```
In [2]: ! pip install nbconvert
```

```
Requirement already satisfied: nbconvert in c:\users\kisho\appdata\local\programs
\python\python312\lib\site-packages (7.16.4)
Requirement already satisfied: beautifulsoup4 in c:\users\kisho\appdata\local\pro
grams\python\python312\lib\site-packages (from nbconvert) (4.12.3)
Requirement already satisfied: bleach!=5.0.0 in c:\users\kisho\appdata\local\prog
rams\python\python312\lib\site-packages (from nbconvert) (6.1.0)
Requirement already satisfied: defusedxml in c:\users\kisho\appdata\local\program
s\python\python312\lib\site-packages (from nbconvert) (0.7.1)
Requirement already satisfied: jinja2>=3.0 in c:\users\kisho\appdata\local\progra
ms\python\python312\lib\site-packages (from nbconvert) (3.1.4)
Requirement already satisfied: jupyter-core>=4.7 in c:\users\kisho\appdata\local
\programs\python\python312\lib\site-packages (from nbconvert) (5.7.2)
Requirement already satisfied: jupyterlab-pygments in c:\users\kisho\appdata\loca
l\programs\python\python312\lib\site-packages (from nbconvert) (0.3.0)
Requirement already satisfied: markupsafe>=2.0 in c:\users\kisho\appdata\local\pr
ograms\python\python312\lib\site-packages (from nbconvert) (2.1.5)
Requirement already satisfied: mistune<4,>=2.0.3 in c:\users\kisho\appdata\local
\programs\python\python312\lib\site-packages (from nbconvert) (3.0.2)
Requirement already satisfied: nbclient>=0.5.0 in c:\users\kisho\appdata\local\pr
ograms\python\python312\lib\site-packages (from nbconvert) (0.10.0)
Requirement already satisfied: nbformat>=5.7 in c:\users\kisho\appdata\local\prog
rams\python\python312\lib\site-packages (from nbconvert) (5.10.4)
Requirement already satisfied: packaging in c:\users\kisho\appdata\local\programs
\python\python312\lib\site-packages (from nbconvert) (24.0)
Requirement already satisfied: pandocfilters>=1.4.1 in c:\users\kisho\appdata\loc
al\programs\python\python312\lib\site-packages (from nbconvert) (1.5.1)
Requirement already satisfied: pygments>=2.4.1 in c:\users\kisho\appdata\local\pr
ograms\python\python312\lib\site-packages (from nbconvert) (2.18.0)
Requirement already satisfied: tinycss2 in c:\users\kisho\appdata\local\programs
\python\python312\lib\site-packages (from nbconvert) (1.3.0)
Requirement already satisfied: traitlets>=5.1 in c:\users\kisho\appdata\local\pro
grams\python\python312\lib\site-packages (from nbconvert) (5.14.3)
Requirement already satisfied: six>=1.9.0 in c:\users\kisho\appdata\local\program
s\python\python312\lib\site-packages (from bleach!=5.0.0->nbconvert) (1.16.0)
Requirement already satisfied: webencodings in c:\users\kisho\appdata\local\progr
ams\python\python312\lib\site-packages (from bleach!=5.0.0->nbconvert) (0.5.1)
Requirement already satisfied: platformdirs>=2.5 in c:\users\kisho\appdata\local
\programs\python\python312\lib\site-packages (from jupyter-core>=4.7->nbconvert)
(4.2.1)
Requirement already satisfied: pywin32>=300 in c:\users\kisho\appdata\local\progr
ams\python\python312\lib\site-packages (from jupyter-core>=4.7->nbconvert) (306)
Requirement already satisfied: jupyter-client>=6.1.12 in c:\users\kisho\appdata\l
ocal\programs\python\python312\lib\site-packages (from nbclient>=0.5.0->nbconver
t) (8.6.1)
Requirement already satisfied: fastjsonschema>=2.15 in c:\users\kisho\appdata\loc
al\programs\python\python312\lib\site-packages (from nbformat>=5.7->nbconvert)
(2.19.1)
Requirement already satisfied: jsonschema>=2.6 in c:\users\kisho\appdata\local\pr
ograms\python\python312\lib\site-packages (from nbformat>=5.7->nbconvert) (4.22.
0)
Requirement already satisfied: soupsieve>1.2 in c:\users\kisho\appdata\local\prog
rams\python\python312\lib\site-packages (from beautifulsoup4->nbconvert) (2.5)
Requirement already satisfied: attrs>=22.2.0 in c:\users\kisho\appdata\local\prog
rams\python\python312\lib\site-packages (from jsonschema>=2.6->nbformat>=5.7->nbc
onvert) (23.2.0)
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in c:\users\k
isho\appdata\local\programs\python\python312\lib\site-packages (from jsonschema>=
2.6->nbformat>=5.7->nbconvert) (2023.12.1)
Requirement already satisfied: referencing>=0.28.4 in c:\users\kisho\appdata\loca
l\programs\python\python312\lib\site-packages (from jsonschema>=2.6->nbformat>=5.
```

```
7->nbconvert) (0.35.1)
Requirement already satisfied: rpds-py>=0.7.1 in c:\users\kisho\appdata\local\pro
grams\python\python312\lib\site-packages (from jsonschema>=2.6->nbformat>=5.7->nb
convert) (0.18.1)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\kisho\appdata\l
ocal\programs\python\python312\lib\site-packages (from jupyter-client>=6.1.12->nb
client>=0.5.0->nbconvert) (2.9.0.post0)
Requirement already satisfied: pyzmq>=23.0 in c:\users\kisho\appdata\local\progra
ms\python\python312\lib\site-packages (from jupyter-client>=6.1.12->nbclient>=0.
5.0->nbconvert) (26.0.3)
Requirement already satisfied: tornado>=6.2 in c:\users\kisho\appdata\local\progr
ams\python\python312\lib\site-packages (from jupyter-client>=6.1.12->nbclient>=0.
5.0->nbconvert) (6.4)
```

