

# WEB MINING

## LAB ASSESSMENT-1

**NAME:ISHAN ARORA**  
**REGNO-17BCE0972**

First we will import libraries

In [3]:

```
import nltk
import numpy as np
import pandas as pd
import os
```

Now we will import the text files.Each file is approx 10 lines long

In [5]:

```
movie_review=['doc1.txt','doc2.txt','doc3.txt','doc4.txt','doc5.txt','doc6.txt','doc
```

In [6]:

```
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
```

In [7]:

```
stop_word=set(stopwords.words('english'))
extra=[',','.', '(', ')', '[', ']']
for i in extra:
    stop_word.add(i)
```

In [8]:

```
doc=[]
for docs in movie_review:
    f=open(docs, 'r')
    doc.append(f.read())
```

In [9]:

```
from nltk.tokenize import word_tokenize
import collections
```

In [ ]:

```
**Now we will do tokenization and stop words in this step**
```

In [10]:

```
frequency=[]
for i in doc:
    word_tokens = word_tokenize(i)
    filtered_sent = [w for w in word_tokens if not w in stop_word]
    frequency.append(collections.Counter(filtered_sent))
```

In [11]:

```
frequency
```

...

In [12]:

```
s=[]
for i in frequency:
    d_dict=dict(i)
    a=d_dict.keys()
    for j in a:
        if j not in s:
            s.append(j)
```

In [13]:

```
f=[]
for i in frequency:
    f.append(dict(i))
```

In [14]:

```
data_table=[]
data_table.append(['Words','doc1','doc2','doc3','doc4','doc5','doc6','doc7','doc8',
for i in s:
    tmp=[]
    tmp.append(i)
    for j in range(10):
        try:
            tmp.append(f[j][i])
        except:
            tmp.append(0)
    data_table.append(tmp)
data_table=(np.array(data_table))
print(data_table)
```

```
[['Words' 'doc1' 'doc2' ... 'doc8' 'doc 9' 'doc10']
['As' '1' '0' ... '0' '0' '0']
['child' '1' '0' ... '0' '0' '0']
...
['encouraged' '0' '0' ... '0' '0' '1']
['class' '0' '0' ... '0' '0' '1']
['warfare' '0' '0' ... '0' '0' '1']]
```

In [15]:

```
df=pd.DataFrame(data_table)
```

## COMPLETE DATA FRAME

In [16]:

df

Out[16]:

		0	1	2	3	4	5	6	7	8	9	10
0	Words	doc1	doc2	doc3	doc4	doc5	doc6	doc7	doc8	doc9	doc10	
1	As	1	0	0	0	0	0	0	0	0	0	0
2	child	1	0	0	0	0	0	0	0	0	0	0
3	I	5	4	1	0	0	1	0	1	0	0	0
4	simply	1	0	0	0	0	0	0	0	0	0	0
5	notice	1	0	0	0	0	0	0	0	0	0	0
6	whether	1	0	0	0	0	0	0	0	0	0	0
7	movie	2	1	0	1	0	0	0	0	1	0	0
8	color	6	0	0	0	0	0	0	0	0	0	0
9	The	3	1	5	1	1	4	0	2	5	1	1
10	movies	1	0	0	1	0	0	0	2	0	0	0
11	overwhelming	1	0	0	0	0	0	0	0	0	0	0
12	mystery	1	1	0	0	0	0	0	0	0	0	0
13	wanted	1	0	0	0	1	1	0	0	0	0	0
14	black	3	0	0	0	0	0	0	0	0	0	0
15	white	3	0	0	0	1	0	0	0	0	0	0
16	business	1	0	0	1	0	0	0	0	0	0	0
17	It	2	2	1	1	0	0	0	1	0	1	1
18	saw	1	0	0	0	0	0	0	0	0	0	0
19	"	2	0	2	1	0	1	4	2	0	0	0
20	Wizard	1	1	0	0	0	0	0	0	0	0	0
21	Oz	2	0	0	0	0	0	0	0	0	0	0
22	"	2	0	2	1	0	1	4	4	0	0	0
23	first	1	0	0	1	0	1	0	0	0	0	0
24	time	1	1	0	0	0	0	0	0	0	0	1
25	consciously	1	0	0	0	0	0	0	0	0	0	0
26	noticed	1	0	0	0	0	0	0	0	0	0	0
27	B	1	0	0	0	0	0	0	0	0	0	0
28	&	1	0	0	1	0	0	0	0	0	0	0
29	W	1	0	0	0	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...
876	Olympian	0	0	0	0	0	0	0	0	0	0	1

