

How to make money with movies

Text Analytics

Cohort 3

Professor: Thomas Kurnicki

Student: Thiago Marcondes

February 13th, 2020

This sentiment analysis main objective is to further understand the movies industry towards the content of the movies that generated the higher revenue in the last decades. This way, it will be possible to check if the content changed over time, and what were the key sentiments these movies explored to become so attractive to the audience.

To develop the analysis, it was selected the top 5 movies considering their worldwide grossing revenue in the past 4 decades (2010s, 2000s, 1990s, and 1980s). Looking to the list, it is possible to see that these are very remarkable movies, even if watching it is not one of your highest hobbies.

Ranking	Movie	Year	Decade	Grossing Revenue (\$)
1	Avengers: Endgame	2019	2010	\$2,797,800,564
2	Star Wars: Episode VII - The Force Awakens	2015		\$2,068,223,624
3	Avengers: Infinity War	2018		\$2,048,359,754
4	Jurassic World	2015		\$1,671,713,208
5	The Lion King	2019		\$1,656,943,394
1	Avatar	2009	2000	\$2,789,679,794
2	The Lord of the Rings: The Return of the King	2003		\$1,142,219,401
3	Pirates of the Caribbean: Dead Man's Chest	2006		\$1,066,179,725
4	Batman: The Dark Knight	2008		\$1,004,934,033
5	Harry Potter and the Sorcerer's Stone	2001		\$978,087,613
1	Titanic	1997	1990	\$2,187,463,944
2	Jurassic Park	1993		\$1,029,939,903
3	Star Wars: Episode I - The Phantom Menace	1999		\$1,027,044,677
4	The Lion King	1994		\$968,483,777
5	Independence Day	1996		\$817,400,891
1	ET: The Extra-Terrestrial	1982	1980	\$793,482,178
2	Star Wars: Episode V - The Empire Strikes Back	1980		\$547,897,454
3	Star Wars: Episode VI - Return of the Jedi	1983		\$475,347,111
4	Indiana Jones and the Last Crusade	1989		\$474,310,887
5	Batman	1989		\$411,556,825



There is also a prevalence of action words, but it is very interesting how the movies emphasize “death” and its variances, taking the narrative to a more dramatic level.

1990's top 5 movies:



This decade explores a lot more romance. Words like “love”, “feel”, “heart”, “promise”, “listen”, give us the idea that, even we are also talking about action movies (except Titanic), they also try to explore love stories in the narrative.

1980's top 5 movies:



Lastly, in the 1980's, we can see a lot of words related to family, e.g., "home", "father", "dad", "mom", even we still have the essence of an action movie.

There are also words that are extremely relevant in all decades. “Time” is one of the top words in all charts, what lead us to realize that these movies try to give a higher sense of urgency to the audience. The idea that the characters are running against time and need to solve the problems

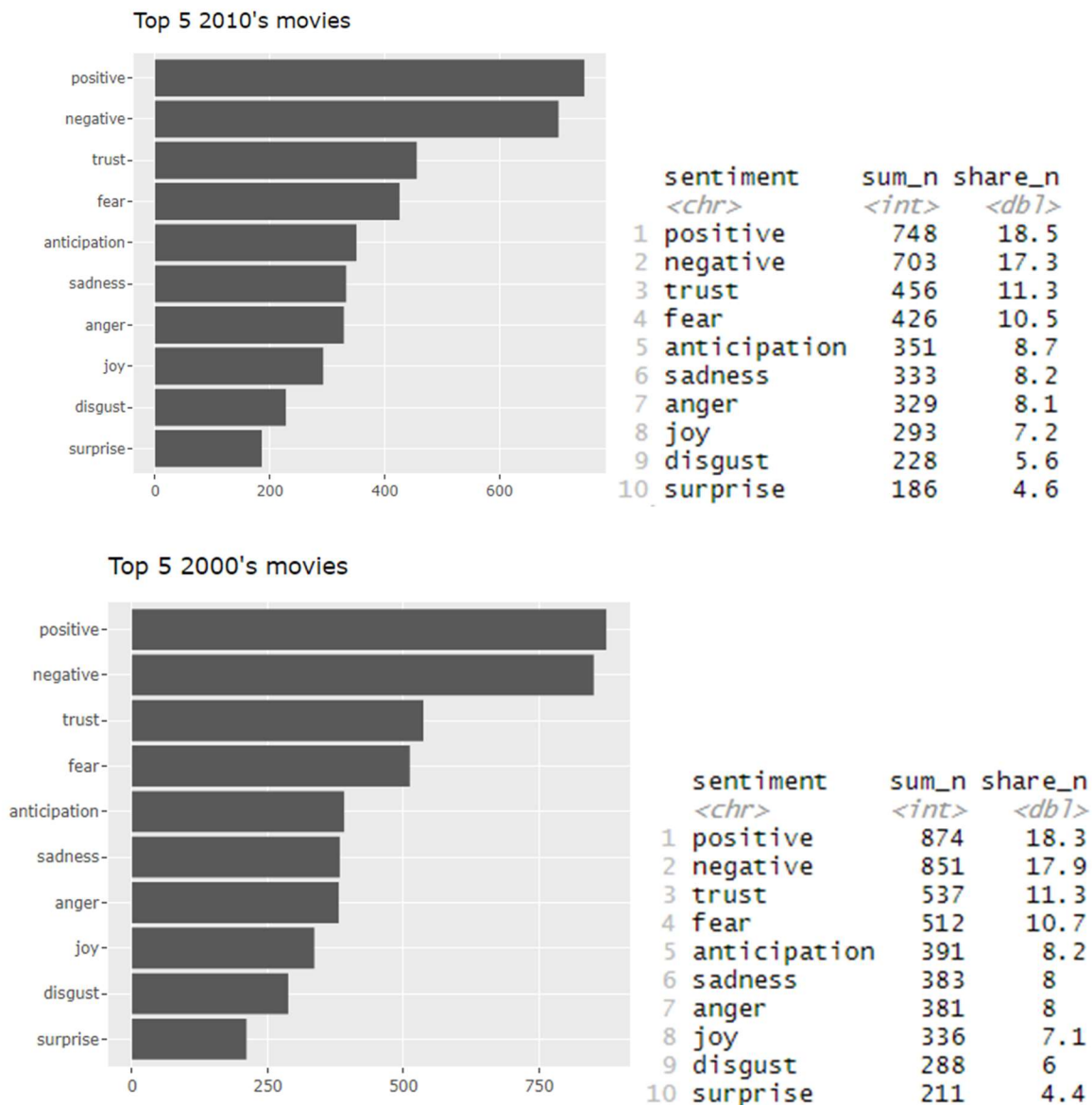
before it is too late. Checking the stories, it happens in all the movies, even Titanic, that is a drama and not exactly an action.

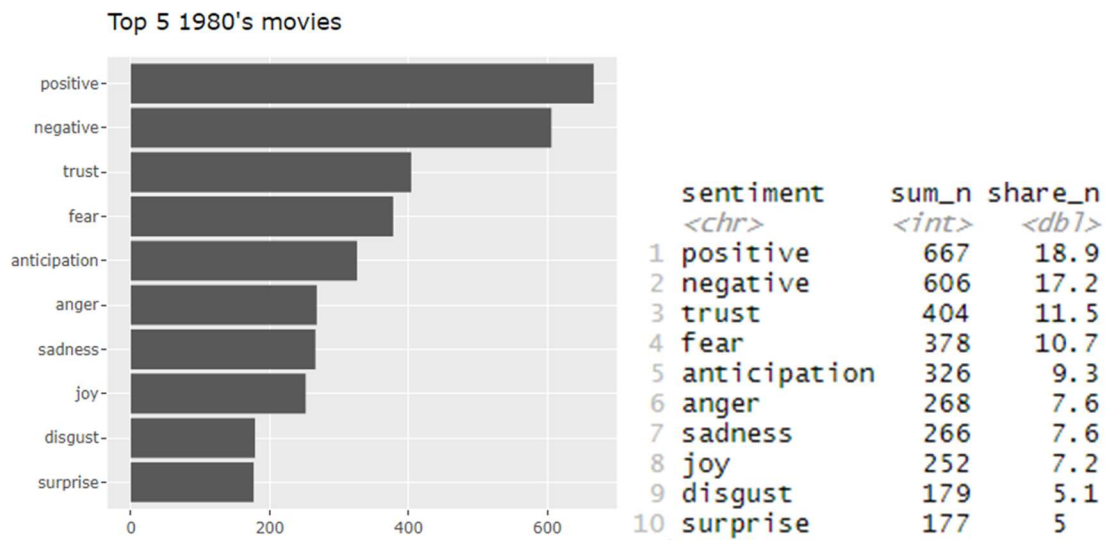
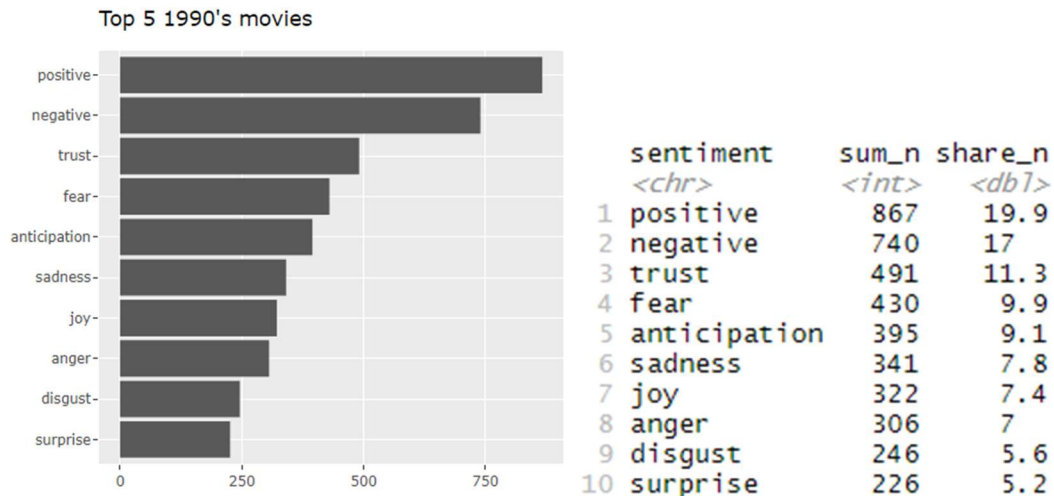
Sentiment analysis

