CONTENT BASED RECOMMENDER SYSTEM

BASED ON KNN

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

In [2]:
```python
movies=pd.read_csv('E:/movies.csv')
credits=pd.read_csv('E:/credits.csv')
```

In [3]:
```python
movies.head(3)
```

Out[3]:

	budget	genres	homepage	id	keywords	original_language	original_title	overview	popularity
0	237000000	[{"id": 28, "name": "Action"}, {"id": 12, "nam...	http://www.avatarmovie.com/	19995	[{"id": 1463, "name": "culture clash"}, {"id":...	en	Avatar	In the 22nd century, a paraplegic Marine is di...	150.437577
1	300000000	[{"id": 12, "name": "Adventure"}, {"id": 14, "...	http://disney.go.com/disneypictures/pirates/	285	[{"id": 270, "name": "ocean"}, {"id": 726, "na...	en	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...	139.082615
2	245000000	[{"id": 28, "name": "Action"}, {"id": 12, "nam...	http://www.sonypictures.com/movies/spectre	206647	[{"id": 470, "name": "spy"}, {"id": 818, "name...	en	Spectre	A cryptic message from Bond's past sends him o...	107.376788

In [4]:
```python
credits.head(1)
```

Out[4]:

	movie_id	title	cast	crew
0	19995	Avatar	[{"cast_id": 242, "character": "Jake Sully", "...	[{"credit_id": "52fe48009251416c750aca23", "de...

In [5]:
```python
credits.cast.values
```

Out[5]:
array(['[{"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": "52fe480092514 16c750ac9cb", "gender": 1, "id": 8691, "name": "Zoe Saldana", "order": 1}, {"cast_id": 25, "character": "Dr. Grac e Augustine", "credit_id": "52fe48009251416c750aca39", "gender": 1, "id": 10205, "name": "Sigourney Weaver", "ord er": 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": "52fe4800925 1416c750ac9d3", "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": 59 231, "name": "Joel David Moore", "order": 6}, {"cast_id": 9, "character": "Moat", "credit_id": "52fe48009251416c7 50ac9e5", "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": 1 0, "character": "Tsu\'Tey", "credit_id": "52fe48009251416c750ac9e9", "gender": 2, "id": 10964, "name": "Laz Alons o", "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, "charac ter": "Private Fike", "credit_id": "52fe48009251416c750aca5b", "gender": 2, "id": 154153, "name": "Sean Anthony M oran", "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 C hief", "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": 98673 4, "name": "Kelly Kilgour", "order": 15}, {"cast_id": 36, "character": "Shuttle Pilot", "credit_id": "52fe4800925 1416c750aca6b", "gender": 0, "id": 1207227, "name": "James Patrick Pitt", "order": 16}, {"cast_id": 37, "characte r": "Shuttle Co-Pilot", "credit_id": "52fe48009251416c750aca6f", "gender": 0, "id": 1180936, "name": "Sean Patric k Murphy", "order": 17}, {"cast_id": 38, "character": "Shuttle Crew Chief", "credit_id": "52fe48009251416c750aca7 3", "gender": 2, "id": 1019578, "name": "Peter Dillon", "order": 18}, {"cast_id": 39, "character": "Tractor Opera tor / Troupe", "credit_id": "52fe48009251416c750aca77", "gender": 2, "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", "cr edit_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": 215 913, "name": "Jacob Tomuri", "order": 22}, {"cast_id": 43, "character": "Suit #1", "credit_id": "52fe48009251416c

750aca87", "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": "Sonia 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": "52fe48019251416c750acaeb", "gender": 0, "id": 1207265, "name": "Paul Yates", "order": 48}, {"cast_id": 70, "character": "Op Center Duty Officer", "credit_id": "52fe48019251416c750acaef", "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": "Rhian 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}]',
        '[{"cast_id": 4, "character": "Captain Jack Sparrow", "credit_id": "52fe4232c3a36847f800b50d", "gender": 2, "id": 85, "name": "Johnny Depp", "order": 0}, {"cast_id": 5, "character": "Will Turner", "credit_id": "52fe4232c3a36847f800b511", "gender": 2, "id": 114, "name": "Orlando Bloom", "order": 1}, {"cast_id": 6, "character": "Elizabeth Swann", "credit_id": "52fe4232c3a36847f800b515", "gender": 1, "id": 116, "name": "Keira Knightley", "order": 2}, {"cast_id": 12, "character": "William \\"Bootstrap Bill\\" Turner", "credit_id": "52fe4232c3a36847f800b52d", "gender": 2, "id": 1640, "name": "Stellan Skarsg\u00e5rd", "order": 3}, {"cast_id": 10, "character": "Captain

Sao Feng", "credit_id": "52fe4232c3a36847f800b525", "gender": 2, "id": 1619, "name": "Chow Yun-fat", "order": 4},
{"cast_id": 9, "character": "Captain Davy Jones", "credit_id": "52fe4232c3a36847f800b521", "gender": 2, "id": 244
0, "name": "Bill Nighy", "order": 5}, {"cast_id": 7, "character": "Captain Hector Barbossa", "credit_id": "52fe42
32c3a36847f800b519", "gender": 2, "id": 118, "name": "Geoffrey Rush", "order": 6}, {"cast_id": 14, "character": "
Admiral James Norrington", "credit_id": "52fe4232c3a36847f800b535", "gender": 2, "id": 1709, "name": "Jack Davenp
ort", "order": 7}, {"cast_id": 13, "character": "Joshamee Gibbs", "credit_id": "52fe4232c3a36847f800b531", "gende
r": 2, "id": 2449, "name": "Kevin McNally", "order": 8}, {"cast_id": 11, "character": "Lord Cutler Beckett", "cre
dit_id": "52fe4232c3a36847f800b529", "gender": 2, "id": 2441, "name": "Tom Hollander", "order": 9}, {"cast_id": 1
9, "character": "Tia Dalma", "credit_id": "52fe4232c3a36847f800b549", "gender": 1, "id": 2038, "name": "Naomie Ha
rris", "order": 10}, {"cast_id": 8, "character": "Governor Weatherby Swann", "credit_id": "52fe4232c3a36847f800b5
1d", "gender": 2, "id": 378, "name": "Jonathan Pryce", "order": 11}, {"cast_id": 37, "character": "Captain Teague
Sparrow", "credit_id": "52fe4232c3a36847f800b5b3", "gender": 2, "id": 1430, "name": "Keith Richards", "order": 12
}, {"cast_id": 16, "character": "Pintel", "credit_id": "52fe4232c3a36847f800b53d", "gender": 2, "id": 1710, "name
": "Lee Arenberg", "order": 13}, {"cast_id": 15, "character": "Ragetti", "credit_id": "52fe4232c3a36847f800b539",
"gender": 2, "id": 1711, "name": "Mackenzie Crook", "order": 14}, {"cast_id": 18, "character": "Lieutenant Theodo
re Groves", "credit_id": "52fe4232c3a36847f800b545", "gender": 2, "id": 4031, "name": "Greg Ellis", "order": 15},
{"cast_id": 55, "character": "Cotton", "credit_id": "57e28d2ec3a3681a01005b5c", "gender": 2, "id": 1715, "name":
"David Bailie", "order": 16}, {"cast_id": 17, "character": "Marty", "credit_id": "52fe4232c3a36847f800b541", "gen
der": 2, "id": 4030, "name": "Martin Klebba", "order": 17}, {"cast_id": 57, "character": "Ian Mercer", "credit_id
": "57e28d78c3a36808b900bf4f", "gender": 0, "id": 939, "name": "David Schofield", "order": 18}, {"cast_id": 62, "
character": "Scarlett", "credit_id": "57e28ec5c3a3681a50005855", "gender": 1, "id": 2450, "name": "Lauren Maher",
"order": 19}, {"cast_id": 63, "character": "Giselle", "credit_id": "57e28ed692514123f5005635", "gender": 1, "id":
2452, "name": "Vanessa Branch", "order": 20}, {"cast_id": 60, "character": "Mullroy", "credit_id": "57e28db2c3a36
81a01005bc7", "gender": 2, "id": 1714, "name": "Angus Barnett", "order": 21}, {"cast_id": 59, "character": "Murto
gg", "credit_id": "57e28da192514118f7006008", "gender": 0, "id": 1713, "name": "Giles New", "order": 22}, {"cast_
id": 58, "character": "Tai Huang", "credit_id": "57e28d8ec3a3681a01005bab", "gender": 2, "id": 22075, "name": "Re
ggie Lee", "order": 23}, {"cast_id": 64, "character": "Henry Turner", "credit_id": "57e29119925141151100a6cc", "g
ender": 2, "id": 61259, "name": "Dominic Scott Kay", "order": 24}, {"cast_id": 39, "character": "Mistress Ching",
"credit_id": "52fe4232c3a36847f800b5bd", "gender": 1, "id": 33500, "name": "Takayo Fischer", "order": 25}, {"cast
_id": 40, "character": "Lieutenant Greitzer", "credit_id": "52fe4232c3a36847f800b5c1", "gender": 2, "id": 1224149
, "name": "David Meunier", "order": 26}, {"cast_id": 49, "character": "Hadras", "credit_id": "56d1871c92514174680
010cf", "gender": 2, "id": 429401, "name": "Ho-Kwan Tse", "order": 27}, {"cast_id": 56, "character": "Clacker", "
credit_id": "57e28d4b92514125710055cb", "gender": 0, "id": 1123, "name": "Andy Beckwith", "order": 28}, {"cast_id
": 51, "character": "Penrod", "credit_id": "56ec8c14c3a3682260003c53", "gender": 2, "id": 1056117, "name": "Peter
Donald Badalamenti II", "order": 29}, {"cast_id": 61, "character": "Cotton\'s Parrot (voice)", "credit_id": "57e2
8dcc9251412463005678", "gender": 2, "id": 21700, "name": "Christopher S. Capp", "order": 30}, {"cast_id": 65, "ch
aracter": "Captain Teague", "credit_id": "58bc2a37c3a368663003740b", "gender": 2, "id": 1430, "name": "Keith Rich
ards", "order": 31}, {"cast_id": 66, "character": "Captain Jocard", "credit_id": "58bc2a8e925141609e03a179", "gen
der": 2, "id": 2603, "name": "Hakeem Kae-Kazim", "order": 32}, {"cast_id": 67, "character": "Captain Ammand", "cr
edit_id": "58e2a21ac3a36872af00f9c2", "gender": 0, "id": 70577, "name": "Ghassan Massoud", "order": 33}]',
        '[{"cast_id": 1, "character": "James Bond", "credit_id": "52fe4d22c3a368484e1d8d6b", "gender": 2, "id": 87
84, "name": "Daniel Craig", "order": 0}, {"cast_id": 14, "character": "Blofeld", "credit_id": "54805866c3a36829ab
002592", "gender": 2, "id": 27319, "name": "Christoph Waltz", "order": 1}, {"cast_id": 13, "character": "Madelein
e", "credit_id": "546f934fc3a3682f9a002ca5", "gender": 1, "id": 121529, "name": "L\u00e9a Seydoux", "order": 2},
{"cast_id": 10, "character": "M", "credit_id": "53e86503c3a368399c0031f0", "gender": 2, "id": 5469, "name": "Ralp
h Fiennes", "order": 3}, {"cast_id": 17, "character": "Lucia", "credit_id": "54805920c3a36829ae0022c5", "gender":
1, "id": 28782, "name": "Monica Bellucci", "order": 4}, {"cast_id": 8, "character": "Q", "credit_id": "52fe4d22c3
a368484e1d8d87", "gender": 2, "id": 17064, "name": "Ben Whishaw", "order": 5}, {"cast_id": 11, "character": "Mone
ypenny", "credit_id": "53e8650cc3a368399c0031f4", "gender": 1, "id": 2038, "name": "Naomie Harris", "order": 6},
{"cast_id": 16, "character": "Hinx", "credit_id": "5480590f9251416e7b0022f4", "gender": 2, "id": 543530, "name":
"Dave Bautista", "order": 7}, {"cast_id": 15, "character": "C", "credit_id": "548058ffc3a36829a7002511", "gender"
: 2, "id": 125660, "name": "Andrew Scott", "order": 8}, {"cast_id": 19, "character": "Tanner", "credit_id": "5480
593cc3a36829ae0022ca", "gender": 2, "id": 139549, "name": "Rory Kinnear", "order": 9}, {"cast_id": 23, "character
": "Mr. White", "credit_id": "5517b3849251416f0d0032c8", "gender": 2, "id": 2244, "name": "Jesper Christensen", "
order": 10}, {"cast_id": 24, "character": "Marco Sciarra", "credit_id": "5517b398c3a36862fc002df6", "gender": 0,
"id": 1381616, "name": "Alessandro Cremona", "order": 11}, {"cast_id": 25, "character": "Estrella", "credit_id":
"5517b3c49251412c1e00180b", "gender": 1, "id": 929937, "name": "Stephanie Sigman", "order": 12}, {"cast_id": 49,
"character": "Mexican Man in Lift", "credit_id": "5664561cc3a3682bd0007fc8", "gender": 0, "id": 87265, "name": "T
enoch Huerta", "order": 13}, {"cast_id": 50, "character": "Mexican Woman in Lift", "credit_id": "5664561dc3a3682b
e9008dc6", "gender": 1, "id": 983710, "name": "Adriana Paz", "order": 14}, {"cast_id": 180, "character": "Gallo",
"credit_id": "56fddbacc3a36821b6000aca", "gender": 2, "id": 1385592, "name": "Domenico Fortunato", "order": 15},
{"cast_id": 51, "character": "Gallo\'s Accomplice", "credit_id": "5664561dc3a3682bd8008fbe", "gender": 0, "id": 1
545546, "name": "Marco Zingaro", "order": 16}, {"cast_id": 52, "character": "Gallo\'s Accomplice", "credit_id": "
5664561d9251412ad300870b", "gender": 2, "id": 1545547, "name": "Stefano Elfi DiClaudia", "order": 17}, {"cast_id"
: 181, "character": "Q\'s Assistant", "credit_id": "56fddc319251412a4d0009f9", "gender": 0, "id": 122556, "name":
"Ian Bonar", "order": 18}, {"cast_id": 182, "character": "Moneypenny\'s Boyfriend", "credit_id": "56fddc66c3a3682
1b6000b04", "gender": 2, "id": 1599239, "name": "Tam Williams", "order": 19}, {"cast_id": 183, "character": "Blof
eld\'s London Helicopter Pilot", "credit_id": "56fddc989251412a56000a85", "gender": 2, "id": 1599241, "name": "Ri
chard Banham", "order": 20}, {"cast_id": 184, "character": "SC019 Police Commander", "credit_id": "56fddccf925141
2a4b000b5d", "gender": 2, "id": 230681, "name": "Pip Carter", "order": 21}, {"cast_id": 185, "character": "SC019
Police Officer", "credit_id": "56fddce6c3a36821ae000b45", "gender": 2, "id": 30446, "name": "Simon Lenagan", "ord
er": 22}, {"cast_id": 27, "character": "Priest", "credit_id": "5517b4f69251416efe003370", "gender": 2, "id": 4837
7, "name": "Alessandro Bressanello", "order": 23}, {"cast_id": 38, "character": "Moreau", "credit_id": "56486811c
3a3682605006adc", "gender": 0, "id": 578512, "name": "Marc Zinga", "order": 24}, {"cast_id": 39, "character": "Vo
gel", "credit_id": "56486820c3a368261a006de7", "gender": 0, "id": 1093974, "name": "Brigitte Millar", "order": 25
}, {"cast_id": 41, "character": "Abrika", "credit_id": "564868a09251413ad7006da5", "gender": 2, "id": 89624, "nam
e": "Adel Bencherif", "order": 26}, {"cast_id": 186, "character": "Blofeld\'s Right Hand Man", "credit_id": "56fd
dd5a9251412a42000b52", "gender": 2, "id": 1129783, "name": "Gediminas Adomaitis", "order": 27}, {"cast_id": 42, "
character": "Lorenzo", "credit_id": "564868c5c3a368260c006ee4", "gender": 0, "id": 226388, "name": "Peppe Lanzett
a", "order": 28}, {"cast_id": 187, "character": "Francesco", "credit_id": "56fddd9cc3a36821a4000b5b", "gender": 2
, "id": 1299312, "name": "Francesco Arca", "order": 29}, {"cast_id": 188, "character": "Marco", "credit_id": "56f
dddb79251412a50000aad", "gender": 2, "id": 124628, "name": "Matteo Taranto", "order": 30}, {"cast_id": 189, "char
acter": "Palazzo Security", "credit_id": "56fdddd69251412a56000ac6", "gender": 2, "id": 1562103, "name": "Emilio
Aniba", "order": 31}, {"cast_id": 190, "character": "Guerra", "credit_id": "56fdddf19251412a50000ac1", "gender":
2, "id": 1474616, "name": "Benito Sagredo", "order": 32}, {"cast_id": 191, "character": "Businessman", "credit_id
": "56fdde0d9251416cc200004c", "gender": 2, "id": 1247684, "name": "Dai Tabuchi", "order": 33}, {"cast_id": 192,
"character": "Businessman", "credit_id": "56fdde1bc3a36821b1000ba8", "gender": 2, "id": 1599245, "name": "George
Lasha", "order": 34}, {"cast_id": 193, "character": "Businessman", "credit_id": "56fdde259251412a42000b93", "gend
er": 0, "id": 1599247, "name": "Sargon Yelda", "order": 35}, {"cast_id": 194, "character": "Businessman", "credit

