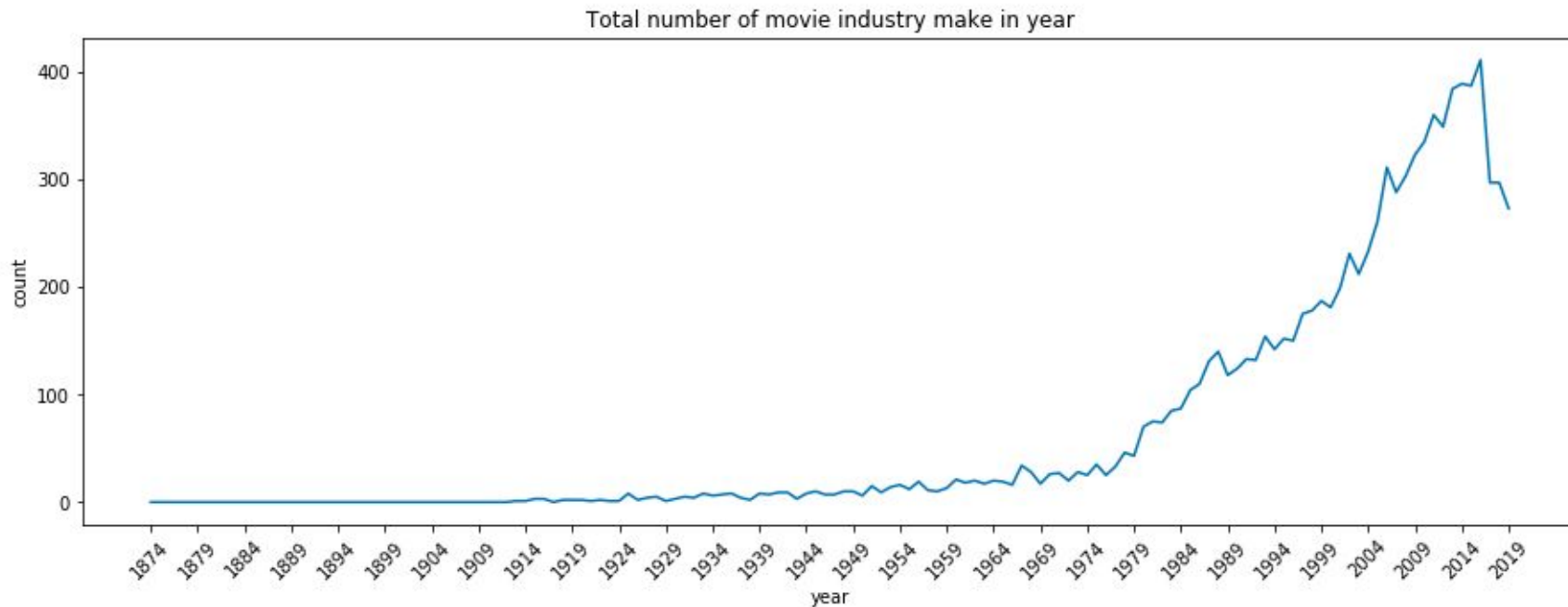


MOVIE PREDICT AND RECOMENDATION

Total number movie over year!



Total, Avg. revenue and number movie each year

	Average_revenue	Total_revenue	count
year			
2016	7.627827e+07	3.135037e+10	411
2015	7.414786e+07	2.869522e+10	387
2014	7.113407e+07	2.767116e+10	389
2013	7.039734e+07	2.703258e+10	384
2012	7.569661e+07	2.641812e+10	349
2011	6.939981e+07	2.498393e+10	360
2010	7.205477e+07	2.413835e+10	335
2009	7.440009e+07	2.403123e+10	323
2008	6.991971e+07	2.118567e+10	303
2007	7.076921e+07	2.038153e+10	288

Year was first movie

	year	title	runtime
34693	1874	Passage of Venus	1.0
34690	1878	Sallie Gardner at a Gallop	1.0
70946	1881	Athlete Swinging a Pick	1.0
41270	1883	Buffalo Running	1.0
135279	1885	L'homme machine	1.0

