

```
In [2]: import pandas as pd
import numpy as np
```

```
In [3]: df = pd.read_csv('Desktop/Projects/Movie Recommender /archive (1)/tmdb_5000_credits.csv')
df
```

Out[3]:

	movie_id		title	cast	crew
0	19995		Avatar	[{"cast_id": 242, "character": "Jake Sully", "de...	[{"credit_id": "52fe48009251416c750aca23", "de...
1	285	Pirates of the Caribbean: At World's End		[{"cast_id": 4, "character": "Captain Jack Spa...	[{"credit_id": "52fe4232c3a36847f800b579", "de...
2	206647		Spectre	[{"cast_id": 1, "character": "James Bond", "cr...	[{"credit_id": "54805967c3a36829b5002c41", "de...
3	49026	The Dark Knight Rises		[{"cast_id": 2, "character": "Bruce Wayne / Ba...	[{"credit_id": "52fe4781c3a36847f81398c3", "de...
4	49529		John Carter	[{"cast_id": 5, "character": "John Carter", "c...	[{"credit_id": "52fe479ac3a36847f813eaa3", "de...
...	...		...	...	...
4798	9367		El Mariachi	[{"cast_id": 1, "character": "El Mariachi", "c...	[{"credit_id": "52fe44eec3a36847f80b280b", "de...
4799	72766		Newlyweds	[{"cast_id": 1, "character": "Buzzy", "credit...	[{"credit_id": "52fe487dc3a368484e0fb013", "de...
4800	231617	Signed, Sealed, Delivered		[{"cast_id": 8, "character": "Oliver O\u002019To...	[{"credit_id": "52fe4df3c3a36847f8275ecf", "de...
4801	126186		Shanghai Calling	[{"cast_id": 3, "character": "Sam", "credit_id...	[{"credit_id": "52fe4ad9c3a368484e16a36b", "de...
4802	25975		My Date with Drew	[{"cast_id": 3, "character": "Herself", "credi...	[{"credit_id": "58ce021b9251415a390165d9", "de...

4803 rows x 4 columns

```
In [4]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4803 entries, 0 to 4802
Data columns (total 4 columns):
#   Column      Non-Null Count  Dtype
---  -
0   movie_id    4803 non-null   int64
1   title       4803 non-null   object
2   cast        4803 non-null   object
3   crew        4803 non-null   object
dtypes: int64(1), object(3)
memory usage: 150.2+ KB
```

```
In [5]: df2 = pd.read_csv('Desktop/Projects/Movie Recommender /archive (1)/tmdb_5000_movies.csv')
df2
```

Out[5]:

	budget	genres	homepage	id	keywords	original_language	origi
0	237000000	[{"id": 28, "name": "Action"}, {"id": 12, "nam...	http://www.avatarmovie.com/	19995	[{"id": 1463, "name": "culture clash"}, {"id": ....	en	
1	300000000	[{"id": 12, "name": "Adventure"}, {"id": 14, "...	http://disney.go.com/disneypictures/pirates/	285	[{"id": 270, "name": "ocean"}, {"id": 726, "na...	en	Pirat Cæ A
2	245000000	[{"id": 28, "name": "Action"}, {"id": 12, "nam...	http://www.sonypictures.com/movies/spectre/	206647	[{"id": 470, "name": "spy"}, {"id": 818, "name...	en	
3	250000000	[{"id": 28, "name": "Action"}, {"id": 80, "nam...	http://www.thedarkknighttrises.com/	49026	[{"id": 849, "name": "dc comics"}, {"id": 853, ...	en	1 Knig
4	260000000	[{"id": 28, "name": "Action"}, {"id": 12, "nam...	http://movies.disney.com/john-carter	49529	[{"id": 818, "name": "based on novel"}, {"id": ....	en	Jof
...	...	...	...	...	...	...	...
4798	220000	[{"id": 28, "name": "Action"}, {"id": 80, "nam...	NaN	9367	[{"id": 5616, "name": "united states\u2013mexi...	es	El
4799	9000	[{"id": 35, "name": "Comedy"}, {"id": 10749, "...	NaN	72766	[]	en	Ne
4800	0	[{"id": 35, "name": "Comedy"}, {"id": 18, "nam...	http://www.hallmarkchannel.com/signedsealeddel...	231617	[{"id": 248, "name": "date"}, {"id": 699, "nam...	en	C
4801	0	[]	http://shanghaicalling.com/	126186	[]	en	S
4802	0	[{"id": 99, "name": "Documentary"}]	NaN	25975	[{"id": 1523, "name": "obsession"}, {"id": 224...	en	My I

4803 rows × 20 columns

In [6]: df2.info()

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4803 entries, 0 to 4802
Data columns (total 20 columns):
#   Column                Non-Null Count  Dtype
---  -
0   budget                4803 non-null   int64
1   genres                4803 non-null   object
2   homepage              1712 non-null   object
3   id                    4803 non-null   int64
4   keywords              4803 non-null   object
5   original_language     4803 non-null   object
6   original_title        4803 non-null   object
7   overview              4800 non-null   object
8   popularity            4803 non-null   float64
9   production_companies  4803 non-null   object
10  production_countries  4803 non-null   object
11  release_date          4802 non-null   object
12  revenue               4803 non-null   int64
13  runtime               4801 non-null   float64
14  spoken_languages      4803 non-null   object
15  status                4803 non-null   object
16  tagline               3959 non-null   object
17  title                 4803 non-null   object
18  vote_average          4803 non-null   float64
19  vote_count            4803 non-null   int64
dtypes: float64(3), int64(4), object(13)
memory usage: 750.6+ KB

```

```

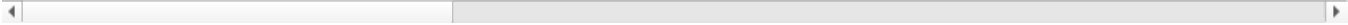
In [7]: movies = df.merge(df2, on = 'title')
        movies

```

Out[7]:

	movie_id	title	cast	crew	budget	genres	
0	19995	Avatar	[{"cast_id": 242, "character": "Jake Sully", "credit_id": "52fe48009251416c750aca23", "de...}]	[{"credit_id": "52fe48009251416c750aca23", "de...}]	237000000	[{"id": 28, "name": "Action"}, {"id": 12, "name": "Adventure"}]	http://www.av
1	285	Pirates of the Caribbean: At World's End	[{"cast_id": 4, "character": "Captain Jack Sparrow", "credit_id": "52fe4232c3a36847f800b579", "de...}]	[{"credit_id": "52fe4232c3a36847f800b579", "de...}]	300000000	[{"id": 12, "name": "Adventure"}, {"id": 14, "name": "Action"}]	http://disney.go.com/disney
2	206647	Spectre	[{"cast_id": 1, "character": "James Bond", "credit_id": "54805967c3a36829b5002c41", "de...}]	[{"credit_id": "54805967c3a36829b5002c41", "de...}]	245000000	[{"id": 28, "name": "Action"}, {"id": 12, "name": "Adventure"}]	http://www.sonypictures.com/
3	49026	The Dark Knight Rises	[{"cast_id": 2, "character": "Bruce Wayne / Batman", "credit_id": "52fe4781c3a36847f81398c3", "de...}]	[{"credit_id": "52fe4781c3a36847f81398c3", "de...}]	250000000	[{"id": 28, "name": "Action"}, {"id": 80, "name": "Drama"}]	http://www.thedarkk
4	49529	John Carter	[{"cast_id": 5, "character": "John Carter", "credit_id": "52fe479ac3a36847f813eaa3", "de...}]	[{"credit_id": "52fe479ac3a36847f813eaa3", "de...}]	260000000	[{"id": 28, "name": "Action"}, {"id": 12, "name": "Adventure"}]	http://movies.disney.
...	...	...	...	...	...	...	...
4804	9367	El Mariachi	[{"cast_id": 1, "character": "El Mariachi", "credit_id": "52fe44eec3a36847f80b280b", "de...}]	[{"credit_id": "52fe44eec3a36847f80b280b", "de...}]	220000	[{"id": 28, "name": "Action"}, {"id": 80, "name": "Drama"}]	
4805	72766	Newlyweds	[{"cast_id": 1, "character": "Buzzy", "credit_id": "52fe487dc3a368484e0fb013", "de...}]	[{"credit_id": "52fe487dc3a368484e0fb013", "de...}]	9000	[{"id": 35, "name": "Comedy"}, {"id": 10749, "name": "Drama"}]	
4806	231617	Signed, Sealed, Delivered O\u0026019To...	[{"cast_id": 8, "character": "Oliver", "credit_id": "52fe4df3c3a36847f8275ecf", "de...}]	[{"credit_id": "52fe4df3c3a36847f8275ecf", "de...}]	0	[{"id": 35, "name": "Comedy"}, {"id": 18, "name": "Drama"}]	http://www.hallmarkchannel.com/sig
4807	126186	Shanghai Calling	[{"cast_id": 3, "character": "Sam", "credit_id": "52fe4ad9c3a368484e16a36b", "de...}]	[{"credit_id": "52fe4ad9c3a368484e16a36b", "de...}]	0	[]	http://shang
4808	25975	My Date with Drew	[{"cast_id": 3, "character": "Herself", "credit_id": "58ce021b9251415a390165d9", "de...}]	[{"credit_id": "58ce021b9251415a390165d9", "de...}]	0	[{"id": 99, "name": "Documentary"}]	

4809 rows × 23 columns



In [8]: `movies.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4809 entries, 0 to 4808
Data columns (total 23 columns):
#   Column                Non-Null Count  Dtype
---  -
0   movie_id              4809 non-null   int64
1   title                 4809 non-null   object
2   cast                  4809 non-null   object
3   crew                  4809 non-null   object
4   budget                4809 non-null   int64
5   genres                4809 non-null   object
6   homepage              1713 non-null   object
7   id                    4809 non-null   int64
8   keywords              4809 non-null   object
9   original_language     4809 non-null   object
10  original_title         4809 non-null   object
11  overview               4806 non-null   object
12  popularity             4809 non-null   float64
13  production_companies   4809 non-null   object
14  production_countries   4809 non-null   object
15  release_date           4808 non-null   object
16  revenue                4809 non-null   int64
17  runtime                4807 non-null   float64
18  spoken_languages       4809 non-null   object
19  status                 4809 non-null   object
20  tagline                3965 non-null   object
21  vote_average           4809 non-null   float64
22  vote_count             4809 non-null   int64
dtypes: float64(3), int64(5), object(15)
memory usage: 864.2+ KB
```

```
In [9]: movies = movies[['movie_id','title','overview','genres','keywords','cast','crew']]
```

```
In [10]: movies
```

Out[10]:

	movie_id	title	overview	genres	keywords	cast	crew
0	19995	Avatar	In the 22nd century, a paraplegic Marine is di...	[{"id": 28, "name": "Action"}, {"id": 12, "nam...	[{"id": 1463, "name": "culture clash"}, {"id": ...	[{"cast_id": 242, "character": "Jake Sully", "...	[{"credit_id": "52fe48009251416c750aca23", "de...
1	285	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...	[{"id": 12, "name": "Adventure"}, {"id": 14, "...	[{"id": 270, "name": "ocean"}, {"id": 726, "na...	[{"cast_id": 4, "character": "Captain Jack Spa...	[{"credit_id": "52fe4232c3a36847f800b579", "de...
2	206647	Spectre	A cryptic message from Bond's past sends him o...	[{"id": 28, "name": "Action"}, {"id": 12, "nam...	[{"id": 470, "name": "spy"}, {"id": 818, "name...	[{"cast_id": 1, "character": "James Bond", "cr...	[{"credit_id": "54805967c3a36829b5002c41", "de...
3	49026	The Dark Knight Rises	Following the death of District Attorney Harve...	[{"id": 28, "name": "Action"}, {"id": 80, "nam...	[{"id": 849, "name": "dc comics"}, {"id": 853, ...	[{"cast_id": 2, "character": "Bruce Wayne / Ba...	[{"credit_id": "52fe4781c3a36847f81398c3", "de...
4	49529	John Carter	John Carter is a war-weary, former military ca...	[{"id": 28, "name": "Action"}, {"id": 12, "nam...	[{"id": 818, "name": "based on novel"}, {"id": ...	[{"cast_id": 5, "character": "John Carter", "c...	[{"credit_id": "52fe479ac3a36847f81398c3", "de...
...	...	...	...	...	...	...	...
4804	9367	El Mariachi	El Mariachi just wants to play his guitar and ...	[{"id": 28, "name": "Action"}, {"id": 80, "nam...	[{"id": 5616, "name": "united states\u2013mexi...	[{"cast_id": 1, "character": "El Mariachi", "c...	[{"credit_id": "52fe44eec3a36847f80b280b", "de...
4805	72766	Newlyweds	A newlywed couple's honeymoon is upended by th...	[{"id": 35, "name": "Comedy"}, {"id": 10749, "...		[{"cast_id": 1, "character": "Buzzy", "credit_...	[{"credit_id": "52fe487dc3a368484e0fb013", "de...
4806	231617	Signed, Sealed, Delivered	"Signed, Sealed, Delivered" introduces a dedic...	[{"id": 35, "name": "Comedy"}, {"id": 18, "nam...	[{"id": 248, "name": "date"}, {"id": 699, "nam...	[{"cast_id": 8, "character": "Oliver Olu2019To...	[{"credit_id": "52fe4df3c3a36847f8275ecf", "de...
4807	126186	Shanghai Calling	When ambitious New York attorney Sam is sent t...			[{"cast_id": 3, "character": "Sam", "credit_id...	[{"credit_id": "52fe4ad9c3a368484e16a36b", "de...
4808	25975	My Date with Drew	Ever since the second grade when he first saw ...	[{"id": 99, "name": "Documentary"}]	[{"id": 1523, "name": "obsession"}, {"id": 224...	[{"cast_id": 3, "character": "Herself", "credi...	[{"credit_id": "58ce021b9251415a390165d9", "de...

4809 rows × 7 columns

In [11]: # Data Preprocessing

In [12]: backup = movies.copy()

In [13]: 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
```

In [14]: movies.dropna(inplace=True)

/var/folders/qg/km03k0gn1yn690bk\_brhvb3c0000gn/T/ipykernel\_62108/3786870272.py:1: SettingWithCopyWarning: A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: [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)

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

In [15]: movies.isnull().sum()

```
Out[15]: movie_id    0
         title      0
         overview   0
         genres     0
         keywords   0
         cast       0
         crew       0
         dtype: int64
```

```
In [16]: movies.duplicated().sum()
```

```
Out[16]: 0
```

```
In [17]: # Genres
         # Sorting everything in the form ['Action','Adventure','Thriller','SciFi']
```

```
In [18]: movies.iloc[0].genres
```

```
Out[18]: '[{"id": 28, "name": "Action"}, {"id": 12, "name": "Adventure"}, {"id": 14, "name": "Fantasy"}, {"id": 878, "name": "Science Fiction"}]'
```

```
In [19]: import ast
         def convert(obj):
             L = []
             for i in ast.literal_eval(obj):
                 L.append(i['name'])
             return L
```

```
In [20]: movies['genres'] = movies['genres'].apply(convert)
```

```
/var/folders/qg/km03k0gnlyn690bk_brhvb3c0000gn/T/ipykernel_62108/1189516248.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
movies['genres'] = movies['genres'].apply(convert)
```

```
In [21]: movies.head(10)
```

Out[21]:	movie_id	title	overview	genres	keywords	cast	crew
0	19995	Avatar	In the 22nd century, a paraplegic Marine is di...	[Action, Adventure, Fantasy, Science Fiction]	{{"id": 1463, "name": "culture clash"}, {"id": ...	{{"cast_id": 242, "character": "Jake Sully", "...	{{"credit_id": "52fe48009251416c750aca23", "de...
1	285	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...	[Adventure, Fantasy, Action]	{{"id": 270, "name": "ocean"}, {"id": 726, "na...	{{"cast_id": 4, "character": "Captain Jack Spa...	{{"credit_id": "52fe4232c3a36847f800b579", "de...
2	206647	Spectre	A cryptic message from Bond's past sends him o...	[Action, Adventure, Crime]	{{"id": 470, "name": "spy"}, {"id": 818, "name...	{{"cast_id": 1, "character": "James Bond", "cr...	{{"credit_id": "54805967c3a36829b5002c41", "de...
3	49026	The Dark Knight Rises	Following the death of District Attorney Harve...	[Action, Crime, Drama, Thriller]	{{"id": 849, "name": "dc comics"}, {"id": 853, ...	{{"cast_id": 2, "character": "Bruce Wayne / Ba...	{{"credit_id": "52fe4781c3a36847f81398c3", "de...
4	49529	John Carter	John Carter is a war-weary, former military ca...	[Action, Adventure, Science Fiction]	{{"id": 818, "name": "based on novel"}, {"id": ...	{{"cast_id": 5, "character": "John Carter", "c...	{{"credit_id": "52fe479ac3a36847f813eaa3", "de...
5	559	Spider-Man 3	The seemingly invincible Spider-Man goes up ag...	[Fantasy, Action, Adventure]	{{"id": 851, "name": "dual identity"}, {"id": ...	{{"cast_id": 30, "character": "Peter Parker / ...	{{"credit_id": "52fe4252c3a36847f80151a5", "de...
6	38757	Tangled	When the kingdom's most wanted-and most charmi...	[Animation, Family]	{{"id": 1562, "name": "hostage"}, {"id": 2343, ...	{{"cast_id": 34, "character": "Flynn Rider (vo...	{{"credit_id": "52fe46db9251416c91062101", "de...
7	99861	Avengers: Age of Ultron	When Tony Stark tries to jumpstart a dormant p...	[Action, Adventure, Science Fiction]	{{"id": 8828, "name": "marvel comic"}, {"id": ...	{{"cast_id": 76, "character": "Tony Stark / Ir...	{{"credit_id": "55d5f7d4c3a3683e7e0016eb", "de...
8	767	Harry Potter and the Half-Blood Prince	As Harry begins his sixth year at Hogwarts, he...	[Adventure, Fantasy, Family]	{{"id": 616, "name": "witch"}, {"id": 2343, "n...	{{"cast_id": 3, "character": "Harry Potter", "...	{{"credit_id": "52fe4273c3a36847f801fab1", "de...
9	209112	Batman v Superman: Dawn of Justice	Fearing the actions of a god-like Super Hero l...	[Action, Adventure, Fantasy]	{{"id": 849, "name": "dc comics"}, {"id": 7002...	{{"cast_id": 18, "character": "Bruce Wayne / B...	{{"credit_id": "553bf23692514135c8002886", "de...

