UNIVERSIDAD TORCUATO DI TELLA

Master in Management + Analytics

Psychological Reactance: Background music when promoting Spotify brand and consumer behaviour

Juan Martín Leguizamón Maurette

Tutora: Elena Fumagalli

ABSTRACT

During the 21st century, the way in which music is commercialized and consumed has been completely reshaped. In light of the technological advances taking place during the last decades, music consumption has become ubiquitous, and people can now listen to their favorite artists at the same time they perform multiple activities. This translated to hundreds of millions of people listening to music every day, as well as an intense competition between digital platforms providing audio streaming services. Of these platforms, Spotify appears as one of the leading companies, with a total of 345 million active users and 155 million paid subscribers. Spotify's business model, defined as *freemium*, combines a paid service, unlimited and free of interruptions, with a free one, which offers the same functions as the paid service but with lower audio quality and ad interruptions. These announcements are formed by background music and a voice over promoting the paid service. The core motivation of the present study is to shed light upon the effects of the election of background music when promoting the Spotify brand. Is it beneficial for the company to use background music that the user isn't accustomed to listening in order to catch her attention and persuade her to go premium? Or will this approach generate Psychological Reactance (i.e., an "unpleasant motivational arousal that emerges when people experience a threat to or loss of their free behaviors", Steindl, Jonas, Traut-Mattausch, Sittenthaler & Greenber, 2015) in the consumer, undermining her willingness to go premium and her behavioral intention towards the brand? Although the literature on Reactance is vaste, there is no evidence of empirical studies testing the effectiveness of audio advertising interruptions in the context of a music streaming service. Particularly, we suggest a combined approach to detect "opposite" music genres to be used as background music in a simulacrum of an audio streaming company providing it's service. By means of a K-means clustering and an experimental pretest, we identified *Rock* and *Reggaeton* as the most different/opposite music genres as perceived by the consumers. Later, we created a fictional company and ran an experiment (N = 150) where respondents had to listen to a succession of songs from their preferred music genre to be interrupted with an ad containing either no background music, background music from the same genre, or background music from the opposite genre in the tuple. From the main

experiment, we concluded that for the *Reggaeton* sample, the assignment of people to the experimental groups had a significant effect in user's *Behavioral Intention* towards the brand (i.e., people's disposition to consider the company when subscribing to an audio streaming service) and in overall feelings of *Reactance.*

INDEX

1. Introduction

1.1. Motivation

Music as an art form and cultural activity has existed since prehistoric times. The earliest and largest collection of prehistoric musical instruments was found in China and dates back to between 7000 and 6600 BC (Wilkinson, 2000). But although music has been with us since a long time ago, it is a phenomenon in constant change. Not only new forms and types of music are constantly created, but also the way in which music is spread among people is continuously affected by technology. Even along the second half of the 21st century music commercialization has reshaped completely. If people bought their favorite artists albums in vinyl format in the 50s and 60s, these were replaced by cassette tapes between the 70s and 90s, which were later replaced by CDs by the end of the 90s. Nowadays, there is no doubt that music commercialization doesn’t occur in physical form, and that most people listen to music from audio streaming platforms like Spotify or online video-sharing platforms like Youtube.

What’s interesting is that not only technology has affected music regarding the means by which songs are listened to, but that this also permitted the capturing and storage of huge amounts of data about the way in which this music is listened to, letting us formulate questions such as: which music genres are currently the most popular? which artists are currently being listened to? which songs from these artists are the most viral?, etc. Everyday millions of human interactions with audio streaming platforms are stored, and access to this information is becoming faster and easier overtime. If a teenager in the 1960s bought a vinyl from The Beatles, at most he could access the band’s name, the songs contained in the album, the lyrics of these songs, and a couple of pictures of the group. Nowadays, one can download the Spotify app, search for The Beatles in its search bar, and access all of its songs, lyrics, know how many times the songs have been played in the platform, know how many monthly listeners the band has, look for similar artists, etc. In addition, if one knows how to use programming tools such as Python, the access to information is even greater.