	0	1	2	3	4	5	6	7	8	9	10
877	stadium	0	0	0	0	0	0	0	0	0	1
878	Pleasure	0	0	0	0	0	0	0	0	0	1
879	Gardens	0	0	0	0	0	0	0	0	0	1
880	Below	0	0	0	0	0	0	0	0	0	1
881	workers	0	0	0	0	0	0	0	0	0	2
882	clocks	0	0	0	0	0	0	0	0	0	1
883	10	0	0	0	0	0	0	0	0	0	1
884	hours	0	0	0	0	0	0	0	0	0	1
885	squeeze	0	0	0	0	0	0	0	0	0	1
886	live	0	0	0	0	0	0	0	0	0	1
887	tenement	0	0	0	0	0	0	0	0	0	1
888	housing	0	0	0	0	0	0	0	0	0	1
889	consists	0	0	0	0	0	0	0	0	0	1
890	unrelenting	0	0	0	0	0	0	0	0	0	1
891	service	0	0	0	0	0	0	0	0	0	1
892	machine	0	0	0	0	0	0	0	0	0	1
893	plutocracy	0	0	0	0	0	0	0	0	0	1
894	vs.	0	0	0	0	0	0	0	0	0	1
895	labor	0	0	0	0	0	0	0	0	0	1
896	powerful	0	0	0	0	0	0	0	0	0	1
897	era	0	0	0	0	0	0	0	0	0	1
898	assembly	0	0	0	0	0	0	0	0	0	1
899	line	0	0	0	0	0	0	0	0	0	1
900	large	0	0	0	0	0	0	0	0	0	1
901	scale	0	0	0	0	0	0	0	0	0	1
902	Marx	0	0	0	0	0	0	0	0	0	1
903	encouraged	0	0	0	0	0	0	0	0	0	1
904	class	0	0	0	0	0	0	0	0	0	1
905	warfare	0	0	0	0	0	0	0	0	0	1

906 rows × 11 columns

**Total number of term present in each document (row sum)**

In [17]:

```

data_table=[]
data_table.append(['Word','doc1','doc2','doc3','doc4','doc5','doc6','doc7','doc8','doc9','doc10'])
for i in s:
    tmp=[]
    tmp.append(i)
    for j in range(10):
        try:
            if f[j][i]!=0:
                tmp.append(1)
            else:
                tmp.append(0)
        except:
            tmp.append(0)
    data_table.append(tmp)
data_table=(np.array(data_table))
import pandas as pd
df=pd.DataFrame(data_table)
print('If word is found - 0, if not found - 1')
df

```

If word is found - 0, if not found - 1

Out[17]:

		0	1	2	3	4	5	6	7	8	9	10
0	Word	doc1	doc2	doc3	doc4	doc5	doc6	doc7	doc8	doc9	doc10	
1	As	1	0	0	0	0	0	0	0	0	0	
2	child	1	0	0	0	0	0	0	0	0	0	
3	I	1	1	1	0	0	1	0	1	0	0	
4	simply	1	0	0	0	0	0	0	0	0	0	
5	notice	1	0	0	0	0	0	0	0	0	0	
6	whether	1	0	0	0	0	0	0	0	0	0	
7	movie	1	1	0	1	0	0	0	0	1	0	
8	color	1	0	0	0	0	0	0	0	0	0	
9	The	1	1	1	1	1	1	0	1	1	1	
10	movies	1	0	0	1	0	0	0	1	0	0	
11	overwhelming	1	0	0	0	0	0	0	0	0	0	
12	mystery	1	1	0	0	0	0	0	0	0	0	
13	wanted	1	0	0	0	1	1	0	0	0	0	
14	black	1	0	0	0	0	0	0	0	0	0	
15	white	1	0	0	0	1	0	0	0	0	0	
16	business	1	0	0	1	0	0	0	0	0	0	

	0	1	2	3	4	5	6	7	8	9	10
17	It	1	1	1	1	0	0	0	1	0	1
18	saw	1	0	0	0	0	0	0	0	0	0
19	“	1	0	1	1	0	1	1	1	0	0
20	Wizard	1	1	0	0	0	0	0	0	0	0
21	Oz	1	0	0	0	0	0	0	0	0	0
22	"	1	0	1	1	0	1	1	1	0	0
23	first	1	0	0	1	0	1	0	0	0	0
24	time	1	1	0	0	0	0	0	0	0	1
25	consciously	1	0	0	0	0	0	0	0	0	0
26	noticed	1	0	0	0	0	0	0	0	0	0
27	B	1	0	0	0	0	0	0	0	0	0
28	&	1	0	0	1	0	0	0	0	0	0
29	W	1	0	0	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...
876	Olympian	0	0	0	0	0	0	0	0	0	1
877	stadium	0	0	0	0	0	0	0	0	0	1
878	Pleasure	0	0	0	0	0	0	0	0	0	1
879	Gardens	0	0	0	0	0	0	0	0	0	1
880	Below	0	0	0	0	0	0	0	0	0	1
881	workers	0	0	0	0	0	0	0	0	0	1
882	clocks	0	0	0	0	0	0	0	0	0	1
883	10	0	0	0	0	0	0	0	0	0	1
884	hours	0	0	0	0	0	0	0	0	0	1
885	squeeze	0	0	0	0	0	0	0	0	0	1
886	live	0	0	0	0	0	0	0	0	0	1
887	tenement	0	0	0	0	0	0	0	0	0	1
888	housing	0	0	0	0	0	0	0	0	0	1
889	consists	0	0	0	0	0	0	0	0	0	1
890	unrelenting	0	0	0	0	0	0	0	0	0	1
891	service	0	0	0	0	0	0	0	0	0	1
892	machine	0	0	0	0	0	0	0	0	0	1
893	plutocracy	0	0	0	0	0	0	0	0	0	1
894	vs.	0	0	0	0	0	0	0	0	0	1
895	labor	0	0	0	0	0	0	0	0	0	1
896	powerful	0	0	0	0	0	0	0	0	0	1
897	era	0	0	0	0	0	0	0	0	0	1
898	assembly	0	0	0	0	0	0	0	0	0	1

		0	1	2	3	4	5	6	7	8	9	10
899	line	0	0	0	0	0	0	0	0	0	0	1
900	large	0	0	0	0	0	0	0	0	0	0	1
901	scale	0	0	0	0	0	0	0	0	0	0	1
902	Marx	0	0	0	0	0	0	0	0	0	0	1
903	encouraged	0	0	0	0	0	0	0	0	0	0	1
904	class	0	0	0	0	0	0	0	0	0	0	1
905	warfare	0	0	0	0	0	0	0	0	0	0	1

906 rows × 11 columns

In [18]:

```
new_head=df.iloc[0]
df=df[1:]
df.columns=new_head
columns=['doc1','doc2','doc3','doc4','doc5','doc6','doc7','doc8','doc9','doc10']
df[columns]=df[columns].astype(int)
pd.to_numeric(columns,errors='coerce')
```

/Users/ishanarora/anaconda3/lib/python3.7/site-packages/pandas/core/frame.py:3381: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy> (<http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>)  
self[k1] = value[k2]

Out[18]:

```
array([nan, nan, nan, nan, nan, nan, nan, nan, nan, nan])
```

In [19]:

```
df['sum_rows']=df[columns].sum(axis=1)
```

/Users/ishanarora/anaconda3/lib/python3.7/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy> (<http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>)  
"""Entry point for launching an IPython kernel.



In [20]:

```
summation=['sum_cols']
for i in columns:
    summation.append(df[i].sum())
summation.append('')
```

In [21]:

```
df.loc[-1]=summation
```

/Users/ishanarora/anaconda3/lib/python3.7/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy> (<http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>)  
 """Entry point for launching an IPython kernel.

