TMDb Movie Profitability Analysis

Background

What can we say about the success of a movie before it is released? Are there certain companies (Pixar?) that have found a consistent formula? Given that major films costing over $100 million to produce can still flop, this question is more important than ever to the industry. Film aficionados might have different interests. Can we predict which films will be highly rated, whether or not they are a commercial success?

This is a great place to start digging in to those questions, with data on the plot, cast, crew, budget, and revenues of several thousand films.

Click *here* to check out data source from Kaggle.

1 Summary Statistic

Below are statistics for quantitative variables and release dates.

```
##   release_date          budget          revenue        profitability
## Min.   :1916-09-04  Min.   :  0.00  Min.   :   0.00  Min.   :-165.71
## 1st Qu.:1999-07-14  1st Qu.:  0.80  1st Qu.:   0.00  1st Qu.:  -0.80
## Median :2005-10-03  Median : 15.00  Median :  19.18  Median :   2.53
## Mean   :2002-12-27  Mean   : 29.05  Mean   :  82.28  Mean   :  53.23
## 3rd Qu.:2011-02-16  3rd Qu.: 40.00  3rd Qu.:  92.92  3rd Qu.:  55.34
## Max.   :2017-02-03  Max.   :380.00  Max.   :2787.97  Max.   :2550.97
##     runtime        popularity       vote_average     vote_count
## Min.   :  0.0   Min.   :  0.0004   Min.   : 0.000   Min.   :    0.0
## 1st Qu.: 94.0   1st Qu.:  4.6717   1st Qu.: 5.600   1st Qu.:   54.0
## Median :103.0   Median : 12.9249   Median : 6.200   Median :  235.5
## Mean   :106.9   Mean   : 21.4968   Mean   : 6.093   Mean   :  690.4
## 3rd Qu.:117.8   3rd Qu.: 28.3320   3rd Qu.: 6.800   3rd Qu.:  737.0
## Max.   :338.0   Max.   :875.5813   Max.   :10.000   Max.   :13752.0
```

The data sets include movies from Sept 1916 to Feb 2017. Highest budget is $ 380 million, revenue $ 2,788 million and the most profitable movie was $ 2,551 million in profit.

2 Qualitative Analysis

2.1 Director Effect

For a start, a movie is born out of wisdom / ideas of a director. If wisdom or brilliant ideas of a director give rise to successful movie, I call this, the ***Director Effect***!!

Below tables are Top 20 directors of profitable movies (at least $ 50 million) & loss-making movies (at least $ 1 million).

```
## # A tibble: 20 x 3
##    `Director Name` Title                                       Profitability
##    <chr>          <chr>                                               <dbl>
##  1 James Cameron   Avatar                                              2551.
##  2 James Cameron   Titanic                                             1645.
##  3 Colin Trevorrow Jurassic World                                      1364.
##  4 James Wan       Furious 7                                           1316.
##  5 Joss Whedon     The Avengers                                        1300.
##  6 Joss Whedon     Avengers: Age of Ultron                             1125.
##  7 Chris Buck      Frozen                                              1124.
##  8 Jennifer Lee    Frozen                                              1124.
##  9 Kyle Balda      Minions                                             1083.
## 10 Pierre Coffin   Minions                                             1083.
## 11 Peter Jackson   The Lord of the Rings: The Return of the King       1025.
## 12 Shane Black     Iron Man 3                                          1015.
## 13 Michael Bay     Transformers: Dark of the Moon                       929.
## 14 Sam Mendes      Skyfall                                              909.
## 15 Anthony Russo   Captain America: Civil War                           903.
## 16 Joe Russo       Captain America: Civil War                           903.
## 17 Pierre Coffin   Despicable Me 2                                       895.
## 18 Chris Renaud    Despicable Me 2                                       895.
## 19 Michael Bay     Transformers: Age of Extinction                      881.
## 20 Lee Unkrich     Toy Story 3                                          867.
```

```
## # A tibble: 20 x 3
##    `Director Name`   Title                   Profitability
##    <chr>             <chr>                           <dbl>
##  1 Gore Verbinski    The Lone Ranger                 -166.
##  2 Joe Johnston      The Wolfman                     -150
##  3 John Lee Hancock  The Alamo                       -119.
##  4 Simon Wells       Mars Needs Moms                 -111.
##  5 James Wong        Dragonball Evolution            -100
##  6 John McTiernan    The 13th Warrior                -98.3
##  7 Ron Underwood     The Adventures of Pluto Nash    -92.9
##  8 Fritz Lang        Metropolis                      -92.0
##  9 Mick Jackson      Volcano                         -90
## 10 Ron Underwood     Mighty Joe Young                -90
## 11 Renny Harlin      Cutthroat Island                -88.0
## 12 Steven Brill      Little Nicky                    -85
## 13 David Bowers      Flushed Away                    -84.5
## 14 Sam Fell          Flushed Away                    -84.5
## 15 Lawrence Guterman Son of the Mask                 -84
## 16 Dean Parisot      RED 2                           -84
## 17 Baz Luhrmann      Australia                       -80.4
## 18 Ron Howard        Edtv                            -80
## 19 Jon Turteltaub    Instinct                        -80
## 20 Peter Chelsom     Town & Country                  -79.6
```

Notice that same name shows up again and again for few block buster movies, just like a top student perform well consistently in a class. For example, **James Cameron** for *Avatar* & *Titanic*, **Michael Bay** for *Transformer* series and **Joss Whedon** for *Marvel Avenger* series.

On the other hand, the name of directors for movies that suffered huge losses is hardly repeatable.