_id": "56fdde9fc3a36821b6000ba0", "gender": 2, "id": 1599251, "name": "Andy Cheung", "order": 36}, {"cast_id": 195, "character": "Marshall", "credit_id": "56fddeb29251412a50000af4", "gender": 2, "id": 1599252, "name": "Erick Hayden", "order": 37}, {"cast_id": 196, "character": "Valerian", "credit_id": "56fddec8c3a36821ab000c19", "gender": 2, "id": 1599253, "name": "Oleg Mirochnikov", "order": 38}, {"cast_id": 197, "character": "Fiat Driver", "credit_id": "56fddef89251412a47000a79", "gender": 2, "id": 123548, "name": "Antonio Salines", "order": 39}, {"cast_id": 198, "character": "Street Sweeper", "credit_id": "56fddf109251416cc2000086", "gender": 2, "id": 1599254, "name": "Miloud Mourad Benamara", "order": 40}, {"cast_id": 199, "character": "Chairman", "credit_id": "56fddf4a9251412a42000bcf", "gender": 2, "id": 1599255, "name": "Gido Schimanski", "order": 41}, {"cast_id": 200, "character": "Head of Nation", "credit_id": "56fddf69c3a36821a7000b12", "gender": 0, "id": 1404403, "name": "Nigel Barber", "order": 42}, {"cast_id": 201, "character": "Head of Nation", "credit_id": "56fddf7fc3a36821b1000c0d", "gender": 2, "id": 74415, "name": "Patrice Naiambana", "order": 43}, {"cast_id": 202, "character": "Head of Nation", "credit_id": "56fddf89c3a368219d000adb", "gender": 2, "id": 1599256, "name": "Stephane Cornicard", "order": 44}, {"cast_id": 203, "character": "Head of Nation", "credit_id": "56fddf929251412a56000b29", "gender": 2, "id": 1599257, "name": "Gary Fannin", "order": 45}, {"cast_id": 204, "character": "Head of Nation", "credit_id": "56fddf9cc3a36821ae000be1", "gender": 2, "id": 1599258, "name": "Sadao Ueda", "order": 46}, {"cast_id": 205, "character": "Head of Nation", "credit_id": "56fddfb99251412a50000b30", "gender": 0, "id": 1599259, "name": "Phillip Law", "order": 47}, {"cast_id": 206, "character": "Head of Nation", "credit_id": "56fddfef9251412a54000ba8", "gender": 2, "id": 1599260, "name": "Wai Wong", "order": 48}, {"cast_id": 207, "character": "Head of Nation", "credit_id": "56fde017925141 2a47000ac2", "gender": 2, "id": 1477143, "name": "Joseph Balderrama", "order": 49}, {"cast_id": 208, "character": "Aide", "credit_id": "56fde0809251416cc20000cb", "gender": 2, "id": 1599261, "name": "Eiji Mihara", "order": 50}, {"cast_id": 209, "character": "Aide", "credit_id": "56fde0a49251412a47000ae0", "gender": 2, "id": 1168129, "name": "Junichi Kajioka", "order": 51}, {"cast_id": 40, "character": "Clinic Barman", "credit_id": "56486850c3a368260c006ecf", "gender": 0, "id": 48393, "name": "Victor Schef\\u00e9", "order": 52}, {"cast_id": 210, "character": "Clinic Security Guard", "credit_id": "56fde0f19251412a42000c28", "gender": 2, "id": 1437333, "name": "Harald Windisch", "order": 53}, {"cast_id": 211, "character": "Clinic Security Guard", "credit_id": "56fde118c3a36821a4000c3c", "gender": 2, "id": 1599262, "name": "Tristan Matthiae", "order": 54}, {"cast_id": 177, "character": "Cable Car Heavy", "credit_id": "5696db21c3a3683af90002d6", "gender": 2, "id": 14637, "name": "Detlef Bothe", "order": 55}, {"cast_id": 212, "character": "Cable Car Heavy", "credit_id": "56fde128c3a368219d000b4f", "gender": 2, "id": 1599263, "name": "Bodo Friesecke", "order": 56}, {"cast_id": 213, "character": "Syringe Heavy", "credit_id": "56fde1769251412a54000c0b", "gender": 2, "id": 1599264, "name": "Wilhem Iben", "order": 57}, {"cast_id": 214, "character": "Snowboarder", "credit_id": "56fde1849251412a42000c4d", "gender": 1, "id": 1599265, "name": "Noemi Krausz", "order": 58}, {"cast_id": 215, "character": "Snowboarder", "credit_id": "56fde18dc3a36821b6000c4f", "gender": 2, "id": 1599266, "name": "Noah Saavedra", "order": 59}, {"cast_id": 216, "character": "Snowboarder", "credit_id": "56fde1989251412a54000c16", "gender": 0, "id": 1599267, "name": "Francis Attakpah", "order": 60}, {"cast_id": 217, "character": "Snowboarder", "credit_id": "56fde1a39251412a42000c58", "gender": 2, "id": 1599268, "name": "Michael Glantschnig", "order": 61}, {"cast_id": 218, "character": "Snowboarder", "credit_id": "56fde1aec3a36821ab000cbf", "gender": 2, "id": 1599269, "name": "Marlon Boess", "order": 62}, {"cast_id": 219, "character": "Snowboarder", "credit_id": "56fde1b8c3a36821ae000c62", "gender": 1, "id": 1599270, "name": "Marie Wohlmuth", "order": 63}, {"cast_id": 220, "character": "Snowboarder", "credit_id": "56fde1c1c3a36821b1000c9c", "gender": 1, "id": 1599271, "name": "Lili Epply", "order": 64}, {"cast_id": 221, "character": "Snowboarder", "credit_id": "56fde1cbc3a36821a7000baf", "gender": 2, "id": 1599272, "name": "Konstantin Gerlach", "order": 65}, {"cast_id": 222, "character": "L\'Americain Manater", "credit_id": "56fde24c9251412a45000bfb", "gender": 1, "id": 1102369, "name": "Lara Parmiani", "order": 66}, {"cast_id": 223, "character": "Train Guard", "credit_id": "56fde2679251412a54000c42", "gender": 2, "id": 1372694, "name": "Umit Ulgen", "order": 67}, {"cast_id": 224, "character": "Train Waiter", "credit_id": "56fde2869251412a50000bf1", "gender": 2, "id": 1363060, "name": "Amra Mallassi", "order": 68}, {"cast_id": 225, "character": "Train Barman", "credit_id": "56fde2969251412a45000c15", "gender": 2, "id": 1599275, "name": "Ziad Abaza", "order": 69}, {"cast_id": 226, "character": "Blofeld\'s Chauffeur", "credit_id": "56fde2a79251412a54000c55", "gender": 2, "id": 1599276, "name": "Walid Mumuni", "order": 70}, {"cast_id": 227, "character": "Blofeld\'s Guard", "credit_id": "56fde2eb9251416cc200015b", "gender": 0, "id": 1599277, "name": "Derek Horsham", "order": 71}, {"cast_id": 228, "character": "Blofeld\'s Guard", "credit_id": "56fde3059251412a47000b6a", "gender": 2, "id": 229320, "name": "Nari Blair-Mangat", "order": 72}, {"cast_id": 229, "character": "Blofeld\'s Butler", "credit_id": "56fde321c3a36821ab000d32", "gender": 0, "id": 1599278, "name": "Michael White", "order": 73}, {"cast_id": 230, "character": "Blofeld\'s Waiter", "credit_id": "56fde34c9251412a56000c21", "gender": 0, "id": 1599279, "name": "Adam McGrady", "order": 74}, {"cast_id": 231, "character": "Blofeld\'s Analyst", "credit_id": "56fde36d9251412a50000c2b", "gender": 2, "id": 1436816, "name": "Nader Dernaika", "order": 75}, {"cast_id": 232, "character": "Blofeld\'s Chief Analyst", "credit_id": "56fde37f9251412a4b000cf7", "gender": 2, "id": 1599280, "name": "Pezhmaan Alinia", "order": 76}, {"cast_id": 31, "character": "M (uncredited)", "credit_id": "562f218fc3a3681b5200c9e5", "gender": 1, "id": 5309, "name": "Judi Dench", "order": 77}, {"cast_id": 26, "character": "Clinic Patron (uncredited)", "credit_id": "5517b4bbc3a3686302002f36", "gender": 1, "id": 1186597, "name": "Neve Gachev", "order": 78}, {"cast_id": 245, "character": "Businessman (uncredited)", "credit_id": "57e11a3e92514104570052ba", "gender": 2, "id": 1654739, "name": "Karl Farrer", "order": 79}, {"cast_id": 246, "character": "Lair Staff (uncredited)", "credit_id": "57e124d992514104530055ec", "gender": 1, "id": 1682535, "name": "Kim Adis", "order": 80}, {"cast_id": 247, "character": "Mexican Sprite (uncredited)", "credit_id": "57e125ccc3a36853f2005365", "gender": 1, "id": 1682536, "name": "Maurisa Selene Coleman", "order": 81}, {"cast_id": 249, "character": "Pilot David", "credit_id": "585ab69fc3a3682fb8091219", "gender": 2, "id": 1725801, "name": "Matija Mondi Matovi\\u0107", "order": 82}]',
        ...,
        '[{"cast_id": 8, "character": "Oliver O\\u2019Toole", "credit_id": "52fe4df3c3a36847f8275ed3", "gender": 2, "id": 11867, "name": "Eric Mabius", "order": 0}, {"cast_id": 9, "character": "Shane McInerney", "credit_id": "52fe4df3c3a36847f8275ed7", "gender": 1, "id": 34915, "name": "Kristin Booth", "order": 1}, {"cast_id": 11, "character": "Rita Haywith", "credit_id": "52fe4df3c3a36847f8275edf", "gender": 1, "id": 58393, "name": "Crystal Lowe", "order": 2}, {"cast_id": 12, "character": "Norman Dorman", "credit_id": "52fe4df3c3a36847f8275ee3", "gender": 0, "id": 119810, "name": "Geoff Gustafson", "order": 3}, {"cast_id": 14, "character": "Charlie Riggs", "credit_id": "52fe4df3c3a36847f8275eeb", "gender": 2, "id": 77865, "name": "Benjamin Hollingsworth", "order": 4}, {"cast_id": 13, "character": "Kelly", "credit_id": "52fe4df3c3a36847f8275ee7", "gender": 1, "id": 1037844, "name": "Laci J Mailey", "order": 5}, {"cast_id": 10, "character": "Andrea Shmeckle", "credit_id": "52fe4df3c3a36847f8275edb", "gender": 1, "id": 14668, "name": "Daphne Zuniga", "order": 6}]',
        '[{"cast_id": 3, "character": "Sam", "credit_id": "52fe4ad9c3a368484e16a375", "gender": 2, "id": 82093, "name": "Daniel Henney", "order": 0}, {"cast_id": 4, "character": "Amanda", "credit_id": "52fe4ad9c3a368484e16a379", "gender": 1, "id": 210050, "name": "Eliza Coupe", "order": 1}, {"cast_id": 5, "character": "Donald", "credit_id": "52fe4ad9c3a368484e16a37d", "gender": 2, "id": 2053, "name": "Bill Paxton", "order": 2}, {"cast_id": 6, "character": "Marcus Groff", "credit_id": "52fe4ad9c3a368484e16a381", "gender": 2, "id": 2394, "name": "Alan Ruck", "order": 3}, {"cast_id": 7, "character": "", "credit_id": "548347219251416e7b0081f7", "gender": 0, "id": 1395490, "name": "Zhu Shimao", "order": 4}]',
        '[{"cast_id": 3, "character": "Herself", "credit_id": "52fe44e8c3a368484e03da91", "gender": 1, "id": 69597, "name": "Drew Barrymore", "order": 0}, {"cast_id": 5, "character": "Himself", "credit_id": "58ce01169251415a3901648f", "gender": 2, "id": 85563, "name": "Brian Herzlinger", "order": 1}, {"cast_id": 6, "character": "Himself", "credit_id": "58ce01339251415a410167f0", "gender": 2, "id": 3034, "name": "Corey Feldman", "order": 2}, {"cast_id": 8, "character": "Himself", "credit_id": "58ce018c9251415a7d016e36", "gender": 2, "id": 21315, "name": "Eric Roberts", "order": 3}, {"cast_id": 9, "character": "Himself", "credit_id": "58ce01b99251415a7d016e7d", "gender": 0,