## Single Term present in set of documents (column sum)

In [22]:

```
f.set_index('Word',inplace=True)
print('Summation of rows(frequency a word in any document )and Summation of columns(frequency of words without stopwords in each document)')
f
```

Summation of rows(frequency a word in any document )and Summation of columns(frequency of words without stopwords in each document)

## Bag of words – complete set of terms

In [23]:

```
bag_words=set(s)
print(bag_words)
```

```
{'word', 'rival', 'Victor', 'new', 'prostitution', 'enough', 'personal
ity', 'essential', 'plays', 'restored', 'unrelenting', 'And', 'Russia
n', 'filmmaker', 'based', 'bulk', 'star', 'friends', 'Three', 'pier',
'Citizen', 'Buenos', 'alcoholic', 'hysterical', 'Martins', 'cumbersom
e', 'associations', 'could', 'trees', 'Gibson', 'Sadness', 'Kilimanjar
o', 'way', 'whose', 'Mean', 'filled', 'Puzo', 'talent', 'business', 'f
eelings', 'treasures', 'pushing', 'Amy', 'performed', 'always', 'simpl
y', 'series', 'think', 'actual', 'speaking', 'bullshit', 'shattered',
'Joseph', 'characters', 'family', 'officials', 'emotions', 'vibrates',
'slender', 'childhood', 'Cotton', 'directing', 'appearances', 'patriar
chy', 'suppress', 'perhaps', 'desolate', 'Gloria', 'short-sleeved', 'G
odfather', 'line', 'thing', 'material', 'avoids', 'record', 'card', 't
itle', 'premiered', 'stakes', 'parents', 'satisfactory', 'glorious',
'Richard', 'charm', 'war', 'look', 'missing', 'attempt', 'Inside', 'pr
oduction', 'Get', 'infuses', 'machine', 'doubt', 'Seven', 'see', 'char
ged', 'make', 'sets', 'Shooting', 'MGM', 'transience', 'classics', 'ga
mbling', 'meet', 'spend', 'finally', 'realistically', 'Norma', 'stadiu
m', 'mom', 'vision', 'really', 'sides', 'Harry', 'CGI', 'Technicolor',
'together', 'oft-maligned', 'Munchkin', 'installment', 'aging', 'stor
y', 'tossed', 'explained', '12', 'overwhelming', 'resonance', 'often',
'girlfriend', 'response', 'mystery', 'police', '--', '&', 'believe',
'perfectly', 'Desmond', 'groups', 'With', 'racist', 'plutocracy', 'lif
elong', 'shirt', 'would', 'perception', 'determined', 'works', 'victi
m', 'opening', 'Her', 'tilted', 'growing', 'remembers', 'lark', 'spark
ing', 'survivor', 'movie', 'sinkhole', "'ve", 'posture', 'grass', 'leo
pard', 'belong', 'walls', 'levels', 'degree', 'soft', 'lingers', 'decisi
on', 'original', 'undergraduate', 'Swanson', 'somewhere', 'Wind', 'f
ootsteps', 'performance', 'edge', 're-wrote', 'Dr.', 'channel', 'Bordw
ell', 'sympathetic', 'marks', 'Lime', 'paint', 'organized', 'expect',
'comedies', 'Garland', 'higher', 'directorial', 'originally', 'Wives',
'Karas', 'searchers', 'Stepford', 'assembly', 'About', 'someday', 'yea
rning', 'slacks', '?', 'consultant', 'mind', 'climbing', 'summer', 'wo
rld', 'fraud', 'feeling', 'real', 'predator', 'directors', 'tie', 'for
ward', 'justice', 'views', 'inviting', 'analyzed', 'cameras', 'Sunse
t', 'racial', 'obviously', 'suggested', 'evil', 'five', 'class', 'pape
r-thin', 'upping', 'Kansas', 'invitation', 'gothic', 'black', 'Kaitly
n', 'powerful', 'wrecked', 'King', 'comedy-adventure', 'chum', 'teneme
nt', 'know', 'Hayao', 'paradoxes', 'explore', 'character', 'expert',
'30', 'believed', 'Hardy', 'angles', 'Has', 'pretty', 'Wiene', '"', 't
wist', 'likes', 'It', 'Thorpe', 'life-threatening', 'warfare', 'actio
n', 'But', 'Michael', 'workers', 'smart', 'inside', 'Mel', 'traitors',
'split', 'sound', 'True', 'author', 'better', 'Dias', 'noticed', 'To',
'defeated', 'live', 'made', 'Never', 'Cabinet', 'post-apocalyptic', 's
ingular', 'Bette', 'remarkably', 'crisp', 'sympathize', 'corrupt', 'fl
ow', 'introduced', 'realistic', 'Olympian', 'mistreated', 'irony', 'an
imated', 'great', 'All', 'He', 'highly-anticipated', 'bitchy', 'actor
s', 'Disney', 'never', 'I', 'flat-topped', 'green', 'Francis', 'challe
nging', 'commensurate', 'act', 'unafraid', 'films', '2008', 'good', 'G
eorge', 'approaches', 'come', 'special', 'nominated', 'depressed', '19
39', 'hope', 'fireplug', 'returned', 'write', 'scrawny', '"', 'disapp
rove', 'Maybe', 'security', 'stirred', 'also', 'Out', 'brain', 'featur
e', 'fair', 'need', 'child', 'Carol', 'undertones', 'move', 'beauty',
'Yesterday', 'ones', 'superficial', 'windows', 'Wizard', 'American',
'home', 'must', 'footage', 'stumbling', 'large', '2001.', 'ambitious',
```

