

```
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
```

```
movies=pd.read_csv("tmdb_5000_movies.csv")
credit=pd.read_csv("tmdb_5000_credits.csv")
```

```
movies.head()
```

	budget	genres \
0	237000000	[{"id": 28, "name": "Action"}, {"id": 12, "nam...
1	300000000	[{"id": 12, "name": "Adventure"}, {"id": 14, "...
2	245000000	[{"id": 28, "name": "Action"}, {"id": 12, "nam...
3	250000000	[{"id": 28, "name": "Action"}, {"id": 80, "nam...
4	260000000	[{"id": 28, "name": "Action"}, {"id": 12, "nam...

	homepage	id \
0	http://www.avatarmovie.com/	19995
1	http://disney.go.com/disneypictures/pirates/	285
2	http://www.sonypictures.com/movies/spectre/	206647
3	http://www.thedarkknightriser.com/	49026
4	http://movies.disney.com/john-carter	49529

	keywords	original_language
0	[{"id": 1463, "name": "culture clash"}, {"id":...	en
1	[{"id": 270, "name": "ocean"}, {"id": 726, "na...	en
2	[{"id": 470, "name": "spy"}, {"id": 818, "name...	en
3	[{"id": 849, "name": "dc comics"}, {"id": 853,...	en
4	[{"id": 818, "name": "based on novel"}, {"id":...	en

	original_title \
0	Avatar
1	Pirates of the Caribbean: At World's End
2	Spectre
3	The Dark Knight Rises
4	John Carter

	overview	popularity \
0	In the 22nd century, a paraplegic Marine is di...	150.437577
1	Captain Barbossa, long believed to be dead, ha...	139.082615
2	A cryptic message from Bond's past sends him o...	107.376788
3	Following the death of District Attorney Harve...	112.312950
4	John Carter is a war-weary, former military ca...	43.926995

production_companies \

```

0 [{"name": "Ingenious Film Partners", "id": 289...
1 [{"name": "Walt Disney Pictures", "id": 2}, {""...
2 [{"name": "Columbia Pictures", "id": 5}, {"nam...
3 [{"name": "Legendary Pictures", "id": 923}, {""...
4 [{"name": "Walt Disney Pictures", "id": 2}]

```

```

                                production_countries release_date
revenue \
0 [{"iso_3166_1": "US", "name": "United States o... 2009-12-10
2787965087
1 [{"iso_3166_1": "US", "name": "United States o... 2007-05-19
961000000
2 [{"iso_3166_1": "GB", "name": "United Kingdom"... 2015-10-26
880674609
3 [{"iso_3166_1": "US", "name": "United States o... 2012-07-16
1084939099
4 [{"iso_3166_1": "US", "name": "United States o... 2012-03-07
284139100

```

```

runtime                                spoken_languages
status \
0 162.0 [{"iso_639_1": "en", "name": "English"}, {"iso...
Released
1 169.0 [{"iso_639_1": "en", "name": "English"}]
Released
2 148.0 [{"iso_639_1": "fr", "name": "Fran\u00e7ais"},...
Released
3 165.0 [{"iso_639_1": "en", "name": "English"}]
Released
4 132.0 [{"iso_639_1": "en", "name": "English"}]
Released

```

```

                                tagline \
0 Enter the World of Pandora.
1 At the end of the world, the adventure begins.
2 A Plan No One Escapes
3 The Legend Ends
4 Lost in our world, found in another.

```

```

                                title vote_average vote_count
0 Avatar 7.2 11800
1 Pirates of the Caribbean: At World's End 6.9 4500
2 Spectre 6.3 4466
3 The Dark Knight Rises 7.6 9106

```

4	John Carter	6.1	2124
---	-------------	-----	------

```
credit.head()
```

	movie_id		title	\
0	19995		Avatar	
1	285	Pirates of the Caribbean: At World's End		
2	206647		Spectre	
3	49026	The Dark Knight Rises		
4	49529		John Carter	

		cast	\
0	[{"cast_id": 242, "character": "Jake Sully", "...		
1	[{"cast_id": 4, "character": "Captain Jack Spa...		
2	[{"cast_id": 1, "character": "James Bond", "cr...		
3	[{"cast_id": 2, "character": "Bruce Wayne / Ba...		
4	[{"cast_id": 5, "character": "John Carter", "c...		

		crew
0	[{"credit_id": "52fe48009251416c750aca23", "de...	
1	[{"credit_id": "52fe4232c3a36847f800b579", "de...	
2	[{"credit_id": "54805967c3a36829b5002c41", "de...	
3	[{"credit_id": "52fe4781c3a36847f81398c3", "de...	
4	[{"credit_id": "52fe479ac3a36847f813eaa3", "de...	

```
movies.isnull().sum()
```

budget	0
genres	0
homepage	3091
id	0
keywords	0
original_language	0
original_title	0
overview	3
popularity	0
production_companies	0
production_countries	0
release_date	1
revenue	0
runtime	2
spoken_languages	0
status	0
tagline	844
title	0
vote_average	0
vote_count	0
dtype: int64	

```
credit.isnull().sum()
```

```
movie_id    0
title       0
cast        0
crew        0
dtype: int64
```

#MERGING MOVIE AND CREDIT TOGETHER ON THE BASIS OF TITLE

```
movies=movies.merge(credit,on='title')
```

#REMOVING WASTE COLUMNS THAT WE ARE NOT USING FOR THE RECOMMENDATION SYSTEM AND KEEPING ONLY USED COLUMNS

```
movies=movies[["movie_id","title","overview","genres","keywords","cast","crew"]]
```

```
movies.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4809 entries, 0 to 4808
Data columns (total 7 columns):
#   Column      Non-Null Count  Dtype
---  -
0   movie_id    4809 non-null   int64
1   title       4809 non-null   object
2   overview    4806 non-null   object
3   genres      4809 non-null   object
4   keywords    4809 non-null   object
5   cast        4809 non-null   object
6   crew        4809 non-null   object
dtypes: int64(1), object(6)
memory usage: 263.1+ KB
```

```
movies.isnull().sum()
```

```
movie_id    0
title       0
overview     3
genres       0
keywords     0
cast         0
crew         0
dtype: int64
```

#DROPPING NULL VALUES FROM DATASET

```
movies.dropna(inplace=True)
```

```
movies.duplicated().sum()
```

```
0
```

```
movies.isnull().sum()
```

```
movie_id    0
title       0
overview    0
genres      0
keywords    0
cast        0
crew        0
dtype: int64
```

```
movies.head()
```

```
   movie_id  title \
0    19995  Avatar
1     285  Pirates of the Caribbean: At World's End
2   206647  Spectre
3    49026  The Dark Knight Rises
4    49529  John Carter
```

```
   overview \
0  In the 22nd century, a paraplegic Marine is di...
1  Captain Barbossa, long believed to be dead, ha...
2  A cryptic message from Bond's past sends him o...
3  Following the death of District Attorney Harve...
4  John Carter is a war-weary, former military ca...
```

```
   genres \
0  [{"id": 28, "name": "Action"}, {"id": 12, "nam...
1  [{"id": 12, "name": "Adventure"}, {"id": 14, "...
2  [{"id": 28, "name": "Action"}, {"id": 12, "nam...
3  [{"id": 28, "name": "Action"}, {"id": 80, "nam...
4  [{"id": 28, "name": "Action"}, {"id": 12, "nam...
```

```
   keywords \
0  [{"id": 1463, "name": "culture clash"}, {"id":...
1  [{"id": 270, "name": "ocean"}, {"id": 726, "na...
2  [{"id": 470, "name": "spy"}, {"id": 818, "name...
3  [{"id": 849, "name": "dc comics"}, {"id": 853,...
4  [{"id": 818, "name": "based on novel"}, {"id":...
```

```
   cast \
0  [{"cast_id": 242, "character": "Jake Sully", "...
1  [{"cast_id": 4, "character": "Captain Jack Spa...
2  [{"cast_id": 1, "character": "James Bond", "cr...
3  [{"cast_id": 2, "character": "Bruce Wayne / Ba...
4  [{"cast_id": 5, "character": "John Carter", "c...
```

```
   crew
0  [{"credit_id": "52fe48009251416c750aca23", "de...
1  [{"credit_id": "52fe4232c3a36847f800b579", "de...
2  [{"credit_id": "54805967c3a36829b5002c41", "de...
```

```

3 [{"credit_id": "52fe4781c3a36847f81398c3", "de...
4 [{"credit_id": "52fe479ac3a36847f813eaa3", "de...

#CONVERTING LIST DICTIONARY INTO SIMPLE LIST #GENRES #KEYWORDS AND
#CAST

import ast
def convert(obj):
    L = []
    for i in ast.literal_eval(obj):
        L.append(i["name"])
    return L

movies['genres']=movies['genres'].apply(convert)

movies["genres"]

0          [Action]
1          [Adventure]
2          [Action]
3          [Action]
4          [Action]
...
4804        [Action]
4805        [Comedy]
4806        [Comedy]
4807        None
4808    [Documentary]
Name: genres, Length: 4806, dtype: object

movies.head()

   movie_id      title \
0      19995      Avatar
1       285  Pirates of the Caribbean: At World's End
2     206647      Spectre
3      49026  The Dark Knight Rises
4      49529    John Carter

   overview      genres \
0  In the 22nd century, a paraplegic Marine is di...    [Action]
1  Captain Barbossa, long believed to be dead, ha...  [Adventure]
2  A cryptic message from Bond's past sends him o...    [Action]
3  Following the death of District Attorney Harve...    [Action]
4  John Carter is a war-weary, former military ca...    [Action]

   keywords \
0  [{"id": 1463, "name": "culture clash"}, {"id":...
1  [{"id": 270, "name": "ocean"}, {"id": 726, "na...
2  [{"id": 470, "name": "spy"}, {"id": 818, "name...
3  [{"id": 849, "name": "dc comics"}, {"id": 853,...

```

```
4 [{"id": 818, "name": "based on novel"}, {"id":...
```

```
cast \
0 [{"cast_id": 242, "character": "Jake Sully", "...
1 [{"cast_id": 4, "character": "Captain Jack Spa...
2 [{"cast_id": 1, "character": "James Bond", "cr...
3 [{"cast_id": 2, "character": "Bruce Wayne / Ba...
4 [{"cast_id": 5, "character": "John Carter", "c...
```

```
crew
0 [{"credit_id": "52fe48009251416c750aca23", "de...
1 [{"credit_id": "52fe4232c3a36847f800b579", "de...
2 [{"credit_id": "54805967c3a36829b5002c41", "de...
3 [{"credit_id": "52fe4781c3a36847f81398c3", "de...
4 [{"credit_id": "52fe479ac3a36847f813eaa3", "de...
```

```
movies['keywords']=movies['keywords'].apply(convert)
```

```
movies.head()
```

```
movie_id title \
0 19995 Avatar
1 285 Pirates of the Caribbean: At World's End
2 206647 Spectre
3 49026 The Dark Knight Rises
4 49529 John Carter
```

```
overview genres \
0 In the 22nd century, a paraplegic Marine is di... [Action]
1 Captain Barbossa, long believed to be dead, ha... [Adventure]
2 A cryptic message from Bond's past sends him o... [Action]
3 Following the death of District Attorney Harve... [Action]
4 John Carter is a war-weary, former military ca... [Action]
```

```
keywords cast
\
0 [culture clash] [{"cast_id": 242, "character": "Jake Sully", "...
1 [ocean] [{"cast_id": 4, "character": "Captain Jack Spa...
2 [spy] [{"cast_id": 1, "character": "James Bond", "cr...
3 [dc comics] [{"cast_id": 2, "character": "Bruce Wayne / Ba...
4 [based on novel] [{"cast_id": 5, "character": "John Carter", "c...
```

```
crew
0 [{"credit_id": "52fe48009251416c750aca23", "de...
1 [{"credit_id": "52fe4232c3a36847f800b579", "de...
2 [{"credit_id": "54805967c3a36829b5002c41", "de...
```

```
3 [{"credit_id": "52fe4781c3a36847f81398c3", "de...
4 [{"credit_id": "52fe479ac3a36847f813eaa3", "de...
```

```
movies["keywords"]
```

```
0          [culture clash]
1          [ocean]
2          [spy]
3          [dc comics]
4          [based on novel]
...
4804      [united states-mexico barrier]
4805          None
4806          [date]
4807          None
4808          [obsession]
Name: keywords, Length: 4806, dtype: object
```

```
import ast
def conv(obj):
    L=[]
    count=0
    for i in ast.literal_eval(obj):
        if count!=3:
            L.append(i["name"])
            count=count+1
        else:
            break
    return L
```

```
movies["cast"]=movies["cast"].apply(conv)
```

```
movies["cast"]
```

```
0          [Sam Worthington, Zoe Saldana, Sigourney Weaver]
1          [Johnny Depp, Orlando Bloom, Keira Knightley]
2          [Daniel Craig, Christoph Waltz, Léa Seydoux]
3          [Christian Bale, Michael Caine, Gary Oldman]
4          [Taylor Kitsch, Lynn Collins, Samantha Morton]
...
4804      [Carlos Gallardo, Jaime de Hoyos, Peter Marqua...
4805          [Edward Burns, Kerry Bishé, Marsha Dietlein]
4806          [Eric Mabius, Kristin Booth, Crystal Lowe]
4807          [Daniel Henney, Eliza Coupe, Bill Paxton]
4808      [Drew Barrymore, Brian Herzlinger, Corey Feldman]
Name: cast, Length: 4806, dtype: object
```

```
movies.head()
```

	movie_id	title \
0	19995	Avatar

1	285	Pirates of the Caribbean: At World's End
2	206647	Spectre
3	49026	The Dark Knight Rises
4	49529	John Carter

	overview	genres \
0	In the 22nd century, a paraplegic Marine is di...	[Action]
1	Captain Barbossa, long believed to be dead, ha...	[Adventure]
2	A cryptic message from Bond's past sends him o...	[Action]
3	Following the death of District Attorney Harve...	[Action]
4	John Carter is a war-weary, former military ca...	[Action]

	keywords	cast \
0	[culture clash]	[Sam Worthington, Zoe Saldana, Sigourney Weaver]
1	[ocean]	[Johnny Depp, Orlando Bloom, Keira Knightley]
2	[spy]	[Daniel Craig, Christoph Waltz, Léa Seydoux]
3	[dc comics]	[Christian Bale, Michael Caine, Gary Oldman]
4	[based on novel]	[Taylor Kitsch, Lynn Collins, Samantha Morton]

	crew
0	[{"credit_id": "52fe48009251416c750aca23", "de...
1	[{"credit_id": "52fe4232c3a36847f800b579", "de...
2	[{"credit_id": "54805967c3a36829b5002c41", "de...
3	[{"credit_id": "52fe4781c3a36847f81398c3", "de...
4	[{"credit_id": "52fe479ac3a36847f813eaa3", "de...

movies["crew"][0]

```
[{"credit_id": "52fe48009251416c750aca23", "department": "Editing",
"gender": 0, "id": 1721, "job": "Editor", "name": "Stephen E.
Rivkin"}, {"credit_id": "539c47ecc3a36810e3001f87", "department":
"Art", "gender": 2, "id": 496, "job": "Production Design", "name":
"Rick Carter"}, {"credit_id": "54491c89c3a3680fb4001cf7",
"department": "Sound", "gender": 0, "id": 900, "job": "Sound
Designer", "name": "Christopher Boyes"}, {"credit_id":
"54491cb70e0a267480001bd0", "department": "Sound", "gender": 0, "id":
900, "job": "Supervising Sound Editor", "name": "Christopher Boyes"},
{"credit_id": "539c4a4cc3a36810c9002101", "department": "Production",
"gender": 1, "id": 1262, "job": "Casting", "name": "Mali Finn"},
{"credit_id": "5544ee3b925141499f0008fc", "department": "Sound",
"gender": 2, "id": 1729, "job": "Original Music Composer", "name":
"James Horner"}, {"credit_id": "52fe48009251416c750ac9c3",
"department": "Directing", "gender": 2, "id": 2710, "job": "Director",
"name": "James Cameron"}, {"credit_id": "52fe48009251416c750ac9d9",
```

"department": "Writing", "gender": 2, "id": 2710, "job": "Writer",
"name": "James Cameron"}, {"credit_id": "52fe48009251416c750aca17",
"department": "Editing", "gender": 2, "id": 2710, "job": "Editor",
"name": "James Cameron"}, {"credit_id": "52fe48009251416c750aca29",
"department": "Production", "gender": 2, "id": 2710, "job":
"Producer", "name": "James Cameron"}, {"credit_id":
"52fe48009251416c750aca3f", "department": "Writing", "gender": 2,
"id": 2710, "job": "Screenplay", "name": "James Cameron"},
{"credit_id": "539c4987c3a36810ba0021a4", "department": "Art",
"gender": 2, "id": 7236, "job": "Art Direction", "name": "Andrew
Menzies"}, {"credit_id": "549598c3c3a3686ae9004383", "department":
"Visual Effects", "gender": 0, "id": 6690, "job": "Visual Effects
Producer", "name": "Jill Brooks"}, {"credit_id":
"52fe48009251416c750aca4b", "department": "Production", "gender": 1,
"id": 6347, "job": "Casting", "name": "Margery Simkin"}, {"credit_id":
"570b6f419251417da70032fe", "department": "Art", "gender": 2, "id":
6878, "job": "Supervising Art Director", "name": "Kevin Ishioka"},
{"credit_id": "5495a0fac3a3686ae9004468", "department": "Sound",
"gender": 0, "id": 6883, "job": "Music Editor", "name": "Dick
Bernstein"}, {"credit_id": "54959706c3a3686af3003e81", "department":
"Sound", "gender": 0, "id": 8159, "job": "Sound Effects Editor",
"name": "Shannon Mills"}, {"credit_id": "54491d58c3a3680fb1001ccb",
"department": "Sound", "gender": 0, "id": 8160, "job": "Foley",
"name": "Dennie Thorpe"}, {"credit_id": "54491d6cc3a3680fa5001b2c",
"department": "Sound", "gender": 0, "id": 8163, "job": "Foley",
"name": "Jana Vance"}, {"credit_id": "52fe48009251416c750aca57",
"department": "Costume & Make-Up", "gender": 1, "id": 8527, "job":
"Costume Design", "name": "Deborah Lynn Scott"}, {"credit_id":
"52fe48009251416c750aca2f", "department": "Production", "gender": 2,
"id": 8529, "job": "Producer", "name": "Jon Landau"}, {"credit_id":
"539c4937c3a36810ba002194", "department": "Art", "gender": 0, "id":
9618, "job": "Art Direction", "name": "Sean Haworth"}, {"credit_id":
"539c49b6c3a36810c10020e6", "department": "Art", "gender": 1, "id":
12653, "job": "Set Decoration", "name": "Kim Sinclair"}, {"credit_id":
"570b6f2f9251413a0e00020d", "department": "Art", "gender": 1, "id":
12653, "job": "Supervising Art Director", "name": "Kim Sinclair"},
{"credit_id": "54491a6c0e0a26748c001b19", "department": "Art",
"gender": 2, "id": 14350, "job": "Set Designer", "name": "Richard F.
Mays"}, {"credit_id": "56928cf4c3a3684cff0025c4", "department":
"Production", "gender": 1, "id": 20294, "job": "Executive Producer",
"name": "Laeta Kalogridis"}, {"credit_id": "52fe48009251416c750aca51",
"department": "Costume & Make-Up", "gender": 0, "id": 17675, "job":
"Costume Design", "name": "Mayes C. Rubeo"}, {"credit_id":
"52fe48009251416c750aca11", "department": "Camera", "gender": 2, "id":
18265, "job": "Director of Photography", "name": "Mauro Fiore"},
{"credit_id": "5449194d0e0a26748f001b39", "department": "Art",
"gender": 0, "id": 42281, "job": "Set Designer", "name": "Scott
Herbertson"}, {"credit_id": "52fe48009251416c750aca05", "department":
"Crew", "gender": 0, "id": 42288, "job": "Stunts", "name": "Woody

Schultz"}, {"credit_id": "5592aefb92514152de0010f5", "department": "Costume & Make-Up", "gender": 0, "id": 29067, "job": "Makeup Artist", "name": "Linda DeVetta"}, {"credit_id": "5592afa492514152de00112c", "department": "Costume & Make-Up", "gender": 0, "id": 29067, "job": "Hairstylist", "name": "Linda DeVetta"}, {"credit_id": "54959ed592514130fc002e5d", "department": "Camera", "gender": 2, "id": 33302, "job": "Camera Operator", "name": "Richard Bluck"}, {"credit_id": "539c4891c3a36810ba002147", "department": "Art", "gender": 2, "id": 33303, "job": "Art Direction", "name": "Simon Bright"}, {"credit_id": "54959c069251417a81001f3a", "department": "Visual Effects", "gender": 0, "id": 113145, "job": "Visual Effects Supervisor", "name": "Richard Martin"}, {"credit_id": "54959a0dc3a3680ff5002c8d", "department": "Crew", "gender": 2, "id": 58188, "job": "Visual Effects Editor", "name": "Steve R. Moore"}, {"credit_id": "52fe48009251416c750acald", "department": "Editing", "gender": 2, "id": 58871, "job": "Editor", "name": "John Refoua"}, {"credit_id": "54491a4dc3a3680fc30018ca", "department": "Art", "gender": 0, "id": 92359, "job": "Set Designer", "name": "Karl J. Martin"}, {"credit_id": "52fe48009251416c750aca35", "department": "Camera", "gender": 1, "id": 72201, "job": "Director of Photography", "name": "Chiling Lin"}, {"credit_id": "52fe48009251416c750ac9ff", "department": "Crew", "gender": 0, "id": 89714, "job": "Stunts", "name": "Ilram Choi"}, {"credit_id": "54959c529251416e2b004394", "department": "Visual Effects", "gender": 2, "id": 93214, "job": "Visual Effects Supervisor", "name": "Steven Quale"}, {"credit_id": "54491edf0e0a267489001c37", "department": "Crew", "gender": 1, "id": 122607, "job": "Dialect Coach", "name": "Carla Meyer"}, {"credit_id": "539c485bc3a368653d001a3a", "department": "Art", "gender": 2, "id": 132585, "job": "Art Direction", "name": "Nick Bassett"}, {"credit_id": "539c4903c3a368653d001a74", "department": "Art", "gender": 0, "id": 132596, "job": "Art Direction", "name": "Jill Cormack"}, {"credit_id": "539c4967c3a368653d001a94", "department": "Art", "gender": 0, "id": 132604, "job": "Art Direction", "name": "Andy McLaren"}, {"credit_id": "52fe48009251416c750aca45", "department": "Crew", "gender": 0, "id": 236696, "job": "Motion Capture Artist", "name": "Terry Notary"}, {"credit_id": "54959e02c3a3680fc60027d2", "department": "Crew", "gender": 2, "id": 956198, "job": "Stunt Coordinator", "name": "Garrett Warren"}, {"credit_id": "54959ca3c3a3686ae300438c", "department": "Visual Effects", "gender": 2, "id": 957874, "job": "Visual Effects Supervisor", "name": "Jonathan Rothbart"}, {"credit_id": "570b6f519251412c74001b2f", "department": "Art", "gender": 0, "id": 957889, "job": "Supervising Art Director", "name": "Stefan Dechant"}, {"credit_id": "570b6f62c3a3680b77007460", "department": "Art", "gender": 2, "id": 959555, "job": "Supervising Art Director", "name": "Todd Cherniawsky"}, {"credit_id": "539c4a3ac3a36810da0021cc", "department": "Production", "gender": 0, "id": 1016177, "job": "Casting", "name": "Miranda Rivers"}, {"credit_id": "539c482cc3a36810c1002062", "department": "Art", "gender": 0, "id": 1032536, "job": "Production Design", "name":

"Robert Stromberg"}, {"credit_id": "539c4b65c3a36810c9002125",
"department": "Costume & Make-Up", "gender": 2, "id": 1071680, "job":
"Costume Design", "name": "John Harding"}, {"credit_id":
"54959e6692514130fc002e4e", "department": "Camera", "gender": 0, "id":
1177364, "job": "Steadicam Operator", "name": "Roberto De Angelis"},
{"credit_id": "539c49f1c3a368653d001aac", "department": "Costume &
Make-Up", "gender": 2, "id": 1202850, "job": "Makeup Department Head",
"name": "Mike Smithson"}, {"credit_id": "5495999ec3a3686ae100460c",
"department": "Visual Effects", "gender": 0, "id": 1204668, "job":
"Visual Effects Producer", "name": "Alain Lalanne"}, {"credit_id":
"54959cdfc3a3681153002729", "department": "Visual Effects", "gender":
0, "id": 1206410, "job": "Visual Effects Supervisor", "name": "Lucas
Salton"}, {"credit_id": "549596239251417a81001eae", "department":
"Crew", "gender": 0, "id": 1234266, "job": "Post Production
Supervisor", "name": "Janace Tashjian"}, {"credit_id":
"54959c859251416ele003efe", "department": "Visual Effects", "gender":
0, "id": 1271932, "job": "Visual Effects Supervisor", "name": "Stephen
Rosenbaum"}, {"credit_id": "5592af28c3a368775a00105f", "department":
"Costume & Make-Up", "gender": 0, "id": 1310064, "job": "Makeup
Artist", "name": "Frankie Karena"}, {"credit_id":
"539c4adfc3a36810e300203b", "department": "Costume & Make-Up",
"gender": 1, "id": 1319844, "job": "Costume Supervisor", "name": "Lisa
Lovaas"}, {"credit_id": "54959b579251416e2b004371", "department":
"Visual Effects", "gender": 0, "id": 1327028, "job": "Visual Effects
Supervisor", "name": "Jonathan Fawcner"}, {"credit_id":
"539c48a7c3a36810b5001fa7", "department": "Art", "gender": 0, "id":
1330561, "job": "Art Direction", "name": "Robert Bavin"},
{"credit_id": "539c4a71c3a36810da0021e0", "department": "Costume &
Make-Up", "gender": 0, "id": 1330567, "job": "Costume Supervisor",
"name": "Anthony Almaraz"}, {"credit_id": "539c4a8ac3a36810ba0021e4",
"department": "Costume & Make-Up", "gender": 0, "id": 1330570, "job":
"Costume Supervisor", "name": "Carolyn M. Fenton"}, {"credit_id":
"539c4ab6c3a36810da0021f0", "department": "Costume & Make-Up",
"gender": 0, "id": 1330574, "job": "Costume Supervisor", "name": "Beth
Koenigsberg"}, {"credit_id": "54491ab70e0a267480001ba2", "department":
"Art", "gender": 0, "id": 1336191, "job": "Set Designer", "name": "Sam
Page"}, {"credit_id": "544919d9c3a3680fc30018bd", "department": "Art",
"gender": 0, "id": 1339441, "job": "Set Designer", "name": "Tex
Kadonaga"}, {"credit_id": "54491cf50e0a267483001b0c", "department":
"Editing", "gender": 0, "id": 1352422, "job": "Dialogue Editor",
"name": "Kim Foscato"}, {"credit_id": "544919f40e0a26748c001b09",
"department": "Art", "gender": 0, "id": 1352962, "job": "Set
Designer", "name": "Tammy S. Lee"}, {"credit_id":
"5495a115c3a3680ff5002d71", "department": "Crew", "gender": 0, "id":
1357070, "job": "Transportation Coordinator", "name": "Denny Caira"},
{"credit_id": "5495a12f92514130fc002e94", "department": "Crew",
"gender": 0, "id": 1357071, "job": "Transportation Coordinator",
"name": "James Waitkus"}, {"credit_id": "5495976fc3a36811530026b0",
"department": "Sound", "gender": 0, "id": 1360103, "job": "Supervising

Sound Editor", "name": "Addison Teague"}, {"credit_id": "54491837c3a3680fb1001c5a", "department": "Art", "gender": 2, "id": 1376887, "job": "Set Designer", "name": "C. Scott Baker"}, {"credit_id": "54491878c3a3680fb4001c9d", "department": "Art", "gender": 0, "id": 1376888, "job": "Set Designer", "name": "Luke Caska"}, {"credit_id": "544918dac3a3680fa5001ae0", "department": "Art", "gender": 0, "id": 1376889, "job": "Set Designer", "name": "David Chow"}, {"credit_id": "544919110e0a267486001b68", "department": "Art", "gender": 0, "id": 1376890, "job": "Set Designer", "name": "Jonathan Dyer"}, {"credit_id": "54491967c3a3680faa001b5e", "department": "Art", "gender": 0, "id": 1376891, "job": "Set Designer", "name": "Joseph Hiura"}, {"credit_id": "54491997c3a3680fb1001c8a", "department": "Art", "gender": 0, "id": 1376892, "job": "Art Department Coordinator", "name": "Rebecca Jellie"}, {"credit_id": "544919ba0e0a26748f001b42", "department": "Art", "gender": 0, "id": 1376893, "job": "Set Designer", "name": "Robert Andrew Johnson"}, {"credit_id": "54491b1dc3a3680faa001b8c", "department": "Art", "gender": 0, "id": 1376895, "job": "Assistant Art Director", "name": "Mike Stassi"}, {"credit_id": "54491b79c3a3680fbb001826", "department": "Art", "gender": 0, "id": 1376897, "job": "Construction Coordinator", "name": "John Villarino"}, {"credit_id": "54491baec3a3680fb4001ce6", "department": "Art", "gender": 2, "id": 1376898, "job": "Assistant Art Director", "name": "Jeffrey Wisniewski"}, {"credit_id": "54491d2fc3a3680fb4001d07", "department": "Editing", "gender": 0, "id": 1376899, "job": "Dialogue Editor", "name": "Cheryl Nardi"}, {"credit_id": "54491d86c3a3680fa5001b2f", "department": "Editing", "gender": 0, "id": 1376901, "job": "Dialogue Editor", "name": "Marshall Winn"}, {"credit_id": "54491d9dc3a3680faa001bb0", "department": "Sound", "gender": 0, "id": 1376902, "job": "Supervising Sound Editor", "name": "Gwendolyn Yates Whittle"}, {"credit_id": "54491dc10e0a267486001bce", "department": "Sound", "gender": 0, "id": 1376903, "job": "Sound Re-Recording Mixer", "name": "William Stein"}, {"credit_id": "54491f500e0a26747c001c07", "department": "Crew", "gender": 0, "id": 1376909, "job": "Choreographer", "name": "Lula Washington"}, {"credit_id": "549599239251412c4e002a2e", "department": "Visual Effects", "gender": 0, "id": 1391692, "job": "Visual Effects Producer", "name": "Chris Del Conte"}, {"credit_id": "54959d54c3a36831b8001d9a", "department": "Visual Effects", "gender": 2, "id": 1391695, "job": "Visual Effects Supervisor", "name": "R. Christopher White"}, {"credit_id": "54959bdf9251412c4e002a66", "department": "Visual Effects", "gender": 0, "id": 1394070, "job": "Visual Effects Supervisor", "name": "Dan Lemmon"}, {"credit_id": "5495971d92514132ed002922", "department": "Sound", "gender": 0, "id": 1394129, "job": "Sound Effects Editor", "name": "Tim Nielsen"}, {"credit_id": "5592b25792514152cc0011aa", "department": "Crew", "gender": 0, "id": 1394286, "job": "CG Supervisor", "name": "Michael Mulholland"}, {"credit_id": "54959a329251416e2b004355", "department": "Crew", "gender": 0, "id": 1394750, "job": "Visual Effects Editor",

"name": "Thomas Nittmann"}, {"credit_id": "54959d6dc3a3686ae9004401", "department": "Visual Effects", "gender": 0, "id": 1394755, "job": "Visual Effects Supervisor", "name": "Edson Williams"}, {"credit_id": "5495a08fc3a3686ae300441c", "department": "Editing", "gender": 0, "id": 1394953, "job": "Digital Intermediate", "name": "Christine Carr"}, {"credit_id": "55402d659251413d6d000249", "department": "Visual Effects", "gender": 0, "id": 1395269, "job": "Visual Effects Supervisor", "name": "John Bruno"}, {"credit_id": "54959e7b9251416e1e003f3e", "department": "Camera", "gender": 0, "id": 1398970, "job": "Steadicam Operator", "name": "David Emmerichs"}, {"credit_id": "54959734c3a3686ae10045e0", "department": "Sound", "gender": 0, "id": 1400906, "job": "Sound Effects Editor", "name": "Christopher Scarabosio"}, {"credit_id": "549595dd92514130fc002d79", "department": "Production", "gender": 0, "id": 1401784, "job": "Production Supervisor", "name": "Jennifer Teves"}, {"credit_id": "549596009251413af70028cc", "department": "Production", "gender": 0, "id": 1401785, "job": "Production Manager", "name": "Brigitte Yorke"}, {"credit_id": "549596e892514130fc002d99", "department": "Sound", "gender": 0, "id": 1401786, "job": "Sound Effects Editor", "name": "Ken Fischer"}, {"credit_id": "549598229251412c4e002a1c", "department": "Crew", "gender": 0, "id": 1401787, "job": "Special Effects Coordinator", "name": "Iain Hutton"}, {"credit_id": "549598349251416e2b00432b", "department": "Crew", "gender": 0, "id": 1401788, "job": "Special Effects Coordinator", "name": "Steve Ingram"}, {"credit_id": "54959905c3a3686ae3004324", "department": "Visual Effects", "gender": 0, "id": 1401789, "job": "Visual Effects Producer", "name": "Joyce Cox"}, {"credit_id": "5495994b92514132ed002951", "department": "Visual Effects", "gender": 0, "id": 1401790, "job": "Visual Effects Producer", "name": "Jenny Foster"}, {"credit_id": "549599cbc3a3686ae1004613", "department": "Crew", "gender": 0, "id": 1401791, "job": "Visual Effects Editor", "name": "Christopher Marino"}, {"credit_id": "549599f2c3a3686ae100461e", "department": "Crew", "gender": 0, "id": 1401792, "job": "Visual Effects Editor", "name": "Jim Milton"}, {"credit_id": "54959a51c3a3686af3003eb5", "department": "Visual Effects", "gender": 0, "id": 1401793, "job": "Visual Effects Producer", "name": "Cyndi Ochs"}, {"credit_id": "54959a7cc3a36811530026f4", "department": "Crew", "gender": 0, "id": 1401794, "job": "Visual Effects Editor", "name": "Lucas Putnam"}, {"credit_id": "54959b91c3a3680ff5002cb4", "department": "Visual Effects", "gender": 0, "id": 1401795, "job": "Visual Effects Supervisor", "name": "Anthony 'Max' Ivins"}, {"credit_id": "54959bb69251412c4e002a5f", "department": "Visual Effects", "gender": 0, "id": 1401796, "job": "Visual Effects Supervisor", "name": "John Knoll"}, {"credit_id": "54959cbbc3a3686ae3004391", "department": "Visual Effects", "gender": 2, "id": 1401799, "job": "Visual Effects Supervisor", "name": "Eric Saindon"}, {"credit_id": "54959d06c3a3686ae90043f6", "department": "Visual Effects", "gender": 0, "id": 1401800, "job": "Visual Effects Supervisor", "name": "Wayne

Stables"}, {"credit_id": "54959d259251416ele003f11", "department": "Visual Effects", "gender": 0, "id": 1401801, "job": "Visual Effects Supervisor", "name": "David Stinnett"}, {"credit_id": "54959db49251413af7002975", "department": "Visual Effects", "gender": 0, "id": 1401803, "job": "Visual Effects Supervisor", "name": "Guy Williams"}, {"credit_id": "54959de4c3a3681153002750", "department": "Crew", "gender": 0, "id": 1401804, "job": "Stunt Coordinator", "name": "Stuart Thorp"}, {"credit_id": "54959ef2c3a3680fc60027f2", "department": "Lighting", "gender": 0, "id": 1401805, "job": "Best Boy Electric", "name": "Giles Coburn"}, {"credit_id": "54959f07c3a3680fc60027f9", "department": "Camera", "gender": 2, "id": 1401806, "job": "Still Photographer", "name": "Mark Fellman"}, {"credit_id": "54959f47c3a3681153002774", "department": "Lighting", "gender": 0, "id": 1401807, "job": "Lighting Technician", "name": "Scott Sprague"}, {"credit_id": "54959f8cc3a36831b8001df2", "department": "Visual Effects", "gender": 0, "id": 1401808, "job": "Animation Director", "name": "Jeremy Hollobon"}, {"credit_id": "54959fa0c3a36831b8001dfb", "department": "Visual Effects", "gender": 0, "id": 1401809, "job": "Animation Director", "name": "Orlando Meunier"}, {"credit_id": "54959fb6c3a3686af3003f54", "department": "Visual Effects", "gender": 0, "id": 1401810, "job": "Animation Director", "name": "Taisuke Tanimura"}, {"credit_id": "54959fd2c3a36831b8001e02", "department": "Costume & Make-Up", "gender": 0, "id": 1401812, "job": "Set Costumer", "name": "Lilia Mishel Acevedo"}, {"credit_id": "54959ff9c3a3686ae300440c", "department": "Costume & Make-Up", "gender": 0, "id": 1401814, "job": "Set Costumer", "name": "Alejandro M. Hernandez"}, {"credit_id": "5495a0ddc3a3686ae10046fe", "department": "Editing", "gender": 0, "id": 1401815, "job": "Digital Intermediate", "name": "Marvin Hall"}, {"credit_id": "5495alf7c3a3686ae3004443", "department": "Production", "gender": 0, "id": 1401816, "job": "Publicist", "name": "Judy Alley"}, {"credit_id": "5592b29fc3a36869d100002f", "department": "Crew", "gender": 0, "id": 1418381, "job": "CG Supervisor", "name": "Mike Perry"}, {"credit_id": "5592b23a9251415df8001081", "department": "Crew", "gender": 0, "id": 1426854, "job": "CG Supervisor", "name": "Andrew Morley"}, {"credit_id": "55491e1192514104c40002d8", "department": "Art", "gender": 0, "id": 1438901, "job": "Conceptual Design", "name": "Seth Engstrom"}, {"credit_id": "5525d5809251417276002b06", "department": "Crew", "gender": 0, "id": 1447362, "job": "Visual Effects Art Director", "name": "Eric Oliver"}, {"credit_id": "554427ca925141586500312a", "department": "Visual Effects", "gender": 0, "id": 1447503, "job": "Modeling", "name": "Matsune Suzuki"}, {"credit_id": "551906889251415aab001c88", "department": "Art", "gender": 0, "id": 1447524, "job": "Art Department Manager", "name": "Paul Tobin"}, {"credit_id": "5592af8492514152cc0010de", "department": "Costume & Make-Up", "gender": 0, "id": 1452643, "job": "Hairstylist", "name": "Roxane Griffin"}, {"credit_id": "553d3c109251415852001318", "department": "Lighting", "gender": 0, "id": 1453938, "job": "Lighting Artist",

```

"name": "Arun Ram-Mohan"}, {"credit_id": "5592af4692514152d5001355",
"department": "Costume & Make-Up", "gender": 0, "id": 1457305, "job":
"Makeup Artist", "name": "Georgia Lockhart-Adams"}, {"credit_id":
"5592b2eac3a36877470012a5", "department": "Crew", "gender": 0, "id":
1466035, "job": "CG Supervisor", "name": "Thrain Shadbolt"},
{"credit_id": "5592b032c3a36877450015f1", "department": "Crew",
"gender": 0, "id": 1483220, "job": "CG Supervisor", "name": "Brad
Alexander"}, {"credit_id": "5592b05592514152d80012f6", "department":
"Crew", "gender": 0, "id": 1483221, "job": "CG Supervisor", "name":
"Shadi Almassizadeh"}, {"credit_id": "5592b090c3a36877570010b5",
"department": "Crew", "gender": 0, "id": 1483222, "job": "CG
Supervisor", "name": "Simon Clutterbuck"}, {"credit_id":
"5592b0dbc3a368774b00112c", "department": "Crew", "gender": 0, "id":
1483223, "job": "CG Supervisor", "name": "Graeme Demmocks"},
{"credit_id": "5592b0fe92514152db0010c1", "department": "Crew",
"gender": 0, "id": 1483224, "job": "CG Supervisor", "name": "Adrian
Fernandes"}, {"credit_id": "5592b11f9251415df8001059", "department":
"Crew", "gender": 0, "id": 1483225, "job": "CG Supervisor", "name":
"Mitch Gates"}, {"credit_id": "5592b15dc3a3687745001645",
"department": "Crew", "gender": 0, "id": 1483226, "job": "CG
Supervisor", "name": "Jerry Kung"}, {"credit_id":
"5592b18e925141645a0004ae", "department": "Crew", "gender": 0, "id":
1483227, "job": "CG Supervisor", "name": "Andy Lomas"}, {"credit_id":
"5592b1bfc3a368775d0010e7", "department": "Crew", "gender": 0, "id":
1483228, "job": "CG Supervisor", "name": "Sebastian Marino"},
{"credit_id": "5592b2049251415df8001078", "department": "Crew",
"gender": 0, "id": 1483229, "job": "CG Supervisor", "name": "Matthias
Menz"}, {"credit_id": "5592b27b92514152d800136a", "department":
"Crew", "gender": 0, "id": 1483230, "job": "CG Supervisor", "name":
"Sergei Nevshupov"}, {"credit_id": "5592b2c3c3a36869e800003c",
"department": "Crew", "gender": 0, "id": 1483231, "job": "CG
Supervisor", "name": "Philippe Rebours"}, {"credit_id":
"5592b317c3a36877470012af", "department": "Crew", "gender": 0, "id":
1483232, "job": "CG Supervisor", "name": "Michael Takarangi"},
{"credit_id": "5592b345c3a36877470012bb", "department": "Crew",
"gender": 0, "id": 1483233, "job": "CG Supervisor", "name": "David
Weitzberg"}, {"credit_id": "5592b37cc3a368775100113b", "department":
"Crew", "gender": 0, "id": 1483234, "job": "CG Supervisor", "name":
"Ben White"}, {"credit_id": "573c8e2f9251413f5d000094", "department":
"Crew", "gender": 1, "id": 1621932, "job": "Stunts", "name": "Min
Windle"]}']

```

```

import ast
def dir(obj):
    L=[]
    for i in ast.literal_eval(obj):
        if i["job"]=="Director":
            L.append(i["name"])
            break
    return L

```



```
movies["crew"]=movies["crew"].apply(dir)
```

```
movies.crew[0]
```

```
['James Cameron']
```

```
movies.head()
```

	movie_id	title \
0	19995	Avatar
1	285	Pirates of the Caribbean: At World's End
2	206647	Spectre
3	49026	The Dark Knight Rises
4	49529	John Carter

	overview	genres \
0	In the 22nd century, a paraplegic Marine is di...	[Action]
1	Captain Barbossa, long believed to be dead, ha...	[Adventure]
2	A cryptic message from Bond's past sends him o...	[Action]
3	Following the death of District Attorney Harve...	[Action]
4	John Carter is a war-weary, former military ca...	[Action]

	keywords	cast
0	[culture clash]	[Sam Worthington, Zoe Saldana, Sigourney Weaver]
1	[ocean]	[Johnny Depp, Orlando Bloom, Keira Knightley]
2	[spy]	[Daniel Craig, Christoph Waltz, Léa Seydoux]
3	[dc comics]	[Christian Bale, Michael Caine, Gary Oldman]
4	[based on novel]	[Taylor Kitsch, Lynn Collins, Samantha Morton]

	crew
0	[James Cameron]
1	[Gore Verbinski]
2	[Sam Mendes]
3	[Christopher Nolan]
4	[Andrew Stanton]

#Now Converting Overview Into List

```
movies["overview"]=movies["overview"].apply(lambda x:x.split())
```

```
movies.head()
```

	movie_id	title \
0	19995	Avatar
1	285	Pirates of the Caribbean: At World's End
2	206647	Spectre

```

3      49026      The Dark Knight Rises
4      49529      John Carter

```

```

                                overview      genres \
0 [In, the, 22nd, century,, a, paraplegic, Marin... [Action]
1 [Captain, Barbossa,, long, believed, to, be, d... [Adventure]
2 [A, cryptic, message, from, Bond's, past, send... [Action]
3 [Following, the, death, of, District, Attorney... [Action]
4 [John, Carter, is, a, war-weary,, former, mili... [Action]

```

```

                                keywords      cast
\
0 [culture clash] [Sam Worthington, Zoe Saldana, Sigourney Weaver]
1 [ocean] [Johnny Depp, Orlando Bloom, Keira Knightley]
2 [spy] [Daniel Craig, Christoph Waltz, Léa Seydoux]
3 [dc comics] [Christian Bale, Michael Caine, Gary Oldman]
4 [based on novel] [Taylor Kitsch, Lynn Collins, Samantha Morton]

```

```

                                crew
0 [James Cameron]
1 [Gore Verbinski]
2 [Sam Mendes]
3 [Christopher Nolan]
4 [Andrew Stanton]

```

```
movies.head()
```

```

movie_id      title \
0      19995      Avatar
1      285      Pirates of the Caribbean: At World's End
2      206647      Spectre
3      49026      The Dark Knight Rises
4      49529      John Carter

```

```

                                overview      genres \
0 [In, the, 22nd, century,, a, paraplegic, Marin... [Action]
1 [Captain, Barbossa,, long, believed, to, be, d... [Adventure]
2 [A, cryptic, message, from, Bond's, past, send... [Action]
3 [Following, the, death, of, District, Attorney... [Action]
4 [John, Carter, is, a, war-weary,, former, mili... [Action]

```

```

                                keywords      cast
\
0 [culture clash] [Sam Worthington, Zoe Saldana, Sigourney Weaver]
1 [ocean] [Johnny Depp, Orlando Bloom, Keira Knightley]

```

```

2          [spy]          [Daniel Craig, Christoph Waltz, Léa Seydoux]
3      [dc comics]          [Christian Bale, Michael Caine, Gary Oldman]
4  [based on novel]      [Taylor Kitsch, Lynn Collins, Samantha Morton]

```

```

                crew
0      [James Cameron]
1      [Gore Verbinski]
2          [Sam Mendes]
3  [Christopher Nolan]
4      [Andrew Stanton]

```

#Now Merging overview,genres,keywords, cast And crew In 1 Single Column Named As Tag

```

movies["tags"]=movies["overview"]+movies["genres"]+movies["keywords"]
+movies["cast"]+movies["crew"]

```

```

movies.head()

```

```

   movie_id      title \
0      19995      Avatar
1       285  Pirates of the Caribbean: At World's End
2    206647      Spectre
3     49026  The Dark Knight Rises
4     49529    John Carter

```

```

                overview      genres \
0  [In, the, 22nd, century,, a, paraplegic, Marin...  [Action]
1  [Captain, Barbossa,, long, believed, to, be, d...  [Adventure]
2  [A, cryptic, message, from, Bond's, past, send...  [Action]
3  [Following, the, death, of, District, Attorney...  [Action]
4  [John, Carter, is, a, war-weary,, former, mili...  [Action]

```

```

                keywords      cast
\
0  [culture clash]  [Sam Worthington, Zoe Saldana, Sigourney Weaver]
1          [ocean]    [Johnny Depp, Orlando Bloom, Keira Knightley]
2          [spy]      [Daniel Craig, Christoph Waltz, Léa Seydoux]
3      [dc comics]      [Christian Bale, Michael Caine, Gary Oldman]
4  [based on novel]  [Taylor Kitsch, Lynn Collins, Samantha Morton]

```

```

                crew

```

```
tags
0      [James Cameron] [In, the, 22nd, century,, a, paraplegic,
Marin...
1      [Gore Verbinski] [Captain, Barbossa,, long, believed, to, be,
d...
2      [Sam Mendes] [A, cryptic, message, from, Bond's, past,
send...
3      [Christopher Nolan] [Following, the, death, of, District,
Attorney...
4      [Andrew Stanton] [John, Carter, is, a, war-weary,, former,
mili...
```

```
movies["tags"]
```

```
0      [In, the, 22nd, century,, a, paraplegic, Marin...
1      [Captain, Barbossa,, long, believed, to, be, d...
2      [A, cryptic, message, from, Bond's, past, send...
3      [Following, the, death, of, District, Attorney...
4      [John, Carter, is, a, war-weary,, former, mili...
...
4804    [El, Mariachi, just, wants, to, play, his, gui...
4805                                           NaN
4806    ["Signed,, Sealed,, Delivered", introduces, a,...
4807                                           NaN
4808    [Ever, since, the, second, grade, when, he, fi...
Name: tags, Length: 4806, dtype: object
```

```
mrs=movies[["movie_id","title","tags"]]
```

```
mrs
```

	movie_id	title \
0	19995	Avatar
1	285	Pirates of the Caribbean: At World's End
2	206647	Spectre
3	49026	The Dark Knight Rises
4	49529	John Carter
...
4804	9367	El Mariachi
4805	72766	Newlyweds
4806	231617	Signed, Sealed, Delivered
4807	126186	Shanghai Calling
4808	25975	My Date with Drew

	tags
0	[In, the, 22nd, century,, a, paraplegic, Marin...
1	[Captain, Barbossa,, long, believed, to, be, d...
2	[A, cryptic, message, from, Bond's, past, send...
3	[Following, the, death, of, District, Attorney...
4	[John, Carter, is, a, war-weary,, former, mili...

```

...
4804 [El, Mariachi, just, wants, to, play, his, gui...
4805                                     NaN
4806 ["Signed,, Sealed,, Delivered", introduces, a,...
4807                                     NaN
4808 [Ever, since, the, second, grade, when, he, fi...

```

[4806 rows x 3 columns]

```
mrs["tags"]=mrs["tags"].fillna("No_Movies")
```

C:\Users\ACER\AppData\Local\Temp\ipykernel_14852\559777636.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

```
mrs["tags"]=mrs["tags"].fillna("No_Movies")
```

```
mrs["tags"] = mrs["tags"].apply(lambda x: " ".join(x))
```

C:\Users\ACER\AppData\Local\Temp\ipykernel_14852\1121504557.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

```
mrs["tags"] = mrs["tags"].apply(lambda x: " ".join(x))
```

```
mrs["tags"] = mrs["tags"].apply(lambda x:x.lower())
```

C:\Users\ACER\AppData\Local\Temp\ipykernel_14852\3332086352.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

```
mrs["tags"] = mrs["tags"].apply(lambda x:x.lower())
```

mrs

	movie_id	title \
0	19995	Avatar
1	285	Pirates of the Caribbean: At World's End
2	206647	Spectre

3	49026	The Dark Knight Rises
4	49529	John Carter
...
4804	9367	El Mariachi
4805	72766	Newlyweds
4806	231617	Signed, Sealed, Delivered
4807	126186	Shanghai Calling
4808	25975	My Date with Drew

		tags
0	in the 22nd century, a paraplegic marine is di...	
1	captain barbossa, long believed to be dead, ha...	
2	a cryptic message from bond's past sends him o...	
3	following the death of district attorney harve...	
4	john carter is a war-weary, former military ca...	
...		...
4804	el mariachi just wants to play his guitar and ...	
4805		n o _ m o v i e s
4806	"signed, sealed, delivered" introduces a dedic...	
4807		n o _ m o v i e s
4808	ever since the second grade when he first saw ...	

[4806 rows x 3 columns]

```
mrs["tags"]=mrs["tags"].fillna("No_Movies")
```

C:\Users\ACER\AppData\Local\Temp\ipykernel_14852\559777636.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

```
mrs["tags"]=mrs["tags"].fillna("No_Movies")
```

```
mrs.head()
```

	movie_id	title \
0	19995	Avatar
1	285	Pirates of the Caribbean: At World's End
2	206647	Spectre
3	49026	The Dark Knight Rises
4	49529	John Carter

		tags
0	in the 22nd century, a paraplegic marine is di...	
1	captain barbossa, long believed to be dead, ha...	
2	a cryptic message from bond's past sends him o...	
3	following the death of district attorney harve...	
4	john carter is a war-weary, former military ca...	

```

#Sklearn is a module pack of CountVectorizer Package
from sklearn.feature_extraction.text import CountVectorizer

cv=CountVectorizer(max_features=5000, stop_words="english")
vect=cv.fit_transform(mrs["tags"]).toarray()
cv.get_feature_names_out()

array(['000', '007', '10', ..., 'zooey', 'zucker', 'zwick'],
dtype=object)

#nltk natural language toolkit
#stem used to Give Only 1 Meaning Word And Removes Remaining Synonyms
import nltk
from nltk.stem.porter import PorterStemmer
ps=PorterStemmer()
ps.stem("Loving")

'love'

mrs

```

	movie_id	title \
0	19995	Avatar
1	285	Pirates of the Caribbean: At World's End
2	206647	Spectre
3	49026	The Dark Knight Rises
4	49529	John Carter
...
4804	9367	El Mariachi
4805	72766	Newlyweds
4806	231617	Signed, Sealed, Delivered
4807	126186	Shanghai Calling
4808	25975	My Date with Drew

```

tags
0    in the 22nd century, a paraplegic marine is di...
1    captain barbossa, long believed to be dead, ha...
2    a cryptic message from bond's past sends him o...
3    following the death of district attorney harve...
4    john carter is a war-weary, former military ca...
...
4804  el mariachi just wants to play his guitar and ...
4805                                     n o _ m o v i e s
4806  "signed, sealed, delivered" introduces a dedic...
4807                                     n o _ m o v i e s
4808  ever since the second grade when he first saw ...

[4806 rows x 3 columns]

def stem(text):
    y=[]

```

```

    for i in text.split():
        y.append(ps.stem(i))
    return " ".join(y)

mrs["tags"]=mrs["tags"].apply(stem)

mrs

#Finding The Angles Between 2 movies for that-:

from sklearn.metrics.pairwise import cosine_similarity
similar=cosine_similarity(vect)

similar

sorted([similar],reverse=True)

list(enumerate(similar[0]))

def recommend(movie):
    movie_index=mrs[mrs["title"]==movie].index[0]
    distance=similar[movie_index]

movie_list=sorted(list(enumerate(distance)),reverse=True,key=lambda
x:x[1])[1:6]

    for i in movie_list:
        print(mrs.iloc[i[0]].title)

sorted(list(enumerate(similar[0])),reverse=True,key=lambda x:x[1])
[1:6]

recommend("Iron Man")

import pickle
pickle.dump(mrs.to_dict(),open("movie_dic.pkl","wb"))

pickle.dump(similar,open("similar.pkl","wb"))

```