In [6]: `! pip install nbconvert[webpdf]`

```
Requirement already satisfied: nbconvert[webpdf] in c:\users\kisho\appdata\local
\programs\python\python312\lib\site-packages (7.16.4)
Requirement already satisfied: beautifulsoup4 in c:\users\kisho\appdata\local\pro
grams\python\python312\lib\site-packages (from nbconvert[webpdf]) (4.12.3)
Requirement already satisfied: bleach!=5.0.0 in c:\users\kisho\appdata\local\prog
rams\python\python312\lib\site-packages (from nbconvert[webpdf]) (6.1.0)
Requirement already satisfied: defusedxml in c:\users\kisho\appdata\local\program
s\python\python312\lib\site-packages (from nbconvert[webpdf]) (0.7.1)
Requirement already satisfied: jinja2>=3.0 in c:\users\kisho\appdata\local\progra
ms\python\python312\lib\site-packages (from nbconvert[webpdf]) (3.1.4)
Requirement already satisfied: jupyter-core>=4.7 in c:\users\kisho\appdata\local
\programs\python\python312\lib\site-packages (from nbconvert[webpdf]) (5.7.2)
Requirement already satisfied: jupyterlab-pygments in c:\users\kisho\appdata\loca
l\programs\python\python312\lib\site-packages (from nbconvert[webpdf]) (0.3.0)
Requirement already satisfied: markupsafe>=2.0 in c:\users\kisho\appdata\local\pr
ograms\python\python312\lib\site-packages (from nbconvert[webpdf]) (2.1.5)
Requirement already satisfied: mistune<4,>=2.0.3 in c:\users\kisho\appdata\local
\programs\python\python312\lib\site-packages (from nbconvert[webpdf]) (3.0.2)
Requirement already satisfied: nbclient>=0.5.0 in c:\users\kisho\appdata\local\pr
ograms\python\python312\lib\site-packages (from nbconvert[webpdf]) (0.10.0)
Requirement already satisfied: nbformat>=5.7 in c:\users\kisho\appdata\local\prog
rams\python\python312\lib\site-packages (from nbconvert[webpdf]) (5.10.4)
Requirement already satisfied: packaging in c:\users\kisho\appdata\local\programs
\python\python312\lib\site-packages (from nbconvert[webpdf]) (24.0)
Requirement already satisfied: pandocfilters>=1.4.1 in c:\users\kisho\appdata\loc
al\programs\python\python312\lib\site-packages (from nbconvert[webpdf]) (1.5.1)
Requirement already satisfied: pygments>=2.4.1 in c:\users\kisho\appdata\local\pr
ograms\python\python312\lib\site-packages (from nbconvert[webpdf]) (2.18.0)
Requirement already satisfied: tinycss2 in c:\users\kisho\appdata\local\programs
\python\python312\lib\site-packages (from nbconvert[webpdf]) (1.3.0)
Requirement already satisfied: traitlets>=5.1 in c:\users\kisho\appdata\local\pro
grams\python\python312\lib\site-packages (from nbconvert[webpdf]) (5.14.3)
Collecting playwright (from nbconvert[webpdf])
  Downloading playwright-1.50.0-py3-none-win_amd64.whl.metadata (3.5 kB)
Requirement already satisfied: six>=1.9.0 in c:\users\kisho\appdata\local\program
s\python\python312\lib\site-packages (from bleach!=5.0.0->nbconvert[webpdf]) (1.1
6.0)