'masterpiece', 'within', 'elevated', 'During', 'Minnesota', 'arrives', 'crime', 'lot', 'entirely', 'college', 'neither', 'Foster', 'Mindy', 'jolt', 'nondescript', 'warehouse', 'overtly', 'Snow', 'tell', 'increasingly', 'None', 'Pixar', 'Pacino', 'postwar', 'right', 'hard-edged', 'adults', 'crazy', 'discovered', 'two-dimensional', 'Holly', 'kind', 'As', 'Boulevard', 'years', 'Don', 'silent', 'studio', 'monsters', 'assigned', '"', 'people', "n't", 'playing', 'returns', 'clearly', 'divided', 'truly', 'sent', 'No', 'vicious', 'older', 'dark', 'triumph', 'We', 'sterns', 'visceral', 'whistling', '11-year', 'behind', 'purple', 'atop', 'nothing', 'satirical', 'Anne', 'Because', 'spires', 'surface', 'suspicious', 'riffing', 'manifestation', 'stunt', 'devilish', 'first', 'Sundance', 'went', 'actress', 'Oscar', 'Cukor', 'jaunty', 'played', 'success', 'Davis', 'Hader', ':', 'almost', 'Its', 'We', 'taken', 'audiences', 'age-old', 'tells', 'long', 'Key', 'debut', 'fascinating', 'animation', 'explain', 'Riley', 'B', 'horror', 'essentially', 'woman', 'Metropolis', 'reveals', 'little', 'vote', 'women', 'clear', 'begins', 'took', 'civilian', 'torn', 'skin', 'labor', 'demonstration', 'innocent', 'immediately', 'future', 'effects', 'promises', 'level', 'jagged', 'Holiday', 'battered', 'modern', 'screen', 'deceptive', 'complex', 'Pelee', 'competitors', 'tone', 'alter', 'Pleasure', 'time', 'money-making', 'When', 'Vienna', 'veteran', 'Born', 'different', 'Joy', 'entire', 'greatest', 'officer', 'distortions', 'another', 'Below', 'reporter', 'foundation', 'Brando', 'question-mark', 'unwanted', 'These', 'Anger', 'anything.', 'totally', 'uncomfortable', 'film', 'W', 'school', 'Tinke', 'rbell', 'spell', 'younger', 'Third', '1950', 'completely', 'shot', 'mired', 'doldrums', 'role', 'Gone', 'curling', 'buried', 'night', 'since', 'still', 'consists', 'regain', 'last', 'Bill', 'wig', 'conjuring', 'commandment', 'remember', 'surpass', 'trapped', 'movies', 'Anyway', 'Beyond', 'protection', 'everyone', 'fact', 'Tom', 'ingenue', 'stands', 'Margo', 'Mafia', 'print', 'flames', 'scale', 'continued', 'Francisco', 'game', 'notice', 'revealed', 'fans', 'compare', 'exist', 'Max', 'towering', 'may', 'The', 'radical', 'seen', 'notices', 'Into', 'Oz', 'sharp', 'rules', '16mm', 'blue', 'interplay', 'Mad', 'brilliant', 'started', 'earlier', 'blown', 's', 'landscape', 'transformed', 'lose', 'risks', 'although', 'sacred', 'spoken', 'screening', 'cheap', 'grizzled', 'playful', 'professional', 'Channing', 'furnace', 'cadre', 'fun', 'Aires', 'city', 'began', 'zither', 'multiple', 'previous', 'cinematographer', 'Miller', 'desperate', 'vibe', 'Disgust', 'Gatsby', 'emerges', 'Almost', 'might', 'hours', 'depths', 'mostly', 'vs.', 'theft', 'This', 'Thompson', 'intrigue', 'Growing', 'history', 'Caligari', 'Film', 'Festival', 