"id": 2171, "name": "Griffin Dunne", "order": 4}, {"cast_id": 10, "character": "Himself", "credit_id": "58ce01d19
251415a8b0168be", "gender": 2, "id": 2231, "name": "Samuel L. Jackson", "order": 5}, {"cast_id": 11, "character":
"Himself", "credit_id": "58ce01dd9251415a39016580", "gender": 2, "id": 14407, "name": "Matt LeBlanc", "order": 6}
, {"cast_id": 7, "character": "Himself", "credit_id": "58ce0164c3a3685104015b28", "gender": 2, "id": 155007, "nam
e": "Bill D\'Elia", "order": 7}]'],
      dtype=object)

In [6]:
```python
movies=movies.merge(credits,on='title') #Merging two data sets can be done by cocat also
```

In [7]:
```python
movies.head(1)
```

Out[7]:

	budget	genres	homepage	id	keywords	original_language	original_title	overview	popularity	production_compan
0	237000000	[{"id": 28, "name": "Action"}, {"id": 12, "nam...	http://www.avatarmovie.com/	19995	[{"id": 1463, "name": "culture clash"}, {"id":...	en	Avatar	In the 22nd century, a paraplegic Marine is di...	150.437577	[{"name": "Ingeni Film Partners", " 28

1 rows × 23 columns

In [8]:
```python
movies.original_language.value_counts()
```

Out[8]:
```
en    4510
fr      70
es      32
zh      27
de      27
hi      19
ja      16
it      14
ko      12
cn      12
ru      11
pt       9
da       7
sv       5
nl       4
fa       4
th       3
he       3
ta       2
cs       2
ro       2
id       2
ar       2
vi       1
sl       1
ps       1
no       1
ky       1
hu       1
pl       1
af       1
nb       1
tr       1
is       1
xx       1
te       1
el       1
Name: original_language, dtype: int64
```

In [9]:
```python
movies.columns
```

Out[9]:
```
Index(['budget', 'genres', 'homepage', 'id', 'keywords', 'original_language',
       'original_title', 'overview', 'popularity', 'production_companies',
       'production_countries', 'release_date', 'revenue', 'runtime',
       'spoken_languages', 'status', 'tagline', 'title', 'vote_average',
       'vote_count', 'movie_id', 'cast', 'crew'],
      dtype='object')
```

In [10]:
```python
movies=movies[['movie_id','title','overview','genres','keywords','cast','crew']]
movies.head(1)
```

	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...

```python
movies.isna().sum()
```

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

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

```python
movies.head(1)
```

	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...

```python
movies.isna().sum()
```

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

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

```
0
```

```python
movies.iloc[0].genres
```

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

```python
#We want a lits of genres name out of dict
import ast
def convert(obj):
    L=[]
    for i in ast.literal_eval(obj): #To covert string to list
        L.append(i['name'])
    return L
```

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

```python
movies.head(1)
```

	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... |

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

In [21]:
```python
movies.head(1)
```

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]	[culture clash, future, space war, space colon...	[{"cast_id": 242, "character": "Jake Sully", "...	[{"credit_id": "52fe48009251416c750aca23", "de...