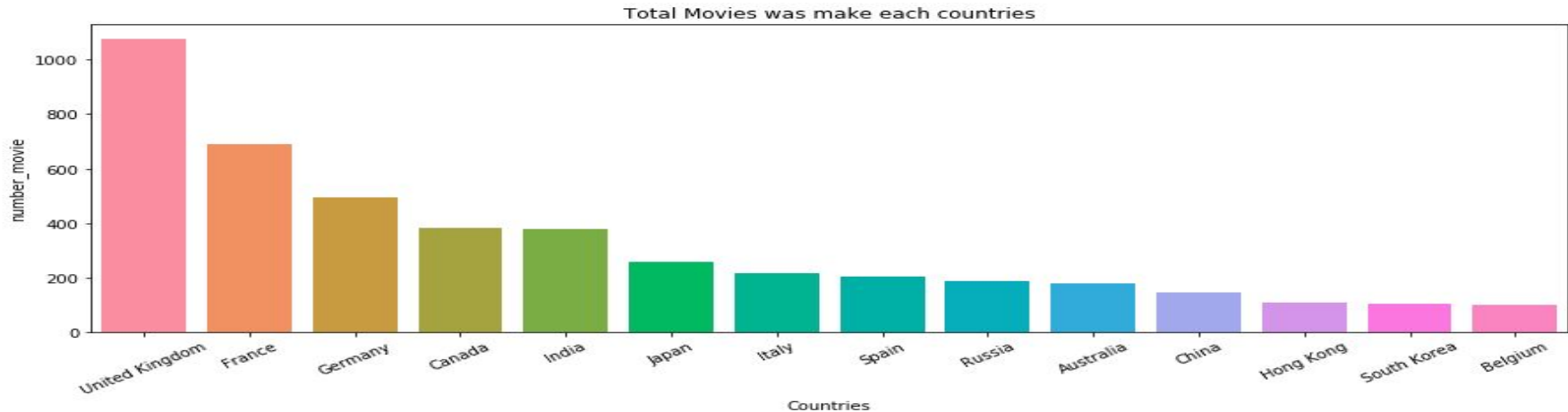
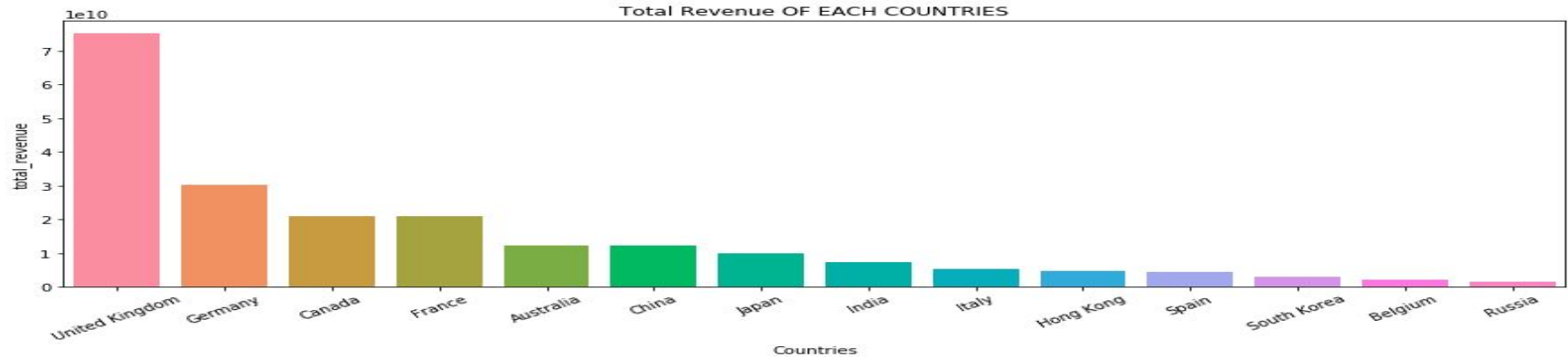
Year last movie in our data set

	year	title	runtime
174446	2019	Sunday	13.0
174441	2019	Queen + Béjart - Ballet For Life	58.0
126731	2019	The Ocean Washed Open Your Grave	3.0
174467	2019	Entropia	28.0
210551	2019	Jorge	20.0

Total, Avg. revenue of each countries

	average_revenue	total_revenue	number_movie
Countries			
United States of America	7.714978e+07	4.976932e+11	6451
United Kingdom	6.990096e+07	7.514353e+10	1075
France	3.021624e+07	2.084921e+10	690
Germany	6.078097e+07	3.014736e+10	496
Canada	5.485054e+07	2.089806e+10	381
India	1.911960e+07	7.246328e+09	379
Japan	3.894279e+07	9.969354e+09	256
Italy	2.382187e+07	5.169346e+09	217
Spain	2.157035e+07	4.378782e+09	203
Russia	8.155989e+06	1.517014e+09	186