'Ford', 'filmmakers', 'unsettling', 'familiar', 'much', 'seeking', 'That', 'huge', 'Marlon', 'admirable', ';', 'bizarre', 'Dorothy', 'If', 'Coppola', '2017', 'consciously', 'switch', 'doors', 'camera', 'Judy', 'innate', 'adult', 'next', 'feels', 'fits', 'delusions', 'squeeze', 'autobiographical', '1927', 'significant', 'Lang', 'visuals', 'towers', 'looks', 'less', 'nobody', 'moment', 'Thunderdome', 'recessive', 'understanding', 'energy', 'Gardens', 'inhabit', 'career', 'ever', 'girl', '1937', 'famous', 'victory', 'today', 'personal', 'attempts', 'Kane', 'days', 'maybe', 'fat', 'Rosemary', 'bug-eyed', 'Robert', 'heard', 'criminal', 'wrong', 'best', 'San', 'Al', 'officially', 'life', 'Fritz', 'Above', 'knew', 'closeup', 'decades', 'terms', 'making', 'whether', 'ironically', 'genre', 'came', 'makeup', 'sledded', 'pulp', 'generation', 'wives', 'housing', 'physical', 'era', 'follow', 'dying', 'Then', 'reality', 'David', 'rackets', 'defined', 'storytelling', 'lost', 'wanted', 'white', 'consider', 'writer', 'color', 'music', 'fresh', 'told', 'factory', 'air', 'around', 'comfortable', 'single', 'clocks', 'Miyazaki', 'shaped', 'scenes', 'City', 'middle-manager', 'us', 'left', 'Lewis', 'experienced', 'Road', 'stood', 'young', 'replaced', 'There', 'parts', 'Vidor', 'Girls', 'White', 'get', 'Poe', 'hler', 'Corleone', 'From', 'French', 'separating', 'without', 'saw', 'three', '1920', 'Marx', 'staircases', 'light', 'fired', 'locations',

```
'`', 'east', 'puzzle', 'Anton', 'innocence', 'Rosebud', 'beats', 'indicated', 'What', 'one', 'wings', 'highways', 'entertainment', 'Fleming', 'apart', 'threat', 'Vito', 'something', 'cliched', 'secret', 'rich', 'hybrid', 'paired', 'says', 'madmen', 'quite', 'successful', 'diagonals', 'score', 'recasting', 'cartoonish', 'take', 'bone', 'play', 'leaves', 'feel', 'She', 'versus', 'villains', 'joy', 'boarding', 'accomplished', 'Man', 'sprite-type', 'age', 'suspect', 'blockbuster', 'British', 'learn', 'set', 'stylized', 'Eve', 'prints', 'mayhem', 'spiky', 'end', 'dad', 'closed', 'beerhouse', 'authoritarian', 'actor', '10', 'classic', 'cinema.', 'Charles', 'everything', 'His', 'like', 'willing', 'Jordan', 'performances', 'zones', 'Fear', 'Kaling', 'going', 'tendency', 'knows', 'vehicular', 'including', 'victims', 'prologue', 'pioneering', 'falls', 'Fury', 'explains', 'old', 'leading', 'Phyllis', 'Smith', 'Black', 'even', 'sequences', 'situations', 'public', 'service', 'red', 'claimed', 'stars', 'bit', 'knives', 'lives', 'Both', 'call', 'power', 'Baxter', 'franchise', 'murder', 'encouraged', 'top', 'director', 'Mario', 'Dwarfs', 'Park', 'Baby', 'similar', 'Reed', 'seems', 'work', 'expensive', '"', 'emblem', 'working', 'sense', 'show', 'knowing', 'man', 'suited', 'visage', 'waxwork', 'star-they', 'shots', 'downright', 'shake', 'ground'}
```

In [ ]: