Out[9]:	budget 0 237000000	genres homepage id keywords original_language original_title overview popularity production_companies production_countries release_date revenue runtime spreading in the spreadi
		"nam" Tid": Gi
	2 245000000	Find dead, ha ha ha
	3 250000000	Following Foll
	4 260000000	Second
In [10]: Out[10]: .	credits.head() movie_id 0 19995	Ca
	2 2066473 490264 49529	Spectre [{"cast_id": 4, "character": "Captain Jack Spa [{"credit_id": "52fe4232c3a36847f800b579", "de Spectre [{"cast_id": 1, "character": "James Bond", "cr [{"credit_id": "54805967c3a36829b5002c41", "de The Dark Knight Rises [{"cast_id": 2, "character": "Bruce Wayne / Ba [{"credit_id": "52fe4781c3a36847f81398c3", "de John Carter [{"cast_id": 5, "character": "John Carter", "c [{"credit_id": "52fe479ac3a36847f813eaa3", "de
	movies.head() budget	genres homepage id keywords original_language original_title overview popularity production_companies runtime spoken_languages status tagline [{"id": 28,
	0 237000000	"Action"}, http://www.avatarmovie.com/ 1995
		Adventure"}, http://disney.go.com/disneypictures/pirates/ 285 "ocean"}, en Caribbean: believed 139.082615 Disney Pictures", "id": 169.0 "en", "name": Released the At World's to be 2}, {" "English"}] adventure begins. Table
	2 245000000	"Action"}, http://www.sonypictures.com/movies/spectre/ 20647 "spy"}, en Spectre Bond's 107.376788 Pictures", "id": 5}, 148.0 "name": Released No One Spectre Past ("nam "Fran\u00e7ais"}, Escapes "nam "Fran\u00e7ais"}, Escapes "nam him o "Fran\u00e7ais"}, Escapes "nam him o "Fran\u00e7ais"}, The Dark the death [{"name": "Legendary [{"id": 5}, 148.0 "name": Released No One Spectre Past ("name": Released No One Spectre Past ("name": Released No One Spectre Past ("name": Name": Released No One Spectre Past ("name": Tran\u00e7ais"), Escapes "name": Name": Tran\u00e7ais"), Following [{"id": 5}, 148.0 "name": Released No One Spectre Past ("name": Name": Name": Name": Name": Name": Name": Name No One Spectre Past ("name": Spectre Past ("name": Name": Name": Name No One Spectre Past ("name": Name No One Spectre Past ("name": Name": Name No One Spectre Past ("name": Name": Name No One Spectre Past ("name": Name": Name No One Spectre Past ("name": Name No
	3 250000000	"Action"}, http://www.thedarkknightrises.com/ 49026 "dc en Knight of District 112.312950 Pictures", "id": 923}, 165.0 "en", "name": Released Legend Knight en Knight of District 112.312950 Pictures", "id": 923}, 165.0 "en", "name": Released Legend Knight en K
	4 260000000 5 rows × 23 column movies.descrik	"Action"}, http://movies.disney.com/john-carter 49529 "based en John Carter weary, 43.926995 Disney Pictures", "id": 132.0 "en", "name": Released world, C {"id": 12, on former 2}] "English"}] found in "nam military another. {"id": ca
Out[15]:	count 4.8090006 mean 2.9027806 std 4.0704736 min 0.0000006 25% 7.8000006 50% 1.5000006	21.491664 8.227511e+07 106.882255 6.092514 690.331670 57120.571429 21.491664 8.227511e+07 106.882255 6.092514 690.331670 57120.571429 21.491664 8.227511e+07 106.882255 6.092514 690.331670 57120.571429 21.491664 8.227511e+07 106.882255 6.092514 690.331670 57120.571429 22.602535 1.193989 1234.187111 88653.369849 23.602535 1.193989 1234.187111 88653.369849 24.000000 0.000000 0.000000e+00 0.000000 0.000000 0.000000 5.000000 24.000000 0.000000 0.000000e+00 0.000000 0.000000 0.000000 0.000000 24.000000 0.000000 0.000000e+00 0.000000 0.000000 0.000000 0.000000 24.000000 0.000000 0.000000e+00 0.000000 0.000000 0.000000 0.000000 0.000000
	movies=movies[movies.head() movie_id 19995	['movie_id', 'title', 'overview', 'genres', 'keywords', 'cast', 'crew']] title overview genres keywords cast crew Avatar In the 22nd century, a paraplegic [{"id": 28, "name": "Action"}, {"id": [{"id": 1463, "name": "culture [{"cast_id": 242, "character": "Jake [{"credit_id": "52fe48009251416c750aca23", Marine is di 12, "nam clash"}, {"id": Sully", " Sully", " "de
		Marine is di Pirates of the Caribbean: At World's End Spectre A cryptic message from Bond's past sends him o The Dark Knight Rises Marine is di Marine is di 12, "nam 13, "id": 270, "name": "ocean"}, {"id": 17, "character": "Captain and Jack Spa 14, "character": "Captain and Spa 15, "id": 270, "name": "ocean"}, {"id": 17, "character": "Captain and Jack Spa 15, "id": 270, "name": "ocean"}, {"id": 17, "character": "Captain and Jack Spa 16, "credit_id": "54805967c3a36829b5002c41", and medical an
<		John Carter John Carter is a war-weary, former ["id": 28, "name": "Action"], {"id": ["id": 818, "name": "based on ["cast_id": 5, "character": "John ["credit_id": "52fe479ac3a36847f813eaa3", military ca 12, "nam novel"], {"id": Carter", "c Carter", "c "de core.frame.DataFrame'> 9 entries, 0 to 4808
-	# Columns (t # Column 0 movie_id 1 title 2 overview 3 genres 4 keywords 5 cast	9 entries, 0 to 4808 otal 7 columns): Non-Null Count Dtype
ć m In [26]:		4809 non-null object), object(6) 63.1+ KB
	keywords 0 cast 0 crew 0 dtype: int64 movies.dropna(import warning warnings.filte	gs erwarnings(<mark>'ignore'</mark>)
Out[30]: In [32]: In [34]:	<pre># listing genr movies.iloc[0]</pre>	res
In [36]:	<pre>import ast def convert(ob L=[] for i in a</pre>	
		title overview genres keywords cast crew Avatar In the 22nd century, a paraplegic [Action, Adventure, Fantasy, ["id": 1463, "name": "culture [("cast_id": 242, "character": "Jake [("credit_id": "52fe48009251416c750aca23", Marine is di Science Fiction] clash"}, ["id": Sully", " Sully", " "de
		Avatar Marine is di Science Fiction] Clash"}, {"id": Sully", " Sully", " Sully", " Sully", " Sully", " Sully", " ("credit_id": "52fe4232c3a36847f800b579", "de [adventure, Fantasy, Action] Spectre A cryptic message from Bond's past sends him o Spectre A cryptic message from Bond's past sends him o The Dark Knight Rises Following the death of District Attorney [Action, Crime, Drama, Thriller] ["id": 849, "name": "dc comics"}, ["id": 270, "name": "coean"}, {"id": 270, "name": "coean"}, {"id": 270, "name": "coean"}, {"id": 270, "name": "coean"}, {"id": 270, "name": "spy"}, {"id": 270, "name"
In [42]:	4 49529	The Dark Knight Rises Harve John Carter John Carter military ca The Dark Knight Rises Harve John Carter is a war-weary, former military ca The Dark Knight Rises Harve John Carter is a war-weary, former military ca The Dark Knight Rises Harve [Action, Crime, Drama, Thriller] [Action, Crime, Drama, Thriller] [Action, Adventure, Science ["id": 818, "name": "based on novel"], "id": Fiction] The Dark Knight Rises [Action, Crime, Drama, Thriller] [Action, Adventure, Science ["id": 818, "name": "based on novel"], "id": Carter", "c Carter", "c The Dark Knight Rises [Action, Crime, Drama, Thriller] [Action, Crime, Drama, Thriller] [Action, Adventure, Science ["id": 818, "name": "based on novel"], "id": Carter", "c The Dark Knight Rises [Action, Crime, Drama, Thriller] [Action, Adventure, Science ["id": 818, "name": "based on novel"], "id": Carter", "c The Dark Knight Rises [Action, Crime, Drama, Thriller] [Action, Crime, Drama, Thriller] [Action, Crime, Drama, Thriller] [Action, Crime, Drama, Thriller] ["id": 849, "flaffed de d
In [45]: Out[45]:	movie_id 0 19995	title overview genres keywords cast crew Avatar Avatar Pirates of the Caribbean: At World's End World's End World's End Keywords words title overview genres keywords (acast crew genres keywords) (acast cast crew genres cast cast crew genres war, space colon Sully", " [4"cast_id": 242, "character": "Jake [4"credit_id": "52fe48009251416c750aca23", "de
	 2 206647 3 49026 4 49529 	Spectre Spectre A cryptic message from Bond's past sends him o The Dark Knight Rises John Carter Spectre John Carter is a war-weary, former military ca The Dark Knight Rises Spectre Spectra