In [22]:
```python
movies.cast.values
```

Out[22]:
array(['[{"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": "52fe480092514 16c750ac9cb", "gender": 1, "id": 8691, "name": "Zoe Saldana", "order": 1}, {"cast_id": 25, "character": "Dr. Grac e Augustine", "credit_id": "52fe48009251416c750aca39", "gender": 1, "id": 10205, "name": "Sigourney Weaver", "ord er": 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": "52fe4800925 1416c750ac9d3", "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": 59 231, "name": "Joel David Moore", "order": 6}, {"cast_id": 9, "character": "Moat", "credit_id": "52fe48009251416c7 50ac9e5", "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": 1 0, "character": "Tsu\'Tey", "credit_id": "52fe48009251416c750ac9e9", "gender": 2, "id": 10964, "name": "Laz Alons o", "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, "charac ter": "Private Fike", "credit_id": "52fe48009251416c750aca5b", "gender": 2, "id": 154153, "name": "Sean Anthony M oran", "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 C hief", "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": 98673 4, "name": "Kelly Kilgour", "order": 15}, {"cast_id": 36, "character": "Shuttle Pilot", "credit_id": "52fe4800925 1416c750aca6b", "gender": 0, "id": 1207227, "name": "James Patrick Pitt", "order": 16}, {"cast_id": 37, "characte r": "Shuttle Co-Pilot", "credit_id": "52fe48009251416c750aca6f", "gender": 0, "id": 1180936, "name": "Sean Patric k Murphy", "order": 17}, {"cast_id": 38, "character": "Shuttle Crew Chief", "credit_id": "52fe48009251416c750aca7 3", "gender": 2, "id": 1019578, "name": "Peter Dillon", "order": 18}, {"cast_id": 39, "character": "Tractor Opera tor / 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", "cr edit_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": 215 913, "name": "Jacob Tomuri", "order": 22}, {"cast_id": 43, "character": "Suit #1", "credit_id": "52fe48009251416c 750aca87", "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": 104 8610, "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, "c haracter": "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", "g ender": 0, "id": 1207247, "name": "Sonia Yee", "order": 28}, {"cast_id": 50, "character": "Basketball Avatar / Tr oupe", "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": "52fe4800925 1416c750acaa7", "gender": 0, "id": 1207249, "name": "Kyla Warren", "order": 31}, {"cast_id": 53, "character": "Tr oupe", "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", "ge nder": 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, "ch aracter": "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": "52fe48009251416 c750acac3", "gender": 0, "id": 1207257, "name": "Cullen B. Madden", "order": 38}, {"cast_id": 60, "character": "T roupe", "credit_id": "52fe48009251416c750acac7", "gender": 0, "id": 1207259, "name": "Joseph Brady Madden", "orde r": 39}, {"cast_id": 61, "character": "Troupe", "credit_id": "52fe48009251416c750acacb", "gender": 0, "id": 12072 62, "name": "Frankie Torres", "order": 40}, {"cast_id": 62, "character": "Troupe", "credit_id": "52fe48009251416c 750acacf", "gender": 1, "id": 1158600, "name": "Austin Wilson", "order": 41}, {"cast_id": 63, "character": "Troup e", "credit_id": "52fe48019251416c750acad3", "gender": 1, "id": 983705, "name": "Sara Wilson", "order": 42}, {"ca st_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": "52fe480192 51416c750acadb", "gender": 1, "id": 1145098, "name": "Lucy Briant", "order": 44}, {"cast_id": 66, "character": "O p Center Staff", "credit_id": "52fe48019251416c750acadf", "gender": 2, "id": 33305, "name": "Nathan Meister", "or der": 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": "52fe48019251416c750acaeb", "gender": 0, "id": 1207265, "name": "Paul Yates", "order": 48}, {"cast_id": 70, "character": "Op Center Duty Officer", "credit_id": "52fe48019251416c750acaef", "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": "Rhian 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}]',
    '[{"cast_id": 4, "character": "Captain Jack Sparrow", "credit_id": "52fe4232c3a36847f800b50d", "gender": 2, "id": 85, "name": "Johnny Depp", "order": 0}, {"cast_id": 5, "character": "Will Turner", "credit_id": "52fe4232c3a36847f800b511", "gender": 2, "id": 114, "name": "Orlando Bloom", "order": 1}, {"cast_id": 6, "character": "Elizabeth Swann", "credit_id": "52fe4232c3a36847f800b515", "gender": 1, "id": 116, "name": "Keira Knightley", "order": 2}, {"cast_id": 12, "character": "William \\"Bootstrap Bill\\\" Turner", "credit_id": "52fe4232c3a36847f800b52d", "gender": 2, "id": 1640, "name": "Stellan Skarsg\\u00e5rd", "order": 3}, {"cast_id": 10, "character": "Captain Sao Feng", "credit_id": "52fe4232c3a36847f800b525", "gender": 2, "id": 1619, "name": "Chow Yun-fat", "order": 4}, {"cast_id": 9, "character": "Captain Davy Jones", "credit_id": "52fe4232c3a36847f800b521", "gender": 2, "id": 2440, "name": "Bill Nighy", "order": 5}, {"cast_id": 7, "character": "Captain Hector Barbossa", "credit_id": "52fe4232c3a36847f800b519", "gender": 2, "id": 118, "name": "Geoffrey Rush", "order": 6}, {"cast_id": 14, "character": "Admiral James Norrington", "credit_id": "52fe4232c3a36847f800b535", "gender": 2, "id": 1709, "name": "Jack Davenport", "order": 7}, {"cast_id": 13, "character": "Joshamee Gibbs", "credit_id": "52fe4232c3a36847f800b531", "gender": 2, "id": 2449, "name": "Kevin McNally", "order": 8}, {"cast_id": 11, "character": "Lord Cutler Beckett", "credit_id": "52fe4232c3a36847f800b529", "gender": 2, "id": 2441, "name": "Tom Hollander", "order": 9}, {"cast_id": 19, "character": "Tia Dalma", "credit_id": "52fe4232c3a36847f800b549", "gender": 1, "id": 2038, "name": "Naomie Harris", "order": 10}, {"cast_id": 8, "character": "Governor Weatherby Swann", "credit_id": "52fe4232c3a36847f800b51d", "gender": 2, "id": 378, "name": "Jonathan Pryce", "order": 11}, {"cast_id": 37, "character": "Captain Teague Sparrow", "credit_id": "52fe4232c3a36847f800b5b3", "gender": 2, "id": 1430, "name": "Keith Richards", "order": 12}, {"cast_id": 16, "character": "Pintel", "credit_id": "52fe4232c3a36847f800b53d", "gender": 2, "id": 1710, "name": "Lee Arenberg", "order": 13}, {"cast_id": 15, "character": "Ragetti", "credit_id": "52fe4232c3a36847f800b539", "gender": 2, "id": 1711, "name": "Mackenzie Crook", "order": 14}, {"cast_id": 18, "character": "Lieutenant Theodore Groves", "credit_id": "52fe4232c3a36847f800b545", "gender": 2, "id": 4031, "name": "Greg Ellis", "order": 15}, {"cast_id": 55, "character": "Cotton", "credit_id": "57e28d2ec3a3681a01005b5c", "gender": 2, "id": 1715, "name": "David Bailie", "order": 16}, {"cast_id": 17, "character": "Marty", "credit_id": "52fe4232c3a36847f800b541", "gender": 2, "id": 4030, "name": "Martin Klebba", "order": 17}, {"cast_id": 57, "character": "Ian Mercer", "credit_id": "57e28d78c3a36808b900bf4f", "gender": 0, "id": 939, "name": "David Schofield", "order": 18}, {"cast_id": 62, "character": "Scarlett", "credit_id": "57e28ec5c3a3681a50005855", "gender": 1, "id": 2450, "name": "Lauren Maher", "order": 19}, {"cast_id": 63, "character": "Giselle", "credit_id": "57e28ed692514123f5005635", "gender": 1, "id": 2452, "name": "Vanessa Branch", "order": 20}, {"cast_id": 60, "character": "Mullroy", "credit_id": "57e28db2c3a3681a01005bc7", "gender": 2, "id": 1714, "name": "Angus Barnett", "order": 21}, {"cast_id": 59, "character": "Murtogg", "credit_id": "57e28da192514118f7006008", "gender": 0, "id": 1713, "name": "Giles New", "order": 22}, {"cast_id": 58, "character": "Tai Huang", "credit_id": "57e28d8ec3a3681a01005bab", "gender": 2, "id": 22075, "name": "Reggie Lee", "order": 23}, {"cast_id": 64, "character": "Henry Turner", "credit_id": "57e2911992514115110a6cc", "gender": 2, "id": 61259, "name": "Dominic Scott Kay", "order": 24}, {"cast_id": 39, "character": "Mistress Ching", "credit_id": "52fe4232c3a36847f800b5bd", "gender": 1, "id": 33500, "name": "Takayo Fischer", "order": 25}, {"cast_id": 40, "character": "Lieutenant Greitzer", "credit_id": "52fe4232c3a36847f800b5c1", "gender": 2, "id": 1224149, "name": "David Meunier", "order": 26}, {"cast_id": 49, "character": "Hadras", "credit_id": "56d1871c92514174680010cf", "gender": 2, "id": 429401, "name": "Ho-Kwan Tse", "order": 27}, {"cast_id": 56, "character": "Clacker", "credit_id": "57e28d4b92514125710055cb", "gender": 0, "id": 1123, "name": "Andy Beckwith", "order": 28}, {"cast_id

": 51, "character": "Penrod", "credit_id": "56ec8c14c3a3682260003c53", "gender": 2, "id": 1056117, "name": "Peter Donald Badalamenti II", "order": 29}, {"cast_id": 61, "character": "Cotton\'s Parrot (voice)", "credit_id": "57e2 8dcc9251412463005678", "gender": 2, "id": 21700, "name": "Christopher S. Capp", "order": 30}, {"cast_id": 65, "ch aracter": "Captain Teague", "credit_id": "58bc2a37c3a368663003740b", "gender": 2, "id": 1430, "name": "Keith Rich ards", "order": 31}, {"cast_id": 66, "character": "Captain Jocard", "credit_id": "58bc2a8e925141609e03a179", "gen der": 2, "id": 2603, "name": "Hakeem Kae-Kazim", "order": 32}, {"cast_id": 67, "character": "Captain Ammand", "cr edit_id": "58e2a21ac3a36872af00f9c2", "gender": 0, "id": 70577, "name": "Ghassan Massoud", "order": 33}]',
       '[{"cast_id": 1, "character": "James Bond", "credit_id": "52fe4d22c3a368484e1d8d6b", "gender": 2, "id": 87 84, "name": "Daniel Craig", "order": 0}, {"cast_id": 14, "character": "Blofeld", "credit_id": "54805866c3a36829ab 002592", "gender": 2, "id": 27319, "name": "Christoph Waltz", "order": 1}, {"cast_id": 13, "character": "Madelein e", "credit_id": "546f934fc3a3682f9a002ca5", "gender": 1, "id": 121529, "name": "L\\u00e9a Seydoux", "order": 2}, {"cast_id": 10, "character": "M", "credit_id": "53e86503c3a368399c0031f0", "gender": 2, "id": 5469, "name": "Ralp h Fiennes", "order": 3}, {"cast_id": 17, "character": "Lucia", "credit_id": "54805920c3a36829ae0022c5", "gender": 1, "id": 28782, "name": "Monica Bellucci", "order": 4}, {"cast_id": 8, "character": "Q", "credit_id": "52fe4d22c3 a368484e1d8d87", "gender": 2, "id": 17064, "name": "Ben Whishaw", "order": 5}, {"cast_id": 11, "character": "Mone ypenny", "credit_id": "53e8650cc3a368399c0031f4", "gender": 1, "id": 2038, "name": "Naomie Harris", "order": 6}, {"cast_id": 16, "character": "Hinx", "credit_id": "5480590f9251416e7b0022f4", "gender": 2, "id": 543530, "name": "Dave Bautista", "order": 7}, {"cast_id": 15, "character": "C", "credit_id": "548058ffc3a36829a7002511", "gender" : 2, "id": 125660, "name": "Andrew Scott", "order": 8}, {"cast_id": 19, "character": "Tanner", "credit_id": "5480 593cc3a36829ae0022ca", "gender": 2, "id": 139549, "name": "Rory Kinnear", "order": 9}, {"cast_id": 23, "character ": "Mr. White", "credit_id": "5517b3849251416f0d0032c8", "gender": 2, "id": 2244, "name": "Jesper Christensen", " order": 10}, {"cast_id": 24, "character": "Marco Sciarra", "credit_id": "5517b398c3a36862fc002df6", "gender": 0, "id": 1381616, "name": "Alessandro Cremona", "order": 11}, {"cast_id": 25, "character": "Estrella", "credit_id": "5517b3c49251412c1e00180b", "gender": 1, "id": 929937, "name": "Stephanie Sigman", "order": 12}, {"cast_id": 49, "character": "Mexican Man in Lift", "credit_id": "5664561cc3a3682bd0007fc8", "gender": 0, "id": 87265, "name": "T enoch Huerta", "order": 13}, {"cast_id": 50, "character": "Mexican Woman in Lift", "credit_id": "5664561dc3a3682b e9008dc6", "gender": 1, "id": 983710, "name": "Adriana Paz", "order": 14}, {"cast_id": 180, "character": "Gallo", "credit_id": "56fddbacc3a36821b6000aca", "gender": 2, "id": 1385592, "name": "Domenico Fortunato", "order": 15}, {"cast_id": 51, "character": "Gallo\'s Accomplice", "credit_id": "5664561dc3a3682bd8008fbe", "gender": 0, "id": 1 545546, "name": "Marco Zingaro", "order": 16}, {"cast_id": 52, "character": "Gallo\'s Accomplice", "credit_id": " 5664561d9251412ad300870b", "gender": 2, "id": 1545547, "name": "Stefano Elfi DiClaudia", "order": 17}, {"cast_id" : 181, "character": "Q\'s Assistant", "credit_id": "56fddc319251412a4d0009f9", "gender": 0, "id": 122556, "name": "Ian Bonar", "order": 18}, {"cast_id": 182, "character": "Moneypenny\'s Boyfriend", "credit_id": "56fddc66c3a3682 1b6000b04", "gender": 2, "id": 1599239, "name": "Tam Williams", "order": 19}, {"cast_id": 183, "character": "Blof eld\'s London Helicopter Pilot", "credit_id": "56fddc989251412a56000a85", "gender": 2, "id": 1599241, "name": "Ri chard Banham", "order": 20}, {"cast_id": 184, "character": "SC019 Police Commander", "credit_id": "56fddccf925141 2a4b000b5d", "gender": 2, "id": 230681, "name": "Pip Carter", "order": 21}, {"cast_id": 185, "character": "SC019 Police Officer", "credit_id": "56fddce6c3a36821ae000b45", "gender": 2, "id": 30446, "name": "Simon Lenagan", "ord er": 22}, {"cast_id": 27, "character": "Priest", "credit_id": "5517b4f69251416efe003370", "gender": 2, "id": 4837 7, "name": "Alessandro Bressanello", "order": 23}, {"cast_id": 38, "character": "Moreau", "credit_id": "56486811c 3a3682605006adc", "gender": 0, "id": 578512, "name": "Marc Zinga", "order": 24}, {"cast_id": 39, "character": "Vo gel", "credit_id": "56486820c3a368261a006de7", "gender": 0, "id": 1093974, "name": "Brigitte Millar", "order": 25 }, {"cast_id": 41, "character": "Abrika", "credit_id": "564868a09251413ad7006da5", "gender": 2, "id": 89624, "nam e": "Adel Bencherif", "order": 26}, {"cast_id": 186, "character": "Blofeld\'s Right Hand Man", "credit_id": "56fd dd5a9251412a42000b52", "gender": 2, "id": 1129783, "name": "Gediminas Adomaitis", "order": 27}, {"cast_id": 42, " character": "Lorenzo", "credit_id": "564868c5c3a368260c006ee4", "gender": 0, "id": 226388, "name": "Peppe Lanzett a", "order": 28}, {"cast_id": 187, "character": "Francesco", "credit_id": "56fddd9cc3a36821a4000b5b", "gender": 2 , "id": 1299312, "name": "Francesco Arca", "order": 29}, {"cast_id": 188, "character": "Marco", "credit_id": "56f ddddb79251412a50000aad", "gender": 2, "id": 124628, "name": "Matteo Taranto", "order": 30}, {"cast_id": 189, "char acter": "Palazzo Security", "credit_id": "56fdddd69251412a56000ac6", "gender": 2, "id": 1562103, "name": "Emilio Aniba", "order": 31}, {"cast_id": 190, "character": "Guerra", "credit_id": "56fdddf19251412a50000ac1", "gender": 2, "id": 1474616, "name": "Benito Sagredo", "order": 32}, {"cast_id": 191, "character": "Businessman", "credit_id ": "56fdde0d9251416cc200004c", "gender": 2, "id": 1247684, "name": "Dai Tabuchi", "order": 33}, {"cast_id": 192, "character": "Businessman", "credit_id": "56fdde1bc3a36821b1000ba8", "gender": 2, "id": 1599245, "name": "George Lasha", "order": 34}, {"cast_id": 193, "character": "Businessman", "credit_id": "56fdde259251412a42000b93", "gend er": 0, "id": 1599247, "name": "Sargon Yelda", "order": 35}, {"cast_id": 194, "character": "Businessman", "credit _id": "56fdde9fc3a36821b6000ba0", "gender": 2, "id": 1599251, "name": "Andy Cheung", "order": 36}, {"cast_id": 19 5, "character": "Marshall", "credit_id": "56fddeb29251412a50000af4", "gender": 2, "id": 1599252, "name": "Erick H ayden", "order": 37}, {"cast_id": 196, "character": "Valerian", "credit_id": "56fddec8c3a36821ab000c19", "gender" : 2, "id": 1599253, "name": "Oleg Mirochnikov", "order": 38}, {"cast_id": 197, "character": "Fiat Driver", "credi t_id": "56fddef89251412a47000a79", "gender": 2, "id": 123548, "name": "Antonio Salines", "order": 39}, {"cast_id" : 198, "character": "Street Sweeper", "credit_id": "56fddf109251416cc2000086", "gender": 2, "id": 1599254, "name" : "Miloud Mourad Benamara", "order": 40}, {"cast_id": 199, "character": "Chairman", "credit_id": "56fddf4a9251412 a42000bcf", "gender": 2, "id": 1599255, "name": "Gido Schimanski", "order": 41}, {"cast_id": 200, "character": "H ead of Nation", "credit_id": "56fddf69c3a36821a7000b12", "gender": 0, "id": 1404403, "name": "Nigel Barber", "ord er": 42}, {"cast_id": 201, "character": "Head of Nation", "credit_id": "56fddf7fc3a36821b1000c0d", "gender": 2, " id": 74415, "name": "Patrice Naiambana", "order": 43}, {"cast_id": 202, "character": "Head of Nation", "credit_id ": "56fddf89c3a368219d000adb", "gender": 2, "id": 1599256, "name": "Stephane Cornicard", "order": 44}, {"cast_id" : 203, "character": "Head of Nation", "credit_id": "56fddf929251412a56000b29", "gender": 2, "id": 1599257, "name" : "Gary Fannin", "order": 45}, {"cast_id": 204, "character": "Head of Nation", "credit_id": "56fddf9cc3a36821ae00 0be1", "gender": 2, "id": 1599258, "name": "Sadao Ueda", "order": 46}, {"cast_id": 205, "character": "Head of Nat ion", "credit_id": "56fddfb99251412a50000b30", "gender": 0, "id": 1599259, "name": "Phillip Law", "order": 47}, { "cast_id": 206, "character": "Head of Nation", "credit_id": "56fddfef9251412a54000ba8", "gender": 2, "id": 159926 0, "name": "Wai Wong", "order": 48}, {"cast_id": 207, "character": "Head of Nation", "credit_id": "56fde017925141 2a47000ac2", "gender": 2, "id": 1477143, "name": "Joseph Balderrama", "order": 49}, {"cast_id": 208, "character": "Aide", "credit_id": "56fde0809251416cc20000cb", "gender": 2, "id": 1599261, "name": "Eiji Mihara", "order": 50}, {"cast_id": 209, "character": "Aide", "credit_id": "56fde0a49251412a47000ae0", "gender": 2, "id": 1168129, "name ": "Junichi Kajioka", "order": 51}, {"cast_id": 40, "character": "Clinic Barman", "credit_id": "56486850c3a368260c 006ecf", "gender": 0, "id": 48393, "name": "Victor Schef\\u00e9", "order": 52}, {"cast_id": 210, "character": "Cl inic Security Guard", "credit_id": "56fde0f19251412a42000c28", "gender": 2, "id": 1437333, "name": "Harald Windis ch", "order": 53}, {"cast_id": 211, "character": "Clinic Security Guard", "credit_id": "56fde118c3a36821a4000c3c" , "gender": 2, "id": 1599262, "name": "Tristan Matthiae", "order": 54}, {"cast_id": 177, "character": "Cable Car Heavy", "credit_id": "5696db21c3a3683af90002d6", "gender": 2, "id": 14637, "name": "Detlef Bothe", "order": 55}, {"cast_id": 212, "character": "Cable Car Heavy", "credit_id": "56fde128c3a368219d000b4f", "gender": 2, "id": 1599 263, "name": "Bodo Friesecke", "order": 56}, {"cast_id": 213, "character": "Syringe Heavy", "credit_id": "56fde17 69251412a54000c0b", "gender": 2, "id": 1599264, "name": "Wilhem Iben", "order": 57}, {"cast_id": 214, "character" : "Snowboarder", "credit_id": "56fde1849251412a42000c4d", "gender": 1, "id": 1599265, "name": "Noemi Krausz", "or der": 58}, {"cast_id": 215, "character": "Snowboarder", "credit_id": "56fde18dc3a36821b6000c4f", "gender": 2, "id ": 1599266, "name": "Noah Saavedra", "order": 59}, {"cast_id": 216, "character": "Snowboarder", "credit_id": "56f