Life's Work of James Cameron, Michael Bay & Joss Whedon
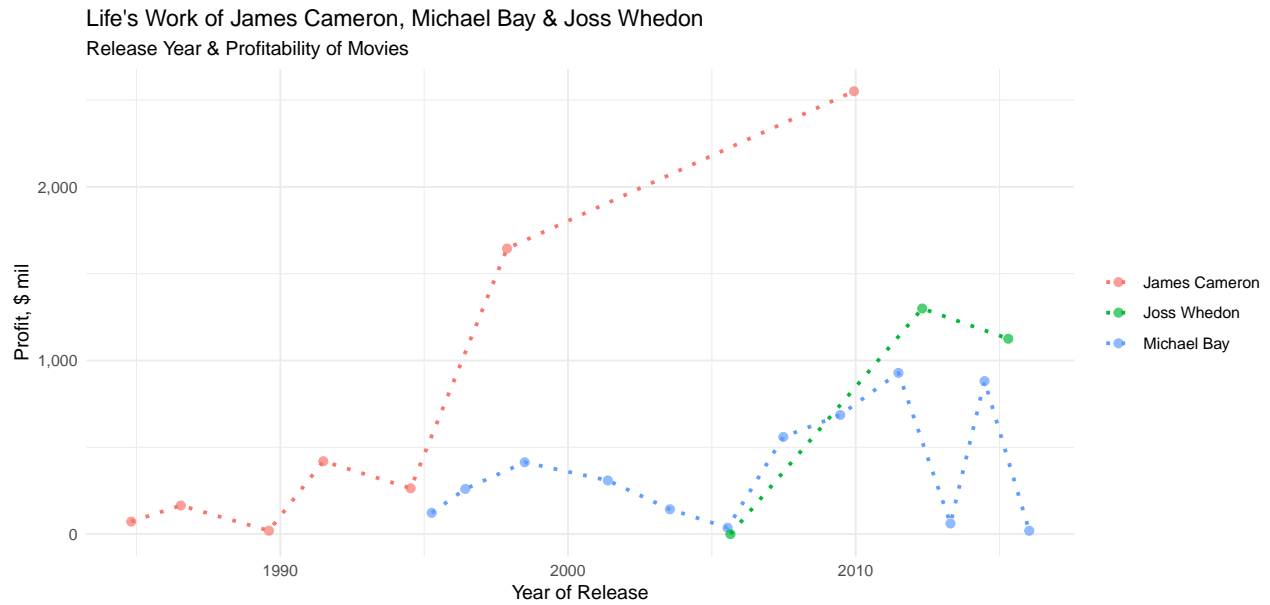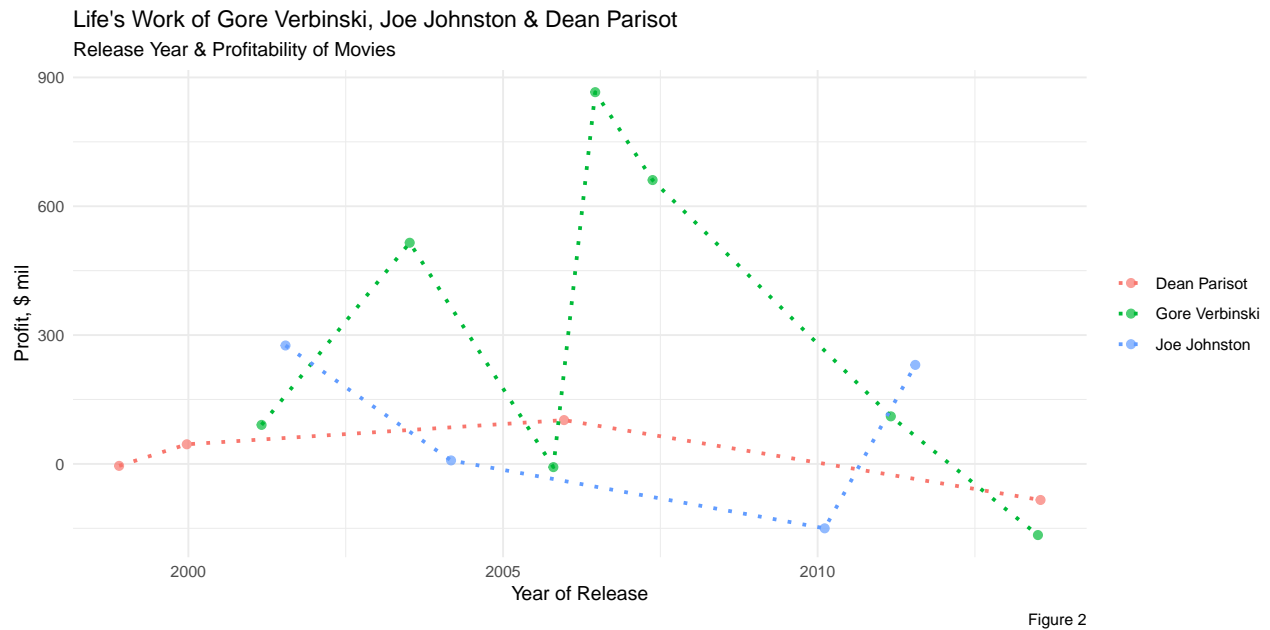Release Year & Profitability of Movies

Figure 1

Figure 1 presented time line of movies for James Cameron, Joss Whedon & Michael Bay. Their brilliant ideas led to production of a couple of block buster movies that made billion of dollars and some movies that barely break even.

Life's Work of Gore Verbinski, Joe Johnston & Dean Parisot
Release Year & Profitability of Movies

Figure 2

When life's work of Gore Verbinski (Figure 2) is presented in similar context, his ideas have been doing quite well actually given the chart shows 3 highly profitable movies.

2.2 Keyword / Idea / Story Line

Let's identify keywords for movies that recorded profit of $ 50 million and above. 75% (3rd Quartile) of movies in the data set recorded profitability of $ 55 million and above, hence I am doing deep dive into keywords of Top 25% of movies.

⊐

Word cloud above highlights *during credits stringer*, *after credits stinger* & *based on novel* as Top 3 keywords out of Top 100 most frequent keywords. This is followed by keywords such as *sequel*, *super hero*, *violence*, *marvel comic*, *dystopia*, *based on comic book*, *women director*, *revenge*, *friendship* and *alien*.

