

Data:

We retrieve the data from:

<https://www.kaggle.com/deepmatrix/imdb-5000-movie-dataset>

The data is in a csv format, it contains variables for 5043 movies. Below are the 28 variables:

"movie_title" "color" "num_critic_for_reviews" "movie_facebook_likes"
"duration" "director_name" "director_facebook_likes" "actor_3_name"
"actor_3_facebook_likes" "actor_2_name" "actor_2_facebook_likes"
"actor_1_name" "actor_1_facebook_likes" "gross" "genres"
"num_voted_users" "cast_total_facebook_likes" "facenumber_in_poster"
"plot_keywords" "movie_imdb_link" "num_user_for_reviews" "language"
"country" "content_rating" "budget" "title_year" "imdb_score"
"aspect_ratio"

The marked data attributes were used in our visualization the ones we used in order to show a simpler and comprehensible visualization for the users. These several attributes were used in multiple visualizations, in which we will discuss later in the next sections.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	color	director_name	num_critic_for_reviews	duration	director_facebook_likes	actor_3_name	actor_3_facebook_likes	actor_2_name	actor_2_facebook_likes	actor_1_name	actor_1_facebook_likes	gross	genres	num_voted_users	cast_total_facebook_likes	facenumber_in_poster	plot_keywords	movie_imdb_link	num_user_for_reviews	language	country
2	Color	James Cameron	723	178	0	855 Joel David Moore	1000	7.61E+08	Action AcCCH Poun	Avatar	886204	4834 Wes Studi	0 avatar ful	http://www	3054	English	USA				
3	Color	Gore Verbinski	302	169	563	1000 Orlando Bloom	40000	3.09E+08	Action AcJohnny De	Pirates of the Caribbean: The Curse of the Black Pearl	471220	48350 Jack Davenport	0 goddess http://www	1238	English	USA					
4	Color	Sam Mendes	602	148	0	161 Rory Kinnear	11000	2E+08	Action AcChristoph	Spectre	275868	11700 Stephanie	1 bomb es http://www	994	English	UK					
5	Color	Christopher Nolan	813	164	22000	23000 Christian Bale	27000	4.48E+08	Action ThTom Hard	The Dark Knight	1144337	106759 Joseph Gordon-Levitt	0 deception http://www	2701	English	USA					
6		Doug Walker			131	Rob Walker	131		Document	Doug Wall Star Wars	8	143	0	http://www.imdb.com/title/tt52899							
7	Color	Andrew Stanton	462	132	475	530 Samantha Morton	640	73058679	Action AcDaryl Sabz	John Carter	212204	1873 Polly Walker	1 alien ame http://www	738	English	USA					
8	Color	Sam Raimi	392	156	0	4000 James Franco	24000	3.37E+08	Action AcJ.K. Simm	Spider-Man 3	383056	46055 Kirsten Dunst	0 sandman http://www	1902	English	USA					
9	Color	Nathan Grey	324	100	15	284 Donna Murphy	799	2.01E+08	Adventure Brad Garre	Tangled	294810	2036 M.C. Gainey	1 17th centu http://www	387	English	USA					
10	Color	Joss Whedon	635	141	0	19000 Robert Downey Jr.	26000	4.59E+08	Action AcChris Hem	Avengers: The Avengers	462669	92000 Scarlett Johansson	4 artificial i http://www	1117	English	USA					
11	Color	David Yates	375	153	282	10000 Daniel Radcliffe	25000	3.02E+08	Adventure Alan Rickr	Harry Potter and the Prisoner of Azkaban	321795	58753 Rupert Grint	3 blood boi http://www	973	English	UK					
12	Color	Zack Snyder	673	183	0	2000 Lauren Cohan	15000	3.3E+08	Action AcHenry Cav	Batman v Superman: Dawn of Justice	371639	24450 Alan D. Poo	0 based on http://www	3018	English	USA					
13	Color	Bryan Singer	434	169	0	903 Marlon Brando	18000	2E+08	Action AcKevin Spa	Superman Returns	240396	29991 Frank Langella	0 crystal ep http://www	2367	English	USA					
14	Color	Marc Forster	403	106	395	393 Mathieu Amalric	451	1.68E+08	Action AcGiancarlo	Quantum of Solace	330784	2023 Rory Kinnear	1 action her http://www	1243	English	UK					
15	Color	Gore Verbinski	313	151	563	1000 Orlando Bloom	40000	4.23E+08	Action AcJohnny De	Pirates of the Caribbean: On Stranger Tides	522040	48486 Jack Davenport	2 box office http://www	1832	English	USA					
16	Color	Gore Verbinski	450	150	563	1000 Ruth Wils	40000	89289910	Action AcJohnny De	The Lone Ranger	181792	45757 Tom Wilki	1 horse out http://www	711	English	USA					
17	Color	Zack Snyder	733	143	0	748 Christoph Waltz	15000	2.91E+08	Action AcHenry Cav	Man of Steel	548573	20495 Harry Leni	0 based on http://www	2536	English	USA					
18	Color	Andrew A. Kosove	258	150	80	201 Pierfrancesco Favino	22000	1.42E+08	Action AcPeter Dink	The Chron	149922	22697 Damián Marín	4 brother br http://www	438	English	USA					
19	Color	Joss Whedon	703	173	0	19000 Robert Downey Jr.	26000	6.23E+08	Action AcChris Hem	The Avengers	995415	87697 Scarlett Johansson	3 alien inva http://www	1722	English	USA					
20	Color	Rob Marshall	448	136	252	1000 Sam Claflin	40000	2.41E+08	Action AcJohnny De	Pirates of the Caribbean: The Curse of the Black Pearl	370704	54083 Stephen Gately	4 blackbear http://www	484	English	USA					
21	Color	Barry Sonnenfeld	451	106	188	718 Michael St. Louis	10000	1.79E+08	Action AcWill Smith	Men in Black II	268154	12572 Nicole Scherzinger	1 alien crin http://www	341	English	USA					
22	Color	Peter Jackson	422	164	0	773 Adam Brown	5000	2.55E+08	Adventure Aidan Turi	The Hobbit	354228	9152 James Nesbitt	0 army elf http://www	802	English	New Zealand					
23	Color	Marc Webb	599	153	464	963 Andrew Garfield	15000	2.62E+08	Action AcEmma Sto	The Amazing Spider-Man	451803	28489 Chris Zylka	0 lizard out http://www	1225	English	USA					

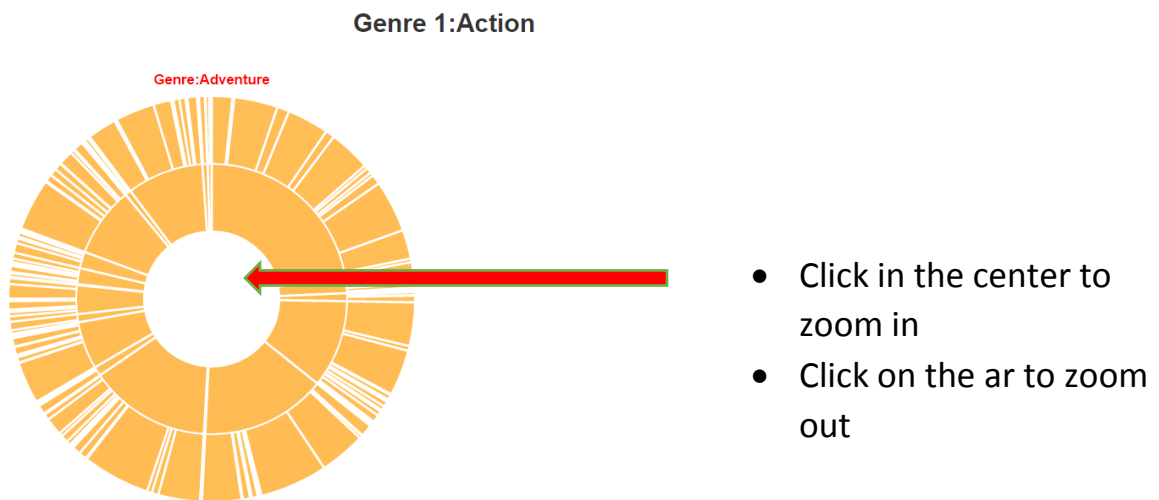
Problem number 1:

The Variables that we are trying to plot is the related movies to the three genres and there scores

So the problem is that graphs that use data points or lines to encode data, multiple objects ends up sharing the same space and then the data is no more comprehensive, because one can't see individual values.

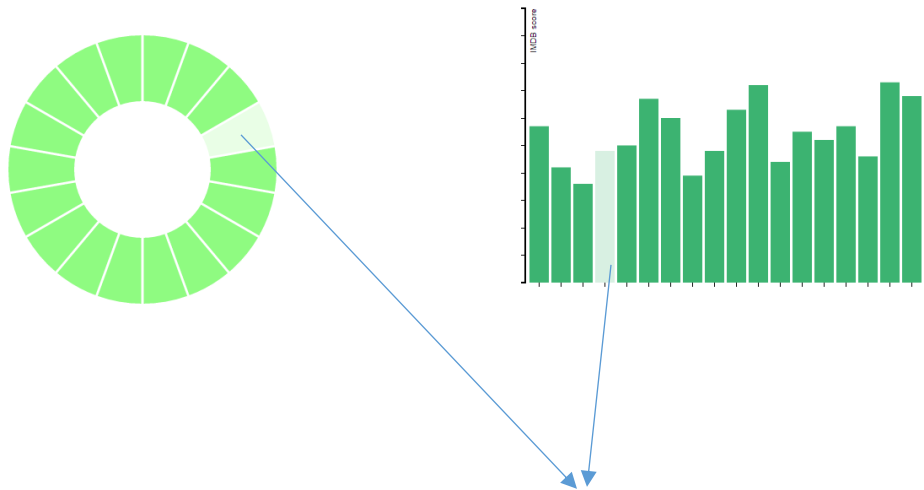
Solution:

We Change the shape of data objects, by categorizing the data regarding there 3 first genres, and then we plotted the movies that related to these genres, and there IMDB scores in a bar-chart graph.



Genre 1:Horror, Genre 2:Thriller, Genre 3:Sci-Fi

Containment
SCORE:4.8



The same movie

Problem number 2:

We are trying to plot the distribution of the genres for a specific year

Solution:

We used pie chart to show the genres distribution of the year, and a different color for each genre. And a histogram chart to choose a year on hovering over it, the height of the bar in the histogram is the number of movies in that year.