In [46]:	<pre>def convert3(c L=[] counter=0 for i in a if counter=0</pre>	military ca Fiction] space travel Carter", "c "de
	else: br return L	title overview genres keywords cast crew
	0 19995	Avatar In the 22nd century, a paraplegic Marine is di Science Fiction] [Action, Adventure, Fantasy, Science Fiction] [Coean, drug abuse, exotic island, World's End Spectre A cryptic message from Bond's past [Action, Adventure, Crime] [Action, Adventure, Fantasy, Action] [Sam Worthington, Zoe Saldana, Sigourney Weave [Sam Worthington, Zoe Saldana, Sigourney Weave [It "credit_id": "52fe48009251416c750aca23", Sigourney Weave [It "cre
	3 490264 49529	Spectre sends him o The Dark Knight Rises John Carter John Carter Sends him o The Dark Knight Rises Following the death of District Attorney Harve John Carter is a war-weary, former military ca The Dark Knight Rises Following the death of District Attorney Harve Attorney Harve John Carter is a war-weary, former military ca Following the death of District Attorney Drama, Thriller [Action, Crime, Drama, Thriller] [Action, Crime, Drama, Thriller] [Action, Crime, Drama, Thriller] [Action, Crime, Drama, Thriller] [Action, Adventure, Science [based on novel, mars, medallion, space travel [Taylor Kitsch, Lynn Collins, Samantha Morton, ["credit_id": "52fe479ac3a36847f813eaa3", Morton, "de "de
	if i['L. br return L	<pre>ast.literal_eval(obj): job'] == 'Director': append(i['name']) reak</pre>
	movies['crew'] movies.head() movie_id 19995	= movies['crew'].apply(Fetch_director) title overview genres keywords cast crew Avatar In the 22nd century, a paraplegic Marine is di [Action, Adventure, Fantasy, Science Fiction] [culture clash, future, space war, space colon] [Sam Worthington, Zoe Saldana, Sigourney Weave] [James Cameron]
	 285 206647 49026 	Pirates of the Caribbean: At World's End Spectre The Dark Knight Rises Only Pirates of the Caribbean: At World's End Only Pirates of the Caribbean: At World's End Only Pirates of the Caribbean: At Captain Barbossa, long believed to be dead, ha (Adventure, Fantasy, Action) (Coean, drug abuse, exotic island, east india india india Knightley, (Sore Verbinski) (Borie Verbinski) (Coribe, Drama, Thriller) (Comics, crime fighter, terrorist, secret agent, sequel, mi (Christian Bale, Michael Caine, Gary Oldman, Nolan) Only Pirates of the Caribbean: At Captain Barbossa, long believed to be dead, ha (Gore Verbinski) (Borie Verbinski) (Coribe, Drama, Thriller) (Comics, crime fighter, terrorist, secret in (Christian Bale, Michael Caine, Gary Oldman, Nolan)
Out[55]:	4 49529 movies['overvi 'In the 22nd o	John Carter John Carter is a war-weary, former military ca [Action, Adventure, Science Fiction] [based on novel, mars, medallion, space travel [Taylor Kitsch, Lynn Collins, Samantha Morton, [Andrew Stanton] [Andrew Stanton] [Andrew Stanton] [Description] [based on novel, mars, medallion, space travel [Taylor Kitsch, Lynn Collins, Samantha Morton, [Andrew Stanton]
In [56]:	0 [In, 1] 1 [Capta 2 [A, c: 3 [Folio 4 [John] 4804 [El, 1]	the, 22nd, century, a, paraplegic, Marin ain, Barbossa, long, believed, to, be, d ryptic, message, from, Bond's, past, send wwing, the, death, of, District, Attorney Carter, is, a, war-weary, former, mili Mariachi, just, wants, to, play, his, gui ewlywed, couple's, honeymoon, is, upended
In [57]: Out[57]:	4805 [A, ne 4806 ["Sign 4807 [When, 4808 [Ever,	title overview genres keywords tasts crew
	0 199951 2852 206647	Avatar Avatar In the 22nd century, a paraplegic Marine is di Pirates of the Caribbean: At World's End Spectre Spectre A cryptic message from Bond's past sends him o [Action, Adventure, Fantasy, Science Fiction] [Action, Adventure, Fantasy, Science Fiction] [Culture clash, future, space war, space colon [Colon [Collure clash, future, space war, space colon [Colon [Colon [Collure Clash, future, space war, space colon [Colon