Requirement already satisfied: webencodings in c:\users\kisho\appdata\local\progr
ams\python\python312\lib\site-packages (from bleach!=5.0.0->nbconvert[webpdf])
(0.5.1)
Requirement already satisfied: platformdirs>=2.5 in c:\users\kisho\appdata\local
\programs\python\python312\lib\site-packages (from jupyter-core>=4.7->nbconvert[w
ebpdf]) (4.2.1)
Requirement already satisfied: pywin32>=300 in c:\users\kisho\appdata\local\progr
ams\python\python312\lib\site-packages (from jupyter-core>=4.7->nbconvert[webpd
f]) (306)
Requirement already satisfied: jupyter-client>=6.1.12 in c:\users\kisho\appdata\l
ocal\programs\python\python312\lib\site-packages (from nbclient>=0.5.0->nbconvert
[webpdf]) (8.6.1)
Requirement already satisfied: fastjsonschema>=2.15 in c:\users\kisho\appdata\loc
al\programs\python\python312\lib\site-packages (from nbformat>=5.7->nbconvert[web
pdf]) (2.19.1)
Requirement already satisfied: jsonschema>=2.6 in c:\users\kisho\appdata\local\pr
ograms\python\python312\lib\site-packages (from nbformat>=5.7->nbconvert[webpdf])
(4.22.0)
Requirement already satisfied: soupsieve>1.2 in c:\users\kisho\appdata\local\prog
rams\python\python312\lib\site-packages (from beautifulsoup4->nbconvert[webpdf])
(2.5)
Collecting pyee<13,>=12 (from playwright->nbconvert[webpdf])
  Downloading pyee-12.1.1-py3-none-any.whl.metadata (2.9 kB)
```

```
Collecting greenlet<4.0.0,>=3.1.1 (from playwright->nbconvert[webpdf])
  Downloading greenlet-3.1.1-cp312-cp312-win_amd64.whl.metadata (3.9 kB)
Requirement already satisfied: attrs>=22.2.0 in c:\users\kisho\appdata\local\prog
rams\python\python312\lib\site-packages (from jsonschema>=2.6->nbformat>=5.7->nbc
onvert[webpdf]) (23.2.0)
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in c:\users\k
isho\appdata\local\programs\python\python312\lib\site-packages (from jsonschema>=
2.6->nbformat>=5.7->nbconvert[webpdf]) (2023.12.1)
Requirement already satisfied: referencing>=0.28.4 in c:\users\kisho\appdata\loca
l\programs\python\python312\lib\site-packages (from jsonschema>=2.6->nbformat>=5.
7->nbconvert[webpdf]) (0.35.1)
Requirement already satisfied: rpds-py>=0.7.1 in c:\users\kisho\appdata\local\pro
grams\python\python312\lib\site-packages (from jsonschema>=2.6->nbformat>=5.7->nb
convert[webpdf]) (0.18.1)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\kisho\appdata\l
ocal\programs\python\python312\lib\site-packages (from jupyter-client>=6.1.12->nb
client>=0.5.0->nbconvert[webpdf]) (2.9.0.post0)
Requirement already satisfied: pyzmq>=23.0 in c:\users\kisho\appdata\local\progra
ms\python\python312\lib\site-packages (from jupyter-client>=6.1.12->nbclient>=0.
5.0->nbconvert[webpdf]) (26.0.3)
Requirement already satisfied: tornado>=6.2 in c:\users\kisho\appdata\local\progr
ams\python\python312\lib\site-packages (from jupyter-client>=6.1.12->nbclient>=0.
5.0->nbconvert[webpdf]) (6.4)