```
In [22]: # Keywords
# Doing the same thing on keywords column
```

```
In [23]: movies.iloc[0].keywords
```

```
Out[23]: '[{"id": 1463, "name": "culture clash"}, {"id": 2964, "name": "future"}, {"id": 3386, "name": "space war"}, {"id": 3388, "name": "space colony"}, {"id": 3679, "name": "society"}, {"id": 3801, "name": "space travel"}, {"id": 9685, "name": "futuristic"}, {"id": 9840, "name": "romance"}, {"id": 9882, "name": "space"}, {"id": 9951, "name": "alien"}, {"id": 10148, "name": "tribe"}, {"id": 10158, "name": "alien planet"}, {"id": 10987, "name": "cgi"}, {"id": 11399, "name": "marine"}, {"id": 13065, "name": "soldier"}, {"id": 14643, "name": "battle"}, {"id": 14720, "name": "love affair"}, {"id": 165431, "name": "anti war"}, {"id": 193554, "name": "power relations"}, {"id": 206690, "name": "mind and soul"}, {"id": 209714, "name": "3d"}]'
```

```
In [24]: movies['keywords'] = movies['keywords'].apply(convert)

/var/folders/qg/km03k0gn1yn690bk_brhvb3c0000gn/T/ipykernel_62108/3328512696.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
movies['keywords'] = movies['keywords'].apply(convert)
```

```
In [25]: movies.head(10)
```



Out[25]:

	movie_id	title	overview	genres	keywords	cast	crew
0	19995	Avatar	In the 22nd century, a paraplegic Marine is di...	[Action, Adventure, Fantasy, Science Fiction]	[culture clash, future, space war, space colon...	{{"cast_id": 242, "character": "Jake Sully", "...	{{"credit_id": "52fe48009251416c750aca23", "de...
1	285	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...	[Adventure, Fantasy, Action]	[ocean, drug abuse, exotic island, east india ...	{{"cast_id": 4, "character": "Captain Jack Spa...	{{"credit_id": "52fe4232c3a36847f800b579", "de...
2	206647	Spectre	A cryptic message from Bond's past sends him o...	[Action, Adventure, Crime]	[spy, based on novel, secret agent, sequel, mi...	{{"cast_id": 1, "character": "James Bond", "cr...	{{"credit_id": "54805967c3a36829b5002c41", "de...
3	49026	The Dark Knight Rises	Following the death of District Attorney Harve...	[Action, Crime, Drama, Thriller]	[dc comics, crime fighter, terrorist, secret i...	{{"cast_id": 2, "character": "Bruce Wayne / Ba...	{{"credit_id": "52fe4781c3a36847f81398c3", "de...
4	49529	John Carter	John Carter is a war-weary, former military ca...	[Action, Adventure, Science Fiction]	[based on novel, mars, medallion, space travel...	{{"cast_id": 5, "character": "John Carter", "c...	{{"credit_id": "52fe479ac3a36847f813eaa3", "de...
5	559	Spider-Man 3	The seemingly invincible Spider-Man goes up ag...	[Fantasy, Action, Adventure]	[dual identity, amnesia, sandstorm, love of on...	{{"cast_id": 30, "character": "Peter Parker / ...	{{"credit_id": "52fe4252c3a36847f80151a5", "de...
6	38757	Tangled	When the kingdom's most wanted-and most charmi...	[Animation, Family]	[hostage, magic, horse, fairy tale, musical, p...	{{"cast_id": 34, "character": "Flynn Rider (vo...	{{"credit_id": "52fe46db9251416c91062101", "de...
7	99861	Avengers: Age of Ultron	When Tony Stark tries to jumpstart a dormant p...	[Action, Adventure, Science Fiction]	[marvel comic, sequel, superhero, based on com...	{{"cast_id": 76, "character": "Tony Stark / Ir...	{{"credit_id": "55d5f7d4c3a3683e7e0016eb", "de...
8	767	Harry Potter and the Half-Blood Prince	As Harry begins his sixth year at Hogwarts, he...	[Adventure, Fantasy, Family]	[witch, magic, broom, school of witchcraft, wi...	{{"cast_id": 3, "character": "Harry Potter", "...	{{"credit_id": "52fe4273c3a36847f801fab1", "de...
9	209112	Batman v Superman: Dawn of Justice	Fearing the actions of a god-like Super Hero l...	[Action, Adventure, Fantasy]	[dc comics, vigilante, superhero, based on com...	{{"cast_id": 18, "character": "Bruce Wayne / B...	{{"credit_id": "553bf23692514135c8002886", "de...

In [26]:

```
# cast
# heere i only want the name of first 3 actors as they are the main actors
```

In [27]:

```
movies['cast'][0]
```

Out[27]:

```
'[{"cast_id": 242, "character": "Jake Sully", "credit_id": "5602a8a7c3a3685532001c9a", "gender": 2, "id": 65731, "name": "Sam Worthington", "order": 0}, {"cast_id": 3, "character": "Neytiri", "credit_id": "52fe48009251416c750ac9cb", "gender": 1, "id": 8691, "name": "Zoe Saldana", "order": 1}, {"cast_id": 25, "character": "Dr. Grace Augustine", "credit_id": "52fe48009251416c750aca39", "gender": 1, "id": 10205, "name": "Sigourney Weaver", "order": 2}, {"cast_id": 4, "character": "Col. Quaritch", "credit_id": "52fe48009251416c750ac9cf", "gender": 2, "id": 32747, "name": "Stephen Lang", "order": 3}, {"cast_id": 5, "character": "Trudy Chacon", "credit_id": "52fe48009251416c750ac9d3", "gender": 1, "id": 17647, "name": "Michelle Rodriguez", "order": 4}, {"cast_id": 8, "character": "Selfridge", "credit_id": "52fe48009251416c750ac9e1", "gender": 2, "id": 1771, "name": "Giovanni Ribisi", "order": 5}, {"cast_id": 7, "character": "Norm Spellman", "credit_id": "52fe48009251416c750ac9dd", "gender": 2, "id": 59231, "name": "Joel David Moore", "order": 6}, {"cast_id": 9, "character": "Moat", "credit_id": "52fe48009251416c750ac9e5", "gender": 1, "id": 30485, "name": "CCH Pounder", "order": 7}, {"cast_id": 11, "character": "Eytukan", "credit_id": "52fe48009251416c750ac9ed", "gender": 2, "id": 15853, "name": "Wes Studi", "order": 8}, {"cast_id": 10, "character": "Tsu'Tey", "credit_id": "52fe48009251416c750ac9e9", "gender": 2, "id": 10964, "name": "Laz Alonso", "order": 9}, {"cast_id": 12, "character": "Dr. Max Patel", "credit_id": "52fe48009251416c750ac9f1", "gender": 2, "id": 95697, "name": "Dileep Rao", "order": 10}, {"cast_id": 13, "character": "Lyle Wainfleet", "credit_id": "52fe48009251416c750ac9f5", "gender": 2, "id": 98215, "name": "Matt Gerald", "order": 11}, {"cast_id": 32, "character": "Private Fike", "credit_id": "52fe48009251416c750aca5b", "gender": 2, "id": 154153, "name": "Sean Anthony Moran", "order": 12}, {"cast_id": 33, "character": "Cryo Vault Med Tech", "credit_id": "52fe48009251416c750aca5f", "gender": 2, "id": 397312, "name": "Jason Whyte", "order": 13}, {"cast_id": 34, "character": "Venture Star Crew Chief", "credit_id": "52fe48009251416c750aca63", "gender": 2, "id": 42317, "name": "Scott Lawrence", "order": 14}, {"cast_id": 35, "character": "Lock Up Trooper", "credit_id": "52fe48009251416c750aca67", "gender": 2, "id": 986734, "name": "Kelly Kilgour", "order": 15}, {"cast_id": 36, "character": "Shuttle Pilot", "credit_id": "52fe48009251416c750aca6b", "gender": 0, "id": 1207227, "name": "James Patrick Pitt", "order": 16}, {"cast_id": 37, "character": "Shuttle Co-Pilot", "credit_id": "52fe48009251416c750aca6f", "gender": 0, "id": 1180936, "name": "Sean Patrick Murphy", "order": 17}, {"cast_id": 38, "character": "Shuttle Crew Chief", "credit_id": "52fe48009251416c750aca73", "gender": 2, "id": 1019578, "name": "Peter Dillon", "order": 18}, {"cast_id": 39, "character": "Tractor Operator / Troupe", "credit_id": "52fe48009251416c750aca77", "gender": 0, "id": 91443, "name": "Kevin Dorman", "order": 19}, {"cast_id": 40, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750aca7b", "gender": 2, "id": 173391, "name": "Kelson Henderson", "order": 20}, {"cast_id": 41, "character": "Dragon Gunship Gunner", "credit_id": "52fe48009251416c750aca7f", "gender": 0, "id": 1207236, "name": "David Van Horn", "order": 21}, {"cast_id": 42, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750aca83", "gender": 0, "id": 215913, "name": "Jacob Tomuri", "order": 22}, {"cast_id": 43, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750aca87", "gender": 0, "id": 215914, "name": "Jacob Tomuri", "order": 23}, {"cast_id": 44, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750aca8b", "gender": 0, "id": 215915, "name": "Jacob Tomuri", "order": 24}, {"cast_id": 45, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750aca8f", "gender": 0, "id": 215916, "name": "Jacob Tomuri", "order": 25}, {"cast_id": 46, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750aca93", "gender": 0, "id": 215917, "name": "Jacob Tomuri", "order": 26}, {"cast_id": 47, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750aca97", "gender": 0, "id": 215918, "name": "Jacob Tomuri", "order": 27}, {"cast_id": 48, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750aca9b", "gender": 0, "id": 215919, "name": "Jacob Tomuri", "order": 28}, {"cast_id": 49, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750aca9f", "gender": 0, "id": 215920, "name": "Jacob Tomuri", "order": 29}, {"cast_id": 50, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaab", "gender": 0, "id": 215921, "name": "Jacob Tomuri", "order": 30}, {"cast_id": 51, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215922, "name": "Jacob Tomuri", "order": 31}, {"cast_id": 52, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215923, "name": "Jacob Tomuri", "order": 32}, {"cast_id": 53, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215924, "name": "Jacob Tomuri", "order": 33}, {"cast_id": 54, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215925, "name": "Jacob Tomuri", "order": 34}, {"cast_id": 55, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215926, "name": "Jacob Tomuri", "order": 35}, {"cast_id": 56, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215927, "name": "Jacob Tomuri", "order": 36}, {"cast_id": 57, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215928, "name": "Jacob Tomuri", "order": 37}, {"cast_id": 58, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215929, "name": "Jacob Tomuri", "order": 38}, {"cast_id": 59, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215930, "name": "Jacob Tomuri", "order": 39}, {"cast_id": 60, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215931, "name": "Jacob Tomuri", "order": 40}, {"cast_id": 61, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215932, "name": "Jacob Tomuri", "order": 41}, {"cast_id": 62, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215933, "name": "Jacob Tomuri", "order": 42}, {"cast_id": 63, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215934, "name": "Jacob Tomuri", "order": 43}, {"cast_id": 64, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215935, "name": "Jacob Tomuri", "order": 44}, {"cast_id": 65, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215936, "name": "Jacob Tomuri", "order": 45}, {"cast_id": 66, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215937, "name": "Jacob Tomuri", "order": 46}, {"cast_id": 67, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215938, "name": "Jacob Tomuri", "order": 47}, {"cast_id": 68, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215939, "name": "Jacob Tomuri", "order": 48}, {"cast_id": 69, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215940, "name": "Jacob Tomuri", "order": 49}, {"cast_id": 70, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215941, "name": "Jacob Tomuri", "order": 50}, {"cast_id": 71, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215942, "name": "Jacob Tomuri", "order": 51}, {"cast_id": 72, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215943, "name": "Jacob Tomuri", "order": 52}, {"cast_id": 73, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215944, "name": "Jacob Tomuri", "order": 53}, {"cast_id": 74, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215945, "name": "Jacob Tomuri", "order": 54}, {"cast_id": 75, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215946, "name": "Jacob Tomuri", "order": 55}, {"cast_id": 76, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215947, "name": "Jacob Tomuri", "order": 56}, {"cast_id": 77, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215948, "name": "Jacob Tomuri", "order": 57}, {"cast_id": 78, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215949, "name": "Jacob Tomuri", "order": 58}, {"cast_id": 79, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215950, "name": "Jacob Tomuri", "order": 59}, {"cast_id": 80, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215951, "name": "Jacob Tomuri", "order": 60}, {"cast_id": 81, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215952, "name": "Jacob Tomuri", "order": 61}, {"cast_id": 82, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215953, "name": "Jacob Tomuri", "order": 62}, {"cast_id": 83, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215954, "name": "Jacob Tomuri", "order": 63}, {"cast_id": 84, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215955, "name": "Jacob Tomuri", "order": 64}, {"cast_id": 85, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215956, "name": "Jacob Tomuri", "order": 65}, {"cast_id": 86, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215957, "name": "Jacob Tomuri", "order": 66}, {"cast_id": 87, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215958, "name": "Jacob Tomuri", "order": 67}, {"cast_id": 88, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215959, "name": "Jacob Tomuri", "order": 68}, {"cast_id": 89, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215960, "name": "Jacob Tomuri", "order": 69}, {"cast_id": 90, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215961, "name": "Jacob Tomuri", "order": 70}, {"cast_id": 91, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215962, "name": "Jacob Tomuri", "order": 71}, {"cast_id": 92, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215963, "name": "Jacob Tomuri", "order": 72}, {"cast_id": 93, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215964, "name": "Jacob Tomuri", "order": 73}, {"cast_id": 94, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215965, "name": "Jacob Tomuri", "order": 74}, {"cast_id": 95, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215966, "name": "Jacob Tomuri", "order": 75}, {"cast_id": 96, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215967, "name": "Jacob Tomuri", "order": 76}, {"cast_id": 97, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215968, "name": "Jacob Tomuri", "order": 77}, {"cast_id": 98, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215969, "name": "Jacob Tomuri", "order": 78}, {"cast_id": 99, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215970, "name": "Jacob Tomuri", "order": 79}, {"cast_id": 100, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215971, "name": "Jacob Tomuri", "order": 80}, {"cast_id": 101, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215972, "name": "Jacob Tomuri", "order": 81}, {"cast_id": 102, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215973, "name": "Jacob Tomuri", "order": 82}, {"cast_id": 103, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215974, "name": "Jacob Tomuri", "order": 83}, {"cast_id": 104, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215975, "name": "Jacob Tomuri", "order": 84}, {"cast_id": 105, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215976, "name": "Jacob Tomuri", "order": 85}, {"cast_id": 106, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215977, "name": "Jacob Tomuri", "order": 86}, {"cast_id": 107, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215978, "name": "Jacob Tomuri", "order": 87}, {"cast_id": 108, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215979, "name": "Jacob Tomuri", "order": 88}, {"cast_id": 109, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215980, "name": "Jacob Tomuri", "order": 89}, {"cast_id": 110, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215981, "name": "Jacob Tomuri", "order": 90}, {"cast_id": 111, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215982, "name": "Jacob Tomuri", "order": 91}, {"cast_id": 112, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215983, "name": "Jacob Tomuri", "order": 92}, {"cast_id": 113, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215984, "name": "Jacob Tomuri", "order": 93}, {"cast_id": 114, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215985, "name": "Jacob Tomuri", "order": 94}, {"cast_id": 115, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215986, "name": "Jacob Tomuri", "order": 95}, {"cast_id": 116, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215987, "name": "Jacob Tomuri", "order": 96}, {"cast_id": 117, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215988, "name": "Jacob Tomuri", "order": 97}, {"cast_id": 118, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215989, "name": "Jacob Tomuri", "order": 98}, {"cast_id": 119, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215990, "name": "Jacob Tomuri", "order": 99}, {"cast_id": 120, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215991, "name": "Jacob Tomuri", "order": 100}, {"cast_id": 121, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215992, "name": "Jacob Tomuri", "order": 101}, {"cast_id": 122, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215993, "name": "Jacob Tomuri", "order": 102}, {"cast_id": 123, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215994, "name": "Jacob Tomuri", "order": 103}, {"cast_id": 124, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215995, "name": "Jacob Tomuri", "order": 104}, {"cast_id": 125, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215996, "name": "Jacob Tomuri", "order": 105}, {"cast_id": 126, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215997, "name": "Jacob Tomuri", "order": 106}, {"cast_id": 127, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 215998, "name": "Jacob Tomuri", "order": 107}, {"cast_id": 128, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 215999, "name": "Jacob Tomuri", "order": 108}, {"cast_id": 129, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216000, "name": "Jacob Tomuri", "order": 109}, {"cast_id": 130, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216001, "name": "Jacob Tomuri", "order": 110}, {"cast_id": 131, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216002, "name": "Jacob Tomuri", "order": 111}, {"cast_id": 132, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216003, "name": "Jacob Tomuri", "order": 112}, {"cast_id": 133, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216004, "name": "Jacob Tomuri", "order": 113}, {"cast_id": 134, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216005, "name": "Jacob Tomuri", "order": 114}, {"cast_id": 135, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216006, "name": "Jacob Tomuri", "order": 115}, {"cast_id": 136, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216007, "name": "Jacob Tomuri", "order": 116}, {"cast_id": 137, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216008, "name": "Jacob Tomuri", "order": 117}, {"cast_id": 138, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216009, "name": "Jacob Tomuri", "order": 118}, {"cast_id": 139, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216010, "name": "Jacob Tomuri", "order": 119}, {"cast_id": 140, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216011, "name": "Jacob Tomuri", "order": 120}, {"cast_id": 141, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216012, "name": "Jacob Tomuri", "order": 121}, {"cast_id": 142, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216013, "name": "Jacob Tomuri", "order": 122}, {"cast_id": 143, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216014, "name": "Jacob Tomuri", "order": 123}, {"cast_id": 144, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216015, "name": "Jacob Tomuri", "order": 124}, {"cast_id": 145, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216016, "name": "Jacob Tomuri", "order": 125}, {"cast_id": 146, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216017, "name": "Jacob Tomuri", "order": 126}, {"cast_id": 147, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216018, "name": "Jacob Tomuri", "order": 127}, {"cast_id": 148, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216019, "name": "Jacob Tomuri", "order": 128}, {"cast_id": 149, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216020, "name": "Jacob Tomuri", "order": 129}, {"cast_id": 150, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216021, "name": "Jacob Tomuri", "order": 130}, {"cast_id": 151, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216022, "name": "Jacob Tomuri", "order": 131}, {"cast_id": 152, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216023, "name": "Jacob Tomuri", "order": 132}, {"cast_id": 153, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216024, "name": "Jacob Tomuri", "order": 133}, {"cast_id": 154, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216025, "name": "Jacob Tomuri", "order": 134}, {"cast_id": 155, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216026, "name": "Jacob Tomuri", "order": 135}, {"cast_id": 156, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216027, "name": "Jacob Tomuri", "order": 136}, {"cast_id": 157, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216028, "name": "Jacob Tomuri", "order": 137}, {"cast_id": 158, "character": "Dragon Gunship Navigator", "credit_id": "52fe48009251416c750acaa3", "gender": 0, "id": 216029, "name": "Jacob Tomuri", "order": 138}, {"cast_id": 159, "character": "Dragon Gunship Pilot", "credit_id": "52fe48009251416c750acaa7", "gender": 0, "id": 216030, "name": "Jacob Tomuri", "order": 139}, {"cast_id": 160, "character": "Dragon Gunship Navigator", "credit_id": "52fe48
```

: 43, "character": "Suit #1", "credit\_id": "52fe48009251416c750aca87", "gender": 0, "id": 143206, "name": "Michael Blain-Rozgay", "order": 23}, {"cast\_id": 44, "character": "Suit #2", "credit\_id": "52fe48009251416c750aca8b", "gender": 2, "id": 169676, "name": "Jon Curry", "order": 24}, {"cast\_id": 46, "character": "Ambient Room Tech", "credit\_id": "52fe48009251416c750aca8f", "gender": 0, "id": 1048610, "name": "Luke Hawker", "order": 25}, {"cast\_id": 47, "character": "Ambient Room Tech / Troupe", "credit\_id": "52fe48009251416c750aca93", "gender": 0, "id": 42288, "name": "Woody Schultz", "order": 26}, {"cast\_id": 48, "character": "Horse Clan Leader", "credit\_id": "52fe48009251416c750aca97", "gender": 2, "id": 68278, "name": "Peter Mensah", "order": 27}, {"cast\_id": 49, "character": "Link Room Tech", "credit\_id": "52fe48009251416c750aca9b", "gender": 0, "id": 1207247, "name": "Sonnia Yee", "order": 28}, {"cast\_id": 50, "character": "Basketball Avatar / Troupe", "credit\_id": "52fe48009251416c750aca9f", "gender": 1, "id": 1207248, "name": "Jahnel Curfman", "order": 29}, {"cast\_id": 51, "character": "Basketball Avatar", "credit\_id": "52fe48009251416c750acaa3", "gender": 0, "id": 89714, "name": "Ilram Choi", "order": 30}, {"cast\_id": 52, "character": "Na\\'vi Child", "credit\_id": "52fe48009251416c750acaa7", "gender": 0, "id": 1207249, "name": "Kyla Warren", "order": 31}, {"cast\_id": 53, "character": "Troupe", "credit\_id": "52fe48009251416c750acaab", "gender": 0, "id": 1207250, "name": "Lisa Roumain", "order": 32}, {"cast\_id": 54, "character": "Troupe", "credit\_id": "52fe48009251416c750acaaf", "gender": 1, "id": 83105, "name": "Debra Wilson", "order": 33}, {"cast\_id": 57, "character": "Troupe", "credit\_id": "52fe48009251416c750acabb", "gender": 0, "id": 1207253, "name": "Chris Mala", "order": 34}, {"cast\_id": 55, "character": "Troupe", "credit\_id": "52fe48009251416c750acab3", "gender": 0, "id": 1207251, "name": "Taylor Kibby", "order": 35}, {"cast\_id": 56, "character": "Troupe", "credit\_id": "52fe48009251416c750acab7", "gender": 0, "id": 1207252, "name": "Jodie Landau", "order": 36}, {"cast\_id": 58, "character": "Troupe", "credit\_id": "52fe48009251416c750acabf", "gender": 0, "id": 1207254, "name": "Julie Lamm", "order": 37}, {"cast\_id": 59, "character": "Troupe", "credit\_id": "52fe48009251416c750acac3", "gender": 0, "id": 1207257, "name": "Cullen B. Madden", "order": 38}, {"cast\_id": 60, "character": "Troupe", "credit\_id": "52fe48009251416c750acac7", "gender": 0, "id": 1207259, "name": "Joseph Brady Madden", "order": 39}, {"cast\_id": 61, "character": "Troupe", "credit\_id": "52fe48009251416c750acacb", "gender": 0, "id": 1207262, "name": "Frankie Torres", "order": 40}, {"cast\_id": 62, "character": "Troupe", "credit\_id": "52fe48009251416c750acacf", "gender": 1, "id": 1158600, "name": "Austin Wilson", "order": 41}, {"cast\_id": 63, "character": "Troupe", "credit\_id": "52fe48019251416c750acad3", "gender": 1, "id": 983705, "name": "Sara Wilson", "order": 42}, {"cast\_id": 64, "character": "Troupe", "credit\_id": "52fe48019251416c750acad7", "gender": 0, "id": 1207263, "name": "Tamica Washington-Miller", "order": 43}, {"cast\_id": 65, "character": "Op Center Staff", "credit\_id": "52fe48019251416c750acadb", "gender": 1, "id": 1145098, "name": "Lucy Briant", "order": 44}, {"cast\_id": 66, "character": "Op Center Staff", "credit\_id": "52fe48019251416c750acadf", "gender": 2, "id": 33305, "name": "Nathan Meister", "order": 45}, {"cast\_id": 67, "character": "Op Center Staff", "credit\_id": "52fe48019251416c750acae3", "gender": 0, "id": 1207264, "name": "Gerry Blair", "order": 46}, {"cast\_id": 68, "character": "Op Center Staff", "credit\_id": "52fe48019251416c750acae7", "gender": 2, "id": 33311, "name": "Matthew Chamberlain", "order": 47}, {"cast\_id": 69, "character": "Op Center Staff", "credit\_id": "52fe48019251416c750acae9", "gender": 0, "id": 1207265, "name": "Paul Yates", "order": 48}, {"cast\_id": 70, "character": "Op Center Duty Officer", "credit\_id": "52fe48019251416c750acae7", "gender": 0, "id": 1207266, "name": "Wray Wilson", "order": 49}, {"cast\_id": 71, "character": "Op Center Staff", "credit\_id": "52fe48019251416c750acaf3", "gender": 2, "id": 54492, "name": "James Gaylyn", "order": 50}, {"cast\_id": 72, "character": "Dancer", "credit\_id": "52fe48019251416c750acaf7", "gender": 0, "id": 1207267, "name": "Melvin Leno Clark III", "order": 51}, {"cast\_id": 73, "character": "Dancer", "credit\_id": "52fe48019251416c750acafb", "gender": 0, "id": 1207268, "name": "Carvon Futrell", "order": 52}, {"cast\_id": 74, "character": "Dancer", "credit\_id": "52fe48019251416c750acaff", "gender": 0, "id": 1207269, "name": "Brandon Jelkes", "order": 53}, {"cast\_id": 75, "character": "Dancer", "credit\_id": "52fe48019251416c750acb03", "gender": 0, "id": 1207270, "name": "Micah Moch", "order": 54}, {"cast\_id": 76, "character": "Dancer", "credit\_id": "52fe48019251416c750acb07", "gender": 0, "id": 1207271, "name": "Hanniyah Muhammad", "order": 55}, {"cast\_id": 77, "character": "Dancer", "credit\_id": "52fe48019251416c750acb0b", "gender": 0, "id": 1207272, "name": "Christopher Nolen", "order": 56}, {"cast\_id": 78, "character": "Dancer", "credit\_id": "52fe48019251416c750acb0f", "gender": 0, "id": 1207273, "name": "Christa Oliver", "order": 57}, {"cast\_id": 79, "character": "Dancer", "credit\_id": "52fe48019251416c750acb13", "gender": 0, "id": 1207274, "name": "April Marie Thomas", "order": 58}, {"cast\_id": 80, "character": "Dancer", "credit\_id": "52fe48019251416c750acb17", "gender": 0, "id": 1207275, "name": "Bravita A. Threatt", "order": 59}, {"cast\_id": 81, "character": "Mining Chief (uncredited)", "credit\_id": "52fe48019251416c750acb1b", "gender": 0, "id": 1207276, "name": "Colin Bleasdale", "order": 60}, {"cast\_id": 82, "character": "Veteran Miner (uncredited)", "credit\_id": "52fe48019251416c750acb1f", "gender": 0, "id": 107969, "name": "Mike Bodnar", "order": 61}, {"cast\_id": 83, "character": "Richard (uncredited)", "credit\_id": "52fe48019251416c750acb23", "gender": 0, "id": 1207278, "name": "Matt Clayton", "order": 62}, {"cast\_id": 84, "character": "Nav\\'i (uncredited)", "credit\_id": "52fe48019251416c750acb27", "gender": 1, "id": 147898, "name": "Nicole Dionne", "order": 63}, {"cast\_id": 85, "character": "Trooper (uncredited)", "credit\_id": "52fe48019251416c750acb2b", "gender": 0, "id": 1207280, "name": "Jamie Harrison", "order": 64}, {"cast\_id": 86, "character": "Trooper (uncredited)", "credit\_id": "52fe48019251416c750acb2f", "gender": 0, "id": 1207281, "name": "Allan Henry", "order": 65}, {"cast\_id": 87, "character": "Ground Technician (uncredited)", "credit\_id": "52fe48019251416c750acb33", "gender": 2, "id": 1207282, "name": "Anthony Ingruber", "order": 66}, {"cast\_id": 88, "character": "Flight Crew Mechanic (uncredited)", "credit\_id": "52fe48019251416c750acb37", "gender": 0, "id": 1207283, "name": "Ashley Jeffery", "order": 67}, {"cast\_id": 14, "character": "Samson Pilot", "credit\_id": "52fe48009251416c750ac9f9", "gender": 0, "id": 98216, "name": "Dean Knowsley", "order": 68}, {"cast\_id": 89, "character": "Trooper (uncredited)", "credit\_id": "52fe48019251416c750acb3b", "gender": 0, "id": 1201399, "name": "Joseph Mika-Hunt", "order": 69}, {"cast\_id": 90, "character": "Banshee (uncredited)", "credit\_id": "52fe48019251416c750acb3f", "gender": 0, "id": 236696, "name": "Terry Notary", "order": 70}, {"cast\_id": 91, "character": "Soldier (uncredited)", "credit\_id": "52fe48019251416c750acb43", "gender": 0, "id": 1207287, "name": "Kai Pantano", "order": 71}, {"cast\_id": 92, "character": "Blast Technician (uncredited)", "credit\_id": "52fe48019251416c750acb47", "gender": 0, "id": 1207288, "name": "Logan Pithyou", "order": 72}, {"cast\_id": 93, "character": "Vindum Raah (uncredited)", "credit\_id": "52fe48019251416c750acb4b", "gender": 0, "id": 1207289, "name": "Stuart Pollock", "order": 73}, {"cast\_id": 94, "character": "Hero (uncredited)", "credit\_id": "52fe48019251416c750acb4f", "gender": 0, "id": 584868, "name": "Raja", "order": 74}, {"cast\_id": 95, "character": "Ops Centreworker (uncredited)", "credit\_id": "52fe48019251416c750acb53", "gender": 0, "id": 1207290, "name": "Gareth Ruck", "order": 75}, {"cast\_id": 96, "character": "Engineer (uncredited)", "credit\_id": "52fe48019251416c750acb57", "gender": 0, "id": 1062463, "name": "Rohan Sheehan", "order": 76}, {"cast\_id": 97, "character": "Col. Quaritch\\'s Mech Suit (uncredited)", "credit\_id": "52fe48019251416c750acb5b", "gender": 0, "id": 60656, "name": "T. J. Storm", "order": 77}, {"cast\_id": 98, "character": "Female Marine (uncredited)", "credit\_id": "52fe48019251416c750acb5f", "gender": 0, "id": 1207291, "name": "Jodie Taylor", "order": 78}, {"cast\_id": 99, "character": "Ikran Clan Leader (uncredited)", "credit\_id": "52fe48019251416c750acb63", "gender": 1, "id": 1186027, "name": "Alicia Vela-Bailey", "order": 79}, {"cast\_id": 100, "character": "Geologist (uncredited)", "credit\_id": "52fe48019251416c750acb67", "gender": 0, "id": 1207292, "name": "Richard Whiteside", "order": 80}, {"cast\_id": 101, "character": "Na\\'vi (uncredited)", "credit\_id": "52fe48019251416c750acb6b", "gender": 0, "id": 103259, "name": "Nikie Zambo", "order": 81}, {"cast\_id": 102

```
, "character": "Ambient Room Tech / Troupe", "credit_id": "52fe48019251416c750acb6f", "gender": 1, "id": 42286, "name": "Julene Renee", "order": 82}]'
```

```
In [28]: def converted3(obj):
        L = []
        counter = 0
        for i in ast.literal_eval(obj):
            if counter!=3:
                L.append(i['name'])
                counter+=1
            else:
                break
        return L
```

```
In [29]: movies['cast'] = movies['cast'].apply(converted3)
```

```
/var/folders/qg/km03k0gn1yn690bk_brhvb3c0000gn/T/ipykernel_62108/1808302032.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
movies['cast'] = movies['cast'].apply(converted3)
```

```
In [30]: movies.head()
```

	movie_id	title	overview	genres	keywords	cast	crew
0	19995	Avatar	In the 22nd century, a paraplegic Marine is di...	[Action, Adventure, Fantasy, Science Fiction]	[culture clash, future, space war, space colon...	[Sam Worthington, Zoe Saldana, Sigourney Weaver]	["credit_id": "52fe48009251416c750aca23", "de...
1	285	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...	[Adventure, Fantasy, Action]	[ocean, drug abuse, exotic island, east india ...	[Johnny Depp, Orlando Bloom, Keira Knightley]	["credit_id": "52fe4232c3a36847f800b579", "de...
2	206647	Spectre	A cryptic message from Bond's past sends him o...	[Action, Adventure, Crime]	[spy, based on novel, secret agent, sequel, mi...	[Daniel Craig, Christoph Waltz, Léa Seydoux]	["credit_id": "54805967c3a36829b5002c41", "de...
3	49026	The Dark Knight Rises	Following the death of District Attorney Harve...	[Action, Crime, Drama, Thriller]	[dc comics, crime fighter, terrorist, secret i...	[Christian Bale, Michael Caine, Gary Oldman]	["credit_id": "52fe4781c3a36847f81398c3", "de...
4	49529	John Carter	John Carter is a war-weary, former military ca...	[Action, Adventure, Science Fiction]	[based on novel, mars, medallion, space travel...	[Taylor Kitsch, Lynn Collins, Samantha Morton]	["credit_id": "52fe479ac3a36847f81398c3", "de...

