Exploring the key success factors of films: a survival analysis approach

Ahyun Kim¹ · Silvana Trimi² · Sang-Gun Lee³

Received: 17 March 2021 / Accepted: 27 September 2021 / Published online: 29 November 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

This paper investigates the key factors that contribute to the success of movies. By using sentiment and survival analysis, this study classified 1,038 movies according to the customer comments and movie characteristics and compared the number of screening days, the primary measure of success of movies, between the groups. Based on the analysis of film reviews (i.e., positive, negative, and neutral), screening days showed significant differences between (1) the positive and neutral groups, negative and neutral groups, (2) the density (positive—negative comments) of the positive and negative groups, (3) drama and action, drama and comedy, (4) domestic and foreign films, (5) G-Rated and R-Rated, R-Rated and X-rated films.

Keywords Key Factors of Movie Success · Screening Days · Customer Reviews · Film Characteristics · Survival Analysis · Sentiment Analysis

1 Introduction

An easy access to the Internet in the digital age has enabled the global communication integration, which in turn facilitates cultural convergence. As a result, the digital environment enables the capture of peoples' opinions for much greater reach and richness than traditional media could (Kaplan and Haenlein 2010). For most people

✉ Sang-Gun Lee
slee1028@sogang.ac.kr

Ahyun Kim
ahyunk@dongyang.ac.kr

Silvana Trimi
silvana@unl.edu

¹ Department of Business Administration, Dongyang Mirae University, 445 Gyeongin-ro, Seoul 08221, South Korea

² College of Business administration, University of Nebraska-Lincoln, Lincoln, NE 68514, USA

³ Sogang Business School, Sogang University, 35 Baekbeom-ro, Seoul 04107, South Korea

around the world, watching movies and documentaries is a popular means of entertainment (Tisdell 2008). In addition to the traditional means of watching movies in theaters, the exploding popularity of digital streaming services (to home TVs, computers, or mobile devices) has made movie watching a major part of home entertainment (Chopra and Veeraiyan 2017; Antonello et al. 2020; Zhang et al. 2020). Furthermore, the recent COVID-19 pandemic accelerated the growth of streaming services (Kang et al. 2020; Simoes et al. 2020).

The average time to produce a film is about one year with a substantial amount of investment (KOBIS, Korean Box Office Information System, www.kobis.or.kr). The financial success of any new production depends on the size of audience, often translated into the number of screening days. Film-producing firms promote the release of movies in several ways, such as highlighting the leading actors in the film in various media programs, advertising the film itself, releasing stories about the filmmaking process, etc. Today, consumers read other consumers' reviews or experiences of products/services before making a purchase. Thus, marketing strategies emphasize creating good relationships with the public through leveraging both positive and negative comments about a product or service for the maximum effect (Menon et al. 1996). Online communication in the form of customer reviews (comments) can be considered as media for expressing an individual's point of view or an opinion on a particular topic. The fact that anyone can write and read reviews means that people can more easily access what others think. This immersive phenomenon that occurs around a particular opinion is called fandom (Gray and Sandvoss 2020).

Film production companies use buzz marketing among the various promotion strategies. Buzz marketing is an advertising technique that focuses on creating intriguing issues or events to capture social attention and create broad visibility for a film (Kuruca and Akjol 2014). From a psychological perspective, buzz marketing can be referred to as the halo effect, a phenomenon where accentuating an event would lead to a greater positive effect than it is originally expected to (Notarantonio and Quigley 2009). However, buzz marketing is like a double-edged sword, because while it is an effective method for reaching out to a large audience at a relatively low cost, it may cause loss of trust if pushed too aggressively. According to theories of consumer behavior, negative information about a product or service tends to be more compelling to potential customers than positive information (Lee et al. 2008; Hussain et al. 2018; Moliner et al. 2018). For example, negative rumors are more powerful and spread much faster than positive information (Smith 2011; Yen 2016).

Thus, if the fate of a product/service is determined by human opinions (either positive or negative), it is possible to predict how the product's/service's outcome will be (Sweeney et al. 2008). While the impact of online reviews on the success of products/services has been researched extensively in the literature, there is a paucity of empirical studies on art related products, especially about films. Many studies have examined the degree of influence that negative or positive reviews have on movie success, but it is possible to investigate the influence and intensity of reviews in terms of detailed breakdown of different review types or the characteristics (e.g., actors, etc.) that make up the movie itself. Previous studies used factors like genres, production, and cast to determine their impact on gross revenue, utilizing a limited amount of data or no empirical evidence (Brewer et al. 2009; Simonton. 2009). To

address these limitations, we used 10 years of film data in this study. This study intends to fill this gap in the literature. Unlike ordinary products or services, the success of a film is influenced not only by social media and review comments, but also by its unique characteristics (Zhuang et al. 2006; Visch et al. 2010; Shen 2020), such as the film genre, main characters, the director, the distributor, and the main target audience. This study investigates the key factors that influence the success for films, measured with number of screening days. We collected secondary data on 1,038 films from the Korean Box Office Information Systems (KOBIS) and applied sentiment and survival analysis to determine the key factors that influence the number of screening days. We seek to explore the following two research questions:

- RQ1. Do the types of review (positive/neutral/negative) make a difference in the number of screening days of a film? Which type of reviews (positive or negative) has the most influence on screening days?
- RQ2. Do the characteristics of a film (i.e., genre, the country of production, and the movie ratings) make a difference in the movie's number of screening days?

The rest of the paper is organized as follows. In Section [Literature review and hypotheses development](#), we review relevant literature to support our study and develop research hypotheses. Section [Research methodology](#) presents research methodology, including survival analysis and Section [Results](#) describes the results of the study including hypotheses tests. Section [Conclusion](#) presents discussion of the study results and Sect. 6 concludes the study with a summary, implications, limitations, and future research needs.

2 Literature review and hypotheses development

2.1 Literature review

Movie is a collection of screenplays, actors, filming, and video productions. In order to make a movie a box office success, online or social media word-of-mouth is very important as people's comments generate the interests of the masses (Harris et al. 2016). In addition, the story, distributor, producer, promotion frequency, genre, movie rating, and leading actors have varying degrees of influence on the success of a movie. As pointed out earlier, buzz marketing is often used as a strategy to build brand awareness and word of mouth (WOM, electronic-WOM) marketing for promoting brands (Pollak et al. 2017). Table 1 presents a list of several previous studies on the keywords buzz marketing and WOM. In most of previous studies, buzz marketing and WOM were treated as closely related concepts (Mourdoukoutas and Siomkos 2009). In recent studies, however, WOM has gradually transitioned toward eWOM (Warne and Drake-Brooks 2016; Chiu et al. 2019).

As shown in Table 1, many studies measured the effect of positive or negative assessments on companies by using methodologies such as the structural equation modeling, linear regression, or case study research. In this study, we searched for

Table 1 A sample of previous studies on buzz marketing and WOM marketing

Year	Author	Purpose	Analysis method
2020	Doyle and Campbell	Investigated the patterns of language in 143 videos produced by content creators who run YouTube channels	Omnibus effect and linear mixed model
2020	Herrero and Martinez	Examined the impact of information on companies due to social network expansion	Structural equation
2016	Maisam and Mahsa	Investigated the factors that enhance the effect of positive WOM	Structural equation
2012	Sweeney et al	Studied the magnitude of positive and negative WOM	Survey
2010	Bughin et al	Discussed the effect of positive/negative WOM on a company's products and services in an online community	Study of each situation

studies that used “movies” as the keyword, rather than evaluating products or services. Table 2 provides a list of important previous studies on films that were based mainly on reviews or the movie market.

As shown in Table 2, previous studies found that when movies are searched by keywords, they are sorted into several categories. The previous studies in Table 2 can be broadly classified based on the level of research: industry, enterprise, or movie. At the industry level, research was conducted with a focus on either success strategies or phenomena occurring in a country’s film industry (Bartosiewicz and Orankiewicz 2020; Walls and McKenzie 2020; Yue 2020; Jang et al. 2021). At the enterprise level, research was conducted in connection with strategies related to marketing in order to improve the performance of movies (Szomszor et al. 2007; Weng et al. 2009; Gao et al. 2020). It has been confirmed that a movie itself is studied based on its basic information (character characteristics, reviews, etc.) (Zhuang et al. 2006; Thet et al. 2010; Lash and Zhao 2016; Viahapsari and Parmawat 2020). As shown in Table 2, the main research method employed on films was sentiment analysis, which focuses on the analysis of sentiment and recommendations on social networks by using data crawling or data mining methodologies (Zhuang et al. 2006; Szomszor et al. 2007; Weng et al. 2009; Thet et al. 2010; Viachapsari and Parmawat 2020). eSentiment analysis, also referred to as opinion mining, is a technique for analyzing individuals’ subjective opinions, attitudes, and tendencies (Nasukawa and Yi. 2003; Thet et al. 2010). The most important feature of this technique is its ability to determine whether words and contexts are positive or negative. However, an objective analysis is difficult to be performed if the combination of words in the statement is less emotional. Therefore, the subjects of this analysis are mostly reviews and blog posts that contain a significant amount of expressive and emotional words (Thet et al. 2010). Another analysis method used is classification through machine learning (Ahmed and Jaber 2020). In this study, we used survival analysis to examine the emotional characteristics of films, which is one of the major contributions of this study.

2.2 Hypotheses development

In general, when a new film is about to be released, we could likely find a large portion of reviews containing positive words in anticipation. Also, the larger the size of film’s audience, the more diverse its reviews. Due to digitalization, which provides equal access of information to everyone, people’s intention to watch a movie is influenced by a variety of ways (Chowdhury 2012; Chopra and Veeraiyan 2017). Some consumers prefer watching a film without the influence of others’ opinions. Such consumers can obtain general information about a film directly from its website, which does not provide subjective opinions or people’s reviews. However, other consumers may prefer to know other people’s opinions or reviews about a movie before deciding to watch it. Thus, consumers in general tend to obtain information (general or opinions) prior to watching a film. The sentiment of information obtained (positive, negative, or neutral) will affect the number of the film’s screening days. Thus,

Table 2 Previous studies on films

Year	Author	Category	Purpose	Data and Analysis method	Result
2021	Jang et al	Industry (movie success strategy)	Analysis of appropriate service models in the film industry	Secondary data and multiple regression analysis	G-rated, sequel, fandom, OST etc. affect download-to-own
2020	Gao et al	Enterprise (movie marketing strategy)	Envisioning a brand name strategy when cultural products are sold abroad	Survey and regression analysis	Similarity and informativeness affect sales
2020	Walls and Mckenzie	Industry (movie success strategy)	Analyzing a small number of high-impact films to confirm accurate statistics of the film industry	Secondary data and regression analysis	Budget, in open screens, sequel, star, genre, and rating affect revenue
2020	Yue	Industry (the movie industry in China)	Analysis of institutional and economic impact of pirated movies in China	Secondary data and regression equation	Copyright infringement, intellectual property protection, Internet piracy, and movie box office affect revenue
2020	Bartosiewicz and Orankiewicz	Industry (the movie industry in Poland)	To confirm the degree of concentration Poland has in the film distribution market and how its position affects corporate performance	Survey and secondary data, and basic calculating	Calculate absolute value share, absolute volume share, relative value share, and relative volume share
2020	Viahapsari and Parmawat	Movie (movie's content)	Analysis of film characters	Collection and qualitative research	–
2016	Lash and Zhao	Movie (investment in movies)	Proposal of a decision support system to assist in making investment decisions in the early stages of production	Machine learning	Create a prediction model using star power, genre, rating, profit, and release date
2010	Thet et al	Movie (movie review)	A sentiment analysis of film reviews	Opinion mining	Examine the correlation of words used in movie reviews and judging the characteristics of the movie

Table 2 (continued)

Year	Author	Category	Purpose	Data and Analysis method	Result
2009	Weng et al	Enterprise (movie marketing)	A film analysis from the perspective of social networks	Community analysis	Understand the flow of the film based on the actors' facial expressions and relationships by role
2007	Szomszor et al	Enterprise (movie marketing)	Constructing a film recommendation system	Model design	Analyze the existing recommendation system and propose a better system
2006	Zhuang et al	Movie (movie review)	A sentiment analysis of film reviews	Mining analysis	Suggest a process for grouping opinions through cast and training data

according to negativity bias theory (Baumeister et al. 2001), people tend to give greater importance and velocity to negative reviews than to positive reviews, even when a film received the same number of positive and negative reviews. Neutral reviews are less likely to influence people's intention to watch a movie than either negative or positive reviews would. Thus, we propose the following hypothesis.

H1 The number of screening days of films differs across positive, neutral, or negative reviews.

When exposed to information or objects such as films, books, or news, individuals react differently. People tend to share their feelings and opinions randomly offline (WOM) and/or online (reviews, comments, replies, etc.: eWOM). Thus, a phenomenon of “emotional contagion” occurs where people would imitate and synchronize with others' behavioral or emotional characteristics (Hatfield et al. 1993). Emotional contagion, similar with social contagion, signifies social transmission of thoughts or emotions to others. People's word choice in writing is based on their thoughts, values, and experiences. On certain objects, in the absence of unique personal motives, people would most likely have similar opinions to the prevailing ones. In this case, there is a high probability that most reviews will contain the same type of words. For example, people choose negative words if they reached a negative conclusion about something, whereas they choose more positive words if they inclined to a positive conclusion. However, Landman (1996) argued that in reviews people would most likely choose negative words than positive ones. Theories of consumer behavior also have proved that negative information about a product or service is more influential to potential customers than positive information (Lee et al. 2008; Smith 2011; Yen 2016; Hussain et al. 2018; Moliner et al. 2018). According to Prospect Theory (Kahneman and Tversky (1979), people are more sensitive to negativity than positivity. Therefore, we proposed the hypothesis that the negative density (negative – positive comments) in the negative group is greater than positive density (positive–negative comments) in the positive group.

