DATA ANALYSIS TO FIGURE OUT THE DISTRIBUTION OF ENGLISH WORDS

In [1]:
```python
import pandas as pd
import numpy as np
```

In [2]:
```python
# Read the table of dataset 1
df1 = pd.read_csv('words_fre.csv')
```

In [3]:
```python
# Read the table of dataset 2
df2 = pd.read_csv('Phrases_Fre.csv')
```
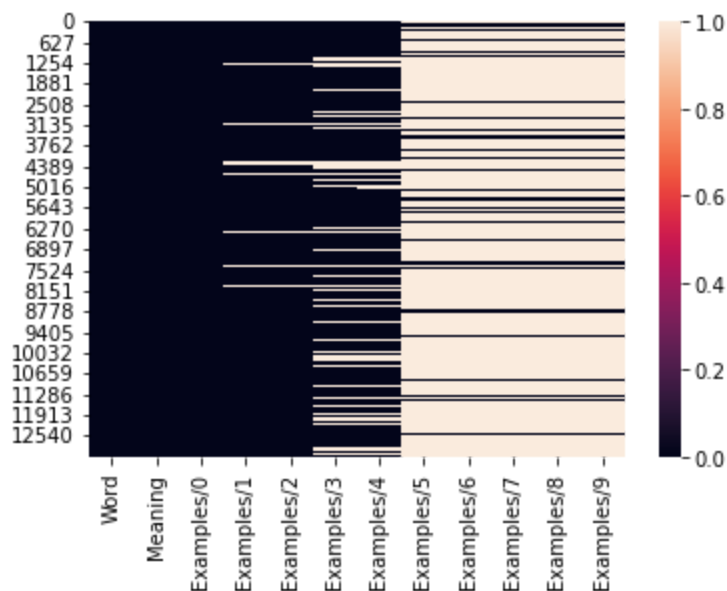
1. Check the data

In [4]:
```python
df2.isnull().sum()
```

Out[4]:
```
Word              0
Meaning          16
Examples/0        0
Examples/1      375
Examples/2      382
Examples/3     1721
Examples/4     1733
Examples/5    11475
Examples/6    11475
Examples/7    11475
Examples/8    11475
Examples/9    11479
dtype: int64
```

In [5]:
```python
import seaborn as sns
sns.heatmap(df2.isnull())
```
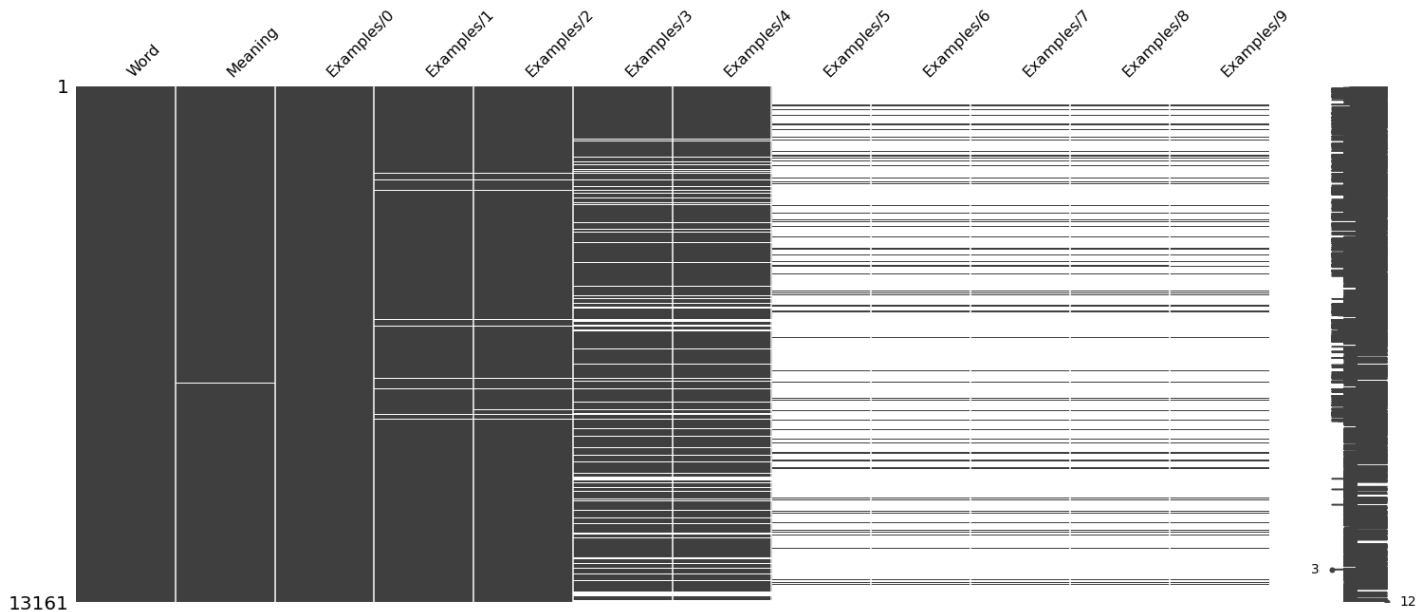
Out[5]:
```
<AxesSubplot:>
```


In [6]:
```python
!pip install missingno
```

In [7]:
```python
import missingno as msno
msno.matrix(df2)
```

Out[7]: `<AxesSubplot:>`

In [8]:
```python
msno.heatmap(df2)
```

Out[8]: `<AxesSubplot:>`

Meaning

	Meaning	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5	Examples/6	Examples/7	Examples/8	Examples/9
Examples/1										
Examples/2		<1								
Examples/3		0.4	0.5							
Examples/4		0.4	0.4	1						
Examples/5	-0.1	0.1	0.1	0.1	0.1					
Examples/6	-0.1	0.1	0.1	0.1	0.1	1				
Examples/7	-0.1	0.1	0.1	0.1	0.1	1	1			
Examples/8	-0.1	0.1	0.1	0.1	0.1	1	1	1		
Examples/9	-0.1	0.1	0.1	0.1	0.1	1	1	1	1	
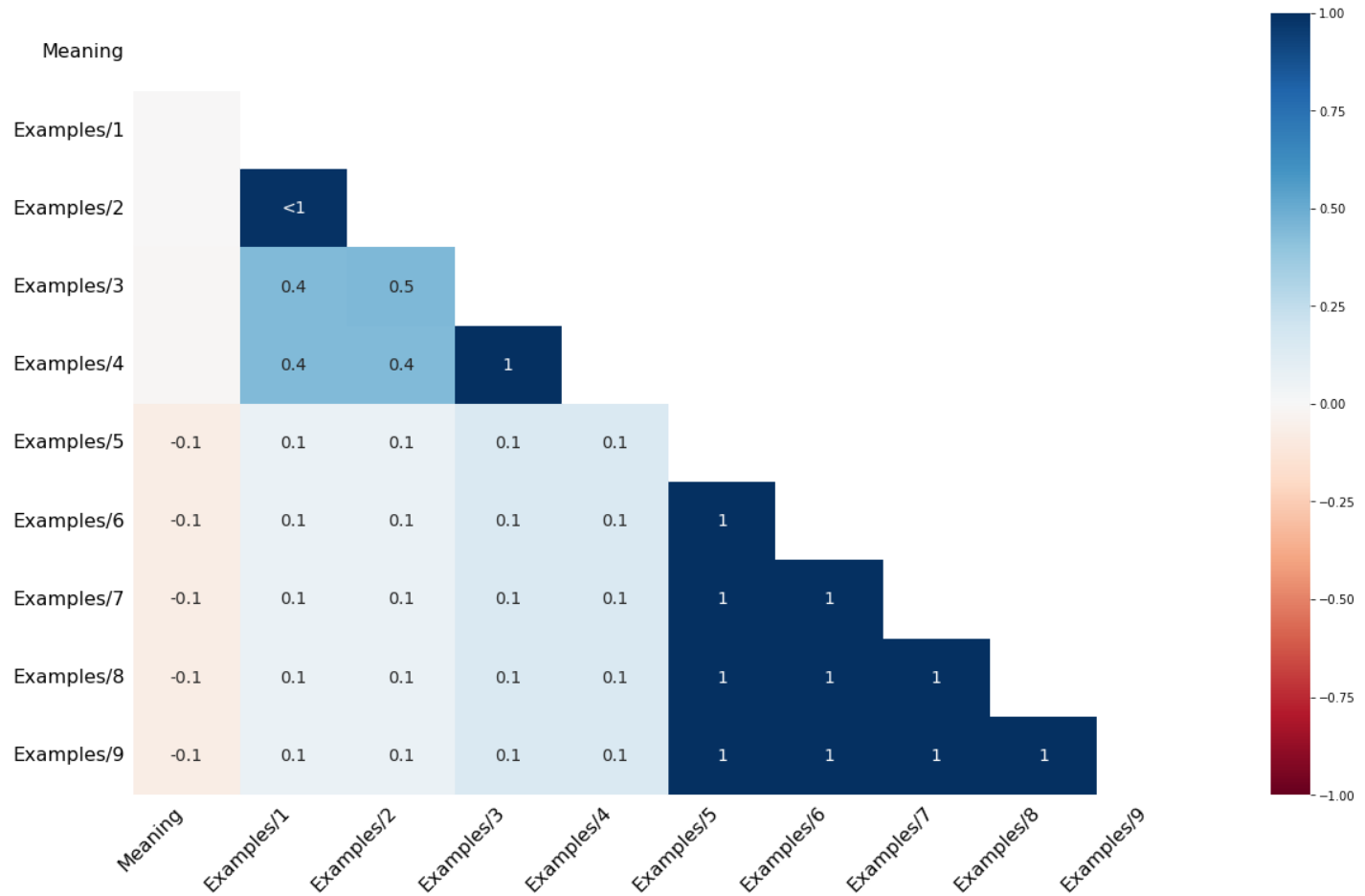
```
In [9]:   df1.head(100)
```

Out[9]:

	word	count
0	the	125971793511
1	of	77466218166
2	and	55607010885
3	to	46622823631
4	in	40069476427
...	...	...
95	me	1567968527
96	between	1545114434
97	those	1532715465
98	same	1530696473
99	now	1520107899