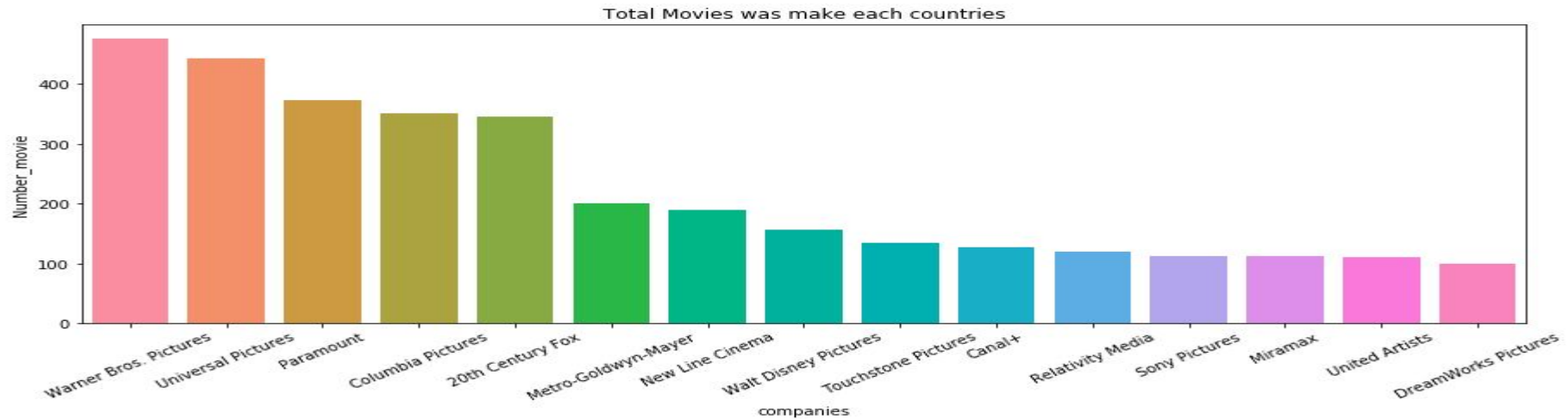
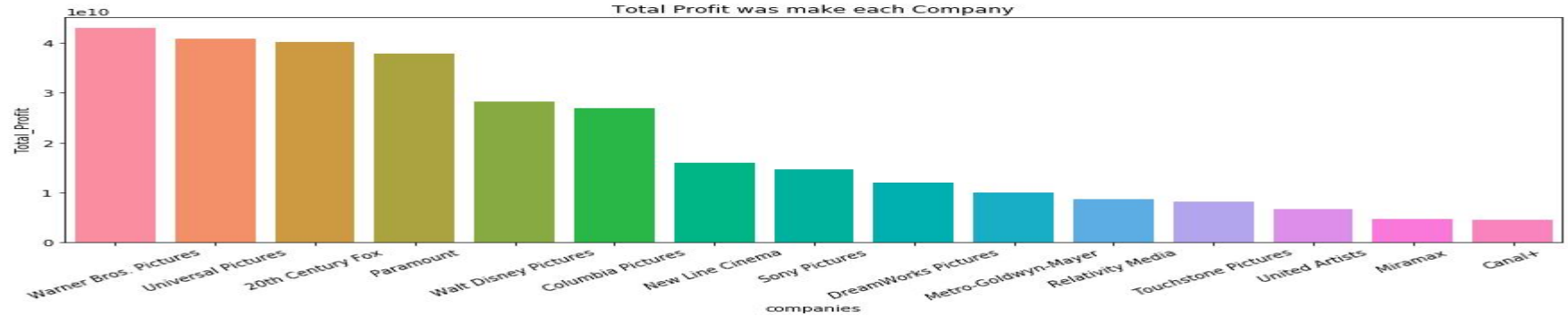
Total, Avg. revenue of each countries (continous)



Total, Avg profit of each company

	Avergae_profit	Total_Profit	Number_movie
companies			
Warner Bros. Pictures	9.032809e+07	4.299617e+10	476
Universal Pictures	9.229253e+07	4.088559e+10	443
Paramount	1.014296e+08	3.793466e+10	374
Columbia Pictures	7.684656e+07	2.697314e+10	351
20th Century Fox	1.164543e+08	4.017672e+10	345
Metro-Goldwyn-Mayer	5.015126e+07	1.003025e+10	200
New Line Cinema	8.415754e+07	1.598993e+10	190
Walt Disney Pictures	1.797291e+08	2.821747e+10	157
Touchstone Pictures	6.081461e+07	8.209972e+09	135
Canal+	3.521476e+07	4.437060e+09	126