		The Dark Knight Rises Following the death of District Attorney Harve John Carter John Cart
	movies['cast'] movies['crew'] movies.head() movie_id	title overview genres keywords 'lambda x: [i.replace (" ","") for i in x]) title overview genres keywords cast crew In the 22nd century a paraplegic Marine is [Action Adventure Fantasy [Cultureclash future spacewar [SamWorthington ZoeSaldana]]
	19995285206647	Avatar di ScienceFiction] spacecolony, Pirates of the Caribbean: At World's End World's End A cryptic message from Bond's past sends him o Spectre A cryptic message from Bond's past sends him o Spectre A cryptic message from Bond's past sends him o [Action, Adventure, Crime] [Action
In [60]:	<pre>3 49026 4 49529 def combine_fereturn f"{</pre>	The Dark Knight Rises Following the death of District Attorney Harve John Carter John Carter is a war-weary, former military ca [Action, Adventure, ScienceFiction] John Carter (Frow['crew']) {row['genres']} {row['werview']} {row['verview']} {row['crew']} {r
		= movies.apply(combine_features, axis=1) title overview genres keywords cast crew tags Avatar In the 22nd century, a paraplegic [Action, Adventure, Fantasy, [cultureclash, future, spacewar, [SamWorthington, ZoeSaldana, [JamesCameron]] In the 22nd century, a paraplegic
		Marine is di Marine is di ScienceFiction] spacecolony, SigourneyWeaver, SigourneyWeaver, SigourneyWeaver, SigourneyWeaver, SigourneyWeaver, SigourneyWeaver, SigourneyWeaver, SigourneyWeaver, SigourneyWeaver, [JohnnyDepp, OrlandoBloom, KeiraKnightley, Ste [GoreVerbinski] Captain Barbossa, long believed to be dead, ha Spectre A cryptic message from Bond's past sends him o [Action, Adventure, Crime] Spectre sequel, mi6, The Dark Knight Rises Following the death of District [Action, Crime, Drama, Thriller] Science Fiction] spacecolony, SigourneyWeaver, SigourneyWeaver, SigourneyWeaver, SigourneyWeaver, SigourneyWeaver, SigourneyWeaver, [JohnnyDepp, OrlandoBloom, KeiraKnightley, Ste [Action Adventure, Crime] Sequel, mid, [SamMendes] A cryptic message from Bond's past sends him o [SamMendes] Following the death of District Politority (ChristopherNolan) Following the District Politority (ChristopherNolan) Following the Di
In [62]:	4 49529 data=movies[['	The Dark Knight Rises Attorney Harve John Carter John Carter is a war-weary, former military ca Movie_id', 'tittle', 'tags']] Attorney Harve [Action, Crime, Drama, Thriller] Secretiden [Action, Crime, Drama, Thriller] Secretiden [Action, Adventure, [basedonnovel, mars, medallion, spacetravel, p ScienceFiction] Secretiden [ChristopherNolan] Attorney Harve [AndrewStanton] John Carter is a war-weary, former military ca ScienceFiction] Secretiden SamanthaMorton, Wi Movie_id', 'tittle', 'tags']]
In [63]:	0 In 1 Cap 2 A c 3 Fol 4 Joh 4804 El 4805 An	the 22nd century, a tain Barbossa, long ryptic message from lowing the death of n Carter is a war-we Mariachi just wants ewlywed couple's hon
In [64]: Out[64]:	4806 "Si 4807 Whe 4808 Eve Name: tags, Le data.head()	gned, Sealed, Delive n ambitious New York r since the second g ength: 4806, dtype: object title tags
	 19995 285 Pira 206647 49026 49529 	Avatar In the 22nd century, a paraplegic Marine is di Attes of the Caribbean: At World's End Captain Barbossa, long believed to be dead, ha Spectre A cryptic message from Bond's past sends him o The Dark Knight Rises Following the death of District Attorney Harve John Carter S a war-weary, former military ca
In [66]:	vectoriza import nltk from nltk.stem	ation approximate import PorterStemmer
	<pre>ps=PorterStemm def stem(text) y = [] for i in t</pre>	<pre>iner() cext.split(): cext.split(): cend(ps.stem(i))</pre>
In [70]:	data['tags'] = data['tags'] [0 "in the 22nd of tasy', 'sciency', 'sciency', 'beldavidmoory', 'peterdil', 'ilramchoory', 'ilramchoory', 'brandonjelkes' jamieharrisonansheehan', 'sciency', 'sciency', 'sciency', 'illower', 'illower'	data['tags'].apply(stem)
Out[72]:	0 19995 1 285 Pira 2 206647 3 49026 4 49529 data['tags'][0	century, a parapleg marin is dispatch to the moon pandora on a uniqu mission, but becom torn between follow order and protect an alien civilization. ['action', 'adventure', 's
In [74]: In [75]: In [76]:	tasy', 'sciend' 'soldier', 'ba' 'joeldavidmoory', 'peterdil' n', 'ilramchor' sarawilson', 'brandonjelker' jamieharrisonansheehan', 'tanker' sarawilson', 'tanker' tanker' tank	century, a parapleg marin is dispatch to the moon pandora on a uniqu mission, but becom torn between follow order and protect an alien civilization. ['action', 'adventure', 'refiction', '[cultureclash', 'future', 'spacewar', 'spacecolony', 'society', 'spacetravel', 'futuristic', 'romance', 'space', 'alien', 'tribe', 'alienplanet', 'cgi', 'marine', attle', 'loveaffair', 'antiwar', 'powerrelations', 'mindandsoul', '3d'] ['samworthinţton', 'zoesaldana', 'sigounneyweaver', 'stephenlang', 'michellerodriguez', 'giovanniribisis' re', 'cchpounder', 'wesstudi', 'lazalonso', 'dileeprao', 'mattgerald', 'seananthonymoran', 'jasonwhyte', 'scottlawrence', 'kellykilgour', 'jamespatrickpitt', 'seanpatrickmurp'lon', 'kevindorman', 'kelsonhenderson', 'davidvanhorn', 'jacobtomuri', 'michaelblain-rozgay', 'joncurry', 'lukehawker', 'woodyschultz', 'petermensah', 'sonlayee', 'jahnelcurfn', 'kylwarren', 'lisaroumanin', 'debrawilson', 'chrismala', 'taylorkhby', 'joidelandau', 'julielamm', 'cullenb.madden', 'josephbradymadden', 'frankierores', 'austinwilson', 'tamicawashington-miller', 'lucybriant', 'nathanmeister', 'gerryblair', 'matthewchamberlain', 'paulyates', 'wraywilson', 'jamesgaylyn', 'melvinlenoclarkiii', 'carvonfutrell', 's', 'micahmoch', 'hanniyahmuhammad', 'christophernolen', 'christaoliver', 'aprilmariethomas', 'bravitaa.threatt', 'colinbleasdale', 'mikebodnar', 'mattolayton', 'nicoledionne' n', 'alienhenry, 'anthonyingruber', 'ashleyjeffery, 'deanknowsley', 'josephmika-hunt', 'terrynotary', 'kaipantano', 'loganpithyou', 'stuartpollock', 'raja', 'garethruck', 'st.j.storm', 'jodietaylor', 'aliciavela-bailey', 'richardwhiteside', 'nikiezambo', 'julenerenee'] ['jamescameron']" **Reature_extraction.text import CountVectorizer orizer(max_features=5000, stop_words='english') **Int_transform(data['tags']).toarray() **O, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
In [77]:	[0, 0, [0, 0,, [0, 0, [0, 0, vectors[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 [78]: In [79]: In [83]:	<pre>from sklearn.m similarity=cos similarity array([[1.</pre>	netrics.pairwise import cosine_similarity sine_similarity(vectors) , 0.07325794, 0.0541332 ,, 0.02311251, 0.02441931,
.0[83]:	0. [0.073; 0.0224] [0.054] 0, [0.023; 0.042]], 25794, 1. , 0.07733089,, 0.02201127, 0. , 48469], 1332 , 0.07733089, 1. ,, 0.0243975 , 0. ,], 11251, 0.02201127, 0.0243975 ,, 1. , 0.06603382, 56283], 41931, 0. , 0. ,, 0.06603382, 1. ,
Out[84]:	0.089 [0. 1. similarity[0] array([1. 0.	41931, 0. , 0. ,, 0.06603382, 1. , , 93875], , 0.02248469, 0. ,, 0.04256283, 0.08993875,]]) , 0.07325794, 0.0541332 ,, 0.02311251, 0.02441931,])
Out[85]:	[(1214, 0.275) (2329, 0.246)	
	(507, 0.24602 (1916, 0.2433 (539, 0.23692 Main fur def recommend (movie_inde	2771043141894), 599382882345), 241569216954)] Inction Imovie): ex = data[data['title'] == movie].index[0]
A	<pre>distance = movie_list for i in m</pre>	<pre>s similarity[movie_index]</pre>
9 I I	Predator Independence Da Lifeforce Titan A.E. import pickle pickle.dump(da	y ita,open('movies.pkl','wb'))
		values r', "Pirates of the Caribbean: At World's End", 'Spectre', Signed, Sealed, Delivered', 'Shanghai Calling', te with Drew'], dtype=object)
In [94]: Out[94]: In [96]:	pickle.dump(da	te with Drew'], dtype=object) ta.to_dict(),open('moviesdict.pkl','wb')) milarity,open('similarity.pkl','wb'))