```
In [31]: # Crew
# i only want the jobs where job = director and ignore everything else
```

```
In [32]: movies.iloc[0].crew
```

```
Out[32]: [{"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 Decor
```

ation", "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": "52fe48009251416c750acall", "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 Robertson"}, {"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": "54491a4dc3a3680ffc30018ca", "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": "54959e02c3a3680ffc60027d2", "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 Chermiawsky"}, {"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": "54959c859251416e1e003efe", "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 Karen"}, {"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": "544919d9c3a3680ffc30018bd", "department": "Art", "gender": 0, "id": 1339441, "job": "Set Designer", "name": "Tex Kadosaga"}, {"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 Cairra"}, {"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": 1376900, "job": "Dialogue Editor", "name": "Cheryl Nardi"}]

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": "549599cb3a3686ae1004613", "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 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": "54959d259251416e1e003f11", "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": "5495a1f7c3a3686ae3004443", "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": "5592b3ca9251415df8001081", "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": "5592b0db3a368774b00112c", "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"}]



```
rvisor", "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"}]
```

```
In [33]: def call_dir(obj):
        L = []
        for i in ast.literal_eval(obj):
            if i['job'] == 'Director':
                L.append(i['name'])
                break
        return L
```

```
In [34]: movies['crew'] = movies['crew'].apply(call_dir)
```

/var/folders/qg/km03k0gn1yn690bk\_brhvb3c0000gn/T/ipykernel\_62108/2114321798.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)

```
movies['crew'] = movies['crew'].apply(call_dir)
```

```
In [35]: movies.head(10)
```

	movie_id		title	overview	genres	keywords	cast	crew
0	19995		Avatar	In the 22nd century, a paraplegic Marine is di...	[Action, Adventure, Fantasy, Science Fiction]	[culture clash, future, space war, space colon...	[Sam Worthington, Zoe Saldana, Sigourney Weaver]	[James Cameron]
1	285		Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...	[Adventure, Fantasy, Action]	[ocean, drug abuse, exotic island, east india ...	[Johnny Depp, Orlando Bloom, Keira Knightley]	[Gore Verbinski]
2	206647		Spectre	A cryptic message from Bond's past sends him o...	[Action, Adventure, Crime]	[spy, based on novel, secret agent, sequel, mi...	[Daniel Craig, Christoph Waltz, Léa Seydoux]	[Sam Mendes]
3	49026	The Dark Knight Rises		Following the death of District Attorney Harve...	[Action, Crime, Drama, Thriller]	[dc comics, crime fighter, terrorist, secret i...	[Christian Bale, Michael Caine, Gary Oldman]	[Christopher Nolan]
4	49529	John Carter		John Carter is a war-weary, former military ca...	[Action, Adventure, Science Fiction]	[based on novel, mars, medallion, space travel...	[Taylor Kitsch, Lynn Collins, Samantha Morton]	[Andrew Stanton]
5	559	Spider-Man 3		The seemingly invincible Spider-Man goes up ag...	[Fantasy, Action, Adventure]	[dual identity, amnesia, sandstorm, love of on...	[Tobey Maguire, Kirsten Dunst, James Franco]	[Sam Raimi]
6	38757	Tangled		When the kingdom's most wanted-and most charmi...	[Animation, Family]	[hostage, magic, horse, fairy tale, musical, p...	[Zachary Levi, Mandy Moore, Donna Murphy]	[Byron Howard]
7	99861	Avengers: Age of Ultron		When Tony Stark tries to jumpstart a dormant p...	[Action, Adventure, Science Fiction]	[marvel comic, sequel, superhero, based on com...	[Robert Downey Jr., Chris Hemsworth, Mark Ruff...	[Joss Whedon]
8	767	Harry Potter and the Half-Blood Prince		As Harry begins his sixth year at Hogwarts, he...	[Adventure, Fantasy, Family]	[witch, magic, broom, school of witchcraft, wi...	[Daniel Radcliffe, Rupert Grint, Emma Watson]	[David Yates]
9	209112	Batman v Superman: Dawn of Justice		Fearing the actions of a god-like Super Hero l...	[Action, Adventure, Fantasy]	[dc comics, vigilante, superhero, based on com...	[Ben Affleck, Henry Cavill, Gal Gadot]	[Zack Snyder]

```
In [36]: # overview
        # now overview is in the form of a string we just want to convert it into list
```

```
In [37]: movies.iloc[0].overview
```

```
Out[37]: 'In the 22nd century, a paraplegic Marine is dispatched to the moon Pandora on a unique mission, but becomes torn between following orders and protecting an alien civilization.'
```

```
In [38]: movies['overview'] = movies['overview'].apply(lambda x:x.split())
```

```

/var/folders/qg/km03k0gn1yn690bk_brhvb3c0000gn/T/ipykernel_62108/3942134491.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
movies['overview'] = movies['overview'].apply(lambda x:x.split())

```

In [39]: `movies.head()`

Out[39]:

	movie_id	title	overview	genres	keywords	cast	crew
0	19995	Avatar	[In, the, 22nd, century,, a, paraplegic, Marin...	[Action, Adventure, Fantasy, Science Fiction]	[culture clash, future, space war, space colon...	[Sam Worthington, Zoe Saldana, Sigourney Weaver]	[James Cameron]
1	285	Pirates of the Caribbean: At World's End	[Captain, Barbossa,, long, believed, to, be, d...	[Adventure, Fantasy, Action]	[ocean, drug abuse, exotic island, east india ...	[Johnny Depp, Orlando Bloom, Keira Knightley]	[Gore Verbinski]
2	206647	Spectre	[A, cryptic, message, from, Bond's, past, send...	[Action, Adventure, Crime]	[spy, based on novel, secret agent, sequel, mi...	[Daniel Craig, Christoph Waltz, Léa Seydoux]	[Sam Mendes]
3	49026	The Dark Knight Rises	[Following, the, death, of, District, Attorney...	[Action, Crime, Drama, Thriller]	[dc comics, crime fighter, terrorist, secret i...	[Christian Bale, Michael Caine, Gary Oldman]	[Christopher Nolan]
4	49529	John Carter	[John, Carter, is, a, war-weary,, former, mili...	[Action, Adventure, Science Fiction]	[based on novel, mars, medallion, space travel...	[Taylor Kitsch, Lynn Collins, Samantha Morton]	[Andrew Stanton]

In [40]: `# Now i want to concat all the lists together in 1 big list with column name tags`

In [41]: `movies['genres'] = movies['genres'].apply(lambda x:[i.replace(" ",",") for i in x])`

```

/var/folders/qg/km03k0gn1yn690bk_brhvb3c0000gn/T/ipykernel_62108/3381874298.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
movies['genres'] = movies['genres'].apply(lambda x:[i.replace(" ",",") for i in x])

```

In [42]: `movies['keywords'] = movies['keywords'].apply(lambda x:[i.replace(" ",",") for i in x])`  
`movies['cast'] = movies['cast'].apply(lambda x:[i.replace(" ",",") for i in x])`  
`movies['crew'] = movies['crew'].apply(lambda x:[i.replace(" ",",") for i in x])`

```

/var/folders/qg/km03k0gn1yn690bk_brhvb3c0000gn/T/ipykernel_62108/3553446180.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
movies['keywords'] = movies['keywords'].apply(lambda x:[i.replace(" ",",") for i in x])
/var/folders/qg/km03k0gn1yn690bk_brhvb3c0000gn/T/ipykernel_62108/3553446180.py:2: 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
movies['cast'] = movies['cast'].apply(lambda x:[i.replace(" ",",") for i in x])
/var/folders/qg/km03k0gn1yn690bk_brhvb3c0000gn/T/ipykernel_62108/3553446180.py:3: 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
movies['crew'] = movies['crew'].apply(lambda x:[i.replace(" ",",") for i in x])

```

In [43]: `movies`

Out [43]:

	movie_id		title	overview	genres	keywords	cast	crew
0	19995		Avatar	[In, the, 22nd, century,, a, paraplegic, Marin...	[Action, Adventure, Fantasy, ScienceFiction]	[cultureclash, future, spacewar, spacecolony, ...	[SamWorthington, ZoeSaldana, SigourneyWeaver]	[JamesCameron]
1	285		Pirates of the Caribbean: At World's End	[Captain, Barbossa,, long, believed, to, be, d...	[Adventure, Fantasy, Action]	[ocean, drugabuse, exoticisland, eastindiatrad...	[JohnnyDepp, OrlandoBloom, KeiraKnightley]	[GoreVerbinski]
2	206647		Spectre	[A, cryptic, message, from, Bond's, past, send...	[Action, Adventure, Crime]	[spy, basedonnovel, secretagent, sequel, mi6, ...	[DanielCraig, ChristophWaltz, LéaSeydoux]	[SamMendes]
3	49026		The Dark Knight Rises	[Following, the, death, of, District, Attorney...	[Action, Crime, Drama, Thriller]	[dccomics, crimefighter, terrorist, secretiden...	[ChristianBale, MichaelCaine, GaryOldman]	[ChristopherNolan]
4	49529		John Carter	[John, Carter, is, a, war-weary,, former, mili...	[Action, Adventure, ScienceFiction]	[basedonnovel, mars, medallion, spacetravel, p...	[TaylorKitsch, LynnCollins, SamanthaMorton]	[AndrewStanton]
...	...		...	...	...	...	...	...
4804	9367		El Mariachi	[El, Mariachi, just, wants, to, play, his, gui...	[Action, Crime, Thriller]	[unitedstates-mexicobarrier, legs, arms, paper...	[CarlosGallardo, JaimeHoyos, PeterMarquardt]	[RobertRodriguez]
4805	72766		Newlyweds	[A, newlywed, couple's, honeymoon, is, upended...	[Comedy, Romance]	[]	[EdwardBurns, KerryBishé, MarshaDietlein]	[EdwardBurns]
4806	231617		Signed, Sealed, Delivered	["Signed,, Sealed,, Delivered", introduces, a,...	[Comedy, Drama, Romance, TVMovie]	[date, loveatfirstsight, narration, investigat...	[EricMabius, KristinBooth, CrystalLowe]	[ScottSmith]
4807	126186		Shanghai Calling	[When, ambitious, New, York, attorney, Sam, is...	[]	[]	[DanielHenney, ElizaCoupe, BillPaxton]	[DanielHsia]
4808	25975		My Date with Drew	[Ever, since, the, second, grade, when, he, fi...	[Documentary]	[obsession, camcorder, crush, dreamgirl]	[DrewBarrymore, BrianHerzlinger, CoreyFeldman]	[BrianHerzlinger]

4806 rows × 7 columns

```
In [44]: movies['tags'] = movies['overview'] + movies['genres'] + movies['keywords'] + movies['cast'] + movies['crew']
```

/var/folders/qg/km03k0gn1yn690bk\_brhvb3c0000gn/T/ipykernel\_62108/603755214.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)

```
movies['tags'] = movies['overview'] + movies['genres'] + movies['keywords'] + movies['cast'] + movies['crew']
```

```
In [45]: movies.head()
```



Out [45]:

	movie_id	title	overview	genres	keywords	cast	crew	tags
0	19995	Avatar	[In, the, 22nd, century,, a, paraplegic, Marin...	[Action, Adventure, Fantasy, ScienceFiction]	[cultureclash, future, spacewar, spacecolony, ...	[SamWorthington, ZoeSaldana, SigourneyWeaver]	[JamesCameron]	[In, the, 22nd, century,, a, paraplegic, Marin...
1	285	Pirates of the Caribbean: At World's End	[Captain, Barbossa,, long, believed, to, be, d...	[Adventure, Fantasy, Action]	[ocean, drugabuse, exoticisland, eastindiatrad...	[JohnnyDepp, OrlandoBloom, KeiraKnightley]	[GoreVerbinski]	[Captain, Barbossa,, long, believed, to, be, d...
2	206647	Spectre	[A, cryptic, message, from, Bond's, past, send...	[Action, Adventure, Crime]	[spy, basedonnovel, secretagent, sequel, mi6, ...	[DanielCraig, ChristophWaltz, LéaSeydoux]	[SamMendes]	[A, cryptic, message, from, Bond's, past, send...
3	49026	The Dark Knight Rises	[Following, the, death, of, District, Attorney...	[Action, Crime, Drama, Thriller]	[dccomics, crimefighter, terrorist, secretiden...	[ChristianBale, MichaelCaine, GaryOldman]	[ChristopherNolan]	[Following, the, death, of, District, Attorney...
4	49529	John Carter	[John, Carter, is, a, war-weary,, former, mili...	[Action, Adventure, ScienceFiction]	[basedonnovel, mars, medallion, spacetravel, p...	[TaylorKitsch, LynnCollins, SamanthaMorton]	[AndrewStanton]	[John, Carter, is, a, war-weary,, former, mili...

In [46]:

```
new_moviesdf = movies[['movie_id','title','tags']]
```

In [47]:

```
# converting tags to string
new_moviesdf['tags'] = new_moviesdf['tags'].apply(lambda x: " ".join(x))
```

/var/folders/qg/km03k0gn1yn690bk\_brhvb3c0000gn/T/ipykernel\_62108/1336399018.py:2: 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
new\_moviesdf['tags'] = new\_moviesdf['tags'].apply(lambda x: " ".join(x))

In [48]:

```
new_moviesdf.head(10)
```

Out [48]:

	movie_id	title	tags
0	19995	Avatar	In the 22nd century, a paraplegic Marine is di...
1	285	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...
2	206647	Spectre	A cryptic message from Bond's past sends him o...
3	49026	The Dark Knight Rises	Following the death of District Attorney Harve...
4	49529	John Carter	John Carter is a war-weary, former military ca...
5	559	Spider-Man 3	The seemingly invincible Spider-Man goes up ag...
6	38757	Tangled	When the kingdom's most wanted-and most charmi...
7	99861	Avengers: Age of Ultron	When Tony Stark tries to jumpstart a dormant p...
8	767	Harry Potter and the Half-Blood Prince	As Harry begins his sixth year at Hogwarts, he...
9	209112	Batman v Superman: Dawn of Justice	Fearing the actions of a god-like Super Hero l...

In [49]:

```
new_moviesdf['tags'][0]
```

Out [49]:

```
'In the 22nd century, a paraplegic Marine is dispatched to the moon Pandora on a unique mission, but becomes torn between following orders and protecting an alien civilization. Action Adventure Fantasy ScienceFiction cultureclash future spacewar spacecolony society spacetravel futuristic romance space alien tribe alienplanet cgi marine soldier battle loveaffair antiwar powerrelations mindandsoul 3d SamWorthington ZoeSaldana SigourneyWeaver JamesCameron'
```

In [50]:

```
new_moviesdf['tags'] = new_moviesdf['tags'].apply(lambda x:x.lower())
```

/var/folders/qg/km03k0gn1yn690bk\_brhvb3c0000gn/T/ipykernel\_62108/3736736714.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
new\_moviesdf['tags'] = new\_moviesdf['tags'].apply(lambda x:x.lower())

```

In [51]: new_moviesdf.head()

```

	movie_id	title	tags
0	19995	Avatar	in the 22nd century, a paraplegic marine is di...
1	285	Pirates of the Caribbean: At World's End	captain barbossa, long believed to be dead, ha...
2	206647	Spectre	a cryptic message from bond's past sends him o...
3	49026	The Dark Knight Rises	following the death of district attorney harve...
4	49529	John Carter	john carter is a war-weary, former military ca...

```

In [52]: import nltk
from nltk.stem.porter import PorterStemmer

In [53]: ps = PorterStemmer()
def stem(text):
    y = []
    for i in text.split():
        y.append(ps.stem(i))
    return " ".join(y)

In [54]: new_moviesdf['tags'] = new_moviesdf['tags'].apply(stem)

/var/folders/qg/km03k0gn1yn690bk_brhvb3c0000gn/T/ipykernel_62108/1731500766.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
new_moviesdf['tags'] = new_moviesdf['tags'].apply(stem)

In [55]: # Vectorization
# calculate similarity score
# convert the text into vectors using techniques like bag of words, tfidf, word2vec
# we are using bag of words

In [56]: from sklearn.feature_extraction.text import CountVectorizer
cv = CountVectorizer(max_features=5000, stop_words='english')

In [57]: vectors = cv.fit_transform(new_moviesdf['tags']).toarray()

In [58]: vectors

Out[58]: array([[0, 0, 0, ..., 0, 0, 0],
[0, 0, 0, ..., 0, 0, 0],
[0, 0, 0, ..., 0, 0, 0],
...,
[0, 0, 0, ..., 0, 0, 0],
[0, 0, 0, ..., 0, 0, 0],
[0, 0, 0, ..., 0, 0, 0]])

In [59]: cv.get_feature_names_out()[30:100]

Out[59]: array(['1980', '1985', '1990', '1999', '19th', '19thcenturi', '20', '200',
'2003', '2009', '20th', '21st', '23', '24', '25', '30', '300',
'3d', '40', '50', '500', '60', '70', '80', 'aaron', 'aaroneckhart',
'abandon', 'abduct', 'abigailbreslin', 'abil', 'abl', 'aboard',
'abov', 'abus', 'academ', 'academi', 'accept', 'access', 'accid',
'accident', 'acclaim', 'accompani', 'accomplish', 'account',
'accus', 'ace', 'achiev', 'acquaint', 'act', 'action',
'actionhero', 'activ', 'activist', 'activities', 'actor',
'actress', 'actual', 'ad', 'adam', 'adamsandl', 'adamshankman',
'adapt', 'add', 'addict', 'adjust', 'admir', 'admit', 'adolesc',
'adopt', 'ador'], dtype=object)

In [60]: # Calculating the distance between the vectors
# We will be using cosine distance instead of euclidean distance as euclidean distance is not a good major when
# More the distance less the similarity

In [61]: from sklearn.metrics.pairwise import cosine_similarity

In [62]: similarity = cosine_similarity(vectors)

In [63]: # Calculating the index of the movies
similarity[1]

Out[63]: array([0.08346223, 1.          , 0.06063391, ..., 0.02378257, 0.
0.02615329])

```

```
In [64]: new_moviesdf[new_moviesdf['title']=='Avatar']
```

```
Out[64]:
```

	movie_id	title	tags
0	19995	Avatar	in the 22nd century, a parapleg marin is dispa...

```
In [65]: new_moviesdf[new_moviesdf['title']=='Avatar'].index[0]
```

```
Out[65]: 0
```

```
In [66]: sorted(list(enumerate(similarity[0])),reverse=True)
```

```
Out[66]: [(4805, 0.0),
(4804, 0.0),
(4803, 0.04499212706658476),
(4802, 0.046829290579084706),
(4801, 0.019252140716412975),
(4800, 0.0),
(4799, 0.052631578947368425),
(4798, 0.04223886030955117),
(4797, 0.0),
(4796, 0.0),
(4795, 0.0),
(4794, 0.0),
(4793, 0.05407380704358751),
(4792, 0.0),
(4791, 0.0),
(4790, 0.0582716546748065),
(4789, 0.060833032924035954),
(4788, 0.0),
(4787, 0.019672236884115842),
(4786, 0.0),
(4785, 0.019672236884115842),
(4784, 0.043355498476206004),
(4783, 0.0),
(4782, 0.025649458802128853),
(4781, 0.0582716546748065),
(4780, 0.0),
(4779, 0.0),
(4778, 0.0),
(4777, 0.10814761408717502),
(4776, 0.0),
(4775, 0.05647824947249051),
(4774, 0.0),
(4773, 0.029617443887954616),
(4772, 0.023918243661746996),
(4771, 0.039344473768231684),
(4770, 0.0),
(4769, 0.0),
(4768, 0.0),
(4767, 0.03627381250550058),
(4766, 0.017195436249022724),
(4765, 0.0),
(4764, 0.0),
(4763, 0.0),
(4762, 0.0),
(4761, 0.02742042485535409),
(4760, 0.025334729596907),
(4759, 0.019389168358237032),
(4758, 0.0),
(4757, 0.046829290579084706),
(4756, 0.0),
(4755, 0.026315789473684213),
(4754, 0.08603090020146065),
(4753, 0.0),
(4752, 0.029617443887954616),
(4751, 0.021119430154775586),
(4750, 0.0),
(4749, 0.0),
(4748, 0.01981850088223545),
(4747, 0.0),
(4746, 0.0),
(4745, 0.0),
(4744, 0.03065696697424829),
(4743, 0.014627012952620635),
(4742, 0.040555355282690636),
(4741, 0.0),
(4740, 0.0),
(4739, 0.025649458802128853),
(4738, 0.0),
(4737, 0.0837707816583391),
(4736, 0.0),
```

(4735, 0.0),  
(4734, 0.024115646544478436),  
(4733, 0.0),  
(4732, 0.0),  
(4731, 0.02782074420373286),  
(4730, 0.02294157338705618),  
(4729, 0.0),  
(4728, 0.0),  
(4727, 0.0533380747062665),  
(4726, 0.0),  
(4725, 0.0),  
(4724, 0.0),  
(4723, 0.0),  
(4722, 0.0415406492397264),  
(4721, 0.0),  
(4720, 0.028676966733820218),  
(4719, 0.0),  
(4718, 0.040555355282690636),  
(4717, 0.0),  
(4716, 0.0),  
(4715, 0.0),  
(4714, 0.0),  
(4713, 0.023174488732966077),  
(4712, 0.02271554252121273),  
(4711, 0.056573571480353566),  
(4710, 0.0),  
(4709, 0.0),  
(4708, 0.0),  
(4707, 0.025649458802128853),  
(4706, 0.011302501972710611),  
(4705, 0.0),  
(4704, 0.06362847629757779),  
(4703, 0.03382550457458692),  
(4702, 0.0),  
(4701, 0.0),  
(4700, 0.03244428422615251),  
(4699, 0.05564148840746572),  
(4698, 0.0),  
(4697, 0.0),  
(4696, 0.014567913668701625),  
(4695, 0.04024218182927669),  
(4694, 0.022282825891079324),  
(4693, 0.0602475233128778),  
(4692, 0.0),  
(4691, 0.052631578947368425),  
(4690, 0.023918243661746996),  
(4689, 0.0),  
(4688, 0.0),  
(4687, 0.0),  
(4686, 0.0),  
(4685, 0.0),  
(4684, 0.0),  
(4683, 0.0),  
(4682, 0.0),  
(4681, 0.0),  
(4680, 0.05923488777590923),  
(4679, 0.04499212706658476),  
(4678, 0.0),  
(4677, 0.0),  
(4676, 0.0),  
(4675, 0.027036903521793755),  
(4674, 0.02249606353329238),  
(4673, 0.11846977555181847),  
(4672, 0.04588314677411236),  
(4671, 0.028239124736245257),  
(4670, 0.03673591791853225),  
(4669, 0.028676966733820218),  
(4668, 0.0),  
(4667, 0.02742042485535409),  
(4666, 0.03181423814878889),  
(4665, 0.0),  
(4664, 0.06131393394849658),  
(4663, 0.0),  
(4662, 0.0207703246198632),  
(4661, 0.03181423814878889),  
(4660, 0.0),  
(4659, 0.031219527052723135),  
(4658, 0.0),  
(4657, 0.08111071056538127),  
(4656, 0.019389168358237032),  
(4655, 0.039344473768231684),  
(4654, 0.08111071056538127),  
(4653, 0.0),

(4652, 0.06131393394849658),  
(4651, 0.040555355282690636),  
(4650, 0.0),  
(4649, 0.0),  
(4648, 0.018731716231633878),  
(4647, 0.021119430154775586),  
(4646, 0.039344473768231684),  
(4645, 0.0),  
(4644, 0.0),  
(4643, 0.04188539082916955),  
(4642, 0.0331133089266261),  
(4641, 0.08111071056538127),  
(4640, 0.07254762501100118),  
(4639, 0.0),  
(4638, 0.07079923254047887),  
(4637, 0.0),  
(4636, 0.057353933467640436),  
(4635, 0.039344473768231684),  
(4634, 0.025976216673306556),  
(4633, 0.02666903735313325),  
(4632, 0.03458572319330373),  
(4631, 0.0),  
(4630, 0.03419927840283847),  
(4629, 0.02913582733740325),  
(4628, 0.0),  
(4627, 0.04024218182927669),  
(4626, 0.08111071056538127),  
(4625, 0.0),  
(4624, 0.07694837640638656),  
(4623, 0.06488856845230502),  
(4622, 0.03458572319330373),  
(4621, 0.0),  
(4620, 0.025031308716087945),  
(4619, 0.04499212706658476),  
(4618, 0.029617443887954616),  
(4617, 0.0),  
(4616, 0.0),  
(4615, 0.0),  
(4614, 0.0),  
(4613, 0.0),  
(4612, 0.0),  
(4611, 0.03382550457458692),  
(4610, 0.0),  
(4609, 0.0),  
(4608, 0.0),  
(4607, 0.0415406492397264),  
(4606, 0.04188539082916955),  
(4605, 0.02666903735313325),  
(4604, 0.03382550457458692),  
(4603, 0.0),  
(4602, 0.05407380704358751),  
(4601, 0.0),  
(4600, 0.01981850088223545),  
(4599, 0.0),  
(4598, 0.0),  
(4597, 0.021486752129677003),  
(4596, 0.051298917604257706),  
(4595, 0.0),  
(4594, 0.0),  
(4593, 0.0),  
(4592, 0.0),  
(4591, 0.057353933467640436),  
(4590, 0.04783648732349399),  
(4589, 0.02666903735313325),  
(4588, 0.0),  
(4587, 0.08671099695241201),  
(4586, 0.0),  
(4585, 0.0),  
(4584, 0.0),  
(4583, 0.0),  
(4582, 0.052786150730834015),  
(4581, 0.0331133089266261),  
(4580, 0.03181423814878889),  
(4579, 0.029617443887954616),  
(4578, 0.0582716546748065),  
(4577, 0.0),  
(4576, 0.02294157338705618),  
(4575, 0.0),  
(4574, 0.017005425596285897),  
(4573, 0.03721614637823934),  
(4572, 0.0),  
(4571, 0.0),  
(4570, 0.0301237616564389),

(4569, 0.028676966733820218),  
(4568, 0.08111071056538127),  
(4567, 0.0),  
(4566, 0.06131393394849658),  
(4565, 0.014686837386151168),  
(4564, 0.04188539082916955),  
(4563, 0.0),  
(4562, 0.0),  
(4561, 0.0),  
(4560, 0.08471737420873576),  
(4559, 0.03673591791853225),  
(4558, 0.076004188790721),  
(4557, 0.03823595564509363),  
(4556, 0.0),  
(4555, 0.040555355282690636),  
(4554, 0.03627381250550058),  
(4553, 0.024455799402225926),  
(4552, 0.04120428217151646),  
(4551, 0.0),  
(4550, 0.0),  
(4549, 0.023918243661746996),  
(4548, 0.0),  
(4547, 0.0),  
(4546, 0.0),  
(4545, 0.09558988911273408),  
(4544, 0.0),  
(4543, 0.02742042485535409),  
(4542, 0.07443229275647868),  
(4541, 0.0),  
(4540, 0.05407380704358751),  
(4539, 0.0),  
(4538, 0.13629325512727639),  
(4537, 0.0),  
(4536, 0.02913582733740325),  
(4535, 0.03382550457458692),  
(4534, 0.0),  
(4533, 0.051298917604257706),  
(4532, 0.033463724070512735),  
(4531, 0.0),  
(4530, 0.050669459193814),  
(4529, 0.01255307044856882),  
(4528, 0.04891159880445185),  
(4527, 0.0),  
(4526, 0.0),  
(4525, 0.04456565178215865),  
(4524, 0.023918243661746996),  
(4523, 0.0),  
(4522, 0.013809204416652687),  
(4521, 0.0),  
(4520, 0.0),  
(4519, 0.05006261743217589),  
(4518, 0.0),  
(4517, 0.0),  
(4516, 0.08671099695241201),  
(4515, 0.0),  
(4514, 0.02913582733740325),  
(4513, 0.0),  
(4512, 0.05006261743217589),  
(4511, 0.027036903521793755),  
(4510, 0.025649458802128853),  
(4509, 0.0),  
(4508, 0.0),  
(4507, 0.0),  
(4506, 0.07987230638308718),  
(4505, 0.029617443887954616),  
(4504, 0.0),  
(4503, 0.0),  
(4502, 0.09365858115816941),  
(4501, 0.0),  
(4500, 0.0),  
(4499, 0.0),  
(4498, 0.025031308716087945),  
(4497, 0.0),  
(4496, 0.0),  
(4495, 0.0),  
(4494, 0.0),  
(4493, 0.0),  
(4492, 0.08111071056538127),  
(4491, 0.0),  
(4490, 0.09933992677987831),  
(4489, 0.015682536711195138),  
(4488, 0.0),  
(4487, 0.021873931962990357),

(4486, 0.03382550457458692),  
(4485, 0.029617443887954616),  
(4484, 0.020942695414584774),  
(4483, 0.07098727864204515),  
(4482, 0.03244428422615251),  
(4481, 0.0),  
(4480, 0.0),  
(4479, 0.0),  
(4478, 0.0),  
(4477, 0.03823595564509363),  
(4476, 0.0),  
(4475, 0.0),  
(4474, 0.03627381250550058),  
(4473, 0.07868894753646337),  
(4472, 0.09733285267845754),  
(4471, 0.03797316245786213),  
(4470, 0.0),  
(4469, 0.025649458802128853),  
(4468, 0.11128297681493143),  
(4467, 0.028676966733820218),  
(4466, 0.03905832834322535),  
(4465, 0.0),  
(4464, 0.0),  
(4463, 0.03627381250550058),  
(4462, 0.0),  
(4461, 0.0),  
(4460, 0.03721614637823934),  
(4459, 0.05195243334661311),  
(4458, 0.04891159880445185),  
(4457, 0.0),  
(4456, 0.0),  
(4455, 0.0),  
(4454, 0.0),  
(4453, 0.0415406492397264),  
(4452, 0.0),  
(4451, 0.04499212706658476),  
(4450, 0.025334729596907),  
(4449, 0.0),  
(4448, 0.04461243988090516),  
(4447, 0.0),  
(4446, 0.023662426214015053),  
(4445, 0.0),  
(4444, 0.0),  
(4443, 0.07254762501100116),  
(4442, 0.0),  
(4441, 0.0),  
(4440, 0.0),  
(4439, 0.0),  
(4438, 0.0),  
(4437, 0.0331133089266261),  
(4436, 0.0),  
(4435, 0.0),  
(4434, 0.03244428422615251),  
(4433, 0.03181423814878889),  
(4432, 0.0),  
(4431, 0.07254762501100116),  
(4430, 0.0),  
(4429, 0.05546053999849018),  
(4428, 0.0),  
(4427, 0.04456565178215865),  
(4426, 0.0),  
(4425, 0.0),  
(4424, 0.0),  
(4423, 0.0),  
(4422, 0.025031308716087945),  
(4421, 0.03382550457458692),  
(4420, 0.029617443887954616),  
(4419, 0.0),  
(4418, 0.0),  
(4417, 0.0),  
(4416, 0.0),  
(4415, 0.0),  
(4414, 0.03382550457458692),  
(4413, 0.023174488732966077),  
(4412, 0.07443229275647868),  
(4411, 0.0),  
(4410, 0.025649458802128853),  
(4409, 0.0),  
(4408, 0.0602475233128778),  
(4407, 0.0),  
(4406, 0.028239124736245257),  
(4405, 0.18731716231633883),  
(4404, 0.0),

(4403, 0.09733285267845754),  
(4402, 0.0301237616564389),  
(4401, 0.024738534799764674),  
(4400, 0.039344473768231684),  
(4399, 0.04588314677411236),  
(4398, 0.02294157338705618),  
(4397, 0.021300716142115247),  
(4396, 0.0),  
(4395, 0.0),  
(4394, 0.0),  
(4393, 0.0),  
(4392, 0.06814662756363819),  
(4391, 0.0),  
(4390, 0.03823595564509363),  
(4389, 0.02782074420373286),  
(4388, 0.01813690625275029),  
(4387, 0.0),  
(4386, 0.0),  
(4385, 0.0),  
(4384, 0.043747863925980714),  
(4383, 0.03498557142706598),  
(4382, 0.03993615319154359),  
(4381, 0.0),  
(4380, 0.025334729596907),  
(4379, 0.0),  
(4378, 0.0),  
(4377, 0.05484084971070818),  
(4376, 0.0),  
(4375, 0.040555355282690636),  
(4374, 0.0),  
(4373, 0.0),  
(4372, 0.046829290579084706),  
(4371, 0.023662426214015053),  
(4370, 0.0),  
(4369, 0.052631578947368425),  
(4368, 0.03823595564509363),  
(4367, 0.0),  
(4366, 0.0),  
(4365, 0.043355498476206004),  
(4364, 0.0),  
(4363, 0.0),  
(4362, 0.04188539082916955),  
(4361, 0.0),  
(4360, 0.0),  
(4359, 0.021300716142115247),  
(4358, 0.05990422978731538),  
(4357, 0.06243905410544627),  
(4356, 0.0),  
(4355, 0.06362847629757779),  
(4354, 0.0),  
(4353, 0.08998425413316952),  
(4352, 0.0),  
(4351, 0.0662266178532522),  
(4350, 0.0),  
(4349, 0.02913582733740325),  
(4348, 0.13497638119975428),  
(4347, 0.0),  
(4346, 0.0),  
(4345, 0.0),  
(4344, 0.0),  
(4343, 0.08998425413316952),  
(4342, 0.0),  
(4341, 0.07792865001991967),  
(4340, 0.0),  
(4339, 0.05407380704358751),  
(4338, 0.0),  
(4337, 0.0),  
(4336, 0.17521916101261562),  
(4335, 0.022282825891079324),  
(4334, 0.0),  
(4333, 0.04188539082916955),  
(4332, 0.03627381250550058),  
(4331, 0.09365858115816941),  
(4330, 0.05735393346764045),  
(4329, 0.050669459193814),  
(4328, 0.0),  
(4327, 0.025976216673306556),  
(4326, 0.03721614637823934),  
(4325, 0.046348977465932154),  
(4324, 0.03065696697424829),  
(4323, 0.025976216673306556),  
(4322, 0.0),  
(4321, 0.0),



(4320, 0.06131393394849658),  
(4319, 0.03166237934306518),  
(4318, 0.0),  
(4317, 0.02249606353329238),  
(4316, 0.07894736842105264),  
(4315, 0.0),  
(4314, 0.019117977822546817),  
(4313, 0.028239124736245257),  
(4312, 0.06362847629757779),  
(4311, 0.0),  
(4310, 0.038778336716474064),  
(4309, 0.0),  
(4308, 0.0),  
(4307, 0.0),  
(4306, 0.0),  
(4305, 0.039344473768231684),  
(4304, 0.0),  
(4303, 0.0),  
(4302, 0.0),  
(4301, 0.0),  
(4300, 0.040555355282690636),  
(4299, 0.024182541670333724),  
(4298, 0.0),  
(4297, 0.03181423814878889),  
(4296, 0.0),  
(4295, 0.0),  
(4294, 0.0),  
(4293, 0.057353933467640436),  
(4292, 0.0),  
(4291, 0.0),  
(4290, 0.06917144638660747),  
(4289, 0.0),  
(4288, 0.025976216673306556),  
(4287, 0.023174488732966077),  
(4286, 0.023414645289542353),  
(4285, 0.0),  
(4284, 0.024738534799764674),  
(4283, 0.027036903521793755),  
(4282, 0.03627381250550058),  
(4281, 0.051298917604257706),  
(4280, 0.023662426214015053),  
(4279, 0.07868894753646337),  
(4278, 0.0),  
(4277, 0.018731716231633878),  
(4276, 0.0),  
(4275, 0.0),  
(4274, 0.0),  
(4273, 0.0),  
(4272, 0.0),  
(4271, 0.06362847629757779),  
(4270, 0.057353933467640436),  
(4269, 0.0),  
(4268, 0.0),  
(4267, 0.0),  
(4266, 0.07694837640638656),  
(4265, 0.0207703246198632),  
(4264, 0.0),  
(4263, 0.04891159880445185),  
(4262, 0.025031308716087945),  
(4261, 0.024182541670333724),  
(4260, 0.0),  
(4259, 0.039344473768231684),  
(4258, 0.026315789473684213),  
(4257, 0.057353933467640436),  
(4256, 0.0),  
(4255, 0.0),  
(4254, 0.09733285267845754),  
(4253, 0.0),  
(4252, 0.0),  
(4251, 0.0),  
(4250, 0.012988108336653278),  
(4249, 0.0),  
(4248, 0.024738534799764674),  
(4247, 0.019968076595771794),  
(4246, 0.039344473768231684),  
(4245, 0.0),  
(4244, 0.05407380704358751),  
(4243, 0.0),  
(4242, 0.029494803841956827),  
(4241, 0.040555355282690636),  
(4240, 0.0),  
(4239, 0.0),  
(4238, 0.023174488732966077),

(4237, 0.0),  
(4236, 0.05006261743217589),  
(4235, 0.042973504259354006),  
(4234, 0.04188539082916955),  
(4233, 0.016303866268150616),  
(4232, 0.05195243334661311),  
(4231, 0.0),  
(4230, 0.0),  
(4229, 0.0),  
(4228, 0.12824729401064427),  
(4227, 0.0),  
(4226, 0.03181423814878889),  
(4225, 0.022282825891079324),  
(4224, 0.0),  
(4223, 0.041278324068019595),  
(4222, 0.0),  
(4221, 0.0),  
(4220, 0.0),  
(4219, 0.0),  
(4218, 0.03065696697424829),  
(4217, 0.0),  
(4216, 0.027036903521793755),  
(4215, 0.02271554252121273),  
(4214, 0.06765100914917384),  
(4213, 0.024738534799764674),  
(4212, 0.023174488732966077),  
(4211, 0.05407380704358751),  
(4210, 0.0),  
(4209, 0.03065696697424829),  
(4208, 0.0),  
(4207, 0.027036903521793755),  
(4206, 0.0),  
(4205, 0.02043797798283219),  
(4204, 0.0331133089266261),  
(4203, 0.07509392614826384),  
(4202, 0.03065696697424829),  
(4201, 0.09544271444636668),  
(4200, 0.0331133089266261),  
(4199, 0.0),  
(4198, 0.0),  
(4197, 0.0),  
(4196, 0.09933992677987831),  
(4195, 0.0),  
(4194, 0.06131393394849658),  
(4193, 0.03382550457458692),  
(4192, 0.02249606353329238),  
(4191, 0.02913582733740325),  
(4190, 0.2029530274475215),  
(4189, 0.0),  
(4188, 0.0),  
(4187, 0.0),  
(4186, 0.0),  
(4185, 0.024455799402225926),  
(4184, 0.0),  
(4183, 0.0),  
(4182, 0.03244428422615251),  
(4181, 0.0),  
(4180, 0.0),  
(4179, 0.04120428217151646),  
(4178, 0.06952346619889824),  
(4177, 0.02271554252121273),  
(4176, 0.04087595596566438),  
(4175, 0.0),  
(4174, 0.023174488732966077),  
(4173, 0.0),  
(4172, 0.0),  
(4171, 0.0),  
(4170, 0.0),  
(4169, 0.0),  
(4168, 0.0),  
(4167, 0.04120428217151646),  
(4166, 0.024738534799764674),  
(4165, 0.0),  
(4164, 0.0602475233128778),  
(4163, 0.07647191129018727),  
(4162, 0.0),  
(4161, 0.021677749238103002),  
(4160, 0.0),  
(4159, 0.03382550457458692),  
(4158, 0.044151078568834795),  
(4157, 0.0),  
(4156, 0.0),  
(4155, 0.0),

(4154, 0.026315789473684213),  
(4153, 0.0),  
(4152, 0.03993615319154359),  
(4151, 0.0),  
(4150, 0.0),  
(4149, 0.04204138910034079),  
(4148, 0.0),  
(4147, 0.0),  
(4146, 0.04891159880445185),  
(4145, 0.0),  
(4144, 0.03458572319330373),  
(4143, 0.0),  
(4142, 0.0),  
(4141, 0.02782074420373286),  
(4140, 0.0),  
(4139, 0.028676966733820218),  
(4138, 0.0),  
(4137, 0.040555355282690636),  
(4136, 0.04024218182927669),  
(4135, 0.03382550457458692),  
(4134, 0.0),  
(4133, 0.10619884881071831),  
(4132, 0.052631578947368425),  
(4131, 0.0),  
(4130, 0.03458572319330373),  
(4129, 0.03721614637823934),  
(4128, 0.03823595564509363),  
(4127, 0.014870813293635054),  
(4126, 0.0),  
(4125, 0.02666903735313325),  
(4124, 0.0),  
(4123, 0.05564148840746572),  
(4122, 0.025031308716087945),  
(4121, 0.06814662756363819),  
(4120, 0.0),  
(4119, 0.0),  
(4118, 0.05990422978731538),  
(4117, 0.0301237616564389),  
(4116, 0.0),  
(4115, 0.04543108504242546),  
(4114, 0.02878964680142297),  
(4113, 0.02666903735313325),  
(4112, 0.029617443887954616),  
(4111, 0.02913582733740325),  
(4110, 0.05923488777590923),  
(4109, 0.0),  
(4108, 0.02742042485535409),  
(4107, 0.027036903521793755),  
(4106, 0.0),  
(4105, 0.03065696697424829),  
(4104, 0.0),  
(4103, 0.02043797798283219),  
(4102, 0.0),  
(4101, 0.037463432463267755),  
(4100, 0.0),  
(4099, 0.0),  
(4098, 0.055824219567359015),  
(4097, 0.05923488777590923),  
(4096, 0.0),  
(4095, 0.10259783520851541),  
(4094, 0.0),  
(4093, 0.0),  
(4092, 0.03244428422615251),  
(4091, 0.0),  
(4090, 0.0),  
(4089, 0.0301237616564389),  
(4088, 0.0533380747062665),  
(4087, 0.0),  
(4086, 0.0),  
(4085, 0.04947706959952935),  
(4084, 0.0),  
(4083, 0.028676966733820218),  
(4082, 0.04543108504242546),  
(4081, 0.07175473098524099),  
(4080, 0.0),  
(4079, 0.0),  
(4078, 0.04783648732349399),  
(4077, 0.13157894736842107),  
(4076, 0.026315789473684213),  
(4075, 0.05990422978731538),  
(4074, 0.039344473768231684),  
(4073, 0.04188539082916955),  
(4072, 0.0),

(4071, 0.0),  
(4070, 0.03382550457458692),  
(4069, 0.0),  
(4068, 0.0),  
(4067, 0.025649458802128853),  
(4066, 0.044151078568834795),  
(4065, 0.02294157338705618),  
(4064, 0.03721614637823934),  
(4063, 0.02666903735313325),  
(4062, 0.0),  
(4061, 0.023918243661746996),  
(4060, 0.04499212706658476),  
(4059, 0.0662266178532522),  
(4058, 0.0),  
(4057, 0.021873931962990357),  
(4056, 0.0),  
(4055, 0.0),  
(4054, 0.0),  
(4053, 0.05407380704358751),  
(4052, 0.0),  
(4051, 0.0),  
(4050, 0.043355498476206004),  
(4049, 0.0207703246198632),  
(4048, 0.0),  
(4047, 0.0),  
(4046, 0.2278389747471728),  
(4045, 0.025031308716087945),  
(4044, 0.0602475233128778),  
(4043, 0.07509392614826384),  
(4042, 0.03721614637823934),  
(4041, 0.019968076595771794),  
(4040, 0.05923488777590923),  
(4039, 0.0),  
(4038, 0.0),  
(4037, 0.028676966733820218),  
(4036, 0.03181423814878889),  
(4035, 0.028239124736245257),  
(4034, 0.046829290579084706),  
(4033, 0.0),  
(4032, 0.0),  
(4031, 0.0),  
(4030, 0.0),  
(4029, 0.0),  
(4028, 0.03627381250550058),  
(4027, 0.03244428422615251),  
(4026, 0.01391037210186643),  
(4025, 0.0),  
(4024, 0.043355498476206004),  
(4023, 0.03539961627023944),  
(4022, 0.016643566632465155),  
(4021, 0.029617443887954616),  
(4020, 0.019968076595771794),  
(4019, 0.021300716142115247),  
(4018, 0.0),  
(4017, 0.06554735253444621),  
(4016, 0.09933992677987831),  
(4015, 0.020121090914638345),  
(4014, 0.0),  
(4013, 0.0),  
(4012, 0.0),  
(4011, 0.021486752129677003),  
(4010, 0.0415406492397264),  
(4009, 0.07443229275647868),  
(4008, 0.15737789507292674),  
(4007, 0.02742042485535409),  
(4006, 0.031219527052723135),  
(4005, 0.0),  
(4004, 0.06362847629757779),  
(4003, 0.0),  
(4002, 0.0),  
(4001, 0.019529164171612674),  
(4000, 0.03498557142706598),  
(3999, 0.0),  
(3998, 0.11803342130469505),  
(3997, 0.18074256993863339),  
(3996, 0.0),  
(3995, 0.04223886030955117),  
(3994, 0.03627381250550058),  
(3993, 0.0582716546748065),  
(3992, 0.0662266178532522),  
(3991, 0.01813690625275029),  
(3990, 0.04588314677411236),  
(3989, 0.024738534799764674),

(3988, 0.02294157338705618),  
(3987, 0.0301237616564389),  
(3986, 0.04188539082916955),  
(3985, 0.029617443887954616),  
(3984, 0.0),  
(3983, 0.0),  
(3982, 0.0),  
(3981, 0.031219527052723135),  
(3980, 0.13530201829834768),  
(3979, 0.0),  
(3978, 0.0),  
(3977, 0.03181423814878889),  
(3976, 0.0837707816583391),  
(3975, 0.02742042485535409),  
(3974, 0.07254762501100116),  
(3973, 0.0),  
(3972, 0.0),  
(3971, 0.0),  
(3970, 0.03627381250550058),  
(3969, 0.03539961627023944),  
(3968, 0.0),  
(3967, 0.0),  
(3966, 0.031219527052723135),  
(3965, 0.0),  
(3964, 0.0),  
(3963, 0.0),  
(3962, 0.0),  
(3961, 0.019389168358237032),  
(3960, 0.03823595564509363),  
(3959, 0.01749278571353299),  
(3958, 0.0),  
(3957, 0.0),  
(3956, 0.0),  
(3955, 0.0),  
(3954, 0.02913582733740325),  
(3953, 0.03382550457458692),  
(3952, 0.10147651372376076),  
(3951, 0.02782074420373286),  
(3950, 0.0),  
(3949, 0.030254451040802585),  
(3948, 0.021486752129677003),  
(3947, 0.043355498476206004),  
(3946, 0.03797316245786213),  
(3945, 0.03582871819500093),  
(3944, 0.03181423814878889),  
(3943, 0.0),  
(3942, 0.0),  
(3941, 0.0),  
(3940, 0.0),  
(3939, 0.04188539082916955),  
(3938, 0.0),  
(3937, 0.0),  
(3936, 0.043355498476206004),  
(3935, 0.022075539284417398),  
(3934, 0.0),  
(3933, 0.0989541391990587),  
(3932, 0.024182541670333724),  
(3931, 0.0),  
(3930, 0.024455799402225926),  
(3929, 0.06362847629757779),  
(3928, 0.04456565178215865),  
(3927, 0.06488856845230502),  
(3926, 0.01860807318911967),  
(3925, 0.0),  
(3924, 0.08111071056538127),  
(3923, 0.023174488732966077),  
(3922, 0.03181423814878889),  
(3921, 0.051298917604257706),  
(3920, 0.0),  
(3919, 0.03498557142706598),  
(3918, 0.0),  
(3917, 0.05407380704358751),  
(3916, 0.03823595564509363),  
(3915, 0.02913582733740325),  
(3914, 0.027036903521793755),  
(3913, 0.0),  
(3912, 0.04188539082916955),  
(3911, 0.017806114244894068),  
(3910, 0.025031308716087945),  
(3909, 0.0),  
(3908, 0.06748819059987714),  
(3907, 0.0),  
(3906, 0.01981850088223545),

(3905, 0.06952346619889824),  
(3904, 0.027036903521793755),  
(3903, 0.0830812984794528),  
(3902, 0.022075539284417398),  
(3901, 0.0),  
(3900, 0.0),  
(3899, 0.06488856845230502),  
(3898, 0.03181423814878889),  
(3897, 0.0),  
(3896, 0.0),  
(3895, 0.030254451040802585),  
(3894, 0.03823595564509363),  
(3893, 0.0),  
(3892, 0.026315789473684213),  
(3891, 0.0),  
(3890, 0.03721614637823934),  
(3889, 0.06917144638660747),  
(3888, 0.0582716546748065),  
(3887, 0.0),  
(3886, 0.046829290579084706),  
(3885, 0.02913582733740325),  
(3884, 0.03382550457458692),  
(3883, 0.04499212706658476),  
(3882, 0.0),  
(3881, 0.03244428422615251),  
(3880, 0.0),  
(3879, 0.09086217008485092),  
(3878, 0.0),  
(3877, 0.06684847767323797),  
(3876, 0.024738534799764674),  
(3875, 0.0331133089266261),  
(3874, 0.0),  
(3873, 0.0),  
(3872, 0.10259783520851541),  
(3871, 0.019529164171612674),  
(3870, 0.021486752129677003),  
(3869, 0.0),  
(3868, 0.0),  
(3867, 0.021486752129677003),  
(3866, 0.0),  
(3865, 0.0),  
(3864, 0.02913582733740325),  
(3863, 0.0),  
(3862, 0.051298917604257706),  
(3861, 0.03627381250550058),  
(3860, 0.056573571480353566),  
(3859, 0.0),  
(3858, 0.05647824947249051),  
(3857, 0.0),  
(3856, 0.0),  
(3855, 0.03181423814878889),  
(3854, 0.0),  
(3853, 0.04836508334066745),  
(3852, 0.043355498476206004),  
(3851, 0.0),  
(3850, 0.025976216673306556),  
(3849, 0.08000711205939975),  
(3848, 0.0),  
(3847, 0.0),  
(3846, 0.043355498476206004),  
(3845, 0.027036903521793755),  
(3844, 0.0),  
(3843, 0.024455799402225926),  
(3842, 0.0),  
(3841, 0.02294157338705618),  
(3840, 0.03382550457458692),  
(3839, 0.0),  
(3838, 0.0),  
(3837, 0.1037571695799112),  
(3836, 0.04891159880445185),  
(3835, 0.0),  
(3834, 0.025649458802128853),  
(3833, 0.0),  
(3832, 0.06243905410544627),  
(3831, 0.0),  
(3830, 0.07868894753646337),  
(3829, 0.0),  
(3828, 0.0),  
(3827, 0.0),  
(3826, 0.0),  
(3825, 0.016821577756312225),  
(3824, 0.03458572319330373),  
(3823, 0.05006261743217589),

