In [86]:
```
tup = 1, 2, 3
```

In [87]:
```
print(type(tup))
```

```
<class 'tuple'>
```

In [142…
```
import pandas as pd
```

In [143…
```
df = pd.read_csv(r"C:\Users\kajal\Downloads\movie_data\movies.csv")
```

In [144…
```
df
```

Out[144]:

	movieId	title	genres
0	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy
1	2	Jumanji (1995)	Adventure|Children|Fantasy
2	3	Grumpier Old Men (1995)	Comedy|Romance
3	4	Waiting to Exhale (1995)	Comedy|Drama|Romance
4	5	Father of the Bride Part II (1995)	Comedy
...	...	...	...
9737	193581	Black Butler: Book of the Atlantic (2017)	Action|Animation|Comedy|Fantasy
9738	193583	No Game No Life: Zero (2017)	Animation|Comedy|Fantasy
9739	193585	Flint (2017)	Drama
9740	193587	Bungo Stray Dogs: Dead Apple (2018)	Action|Animation
9741	193609	Andrew Dice Clay: Dice Rules (1991)	Comedy

9742 rows × 3 columns

In [145…
```
df1 = pd.read_csv(r"C:\Users\kajal\Downloads\movie_data\ratings.csv")
```

In [146…
```
df1
```

Out[146]:

	userId	movieId	rating	timestamp
0	1	1	4.0	964982703
1	1	3	4.0	964981247
2	1	6	4.0	964982224
3	1	47	5.0	964983815
4	1	50	5.0	964982931
...	...	...	...	...
100831	610	166534	4.0	1493848402
100832	610	168248	5.0	1493850091
100833	610	168250	5.0	1494273047
100834	610	168252	5.0	1493846352
100835	610	170875	3.0	1493846415

100836 rows × 4 columns

In [147…
```
df1['userId'].nunique()
```

Out[147]:
```
610
```

In [148…
```
df1['movieId'].nunique()
```

Out[148]:
```
9724
```

In [149…
```
df1['movieId'].max()
```

Out[149]:
```
193609
```

In [150…
```
merged_df = pd.merge(df, df1, on='movieId', how='inner')
```

In [151…
```
merged_df
```

Out[151]:

	movieId	title	genres	userId	rating	timestamp
0	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy	1	4.0	964982703
1	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy	5	4.0	847434962
2	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy	7	4.5	1106635946
3	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy	15	2.5	1510577970
4	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy	17	4.5	1305696483
...	...	...	...	...	...	...
100831	193581	Black Butler: Book of the Atlantic (2017)	Action|Animation|Comedy|Fantasy	184	4.0	1537109082
100832	193583	No Game No Life: Zero (2017)	Animation|Comedy|Fantasy	184	3.5	1537109545
100833	193585	Flint (2017)	Drama	184	3.5	1537109805
100834	193587	Bungo Stray Dogs: Dead Apple (2018)	Action|Animation	184	3.5	1537110021
100835	193609	Andrew Dice Clay: Dice Rules (1991)	Comedy	331	4.0	1537157606

100836 rows × 6 columns

In [152…
```
most_frequent_value = merged_df['title'].mode()
```

In [153…
```
most_frequent_value
```

Out[153]:
```
0    Forrest Gump (1994)
Name: title, dtype: object
```

In [155…
```
df2 = pd.read_csv(r"C:\Users\kajal\Downloads\movie_data\tags.csv")
```

In [156…
```
df2
```

Out[156]:

	userId	movieId	tag	timestamp
0	2	60756	funny	1445714994
1	2	60756	Highly quotable	1445714996
2	2	60756	will ferrell	1445714992
3	2	89774	Boxing story	1445715207
4	2	89774	MMA	1445715200
...	...	...	...	...
3678	606	7382	for katie	1171234019
3679	606	7936	austere	1173392334
3680	610	3265	gun fu	1493843984
3681	610	3265	heroic bloodshed	1493843978
3682	610	168248	Heroic Bloodshed	1493844270