de1989251412a54000c16", "gender": 0, "id": 1599267, "name": "Francis Attakpah", "order": 60}, {"cast_id": 217, "character": "Snowboarder", "credit_id": "56fde1a39251412a42000c58", "gender": 2, "id": 1599268, "name": "Michael Glantschnig", "order": 61}, {"cast_id": 218, "character": "Snowboarder", "credit_id": "56fde1aec3a36821ab000cbf", "gender": 2, "id": 1599269, "name": "Marlon Boess", "order": 62}, {"cast_id": 219, "character": "Snowboarder", "credit_id": "56fde1b8c3a36821ae000c62", "gender": 1, "id": 1599270, "name": "Marie Wohlmuth", "order": 63}, {"cast_id": 220, "character": "Snowboarder", "credit_id": "56fde1c1c3a36821b1000c9c", "gender": 1, "id": 1599271, "name": "Lili Epply", "order": 64}, {"cast_id": 221, "character": "Snowboarder", "credit_id": "56fde1cbc3a36821a7000baf", "gender": 2, "id": 1599272, "name": "Konstantin Gerlach", "order": 65}, {"cast_id": 222, "character": "L\'Americain Manater", "credit_id": "56fde24c9251412a45000bfb", "gender": 1, "id": 1102369, "name": "Lara Parmiani", "order": 66}, {"cast_id": 223, "character": "Train Guard", "credit_id": "56fde2679251412a54000c42", "gender": 2, "id": 1372694, "name": "Umit Ulgen", "order": 67}, {"cast_id": 224, "character": "Train Waiter", "credit_id": "56fde2869251412a50000bf1", "gender": 2, "id": 1363060, "name": "Amra Mallassi", "order": 68}, {"cast_id": 225, "character": "Train Barman", "credit_id": "56fde2969251412a45000c15", "gender": 2, "id": 1599275, "name": "Ziad Abaza", "order": 69}, {"cast_id": 226, "character": "Blofeld\'s Chauffeur", "credit_id": "56fde2a79251412a54000c55", "gender": 2, "id": 1599276, "name": "Walid Mumuni", "order": 70}, {"cast_id": 227, "character": "Blofeld\'s Guard", "credit_id": "56fde2eb9251416cc200015b", "gender": 0, "id": 1599277, "name": "Derek Horsham", "order": 71}, {"cast_id": 228, "character": "Blofeld\'s Guard", "credit_id": "56fde3059251412a47000b6a", "gender": 2, "id": 229320, "name": "Nari Blair-Mangat", "order": 72}, {"cast_id": 229, "character": "Blofeld\'s Butler", "credit_id": "56fde321c3a36821ab000d32", "gender": 0, "id": 1599278, "name": "Michael White", "order": 73}, {"cast_id": 230, "character": "Blofeld\'s Waiter", "credit_id": "56fde34c9251412a56000c21", "gender": 0, "id": 1599279, "name": "Adam McGrady", "order": 74}, {"cast_id": 231, "character": "Blofeld\'s Analyst", "credit_id": "56fde36d9251412a50000c2b", "gender": 2, "id": 1436816, "name": "Nader Dernaika", "order": 75}, {"cast_id": 232, "character": "Blofeld\'s Chief Analyst", "credit_id": "56fde37f9251412a4b000cf7", "gender": 2, "id": 1599280, "name": "Pezhmaan Alinia", "order": 76}, {"cast_id": 31, "character": "M (uncredited)", "credit_id": "562f218fc3a3681b5200c9e5", "gender": 1, "id": 5309, "name": "Judi Dench", "order": 77}, {"cast_id": 26, "character": "Clinic Patron (uncredited)", "credit_id": "5517b4bbc3a3686302002f36", "gender": 1, "id": 1186597, "name": "Neve Gachev", "order": 78}, {"cast_id": 245, "character": "Businessman (uncredited)", "credit_id": "57e11a3e92514104570052ba", "gender": 2, "id": 1654739, "name": "Karl Farrer", "order": 79}, {"cast_id": 246, "character": "Lair Staff (uncredited)", "credit_id": "57e124d992514104530055ec", "gender": 1, "id": 1682535, "name": "Kim Adis", "order": 80}, {"cast_id": 247, "character": "Mexican Sprite (uncredited)", "credit_id": "57e125ccc3a36853f2005365", "gender": 1, "id": 1682536, "name": "Maurisa Selene Coleman", "order": 81}, {"cast_id": 249, "character": "Pilot David", "credit_id": "585ab69fc3a3682fb8091219", "gender": 2, "id": 1725801, "name": "Matija Mondi Matovi\\u0107", "order": 82}]',
       ...,
       '[{"cast_id": 8, "character": "Oliver O\\u2019Toole", "credit_id": "52fe4df3c3a36847f8275ed3", "gender": 2, "id": 11867, "name": "Eric Mabius", "order": 0}, {"cast_id": 9, "character": "Shane McInerney", "credit_id": "52fe4df3c3a36847f8275ed7", "gender": 1, "id": 34915, "name": "Kristin Booth", "order": 1}, {"cast_id": 11, "character": "Rita Haywith", "credit_id": "52fe4df3c3a36847f8275edf", "gender": 1, "id": 58393, "name": "Crystal Lowe", "order": 2}, {"cast_id": 12, "character": "Norman Dorman", "credit_id": "52fe4df3c3a36847f8275ee3", "gender": 0, "id": 119810, "name": "Geoff Gustafson", "order": 3}, {"cast_id": 14, "character": "Charlie Riggs", "credit_id": "52fe4df3c3a36847f8275eeb", "gender": 2, "id": 77865, "name": "Benjamin Hollingsworth", "order": 4}, {"cast_id": 13, "character": "Kelly", "credit_id": "52fe4df3c3a36847f8275ee7", "gender": 1, "id": 1037844, "name": "Laci J Mailey", "order": 5}, {"cast_id": 10, "character": "Andrea Shmeckle", "credit_id": "52fe4df3c3a36847f8275edb", "gender": 1, "id": 14668, "name": "Daphne Zuniga", "order": 6}]',
       '[{"cast_id": 3, "character": "Sam", "credit_id": "52fe4ad9c3a368484e16a375", "gender": 2, "id": 82093, "name": "Daniel Henney", "order": 0}, {"cast_id": 4, "character": "Amanda", "credit_id": "52fe4ad9c3a368484e16a379", "gender": 1, "id": 210050, "name": "Eliza Coupe", "order": 1}, {"cast_id": 5, "character": "Donald", "credit_id": "52fe4ad9c3a368484e16a37d", "gender": 2, "id": 2053, "name": "Bill Paxton", "order": 2}, {"cast_id": 6, "character": "Marcus Groff", "credit_id": "52fe4ad9c3a368484e16a381", "gender": 2, "id": 2394, "name": "Alan Ruck", "order": 3}, {"cast_id": 7, "character": "", "credit_id": "548347219251416e7b0081f7", "gender": 0, "id": 1395490, "name": "Zhu Shimao", "order": 4}]',
       '[{"cast_id": 3, "character": "Herself", "credit_id": "52fe44e8c3a368484e03da91", "gender": 1, "id": 69597, "name": "Drew Barrymore", "order": 0}, {"cast_id": 5, "character": "Himself", "credit_id": "58ce01169251415a3901648f", "gender": 2, "id": 85563, "name": "Brian Herzlinger", "order": 1}, {"cast_id": 6, "character": "Himself", "credit_id": "58ce01339251415a410167f0", "gender": 2, "id": 3034, "name": "Corey Feldman", "order": 2}, {"cast_id": 8, "character": "Himself", "credit_id": "58ce018c9251415a7d016e36", "gender": 2, "id": 21315, "name": "Eric Roberts", "order": 3}, {"cast_id": 9, "character": "Himself", "credit_id": "58ce01b99251415a7d016e7d", "gender": 0, "id": 2171, "name": "Griffin Dunne", "order": 4}, {"cast_id": 10, "character": "Himself", "credit_id": "58ce01d19251415a8b0168be", "gender": 2, "id": 2231, "name": "Samuel L. Jackson", "order": 5}, {"cast_id": 11, "character": "Himself", "credit_id": "58ce01dd9251415a39016580", "gender": 2, "id": 14407, "name": "Matt LeBlanc", "order": 6}, {"cast_id": 7, "character": "Himself", "credit_id": "58ce0164c3a3685104015b28", "gender": 2, "id": 155007, "name": "Bill D\'Elia", "order": 7}]'],
      dtype=object)