H2 The density of sentiment words (lpositive—negative comments) differs between the positive and negative groups.

Films are often categorized by their genres (similarity in the subject). However, a film may have multiple genres that overlap rather than just one genre. These overlapping genres of a film, however, are often similar rather than completely different. As such the characters and the overall story of a film can be used to predict the age group of the potential audience. For example, a family with young children would possibly watch a drama, comedy, or animation rather than an action or horror flick. Thus, while there may be some variations from one person/family to another, the age group related to a film's genre can help predict the number of screening days (Brown 2017).

Furthermore, a film often reflects the spirit and culture of its production country. For example, the most common film genres of U.S. producers are large-scale

action and animation oriented. Europe is dominated by relaxing films and Korean films are mostly comedy, action, or drama types (d'Astous et al. 2007). The different film genres preferred by each country reveal the diversity of national sentiment. For instance, if a documentary film is released in a country where the public tends to prefer horror films, the box office numbers would be disappointing.

Thus, genre is as important as a film's main audience age group and national preferences, therefore, it exerts a strong influence on the film's success. There are about 19 genres for movies. However, in the order of their sample size, the three dominant genres are drama, action, and comedy. Thus, in this study, we focus on these three genres. We develop the following hypothesis.

H3 The number of screening days of a film differs depending on the film's genre (drama, action, or comedy) classified by technical characteristics.

There are many different types of movies produced: short films, independent films, or feature films. Short and independent films have recently become less popular, and consequently, film producers have placed higher priorities on films that are more favored by audiences and distributors. Specifically, recently movie companies have been aiming to produce blockbusters, which typically require a large budget, big name stars, and sophisticated technologies (Chiu et al. 2019). Blockbusters caused a stir in the U.S. that now people expect the same type of movies. Thus, the audience expects a movie to be a high-quality blockbuster which is promoted as a very expensive film to produce (Ahuja and Novelli 2016).

The themes of most films usually reflect characteristics of the country where they are produced (d'Astous et al. 2007). For example, film *Parasite*, the winner of the 2019 Film of the Year Oscar Award, is known as a movie that reflects today's social norms around the world, but especially in South Korea where it is produced. Capernaum (2018), a film about human rights for children, is based on a theme that is unique in a certain country or region, and therefore, its audience would be limited. In sum, the number of a film's screening days is influenced by the producer's national identity, either domestic or foreign. We propose the following hypothesis.

H4 The number of a film's screening days differs depending on the domestic or foreign identity of the film producer.

The suitability of a film for different ages of the audience is assessed by a regulating organization before the film is released to theaters. Although film rating procedures and regulations vary from country to country, film are generally rated by the following classifications: G (all ages are admitted), PG-13 (parental guidance for children aged 13 or under), R (under the high school age not admitted), and X (no teenagers admitted). The target audience of each film rating classification is different based on the demographic characteristics of the population and the popularity of films to different age groups (Mulay et al. 2020). Thus, the target audiences of X-rated and G-rated films are significantly different. Theaters also use film

Fig. 1 The research model

ratings to determine the show times (e.g., G-rated films are shown until 10 PM and X-rated films begin show times after 10 PM). Therefore, we suggest the following hypothesis.

H5 The number of screening days of films differs depending on their rating classifications.

Figure 1 presents the research model with associated hypotheses.

3 Research methodology

In this study, we used data of films released in South Korea. Film consumption in Korea is high, showing one of the highest numbers of cinema admissions per capita in the world (Moon et al. 2015). the box office revenue in South Korea amounted to over \$1.6 billion in 2019, indicating that the film market of Korea is sufficiently large and robust to conduct this research (MPA 2020).

This study applied survival analysis to identify factors that affect the occurrence of a particular event, and the subject of the event is death (the terminal point). The process of survival analysis can be explored in four different ways. The most basic method is to use a life table to determine whether a group of people with certain characteristics differs from other groups within a given time period. If a more in-depth analysis is required, the Kaplan–Meier curve can be used as it supports a comparative analysis among groups. The third method is based on the Cox regression model, which suggests that the probability of an event occurrence depends on the characteristics of the subject (Cox 1972). Lastly, the time-dependent Cox model can

Table 3 Previous studies on survival analysis applications

Year	Author	Purpose
2017	Yang et al	Analyzed the survival strategies for ICT and automobile industries
2012	Cho et al	Examined occupational turnover of nurses
2010	Argyres and Bigelow	Empirically investigated the effect of modularity on a firm's vertical integration choices
2002	Stepanova and Thomas	Survival analysis methods were applied for personal loan data
2001	Gudmundsson and Rhoades	Predicted the survival and duration of airline alliances

Fig. 2 The research process

be applied. This model considers the proportion of variables that affect survival and the likelihood that they may rise proportionally over time in the analysis process.

Survival analysis has been widely applied in medical research. For example, it has been used to analyze the effect of a certain medical treatment on patients (Kurian et al. 2010). Survival analysis has also been applied in business research (Yang et al. 2017). Nonetheless, only a limited number of studies were conducted in business because most firms are reluctant to disclose failures and as such it is very difficult to gather sufficient data (see Table 3). This study applied survival analysis to examine the factors that affect how long a movie would survive.

In this study, two types of data were used. The first type was the data provided by KOBIS. We disregarded the 2020 data as the COVID-19 pandemic has greatly affected theater operations. Thus, this study analyzed films released between 2017 and 2019. We used secondary data from the Korean Film Council, which released date including film title, country of production, genre, number of showtimes during 2010–2019. We calculated the number of screening days (days between the release date and close date) and identified the nationality of the distributor. Additionally, this study used R to crawl data of film reviews written by customers who watched the films from Korean portal Daum. Afterwards, positive and negative words were sorted out to proceed with sentiment analysis. The research process undertaken is shown in Fig. 2.

