Electronic Commerce Research and Applications 53 (2022) 101145

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/elerap

Understanding music streaming services via text mining of online customer reviews

Jaemin Chung [a], Jiho Lee [b], Janghyeok Yoon [b,*]

[a] Korea Institute of Science and Technology Information, 66 Hoegi-ro, Dongdaemun-gu, Seoul 02456, Republic of Korea
[b] Department of Industrial Engineering, Konkuk University, 120 Neungdong-ro, Gwangjin-gu, Seoul 05029, Republic of Korea

ARTICLE INFO

Keywords:
Music streaming service
Customer satisfaction
Social media mining
Online app review
Text mining

ABSTRACT

With the development of information technology, the subscription economy and streaming service market have grown rapidly. In particular, music streaming services have disrupted the traditional music industry. Nevertheless, few attempts have been made to understand music streaming services in terms of overall customer satisfaction. This study analyzes social media data to investigate the determinants of customer satisfaction in music streaming services. Topic modeling and text regression were applied to online app reviews for five music streaming services. This study finds that customers comment on factors related to usage environments, price plans, and content. All environment-related factors, some pricing-related factors, and content-related factors have a significant effect on customer satisfaction. In addition, the satisfaction determinants differ for each service. This study is an early attempt to analyze music streaming services from a data-driven perspective and contributes to a comprehensive understanding of music streaming services from the customer's point of view.

1. Introduction

Recently, subscription business models and streaming services have grown rapidly owing to the progress in information technology (Danaher, 2002; Pauwels and Weiss, 2008). In particular, music streaming has disrupted the traditional music industry, which has evolved from vinyl records to MP3s (Webster, 2020). Music consumers can access vast amounts of high-quality music through the web or mobile applications, even if they no longer own music physically or digitally. In addition, music producers have been able to monetize their intellectual property by allowing the public to legally listen to their music (Hracs and Webster, 2021). On the other hand, music streaming services, such as Spotify, YouTube Music, and Amazon Music, connect consumers and producers and monetize through commissions or subscription fees. The more customers consume music through a music streaming service, the more revenue that service earns. Therefore, music streaming services are employing various strategies, such as developing cutting-edge curation systems to attract more customers (Morris and Powers, 2015).

In addition to providing satisfactory technology, market-pull innovation, also known as incremental innovation, could be another solution to gaining a competitive edge (Gerpott, 2005). From a market point of view, innovation occurs by addressing customers dissatisfaction (Brem and Voigt, 2009). Therefore, firms need to identify and address problems their customers are experiencing (Zuo et al., 2019). Prior studies on music streaming services have interviewed experts and users to analyze experiences related to the curation systems (Hagen, 2015), to investigate the impact of differences in customer preferences by country (Kim et al., 2017), or to analyze service usage experiences (Hracs and Webster, 2021). These studies evaluate music streaming services from different perspectives. Nevertheless, the overall customer satisfaction was not analyzed due to the limitations in the survey-based method and the amount of data available. In addition, a comparison of customer opinions for each music streaming service has not been made.

To overcome these limitations, this study aims to understand music streaming services by exploring the determinants of customer satisfaction via text mining of online customer reviews. With the advent of big data, many efforts have been made to identify business opportunities and their implications by analyzing online reviews or product reviews based on social media mining (Choi et al., 2020a; Liu et al., 2019b; Stamolampros et al., 2019). These systematic and analytic studies used data analytics techniques to identify customers' overall interest, satisfaction, and latent requirements that could not be obtained at the survey level (Choi et al., 2020b). In this study, online app reviews written by

* Corresponding author.
E-mail addresses: jmchung@kisti.re.kr (J. Chung), jiholee255@konkuk.ac.kr (J. Lee), janghyoon@konkuk.ac.kr (J. Yoon).

https://doi.org/10.1016/j.elerap.2022.101145
Received 5 August 2021; Received in revised form 7 March 2022; Accepted 24 March 2022
Available online 26 March 2022
1567-4223/

service users are analyzed. Many customers use music streaming services through mobile applications. In addition, application users disclose their satisfaction with the services provided publicly through online reviews (Qiao et al., 2018). Therefore, it is possible to identify common needs and attributes of customers with reference to music streaming services by analyzing the ratings and text data shown in the app reviews.

In general, there are three main streams in social media mining research: analysis of topics or customer requirements (Jeong et al., 2019; Misuraca et al., 2020; Wang et al., 2020); investigation of the determinants of customer satisfaction (Alrawadieh and Law, 2019; Xu, 2020); competitive analysis from the customer's point of view (Liu et al., 2017; Wang et al., 2018). Following these three research streams, this study raises three research questions:

- RQ1: What do customers say about music streaming services?
- RQ2: Which factors affect overall customer satisfaction with music streaming services?
- RQ3: How do customer satisfaction determinants differ for each service?

To answer these questions, this study first collected app reviews of major music streaming services. Topic modeling was then applied to the textual data of app reviews to identify latent factors considered by customers regarding the music streaming services. Then, text regression was applied to investigate the determinants that affect overall customer satisfaction. Text regression was also applied service-by-service to analyze service-specific determinants.

Main findings of this study include:

1) Customers comment on the environment-related, pricing-related, and content-related factors they experience while using music streaming services.
2) The factors related to the environment and price plans generally have negative effects on overall customer satisfaction, while the factor related to content has a positive effect.
3) Customers of music streaming services also want to have a great consumption experience for what they pay for.
4) Music streaming services should secure their music content pool of various countries and genres.
5) Reviews with a high number of thumps-up of long reviews are posted by customers with low satisfaction.

The contributions of this study are threefold. First, this study is an early attempt to analyze large amounts of customer review data for music streaming services through a data-driven method. Therefore, this study presents more systematic and objective results than those of previous survey-based studies. Second, this study provides a comprehensive understanding of music streaming services from a customer-centric perspective. The factors that music streaming users actually experience and comment on are identified and investigated. Finally, this study provides the results of a competitive analysis of music streaming services. Music streaming services will be able to use the results of this study to understand their customer for customer-driven innovation.

The remainder of this paper is organized as follows. Section 2 reviews related literature. Section 3 introduces the data and methodology used. Section 4 presents the data analysis and results, and Section 5 discusses the implications and limitations of this study. Finally, Section 6 provides the conclusion.

2. Literature review

2.1. Music streaming service

Technological developments, such as data communication and the spread of mobile devices, have brought major changes to the music industry (Borja et al., 2015). People no longer need to own music to listen to it. Instead, they simply access a music streaming service via their smartphones or computers and select a song they want to listen to (Maasø, 2018; Wlömert and Papies, 2016). Music streaming services have disrupted the music industry, which has evolved from vinyl records to MP3s, by completely changing the experience of music consumers (Webster, 2020). According to the RIAA report, the share of music streaming of the US music industry has grown steadily, reaching 83 % by 2020 (Friedlander, 2021). Various services such as Spotify, YouTube Music, and Amazon Music not only lead the current global music industry but also fiercely compete for market dominance. For example, they develop recommendation algorithms to curate playlists customers may want to listen to (Barna, 2017; Joosse and Hracs, 2015; Morris and Powers, 2015; Smits and Nikdel, 2019; Wlömert and Papies, 2016). Even recently, a lot of effort is put into developing cutting-edge recommendation system that consider the context information of customers, such as location, motion, and time of day (Prey, 2018).