In [23]:
```python
import ast
def convert3(obj): #we want first 3 dict from cast
    L=[]
    count=0
    for i in ast.literal_eval(obj): #To covert string to list
        if count!=3:
            L.append(i['name'])
        else:
            break
    return L
```

In [24]:
```python
movies['cast']=movies['cast'].apply(convert3)
```

In [25]:
```python
movies.cast.values
```

Out[25]:
```
array([list(['Sam Worthington', 'Zoe Saldana', 'Sigourney Weaver', 'Stephen Lang', 'Michelle Rodriguez', 'Giovanni Ribisi', 'Joel David Moore', 'CCH Pounder', 'Wes Studi', 'Laz Alonso', 'Dileep Rao', 'Matt Gerald', 'Sean Anthony Moran', 'Jason Whyte', 'Scott Lawrence', 'Kelly Kilgour', 'James Patrick Pitt', 'Sean Patrick Murphy', 'Peter Dillon', 'Kevin Dorman', 'Kelson Henderson', 'David Van Horn', 'Jacob Tomuri', 'Michael Blain-Rozgay', 'Jon Curry
```

```
', 'Luke Hawker', 'Woody Schultz', 'Peter Mensah', 'Sonia Yee', 'Jahnel Curfman', 'Ilram Choi', 'Kyla Warren', 'L
isa Roumain', 'Debra Wilson', 'Chris Mala', 'Taylor Kibby', 'Jodie Landau', 'Julie Lamm', 'Cullen B. Madden', 'Jo
seph Brady Madden', 'Frankie Torres', 'Austin Wilson', 'Sara Wilson', 'Tamica Washington-Miller', 'Lucy Briant',
'Nathan Meister', 'Gerry Blair', 'Matthew Chamberlain', 'Paul Yates', 'Wray Wilson', 'James Gaylyn', 'Melvin Leno
Clark III', 'Carvon Futrell', 'Brandon Jelkes', 'Micah Moch', 'Hanniyah Muhammad', 'Christopher Nolen', 'Christa
Oliver', 'April Marie Thomas', 'Bravita A. Threatt', 'Colin Bleasdale', 'Mike Bodnar', 'Matt Clayton', 'Nicole Di
onne', 'Jamie Harrison', 'Allan Henry', 'Anthony Ingruber', 'Ashley Jeffery', 'Dean Knowsley', 'Joseph Mika-Hunt'
, 'Terry Notary', 'Kai Pantano', 'Logan Pithyou', 'Stuart Pollock', 'Raja', 'Gareth Ruck', 'Rhian Sheehan', 'T. J
. Storm', 'Jodie Taylor', 'Alicia Vela-Bailey', 'Richard Whiteside', 'Nikie Zambo', 'Julene Renee']),
       list(['Johnny Depp', 'Orlando Bloom', 'Keira Knightley', 'Stellan Skarsgård', 'Chow Yun-fat', 'Bill Nighy'
, 'Geoffrey Rush', 'Jack Davenport', 'Kevin McNally', 'Tom Hollander', 'Naomie Harris', 'Jonathan Pryce', 'Keith
Richards', 'Lee Arenberg', 'Mackenzie Crook', 'Greg Ellis', 'David Bailie', 'Martin Klebba', 'David Schofield', '
Lauren Maher', 'Vanessa Branch', 'Angus Barnett', 'Giles New', 'Reggie Lee', 'Dominic Scott Kay', 'Takayo Fischer
', 'David Meunier', 'Ho-Kwan Tse', 'Andy Beckwith', 'Peter Donald Badalamenti II', 'Christopher S. Capp', 'Keith
Richards', 'Hakeem Kae-Kazim', 'Ghassan Massoud']),
       list(['Daniel Craig', 'Christoph Waltz', 'Léa Seydoux', 'Ralph Fiennes', 'Monica Bellucci', 'Ben Whishaw',
'Naomie Harris', 'Dave Bautista', 'Andrew Scott', 'Rory Kinnear', 'Jesper Christensen', 'Alessandro Cremona', 'St
ephanie Sigman', 'Tenoch Huerta', 'Adriana Paz', 'Domenico Fortunato', 'Marco Zingaro', 'Stefano Elfi DiClaudia',
'Ian Bonar', 'Tam Williams', 'Richard Banham', 'Pip Carter', 'Simon Lenagan', 'Alessandro Bressanello', 'Marc Zin
ga', 'Brigitte Millar', 'Adel Bencherif', 'Gediminas Adomaitis', 'Peppe Lanzetta', 'Francesco Arca', 'Matteo Tara
nto', 'Emilio Aniba', 'Benito Sagredo', 'Dai Tabuchi', 'George Lasha', 'Sargon Yelda', 'Andy Cheung', 'Erick Hayd
en', 'Oleg Mirochnikov', 'Antonio Salines', 'Miloud Mourad Benamara', 'Gido Schimanski', 'Nigel Barber', 'Patrice
Naiambana', 'Stephane Cornicard', 'Gary Fannin', 'Sadao Ueda', 'Phillip Law', 'Wai Wong', 'Joseph Balderrama', 'E
iji Mihara', 'Junichi Kajioka', 'Victor Schefé', 'Harald Windisch', 'Tristan Matthiae', 'Detlef Bothe', 'Bodo Fri
esecke', 'Wilhem Iben', 'Noemi Krausz', 'Noah Saavedra', 'Francis Attakpah', 'Michael Glantschnig', 'Marlon Boess
', 'Marie Wohlmuth', 'Lili Epply', 'Konstantin Gerlach', 'Lara Parmiani', 'Umit Ulgen', 'Amra Mallassi', 'Ziad Ab
aza', 'Walid Mumuni', 'Derek Horsham', 'Nari Blair-Mangat', 'Michael White', 'Adam McGrady', 'Nader Dernaika', 'P
ezhmaan Alinia', 'Judi Dench', 'Neve Gachev', 'Karl Farrer', 'Kim Adis', 'Maurisa Selene Coleman', 'Matija Mondi
Matović']),
       ...,
       list(['Eric Mabius', 'Kristin Booth', 'Crystal Lowe', 'Geoff Gustafson', 'Benjamin Hollingsworth', 'Laci J
Mailey', 'Daphne Zuniga']),
       list(['Daniel Henney', 'Eliza Coupe', 'Bill Paxton', 'Alan Ruck', 'Zhu Shimao']),
       list(['Drew Barrymore', 'Brian Herzlinger', 'Corey Feldman', 'Eric Roberts', 'Griffin Dunne', 'Samuel L. J
ackson', 'Matt LeBlanc', "Bill D'Elia"])],
      dtype=object)
```

In [26]:
```python
import ast
def fetchdirector(obj): #we want director names from crew feature
    L=[]

    for i in ast.literal_eval(obj): #To covert string of list to list
        if i['job']=='Director':
            L.append(i['name'])
            break
    return L
```

In [27]:
```python
movies['crew']=movies['crew'].apply(fetchdirector)
```

In [28]:
```python
movies.crew.value_counts
```

Out[28]:
```
<bound method IndexOpsMixin.value_counts of 0           [James Cameron]
1          [Gore Verbinski]
2             [Sam Mendes]
3       [Christopher Nolan]
4          [Andrew Stanton]
                ...
4804     [Robert Rodriguez]
4805          [Edward Burns]
4806           [Scott Smith]
4807           [Daniel Hsia]
4808       [Brian Herzlinger]
Name: crew, Length: 4806, dtype: object>
```

In [29]:
```python
movies.overview.dtypes
```

Out[29]: dtype('O')

In [30]:
```python
movies.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 4806 entries, 0 to 4808
Data columns (total 7 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   movie_id  4806 non-null   int64
```

```
1   title     4806 non-null   object
2   overview  4806 non-null   object
3   genres    4806 non-null   object
4   keywords  4806 non-null   object
5   cast      4806 non-null   object
6   crew      4806 non-null   object
dtypes: int64(1), object(6)
memory usage: 300.4+ KB
```

In [31]:
```python
#To covert string to list
def overview(obj):
    L=obj.split()
    return L
```

In [32]:
```python
movies['overview']=movies['overview'].apply(overview)
```

In [33]:
```python
movies.head(1)
```

Out[33]:

	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 Weave...	[James Cameron]

In [34]:
```python
movies['genres']=movies['genres'].apply(lambda x:[i.replace(" ","") for i in x])
```

In [35]:
```python
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])
```

In [36]:
```python
movies['tags']=movies['overview']+movies['genres']+movies['keywords']+movies['cast']+movies['crew']
```

In [37]:
```python
movies.head(1)
```

Out[37]:

	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...

In [38]:
```python
new_df=movies[['movie_id','title','tags']]
```

In [39]:
```python
new_df.head()
```

Out[39]:

	movie_id	title	tags
0	19995	Avatar	[In, the, 22nd, century,, a, paraplegic, Marin...
1	285	Pirates of the Caribbean: At World's End	[Captain, Barbossa,, long, believed, to, be, d...
2	206647	Spectre	[A, cryptic, message, from, Bond's, past, send...
3	49026	The Dark Knight Rises	[Following, the, death, of, District, Attorney...
4	49529	John Carter	[John, Carter, is, a, war-weary,, former, mili...