This study used sentiment and survival analysis to test Hypothesis 1, calculated the density of positive and negative customer comments to test Hypothesis 2, and used survival analysis to verify Hypotheses 3, 4, and 5. We used R to crawl up to 300 comments on each film. The number of pre-classified positive and negative words in each review were counted in order to identify the sentiment classification for each film. The dictionary of emotional words for the sentiment analysis consisted of 4868 positive words and 9827 negative words. Previous research sorted through the total of approximately 510,000 words in the Korean language and found that about 70% of the emotional words are negative (Park and Min 2005). A review is considered positive if it contains two or more positive words than negative words, and negative if there are two or more negative words than positive ones. If the difference between positive and negative words is between -1 and 1 , the review is considered neutral. In sum, this study counted and compared the number of positive and negative words in reviews to classify reviews into positive, negative, or neutral groups. Excluding reviews that were not able to be crawled, the analysis was conducted on 746 reviews. The nine items we analyzed in this study are shown in Appendix A.

For testing Hypothesis 2, a comparison of density between the positive and negative word groups was carried out. A total of 746 films was sampled because the group comparison is needed for only films whose sentiment words were affected by positive or negative reviews. Unlike Hypothesis 1 that compared the number of screening days, this hypothesis compared the density of positive and negative words. Because Hypothesis 1 sets the number of screening days as the dependent variable, it was difficult to specifically identify frequency for emotional words. In order to overcome this limitation, this hypothesis examined the absolute values of the difference between the numbers of positive and negative words in both the positive and the negative word groups. For testing Hypotheses 3, 4, and 5, only survival analysis was performed. Films released over the past 10 years were classified by genre in order to test Hypothesis 3. There was a total of 19 genre categories including Sci-fi, Horror, Documentary, Dramas, Melo/Romance, Musical, Masters, Criminal, Historic, Thriller, Animation, Action, Adventure, Independent Film, War, Comedy, Fantasy, Western, and Others. 504 films belong to the top three genres (dramas, actions, and comedies) over the study period. In order to test Hypothesis 4, we categorized films based on the country of ownership and extracted data from representative countries. We classified 1038 films into two groups, domestic and foreign, for analysis.

Finally, to test Hypothesis 5 we applied the Korean film rating classification standards and divided them into four categories for analysis. The Korean film rating system has five different levels. However, we disregarded “Limited screening” (films with limited number of screenings) because there were less than 30 films in this category during the past 10 years according to the Korea Media Rating Board (www.kofic.or.kr). In addition, this study used the Korean film rating system as it is quite similar to that of the U.S. The four categories of the Korean film rating system include: G, PG-13, R, and X. Our data analysis included 1,038 films in the four ratings categories. Hypotheses H1, H3, H4, and H5 were tested using the Kaplan–Meier method because this study attempted to identify specific

characteristics that cause “death” (Stel et al. 2011). As we presented in the literature review section, the probability that a subject will survive past a particular time (i.e., the probability that no event occurs until a specified time) is computed by the survival function. The survival function could be solved by the following procedures.

Equation (1) represents the survival function (S), which indicates the probability that an event has not occurred by time t .

$$S(t) = \Pr(T \geq t) \quad (1)$$

where T denotes time until death.

Equation (2) shows the cumulative distribution function or simply the cumulative probability, which computes the probability of death $F(t)$ by time t .

$$F(t) = \Pr(T \leq t) = 1 - S(t) \quad (2)$$

No event has yet occurred at the starting point of observation, meaning that $S(t) = 1$ at time $t = 0$. At time $t = \infty$, $S(t) = 0$ because an event would have occurred to all subjects as time approaches infinity. A function from 1 to infinity can be drawn as a curve, and the slope indicates the rate of death. The survival function can be illustrated with a curve that the survival probability decreases gradually from 1 to 0 as t increases from 0 to ∞ .

The Kaplan–Meier estimator covers both the cause of death and associated consequences (Stel et al. 2011). Referring to equation (3), denotes the number of deaths occurs at point(i) (i.e., the aggregation of exposed to the hazard occurring at a given point of time).

$$p(i) = 1 - \frac{\text{deathatd}(i)}{\text{hazardatn}(i)} \quad (3)$$

Once $p(i)$ is computed, the cumulative survival probability $S(t)$ can be estimated by multiplying the whole set of $p(i)$ by each set of $n(i)$ and $d(i)$. Although the survival probability of any time interval is presented by formula (3) regardless of censoring, hazard indicates the difference between the number of subjects survived and the number of objects censored at the time death occurs, in case there are censorings. Therefore, even if there is a censoring, the conditional survival probability $p(i)$ for a time interval is $1 - d(i)/n(i)$, as presented in formula (4).

$$S(t) = \prod_{i=1}^t \left(1 - \frac{d_i}{n_i}\right)^{C^i} \quad (4)$$

$S(t)$ is the estimated survival function at time t ; $\prod_{i=1}^t$ is calculated by multiplying all subjects' life durations at a point in time, t ; n is the total number of subjects in the sample; i is the total number of surviving subjects at time t ; and C^i is a constant with value 1 or 0.

This study chose the number of screening days as the dependent variable because it accurately reflects a film's sales and audience numbers. In most previous studies, the success of a film was measured by total amount of sales or

revenue (Mishne and Glance 2006; Nelson and Glotfelty 2012). Sales or revenue is closely related to the number of screening days (Ciciretti et al. 2015). Thus, this study used the number of screening days as the dependent variable as it is directly related to sales.

4 Results

Producing films is an arduous journey that requires substantial amounts of financial investments, time, marketing efforts, and artistic talents on the part of producer/directors, actors, and technical specialists. Measuring a film's success is of utmost importance to its stakeholders and its marketing for designing right promotional strategies. This study used the number of screening days as the primary criterion to measure a film's success. For that end, we proposed five hypotheses to extract most important factors that contribute to the number of screening days. Table 4 shown the overall results. Among the five hypotheses we tested, two were fully supported and three were partially supported.

First, we used the Kaplan–Meier estimator to calculate the significance level for each review category to test Hypothesis 1. Analytics methods included the log-rank test, Breslow test, and Tarone–Ware test. In this study, we used log ranking as it does not provide weights to time intervals. The results indicated that the differences between positive and negative comments (0.000) and between positive and neutral comments (0.008) were statistically significant, while the difference between neutral and negative comments were not (0.814). In general, if reviews of a product or service are classified into either the positive, negative, or neutral group, the order of preference would most likely be positive, neutral, and negative. However, the results proved that the positive reviews group led to the longest screening days, followed by negative, and then neutral reviews group. More specifically, on average, movies that received more positive comments had 29 more screening days than movies with more negative comments and 62 more screening days than those with neutral comments. The films in the neutral review group stayed in theaters shorter (average of 61 days) than those in the negative reviews group (average of 114 days); however, this difference was not statistically significant. Thus, H1 was only partially supported. It is evident that negative review comments are better than neutral comments for the number of screening day of films. The test results showed that film audiences used other customers' positive comments are more effective than negative comments on screening days. The life table of this hypothesis shows that the screening days of movies belong to the positive reviews group were the longest, as illustrated in Fig. 3.

While H1 used screening day as the dependent variable, the density of sentiment words was the measure for the dependent variable. To test H2, the difference in the number of sentiment words between the positive and the negative groups was analyzed by the independent sample T test. The result indicated that in the sample that considered only positive words, the differences between the number of positive and negative words were both significant at the <0.000 level in both groups. Nevertheless, contrary to what the prospect theory advocate, the

Fig. 3 The Survival Function for H1

number of positive comments in the positive group were more than negative comments in the negative group.

To test Hypothesis 3, we used the log-rank test of the Kaplan–Meier estimator. As mentioned earlier, there was a total of 19 genres in the study sample, with three prominent genres on the top: drama, action, and comedy. We tested the statistical significance of only these top three genres. The results indicated that the differences between drama and action (0.000) and between drama and comedy were significant (0.004), while the difference between drama and comedy was not (0.296). Thus, H3 was partially supported. The test results of the third hypothesis indicate that genre does affect the number of screening days. The results showed that genres that attracted the audience’s interest the longest were in the order of action, drama and comedy: the average screening days of dramas was 60 days shorter than actions but 159 days longer than comedies. While the number of screening days of films between drama and action (0.000), as well as that between drama and comedy (0.004), was statistically significant, the difference between action and comedy films was not statistically significant (0.296). Thus, H3 was partially supported. The results also indicated that each film genre achieve a very different outcome partially because of the difference in their targeted audience. Thus, action films and dramas attract different audience due to their different naturals, indicating why actions films have longer screening days than dramas. This result could be closely linked to the Korean culture of collectivism. Korea has a very family-oriented culture. For a typical Korean family, a favorite family time is to gather and watch a movie together, usually a drama or comedy rather than a violent action flick. Figure 4 presents the life table of this hypothesis, revealing

Fig. 4 The Survival Function for H3

Drama had the longest screening days that differs greatly from other genres since about the 100th screening day.

To test Hypothesis 4, we again used the log-rank test of the Kaplan–Meier estimator. We examined whether the ownership nationality of films affects the screening days. We found there were statistically significant differences in the number of screening days of films between domestic and foreign films (0.022). Thus, H4 was supported. The results of H4 showed domestic films were screened about twice longer than foreign ones, meaning that the country of movie ownership does have an effect on the number of screening days, even though the reason could be partially due to the screen quota regulation for foreign films. The screen quota regulation currently exists in Korea to protect the domestic film industry. Although the screen quota regulation is not strictly enforced, for foreign films, it is important to have the initial box office success for sufficiently long screening days. As shown in Fig. 5, the total screening days are similar between the domestic and foreign ownership movies while the difference occurs from about the 100th day.

To test Hypothesis 5, we used the same log-rank test of the Kaplan–Meier estimator. The differences in screening days by film rating categories were examined. Among the four classifications (G, PG-13, R and X), the results showed significant differences only between G and R (0.045) and between R and X (0.027). Therefore, H5 was partially supported. The test results of H5 showed that the rating classification (G, PG, R, and X) indeed have an effect on the number of screening days. While most films were R-rated, the average number of screening days was the lowest. The number of screening days for R movies differed significantly from that of G and X

Fig. 5 The Survival Function for H4

Fig. 6 The Survival Function for H5

Table 4 Summarizes the results of hypotheses tests

H#	Contents	(N)/AVG(Days)	Group α (Sig)	Result
H1 ^a (N = 746)	The number of screening days of films differs across positive, neutral, or negative reviews	P (566) N (124) NE (56)	NE&N P&NE P&N	0.008 0.814 0.000
	Contents	(N)/AVG(Words frequency)	Group α (Sig)	Result
H2* (N = 746)	The density of sentiment words (positive—negative comments) differs between the positive and negative groups	P (566) N (124)	P&N	0.000 Not Supported
	Contents	(N)/AVG(Days)	Group α (Sig)	Result
H3** (N = 504)	The number of screening days of a film differs depending on the film's genre (drama, action, or comedy) classified by technical characteristics	D (181) A (241) C (82)	D&A D&C A&C	0.000 0.004 0.296
H4*** (N = 1038)	The number of a film's screening days differs depending on the domestic or foreign identity of the film producer	D (467) F (571)	D&F	0.022 Supported
H5**** (N = 1038)	The number of screening days of films differs depending on their rating classifications	G (195) PG-13 (312) R (400) X (131)	G&R R&X Other	0.045 0.027 NA

*N (Negative), P (Positive), NE (Neutrality)

**D (Drama), A (Action), C (Comedy)

***D (Domestic, Korea), F (Foreign, other country)

****G (G-rated), PG-13 (Parental Guidance), R (Restricted), X (X-rated)

Table 5 Additional test results

Genre			Distributor's country		Film rating	
Pos	Drama & Action	0.006	Domestic & Foreign	0.002	G-rated & PG-13	1.000
	Drama & Comedy	0.087			G-rated & R-rated	0.903
	Action & Comedy	0.960			G-rated & X-rated	0.709
					R-rated & PG-13	0.892
					PG-13 & X-rated	0.640
					R-rated & X-rated	0.294
Neg	Drama & Action	0.089	Domestic & Foreign	0.967	G-rated & PG-13	0.533
	Drama & Comedy	0.270			G-rated & R-rated	0.832
	Action & Comedy	0.994			G-rated & X-rated	0.999
					R-rated & PG-13	0.934
					PG-13 & X-rated	0.449
					R-rated & X-rated	0.751

rated movies. For G-rated movies, even though the number of theaters that show the films may be small, the screening days tend to be very long, while X-rated films are usually shown by a very small number of theaters for allowed time slots. Figure 6 shows the differences between the degrees of film ratings occur between the 100th to 200th day of screening, while the total screen days are not significantly different.

Regarding Hypothesis 3, we tested whether there was a moderating effect in the relationship with other factors in each group. As shown in Table 5, the results indicate that no significant difference was found between drama and action films in the positive and negative review groups, and between movies produced by domestic and foreign distributors in the positive group. Other movie categories showed meaningful moderating effects.

5 Conclusion

Comparing with previous studies, this study examined the effects of additional factors on sales and revenue of films (Lash and Zhao 2016; Wall and McKenzie 2020; Jang et al. 2021). For example, previous studies targeted mostly R-rated movies released in 2019 in order to reduce bias, but they did not examine the movies with other properties. Second, there were studies that focused on searching for new indicators that could be helpful in generating predictive models. In our study, we expanded the factors or indicators suggested by other studies (Szomszor et al. 2007; Lash and Zhao 2016; Bartosiewicz and Orankiewicz 2020).