Nevertheless, since the music streaming has a unique characteristic that is different from the video or live streaming, it is not easy for a music streaming service to dominate the market. Video streaming services such as Netflix and Disney + typically differentiate themselves by monopolizing their content (Raats and Evens, 2021). These services try to attract customers by protecting their content or even creating own content, such as Netflix Originals. On the other hand, live streaming services such as Twitch and YouTube Gaming attract customers through streamers who utilize the platform. However, all the leading music streaming services have nearly similar content and artist pools. Thus, music streaming services are difficult to differentiate content (Hracs and Webster, 2021). An example of this unique characteristics of the music streaming is Spotify's entry into Republic of Korea. When Spotify entered the Korean market, Kakao Entertainment, a major Korean music distribution company, refused to distribute its music to Spotify. Some view this as an effort to keep the market share of Melon, one of the leading music services of Korea as well as its own music service. However, due to criticism from K-pop fans around the world, Kakao Entertainment eventually agreed to a global licensing renewal agreement to provide content to Spotify on March 11, 2021. Consequently, customers can consume music owned by Kakao Entertainment on any music streaming service. As can be seen in this case, a streaming service can gain competitiveness through content protection, but if the service is in the music streaming industry, monopoly of content is not easy. Thus, even with the state-of-the-art recommendation system, it is difficult for a music streaming service to curate completely new music.

Thus, music streaming services now need to consider market-pull innovation that incrementally innovates by finding and solving customer dissatisfaction (Brem and Voigt, 2009; Gerpott, 2005). In that sense, many studies have been conducted to understand music streaming services by interviewing users or domain experts. Some researchers have analyzed the usage habits and practices of music streaming service users. For instance, Hagen (2015) explained the behavior of personal playlist users and Weinberger and Bouhnik (2020) analyzed the effects of streaming usage habits on personal music management, sense of ownership, and privacy concerns. Other researchers have investigated the experience of music discovery (Garcia-Gathright et al., 2018), or musical exploration (Kjus, 2016). In addition, many efforts have been made to understand music streaming services and their users from a variety of perspectives. Kim et al. (2017) estimated and compared the marginal willingness to pay for music streaming services by US and Korean consumers, Hracs and Webster (2021) analyzed the characteristics of platform competition and user experience, and Lüders (2020) investigated the effects of price, convenience, and achievement on the habitual use of streaming services.

Despite various attempts to understand customers of music streaming services, to the best of the authors' knowledge, few attempts have been made to use large-scale social media data to evaluate music streaming services. Analyzing social media data has the advantage of being able to identify latent customer needs or dissatisfaction and uncover insights that can be developed into product or service opportunities (Choi et al., 2020b; Jeong et al., 2019). Thus, this study aims to understand music streaming services from a customer's point of view by investigating customer satisfaction determinants for music streaming services through social media mining. Thus, this study will be able to identify factors that statistically affect customer satisfaction with music streaming services as well as customers' overall interest or hidden requirements. Consequently, this study may provide clues to the market-pull innovation of music streaming services.

2.2. Social media mining

Social media is an important communication channel for firms to understand their customers (Sarker et al., 2016) and serves to provide the possibility of value co-creation through interaction with customers (Silverstein et al., 2013). Studies that analyze large-scale social media data for business intelligence research generally aim to identify and analyze customer requirements for firms, products, or services (Choi et al., 2020b). In this subsection, the previous studies on social media mining are reviewed in detail by classifying them into three categories according to the research purpose.

First, some studies analyzed topics or requirements for specific targets. These studies have provided insights, based on text mining, into what customers say on social media. For example, Jeong et al. (2019) identified product opportunities using topic modeling and sentiment analysis of online product reviews. Misuraca et al. (2020) proposed a network-based method for managing customer requests sent to a firm's customer support service. Choi et al. (2020a) proposed an approach to identifying time-evolving product opportunities by applying an aging theory-based event detection and tracking algorithm and opportunity algorithm. In addition, Wang et al. (2020) used predictive analysis to create a product design specification, through customer requirements, based on deep-learning techniques.

Second, some studies have investigated the determinants of customer satisfaction using a large amount of customer data. These studies analyzed customer review data to investigate the factors that significantly affect customer satisfaction. Customer demographic data, ratings by category, and even topics extracted from the review text were used as independent variables. The overall ratings of reviews are usually utilized as dependent variables. The studies are not limited to any specific industry but are extensive, including hotels and accommodation (Alrawadieh and Law, 2019; Xu, 2020), restaurants (Park et al., 2020), and on-demand food delivery services (Xu, 2021). These studies provide in-depth implications for target domains, based on significant factors.

Finally, some prior studies conducted a competitive analysis from the customer's point of view. These studies selected a target (e.g., firm, product, or industry) and comparatively analyzed the target with competitors. For example, Liu et al. (2019b) and Wang et al. (2018) presented methods to assess the competitive advantages and disadvantages of a target product over its competitors through online reviews. Liu et al. (2017) prioritized competing products by applying sentiment analysis and intuitionistic fuzzy set theory to support customers' product purchase decisions. Jin et al. (2016) proposed a method for selecting competitively comparable sentences, along with product features, in online reviews of competing products.

3. Data and methodology

This study follows the analytical procedure shown in Fig. 1. RQ1 was addressed by applying a topic modeling technique to the collected online app reviews. RQ2 and RQ3 were answered using text regression results. The following subsections describe the data and methodology used in this study.

3.1. Data

This study collects app review data from Google Play, a digital application distribution service developed and operated by Google (McIlroy et al., 2015). Since mobile applications are the only channel through which mobile customers can use music streaming services, customers tend to post app reviews about what they are satisfied with of dissatisfied with while using the service. Thus, in this study, it was determined that it would be most suitable to analyze app reviews to provide an understanding of music streaming service customers. Among the various music streaming services, five major streaming services in the US are selected: Amazon Music, Deezer, Spotify, Tidal, and YouTube Music. This study crawled review data of the five target music streaming services from the US Google Play. The collected data included features such as overall rating, review comments, time of review, replies, and thumb-up counts.

In this study, the text data of app reviews were mainly used. Thus, when data collection was completed, key phrases were extracted from the review comments to structure each review. Key phrases were obtained through advanced text-mining tools. However, some phrases may not be related to music streaming services (e.g., 'hey guys', 'every day'), or may be too general (e.g., 'google play', 'amazon music', 'music streaming'). Previous studies using topic modeling excluded stopwords such as phrases that describe the object of analysis (Stamolampros et al., 2019), common phrases that appear too often (Choi et al., 2018), emotional phrases that express emotions (He et al., 2020), or phrases that do not contain any meaning (Jeong et al., 2019). These phrases can act as noise in the analysis. For example, common phrases that appear in too many documents can hinder the clear identification of topics. In addition, since phrases expressing emotions can be divided into emotion-related topics, results that do not meet the purpose of this study may be derived. Thus, in this study, key phrases and reviews were screened by removing stopwords as in previous studies. The sentiment polarity of each review was then calculated. Finally, each app review is structured with the collected numerical features, key phrases and their frequencies, and sentiment polarity.

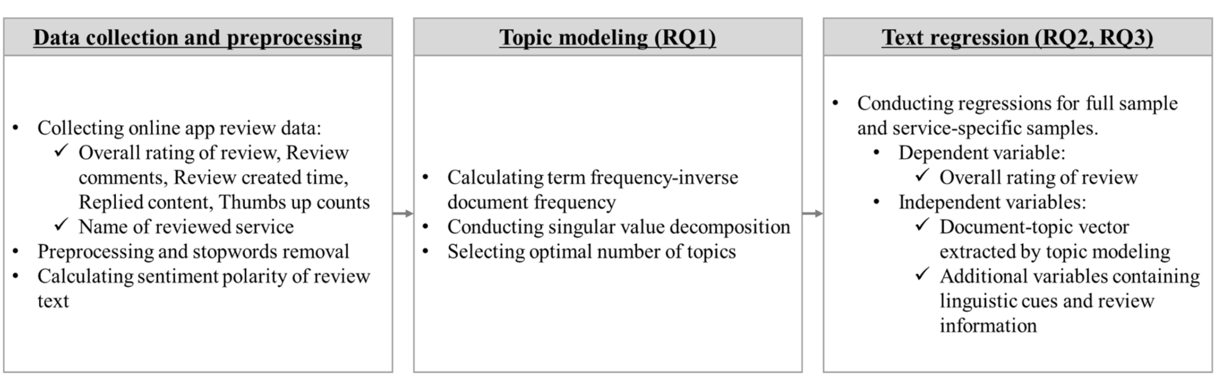

Fig. 1. Overview of data analytics procedure.

3.2. Topic modeling

To answer the research questions, topic modeling is used as the core technique of this study. Topic modeling is a technique for identifying latent topics hidden in large corpora (Hu et al., 2014; Vayansky and Kumar, 2020). It is used in conjunction with a variety of data sources, such as patents (Choi et al., 2018; Erzurumlu and Pachamanova, 2020; Venugopalan and Rai, 2015; Yun et al., 2021) and news (Liu et al., 2019a; Pinto et al., 2019), and is often used to identify customer requirements, especially in data from social media (Hu et al., 2019; Stamolampros et al., 2019). In addition, topic modeling is used to identify latent factors affecting customer satisfaction through text regression (Park et al., 2020; Xu, 2020, 2021) or to identify opportunities for products or services (Choi et al., 2020a; Jeong et al., 2019). This study utilizes latent semantic analysis (LSA) to identify latent factors hidden in app reviews.

To perform the LSA, a document-term matrix should first be prepared. In this study, the matrix is constructed based on the term frequency-inverse document frequency (TF-IDF) weights, which is widely used in the field of information retrieval (Aizawa, 2003). This weight assigns high values to terms that occur frequently within a document but that rarely appear within the entire document set (Sidorova et al., 2008). Therefore, terms with high TF-IDF weights are considered important terms within a document (Choi et al., 2018). The TF-IDF weights can be calculated using the formula as shown in Eq. (1), where w_{ab} is the TF-IDF weight of term a in document b, tf_{ab} is the occurrence frequency of term a in document b, df_a is the number of documents containing term a, and N is the number of documents in the corpus.

$$w_{ab} = tf_{ab} \times \log(N/df_a) \tag{1}$$

Then, we decompose the document-term matrix X (Landauer et al., 1998). Through singular vector decomposition (SVD), matrix X can be decomposed into the product of three other matrices (Eq. (2)) (Park et al., 2015). Specifically, matrices U and V are orthogonal matrices of the left and right singular vectors, respectively. Matrix Σ is a diagonal matrix of singular values, which is the square root of the eigenvalues from the matrix U or V. For the LSA, truncated SVD is used to select the first k columns, with large singular values, from the decomposed matrices. In general, k is selected as the point at which the singular values are stabilized (Kulkarni et al., 2014) or empirically selected by an analyst based on the performance of the analysis results (Deerwester et al., 1990; Park et al., 2015). Finally, the LSA is completed by obtaining a document-topic matrix and a topic-term matrix based on k latent factors through truncated SVD.

$$X = U \bullet \Sigma \bullet V^T \tag{2}$$

3.3. Text regression

After the latent factors of the music streaming services, commented on by the customers, were identified, regression analysis was used to investigate how the latent factors affect overall customer satisfaction. Multiple linear regression was used in this study (Park et al., 2020). As shown in Eq. (3), the dependent variable of the regression is the overall rating of each review, and there are k +5 independent variables. Overall ratings are often used as indicators of overall customer satisfaction (Park et al., 2020; Xu, 2020, 2021). LF_i is the value of the i th latent factor of each review derived from the LSA. The value in the document-topic matrix indicates how much a review mentions a specific latent factor. Thus, a high LF_i indicates that a customer is commenting more on the i th latent factor. This study aims to examine the significance of the extent of mentions for each latent factor on overall satisfaction.

$$OverallRating = \beta_0 + \beta_1 LF_1 + \beta_2 LF_2 + \cdots + \beta_k LF_k + \beta_{k+1} ThumbsUps + \beta_{k+2} ReplyOrNot + \beta_{k+3} Sentiment + \beta_{k+4} ReviewLength \tag{3}$$

In addition to the latent factors, this study adds four more independent variables, including some information from the reviews. In social media, a “thumbs-up” or “like” usually means that others like or agree with specific content (Lowe-Calverley and Grieve, 2018). Similarly, if someone clicks “thumb-up” for a particular app review, it is likely that the person has the same feeling or shares the experience. This study counts the number of thumbs-ups each review receives and analyzes the relationship between *ThumbsUps* and overall ratings. Replies to app reviews indicate that the application provider responds to customers. Providers may leave a thank you note or an apology letter in response to a customer review. In this study, the ratings of app reviews both with replies and without replies were analyzed. The *ReplyOrNot* variable is coded to have a value of 1 if there is a reply, and −1 if not. In addition, according to previous studies, linguistic cues in online reviews may affect overall customer satisfaction (Biswas et al., 2020; Deng et al., 2021; Siering et al., 2016; Xu, 2021). Thus, referring to a previous study (Xu, 2021), this study adds linguistic variables: *Sentiment*, *ReviewLength*. *Sentiment* is the sentiment polarity of a review text and has a range of –1 and 1. If the *Sentiment* is close to 1, the review text is positive, and if it is close to –1, the review is negative. On the other hand, a value of 0 indicates that the review has a neutral sentiment. *ReviewLength* is measured by the number of characters in the review text.

4. Data analysis and results

4.1. Data analysis

App reviews posted between January 1, 2016 and December 31, 2020 for the five target services (Amazon Music, Deezer, Spotify, Tidal and YouTube Music) were crawled from Google Play. Because there were too many reviews in various languages, only the reviews written in English were selected. Among various natural language processing tools, spaCy (https://spacy.io/) was used to extract noun phrases from the review text. This study applied strict preprocessing to remove data that could act as noise in the analysis. Thus, in this study, noun phrases that met the following five conditions were removed from the extracted phrase set: 1) phrases containing the name of the target services or other services (e.g., ‘spotify’, ‘youtube’, ‘pandora’, ‘apple’); 2) phrases with a document frequency less than five or more than 2000; 3) phrases with a length of four characters or less; 4) phrases containing adjectives expressing feelings (e.g., ‘nice’, ‘awesome’, ‘terrible’); and 5) phrases without specific meaning (e.g., ‘one thing’, ‘something else’). As a result, this study used 1,091 noun phrases and 22,819 reviews, including preprocessed noun phrases. Finally, the sentiment polarities of the selected app reviews were obtained through VADER (Valence Aware Dictionary for sEntiment Reasoning), included in the Python Natural Language ToolKit.

Fig. 2 shows graphs visualizing the number of reviews, after preprocessing, by service and by year. According to the graph on the left, almost half of all reviews are from Spotify. YouTube Music and Amazon Music ranked next, in that order. On the other hand, Deezer and Tidal have fewer reviews. According to the graph on the right, the number of reviews by year has been steadily increasing since 2018. As the number

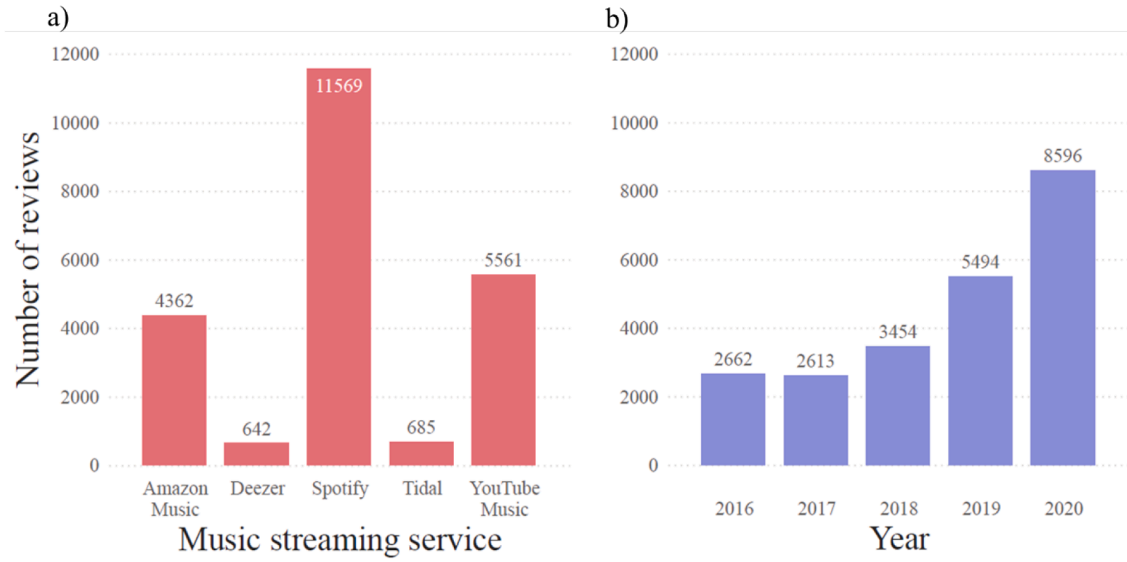

Fig. 2. a) Number of reviews for different music streaming provider and b) Changes in volume of reviews with time.

Table 1
Latent factors for music streaming services.

Latent factor #	Label of latent factor	Major phrases	Subject of comments
LF 1	Playing environment	car mode, landscape mode, landscape view, car view, sleep mode, regular mode	Environment
LF 2	Internet Connection	offline mode, online mode, downloaded album, airplane mode, wi-fi	Environment
LF 3	Nonpremium user	free play, free skip, free version, free user, free mode, non premium user, free app	Price plan
LF 4	Android device memory	sd card, phone memory, offline use, android marshmallow, device memory, android device	Environment
LF 5	Premium membership	free version, paid version, premium version, premium user, premium service	Price plan
LF 6	Local file usage	free version, paid version, premium version, premium user, premium service	Environment
LF 7	Amazon Prime membership	prime member, prime membership, prime video, non prime member, prime user	Price plan
LF 8	Genre & contents	k pop, hip hop, j pop, j rock, k drama, classic rock, anime opening	Content
LF 9	Features by price plan	free user, premium user, background play, shuffle mode, shuffle play, audio mode	Price plan

Table 2
Brief statistics of the variables.

Variables	Mean	Std	Min	Max
Overall rating	2.857	1.642	1.000	5.000
Thumbs-up	7.373	60.666	0.000	2863.000
Reply or not	−0.430	0.903	−1.000	1.000
Sentiment	0.339	0.575	−0.989	0.999
Review length	283.506	208.751	16.000	4095.000

of users increases and the proportion of music streaming increases in the music industry, the number of reviews for music streaming services also seemingly increases.

LSA was performed based on the TF-IDF weights for the extracted noun phrases. In this study, the optimal *k* was selected as 9, at which the trend of the singular value was stabilized and interpretable results were derived. Using the topic-term matrix, which is an output of the LSA, latent factors were labeled, based on the major noun phrases of each factor. The labels and major noun phrases for the nine latent factors are summarized in Table 1. Through the identified latent factors, it was found that music streaming service customers comment on the usage environment, price plans, and content.

The dataset for text regression was constructed based on the overall ratings of app reviews, nine latent factors, and four additional independent variables. Table 2 shows the brief statistics of each variable, excluding latent factors. Before performing the text regression, the correlation between the variables was checked (Table 3). In addition, all of the variance inflation factor (VIF) values were not higher than 2; thus, it was confirmed that there was no multicollinearity.

The regression analysis results and the standardized coefficients are summarized in Table 4. Seven different regression analyses were conducted, and the p-values of all regression models were significant. The first model (Model 1) was the basic model, which was used to determine the influence of the nine latent factors in Eq. (7). The second model (Model 2), including four more independent variables related to the additional information and linguistic characteristics of reviews, was tested based on Eq. (8). In Models 1 and 2, all of the review data screened in this study were analyzed. According to the adjusted R-squared, the model with the additional variables explains customer satisfaction better (increase from 0.011 to 0.253). The third to seventh models analyzed the reviews of each streaming service based on Eq. (8): Models 3, 4, 5, 6, and 7 targeted Amazon Music, Deezer, Spotify, Tidal, and YouTube Music, respectively. These models were constructed to analyze how the factors influencing customer satisfaction differ between streaming services. Excluding Model 7, which excludes the *ReplyOrNot* variable, the other models showed slightly higher explanatory power than Model 2. When using the full sample, the service-specific characteristics were squashed. On the other hand, if regressions were performed by separating the data for each service, the characteristics of each service were prominent and appeared to have high explanatory power.

Table 3
Correlation matrix of independent variables.

	LF 1	LF 2	LF 3	LF 4	LF 5	LF 6	LF 7	LF 8	LF 9	Thumbs-up	Reply or not	Sentiment	Review length	Number of phrases
LF 1	1.00													
LF 2	−0.01	1.00												
LF 3	0.00	0.00	1.00											
LF 4	0.00	−0.02	0.00	1.00										
LF 5	−0.01	−0.03	0.00	−0.02	1.00									
LF 6	0.00	−0.02	0.00	−0.01	−0.02	1.00								
LF 7	−0.01	−0.03	0.00	−0.02	−0.03	−0.02	1.00							
LF 8	0.00	−0.01	0.00	0.00	−0.01	0.00	−0.01	1.00						
LF 9	0.00	−0.03	0.00	−0.02	−0.02	−0.02	−0.03	0.00	1.00					
Thumbs-up	0.00	0.01	0.00	−0.01	0.01	0.02	−0.01	0.00	0.00	1.00				
Reply or not	−0.01	−0.04	−0.01	−0.03	0.01	0.07	−0.10	−0.01	0.00	0.06	1.00			
Sentiment	0.00	−0.09	0.01	−0.01	0.07	−0.01	0.04	0.02	0.04	0.02	−0.11	1.00		
Review length	0.01	0.06	0.05	0.02	−0.02	0.04	−0.03	−0.01	0.02	0.11	0.14	0.02	1.00	
Number of phrases	0.01	0.10	0.00	0.04	0.09	0.02	0.04	0.00	0.09	0.01	−0.06	−0.03	0.08	1.00

Table 4
Text regression results.

	Dependent variable: Overall rating						
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
(LF 1) Playing environment	–0.014**	–0.010*	–0.012	–0.026	–0.020**	–0.079*	0.016
(LF 2) Internet connection	–0.051***	0.002	–0.021*	0.010	–0.024***	0.078*	0.014
(LF 3) Nonpremium user	–0.003	0.003	0.013	–0.112***	0.001	–0.097	–0.023*
(LF 4) Android device memory	–0.022***	–0.014**	–0.046***	0.007	–0.014*	–0.044	–0.010
(LF 5) Premium membership	–0.011	–0.044***	–0.004	–0.087***	–0.043***	–0.128*	–0.034***
(LF 6) Local file usage	–0.058***	–0.045***	0.015	0.016	–0.017**	0.036	–0.061***
(LF 7) Amazon prime membership	0.051***	0.024***	–0.004	0.040	0.014	–0.125	0.008
(LF 8) Genre & content	0.017**	0.008	0.016	–0.036	–0.006	0.152	0.038***
(LF 9) Features by price plan	–0.039***	–0.049***	–0.015	–0.092	–0.064***	0.201*	–0.039***
Thumbs–up		–0.018***	–0.020	0.046	–0.017**	0.004	0.014
Reply or not		–0.059***	–0.031**	–0.375***	–0.123***	–0.183***	N/A[a]
Sentiment		0.436***	0.505***	0.353***	0.443***	0.429***	0.338***
Review length		–0.228***	–0.249***	–0.145***	–0.218***	–0.233***	–0.133***
Adjusted R-squared	0.011	0.253	0.325	0.388	0.276	0.301	0.131
F-statistics	28.664	597.021	162.584	32.283	339.837	23.687	70.681
p-value	<0.001	0.000	0.000	<0.001	0.000	<0.001	<0.001
Log-likelihood	–43,562	–40,348	–7,263	–1,048	–20,108	–1,139	–8,986
AIC	87,140	80,720	14,550	2,125	40,240	2,306	18,000
Observations	22,819	22,819	4,362	642	11,569	685	5,561
Sample Segment	Full sample	Full sample	Amazon Music only	Deezer only	Spotify only	Tidal only	YouTube Music only

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, [a] N/A shows that this factor is excluded from the analysis because it is a constant in the dataset.

$$OverallRating = \beta_0 + \beta_1 PlayingEnvironment + \beta_2 InternetConnection + \beta_3 NonPremiumUser + \beta_4 AndroidDeviceMemory + \beta_5 PremiumMembership + \beta_6 LocalFileUsage + \beta_7 AmazonPrimeMembership + \beta_8 GenreContents + \beta_9 FeaturesByPricePlan \tag{7}$$

$$OverallRating = \beta_0 + \beta_1 PlayingEnvironment + \beta_2 InternetConnection + \beta_3 NonPremiumUser + \beta_4 AndroidDeviceMemory + \beta_5 PremiumMembership + \beta_6 LocalFileUsage + \beta_7 AmazonPrimeMembership + \beta_8 GenreContents + \beta_9 FeaturesByPricePlan + \beta_{10} ThumbsUps + \beta_{11} ReplyOrNot + \beta_{12} Sentiment + \beta_{13} ReviewLength \tag{8}$$

4.2. Results

4.2.1. RQ1: What do customers say about music streaming services?

To answer RQ1, the detailed results of topic modeling (Table 1) are discussed in this subsection. The latent factors for music streaming services fall into three broad categories: LF 1, 2, 4, and 6 are factors related to the environment using the services, LF 3, 5, 7, and 9 are factors related to price plans, also called subscription type, and LF 8 is a factor related to the content.

As factors related to the usage environment, "Playing environment" (LF 1) describes how customers use the music streaming services. For example, when a car and an Android device are connected, a driver can use some applications such as music streaming service or navigation stored on the mobile phone through the Android Auto function. Customers share various experiences of using music streaming services in their car through the Android Auto. In addition, customers comment on functions related to sleep mode as well as the landscape or portrait mode. "Internet connection" (LF 2) is a factor that refers to cellular and Wi-Fi environments. In general, the use of music streaming services on mobile devices presupposes a good Internet connection. Nevertheless, the services provide functions that allow customers to enjoy music, even in environments with poor Internet access, such as in airplanes. Customers can store music so that they can listen to it anytime and anywhere without any data consumption or internet connection. Customers comment on app reviews on their experiences of using the services with or without an Internet connection. "Android device memory" (LF 4) and "Local file usage" (LF 6) are also factors related to the usage environment of the services. The music streaming services analyzed in this study are "on-demand" services that allow customers to listen to the music they want at any time without owing the music (Marshall, 2015). Nevertheless, some Android users often want to play music files or download music on their devices through these on-demand services. These factors relate to the customers' experience with utilizing the files stored on the device.

Customers of music streaming services also mention price plans. The five music streaming services, analyzed in this study, have a variety of price plans that offer different service features depending on whether the customer pays or not (Appendix). "Premium membership" (LF 5) and "Amazon prime membership" (LF 7) refer to user experience associated with the paid plans. By paying for a music streaming service, customers can enjoy almost all the features of a music streaming service through price plans such as "Premium" or "Prime". These factors refer to the experiences of customers who have paid for the services. On the other hand, "Nonpremium user" (LF 3) is related to ad-supported plans. In fact, some services offer ad-supported plans that allow customers to use services even without paying. These free plans are generally limited by commercial advertisements interspersed with music, not being available offline, or, for some services, offering only low-quality music. This factor refers to the customer experience for these free services. As the last factor related to the price plan, "Features by price plan" (LF 9) is related to the features of mobile applications provided by the services. As mentioned earlier, music streaming services offer different service features to their customers depending on their plan. For example, some plans offer higher quality music than "Premium" or "Prime" plans. As such, customers compare the features of different plans of music streaming services.

"Genre & content" (LF 8), which is not categorized with any other factors, is a factor related to the music content provided by the service. Like other streaming services such as video or live, customers want to be able to enjoy a variety of content from the streaming service they choose. As mentioned in previous studies (Hracs and Webster, 2021; Raats and Evens, 2021), the importance of the content was also derived from the result of this study. In addition, customers often express their discomfort when the music of a genre they do not enjoy is recommended.

4.2.2. RQ2: Which factors affect overall customer satisfaction with music streaming services?

RQ2 is addressed by the results of the regression analysis of Models 1 and 2 (Table 4). The results of the regression analysis can be useful in determining whether the mention of a particular factor has a positive or negative effect on overall satisfaction. For example, if the standardized coefficient of a factor is negative and the factor is statistically significant, it means that the more mentions of that factor in a review the lower the star rating.

All four factors (LF 1, 2, 4, and 6) related to the usage environment were found to have a significant effect on customer satisfaction in at least one regression analysis. The mention of these factors negatively affects overall customer satisfaction. Customers left many unsatisfactory reviews, especially regarding the in-vehicle mode ("…car mode with the larger button for safer driving…," "…App has a couple of small issues, especially when used in car…"), landscape mode ("…the provided widget just doesn't work in landscape mode, please fix it"), on/offline mode ("Some of the offline tracks skip or just play weird…," "…It is unresponsive most of the time when in online mode and I have to force close the app in order to switch songs…"), and the use of SD card or local files ("…Tried many sd cards, all returned with same error…", "When using local files, the app REQUIRES that I turn on my screen or even at times log into my phone in order to start my music."). As mentioned in a previous study (Qiao et al., 2018), customers of music streaming services also specifically described errors and inconveniences that occurred in their environment and asked that they be fixed.

Among the factors related to the price plan, some factors (LFs 5, 7, and 9) have a significant effect on customer satisfaction in at least one regression. On the other hand, the "Nonpremium user" (LF 3) factor was found to have no significant effect on customer satisfaction. According to previous studies (Homburg et al., 2005; Xu, 2021), the more money customers pay, the higher the quality of service they expect. In addition, service expectations significantly affect customer satisfaction (Nam et al., 2020). In this study, the same results as those of previous studies were derived. The premium-related factors have a significant effect on customer satisfaction, but the nonpremium-related factor has no significant effect on customer satisfaction. Customers expect music streaming services to be worth the money they pay, and they evaluate service quality—the gap between the expected value and the perceived value (Grönroos, 1984; Zuo et al., 2019). If the service quality is poor, customers post negative reviews about their unsatisfactory experience ("…but now most options are reserved for premium users, which makes sense Spotify is a business after all…," "I'm poor. I can't afford premium…").

The factor related to content significantly and positively affects customer satisfaction. In other words, reviews mention a lot of content or genre of music streaming services that have high ratings. Customers admire a variety of genres in the services offered ("A really wonderful app where there is a wide variety of music, bachata, rap, hip-hop, rock, k-pop, pop en español and in english, etc…," "…now I can listen to every Punjabi songs, Hindi songs, and English songs and of course K-pop songs."). In addition, many customers comment on positive reviews about the K-pop content of the services ("Pretty cool and really good for listening to K-pop…," "Love this app. Because I love K-Pop and k- drama, I love the OST of the dramas. In this app all songs are available…"). This result suggests that music streaming services, such as streaming services in other industries such as Netflix and Twitch, can gain a competitive advantage through content diversity (Raats and Evens, 2021).

All four additional independent variables had significant effects on overall customer satisfaction. Reviews with many thumbs ups typically have low ratings. This suggests that music streaming customers agree with reviews stating complaints similar to theirs. In addition, replies negatively affect customer satisfaction. In general, replies to app reviews are posted by customer service. They often post replies to apologize for negative reviews and promise some improvements. The results of sentiment polarity and the length of a review are the same as in a previous study (Deng et al., 2021; Xu, 2021): sentiment polarity has a positive effect, and the length of a review has a negative effect on overall satisfaction. Customers generally post positive (or negative) comments when they are satisfied (or dissatisfied) with music streaming services. In addition, dissatisfied customers use more words to express their detailed grievances.

4.2.3. RQ3: How do the customer satisfaction determinants differ from each service?

The answer to RQ3 is described by analyzing the results of Models 3, 4, 5, 6, and 7 (Table 4). These regression models analyze Amazon Music, Deezer, Spotify, Tidal, and YouTube Music.

For Amazon Music, only some of the factors related to the environment were found to have significant negative impacts. Some low-rated reviews explain that downloaded songs disappear in offline mode. In addition, there are a number of low-rated reviews stating that files downloaded to SD cards cannot be executed. In summary, customers want Amazon Music to solve errors that occur when playing music in an offline environment or the SD card environment. Unlike other services, only Amazon Music does not have a statistically significant effect on customer satisfaction for premium membership-related factors. This is because Amazon Music does not use any price plan with a name that includes the word "Premium".

On the other hand, for Deezer, only some of the factors related to price plans were found to have significant negative impacts on customer satisfaction. Interestingly, the nonpremium-related factor left by Deezer's customers negatively affects overall satisfaction significantly. Unlike other services, Deezer's customers mention a wide variety of complaints about using the ad-supported plan. For example, there are reviews that say there are too many ads, music is too difficult to find, or the ad-supported plan has too few features. Therefore, Deezer should review the policy on nonpremium users and try to provide a better free experience.

Spotify has the most factors influencing on customer satisfaction. In particular, it was found that all factors related to the usage environment had negative effects. There are reviews expressing dissatisfaction with the disappearance of the car mode around the 8.5 version of Spotify, and there are reviews mentioning bugs in which songs are not sorted based on the spelling order, or the login is canceled in the offline mode. In addition, there are reviews expressing inconvenience in the interoperability of SD cards and local files. Factors related to the price plan also have a significant negative effect. As interpreted in the previous subsection, there are negative reviews where complaints are made on the high price of premium plans or the difference between premium and nonpremium features. Therefore, Spotify needs to address the factors that negatively affect customer satisfaction to protect its wide customer pool.

Tidal uses master quality authenticated (MQA) to provide high-quality music streaming to its customers who subscribe to a HiFi plan. Although some other music streaming services also offer high-quality music with specific price plans, Tidal is known for being most specialized in listening to high-quality music. Thus, despite the expensive subscription fee of the HiFi plan, Tidal has many customers who want to enjoy lossless music. In addition, customers already know that lossless music is large and requires extra memory. Therefore, it can be interpreted that customers tend not to express issues related to the SD card or device memory that occur when they enjoy lossless music because they are already aware of the inconveniences related to memory. In addition, among the factors influencing customer satisfaction with Tidal, the factors related to Internet connection and features by price plan are characteristic. Across all services, Tidal is the only service in which these two factors have positive effects on customer satisfaction. In fact, customers leave positive comments about these factors, along with mentions of excellent music quality. As a result, Tidal will have to improve the negative factors without losing the competitive advantage of high sound quality.

YouTube Music is the only service where the content-related factor has a significant effect on customer satisfaction. As in the preceding subsection, mentions about content in reviews of YouTube Music have a positive effect. YouTube Music shares YouTube's vast pool of video data. Therefore, YouTube Music has the advantage of offering music that has not been released as an official album among videos uploaded to YouTube. Even in actual reviews, there are many positive comments that customers can listen to all the music they are looking for. Other negative influencing factors are similar to those of other services. In summary, YouTube Music should be improved to provide a better customer experience while emphasizing the advantages of having a variety of content.

In all models, the statistical effects of reply, sentiment polarity, and review length were identical to those obtained in Model 2. Interestingly, the thumbs-up count has a significant negative effect only on Spotify, and no significant effect on the other services. It can be inferred that customers using services other than Spotify share sentiments with both satisfactory and unsatisfactory experiences.

5. Discussion

5.1. Academic contributions

To the best of the authors' knowledge, this study is an early attempt to analyze large amounts of data to understand music streaming services from a customer's point of view. Previous studies have attempted to analyze the characteristics of music streaming services through survey-based research. However, due to the limitations of the survey, the results may have been biased or the scope of the study limited. This study contributes to analyzing the user-generated content of music streaming services by introducing quantitative and systematic data analysis techniques.

Second, this study provides a comprehensive understanding of music streaming services and their customers. Through the text mining technique, it was found that customers comment on the environment-related, pricing-related, and content-related factors experienced while using music streaming services. This study also identified factors significantly affecting customer satisfaction and how significant factors differed across services. Therefore, this study extended the perspective of the music streaming industry by investigating the review writing behavior of customers of music streaming services.

Third, this study provided analysis results on the relationship between app review characteristics and customer satisfaction. Four variables, extracted from app reviews, were used as the independent variables. Some of these variables had already been used in previous studies but this study was the first to analyze app review data. Therefore, this study confirms the characteristics of social media data in app reviews and contributes to expanding the understanding of social media data.

Finally, this study enhances the understanding of the relationship between the price of products or services and customer expectations and satisfaction. Customers who pay more expect more satisfactory services. When the service is poor, customers with high expectations are more dissatisfied than customers with low expectations. The nonpremium-related factor usually did not significantly affect customer satisfaction, whereas other premium-related factors had significant effects on customer satisfaction. Customers who pay for music streaming express their dissatisfaction more when they do not receive services that meet their expectations. In other words, like any customer in any other industry or service, customers of music streaming services also want a good experience, commensurate to what they paid.

5.2. Managerial implications

This study may reveal insights into the customer-centric incremental innovation of music streaming services. This study determined how the latent factors mentioned by customers affect satisfaction and how determinants differ from service to service. The factors related to the environment of the customers when using the services and the factors related to the price plans usually have negative impacts on overall customer satisfaction. On the other hand, the factor related to the music content has a positive impact. In addition, this study provides an interpretation of the regression analysis results for each service. Music streaming service managers will be able to plan service improvements based on the results of this study. For example, managers may focus on factors with negative coefficients. It may also be possible to offset the shortcomings by highlighting the unique advantages of the services, such as the wide content pool of YouTube Music or the high-quality music of Tidal.

In the streaming industry, firms gain competitive advantages through their content. Exclusive content attracts customers and is the driving force behind the growth of a service (Raats and Evens, 2021). It is not just about video streaming services or live streaming services. This study concluded that the influence of content is also present in music streaming services. Customers generally leave positive ratings when they comment on the various genres and content of the services in their reviews. According to a previous study, music streaming services have difficulty gaining competitive advantages with reference to content (Hracs and Webster, 2021). Nevertheless, service managers should expand their content pool by obtaining licensing agreements for content in various countries and genres.

The results of the regression analysis in this study can reduce the qualitative efforts of managers. Because too many reviews are posted, managers will not be able to see all the reviews. This study found that reviews with replies or relatively long reviews tended to be low rated and unsatisfactory. Therefore, managers can be able to target these reviews. In other words, the results of this study could set criteria for sampling key reviews that managers should check.

5.3. Limitations and future research

This study has several limitations, as well as directions for future research. First, this study selected five major music streaming services for analysis and used various data analysis techniques to provide insights into music streaming services. However, there are various music streaming services around the world. For instance, Apple Music is also a major service with a large market share. However, because Apple Music is not compatible with Android devices using Google Play, it is difficult to obtain high-quality customer opinions on music streaming from the Apple Music app reviews that are written on the Google Play app. In future research, if various music streaming services are included as targets, more diverse insights could be obtained. Second, the purpose of this study is to provide an understanding of music streaming services by analyzing customer opinions expressed in app reviews through a data-driven method. Therefore, this study focuses mostly on correlation between independent variables extracted from app reviews and overall rating. However, correlation cannot fully represent causation. Therefore, in order to derive more in-depth managerial implications, it is necessary to conduct research that can estimate reliable causal effect. Third, the five services analyzed in this study were global services. Nevertheless, this study only selected English reviews from the US Google Play. Future research can consider using the latest multilingual language models to analyze reviews written by customers of music streaming services around the world. Fourth, this study only considered app reviews to analyze music streaming service customers. However, since there may be various social media platforms where customers can post opinions about music streaming services, future research can integrate and analyze various social media data. Fifth, this study is a data-driven study, focusing on researching and analyzing reviews posted by customers who use music streaming services. Thus, future research will be able to combine the results of this study with various theories related to consumers and customers established by eminent scholars. It is expected that the theoretical implications of the results presented in this study can be further expanded. Finally, some customers may misunderstand the device problems as a service problem. For example, the loading speed could be a problem of Internet connection or server optimization, or it could pertain to using an aged device. If the device information of each customer can be collected, future research may be able to identify problems with the services by minimizing noise.

6. Conclusion

This study, involving social media mining, provides a comprehensive understanding of music streaming services from a customer-centric perspective. With advances in information technology, music streaming services have completely upended the music industry. In particular, music streaming services are often used on mobile devices; thus, there are plenty of app reviews that allow people to understand their customers' needs. Nevertheless, previous studies simply tried to do this by surveys. This study analyzed large-scale app review data for five music streaming services to overcome the limitations of previous survey-based studies. Three research questions were used, reflecting the main research streams in the field of social media mining and were addressed based on topic modeling and text regression.

As a result, this study identified nine latent factors that customers commented on with regards to music streaming services and grouped them into three categories: environment-related, pricing-related, and content-related factors. In addition, the factors that affect overall customer satisfaction were investigated from a global perspective and from a service-specific perspective. This study identified whether each factor had a significant effect on customer satisfaction and interpreted the results by reflecting the actual reviews and characteristics of each service. This study contributes to providing a comprehensive understanding of music streaming services, evaluating the results of cross-service competitive analysis from a customer's point of view.

CRediT authorship contribution statement

Jaemin Chung: Conceptualization, Data curation, Formal analysis, Writing – original draft. **Jiho Lee:** Conceptualization, Methodology, Resources, Software. **Janghyeok Yoon:** Conceptualization, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2021R1A2C101002711). This work was supported by the Human Resources Program in Energy Technology of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No. 20204010600220).

Appendix

Appendix. Price plans for music streaming services (plans as of the end of 2020).

Music streaming Service	Name of plan	Fee (per month)	Description
Amazon Music	Prime (for Amazon Prime user)	$0.00	
	Unlimited (for Amazon Prime user)	$7.99	lossless HD songs; Ultra HD tracks; expert-programmed playlists and stations
	Unlimited for family (for Amazon Prime user)	$14.99	Six different accounts can use Unlimited plan
	Free (for non-Prime user)	$0.00	Commercial breaks; Does not offer offline playback
	Unlimited (for non-Prime user)	$9.99	lossless HD songs; Ultra HD tracks; expert-programmed playlists and stations
Deezer	Free	$0.00	Commercial breaks; Only shuffle mode available; Does not offer offline mode
	Premium	$9.99	
	Family	$14.99	Six different accounts can use Premium plan
	Hi-Fi	$14.99	Premium plan with Access to lossless audio (FLAC) cross device
Spotify	Free	$0.00	Commercial breaks; Does not offer offline mode; sound quality up to 96Kbits/s
	Premium	$9.99	
	Family	$14.99	Five different accounts can use Premium plan
Tidal	Premium	$9.99	
	Hi-Fi	$19.99	Premium plan with lossless High Fidelity sound quality, Master Quality audio
	Family Premium	$14.99	Six different accounts can use Premium plan
	Family Hi-Fi	$29.99	Six different accounts can use Premium plan with lossless High Fidelity sound quality, Master Quality audio
YouTube Music	Premium (for YouTube Premium user)	$0.00	
	Free (for non-Premium user)	$0.00	Commercial breaks; Does not offer offline mode
	Premium (for non-Premium user)	$11.99	
	Family (for non-Premium user)	$17.99	Six different accounts can use Premium plan

Note: Discounted plans (e.g., plans for students and military, single device plan) are not listed in the table.

References

Aizawa, A., 2003. An information-theoretic perspective of tf–idf measures. Inf. Process. Manage. 39, 45–65.

Alrawadieh, Z., Law, R., 2019. Determinants of hotel guests' satisfaction from the perspective of online hotel reviewers. Int. J. Cult., Tour. Hospitality Res. 13, 84–97.

Barna, E., 2017. "The perfect guide in a crowded musical landscape:" Online music platforms and curatorship. First Monday.

Biswas, B., Sengupta, P., Chatterjee, D., 2020. Examining the determinants of the count of customer reviews in peer-to-peer home-sharing platforms using clustering and count regression techniques. Decis. Support Syst. 135, 113324.

Borja, K., Dieringer, S., Daw, J., 2015. The effect of music streaming services on music piracy among college students. Comput. Hum. Behav. 45, 69–76.

Brem, A., Voigt, K.-I., 2009. Integration of market pull and technology push in the corporate front end and innovation management—insights from the German software industry. Technovation 29, 351–367.

Choi, H., Oh, S., Choi, S., Yoon, J., 2018. Innovation topic analysis of technology: The case of augmented reality patents. IEEE Access 6, 16119–16137.

Choi, J., Oh, S., Yoon, J., Lee, J.-M., Coh, B.-Y., 2020a. Identification of time-evolving product opportunities via social media mining. Technol. Forecast. Soc. Chang. 156, 120045.

Choi, J., Yoon, J., Chung, J., Coh, B.-Y., Lee, J.-M., 2020b. Social media analytics and business intelligence research: A systematic review. Inf. Process. Manage. 57, 102279.

Danaher, P.J., 2002. Optimal pricing of new subscription services: Analysis of a market experiment. Mark. Sci. 21, 119–138.

Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R., 1990. Indexing by latent semantic analysis. J. Am. Soc. Inf. Sci. 41, 391–407.

Deng, Q., Hine, M.J., Ji, S., Wang, Y., 2021. Understanding consumer engagement with brand posts on social media: the effects of post linguistic styles. Electron. Commer. Res. Appl. 48, 101068.

Erzurumlu, S.S., Pachamanova, D., 2020. Topic modeling and technology forecasting for assessing the commercial viability of healthcare innovations. Technol. Forecast. Soc. Chang. 156, 120041.

Friedlander, J.P., 2021. 2020 Year-End Music Industry Revenue Report.

Garcia-Gathright, J., St. Thomas, B., Hosey, C., Nazari, Z., Diaz, F., 2018. Understanding and evaluating user satisfaction with music discovery. In: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, pp. 55–64.

Gerpott, T.J., 2005. Strategisches technologie-und innovationsmanagement. Schäffer-Poeschel.

Grönroos, C., 1984. A service quality model and its marketing implications. European. J. Mark.

Hagen, A.N., 2015. The playlist experience: personal playlists in music streaming services. Popular Music Society 38, 625–645.

He, L., Han, D., Zhou, X., Qu, Z., 2020. The voice of drug consumers: online textual review analysis using structural topic model. Int. J. Environ. Res. Public Health 17, 3648.

Homburg, C., Hoyer, W.D., Koschate, N., 2005. Customers' reactions to price increases: do customer satisfaction and perceived motive fairness matter? J. Acad. Mark. Sci. 33, 36–49.

Hracs, B.J., Webster, J., 2021. From selling songs to engineering experiences: exploring the competitive strategies of music streaming platforms. J. Cult. Econ. 14, 240–257.

Hu, Y., Boyd-Graber, J., Satinoff, B., Smith, A., 2014. Interactive topic modeling. Mach. Learning 95, 423–469.

Hu, N., Zhang, T., Gao, B., Bose, I., 2019. What do hotel customers complain about? Text analysis using structural topic model. Tour. Manage. 72, 417–426.

Jeong, B., Yoon, J., Lee, J.-M., 2019. Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis. Int. J. Inf. Manage. 48, 280–290.

Jin, J., Ji, P., Gu, R., 2016. Identifying comparative customer requirements from product online reviews for competitor analysis. Eng. Appl. Artif. Intell. 49, 61–73.

Joosse, S., Hracs, B.J., 2015. Curating the quest for 'good food': The practices, spatial dynamics and influence of food-related curation in Sweden. Geoforum 64, 205–216.

Kim, J., Nam, C., Ryu, M.H., 2017. What do consumers prefer for music streaming services?: A comparative study between Korea and US. Telecommun. Policy 41, 263–272.

Kjus, Y., 2016. Musical exploration via streaming services: The Norwegian experience. Popular Commun. 14, 127–136.

Kulkarni, S.S., Apte, U.M., Evangelopoulos, N.E., 2014. The use of latent semantic analysis in operations management research. Decis. Sci. 45, 971–994.

Landauer, T.K., Foltz, P.W., Laham, D., 1998. An introduction to latent semantic analysis. Discourse Processes 25, 259–284.

Liu, Y., Bi, J.-W., Fan, Z.-P., 2017. Ranking products through online reviews: A method based on sentiment analysis technique and intuitionistic fuzzy set theory. Inf. Fusion 36, 149–161.

Liu, Y., Jiang, C., Zhao, H., 2019b. Assessing product competitive advantages from the perspective of customers by mining user-generated content on social media. Decis. Support Syst. 123, 113079.

Liu, D.-R., Liao, Y.-S., Lu, J.-Y., 2019a. Online news recommendations based on topic modeling and online interest adjustment. Ind. Manage. Data Syst. 119, 1802–1818.

Lowe-Calverley, E., Grieve, R., 2018. Thumbs up: A thematic analysis of image-based posting and liking behaviour on social media. Telematics Inform. 35, 1900–1913.

Lüders, M., 2020. Ubiquitous tunes, virtuous archiving and catering for algorithms: the tethered affairs of people and music streaming services. Inf., Commun. Society 1–17.

Maasø, A., 2018. Music streaming, festivals, and the eventization of music. Popular Music Society 41, 154–175.

Marshall, L., 2015. 'Let's keep music special. F—Spotify': on-demand streaming and the controversy over artist royalties. Creative Ind. J. 8, 177–189.

McIlroy, S., Shang, W., Ali, N., Hassan, A.E., 2015. Is it worth responding to reviews? Studying the top free apps in google play. IEEE Softw. 34, 64–71.

Misuraca, M., Scepi, G., Spano, M., 2020. A network-based concept extraction for managing customer requests in a social media care context. Int. J. Inf. Manage. 51, 101956.

Morris, J.W., Powers, D., 2015. Control, curation and musical experience in streaming music services. Creative Ind. J. 8, 106–122.

Nam, K., Baker, J., Ahmad, N., Goo, J., 2020. Determinants of writing positive and negative electronic word-of-mouth: empirical evidence for two types of expectation confirmation. Decis. Support Syst. 129, 113168.

Park, E.O., Chae, B.K., Kwon, J., Kim, W.-H., 2020. The effects of green restaurant attributes on customer satisfaction using the structural topic model on online customer reviews. Sustainability 12, 2843.

Park, I., Jeong, Y., Yoon, B., Mortara, L., 2015. Exploring potential R&D collaboration partners through patent analysis based on bibliographic coupling and latent semantic analysis. Technol. Anal. Strategic Manage. 27, 759–781.

Pauwels, K., Weiss, A., 2008. Moving from free to fee: How online firms market to change their business model successfully. J. Mark. 72, 14–31.

Pinto, S., Albanese, F., Dorso, C.O., Balenzuela, P., 2019. Quantifying time-dependent Media Agenda and public opinion by topic modeling. Physica A 524, 614–624.

Prey, R., 2018. Nothing personal: Algorithmic individuation on music streaming platforms. Media Cult. Soc. 40, 1086–1100.

Qiao, Z., Wang, G.A., Zhou, M., Fan, W., 2018. The impact of customer reviews on product innovation: empirical evidence in mobile apps. Anal. Data Sci. 95–110.

Raats, T., Evens, T., 2021. 'If you can't beat them, be them': A critical analysis of local streaming platform and Netflix alternative Streamz. MedieKultur: J. Media Commun. Res. 37, 050–065.

Sarker, A., O'connor, K., Ginn, R., Scotch, M., Smith, K., Malone, D., Gonzalez, G., 2016. Social media mining for toxicovigilance: automatic monitoring of prescription medication abuse from Twitter. Drug Saf. 39, 231–240.

Sidorova, A., Evangelopoulos, N., Valacich, J.S., Ramakrishnan, T., 2008. Uncovering the intellectual core of the information systems discipline. MIS Q. 467–482.

Siering, M., Koch, J.-A., Deokar, A.V., 2016. Detecting fraudulent behavior on crowdfunding platforms: The role of linguistic and content-based cues in static and dynamic contexts. J. Manage. Inf. Syst. 33, 421–455.

Silverstein, D., Samuel, P., DeCarlo, N., 2013. The innovator's toolkit: 50+ techniques for predictable and sustainable organic growth. John Wiley & Sons.

Smits, R., Nikdel, E., 2019. Beyond Netflix and Amazon: MUBI and the curation of on-demand film. Stud. Eur. Cinema 16, 22–37.

Stamolampros, P., Korfiatis, N., Chalvatzis, K., Buhalis, D., 2019. Job satisfaction and employee turnover determinants in high contact services: Insights from Employees' Online reviews. Tour. Manage. 75, 130–147.

Vayansky, I., Kumar, S.A., 2020. A review of topic modeling methods. Inf. Syst. 94, 101582.

Venugopalan, S., Rai, V., 2015. Topic based classification and pattern identification in patents. Technol. Forecast. Soc. Chang. 94, 236–250.

Wang, W., Feng, Y., Dai, W., 2018. Topic analysis of online reviews for two competitive products using latent Dirichlet allocation. Electron. Commer. Res. Appl. 29, 142–156.

Wang, Y., Luo, L., Liu, H., 2020. Bridging the semantic gap between customer needs and design specifications using user-generated content. IEEE Trans. Eng. Manage. 1–13.

Webster, J., 2020. Taste in the platform age: music streaming services and new forms of class distinction. Inf., Commun. Society 23, 1909–1924.

Weinberger, M., Bouhnik, D., 2020. Various information aspects following the emergence of music streaming applications. Online Inf. Rev. 45, 118–137.

Wlömert, N., Papies, D., 2016. On-demand streaming services and music industry revenues — insights from Spotify's market entry. Int. J. Res. Mark. 33, 314–327.

Xu, X., 2020. How do consumers in the sharing economy value sharing? Evidence from online reviews. Decis. Support Syst. 128, 113162.

Xu, X., 2021. What are customers commenting on, and how is their satisfaction affected? Examining online reviews in the on-demand food service context. Decis. Support Syst. 142, 113467.

Yun, S., Song, K., Kim, C., Lee, S., 2021. From stones to jewellery: Investigating technology opportunities from expired patents. Technovation 103, 102235.

Zuo, W., Zhu, W., Chen, S., He, X., 2019. Service quality management of online car-hailing based on PCN in the sharing economy. Electron. Commer. Res. Appl. 34, 100827.