

```
import pandas as pd

import numpy as np

df =
pd.read_csv(r'https://raw.githubusercontent.com/Sarvesh-S-Patil/Dataset/main/Movies%20Recommendation.csv')

```

```
df.head()
```

	Movie_ID	Movie_Title	Movie_Genre
0	1	Four Rooms	Crime Comedy
1	2	Star Wars	Adventure Action Science Fiction
2	3	Finding Nemo	Animation Family
3	4	Forrest Gump	Comedy Drama Romance
4	5	American Beauty	Drama

	Movie_Budget	Movie_Popularity	Movie_Release_Date	Movie_Revenue
0	4000000	22.876230	09-12-1995	4300000
1	11000000	126.393695	25-05-1977	775398007
2	94000000	85.688789	30-05-2003	940335536
3	55000000	138.133331	06-07-1994	677945399
4	15000000	80.878605	15-09-1999	356296601

	Movie_Runtime	Movie_Vote	...
0	98.0	6.5	...
1	121.0	8.1	...
2	100.0	7.6	...
3	142.0	8.2	...
4	122.0	7.9	...

	Movie_Homepage
0	NaN
1	<a href="http://www.starwars.com/films/star-wars-episod...">http://www.starwars.com/films/star-wars-episod...</a>
2	<a href="http://movies.disney.com/finding-nemo">http://movies.disney.com/finding-nemo</a>
3	NaN
4	<a href="http://www.dreamworks.com/ab/">http://www.dreamworks.com/ab/</a>

	Movie_Keywords
0	hotel new year's eve witch bet hotel room
1	android galaxy hermit death star lightsaber
2	father son relationship harbor underwater fish...
3	vietnam veteran hippie mentally disabled runni...
4	male nudity female nudity adultery midlife cri...

```

Movie_Overview \
0 It's Ted the Bellhop's first night on the job....
1 Princess Leia is captured and held hostage by ...
2 Nemo, an adventurous young clownfish, is unexp...
3 A man with a low IQ has accomplished great thi...
4 Lester Burnham, a depressed suburban father in...

```

```

Movie_Production_House \
0 [{"name": "Miramax Films", "id": 14}, {"name": "...
1 [{"name": "Lucasfilm", "id": 1}, {"name": "Twe...
2 [{"name": "Pixar Animation Studios", "id": 3}]
3 [{"name": "Paramount Pictures", "id": 4}]
4 [{"name": "DreamWorks SKG", "id": 27}, {"name": "...

```

```

Movie_Production_Country \
0 [{"iso_3166_1": "US", "name": "United States o...
1 [{"iso_3166_1": "US", "name": "United States o...
2 [{"iso_3166_1": "US", "name": "United States o...
3 [{"iso_3166_1": "US", "name": "United States o...
4 [{"iso_3166_1": "US", "name": "United States o...

```

```

Movie_Spoken_Language \
0 [{"iso_639_1": "en", "name": "English"}]
1 [{"iso_639_1": "en", "name": "English"}]
2 [{"iso_639_1": "en", "name": "English"}]
3 [{"iso_639_1": "en", "name": "English"}]
4 [{"iso_639_1": "en", "name": "English"}]

```

```

Movie_Tagline \
0 Twelve outrageous guests. Four scandalous requ...
1 A long time ago in a galaxy far, far away...
2 There are 3.7 trillion fish in the ocean, they...
3 The world will never be the same, once you've ...
4 Look closer.

```

```

Movie_Cast \
0 Tim Roth Antonio Banderas Jennifer Beals Madōn...
1 Mark Hamill Harrison Ford Carrie Fisher Peter ...
2 Albert Brooks Ellen DeGeneres Alexander Gould ...
3 Tom Hanks Robin Wright Gary Sinise Mykelti Wil...
4 Kevin Spacey Annette Bening Thora Birch Wes Be...

```

	Movie_Crew	Movie_Director
0	[{'name': 'Allison Anders', 'gender': 1, 'depa...	Allison Anders
1	[{'name': 'George Lucas', 'gender': 2, 'depart...	George Lucas
2	[{'name': 'Andrew Stanton', 'gender': 2, 'depa...	Andrew Stanton

```
3  [{'name': 'Alan Silvestri', 'gender': 2, 'depa... Robert Zemeckis
```

```
4  [{'name': 'Thomas Newman', 'gender': 2, 'depar... Sam Mendes
```

```
[5 rows x 21 columns]
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 4760 entries, 0 to 4759
```

```
Data columns (total 21 columns):
```

#	Column	Non-Null Count	Dtype
0	Movie_ID	4760 non-null	int64
1	Movie_Title	4760 non-null	object
2	Movie_Genre	4760 non-null	object
3	Movie_Language	4760 non-null	object
4	Movie_Budget	4760 non-null	int64
5	Movie_Popularity	4760 non-null	float64
6	Movie_Release_Date	4760 non-null	object
7	Movie_Revenue	4760 non-null	int64
8	Movie_Runtime	4758 non-null	float64
9	Movie_Vote	4760 non-null	float64
10	Movie_Vote_Count	4760 non-null	int64
11	Movie_Homepage	1699 non-null	object
12	Movie_Keywords	4373 non-null	object
13	Movie_Overview	4757 non-null	object
14	Movie_Production_House	4760 non-null	object
15	Movie_Production_Country	4760 non-null	object
16	Movie_Spoken_Language	4760 non-null	object
17	Movie_Tagline	3942 non-null	object
18	Movie_Cast	4733 non-null	object
19	Movie_Crew	4760 non-null	object
20	Movie_Director	4738 non-null	object

```
dtypes: float64(3), int64(4), object(14)
```

```
memory usage: 781.1+ KB
```

```
df
```

	Movie_ID	Movie_Title	
Movie_Genre \			
0	1	Four Rooms	Crime
Comedy			
1	2	Star Wars	Adventure Action Science
Fiction			
2	3	Finding Nemo	Animation
Family			
3	4	Forrest Gump	Comedy Drama
Romance			

4	5	American Beauty	
Drama			
...	...	...	
...			
4755	4784	Midnight Cabaret	
Horror			
4756	4785	Growing Up Smith	Comedy
Family Drama			
4757	4786	8 Days	
Thriller Drama			
4758	4787	Running Forever	
Family			
4759	4788	To Be Frank, Sinatra at 100	
Documentary			

	Movie_Language	Movie_Budget	Movie_Popularity	Movie_Release_Date
\				
0	en	4000000	22.876230	09-12-1995
1	en	11000000	126.393695	25-05-1977
2	en	94000000	85.688789	30-05-2003
3	en	55000000	138.133331	06-07-1994
4	en	15000000	80.878605	15-09-1999
...	...	...	...	...
4755	en	0	0.001389	01-01-1990
4756	en	0	0.710870	03-02-2017
4757	en	0	0.015295	15-06-2014
4758	en	0	0.028756	27-10-2015
4759	en	2	0.050625	12-12-2015

	Movie_Revenue	Movie_Runtime	Movie_Vote	...	\
0	4300000	98.0	6.5	...	
1	775398007	121.0	8.1	...	
2	940335536	100.0	7.6	...	
3	677945399	142.0	8.2	...	
4	356296601	122.0	7.9	...	
...	...	...	...	...	
4755	0	94.0	0.0	...	
4756	0	102.0	7.4	...	

4757	0	90.0	0.0	...
4758	0	88.0	0.0	...
4759	0	NaN	0.0	...

		Movie_Homepage	\
0		NaN	
1	http://www.starwars.com/films/star-wars-episod...		
2	http://movies.disney.com/finding-nemo		
3		NaN	
4	http://www.dreamworks.com/ab/		
...		...	
4755		NaN	
4756	http://www.growingupsmithmovie.com		
4757		NaN	
4758		NaN	
4759		NaN	

		Movie_Keywords	\
0	hotel new year's eve witch bet hotel room		
1	android galaxy hermit death star lightsaber		
2	father son relationship harbor underwater fish...		
3	vietnam veteran hippie mentally disabled runni...		
4	male nudity female nudity adultery midlife cri...		
...		...	
4755		NaN	
4756		NaN	
4757	christian film sex trafficking		
4758		NaN	
4759	music actors legendary perfomer classic hollyw...		

		Movie_Overview	\
0	It's Ted the Bellhop's first night on the job....		
1	Princess Leia is captured and held hostage by ...		
2	Nemo, an adventurous young clownfish, is unexp...		
3	A man with a low IQ has accomplished great thi...		
4	Lester Burnham, a depressed suburban father in...		
...		...	
4755	A Broadway producer puts on a play with a Devi...		
4756	In 1979, an Indian family moves to America wit...		
4757	After sneaking to a party with her friends, 16...		
4758	After being estranged since her mother's death...		
4759		NaN	

		Movie_Production_House	\
0	[{"name": "Miramax Films", "id": 14}, {"name": ...		
1	[{"name": "Lucasfilm", "id": 1}, {"name": "Twe...		
2	[{"name": "Pixar Animation Studios", "id": 3}]		
3	[{"name": "Paramount Pictures", "id": 4}]		
4	[{"name": "DreamWorks SKG", "id": 27}, {"name": ...		
...		...	

```

4755 []
4756 []
4757 [{"name": "After Eden Pictures", "id": 85248}]
4758 [{"name": "New Kingdom Pictures", "id": 41671}]
4759 [{"name": "Eyeline Entertainment", "id": 60343}]

```

```

Movie_Production_Country \
0 [{"iso_3166_1": "US", "name": "United States o...
1 [{"iso_3166_1": "US", "name": "United States o...
2 [{"iso_3166_1": "US", "name": "United States o...
3 [{"iso_3166_1": "US", "name": "United States o...
4 [{"iso_3166_1": "US", "name": "United States o...
...
4755 []
4756 []
4757 [{"iso_3166_1": "US", "name": "United States o...
4758 [{"iso_3166_1": "US", "name": "United States o...
4759 [{"iso_3166_1": "GB", "name": "United Kingdom"}]

```

```

Movie_Spoken_Language \
0 [{"iso_639_1": "en", "name": "English"}]
1 [{"iso_639_1": "en", "name": "English"}]
2 [{"iso_639_1": "en", "name": "English"}]
3 [{"iso_639_1": "en", "name": "English"}]
4 [{"iso_639_1": "en", "name": "English"}]
...
4755 [{"iso_639_1": "en", "name": "English"}]
4756 [{"iso_639_1": "en", "name": "English"}]
4757 [{"iso_639_1": "en", "name": "English"}]
4758 []
4759 []

```

```

Movie_Tagline \
0 Twelve outrageous guests. Four scandalous requ...
1 A long time ago in a galaxy far, far away...
2 There are 3.7 trillion fish in the ocean, they...
3 The world will never be the same, once you've ...
4 Look closer.
...
4755 The hot spot where Satan's waitin'.
4756 It's better to stand out than to fit in.
4757 She never knew it could happen to her...
4758 NaN
4759 NaN

```

```

Movie_Cast \
0 Tim Roth Antonio Banderas Jennifer Beals Madon...
1 Mark Hamill Harrison Ford Carrie Fisher Peter ...
2 Albert Brooks Ellen DeGeneres Alexander Gould ...
3 Tom Hanks Robin Wright Gary Sinise Mykelti Wil...

```

```

4      Kevin Spacey Annette Bening Thora Birch Wes Be...
...
4755 Lisa Hart Carroll Michael Des Barres Paul Drak...
4756 Roni Akurati Brighton Sharbino Jason Lee Anjul...
4757 Nicole Smolen Kim Baldwin Ariana Stephens Brys...
4758                                     NaN
4759                                Tony Oppedisano

```

```

Movie_Crew
Movie_Director
0      [{'name': 'Allison Anders', 'gender': 1, 'depa... Allison
Anders
1      [{'name': 'George Lucas', 'gender': 2, 'depart... George
Lucas
2      [{'name': 'Andrew Stanton', 'gender': 2, 'depa... Andrew
Stanton
3      [{'name': 'Alan Silvestri', 'gender': 2, 'depa... Robert
Zemeckis
4      [{'name': 'Thomas Newman', 'gender': 2, 'depar... Sam
Mendes
...
...
4755  [{'name': 'Pece Dingo', 'gender': 0, 'departme... Pece
Dingo
4756  [{'name': 'Frank Lotito', 'gender': 0, 'depart... Frank
Lotito
4757  [{'name': 'Jaco Booyens', 'gender': 0, 'depart... Jaco
Booyens
4758                                     []
NaN
4759  [{'name': 'Simon Napier-Bell', 'gender': 0, 'd... Simon Napier-
Bell

```

```
[4760 rows x 21 columns]
```

```
df.shape
```

```
(4760, 21)
```

```
df.columns
```

```

Index(['Movie_ID', 'Movie_Title', 'Movie_Genre', 'Movie_Language',
      'Movie_Budget', 'Movie_Popularity', 'Movie_Release_Date',
      'Movie_Revenue', 'Movie_Runtime', 'Movie_Vote',
      'Movie_Vote_Count',
      'Movie_Homepage', 'Movie_Keywords', 'Movie_Overview',
      'Movie_Production_House', 'Movie_Production_Country',
      'Movie_Spoken_Language', 'Movie_Tagline', 'Movie_Cast',
      'Movie_Crew',
      'Movie_Director'],
      dtype='object')

```

```

df_features = df[[ 'Movie_Genre', 'Movie_Keywords' ,
'Movie_Tagline' , 'Movie_Cast' , 'Movie_Director' ]].fillna('')

df_features.shape

(4760, 5)

df_features

      Movie_Genre \
0              Crime Comedy
1  Adventure Action Science Fiction
2              Animation Family
3      Comedy Drama Romance
4              Drama
...
4755              Horror
4756      Comedy Family Drama
4757      Thriller Drama
4758              Family
4759      Documentary

      Movie_Keywords \
0      hotel new year's eve witch bet hotel room
1      android galaxy hermit death star lightsaber
2      father son relationship harbor underwater fish...
3      vietnam veteran hippie mentally disabled runni...
4      male nudity female nudity adultery midlife cri...
...
4755
4756
4757      christian film sex trafficking
4758
4759 music actors legendary perfomer classic hollyw...

      Movie_Tagline \
0      Twelve outrageous guests. Four scandalous requ...
1      A long time ago in a galaxy far, far away...
2      There are 3.7 trillion fish in the ocean, they...
3      The world will never be the same, once you've ...
4              Look closer.
...
4755              The hot spot where Satan's waitin'.
4756      It's better to stand out than to fit in.
4757      She never knew it could happen to her...
4758
4759

      Movie_Cast
Movie_Director
0      Tim Roth Antonio Banderas Jennifer Beals Madon...      Allison

```



```

Anders
1      Mark Hamill Harrison Ford Carrie Fisher Peter ...      George
Lucas
2      Albert Brooks Ellen DeGeneres Alexander Gould ...      Andrew
Stanton
3      Tom Hanks Robin Wright Gary Sinise Mykelti Wil...      Robert
Zemeckis
4      Kevin Spacey Annette Bening Thora Birch Wes Be...      Sam
Mendes
...
...
4755   Lisa Hart Carroll Michael Des Barres Paul Drak...      Pece
Dingo
4756   Roni Akurati Brighton Sharbino Jason Lee Anjul...      Frank
Lotito
4757   Nicole Smolen Kim Baldwin Ariana Stephens Brys...      Jaco
Booyens
4758

4759                                     Tony Oppedisano Simon Napier-
Bell

```

```
[4760 rows x 5 columns]
```

```
df_features.loc[4759]
```

```

Movie_Genre                                Documentary
Movie_Keywords    music actors legendary perfomer classic hollyw...
Movie_Tagline
Movie_Cast                                Tony Oppedisano
Movie_Director                                Simon Napier-Bell
Name: 4759, dtype: object

```

```

X = df_features['Movie_Genre'] + ' ' + df_features['Movie_Keywords'] +
' ' + df_features['Movie_Tagline'] + ' ' + df_features['Movie_Cast'] +
' ' + df_features['Movie_Director']

```

```
X
```

```

0      Crime Comedy hotel new year's eve witch bet ho...
1      Adventure Action Science Fiction android galax...
2      Animation Family father son relationship harbo...
3      Comedy Drama Romance vietnam veteran hippie me...
4      Drama male nudity female nudity adultery midli...
...
4755   Horror The hot spot where Satan's waitin'. Li...
4756   Comedy Family Drama It's better to stand out ...
4757   Thriller Drama christian film sex trafficking ...
4758                                     Family
4759   Documentary music actors legendary perfomer cl...
Length: 4760, dtype: object

```

```
X.shape
```

```
(4760,)
```

```
from sklearn.feature_extraction.text import TfidfVectorizer
```

```
tfidf = TfidfVectorizer()
```

```
X=tfidf.fit_transform(X)
```

```
print(X)
```

```
(0, 617) 0.1633382144407513
(0, 492) 0.1432591540388685
(0, 15413) 0.1465525095337543
(0, 9675) 0.14226057295252661
(0, 9465) 0.1659841367820977
(0, 1390) 0.16898383612799558
(0, 7825) 0.09799561597509843
(0, 1214) 0.13865857545144072
(0, 729) 0.13415063359531618
(0, 13093) 0.1432591540388685
(0, 15355) 0.10477815972666779
(0, 9048) 0.0866842116160778
(0, 11161) 0.06250380151644369
(0, 16773) 0.17654247479915475
(0, 5612) 0.08603537588547631
(0, 16735) 0.10690083751525419
(0, 7904) 0.13348000542112332
(0, 15219) 0.09800472886453934
(0, 11242) 0.07277788238484746
(0, 3878) 0.11998399582562203
(0, 5499) 0.11454057510303811
(0, 7071) 0.19822417598406614
(0, 7454) 0.14745635785412262
(0, 1495) 0.19712637387361423
(0, 9206) 0.15186283580984414
:
(4757, 5455) 0.12491480594769522
(4757, 2967) 0.16273475835631626
(4757, 8464) 0.23522565554066333
(4757, 6938) 0.17088173678136628
(4757, 8379) 0.17480603856721913
(4757, 15303) 0.07654356007668191
(4757, 15384) 0.09754322497537371
(4757, 7649) 0.11479421494340192
(4757, 10896) 0.14546473055066447
(4757, 4494) 0.05675298448720501
(4758, 5238) 1.0
(4759, 11264) 0.33947721804318337
(4759, 11708) 0.33947721804318337
```

```

(4759, 205)    0.3237911628497312
(4759, 8902)   0.3040290704566037
(4759, 14062)  0.3237911628497312
(4759, 3058)   0.2812896191863103
(4759, 7130)   0.26419662449963793
(4759, 10761)  0.3126617295732147
(4759, 4358)   0.18306542312175342
(4759, 14051)  0.20084315377640435
(4759, 5690)   0.19534291014627303
(4759, 15431)  0.19628653185946862
(4759, 1490)   0.21197258705292082
(4759, 10666)  0.15888268987343043

```

```
from sklearn.metrics.pairwise import cosine_similarity
```

```
Similarity_Score = cosine_similarity(X)
```

```
Similarity_Score
```

```

array([[1.          , 0.01351235, 0.03570468, ..., 0.          , 0.
,
      0.          ],
      [0.01351235, 1.          , 0.00806674, ..., 0.          , 0.
,
      0.          ],
      [0.03570468, 0.00806674, 1.          , ..., 0.          ,
0.08014876,
      0.          ],
      ...,
      [0.          , 0.          , 0.          , ..., 1.          , 0.
,
      0.          ],
      [0.          , 0.          , 0.08014876, ..., 0.          , 1.
,
      0.          ],
      [0.          , 0.          , 0.          , ..., 0.          , 0.
,
      1.          ]])

```

```
Similarity_Score.shape
```

```
(4760, 4760)
```

```
Favourite_Movie_Name = input( "Enter your favourite Movie Name : " )
```

```
Enter your favourite Movie Name : avtaar
```

```
All_Movie_Title_List = df['Movie_Title'].tolist()
```

```
import difflib
```

```
Movie_Recommendation =
```

```
difflib.get_close_matches(Favourite_Movie_Name,All_Movie_Title_List)
```

```

print(Movie_Recommendation)

['Avatar', 'Gattaca']

Close_Match = Movie_Recommendation[0]

print(Close_Match)

Avatar

Index_of_Close_Match_Movie = df[df.Movie_Title == Close_Match]
['Movie_ID'].values[0]

print(Index_of_Close_Match_Movie)

2692

Recommendation_Score =
list(enumerate(Similarity_Score[Index_of_Close_Match_Movie]))
# print(Recommendation_Score)

len(Recommendation_Score)

4760

Sorted_Similar_Movies = sorted(Recommendation_Score , key = lambda
x:x[1],reverse = True)
# print(Sorted_Similar_Movies)

print("Top 30 Movies suggested for you are : ")

i= 1
for movie in Sorted_Similar_Movies:
    index = movie[0];
    title_from_index = df[df.index == index]['Movie_Title'].values[0]
    if(i<31):
        print(i, '.',title_from_index)
        i+=1

Top 30 Movies suggested for you are :
1 . Niagara
2 . Caravans
3 . My Week with Marilyn
4 . Brokeback Mountain
5 . Harry Brown
6 . Night of the Living Dead
7 . The Curse of Downers Grove
8 . The Boy Next Door
9 . Back to the Future
10 . The Juror
11 . Some Like It Hot
12 . Enough
13 . The Kentucky Fried Movie

```

- 14 . Eye for an Eye
- 15 . Welcome to the Sticks
- 16 . Alice Through the Looking Glass
- 17 . Superman III
- 18 . The Misfits
- 19 . Premium Rush
- 20 . Duel in the Sun
- 21 . Sabotage
- 22 . Small Soldiers
- 23 . All That Jazz
- 24 . Camping Sauvage
- 25 . The Raid
- 26 . Beyond the Black Rainbow
- 27 . To Kill a Mockingbird
- 28 . World Trade Center
- 29 . The Dark Knight Rises
- 30 . Tora! Tora! Tora!