



Exploratory Data Analysis Movie Data (2010-2018)

Donnell Jackson, 2020

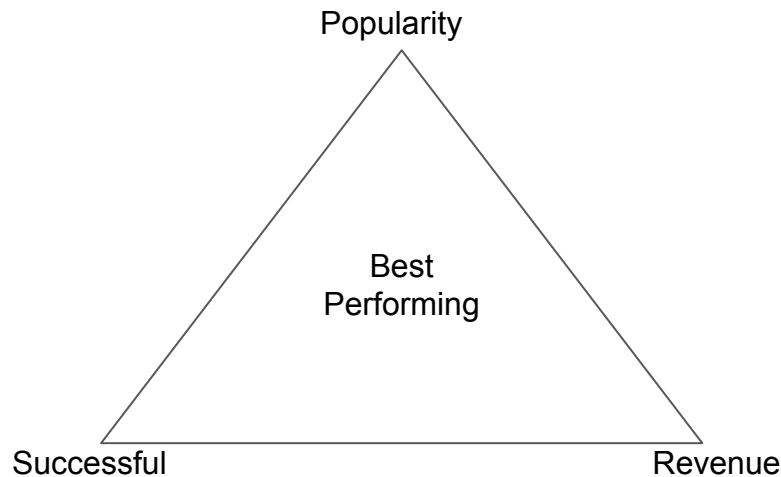


Analysis Overview

Objective: Find out what movies are doing the best in the box office.

My Questions to derive a reliable answer through analyzing the movie data:

1. What movies are most popular?
2. What movies produced the most revenue?
3. What movies are most successful?



The Movie Data

- Collection of various data (IMDB,TN)
- Combined data set (14 columns, 121825 rows)
- Data Categories: Movie Title, Year, Runtime, Ratings, Budget, Gross, Revenue

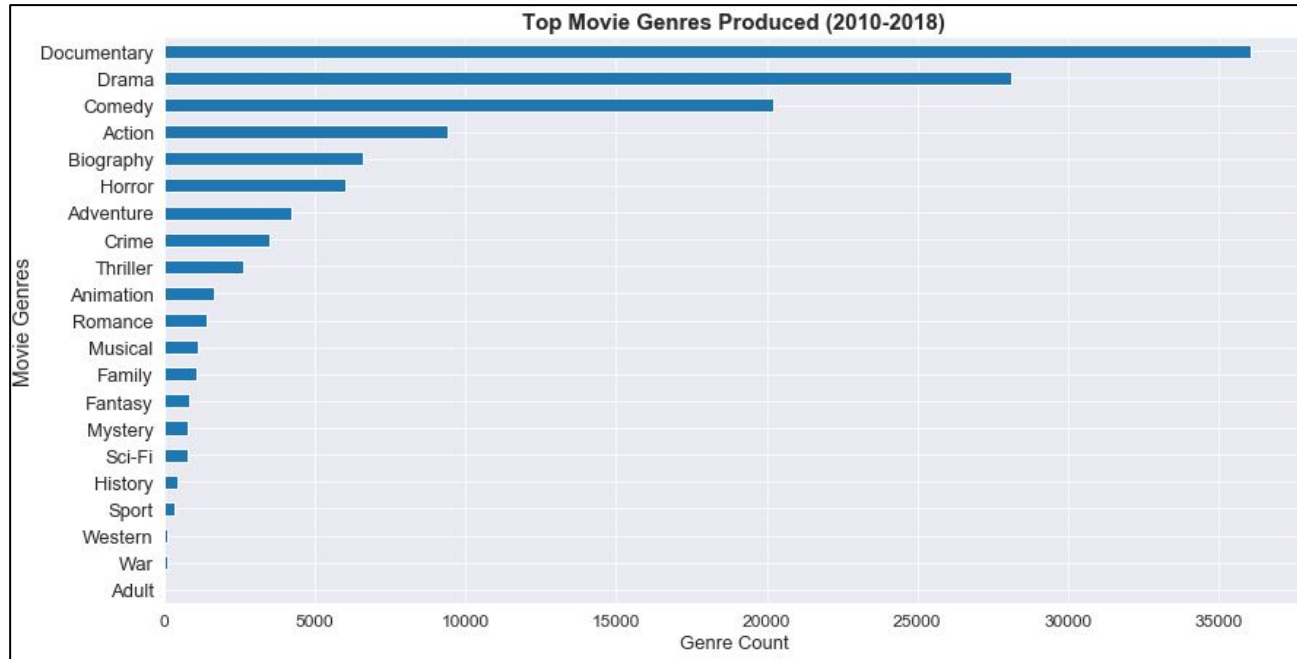
	title	original_title	year	runtime_minutes	genres	averagerating
73	John Carter	John Carter	2012	132.0	Action,Adventure,Sci-Fi	6.6
76	Action Jackson	Action Jackson	2014	144.0	Action,Thriller	3.3
92	Dinner for Schmucks	Dinner for Schmucks	2010	114.0	Comedy	5.9
95	The A-Team	The A-Team	2010	117.0	Action,Adventure,Thriller	6.8
97	The Possession	The Possession	2012	92.0	Horror,Mystery,Thriller	5.9
99	Real Steel	Real Steel	2011	127.0	Action,Drama,Family	7.1
103	The Giver	The Giver	2014	97.0	Drama,Romance,Sci-Fi	6.5
104	Toy Story 3	Toy Story 3	2010	103.0	Adventure,Animation,Comedy	8.3
115	Lincoln	Lincoln	2012	150.0	Biography,Drama,History	7.4
117	Before We Go	Before We Go	2014	95.0	Comedy,Drama,Romance	6.8
118	The History of Love	The History of Love	2016	134.0	Drama,Romance,War	6.3
126	Puss in Boots	Puss in Boots	2011	90.0	Action,Adventure,Animation	6.6
131	Wonder Woman	Wonder Woman	2017	141.0	Action,Adventure,Fantasy	7.5
134	42	42	2013	128.0	Biography,Drama,Sport	7.5
136	Life of Pi	Life of Pi	2012	127.0	Adventure,Drama,Fantasy	7.9
138	The Crazies	The Crazies	2010	101.0	Horror,Thriller	6.5
139	The Equalizer	The Equalizer	2014	132.0	Action,Crime,Thriller	7.2
141	Captain America: The First Avenger	Captain America: The First Avenger	2011	124.0	Action,Adventure,Sci-Fi	6.9
143	Sin City: A Dame to Kill For	Sin City: A Dame to Kill For	2014	102.0	Crime,Thriller	6.5

What movies are most popular?

Data Categories: Genre

Methods: Summation, Grouping

Findings: Movies Genres - Documentary, Drama, Comedy, Action, Biography



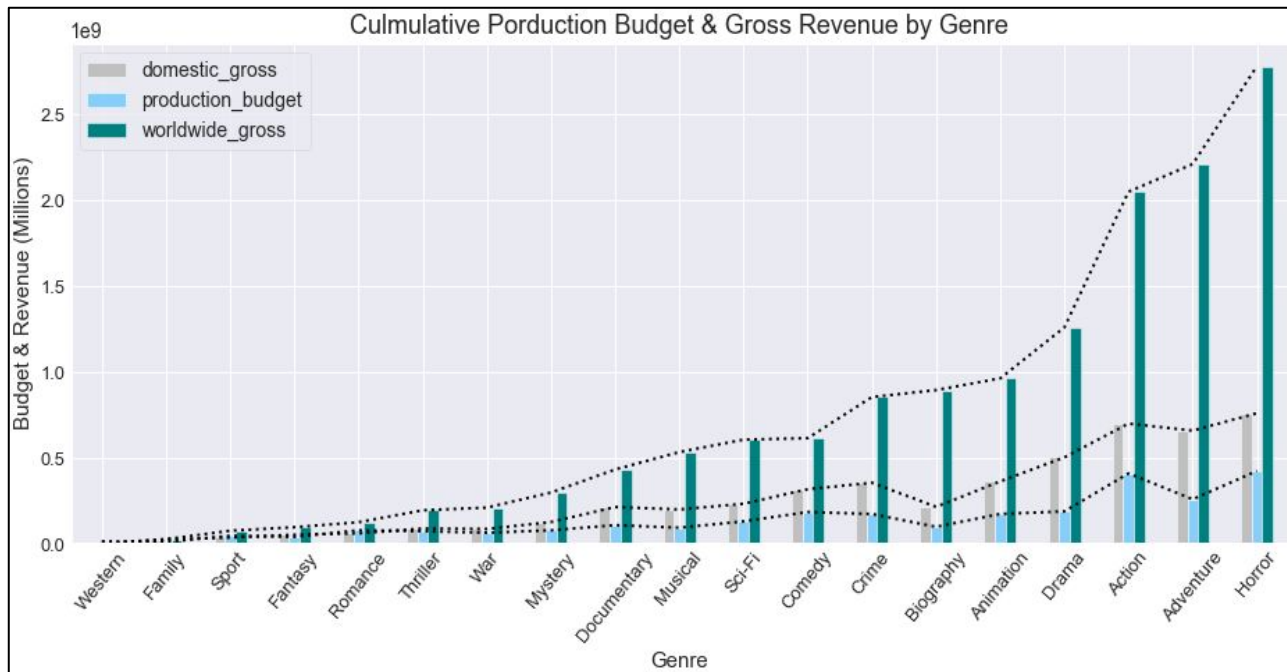
prime_genre	title_count	percentage
Documentary	34158	28.04
Drama	26229	21.53
Comedy	18923	15.53
Action	8675	7.12
Biography	6341	5.21
Horror	5480	4.50
Adventure	3950	3.24
Crime	3216	2.64
Thriller	2304	1.89
Animation	1524	1.25
Romance	1311	1.08
Musical	1058	0.87
Family	957	0.79
Fantasy	762	0.63
Mystery	720	0.59
Sci-Fi	680	0.56
History	418	0.34
Sport	333	0.27
War	113	0.09
Western	112	0.09
Adult	22	0.02

What movies produced the most revenue?

Data: Financial(Budget,Domestic Gross, International Gross)

Methods: Grouping, Summation

Findings: Movie Genres: Horror, Adventure, Action, Drama, Animation



What movies are most successful?

Top 5 Ratings by Genre for 2010



Top 5 Ratings by Genre for 2011



Top 5 Ratings by Genre for 2012



Top 5 Ratings by Genre for 2013



Top 5 Ratings by Genre for 2014



Top 5 Ratings by Genre for 2015



Top 5 Ratings by Genre for 2016



Top 5 Ratings by Genre for 2017



Top 5 Ratings by Genre for 2018

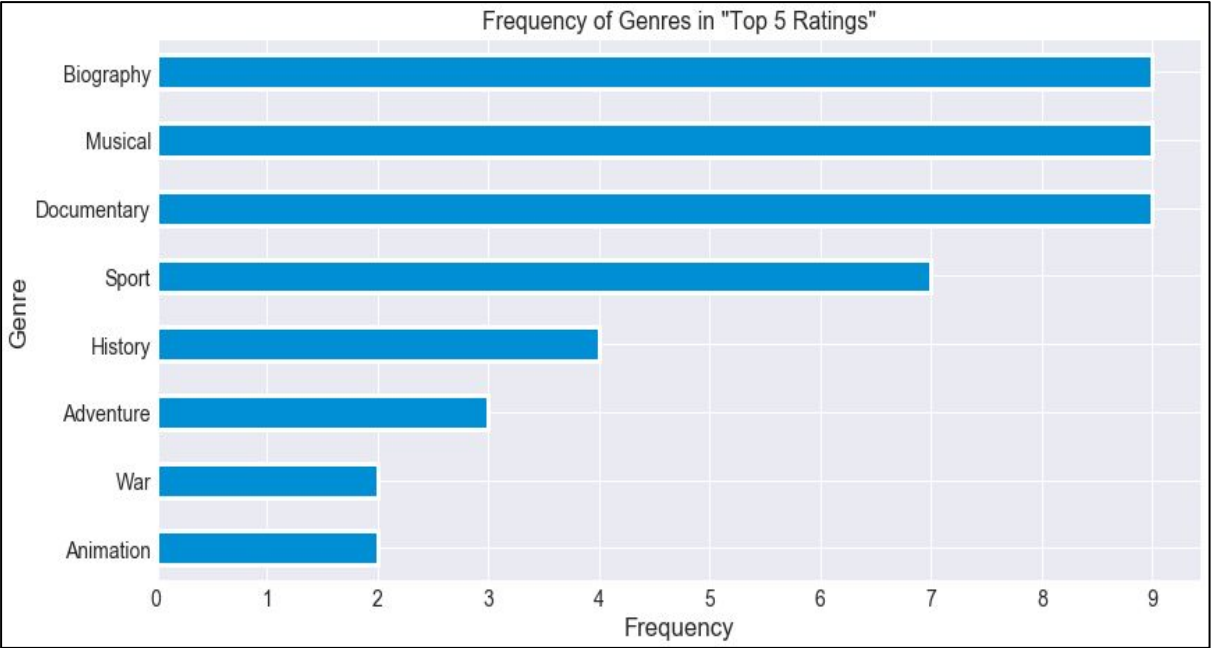


Data: Genre, Average Rating

Methods: Averaging, Summation

Findings: Documentary, Adventure, Musical, Animation, Biography, Sport, History, Animation, Action

prime_genre	domestic_gross	production_budget	worldwide_gross
Western	1686429.0	1600000	2684207
Family	27467564.0	1200000	27936778
Sport	37562568.0	4800000	77562568
Fantasy	54239856.0	4200000	97778356
Romance	62495645.0	8000000	126931325
Thriller	90380162.0	7500000	196439693
War	87707396.0	6600000	211700000
Mystery	128012934.0	8000000	299461782
Documentary	215288866.0	11000000	433449571
Musical	201151353.0	9500000	534551353
Sci-Fi	234280354.0	13200000	606836535
Comedy	318604126.0	18600000	615461394
Crime	356461711.0	17500000	854235992
Biography	216303339.0	10000000	894985342
Animation	364001123.0	17500000	962854547
Drama	504014165.0	19000000	1259199706
Action	700059566.0	41060000	2048134200
Adventure	659363944.0	26000000	2147483648
Horror	760507625.0	42500000	2147483648



Recommendations

Focus movies involving the following genres:

- Documentary
- Adventure
- Musical
- Animation
- Biography
- History
- Action

Future Work

- Further analyzing financial data
- Relationship between movie revenue and genres
- Relationship between revenue and runtime
- Deeper analysis of runtime dispersion in relation to ratings

Thank You

Your time taken to review my analysis and findings is greatly appreciated.