Even that it was possible to have an understanding of how was the approach for the narrative per decade with the word cloud, the sentiment analysis is the ideal to classify the words per sentiment, giving us the frequency of each sentiment explored in the movies.

In this analyze, we decided to focus on the NRC sentiments library, which divides the words in 10 different possible sentiments: *trust, fear, negative, sadness, anger, surprise, positive, disgust, joy, and anticipation*.

By running the analysis, these are the sentiments frequency per decade:





It is very interesting to check how we can visualize a pattern in all the decades. Despite the subtle differences pointed in the word cloud, there are a lot of similarities in all these movies, and the decades do not change it substantially.

Essentially, these movies are very divided with positive and negative terms. That means that the narrative constantly changes from positive and negative outcomes to bring emotion to the audience.

Trust, fear, and anticipation are the next most important sentiments that are covered in these narratives. There is a lot of room in these movies for betrayal and fear when it comes to the fight with the “enemy”, but also a process of building trust to work in partnership and achieve their goals.

On the other hand, these movies do not have a lot of surprise content. It is interesting to think about it, because there are some very classic scenes in these movies that surprised the audience. However, those were very punctual scenes. Then, the insight is that it may be interesting to

include a plot twist in your movie, surprising your audience, but you must build an entire narrative before that. These movies do not focus on constantly surprise people.

Per decade, the main difference around sentiments are in the middle-ranked sentiments. In the 1980's, there were a big focus in exploring anger, while the 1990's wanted to explore a little more of joy. In the most recent years (2000's and 2010's), on the other hand, directors and writers prioritized sadness.

Overall, it is possible to conclude that we have a clear pattern analyzing the top 5 highest grossing revenues per decade. There are some subtle differences, showing how the narrative has been adapted to continually entertain different generations, but we still have a lot of similarities. It is almost a success formula to make money with movies. A very interesting example of company who has been exploring very successfully these types of narratives is Marvel and its universe.

"In just a decade Marvel Studios has redefined the franchise movie. Its 22 films have grossed \$ 17 billion - more than any other movie franchise in history." (Harrison, Carlsen & Skerlavaj, 2019)

A next step in this type of analysis is to compare the content of these successful movies with similar but unsuccessful films. Comparing the content can give us some indications of bad approaches that these other movies decided to explore.

Reference:

Top Lifetime Grosses - Box Office Mojo. (2020). Retrieved 14 February 2020, from https://www.boxofficemojo.com/chart/top_lifetime_gross/?area=XWW

Harrison, S., Carlsen, A., & Skerlavaj, M. (2019). Marvel's Blockbuster Machine. Retrieved 14 February 2020, from <https://hbr.org/2019/07/marvels-blockbuster-machine>