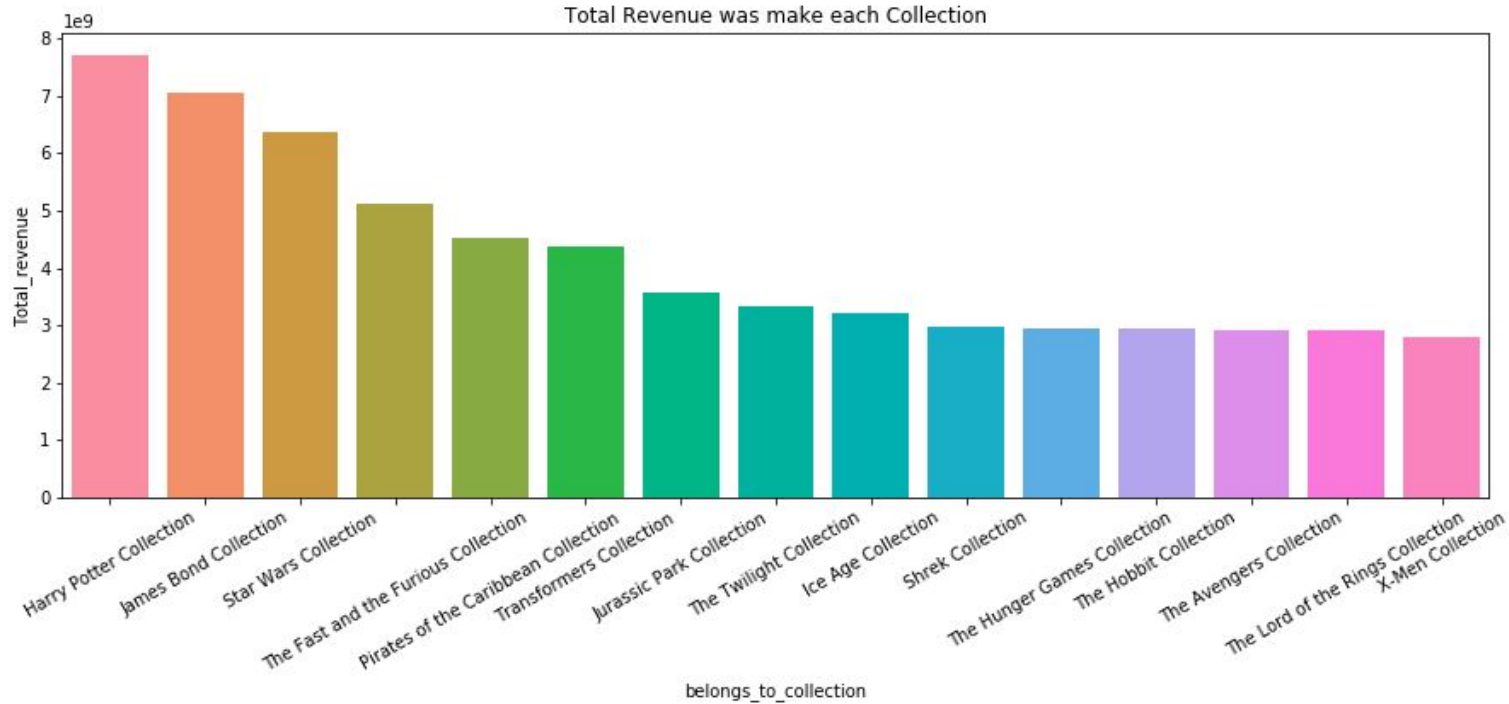
Total, Avg profit of each company (continous)



Total, Avg. of each collection of movie

	Average_revenue	Total_revenue	Total_movie
belongs_to_collection			
Harry Potter Collection	9.633598e+08	7.706879e+09	8
James Bond Collection	2.827486e+08	7.068715e+09	25
Star Wars Collection	9.112054e+08	6.378438e+09	7
The Fast and the Furious Collection	6.406373e+08	5.125099e+09	8
Pirates of the Caribbean Collection	9.043154e+08	4.521577e+09	5
Transformers Collection	8.758574e+08	4.379287e+09	5
Jurassic Park Collection	8.948083e+08	3.579233e+09	4
The Twilight Collection	6.686215e+08	3.343107e+09	5
Ice Age Collection	6.433533e+08	3.216767e+09	5
Shrek Collection	7.411823e+08	2.964729e+09	4

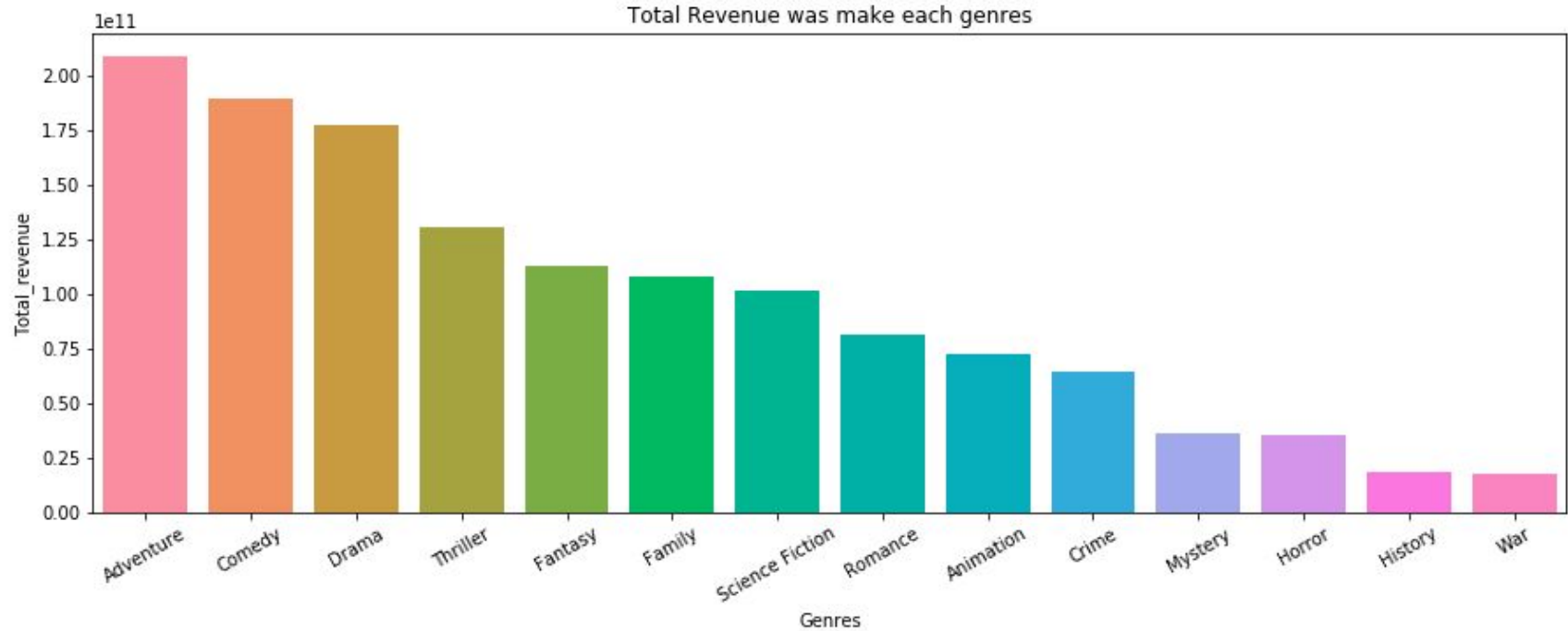
Total, Avg. of each collection of movie (continous)



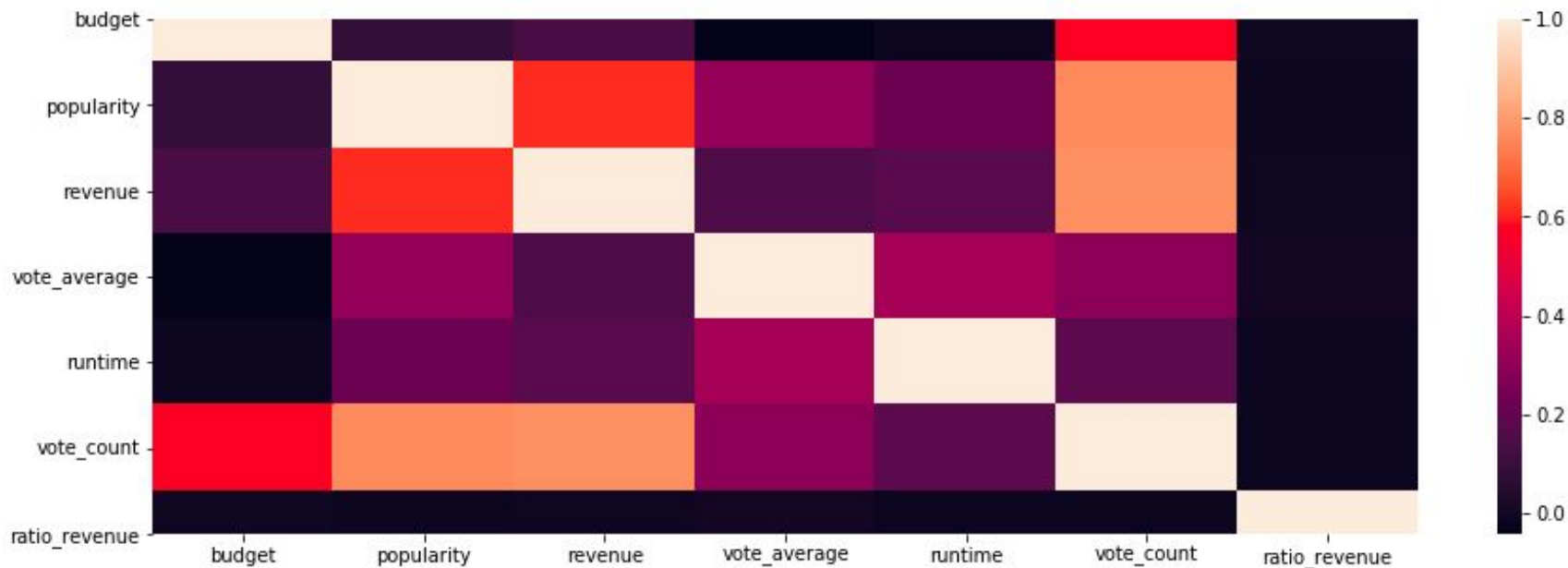
Total, Avg. of Genres of movie

	Avergae_revenue	Total_revenue	Total_movie
Genres			
Action	1.025900e+08	2.134897e+11	2081
Adventure	1.640720e+08	2.086996e+11	1272
Comedy	5.798689e+07	1.889213e+11	3258
Drama	3.835340e+07	1.773078e+11	4623
Thriller	6.120269e+07	1.302393e+11	2128
Fantasy	1.463286e+08	1.122341e+11	767
Family	1.299991e+08	1.080293e+11	831
Science Fiction	1.181741e+08	1.016297e+11	860
Romance	4.568787e+07	8.073046e+10	1767
Animation	1.467283e+08	7.189686e+10	490