The collection and fast and easy access to this type of information represents a gold mine for the companies that own such streaming platforms. Digital platforms, including Spotify, have come to dominate how music is circulated and consumed (Morris and Powers 2015, Webster 2019, Eriksson 2020). Competition between platforms is intense, as all of the leading platforms provide access to similar content at a similar price point, and they all provide basic forms of curation. Therefore, these firms need to develop innovative ways to attract and engage users (Hracs & Webster, 2020). Recorded music is ubiquitous in people's everyday lives, as people listen to music when they travel, work and perform social and recreational activities (Fuentes et. al, 2019). As such, if companies can understand what people like, and how to convince them to use their platforms to access what they like, economic benefits are endless.

This investigation develops in the conjunction of music and technology, as it is in the union of these two worlds that valuable insights for companies offering audio streaming services can be made. Using Spotify's API for Python (*Spotipy*), we will access information regarding the audio attributes that define a song: *danceability*, *energy*, *key*, *loudness*, *mode*, *speechiness*, *acousticness, instrumentalness, liveness, valence and tempo*[1]. Each song contains a numeric value for each of these attributes, so in mathematical terms each song can be represented as a vector of size 11.

Assuming that songs from the same music genre share similar audio attributes, genres could then be mapped into clusters of songs. Each cluster would represent a music genre in a 11- dimension space, and Euclidean measures between clusters centroids could then be computed. Therefore, we could select a tuple of "opposite" music genres as those with the largest distance between centroids. Knowing which tuple of genres are the most opposed, we will try to understand if the songs genre used in Spotify's ads background music (which are played when using the app for free) affect the consumer's decision to subscribe to the platform or to keep using it for free. If a consumer using Spotify for free is interrupted by an ad with background music, how does the genre of the song used in the ad influences the consumer's behaviour?

[1] The full outline of Spotify Audio Features can be found in Appendix 8.1

If the assumption that songs from the same music genre share similar audio attributes proves to be incorrect, and songs can't be mapped into different clusters representing each music genre, an alternative approach will be taken. Opposition between genres won't be defined depending on the song's audio attributes, but on people's perception of what music genres they consider similar, and which they consider the most different.

In addition to managerial relevance, the present work strives to make an academic contribution as well. After reviewing the relevant literature, no study was found that tests the effectiveness of audio advertising interruptions in the context of a music streaming service.

1.2. Spotify

Founded in 2006, Spotify is now the global music streaming service leader and also has the highest share of music streaming subscribers in the world, outperforming its competitors by some margin. With a wide variety of music content as well as audiobooks and podcasts, Spotify has established itself as the “go-to” option for many. Launched in 2008, Spotify breathed new life into the music streaming industry by providing listeners with an almost endless catalog of music to enjoy. The company went public in April 2018 and, by the end of the first day’s trading, it was already worth $26.5 billion.

As a freemium service, Spotify offers basic features for free with advertisement interruptions and limited user control whereas in its premium version (paid subscription), it offers additional features, such as offline listening and commercial-free listening.

Spotify has 320 million total subscribers, of which 144 million are premium (paying) subscribers, while the other 176 million use the app for free. As of today, 91% of the company’s revenue comes from its premium service, so there’s no doubt about the importance for the company to understand whether the advertising interruptions displayed in the free version are effective when trying to convince consumers to go premium.

Latin America in particular is Spotify's third largest market and accounts for 22 percent of its monthly active users: it had 66 million monthly active users in this region in the second quarter of 2020. In Argentina, Spotify ranked as the second choice, behind YouTube, when it comes to music streaming services. During a survey carried out by "Carrier & Asociados"[2] among internet users in Argentina in 2020, around 26 percent of respondents stated using the Spotify free service, whereas 20 percent stated they used Spotify Premium.

1.3. Music genres

Although the unit of analysis of the investigation are songs, we will group them in genres in order to detect which genres are most similar to others and vice versa. Music genre is a conventional category that identifies some pieces of music as belonging to a shared tradition or set of conventions. It is to be distinguished from musical form and musical style, although in practice these terms are sometimes used interchangeably (Wong, 2011). Music can be divided into genres in varying ways, such as into popular music and art music, or religious music and secular music. The artistic nature of music means that these classifications are often subjective and controversial, and some genres may overlap. Academic definitions of the term *genre* itself vary. This is why in this study we have adopted the traditionally used classifications of genres, but we took the freedom to adjust them in some cases. For example, although Hip Hop and Rap are usually considered as being two separate genres, in this study we considered them as one due to their similarities in sound attributes.

Grouping songs into genres is a difficult task not only because of the diverse and subjective classifications made in the literature, but also because of the way in which genres are labeled in Spotify. Some clarifications on this matter are necessary, in Spotify:

---

2 https://noticias.perfil.com/noticias/empresas/youtube-le-gana-a-spotify-la-batalla-local-por-el-streaming-de-audio.phtml

- Music genres are associated with artists and not songs. Therefore, we understand the genre of a song as the genre of the artist to which the song is related.

- Artists are almost always labelled with subgenres instead of genres. The list of subgenres in Spotify is huge, and even Spotify creates its own subgenres for some artists. This is why for this study we had to group these subgenres into genres. In doing so, some subjective decisions were made. For example, when defining the core genre of the subgenres created by Spotify, or when defining the core genre of subgenres that resulted from a combination of different genres, such as “Blues rock”.

The procedure and detailed description of the genre classification made and how specific cases were handled is detailed in “*section 3: Dataset Creation*”.

1.4. Problem & Objective

Past research papers studying the concept of Psychological Reactance to explain the effects of ad's interruptions on consumer behavior do exist, as well as a great contribution of data scientists that used *Spotipy* to gain some insights on Spotify's audio features. However, to our knowledge, to date there isn't a study that combines both tools to investigate the effects of voice over ad interruptions in audio streaming services. Specifically, regarding the impact in consumer's behaviour when manipulating the background music of the ad.

The business question we will be trying to answer in this thesis is the following: if I give a Spotify non premium user an ad with a song of a music genre opposite to what they usually and currently listen to, will they be more or less willing to go premium, and what will be the effect on their behavioral intentions towards the brand providing the service?

Whereas there is no managerial reason to believe that the genre of the song used in an advertisement will influence the consumer’s choice of going premium or not, consumer psychology says otherwise. Advertising features, such as songs, do persuade consumers differently. Previous research has shown that music acts as a prominent advertising cue in broadcast media (Alpert and Alpert 1990), as it is considered an efficient and valuable way to trigger moods and communicate nonverbally. The type of music used in an ad is also a way to adjust the message being communicated. In fact, previous studies have shown that message reception is influenced by the interplay of two musical properties: attention-gaining value and music-message congruency (Kellaris et al. 1993). Based on these concepts, we hypothesize that Psychological Reactance is the mediator by which the song used in an advertisement influences the consumer's choice.