3683 rows × 4 columns

In [157…
```
df_final = pd.merge(merged_df,df2, on='movieId', how='inner')
```

In [158…
```
df_final
```

Out[158]:

	movieId	title	genres	userId_x	rating	timestamp_x	userId_y	tag	timestamp_y
0	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy	1	4.0	964982703	336	pixar	1139045764
1	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy	1	4.0	964982703	474	pixar	1137206825
2	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy	1	4.0	964982703	567	fun	1525286013
3	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy	5	4.0	847434962	336	pixar	1139045764
4	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy	5	4.0	847434962	474	pixar	1137206825
...	...	...	...	...	...	...	...	...	...
233208	187595	Solo: A Star Wars Story (2018)	Action|Adventure|Children|Sci-Fi	586	5.0	1529899556	62	star wars	1528934552
233209	193565	Gintama: The Movie (2010)	Action|Animation|Comedy|Sci-Fi	184	3.5	1537098554	184	anime	1537098582
233210	193565	Gintama: The Movie (2010)	Action|Animation|Comedy|Sci-Fi	184	3.5	1537098554	184	comedy	1537098587
233211	193565	Gintama: The Movie (2010)	Action|Animation|Comedy|Sci-Fi	184	3.5	1537098554	184	gintama	1537098603
233212	193565	Gintama: The Movie (2010)	Action|Animation|Comedy|Sci-Fi	184	3.5	1537098554	184	remaster	1537098592

233213 rows × 9 columns

In [160…
```
filtered_df = df_final[df_final['title'] == "Matrix, The (1999)"]
```

In [161…
```
filtered_df
```

Out[161]:

	movieId	title	genres	userId_x	rating	timestamp_x	userId_y	tag	timestamp_y
142140	2571	Matrix, The (1999)	Action|Sci-Fi|Thriller	1	5.0	964981888	424	martial arts	1457842912
142141	2571	Matrix, The (1999)	Action|Sci-Fi|Thriller	1	5.0	964981888	424	sci-fi	1457842899
142142	2571	Matrix, The (1999)	Action|Sci-Fi|Thriller	1	5.0	964981888	474	alternate universe	1137204991
142143	2571	Matrix, The (1999)	Action|Sci-Fi|Thriller	1	5.0	964981888	537	philosophy	1424141098
142144	2571	Matrix, The (1999)	Action|Sci-Fi|Thriller	1	5.0	964981888	537	post apocalyptic	1424141101
...	...	...	...	...	...	...	...	...	...
143525	2571	Matrix, The (1999)	Action|Sci-Fi|Thriller	610	5.0	1479545822	424	martial arts	1457842912
143526	2571	Matrix, The (1999)	Action|Sci-Fi|Thriller	610	5.0	1479545822	424	sci-fi	1457842899
143527	2571	Matrix, The (1999)	Action|Sci-Fi|Thriller	610	5.0	1479545822	474	alternate universe	1137204991
143528	2571	Matrix, The (1999)	Action|Sci-Fi|Thriller	610	5.0	1479545822	537	philosophy	1424141098
143529	2571	Matrix, The (1999)	Action|Sci-Fi|Thriller	610	5.0	1479545822	537	post apocalyptic	1424141101

1390 rows × 9 columns

In [162…
```
filtered_df["tag"].unique()
```

Out[162]:
```
array(['martial arts', 'sci-fi', 'alternate universe', 'philosophy',
       'post apocalyptic'], dtype=object)
```

In [163…
```
filtered_df1 = df_final[df_final['title'] == "Terminator 2: Judgment Day (1991)"]
```

In [164…
```
filtered_df1
```

Out[164]:

	movieId	title	genres	userId_x	rating	timestamp_x	userId_y	tag	timestamp_y
93035	589	Terminator 2: Judgment Day (1991)	Action|Sci-Fi	5	3.0	847435258	424	apocalypse	1457844854
93036	589	Terminator 2: Judgment Day (1991)	Action|Sci-Fi	5	3.0	847435258	424	Arnold Schwarzenegger	1457844841
93037	589	Terminator 2: Judgment Day (1991)	Action|Sci-Fi	5	3.0	847435258	424	nuclear war	1457844861
93038	589	Terminator 2: Judgment Day (1991)	Action|Sci-Fi	5	3.0	847435258	424	sci-fi	1457844847
93039	589	Terminator 2: Judgment Day (1991)	Action|Sci-Fi	5	3.0	847435258	424	Suspense	1457844864
...	...	...	...	...	...	...	...	...	...
94822	589	Terminator 2: Judgment Day (1991)	Action|Sci-Fi	610	5.0	1479542983	424	sci-fi	1457844847
94823	589	Terminator 2: Judgment Day (1991)	Action|Sci-Fi	610	5.0	1479542983	424	Suspense	1457844864
94824	589	Terminator 2: Judgment Day (1991)	Action|Sci-Fi	610	5.0	1479542983	424	time travel	1457844843
94825	589	Terminator 2: Judgment Day (1991)	Action|Sci-Fi	610	5.0	1479542983	474	robots	1137206517
94826	589	Terminator 2: Judgment Day (1991)	Action|Sci-Fi	610	5.0	1479542983	477	Scifi masterpiece	1201166474

1792 rows × 9 columns

In [165…
```
filtered_df1["rating"].mean()
```

Out[165]:
```
3.970982142857143
```

In [167…
```
filtered_df2 = df_final[df_final['title'] == "Fight Club (1999)"]
```

In [168…
```
filtered_df2
```

Out[168]:

	movieId	title	genres	userId_x	rating	timestamp_x	userId_y	tag	timestamp_y
148882	2959	Fight Club (1999)	Action|Crime|Drama|Thriller	1	5.0	964983282	424	dark comedy	1457842797
148883	2959	Fight Club (1999)	Action|Crime|Drama|Thriller	1	5.0	964983282	424	psychology	1457842802
148884	2959	Fight Club (1999)	Action|Crime|Drama|Thriller	1	5.0	964983282	424	thought-provoking	1457842786
148885	2959	Fight Club (1999)	Action|Crime|Drama|Thriller	1	5.0	964983282	424	twist ending	1457842777
148886	2959	Fight Club (1999)	Action|Crime|Drama|Thriller	1	5.0	964983282	435	dark comedy	1366676088
...	...	...	...	...	...	...	...	...	...
160649	2959	Fight Club (1999)	Action|Crime|Drama|Thriller	610	5.0	1479541966	599	thought-provoking	1498456901
160650	2959	Fight Club (1999)	Action|Crime|Drama|Thriller	610	5.0	1479541966	599	twist	1498456943
160651	2959	Fight Club (1999)	Action|Crime|Drama|Thriller	610	5.0	1479541966	599	twist ending	1498456888
160652	2959	Fight Club (1999)	Action|Crime|Drama|Thriller	610	5.0	1479541966	599	violence	1498456904
160653	2959	Fight Club (1999)	Action|Crime|Drama|Thriller	610	5.0	1479541966	599	violent	1498456914

11772 rows × 9 columns

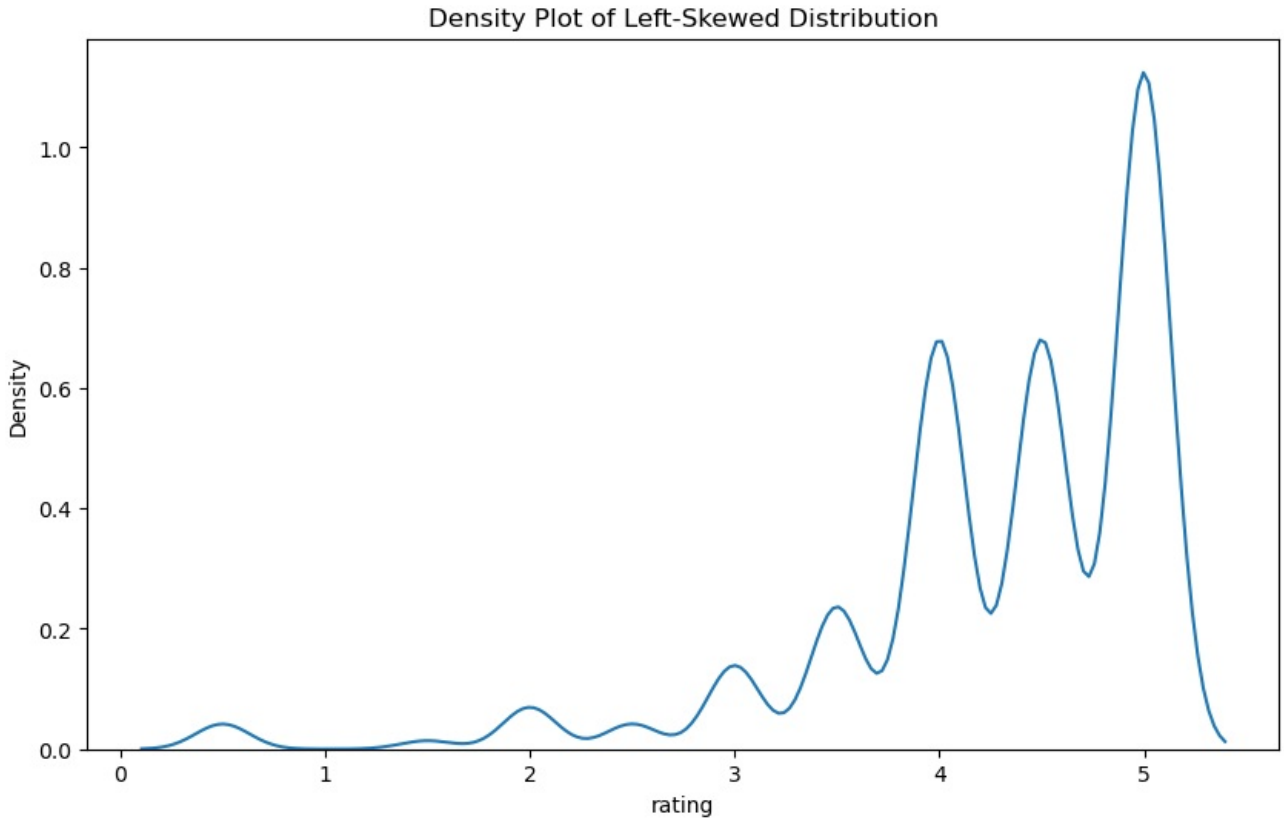
In [170…
```
import matplotlib.pyplot as plt
import seaborn as sns
```

In [171…
```
skewness = filtered_df2['rating'].skew()
```

In [113…
```
skewness
```

Out[113]:
```
-1.8477291838682013
```

In [172…
```
plt.figure(figsize=(10, 6))
sns.kdeplot(filtered_df2['rating'])
plt.title('Density Plot of Left-Skewed Distribution')
plt.xlabel('rating')
plt.ylabel('Density')
plt.show()
```



In [237…
```
grouped = df_final.groupby('movieId')['rating'].agg(["count","mean"])
grouped
```

Out[237]:

movieId	count	mean
1	645	3.920930
2	440	3.431818
3	104	3.259615
5	98	3.071429
7	54	3.185185
...	...	...
183611	3	4.000000
184471	12	2.500000
187593	36	3.875000
187595	10	3.900000
193565	4	3.500000

1554 rows × 2 columns

In [180…
```
merged_df = pd.merge(df, grouped, on='movieId', how='inner')
```

In [181…
```
merged_df
```

Out[181]:

	movieId	title	genres	count	mean
0	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy	645	3.920930
1	2	Jumanji (1995)	Adventure|Children|Fantasy	440	3.431818
2	3	Grumpier Old Men (1995)	Comedy|Romance	104	3.259615
3	5	Father of the Bride Part II (1995)	Comedy	98	3.071429
4	7	Sabrina (1995)	Comedy|Romance	54	3.185185
...	...	...	...	...	...
1549	183611	Game Night (2018)	Action|Comedy|Crime|Horror	3	4.000000
1550	184471	Tomb Raider (2018)	Action|Adventure|Fantasy	12	2.500000
1551	187593	Deadpool 2 (2018)	Action|Comedy|Sci-Fi	36	3.875000
1552	187595	Solo: A Star Wars Story (2018)	Action|Adventure|Children|Sci-Fi	10	3.900000
1553	193565	Gintama: The Movie (2010)	Action|Animation|Comedy|Sci-Fi	4	3.500000

1554 rows × 5 columns

In [182…
```
filtered_50 = (merged_df['count'] >50).value_counts()
filtered_50
```

Out[182]:
```
count
False    1074
True      480
Name: count, dtype: int64
```

In [192…
```
filtered_50_new = merged_df[merged_df['count'] > 50]
filtered_50_new
```

Out[192]:

	movieId	title	genres	count	mean
0	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy	645	3.920930
1	2	Jumanji (1995)	Adventure|Children|Fantasy	440	3.431818
2	3	Grumpier Old Men (1995)	Comedy|Romance	104	3.259615
3	5	Father of the Bride Part II (1995)	Comedy	98	3.071429
4	7	Sabrina (1995)	Comedy|Romance	54	3.185185
...	...	...	...	...	...
1525	148626	Big Short, The (2015)	Drama	78	3.961538
1536	164179	Arrival (2016)	Sci-Fi	182	3.980769
1539	168248	John Wick: Chapter Two (2017)	Action|Crime|Thriller	56	4.142857
1540	168252	Logan (2017)	Action|Sci-Fi	125	4.280000
1544	176371	Blade Runner 2049 (2017)	Sci-Fi	144	3.805556

480 rows × 5 columns

In [ ]:
```

```

In [195…
```
filtered_50_new[filtered_50_new['mean'] == filtered_50_new["mean"].max() ]
```

Out[195]:

	movieId	title	genres	count	mean
83	318	Shawshank Redemption, The (1994)	Crime|Drama	1268	4.429022

In [220…
```
top_5_nlargest = filtered_50_new.nlargest(5, 'count')
top_5_nlargest
```

Out[220]:

	movieId	title	genres	count	mean
77	296	Pulp Fiction (1994)	Comedy|Crime|Drama|Thriller	55567	4.197068
637	2959	Fight Club (1999)	Action|Crime|Drama|Thriller	11772	4.272936
65	260	Star Wars: Episode IV - A New Hope (1977)	Action|Adventure|Sci-Fi	6526	4.231076
76	293	Léon: The Professional (a.k.a. The Professiona...	Action|Crime|Drama|Thriller	4655	4.018797
206	924	2001: A Space Odyssey (1968)	Adventure|Drama|Sci-Fi	4469	3.894495

In [ ]:
```

```

In [240…
```
sci_fi_df = filtered_50_new[filtered_50_new['genres'].str.contains ("Sci-Fi")]
sci_fi_df
```

Out[240]:

	movieId	title	genres	count	mean
16	32	Twelve Monkeys (a.k.a. 12 Monkeys) (1995)	Mystery|Sci-Fi|Thriller	1770	3.983051
44	160	Congo (1995)	Action|Adventure|Mystery|Sci-Fi	57	2.877193
65	260	Star Wars: Episode IV - A New Hope (1977)	Action|Adventure|Sci-Fi	6526	4.231076
81	316	Stargate (1994)	Action|Adventure|Sci-Fi	140	3.375000
85	329	Star Trek: Generations (1994)	Adventure|Drama|Sci-Fi	108	3.393519
...	...	...	...	...	...
1516	135536	Suicide Squad (2016)	Action|Crime|Sci-Fi	228	2.916667
1517	136864	Batman v Superman: Dawn of Justice (2016)	Action|Adventure|Fantasy|Sci-Fi	144	2.343750
1536	164179	Arrival (2016)	Sci-Fi	182	3.980769
1540	168252	Logan (2017)	Action|Sci-Fi	125	4.280000
1544	176371	Blade Runner 2049 (2017)	Sci-Fi	144	3.805556

93 rows × 5 columns

In [247…
```
top_3_nlargest = sci_fi_df.nlargest(10,'count')
top_3_nlargest
```

Out[247]:

	movieId	title	genres	count	mean
65	260	Star Wars: Episode IV - A New Hope (1977)	Action|Adventure|Sci-Fi	6526	4.231076
206	924	2001: A Space Odyssey (1968)	Adventure|Drama|Sci-Fi	4469	3.894495
1140	7361	Eternal Sunshine of the Spotless Mind (2004)	Drama|Romance|Sci-Fi	4454	4.160305
1423	79132	Inception (2010)	Action|Crime|Drama|Mystery|Sci-Fi|Thriller|IMAX	3718	4.066434
859	4878	Donnie Darko (2001)	Drama|Mystery|Sci-Fi|Thriller	3161	3.981651
287	1196	Star Wars: Episode V - The Empire Strikes Back...	Action|Adventure|Sci-Fi	2110	4.215640
143	589	Terminator 2: Judgment Day (1991)	Action|Sci-Fi	1792	3.970982
16	32	Twelve Monkeys (a.k.a. 12 Monkeys) (1995)	Mystery|Sci-Fi|Thriller	1770	3.983051
1420	72998	Avatar (2009)	Action|Adventure|Sci-Fi|IMAX	1746	3.603093
134	541	Blade Runner (1982)	Action|Sci-Fi|Thriller	1612	4.100806

In [245…
```
top_3_nsmall = sci_fi_df.nsmallest(5,'count')
top_3_nsmall
```

Out[245]:

	movieId	title	genres	count	mean
160	673	Space Jam (1996)	Adventure|Animation|Children|Comedy|Fantasy|Sc...	53	2.707547
812	4343	Evolution (2001)	Comedy|Sci-Fi	54	3.037037
1443	91500	The Hunger Games (2012)	Action|Adventure|Drama|Sci-Fi|Thriller	54	3.435185
572	2529	Planet of the Apes (1968)	Action|Drama|Sci-Fi	56	3.803571
44	160	Congo (1995)	Action|Adventure|Mystery|Sci-Fi	57	2.877193

In [246…
```
sci_fi_df.sort_values(by=['count', 'mean'], ascending=[True, True])
```

Out[246]:

	movieId	title	genres	count	mean
160	673	Space Jam (1996)	Adventure|Animation|Children|Comedy|Fantasy|Sc...	53	2.707547
812	4343	Evolution (2001)	Comedy|Sci-Fi	54	3.037037
1443	91500	The Hunger Games (2012)	Action|Adventure|Drama|Sci-Fi|Thriller	54	3.435185
572	2529	Planet of the Apes (1968)	Action|Drama|Sci-Fi	56	3.803571
44	160	Congo (1995)	Action|Adventure|Mystery|Sci-Fi	57	2.877193
...	...	...	...	...	...
859	4878	Donnie Darko (2001)	Drama|Mystery|Sci-Fi|Thriller	3161	3.981651
1423	79132	Inception (2010)	Action|Crime|Drama|Mystery|Sci-Fi|Thriller|IMAX	3718	4.066434
1140	7361	Eternal Sunshine of the Spotless Mind (2004)	Drama|Romance|Sci-Fi	4454	4.160305
206	924	2001: A Space Odyssey (1968)	Adventure|Drama|Sci-Fi	4469	3.894495
65	260	Star Wars: Episode IV - A New Hope (1977)	Action|Adventure|Sci-Fi	6526	4.231076

93 rows × 5 columns

In [253…
```
sci_fi_df[sci_fi_df['title'] == "Terminator 2: Judgment Day (1991)"]
```

Out[253]:

	movieId	title	genres	count	mean
143	589	Terminator 2: Judgment Day (1991)	Action|Sci-Fi	1792	3.970982

In [255…
```
sci_fi_df[sci_fi_df['title'] == "X-Men: The Last Stand (2006)"]
```

Out[255]:

	movieId	title	genres	count	mean
1345	45499	X-Men: The Last Stand (2006)	Action|Sci-Fi|Thriller	104	3.355769

In [256…
```
sci_fi_df[sci_fi_df['title'] == "Jurassic Park (1993)"]
```

Out[256]:

	movieId	title	genres	count	mean
116	480	Jurassic Park (1993)	Action|Adventure|Sci-Fi|Thriller	238	3.75

In [257…
```
df4 = pd.read_csv(r"C:\Users\kajal\Downloads\movie_data\links.csv")
```

In [258…
```
df4
```

Out[258]:

	movieId	imdbId	tmdbId
0	1	114709	862.0
1	2	113497	8844.0
2	3	113228	15602.0
3	4	114885	31357.0
4	5	113041	11862.0
...	...	...	...
9737	193581	5476944	432131.0
9738	193583	5914996	445030.0
9739	193585	6397426	479308.0
9740	193587	8391976	483455.0
9741	193609	101726	37891.0

9742 rows × 3 columns

In [259…
```
df_newfinal = pd.merge(df4, filtered_50_new, on='movieId', how='inner')
```

In [260…
```
df_newfinal
```

Out[260]:

	movieId	imdbId	tmdbId	title	genres	count	mean
0	1	114709	862.0	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy	645	3.920930
1	2	113497	8844.0	Jumanji (1995)	Adventure|Children|Fantasy	440	3.431818
2	3	113228	15602.0	Grumpier Old Men (1995)	Comedy|Romance	104	3.259615
3	5	113041	11862.0	Father of the Bride Part II (1995)	Comedy	98	3.071429
4	7	114319	11860.0	Sabrina (1995)	Comedy|Romance	54	3.185185
...	...	...	...	...	...	...	...
475	148626	1596363	318846.0	Big Short, The (2015)	Drama	78	3.961538
476	164179	2543164	329865.0	Arrival (2016)	Sci-Fi	182	3.980769
477	168248	4425200	324552.0	John Wick: Chapter Two (2017)	Action|Crime|Thriller	56	4.142857
478	168252	3315342	263115.0	Logan (2017)	Action|Sci-Fi	125	4.280000
479	176371	1856101	335984.0	Blade Runner 2049 (2017)	Sci-Fi	144	3.805556

480 rows × 7 columns

In [263…
```
df_newfinal['imdbId'].max()
```

Out[263]:
```
4425200
```

In [264…
```
df_newfinal[df_newfinal['imdbId'] == 4425200]
```

Out[264]:

	movieId	imdbId	tmdbId	title	genres	count	mean
477	168248	4425200	324552.0	John Wick: Chapter Two (2017)	Action|Crime|Thriller	56	4.142857

In [265…
```
sci_fi_imdb = df_newfinal[df_newfinal['genres'].str.contains ("Sci-Fi")]
sci_fi_imdb
```

Out[265]:

	movieId	imdbId	tmdbId	title	genres	count	mean
11	32	114746	63.0	Twelve Monkeys (a.k.a. 12 Monkeys) (1995)	Mystery|Sci-Fi|Thriller	1770	3.983051
26	160	112715	10329.0	Congo (1995)	Action|Adventure|Mystery|Sci-Fi	57	2.877193
35	260	76759	11.0	Star Wars: Episode IV - A New Hope (1977)	Action|Adventure|Sci-Fi	6526	4.231076
43	316	111282	2164.0	Stargate (1994)	Action|Adventure|Sci-Fi	140	3.375000
46	329	111280	193.0	Star Trek: Generations (1994)	Adventure|Drama|Sci-Fi	108	3.393519
...	...	...	...	...	...	...	...
472	135536	1386697	297761.0	Suicide Squad (2016)	Action|Crime|Sci-Fi	228	2.916667
473	136864	2975590	209112.0	Batman v Superman: Dawn of Justice (2016)	Action|Adventure|Fantasy|Sci-Fi	144	2.343750
476	164179	2543164	329865.0	Arrival (2016)	Sci-Fi	182	3.980769
478	168252	3315342	263115.0	Logan (2017)	Action|Sci-Fi	125	4.280000
479	176371	1856101	335984.0	Blade Runner 2049 (2017)	Sci-Fi	144	3.805556

93 rows × 7 columns

In [266…
```
sci_fi_imdb[sci_fi_imdb['imdbId'] == sci_fi_imdb['imdbId'].max() ]
```

Out[266]:

	movieId	imdbId	tmdbId	title	genres	count	mean
468	122912	4154756	299536.0	Avengers: Infinity War - Part I (2018)	Action|Adventure|Sci-Fi	195	4.0

In [ ]:

In [297…
```
from bs4 import BeautifulSoup
import requests
```

URL = "https://www.imdb.com/chart/top/" page = requests.get(URL) page

In [312…
```
import requests
from bs4 import BeautifulSoup
import numpy as np

def scrapper(imdbId):
    id = str(int(imdbId))
    n_zeroes = 7 - len(id)
    new_id = "0"*n_zeroes + id
    URL = "https://www.imdb.com/list/ls053181721/"
    request_header = {'Content-Type': 'text/html; charset=UTF-8',
                      'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:109.0) Gecko/20100101 Firefox
                      'Accept-Encoding': 'gzip, deflate, br'}
    response = requests.get(URL, headers=request_header)
    soup = BeautifulSoup(response.text, 'html.parser')
    imdb_rating = soup.find('span', class_ = "sc-b189961a-0 iqHBGn")
    return imdb_rating.text if imdb_rating else np.nan
```

In [292…

In [ ]:

In [ ]:

In [313…
```
scrapper(122912)
```

Out[313]:
```
nan
```

In [ ]: