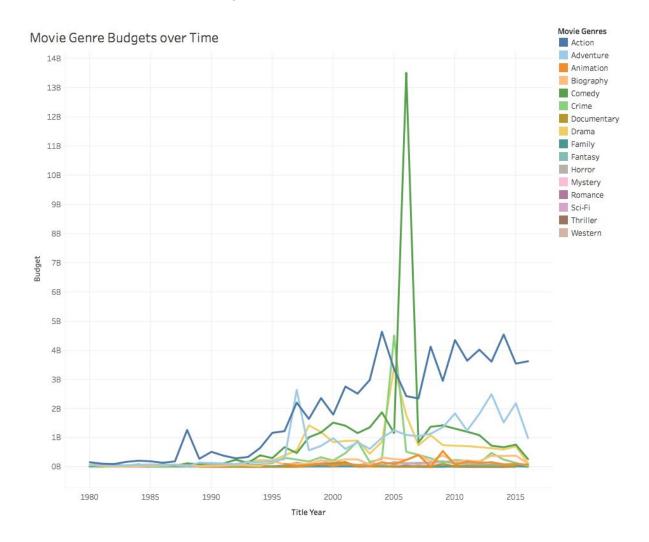
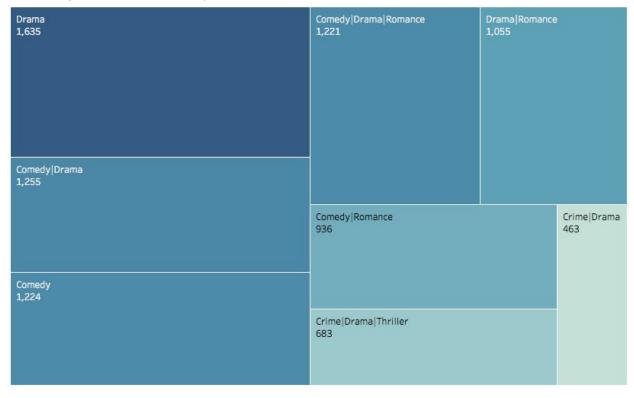
Some background on the IMDB 5000 dataset. Includes data on movies from 1916-2016.

Movie genre budget have seen interesting trends over time-- noticeably, there's a spike near 2007 for comedy movies. Meanwhile Dramas budgets saw a spike near 2005, but have been in decline since. Action movie budgets are on the rise...



In contrast to the budget observation for the Drama genre, it is the highest rated genre amongst the top 8 genres as summed by IMDB Scores



Apart from genre related exploration: sum of IMDB Scores colored and sorted by keywords of a movie plot; most popular movies are based on books or novels. This query filters out results that were Null.

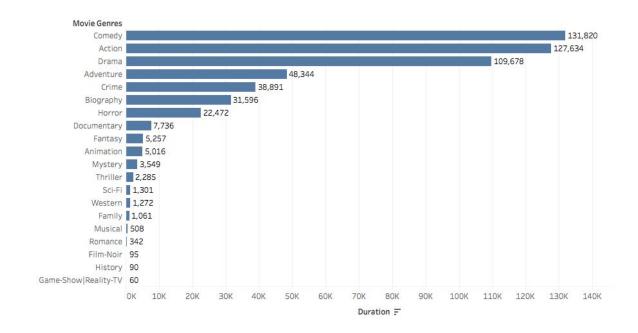
Plot Keywords	
based on novel	25.40
halloween masked killer michael myers slasher trick or treat	23.70
animal name in title ape abducts a woman gorilla island king kong	
alien friendship alien invasion australia flying car mother daughter relationship	
eighteen wheeler illegal street racing truck trucker undercover cop	20.10
meaning of life moral challenge morality search for meaning ten commandments	18.20
assistant experiment frankenstein medical student scientist	18.00
one word title	17.80
1940s child hero fantasy world orphan reference to peter pan	17.40
boxer boxing diamond fight gypsy	16.60
blood godzilla monster sequel	16.40
alien invasion assassin battle iron man soldier	16.20
casino espionage free running james bond terrorist	16.00
cemetery farmhouse radiation running out of gas zombie	16.00
army black soldier military uniform prejudice u.s. civil war	
china flying king palace sword	
race relations racism racist social problem stereotype	
brawl childhood home computer cracker intelligence agency terrorist cell	
classic literature denmark funeral madness prince	15.60
coroner death gay son love mistaken identity	15.60
crime boss drug smuggling heroin mafia boss organized crime	15.60
fire jungle remake tiger wolf	15.60
1980s colonel deputy rambo sheriff	15.40
catholic corruption emotional abuse france seduction	15.40
lolita nymphet older man young girl relationship professor writer	15.40

The bottom-tail of data for the same query as above:

Plot Keywords

Plot Reywords	
cancun mexico mexico spring break vacation wisconsin	2.70
female protagonist greek overweight spoof wedding planner	2.70
king persian spartan training xerxes	
masturbation porn actor in mainstream movie sex spoof virginity	
exorcism female nudity haunted house lesbian kiss masturbation	
baby boy science scientist toddler	2.50
church competition high school new school pastor	2.50
alien bad acting pulp fiction resistance fighter voyager golden record	2.40
gangster hospital lesbian mob thug	2.40
artifact creature darkness demon museum	2.30
box office flop company invention roommate surfer	2.30
chocolate chocolate factory factory orphan student	2.30
baby cartoon on tv cartoonist dog mask	2.20
military submarine suspense time travel world war two	2.20
reference to snow white sleeping bag stepmother stepdaughter relationship stra	2.20
african american basketball basketball movie critically bashed tv commercial	2.10
bar florida love spring break waitress	2.10
cult film disc jockey female protagonist music business one word title	2.10
2000s golf caddy golf course golfer question mark in title	2.00
20th century association fifa football world cup	2.00
cnn reporter ellipsis in title matrix punctuation in title secret agent	1.90
computer good versus evil mind control science experiment sequel	1.90
disaster glass mousetrap natural disaster scatological humor	1.90
femme fatale food pun supermarket troubled production	1.70
boyhood friend manager plasma tv prodigy star	1.60

Back to genres. A comparison of movie genres by duration shows that comedy movies are typically the longest, while strictly historical films are the shortest in duration. The x-scale is in minutes.

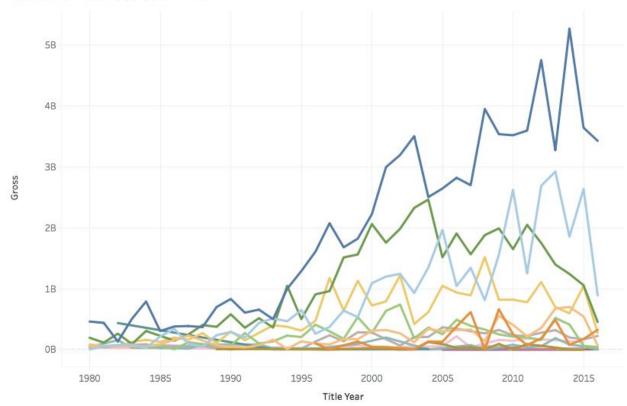


Action and Comedy movies seem to make the most money, as measured by gross revenue generated over opening weekend. This explains the increase in their respective budgets over time.

Action	75,145,540,469
Comedy	40,249,689,383
Adventure	33,580,719,712
Drama	21,817,363,043
Crime	8,517,148,868
Biography	7,885,126,728
Horror	5,851,150,828
Animation	4,265,703,221
Fantasy	1,511,309,412
Mystery	1,171,735,110
Documentary	626,566,503
Family	447,481,433
Sci-Fi	208,574,725
Musical	188,119,773
Romance	62,659,016
Western	47,972,862
Thriller	3,245,818
Film-Noir	
Game-Show Reality-TV	
History	

This plot may as well be a self-fulfilling prophecy: if Action movies are believed to make the most, they will be allocated the most budgets, and thus also make more money over time.

Movie Genre Gross over Time



Based on this EDA it will be interesting to explore predictive algorithms that can determine a movie genre by either duration or budget, or a combination of the two.