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

```python
In [2]:  1  df = pd.read_csv('1000_popular_hindi_spotify.csv')
         2  print(df)
         3
```

```
                          ID                          Name  Popularity  \
0      4nc6XiUze2Yh7wFueGOPv7         Chaleya (From "Jawan")          83
1      2wAiFWjRupWmnDkQcu91MF                Rumah Ke Rumah          80
2      7J0isBrUxhIYZVdrBOOlIh                        Cincin          78
3      2dIBMHByUGcNPzmYBJ6OAj                      Evaluasi          78
4      5drW6PGRxkE6MxttzVLNk5                    Secukupnya          76
..                        ...                           ...         ...
945    2acarfkiqwEbrGR4Atben0              Deja Vu - Remix            0
946    6iqx6vZNsbf5CzbWeYE8Sa            Halla Macha - Remix          0
947    5HphButx0MoTcdGqWmpV2R               Laadla - Remix            0
948    7e1cCE3J68Tpc5MIczlWY8  Gabbar Bhi Nachega - Remix            0
949    1SfHL4otGzy02RtBcetlyi                Da One - Remix            0

                                       Artists
0      Anirudh Ravichander, Arijit Singh, Shilpa Rao,...
1                                         Hindia
2                                         Hindia
3                                         Hindia
4                                         Hindia
..                                          ...
945        Hindi Song, Hindi Music, Hindi Remix Group
946        Hindi Song, Hindi Music, Hindi Remix Group
947        Hindi Song, Hindi Music, Hindi Remix Group
948        Hindi Song, Hindi Music, Hindi Remix Group
949        Hindi Song, Hindi Music, Hindi Remix Group

[950 rows x 4 columns]
```

```python
In [3]:  1
         2  df.head()
```

Out[3]:

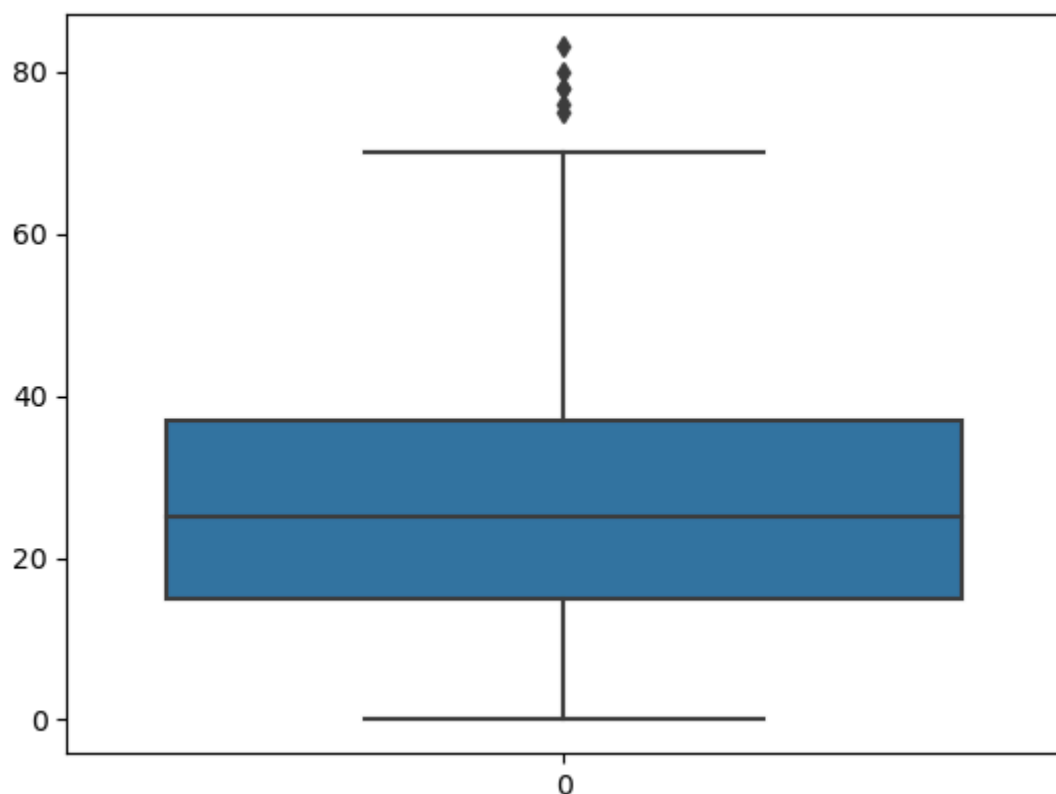|	ID	Name	Popularity	Artists
0	4nc6XiUze2Yh7wFueGOPv7	Chaleya (From "Jawan")	83	Anirudh Ravichander, Arijit Singh, Shilpa Rao,...
1	2wAiFWjRupWmnDkQcu91MF	Rumah Ke Rumah	80	Hindia
2	7J0isBrUxhIYZVdrBOOlIh	Cincin	78	Hindia
3	2dIBMHByUGcNPzmYBJ6OAj	Evaluasi	78	Hindia
4	5drW6PGRxkE6MxttzVLNk5	Secukupnya	76	Hindia

```python
In [4]:  1  df.dtypes
```

```
Out[4]:  ID             object
         Name           object
         Popularity      int64
         Artists        object
         dtype: object
```

```
In [5]:   1  df.head(3)
          2
```

Out[5]:

	ID	Name	Popularity	Artists
0	4nc6XiUze2Yh7wFueGOPv7	Chaleya (From "Jawan")	83	Anirudh Ravichander, Arijit Singh, Shilpa Rao,...
1	2wAiFWjRupWmnDkQcu91MF	Rumah Ke Rumah	80	Hindia
2	7J0isBrUxhIYZVdrBOOllh	Cincin	78	Hindia

```
In [6]:   1  print(df.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 950 entries, 0 to 949
Data columns (total 4 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   ID          950 non-null    object
 1   Name        950 non-null    object
 2   Popularity  950 non-null    int64
 3   Artists     950 non-null    object
dtypes: int64(1), object(3)
memory usage: 29.8+ KB
None
```

```
In [7]:   1
          2  import seaborn as sns
```

```
In [8]:   1  sns.boxplot(df['Popularity'])
```

Out[8]: <Axes: >

```
In [9]:    1  from scipy.stats import iqr
           2  iqr_value = iqr(df['Popularity'])
           3
```

```
In [10]:   1
           2  iqr_value
```

Out[10]: 22.0

```
In [11]:   1  df.isnull().sum()
           2
```

Out[11]: ID   0
         Name   0
         Popularity   0
         Artists   0
         dtype: int64

```
In [12]:   1
           2  df['Popularity']
           3
```

Out[12]: 0   83
         1   80
         2   78
         3   78
         4   76
                ..
         945   0
         946   0
         947   0
         948   0
         949   0
         Name: Popularity, Length: 950, dtype: int64

```
In [13]:   1  new_df=df[['Name','Artists']]
           2  new_df
```

Out[13]:

	Name	Artists
0	Chaleya (From "Jawan")	Anirudh Ravichander, Arijit Singh, Shilpa Rao,...
1	Rumah Ke Rumah	Hindia
2	Cincin	Hindia
3	Evaluasi	Hindia
4	Secukupnya	Hindia
...	...	...
945	Deja Vu - Remix	Hindi Song, Hindi Music, Hindi Remix Group
946	Halla Macha - Remix	Hindi Song, Hindi Music, Hindi Remix Group
947	Laadla - Remix	Hindi Song, Hindi Music, Hindi Remix Group
948	Gabbar Bhi Nachega - Remix	Hindi Song, Hindi Music, Hindi Remix Group
949	Da One - Remix	Hindi Song, Hindi Music, Hindi Remix Group

950 rows × 2 columns

```
In [14]:  1
          2  new_df['Artists']=new_df['Artists'].apply(lambda x:x.lower())
          3  new_df
```

C:\Users\user\AppData\Local\Temp\ipykernel_7684\3266672334.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: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)
  new_df['Artists']=new_df['Artists'].apply(lambda x:x.lower())

Out[14]:

	Name	Artists
0	Chaleya (From "Jawan")	anirudh ravichander, arijit singh, shilpa rao,...
1	Rumah Ke Rumah	hindia
2	Cincin	hindia
3	Evaluasi	hindia
4	Secukupnya	hindia
...	...	...
945	Deja Vu - Remix	hindi song, hindi music, hindi remix group
946	Halla Macha - Remix	hindi song, hindi music, hindi remix group
947	Laadla - Remix	hindi song, hindi music, hindi remix group
948	Gabbar Bhi Nachega - Remix	hindi song, hindi music, hindi remix group
949	Da One - Remix	hindi song, hindi music, hindi remix group

950 rows × 2 columns

```
In [15]:  1  from sklearn.feature_extraction.text import CountVectorizer
          2  cv=CountVectorizer(max_features=2000)
```

```
In [16]:  1  vectors=cv.fit_transform(new_df['Artists']).toarray()
          2
```

```
In [17]:  1  vectors.shape
          2
```

Out[17]: (950, 1113)

```
In [18]:  1  from sklearn.metrics.pairwise import cosine_similarity
          2  similarity=cosine_similarity(vectors)
```

```
In [19]:   1  sorted(list(enumerate(similarity[0])),reverse=True,key=lambda x:x[1])
```

```
Out[19]: [(0, 0.9999999999999998),
          (5, 0.5345224838248487),
          (320, 0.5345224838248487),
          (331, 0.5345224838248487),
          (475, 0.5345224838248487),
          (419, 0.4364357804719848),
          (129, 0.4285714285714285),
          (295, 0.4285714285714285),
          (225, 0.3779644730092272),
          (343, 0.3779644730092272),
          (361, 0.3779644730092272),
          (397, 0.3779644730092272),
          (517, 0.3779644730092272),
          (86, 0.3380617018914066),
          (294, 0.3380617018914066),
          (239, 0.3086066999241838),
          (242, 0.3086066999241838),
          (164, 0.28571428571428564),
          (155, 0.26726124191242434),
```

```
In [20]:   1  new_df.rename(columns={'Name':'Artists'},inplace=True)
```

C:\Users\user\AppData\Local\Temp\ipykernel_7684\4270318282.py:1: SettingWi
thCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-doc
s/stable/user_guide/indexing.html#returning-a-view-versus-a-copy (https://
pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-
view-versus-a-copy)
  new_df.rename(columns={'Name':'Artists'},inplace=True)

```
In [21]:   1  def recommend(music):
           2      music_index=new_df[new_df['Artists']==music].index[0]
           3      distances=similarity[music_index]
           4      music_list=sorted(list(enumerate(distances)),reverse=True,key=lambd
           5      for i in music_list:
           6          print(new_df.iloc[i[0]].Artists)
           7
           8
```

```
In [22]:  1  recommend('Gabbar Bhi Nachega - Remix    ')
```

```
Artists         Tum Hi Ho
Artists     arijit singh
Name: 5, dtype: object
Artists         Lokiverse 2.0
Artists     anirudh ravichander
Name: 320, dtype: object
Artists     Tum Hi Ho (From "Aashiqui 2")
Artists                        arijit singh
Name: 331, dtype: object
Artists     Dher Tak Chala Hindi
Artists             arijit singh
Name: 475, dtype: object
Artists     Sanam Re (From "Sanam Re")
Artists           mithoon, arijit singh
Name: 419, dtype: object
```

```
In [23]:    1  df.head(50)
```

	ID	Name	Popularity	Artists
0	4nc6XiUze2Yh7wFueGOPv7	Chaleya (From "Jawan")	83	Anirudh Ravichander, Arijit Singh, Shilpa Rao,...
1	2wAiFWjRupWmnDkQcu91MF	Rumah Ke Rumah	80	Hindia
2	7J0isBrUxhIYZVdrBOOIlh	Cincin	78	Hindia
3	2dIBMHByUGcNPzmYBJ6OAj	Evaluasi	78	Hindia
4	5drW6PGRxkE6MxttzVLNk5	Secukupnya	76	Hindia
5	56zZ48jdyY2oDXHVnwg5Di	Tum Hi Ho	75	Arijit Singh
6	3CBxVM0zGj41BQtE6m7gwb	Membasuh	70	Hindia, Rara Sekar
7	1bcIiU5FWRa2IYBqx0bk67	Evakuasi	69	Hindia
8	5d6K5YMfby3mRfsTeTxUV3	Laagi Lagan Shankara - Hindi	68	Hansraj Raghuwanshi
9	3kj14iUwv0rvTweOr9P2ut	Mata Air	68	Hindia, Natasha Udu, Kamga
10	3F0I9AkbRwz12yz7Hxn0bF	Oh Ho Ho Ho (Remix)	66	Sukhbir, Ikka
11	2SmKwCCSiUTjtR8I8rv6MS	Berdansalah, Karir Ini Tak Ada Artinya	66	Hindia
12	46ION8zSSaelA2XI44HxRL	Hindi Na Nga	65	This Band
13	64G8ywuyGGQl4S7U2a6XJf	Hindi Tayo Pwede	64	The Juans
14	7CXBhqBSbHNgecVX6aguUp	Sa Hindi Pag-Alala	64	Munimuni
15	08OPqLv99g8avzmxQepmiw	Besok Mungkin Kita Sampai	64	Hindia
16	2LQZ32XuFlkcoEUDSG6BMT	Untuk Apa / Untuk Apa?	64	Hindia
17	1XabvPK1VQEH4YqzDovs46	Forgot Password	63	Hindia, Nadin Amizah
18	1AAmQJO43qsQGc1pWRmTxN	Beautiful Tango - Unplugged	62	Hindi Zahra
19	3QZgiXw3VyeSfFnhY7zpj1	Bayangkan Jika Kita Tidak Menyerah	62	Hindia
20	44QCycNgqPxNaij2BrA9H3	Chhupana Bhi Nahin Aata - From "Baazigar-Hindi"	62	Vinod Rathod
21	5Nd4DTlDWYkkZfi1NpgvJk	Suit Suit	61	Guru Randhawa, Arjun
22	70fA3vjK2r6kpHrN6adr9q	Ride It (Kya Yehi Pyaar Hai) - Hindi Version	61	Jay Sean
23	1ZGLauVoD74BU8y8O5btcO	Zakir	59	NAALAYAK
24	5DqSZFvWvSWg5eyCDi0CuW	Kita Ke Sana	58	Hindia
25	4FERIZkT9xabNYCgTyMErn	Hindi Magbabago	57	Randy Santiago
26	26iZgiCSejmjrSBoimJbxu	Imik Si Mik - Remastered Version	57	Hindi Zahra
27	0JkJLhizW12Ii6KJNFDryx	Nabi Palsu	57	Hindia
28	1VTjpFsZhnKvEgZhlzPIzb	Ramai Sepi Bersama	57	Hindia
29	4RGeIxGrN9VH1TJ94ppyje	Secukupnya	57	Hindia
30	4mjTMmz0j6WI3GhODoPqTx	Hoor	57	Atif Aslam, Sachin-Jigar
31	7eA7fM538sgco6UUAQcnYA	The Man I Love - Unplugged	56	Hindi Zahra

	ID	Name	Popularity	Artists
32	5zzDbrXfdajBqVt1jMAMNM	Hanuman Chalisa (From "HanuMan") [Hindi]	56	GowraHari, Sai Charan
33	4qOpQXiEDEvt8MhYCGcwB6	Dehidrasi	56	Hindia, Petra Sihombing
34	4YSOIJJoFu4vfZYJUiewl9	Apapun Yang Terjadi	56	Hindia
35	5dFyHpANHOHWdAWQDboy0w	Janji Palsu	56	Hindia
36	0vPisMufQ5Ht1M5LaUlohP	Matahari Tenggelam	56	Hindia
37	5iezEsWCl1JIyvosnI7vx8	Evaluasi	55	Hindia
38	7BkTkOU0JJg4KklnNRmTXu	Janji Palsu	55	Hindia
39	34dTSIA1q3JMg59gMwQeOb	Masalah Masa Depan	55	Hindia
40	2w9GgxOBuUHJEnRuQBdpWA	O Lal Dupatte Wali	55	Kumar Sanu, Sudesh Bhosle, Kavita Krishnamurthy
41	6lhtQoh8eIVn3ztvcE1axU	Any Story	54	Hindi Zahra
42	1A0SVbPYh76qzwtnlAkgH5	Alexandra	54	Hindia
43	5GMJVKfh0TbV1TjqBUa2et	Masalah Masa Depan	54	Hindia
44	6BijMn5VMOaOBuUMKbQS1W	Hindi Ikaw	53	Kenaniah
45	47DgfSmVQVwHkJJct5c3sV	Hindi Tayo Pwede	53	The Juans
46	0ufk0A6oErK7oAgeMZWJFg	Belum Tidur	53	Hindia, Sal Priadi
47	00kuZKKbta1ZqQF9JaK3Ep	Matahari Tenggelam	53	Hindia
48	3SNRMAmRlwswe6k8i5ARiC	Membasuh	53	Hindia, Rara Sekar
49	0LtToIySj2geRJFEKcMsjp	Dehidrasi	53	Hindia, Petra Sihombing

In [ ]: `1`