Research Article

TUJOM
ISSN:2458-9748

TURKISH JOURNAL OF MARKETING
Vol.: 4 Issue: 3 Year: 2019, pp.265-279

Citation: Pınarbaşı, F. (2019), Demystifying Musical Preferences At Turkish Music Market Through Audio Features Of Spotify Charts, TUJOM, (2019), 4(3): 265-279 doi: http://dx.doi.org/10.30685/tujom.v4i3.62

DEMYSTIFYING MUSICAL PREFERENCES AT TURKISH MUSIC MARKET THROUGH AUDIO FEATURES OF SPOTIFY CHARTS

Fatih PINARBAŞI[1]

Received Date (Başvuru Tarihi): 30/10/2019
Accepted Date (Kabul Tarihi): 22/11/2019
Published Date (Yayın Tarihi): 26/12/2019

ABSTRACT

Online music streaming services are one of the important actors in music consumption for today's consumers. In addition to widespread use of mobile devices, many changes in the patterns of music consumption are witnessed such as the purchase of single tracks instead of albums, listening to music on different platforms, and personalized music consumption options. This study aims to examine the concept of music consumption in Turkey through audio characteristics of popular songs. Top 200 popular song-lists for 6 months period are chosen as sample and audio characteristics provided by Spotify API service regarding 676 unique songs are analyzed. Following descriptive statistics of Turkey Music Market, clustering methodology is employed and three different clusters for songs are concluded. Finally, decision tree methodology is employed to classify the dataset with popularity scores and audio characteristics together, while loudness and energy characteristics are found as significant classifiers.

Keywords: *Music Marketing, Consumer Preferences, Spotify*

JEL Codes: *M31, M30*

TÜRK MÜZİK PAZARINDAKİ MÜZİKAL TERCİHLERİN SPOTIFY MÜZİK LİSTELERİNDEKİ PARÇALARIN SES ÖZELLİKLERİ ARACILIĞIYLA BELİRLENMESİ

ÖZ

Çevrimiçi müzik yayını hizmetleri, günümüz tüketicileri için müzik tüketiminde önemli aktörlerden biridir. Mobil cihazların yaygın olarak kullanılmasına ek olarak müzik tüketiminde; albüm yerine tek parça satın alma, farklı platformlarda müzik dinleme ve kişiselleştirilmiş müzik tüketimi seçenekleri gibi birçok değişiklik yaşanmıştır. Bu çalışma Türkiye'deki müzik tüketimini, popular şarkıların ses özellikleri üzerinden incelemeyi amaçlamaktadır. 6 aylık süreçte yer alan popüler 200 şarkı listeleri örneklem kitle olarak seçilmiş ve 676 parçaya dair Spotify API hizmetinin sağladığı ses özellikleri analiz edilmiştir. Türkiye müzik piyasasına dair tanımlayıcı bilgilerin ardından, araştırmada kümeleme analizi gerçekleştirilmiş ve üç farklı kümeye ulaşılmıştır. Son olarak, ses özelliklerine ilaveten, Spotify API tarafından sağlanan popülerlik değerleri üzerinden karar ağacı yöntemi kullanılarak sınıflandırma yapılmış, gürültü (loudness) ve enerji değerlerinin popülerlik açısından önemli ayırt edici özellikler olduğu sonucuna ulaşılmıştır.

Anahtar Kelimeler: *Müzik Pazarlaması, Tüketici Tercihleri, Spotify*

JEL Kodları: *M31, M30*

[1] Research Assistant, Istanbul Medipol University, fpinarbasi@medipol.edu.tr https://orcid.org/0000-0001-9005-0324

1. INTRODUCTION

Online music platforms have crucial roles in today's consumers' consumption processes. As distribution of content evolves, several business models emerge and new research questions regarding to these concepts arise. From historical perspective, music consumption history have many breakthroughs; including vinyl period, tape and CD's period and online music age. This study focuses on last period of online music age which reflects streaming services beyond music downloading or individual sharing.

Online streaming services become popular in recent years especially with brands like Netflix and Spotify. Spotify service which consumers can listen songs from their various devices and make playlists is an online streaming service for audio content. It is launched in 2008 and has more than 50 million songs and available in 79 markets (Spotify, 2019b). The market of streaming services in terms of Spotify is notable for music and entertainment marketing. According to Spotify (2019), service has 100 million subscribers in 2019 Q1, while it has 217 million monthly active users. This study considers potential of online streaming services and Spotify platform and examine market with songs-based perspective.

The scope of previous studies regarding to Spotify and music concept can be divided into three parts including; popularity-based approach for music markets, individuals' music preferences and psychologies and business-model approach for streaming services. Instead of individual-focused music consumption behavior, this study has market-based perspective. This study has popularity-based approach as main perspective and examines music market through popular charts and audio features of songs in these charts. In addition, there is lack of studies in Turkey music market context, as previous studies mostly evaluate other countries' songs. Therefore, this study aims to fill local context gap with popularity-based approach. Research questions of study refer to;

- What are the main musical characteristics of popular songs in Turkey market?
- How can songs be segmented into groups by audio features?
- How can songs can be segmented with audio features and popularity ranking?

The study starts with theoretical background section including online media streaming and Spotify platform concepts. The methodology section follows theoretical background with descriptive statistics, continues with cluster analysis and decision tree classification methodologies. Finally, conclusion and future research directions are indicated.

2. THEORETICAL BACKGROUND

2.1. Music Consumption

Music concept is a significant factor for today's consumers as they face musical elements in daily lives consciously or unconsciously. Bruner (1990) examine music, mood and marketing topics together in his study and conclude a table of elements (based on information given by Cooke (1962) and Zettl (1973)) regarding to emotional expressions ascribed to various components of music. This summative table has three main segments; time related expressions, pitch-related expressions and texture-related expressions. For example, time related expressions indicate that faster tempo implies expression of more animation and happiness. Pitch-related expressions indicates that higher keys songs are considered as happier than lower keys songs. This table of element is starting point of this study, as the audio features includes important insights for consumers.

Music consumption phenomenon can be examined in two main concepts; personal perspective, market perspective. These perspectives reflect various sides of consumption behaviour with different scopes. For example, Larsen et al. (2010) focus on symbolic music consumption and propose a framework including people-based factors and music-based factors as situation group category. Music consumption phenomenon can be portrayed as meeting up of music meta (song, playlist etc.) and consumer in various ways including vinyl records, cassettes, CDs or digital platforms.

First perspective includes personal factors affecting consumption; psychological variables and personal variables. Rentfrow and Gosling (2003) examine personality with musical preference in their study and they conclude four music-preference dimensions. These dimensions are reflective and complex, intense and rebellious, upbeat and conventional, energetic and rhythmic. In another study, Delsing et al. (2008) focus on adolescents' music preferences with personality dimensions and reveal four music-preference dimensions; rock, elite, urban and pop/dance. Personal attitudes and preferences plays a key role in musical consumption in terms of individual perspective.

Music consumption is also related to market elements, beyond personal factors. Music consumption has lived a revolution through digitalization and MP3 technology since accessibility to preferred songs at any time and any place is available for consumers (Cockrill et al., 2011). New music consumer does not wait to listen his/her favourite songs to play on radio or TV, instead consumer play any song in mobile device in seconds. The availability is one of important factors for today's market for any type of content; music, video etc. Other aspect of new changes refers to singularity of purchasing behavior. New consumers do not have to purchase entire albums, they can choose any song they want to purchase without buying entire album (Cockrill et al., 2011). These factors create a new way of music consumption which has new research questions due to new conditions.

Beyond perspectives of phenomenon, music consumption is studied in several new contexts including; digital music consumption (Magaudda, 2011), vinyl record in digital age (Bartmanski & Woodward, 2015), symbolic music consumption (Larsen et al., 2010) etc. As advancements and new consumption concepts have many sides to evaluate, the perspective for music consumption of this study refers to Turkey music market consists of songs. Consistent to market-based perspective, next section provides theoretical information about online streaming services and Spotify platform.

2.2 Spotify and Music Streaming

Digital revenues in global music market has crucial role for music marketing. According to IFPI Global Music Report (2019), digital music has 58.9% of global music revenue. Subscription-based business models are critical for today's digital music market, as purchasing behaviour patterns have changed. In this new music market, music streaming services like Spotify has a bridge role between music producers and music consumers.

Online streaming services and music streaming concepts are studied in various contexts including; music industry revenue (Wlömert & Papies, 2016), music piracy (Borja et al., 2015), ethics (Weijters et al., 2014) etc. Music streaming services have various roles for music consumption; i) gathering of musical meta from artists/ producers, ii) delivering audio content to end users, iii) improving user experience with personalization. First and second roles are mostly related to economical side of music consumption, while third part is related to preference part of music consumption.

Understanding consumer experience, improving experience and long-term sustainability are important factors for online music services. Music streaming services like Spotify offer personalized playlists for their consumers and the playlist feature is one of important concepts in musical consumption. Spotify creates a consuming experience which let users to discover new songs similar to their current interests or provide them dashboards with various genres of music.

Evaluating what music consumers really want or would like to listen is important question for online streaming services and technical features have critical roles for this question.

Preferences are main object of this study and one of significant factors for preferences of music consumption is related to attributes of songs. Greenberg et al. (2016) conclude complexity of music and suggest evaluating sonic attributes (like timbre and instrumentation) in music regarding to perceived attributes, for future researchers. Since, Spotify API has various audio features for musical attributes including; acousticness, danceability, energy, instrumentalness, liveness, loudness etc., (Spotify, 2019c), consistent to Greenberg et al. (2016)'s suggestion and Bruner's (1990) emotional expressions summary related to components of music, this study employs audio features for evaluating music market and next section includes methodology section of study. This study evaluates online streaming services in music streaming context while focusing on market by Spotify charts.

3. RESEARCH

The aim of study is two-folds as first part includes detection of popular song characteristics and their cluster groups, while other part refers to examine them with their significant characteristics with decision tree methodology. The study inspects Turkey market as context and evaluates popular songs in Turkey market as sample of study. Six months period from January 2019 to July 2019 is selected as time frame and audio features provided by Spotify API is employed for data of methodology. Data collection starts with scraping popular songs chart data of Turkey from spotifycharts.com for the 6 months-time period. Methodology section employs 676 unique songs data from 36200 chart positions.

3.1. Methodology

The methodology section of study consists of five sections; gathering of popular songs charts data, unique songs filtering, gathering audio features of selected 676 popular songs, descriptive stats of 676 popular songs, cluster analysis of audio features and employing decision tree methodology for popularity classification. This study mainly uses R programming language (R Core Team, 2018) and RStudio software (R Studio Team, 2015) for methodology and analysis.

Gathering of data starts with Spotify API (Spotify, 2019d) as Spotify developer platform makes audio features of tracks (like danceability, energy, valence) available to developers. Spotifyr package (Thompson et al., 2019) is employed for data scraping of audio features, while rvest package (Wickham, 2019) packages for scraping data of www.spotifycharts.com website. Other data manipulation and editing code packages like dplyr (Wickham et al., 2019), purrr (Henry & Wickham, 2019), tibble (Müller & Wickham, 2019), magrittr (Bache & Wickham, 2014), ggplot (Wickham, 2016) and xlsx (Dragulescu & Arendt, 2018) are also employed in this study.

This study examines Top 200 list for each day in 6 months period and chart information for 36200 songs are gathered for study. As same songs remain in top list for several days in different rankings, duplicate values are removed and only 676 songs remained. Next section includes descriptive statistics regarding to audio features provided by Spotify API (Spotify, 2019c).

3.1.1. Descriptive Information

Table 1 summarizes information of audio features from Spotify (2019c).

Table 1. Descriptions of Audio Features

Feature	Description
Duration_ms	Duration of songs in millisecond format.
Danceability	Suitability value of songs to dance. This includes several musical elements like tempo, beat strength etc.
Energy	Perceptual measure of intensity and activity. This includes several musical elements dynamic range, timbre, perceived loudness etc.
Key	Overall key of track. 0=C, 1=C#, 2=D etc.
Loudness	Decibel values of track. Values range between -60 and 0db.
Mode	Modality of tracks which indicates major (1) or minor (0) values.
Speechiness	Presence of spoken words in track. 0.66 and 0.33 are segmenting values. Typical songs have lower than 0.33 values.
Acousticness	Indicates acousticness level of songs and have values between 0 and 1.
Instumentalness	Related to amount of vocals in track. Songs have higher than 0.5 instrumentalness values are tend to be instrumental tracks.
Liveness	Measures presence of audience in track. Values higher than 0.8 tend to be live performance.
Valence	Musical positiveness transferred by track. Higher value signals positive while lower values signal negative.

Tempo	Estimated beats per minute value.
Time_Signature	Estimated overall time signatue value.

Source: https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/

Second step of descriptive stats refers to frequency distributions of 676 songs regarding to audio features provided by Spotify API. Black lines indicate mean values of distributions in Figure 3. Multiple graphs are created with Chang (2019)'s multiple plot function which includes grid (R Core Team, 2018) and ggplot2 (Wickham, 2016) code packages.

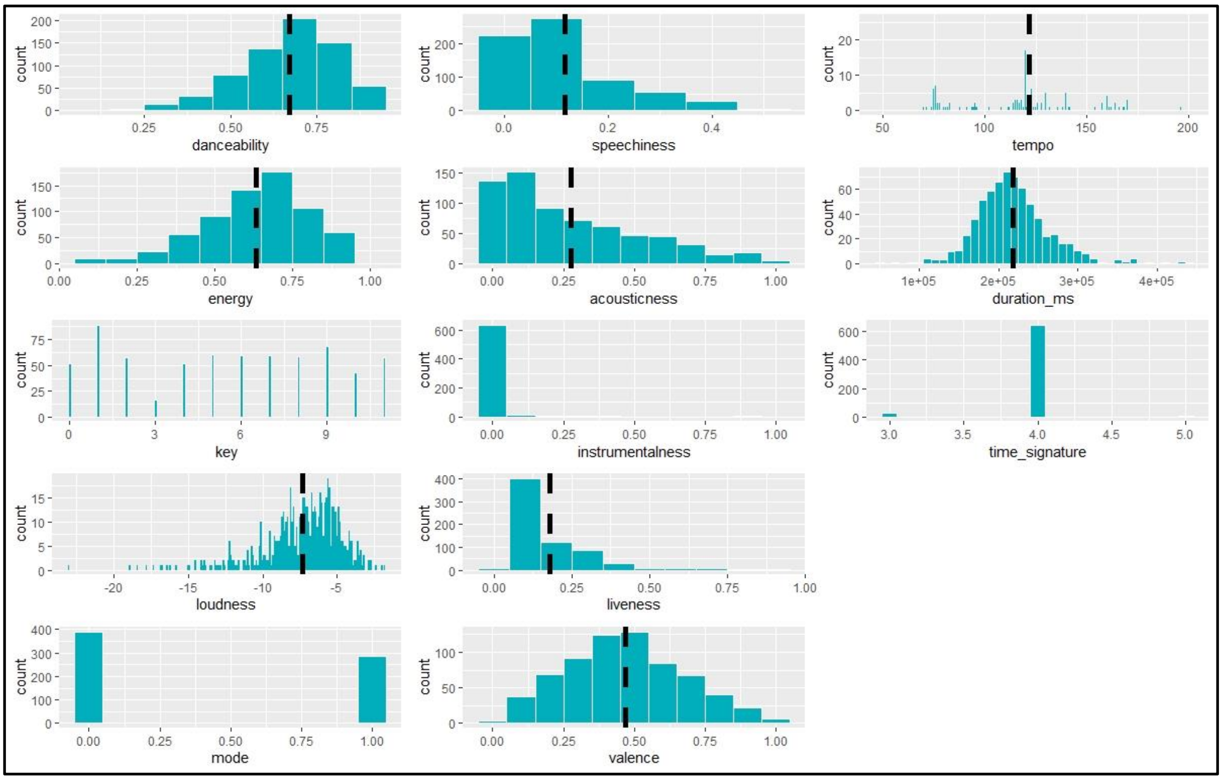

Figure 1. Frequency of Audio Features

Figure 1 indicates that danceability, energy, loudness features have left-skewed distribution, as mean values are cumulated to higher values. On the other hand, speechiness, acousticness, and liveness features have right-skewed distribution, as mean values are cumulated to lower values. Valence, duration_ms and tempo audio features are like normal distribution. In addition to three distribution types, there are various audio features have only limited values like mode, instrumentalness, time_signature.

The distributions shown in Figure 1 point out overall features of Turkey market popular songs as they show summary information. Most popular songs are generally have higher energy, loudness and danceability features, while they have lower speechiness, acousticness and liveness values. Last conclusion of descriptive statistics refers to valence distribution of popular songs. As valence feature indicates positivity/negativity sides of songs, popular songs of Turkey market do not have skewed distribution. The equilibrium of valence values shows the overall preference of Turkish music consumers regarding to positivity/negativity sides of songs.

Following summary of audio features, correlation relationships between features would contribute to understanding of market. Correlation values are calculated between audio features, while corrplot (Wei & Simko, 2017) package is employed for graph generation.

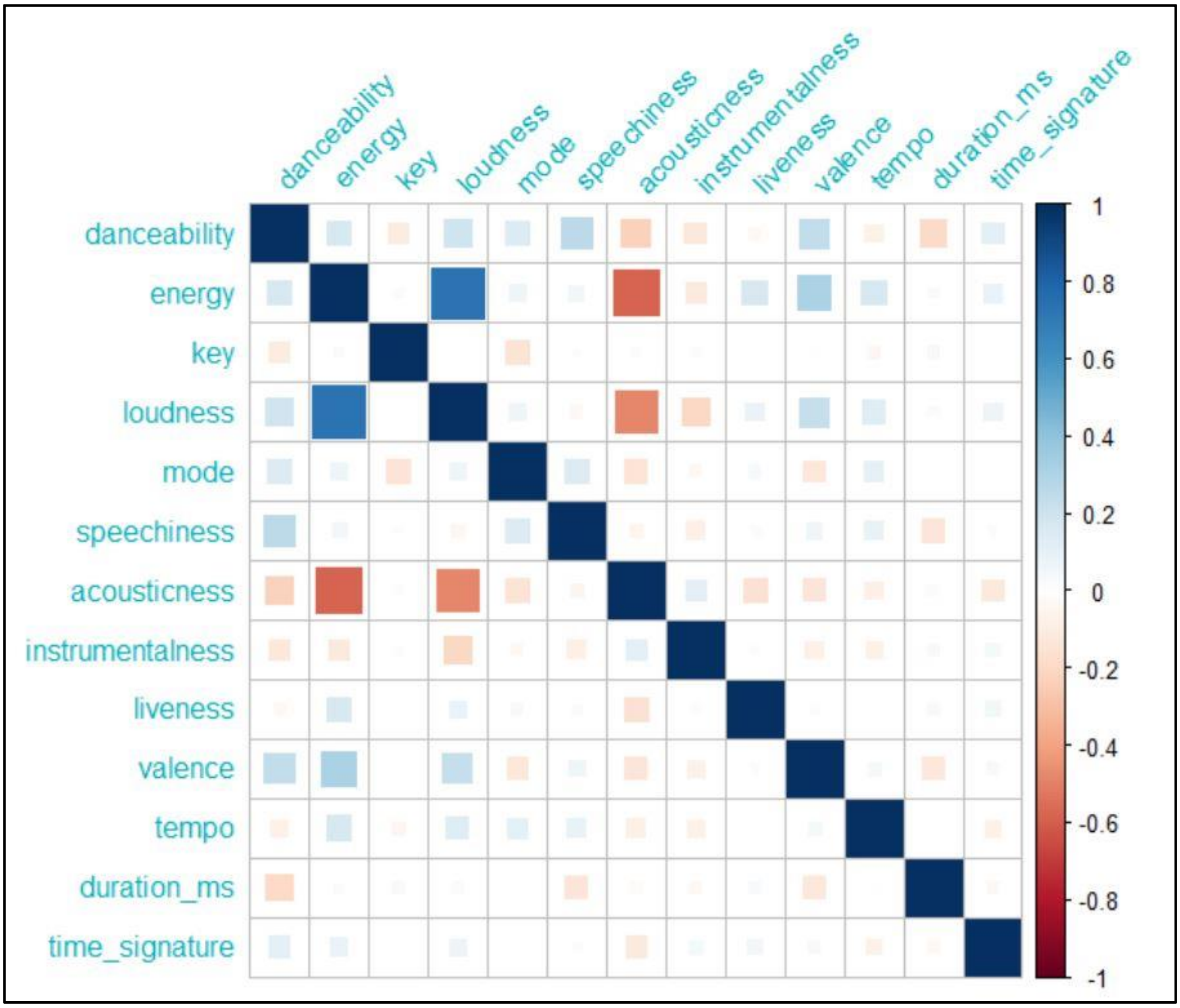

Figure 2. Correlation Table of Audio Features

Figure 2 indicates correlation values with color tones and sizes of squares, darker blue colors refer to positive correlation, while dark orange color tones refer to negative correlation. Sizes of square point out correlation power. In conclusion, there are strong positive correlations between energy and loudness features and there are medium positive correlations between valence feature and danceability, energy and loudness features. On the other hand, there are negative correlations between energy and acousticness, loudness and acousticness features.

First section includes descriptive information about popular songs in Turkey market. Next section examines audio features of songs with clustering approach to have better understanding.

3.1.2. Clustering of Songs

Unique songs of dataset is examined with clustering approach to segment songs with features. RapidMiner Studio software (Mierswa et al., 2006) is employed for cluster methodology. Time_signature audio feature is removed from cluster analysis, since it includes commonly same values. Methodology part of clustering employs X-means algorithm. X-means algorithm is developed for problems of k-means algorithm and it has an algorithm which quickly estimates k (number of clusters) (Pelleg & Moore, 2000). X-means algorithm is employed with minimum 2, maximum 20 clusters options and it concludes three different clusters for audio features.

Table 2. Clusters of Audio Features

Clusters	N/ %	Feature Average Values		
Cluster 0	484 - 71.60%	Instrumentalness 65.10% smaller	Acousticness 34.17% smaller	Speechiness 17.62% larger
Cluster 1	178 - 26.33%	Acousticness 88% larger	Speechiness 43.94% smaller	Mode 36.66 % smaller
Cluster 2	14 - 2.7%	Instrumentalness 2896.97% larger	Acousticness 88.66% larger	Speechiness 65.16% smaller
Distance Measure: Euclidean Distance Average Cluster Distance: 9.586 Davies-Bouldin Index: 1.808				

In addition to values of audio features, cluster centroid chart is included in Figure 3.

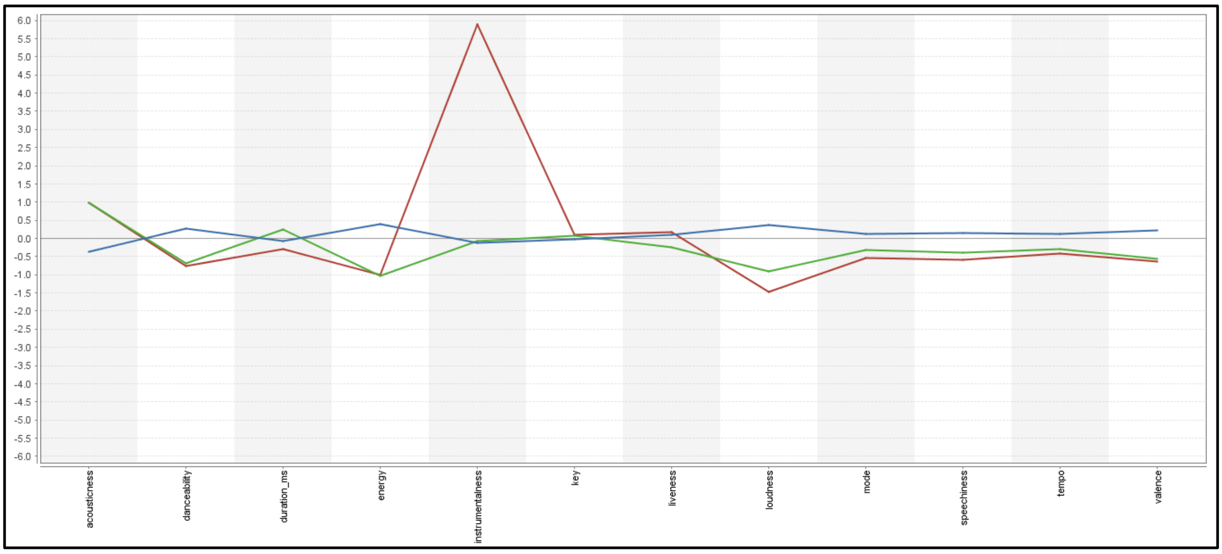

Figure 3: Centroid Chart for Clusters

Table 2 and Figure 3 show that cluster named “cluster 2” is different from other clusters with instrumentalness values. Cluster 2 has 14 songs inside and it includes 2.7% of all songs. Cluster 0 and Cluster 1 has similar feature values for duration_ms, instrumentalness, key, liveness and mode features. Different part of these two clusters refer to loudness, energy and danceability audio features. As Figure 3 indicates, cluster 0 (blue line) has higher danceability, energy and loudness levels than cluster 1 (red line). Therefore, it can be concluded that there are two main clusters in Turkey market popular songs and they differentiate by three audio features. Turkey music market also has a marginal cluster of songs that differ by instrumentallnes values.

First two sections of methodology include descriptive statistics regarding to audio features of Turkey market popular songs. Third section focuses on popularity concept with audio features while it combines popularity attribute scores provided by Spotify.

3.1.3. Decision Tree for Popularity Classification

This section includes popularity scores in addition to audio features; and examine the popularity concept with audio features. Spotify (2019e) indicates that popularity scores calculated by an algorithm which includes total numbers of play and recentness of plays, it ranges from 0 to 100. Classification nature of this section employs decision tree methodology. Decision tree refers to classification procedure which recursively partitions a dataset to smaller groups by using set of tests which are defined at each branch (Friedl & Brodley, 1997). Consistent to Oklap (2018)’s approach, all songs are ordered by popularity scores, as maximum popularity scores indicate top rankings and minimum popularity scores indicate bottom rankings.

Decision tree section uses rpart (Therneau & Atkinson, 2019) and rpart.plot (Milborrow, 2019) R code packages for methodology.

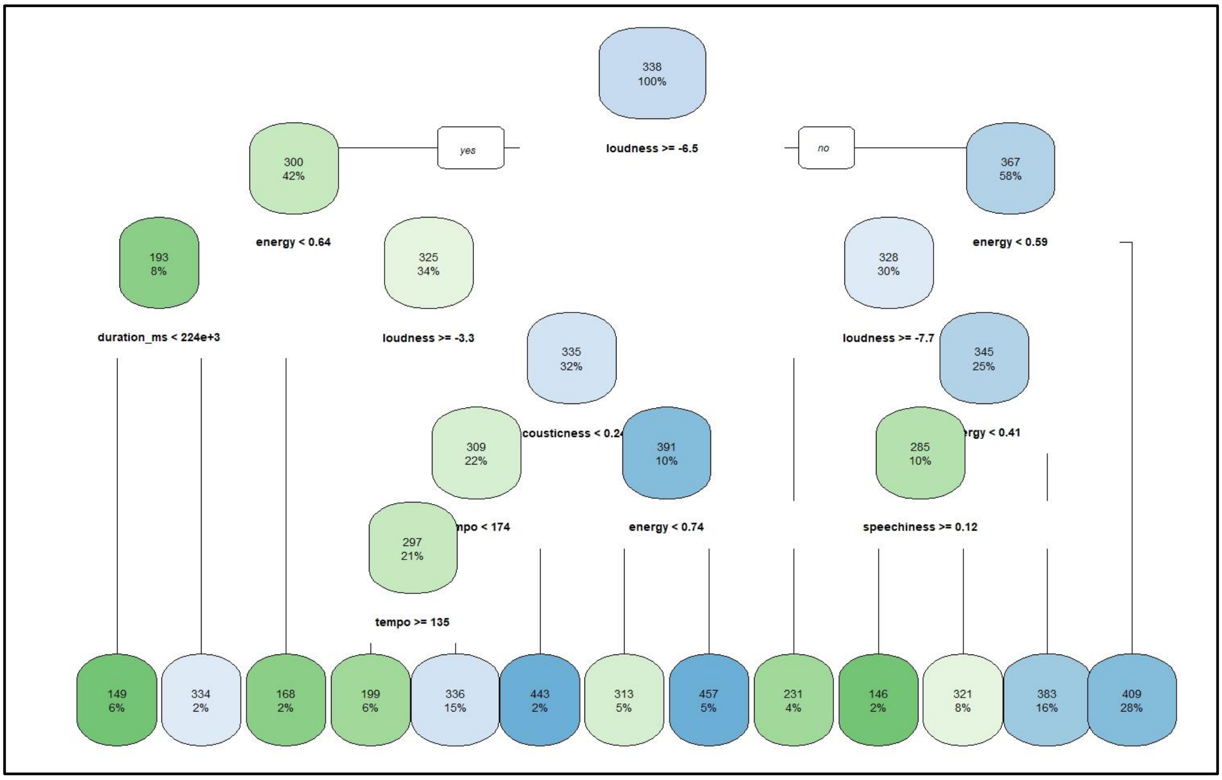

Figure 4. Decision Tree Classification of Popularity Rankings

The importance ranking of variables for dataset refers to; energy (29), loudness (23), acousticness (12), tempo (10), danceability (6).

Figure 4 classify popularity of songs left to right, as green leaf at bottom left indicates most popular songs and blue leaf at bottom right indicates least popular songs. Figure 4 also indicates that popularity rankings of Turkish market's popular songs are first classified by loudness level (-6.5) and 58% of songs have < -6.5 loudness levels. Loudness levels provided by Spotify API (Spotify, 2019c) are ranged from -60db to 0db. Second level significant classifier variables refer to energy and it divides two main nodes. Third level significant classifier variables refer to duration_ms and loudness, while other levels include acousticness, energy, tempo and speechiness variables.

Conclusion part of classification includes two main results; highest probability of top rankings indicates features (higher than) >=-6.4 loudness level, (lower than) <0.64 energy level and (lower than) duration_ms < 224000 (224 seconds). Highest probability of lowest rankings indicates features (higher than) < -6.4 loudness level and (higher than) >0.59 energy level.

Example of most popular song feature values refer to -6.039 loudness level and 0.54 energy level, while least popular song feature value refers to -7.895 loudness level and 0.617 energy level.

4. CONCLUSION

This study examines Turkish music market by popular songs and their audio features; and evaluates three main research topics and questions. First topic refers to descriptive statistics of songs and second topic refers to clustering of them, while last question employs decision tree methodology for classification of popular songs' characteristics.

Descriptive statistics conclude 676 unique songs from top 200 charts for six months period. As top 200 charts for six months period includes 36200 song listings, variety of songs for Turkish market is limited. In second step of descriptive analysis, frequencies of various audio features related to song data set is included. Higher values for danceability, energy, loudness are concluded, while lower values for speechiness, acousticness and instrumentalness and liveness are found. It is also found that there are negative correlations between energy/acousticness, loudness/acousticness. Therefore, the musical preference of Turkish market signals main polarities between two segments of audio features.

Clustering approach is employed for clustering of song dataset by audio features. Conclusion of X-means clustering indicates that there are three different song clusters in Turkish market and main significant features for clustering are instrumentalness, acousticness, speechiness and mode. Cluster 0 has 484 songs and reflects low acousticness values and higher energy and danceability songs. On the other hand, Cluster 1 has 178 songs has higher acousticness values and lower mode values. Cluster 1 can be interpreted as "emotion-based" songs. Lastly, Cluster 2 has only 14 songs and they have significantly higher instrumentalness values.

Decision tree approach uses not only audio features, but also popularity scores for classification of songs. According to results, loudness is the main significant classifier for songs regarding to popularity scores. Songs are divided into two parts mainly -6.5 loudness level. Second significant classifier refers to energy in decision tree. Greater equal values for -6.5 loudness level, lower values than 0.64 energy level and lower values than 224 seconds are best combination for most popular songs in song data set.

Theoretical implication of this study mainly refers to better understanding of musical marketing concept with audio features topic which has only a few studies. This study contributes to current understanding of music marketing by implementing audio-features perspective to local market. Managerial implication of the study refers to a detailed perspective for decision making regarding to musical elements. Deciding a sponsorship agreement with singer includes various decision conditions like view counts, social media statistics etc. This study indicates that markets have common preferences for music market and this insight can provide a wider perspective for businesses. Companies can use macro trends for target market and evaluate possible decisions with these trends.

5. FUTURE RESEARCH DIRECTIONS

Future research directions of this study indicate two main areas; scope of market, extension of features. Scope of market for this study refers to popular songs of Turkish market. Including various countries or sub-markets into analysis could result in several contribution since variety in global market contains significant conclusions. Therefore, first research direction signals adding other countries, regions or markets like worldwide charts, Asian countries charts, Korean songs in western culture. Extension of features refer to adding additional music related elements for evaluating musical marketing. For example, this study focuses on audio features and popularity positions, but new studies can include lyrics of songs, social media promotion elements etc. to provide better understanding about music market popularity.

REFERENCES

Bache, S.M. & Wickham, H. (2014). magrittr: A Forward-Pipe Operator for R. R package version 1.5. https://CRAN.R-project.org/package=magrittr

Bartmanski, D., & Woodward, I. (2015). The vinyl: The analogue medium in the age of digital reproduction. Journal of consumer culture, 15(1), 3-27.

Borja, K., Dieringer, S., & Daw, J. (2015). The effect of music streaming services on music piracy among college students. Computers in Human Behavior, 45, 69-76.

Bruner, G. C. (1990). Music, mood, and marketing. Journal of marketing, 54(4), 94-104.

Chang, W. (2019). Multiple graphs on one page (ggplot2). Retrieved from http://www.cookbook-r.com/Graphs/Multiple_graphs_on_one_page_(ggplot2)/

Cockrill, A., Sullivan, M., & Norbury, H. L. (2011). Music consumption: Lifestyle choice or addiction. Journal of Retailing and Consumer Services, 18(2), 160-166.

Cooke, D. (1962). The Language of Music, London:Oxford University Press.

Delsing, M. J., Ter Bogt, T. F., Engels, R. C., & Meeus, W. H. (2008). Adolescents' music preferences and personality characteristics. European Journal of Personality: Published for the European Association of Personality Psychology, 22(2), 109-130.

Dragulescu, A.A., & Arendt, C. (2018). xlsx: Read, Write, Format Excel 2007 and Excel 97/2000/XP/2003 Files. R package version 0.6.1. https://CRAN.R-project.org/package=xlsx

Friedl, M. A., & Brodley, C. E. (1997). Decision tree classification of land cover from remotely sensed data. Remote sensing of environment, 61(3), 399-409.

Greenberg, D. M., Kosinski, M., Stillwell, D. J., Monteiro, B. L., Levitin, D. J., & Rentfrow, P. J. (2016). The song is you: Preferences for musical attribute dimensions reflect personality. Social Psychological and Personality Science, 7(6), 597-605.

Henry, L., & Wickham, H. (2019). purrr: Functional Programming Tools. R package version 0.3.2. https://CRAN.R-project.org/package=purrr

IFPI. (2019). IFPI Global Music Report 2019. Retrieved from https://ifpi.org/news/IFPI-GLOBAL-MUSIC-REPORT-2019

Larsen, G., Lawson, R., & Todd, S. (2010). The symbolic consumption of music. Journal of Marketing Management, 26(7-8), 671-685.

Magaudda, P. (2011). When materiality 'bites back': Digital music consumption practices in the age of dematerialization. Journal of Consumer Culture, 11(1), 15-36.

Mierswa, I., Wurst, M., Klinkenberg, R., Scholz, M., & Euler, T. (2006, August). Yale: Rapid prototyping for complex data mining tasks. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 935-940). ACM.

Milborrow, S. (2019). rpart.plot: Plot 'rpart' Models: An Enhanced Version of 'plot.rpart'. R package version 3.0.7. https://CRAN.R-project.org/package=rpart.plot

Müller, K. & Wickham, H. (2019). tibble: Simple Data Frames. R package version 2.1.3. https://CRAN.R-project.org/package=tibble

Oklap, C. (2018). Top Songs on Spotify: What makes them popular?. Retrieved from https://www.kaggle.com/cihanoklap/top-songs-on-spotify-what-makes-them-popular

Pelleg, D., & Moore, A. W. (2000, June). X-means: Extending k-means with efficient estimation of the number of clusters. In Icml (Vol. 1, pp. 727-734).

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: the structure and personality correlates of music preferences. Journal of personality and social psychology, 84(6), 1236.

Spotify (2019, April 29). Spotify Reports First Quarter 2019 Earnings. Retrieved from https://newsroom.spotify.com/2019-04-29/spotify-reports-first-quarter-2019-earnings/

Spotify (2019b). Company Info. Retrieved from https://newsroom.spotify.com/company-info/

Spotify (2019c). Get Audio Features for a Track. Retrieved from https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/

Spotify (2019d). Features - Spotify for Developers. Retrieved from https://developer.spotify.com/discover/

Spotify (2019e). Get a Track - Spotify for Developers. Retrieved from https://developer.spotify.com/documentation/web-api/reference/tracks/get-track/

Therneau, T., & Atkinson, B. (2019). rpart: Recursive Partitioning and Regression Trees. R package version 4.1-15. https://CRAN.R-project.org/package=rpart

Thompson, C., Parry, J., Phipps, D. & Wolff, T. (2019). spotifyr: R Wrapper for the 'Spotify' Web API. R package version 2.1.1. https://CRAN.R-project.org/package=spotifyr

Wei, T & Simko, V. (2017). R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). Available from https://github.com/taiyun/corrplot

Weijters, B., Goedertier, F., & Verstreken, S. (2014). Online music consumption in today’s technological context: Putting the influence of ethics in perspective. Journal of Business Ethics, 124(4), 537-550.

Wickham, H. (2016). ggplot2: elegant graphics for data analysis. Springer.

Wickham, H. (2019). rvest: Easily Harvest (Scrape) Web Pages. R package version 0.3.3. https://CRAN.R-project.org/package=rvest

Wickham, H., François, R., Henry, L., & Müller, K. (2019). dplyr: A Grammar of Data Manipulation. R package version 0.8.0.1. https://CRAN.R-project.org/package=dplyr

Wlömert, N., & Papies, D. (2016). On-demand streaming services and music industry revenues—Insights from Spotify's market entry. International Journal of Research in Marketing, 33(2), 314-327.

Zettl, H. (1973). Sight, Sound, Motion; Applied Media Aesthetics. Belmont, CA: Wadsworth Publishing Co., Inc.