100 rows × 2 columns

```
In [10]:   df2.head(20)
```

Out[10]:

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5
0	A bed of roses	a phrase used to reference a situation that is...	The poor orphan's life hasn't been a bed of ro...	Because Mark studied for many weeks, taking th...	The stressed business owner quickly realized t...	NaN	NaN	NaN
1	A Priori	something that is believed to be true without ...	Religious people have the a priori belief that...	The jaded woman made a priori assumptions that...	Christopher Columbus had the a priori belief t...	People make a priori assumptions that the sun ...	Lawyers use a priori arguments about their def...	NaN
2	Aardvark	a large, long-eared nocturnal animal that uses...	The long-eared aardvark slept all day and ate ...	With its rabbit like ears and piglike snout, t...	Living a solitary nightlife, the African aardv...	The aardvark is a snouty mammal that dines hea...	Though it is a small mammal, the closest relat...	NaN
3	Aback	in amazement; shocked	I was taken aback by the server's rude comment...	While Amy enjoyed the film as a whole, as a mi...	Ken believed he had a great marriage so he was...	At seventy years of age, Wanda was taken aback...	Although I had been exercising daily and eatin...	NaN
4	Abaft	toward the stern; behind	The wind was abaft which made the ship glide f...	They placed the lifeboats in the rear part of ...	My first sailing lesson was learning that abaf...	The sea spray was coming from abaft, making my...	I was told to sit abaft to lessen my seasickne...	NaN
5	Abandon	to stop caring and leave or to desert	Since the young girl was unable to care for th...	Carl rarely commits to anything so everyone ex...	During the snow storm, many people made the de...	The military police is currently looking for t...	Because this plan is not feasible, we need to ...	NaN
6	Abase	to lower so as to hurt feelings	My stepmother is an evil woman who likes to ab...	Because Mark hated himself, he often tried to ...	I will not remain married to a man who tries t...	Although I need money badly, I will not abase ...	My jealous sister tried to abase me by making ...	When my brother behaved badly, my mother would...
7	Abaser	someone who degrades or lowers someone	After the abaser would make fun of her daughte...	When everyone in the grocery store heard the a...	Even the abaser couldn't think of something me...	People who have low self-esteem will likely be...	Usually the classmates would avoid the abaser ...	NaN
8	Abash	to humiliate someone and make them feel embarr...	The officer tried to abash the inmates, humili...	Daniel's tendency to abash his wife left her e...	Though he claims he didn't mean to abash his b...	Ross vowed to humiliate his opponent and abash...	Because he doesn't have any empathy, the evil ...	NaN
9	Abate	to reduce in degree or intensity	I hope this medicine will abate the pain in my...	In time, his anger will abate, and he will for...	With any luck, the storm will abate soon so we...	The firemen sprayed water on the burning house...	After my massage, I felt my stress level abate...	Even though the unemployment rate continues to...

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5
10	Abatement	reduction or weakening or something	Natalie longed for the abatement of the scream...	The abatement of the prisoner's sentence anger...	The college students hoped for the abatement o...	The abatement of her headache filled Jennifer ...	Despite the fact that the joke finished over a...	NaN
11	Abattoir	a place where animals are taken to slaughter	Although many farmers own cattle, they don't s...	Butchers get veal, lamb, and other meats from ...	After they fattened the pig for several months...	The Kosher abattoir forbids pork on the premis...	The abattoir had several buildings labeled for...	NaN
12	Abbess	the female superior of a community of nuns	An Abbess is the female leader of a small comm...	You have to be a nun for a certain amount of t...	An Abbess is voted into office by secret votes...	A woman of illegitimate birth, who is not a vi...	The sisters looked to their Abbess for encoura...	NaN
13	Abbreviate	to shorten a word or phrase	Learning to abbreviate long words will help th...	Please do not abbreviate your name, but instea...	Some packages of food abbreviate teaspoon to t...	When writing the date, you can abbreviate Sept...	Because we tend to abbreviate while texting, i...	NaN
14	Abdicate	to stop accepting a particular responsibility ...	The young prince became king after his older b...	Because the mother chose to abdicate her role ...	Of course, it is easier to abdicate your respo...	Since the dictator of that country will not si...	When his company's stock price dropped under a...	Because of a sex scandal, the president had no...
15	Abdominal	related to the abdomen/stomach, or the part be...	My severe abdominal pains sent me to the hospi...	A hole in Misty's abdominal wall means that sh...	Wearing an abdominal girdle, the dancer was ab...	Doing abdominal crunches, the soldier hoped to...	Holding on to her abdominal area, the little g...	NaN
16	Abduct	to grab an individual by trickery or force	The kidnappers planned to abduct the millionai...	As a prank, members of our fraternity usually ...	The child predator hangs out in the park looki...	Since the police will not help her, Helen is g...	There is a nationwide search for the woman who...	NaN
17	Abecedarian	alphabetical	We decided to organize the files in an abeceda...	When something is abecedarian, it is alphabeti...	I like to put my books in abecedarian order, b...	Many people choose to sort things alphabetical...	If you sort things with the abecedarian method...	NaN
18	Aberrant	straying from the normal or right way	John's aberrant behavior is going to get him i...	When the astronomer looked into the telescope,...	If my aunt does not take her medication, she w...	Sarah's aberrant manners led to her being kick...	My unusual uncle has an aberrant way of tying ...	Because my daughter has aberrant eating habits...

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5
19	Aberration	deviating from what is normal or desirable, no...	Shelley's angry retort was an aberration from ...	Since I did not properly adjust my camera sett...	If your goal is to reduce the likelihood of an...	A person with one blue eye and one green eye i...	Even though Janice has a medical aberration wh...	The presence of cancer cells under a microscop...

In [11]:
```python
df2.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 13161 entries, 0 to 13160
Data columns (total 12 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   Word        13161 non-null  object
 1   Meaning     13145 non-null  object
 2   Examples/0  13161 non-null  object
 3   Examples/1  12786 non-null  object
 4   Examples/2  12779 non-null  object
 5   Examples/3  11440 non-null  object
 6   Examples/4  11428 non-null  object
 7   Examples/5  1686 non-null   object
 8   Examples/6  1686 non-null   object
 9   Examples/7  1686 non-null   object
 10  Examples/8  1686 non-null   object
 11  Examples/9  1682 non-null   object
dtypes: object(12)
memory usage: 1.2+ MB
```

In [12]:
```python
df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 313282 entries, 0 to 313281
Data columns (total 2 columns):
 #   Column  Non-Null Count   Dtype
---  ------  --------------   -----
 0   word    313280 non-null  object
 1   count   313282 non-null  int64
dtypes: int64(1), object(1)
memory usage: 4.8+ MB
```

Process Steps

Cleaning data and correct the data formats
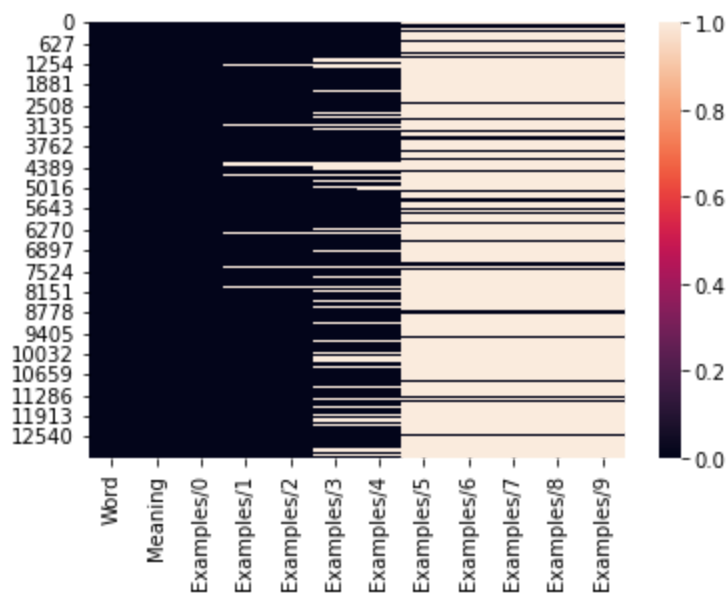
In [13]:
```python
# import pyplot and seaborn libraries
from matplotlib import pyplot as plt
import seaborn as sns
```

In [14]:
```python
# the best way to find null values
sns.heatmap(data=df2.isnull())
```

Out[14]:
```
<AxesSubplot:>
```

In [15]:
```python
df2.isnull().sum()
```

Out[15]:
```
Word             0
Meaning         16
Examples/0       0
Examples/1     375
Examples/2     382
Examples/3    1721
Examples/4    1733
Examples/5   11475
Examples/6   11475
Examples/7   11475
Examples/8   11475
Examples/9   11479
dtype: int64
```

In [16]:
```python
# Lower strings in Word column
df2['Word']=df2['Word'].str.lower()
```

In [17]:
```python
df2[df2['Word'].str.contains('a ')]
```

Out[17]:

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5
0	a bed of roses	a phrase used to reference a situation that is...	The poor orphan's life hasn't been a bed of ro...	Because Mark studied for many weeks, taking th...	The stressed business owner quickly realized t...	NaN	NaN	NaN
1	a priori	something that is believed to be true without ...	Religious people have the a priori belief that...	The jaded woman made a priori assumptions that...	Christopher Columbus had the a priori belief t...	People make a priori assumptions that the sun ...	Lawyers use a priori arguments about their def...	NaN
372	aha moment	a moment of sudden insight	The scientist had an aha moment and figured ou...	Even though the breakup was painful, it was an...	While working on the difficult math problem, t...	NaN	NaN	NaN