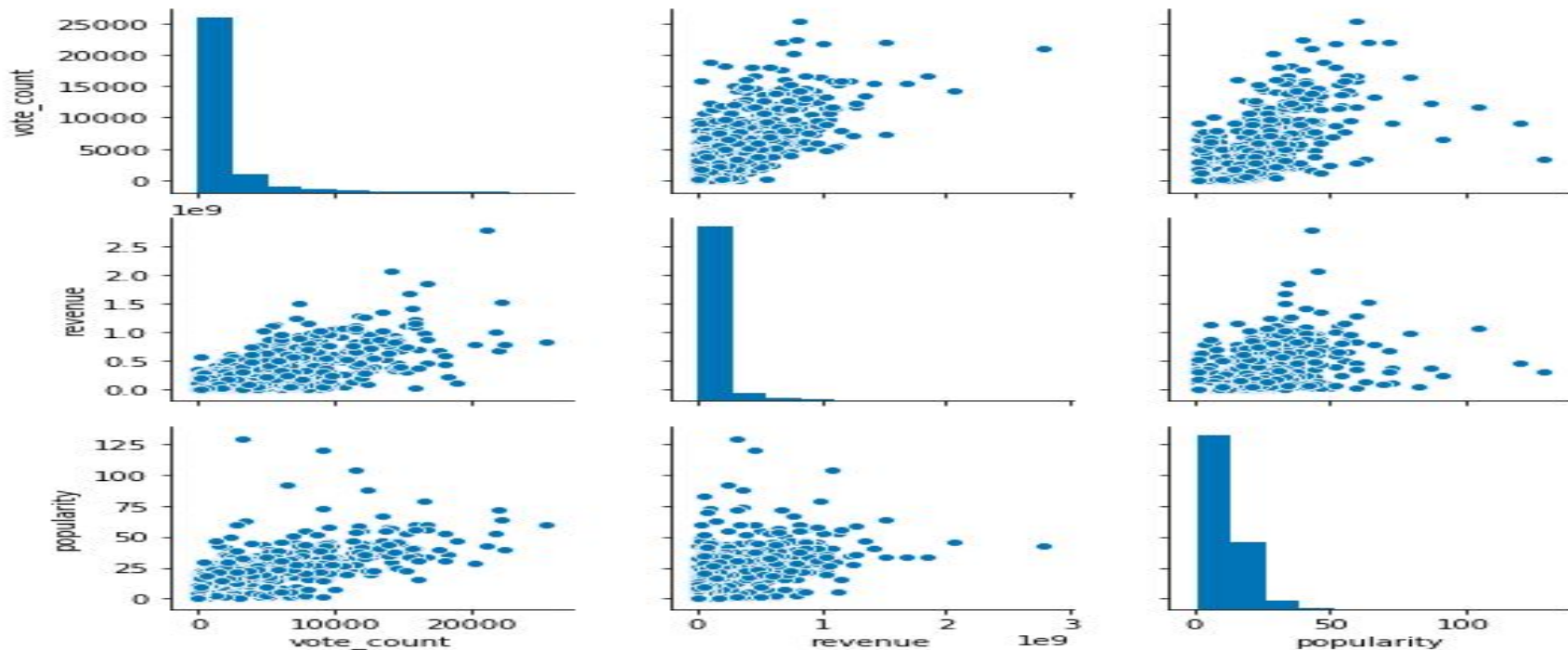
Total, Avg. of Genres of movie (continuous)