Watching movies is an important part of entertainment for most people around the world. The success of any movie or film depends primarily on audience size, traditionally measured by the number of screening days in theaters. In addition to figuring out the factors that contribute to film success, there are two unique academic contributions of this study. First, while analyzing the effect of online reviews on the screening days, previous studies focused primarily on the comparison of positive

and negative comments on the film. In this study, we added another dimension, the neutral comments, in measuring the effect of online reviews. Second, to evaluate the effects of each of the factors (online reviews, genre, rating, and country of the film ownership) on the screening days (the success measure of movies), we applied a methodology that has not been tried before in such studies, survival analysis. This methodological approach should open new possibilities for rigorous research for making strategic decisions on film production and promotion.

The results of this study have several important practical implications. First, both positive and negative words needed to be focused. According to previous research (Park and Min 2005), approximately 70% of Korean words are negative. If we simply look at the average between positive and negative words, the difference is about 2 times. However, if we analyze them by groups, the difference occurs reversely. Overall, this phenomenon shows that the number of negative words is less mentioned than the number of positive words, leading to a different result from the previously mentioned Prospect theory. The reason may be that although there are reviews written by audience who are simply interested in the movie, invisible reviews by film companies for the purpose of viral marketing may also had an impact. Therefore, in order to assure the success of a film, it is necessary to constantly examine audience opinions/responses and establish strategies to promote the movie to alleviate some of the effects of comments. Second, it is important to use film genre strategically. Genre is difficult to use for a clear emotional categorization because the typical film involves a combination of two or more emotions. Thus, instead of using one main genre for a film to provide clues about its content and style, it is a better strategy to offer a combination of subgenres for a more vivid inclusive story. That is to say, producing only one genre type movies may not be an effective strategy.

Third, the decision to import overseas films requires careful planning. As discussed earlier, Korea has a screen quota system in place for foreign films to be successful. Foreign films must perform well in the initial screening stage. If they fail to attract the public interest fast, a decision needs to be made about the screening strategies, either to continue showing the movie for a decreasing size of the audience or terminate the screening and send it to theaters in smaller cities. While the film market in the U.S. has grown tremendously over the years due to its entrepreneurial spirit and financial resources, the demand for films in Korea has increased due to the people's love for visual entertainment. Furthermore, the relative position of Korean culture has continuously enhanced over the past several decades (e.g., K-pop) and it is important to investigate whether this makes a difference in Koreans' preference for films produced either by domestic or overseas film makers. The results of this study provide valuable insights to Korean film distributors when they consider importing foreign films, the genre, the expected rating, and other characteristics. Fourth, the number of screening days varies depending on film ratings because the rating affects not only the target audience but also the film's box office success (Vany and Walls 1999). Finally, to assure the success of a film, the production firm should consider the target audience and its size (film ratings). For example, X-rated movies are targeted to adults, which are clearly defined with public awareness about what to expect from the films' contents. However, when it comes to PG-13 films, a close examination may be required to make sure that has the movie scenes and language

are proper for the rating. These are very important for the promotion and screenings of the movies. Based on the findings of this study, the number of screening days of PG-13 movies differed significantly from that of G- and X-rated films. Thus, we can conclude that the main criterion that determines a film's target audience is its suitability for people under the age of 15. Considering film's rating is important for the target audience size, therefore the revenue, the production decision for a film should start with the target audience. However, strategies can be used to prolong screening days for films that receive a rating not suitable for people under the age 15. Thus, G-rated movies, after taken off from big theatres, should continue to be shown in smaller theaters to capture more audience.

This study has some limitations. First, the generalizability of the study results is limited. The characteristics and scale of Korean movie market are unique. The film culture in Korea is more active than in many other countries, as witnesses by the flourishing K-culture. Thus, applying the results of this study to movie markets in other countries need careful considerations of environmental differences. It may be difficult to conduct a similar study in other countries, where film culture is not as robust and film-related data is difficult to collect on a continuous basis. The second limitation deals with the difficulties involved with sentiment analysis. Although previous research put great efforts into determining the nuances of film reviews, there is no precise method to analyze the significance of every single sentence of reviews. Therefore, although many ambiguous and long sentences need to be analyzed, it is still difficult to make detailed judgments.

Finally, diffusion of the Internet to even remote areas of the world has made movies available through digital streaming. During the pandemic, movie watching at home has become more prevalent. Some of these changes are temporary while the lockdown lasts. When pandemic is over, many people would go back to movie theatres for entertainment and socializing. However, there may be permanent changes in movie watching behaviors that has yet to be ascertained when the pandemic is under control. Therefore, a multi-channel approach is important to make effective decisions to ensure a movie's success. Our data were collected before the pandemic and does not include online channels data. These limitations provide opportunities to develop future research in this unique industry.

Appendix A

Appendix: Data Summary

No	Genre*	Film Rating**	Name	National-Open ity***	Open date	Close Date	Screen- ing days	Neg****	Pos****
1	12	PG	Your name	F	2017– 01-04	On screen	–	–	–
2	12	G-rated	The Snow Queen 3: Fire and Ice	F	2017– 01-04	2018– 05-13	494	23	4
3	17	PG	BECAUSE I LOVE YOU	D	2017– 01-04	2017– 06-07	154	71	14
4	14	PG	Passengers	F	2017– 01-04	2017– 02-15	42	21	8
5	13	PG-13	Assassin's Creed	F	2017– 01-11	2017– 03-14	62	24	7
6	11	PG-13	Allied	F	2017– 01-11	2017– 04-28	107	44	21
7	12	G-rated	Moana	F	2017– 01-12	2020– 03-01	1144	50	14
8	13	PG-13	Confidential Assign- ment	D	2017– 01-18	2018– 10-28	648	46	13
9	9	PG-13	The King	D	2017– 01-18	2020– 01-18	1095	13	7
10	12	G-rated	Turning Mecards W: Resur- rection of the Black Mirror	D	2017– 01-18	2018– 12-10	691	–	–
500	12	G-rated	The Smurfs	F	2011– 08-11	2011– 10-30	80	–	–
501	13	PG	Rise of the Planet of the Apes	F	2011– 08-17	2011– 10-06	50	–	–
502	17	PG	3 Idiots	F	2011– 08-18	2019– 11-01	2997	169	36

No	Genre*	Film Rating**	Name	Nationality***	Open date	Close Date	Screening days	Neg****	Pos****
503	13	PG-13	Colombiana	F	2011–08-31	2011–09-28	28	45	20
504	13	PG-13	Hindsight	D	2011–08-31	2011–11-15	76	32	20
505	17	PG-13	UNSTOP-PABLE FAMILY	D	2011–09-07	2011–10-26	49	25	21
506	5	PG	CHAMP	D	2011–09-07	2012–03-03	178	54	28
507	6	PG-13	Pain	D	2011–09-07	2011–11-30	84	37	43
508	17	G-rated	Mr. Popper's Penguins	F	2011–09-07	2011–11-16	70	54	10
509	5	X-rated	SILENCED	D	2011–09-22	2018–03-31	2382	160	105
510	11	PG-13	The Client	D	2011–09-29	2012–03-24	177	31	35
1047	13	PG-13	Jack Reacher: Never Go Back	F	2016–11-30	2017–01-11	42	–	–
1048	5	PG	La La Land	F	2016–12-07	On screen	-	–	–
1049	12	G-rated	STORKS	F	2016–12-07	2017–02-12	67	–	–
1050	5	PG	Pandora	D	2016–12-07	2017–11-17	345	86	109
1051	12	PG	ONE PIECE FILM GOLD	F	2016–12-08	2017–09-16	282	–	–
1052	18	PG	Will You Be There?	D	2016–12-14	2017–09-01	261	16	5
1053	13	PG-13	The Master	D	2016–12-21	2017–07-17	208	14	6
1054	12	G-rated	Sing	F	2016–12-21	2017–10-18	301	36	13
1055	12	G-rated	Pokemon the movie XY&Z	F	2016–12-22	2017–05-08	137	–	–
1056	13	PG	Rogue One: A Star Wars Story	F	2016–12-28	2017–03-12	74	–	–

*Genre: 1 (Sci-fi), 2 (Horror), 3 (Other), 4 (Documentary), 5 (Drama), 6 (Melo/Romance), 7 (Musical), 8 (Mystery), 9 (Criminal), 10 (Historical), 11 (Thriller), 12 (Animation), 13 (Action), 14 (Adventure), 15 (Independent film), 16 (War), 17 (Comedy), 18 (Fantasy), 19 (Western)

**G (G-rated), PG, 13 (PG-13), X (X-rated)

***D (Domestic film), F (Foreign film)

****Neg: Negative word, Pos: Positive word

Acknowledgements This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2021-2017-0-01628) supervised by the IITP (Institute for Information & communications Technology Promotion)

References

- Ahmed HM, Jaber HR (2020) Sentiment analysis for movie reviews based on four machine learning techniques. *Diyala Jo Pure Sci* 16(1):65–83
- Ahuja G, Novelli E (2016) Incumbent responses to an entrant with a new business model: resource co-deployment and resource re-deployment strategies. *Resour Redeploy Corp Strategy* 35:125–153
- Antonello VS, Panzenhagen AC, Balanzá-Martínez V, Shansis FM (2020) Virtual meetings and social isolation in COVID-19 times: transposable barriers. *Trends Psychiatr Psychother* 42:221–222
- Argyres N, Bigelow L (2010) Innovation, modularity, and vertical deintegration: evidence from the early US auto industry. *Organ Sci* 21(4):842–853
- Bartosiewicz A, Orankiewicz A (2020) The impact of Hollywood majors on the local film industry. The case of Poland. *Creative Ind J* 13(3):288–302
- Baumeister RF, Bratslavsky E, Finkenauer C, Vohs KD (2001) Bad is stronger than good. *Rev Gen Psychol* 5(4):323–370
- Brewer SM, Kelley JM, Jozefowicz JJ (2009) A blueprint for success in the US film industry. *Appl Econ* 41(5):589–606
- Brown N (2017) The children's film: Genre, nation, and narrative. Columbia University Press
- Bughin J, Doogan J, Vetvik OJ (2010) A new way to measure word-of-mouth marketing. *McKinsey Quarterly* 2(1):113–116
- Chiu YL, Chen KH, Wang JN, Hsu YT (2019) The impact of online movie word-of-mouth on consumer choice. *Int Mark Rev* 36(6):996–1025
- Cho SH, Lee JY, Mark BA, Yun SC (2012) Turnover of new graduate nurses in their first job using survival analysis. *J Nurs Scholarsh* 44(1):63–70
- Chopra S, Veeraiyan M (2017) Movie rental business: blockbuster, netflix, and redbox. *Kellogg Sch Manag Cases* 1(1):1–21
- Chowdhury G (2012) How digital information services can reduce greenhouse gas emissions. *Online Inf Rev* 36(4):489–506
- Ciciretti R, Hasan I, Waisman M (2015) Distribution strategy and movie performance: an empirical note. *Eurasian Econ Rev* 5(1):179–187
- Cox DR (1972) Regression models and life-tables. *J Roy Stat Soc: Ser B (methodol)* 34(2):187–202
- d'Astous A, Colbert F, Nobert V (2007) Effects of country-genre congruence on the evaluation of movies: the moderating role of critical reviews and moviegoers' prior knowledge. *Intern J Arts Manag* 10(1):45–51
- Doyle PC, Campbell WK (2020) Linguistic markers of self-disclosure: using YouTube coming out videos to study disclosure language *PsyArXiv*. <https://doi.org/10.31234/osf.io/tvgs9>
- Gao W, Ji L, Liu Y, Sun Q (2020) Branding cultural products in international markets: a study of hollywood movies in China. *J Mark* 84(3):86–105
- Gray J, Sandvoss C, Harrington CL (2020) Fandom: Identities and communities in a mediated world. New York University Press
- Gudmundsson SV, Rhoades DL (2001) Airline alliance survival analysis: typology, strategy and duration. *Transp Policy* 8(3):209–218
- Harris LC, Fisk RP, Sysalova H (2016) Exposing Pinocchio customers: investigating exaggerated service stories. *J Serv Manag* 27(2):1–38
- Hatfield E, Cacioppo JT, Rapson RL (1993) Emotional Contagion Cambridge University Press 2(3):96–100
- Herrero A, Martínez P (2020) Determinants of electronic word-of-mouth on social networking sites about negative news on CSR. *J Bus Ethics* 163:1–15
- Hussain S, Guangju W, Jafar RMS, Ilyas Z, Mustafa G, Jianzhou Y (2018) Consumers' online information adoption behavior: Motives and antecedents of electronic word of mouth communications. *Comput Hum Behav* 80:22–32

- Jang M, Baek H, Kim S (2021) Movie characteristics as determinants of download-to-own performance in the Korean video-on-demand market. *Telecommunications Policy* 45(7):1–10
- Kahneman D, Tversky A (1979) On the interpretation of intuitive probability: A reply to Jonathan Cohen.
- Kang J, Diao Z, Zanini MT (2020) Business-to-business marketing responses to COVID-19 crisis: a business process perspective. *Mark Intell Plan* 39(3):454–468
- Kaplan AM, Haenlein M (2010) Users of the world, unite! The challenges and opportunities of social media. *Bus Horiz* 53(1):59–68
- Kurian AW, Sigal BM, Plevritis SK (2010) Survival analysis of cancer risk reduction strategies for BRCA1/2 mutation carriers. *J Clin Oncol* 28(2):222–231
- Kuruca Y, Akyol M (2014) Movie promotion through social media. In 12th International Symposium Communication in the Millennium, Anadolu University, Faculty of Communication Sciences, Eskişehir 99–109
- Landman J (1996) Social control of “negative” emotions: the case of regret. In: *The emotions: Social, cultural, and biological dimensions*, ch 5. pp 89–116
- Lash MT, Zhao K (2016) Early predictions of movie success: the who, what, and when of profitability. *J Manag Inf Syst* 33(3):874–903
- Lee J, Park DH, Han I (2008) The effect of negative online consumer reviews on product attitude: an information processing view. *Electron Commer Res Appl* 7(3):341–352
- Maisam S, Mahsa RD (2016) Positive word of mouth marketing: explaining the roles of value congruity and brand love. *Journal of Competitiveness* 8(1):19–37
- Menon A, Bharadwaj SG, Howell R (1996) The quality and effectiveness of marketing strategy: effects of functional and dysfunctional conflict in intraorganizational relationships. *J Acad Mark Sci* 24(4):299–313
- Mishne G, Glance NS (2006) Predicting movie sales from blogger sentiment. In AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs: 155–158
- Moliner MA, Monferrer-Tirado D, Estrada-Guillén M (2018) Consequences of customer engagement and customer self-brand connection. *J Serv Mark* 32(4):387–399
- Moon S, Bayus BL, Yi Y, Kim J (2015) Local consumers’ reception of imported and domestic movies in the Korean movie market. *J Cult Econ* 39(1):99–121
- Mourdoukoutas P, Siomkos GJ (2009) The seven principles of WOM and buzz marketing: Crossing the tipping point. Springer Science & Business Media, Germany
- Mulay R, Pandey S, Patil N, Patil M, Mandlik MS (2020) Movie rating system based on sentiment analysis. *Artificial & Computational Intelligence* 1(2):36–40
- Nasukawa T, Yi J (2003) Sentiment analysis: Capturing favorability using natural language processing. In *Proceedings of the 2nd International Conference on Knowledge Capture*, NY
- Nelson RA, Glotfelty R (2012) Movie stars and box office revenues: an empirical analysis. *J Cult Econ* 36(2):141–166
- Notarantonio EM, Quigley CJ Jr (2009) The effectiveness of a buzz marketing approach compared to traditional advertising: an exploration. *J Promot Manag* 15(4):455–464
- Park IJ, Min KH (2005) Making a list of Korean emotion terms and exploring dimensions underlying them. *Korean J Soc Person Psychol* 19(1):109–129
- Pollak F, Dorcak P, Markovic P (2017) Analytical view on using weblog as a tool of marketing management. *Annals Constantin Brancusi U. Targu Jiu. Lett Soc Sci Series* 2:72–77
- Shen D (2020) Movie box office prediction via joint actor representations and social media sentiment. 34th Conference on Neural Information Processing Systems
- Simoes JC, Ferreira FA, Peris-Ortiz M, Ferreira JJ (2020) A cognition-driven framework for the evaluation of startups in the digital economy: adding value with cognitive mapping and rule-based expert systems. *Manag Decis* 58(11):2327–2347
- Simonton DK (2009) Cinematic success criteria and their predictors: the art and business of the film industry. *Psychol Mark* 26(5):400–420
- Smith KT (2011) Digital marketing strategies that millennials find appealing, motivating, or just annoying. *J Strateg Mark* 19(6):489–499
- Stel VS, Dekker FW, Tripepi G, Zoccali C, Jager KJ (2011) Survival analysis I: the Kaplan-Meier method. *Nephron Clin Pract* 119(1):83–88
- Stepanova M, Thomas L (2002) Survival analysis methods for personal loan data. *Oper Res* 50(2):277–289
- Sweeney JC, Soutar GN, Mazzarol T (2008) Factors influencing word of mouth effectiveness: receiver perspectives. *Eur J Mark* 42(3–4):344–364

- Sweeney JC, Soutar GN, Mazzarol T (2012) Word of mouth: measuring the power of individual messages. *Eur J Mark* 46(1–2):237–257
- Szomszor M, Cattuto C, Alani H, O'Hara K, Baldassarri A, Loreto V, Servedio VD (2007) Folksonomies, the semantic web, and movie recommendation. 4th European Semantic Web Conference, Bridging the Gap between Semantic Web and Web 2.0, Innsbruck, Austria
- The Motion Picture Association (MPA) (2020) Theme report 2019 1–59. <https://www.motionpictures.org/wp-content/uploads/2021/03/MPA-2020-THEME-Report.pdf>
- Thet TT, Na JC, Khoo CS (2010) Aspect-based sentiment analysis of movie reviews on discussion boards. *J Inf Sci* 36(6):823–848
- Tisdell EJ (2008) Critical media literacy and transformative learning: drawing on pop culture and entertainment media in teaching for diversity in adult higher education. *J Transform Educ* 6(1):48–67
- Viahapsari E, Parmawati A (2020) Analysis the type of deixis in the main character on the movie I leave my heart in Lebanon. *Proj (prof J Eng Educ)* 3(3):388–394
- Visch VT, Tan ES, Molenaar D (2010) The emotional and cognitive effect of immersion in film viewing. *Cogn Emot* 24(8):1439–1445
- Walls WD, McKenzie J (2020) Black swan models for the entertainment industry with an application to the movie business. *Empirical Economics* 59(6):3019–3032
- Warne RT, Drake-Brooks MM (2016) Comparing the persuasiveness and professionalism of newspaper, blog, and social media sources of information in marketing and reviewing theatre. *Arts and the Market* 6(2):166–186
- Weng CY, Chu WT, Wu JL (2009) Rolenet: Movie analysis from the perspective of social networks. *IEEE Trans Multimedia* 11(2):256–271
- Yang CG, Trimi S, Lee SG, Yang JS (2017) A survival analysis of business insolvency in ICT and automobile industries. *Int J Inf Technol Decis Mak* 16(6):1523–1548
- Yen YS (2016) Factors enhancing the posting of negative behavior in social media and its impact on venting negative emotions. *Manag Decis* 54(10):2462–2484
- Yue Y (2020) The effects of movie piracy on box-office revenue: an empirical analysis of the Chinese movie market. *J Appl Econ* 23(1):618–655
- Zhang M, Sun L, Qin F, Wang GA (2020) E-service quality on live streaming platforms: swift guanxi perspective. *J Serv Mark* 35(3):321–324
- Zhuang L, Jing F, Zhu XY (2006) Movie review mining and summarization. In *Proceedings of the 15th ACM International Conference on Information and Knowledge Management*. Arlington, Virginia, USA

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.