The effects of social tagging and ratings by consumer segment on Streaming Video on Demand chart success: A quantitative analysis

MASTER THESIS MARKETING ANALYTICS

FALL 2021

Name: Anthony Molina

SNR: 2030852

Supervisor: Dr. George Knox

Second Assessor: Dr. Samuel Stäbler

Date: January 3, 2022

Management Summary

In this thesis, I examined the effects of social tagging and rating by consumer segments in the domain of Streaming Video on Demand with regards to the market leader Netflix. In 2021, streaming platforms were facing a major challenge of subscribers churning. This is based on the individual's variety-seeking tendency. One way to overcome this issue is by shifting subscribers' attention to the Netflix Daily Top 10 via social tagging. This chart captures the ten most trending videos in the United States. Once a video appears on the chart its popularity as well as views increases.

For this purpose, I collected data from a third-party supplier (i.e. IMDB). Social tagging allows consumers to label and categorize digital content as well as express their thoughts, perceptions and feelings with respect to different concepts. In addition, as concluded by extant literature, user ratings play a significant role in product adoption. Thus, I enhanced the dataset with ratings by consumer segment based on demographic characteristics (from IMDB).

To better understand a content's chart success, it is vital to analyze the (total) time it survived on the chart. Hence, I used a multivariate survival analysis model to estimate the effects on chart success. The applied Cox Proportional Hazards regression model allows to handle a set of independent variables while a baseline hazard function is assumed to be model-free.

The results indicate that videos which contain a varied set of social tags such as apology, friend, crying or party have a negative impact on chart survival. Genres – which are also considered as social tags – show a differentiated picture. While Family, Musical or Romance can improve a content's survival time, Adventure or Horror have a vice versa effect. The analysis also reveals that almost all ratings by consumer segment have a negative influence on chart success. Last, the debut ranks 1 to 6 can positively affect a video's survival probability.

The findings of this study show that social tags can serve as a driver for chart success. However, it is important to define the types of keywords a priori in order to improve chart survival. In alignment to previous literature, user ratings should be considered since they significantly affect a content's chart success. Last, practitioners are advised to adopt their post-release advertising activities since the results show that while most of the Netflix videos debuted highly, they exited the chart sooner as expected.

Preface

This thesis would have not been possible without the help and support of many people.

First of all, I would like to thank my supervisor, Professor George Knox, for his expertise and motivating words. He read my numerous revisions and helped me to shed some light with my confusions. I would also like to thank my second assessor, Dr. Samuel Stäbler, for reading and evaluating my thesis. Next, I would like to acknowledge my colleagues, Danae and Sjoerd, for their insightful contributions. To my parents, thank you for supporting me by any means and for believing in me. Also, thank you Filip for being patient with me. Last but not least, I would like to express my deepest gratitude for Anita. Thank you for your unconditional love and support through good and bad times. To you I dedicate this thesis.

Table of contents

1 Introduction

1.1 Business Background

In 2021, the leading Streaming Video On Demand (SVOD) service Netflix (Bean, 2020) launched the new show “Lupin” which was viewed by 76 million households within the first four weeks after its launch. Despite the success of adding 37 million new subscribers to the platform, Netflix’s market share dropped by 11% since it did not meet its projected target of 206 million users (Leonard, 2021). Furthermore, the company is facing a major challenge of customers churning. This is due to a competitive market which allows consumers to choose from a large number of SVOD services. In fact, reducing the total number of subscriptions allows them to focus more on platforms which offer a greater variety of content that better fits their preferences (Paris, 2021). One way to tackle this issue is to proactively shift users’ attention to the broad range of shows and movies that Netflix offers. In particular, the greater variety of content is a main motivation for consumers to subscribe and watch content (Arkenberg et al., 2021). A form of discovering new and popular products is through the “Daily Top 10” which is a chart that shows daily trending content based on actual views (Johnson, 2020). Moreover, once a video appears on this chart its popularity as well as views increases. From a content producer and marketing perspective, it is crucial to understand the drivers of chart success since SVOD platforms barely disclose any viewership data (Garvey, 2021).

When consumers look up a particular video, they enter a query into a search engine. The results are determined and ranked by an algorithm based on how “useful” and “relevant” they are (Lau et al., 2019). These search queries are composed by user-generated keywords which they associate with a certain product. This type of information retrieval is also known as Social Tagging. It allows users to label and categorize digital content as well as express their thoughts, perceptions and feelings with respect to different concepts (Nam et al., 2014). Moreover, these social tags contain information which can support content producers and marketers to better grasp trends and brand equity (Nam et al., 2017). In social media, online videos with up to 17 tags have a positive effect on views, especially when used in moderation (Tafesse, 2020). While content tagging has been mainly studied in the domain of Recommender Systems (Ansari et al., 2018; Lee et al., 2021; Liu et al., 2018), little is known on how consumers anticipate social tags in finding popular products from SVOD charts.

1.2 Problem Statement and Research Question

When consumers search for new content they face a vast number of digital products from which they can choose from. As studied by Datta et al. (2018), the adoption of online streaming services results in an increased and more varied consumption and discovery. This is based on the individual's tendency for seeking variety which can be defined either as direct or derived behavior. The former stems from internal or personal motives such as a desire for change or saturation with product attributes. The latter – which was not directly related to a desire for variety – occurred due to external factors (McAlister et al., 1982). Moreover, when users anticipate variety in future consumption it generates positive thoughts about that future experience (Sevilla et al., 2016). A reason for this behavior is that products which offer a greater variety of options are perceived to have a greater category expertise or core competency in their category (Berger et al., 2007). In addition, demographic factors play a significant role in why certain customer segments have a (stronger) desire for variety in goods (Durante et al., 2015). Thus, increasing product variety can diversify demand which then leads to an overall higher consumption as well as to more market share especially in the domain of motion pictures (Tan et al., 2017).

Nevertheless, when it comes to the drivers of popularity most of the studies in the entertainment industry tend to focus on its effects on box-office performance (Ainslie et al., 2005; Elberse et al., 2003). This approach cannot be applied since SVOD services do not provide such a metric to (future) content creators and marketers. To better understand what leads consumers in finding popular movies and shows is by linking their variety-seeking induced search behavior to different product attributes which are provided by these platforms via labels. These social tags are used to categorize online content in order to briefly indicate what a content is about. This gives users an initial perception whether they can associate with a product according to their varied preferences or not (Nam et al., 2014). In particular, social tags are a type of user-generated content (UGC), which provides value to consumers by reducing the uncertainty about whether the movie/ show is a good match with their preferences (Wen et al., 2014). The advantages of social tagging are two-fold: First, it allows users to quickly find content in accordance with their perceptions of various objects (Nam et al., 2017). Second, industry professionals can cost-effectively gather rich and qualitative information from a "collective intelligence" without the help of domain experts (Lau et al., 2015). That is why it is important to learn what motivates viewers from incorporating their preference for variety via social tags in order to watch a video beforehand (Tafesse, 2020).

One way for consumers to find new digital content is from so-called charts. These ranked lists show how popular certain products are such as movies (Legoux et al., 2016). They serve as a source of information since users are interested in observing these rankings evolve over time and speculating about possible future changes (Bradlow et al., 2001). Moreover, research on motion pictures show that not only the appearance but mostly the perseverance of being on these lists determines the prolonged success (Sochay, 1994). In fact, various predefined and well-known product features such as actors or genre have a significant influence on chart endurance (Ainslie et al., 2005). However, when consumers form their pre-choice expectations about a certain video, they are mainly driven by information about its plot (Neelamegham et al., 1999). As suggested by Mukherjee et al. (2011), information related to a movie or show's plot can be therefore considered as unobserved attributes since they are highly differentiated goods. Understanding latent product attributes from a video's plot is crucial since these features are not determined a priori by content producers and marketers. Instead, they reflect users' perceptions which are revealed after viewing a content (Roos et al., 2014). Social Tagging as such can be therefore considered as an extension of latent product attributes since consumers form these keywords to find, associate and connect with content (Nam et al., 2014), especially after watching a movie or show.

Yet, when it comes to SVOD services and what drives consumers in adopting popular content from charts, most research has tended to focus on user ratings rather than on social tags. For instance, Lee et al. (2015) found out that product ratings can result in a herding tendency of individuals which then increases the popularity of a movie. Moreover, when users evaluate products they first tend to focus on its particular characteristics rather than on the global picture. They therefore apply so-called aspect ratings to express their perceptions towards specific attributes which, in result, they describe in detail through their own corresponding opinion keywords (Li et al., 2020). Social tagging can be considered as such. In essence, the overall rating is significantly impacted by consumers' ratings on particular dimensions and that this effect depends on the dimensions being evaluated (Schneider et al., 2021). That is why content ratings can serve as a proxy measure of viewer satisfaction, especially after controlling for product characteristics (Moon et al., 2010). Nonetheless, most SVOD services do not enable their subscribers to evaluate certain content attributes. Hence, more light needs to be shed on the interaction between social tags and consumer evaluations since the former can result from aspect ratings and therefore affect product adoption.

Most of the studies on social tagging and user ratings have been in the domain of Recommender Systems (Ansari et al., 2018; Lee et al., 2021; Wei et al., 2016) rather than on its effects on chart success. A major disadvantage of this method is the problem of "cold start": Users who have not evaluated any chart content cannot be recommended a video (Mutlu et al., 2015). One way to tackle this issue is by putting viewers' demographics into consideration (Khodabandehlou et al., 2020). As studied by Dogruel (2017), cross-cultural differences have an impact on individuals' decision-making for movies on SVOD platforms. In addition, users have a tendency for product-related cues over crowd-related information. Although research on user demographics and its impact on content ratings exist (Lee et al., 2015; Zhou et al., 2009), gaps on its effects on SVOD chart performance remain. Moreover, even though the influence of content labeling and consumer evaluations has been studied previously (Wei et al., 2016), some questions have been raised as to whether the variety of social tags as well as user ratings based on their demographics have a (simultaneous) moderating effect on chart success.

Since Netflix neither provides content ratings nor enables direct interaction among subscribers, consumers rely on third-party suppliers from which they can obtain information about certain product characteristics or ratings. The Internet Movie Database (IMDB) for example covers a wide-range of topics especially on movie/ show plots. They provide such information to their visitors via UGC in the form of video tag lists (Liu et al., 2016). These keyword lists are user-defined annotations which reflect their perceptions and can be therefore considered as social tags (Nam et al., 2017). In addition, IMDB supplies content ratings in which viewers can evaluate movies or shows on a 1-10 scale with 10 being the most favorable rating. These ratings are also broken down into segments according to users' demographic characteristics (Nelson et al., 2012). However, most studies on UGC and its effects on chart success have put their focal view on reviews (Legoux et al., 2016) rather than on the variety effects of social tags. Furthermore, there is little empirical research on how these heterogeneous tags are linked to product ratings based on different consumer demographic groups as a result of aspect ratings. Hence, little is known on how the variety of social tags is linked to various ratings by consumer segment and how they affect instant chart success as well as the survival of being on such.

In light of the drivers of content popularity on SVOD services, this raises the following problem statement: “To what extent does social tagging along with ratings by consumer segment affect the chart success of online videos?”

The aim of this research is to answer the following research questions:

Theoretical research questions:

- How do social tags affect chart survival of online videos on SVOD platforms and which variables are the driving factors?
- Which information can be derived from social tags in order to explain its influence on chart success?
- How does the effect of social tags on chart performance depend on variety?
- What is the effect of ratings by consumer segment on chart survival of online videos on SVOD platforms and which variables are the driving factors?
- How does the influence of user ratings on chart performance depend on their demographic characteristics?

Practical research questions:

- How can practitioners implement the findings of this study?
- What can be advised to SVOD services, advertisers, marketers, content providers and third-party suppliers in regards to increasing the chart success of movies and shows?

2 Literature Review

This chapter provides an overview of existing literature on the drivers of SVOD chart success. The first stream of literature in section 2.1 is concerned with Social Tagging and the theory of consumer variety-seeking behavior. In section 2.2, previous research on the role of consumer demographic traits on product ratings is outlined. The third part of literature deals with related phenomenon on video chart success in 2.3. The last section 2.4 gives a critical overview on previous literature streams and discusses the contributions of this study.

2.1 Social Tagging and variety

The rise of Streaming Video on Demand services in recent years has led consumers to shift from traditional media (e.g. Television) to these platforms. Netflix for instance offers a wide-range of movies and shows from which their subscribers can choose. With this change in consumer attitude, content creators and marketers face new challenges in terms of effectively targeting (potential) viewers (Schweidel et al., 2016). One reason is that the adoption of online streaming services has led individuals to a greater as well as more varied consumption and product discovery (Datta et al., 2018). This behavior is based on the theory of consumer variety-seeking preference (Farquhar et al., 1976; Kahn, 1995; McAlister et al., 1982). The motivation for this can result from intrinsic as well as from extrinsic factors (McAlister et al., 1982). As viewers' consumption history evolves, their preference for product attributes changes. This leads them to choose services or goods with a greater variety of features instead. In addition, once a consumer has reached an optimal level of an attribute, he may choose to satiate a different attribute afterwards since he seeks a balance of attributes to maximize his utility (McAlister, 1982). Furthermore, when people anticipate variety in prospective consumption, it creates positive thoughts about that future experience. Hence, they satiate at a reduced pace in the present (Sevilla et al., 2016).

Many recent studies document that this phenomenon can be observed in the domain of SVOD services. For instance, Sevilla et al. (2019) found out that users seek variety in a reactive and proactive way with the goal to counter psychological and physiological satiation. These findings are in alignment to the study of Xu et al. (2021) where the authors conclude that a stronger emphasis on viewers' variety-seeking tendency can aid SVOD platforms. In particular, by offering a greater variety of distinct product categories and attributes can help retain subscribers and hinder them from switching to substitutes. In fact, the focus on attribute-based variety can avoid demand uncertainty without automatically raising the amount of products offered (Moreno et al., 2017). In support of Tan et al. (2017), this suggests that product variety has a higher likelihood to increase demand concentration which pushes the interest for popular videos and reduces the demand in niche contents. However, increased variety can lead to “choice overload”: Consumers struggle to choose from a large set of movies/ shows which, in result, makes them less satisfied with their decision. A way of overcoming this issue is through latent feature diversification. This allows individuals to reduce decision uncertainty by diversifying a smaller choice set with latent video attributes (Willemsen et al., 2016).

One school of thought argues that hidden video features can be derived from its plot. This product attribute is an intrinsic cue (Linton et al., 1988) which influences consumers’ pre-choice expectations (Neelamegham et al., 1999). Since movies or shows are highly differentiated products, information linked to its plot can be therefore understood as latent attributes (Mukherjee et al., 2011). From a commercial standpoint, it is significant to understand this aspect. These unknown characteristics are not determined by content producers and marketers beforehand but rather revealed by viewers after watching a video which, in fact, reflects their perceptions (Roos et al., 2014). As studied by Eliashberg et al. (2007), a form of extracting plot information is via movie scripts. The authors applied a Natural Language Processing (NLP) approach where they used bag-of-words to obtain topics, scenes and emotions in a script. However, they also point out that a limitation of their research is the use of human judges. To overcome this drawback, an automated process can be utilized by gathering plot information from sources which stems from a “swarm intelligence” (Lau et al., 2015).

However, SVOD services only offer limited information in regards to a content’s plot. That is why consumers tend to gather information from third-party suppliers such as IMDB. Here, users can match their expectations and perceptions with UGC. In particular, latent product attributes can be linked to video tag lists which contain keywords that are maintained by users (Liu et al., 2016). These user-defined annotations enable consumers to quickly find, connect and adopt products. This concept is known as Social Tagging (Nam et al., 2014). A study by Beenen et al. (2004) shows that individual users on such websites (e.g. IMDB) support others in finding the right content through their tag contributions. In essence, the perceived uniqueness of their social tags leads them to increase the tag quantity they provide to a greater audience. According to Wei (2020), the ideal set of tags consists of a variety of “core” (i.e. well-known) as well as atypical (i.e. latent) video attributes. The author argues that these atypical elements follow a pattern of distinct characteristics outside of popular contents. However, he also suggests that a combination of atypical features can harm the success of a product in the long run.

Another point that needs to be addressed is to what extent distinct information of latent attributes is defined. Moreover, since social tags usually contain one or two words, they are prone to social synonymy and polysemy (Lau et al., 2015). A study by Xu et al. (2008) on the former shows that this phenomenon is not restricted to consumer demographics. The authors find that movie taggers apply rules in indexing which for instance the use of nouns or noun phrases as social tags. On the other hand, social polysemy refers to the same words which express different meanings (Mika, 2005). This often occurs in social tagging websites where annotators apply similar descriptions on heterogeneous concepts (Lau et al., 2015). Although many researchers have applied folksonomy-based ontology methods, results show that they are weak in addressing these problems (Dimitrov et al., 2018). Furthermore, little is known to which degree social synonymy and polysemy influence product adoption when it comes to the variety of social tags. As suggested by Trabelsi et al. (2010), these challenges can be overcome by forming non-taxonomic relationships between tags. Therefore, the flexibility of the authors' approach allows a better distinction between (latent) product attributes reflected via social tagging.

2.2 The role of consumer demographic traits on product ratings

As described in the previous section, websites like IMDB are also rich in user ratings on various movies or shows. Moreover, these evaluations are broken down into segments according to viewers' demographic characteristics (Nelson et al., 2012). Extant literature on video ratings has shown that consumer demographics play a significant role in product adoption (Lee et al., 2015; Li et al., 2020; Roos et al., 2014). For instance, Roos et al. (2014) conclude that the demographic composition of viewers reflects certain social trends. Hence, favorable videos are being consumed which then results in higher ratings of such. However, Lee et al. (2015) argue that the meaning of evaluations can be misleading. This is based on their findings that movie ratings can cause herding behavior of individuals which arises from consumers' tendency for homophily. Thus, the authors suggest to put user demographics as well as product attributes into consideration in order to get a more granular picture of video ratings. This is in support of Li et al. (2020) and Schneider et al. (2021): Demographic traits can influence how consumers perceive particular dimensions of video attributes. In particular, raters tend to assess specific product features through so-called aspect ratings. These aspect ratings on different dimensions then impact the overall product rating. Subsequently, users describe these aspect ratings in their own corresponding opinion (key-) words in detail in which they share publicly. In Ho et al. (2017), the authors conclude that consumers' decision whether to post a rating and what rating to post is influenced by disconfirmation. This relates to the disparity between the expected and experienced evaluation of the same product. In particular, when the degree of disconfirmation is extensive then an individual is more likely to leave a rating. Also, this rating may not neutrally express the consumer's post-purchase assessment. The authors also find that infrequent raters are more prone to disconfirmation than frequent raters. However, some questions of their study have been raised to what extent disconfirmation affects aspect ratings and what role demographic traits play in disconfirmation.

2.3 Video chart success

When consumers search for new online content they retrieve information from video charts. These ranked lists reflect the popularity of movies/ shows. Moreover, consumers are particularly interested in how rankings of items on these charts change over time. They also speculate about possible changes (Bradlow et al., 2001). It is important to understand what drives chart performance since the appearance and especially the extended duration on it determines a video's success (Sochay, 1994). This also applies to SVOD services. For instance, Jang et al. (2021) studied the similarities between countries in online video streaming consumption and the role of cultural factors in determining similarities. The authors find that consumption is analogous when cultural similarity, geographic proximity and linguistic similarity are high. From a business perspective, it is significant for content creators and marketers to learn the drivers of chart success since SVOD companies do not disclose any viewership data (Garvey, 2021). In Netflix, this chart is known as "The Daily Top 10" which lists the ten most popular content on a daily basis. As Lotz (2021) argue, it is difficult to derive how much this chart can explain on how subscribers use Netflix. Therefore, Legoux et al. (2016) suggest linking content characteristics to UGC in the form of reviews. These critics contain information on attribute-based consumer sentiment. This is in line with Ainslie et al. (2005) who find that for example actors have a direct effect on consumers' choice while releasing a movie with other movies of the same genre can harm its success. Nevertheless, while many studies on Recommender Systems have attempted to develop a framework for personalized content (Ansari et al., 2018; Lee et al., 2021; Wei et al., 2016), the findings of Başaran (2021) show that SVOD charts have a higher likelihood in catching viewers' attention on popular and trending videos.

2.4 Research contribution and hypotheses

As described in the previous section 2.3, one way in finding new and popular content from SVOD platforms is via charts. Both schools of thoughts in 2.1 acknowledge consumers' variety-seeking tendency, especially in product characteristics which can be derived from social tags. As argued by Xu et al. (2021), online streaming services can benefit by offering a greater variety of distinct product categories and attributes. In addition, Moreno et al. (2017) point out that increasing variety on the latter can prevent demand uncertainty without naturally raising the quantity of products offered. However, some questions in regard to these studies remain: First, Xu et al. (2021) do not indicate what these attributes are and to what extent distinction is defined. Second, little is known in the SVOD domain to what degree the enlargement of attribute-based variety stops increasing the number of videos offered. Nonetheless, this thesis asserts that a certain number of varied, distinct social tags contributes to the chart success of videos. In particular, the underlying mechanism which supports this assumption can be illustrated as follows: First, since consumers have a variety-seeking tendency, they tend to aim to maximize their utility by balancing product attributes (McAlister, 1982). In addition, when they anticipate variety in the future they satiate slower in the present (Sevilla et al., 2016). Second, videos are differentiated goods and their hidden characteristics can be extracted from the plot (Mukherjee et al., 2011). Latent feature diversification can aid to reduce uncertainty, i.e. "choice overload" (Willemsen et al., 2016). Unknown attributes can be obtained from sources like IMDB where this information is revealed via UCG after viewing a video (Liu et al., 2016). Moreover, IMDB users support others through their contributions and the perceived uniqueness of their tags leads them to even create more social tags (Beenen et al., 2004). The advantage of such "swarm intelligence" (Lau et al., 2015) is that video tag lists are formed in non-taxonomic relationships which helps to overcome the problem of social synonymy and polysemy (Trabelsi et al., 2010). As a result, the ideal set of social tags contains a mix of well-known and latent video attributes (Wei, 2020). Thus, the variety of social tags increases demand concentration which consequently pushes the interest for popular content (Tan et al., 2017), especially on SVOD charts. Reflecting these findings, I therefore define the following terms: First, variety is considered as the total number of distinct social tags per video. Second, social tags (such as Netflix original, murder, violence or friendship) which appear more frequent than others can be viewed as well-known labels. On the other hand, infrequent social tags (like animal rights activist, Oklahoma or polygamist) can be regarded as latent video attributes. Consequently, the hypotheses below can be formulated:

H1: *The variety of (i.e. the total number of distinct) social tags has a positive influence on SVOD chart success.*
H2: *Frequent social tags have a positive impact on SVOD chart survival.*
H3: *Infrequent social tags have a positive influence on SVOD chart survival.*
H4: *The equal proportion of frequent and infrequent social tags leads to a higher SVOD chart survival.*

Section 2.2 discusses the influence of consumers' demographic characteristics on product ratings. In Roos et al. (2014), the authors find that highly rated content results from raters' preferences that match their demographic traits. They also argue that demographic characteristics are a reflection of social trends. Yet, when it comes to SVOD content little is known as to whether they affect video chart success. That is why Lee et al. (2015) suggest putting product attributes into consideration next to user ratings. However, websites like IMDB do not provide aspect ratings but rather social tags which can be considered as a result of (aspect) ratings. There is little empirical research on the connection between social tags and user ratings. Lee et al. (2021) have attempted to study this tradeoff. Yet, their research tends to focus on personalized content recommendations rather on its effects on SVOD chart success. Last, Ho et al. (2017) find that consumers who perceive a greater level of disconfirmation have a higher likelihood to rate a video. They also conclude that infrequent raters are more prone to disconfirmation. However, demographic elements have not been put into attention which can aid to better explain the effects of disconfirmation. Hence, this hypothesis can be expressed:

H4: *Average ratings by consumer segment (as reported in IMDB) each have a positive impact on SVOD chart success.*

In summary, the aim of this thesis is to study the effects of social tags and ratings by consumer segment. First, social tags will be examined according to the theory of consumer variety-seeking tendency. Second, I will analyze video ratings based on various user segments in accordance to their demographic characteristics. Last, social tags and viewer ratings will be studied to conclude its effects on immediate and ongoing SVOD chart success. A schematic representation of the variables I intend to research is represented in Figure 1. A brief overview of relevant literature is shown in Table 1.

Figure 1. Conceptual Model

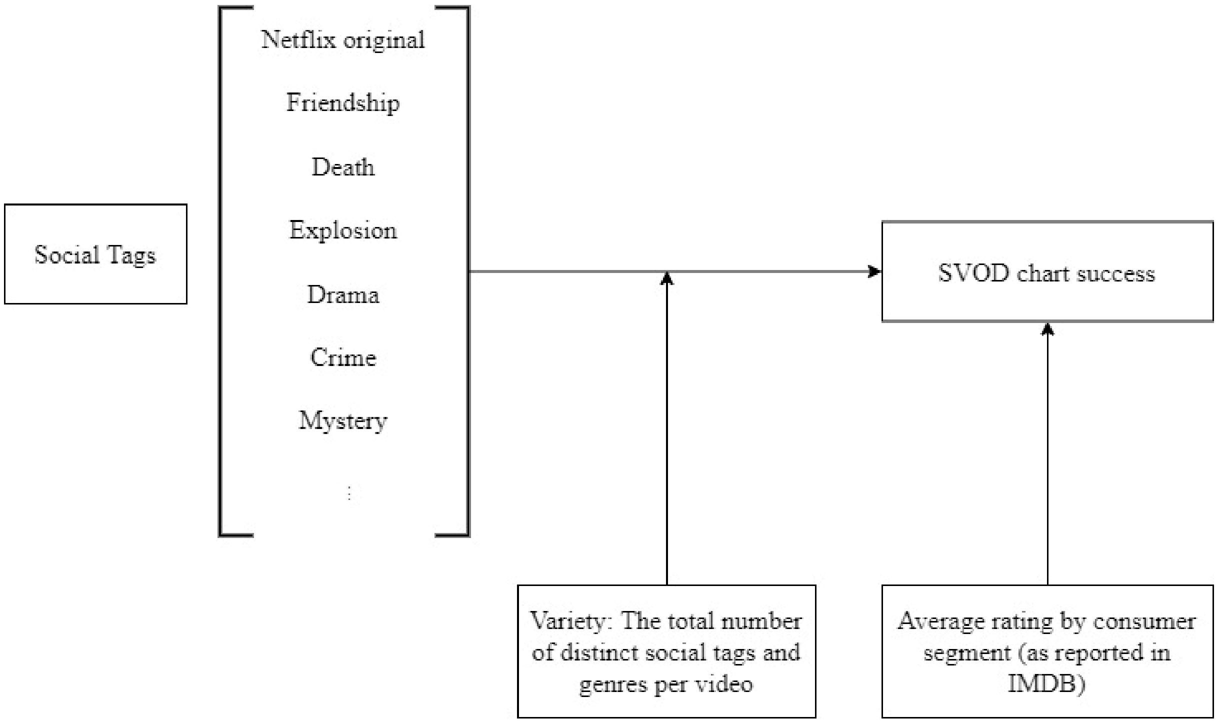


Table 1. Comparison of existing literature on the drivers of SVOD chart success

	Effect of	**Effect on**
Xu et al. (2021)	Users' switching behavior	Online video game streaming
Tan et al. (2017)	Product variety	SVOD demand concentration
Willemsen et al. (2016)	Latent feature diversification	Personalized movie recommendation
Roos et al. (2014)	Latent movie attributes	Local demand for movies
Nam et al. (2014) Nam et al. (2017)	Social Tagging	Brand equity
Wei (2020)	Typical and atypical movie features	Box-office performance
Lau et al. (2015)	Social Tagging (social synonymy and polysemy)	Ontology
Lee et al. (2015)	Social networks	Online movie ratings
Li et al. (2020)	Online reviews, product attributes	Online movie ratings
Schneider et al. (2021)	Multidimensional ratings	Online movie ratings
Wei et al. (2016)	Social Tagging, user ratings	Personalized movie recommendation
Legoux et al. (2016)	Online reviews	Movie survival
This Thesis	**Social Tagging, User ratings**	**SVOD chart success**

3 Methodology

This chapter outlines the following segments: First, a brief overview on the institutional background is provided in 3.1. Next, the data collection and transformation process as well as the variable operationalization is illustrated in 3.2. Section 3.3 describes the model.

3.1 Research context

The goal of this study is to provide insights on how social tagging and ratings by consumer segment affect SVOD chart success. The leading SVOD service Netflix (Bean, 2020) is the focal view of this thesis. Despite its market dominance, Netflix is currently facing a major challenge of customers churning due to a competitive market (Paris, 2021). It is therefore vital to understand the drivers of chart performance for two reasons: First, Netflix barely discloses any viewership data. Second, once content appears on a chart, its popularity as well as views increases (Garvey, 2021). In Netflix, this video chart is known as "The Daily Top 10" (Johnson, 2020).

3.2 Data

In this section I describe the data I have collected to shed more light on how social tagging and ratings by consumer segment influence a video's performance on "The Daily Top 10". The framework I develop for this research is based on an observational approach.

3.2.1 Raw data collection

In general, the data collection process can be divided into two parts: First, data on The Daily Top 10 was obtained via web scraping from the website The Numbers[1] (TN). It gathers Netflix information on the daily ten most trending videos in the United States. The sample covers a period between March 24, 2020 and September 27, 2021. The daily TN list contains the following elements: Rank, YD, LW, title URL, title, type, Netflix exclusive, Netflix release date, days in Top 10, viewership score and chart date. Table 2 provides a detailed overview of these variables. The total dataset contains 5,520 entries. Second, since TN supplies limited video information, I retrieved data from IMDB via its API by using the Python library IMDBpy (Alberani et al., 2021). For this, each individual TN title was matched with IMDB. With the purpose of investigating the research problem, I obtained social tagging data for each movie/ show from its plot keywords. These consist of lists of UCG which reveal video attributes that are maintained by IMDB users. A total of 38,545 social tags with 15,701 distinct labels were gathered. In addition, for each video, I collected ratings by consumer segment which are divided into the total number of votes and the average rating per viewer segment. Demographic characteristics of raters are broken down into the following components: IMDB users, aged under 18, aged 18-29, aged 30-44, aged 45 plus, males, males aged under 18, males aged 18-29, males aged 30-44, males aged 45 plus, females, females aged under 18, females aged 18-29, females aged 30-44, females aged 45 plus, top 1000 voters, US users and Non-US users. Next, in order to get a more granular picture, for each video I gathered genre information from IMDB as well. A total of 1,602 with 25 distinct genres were retrieved. Finally, since the video

[1] See The Numbers Netflix Daily Top 10: https://www.the-numbers.com/netflix-top-10

"Untold" (IMDB ID: 7642104) neither revealed any information to social tagging nor ratings by consumer segment, this content was removed.

Table 2. The Numbers variables

Variable	Description
Rank	The current rank of a video in The Daily Top 10.
YD	This reflects a video's previous rank in the antecedent day.
LW	LW shows a content's highest rank within the last six weeks.
Title URL	The URL of a content that can be found on The Numbers.
Title	The title/ name of a video.
Type	The type of content: TV show, movie, Stand-up comedy and concert/ performance.
Netflix exclusive	This represents if a video is a Netflix exclusive (Yes/ No).
Netflix release date	The date when a content was released on Netflix.
Days in Top 10	The (current) cumulated number of days a video is in The Daily Top 10.
Viewership score	The viewership score is assigned to each show based on its historical daily ranking, assigning 10 points for each No. 1 position, 9 points for each No. 2 position etc.
Chart date	The current date a video is the chart.

3.2.2 Data transformation and variable operationalization

The raw data collection process resulted in two different datasets: TN and IMDB. The following data transformation was done in MS Excel and Python Pandas (McKinney, 2010). In the first step, the TN dataset was examined. Since no faulty entries were found, the dataset was aggregated on a title-level which resulted in 527 distinct videos. Next, for each observation the following variables were assigned: Rank, YD, LW, type, Netflix exclusive, Netflix release date, Viewership Score and days in Top 10.

A quick scan revealed that the Viewership Score was inconsistent due to missing values. In fact, 38% of the total 5,520 entries did not contain a score. Hence, this data was dropped. Since the scope of this study does not cover changes in ranking over time, the variables YD and LW are seen as obsolete. However, to better understand instant chart success, I consider the variable Rank as the position a video displays on its initial appearance on the Daily Top 10. In a similar fashion to other studies (Ahmad et al., 2017; Gopinath et al., 2013; Lee et al., 2018), I transformed the Netflix release date into a categorical variable by determining the day of the week when it was released. For the last variable days in Top 10, the (cumulated) maximum value was determined. For instance, the movie "The Dark Knight" stayed in The Daily Top 10 for eight days, thus, the value 8 was assigned.

In order to match TN with the IMDB dataset, each TN title was used in order to retrieve its IMDB ID. Based on this approach, the IMDB dataset was merged with TN in Pandas. The resulting data frame extended TN with genre, social tags and ratings by consumer segment. The first two variables contain the total number and the latter represents the overall (average) rating. In the next step, for each distinct social tag a column was created. In this, a dummy variable was coded as to whether a particular label is included in the tag list of a video or not. In particular, if for instance the social tag "Netflix-original" is in the list then the value 1 is assigned, otherwise 0. The same procedure is applied to genres. For each demographic characteristic in ratings by consumer segment two columns were set up which include the total number of votes and the average rating. For example, males who evaluated the show "Tiger King" are broken down into two columns: 37,615 votes and the average rating 7.6. In addition, a column with a dummy variable was created whether a content was right-censored or not (i.e. Yes = 0, No = 1).

The measurement of the variety variables is described as follows: In general, the measurement of product variety is a complex and hard task (Alexander, 1997). One approach to operationalize this variable is through quantifying breadth (Lancaster, 1990). Therefore, I measure the breadth of variety by counting the number of distinct social tags per content.

Last, the measurement of ratings by consumer segment is done by taking the average rating per demographic group. Moreover, the demographic information is provided voluntarily by IMDB users and only those who have provided this information are considered. The final dataset along with the operationalization of the model variables is shown in Table 3.

Table 3. Variable Operationalization

Variable	Operationalization
Survival	The total survival time of a video *i* in the Daily Top 10 in days *t*.
Rank	The initial debut rank of a video *i* on its first appearance on the chart. Categorized from 1 to 10.
Variety Breadth	The number of distinct social tags per video *i*.
Type	Categorical: TV Show, Movie, Stand-Up Comedy, Concert/Performance
Netflix exclusive	This variable is dummy coded and shows if a video was exclusively released for Netflix. (Yes = 1, No = 0)
Netflix release day	Categorical: The day of week when the content was released.
Social Tag	If a (distinct) social tag is contained in a video's tags list then the dummy variable is set to 1, otherwise 0. The set displays 15,701 unique social tags.
Genre	This is a set of dummy coded variables for 25 different genres defined by IMDB. If a genre is contained in a video's genre list then the dummy variable is set to 1, otherwise 0.
Rating per consumer segment	The average rating by consumer segment based on its demographic characteristic per video *i* (as reported in IMDB).

3.2.3 Dimensionality reduction

As described previously, the data obtained contain 15,701 unique social tags and 18 consumer segments (as reported in IMDB). Both datasets represent these major characteristics: High dimensionality, sparsity and multicollinearity. In order to optimize computational performance on the consequent analysis, this section therefore applies a dimensionality reduction method. The aim is to project both datasets into a lower dimension space so that it becomes user-readable. In addition, the reduced amount of data stored enhances computation.

One method of dimensionality reduction is Principal Component Analysis (PCA). Unlike other techniques (e.g. Factor Analysis), PCA does not make any assumptions about the underlying data structure. Since it takes multicollinearity into consideration, the transformation of the original variables results into principal components which are independent to each other. These converted variables account for the maximum possible proportion of total variation in the initial variables. This aspect is significant, especially for regression models such as the previously described Cox PH model. In general, PCA is defined as

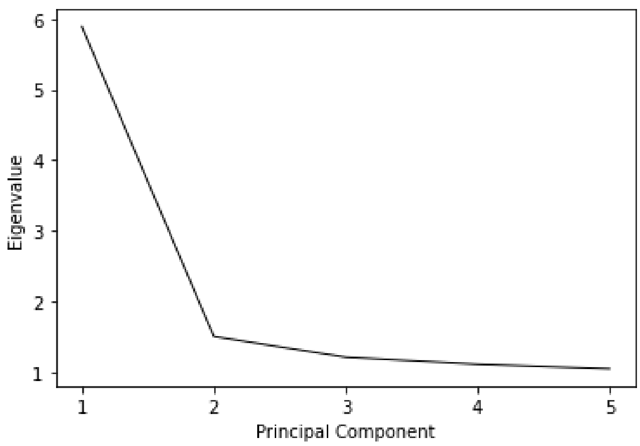
$$X_i = \beta_{i1}P_1 + \beta_{i2}P_2 + \cdots + \beta_{in}P_n$$

where X_i represents the original variable, β_i are the loadings and P_i are the principal components (Cooper, 1983).

Social Tags

In order to apply PCA, I use the Python library "scikit-learn" (SKL) (Pedregosa et al., 2011). First, the values for each observation are standardized with SKL's "StandardScaler" to ensure data consistency. The standard score of a sample x is calculated as $z = \frac{(x-u)}{s}$ where the mean u is set to 0 and the standard deviation s is set to 1 to maintain the sparsity structure of the data. Next, the scaled data is transformed and then fitted for PCA. In PCA, I set the initial number of components equally to the number of observations (i.e. 527). The scree plot in Figure 2 shows that the eigenvalues of the first five principal components are 5.892, 1.503, 1.208, 1.109 and 1.044 respectively. According to the so-called Kaiser criterion, only principal components with eigenvalues greater than one are kept for subsequent analysis (Kaiser, 1960). Thus, the final (survival analysis) model will contain a reduced set of five variables for social tags.

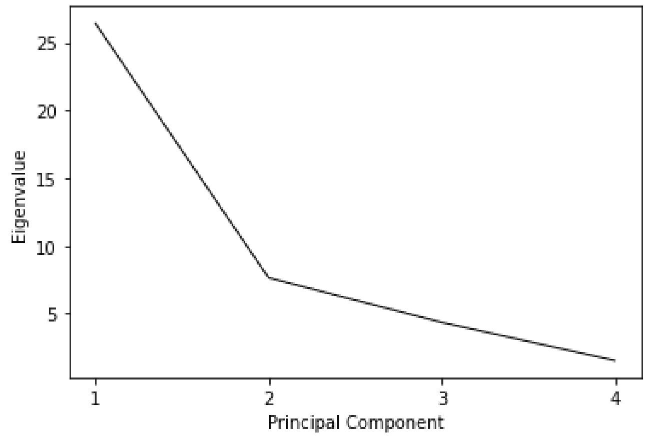
Figure 2. Scree plot of the PCA on social tags



Ratings by consumer segment

For ratings by consumer segment, I apply the same PCA approach like for social tags. Hence, the same SKL steps are implemented (Pedregosa et al., 2011). First, the data is also standardized and scaled where the mean u is 0 and the standard deviation s is 1. Then, the scaled data is transformed and fitted for PCA. The scree plot in Figure 3 shows that the eigenvalues for the first four principal components are 26.438, 7.633, 4.341 and 1.525 respectively. The cumulative sum of these principal components represent around 97% of the total variation. Similar to the previous PCA on social tags, the Kaiser criterion determines the optimal number of principal components based on eigenvalues greater than one. For ratings by consumer segment, the analysis reveals that the final Cox PH model will contain four principal components.

Figure 3. Scree plot of the PCA on ratings by consumer segment



3.3 Model

My objective is to identify the drivers of SVOD chart success, i.e. the effects of social tagging and ratings by consumer segment. In particular, I aim to analyze to what extent variety and demographic traits affect the chart endurance of certain videos. The outcome of interest is the dependent variable (DV) survival time of a content measured in days. As mentioned previously, online videos are highly differentiated products (Mukherjee et al., 2011). Similar to online games, these are considered as experience goods due to their unique attributes which means that users can continue consuming it until they satiate (Heo et al., 2021). Therefore, I define this as time-to-event, i.e. the event occurs when a video exits the Netflix Daily Top 10. For this, I use a multivariate survival analysis model to estimate the effects on chart success.

This approach is mostly applied to predict and explain time until the focal event occurs. Since the dataset in 3.2 may be subject to censoring, it can be assumed that the survival time does not usually follow a normal distribution. Hence, DV should be estimated with a nonparametric or semiparametric method (e.g. hazard function) instead of a parametric approach (e.g. regression analysis). While nonparametric models have the advantage that they can estimate survival time without knowing its distribution, it is difficult to analyze the effect of one independent variable (IV) while controlling for other variables (Im et al., 2018). In contrast, semiparametric methods such as the Cox Proportional Hazards model are able to handle a set of IV's while a baseline hazard function is assumed to be model-free (Cox, 1972). In Im et al. (2018), for instance, the authors applied a Cox PH model and the effects of the covariates were concluded from the relative hazard ratios. For example, they found out that a lower rank debut decreases a song's survival time on a chart. The Cox PH model can account for time-varying predictors and coefficients by applying the counting process formulation of Andersen et al. (1982). However, this aspect does not apply to this study since the covariates in 3.2 are assumed to be time-fixed. In particular, neither TN nor IMDB provide any time information on when certain attributes were added to a video. Similar studies in the motion picture industry have also applied the Cox Proportional Hazards model (e.g. Jang et al. (2021) or Kumar et al. (2014)).

The model of this study can be defined as follows:

$$\begin{aligned} S(t) = S_0(t) \times exp(&\beta_1 Rank_i + \beta_2 Variety\ Breadth_i + \beta_3 Type_i \\ &+ \beta_4 Netflix\ Exclusive_i + \beta_5 Netflix\ Release\ Day_i + \beta_6 Social\ Tag_i \\ &+ \beta_7 Genre_i + \beta_8 ratings\ by\ consumer\ segment_i) \end{aligned}$$

where,

- t denotes the survival time in days,
- $S(t)$ is the hazard function determined by a set of covariates,
- The coefficients β_i measure the effect size of covariates,
- S_0 is known as the baseline hazard. It corresponds to the value of the hazard if all IV's are equal to zero.

The terms $exp(\beta_i X) = \frac{S(t)}{S_0(t)}$ are called hazard ratios (HR). A value of β_i greater than zero, or equivalently a HR greater than one, indicates that as the value of the i^{th} covariate increases, the event hazard increases and thus the length of survival decreases. For instance, if the HR for a particular treatment group (e.g. video type) is above 1, then it is positively associated with the event probability and hence negatively associated with the length of survival.

To check whether a fitted Cox PH model sufficiently describes the data in 3.2, I will test the three following assumptions: First, the "Schoenfeld Residuals" will examine the proportional hazards assumption (Schoenfeld, 1982). Next, the relationship between the IV's and the log hazard will be assessed for its nonlinearity with the "Martingale Residuals" (Barlow et al., 1988). Last, the deviance residuals will test for influential observations or outliers (Davison et al., 1989).

4 Results

In this chapter the following sections are outlined: First, in 4.1 the descriptive statistics provide an initial overview of the variables obtained in chapter 3. Second, 4.2 shows the main results of the analysis along with hypotheses tests. Then, the model in 3.3 is checked for its validity and additional results are given in 4.3. Last, section 4.4 summarizes the results and findings of this study.

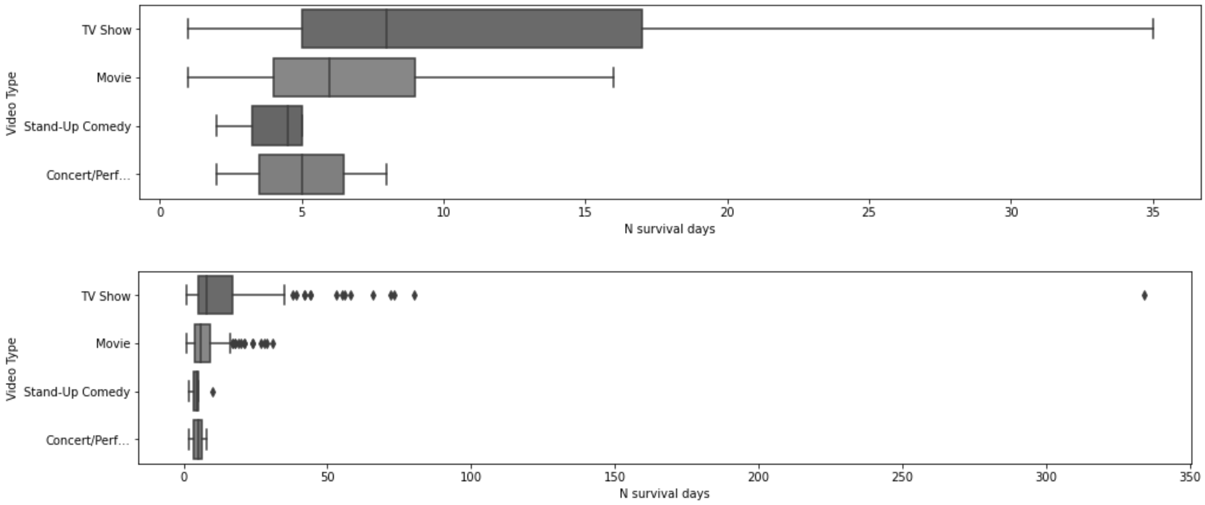
4.1 Descriptive statistics

In this section, Table 5 lists all summary statistics for the relevant variables. For this, I used the Python libraries "Seaborn" (Waskom, 2021) and "Matplotlib" (Hunter, 2007). The sample retained covers a total of 552 days on the Netflix Daily Top 10. As described in Chapter 3, the initial dataset was aggregated on a title-level which resulted in 527 observations.

The focal view of this thesis is the dependent variable, i.e. the number of days a video survived on the chart (see Appendix A). This variable shows that a content survived on average 10.455 days on the chart with a standard deviation of 17.293 days. Identical to social tags and genres, this variable is tested whether the sample follows a normal distribution or not. The Shapiro-Wilk Test results in a test statistic of .337 with a p-value $< .05$. Thus, there is evidence that the data is not normally distributed since the p-value is below the significance threshold. Last, it is important to note that "Cocomelon" (IMDB ID: 12427840) can be considered as an outlier since it stayed for 334 days on the chart.

To get a more differentiated view, the variable survival days is broken down into different video types. For the upper plot in Figure 4, the outliers were excluded while they were included in the plot below. The mean survival for TV show, movie, stand-up comedy and concert/ performance are 14.662, 7.246, 4.833 and 5.000 respectively. This aspect is significant since the Cox PH model allows for analysis of various treatment groups. The following section 4.2 will provide a more in-depth analysis on the chart survival of these different video types.

Figure 4. Box plots: Survival days per video type



Prior to reducing its dimensionality, the variable social tags displays that the mean number of social tags per video is 73.110 with a standard deviation of 108.752. The content "The Dark Knight" (IMDB ID: 0468569) which contains 788 tags can be considered as an outlier. As shown, the data is not normally distributed which is supported by the Shapiro-Wilk Test of Normality (Shapiro et al., 1965). The test statistic is .678 with a corresponding p-value $< .05$. Since the p-value is significant, the null hypothesis is rejected and there is no support that the data is normally distributed (see Appendix B).

The distribution of genres per video shows that on average a video contains 3.034 genres with a standard deviation of 1.481. The Shapiro-Wilk Test of Normality (Shapiro et al., 1965) shows that the test statistic is .919 with a corresponding p-value $< .001$. Thus, the null hypothesis is rejected and there is no evidence that the sample is normally distributed. While the content "Love, Death & Robots" (IMDB ID: 9561862) contains the maximum number of genres (i.e. 11), 79 observations are assigned to at least one genre (see Appendix C).

Netflix exclusive is a dummy coded variable. This displays whether the SVOD platform either has co-produced a content with another network, it has exclusive international streaming rights to a video or it commissioned as well as produced a movie/ show[2]. The majority of videos were Netflix exclusive (i.e. N = 328; 62%) while 199 (38%) were not. In particular, the former includes the following video types: TV Show 178 (33.78%), movie 142 (26.94%), stand-up comedy 6 (1.14%) and concert/ performance 2 (0.38%).

From the data transformation in 3.2.2, the variable Netflix release date was converted into a categorical variable based on the day of a week when they were released. Thus, the number of videos obtained are Monday 64 (12%), Tuesday 63 (12%), Wednesday 96 (18%), Thursday 53 (10%), Friday 189 (36%), Saturday 20 (4%) and Sunday 42 (8%). This shows that most of the content was released on a Friday.

Rank is a variable which describes the debut position when a video initially appeared on the Netflix Daily Top 10. The categories range from 1 to 10 where 1 is the most favorable rank and 10 the least favorable position. It shows that the average debut rank is 5.928 with a standard deviation of 2.965. The total number of videos gathered for this variable is listed in Table 4.

Table 4. The number of videos and its debut rank

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
N	53	32	59	39	43	58	55	44	67	77
Percentage	10%	6%	11%	7%	8%	11%	11%	8%	13%	15%

The last variable contains ratings per consumer segment which was retrieved from IMDB. Demographic characteristics are broken down into gender, age, geographical location as well as whether a rater is considered as one of the top 1,000 voters on IMDB. As mentioned in 3.2.2, the main component of this variable is the average rating by consumer segment per video. The table in Appendix D lists the complete summary statistics for this variable. In general, the average rating of a content is 6.542 with a standard deviation of 1.140.

[2] See: https://help.netflix.com/en/node/4976

Table 5. Summary statistics

Variable	N	Mean	SD	Min	Max
Survival	552	10.455	17.296	1.000	334.000
Per video type:					
• TV Show	230	14.662	24.969	1.000	334.000
• Movie	289	7.246	4.870	1.000	31.000
• Stand-up comedy	6	4.833	2.787	2.000	10.000
• Concert/ performance	2	5.000	4.243	2.000	8.000
Social tags	38,544	73.110	108.752	1.000	788.000
Genre	1,601	3.034	1.481	1.000	11.000
Netflix exclusive	527	.623	.485	0	1.000
Rank	527	5.926	2.967	1.000	10.000
Rating	527	6.542	1.140	2.900	9.300

4.2 Main results and hypothesis tests

In this section, I discuss the main results and hypothesis tests of the analysis. As mentioned in Chapter 3, I use a Cox Proportional Hazards regression model. Therefore, the Python library "lifelines" was implemented for this purpose (Davidson-Pilon, 2021).

4.2.1 Results

Prior to the discussion on the drivers of chart success, it is important to note that I focus on the reduced model (3) where stratification is applied. This allows covariates to be added to the model without estimating its effects. The dataset is partitioned into smaller sets by the stratifying variables to improve the overall model (Anderson et al., 1995). In particular, I stratify the data per video type which leads to an enhanced partial Akaike information criterion (AIC) of 4,573.47 on Model 3. In contrast, the full model (1) which includes all covariates displays a partial AIC of 5,303.46. Hence, the reduced model (3) is more suitable to explain the overall variation. Further details on model comparison are given in 4.3.

4.2.2 Length of chart survival

Figure 5 shows the survival function of Model 1. In particular, the results illustrate that most of the videos survive for 25 days on the Netflix Daily Top 10. The grey dotted line represents the threshold to which the survival probability drops below 1%. This accounts for almost 5% (N = 27) of the videos. In contrast, Figure 6 shows the stratified Model 3 based on the type of content. The black dashed line for TV Show lies above the lines of the other video types. Thus, TV shows outperforms the other forms of content in terms of survival on the Netflix Daily Top 10. On the other hand, 289 movies (grey dashed line) which is around 55% of the total videos survived up to 16 days on the chart until their survival probability dropped near 0%. Concert/ Performance survived 8 days on the chart with a probability of around 46% while Stand-Up Comedy remained 10 days on the chart with a survival probability of almost 5%. A log-rank test was applied to assess the statistical difference of all four distributions. The X^2 value is 8.81. Therefore, the distribution is significantly different (p < .05).

Figure 5. Survival function of Model 2

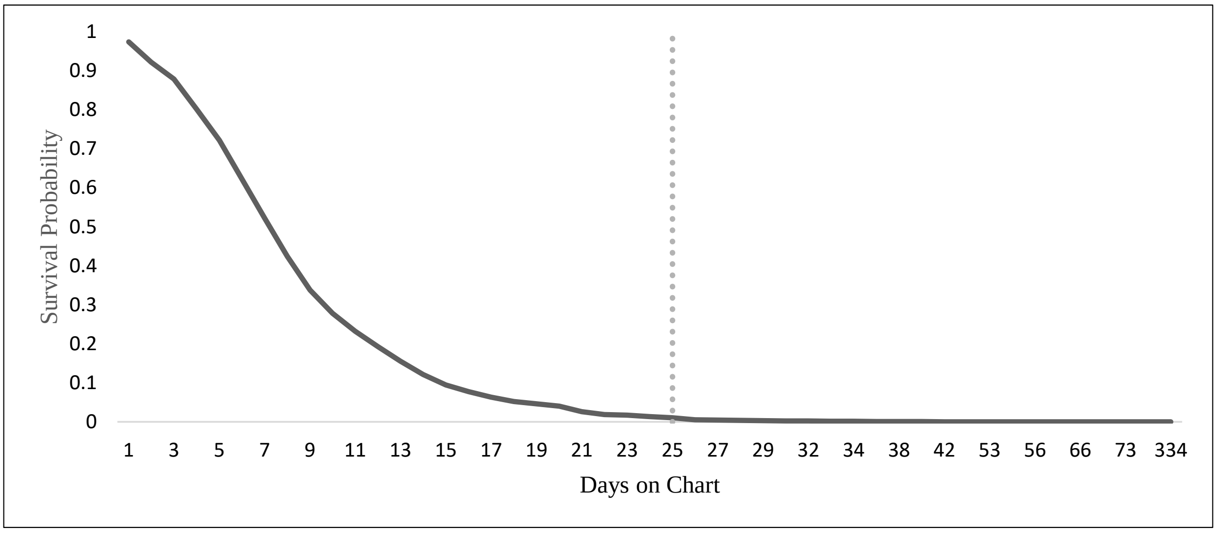
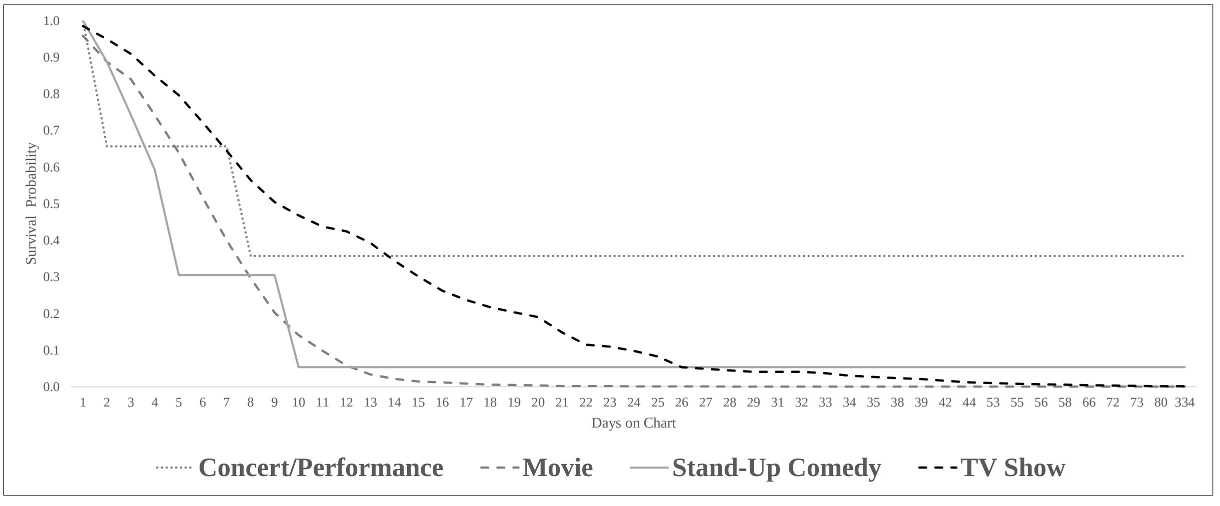


Figure 6. Model 3: Survival probability per video type



4.2.3 The drivers of chart survival

One way to determine the drivers of chart survival is by examining the effect of the covariates. This can be derived from the relative hazard ratios of the Cox Proportional Hazards model (see Appendix E). Table 6 provides an overview of the results for each factor of Model 3. The focal view of the analysis lies on the significant variables with p-value < .05 or those which are considered marginally significant with a p-value between .05 and .09.

Table 6. Covariate results of Model 3

Covariate	β	Hazard Ratio	p-value
Debut Rank: 1	-1.58	0.21	<0.005
Debut Rank: 2	-1.30	0.27	<0.005
Debut Rank: 3	-0.76	0.47	<0.005
Debut Rank: 4	-0.89	0.41	<0.005
Debut Rank: 5	-0.55	0.58	0.02
Debut Rank: 6	-0.40	0.67	0.07
Debut Rank: 7	-0.19	0.83	0.39
Debut Rank: 9	-0.10	0.91	0.66
Debut Rank: 10	0.18	1.19	0.40
Netflix Release Day: Monday	0.29	1.33	0.11
Netflix Release Day: Tuesday	-0.23	0.80	0.18
Netflix Release Day: Wednesday	-0.06	0.94	0.70
Netflix Release Day: Thursday	0.27	1.31	0.14
Netflix Release Day: Saturday	0.38	1.46	0.17
Netflix Release Day: Sunday	0.31	1.36	0.12
Netflix Exclusive	0.24	1.28	0.07
Action	0.10	1.10	0.14
Adventure	0.15	1.17	0.03
Animation	-0.02	0.98	0.81
Biography	0.05	1.05	0.31
Comedy	0.04	1.04	0.55
Crime	0.03	1.04	0.57
Documentary	0.03	1.03	0.65
Drama	-0.08	0.92	0.16
Family	-0.30	0.74	<0.005
Fantasy	0.05	1.05	0.39
Game Show	-0.08	0.92	0.11
History	-0.02	0.98	0.63
Horror	0.10	1.10	0.07
Music	0.04	1.04	0.38
Musical	-0.09	0.91	0.06
Mystery	-0.05	0.95	0.33
Reality TV	0.08	1.08	0.21

Romance	-0.11	0.90	0.04
Sci-Fi	-0.03	0.97	0.61
Short	-0.05	0.95	0.36
Sport	0.01	1.01	0.88
Talk Show	0.05	1.05	0.33
Thriller	-0.09	0.92	0.23
War	-0.04	0.96	0.40
Western	0.06	1.06	0.23
T1	-0.03	0.97	0.22
T2	0.09	1.10	0.04
T3	-0.02	0.98	0.71
T4	-0.003	0.997	0.94
T5	0.03	1.03	0.53
R1	0.03	1.03	<0.005
R2	0.06	1.06	0.02
R3	0.01	1.01	0.76
R4	0.07	1.07	0.09

Note: For each categorical variable, a base reference was excluded for the analysis. These are debut rank 8 and Netflix release day Friday. The principal components derived from 3.2.3 are T1 to T5 for social tags and R1 to R4 for ratings by consumer segment.

To recap, a hazard ratio greater than one is positively associated with the event probability of exiting the Netflix Daily Top 10 and, thus, negatively impacts the survival on the chart. First, the debut ranks one to six are considered (marginally) significant for Model 3. They all display a positive effect on chart survival. For instance, if a video's initial position was equal to one then the HR is .21. This means that a one unit increase in this variable will increase the baseline hazard by a factor of 79%. Analogously, one unit increase in the debut ranks two to six will increase the baseline hazards by .73, .53, .59, .42 and .33 respectively. Moreover, the analysis shows that in terms of chart survival having a debut rank 1 has the highest explanatory power among all variables.

Videos which were considered as Netflix exclusive had a HR of 1.28 and therefore decreased the baseline hazard compared to the control group (i.e. not exclusively produced for Netflix). In essence, being a Netflix exclusive will reduce chart survival by 28%. This covariate was marginally significant due to its p-value of .07.

The ratings by consumer segment show that R1 and R2 are significant (p-value < .05). In order to determine which demographic groups are represented by R1 and R2, the corresponding loadings for each principal component is calculated as follows:

$$Loadings = Eigenvalue \times Explained\,Variance^2$$

Consumer segments which are highly loaded on R1 are “aged under 18” and “females aged under 18”. This covariate shows a HR of 1.03. Thus, this negatively impacts the length of chart survival by decreasing the baseline hazard by 3%. For R2, on the other hand, all other demographic groups (except for “males aged under 18”) load highly on this principal component. Similar to R1, the HR of R2 is 1.06. This increases the event probability of exiting the Netflix Daily Top 10 by 6% and therefore reduces the number of survival days on the chart.

4.2.4 The effects of variety

To better understand the effects of variety on chart performance, I first examine the variables related to social tags. These are the total number of distinct labels of the corresponding tag list per video. As listed in Table 6, the covariate T2 which was obtained from the PCA in 3.2.3 is significant (p-value < .05) for the Cox PH model (3). This variable shows a HR of 1.1 which means that a one unit increase in T2 will lower the baseline hazard by 10%. As a result, social tags that are represented by this principal component have a negative impact on a video’s chart survival. The ten highest loading keywords on T2 are listed in Table 7.

Table 7. The ten highest loadings on T2

Social Tag	Loading
Apology	.131
Friend	.124
Husband wife relationship	.119
Looking at oneself in a mirror	.113
Telephone call	.112
Friendship	.111
Crying	.109
Crying woman	.099
Boyfriend girlfriend relationship	.097
Party	.096

Genres, on the other hand, show a differentiated picture. Comparable to social tags, these variables represent the total number of distinct video categories of the corresponding genre list for each content. The analysis indicates that five categories are considered (marginally) significant for Model 3. First, the genres Family, Musical and Romance all have a negative β-value. In particular, their hazard ratios are .74, .91 and .90 accordingly. Hence, these three genres decrease the event probability of leaving the Netflix Daily Top 10 where Family has the highest impact by 26%. In contrast, the HR of the genres Adventure and Horror are both above 1. They therefore increase the even probability. As a result Adventure reduces the survival on the chart by 17% and Horror by 10%.

In terms of variety, the results show that Netflix videos which contain social tags that are represented in T2 such as apology, friend or crying have a lower survival probability of being on the Daily Top 10. Contents that are associated with either Family, Musical or Romance are more likely to stay longer on the chart. Those that are linked to Adventure or Horror have a higher probability of exiting the chart earlier compared to others.

4.3 Model fit

First, a goodness-of-fit test will examine to what extent the model describes the data. In particular, the focus will be on key metrics which indicate the model fit. For instance, the Concordance index shows the probability that the prediction follows the same direction as the actual data. This index shows the fraction of concordant pairs in relation to the total number of possible evaluation pairs (Harrell et al., 1996). Next, the log-likelihood ratio test (LRT) is used to compute the difference between the log-likelihood statistic (LL) of the reduced versus full model (Cox, 1984). The partial AIC is an estimator of prediction error which compares the model generating the data and a fitted model (Akaike, 1974).

Table 8 shows all model variations with the following parameters:

(1) Full model with all variables,
(2) Full model where rating per consumer segment is excluded,
(3) Extension of Model 1 with stratification on video type,
(4) Extension of Model 3 where ratings by consumer segment is excluded.

Table 8. Variations of the Cox PH model

Metric	(1)*	(2)*	(3)*	(4)*
partial log-likelihood	-2,598.73	-2,608.81	**-2,236.74**	-2,246.24
Concordance	.73	.73	**.72**	.72
Partial AIC	5,303.46	5,315.61	**4,573.47**	4,584.48
log-likelihood ratio test				
(test statistic)	263.81	243.65	**206.15**	187.15
degrees of freedom	53.00	49.00	**50.00**	46.00
-log2(p) of ll-ratio test	96.41	89.37	**66.87**	60.49

*p-value < .05

The main results I discussed in 4.2 are based on the analysis of Model 3. The motivation for this is based on the following key metrics: First, the models are compared based on their partial AIC. This estimator allows me to examine models with different sets of input variables. In Table 8, Model 3 shows the best partial AIC of 4,665.87. Although Model 4 displays the next best partial AIC of 4,584.48, the difference in this metric compared to the previous model shows the necessity for controlling the variable ratings by consumer segment.

For the Concordance index, Model 1 has the highest value of .73. However, as mentioned previously, this model is not considered for further discussion. This is based on the higher partial AIC compared to that of Model 3. Even though the full model (1) includes all variables and the significant variables are identical to Model 3, the covariate video type is not seen as significant (p-values > .05) for the analysis. The stratification on it allows for an improved model (3) which has a higher explanatory power in terms of the overall variation. The LRT for Model 3 shows that the X^2 value is 206.15. With 50 degrees of freedom and a p-value < .05, this model is seen as significant in relation to its reduced version.

As described in 3.3, the fitted Cox PH model will be analyzed if it sufficiently describes the data based on residual metrics. The Schoenfeld residuals, for instance, examine whether the variables suffice the proportional hazard assumption (Schoenfeld, 1982). The results show that all variables exceed the 95% significance threshold and therefore suffice the assumption. However, this does not account for the variables Action, Animation, Netflix release day

Monday, Sunday and debut ranks 1, 2, 4, 7 and 9. For the purpose of this study, as suggested by Stensrud et al. (2020), the violation of proportional hazards can be neglected since this will not affect the explanatory value of the covariates. Next, the Martingale residuals determine to what extent the relationship between the IV's and the log hazard satisfy nonlinearity. It can be applied to assess the true functional form of a certain covariate (Therneau et al., 1990). According to the underlying Cox PH model (i.e. Model 3), positive Martingale values mean that videos exited sooner than expected. In contrast, negative values mean that a Netflix content survived longer than expected which accounts for around 49% of the total observations. In addition, the results also show that there is no linear relationship between the covariates and the log hazard. Last, the Deviance residuals will test for influential observations or outliers. Unlike the Martingale residuals, this metric is symmetric around zero (with a standard deviation of one) which allows for better interpretation in terms of outliers (Davison et al., 1989). For this, observations with values smaller than -1 or greater than 1 are considered. This applies to almost 29% (N=155) of the total observations. Hence, 155 videos can be either seen as influential or as outliers. A detailed overview on the Martingale and Deviance residuals is provided in Appendix F and G.

4.4 Summary

As described previously, there are various factors which significantly drive the success of being on the Netflix Daily Top 10. In general, the analysis unveils that most of the videos stayed on the chart up to 25 days until their survival probability dropped below 1%. However, the comparison of various Cox PH models in 4.3 has shown that the full model (i.e. Model 1) has a lower explanatory power with regards to the total variation. After stratifying the data of Model 1 on video type, Model 3 illustrates that TV shows outperform the other forms of content in terms of survival. When it comes to the debut rank, the results show that the initial positions one to six have a substantial influence on chart performance. In essence, they all have a positive impact where rank 1 enhances the survival probability the most. On the other hand, videos which were exclusively produced for Netflix are less likely to stay on the chart. In the context of variety, social tags such as apology, friend, husband-wife-relationship, looking-at-oneself-in-a-mirror, telephone-call, friendship or crying have a negative influence on a video's chart survival. For genres, while Family, Musical or Romance can improve the chart performance, Adventure or Horror shorten the length of survival days. Last, after controlling for rating by consumer segment, the analysis reveals that it improves the partial AIC (see Model 1 and 3).

5 Conclusion

This chapter will discuss the findings and its implications. First, I will attempt to answer the defined problem statement and research questions in 1.2. In particular, the model results will address its theoretical and practical significance. Next, I will describe the limitations of this thesis. Last, indications for future research will be provided.

5.1 Main findings and its implications

In this thesis, I provided a first overview on the drivers of chart success on the Netflix Daily Top 10. Moreover, I examined to what extent variety in the context of social tags can influence chart survival. In fact, a set of social tags which contain a varied mix of well-known and latent video attributes have a negative influence on chart success. Examples for the former are keywords such as apology, friend or husband-wife-relationship while the latter are labels like listening-to-music, knocking-on-a-door or man-wears-sunglasses. It can be assumed that these social tags reflect a combination of different types of social ties and activities. Therefore, variety can be expressed as such. In addition, tags such as apology or crying can be considered to potentially trigger negative emotions. Thus, viewers might be initially attracted to these type keywords but their interest fades in the long run (Berger et al., 2011). Another form of social tags are genres. The results reveal that video categories can have both effects (i.e. positive or negative) on a content's survival probability. For instance, Family, Musical or Romance can enhance a video's survival chances while Adventure or Horror can have a negative effect on survival duration. One explanation for this difference in genres is the degree of variety that a certain category can have on different viewers (Berger et al., 2007). However, more light needs to be shed on this aspect.

As described previously, the majority of videos survived for almost 25 days on the Netflix Daily Top 10. In a direct comparison, results show that TV shows lasted longer on the chart unlike movies, stand-up comedy or concert/performance. One explanation for this phenomenon is that TV shows usually consist of multiple episodes. This allows viewers to satisfy their variety-seeking tendency in two ways: First, as their consumption history evolves, their preference for product attributes changes (McAlister, 1982). Second, upcoming episodes allow consumers to anticipate variety in future consumption which creates positive thoughts about that prospective experience and they satiate at a slower pace in the present (Sevilla et al., 2016).

The results on debut ranks indicate that positions one to six have a (marginally) positive influence on chart survival. However, movies normally tend to enter the chart on a higher rank (compared to TV shows) and then quickly exit the chart once popularity has declined. For business practitioners, this finding can aid them to adapt their advertising strategies by extending post-release promotional activities.

From a commercial standpoint, content providers as well as marketers and advertisers are advised to act more independently with regards to Netflix. The results indicate that videos which were exclusively produced (or at least created in close proximity) for Netflix have a lower survival probability on the Daily Top 10.

5.2 Limitations and Future Research

The data obtained for this study is based on the Netflix Daily Top 10 from the United States. Thus, the first limitation is based on its geographical as well as language restriction. In fact, most of the SVOD platforms offer their services in multiple countries and languages in which they maintain their local charts. Therefore, one contribution to future research is to extend this research to multiple locations and languages to generate a more in-depth picture on the drivers of chart success.

Another limitation of this thesis is linked to consumers' tendency for variety-seeking. I attempted to explain this phenomenon in the context of SVOD services by analyzing social tags gathered from IMDB. Nonetheless, the examination of social tags can be enhanced by applying techniques from NLP. This allows one to gain a better understanding for instance on how SVOD subscribers express sentiment through social tags (Berger et al., 2019). In addition, social tags as a form of UCG can be expanded to various platforms other than IMDB. One interesting aspect would be to analyze whether viewers apply the same social tags for a specific content in different websites or not.

The Cox PH model I applied for this study is restricted to time-fixed data. Websites such as Netflix or IMDB are characterized by rapid changes in consumer preferences. That is why I suggest for future research to put more emphasis on time-varying components.

References

Ahmad, J., Duraisamy, P., Yousef, A., & Buckles, B. (2017). *Movie success prediction using data mining*. Paper presented at the 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, India.

Ainslie, A., Drèze, X., & Zufryden, F. (2005). Modeling Movie Life Cycles and Market Share. *Marketing Science, 24*(3), 508-517. doi:10.1287/mksc.1040.0106

Akaike, H. (1974). *A new look at the statistical model identification.* Paper presented at the IEEE Transactions on Automatic Control.

Alberani, D., & Uyar, H. T. (2021). IMDbPY Documentation. In *Release 6.8*.

Alexander, P. (1997). Product variety and market structure: A new measure and a simple test. *Journal of Economic Behavior & Organization, 32*(2), 207-214. doi:https://doi.org/10.1016/S0167-2681(96)00902-X

Andersen, P. K., & Gill, R. D. (1982). Cox's Regression Model for Counting Processes: A Large Sample Study. *The Annals of Statistics, 10*(4), 1100-1120.

Anderson, G. L., & Fleming, T. R. (1995). Model Misspecification in Proportional Hazards Regression. *Biometrika, 82*(3), 527-541.

Ansari, A., Li, Y., & Zhang, J. Z. (2018). Probabilistic Topic Model for Hybrid Recommender Systems: A Stochastic Variational Bayesian Approach. *Marketing Science, 37*(6), 987-1008. doi:10.1287/mksc.2018.1113

Arkenberg, C., Ledger, D., Loucks, J., & Westcott, K. (2021). How streaming video services can tackle subscriber churn. *Digital media trends*. Retrieved from https://www2.deloitte.com/us/en/insights/industry/technology/video-streaming-services-churn-rate.html

Barlow, W. E., & Prentice, R. L. (1988). Residuals for relative risk regression. *Biometrika, 75*(1), 65-74.

Başaran, M. (2021). *An Investigation on the Netflix Platform in the Context of Flow Theory*. Paper presented at the Communication and Technology Congress, Istanbul, Turkey.

Bean, T. (2020, April 19, 2021). Yes, Netflix Is The #1 Streaming Platform—But That Is Changing Quickly Thanks To This New Player, Online. *Forbes*. Retrieved from https://www.forbes.com/sites/travisbean/2020/10/16/yes-netflix-is-the-1-streaming-platform-but-that-is-changing-quickly-thanks-to-this-new-player/?sh=7e6938896a7b

Beenen, G., Ling, K., Wang, X., Chang, K., Frankowski, D., Resnick, P., & Kraut, R. (2004). *Using Social Psychology to Motivate Contributions to Online Communities* Chicago, Illinois, USA.

Berger, J., Draganska, M., & Simonson, I. (2007). The Influence of Product Variety on Brand Perception and Choice. *Marketing Science, 26*(4), 460-472. doi:10.1287/mksc.1060.0253

Berger, J., Humphreys, A., Ludwig, S., Moe, W. W., Netzer, O., & Schweidel, D. A. (2019). Uniting the Tribes: Using Text for Marketing Insight. *Journal of Marketing, 84*(1), 1-25. doi:10.1177/0022242919873106

Berger, J., & Schwartz, E. M. (2011). What drives immediate and ongoing word of mouth? *Journal of Marketing Research, 48*(5), 869-880.

Bradlow, E. T., & Fader, P. S. (2001). A Bayesian Lifetime Model for the "Hot 100"BillboardSongs. *Journal of the American Statistical Association, 96*(454), 368-381. doi:10.1198/016214501753168091

Cooper, J. C. B. (1983). Factor Analysis: An Overview. *The American Statistician, 37*(2), 141-147.

Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society, 34*(2), 187-220.
Cox, D. R. (1984). Effective degrees of freedom and the likelihood ratio test. *Biometrika, 71*(3), 487-493.
Datta, H., Knox, G., & Bronnenberg, B. J. (2018). Changing Their Tune: How Consumers' Adoption of Online Streaming Affects Music Consumption and Discovery. *Marketing Science, 37*(1), 5-21. doi:10.1287/mksc.2017.1051
Davidson-Pilon, C. (2021). lifelines, survival analysis in Python (v0.26.3). *Zenodo*. doi:10.5281/zenodo.805993
Davison, A. C., & Gigli, A. (1989). Deviance residuals and normal scores plots. *Biometrika, 76*(2), 211-221.
Dimitrov, D., Helic, D., & Strohmaier, M. (2018). Tag-Based Navigation and Visualization. In P. Brusilovsky & D. He (Eds.), *Social Information Access: Systems and Technologies* (pp. 181-212). Cham: Springer International Publishing.
Dogruel, L. (2017). Cross-cultural Differences in Movie Selection. Decision-making of German, U.S., and Singaporean Media Users for Video-on-Demand Movies. *Journal of International Consumer Marketing, 30*(2), 115-127. doi:10.1080/08961530.2017.1384710
Durante, K. M., & Arsena, A. R. (2015). Playing the Field: The Effect of Fertility on Women's Desire for Variety. *Journal of Consumer Research, 41*(6), 1372-1391. doi:10.1086/679652
Elberse, A., & Eliashberg, J. (2003). Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures. *Marketing Science, 22*(3), 329-354. doi:10.1287/mksc.22.3.329.17740
Eliashberg, J., Hui, S. K., & Zhang, Z. J. (2007). From Story Line to Box Office: A New Approach for Green-Lighting Movie Scripts. *Management Science, 53*(6), 881-893. doi:10.1287/mnsc.1060.0668
Farquhar, P. H., & Rao, V. R. (1976). A Balance Model for Evaluating Subsets of Multiattributed Items. *Management Science, 22*(5), 528-539. Retrieved from https://pubsonline-informs-org.tilburguniversity.idm.oclc.org/doi/pdf/10.1287/mnsc.22.5.528#
Garvey, M. (2021). The Netflix trending list is an algorithmic mystery, Online. *CNN Entertainment*. Retrieved from https://edition.cnn.com/2021/06/30/entertainment/netflix-trending-movies/index.html
Gopinath, S., Chintagunta, P. K., & Venkataraman, S. (2013). Blogs, Advertising, and Local-Market Movie Box Office Performance. *Management Science, 59*(12), 2635-2654. doi:10.1287/mnsc.2013.1732
Harrell, F. E., Lee, K. L., & Mark, D. B. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in medicine, 15*(4), 361-387.
Heo, S., & Park, J. (2021). Are you satisfied or satiated by the games you play? An empirical study about game play and purchase patterns by genres. *Telematics and Informatics, 59*. doi:10.1016/j.tele.2020.101550
Ho, Y.-C. C., Wu, J., & Tan, Y. (2017). Disconfirmation Effect on Online Rating Behavior: A Structural Model. *Information Systems Research, 28*(3), 626-642. doi:10.1287/isre.2017.0694
Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering, 9*(3), 90-95.
Im, H., Song, H., & Jung, J. (2018). A survival analysis of songs on digital music platform. *Telematics and Informatics, 35*(6), 1675-1686. doi:10.1016/j.tele.2018.04.013

Jang, M., Kim, D., & Baek, H. (2021). How do global audiences of TV shows take shape?: Evidence from Netflix. *Applied Economics Letters*, 1-7. doi:10.1080/13504851.2021.1983916
Johnson, C. (2020). Now - for the first time - you can see what's popular on Netflix. Retrieved from https://about.netflix.com/en/news/see-whats-popular-on-netflix
Kahn, B. E. (1995). Consumer variety-seeking among goods and services: An integrative review. *Journal of retailing and consumer services, 2*(3), 139-148. doi:10.1016/0969-6989(95)00038-0
Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement, 20*(1), 141-151. doi:10.1177/001316446002000116
Khodabandehlou, S., Hashemi Golpayegani, S. A., & Zivari Rahman, M. (2020). An effective recommender system based on personality traits, demographics and behavior of customers in time context. *Data Technologies and Applications, 55*(1), 149-174. doi:10.1108/dta-04-2020-0094
Kumar, A., Smith, M., & Telang, R. (2014). Information Discovery and the Long Tail of Motion Picture Content. *MIS Quarterly, 38*(4), 1057-1078.
Lancaster, K. (1990). The Economics of Product Variety: A Survey. *Marketing Science, 9*(3), 189-206.
Lau, R. Y. K., Zhao, J. L., Zhang, W., Cai, Y., & Ngai, E. W. T. (2015). Learning Context-Sensitive Domain Ontologies from Folksonomies: A Cognitively Motivated Method. *INFORMS Journal on Computing, 27*(3), 561-578. doi:10.1287/ijoc.2015.0644
Lau, T., & Akkaraju, U. (2019). When Algorithms Decide Whose Voices Will Be Heard. *Harvard Business Review*. Retrieved from https://hbr.org/2019/11/when-algorithms-decide-whose-voice-will-be-heard
Lee, D., & Hosanagar, K. (2021). How Do Product Attributes and Reviews Moderate the Impact of Recommender Systems Through Purchase Stages? *Management Science, 67*(1), 524-546. doi:10.1287/mnsc.2019.3546
Lee, S.-Y., Qiu, L., & Whinston, A. (2018). Sentiment Manipulation in Online Platforms: An Analysis of Movie Tweets. *Production and Operations Management, 27*(3), 393-416. doi:10.1111/poms.12805
Lee, Y.-J., Hosanagar, K., & Tan, Y. (2015). Do I Follow My Friends or the Crowd? Information Cascades in Online Movie Ratings. *Management Science, 61*(9), 2241-2258. doi:10.1287/mnsc.2014.2082
Legoux, R., Larocque, D., Laporte, S., Belmati, S., & Boquet, T. (2016). The effect of critical reviews on exhibitors' decisions: Do reviews affect the survival of a movie on screen? *International Journal of Research in Marketing, 33*(2), 357-374. doi:https://doi.org/10.1016/j.ijresmar.2015.07.003
Leonard, D. (2021, June 29, 2021). How Netflix's Lupin Pulled Off the Perfect Heist (Show), Online. *Bloomberg Businessweek*. Retrieved from https://www.bloomberg.com/news/features/2021-06-29/how-netflix-s-lupin-and-omar-sy-pulled-off-the-perfect-heist-show
Li, Q., Zeng, D. D., Xu, D. J., Liu, R., & Yao, R. (2020). Understanding and Predicting Users' Rating Behavior: A Cognitive Perspective. *INFORMS Journal on Computing*. doi:10.1287/ijoc.2019.0919
Linton, J. M., & Petrovich, J. A. (1988). The application of the consumer information acquisition approach to movie selection: An exploratory study. *Current research in film: Audiences, economics and law, 4*, 24-44.

Liu, J., & Toubia, O. (2018). A Semantic Approach for Estimating Consumer Content Preferences from Online Search Queries. *Marketing Science*, *37*(6), 930-952. doi:10.1287/mksc.2018.1112

Liu, X., Singh, P. V., & Srinivasan, K. (2016). A Structured Analysis of Unstructured Big Data by Leveraging Cloud Computing. *Marketing Science, 35*(3), 363-388. doi:10.1287/mksc.2015.0972

Lotz, A. D. (2021). Unpopularity and cultural power in the age of Netflix: New questions for cultural studies' approaches to television texts. *European Journal of Cultural Studies, 24*(4), 887-900. doi:10.1177/1367549421994578

McAlister, L. (1982). A Dynamic Attribute Satiation Model of Variety-Seeking Behavior. *Journal of Consumer Research, 9*(2), 141–150. doi:10.1086/208907

McAlister, L., & Pessemier, E. (1982). Variety Seeking Behavior: An Interdisciplinary Review. *Journal of Consumer Research, 9*(3), 311-322.

McKinney, W. (2010). *Data structures for statistical computing in python.* Paper presented at the Proceedings of the 9th Python in Science Conference

Mika, P. (2005). Ontologies Are Us: A Unified Model of Social Networks and Semantics. *Web Semantics: Science, Services and Agents on the World Wide Web, 5*, 5-15. doi:10.1016/j.websem.2006.11.002

Moon, S., Bergey, P. K., & Iacobucci, D. (2010). Dynamic Effects among Movie Ratings, Movie Revenues, and Viewer Satisfaction. *Journal of Marketing, 74*(1), 108-121. Retrieved from https://www.jstor.org/stable/20619083

Moreno, A., & Terwiesch, C. (2017). The Effects of Product Line Breadth: Evidence from the Automotive Industry. *Marketing Science, 36*(2), 254-271. doi:10.1287/mksc.2016.0979

Mukherjee, A., & Kadiyali, V. (2011). Modeling Multichannel Home Video Demand in the U.S. Motion Picture Industry. *Journal of Marketing Research, 48*(6), 985-995. Retrieved from http://www.jstor.org/stable/23142898

Mutlu, B., Veas, E., Trattner, C., & Sabol, V. (2015). *Towards a Recommender Engine for Personalized Visualizations*. Paper presented at the International Conference on User Modeling, Adaptation, and Personalization.

Nam, H., Joshi, Y. V., & Kannan, P. K. (2017). Harvesting Brand Information from Social Tags. *Journal of Marketing, 81*(4), 88-108. doi:10.1509/jm.16.0044

Nam, H., & Kannan, P. K. (2014). The Informational Value of Social Tagging Networks. *Journal of Marketing, 78*(4), 21-40. doi:10.1509/jm.12.0151

Neelamegham, R., & Jain, D. (1999). Consumer Choice Process for Experience Goods: An Econometric Model and Analysis. *Journal of Marketing Research, 36*(3), 373-386. Retrieved from https://journals-sagepub-com.tilburguniversity.idm.oclc.org/doi/10.1177/002224379903600306

Nelson, R. A., & Glotfelty, R. (2012). Movie stars and box office revenues: an empirical analysis. *Journal of Cultural Economics, 36*(2), 141-166. doi:10.1007/s10824-012-9159-5

Paris, M. (2021, April 19, 2021). People are canceling video streaming subscriptions at record rates ahead of Netflix earnings, Online. *Fortune*. Retrieved from https://fortune.com/2021/04/19/netflix-earnings-subscriber-numbers-profit-first-quarter-q1-nflx/

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*(85), 2825−2830.

Roos, J. M. T., & Shachar, R. (2014). When Kerry Met Sally: Politics and Perceptions in the Demand for Movies. *Management Science, 60*(7), 1617-1631. doi:10.1287/mnsc.2013.1834

Schneider, C., Weinmann, M., Mohr, P. N. C., & vom Brocke, J. (2021). When the Stars Shine Too Bright: The Influence of Multidimensional Ratings on Online Consumer Ratings. *Management Science, 67*(6), 3871-3898. doi:10.1287/mnsc.2020.3654

Schoenfeld, D. (1982). Partial Residuals for The Proportional Hazards Regression Model. *Biometrika, 69*(1), 239-241.

Schweidel, D. A., & Moe, W. W. (2016). Binge Watching and Advertising. *Journal of Marketing, 80*(5), 1-19.

Sevilla, J., Lu, J., Kahn, B. E., & John, D. R. (2019). Variety Seeking, Satiation, and Maximizing Enjoyment Over Time. *Journal of Consumer Psychology, 29*(1), 89-103. doi:10.1002/jcpy.1068

Sevilla, J., Zhang, J., & Kahn, B. E. (2016). Anticipation of Future Variety Reduces Satiation from Current Experiences. *Journal of Marketing Research, 53*(6), 954-968. Retrieved from http://www.jstor.org.tilburguniversity.idm.oclc.org/stable/44878474

Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika, 52*(3/4), 591-611.

Sochay, S. (1994). Predicting the Performance of Motion Pictures. *Journal of Media Economics, 7*(4), 1-20. doi:10.1207/s15327736me0704_1

Stensrud, M. J., & Hernán, M. A. (2020). Why Test for Proportional Hazards? *JAMA, 323*(14), 1401-1402.

Tafesse, W. (2020). YouTube marketing: how marketers' video optimization practices influence video views. *Internet Research, 30*(6), 1689-1707. doi:10.1108/intr-10-2019-0406

Tan, T. F., Netessine, S., & Hitt, L. (2017). Is Tom Cruise Threatened? An Empirical Study of the Impact of Product Variety on Demand Concentration. *Information Systems Research, 28*(3), 643-660. doi:10.1287/isre.2017.0712

Therneau, T. M., Grambsch, P. M., & Fleming, T. R. (1990). Martingale-Based Residuals for Survival Models. *Biometrika, 77*(1), 147–160. doi:https://doi.org/10.2307/2336057

Trabelsi, C., Jrad, A. B., & Yahia, S. B. (2010). *Bridging Folksonomies and Domain Ontologies: Getting Out Non-taxonomic Relations*. Paper presented at the 2010 IEEE International Conference on Data Mining Workshops.

Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software, 6*(60), 3021. doi:10.21105/joss.03021

Wei, S., Zheng, X., Chen, D., & Chen, C. (2016). A hybrid approach for movie recommendation via tags and ratings. *Electronic Commerce Research and Applications, 18*, 83-94. doi:10.1016/j.elerap.2016.01.003

Wei, Y. M. (2020). The Similarity Network of Motion Pictures. *Management Science, 66*(4), 1647-1671. doi:10.1287/mnsc.2018.3261

Wen, K., Li, R., Xia, J., & Gu, X. (2014). Optimizing ranking method using social annotations based on language model. *Artificial Intelligence Review, 41*(1), 81-96. Retrieved from https://www.semanticscholar.org/paper/Optimizing-ranking-method-using-social-annotations-Wen-Li/c41227791f6c063544b853f02cc3ee076e6611a3

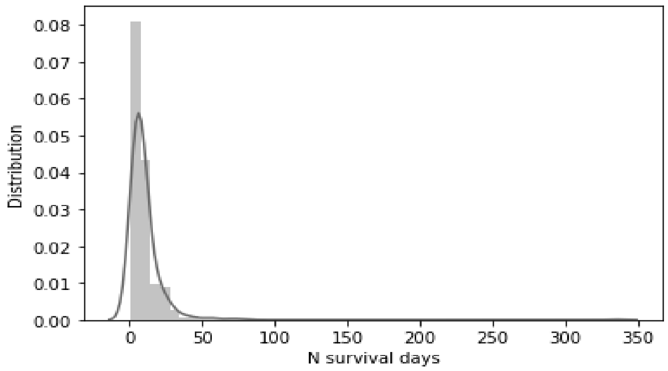
Willemsen, M. C., Graus, M. P., & Knijnenburg, B. P. (2016). Understanding the role of latent feature diversification on choice difficulty and satisfaction. *User Modeling and User-Adapted Interaction, 26*(4), 347-389. doi:10.1007/s11257-016-9178-6

Xu, C., & Chu, H. (2008). Social tagging in China and the USA: A comparative study. *Proceedings of the American Society for Information Science and Technology, 45*(1), 1-9. Retrieved from

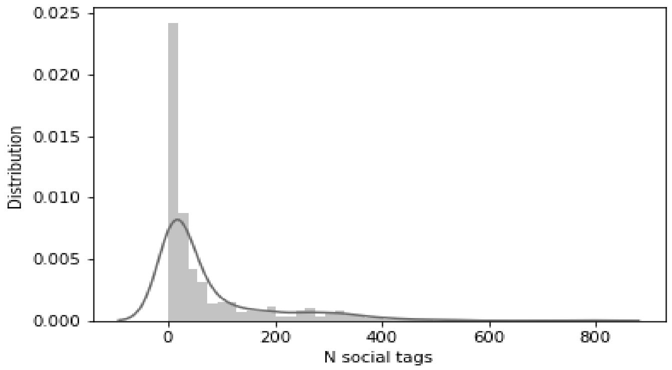
https://asistdl.onlinelibrary.wiley.com/doi/pdf/10.1002/meet.2008.1450450257?casa_token=NlBiKxIgOUoAAAAA:3c8kSbUgpMkigvGiUQClzmM9o3pb4yF34LOqqqJgg_sv2blF1Cf4G-pZQZOFLql8eR7HH_yX9GD6qPQ
Xu, X.-Y., Wang, L.-Y., Zhao, K., & Chang, F.-K. (2021). The Migration of Viewers in Gaming Streaming: The Perspective of a Push-Pull-Mooring Model. *International Journal of Human–Computer Interaction, 37*(14), 1330-1346. doi:10.1080/10447318.2021.1886480
Zhou, H., & Lange, K. (2009). Rating Movies and Rating the Raters Who Rate Them. *The American Statistician*, *63*(4), 297-307. doi:10.1198/tast.2009.08278

Appendices

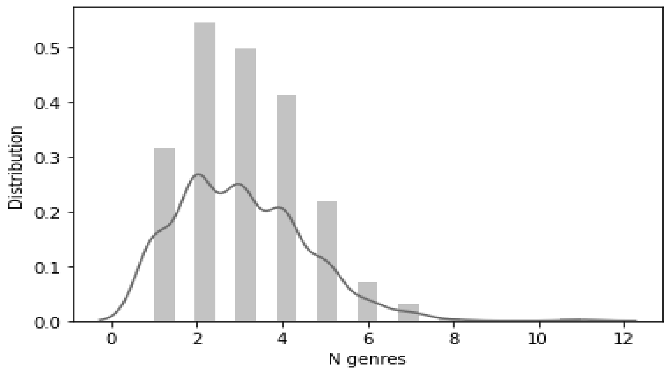
Appendix A: The distribution of survival days on the chart per video



Appendix B: The distribution of social tags per video



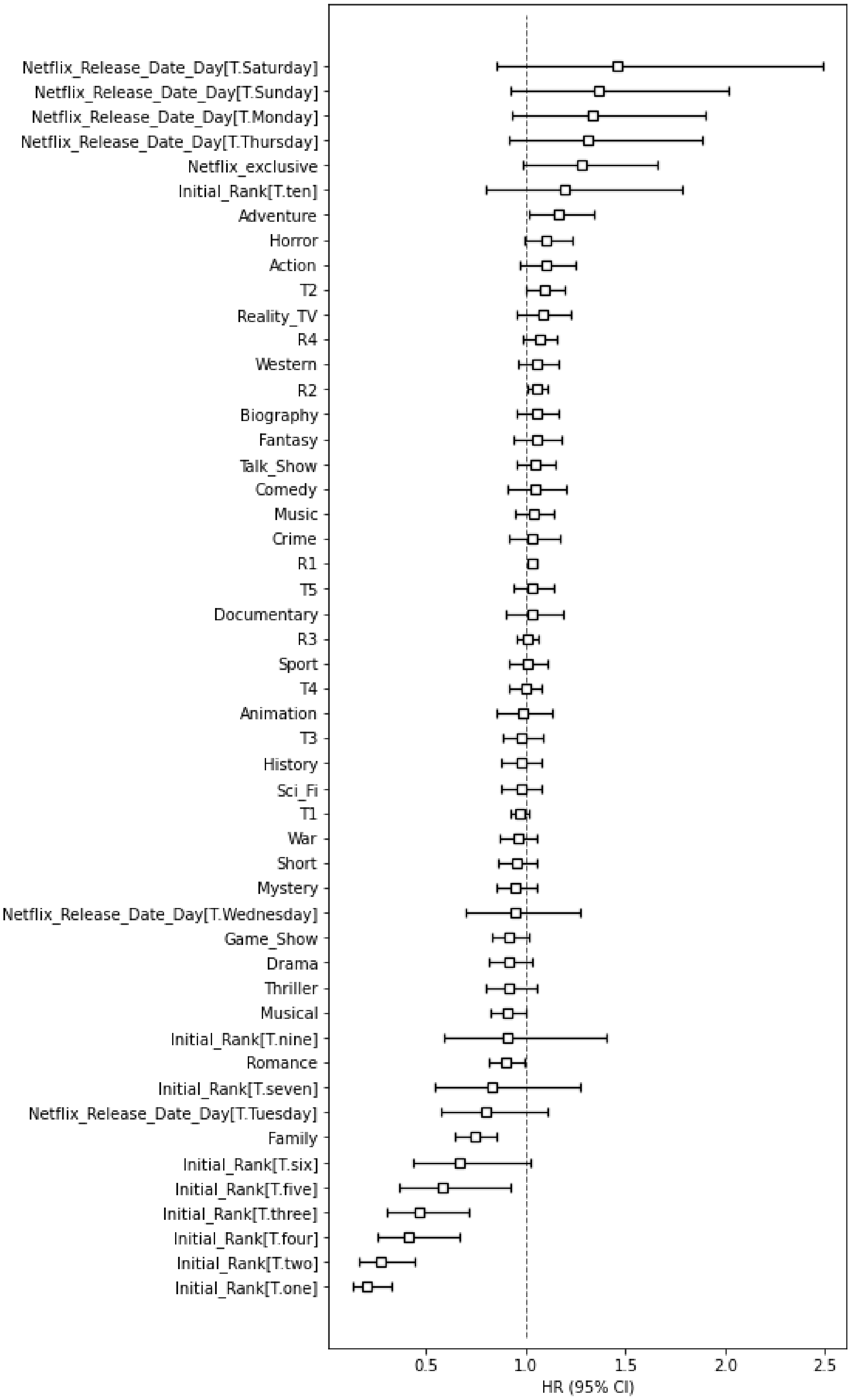
Appendix C: The distribution of genres per video



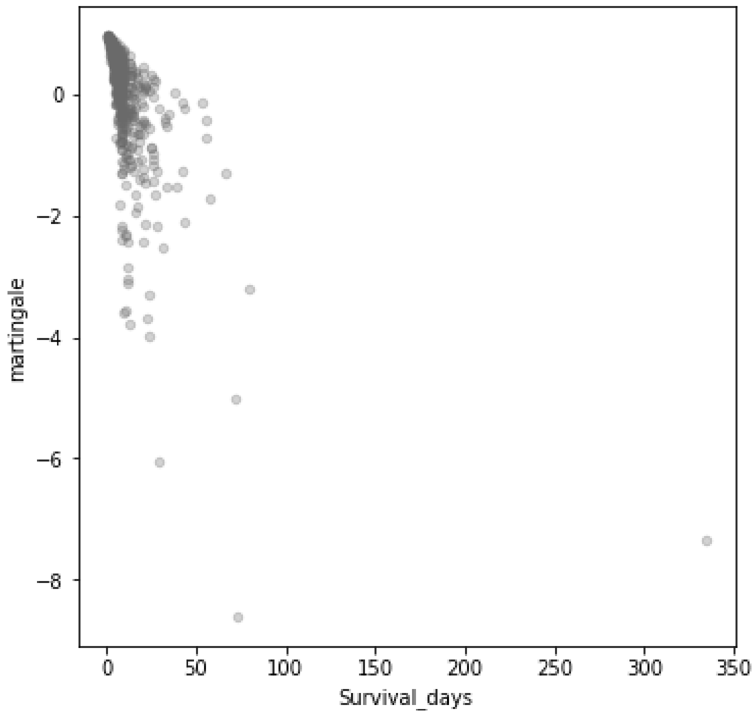
Appendix D: The number of votes and average rating by consumer segment

	Votes			Rating		
	N	**Mean**	**SD**	**N**	**Mean**	**SD**
IMDB users	38,368,730	72,805.9393	172,518.5601	38,368,730	6.542	1.138
aged under 18	31,097	63.2053	130.7347	31,097	6.133	1.773
aged 18-29	6,790,518	12,885.2334	32,141.4974	6,790,518	6.583	1.210
aged 30-44	14,630,800	27,762.4288	69,202.4759	14,630,800	6.483	1.140
aged 45 plus	3,722,378	7,063.3359	14,845.2801	3,722,378	6.420	1.090
Males	20,672,488	39,226.7325	101,506.4283	20,672,488	6.409	1.157
males aged under 18	19,186	43.5057	93.8693	19,186	6.395	1.751
males aged 18-29	4,729,229	8,973.8691	24,724.5266	4,729,229	6.491	1.238
males aged 30-44	11,312,476	21,465.7989	56,876.6658	11,312,476	6.380	1.158
males aged 45 plus	2,910,351	5,522.4877	12,191.1137	2,910,351	6.315	1.112
Females	6,166,139	11,700.4535	24,491.8357	6,166,139	6.712	1.100
females aged under 18	7,425	17.1478	26.8784	7,425	5.969	1.894
females aged 18-29	1,818,879	3,457.9449	7,571.0471	1,818,879	6.691	1.206
females aged 30-44	3,053,823	5,794.7306	12,663.7212	3,053,823	6.678	1.113
females aged 45 plus	725,495	1,376.6509	2,602.1677	725,495	6.718	1.131
top 1000 voters	92,354	175.2448	190.2656	92,354	5.687	1.091
US users	5,180,741	9,830.6281	24,807.9529	5,180,741	6.554	1.159
Non-US users	14,609,492	27,721.9962	65,904.6439	14,609,492	6.413	1.131

Appendix E: All hazard ratios within a 95% confidence interval for Model 3



Appendix F: Martingale residuals



Appendix G: Deviance Residuals