```
(3822, 0.0),
(3821, 0.0),
(3820, 0.065033247714309),
(3819, 0.0),
(3818, 0.03181423814878889),
(3817, 0.0301237616564389),
(3816, 0.0),
(3815, 0.0),
(3814, 0.08000711205939975),
(3813, 0.016643566632465155),
(3812, 0.0),
(3811, 0.0),
(3810, 0.0),
(3809, 0.02666903735313325),
(3808, 0.05757929360284594),
(3807, 0.03627381250550058),
(3806, 0.0),
...]
```

```
In [67]: sorted(list(enumerate(similarity[0])),reverse=True, key = lambda x:x[1])
```

```
Out[67]: [(0, 1.0000000000000002),
(1214, 0.28676966733820225),
(2405, 0.26901379342448517),
(3728, 0.2605130246476754),
(507, 0.255608593705383),
(539, 0.25038669783359574),
(582, 0.24511108480187255),
(1202, 0.24455799402225925),
(1192, 0.2367785320221084),
(61, 0.23179316248638276),
(778, 0.23174488732966073),
(4046, 0.2278389747471728),
(1916, 0.2252817784447915),
(2782, 0.21853668936906193),
(172, 0.21239769762143662),
(972, 0.2108663315950723),
(322, 0.2105263157894737),
(2329, 0.20443988269091456),
(3606, 0.20437977982832192),
(260, 0.20395079136182276),
(151, 0.2029530274475215),
(4190, 0.2029530274475215),
(1440, 0.20277677641345318),
(74, 0.2024645717996314),
(1087, 0.2020475485519274),
(3671, 0.1979082783981174),
(2878, 0.19767387315371682),
(577, 0.1976738731537168),
(47, 0.19672236884115843),
(2969, 0.1925214071641298),
(942, 0.19134594929397597),
(495, 0.19088542889273336),
(1199, 0.19088542889273336),
(305, 0.19007487139298027),
(4, 0.1892994097121204),
(4405, 0.18731716231633883),
(94, 0.1873171623163388),
(27, 0.18693292157876878),
(3325, 0.1860807318911967),
(1273, 0.18394180184548975),
(1327, 0.18394180184548975),
(2997, 0.1813690625275029),
(3997, 0.18074256993863339),
(83, 0.17996850826633903),
(1069, 0.17954621161490197),
(220, 0.1777046633277277),
(184, 0.177343107178349),
(1319, 0.17699808135119718),
(168, 0.17521916101261562),
(2200, 0.17521916101261562),
(4336, 0.17521916101261562),
(931, 0.1748149640244195),
(36, 0.17316974359835272),
(373, 0.17292861596651865),
(300, 0.1711905958155815),
(3536, 0.1711905958155815),
(3160, 0.17040572913692198),
(1978, 0.16943474841747155),
(661, 0.1675415633166782),
(529, 0.16742770563222897),
(1827, 0.16742770563222897),
(91, 0.16692446522239718),
```

(41, 0.16563698349810535),  
(67, 0.16452254913212452),  
(150, 0.16426202236672155),  
(1533, 0.16222142113076254),  
(342, 0.16122923187750418),  
(595, 0.16059101370939324),  
(31, 0.15907119074394446),  
(2511, 0.15907119074394446),  
(1176, 0.15789473684210528),  
(4008, 0.15737789507292674),  
(2727, 0.15609763526361567),  
(149, 0.15511334686589626),  
(838, 0.1540171257313038),  
(26, 0.15389675281277312),  
(187, 0.15389675281277312),  
(1379, 0.152008377581442),  
(111, 0.15018785229652767),  
(1342, 0.15018785229652767),  
(311, 0.15018785229652765),  
(52, 0.14843120879858804),  
(513, 0.14843120879858804),  
(2071, 0.14808721943977307),  
(7, 0.14673479641335554),  
(2756, 0.14659886790209342),  
(127, 0.14601068654819693),  
(579, 0.14567913668701626),  
(581, 0.14567913668701626),  
(108, 0.14509525002200235),  
(1295, 0.14509525002200235),  
(466, 0.14509525002200233),  
(1392, 0.14421498760030763),  
(1475, 0.1442149876003076),  
(1353, 0.14350946197048198),  
(1934, 0.14350946197048198),  
(275, 0.1433848336691011),  
(294, 0.1433848336691011),  
(3352, 0.1433848336691011),  
(1054, 0.1419745572840903),  
(1271, 0.1419745572840903),  
(490, 0.14159846508095775),  
(1484, 0.14159846508095775),  
(413, 0.1411956236812263),  
(461, 0.1411956236812263),  
(568, 0.1411956236812263),  
(2492, 0.1411956236812263),  
(95, 0.14048787173725413),  
(17, 0.13977653617040256),  
(46, 0.1391037210186643),  
(3407, 0.1391037210186643),  
(1961, 0.13872950617564817),  
(56, 0.13834289277321493),  
(249, 0.13834289277321493),  
(1325, 0.13834289277321493),  
(3779, 0.13834289277321493),  
(72, 0.13710212427677043),  
(843, 0.13710212427677043),  
(859, 0.13710212427677043),  
(2165, 0.13710212427677043),  
(4538, 0.13629325512727639),  
(476, 0.13530201829834768),  
(505, 0.13530201829834768),  
(668, 0.13530201829834768),  
(935, 0.13530201829834768),  
(3980, 0.13530201829834768),  
(1804, 0.13518451760896877),  
(4348, 0.13497638119975428),  
(1320, 0.13476498501489081),  
(158, 0.13369695534647594),  
(502, 0.13369695534647594),  
(777, 0.13369695534647594),  
(126, 0.13334518676566626),  
(1556, 0.13334518676566626),  
(1369, 0.1324532357065044),  
(2374, 0.1324532357065044),  
(2580, 0.1324532357065044),  
(4077, 0.13157894736842107),  
(2129, 0.13124359177794215),  
(1013, 0.12988108336653278),  
(2177, 0.12988108336653278),  
(3620, 0.12988108336653278),  
(134, 0.12977713690461004),  
(368, 0.12977713690461004),  
(1770, 0.12977713690461004),



(2869, 0.12892051277806202),  
(14, 0.12824729401064427),  
(138, 0.12824729401064427),  
(369, 0.12824729401064427),  
(1084, 0.12824729401064427),  
(2682, 0.12824729401064427),  
(4228, 0.12824729401064427),  
(487, 0.12725695259515557),  
(773, 0.12725695259515557),  
(2968, 0.12725695259515557),  
(937, 0.126673647984535),  
(2476, 0.12565617248750865),  
(1737, 0.12565617248750863),  
(166, 0.12515654358043973),  
(754, 0.12515654358043973),  
(2446, 0.12515654358043973),  
(3626, 0.12515654358043973),  
(483, 0.12487810821089254),  
(1094, 0.12487810821089254),  
(1156, 0.12487810821089254),  
(2648, 0.12487810821089254),  
(102, 0.12462194771917919),  
(84, 0.12361284651454937),  
(208, 0.12262786789699316),  
(3374, 0.12262786789699316),  
(1874, 0.12262786789699313),  
(438, 0.12227899701112963),  
(557, 0.12227899701112963),  
(1297, 0.12227899701112963),  
(2327, 0.12166606584807191),  
(2670, 0.12166606584807191),  
(3081, 0.12166606584807191),  
(139, 0.12072654548783007),  
(68, 0.1204950466257556),  
(781, 0.1204950466257556),  
(1002, 0.1204950466257556),  
(1272, 0.1204950466257556),  
(1654, 0.1204950466257556),  
(1000, 0.11980845957463077),  
(549, 0.11959121830873498),  
(870, 0.11959121830873498),  
(900, 0.11959121830873498),  
(3233, 0.11959121830873498),  
(1045, 0.11846977555181847),  
(1829, 0.11846977555181847),  
(2056, 0.11846977555181847),  
(2737, 0.11846977555181847),  
(4673, 0.11846977555181847),  
(3719, 0.11803342130469506),  
(1669, 0.11803342130469505),  
(1711, 0.11803342130469505),  
(2785, 0.11803342130469505),  
(3998, 0.11803342130469505),  
(160, 0.11712303424679457),  
(182, 0.11707322644771176),  
(278, 0.11707322644771176),  
(454, 0.11707322644771176),  
(1613, 0.11707322644771176),  
(2280, 0.11707322644771176),  
(28, 0.116543309349613),  
(122, 0.116543309349613),  
(422, 0.116543309349613),  
(678, 0.116543309349613),  
(589, 0.1163350101494222),  
(2543, 0.11587244366483038),  
(1493, 0.11551284429847784),  
(3210, 0.11506355109724453),  
(189, 0.1147078669352809),  
(1218, 0.1147078669352809),  
(236, 0.11470786693528087),  
(1299, 0.11470786693528087),  
(85, 0.1141270638770543),  
(432, 0.11357771260606365),  
(1414, 0.11357771260606365),  
(50, 0.11295649894498103),  
(1612, 0.11295649894498103),  
(123, 0.1124803176664619),  
(420, 0.1124803176664619),  
(536, 0.1124803176664619),  
(889, 0.11243619908556624),  
(3145, 0.11239029738980327),  
(256, 0.11164843913471803),  
(808, 0.11164843913471803),

(1346, 0.11164843913471803),  
(1571, 0.11164843913471803),  
(2368, 0.11141412945539662),  
(3307, 0.11141412945539662),  
(1508, 0.11141412945539661),  
(87, 0.11128297681493143),  
(121, 0.11128297681493143),  
(232, 0.11128297681493143),  
(4468, 0.11128297681493143),  
(206, 0.11092107999698035),  
(238, 0.11092107999698035),  
(228, 0.11037769642208699),  
(81, 0.11020775375559676),  
(1570, 0.11020775375559674),  
(3426, 0.10968169942141637),  
(21, 0.10968169942141635),  
(449, 0.10968169942141635),  
(1029, 0.10968169942141635),  
(1941, 0.10968169942141635),  
(2144, 0.10968169942141635),  
(3006, 0.10968169942141635),  
(692, 0.10936965981495178),  
(200, 0.10882143751650175),  
(627, 0.10882143751650175),  
(653, 0.10882143751650175),  
(1520, 0.10882143751650175),  
(307, 0.10876595882948199),  
(5, 0.10838874619051501),  
(740, 0.10838874619051501),  
(3091, 0.10838874619051501),  
(3628, 0.10838874619051501),  
(786, 0.10814761408717502),  
(3496, 0.10814761408717502),  
(4777, 0.10814761408717502),  
(229, 0.1074861545850028),  
(572, 0.10743376064838502),  
(939, 0.1068366854693644),  
(332, 0.106676149412533),  
(472, 0.106676149412533),  
(1235, 0.106676149412533),  
(170, 0.10650358071057624),  
(262, 0.10650358071057624),  
(1164, 0.10619884881071831),  
(1497, 0.10619884881071831),  
(1741, 0.10619884881071831),  
(2157, 0.10619884881071831),  
(2630, 0.10619884881071831),  
(4133, 0.10619884881071831),  
(467, 0.10559715077387793),  
(541, 0.10559715077387793),  
(1492, 0.10559715077387793),  
(131, 0.10526315789473685),  
(295, 0.10526315789473685),  
(2068, 0.10526315789473685),  
(3338, 0.10526315789473685),  
(2494, 0.10471347707292386),  
(75, 0.10390486669322622),  
(223, 0.10390486669322622),  
(687, 0.10390486669322622),  
(814, 0.10390486669322622),  
(1193, 0.10390486669322622),  
(1767, 0.10390486669322622),  
(2813, 0.10390486669322622),  
(654, 0.10385162309931599),  
(2435, 0.10385162309931599),  
(105, 0.1037571695799112),  
(233, 0.1037571695799112),  
(624, 0.1037571695799112),  
(2319, 0.1037571695799112),  
(2452, 0.1037571695799112),  
(2517, 0.1037571695799112),  
(3053, 0.1037571695799112),  
(3619, 0.1037571695799112),  
(3837, 0.1037571695799112),  
(1451, 0.10301070542879115),  
(10, 0.10259783520851541),  
(1080, 0.10259783520851541),  
(3150, 0.10259783520851541),  
(3632, 0.10259783520851541),  
(3753, 0.10259783520851541),  
(3872, 0.10259783520851541),  
(4095, 0.10259783520851541),  
(2102, 0.10218988991416095),

(2758, 0.10147651372376076),  
(3659, 0.10147651372376076),  
(3952, 0.10147651372376076),  
(3439, 0.10138838820672659),  
(508, 0.101338918387628),  
(1292, 0.101338918387628),  
(2243, 0.10133891838762799),  
(1401, 0.10060545457319173),  
(1435, 0.10060545457319173),  
(2095, 0.10060545457319173),  
(312, 0.10012523486435178),  
(330, 0.10012523486435178),  
(643, 0.10012523486435178),  
(986, 0.10012523486435178),  
(1035, 0.10012523486435178),  
(1610, 0.10012523486435178),  
(35, 0.09933992677987831),  
(157, 0.09933992677987831),  
(319, 0.09933992677987831),  
(325, 0.09933992677987831),  
(715, 0.09933992677987831),  
(1075, 0.09933992677987831),  
(1278, 0.09933992677987831),  
(1880, 0.09933992677987831),  
(2049, 0.09933992677987831),  
(4016, 0.09933992677987831),  
(4196, 0.09933992677987831),  
(4490, 0.09933992677987831),  
(363, 0.09909250441117726),  
(491, 0.09909250441117726),  
(266, 0.0989541391990587),  
(445, 0.0989541391990587),  
(510, 0.0989541391990587),  
(3618, 0.0989541391990587),  
(3781, 0.0989541391990587),  
(3933, 0.0989541391990587),  
(1771, 0.09836118442057921),  
(410, 0.0978231976089037),  
(480, 0.0978231976089037),  
(535, 0.0978231976089037),  
(898, 0.0978231976089037),  
(1107, 0.0978231976089037),  
(1397, 0.0978231976089037),  
(1819, 0.0978231976089037),  
(2249, 0.0978231976089037),  
(33, 0.09733285267845754),  
(148, 0.09733285267845754),  
(263, 0.09733285267845754),  
(878, 0.09733285267845754),  
(1154, 0.09733285267845754),  
(1723, 0.09733285267845754),  
(1938, 0.09733285267845754),  
(2783, 0.09733285267845754),  
(4254, 0.09733285267845754),  
(4403, 0.09733285267845754),  
(4472, 0.09733285267845754),  
(9, 0.0967301666813349),  
(334, 0.0967301666813349),  
(371, 0.09567297464698798),  
(412, 0.09567297464698798),  
(2237, 0.09567297464698798),  
(2552, 0.09567297464698798),  
(4545, 0.09558988911273408),  
(362, 0.09544271444636668),  
(791, 0.09544271444636668),  
(1569, 0.09544271444636668),  
(2558, 0.09544271444636668),  
(4201, 0.09544271444636668),  
(18, 0.09493290614465533),  
(11, 0.09464970485606021),  
(587, 0.09464970485606021),  
(660, 0.09365858115816941),  
(920, 0.09365858115816941),  
(978, 0.09365858115816941),  
(1274, 0.09365858115816941),  
(1355, 0.09365858115816941),  
(1470, 0.09365858115816941),  
(1682, 0.09365858115816941),  
(1688, 0.09365858115816941),  
(1717, 0.09365858115816941),  
(2392, 0.09365858115816941),  
(2484, 0.09365858115816941),  
(2588, 0.09365858115816941),

(2641, 0.09365858115816941),  
(2646, 0.09365858115816941),  
(2816, 0.09365858115816941),  
(3535, 0.09365858115816941),  
(3718, 0.09365858115816941),  
(4331, 0.09365858115816941),  
(4502, 0.09365858115816941),  
(301, 0.09365858115816939),  
(129, 0.09269795493186431),  
(243, 0.09269795493186431),  
(279, 0.09269795493186431),  
(473, 0.09269795493186431),  
(596, 0.09269795493186431),  
(813, 0.09269795493186431),  
(1442, 0.09269795493186431),  
(1697, 0.09269795493186431),  
(3405, 0.09269795493186431),  
(785, 0.09243423333081695),  
(219, 0.09197090092274487),  
(244, 0.09197090092274487),  
(511, 0.09197090092274487),  
(1752, 0.09197090092274487),  
(2089, 0.09197090092274487),  
(2647, 0.09197090092274487),  
(3734, 0.09197090092274487),  
(230, 0.09183979479633063),  
(799, 0.09183979479633063),  
(239, 0.09125667909262308),  
(2566, 0.09125667909262308),  
(43, 0.09086217008485092),  
(602, 0.09086217008485092),  
(995, 0.09086217008485092),  
(1231, 0.09086217008485092),  
(1842, 0.09086217008485092),  
(3879, 0.09086217008485092),  
(12, 0.09037128496931669),  
(88, 0.09037128496931669),  
(167, 0.09037128496931669),  
(356, 0.09037128496931669),  
(400, 0.09037128496931669),  
(1304, 0.09037128496931669),  
(1476, 0.09037128496931669),  
(2247, 0.09037128496931669),  
(2955, 0.09037128496931669),  
(331, 0.08998425413316952),  
(2359, 0.08998425413316952),  
(2792, 0.08998425413316952),  
(2852, 0.08998425413316952),  
(3387, 0.08998425413316952),  
(3441, 0.08998425413316952),  
(3460, 0.08998425413316952),  
(4343, 0.08998425413316952),  
(4353, 0.08998425413316952),  
(274, 0.08977310580745099),  
(1902, 0.0891313035643173),  
(2111, 0.08903057122447033),  
(292, 0.08885233166386385),  
(564, 0.08885233166386385),  
(1025, 0.08885233166386385),  
(2656, 0.08885233166386385),  
(2864, 0.08885233166386385),  
(3320, 0.08885233166386385),  
(169, 0.08830215713766959),  
(270, 0.08830215713766959),  
(618, 0.08830215713766959),  
(1779, 0.08830215713766959),  
(2105, 0.08830215713766959),  
(2709, 0.08830215713766959),  
(2736, 0.08830215713766959),  
(3551, 0.08830215713766959),  
(479, 0.08797691788472335),  
(752, 0.08749572785196143),  
(1828, 0.08749572785196143),  
(2444, 0.08746392856766495),  
(23, 0.08740748201220976),  
(141, 0.08740748201220976),  
(389, 0.08740748201220976),  
(745, 0.08740748201220976),  
(794, 0.08740748201220976),  
(960, 0.08740748201220976),  
(1247, 0.08740748201220976),  
(1257, 0.08740748201220976),  
(3446, 0.08740748201220976),

(3447, 0.08740748201220976),  
(143, 0.08671099695241201),  
(199, 0.08671099695241201),  
(827, 0.08671099695241201),  
(1526, 0.08671099695241201),  
(2876, 0.08671099695241201),  
(3554, 0.08671099695241201),  