These keywords return some blockbuster movies - *Avatar*, *Jurassic World*, *Marvel Avenger* series, *Frozen* (animation), *Minions* (animation) and *Transformer* series.

```
## # A tibble: 20 x 3
##    ‘Movie Title‘                                    ‘Release Date‘ Profitability
##    <chr>                                            <date>              <dbl>
##  1 Avatar                                           2009-12-10          2551.
##  2 Jurassic World                                   2015-06-09          1364.
##  3 Furious 7                                        2015-04-01          1316.
##  4 The Avengers                                     2012-04-25          1300.
##  5 Avengers: Age of Ultron                          2015-04-22          1125.
##  6 Frozen                                           2013-11-27          1124.
##  7 Minions                                          2015-06-17          1083.
##  8 The Lord of the Rings: The Return of the King    2003-12-01          1025.
##  9 Iron Man 3                                       2013-04-18          1015.
## 10 Captain America: Civil War                       2016-04-27           903.
## 11 Transformers: Age of Extinction                  2014-06-25           881.
## 12 Pirates of the Caribbean: Dead Man’s Chest       2006-06-20           866.
## 13 The Lord of the Rings: The Two Towers            2002-12-18           847.
## 14 Finding Nemo                                     2003-05-30           846.
## 15 The Dark Knight Rises                            2012-07-16           835.
## 16 Alice in Wonderland                              2010-03-03           825.
## 17 The Dark Knight                                  2008-07-16           820.
## 18 The Jungle Book                                  2016-04-07           792.
## 19 The Lord of the Rings: The Fellowship of the Ri~ 2001-12-18           778.
## 20 Harry Potter and the Chamber of Secrets          2002-11-13           777.
```

Now, let’s turn our attention to loss-making movies, particularly movie suffered losses of at least $ 1 Million.

It is quite a surprise that loss-making movies share some common top keywords of profit-making movies: *during credits stinger*, *after credits stinger*, *dystopia*, *violence*, *based on novel*, *women director* and *revenge*. However, *women director*, *independent film* (new) & *murder*(new) emerge as the Top 3 most frequent keywords.

```
## # A tibble: 20 x 3
##    `Movie Title`                                    `Release Date` Profitability
##    <chr>                                            <date>                 <dbl>
##  1 Dragonball Evolution                             2009-04-01              -100
##  2 Metropolis                                       1927-01-10             -92.0
##  3 Australia                                        2008-11-18             -80.4
##  4 Instinct                                         1999-06-04               -80
##  5 How Do You Know                                  2010-12-17             -71.3
##  6 The Great Raid                                   2005-08-12             -69.8
##  7 Monkeybone                                       2001-02-23             -69.6
##  8 R.I.P.D.                                         2013-07-18             -68.4
##  9 Proof of Life                                    2000-12-08               -65
## 10 Soldier                                          1998-10-23             -60.4
## 11 This Is It                                       2009-10-28               -60
## 12 In the Name of the King: A Dungeon Siege Tale    2007-11-29               -60
## 13 Stealth                                          2005-07-28             -58.1
## 14 Blackhat                                         2015-01-13             -52.2
## 15 Final Fantasy: The Spirits Within                2001-07-02             -51.9
## 16 The Iron Giant                                   1999-08-06             -46.8
## 17 Gigli                                            2003-08-01             -46.7
## 18 Child 44                                         2015-03-15             -46.7
## 19 Alfie                                            2004-10-22             -46.6
## 20 The Out-of-Towners                               1999-04-02               -46
```

Table above shows Top 20 loss-making movies based on the keywords. *Final Fantasy: The Spirit Within* is based on the famous *Final Fantasy* game and *Dragonball Evolution* is originated from the well known *Dragon Ball* Japanese comic, however, they suffered huge losses.

2.3 Genres

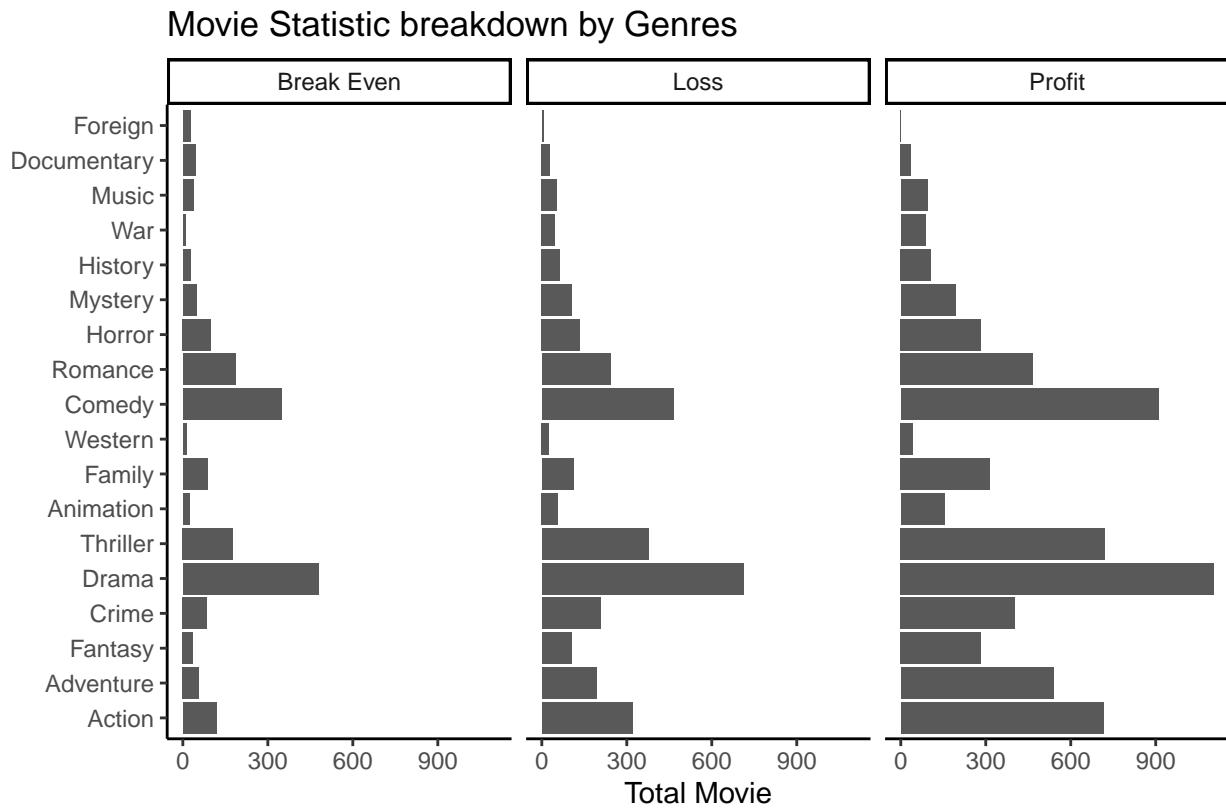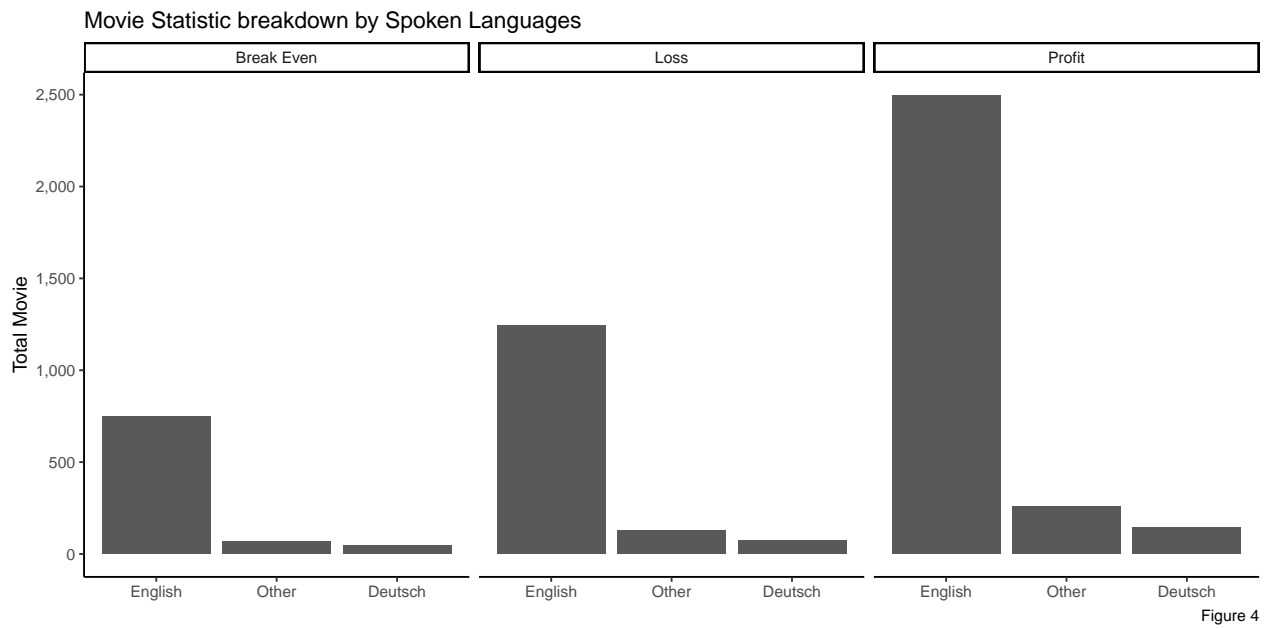Movie Statistic breakdown by Genres

Figure 3

Based on Figure 3, it is clear that genres such as *Comedy*, *Romance*, *Horror*, *Crime*, *Thriller*, *Drama*, *Adventure* and *Action* top the list of profit-making movies. Movies that are loss-making / break-even also share the same genres.

2.4   Language

Movie Statistic breakdown by Spoken Languages

Most movies are English-oriented hence language doesn't matter for profitability either.

2.5 Production Companies

Loss–Making / Break Even / Profit–Making Movie by Production Company
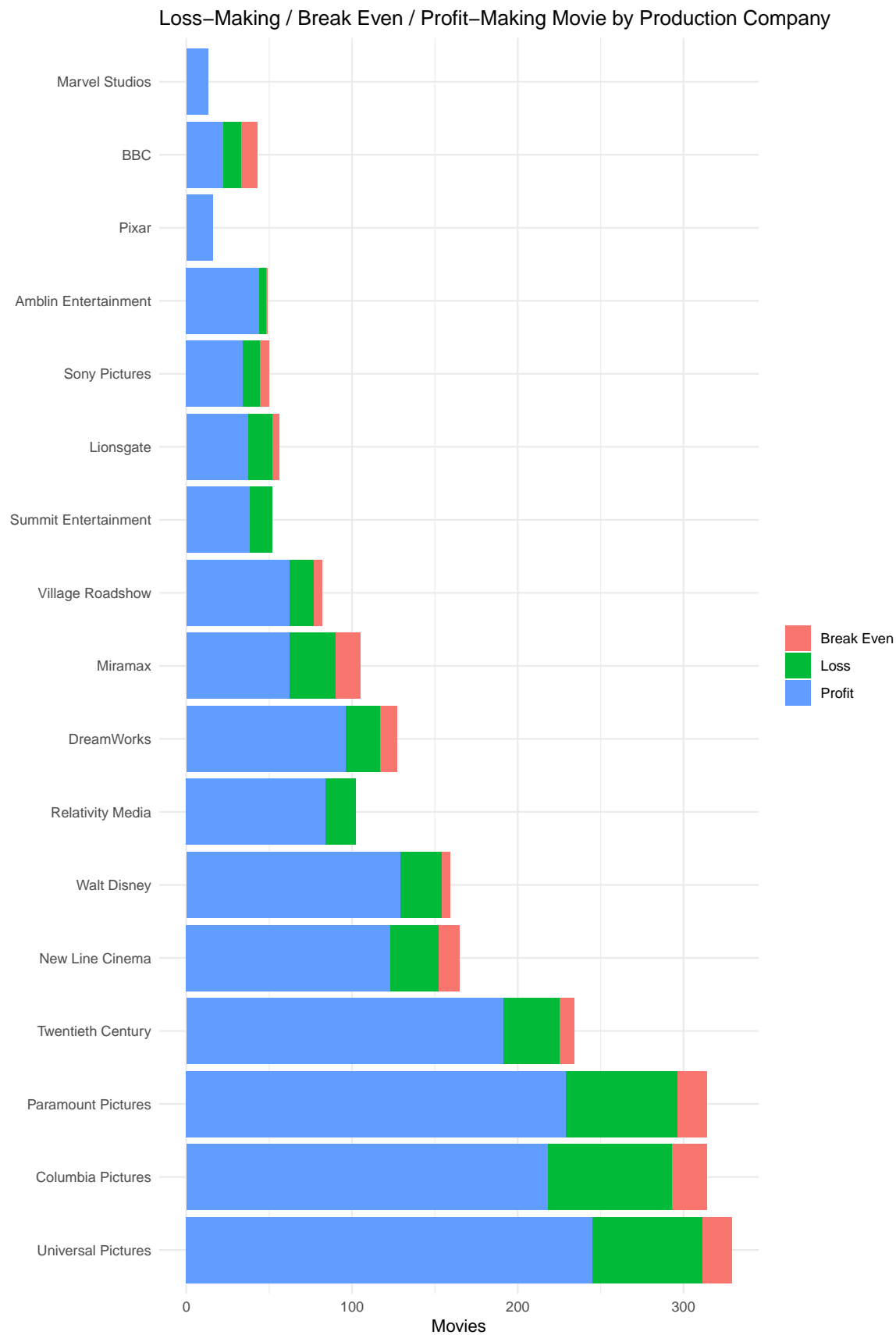
Figure 5

Figure 5 shows globally well-known production companies, for instance, *Universal Pictures*, *Warner Bros* and *Twentieth Century*. These are well established movie producers. Notice all of them produced more profitable movies than loss-making one. *Pixar* and *Marvel Studio* were pretty impressive as all movies produced are profitable, although they produced the least movies among all production companies.

Marvel Studios is well known for the Marvel Avenger series, Iron Man series and other super-hero movies. Pixar have been famous for animation such Wall-E, Toy Story series and Finding Nemo. It is fair to say that even the world's major producers can't guarantee profitability of a movie.

2.6 Casts

Given complexity of a movie, number of actors / actresses involved could be huge. With this in mind, my analysis will only focus on specific individual which has been part of some blockbuster movies, *Tom Cruise* and *Denzel Washington*. Let's find out some of their movies, loss-making or profitable, given their fame. Of course, I believe any actor / actress would have worked for some blockbuster, low quality and okay-okay movies.
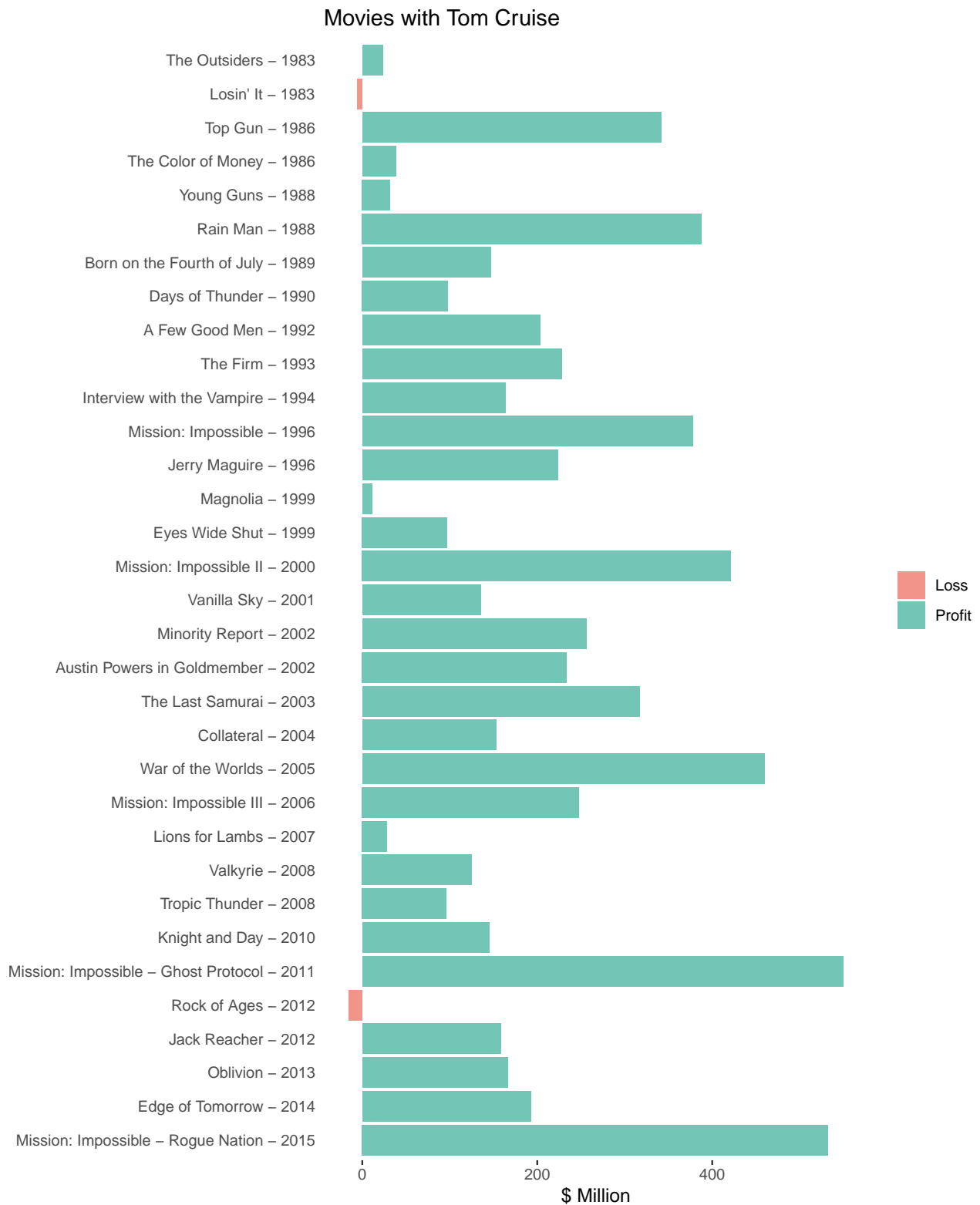
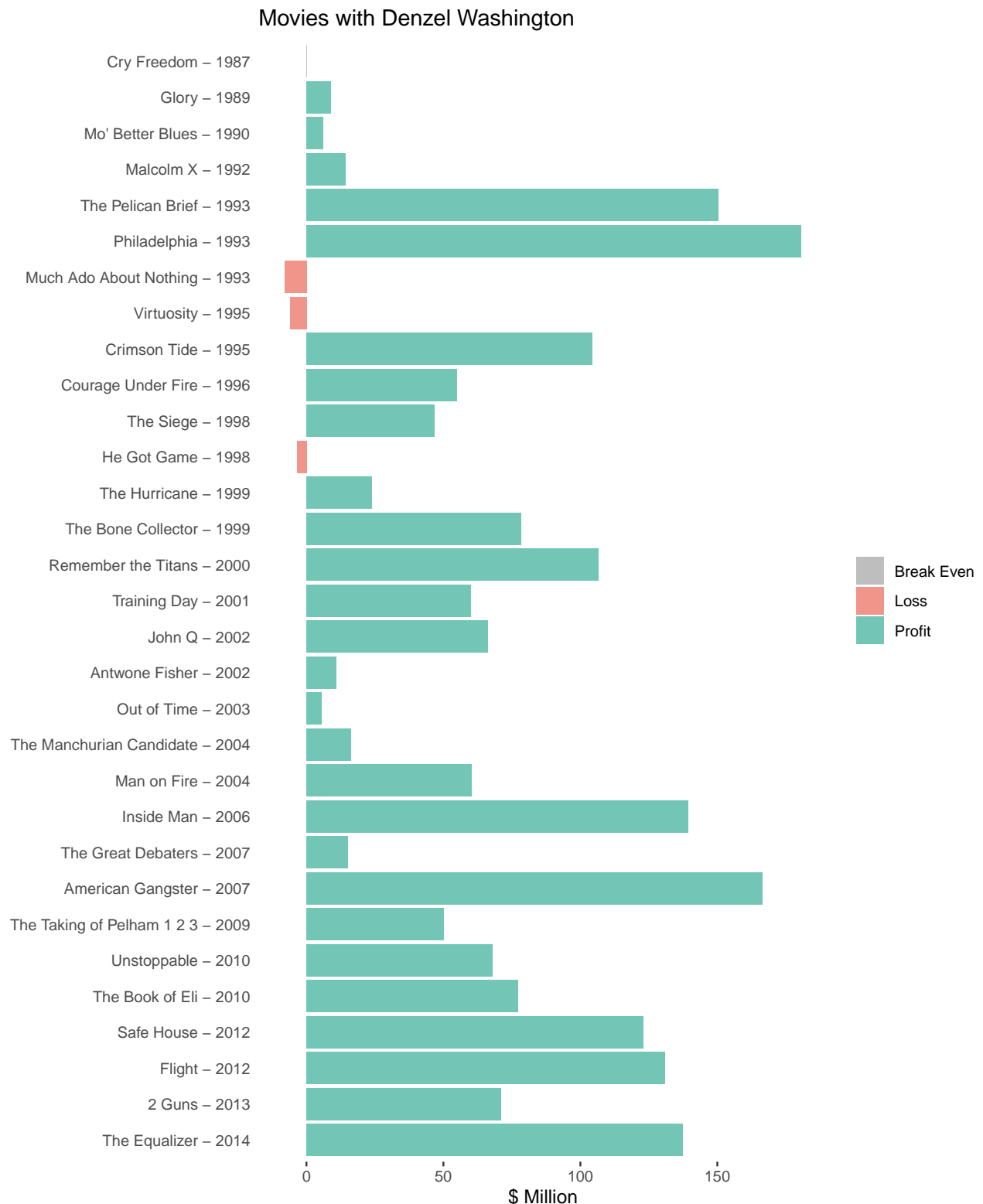# Movies with Tom Cruise

Movie	$ Million
The Outsiders – 1983	
Losin' It – 1983	
Top Gun – 1986	
The Color of Money – 1986	
Young Guns – 1988	
Rain Man – 1988	
Born on the Fourth of July – 1989	
Days of Thunder – 1990	
A Few Good Men – 1992	
The Firm – 1993	
Interview with the Vampire – 1994	
Mission: Impossible – 1996	
Jerry Maguire – 1996	
Magnolia – 1999	
Eyes Wide Shut – 1999	
Mission: Impossible II – 2000	
Vanilla Sky – 2001	
Minority Report – 2002	
Austin Powers in Goldmember – 2002	
The Last Samurai – 2003	
Collateral – 2004	
War of the Worlds – 2005	
Mission: Impossible III – 2006	
Lions for Lambs – 2007	
Valkyrie – 2008	
Tropic Thunder – 2008	
Knight and Day – 2010	
Mission: Impossible – Ghost Protocol – 2011	
Rock of Ages – 2012	
Jack Reacher – 2012	
Oblivion – 2013	
Edge of Tomorrow – 2014	
Mission: Impossible – Rogue Nation – 2015	

Loss
Profit

$ Million

Figure 6

Movies with Denzel Washington

Figure 7

We can observe from Figure 6 (Tom Cruise) & Figure 7 (Denzel Washington) that popular celebrities such as Tom Cruise and Denzel Washington were part of blockbuster movies (Mission Impossible series & Equalizer) that are extremely profitable and also few loss-making movies (Rock of Ages & He Got Game) between 1980s and 2000s. Take Tom Cruise for example. He become a super star after the release of Mission Impossible

series in 1996, 2000 & 2007, but Magnolia (1996), Lions of Lambs (2007) don't share the same fate and Rock of Ages (2012) even suffered losses.

Denzel Washington didn't involved in any movie series like Tom Cruise did between 1980s and 2000s, however, he did deliver consistently profitable movies, except for a few loss-making movies in 1990s.

This mean fame of an actor / actress carry lots of weight but fame alone is not the only factor that gives glory to a movie. Speaking of fame / popularity, how does audience / director / production company evaluates popularity of an actor / actress worth the investment?

3   Quantitative Analysis

3.1   Correlation among key variables

Identify any relationship (positive / negative / none) among the key variables such as *budget*, *revenue*, *runtime*, *vote_average*, *vote_count* with *profitability*. How much each key variable contributes toward *profitability*, if any relationship is identified.

The result is known as **correlation coefficient (r)**. Example: *r = 0.78* indicates **positive** & **strong** relationship.

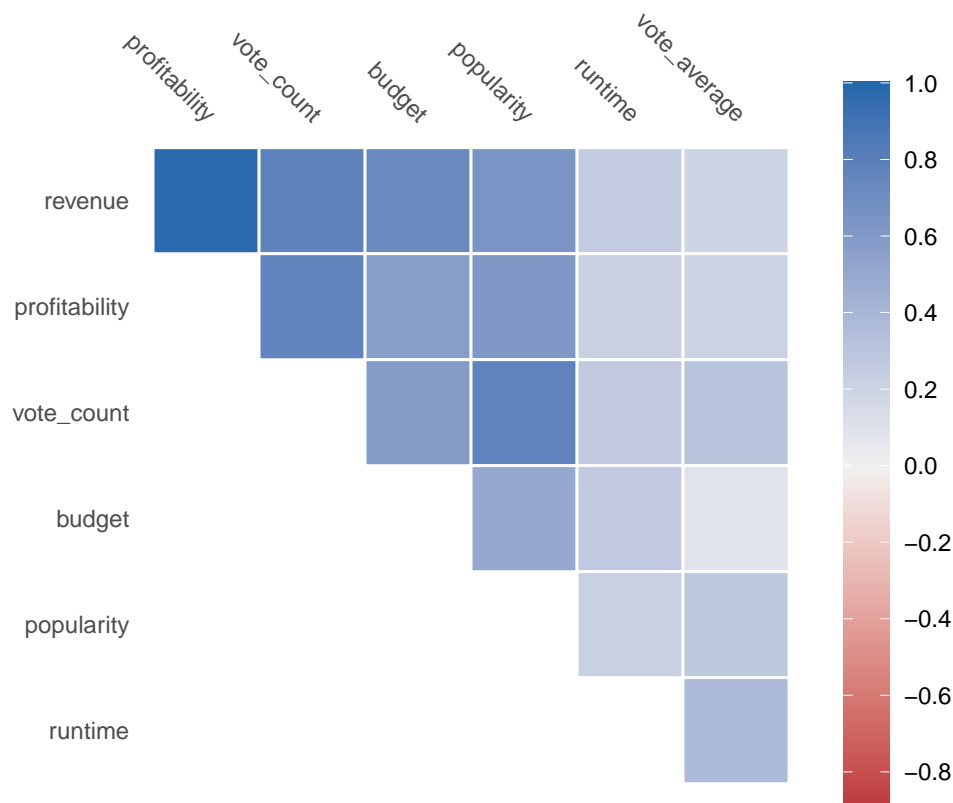## Figure 8: Correlation among Key Variables

Figure 8 above shows budget, revenue, vote count and popularity are the main factors contribute significantly to profitability.

Logically, budget for movie is given & fixed and producers tend to determine if casts, story line and etc of the movie attract movie-goers. In other word, casts (ie: Tom Cruise) and unique idea / story line (ie: mission impossibles) are the lead factors that contribute to vote count and/or popularity and revenue, and

in turn makes a movie a huge success (measured in profitability). Vote count, popularity, revenue and profit are lag factors.

Therefore, I should only focus more on non-financial metrics - vote count & popularity.
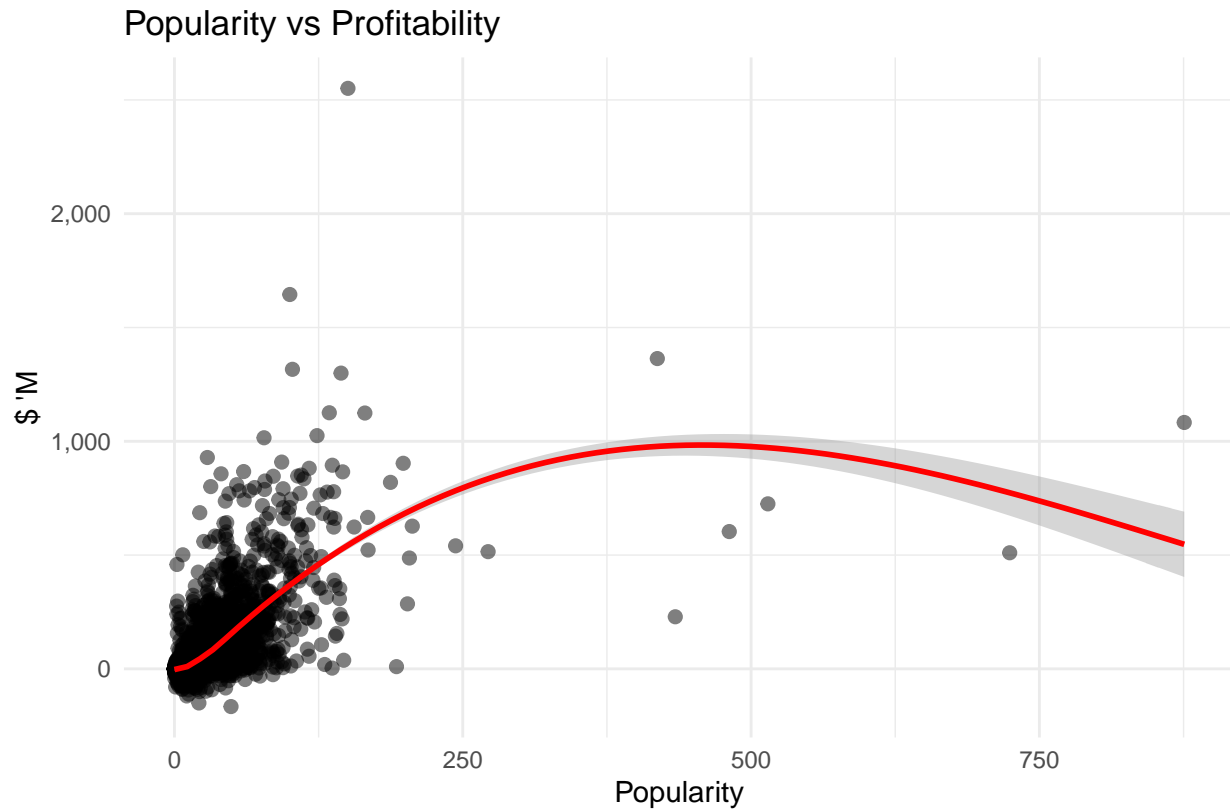
Popularity vs Profitability

Figure 9 produces an interesting finding where popularity of a movie contribute to profitability only to certain extent. Among all movies, a few movies were extremely popular (popularity > 500) and only one hit $ 1 billion profit. In contrary, there are also less popular movies (popularity < 250) and recorded $ 1 billion and more in profit.
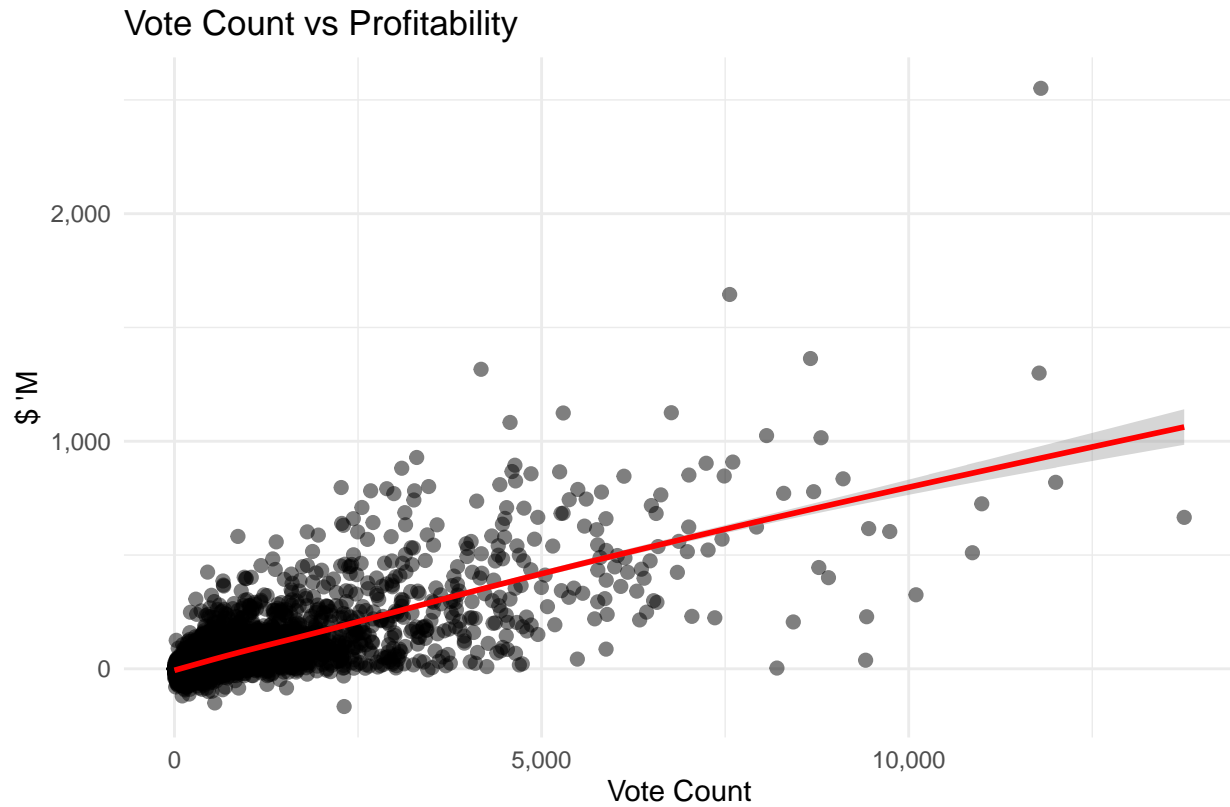
Vote Count vs Profitability

Refer to Figure 10, the trend is more obvious where higher vote count translates to higher profit. This is especially true where vote count is within the range of 7,500.

From quantitative perspective, vote count is better predictor than popularity.

3.2 Conclusion

Based on quantitative and qualitative reasons, can we jump into conclusion that success (measured in revenue & profit) of a movie is the magic combination of the actor / actress and story? Should we attribute the glory to director as well?