attewell', 1.0, 0.09090909090909091),  
 ('@bjaninemorison', 1.0, 0.09090909090909091),  
 ('@laura2qt:', 1.0, 0.09090909090909091),

('dinosaurs', 1.0, 0.019607843137254905),  
 ('dumb???', 1.0, 0.09090909090909091),  
 ('sixers', 1.0, 0.09090909090909091),  
 ('nice', 1.0, 0.02439024390243903),  
 ('@armandodkos', 1.0, 0.09090909090909091),  
 ('beyond', 1.0, 0.047619047619047616),  
 ('belief', 1.0, 0.09090909090909091),  
 ('@lifezette:', 1.0, 0.09090909090909091),  
 ('host', 1.0, 0.03225806451612903),  
 ('@lilsweetstina', 1.0, 0.09090909090909091),  
 ('snakes,', 1.0, 0.09090909090909091),  
 ('norris', 1.0, 0.09090909090909091),  
 ('yawns,', 1.0, 0.09090909090909091),  
 ('#emptysouls', 1.0, 0.09090909090909091),  
 ('filmed', 1.0, 0.09090909090909091),  
 ('@lp62o', 1.0, 0.09090909090909091),  
 ('easily', 1.0, 0.047619047619047616),  
 ('mexican', 1.0, 0.047619047619047616),  
 ('fence', 1.0, 0.047619047619047616),  
 ('hair', 1.0, 0.019607843137254905),  
 ('form', 0.7183098591549295, 0.295774647887324),  
 ('@lunchboxfr', 1.0, 0.09090909090909091),  
 ('turn!', 1.0, 0.09090909090909091),  
 ('ignorant!', 1.0, 0.09090909090909091),  
 ('causes', 1.0, 0.047619047619047616),  
 ('kick', 1.0, 0.03225806451612903),  
 ('rhetoric', 0.5081967213114754, 0.5081967213114754),  
 ('forgive', 1.0, 0.047619047619047616),  
 ('@thom\_not\_a\_bot', 1.0, 0.09090909090909091),  
 ('bro', 1.0, 0.09090909090909091),  
 ('...confusion', 1.0, 0.09090909090909091),  
 ('exercise', 1.0, 0.09090909090909091),  
 ('losses', 1.0, 0.09090909090909091),  
 ('mr.', 1.0, 0.02439024390243903),  
 ('@uberfacts', 1.0, 0.09090909090909091),  
 ('@stevejazz', 1.0, 0.09090909090909091),  
 ('@willkinginri', 1.0, 0.09090909090909091),  
 ('boring', 1.0, 0.09090909090909091),  
 ('politics', 1.0, 0.03225806451612903),  
 ('holds', 1.0, 0.047619047619047616),  
 ('regarded', 1.0, 0.09090909090909091),  
 ('vitamins', 0.5238095238095238, 0.5238095238095238),  
 ('relaxation.\xa0', 0.5238095238095238, 0.5238095238095238),  
 ('bigoted', 1.0, 0.09090909090909091),  
 ('(just', 1.0, 0.09090909090909091),  
 ('tissue)', 1.0, 0.09090909090909091),  
 ('@ellebeaublog', 1.0, 0.09090909090909091),  
 ('@facebook', 1.0, 0.09090909090909091),

('encouraged', 1.0, 0.09090909090909091),  
 ('#juiceplus', 1.0, 0.09090909090909091),  
 ('removed', 0.803921568627451, 0.21568627450980396),  
 ('behaviours', 1.0, 0.09090909090909091),  
 ('@mikeggibbs', 1.0, 0.09090909090909091),  
 ('conservatives.', 1.0, 0.09090909090909091),  
 ('eliminates', 1.0, 0.09090909090909091),  
 ('@mirrorweirdnews:', 0.09090909090909091, 1.0),  
 ('proved', 1.0, 0.09090909090909091),  
 ('2020!', 1.0, 0.09090909090909091),  
 ('@dccra', 1.0, 0.09090909090909091),  
 ('@marilyncapps', 1.0, 0.09090909090909091),  
 ('@downsouth505', 1.0, 0.09090909090909091),  
 ('@girl\_daisey', 1.0, 0.09090909090909091),  
 ('@jahrtt', 1.0, 0.09090909090909091),  
 ('@callisti2024', 1.0, 0.09090909090909091),  
 ('@assemblyroomusa', 1.0, 0.09090909090909091),  
 ('@msgoddessrises', 1.0, 0.09090909090909091),  
 ('@libertyscott63', 1.0, 0.09090909090909091),  
 ('@nude4peace', 1.0, 0.09090909090909091),  
 ('2018/', 1.0, 0.09090909090909091),  
 ('pissed', 1.0, 0.03225806451612903),  
 ('did.', 1.0, 0.047619047619047616),  
 ('@msscha2', 1.0, 0.09090909090909091),  
 ('idiot', 1.0, 0.047619047619047616),  
 ('clinton.', 1.0, 0.09090909090909091),  
 ('shows.', 1.0, 0.09090909090909091),  
 ('views!!!', 1.0, 0.09090909090909091),  
 ('@mussavir', 0.09090909090909091, 1.0),  
 ('useful.', 0.09090909090909091, 1.0),  
 ('indica', 0.09090909090909091, 1.0),  
 ('dhrek', 0.09090909090909091, 1.0),  
 ('useful', 0.6774193548387097, 0.3548387096774194),  
 ('tree.', 0.09090909090909091, 1.0),  
 ('@mkstalin', 1.0, 0.09090909090909091),  
 ('@fouress', 1.0, 0.09090909090909091),  
 ('@lillymarypinto', 1.0, 0.09090909090909091),  
 ('@suryahsg', 1.0, 0.09090909090909091),  
 ('heard', 1.0, 0.01639344262295082),  
 ('published', 0.34426229508196726, 0.6721311475409836),  
 ('@mrdanecw', 1.0, 0.09090909090909091),  
 ('begin?!', 1.0, 0.09090909090909091),  
 ('algorithms', 1.0, 0.09090909090909091),  
 ('source', 0.047619047619047616, 1.0),  
 ('@nc4everblessed', 1.0, 0.09090909090909091),  
 ('@tuckercarlson', 1.0, 0.09090909090909091),  
 ('rug', 1.0, 0.09090909090909091),  
 ('niagara', 1.0, 0.09090909090909091),

('falls!', 1.0, 0.09090909090909091),  
 ('remission', 0.6774193548387097, 0.3548387096774194),  
 ('diseases.', 0.7560975609756099, 0.26829268292682934),  
 ('@nintendoamerica', 1.0, 0.09090909090909091),  
 ('incorrect.', 1.0, 0.09090909090909091),  
 ('alrighty', 1.0, 0.09090909090909091),  
 ('pressure.', 1.0, 0.09090909090909091),  
 ('numpty', 1.0, 0.09090909090909091),  
 ('bureaucrats.', 1.0, 0.09090909090909091),  
 ('@northkoreatimes', 0.09090909090909091, 1.0),  
 ('#northkorea', 0.5238095238095238, 0.5238095238095238),  
 ('please!!!', 0.09090909090909091, 1.0),  
 ('aids,', 0.801980198019802, 0.20792079207920794),  
 ('"they', 1.0, 0.047619047619047616),  
 ('weed', 0.01639344262295082, 1.0),  
 ('@telegraphnews', 1.0, 0.09090909090909091),  
 ('pabgirl', 1.0, 0.09090909090909091),  
 ('"didn\'t', 1.0, 0.09090909090909091),  
 ('scam-artists:', 1.0, 0.09090909090909091),  
 ('@pastormarkburns', 1.0, 0.09090909090909091),  
 ('pastor,', 1.0, 0.09090909090909091),  
 ('cancer...', 1.0, 0.047619047619047616),  
 ('@peacerangel', 1.0, 0.09090909090909091),  
 ('extreme', 0.6774193548387097, 0.3548387096774194),  
 ('honest.', 1.0, 0.03225806451612903),  
 ('liked', 1.0, 0.09090909090909091),  
 ('clinton's", 1.0, 0.09090909090909091),  
 ('common', 1.0, 0.01639344262295082),  
 ('cold.', 1.0, 0.03225806451612903),  
 ('city.', 1.0, 0.09090909090909091),  
 ('@ranjha001', 1.0, 0.09090909090909091),  
 ('abortion', 1.0, 0.03225806451612903),  
 ('burning', 1.0, 0.09090909090909091),  
 ('building...i', 1.0, 0.09090909090909091),  
 ('who"will', 1.0, 0.09090909090909091),  
 ('seed', 0.607843137254902, 0.411764705882353),  
 ('#qanon', 0.7560975609756099, 0.26829268292682934),  
 ('#noko', 1.0, 0.09090909090909091),  
 ('nobel', 1.0, 0.02439024390243903),  
 ('peace.', 1.0, 0.02439024390243903),  
 ('allcharity', 1.0, 0.047619047619047616),  
 ('@just2bizi', 1.0, 0.09090909090909091),  
 ('afraid.', 1.0, 0.09090909090909091),  
 ('https://t.co/td5dddjce0', 1.0, 0.09090909090909091),  
 ('homelessness,', 1.0, 0.09090909090909091),  
 ('change"', 1.0, 0.09090909090909091),  
 ('hospitals', 0.5238095238095238, 0.5238095238095238),  
 ('rain', 1.0, 0.09090909090909091),

('credit!', 1.0, 0.09090909090909091),  
 ('@hrdministry', 1.0, 0.09090909090909091),  
 ('@akashvaniair', 1.0, 0.09090909090909091),  
 ('jio', 1.0, 0.047619047619047616),  
 ('ended', 1.0, 0.03225806451612903),  
 ('reversed', 1.0, 0.09090909090909091),  
 ('climate', 1.0, 0.01639344262295082),  
 ('ask', 1.0, 0.09090909090909091),  
 ('covenant', 1.0, 0.09090909090909091),  
 ('+2349057353987,', 1.0, 0.09090909090909091),  
 ('bollinger', 1.0, 0.09090909090909091),  
 ('cancer.and', 1.0, 0.09090909090909091),  
 ('ann', 1.0, 0.047619047619047616),  
 ('https://t.co/nozmkz5l6h', 0.09090909090909091, 1.0),  
 ('virtue', 1.0, 0.09090909090909091),  
 ('#dems', 1.0, 0.09090909090909091),  
 ('akhi', 1.0, 0.09090909090909091),  
 ('applauding', 1.0, 0.09090909090909091),  
 ('strive', 1.0, 0.09090909090909091),  
 ('#bestpresidentever', 1.0, 0.09090909090909091),  
 ('https://t.co/zywsqvp0bj', 1.0, 0.09090909090909091),  
 ('unemployment".', 1.0, 0.09090909090909091),  
 ('headline', 1.0, 0.03225806451612903),  
 ('#nksummit', 1.0, 0.09090909090909091),  
 ('universe', 0.6774193548387097, 0.3548387096774194),  
 ('daughters', 1.0, 0.09090909090909091),  
 ('neurological', 1.0, 0.09090909090909091),  
 ('sad.', 1.0, 0.09090909090909091),  
 ('tragic.', 1.0, 0.09090909090909091),  
 ('record', 1.0, 0.09090909090909091),  
 ('anymore', 1.0, 0.047619047619047616),  
 ('hdmi', 1.0, 0.09090909090909091),  
 ('@hazyhula', 1.0, 0.09090909090909091),  
 ('@\_iamcannabis', 1.0, 0.09090909090909091),  
 ('@cannilive', 1.0, 0.09090909090909091),  
 ('cute', 1.0, 0.09090909090909091),  
 ('selling', 0.5061728395061729, 0.5061728395061729),  
 ('google,', 1.0, 0.09090909090909091),  
 ('researching!', 1.0, 0.09090909090909091),  
 ('guns...', 1.0, 0.09090909090909091),  
 ('@deeescalante', 1.0, 0.09090909090909091),  
 ('@alexandcohen', 1.0, 0.047619047619047616),  
 ('@blogliberally', 1.0, 0.047619047619047616),  
 ('@emergvictorious', 1.0, 0.047619047619047616),  
 ('@carolyncolson', 1.0, 0.047619047619047616),  
 ('@stopbaddocs', 1.0, 0.047619047619047616),  
 ('@thomasklinemd', 1.0, 0.047619047619047616),  
 ('@michelleziemba', 1.0, 0.047619047619047616),



( '@chrisellington5', 1.0, 0.047619047619047616),  
('space', 1.0, 0.03225806451612903),  
('cousins', 1.0, 0.09090909090909091),  
('dhs,', 1.0, 0.047619047619047616),  
('lying,', 1.0, 0.09090909090909091),  
('father', 0.803921568627451, 0.21568627450980396),  
('ensure', 1.0, 0.02439024390243903),  
('outside,', 1.0, 0.09090909090909091),  
('lenin,', 1.0, 0.09090909090909091),  
('@realdeancain', 1.0, 0.09090909090909091),  
('@stclairashley', 1.0, 0.09090909090909091),  
('technology', 1.0, 0.02439024390243903),  
('turn', 1.0, 0.01639344262295082),  
('child', 0.6183206106870229, 0.3893129770992366),  
('operations', 1.0, 0.09090909090909091),  
('https://t.co/vkftlvxyuf', 1.0, 0.09090909090909091),  
('@saylesjon', 1.0, 0.09090909090909091),  
('says', 0.8347107438016529, 0.17355371900826447),  
('(nay', 1.0, 0.09090909090909091),  
('...', 1.0, 0.09090909090909091),  
('guitars', 1.0, 0.09090909090909091),  
(' ', 1.0, 0.09090909090909091),  
('https://t.co/je3gryeiw7', 1.0, 0.09090909090909091),  
('he'd', 1.0, 0.03225806451612903),  
('soldiers', 1.0, 0.09090909090909091),  
('drs', 1.0, 0.09090909090909091),  
('demokrats', 1.0, 0.09090909090909091),  
('foreign', 1.0, 0.09090909090909091),  
('schools', 1.0, 0.09090909090909091),  
('race', 0.9415204678362573, 0.06432748538011696),  
('hustler', 1.0, 0.047619047619047616),  
('pro-choice', 1.0, 0.09090909090909091),  
('born.', 1.0, 0.09090909090909091),  
('racecar', 1.0, 0.047619047619047616),  
('email.', 1.0, 0.03225806451612903),  
('https://t.co/cosx4qbck3', 1.0, 0.09090909090909091),  
('@rileyjameswarr', 1.0, 0.09090909090909091),  
('scores', 1.0, 0.09090909090909091),  
('one)', 1.0, 0.09090909090909091),  
('touch', 1.0, 0.09090909090909091),  
('politicians...', 1.0, 0.09090909090909091),  
('dummy', 1.0, 0.09090909090909091),  
('america,', 1.0, 0.09090909090909091),  
('fricking', 1.0, 0.09090909090909091),  
('crying', 0.8360655737704918, 0.18032786885245905),  
('supposedly', 0.5121951219512195, 0.5121951219512195),  
('rinos', 1.0, 0.09090909090909091),  
('https://t.co/8bmt73v7hr', 1.0, 0.09090909090909091),

('fail,', 1.0, 0.09090909090909091),  
 ('payments,', 1.0, 0.09090909090909091),  
 ('allow', 1.0, 0.02439024390243903),  
 ('plays', 1.0, 0.047619047619047616),  
 ('vision', 0.8591549295774648, 0.15492957746478875),  
 ('libs,', 1.0, 0.09090909090909091),  
 ('trumpers:', 1.0, 0.09090909090909091),  
 ('@2b7c89526', 1.0, 0.09090909090909091),  
 ('@cmccitizenmedia', 1.0, 0.09090909090909091),  
 ('donate.', 1.0, 0.047619047619047616),  
 ('@elerhinomarch', 0.09090909090909091, 1.0),  
 ('@louisal1000', 0.09090909090909091, 1.0),  
 ('change,', 1.0, 0.09090909090909091),  
 ('perhaps', 1.0, 0.03225806451612903),  
 ('okkkk', 1.0, 0.09090909090909091),  
 ('@supremesaxx:', 1.0, 0.09090909090909091),  
 ('genuis', 1.0, 0.09090909090909091),  
 ('jerk', 1.0, 0.047619047619047616),  
 ('@sxrabit:', 1.0, 0.09090909090909091),  
 ('@grxxvec', 1.0, 0.09090909090909091),  
 ('recognize', 1.0, 0.09090909090909091),  
 ('@teapainusa', 1.0, 0.09090909090909091),  
 ('population!""', 1.0, 0.09090909090909091),  
 ('.no', 0.5238095238095238, 0.5238095238095238),  
 ('#sendsarahskiddies2disney', 1.0, 0.09090909090909091),  
 ('@thehill', 1.0, 0.02439024390243903),  
 ('disappear,', 1.0, 0.09090909090909091),  
 ('determined', 1.0, 0.09090909090909091),  
 ('details.', 1.0, 0.09090909090909091),  
 ('qualities,', 1.0, 0.09090909090909091),  
 ('starting', 1.0, 0.047619047619047616),  
 ('reverse', 1.0, 0.09090909090909091),  
 ('#impeach', 1.0, 0.09090909090909091),  
 ('jackasses', 1.0, 0.09090909090909091),  
 ('happening', 1.0, 0.047619047619047616),  
 ('@deji14\_official', 1.0, 0.09090909090909091),  
 ('@manesti336', 1.0, 0.09090909090909091),  
 ('nightmare', 1.0, 0.09090909090909091),  
 ('brough', 1.0, 0.09090909090909091),  
 ('terminally', 1.0, 0.09090909090909091),  
 ('https://t.co/r4eo6zjya1', 1.0, 0.09090909090909091),  
 ('opinions?', 1.0, 0.09090909090909091),  
 ('@elcontador2000', 1.0, 0.09090909090909091),  
 ('@willfranken', 1.0, 0.09090909090909091),  
 ('@poshfruit', 1.0, 0.09090909090909091),  
 ('spelling', 1.0, 0.09090909090909091),  
 ('already.', 1.0, 0.09090909090909091),  
 ('wasting', 1.0, 0.03225806451612903),

('senate,', 1.0, 0.09090909090909091),  
 ('elites', 1.0, 0.09090909090909091),  
 ('industry,', 1.0, 0.09090909090909091),  
 ('fbi', 1.0, 0.09090909090909091),  
 ('\$3.89', 1.0, 0.09090909090909091),  
 ('higher', 0.7560975609756099, 0.26829268292682934),  
 ('indie-band', 1.0, 0.09090909090909091),  
 ('walk.', 1.0, 0.09090909090909091),  
 ('timeline', 1.0, 0.09090909090909091),  
 ('jetpacks', 1.0, 0.09090909090909091),  
 ('pea,', 0.09090909090909091, 1.0),  
 ('gop', 0.09090909090909091, 1.0),  
 ('@webbgolf:', 1.0, 0.09090909090909091),  
 ('peanut', 1.0, 0.09090909090909091),  
 ('jelly', 1.0, 0.09090909090909091),  
 ('usc', 0.09090909090909091, 1.0),  
 ('ontario,', 0.09090909090909091, 1.0),  
 ('disgraceful.', 1.0, 0.09090909090909091),  
 ('decided', 1.0, 0.03225806451612903),  
 ('@vinayprasamd', 1.0, 0.09090909090909091),  
 ('palliation', 1.0, 0.09090909090909091),  
 ('@vinesauce', 1.0, 0.09090909090909091),  
 ('charity', 1.0, 0.02439024390243903),  
 ('pleased', 1.0, 0.047619047619047616),  
 ('announce', 1.0, 0.03225806451612903),  
 ('sunday,', 1.0, 0.03225806451612903),  
 ('"lol"', 1.0, 0.09090909090909091),  
 ('highschool,', 1.0, 0.09090909090909091),  
 ('blaming', 1.0, 0.047619047619047616),  
 ('rising', 1.0, 0.09090909090909091),  
 ('(weird) ', 1.0, 0.09090909090909091),  
 ('tweet!', 1.0, 0.09090909090909091),  
 ('canada.', 1.0, 0.09090909090909091),  
 ('stance...', 0.09090909090909091, 1.0),  
 ('from...', 0.09090909090909091, 1.0),  
 ('meme...', 0.09090909090909091, 1.0),  
 ('lib?', 0.09090909090909091, 1.0),  
 ('discussing', 0.09090909090909091, 1.0),  
 ('internet', 0.5238095238095238, 0.5238095238095238),  
 ('"cancer', 0.09090909090909091, 1.0),  
 ('experts"', 0.09090909090909091, 1.0),  
 ('start', 0.6183206106870229, 0.3893129770992366),  
 ('exaggerating', 1.0, 0.09090909090909091),  
 ('shelves.', 1.0, 0.09090909090909091),  
 ('high,', 1.0, 0.09090909090909091),  
 ('shkreli', 1.0, 0.09090909090909091),  
 ('bit,', 1.0, 0.09090909090909091),  
 ('hydrogen', 1.0, 0.09090909090909091),

('prevents', 1.0, 0.09090909090909091),  
 ('contributions,', 1.0, 0.09090909090909091),  
 ('actively', 1.0, 0.03225806451612903),  
 ('#scotuspick', 1.0, 0.09090909090909091),  
 ('score', 1.0, 0.01234567901234568),  
 ('certain', 0.09090909090909091, 1.0),  
 ('offers', 0.09090909090909091, 1.0),  
 ('capable', 0.09090909090909091, 1.0),  
 ('esophageal', 0.09090909090909091, 1.0),  
 ('advocate', 1.0, 0.09090909090909091),  
 ('daydream', 1.0, 0.09090909090909091),  
 ('fantasies', 0.5238095238095238, 0.5238095238095238),  
 ('loved', 0.8360655737704918, 0.18032786885245905),  
 ('note,', 1.0, 0.09090909090909091),  
 ('advancement', 1.0, 0.09090909090909091),  
 ('issues', 0.5238095238095238, 0.5238095238095238),  
 ('situations', 0.09090909090909091, 1.0),  
 ('existence:)', 0.09090909090909091, 1.0),  
 ('https://t.co/ymefygzjgx', 0.09090909090909091, 1.0),  
 ('factory', 0.09090909090909091, 1.0),  
 ('poisoning', 0.047619047619047616, 1.0),  
 ('fooled', 0.09090909090909091, 1.0),  
 ('rape', 0.09090909090909091, 1.0),  
 ('times.', 1.0, 0.047619047619047616),  
 ('https://t.co/ya9tolltgr', 1.0, 0.09090909090909091),  
 ('ahh', 1.0, 0.09090909090909091),  
 ('organize', 1.0, 0.09090909090909091),  
 ('#hflpride', 1.0, 0.09090909090909091),  
 ('afterlife', 1.0, 0.09090909090909091),  
 ('theres', 0.09090909090909091, 1.0),  
 ('prone', 1.0, 0.03225806451612903),  
 ('mars', 1.0, 0.09090909090909091),  
 ('https://t.co/bo5rvk0wrm', 1.0, 0.09090909090909091),  
 ('signing', 1.0, 0.09090909090909091),  
 ('store', 0.3548387096774194, 0.6774193548387097),  
 ('raided', 0.09090909090909091, 1.0),  
 ('apricot', 0.01639344262295082, 1.0),  
 ('threatened', 0.09090909090909091, 1.0),  
 ('ladies', 1.0, 0.03225806451612903),  
 ('measles,', 1.0, 0.09090909090909091),  
 ('importance.', 1.0, 0.09090909090909091),  
 ('ian', 0.09090909090909091, 1.0),  
 ('wright', 0.09090909090909091, 1.0),  
 ('world!', 0.09090909090909091, 1.0),  
 ('gratitude', 1.0, 0.09090909090909091),  
 ...]

```
[ ]: # Prepare test data
test_data_rdd = test_data.rdd.map(lambda x: (x[1], x[0]))

[ ]: display(test_data_rdd.collect())

[ ]: # Filter out short words in test data
# filtered_test_data = test_data_rdd.filter(lambda x: all(len(word) > 3 for
↳ word in x[1]))
filtered_test_data = test_data_rdd

[ ]: display(filtered_test_data.collect())

[ ]: # Get a list of words
list_of_words = word_counts.keys().collect()

[ ]: list_of_words
```

```
Out[39]: ['#armyhivestreamingparty',
'purest',
'thing',
'ever',
'experienced.',
'beyonce',
'bts',
'together',
'cure',
'cancer',
'diseases/illnesses',
'bring',
'world',
'peace',
' ',
'#ayurveda_a_panacea',
'"mother',
'healing"',
'@gurmeetramrahim',
'tells',
'ayurveda',
'first',
'scientific',
'medical',
'branch',
'originated',
'india.',
'yoga',
'practices,',
'becomes',
'boon',
```

'cancer,',  
'kidney',  
'disorders,',  
'cardiac',  
'arrest,',  
'etc.',  
'https://t.co/7jdqksviqf',  
" 'mother",  
"healing' .",  
'https://t.co/vkj1uq8sod',  
'#cureforcancer',  
'100',  
'gave',  
'research',  
'teams',  
'back',  
'2014',  
'know',  
'hid',  
'sick',  
'needed',  
'lie',  
'anyone',  
'true',  
'#noname',  
'#executiveorder',  
'sick,',  
'deranged',  
'liberals',  
'happy',  
'president',  
'found',  
'cancer.',  
'morons',  
'say,',  
"no!",  
'want',  
'people',  
'suffer',  
'die! "',  
'unhinged',  
'nothing',  
'please',  
'them.',  
'thrive',  
'something',  
'bitch',  
'about.',

'#garethsouthgatewould',  
'solve',  
'hunger',  
'even',  
'take',  
'credit',  
'say',  
'"the',  
'lads',  
'well"',  
'#worldcup18',  
'#england',  
'#lgbtactionplan',  
'important',  
'step',  
'become',  
'equally',  
'diverse',  
'place',  
'love.',  
'gay,',  
'anything',  
'straight',  
'curable',  
'cure.',  
'work',  
'hiv',  
'#otd',  
'years',  
'ago,',  
'terry',  
'fox',  
'passed',  
'away',  
'new',  
'westminister,',  
'bc.',  
'speaking',  
'run,',  
'said',  
'dreamer,',  
'saying',  
'initiate',  
'kind',  
'definitive',  
'answer',  
'believe',  
'miracles.',

'to."',  
'watch',  
'#heritageminute',  
'here.',  
'https://t.co/oqonptqcs',  
'#predict',  
'&',  
'#cure',  
'next',  
'#cancer',  
'https://t.co/gyztoi46yi',  
'minutes',  
'#cancercure',  
'#healed',  
'#reiki',  
'#rtrrt',  
'#bigpharma',  
'#loa',  
'#fridayreads',  
'#nlp',  
'#acim',  
'https://t.co/uqx20hdk8c',  
'https://t.co/oqh0emyqpk',  
'home',  
'#academiczodiac',  
'#o5iris',  
'predicta...',  
'https://t.co/jvbdy7arqy',  
'https://t.co/4r1zgo8opo',  
'https://t.co/ojby25xlyk',  
'#righttotry',  
'1st',  
'lyrics',  
'trump',  
'released?',  
'give',  
'vote',  
'-not',  
'hillary',  
'dr.',  
'video=',  
'blocked',  
'bill',  
'clinton',  
'since',  
'94',  
'#drburzynskirhymeswithlewinski',  
'#trumptrainjumponit',



' https://t.co/6odiawuwga',  
'via',  
'@youtube',  
'https://t.co/0uta7ccgwk',  
'poor',  
'clinton!',  
'decades',  
'suppressing',  
'cancer-cures',  
'wants',  
'#righttocry',  
'#righttodie',  
'ask',  
'patients',  
'that's',  
'right!',  
'rhymes',  
'mistress',  
'-both',  
'f\*cked',  
'https://t.co/xpikezppcz',  
'#shocking',  
'#government',  
'#petitions',  
'#dogs',  
'get',  
'signatures',  
'life',  
'saving',  
'funding',  
'terminal',  
'diseases',  
'known',  
'cause',  
'kill',  
'them! ',  
'#ipf',  
'#pulmonaryfibrosis',  
'needs',  
'like',  
'#cancer!',  
'#humanity?',  
'show',  
'care ',  
'https://t.co/zzcp7rnld0',  
'#sisterscare:',  
'cbs2's",  
'@burrelltv,',

'@efinchtv',  
'@danatylertv',  
'took',  
'part',  
'sisters',  
'event',  
'today,',  
'addressing',  
'disparities',  
'diagnosis',  
'treatment',  
'breast',  
'african',  
'american',  
'women.',  
'https://t.co/unjsqhcycd6',  
'#teachersleading',  
'come',  
'doctor',  
'revolution',  
'teaching',  
'going',  
'teachers',  
'@nickdonohuenmef',  
'#teamhalon',  
'supports',  
'national',  
'#randomactsoflightday',  
'stands',  
'@llsusa',  
'continued',  
'#cancer.',  
'learn',  
'https://t.co/q2lt7y56km',  
'#trump',  
'eradicate',  
'poverty',  
'still',  
'msm',  
'pull',  
'put',  
'negative',  
'spin',  
'it.',  
'war',  
'conflict',  
'preferable',  
'wrong',

'consistently',  
'are.',  
'#trumpkimsummit',  
'#maga',  
'discover',  
'#trumpderangementsyndrome',  
'sufferers',  
'screaming',  
'injection',  
'stinging.#liberalhypocrisy',  
'#singaporesummit',  
'#trumpkim',  
'#deepstate',  
'#thestorm',  
'self-inflicted',  
'metastasized',  
'full',  
'blown',  
'psychological',  
'insanity',  
'outbreak',  
'across',  
'democrat',  
'party,',  
'liberal',  
'media',  
'nevertrumpers',  
'lead',  
'ultimate',  
'demise.',  
'time',  
'let',  
'go.',  
'#walkaway',  
'\$agen',  
'ceo',  
'@garo\_armen',  
'@agenus\_bio',  
'#wearwhite',  
'@cancerresearch',  
'commitment',  
'fight',  
'#immunotherapy',  
'#oncology',  
'#curecancer',  
'https://t.co/f5vigul6ky',  
'proud',  
'pioneer',

'#immunotherapy',  
'@garo\_armen',  
'stand',  
'courageous',  
'physicians',  
'#inpursuitof',  
'https://t.co/ve97oxqakw',  
'innovation',  
'engine',  
'produced',  
'inds',  
'18mos;',  
'planned',  
'2018',  
'1h2019.',  
'molecules',  
'discovery',  
'ind',  
'yrs.',  
'plus',  
'&gt;100',  
'pts',  
'treated',  
'agen1884',  
'&or',  
'agen2034',  
'#asco2018.',  
'speed',  
'action!',  
'https://t.co/vrwjomjmvo',  
'driven',  
'powered',  
'right',  
'combination',  
'science',  
'people-see',  
'pursuing',  
'passion',  
'advancing',  
'toward',  
'goal',  
'story:',  
'https://t.co/apxu2dplkm',  
'@big3biobos',  
'\$celz',  
'imho',  
'beaten',  
'many',

'countries.;',  
'rate',  
'95%+'.',  
'#fda',  
'big',  
'#pharma',  
'keep',  
'treatments',  
'illegal',  
'u.s',  
'capitalize',  
'\$150,000-\$500,000',  
'unneeded',  
'surgeries',  
'chemo',  
'celz's',  
'#amniostem',  
'help',  
'#cachexia',  
'huge!',  
'&gt;',  
'advice',  
'isn't',  
'helpful.',  
'i'm',  
'sure',  
'it's',  
'wanting',  
'helpful,',  
'sucks.',  
'good',  
'deeds',  
'faith',  
'quinoa',  
'hope',  
'impacted',  
'worse,',  
'none',  
'around',  
'surgery',  
'days.',  
'&gt;"the',  
'idea',  
'tell',  
'story',  
'lesnar',  
'looking',  
'leaving',

'belt',  
'ufc"',  
'garner',  
'sympathy',  
'reigns.',  
'roman',  
'reigns',  
'end',  
'hunger,',  
'find',  
'booed',  
'pushed',  
'hard.',  
'https://t.co/uxful5sx6q',  
'\*breaking',  
'news\*',  
'(hypothetical',  
'situation)',  
'finds',  
'cancer!',  
'everyone',  
'mother:',  
'"what',  
'piece',  
'shit!',  
'probably',  
'knew',  
'whole',  
'time!"',  
'dont',  
'agree',  
'shit',  
'guy',  
'does/says,',  
'ya'll",  
'never',  
'satisfied',  
'\*opens',  
'cab-booking',  
'app\*',  
'cab',  
'\*books',  
'cab\*',  
'years,',  
'months,',  
'weeks,',  
'days',  
'hours.',

'kids',  
'born',  
'grow',  
'knowing',  
'mother,',  
'scientists',  
'volcano',  
'erupt',  
'wipe',  
'hawaii',  
'away.',  
'\*this',  
'chimes\*',  
'positivity',  
'can't',  
'neither',  
'fighting',  
'talk.',  
'"battle"',  
'narrative',  
'created',  
'facing',  
'disease',  
'denied',  
'narratives.',  
'couldn't',  
'more\*',  
'#lifewithcancer',  
'https://t.co/akj8zdd3jz',  
'@abbvie',  
'alphabet-backed',  
'calico',  
'spend',  
'another',  
'\$1b',  
'https://t.co/a7tqwrwtrv',  
'@erictrump:',  
'"we're',  
'raising',  
'money',  
'we're',  
'win',  
'fight.',  
'pediatric',  
'cancer."',  
'https://t.co/7pyfcd7xzy',  
'thank',  
'@erictrump',

'son',  
'alive',  
'due',  
'started',  
'jude.',  
'year',  
'survivor',  
'week.',  
'loose',  
'least',  
'every',  
'month',  
'local',  
'children's',  
'hospital',  
'childhood',  
'https://t.co/gwqtxflqau',  
'11:11',  
'excellent',  
'grades!!!',  
'job',  
'best',  
'friend',  
'happiness',  
'friends',  
'gun',  
'control',  
'@arianagrande',  
'notice',  
'@shawnmendes',  
'follow',  
'bye',  
'anxiety',  
'fun',  
'summer',  
'lion',  
'rid',  
'depression',  
'hug',  
'love',  
'grades',  
'summer!',  
'harry',  
'potter',  
'reunion',  
'weight!',  
'16-year-old',  
'jeremiah',



'thomas,',  
'battling',  
'used',  
'make',  
'wish',  
'abortion.',  
'you,',  
'jeremiah',  
'beautiful',  
'life,',  
'inspiring',  
'many.',  
'https://t.co/bumzmggyay',  
'sad',  
'young,',  
'brainwashed',  
'maybe',  
'health',  
'care',  
'all,',  
'imaginary',  
'exist',  
'https://t.co/27twcb3oek',  
'"selfless',  
'way"',  
'idk',  
'mann',  
'pretty',  
'selfish',  
'try',  
'use',  
'religion',  
'controls',  
'someone',  
'else's',  
'body...',  
'coulda',  
'wished',  
'https://t.co/tyx3hynfa8',  
'17yr',  
'stage',  
'carcinoid',  
'survivor.',  
'longest',  
'surviving',  
'patient',  
'world.',  
'survival',

'5yrs.',  
'plz',  
'read',  
'remarkable',  
'miracle',  
'book',  
'living',  
'silent',  
'enemy',  
'coming',  
'soon.',  
'support',  
'mdacc.',  
'inoperable',  
'neuroendocrine',  
'misdiagnosed',  
'lets',  
'cancer#',  
'17yr.',  
'endcancer#',  
'1965:',  
'bet',  
"there'll",  
'flying',  
'cars',  
'future"',  
'1987:',  
'future',  
'2018:',  
'prototype',  
'potentially',  
'viable',  
'cures',  
'cancers',  
'given',  
'1998:',  
"in",  
'twenty',  
'cancer! "',  
"dennis",  
'rodman',  
'saved',  
'nuclear',  
'celebrity',  
'apprentice,',  
'hosted',  
'united',  
'states,',

'dictator',  
'north',  
'korea! "',  
'@panmass',  
'rider',  
'@danafarber.',  
'fund',  
'retweet!',  
'https://t.co/qdwqn3zzre',  
'https://t.co/lafbipfydc',  
'3mt',  
'semi-finalist',  
'ubc',  
'master',  
'student',  
'evan',  
'warner',  
'presented',  
"shakespeare's",  
'much',  
'thing?',  
'asked.',  
'#tbt',  
'talk:',  
'https://t.co/vjazjwo5em',  
'22,',  
'ethan',  
'hawes',  
'diagnosed',  
'multiple',  
'myeloma,',  
'typically',  
'afflicting',  
'age',  
'65.',  
'hospitalizations,',  
'chemotherapy,',  
'stem',  
'cell',  
'replacement',  
'followed.',  
'https://t.co/f0lztvodcb',  
'5,000',  
'young',  
'women',  
'country',  
'participated',  
'year's',

'lax',  
'raise',  
'funds',  
'awareness',  
'tournament',  
'raised',  
'\$235,000',  
'komen',  
'csnj',  
'@komennja.',  
'more:',  
'https://t.co/4qlw3bei7k',  
'https://t.co/uwdnketoie',  
'6th',  
'40tude',  
'thames',  
'sup',  
'marathon',  
'saturday',  
'already',  
'£63k',  
'marks',  
'colon',  
'add',  
'£180k',  
'past',  
'years.',  
'miles',  
'sups',  
'finishing',  
'putney.',  
'@curebowelcancer',  
'@active360',  
'https://t.co/dn5cjvm9gz',  
'87%',  
'f-ing',  
'republicans',  
'trump.',  
'omg.',  
'how?',  
'children',  
'ripped',  
'mother's',  
'arms.',  
'truly',  
'80%',  
'news',  
'20%',

'don't',  
'on.',  
'#disgusted',  
'must',  
'accept',  
'sinister',  
'brainwashing.',  
'comprehend',  
'maga',  
'grasp',  
'shooting',  
'afternoon',  
'nature.',  
'one',  
'voting',  
'overwhelming',  
'numbers.',  
'https://t.co/s7ctd0lyc6',  
'@1405scituateave',  
'@realjameswoods',  
'funny',  
'is',  
'discovered',  
'tomorrow',  
'donald',  
'telling',  
'somehow',  
'instrumental',  
'it',  
'jackass',  
'trumpuppets',  
'word',  
'it!!!!',  
'...and',  
'you're',  
'saying;',  
'"maybe',  
'be"',  
'now!!',  
'@1776stonewall',  
'pathetic',  
'wasn't',  
'enough!!!',  
'talk',  
'pitiful',  
'souls',  
'way',  
'trump!!!!',

'@1999mcm11',  
'denuclearization',  
'achieved.',  
'gets',  
'taste',  
'western',  
'culture,',  
'mcdonalds,',  
'coke,',  
'xbox,',  
'etc...',  
'over.',  
'achieve',  
'peace...',  
'complain',  
'him.',  
'cuz',  
'he's',  
'orange',  
'tweets',  
'stupid',  
'stuff.',  
'@2mattandgaby2',  
'no.',  
'laetrile',  
'else.',  
'studied,',  
'work.',  
'stop',  
'spreading',  
'dangerous',  
'bullshit.',  
'https://t.co/7z0jpa3osp',  
'@420stonercorner',  
'.@420stonercorner',  
'fascinating',  
'subject',  
'matter',  
'thought',  
'pot',  
'heads',  
'such,',  
'rampant',  
'family',  
'cbd',  
'thc',  
'instead',  
'morphine!'

'#cannabis',  
'#weeds',  
'#pmqs',  
'medically',  
'unqualified',  
'opinion?',  
'#theresamay',  
'@weedseeds\_uk',  
'@ukcannabisclub',  
'@ukcannabisclubs',  
'@younglabouruk',  
'#bbcqt',  
'#bbcaq',  
'assholes',  
'criminalizing',  
'@weedfeed',  
'@norml',  
'#nyc',  
'#nypd',  
'@ukleap',  
'@uklabour',  
'@weednucleus',  
'cannabis',  
'communion',  
'brings',  
'great',  
'forever!',  
'https://t.co/f583xnt3jy',  
'@\_cape\_',  
'@\_\_hunchohoudini',  
'provide',  
'food',  
'shelter',  
'homeless',  
'can't',  
'aids',  
'although',  
'it)',  
'shot',  
'tupac.we',  
'investigators',  
'support.',  
'swear',  
'better',  
'sometimes',  
'always',  
'dumb',  
'ppl',

'fuck',  
'@\_duckyfresh',  
'@babygrotesque',  
'underlying',  
'economic',  
'problems',  
'contribute',  
'gotham's',  
'crime.',  
'beating',  
'mentally',  
'ill',  
'throwing',  
'inhumane',  
'caricature',  
'mental',  
'facility.',  
'bruce',  
'literally',  
'spidey',  
'meme',  
'tea',  
'@\_kingofthehill1:',  
'sexual',  
'https://t.co/qt6n6wrqse',  
'@aacrfoundation:',  
'leading',  
'discoveries,',  
'targeting',  
'lives.',  
'see',  
'helping',  
'@abaabood',  
'@andreas\_krieg',  
'me,',  
'spared',  
'efforts',  
'last',  
'plus',  
'resolve',  
'brotherly',  
'basis',  
'qatar',  
'continues',  
'path',  
'destabilizing',  
'saudi',  
'arabia',



'region.',  
'surgically',  
'remove',  
'regime.',  
'@abby\_higgins5:',  
'government/pharmacies',  
'wanna',  
'money...',  
'idc',  
'https://t.co/ivompefmo0',  
'@ac\_breezy',  
'viruses,',  
'spam/trolls',  
'exact',  
'thing...',  
'yet',  
'enough',  
'realize',  
'earned',  
'millennium',  
'prize',  
'couple',  
'hundred',  
'times',  
'think',  
'https://t.co/29f4udvenv',  
'@acosta',  
'real.',  
'left',  
'vitriol',  
'hatred',  
'wrong.',  
'pity',  
'miserable',  
'people.',  
'@adeabdwbb',  
'tweak',  
'question',  
'little',  
'bit',  
'imagine',  
'youre',  
'dying.',  
'rich',  
'man',  
'giving',  
'choices:',  
'huge',

```
'amount',
'money.',
'basically',
'cant',
'god',
'choose',
'bcause',
'choose.',
'@adevotedyogi:',
...]
```

```
[ ]: # Classify test data
filtered_test_data_list = filtered_test_data.collect()
results = []
```

```
[ ]: display(word_probabilities.collect())
```

```
[ ]: wp = word_probabilities.collect()
```

```
[ ]: print(wp[0][2])
```

0.11602209944751381

```
[ ]: from math import log
epsilon = 1e-10
for query_list in filtered_test_data_list:
    # print(query_list[0])
    query = query_list[1]
    yes_probability = log(1) # Initialize with log(1) to avoid issues with
    ↪ zero probabilities
    no_probability = log(1)
    found = False
    for word in query:
        if word in list_of_words:
            found = True
            word_prob = word_probabilities.filter(lambda x: x[0] == word).
            ↪ collect()
            # print(word_prob)
            # if (word_prob[0][1] == 0 or word_prob[0][2] == 0):
            #     print(word_prob)
            if word_prob[0][1] > 0:
                yes_probability += log(word_prob[0][2])
            if word_prob[0][2] > 0:
                no_probability += log(word_prob[0][1])

    if not found:
        print("Not found")
        results.append((query, "no", query_list[1]))
```

```

else:
    # print(yes_probability, no_probability)
    if yes_probability > no_probability:
        # print(query)
        # print(yes_probability)
        if(query_list[0] == "no"):
            print(yes_probability, no_probability)
            result = (query, "yes", query_list[0])
            results.append(result)
        else:
            # print(query)
            # if(query_list[0] == "yes"):
            #     print(yes_probability, no_probability)
            result = (query, "no", query_list[0])
            results.append(result)

```

```

[ ]: # Create an RDD with the results (1st column: text, 2nd column: prediction, 3rd
    ↪column: label)
results_rdd = sc.parallelize(results)

# Total number of records
total_records = results_rdd.count()
# Correctly predicted records
correct_predictions = results_rdd.filter(lambda x: x[1] == x[2])

# True Positives (TP), True Negatives (TN), False Positives (FP), False
    ↪Negatives (FN)
tp = results_rdd.filter(lambda x: x[1] == 'yes' and x[2] == 'yes').count()
tn = results_rdd.filter(lambda x: x[1] == 'no' and x[2] == 'no').count()
fp = results_rdd.filter(lambda x: x[1] == 'yes' and x[2] == 'no').count()
fn = results_rdd.filter(lambda x: x[1] == 'no' and x[2] == 'yes').count()
# Calculate accuracy
accuracy = (tp + tn) / total_records * 100

# Print the results
print("Accuracy:", accuracy)

```

Accuracy: 88.75968992248062

```

[ ]: # Priori of the classes
print("The prior probability of yes tweets", p_yes)
print("The prior probability of no tweets Texts", p_no)

```

The prior probability of yes tweets 0.14022435897435898

The prior probability of no tweets Texts 0.8589743589743589

```

[ ]:

```