In [40]:
```python
new_df['tags']=new_df['tags'].apply(lambda x:" ".join(x))
```

```
C:\Users\hp\AppData\Local\Temp/ipykernel_12088/487797088.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#retur
ning-a-view-versus-a-copy
  new_df['tags']=new_df['tags'].apply(lambda x:" ".join(x))
```

```
In [41]:   new_df.tags.values
```

```
Out[41]:   array(['In the 22nd century, a paraplegic Marine is dispatched to the moon Pandora on a unique mission, but becom
           es torn between following orders and protecting an alien civilization. Action Adventure Fantasy ScienceFiction cu
           ltureclash future spacewar spacecolony society spacetravel futuristic romance space alien tribe alienplanet cgi m
           arine soldier battle loveaffair antiwar powerrelations mindandsoul 3d SamWorthington ZoeSaldana SigourneyWeaver S
           tephenLang MichelleRodriguez GiovanniRibisi JoelDavidMoore CCHPounder WesStudi LazAlonso DileepRao MattGerald Sea
           nAnthonyMoran JasonWhyte ScottLawrence KellyKilgour JamesPatrickPitt SeanPatrickMurphy PeterDillon KevinDorman Ke
           lsonHenderson DavidVanHorn JacobTomuri MichaelBlain-Rozgay JonCurry LukeHawker WoodySchultz PeterMensah SoniaYee
           JahnelCurfman IlramChoi KylaWarren LisaRoumain DebraWilson ChrisMala TaylorKibby JodieLandau JulieLamm CullenB.Ma
           dden JosephBradyMadden FrankieTorres AustinWilson SaraWilson TamicaWashington-Miller LucyBriant NathanMeister Ger
           ryBlair MatthewChamberlain PaulYates WrayWilson JamesGaylyn MelvinLenoClarkIII CarvonFutrell BrandonJelkes MicahM
           och HanniyahMuhammad ChristopherNolen ChristaOliver AprilMarieThomas BravitaA.Threatt ColinBleasdale MikeBodnar M
           attClayton NicoleDionne JamieHarrison AllanHenry AnthonyIngruber AshleyJeffery DeanKnowsley JosephMika-Hunt Terry
           Notary KaiPantano LoganPithyou StuartPollock Raja GarethRuck RhianSheehan T.J.Storm JodieTaylor AliciaVela-Bailey
           RichardWhiteside NikieZambo JuleneRenee JamesCameron',
                  "Captain Barbossa, long believed to be dead, has come back to life and is headed to the edge of the Earth
           with Will Turner and Elizabeth Swann. But nothing is quite as it seems. Adventure Fantasy Action ocean drugabuse
           exoticisland eastindiatradingcompany loveofone'slife traitor shipwreck strongwoman ship alliance calypso afterlif
           e fighter pirate swashbuckler aftercreditsstinger JohnnyDepp OrlandoBloom KeiraKnightley StellanSkarsgård ChowYun
           -fat BillNighy GeoffreyRush JackDavenport KevinMcNally TomHollander NaomieHarris JonathanPryce KeithRichards LeeA
           renberg MackenzieCrook GregEllis DavidBailie MartinKlebba DavidSchofield LaurenMaher VanessaBranch AngusBarnett G
           ilesNew ReggieLee DominicScottKay TakayoFischer DavidMeunier Ho-KwanTse AndyBeckwith PeterDonaldBadalamentiII Chr
           istopherS.Capp KeithRichards HakeemKae-Kazim GhassanMassoud GoreVerbinski",
                  'A cryptic message from Bond's past sends him on a trail to uncover a sinister organization. While M battl
           es political forces to keep the secret service alive, Bond peels back the layers of deceit to reveal the terrible
           truth behind SPECTRE. Action Adventure Crime spy basedonnovel secretagent sequel mi6 britishsecretservice unitedk
           ingdom DanielCraig ChristophWaltz LéaSeydoux RalphFiennes MonicaBellucci BenWhishaw NaomieHarris DaveBautista And
           rewScott RoryKinnear JesperChristensen AlessandroCremona StephanieSigman TenochHuerta AdrianaPaz DomenicoFortunat
           o MarcoZingaro StefanoElfiDiClaudia IanBonar TamWilliams RichardBanham PipCarter SimonLenagan AlessandroBressanel
           lo MarcZinga BrigitteMillar AdelBencherif GediminasAdomaitis PeppeLanzetta FrancescoArca MatteoTaranto EmilioAnib
           a BenitoSagredo DaiTabuchi GeorgeLasha SargonYelda AndyCheung ErickHayden OlegMirochnikov AntonioSalines MiloudMo
           uradBenamara GidoSchimanski NigelBarber PatriceNaiambana StephaneCornicard GaryFannin SadaoUeda PhillipLaw WaiWon
           g JosephBalderrama EijiMihara JunichiKajioka VictorSchefé HaraldWindisch TristanMatthiae DetlefBothe BodoFrieseck
           e WilhemIben NoemiKrausz NoahSaavedra FrancisAttakpah MichaelGlantschnig MarlonBoess MarieWohlmuth LiliEpply Kons
           tantinGerlach LaraParmiani UmitUlgen AmraMallassi ZiadAbaza WalidMumuni DerekHorsham NariBlair-Mangat MichaelWhit
           e AdamMcGrady NaderDernaika PezhmaanAlinia JudiDench NeveGachev KarlFarrer KimAdis MaurisaSeleneColeman MatijaMon
           diMatović SamMendes',
                  ...,
                  '"Signed, Sealed, Delivered" introduces a dedicated quartet of civil servants in the Dead Letter Office of
           the U.S. Postal System who transform themselves into an elite team of lost-mail detectives. Their determination t
           o deliver the seemingly undeliverable takes them out of the post office into an unpredictable world where letters
           and packages from the past save lives, solve crimes, reunite old loves, and change futures by arriving late, but
           always miraculously on time. Comedy Drama Romance TVMovie date loveatfirstsight narration investigation team post
           alworker EricMabius KristinBooth CrystalLowe GeoffGustafson BenjaminHollingsworth LaciJMailey DaphneZuniga ScottS
           mith',
                  'When ambitious New York attorney Sam is sent to Shanghai on assignment, he immediately stumbles into a le
           gal mess that could end his career. With the help of a beautiful relocation specialist, a well-connected old-time
           r, a clever journalist, and a street-smart legal assistant, Sam might just save his job, find romance, and learn
           to appreciate the beauty and wonders of Shanghai. Written by Anonymous (IMDB.com). DanielHenney ElizaCoupe BillPa
           xton AlanRuck ZhuShimao DanielHsia',
                  "Ever since the second grade when he first saw her in E.T. The Extraterrestrial, Brian Herzlinger has had
           a crush on Drew Barrymore. Now, 20 years later he's decided to try to fulfill his lifelong dream by asking her fo
           r a date. There's one small problem: She's Drew Barrymore and he's, well, Brian Herzlinger, a broke 27-year-old a
           spiring filmmaker from New Jersey. Documentary obsession camcorder crush dreamgirl DrewBarrymore BrianHerzlinger
           CoreyFeldman EricRoberts GriffinDunne SamuelL.Jackson MattLeBlanc BillD'Elia BrianHerzlinger",
                  dtype=object)
```

```
In [42]:   #Coverting to lower case cuz used in general
           new_df['tags']=new_df['tags'].apply(lambda x:x.lower())
```

```
C:\Users\hp\AppData\Local\Temp/ipykernel_12088/2932800441.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#retur
ning-a-view-versus-a-copy
  new_df['tags']=new_df['tags'].apply(lambda x:x.lower())
```

```
In [43]:   new_df.head()
```

```
Out[43]:
```

	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...

Text Vectorization is the process of converting text into numerical representation. Here is some popular methods to accomplish text vectorization: Binary Term Frequency. Bag of Words (BoW) Term Frequency. (L1) Normalized Term Frequency.

In [44]:
```python
from sklearn.feature_extraction.text import CountVectorizer
```

In [45]:
```python
cv=CountVectorizer(max_features=5000,stop_words='english')
#max_features : int, default=None
    #If not None, build a vocabulary that only consider the top
    #max_features ordered by term frequency across the corpus
```

In [47]:
```python
vectors=cv.fit_transform(new_df['tags']).toarray()
```

In [48]:
```python
vectors
```

Out[48]:
```
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]], dtype=int64)
```

In [49]:
```python
vectors.shape
```

Out[49]:
```
(4806, 5000)
```

In [50]:
```python
vectors[0]
```

Out[50]:
```
array([0, 0, 0, ..., 0, 0, 0], dtype=int64)
```

In [52]:
```python
cv.get_feature_names()
```

Out[52]:
```
['000',
 '007',
 '10',
 '100',
 '11',
 '12',
 '13',
 '14',
 '15',
 '16',
 '17',
 '18',
 '18th',
 '19',
 '1930s',
 '1940s',
 '1950s',
 '1960s',
 '1970s',
 '1980s',
 '19th',
 '19thcentury',
 '20',
 '20th',
 '24',
 '25',
 '30',
 '3d',
 '40',
 '50',
 '60',
 '60s',
 'aaron',
 'aaroneckhart',
 'aarontaylor',
 'aasifmandvi',
```

```
'abandoned',
'abducted',
'abigailbreslin',
'ability',
'able',
'abuse',
'abusive',
'academy',
'accept',
'accepts',
'access',
'accident',
'accidentally',
'accompanied',
'account',
'accountant',
'accused',
'ace',
'achieve',
'act',
'action',
'actions',
'activist',
'activities',
'actor',
'actors',
'actress',
'acts',
'actual',
'actually',
'adam',
'adambrody',
'adamgoldberg',
'adamlefevre',
'adams',
'adamsandler',
'adamscott',
'adamshankman',
'adaptation',
'addict',
'addicted',
'addiction',
'adewaleakinnuoye',
'adolescence',
'adopted',
'adoption',
'adrianmartinez',
'adrienbrody',
'adult',
'adultery',
'adulthood',
'adults',
'advantage',
'adventure',
'adventures',
'advertising',
'advice',
'affair',
'affection',
'affections',
'afghanistan',
'africa',
'african',
'aftercreditsstinger',
'afterlife',
'aftermath',
'agbaje',
'age',
'aged',
'agedifference',
'agency',
'agenda',
'agent',
'agents',
'aggressive',
'aging',
'ago',
'agree',
'agrees',
'ahead',
'aid',
'aidanquinn',
'aided',
'aids',
'ailing',
'air',
'airplane',
'airport',
'al',
```

```
'alanalda',
'alanarkin',
'alancumming',
'alanrickman',
'alantudyk',
'alaska',
'albert',
'albertbrooks',
'albertfinney',
'alcohol',
'alcoholic',
'alcoholism',
'alecbaldwin',
'alessandronivola',
'alex',
'alexandersiddig',
'alexapenavega',
'alexborstein',
'alexisbledel',
'alfredhitchcock',
'alfredmolina',
'alfrewoodard',
'ali',
'alice',
'alicedrummond',
'aliciawitt',
'alien',
'alieninvasion',
'alienlife',
'aliens',
'alike',
'alilarter',
'alive',
'allancorduner',
'allen',
'allencovert',
'alleong',
'alliance',
'allied',
'allies',
'allisonjanney',
'allow',
'allows',
'ally',
'alongside',
'alpacino',
'alter',
'alternate',
'alternative',
'alzheimer',
'amanda',
'amandapeet',
'amandaseyfried',
'amazing',
'ambassador',
'amberheard',
'ambition',
'ambitious',
'ambulance',
'ambush',
'america',
'american',
'americanfootball',
'americans',
'amid',
'amidst',
'amnesia',
'amp',
'amy',
'amyadams',
'amypoehler',
'amyryan',
'amysedaris',
'amysmart',
'anagasteyer',
'analyst',
'anarchiccomedy',
'ancient',
'anderson',
'andiemacdowell',
'andreamartin',
'andrebraugher',
'andrew',
'andrewdaly',
'android',
'andy',
'andydick',
'andygarcía',
'andyrichter',
```

```
'andysamberg',
'andyserkis',
'angel',
'angelabassett',
'angeles',
'angelinajolie',
'angelo',
'anger',
'angry',
'animal',
'animalattack',
'animalhorror',
'animals',
'animated',
'animation',
'anjelicahuston',
'ann',
'anna',
'annafaris',
'annakendrick',
'annalevine',
'annapaquin',
'annasophiarobb',
'anndowd',
'anne',
'annebancroft',
'annefletcher',
'annehathaway',
'anneheche',
'annemoss',
'annettebening',
'annie',
'annikapergament',
'annual',
'answer',
'answers',
'ant',
'anthology',
'anthony',
'anthonyanderson',
'anthonyhopkins',
'anthonymackie',
'anthropomorphism',
'anti',
'antics',
'antihero',
'antonicorone',
'antoniobanderas',
'antonyelchin',
'apart',
'apartheid',
'apartment',
'ape',
'apes',
'apocalypse',
'apocalyptic',
'appear',
'appears',
'apple',
'apprentice',
'approaches',
'april',
'arab',
'archaeologist',
'architect',
'arctic',
'area',
'arena',
'arielwinter',
'arise',
'aristocrat',
'armed',
'arminmueller',
'arms',
'armstrong',
'army',
'arnold',
'arnoldschwarzenegger',
'arrest',
'arrested',
'arrival',
'arrive',
'arrives',
'arriving',
'art',
'arthur',
'arthurj',
'arthurtovey',
'artificialintelligence',
```

```
                    'artist',
                    'artists',
                    'arts',
                    'ashley',
                    'ashleyjudd',
                    'ashtonkutcher',
                    'asia',
                    'ask',
                    'asked',
                    'asks',
                    'aspirations',
                    'aspiring',
                    'assassin',
                    'assassination',
                    'assassins',
                    'assault',
                    'assigned',
                    'assignment',
                    'assistant',
                    'asteroid',
                    'astronaut',
                    'asylum',
                    'athlete',
                    'atomicbomb',
                    'attack',
                    'attacked',
                    'attacks',
                    'attempt',
                    'attempting',
                    'attempts',
                    'attending',
                    'attention',
                    'attitude',
                    'attorney',
                    'attraction',
                    'attractive',
                    'audience',
                    'audition',
                    'august',
                    'aunjanueellis',
                    'aunt',
                    'austin',
                    'australia',
                    'australian',
                    'author',
                    'authorities',
                    'authority',
                    'auto',
                    'avaacres',
                    'avenge',
                    'avoid',
                    'awaits',
                    'awakens',
                    'award',
                    'away',
                    'awry',
                    'azizansari',
                    'baby',
                    'bachelor',
                    'backdrop',
                    'bad',
                    'bahamas',
                    'bail',
                    'bailey',
                    'bailing',
                    'baker',
                    'balance',
                    'ball',
                    'ballet',
                    'band',
                    'banished',
                    'bank',
                    'banker',
                    'bankrobbery',
                    'baptiste',
                    'bar',
                    'barely',
                    'bargained',
                    'barn',
                    'barney',
                    'barry',
                    'barrylevinson',
                    'barrypepper',
                    'barryshabakahenley',
                    'base',
                    'baseball',
                    'based',
                    'basedoncomicbook',
                    'basedonnovel',
```

```
'basedonplay',
'basedonstagemusical',
'basedontrueevents',
'basedontruestory',
'basedontvseries',
'basedonvideogame',
'basedonyoungadultnovel',
'basement',
'basketball',
'batman',
'battle',
'battlefield',
'battles',
'battling',
'bay',
'beach',
'bear',
'bears',
'beast',
'beat',
'beautiful',
'beautifulwoman',
'beauty',
'beccasweitzer',
'becominganadult',
'bed',
'befriends',
'began',
'begin',
'beginning',
'begins',
'behavior',
'beliefs',
'believe',
'believed',
'believes',
'believing',
'bell',
'beloved',
'ben',
'benaffleck',
'benchaplin',
'benedictcumberbatch',
'benfoster',
'bengazzara',
'beniciodeltoro',
'benjamin',
'benjaminbratt',
'benkingsley',
'benmendelsohn',
'bennett',
'benstiller',
'bent',
'benwhishaw',
'berlin',
'bernardhill',
'bernardlee',
'berniemac',
'best',
'bestfriend',
'bet',
'beth',
'bethgrant',
'betrayal',
'bettemidler',
'better',
'betty',
'beverly',
'bible',
'big',
'biggest',
'biker',
'billcamp',
'billcobbs',
'billduke',
'billhader',
'billionaire',
'billmoseley',
'billmurray',
'billnighy',
'billnunn',
'billpaxton',
'billpullman',
'billy',
'billybobthornton',
'billyburke',
'billyconnolly',
'billycrudup',
'billycrystal',
```

```
                    'billydeewilliams',
                    'billyslaughter',
                    'billyzane',
                    'biography',
                    'bird',
                    'birth',
                    'birthday',
                    'bishop',
                    'bit',
                    'bite',
                    'bitter',
                    'bizarre',
                    'black',
                    'blackmail',
                    'blackpeople',
                    'blade',
                    'blake',
                    'blakeclark',
                    'blind',
                    'block',
                    'blonde',
                    'blood',
                    'bloodsplatter',
                    'bloodthirsty',
                    'bloody',
                    'blow',
                    'blue',
                    'blythedanner',
                    'board',
                    'boarding',
                    'boat',
                    'bob',
                    'bobbalaban',
                    'bobbergen',
                    'bobby',
                    'bobbycannavale',
                    'bobbyfarrelly',
                    'bobgunton',
                    'bobhoskins',
                    'bobstephenson',
                    'body',
                    'bodyguard',
                    'bokeemwoodbine',
                    'bold',
                    'bollywood',
                    'bomb',
                    'bombing',
                    'bond',
                    'bonniehunt',
                    'book',
                    'books',
                    'border',
                    'bored',
                    'boring',
                    'born',
                    'boss',
                    'boston',
                    'bound',
                    'boundaries',
                    'bounty',
                    'bountyhunter',
                    'box',
                    'boxer',
                    'boxing',
                    'boy',
                    'boydbanks',
                    'boydholbrook',
                    'boyfriend',
                    'boys',
                    'braddourif',
                    'bradgarrett',
                    'bradleycooper',
                    'bradleywhitford',
                    'bradpitt',
                    'bradwilliamhenke',
                    'brain',
                    'brave',
                    'bravery',
                    'brazil',
                    'brazilian',
                    'break',
                    'breaking',
                    'breaks',
                    'breckinmeyer',
                    'brendanfraser',
                    'brendangleeson',
                    'brendansextoniii',
                    'brentbriscoe',
                    'brentspiner',
```

```
'brettcullen',
'brettratner',
'brian',
'briancox',
'briandennehy',
'briandepalma',
'briandoyle',
'brianj',
'bride',
'bridge',
'bridgettewilson',
'brief',
'brielarson',
'brien',
'brilliant',
'bring',
'bringing',
'brings',
'brink',
'britain',
'british',
'britishsecretservice',
'brittanymurphy',
'broadway',
'broke',
'broken',
'brooklyn',
'brooks',
'brothel',
'brother',
'brotherbrotherrelationship',
'brothers',
'brothersisterrelationship',
'brought',
'brown',
'bruce',
'brucecampbell',
'brucedern',
'brucegreenwood',
'brucemcgill',
'brucespence',
'brucewillis',
'brutal',
'brutality',
'brutally',
'bryan',
'bryancranston',
'bryansinger',
'brycedallashoward',
'buck',
'buddies',
'buddy',
'buddycomedy',
'budget',
'build',
'building',
'built',
'bully',
'bullying',
'bumbling',
'buried',
'burns',
'burtreynolds',
'bus',
'bush',
'business',
'businessman',
'busy',
'butler',
'buy',
'byrne',
'cabin',
'california',
'called',
'calls',
'callumkeithrennie',
'calvin',
'came',
'camera',
'cameras',
'cameron',
'cameronbright',
'camerondiaz',
'camgigandet',
'camp',
'campaign',
'campbell',
'camping',
'campus',
```

'canada',
'canadian',
'cancer',
'candidate',
'cannibal',
'capable',
'capitalism',
'captain',
'captive',
'capture',
'captured',
'captures',
'car',
'caraccident',
'caraseymour',
'carchase',
'carcrash',
'card',
'care',
'career',
'carefree',
'careymulligan',
'caribbean',
'carjourney',
'carl',
'carlagallo',
'carlagugino',
'carlosalazraqui',
'carmen',
'carmenelectra',
'carolineaaron',
'carolinegoodall',
'carolkane',
'carpenter',
'carrace',
'carrie',
'carriefisher',
'carry',
'carrying',
'cars',
'carter',
'cary',
'caryelwes',
'case',
'caseyaffleck',
'cash',
'casino',
'cast',
'castle',
'cat',
'catch',
'catches',
'cateblanchett',
'catherinekeener',
'catherineo',
'catherinezeta',
'cathey',
'catholic',
'catholicism',
'caught',
'cause',
'caused',
'causes',
'causing',
'cave',
'cedrictheentertainer',
'celebrate',
'celebrated',
'celebration',
'celebrity',
'celiaimrie',
'celiaweston',
'cellphone',
'cemetery',
'center',
'centered',
'centers',
'central',
'centuries',
'century',
'certain',
'chadmichaelmurray',
'chain',
'chainsaw',
'challenge',
'challenged',
'challenges',
'champion',
'championship',

```
'chance',
'chancekelly',
'change',
'changed',
'changes',
'changing',
'channingtatum',
'chaos',
'chapter',
'character',
'characters',
'charge',
'charged',
'charismatic',
'charles',
'charlesdance',
'charlesdurning',
'charlesnapier',
'charless',
'charlie',
'charliesheen',
'charlizetheron',
'charlotterampling',
'charltonheston',
'charm',
'charming',
'chase',
'chased',
'chazzpalminteri',
'cheating',
'cheechmarin',
'cheerleader',
'chef',
'chelcieross',
'chemical',
'cher',
'chevychase',
'chicago',
'chicken',
'chief',
'child',
'childabuse',
'childhood',
'children',
'chilling',
'chimcbride',
'china',
'chinese',
'chip',
'chiwetelejiofor',
'chloëgracemoretz',
'chloësevigny',
'chocolate',
'choice',
'choices',
'choose',
'chosen',
'chris',
'chriscolumbus',
'chriscooper',
'chrisellis',
'chrisevans',
'chrishemsworth',
'chrisklein',
'chrismessina',
'chrismiller',
'chriso',
'chrisparnell',
'chrispenn',
'chrispine',
'chrispratt',
'chrisrock',
'christ',
'christian',
'christianbale',
'christianity',
'christianslater',
'christinaapplegate',
'christinaricci',
'christinebaranski',
'christmas',
'christmasparty',
'christopher',
'christopherguest',
'christopherlambert',
'christopherlee',
'christopherlloyd',
'christophermcdonald',
'christophermeloni',
```

```
'christophermintz',
'christopherplummer',
'christopherwalken',
'christophwaltz',
'chronicle',
'chronicles',
'church',
'cia',
'ciaránhinds',
'cillianmurphy',
'circle',
'circumstances',
'circus',
'citizens',
'city',
'civil',
'civilization',
'civilwar',
'claim',
'claims',
'claire',
'clairedanes',
'claireforlani',
'clan',
'clancybrown',
'clark',
'clarkgregg',
'clash',
'class',
'classes',
'classic',
'classmate',
'classmates',
'claudevandamme',
'cleaduvall',
'clean',
'clear',
'clerk',
'client',
'cliffcurtis',
'cliftoncollinsjr',
'climbing',
'clinteastwood',
'clinthoward',
'cliveowen',
'clock',
'clone',
'clorisleachman',
'close',
'closer',
'club',
'clues',
'clutches',
'coach',
'coast',
'cocaine',
'code',
'cody',
'cohen',
'cold',
'coldwar',
'cole',
'colehauser',
'colinfarrell',
'colinfirth',
'colinhanks',
'colinsalmon',
'collapse',
'colleague',
'colleagues',
'collect',
'collection',
'collector',
'colleencamp',
'college',
'collide',
'collins',
'collision',
'colmfeore',
'colmmeaney',
'colonel',
'colony',
'colorful',
'coma',
'combat',
'come',
'comeback',
'comedian',
'comedy',
```

```
            'comes',
            'comic',
            'comics',
            'coming',
            'comingofage',
            'comingout',
            'command',
            'commander',
            'commercial',
            'commit',
            'commitment',
            'committed',
            'common',
            'community',
            'companion',
            'company',
            'compete',
            'competing',
            'competition',
            'complete',
            'completely',
            'complex',
            'complicated',
            'complications',
            'compton',
            'computer',
            'concert',
            'condition',
            'confidence',
            'confident',
            'conflict',
            'confront',
            'congress',
            'conman',
            'connected',
            'connection',
            'connell',
            'connienielsen',
            'connieray',
            'connor',
            'conquer',
            'conradvernon',
            'conscience',
            'consequences',
            'conspiracy',
            'constantly',
            'construction',
            'contact',
            'contest',
            'continue',
            'continues',
            'contract',
            'control',
            'controversial',
            'converge',
            'convict',
            'convicted',
            'convince',
            'convinced',
            'convinces',
            'cook',
            'cool',
            'cooper',
            'cop',
            'cope',
            'cops',
            'core',
            'coreyburton',
            'coreystoll',
            'corporate',
            'corporation',
            'corpse',
            'corrupt',
            'corruption',
            ...]
```

We search smiliar vectors(movies),we calculate distance between 2 vectors, more the distance less the smiliarity.

We wont use Eucledian distance cuz its a measure of tip to tip rather we will go for Cosine to measure the angle between 2 vectors.

distance=1/similarity

In [55]:
```python
from sklearn.metrics.pairwise import cosine_similarity
#Cosine similarity, or the cosine kernel, computes similarity as the normalized dot product of X and Y:
```

```
#K(X, Y) = <X, Y> / (||X||*||Y||)
```

In [60]:
```python
similar=cosine_similarity(vectors)
```

In [61]:
```python
similar.shape
```

Out[61]:
```
(4806, 4806)
```

In [62]:
```python
similar[0]
```

Out[62]:
```
array([1.        , 0.07142857, 0.05216405, ..., 0.02326211, 0.02571722,
       0.        ])
```

In [64]:
```python
new_df[new_df['title'] == 'The Lego Movie'].index[0] #At what index movie is present
```

Out[64]:
```
744
```

In [76]:
```python
def recommend_movie(movie):
    index = new_df[new_df['title'] == movie].index[0] #Fetching INDEX of the movie
    distances = sorted(list(enumerate(similar[index])),reverse=True,key = lambda x: x[1])
    for i in distances[1:10]:
        print(new_df.iloc[i[0]].title)
```

In [77]:
```python
recommend_movie('Avatar')
```

```
Lifeforce
Aliens vs Predator: Requiem
Battle: Los Angeles
Titan A.E.
Independence Day
Krull
Apollo 18
Independence Daysaster
Ender's Game
```

In [78]:
```python
recommend_movie('Batman')
```

```
Batman
Batman & Robin
Batman
Batman Returns
Batman Forever
The Dark Knight Rises
Batman Begins
Amidst the Devil's Wings
The Dark Knight
```

In [79]:
```python
recommend_movie('Lifeforce')
```

```
The Thing
Home
Aliens vs Predator: Requiem
Titan A.E.
Star Trek: Insurrection
Area 51
Avatar
Alien
Attack the Block
```

So this a decent recommender system

In [ ]: