MEASURING THE SUCCESS OF MUSICAL PROJECTS IN UKRAINE

by

Eduard Pekach

A thesis submitted in partial fulfillment of the requirements for the degree of

MA in Business and Financial Economics

Kyiv School of Economics

2020

Thesis Supervisor: ________________________________ Professor Tymofiy Brik

Approved by
Head of the KSE Defense Committee, Professor [Type surname, name]

Date ______________________________________

ACKNOWLEDGMENTS

The author wishes to express gratitude to his Thesis Advisor Thymofii Brik, EA20, aspacially to Dmytro Taranenko, Anna Harus, Misha Khaletsky, Aleksandra Serebryannikova, Oleg Pekach, and the whole family for the huge support and inspiring.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

IFPI Representing the recording industry worldwide

GDP Gross Domestic Product

USD United States Dollar

UAH Ukrainian Hryvnia

P2P Pear to Pear

URI Uniform Resource Identifier

GFK Growth From Knowledge

CHAPTER 1. INTRODUCTION

Artists, particularly musicians, might work on their projects for years, investing resources in production. Yet, only a fraction of them become successful. Social media has helped many artists to find their audiences and profit from their art. As we can see on the Figure 1, for the last five years, global recorded music industry revenue has increased by 44.29% (IFPI, 2019). The main reason for this boost was a significant growth of streaming platforms such as Deezer, Spotify, Apple Music. Only in 2019, total streaming revenue jumped by $2.2 billion. This growth is correlated with a decrease in Physical purchases and Downloads.

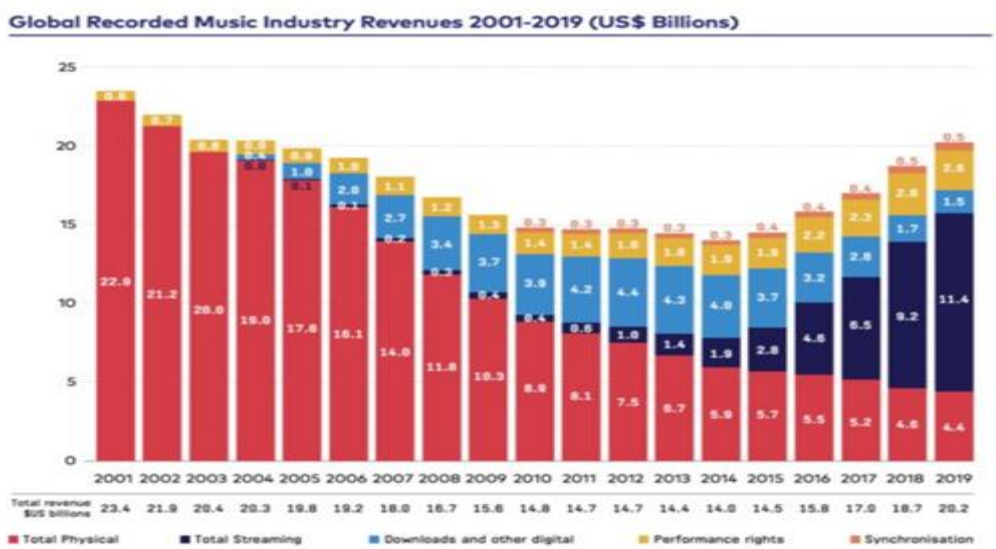

Figure 1. Global Recorded Music Industry Revenues 2001-2019 (USD Billions)[1]

As worldwide revenue grows rapidly, the growth rates of Ukrainian markets are also demonstrating the same trend as the worldwide market overall. According to Statista,

[1] "*Representing the recording industry worldwide report.*" IFPI 4th May 2020: https://www.ifpi.org/wp-content/uploads/2020/07/Global_Music_Report-the_Industry_in_2019-en.pdf

in 2020 expected music streaming revenue is US$16 mln. It is 9.4% higher than it was in 2019 ("*Music Streaming - Ukraine: Statista Market Forecast.*" Statista, July 2020. https://www.statista.com/outlook/209/338/music-streaming/ukraine.). At the same time, Ukrainian artists still struggle financially. According to the Study of the Music Market of Ukraine, more than 60% of artists do not consider music as the main source of income (*Research of the Music Market of Ukraine and Its Foreign Prospects*. Kyiv: Soundbuzz, 2020.). Thus, most of the Ukrainian artists do not take full advantage of the growing streaming industry. Why are some musical products more successful on streaming services than others? This is the central question of my thesis. In what follows, I argue that particular features of music products (music style, way of performance, genre) make them more popular among audiences. Although I do not study the commercial success of artists directly, I use the popularity of their music as the proxy of success.

The empirical part of the paper analyzes data from Spotify to see how such features as tempo, the danceability of song, acousticness, presence of explicit content affect popularity. As Spotify plays an increasingly important role in the way audiences discover music, it is valuable to analyze how the platform organizes and links artists because this information is vital to a musician's potential success (Donker, Silvia. *Networking data. A network analysis of Spotify's socio-technical related artist network*, April 2019.). Spotify provides a variable of popularity with a value between 1 and 100 (100 - the most popular artist). I employ this information to generate a binary variable with 1 for the artist that is considered as popular, and 0 - that is not. Spotify has its relevant metrics of popularity (success).

The data suggest that the most popular songs are the ones that are recorder in the genre of Pop, Rock, and all their directions such as R'n'B, metal, and punk rock. Such songs have to include minimum acoustic and instrumental features but include higher tempo, rhythm stability, beat strength, and overall regularity and are suitable for dancing together with presence of explicit content.

In terms of academic relevance, my thesis contributes to the growing literature on performance and success in arts (Yucesoy, B., & Barabási, 2016; Fraiberger et al., 2018; Lebuda & Karwowski, 2016). Moreover, to the best of my knowledge, there is no quantitative data-driven study of performance in Ukrainian music. Therefore, I provide the first data-driven account of the popularity of Ukrainian bands and variables that predict their popularity on online streaming platforms. Considering practical relevance, music is durable information good that can bring utility to listeners and value to artists (Renn Jing, and Robert Kauffman. *Understanding music track popularity in a social network*, June 5, 2017.). My work is aimed to provide practical knowledge of how Ukrainian artists should plan their activities to gain popularity among Ukrainian industry. Considering the large picture, this project is also relevant for the development of Ukrainian artistic culture. Culture is one of the pillars of national security especially due to the long dominance of the Russian cultural space on the territory of Ukraine. My model provides knowledge on how to make Ukrainian music more popular and thus more resilient to foreign cultural dominance.

CHAPTER 2. INDUSTRY OVERVIEW AND RELATED STUDIES

2.1 Industry review

According to the State Statistical Service, the Ukrainian cultural sector accounted for 4% of Ukrainian GDP in 2019. The employment in this sector accounted for 3% of overall employment in the country, which is an average indicator in Europe. The amount of foreign trade that is generated by the cultural sector is 3.6 billion USD (Economic attractiveness of Ukrainian culture. Analytical report, 2019. http://www.ier.com.ua/ua/projects?pid=6219.).

The Ukrainian music industry, despite its skyrocketing development in music supply for the last 5 years constitutes a relatively small stake in these numbers. Until today, the industry remains underdeveloped in terms of profitability. The main obstacle that hinders the industry development is the problem of musical content piracy exercised by the majority of listeners. Industry estimates show that at an average Ukrainian music project 200 million UAH is spent on video advertising every year, including 70 million UAH on legal services and 130 million is on illegal ones Varenytsya, Sasha. "(*Plan of the Reform: What Have to Be Changed in Ukrainian Music Industry.*" Karabas Live, August 5, 2019. https://karabas.live/music-reform/.).

The industry is characterized by a huge gap in the popularity of pop-music artists and so-called underground music artists. The main reason for this is unequal access to listeners on the market. The whole industry is monopolized by 10 labels that promote their music projects, which mostly play pop-music. In addition to the possession of large financial opportunities, they also have access to the listeners through various channels, such as radio and TV, since most of these labels have direct access or even ownership in these channels. All this undermines fair competition on the market (*Research of the Music Market of Ukraine and Its Foreign Prospects*. Kyiv: Soundbuzz, 2020).

The quotes on the Ukrainian language on the radio and TV contribute to a gradual improvement in the situation (*Results of the monitoring of the language quotes on the radio and television*, 2020. https://www.nrada.gov.ua/wp-content/uploads/2020/07/monitoryng-movnyh-kvot-radio-i-tb_1pivr-2020.pdf.). A drastically increased demand for Ukrainian artists and the ban of broadcasting artists from the aggressor-state Russian Federation lead to the emergence of a new cohort of high-quality Ukrainian artists as well as the revival of already forgotten ones. These new artists are not restricting their activities only to the Ukrainian market but are also actively promoting their art on the foreign market. This, in turn, promotes Ukrainian culture in the world. The process of rapid development is observed not only in the traditional pop-music direction but also almost in all other families of music. Considering these developments, for every Ukrainian artist, both experienced and new ones, the question of what the market needs and expects has become extremely relevant.

2.2 Studies of art: art as a part of a culture

The research on creativity and productivity is vast. There is a long history of studies of cultural taste and cultural consumption in sociology (Goldthorpe, 2007; Lizardo, 2016). These studies usually investigate the attitudes and behavior of individuals to understand their preferences in music or art. Academic scholars argue that such preferences are pivotal for the formation of social ties including friends and partners. Moreover, the distribution of cultural practices widens social inequalities. At the same time, researchers of marketing and consumption argue that personal cultural preferences are important for actual cultural practices (which book to buy, which movie to see, which music to listen to), and therefore the success of a certain band can be explained by the predominant consumer culture in a society.

2.3 Studies of artistic performance: data-driven approaches

Simultaneously, artistic success is addressed by another field of social science which is preoccupied with collaboration and performance of groups. Scholars of social networks, businesses, and organizations study who some groups or projects are successful, and why others fail. For example, these scholars address co-authorship in scientific papers, collaboration in business, or collaboration in artistic projects (Hofstra et al., 2020; Sonnenwald, 2007).

This latter steam of the literature is particularly important for studies of performance and popularity, which is the core facet of my research. Such studies address the success of brands as well as sports and cultural performance as related to their appeal to consumers and targeted audiences (e.g. Zagovora et al, 2018). For instance, Yucesoy and Barabasi (2015) found that a tennis player's successful performance in tournaments can accurately predict an athlete's popularity, both during a player's career and after retirement. In another study, Ren and Kauffman (2017) applied machine-generated music semantics constructs and analyzing data on the performance of music tracks on the ranking charts. Their model determined the popularity and duration of music charts.

The closest research to mine was conducted about the publishing industry by Wang et al. (2019). The authors found that a strong driving factor of book sales across all genres is the publishing house. They also found that for some genres, the publishing history of an author (as measured by previous book sales) is more important, while in others, the author's visibility is crucial. However, the literature on the music industry with similar research questions is scarce.

2.4 How to measure popularity: strategies and definitions

There is a wide range of papers providing different definitions of popularity and respective methods of estimations of the potential popularity of musical projects (Barabási, 2018). There is no one accurate definition of popularity and each scholar that investigates this topic suggests his/her approach in defining popularity. Nunes and Ordanini (2016) assess relative popularity using rankings on Billboard's Hot 100 singles chart and by comparing those songs that reached number 1 (hereafter Top songs) with a comparison set of those songs that, while making it onto the Hot 100, never climbed above number 90 (hereafter

Bottom songs). In their work, they estimate how different instrumental features of the song affect its potential popularity. Singhi and Brown (2015) use a similar method for estimating popularity. They define hits as songs that made it to the Billboard Year-End Hot 100 singles chart in the years 2008-2013 and they analyze the effect of lyrics features of the song on its popularity. Dhanaraj and Logan (2005) combine two methods mentioned above by extracting very general acoustic and lyric-based features from songs then use standard classifiers to separate hits from non-hits. Borg and Hokkanen (2011) measure the popularity of songs by views on youtube which were regressed on music features such as ehonest, loudness, tempo, hotness, danceability. Schedl et al. (201) used social network analysis to measure the degree of "hotness" around a band. They tried to define popularity by different sources: last.FM playcounts, Twitter Posts, shared folders in the P2P network, traditional charts. All of them reflect different customer communities, availability of information, noisiness, time dependence. Donker (2015) conducted a case study of Spotify's related artist network Noisia to determine different powerful artists on Spotify. This social network analysis uncovers how each actor is embedded in networked structures of relationships on Spotify.

2.5 Hypotheses

Drawing on the existing literature, one could say that popularity is affected by various factors such as technical ones, genre, and explicit content. More specifically, this paper provides the following hypotheses.

Ren and Kauffman (2017) found a significant relationship between music popularity and Pop-genre. Therefore, I will test whether the Pop music genre has the highest effect on popularity out of all genres.

Previous studies suggest that resources and infrastructure are pivotal for artistic performance (Barabási, 2018). Therefore, I assume that the origin is one of the important determinants of the popularity of Ukrainian bands. (Hudson, Ray. *Regions and place: Music,* *identity, and place*, October 2006.). Especially if the origin is the Kyiv region due to large opportunities of realization musical potential.

CHAPTER 3. METHODOLOGY

This research consists of the following stages: web scraping the data, creating depending variables, and estimating the model.

3.1 Web scraping.

Firstly, I collected a list of Ukrainian artists, their origin, genre, and year of beginning career from the “Liroom" webpage using the "rvest" R package. "Liroom" is the largest website with a list of Ukrainian bands. Then, I identified the Spotify URI of all bands from this website manually. Using URIs I scraped the data from Spotify with the “spotipy” Python package. These data include such variables as explicit, acousticness, danceability, energy, instrumentalness, liveness, loudness, speechiness, tempo, valence, popularity, and fans. These variables are described in the section "Data and variables".

3.2 Dependent variables

The main dependent variable of this research is the *success* of a song on the streaming platform. This is the dummy variable where 1 means success and 0 means otherwise. Popularity is defined by the number of followers of each artist. In the same vein, I assume that artists who have a popularity coefficient of more than 20 are considered a success. This threshold is based on the exploration analysis of the variable of popularity. As we can observe on the histogram below (Figure 2), the majority of the artists are located below this threshold.

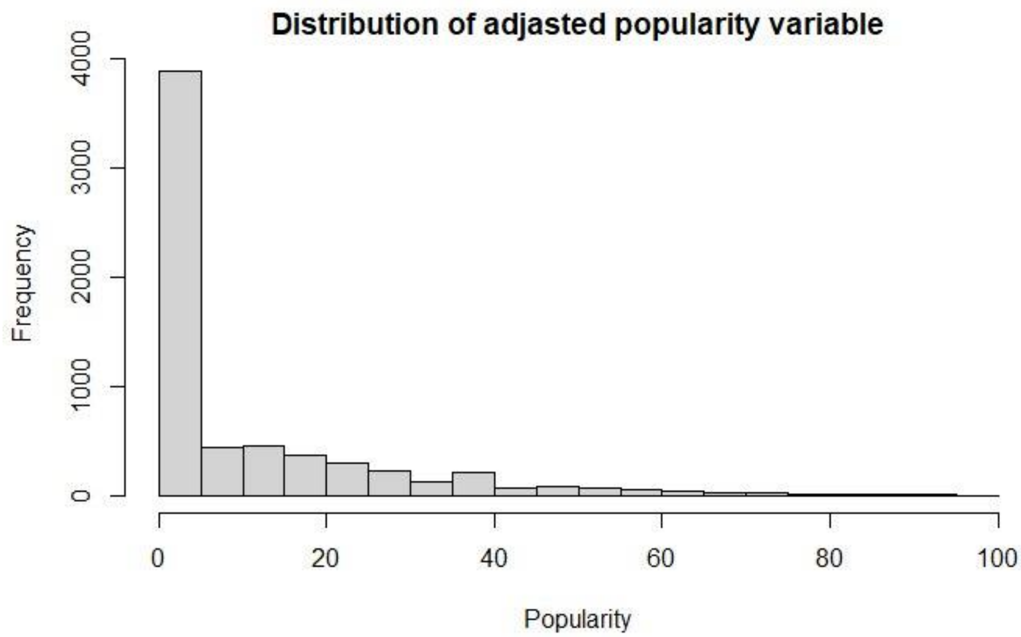

Figure 2. Histogram of distribution of the popularity indicator

The distribution of this variable is showing that overwhelming majority of Ukrainian artist are not considered to be successful. Only sixth part of total dataset have the success value of 1, as it can be seen on the Figure 3:

Figure 3. Distribution of the Success dependent variable values.

3.3 Model.

Taking into account the nature of my data dependent variables (i.e., adjusted popularity), I apply logistic regression in order to quantify the sign and marginal effect of the music features. The estimation is conducted for two models – logit (1) and probit (2):

$$\Lambda(x'\beta) = \beta_0 + \beta_1 explicit + \beta_2 acousticness + \beta_3 danceability + \beta_4 energy + \beta_5 instrumentalness + \beta_6 liveness + \beta_7 loudness + \beta_8 speechiness + \beta_9 tempo + \beta_{10} valence + \beta_{11} City + \beta_{12} Genre + \hat{u} \quad (1)$$

$$\Phi(x'\beta) = \beta_0 + \beta_1 explicit + \beta_2 acousticness + \beta_3 danceability + \beta_4 energy + \beta_5 instrumentalness + \beta_6 liveness + \beta_7 loudness + \beta_8 speechiness + \beta_9 tempo + \beta_{10} valence + \beta_{11} City + \beta_{12} Genre + \hat{u} \quad (2)$$

The marginal effects were calculated for the logit (3) and probit (4) models as following:

$$\frac{\partial p}{\partial x_j} = \frac{e^{x'\beta}}{(1+e^{x'\beta})^2}\beta_j \quad (3)$$

$$\frac{\partial p}{\partial x_j} = \Phi(x'\beta)\beta_j \quad (4)$$

To estimate what model is more appropriate in my case I apply the Mcfadden test which is about calculation pseudo-R squared of logistic regression (5):

$$R^2_{McFadden} = 1 - \frac{\log(L_c)}{\log(L_{null})} \quad (5)$$

Where L_c is the maximized likelihood value and L_{null} is a value for the null model (only with an intercept).

In my case, McFadden's pseudo R squared for logit is 0.197518 and for probit - 0.1975372 which is pretty the same. For both model $R^2_{McFadden}$ is enough small to say that our models have good predictive ability. Thus, further, I use a logit model for estimation.

CHAPTER 4. DATA

The total number of Ukrainian music bands that were scrapped from Liroom was 545. The number of bands that did not have any content on Spotify is 301 which resulted in missing values for this artist in the final dataset. After matching Liroom data with the analogous data from Spotify, the resulting number of songs have reached 6,445.

The final dataset that is employed for the analysis consisted of 6,445 observations (i.e., songs) from 244 Ukrainian music artists. In what follows, I describe the variables that are used in the analysis.

4.1 Independent variables

Acousticness - A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic. ("*Get Audio Features for a Track*." Spotify for Developers, https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/.) The distribution of values for this feature showed that most songs of Ukrainian artists have quite low acoustic in their songs.

Danceability - describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable. ("*Get Audio Features for a Track*." Spotify for Developers, https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/.) The distribution of values for this feature showed quite normal distribution which is implied that most songs are suitable for dancing.

Energy - is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy. ("*Get Audio Features for a Track*." Spotify for Developers, https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/.) The distribution of values for this feature demonstrates that most songs have quite a high energy indicator which implies their high intensity.

Instrumentalness - Predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0. ("*Get Audio Features for a Track*." Spotify for Developers, https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/.) The distribution of values for this feature for Ukrainian artist shows that most of their songs are vocal but not instrumental.

Liveness - Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides a strong likelihood that the track is live. ("Get Audio Features for a Track." Spotify for Developers, https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/.) The distribution of values for this feature shows that most songs were recorded without audience participation.

Loudness - The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing the relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 dB. ("*Get Audio Features for a Track*." Spotify for Developers, https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/.) The distribution of values for this feature that most songs are quite loud.

Speechiness - detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audiobook, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks. ("*Get Audio Features for a Track*." Spotify for Developers, https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/.) The distribution of values for this feature shows the low presence of speech elements in investigated songs.

Valence - A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry). ("*Get Audio Features for a Track*." Spotify for Developers, https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/.) The distribution of values for this feature shows that more artists prefer positive songs than negative ones.

Tempo - The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, the tempo is the speed or pace of a given piece and derives directly from the average beat duration. ("*Get Audio Features for a Track*." Spotify for Developers, https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/.) The distribution of values for this feature shows relatively normal distributions of BPM for investigated songs.

Followers_sp - Information about the followers of the artist on Spotify. The distribution for this feature shows that most Ukrainian artists have on average 30 000 followers on their account on Spotify.

Adj_pop - An adjusted value of popularity for Ukrainian artist, to investigate the level of their popularity among the Ukrainian audience.

Explicit – represent songs with explicit content.

Region – dummy variable that is showing the originated region for each artist from final dataset. Distribution of this variable is depicted in Figure 4

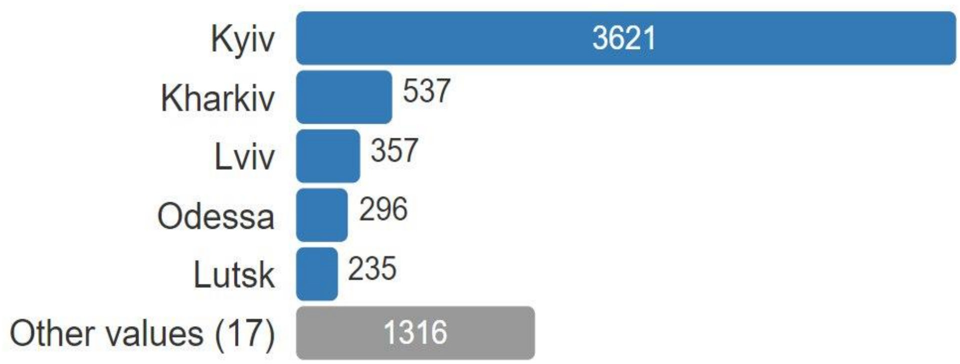

Figure 4. Distribution of the region of origin of the investigated artists

As one may observe more than half of all artists are originated from Kyiv region, which is larger than all other region together, and it could be explained by the fact that Kyiv simply has much more possibilities for cultural development due to its status as capital and center of music industry of Ukraine.

Genre – genre of the particular artist. There are following genres used in the analysis: acoustic, alternative, classic, electronica, folk, funk, hip-hop, house, indie, jazz, metal, pop, punk rock, r'n'b, reggie, rock and soul. The distribution of the variable provided in Figure 5 shows that the most popular genres among the artist are pop, rock and electronica, while the least popular are acoustic, house and soul.

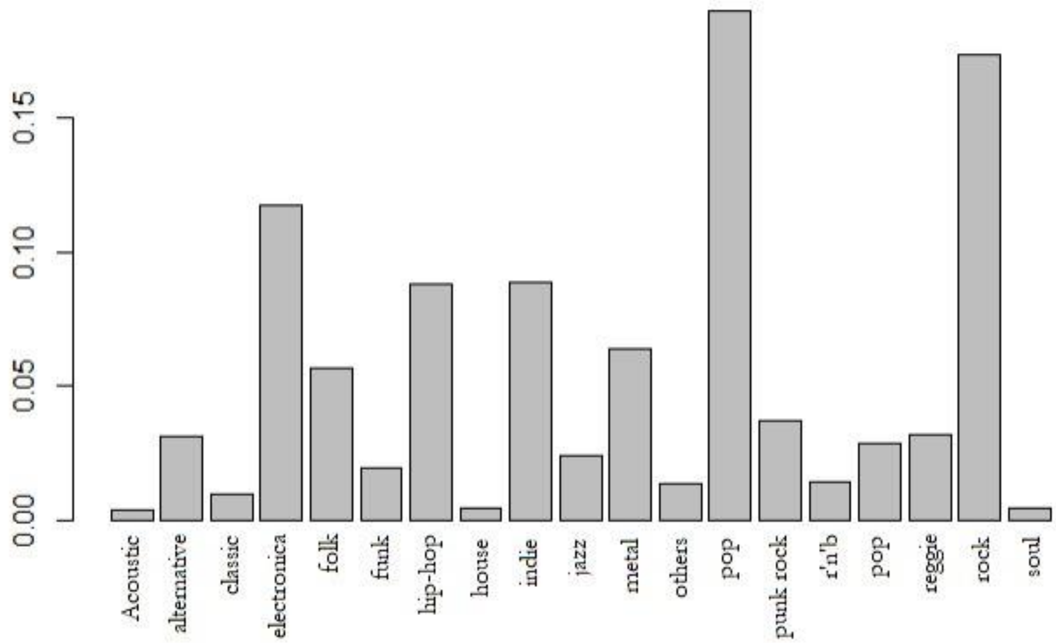

Figure 5. Distribution of variable genre

The distributions of the mentioned variables are displayed below in Figure 6.

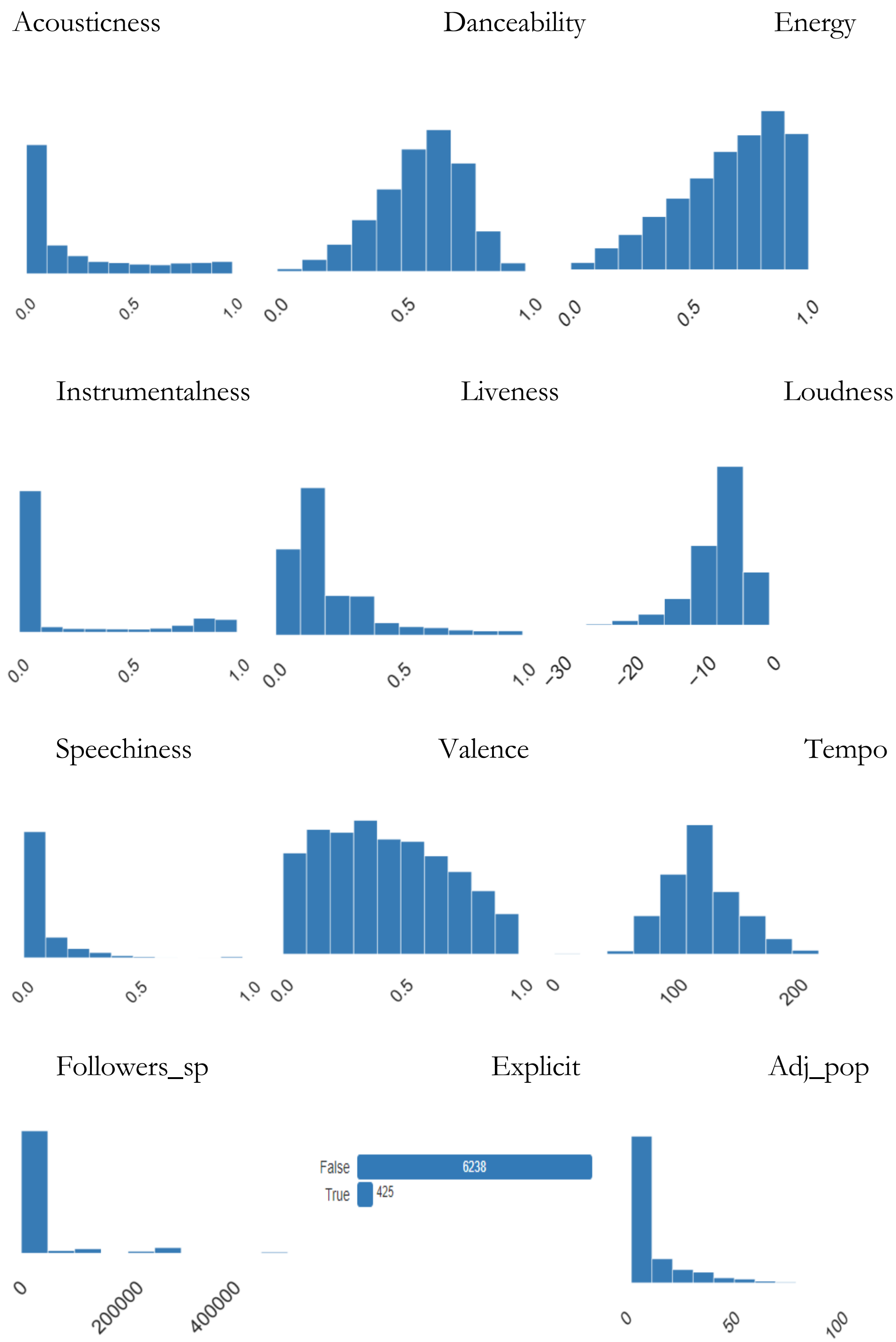

Figure 6. Distributions of the applied features

Descriptive statistics of all numerical variables are provided below in table 1.

Table 1. Descriptive statistics of applied features

Variable	Min	Median	Mean	Max	N
followers_sp	1	2015	32170	548391	6445
acousticness	0.0000011	0.0900000	0.2414136	0.9950000	6445
energy	0.000199	0.712000	0.672007	1.000000	6445
liveness	0.0000	0.1330	0.2138	0.9920	6445
loudness	-37.085	-6.855	-7.770	-0.773	6445
popularity	1.000	2.000	6.388	57.000	6445
adj_pop	2.00	4.00	11.36	100.00	6445

Figure 7 below depicted the correlations among the variables used.

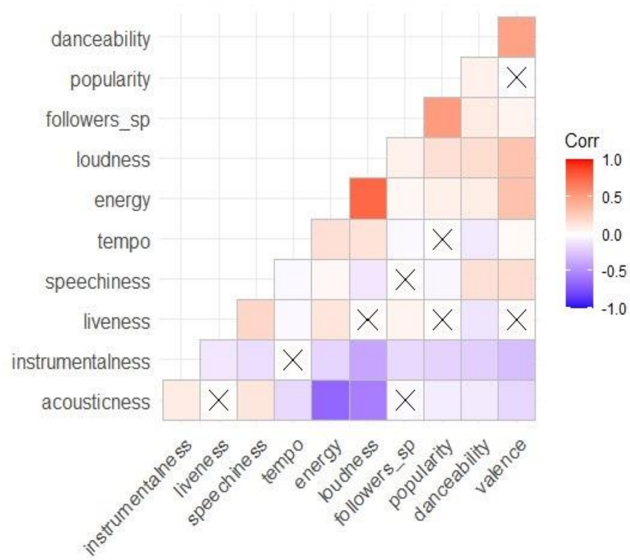

Figure 7. Correlation matrix

As one can observe the variable loudness is highly positively correlated with the variable energy, which can mean that loud songs are more likely to be quite energetic, while the correlation with instrumentalness is highly negatively correlated which means that such songs are very unlikely to be acoustic. Popularity has a high correlation with the number of followers which is quite logical, the more fans you have the more popular you are. Valence has a high correlation with danceability which implies that more dance songs have in general a more positive sound.

CHAPTER 5. RESULTS

5.1. Logit regression model

For the Logit model, the estimation result is provided in Table 2.

Table 2. Logit model estimation

Variable	Estimate	P-value	Variable	Estimate	P-value
Intercept	-1.782	**0.008**	Genre_Pop	1.081	**0.027**
explicit	0.354	**0.017**	Genre_Punk_Rock	0.985	0.084
acousticness	-0.590	**0.001**	Genre_rnb	3.609	**0.000**
danceability	1.009	**0.000**	Genre_Rap	0.859	0.114
energy	-0.362	0.261	Genre_Reggi	0.590	0.267
instrumentalness	-0.939	**0.000**	Genre_Rock	0.666	0.181
liveness	0.476	**0.016**	City_Chernihiv	3.438	**0.000**
loudness	0.106	**0.000**	City_Chernivtsi	-1.346	0.192
speechiness	-1.065	**0.007**	City_Dnipro	-3.198	**0.002**
tempo	0.003	**0.009**	City_Donetsk	-2.025	**0.006**
valence	-1.253	**0.000**	City_Ivano_Frankivsk	-3.488	**0.000**
Genre_alternative	-1.054	0.104	City_Kharkiv	0.007	0.973
Genre_Classic	-2.346	**0.037**	City_Khmelnytsky	-2.351	**0.000**
Genre_electronica	-0.241	0.636	City_Kyiv	0.259	0.068
Genre_Folk	-0.332	0.531	City_Luhansk	-15.588	0.972
Genre_Funk	1.944	**0.000**	City_Lutsk	-1.245	**0.000**
Genre_hip_hop	0.798	0.124	City_Lviv	0.253	0.277
Genre_house	-14.417	0.973	City_Odessa	-0.830	**0.004**
Genre_Indie	0.958	0.053	City_Poltava	0.155	0.563
Genre_Jazz	-2.395	**0.032**	City_Sumy	-15.126	0.972
Genre_Metal	1.026	**0.046**	City_Symferopol	-15.767	0.984
Genre_Other	-1.281	0.094	City_Ternopyl	1.320	**0.000**

Based on estimation results we can observe that our hypothesis is supported in the context of genre attributes of the most popular artist. The hypothesis about the importance of the origin of the artist is not fully supported by our results due to fact that the variable of the Kyiv region has a quite high P-value, so we are not able to report a particular effect and the sign of this effect of the variable for Kyiv region for overall popularity.

5.2. Probit regression model

The results for the probit estimation model is given in the table 3:

Table 3.

Variable	Estimate	P-value	Variable	Estimate	P-value
1	2	3	4	5	6
Intercept	-0.000	**0.003**	Genre_Pop	0.000	**0.013**
explicit	0.207	**0.017**	Genre_Punk_Rock	0.588	0.054
acousticness	-0.325	**0.001**	Genre_rnb	2.15	**0.000**
danceability	0.555	**0.001**	Genre_Rap	0.511	0.079
energy	-0.214	0.237	Genre_Reggi	0.346	0.219
instrumentalness	-0.494	**0.000**	Genre_Rock	0.374	0.151
liveness	2.55	**0.027**	City_Chernihiv	1.99	**0.000**
loudness	0.059	**0.000**	City_Chernivtsi	-0.563	0.190
speechiness	-0.641	**0.004**	City_Dnipro	-1.34	**0.000**
tempo	0.000	**0.012**	City_Donetsk	-0.846	**0.004**
valence	-0.697	**0.000**	City_Ivano_Frankivsk	-1.70	**0.000**
Genre_alternative	-0.480	0.133	City_Kharkiv	-0.001	0.906
Genre_Classic	-1.09	**0.022**	City_Khmelnytsky	-1.07	**0.000**
Genre_electronica	-0.150	0.572	City_Kyiv	0.153	0.063
Genre_Folk	-0.174	0.529	City_Luhansk	-4.67	0.947

1	2	3	4	5	6
Genre_Funk	1.15	**0.000**	City_Lutsk	-0.653	**0.000**
Genre_hip_hop	0.473	0.083	City_Lviv	0.130	0.321
Genre_house	-4.06	0.951	City_Odessa	-0.422	**0.005**
Genre_Indie	0.559	**0.031**	City_Poltava	0.115	0.469
Genre_Jazz	-0.832	**0.039**	City_Sumy	-4.48	0.947
Genre_Metal	0.620	**0.022**	City_Symferopol	-4.77	0.969
Genre_Other	-0.584	0.109	City_Ternopyl	0.800	**0.000**

The probit model results are very similar to the logit one. The signs and significance of the investigated variables are the same as in the previous model for logit. To interpret coefficients, I look at the marginal effect of the logit model based on the result of the McFadden test. The marginal effect of the probit model is given in Table 5.

5.3. Marginal effect for the logit model

Table 4 represents the estimation results of the marginal effect for the logit model. Based on these results we can make the following interpretation. Among the technical features, the largest positive effect of the overall popularity has variables such as "danceability" and “liveness”. In other words, I can say that music with a high tempo, rhythm stability, and beat strength increase the probability that a song will become popular.

Table 4. Marginal effect (logit model)

Variable	(Intercept)	explicit	acousticness	loudness
Marginal effect	-0.223	0.044	-0.074	0.013
Variable	speechiness	tempo	liveness	energy
Marginal effect	-0.133	0.000	0.06	-0.045
Variable	instrumentalness	danceability	valence	Genre_alternative
Marginal effect	-0.118	0.126	-0.157	-0.132

Variable	Genre_Classic	Genre_electronica	Genre_Folk	Genre_Funk
Marginal effect	-0.294	-0.030	-0.042	0.244
Variable	Genre_hip_hop	Genre_Other	Genre_Pop	Genre_Punk_Rock
Marginal effect	0.100	-0.161	0.135	0.123
Variable	Genre_house	Genre_Indie	Genre_Jazz	Genre_rnb
Marginal effect	-1.807	0.120	-0.300	0.452
Variable	Genre_Rap	Genre_Reggi	Genre_Metal	Genre_Rock
Marginal effect	0.108	0.074	0.129	0.083
Variable	City_Chernihiv	City_Chernivtsi	City_Dnipro	City_Kyiv
Marginal effect	0.431	-0.169	-0.401	0.032
Variable	City_Luhansk	City_Lutsk	City_Sumy	City_Symferopol
Marginal effect	-1.954	-0.156	-1.896	-1.976
Variable	City_Ternopyl	City_Donetsk	City_Ivano_Frankivsk	City_Kharkiv
Marginal effect	0.165	-0.254	-0.437	0.001
Variable	City_Lviv	City_Odessa	City_Poltava	City_Khmelnytsky
Marginal effect	0.031	-0.104	0.019	-0.295

Table 5 depicts the marginal effects for probit model, and as one may observe, they are quite similar to the result of logit model, which means that the coefficients of logit link functions and probit inverse normal link function with this data produce almost equivalent cofficients.

Table 5. Marginal effect (probit model)

Variable	(Intercept)	explicit	acousticness	energy
Marginal effect	-0.231	0.045	-0.071	-0.047
Variable	danceability	liveness	loudness	speechiness
Marginal effect	0.121	0.056	0.013	-0.14
Variable	instrumentalness	tempo	valence	Genre_alternative
Marginal effect	-0.108	0.000	-0.152	-0.105
Variable	Genre_Classic	Genre_electronica	Genre_Folk	Genre_Funk
Marginal effect	-0.238	-0.033	-0.038	0.252
Variable	Genre_hip_hop	Genre_Other	Genre_Pop	Genre_Punk_Rock
Marginal effect	0.103	-0.128	0.138	0.128
Variable	Genre_house	Genre_Indie	Genre_Jazz	Genre_rnb
Marginal effect	-0.887	0.122	-0.181	0.470
Variable	Genre_Rap	Genre_Reggi	Genre_Metal	Genre_Rock
Marginal effect	0.111	0.075	0.135	0.082
Variable	City_Chernihiv	City_Chernivtsi	City_Dnipro	City_Kyiv
Marginal effect	0.435	-0.123	-0.292	0.033
Variable	City_Luhansk	City_Lutsk	City_Sumy	City_Symferopol
Marginal effect	-1.017	-0.142	-0.977	-1.039
Variable	City_Ternopyl	City_Donetsk	City_Ivano_Frankivsk	City_Kharkiv
Marginal effect	0.175	-0.185	-0.371	-0.003
Variable	City_Lviv	City_Odessa	City_Poltava	City_Khmelnytsky
Marginal effect	0.028	-0.092	0.025	-0.234

Based on the estimation result for non-technical features we can conclude that the artist that performs in the R'n'B genre has the highest chance to become successful among all genres and taking acoustic genre as a base. This genre has the highest positive marginal effect among the genres and highly statistically significant. R'n'B is commonly known as one of the directions of Pop music, so this result confirms our hypothesis about the high popularity of this genre among the listeners.

The coefficients for the regions were calculated taking Cherkasy as a base we can make the following interpretation: on average the artist that comes from Cherkasy have the highest chance to become popular. Its cause of limitation bias as a large number of young artists in my data are located in other regions, so data distributed unevenly.

CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

This paper is the first attempt to analyze the Ukrainian music industry in the context of determining the factors that have the largest effect on the overall popularity of Ukrainian bands. The analysis was conducted in the following way. First of all, the main part was to collect all the necessary data that was used in the analysis. With this goal, the list of all Ukrainian artist was parsed from the Liroom service, together with the information about their origin and genre. The next step consists form collecting data for these groups regardless of their technical and audience features. To parse such kind of data, the Spotify for Developers service was applied. With help with this service, I was able to scrape data about the relative popularity of each artist which in turn was converted to a binominal variable with two values: 1 for success and 0 – for the opposite. Such a variable was used as a dependent variable, on each the set of independent variables such as valence, loudness, a number of followers, genre, and others were regressed to define the marginal effect of each. With this goal, I applied logistic specifications to understand what features contribute the most to the overall probability of becoming successful.

The logistics models that were used in my analysis were logit and probit specifications. To estimate what specification is the most appropriate for our data, I apply the McFadden test, which is calculated goodness of predictive ability. Based on this test I chose the logit model as my base to interpret the marginal effect of each variable.

The estimation result for both models has shown almost identical results in the context of the sign and statistical significance of the regressors. In particular, all technical features, except *energy* have shown high significance and all features, except *acousticness*, *instrumentalness, speechiness, and valence* have shown a positive sign. For the last two variables, this result is quite interesting because it implies that the songs with positive filling and a large share of lyrics have fewer chances to become popular than more aggressive, melodic songs. In the context of genres the only genres that showed significant results were *classic,* *Pop,* and *R'n'B*. Moreover, the last two variables are shown to contribute to the overall success, while the first one is likely to demonstrate a negative effect.

As was mentioned above, the logit specification was used to calculate the particular marginal effect of each significant variable. In this context, the largest contribution to overall success has variable *danceability* or suitableness of the track for dancing. Such a moment could be logically explained by the number of the target audience for which these tracks were recorder because for dancing music such an audience is much larger than for rock or other genres. Also, worth noting that the variable *explicit*, which measured the presence of explicit content in the songs, has a quite large positive effect, which means that people like more scandal or, as it customary to speak, more hype songs.

Speaking about the genre's preferences of the audience it worth noting that base of the variable was an acoustic genre and regardless of these genres, the contribution of others was calculated. The most significant contributor to the overall success was *the R'n'B* genre, which almost 0.5 more popular than acoustic ones. The following genres with much less but quite a large contribution are *Pop, Punk Rock,* and *Rap,* which supports our initial hypothesis that well investigated in the literature about common more popularity of branches of Pop music, but also its worth mention high level of popularity of Rock music and all its branches such as metal, indie and even reggae.

Our second hypothesis to test was the more chances of success for artists that originated from the Kyiv region. After estimation, I was able to reject this hypothesis due to fact that the largest contribution to the common success was for artists originated from the Chernihiv region with Ternopil on the second stage.

Results of estimation are representing a high value for young Ukrainian artists due to its possibility to help write those songs that have much higher chances to become popular. Also, the created model could be used to predict the success of the new-released tracks.

The overall recommendation based on this research is to an orientation on the genres that have the largest target audience which have such genres as Pop and Rock together with all their directions like metal, Punk Rock and R'n'B. The potential songs have to be suitable for dancing and include minimum acoustic and instrumental features.

Future research on this topic may include expanding the current analysis to control for semantic features of the lyrics, to investigate what words combination contributes to the potential success and in particular in which language these songs were written. Also, it's of high importance to investigate more deeply the audience's features. One of the ways to conduct such research is to analyze the listeners' responses to the tracks, it could be done by using the data for music and video hosting such as YouTube or radio-stations and to investigate the number of likes and dislikes, their ratio, and conduct a semantic analysis of comment section for each song.

REFERENCES

Barabási, Albert-László. The Formula: "*The Five Laws Behind Why People Succeed.*" Macmillan, (2018).

Burakovsky, I. "*Економічна привабливість української культури* [*Economic attractiveness of Ukrainian culture*]". The Institute for Economic Research and Policy Consulting, (2019).

Donker, S. "*Networking data. A network analysis of Spotify's socio-technical related artist network*". International Journal of Music Business Research, 8(1), (2019): 67-101.

Fraiberger, Samuel P., Roberta Sinatra, Magnus Resch, Christoph Riedl, and Albert-László Barabási. "*Quantifying reputation and success in art.*" Science 362, no. 6416 (2018): 825-829.

Goldthorpe, John H. " *"Cultural Capital": Some Critical observations.*" Sociologica 1, no. 2 (2007): 0-0.

Hofstra, Bas, Vivek V. Kulkarni, Sebastian Munoz-Najar Galvez, Bryan He, Dan Jurafsky, and Daniel A. McFarland. "*The Diversity–Innovation Paradox in Science.*" Proceedings of the National Academy of Sciences 117, no. 17 (2020): 9284-9291.

Hudson, Ray. (2006). "*Regions and place: Music, identity and place. Progress in Human Geography*" - PROG HUM GEOGR. 30. 626-634. 10.1177/0309132506070177

"*IFPI Issues Annual Global Music Report.*" IFPI, August 3, 2020. https://www.ifpi.org/ifpi-issues-annual-global-music-report/.

Lebuda, Izabela, and Maciej Karwowski. "*Written on the Writer's Face: Facial Width-to-Height Ratio among Nominees and Laureates of the Nobel Prize in Literature.*" Creativity Research Journal 28, no. 2 (2016): 207-211.

Lizardo, Omar. "*How cultural tastes shape personal networks.*" American sociological review 71, no. 5 (2006): 778-807.

“*Music Streaming - Ukraine: Statista Market Forecast*.” Statista. Accessed September 3, 2020. https://www.statista.com/outlook/209/338/music-streaming/ukraine

Ren, Jing, and Robert John KAUFFMAN. "*Understanding music track popularity in a social network*." AIS, (2017).

Sonnenwald, Diane H. "*Scientific collaboration*." ARIST 41, no. 1 (2007): 643-681.

Schedl, Markus, Tim Pohle, Noam Koenigstein, and Peter Knees. "*What's Hot? Estimating Country-specific Artist Popularity*." In ISMIR, (2010) pp. 117-122.

Wang, Xindi, Burcu Yucesoy, Onur Varol, Tina Eliassi-Rad, and Albert-László Barabási. "*Success in books: predicting book sales before publication*." EPJ Data Science 8, no. 1 (2019): 31.

Yucesoy, Burcu, and Albert-László Barabási. "*Untangling performance from success*." EPJ Data Science 5, no. 1 (2016): 1-10.

Zagovora, Olga, Katrin Weller, Milan Janosov, Claudia Wagner, and Isabella Peters. "*What increases (social) media attention: Research impact, author prominence or title attractiveness?*." In Proceedings of 23rd International Conference on Science and Technology Indicators (STI 2018), pp. 1182-1190.

Дослідження практики споживання цифрового контенту: фільми/серіали, музика, книги [*Digital content consumption research: movies / TV series, music, books*]”, the Internet GfK Ukraine research commissioned by the European Business Association, April-May 2020.

“*Как украинская музыкальная индустрия завоевала Восточную Европу* [*How the Ukrainian music industry conquered Eastern Europe*]”. Fromua.news, 3 November, 2018. Accessed September 3, 2020. https://fromua.news/article/2576938/kak-ukrainskaya-muzikaljnaya-industriya-zavoevala-vostochnuyu-evropu/musicinua.com