(4516, 0.08671099695241201),  
(4587, 0.08671099695241201),  
(2, 0.08603090020146065),  
(457, 0.08603090020146065),  
(775, 0.08603090020146065),  
(1374, 0.08603090020146065),  
(2579, 0.08603090020146065),  
(3133, 0.08603090020146065),  
(4754, 0.08603090020146065),  
(234, 0.08594700851870801),  
(1017, 0.08594700851870801),  
(2350, 0.08594700851870801),  
(32, 0.08520286456846099),  
(217, 0.08520286456846099),  
(142, 0.08471737420873576),  
(894, 0.08471737420873576),  
(1095, 0.08471737420873576),  
(1140, 0.08471737420873576),  
(1585, 0.08471737420873576),  
(2300, 0.08471737420873576),  
(4560, 0.08471737420873576),  
(48, 0.08447772061910234),  
(190, 0.08447772061910234),  
(1303, 0.08447772061910234),  
(1635, 0.08447772061910234),  
(2966, 0.08447772061910234),  
(62, 0.0837707816583391),  
(381, 0.0837707816583391),  
(534, 0.0837707816583391),  
(1694, 0.0837707816583391),  
(3178, 0.0837707816583391),  
(3216, 0.0837707816583391),  
(3976, 0.0837707816583391),  
(4737, 0.0837707816583391),  
(1, 0.08346223261119858),  
(152, 0.08346223261119858),  
(203, 0.08346223261119858),  
(315, 0.08346223261119858),  
(547, 0.08346223261119858),  
(1684, 0.08346223261119858),  
(2124, 0.08346223261119858),  
(2382, 0.08321783316232578),  
(19, 0.0830812984794528),  
(403, 0.0830812984794528),  
(3903, 0.0830812984794528),  
(45, 0.08240856434303293),  
(2199, 0.08240856434303293),  
(39, 0.08226127456606226),  
(227, 0.08226127456606226),  
(341, 0.08226127456606226),  
(626, 0.08226127456606226),  
(787, 0.08226127456606226),  
(1205, 0.08226127456606226),  
(1745, 0.08226127456606226),  
(2163, 0.08226127456606226),  
(2657, 0.08226127456606226),  
(3138, 0.08226127456606226),  
(3299, 0.08226127456606226),  
(3468, 0.08226127456606226),  
(63, 0.08175191193132876),  
(501, 0.08175191193132876),  
(38, 0.08111071056538127),  
(242, 0.08111071056538127),  
(245, 0.08111071056538127),  
(306, 0.08111071056538127),  
(584, 0.08111071056538127),  
(812, 0.08111071056538127),  
(866, 0.08111071056538127),  
(901, 0.08111071056538127),  
(1267, 0.08111071056538127),  
(1515, 0.08111071056538127),  
(1660, 0.08111071056538127),  
(2335, 0.08111071056538127),  
(2640, 0.08111071056538127),  
(2667, 0.08111071056538127),  
(2861, 0.08111071056538127),

(2873, 0.08111071056538127),  
(3525, 0.08111071056538127),  
(3924, 0.08111071056538127),  
(4492, 0.08111071056538127),  
(4568, 0.08111071056538127),  
(4626, 0.08111071056538127),  
(4641, 0.08111071056538127),  
(4654, 0.08111071056538127),  
(4657, 0.08111071056538127),  
(51, 0.08048436365855338),  
(164, 0.08000711205939975),  
(854, 0.08000711205939975),  
(899, 0.08000711205939975),  
(1323, 0.08000711205939975),  
(1743, 0.08000711205939975),  
(2679, 0.08000711205939975),  
(3762, 0.08000711205939975),  
(3814, 0.08000711205939975),  
(3849, 0.08000711205939975),  
(486, 0.07992075789278004),  
(103, 0.07987230638308718),  
(222, 0.07987230638308718),  
(1530, 0.07987230638308718),  
(2788, 0.07987230638308718),  
(4506, 0.07987230638308718),  
(1894, 0.0792740035289418),  
(16, 0.07894736842105264),  
(37, 0.07894736842105264),  
(241, 0.07894736842105264),  
(329, 0.07894736842105264),  
(1078, 0.07894736842105264),  
(1159, 0.07894736842105264),  
(2291, 0.07894736842105264),  
(3231, 0.07894736842105264),  
(4316, 0.07894736842105264),  
(193, 0.07878168645790681),  
(153, 0.07868894753646337),  
(776, 0.07868894753646337),  
(1043, 0.07868894753646337),  
(1824, 0.07868894753646337),  
(1922, 0.07868894753646337),  
(3726, 0.07868894753646337),  
(3830, 0.07868894753646337),  
(4279, 0.07868894753646337),  
(4473, 0.07868894753646337),  
(125, 0.07843304784240032),  
(286, 0.0781166566864507),  
(1655, 0.0781166566864507),  
(174, 0.07792865001991967),  
(402, 0.07792865001991967),  
(922, 0.07792865001991967),  
(1111, 0.07792865001991967),  
(2936, 0.07792865001991967),  
(3026, 0.07792865001991967),  
(3089, 0.07792865001991967),  
(3100, 0.07792865001991967),  
(3162, 0.07792865001991967),  
(4341, 0.07792865001991967),  
(1920, 0.07755667343294813),  
(3558, 0.07755667343294813),  
(1280, 0.0770085628656519),  
(40, 0.07694837640638656),  
(96, 0.07694837640638656),  
(101, 0.07694837640638656),  
(112, 0.07694837640638656),  
(207, 0.07694837640638656),  
(632, 0.07694837640638656),  
(737, 0.07694837640638656),  
(1287, 0.07694837640638656),  
(2215, 0.07694837640638656),  
(4266, 0.07694837640638656),  
(4624, 0.07694837640638656),  
(850, 0.07647191129018727),  
(1003, 0.07647191129018727),  
(1137, 0.07647191129018727),  
(1252, 0.07647191129018727),  
(1617, 0.07647191129018727),  
(2161, 0.07647191129018727),  
(2270, 0.07647191129018727),  
(2708, 0.07647191129018727),  
(3428, 0.07647191129018727),  
(4163, 0.07647191129018727),  
(70, 0.076004188790721),

(82, 0.076004188790721),  
(235, 0.076004188790721),  
(349, 0.076004188790721),  
(492, 0.076004188790721),  
(1288, 0.076004188790721),  
(1519, 0.076004188790721),  
(2445, 0.076004188790721),  
(2914, 0.076004188790721),  
(4558, 0.076004188790721),  
(246, 0.07594632491572426),  
(2001, 0.07543142864047142),  
(3329, 0.07530940414109724),  
(98, 0.07509392614826384),  
(210, 0.07509392614826384),  
(361, 0.07509392614826384),  
(365, 0.07509392614826384),  
(503, 0.07509392614826384),  
(540, 0.07509392614826384),  
(680, 0.07509392614826384),  
(1007, 0.07509392614826384),  
(1291, 0.07509392614826384),  
(2159, 0.07509392614826384),  
(4043, 0.07509392614826384),  
(4203, 0.07509392614826384),  
(2252, 0.07492686492653551),  
(205, 0.07443229275647868),  
(800, 0.07443229275647868),  
(993, 0.07443229275647868),  
(1321, 0.07443229275647868),  
(1517, 0.07443229275647868),  
(1766, 0.07443229275647868),  
(2322, 0.07443229275647868),  
(2547, 0.07443229275647868),  
(2797, 0.07443229275647868),  
(3067, 0.07443229275647868),  
(3115, 0.07443229275647868),  
(3124, 0.07443229275647868),  
(4009, 0.07443229275647868),  
(4412, 0.07443229275647868),  
(4542, 0.07443229275647868),  
(634, 0.07421560439929402),  
(885, 0.07421560439929402),  
(1068, 0.07421560439929402),  
(2391, 0.07421560439929402),  
(3699, 0.07421560439929402),  
(1986, 0.07394738666465357),  
(2186, 0.07394738666465357),  
(3429, 0.07394738666465357),  
(3, 0.0734718358370645),  
(1483, 0.0734718358370645),  
(71, 0.07336739820667779),  
(908, 0.07336739820667779),  
(1619, 0.07336739820667779),  
(2017, 0.07336739820667779),  
(3020, 0.07336739820667779),  
(3283, 0.07336739820667779),  
(3803, 0.07336739820667779),  
(1363, 0.07336739820667777),  
(2052, 0.07336739820667777),  
(728, 0.07254762501100118),  
(833, 0.07254762501100118),  
(977, 0.07254762501100118),  
(1761, 0.07254762501100118),  
(1992, 0.07254762501100118),  
(2986, 0.07254762501100118),  
(3253, 0.07254762501100118),  
(3353, 0.07254762501100118),  
(4640, 0.07254762501100118),  
(59, 0.07254762501100116),  
(179, 0.07254762501100116),  
(1485, 0.07254762501100116),  
(1651, 0.07254762501100116),  
(2110, 0.07254762501100116),  
(2208, 0.07254762501100116),  
(2293, 0.07254762501100116),  
(2313, 0.07254762501100116),  
(2377, 0.07254762501100116),  
(2475, 0.07254762501100116),  
(2768, 0.07254762501100116),  
(3070, 0.07254762501100116),  
(3101, 0.07254762501100116),  
(3153, 0.07254762501100116),  
(3481, 0.07254762501100116),

(3714, 0.07254762501100116),  
(3974, 0.07254762501100116),  
(4431, 0.07254762501100116),  
(4443, 0.07254762501100116),  
(790, 0.07209840939145001),  
(1955, 0.07209840939145001),  
(76, 0.07175473098524099),  
(784, 0.07175473098524099),  
(932, 0.07175473098524099),  
(933, 0.07175473098524099),  
(1166, 0.07175473098524099),  
(1879, 0.07175473098524099),  
(2697, 0.07175473098524099),  
(2848, 0.07175473098524099),  
(2901, 0.07175473098524099),  
(3303, 0.07175473098524099),  
(3421, 0.07175473098524099),  
(3462, 0.07175473098524099),  
(4081, 0.07175473098524099),  
(86, 0.07098727864204515),  
(93, 0.07098727864204515),  
(175, 0.07098727864204515),  
(271, 0.07098727864204515),  
(543, 0.07098727864204515),  
(674, 0.07098727864204515),  
(1143, 0.07098727864204515),  
(1338, 0.07098727864204515),  
(2827, 0.07098727864204515),  
(3204, 0.07098727864204515),  
(3609, 0.07098727864204515),  
(4483, 0.07098727864204515),  
(1238, 0.07079923254047887),  
(1406, 0.07079923254047887),  
(2441, 0.07079923254047887),  
(2759, 0.07079923254047887),  
(2850, 0.07079923254047887),  
(3221, 0.07079923254047887),  
(4638, 0.07079923254047887),  
(55, 0.07038153430777869),  
(183, 0.07038153430777869),  
(1628, 0.07024393586862707),  
(1843, 0.07024393586862707),  
(3483, 0.07024393586862707),  
(97, 0.06952346619889824),  
(254, 0.06952346619889824),  
(302, 0.06952346619889824),  
(782, 0.06952346619889824),  
(1144, 0.06952346619889824),  
(1453, 0.06952346619889824),  
(2099, 0.06952346619889824),  
(3298, 0.06952346619889824),  
(3905, 0.06952346619889824),  
(4178, 0.06952346619889824),  
(336, 0.06917144638660747),  
(1733, 0.06917144638660747),  
(2529, 0.06917144638660747),  
(3889, 0.06917144638660747),  
(4290, 0.06917144638660747),  
(419, 0.06882472016116853),  
(600, 0.06882472016116853),  
(1276, 0.06882472016116853),  
(1998, 0.06882472016116853),  
(2014, 0.06882472016116853),  
(2205, 0.06882472016116853),  
(2239, 0.06882472016116853),  
(665, 0.06814662756363819),  
(671, 0.06814662756363819),  
(704, 0.06814662756363819),  
(732, 0.06814662756363819),  
(1534, 0.06814662756363819),  
(2031, 0.06814662756363819),  
(3018, 0.06814662756363819),  
(3308, 0.06814662756363819),  
(3586, 0.06814662756363819),  
(4121, 0.06814662756363819),  
(4392, 0.06814662756363819),  
(100, 0.06765100914917384),  
(209, 0.06765100914917384),  
(308, 0.06765100914917384),  
(498, 0.06765100914917384),  
(537, 0.06765100914917384),  
(570, 0.06765100914917384),  
(1290, 0.06765100914917384),



(1462, 0.06765100914917384),  
(1856, 0.06765100914917384),  
(2214, 0.06765100914917384),  
(2431, 0.06765100914917384),  
(2618, 0.06765100914917384),  
(2660, 0.06765100914917384),  
(2739, 0.06765100914917384),  
(3396, 0.06765100914917384),  
(3674, 0.06765100914917384),  
(4214, 0.06765100914917384),  
(58, 0.06748819059987714),  
(124, 0.06748819059987714),  
(1441, 0.06748819059987714),  
(2301, 0.06748819059987714),  
(3223, 0.06748819059987714),  
(3908, 0.06748819059987714),  
(1630, 0.0672863110252489),  
(30, 0.06684847767323797),  
(734, 0.06684847767323797),  
(735, 0.06684847767323797),  
(3877, 0.06684847767323797),  
(488, 0.06657426652986062),  
(22, 0.0662266178532522),  
(237, 0.0662266178532522),  
(604, 0.0662266178532522),  
(695, 0.0662266178532522),  
(1083, 0.0662266178532522),  
(1169, 0.0662266178532522),  
(1377, 0.0662266178532522),  
(2246, 0.0662266178532522),  
(2255, 0.0662266178532522),  
(2888, 0.0662266178532522),  
(2974, 0.0662266178532522),  
(3239, 0.0662266178532522),  
(3410, 0.0662266178532522),  
(3992, 0.0662266178532522),  
(4059, 0.0662266178532522),  
(4351, 0.0662266178532522),  
(212, 0.06622661785325219),  
(215, 0.06622661785325219),  
(635, 0.06622661785325219),  
(1016, 0.06622661785325219),  
(1085, 0.06622661785325219),  
(1119, 0.06622661785325219),  
(1428, 0.06622661785325219),  
(2298, 0.06622661785325219),  
(2316, 0.06622661785325219),  
(2653, 0.06622661785325219),  
(304, 0.06578947368421054),  
(1173, 0.06562179588897107),  
(1999, 0.06562179588897107),  
(2131, 0.06562179588897107),  
(2400, 0.06562179588897107),  
(2592, 0.06562179588897107),  
(3388, 0.06562179588897107),  
(3633, 0.06562179588897107),  
(4017, 0.06554735253444621),  
(2828, 0.06521546507260247),  
(29, 0.065033247714309),  
(281, 0.065033247714309),  
(1527, 0.065033247714309),  
(2048, 0.065033247714309),  
(3820, 0.065033247714309),  
(109, 0.06488856845230502),  
(216, 0.06488856845230502),  
(267, 0.06488856845230502),  
(621, 0.06488856845230502),  
(744, 0.06488856845230502),  
(1066, 0.06488856845230502),  
(1074, 0.06488856845230502),  
(1434, 0.06488856845230502),  
(1482, 0.06488856845230502),  
(2264, 0.06488856845230502),  
(2266, 0.06488856845230502),  
(2305, 0.06488856845230502),  
(2575, 0.06488856845230502),  
(2711, 0.06488856845230502),  
(2880, 0.06488856845230502),  
(3179, 0.06488856845230502),  
(3372, 0.06488856845230502),  
(3570, 0.06488856845230502),  
(3899, 0.06488856845230502),  
(3927, 0.06488856845230502),

(4623, 0.06488856845230502),  
(447, 0.06446025638903101),  
(746, 0.06446025638903101),  
(1009, 0.06446025638903101),  
(1844, 0.06446025638903101),  
(3172, 0.06446025638903101),  
(2515, 0.06424925662032356),  
(425, 0.06390214842634574),  
(443, 0.06390214842634574),  
(471, 0.06390214842634574),  
(1975, 0.06390214842634574),  
(2260, 0.06390214842634574),  
(2523, 0.06390214842634574),  
(159, 0.06362847629757779),  
(290, 0.06362847629757779),  
(546, 0.06362847629757779),  
(951, 0.06362847629757779),  
(1260, 0.06362847629757779),  
(1302, 0.06362847629757779),  
(1703, 0.06362847629757779),  
(1850, 0.06362847629757779),  
(1946, 0.06362847629757779),  
(2109, 0.06362847629757779),  
(2224, 0.06362847629757779),  
(3116, 0.06362847629757779),  
(3131, 0.06362847629757779),  
(3251, 0.06362847629757779),  
(3929, 0.06362847629757779),  
(4004, 0.06362847629757779),  
(4271, 0.06362847629757779),  
(4312, 0.06362847629757779),  
(4355, 0.06362847629757779),  
(4704, 0.06362847629757779),  
(612, 0.06335829046432676),  
(1091, 0.06335829046432676),  
(1458, 0.06335829046432676),  
(3342, 0.06335829046432676),  
(3738, 0.06335829046432676),  
(15, 0.06282808624375433),  
(253, 0.06282808624375433),  
(591, 0.06282808624375433),  
(2315, 0.06282808624375433),  
(3296, 0.06282808624375433),  
(3605, 0.06282808624375433),  
(276, 0.06243905410544627),  
(309, 0.06243905410544627),  
(957, 0.06243905410544627),  
(1601, 0.06243905410544627),  
(1602, 0.06243905410544627),  
(1940, 0.06243905410544627),  
(2596, 0.06243905410544627),  
(2946, 0.06243905410544627),  
(3832, 0.06243905410544627),  
(4357, 0.06243905410544627),  
(79, 0.0623109738595896),  
(92, 0.0623109738595896),  
(106, 0.0623109738595896),  
(963, 0.0623109738595896),  
(1189, 0.0623109738595896),  
(1200, 0.0623109738595896),  
(2516, 0.0623109738595896),  
(2886, 0.0623109738595896),  
(1028, 0.06231097385958959),  
(107, 0.061806423257274694),  
(836, 0.061806423257274694),  
(3076, 0.061806423257274694),  
(90, 0.06131393394849658),  
(154, 0.06131393394849658),  
(259, 0.06131393394849658),  
(348, 0.06131393394849658),  
(404, 0.06131393394849658),  
(615, 0.06131393394849658),  
(863, 0.06131393394849658),  
(945, 0.06131393394849658),  
(1145, 0.06131393394849658),  
(1345, 0.06131393394849658),  
(1580, 0.06131393394849658),  
(1603, 0.06131393394849658),  
(2086, 0.06131393394849658),  
(2116, 0.06131393394849658),  
(2123, 0.06131393394849658),  
(2534, 0.06131393394849658),  
(2616, 0.06131393394849658),

```
(3238, 0.06131393394849658),
(3768, 0.06131393394849658),
(4194, 0.06131393394849658),
(4320, 0.06131393394849658),
(4566, 0.06131393394849658),
(4652, 0.06131393394849658),
(4664, 0.06131393394849658),
(375, 0.061313933948496574),
(733, 0.061313933948496574),
(768, 0.061313933948496574),
(1101, 0.061313933948496574),
(224, 0.060833032924035954),
(455, 0.060833032924035954),
(3775, 0.060833032924035954),
(4789, 0.060833032924035954),
(1354, 0.060363272743915036),
(358, 0.0602475233128778),
(533, 0.0602475233128778),
(664, 0.0602475233128778),
(991, 0.0602475233128778),
(996, 0.0602475233128778),
(1062, 0.0602475233128778),
(1090, 0.0602475233128778),
(1310, 0.0602475233128778),
(1404, 0.0602475233128778),
...]
```

```
In [68]: # Main Function()
```

```
In [108]: def recommend(movie):
# Check if the movie exists in the dataset
    if movie in new_moviesdf['title'].values:
        movie_index = new_moviesdf[new_moviesdf['title'] == movie].index[0]
        distances = similarity[movie_index]
        movies_list = sorted(list(enumerate(distances)), reverse=True, key=lambda x: x[1])[1:6]

        for i in movies_list:
            print(new_moviesdf.iloc[i[0]].title)
    else:
        print(f"Movie '{movie}' not found, please type the correct name or full name of the movie.")
```

```
In [121]: recommend('The Avengers')
```

```
Iron Man 3
Avengers: Age of Ultron
Captain America: Civil War
Captain America: The First Avenger
Iron Man
```

```
In [114]: import pickle
```

```
In [117]: pickle.dump(new_moviesdf, open('Movies-Recommender-System.pkl', 'wb'))
```

```
In [120]: pickle.dump(similarity, open('similarity.pkl', 'wb'))
```

```
In [ ]:
```