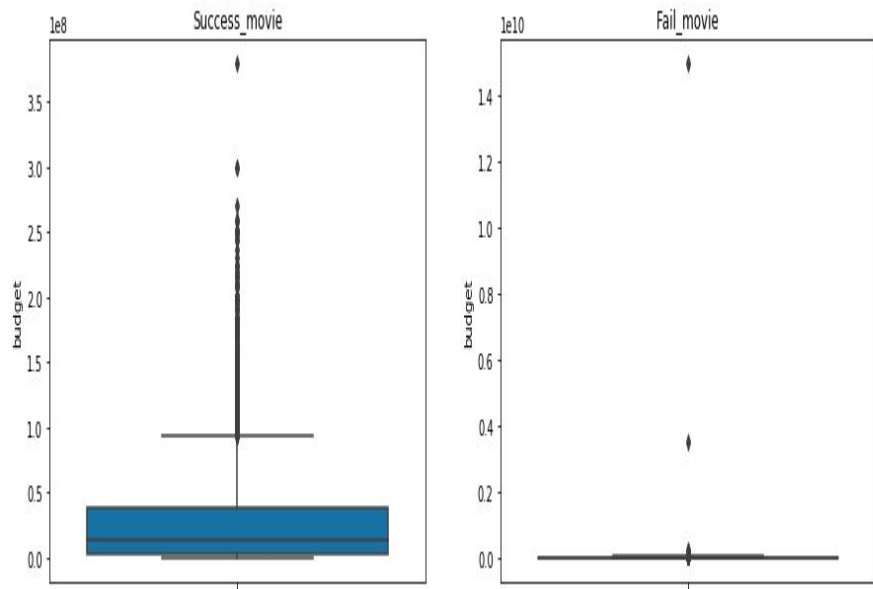
Heatmap relation of each column



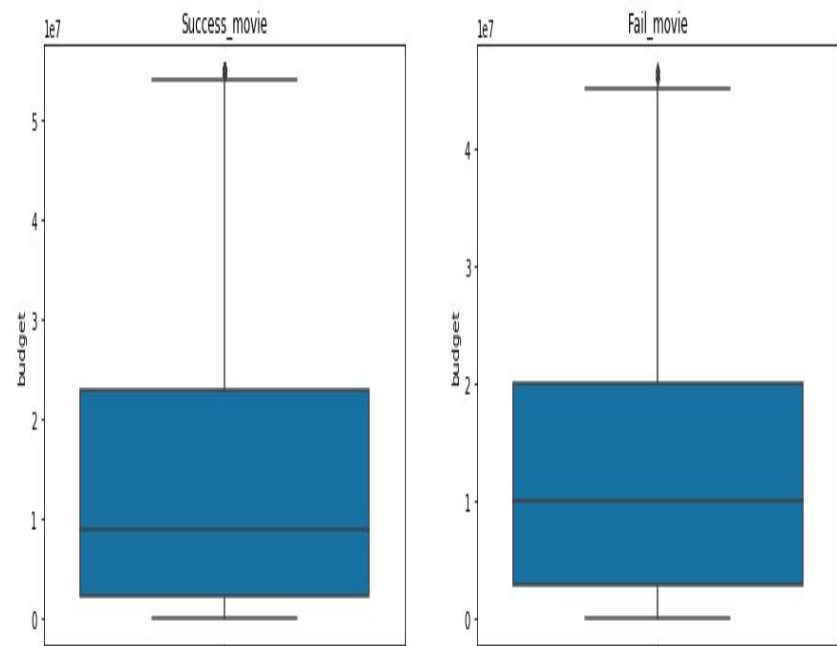
Pairplot for vote_count, revenue, popularity



Boxplot detect outlier and remove



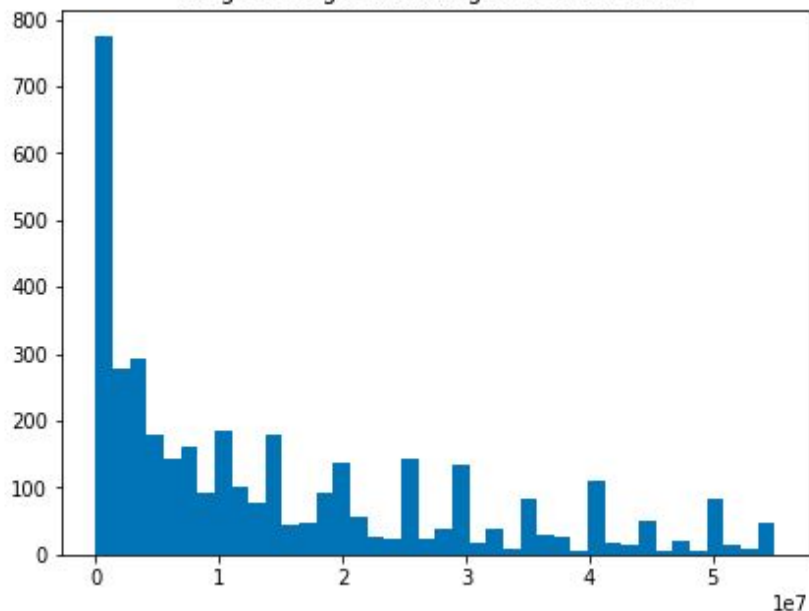
Before remove outlier



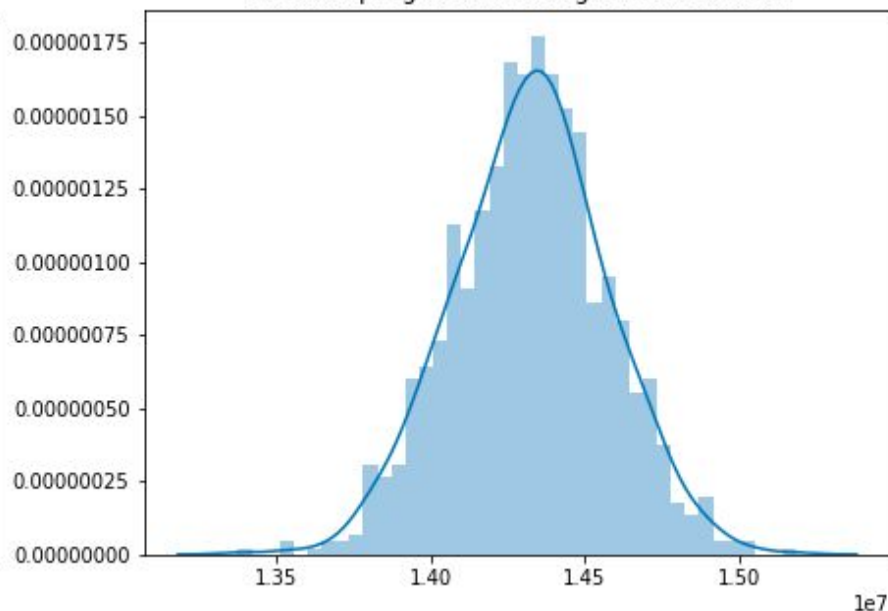
After remove outlier

Resampling mean of success movie

Origin histogram of budget success movie

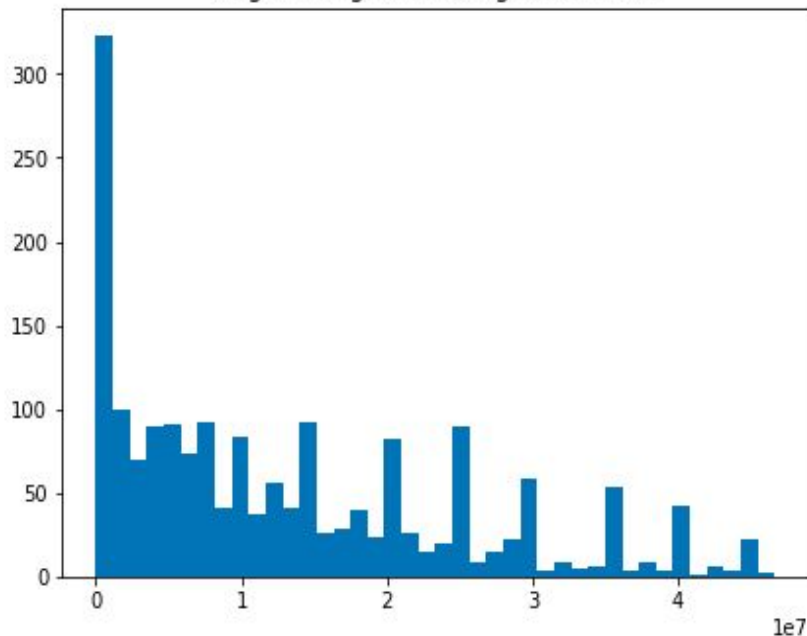


After sampling mean of budget success movie

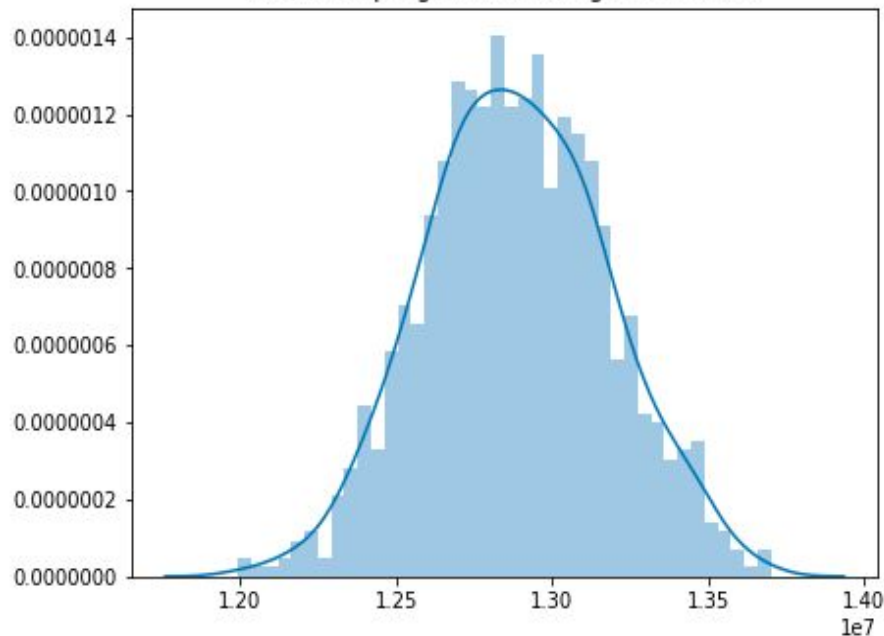


Resampling mean of fail movie

Origin histogram of budget fail movie



After sampling mean of budget fail movie



Different mean of two resampling.