Collecting typing-extensions (from pyee<13,>=12->playwright->nbconvert[webpdf])
  Downloading typing_extensions-4.12.2-py3-none-any.whl.metadata (3.0 kB)
Downloading playwright-1.50.0-py3-none-win_amd64.whl (34.8 MB)
   ------------------------------------- 0.0/34.8 MB ? eta -:--:--
   - ----------------------------------- 1.0/34.8 MB 7.2 MB/s eta 0:00:05
   -- ---------------------------------- 2.1/34.8 MB 6.5 MB/s eta 0:00:06
   --- --------------------------------- 3.4/34.8 MB 6.3 MB/s eta 0:00:05
   ----- ------------------------------- 4.7/34.8 MB 6.2 MB/s eta 0:00:05
   ------ ------------------------------ 6.0/34.8 MB 6.3 MB/s eta 0:00:05
   -------- ---------------------------- 7.3/34.8 MB 6.2 MB/s eta 0:00:05
   --------- --------------------------- 8.4/34.8 MB 6.2 MB/s eta 0:00:05
   ---------- -------------------------- 9.4/34.8 MB 6.0 MB/s eta 0:00:05
   ----------- ------------------------- 10.7/34.8 MB 6.0 MB/s eta 0:00:05
   ------------- ----------------------- 11.5/34.8 MB 5.8 MB/s eta 0:00:05
   ------------- ----------------------- 12.6/34.8 MB 5.7 MB/s eta 0:00:04
   -------------- ---------------------- 13.4/34.8 MB 5.6 MB/s eta 0:00:04
   --------------- --------------------- 14.4/34.8 MB 5.5 MB/s eta 0:00:04
   ---------------- -------------------- 15.7/34.8 MB 5.5 MB/s eta 0:00:04
   ----------------- ------------------- 17.0/34.8 MB 5.6 MB/s eta 0:00:04
   ------------------- ----------------- 18.4/34.8 MB 5.6 MB/s eta 0:00:03
   -------------------- ---------------- 19.4/34.8 MB 5.6 MB/s eta 0:00:03
   --------------------- --------------- 20.7/34.8 MB 5.7 MB/s eta 0:00:03
   ----------------------- ------------- 22.0/34.8 MB 5.7 MB/s eta 0:00:03
   ------------------------ ------------ 23.3/34.8 MB 5.7 MB/s eta 0:00:03
   ------------------------- ----------- 24.6/34.8 MB 5.7 MB/s eta 0:00:02
   -------------------------- ---------- 26.0/34.8 MB 5.8 MB/s eta 0:00:02
   ---------------------------- -------- 27.3/34.8 MB 5.8 MB/s eta 0:00:02
   ----------------------------- ------- 28.6/34.8 MB 5.8 MB/s eta 0:00:02
   ------------------------------ ----- 29.9/34.8 MB 5.8 MB/s eta 0:00:01
   ------------------------------ ---- 30.9/34.8 MB 5.8 MB/s eta 0:00:01
   -------------------------------- -- 32.2/34.8 MB 5.8 MB/s eta 0:00:01
   --------------------------------- - 33.3/34.8 MB 5.8 MB/s eta 0:00:01
   ----------------------------------- 34.6/34.8 MB 5.8 MB/s eta 0:00:01
   ----------------------------------- 34.8/34.8 MB 5.7 MB/s eta 0:00:00
Downloading greenlet-3.1.1-cp312-cp312-win_amd64.whl (299 kB)
Downloading pyee-12.1.1-py3-none-any.whl (15 kB)
Downloading typing_extensions-4.12.2-py3-none-any.whl (37 kB)
```

```
Installing collected packages: typing-extensions, greenlet, pyee, playwright
Successfully installed greenlet-3.1.1 playwright-1.50.0 pyee-12.1.1 typing-extens
ions-4.12.2
```