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5
447	alma mater	the school or college that one attended	My mother wanted me to go to her alma mater, b…	Cole and every male graduate in his family are…	Because my alma mater will give me a discount,…	The high school senior hoped to work for his a…	Honored to teach at the college he attended, P…	NaN
813	as timid as a mouse	describes someone that is easily frightened or…	Charlotte has always been someone that was as …	Most young children are extremely outgoing and…	Someone who is as timid as a mouse is likely s…	If you have a friend that is as timid as a mou…	I used to be as timid as a mouse, but as I got…	NaN
952	aura	a unique sensation or feeling that stems from …	The angel's aura made me feel calm and hopeful.	Because Jim has an aura of negativity surround…	Gina uses a large number of herbal oils to spr…	As I stepped inside the temple, I was embraced…	Under the spotlight, the star's aura seemed ev…	NaN
999	aura	a unique sensation or feeling that stems from …	The angel's aura made me feel calm and hopeful.	Because Jim has an aura of negativity surround…	Gina uses a large number of herbal oils to spr…	As I stepped inside the temple, I was embraced…	Under the spotlight, the star's aura seemed ev…	NaN
1452	bona fide	real or done honestly	The jeweler stated the large diamond was bona …	Because the man had lied about having a medica…	To make sure you are a bona fide buyer, the re…	Unlike the sneaky man at the pawn shop, the ba…	When she married the prince, the commoner beca…	The results of the test were not bona fide bec…
1453	bonanza	a large amount of a thing that is viewed as de…	Once I was a lonely girl, but social media has…	Valentine's Day produces a bonanza of wealth f…	Because football is extremely popular, the Sup…	The reality star knows posting a nude selfie w…	Although we received a bonanza of donations fo…	NaN
5675	iota	a tiny amount	If there is even one iota of doubt, the jury s…	My father says his love for my mother has not …	Although the suspect said he was out of town a…	As a starving writer, all I need is for one pu…	Since Janet had repeatedly warned her son to s…	NaN
5992	lingua franca	a language that is adopted as a common languag…	The Lingua Franca of the country with two fact…	Because the new city was settled by two differ…	Many countries that have two populations of di…	Whenever one country conquers another, that co…	The Lingua Franca of our country is a mix of a…	NaN
6153	magna carta	a formal document signed by King John in 1215 …	By signing the Magna Carta, King John of Engla…	King John's Magna Carta of 1215 was an importa…	After the Magna Carta was signed, the king of …	The Magna Carta of 1215 established constituti…	Signed during the middle ages, the Magna Carta…	NaN
6240	mantra	a phrase or term repeated often that usually e…	As physicians we observe the mantra of "not do…	Cara repeats her mantra about making each day …	When Kurt meditates, he repeats a calming mant…	The Christian man lived by the mantra of becom…	As the small man struggled to move the giganti…	NaN

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5
6311	mea culpa	an acknowledgement of having done wrong (my bad)	The magazine issued a mea culpa after they cit…	After the scandal in the White House, the pres…	Even after the pop singer's mea culpa for his …	The governor's formal mea culpa for stealing f…	If I wanted an allowance, I needed to issue a …	NaN
6927	nirvana	a place of peace and happiness	The monk urged me to sit and meditate until I …	Under the psychologist's care, I was able to l…	The painkillers sent Jason to nirvana and allo…	As the old woman took her final breath, she di…	Before you can enter nirvana, you must first f…	NaN
7111	nirvana	a place of peace and happiness	The monk urged me to sit and meditate until I …	Under the psychologist's care, I was able to l…	The painkillers sent Jason to nirvana and allo…	As the old woman took her final breath, she di…	Before you can enter nirvana, you must first f…	NaN
7951	persona non grata	someone who is unwanted or blacklisted for som…	Alex threw a fit when her former best friend s…	After screaming at the barista and threatening…	Aiden became a persona non grata at family din…	After her troubled cousin stole several hundre…	The senator was considered a persona non grata…	NaN
8142	plasma	the liquid element of blood that transports bl…	As a starving student, Clint often sold his pl…	Ryan is an assistant in a laboratory that eval…	Although we bleed red when we are hurt, the pl…	Did you know liquid plasma is the largest comp…	What is the simplest way to retrieve red blood…	NaN
8156	plea bargain	to agree to plead guilty in return for a lesse…	In order to avoid a long prison sentence, the …	NaN	NaN	NaN	NaN	NaN
8521	prima facie	accepted as true and correct until proven othe…	In court, the story of the plaintiff is Prima …	A gullible person may accept the story of some…	Someone who thinks everything should be taken …	In science, people rarely accept anything Prim…	It would be very easy to trick people if every…	NaN
10124	shake a leg	a phrase used to tell someone to start or do s…	Grandma told the boys to 'shake a leg' and enc…	"Shake a leg!" Erin called as her brother sat …	When urging them to finish their assignments q…	NaN	NaN	NaN
11075	subpoena	a legal document requiring an individual to ap…	As soon as I received the subpoena, I knew I h…	Will they arrest my sister if she ignores the …	If you ignore the judge's subpoena to appear i…	The subpoena will force the unwilling witness …	When the subpoena is issued, James will have n…	NaN
11557	terra firma	dry land, distinguishable from sea and air	The beach is Terra Firma, its sand being insti…	Terra Firma is any type of terrain in this wor…	An island is but one small piece of Terra Firm…	A large continent is no different from a small…	All humans live on Terra Firma, since we need …	NaN

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5
12643	vice versa	with the order reversed	Of what use is a car without the key and vice ...	Because John loves Heather and vice versa, the...	Frank eats at the diner so often he knows all ...	During the dance, the boys talked about the gi...	Cats cannot survive where fish live and vice v...	NaN
12715	vis a vis	face to face; in person	I feel it is better to discuss personal issues...	The couple sat vis-à-vis at the table, holding...	A vis-à-vis confrontation led to a heated argu...	Even though I had met her vis-à-vis once befor...	The vis-à-vis encounter with a serial killer l...	NaN
12746	viva voce	By word of mouth (orally)	Because he could speak the answer but couldn't...	A viva voce project was assigned to the studen...	The professor believes in viva voce methods an...	NaN	NaN	NaN

```
In [18]:    # Check again
            df1.head(100)
```

Out[18]:

	word	count
0	the	125971793511
1	of	77466218166
2	and	55607010885
3	to	46622823631
4	in	40069476427
...	...	...
95	me	1567968527
96	between	1545114434
97	those	1532715465
98	same	1530696473
99	now	1520107899

100 rows × 2 columns

```
In [19]:    # drop 50 initial words
            df1=df1.drop(df1.index[:50]).reset_index(drop=True)
```

```
In [20]:    # Create the copies of two tables
            df1_copy =df1.copy()
            df2_copy = df2.copy()
```

```
In [21]:    # Check the string values in Word columns, then figured out the spacings in the values. Re
            df2_copy['Word']=df2_copy['Word'].apply(lambda x: x.replace(' ',''))
```

```
In [22]:    # Remove duplicates
```

```python
df2_copy=df2_copy.drop_duplicates()
```

In [23]:
```python
df2_copy
```

Out[23]:

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5	Exa
0	abedofroses	a phrase used to reference a situation that is...	The poor orphan's life hasn't been a bed of ro...	Because Mark studied for many weeks, taking th...	The stressed business owner quickly realized t...	NaN	NaN	NaN	
1	apriori	something that is believed to be true without ...	Religious people have the a priori belief that...	The jaded woman made a priori assumptions that...	Christopher Columbus had the a priori belief t...	People make a priori assumptions that the sun ...	Lawyers use a priori arguments about their def...	NaN	
2	aardvark	a large, long-eared nocturnal animal that uses...	The long-eared aardvark slept all day and ate ...	With its rabbit like ears and piglike snout, t...	Living a solitary nightlife, the African aardv...	The aardvark is a snouty mammal that dines hea...	Though it is a small mammal, the closest relat...	NaN	
3	aback	in amazement; shocked	I was taken aback by the server's rude comment...	While Amy enjoyed the film as a whole, as a mi...	Ken believed he had a great marriage so he was...	At seventy years of age, Wanda was taken aback...	Although I had been exercising daily and eatin...	NaN	
4	abaft	toward the stern; behind	The wind was abaft which made the ship glide f...	They placed the lifeboats in the rear part of ...	My first sailing lesson was learning that abaf...	The sea spray was coming from abaft, making my...	I was told to sit abaft to lessen my seasickne...	NaN	
...	...	...	...	...	...	...	...	...	...
13156	zodiac	a divided band in sky along which the sun, the...	Twelve cities in the state were named after as...	The zodiac signs are split over the twelve mon...	Zodiac symbols are based on the constellations...	Each of the zodiac signs covers an equal area ...	The zodiac system was based off of the sun, mo...	NaN	
13157	zone	an area or division	The zone only has houses and no stores or shop...	Walmart workers will place cleaning supplies i...	Before the team warmed up, all of the player's...	Because they live in a quake zone, the couple ...	People must drive slower in the construction z...	NaN	
13158	zoology	that portion of science that deals with the an...	Since Chamberlain had a fierce love for animal...	April used her zoology degree to help ranchers...	Through the school's zoology program, students...	Kids learn about all types of animals when stu...	In order to earn her degree in zoology, Rhonda...	NaN	
13159	zoom	to move very quickly	The cars began to zoom around the racetrack, m...	Police will need to monitor the fast-moving ca...	Reading eagerly, the avid reader can zoom thro...	Hurried shoppers looking for Christmas present...	The track star was able to zoom past all of th...	NaN	

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5	Exa
13160	zygote	the fertilized egg that results from the union...	A zygote is formed when a sperm cell and an eg...	After the zygote goes through a rapid period o...	The production of a zygote is dependent upon a...	Simply stated, a zygote is an egg that has bee...	When a female gamete and a male gamete fuse, t...	NaN	

13102 rows × 12 columns

```python
In [24]:  df_merge= df2_copy.merge(df1_copy, left_on='Word', right_on ='word', how='left')
```

```python
In [25]:  df_merge
```

Out[25]:

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5	Exa
0	abedofroses	a phrase used to reference a situation that is...	The poor orphan's life hasn't been a bed of ro...	Because Mark studied for many weeks, taking th...	The stressed business owner quickly realized t...	NaN	NaN	NaN	
1	apriori	something that is believed to be true without ...	Religious people have the a priori belief that...	The jaded woman made a priori assumptions that...	Christopher Columbus had the a priori belief t...	People make a priori assumptions that the sun ...	Lawyers use a priori arguments about their def...	NaN	
2	aardvark	a large, long-eared nocturnal animal that uses...	The long-eared aardvark slept all day and ate ...	With its rabbit like ears and piglike snout, t...	Living a solitary nightlife, the African aardv...	The aardvark is a snouty mammal that dines hea...	Though it is a small mammal, the closest relat...	NaN	
3	aback	in amazement; shocked	I was taken aback by the server's rude comment...	While Amy enjoyed the film as a whole, as a mi...	Ken believed he had a great marriage so he was...	At seventy years of age, Wanda was taken aback...	Although I had been exercising daily and eatin...	NaN	
4	abaft	toward the stern; behind	The wind was abaft which made the ship glide f...	They placed the lifeboats in the rear part of ...	My first sailing lesson was learning that abaf...	The sea spray was coming from abaft, making my...	I was told to sit abaft to lessen my seasickne...	NaN	
...	...	...	...	...	...	...	...	...	...
13097	zodiac	a divided band in sky along which the sun, the...	Twelve cities in the state were named after as...	The zodiac signs are split over the twelve mon...	Zodiac symbols are based on the constellations...	Each of the zodiac signs covers an equal area ...	The zodiac system was based off of the sun, mo...	NaN	
13098	zone	an area or division	The zone only has houses and no stores or shop...	Walmart workers will place cleaning supplies i...	Before the team warmed up, all of the player's...	Because they live in a quake zone, the couple ...	People must drive slower in the construction z...	NaN	

|	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5	Exa
13099	zoology	that portion of science that deals with the an...	Since Chamberlain had a fierce love for animal...	April used her zoology degree to help ranchers...	Through the school's zoology program, students...	Kids learn about all types of animals when stu...	In order to earn her degree in zoology, Rhonda...	NaN	
13100	zoom	to move very quickly	The cars began to zoom around the racetrack, m...	Police will need to monitor the fast-moving ca...	Reading eagerly, the avid reader can zoom thro...	Hurried shoppers looking for Christmas present...	The track star was able to zoom past all of th...	NaN	
13101	zygote	the fertilized egg that results from the union...	A zygote is formed when a sperm cell and an eg...	After the zygote goes through a rapid period o...	The production of a zygote is dependent upon a...	Simply stated, a zygote is an egg that has bee...	When a female gamete and a male gamete fuse, t...	NaN	

13102 rows × 14 columns

In [26]:
```python
# order data
df_merge.sort_values(by=['Word','count'], ascending =[True,False])
```

Out[26]:

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5	Examp
2	aardvark	a large, long-eared nocturnal animal that uses...	The long-eared aardvark slept all day and ate ...	With its rabbit like ears and piglike snout, t...	Living a solitary nightlife, the African aardv...	The aardvark is a snouty mammal that dines hea...	Though it is a small mammal, the closest relat...	NaN	
3	aback	in amazement; shocked	I was taken aback by the server's rude comment...	While Amy enjoyed the film as a whole, as a mi...	Ken believed he had a great marriage so he was...	At seventy years of age, Wanda was taken aback...	Although I had been exercising daily and eatin...	NaN	
4	abaft	toward the stern; behind	The wind was abaft which made the ship glide f...	They placed the lifeboats in the rear part of ...	My first sailing lesson was learning that abaf...	The sea spray was coming from abaft, making my...	I was told to sit abaft to lessen my seasickne...	NaN	
5	abandon	to stop caring and leave or to desert	Since the young girl was unable to care for th...	Carl rarely commits to anything so everyone ex...	During the snow storm, many people made the de...	The military police is currently looking for t...	Because this plan is not feasible, we need to ...	NaN	
6	abase	to lower so as to hurt feelings	My stepmother is an evil woman who likes to ab...	Because Mark hated himself, he often tried to ...	I will not remain married to a man who tries t...	Although I need money badly, I will not abase ...	My jealous sister tried to abase me by making ...	When my brother behaved badly, my mother would...	The g colu tri abas sir
...	...	...	...	...	...	...	...	...	

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5	Examp
13097	zodiac	a divided band in sky along which the sun, the...	Twelve cities in the state were named after as...	The zodiac signs are split over the twelve mon...	Zodiac symbols are based on the constellations...	Each of the zodiac signs covers an equal area ...	The zodiac system was based off of the sun, mo...	NaN	
13098	zone	an area or division	The zone only has houses and no stores or shop...	Walmart workers will place cleaning supplies i...	Before the team warmed up, all of the player's...	Because they live in a quake zone, the couple ...	People must drive slower in the construction z...	NaN	
13099	zoology	that portion of science that deals with the an...	Since Chamberlain had a fierce love for animal...	April used her zoology degree to help ranchers...	Through the school's zoology program, students...	Kids learn about all types of animals when stu...	In order to earn her degree in zoology, Rhonda...	NaN	
13100	zoom	to move very quickly	The cars began to zoom around the racetrack, m...	Police will need to monitor the fast-moving ca...	Reading eagerly, the avid reader can zoom thro...	Hurried shoppers looking for Christmas present...	The track star was able to zoom past all of th...	NaN	
13101	zygote	the fertilized egg that results from the union...	A zygote is formed when a sperm cell and an eg...	After the zygote goes through a rapid period o...	The production of a zygote is dependent upon a...	Simply stated, a zygote is an egg that has bee...	When a female gamete and a male gamete fuse, t...	NaN	

13102 rows × 14 columns

In [27]:
```python
# Check the N/A values of count column
sns.heatmap(data=df_merge['count'].isna().values.reshape(-1, 1))
```
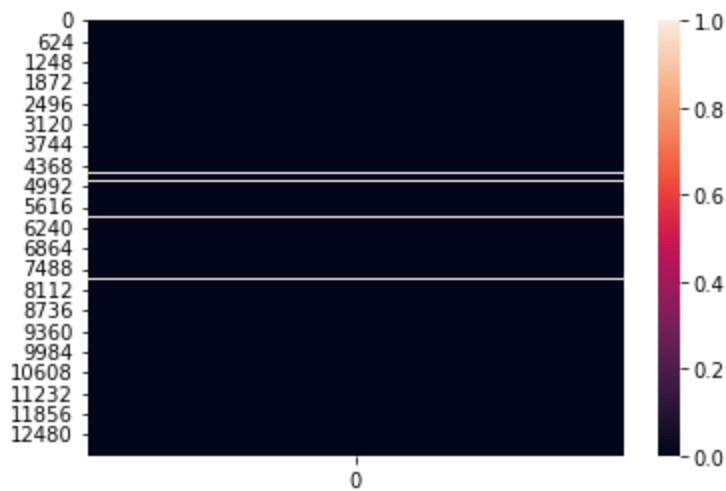
Out[27]: <AxesSubplot:>
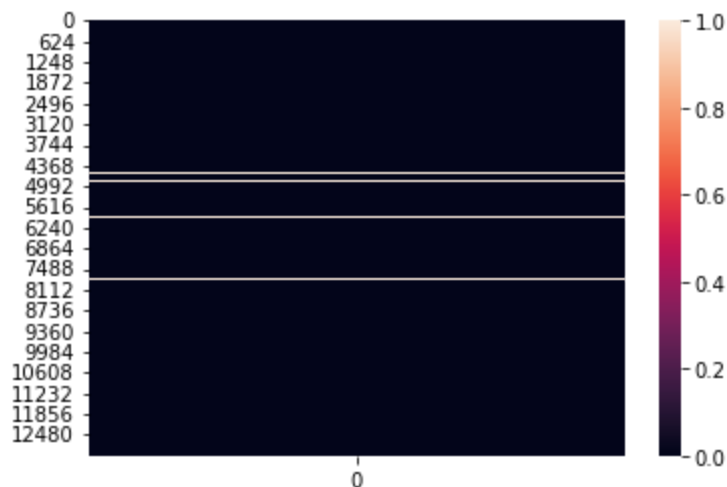
In [28]:
```python
df_merge['count'] =df_merge['count'].dropna()
```

In [29]:
```python
sns.heatmap(data=df_merge['count'].isna().values.reshape(-1, 1))
```

```
Out[29]: <AxesSubplot:>
```



```
In [30]:  # drop na values only consider the 'count' column to determine which rows to drop.
          df_merge.dropna(subset=['count'], inplace=True)
```

```
In [31]:  df_merge.sort_values(by=['count','Word'], ascending =[False,True])
```

Out[31]:

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4
5672	its	related to something that was mentioned previo...	After three days, the sick dog was allowed to ...	The store closed its doors when business came ...	The charity will have to stop its community pr...	The dog took its bone to its favorite burial s...	During Spring Break, the normally quiet town w...
11653	time	the juncture or point at which an event takes ...	When the bell rang, it was time to say good-by...	Due to the low pay, decrease in hours and lack...	"It's time to make the doughnuts," said the Du...	With a growling stomach, the boy knew it was t...	After dating his girlfriend for three years, t...
10769	state	to say something, either out loud or on paper	Eli will state his case to the judge in the mo...	The customer decided to state his complaint in...	Speaking to the witness, the judge asked the w...	While writing her research paper, the student ...	The job application asks the person looking fo...
10770	state	the current condition that someone or somethin...	Mary hasn't been in a very good mental state s...	Our state of economic affairs is rather poor r...	The nation is in a state of war right now, mas...	My state of dress is usually quite casual, but...	The city is in a state of emergency, knowing f...
10771	state	the condition or status of an object or situation	The landlord's last tenants left his rental ho...	Because the antique is in a spotless state, th...	The customer called the company to check on th...	NaN	NaN
...	...	...	...	...	...	...	...
11435	telegnosis	knowledge of events acquired from supernatural...	The psychic must have used telegnosis to read ...	Anyone that claims to read the future also cla...	Some people think an Ouija board can give them...	If you learned something about the future thro...	Because no normal human can learn of things th...
5090	impavid	extreme courage or bravery	Mongols were impavid warriors who rode into da...	Impavid vacationers dare to swim with sharks w...	Only the impavid would choose to jump from an ...	The student wished she was impavid, but truthf...	Ancient Roman gladiators had to be impavid bec...

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4
12046	ultracrepidarian	a person who acts like an expert but rarely kn...	The ultracrepidarian talked about physics as i...	Every university class has an ultracrepidarian...	Bruce knew that the ultracrepidarian liked to ...	Shane was the ultracrepidarian of the family w...	Briana was an ultracrepidarian who acted like ...
12006	turophile	a person who loves cheese	As a connoisseur of fancy cheese, the turophil...	James is a turophile of sorts and enjoys chees...	Brie, pepper jack, and mozzarella are some of ...	NaN	NaN
1548	bumfuzzle	to confuse someone	In an attempt to bumfuzzle his mother, Tony hi...	The twists and turns in the corn maze were con...	Sneaky advertisements were used to bumfuzzle n...	Re-reading the unclear directions, the lost to...	During the first part of the book, the author ...

12717 rows × 14 columns

```python
In [32]:  df_merge = df_merge.sort_values(by='count', ascending =False)
```

```python
In [33]:  # Use syllapy to classify words into syllable groups
          import syllapy
```

```python
In [34]:  # apply to Word column
          df_merge['syllables']= df_merge['Word'].apply(lambda x: syllapy.count(x))
```

```python
In [35]:  # Examine the result, These words below are correctly clasified
          df_merge[df_merge['syllables']==4]
```

Out[35]:

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Example
6967	necessary	being required or vital	For the chocolate cake, the necessary ingredie...	In order to climb the highest mountain, the cl...	Cancer patients must undergo the necessary che...	Since we lived in a bad area of town, we took ...	After writing my sentences, I placed the neces...	N
6784	necessary	being required or vital	For the chocolate cake, the necessary ingredie...	In order to climb the highest mountain, the cl...	Cancer patients must undergo the necessary che...	Since we lived in a bad area of town, we took ...	After writing my sentences, I placed the neces...	N
985	available	accessible and obtainable	When my dream house was put on the market and ...	Since the game was available on the app store,...	Once the employee left the firm, his job becam...	Due to having no events to attend this week, t...	"If you are available on Friday," the mother a...	N
9895	secretary	an office assistant, usually in charge of records	Trying to impress the new secretary caused the...	With no secretary to organize his paperwork, t...	The doctor's secretary had a knack for efficie...	Hiring his wife as a secretary seemed like a m...	Prematurely firing the office secretary proved...	N

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Example
10426	society	an organized group of people	Homeowners contested that the property tax was...	The school society was made up of teachers, pa...	Western society generally holds that polygamy ...	Reintegrating felons into general society is o...	Ordinary people struggle to find their place i...	N
...	...	...	...	...	...	...	...	
7194	obnubilate	to obscure or darken something	If a cloud passes in front of the sun, it will...	I find it difficult to sleep when there is too...	Sunglasses are one of the most common tools us...	Some people use tinted windows to obnubilate t...	The best way to obnubilate someone's vision is...	N
10482	somnambulate	to walk while sleeping	The sleep therapist was studying patients who ...	The parents stayed up all night to make sure t...	In order to somnambulate, one needs to be asleep.	Because he continued to somnambulate, they wan...	He would always end up outside during a somnam...	N
379	ailurophile	a person who loves cats	The neighborhood children only saw the ailurop...	Although she does love cats, the ailurophile d...	The ailurophile loves cats more than people, s...	Linda is a well-known ailurophile who has a fe...	Ann isn't a dog person but an ailurophile whos...	N
1628	callipygous	having a proportionate and well-shaped rear end	The callipygous girls in the videos were hired...	Showing off her callipygous figure, the dancer...	Having a pear-shaped body, the callipygous mod...	Because of her callipygous body-style, it is h...	The callipygous stripper got lots of attention...	N
11435	telegnosis	knowledge of events acquired from supernatural...	The psychic must have used telegnosis to read ...	Anyone that claims to read the future also cla...	Some people think an Ouija board can give them...	If you learned something about the future thro...	Because no normal human can learn of things th...	N

1969 rows × 15 columns

```
In [36]:   # Check N/A values. There is no N/A.
           sns.heatmap(df_merge['Word'].isna().values.reshape(-1, 1))
```

Out[36]: `<AxesSubplot:>`

```
In [37]:    # Drop these column because of useless
            df_merge = df_merge.drop(['Examples/6','Examples/7','Examples/8','Examples/9'], axis=1)
```

```
In [38]:    # drop word column
            df_merge = df_merge.drop('word', axis=1)
```

```
In [39]:    df = df_merge.groupby('syllables')
```

```
In [40]:    # Sum the frequency of each syllable group
            sum_df = df.sum()
            print(sum_df)
```

```
                    count
syllables
1           1.357716e+11
2           1.469363e+11
3           9.142936e+10
4           4.037981e+10
5           1.032895e+10
6           6.671753e+08
7           1.384702e+06
```

df_merge.sort_values(by=['syllables','count'], ascending =[True,False])

Data Visualization

```
In [41]:    # Plot a number words by syllables chart
            plt.figure(figsize=(10, 6))
            sns.barplot(x=df_merge['syllables'].value_counts().index,
                        y=df_merge['syllables'].value_counts())
            plt.title('Frequency of Syllables')
            plt.xlabel('Syllables')
            plt.ylabel('Frequency')
```

Out[41]:   Text(0, 0.5, 'Frequency')
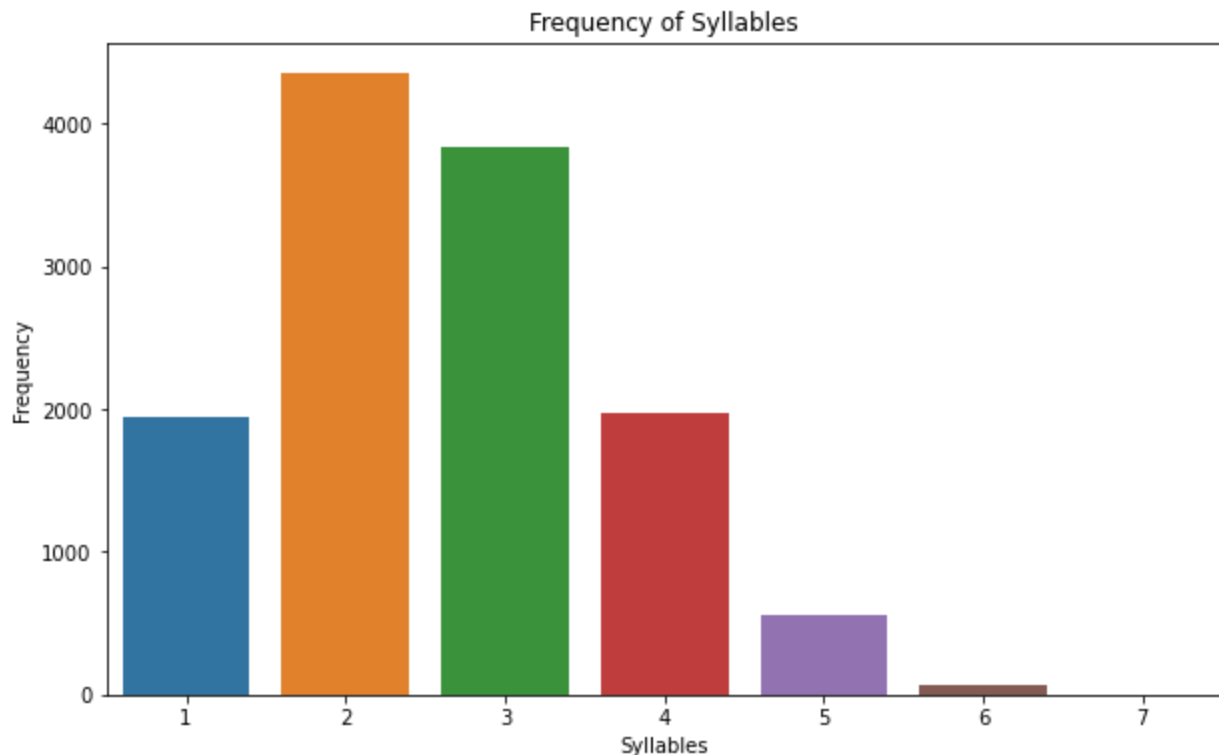
```
In [42]:   # The frequency of each syllable group
           df_merge['syllables'].value_counts()

Out[42]:   2    4349
           3    3831
           4    1969
           1    1941
           5     560
           6      64
           7       3
           Name: syllables, dtype: int64

In [43]:   sum_counts = df_merge.groupby('syllables')['count'].sum()
           print("Sum of counts for each syllable:")
           print(sum_counts)

           Sum of counts for each syllable:
           syllables
           1    1.357716e+11
           2    1.469363e+11
           3    9.142936e+10
           4    4.037981e+10
           5    1.032895e+10
           6    6.671753e+08
           7    1.384702e+06
           Name: count, dtype: float64

In [44]:   # Frequency of Words by Syllables in Use
           # The words with the most frequency are the ones with two syllables.
           sorted_df = df_merge.groupby('syllables')['count'].sum().reset_index().sort_values(by='cou

           # Plot the data
           sns.barplot(x='syllables', y='count', data=sorted_df)

           # Show the plot
           plt.show()
```
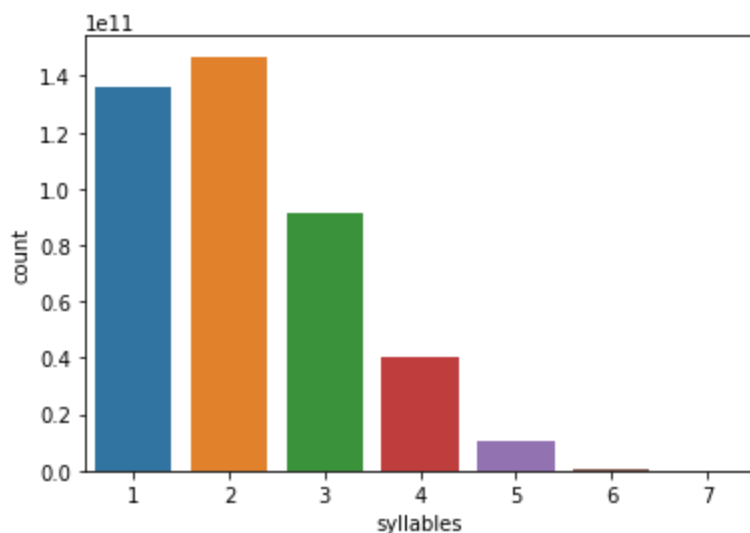


```
In [45]:   import nltk
           from nltk.tokenize import word_tokenize

In [46]:   # Function to categorize words based on part of speech
           def categorize_pos(word):
               pos_tag = nltk.pos_tag([word])[0][1]
```

```
        if pos_tag.startswith('N'):
            return 'N'
        elif pos_tag.startswith('J'):
            return 'Adj'
        elif pos_tag.startswith('V'):
            return 'V'
        elif pos_tag.startswith('R'):
            return 'Adv'
        elif pos_tag.startswith('WDT'):
            return 'wh-determiner'
        elif pos_tag.startswith('MD'):
            return 'modal'
        elif pos_tag.startswith('CC'):
            return 'CC'
        elif pos_tag.startswith('PRP$'):
            return 'PRP'
        elif pos_tag.startswith('WRB'):
            return 'Wh-adverb'
        else:
            return 'other'

    # Assuming df is your DataFrame and 'words' is the column containing the words
    df_merge['pos_category'] = df_merge['Word'].apply(categorize_pos)
```

In [47]:
```
df_merge.sort_values(by=['syllables','count','pos_category'], ascending =[True, False, Tru
```

Out[47]:

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4
5672	its	related to something that was mentioned previo…	After three days, the sick dog was allowed to …	The store closed its doors when business came …	The charity will have to stop its community pr…	The dog took its bone to its favorite burial s…	During Spring Break, the normally quiet town w…
11653	time	the juncture or point at which an event takes …	When the bell rang, it was time to say good-by…	Due to the low pay, decrease in hours and lack…	"It's time to make the doughnuts," said the Du…	With a growling stomach, the boy knew it was t…	After dating his girlfriend for three years, t…
10771	state	the condition or status of an object or situation	The landlord's last tenants left his rental ho…	Because the antique is in a spotless state, th…	The customer called the company to check on th…	NaN	NaN
10769	state	to say something, either out loud or on paper	Eli will state his case to the judge in the mo…	The customer decided to state his complaint in…	Speaking to the witness, the judge asked the w…	While writing her research paper, the student …	The job application asks the person looking fo…
10770	state	the current condition that someone or somethin…	Mary hasn't been in a very good mental state s…	Our state of economic affairs is rather poor r…	The nation is in a state of war right now, mas…	My state of dress is usually quite casual, but…	The city is in a state of emergency, knowing f…
…	…	…	…	…	…	…	…

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4
11669	tintinnabulation	a ringing or tinkling sound	You are more likely to hear tintinnabulation a...	The sound a wind chime makes is akin to a tint...	One might not consider a triangle as a legitim...	Instruments that create the sound of tintinnab...	I am not a fan of tintinnabulation, which is w...
12046	ultracrepidarian	a person who acts like an expert but rarely kn...	The ultracrepidarian talked about physics as i...	Every university class has an ultracrepidarian...	Bruce knew that the ultracrepidarian liked to ...	Shane was the ultracrepidarian of the family w...	Briana was an ultracrepidarian who acted like ...
2067	collaboratively	in a way that involves two or more people work...	Working collaboratively, the two musicians wer...	During the group assignment, several students ...	The staff worked collaboratively to create a c...	NaN	NaN
12141	undesirability	the quality of being unwanted or desired	The product's undesirability caused it to sit ...	The rain coats' undesirability is based on loc...	Looking for a way to rid herself of her undesi...	The southern chef was surprised by the undesir...	Because he was so ugly, the dog's undesirabili...
11131	supererogatory	actions that are considered good but are not s...	A supererogatory act includes extra credit wor...	The medic was a supererogatory hero for runnin...	My supererogatory teacher went out of her way ...	The Medal of Honor is rarely awarded to someon...	A supererogatory hero goes above and beyond th...

12717 rows × 11 columns

In [48]:
```python
plt.figure(figsize=(16,8))
sns.histplot(data=df_merge, x='pos_category', hue='syllables', multiple="stack", palette='
```
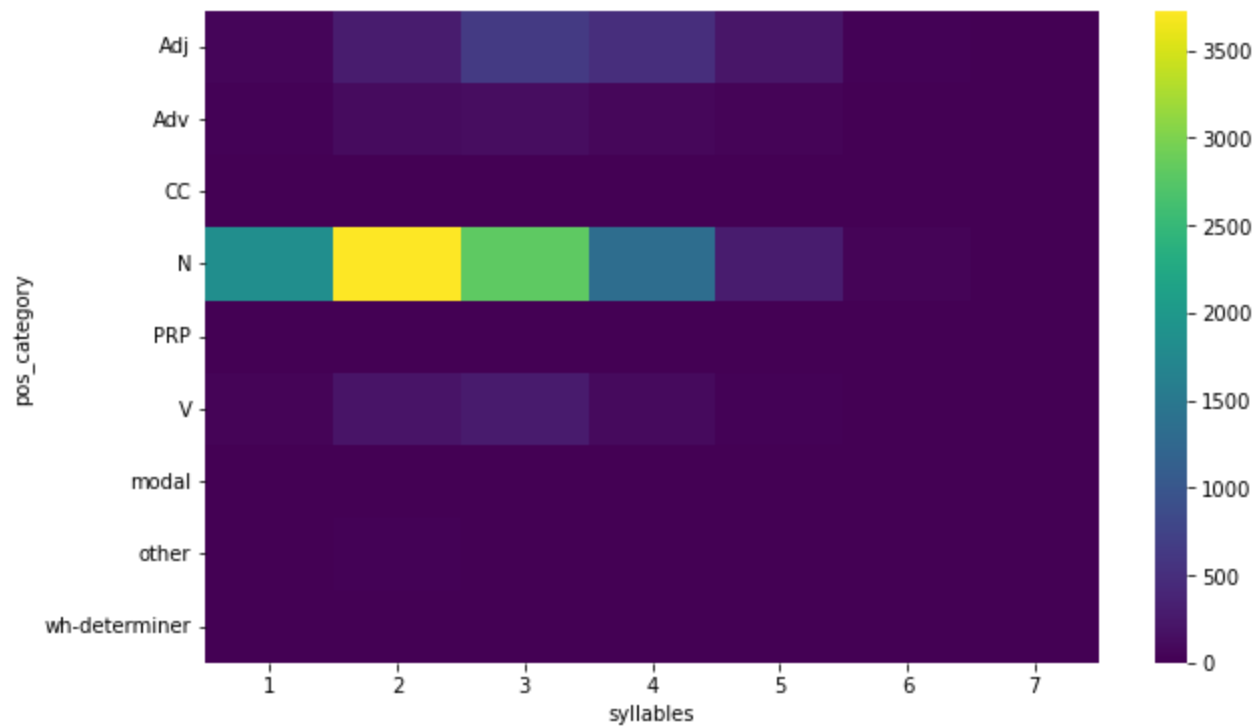
Out[48]:
```
<AxesSubplot:xlabel='pos_category', ylabel='Count'>
```
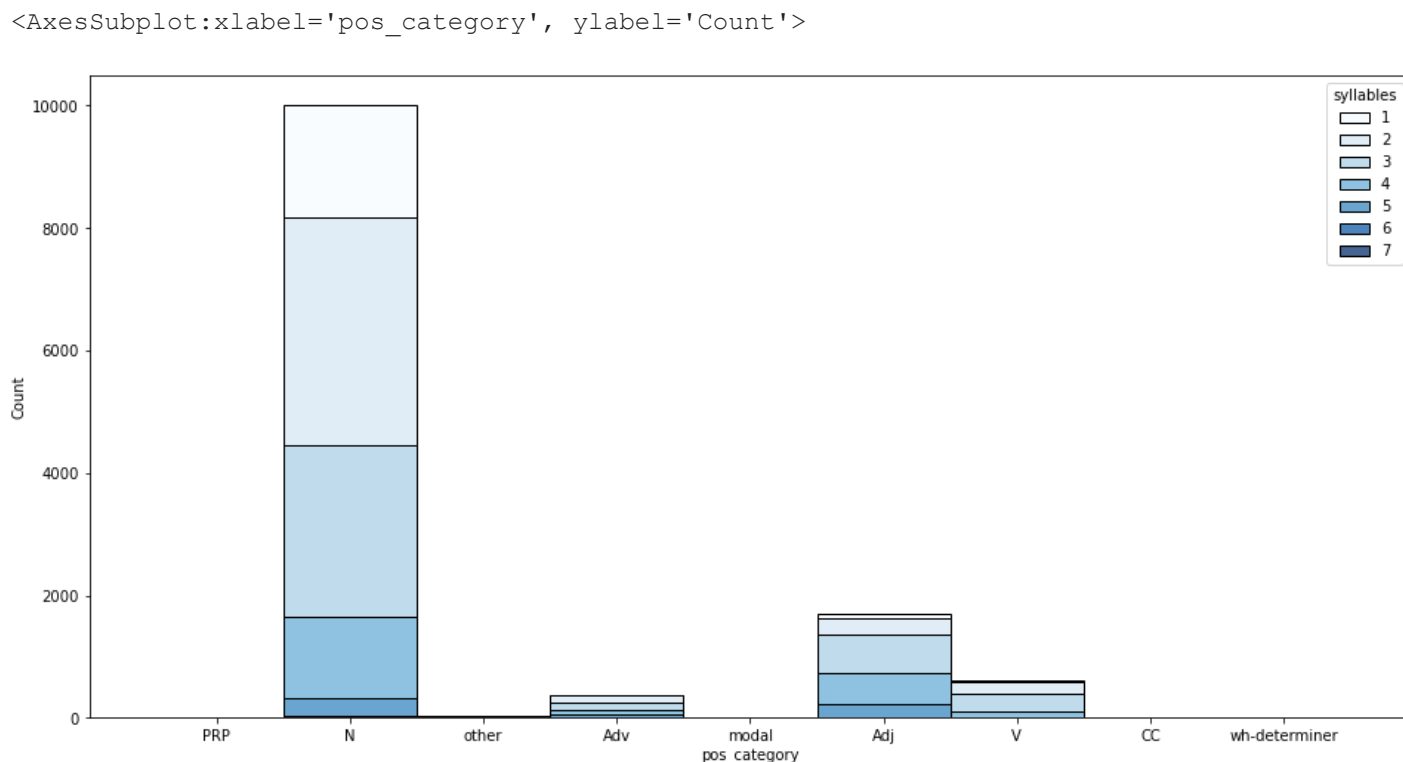

In [49]:
```python
# unstack() method reshapes the Series with a multi-level index into a DataFrame.
# transform the type of data
```

```python
# The size() method counts the number of rows in each group created by groupby.
heatmap_data = df_merge.groupby(['pos_category', 'syllables']).size().unstack(fill_value=(
plt.figure(figsize=(10, 6))
sns.heatmap(data=heatmap_data, cmap='viridis')
```

Out[49]: `<AxesSubplot:xlabel='syllables', ylabel='pos_category'>`

In [50]:
```python
# Ranking the words by their count
df_merge['rank'] = df_merge['count'].rank(ascending=False)
```

In [51]:
```python
df_merge['rank'] = df_merge['rank'].astype(int)
```

In [52]:
```python
df_merge
```

Out[52]:

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4
5672	its	related to something that was mentioned previo...	After three days, the sick dog was allowed to ...	The store closed its doors when business came ...	The charity will have to stop its community pr...	The dog took its bone to its favorite burial s...	During Spring Break, the normally quiet town w...
11653	time	the juncture or point at which an event takes ...	When the bell rang, it was time to say good-by...	Due to the low pay, decrease in hours and lack...	"It's time to make the doughnuts," said the Du...	With a growling stomach, the boy knew it was t...	After dating his girlfriend for three years, t...
10771	state	the condition or status of an object or situation	The landlord's last tenants left his rental ho...	Because the antique is in a spotless state, th...	The customer called the company to check on th...	NaN	NaN
10769	state	to say something, either out loud or on paper	Eli will state his case to the judge in the mo...	The customer decided to state his complaint in...	Speaking to the witness, the judge asked the w...	While writing her research paper, the student ...	The job application asks the person looking fo...

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4
10770	state	the current condition that someone or somethin...	Mary hasn't been in a very good mental state s...	Our state of economic affairs is rather poor r...	The nation is in a state of war right now, mas...	My state of dress is usually quite casual, but...	The city is in a state of emergency, knowing f...
...	...	...	...	...	...	...	...
11435	telegnosis	knowledge of events acquired from supernatural...	The psychic must have used telegnosis to read ...	Anyone that claims to read the future also cla...	Some people think an Ouija board can give them...	If you learned something about the future thro...	Because no normal human can learn of things th...
5090	impavid	extreme courage or bravery	Mongols were impavid warriors who rode into da...	Impavid vacationers dare to swim with sharks w...	Only the impavid would choose to jump from an ...	The student wished she was impavid, but truthf...	Ancient Roman gladiators had to be impavid bec...
12046	ultracrepidarian	a person who acts like an expert but rarely kn...	The ultracrepidarian talked about physics as i...	Every university class has an ultracrepidarian...	Bruce knew that the ultracrepidarian liked to ...	Shane was the ultracrepidarian of the family w...	Briana was an ultracrepidarian who acted like ...
12006	turophile	a person who loves cheese	As a connoisseur of fancy cheese, the turophil...	James is a turophile of sorts and enjoys chees...	Brie, pepper jack, and mozzarella are some of ...	NaN	NaN
1548	bumfuzzle	to confuse someone	In an attempt to bumfuzzle his mother, Tony hi...	The twists and turns in the corn maze were con...	Sneaky advertisements were used to bumfuzzle n...	Re-reading the unclear directions, the lost to...	During the first part of the book, the author ...

12717 rows × 12 columns

```
In [53]:  df_merge.columns
```

```
Out[53]:  Index(['Word', 'Meaning', 'Examples/0', 'Examples/1', 'Examples/2',
                 'Examples/3', 'Examples/4', 'Examples/5', 'count', 'syllables',
                 'pos_category', 'rank'],
                dtype='object')
```

```
In [54]:  df_merge.head()
```

Out[54]:

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5	count
5672	its	related to something that was mentioned previo...	After three days, the sick dog was allowed to ...	The store closed its doors when business came ...	The charity will have to stop its community pr...	The dog took its bone to its favorite burial s...	During Spring Break, the normally quiet town w...	NaN	3.160954e+09
11653	time	the juncture or point at which an event takes ...	When the bell rang, it was time to say good-by...	Due to the low pay, decrease in hours and lack...	"It's time to make the doughnuts," said the Du...	With a growling stomach, the boy knew it was t...	After dating his girlfriend for three years, t...	NaN	2.844215e+09

	Word	Meaning	Examples/0	Examples/1	Examples/2	Examples/3	Examples/4	Examples/5	count
10771	state	the condition or status of an object or situation	The landlord's last tenants left his rental ho...	Because the antique is in a spotless state, th...	The customer called the company to check on th...	NaN	NaN	NaN	2.349553e+09
10769	state	to say something, either out loud or on paper	Eli will state his case to the judge in the mo...	The customer decided to state his complaint in...	Speaking to the witness, the judge asked the w...	While writing her research paper, the student ...	The job application asks the person looking fo...	NaN	2.349553e+09
10770	state	the current condition that someone or somethin...	Mary hasn't been in a very good mental state s...	Our state of economic affairs is rather poor r...	The nation is in a state of war right now, mas...	My state of dress is usually quite casual, but...	The city is in a state of emergency, knowing f...	NaN	2.349553e+09

In [55]:
```python
columns_name_reordered = ['rank', 'Word', 'pos_category','Meaning', 'Examples/0', 'Example
        'Examples/2', 'Examples/3', 'Examples/4', 'Examples/5', 'count',
        'syllables']
```

In [56]:
```python
df_merge=df_merge[columns_name_reordered]
```

In [57]:
```python
df_merge.columns
```

Out[57]:
```
Index(['rank', 'Word', 'pos_category', 'Meaning', 'Examples/0', 'Examples/1',
       'Examples/2', 'Examples/3', 'Examples/4', 'Examples/5', 'count',
       'syllables'],
      dtype='object')
```

In [58]:
```python
columns_renamed = ['Rank', 'Word', 'Lexical words', 'Meaning', 'Examples 1',
        'Examples 2', 'Examples 3', 'Examples 4', 'Examples 5', 'Examples 6',
        'Frequency', 'Syllables']
```

In [59]:
```python
df_merge.columns = columns_renamed
```

In [60]:
```python
df_merge=df_merge.sort_values(by=['Syllables','Rank','Lexical words'], ascending =[True, 
```

In [61]:
```python
df_merge=df_merge.drop(['Frequency','Examples 6'], axis=1)
```

In [62]:
```python
df_merge
```

Out[62]:
	Rank	Word	Lexical words	Meaning	Examples 1	Examples 2	Examples 3	Examples 4

	Rank	Word	Lexical words	Meaning	Examples 1	Examples 2	Examples 3	Examples 4
5672	1	its	PRP	related to something that was mentioned previo...	After three days, the sick dog was allowed to ...	The store closed its doors when business came ...	The charity will have to stop its community pr...	The dog took its bone to its favorite burial s...
11653	2	time	N	the juncture or point at which an event takes ...	When the bell rang, it was time to say good-by...	Due to the low pay, decrease in hours and lack...	"It's time to make the doughnuts," said the Du...	With a growling stomach, the boy knew it was t...
10771	4	state	N	the condition or status of an object or situation	The landlord's last tenants left his rental ho...	Because the antique is in a spotless state, th...	The customer called the company to check on th...	NaN
10769	4	state	N	to say something, either out loud or on paper	Eli will state his case to the judge in the mo...	The customer decided to state his complaint in...	Speaking to the witness, the judge asked the w...	While writing her research paper, the student ...
10770	4	state	N	the current condition that someone or somethin...	Mary hasn't been in a very good mental state s...	Our state of economic affairs is rather poor r...	The nation is in a state of war right now, mas...	My state of dress is usually quite casual, but...
...	...	...	...	...	...	...	...	...
11669	12548	tintinnabulation	N	a ringing or tinkling sound	You are more likely to hear tintinnabulation a...	The sound a wind chime makes is akin to a tint...	One might not consider a triangle as a legitim...	Instruments that create the sound of tintinnab...
12046	12715	ultracrepidarian	Adj	a person who acts like an expert but rarely kn...	The ultracrepidarian talked about physics as i...	Every university class has an ultracrepidarian...	Bruce knew that the ultracrepidarian liked to ...	Shane was the ultracrepidarian of the family w...
2067	9766	collaboratively	Adv	in a way that involves two or more people work...	Working collaboratively, the two musicians wer...	During the group assignment, several students ...	The staff worked collaboratively to create a c...	NaN
12141	11285	undesirability	N	the quality of being unwanted or desired	The product's undesirability caused it to sit ...	The rain coats' undesirability is based on loc...	Looking for a way to rid herself of her undesi...	The southern chef was surprised by the undesir...
11131	12058	supererogatory	N	actions that are considered good but are not s...	A supererogatory act includes extra credit wor...	The medic was a supererogatory hero for runnin...	My supererogatory teacher went out of her way ...	The Medal of Honor is rarely awarded to someon...

In [63]:
```python
import eng_to_ipa as p
```

In [64]:
```python
df_merge=df_merge.drop('Phonetics', axis=1)
```

```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
~\AppData\Local\Temp/ipykernel_5752/3483754962.py in <module>
----> 1 df_merge=df_merge.drop('Phonetics', axis=1)

~\anaconda3\lib\site-packages\pandas\util\_decorators.py in wrapper(*args, **kwargs)
    309                     stacklevel=stacklevel,
    310                 )
--> 311             return func(*args, **kwargs)
    312
    313         return wrapper

~\anaconda3\lib\site-packages\pandas\core\frame.py in drop(self, labels, axis, index, colu
mns, level, inplace, errors)
   4904                 weight  1.0     0.8
   4905         """
-> 4906         return super().drop(
   4907             labels=labels,
   4908             axis=axis,

~\anaconda3\lib\site-packages\pandas\core\generic.py in drop(self, labels, axis, index, co
lumns, level, inplace, errors)
   4148         for axis, labels in axes.items():
   4149             if labels is not None:
-> 4150                 obj = obj._drop_axis(labels, axis, level=level, errors=errors)
   4151
   4152         if inplace:

~\anaconda3\lib\site-packages\pandas\core\generic.py in _drop_axis(self, labels, axis, lev
el, errors)
   4183                 new_axis = axis.drop(labels, level=level, errors=errors)
   4184             else:
-> 4185                 new_axis = axis.drop(labels, errors=errors)
   4186             result = self.reindex(**{axis_name: new_axis})
   4187

~\anaconda3\lib\site-packages\pandas\core\indexes\base.py in drop(self, labels, errors)
   6015         if mask.any():
   6016             if errors != "ignore":
-> 6017                 raise KeyError(f"{labels[mask]} not found in axis")
   6018             indexer = indexer[~mask]
   6019         return self.delete(indexer)

KeyError: "['Phonetics'] not found in axis"
```

In [ ]:
```python
df_merge['Word'][11131][-1:]
```

In [ ]:
```python
def change(word):
    ipa = p.convert(word)
    if ipa[-1:]!= "*":
        return ipa
    else:
        return "unknown"
```

```
df_merge['Phonetics'] = df_merge['Word'].apply(change)
```

```
df_merge
```

```
df_merge=df_merge[df_merge["Phonetics"] != "unknown"]
```

```
## Find the most common range of eng
```

```
df_common = df_merge.copy()
```

```
df_common.sort_values("Rank" ,ascending = True)[:7000][df_common["Syllables"]==6]
```

```
len(df_common.sort_values("Rank" ,ascending = True)[:000][df_common["Syllables"]==2])
```

```
# Create a FacetGrid with Seaborn
g = sns.FacetGrid(df_common, col="Syllables")

# Map the histogram plot onto the grid
g.map(sns.histplot, "Rank", kde=True, bins=20)
```

```
# Show in rank(0-400) how many values have syllables equal 1
df_common.sort_values("Rank" ,ascending = True)[:400][df_common["Syllables"]==1]
```

```
# Show in from lists of Syllables having values =1, sorted by Rank and get 400 individuals
df_common[df_common["Syllables"]==1].sort_values("Rank" ,ascending = True)[:400]
```

```
len(df_common.loc[(df_common["Rank"] <= 300) & (df_common["Syllables"] == 2), :].sort_valu
```

```
df_common.to_csv(r'C:\Users\DELL\Desktop\Words(En) Project\edit_eng_words.csv')
```

```
df_common = pd.read_csv(r'C:\Users\DELL\Desktop\Words(En) Project\edit_eng_words.csv')
```

```
df_common
```

```
one= len(df_common.loc[(400 <= df_common["Rank"]) & (df_common["Rank"] <= 4500) & (df_comm
print(one)
two=len(df_common.loc[(400 <= df_common["Rank"]) & (df_common["Rank"] <= 4500) & (df_commo
print(two)
three=len(df_common.loc[(400 <= df_common["Rank"]) & (df_common["Rank"] <= 4500) & (df_con
print(three)
four= len(df_common.loc[(400 <= df_common["Rank"]) & (df_common["Rank"] <= 4500) & (df_con
print(four)
five=len(df_common.loc[(400 <= df_common["Rank"]) & (df_common["Rank"] <= 4500) & (df_comm
print(five)
```

```python
      six=len(df_common.loc[(400 <= df_common["Rank"]) & (df_common["Rank"] <= 4500) & (df_commo
      print(six)
```

In [ ]:
```python
counts = []
for syllable_count in range(1, 7):
    count = len(df_common.loc[(400 <= df_common["Rank"]) & (df_common["Rank"] <= 4500) &
    counts.append(count)

syllable_counts = range(1, 7)
plt.bar(syllable_counts, counts)
plt.xlabel('Number of Syllables')
plt.ylabel('Count')
plt.title('Count of Words by Syllable Count')
plt.show()
```

In [ ]:
```python
import seaborn as sns
import matplotlib.pyplot as plt

# Filter the DataFrame based on conditions
filtered_df = df_common.loc[(400 <= df_common["Rank"]) & (df_common["Rank"] <= 4500)]

# Create countplot
sns.countplot(x="Syllables", data=filtered_df)
plt.xlabel('Number of Syllables')
plt.ylabel('Count')
plt.title('Count of Words by Syllable Count')
plt.show()
```

In [ ]:
```python
# Filter the DataFrame based on conditions
filtered_df = df_common.loc[(400 <= df_common["Rank"]) & (df_common["Rank"] <= 4500)]

# Create countplot
ax = sns.countplot(x="Syllables", data=filtered_df)

# Add count labels on top of each bar
for p in ax.patches:
    ax.annotate(format(p.get_height(), '.0f'),
                (p.get_x() + p.get_width() / 2., p.get_height()),
                ha = 'center', va = 'center',
                xytext = (0, 5),
                textcoords = 'offset points')

plt.xlabel('Number of Syllables')
plt.ylabel('Count')
plt.title('Count of Words by Syllable Count')
plt.show()
```

In [ ]:
```python
# Define the filter conditions
filter_conditions = [
    ("Rank 400-1000", (400 <= df_common["Rank"]) & (df_common["Rank"] <= 1000)),
    ("Rank 1001-2500", (1001 <= df_common["Rank"]) & (df_common["Rank"] <= 2500)),
    ("Rank 2501-4500", (2501 <= df_common["Rank"]) & (df_common["Rank"] <= 4500))
]

# Create subplots
fig, axes = plt.subplots(nrows=1, ncols=len(filter_conditions), figsize=(15, 5))

# Plot countplots for each filter condition
for i, (title, condition) in enumerate(filter_conditions):
    filtered_df = df_common.loc[condition]
```

```python
        ax = sns.countplot(x="Syllables", data=filtered_df, ax=axes[i])
        ax.set_title(title)
        ax.set_xlabel('Number of Syllables')
        ax.set_ylabel('Count')

plt.tight_layout()
plt.show()
```

```python
# Define the filter conditions
filter_conditions = [
    ("Rank 400-3000", (400 <= df_common["Rank"]) & (df_common["Rank"] <= 3000)),
    ("Rank 3001-5500", (3001 <= df_common["Rank"]) & (df_common["Rank"] <= 5500)),
    ("Rank 5501-7000", (5501 <= df_common["Rank"]) & (df_common["Rank"] <= 7000))
]

# Create a new column to represent the rank range
for title, condition in filter_conditions:
    df_common.loc[condition, "Filter"] = title

# Create a FacetGrid with hue
g = sns.FacetGrid(df_common, col="Filter", hue="Filter", col_wrap=len(filter_conditions),

# Map countplot onto the grid
g.map(sns.countplot, "Syllables", order=[1, 2, 3, 4, 5, 6])

# Annotate each bar with its count
for ax in g.axes.flat:
    for p in ax.patches:
        ax.annotate(format(p.get_height(), '.0f'),
                    (p.get_x() + p.get_width() / 2., p.get_height()),
                    ha = 'center', va = 'center',
                    xytext = (0, 5),
                    textcoords = 'offset points')

# Set labels
g.set_axis_labels("Number of Syllables", "Count")

plt.tight_layout()
plt.show()
```

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Define the filter conditions
filter_conditions = [
    ("Rank 400-3000", (400 <= df_common["Rank"]) & (df_common["Rank"] <= 5000)),
    ("Rank 3001-5500", (3001 <= df_common["Rank"]) & (df_common["Rank"] <= 7500)),
    ("Rank 5501-7000", (5501 <= df_common["Rank"]) & (df_common["Rank"] <= 7000)),
    ("Rank 7001-11000", (7001 <= df_common["Rank"]) & (df_common["Rank"] <= 11000))
]

# Create a new column to represent the rank range
for title, condition in filter_conditions:
    df_common.loc[condition, "Filter"] = title

# Create a FacetGrid with hue
g = sns.FacetGrid(df_common, col="Filter", col_wrap=len(filter_conditions), sharey=False)

# Map countplot onto the grid with custom color palette
g.map(sns.countplot, "Syllables", order=[1, 2, 3, 4, 5, 6], palette="husl")

# Annotate each bar with its count
for ax in g.axes.flat:
```

```
        for p in ax.patches:
            ax.annotate(format(p.get_height(), '.0f'),
                        (p.get_x() + p.get_width() / 2., p.get_height()),
                        ha = 'center', va = 'center',
                        xytext = (0, 5),
                        textcoords = 'offset points')

    # Set labels
    g.set_axis_labels("Number of Syllables", "Count")

    plt.tight_layout()
    plt.show()
```

```
for syllable_count in range(1, 7):
    count = len(df_common.loc[(400 <= df_common["Rank"]) & (df_common["Rank"] <= 5500) &
    print(f"{syllable_count}: {count}")
```

```
df_common[(df_common["Rank"]<=200) & (df_common["Syllables"]==1)]
```

```
df_common.drop(df_common[(df_common["Rank"] <= 200) & (df_common["Syllables"] == 1)].index
```

```
df_common[:5000].drop_duplicates()
```

```
df_common = pd.read_csv(r'C:\Users\DELL\Desktop\Words(En) Project\edit_eng_words.csv')
```

```
df_common.drop
```

```
df_common.drop(df_common[(df_common["Rank"] <= 200) & (df_common["Syllables"] == 1)].index
```

```
df_common = df_common.sort_values("Rank", ascending=True)
```

```
df_common_cut = df_common[:5000]
```

```
df_common_cut
```

```
for syllable_count in range(1, 7):
    count = len(df_common_cut.loc[(400 <= df_common_cut["Rank"]) & (df_common_cut["Rank"]
    print(count)
```

```
df_common_cut.to_csv(r'C:\Users\DELL\Desktop\Words(En) Project\new_cut_eng_words.csv')
```

```
df_common_cut
```