In [9]:
```python
import pandas as pd
```

In [2]:
```python
df = pd.read_csv("spotify_millsongdata.csv")
```

In [3]:
```python
df.head(5)
```

Out[3]:

	artist	song	link	text
0	ABBA	Ahe's My Kind Of Girl	/a/abba/ahes+my+kind+of+girl_20598417.html	Look at her face, it's a wonderful face \r\nA...
1	ABBA	Andante, Andante	/a/abba/andante+andante_20002708.html	Take it easy with me, please \r\nTouch me gen...
2	ABBA	As Good As New	/a/abba/as+good+as+new_20003033.html	I'll never know why I had to go \r\nWhy I had...
3	ABBA	Bang	/a/abba/bang_20598415.html	Making somebody happy is a question of give an...
4	ABBA	Bang-A-Boomerang	/a/abba/bang+a+boomerang_20002668.html	Making somebody happy is a question of give an...

In [38]:
```python
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5000 entries, 0 to 4999
Data columns (total 3 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   artist  5000 non-null   object
 1   song    5000 non-null   object
 2   text    5000 non-null   object
dtypes: object(3)
memory usage: 117.3+ KB
```

In [39]:
```python
df.isnull().sum()
```

Out[39]:
```
artist    0
song      0
text      0
dtype: int64
```

In [40]:
```python
df.describe()
```

	artist	song	text
count	5000	5000	5000
unique	606	4802	4999
top	Helloween	O Come All Ye Faithful	got a good reason \r for taking the easy way ...
freq	24	4	2

In [36]: `df.tail(5)`

Out[36]:

	artist	song	text
4995	Who	How Many Friends	i'm feelin' so good right now \r there's a ha...
4996	George Strait	Living And Living Well	had a nice little life \r a little boat a lit...
4997	Olivia Newton-John	Soul Kiss	soul kiss \r you left me dreaming \r now i w...
4998	Lynyrd Skynyrd	God And Guns	last night i heard this politician \r talking...
4999	W.A.S.P.	Mean Man	a tatooed madman, i'm hell on wheels \r born ...

In [5]: `df.shape`

Out[5]: `(57650, 4)`

In [6]: `df.isnull().sum()`

Out[6]:
```
artist    0
song      0
link      0
text      0
dtype: int64
```

In [41]: `df_cleaned = df.dropna()`

In [7]: `df =df.sample(5000).drop('link', axis=1).reset_index(drop=True)`

In [8]: `df.head(10)`

Out[8]:

	artist	song	text
0	Electric Light Orchestra	10538 Overture	Did you see your friend, crying from his eyes ...
1	Pearl Jam	Last Kiss	Oh, where oh where can my baby be? \r\nThe Lo...
2	Nazareth	Desolation Road	Locked up on Monday \r\nLord I lost my way \...
3	Next To Normal	So Anyway	Diana: \r\nSo anyway I'm leaving \r\nI thoug...
4	Don Henley	Watching Tv	We were watchin' tv...watchin'tv \r\nWe were ...
5	Crowded House	Fall At Your Feet	I'm really close tonight \r\nAnd I feel like ...
6	Ingrid Michaelson	Without You	I did what I did for you I did what I could \...
7	Pharrell Williams	Come Get It Bae	You miss me? \r\nI miss all of y'all \r\nAll...
8	GMB	Holy	Father we've come to seek Your face \r\nDespe...
9	Korn	Make Me Bad	I am watching the rise \r\nAnd fall of my sal...

In [9]:
```python
df['text'][0]
```

Out[9]:
'Did you see your friend, crying from his eyes today? \r\nDid you see him run through the streets and far way? \r\n \r\nDid you see him run? Did you see him fall? \r\nDid his life flash by on the bedroom door? \r\n \r\nDid you hear the news came across the air today? \r\nSomeone has been found on the rocks down in the bay. \r\n \r\nDid you see him hide? Did you see him crawl? \r\nDoes his life mean more than it did before. \r\n \r\nDid you see that man running through the streets today? \r\nDid you catch his face, was it 10538? \r\n\r\n\r\n'

In [10]:
```python
df.shape
```

Out[10]: (5000, 3)

In [11]:
```python
df['text'] = df['text'].str.lower().replace(r'^\w\s', ' ').replace(r'\n', ' ', r
```

In [21]:
```python
import nltk
nltk.download('punkt')
```

Out[21]: True

In [29]:
```python
import nltk
from nltk.tokenize import word_tokenize
from nltk.stem import PorterStemmer

stemmer = PorterStemmer()

def token(txt):
    tokens = word_tokenize(txt)  # Tokenize text
```

```
        stemmed = [stemmer.stem(w) for w in tokens]   # Stem words
        return " ".join(stemmed)
```

In [ ]:
```
df['text'] = df['text'].apply(lambda x: token(x))
```

In [ ]:
```
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics.pairwise import cosine_similarity
```

In [ ]:
```
tfidvector = TfidfVectorizer(analyzer='word',stop_words='english')
matrix = tfidvector.fit_transform(df['text'])
similarity = cosine_similarity(matrix)
```

In [ ]:
```
similer[0]
```

In [ ]:
```
df[df['song'] == 'Crying Over You']
```

In [ ]:
```
def recommendation(song_df):
    idx = df[df['song'] == song_df].index[0]
    distances = sorted(list(enumerate(similarity[idx])),reverse=True,key=lambda

    songs = []
    for m_id in distances[1:21]:
        songs.append(df.iloc[m_id[0]].song)

    return songs
```

In [ ]:
```
recommendation('Crying Over You')
```

In [ ]:
```
import pickle
pickle.dump(similarity,open('similarity.pkl','wb'))
pickle.dump(df,open('df.pkl','wb'))
```