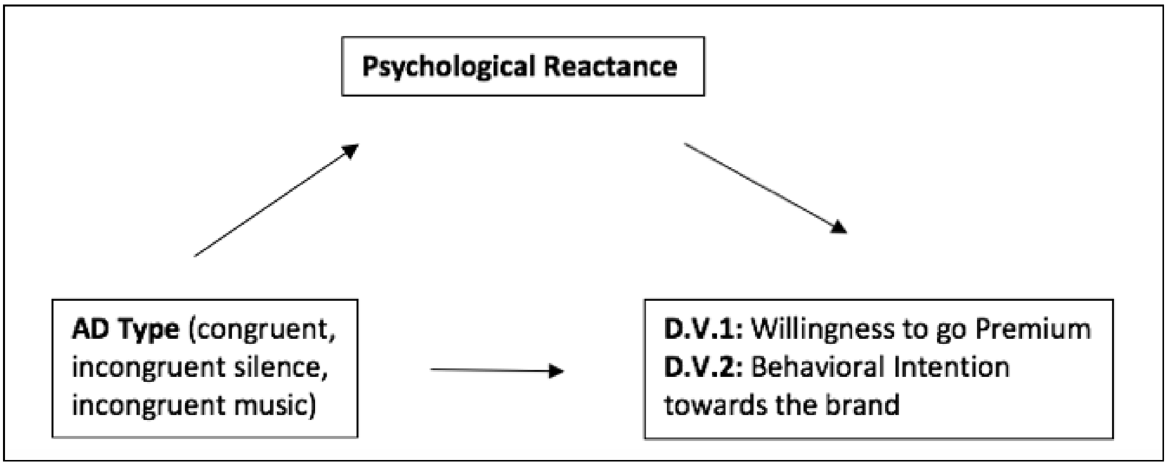
“Psychological Reactance” (Brehm 1966) is a psychology theory which states that individuals have certain freedoms with regard to their behavior. If these behavioral freedoms are reduced or threatened with reduction, the individual will be motivationally aroused to regain them. For example, when children do something especially because it is prohibited or because their parents say so. Among influence mechanisms, reactance is also a variable that is very close to resistance in terms of content. Defined as “a motivational state that is aimed at restoring a threatened or eliminated freedom” (Brehm and Brehm, 1981), it is an effort to regain freedom of choice or options of which the consumer feels he has been deprived. This is sometimes viewed as a trait, thus supporting the hypothesis of a powerful need for freedom on the part of the individual (Shen and Dillard, 2005). Whether considered a trait or motivational state, reactance leads the individual either to avoid what he perceives as an attempt to reduce his choices or to desire even more strongly the option he feels he has lost (Brehm, 1966). We believe that reactance is at the basis of our hypothesized effect. In particular, we hypothesize that when consumers of Spotify’s free service are listening to songs of a certain genre and get interrupted by ads featuring music from an opposite genre, they will perceive that the music of the opposite genre is used as a way to annoy them thus coercing them into subscribing to the premium service, which will eliminate this negative stimulus. As a result, we posit that consumers will be less willing to pay for the premium service when interrupted by ads that are incongruent with their music taste or their current music needs, and that their behavioral intention towards the brand providing the service will deteriorate in this situation. Regardless of whether this is really Spotify’s intention to coerce them by using a dissonant ad, the company would benefit from knowing if this consumers’ reactance reaction leading to negative behavioral intentions is really the case.

Some interesting studies that develop on the basis of psychological reactance have been written. Clee and Wicklund (1980) describe reactance as a boomerang effect in which the perception of coercion is met with an equal but opposite influence, which is used by consumers to restore their freedom of choice. The authors expand the concept of "self-imposed threats to freedom", which explains that prior to a decision, a person has the possibility to move in any of many directions. But, once this person decides to choose, a threat to that freedom ensues. From the theoretical perspective taken by the authors, this should lead to a consequent reassertion of freedom, as reflected in a "sudden desire for decision alternatives that are about to be rejected".

Edwards, Li & Lee (2002) delve into this concept and apply it to internet pop-up ads. *Forced Exposure and Psychological Reactance: Antecedents and Consequences of the Perceived Intrusiveness of Pop-Up Ads*, Edwards explores forced viewing of "pop-up ads" on the Internet to understand the way in which viewers perceive an ad as irritating and decide to avoid them. The authors explain that forced exposure to ads that interrupt a viewer's normal viewing process can lead to a negative perception of the advertising as intrusive. In response to a loss of freedom, viewers will feel uncomfortable and attempt to regain control of their experience. The degree to which viewers perceive an ad as intrusive will be related with feelings of irritation and, consequently, the avoidance of the ad.

This study follows these lines of investigation to propose a novel strategy to investigate the use of pop-up ads in audio streaming services, specifically in the audio-streaming platform Spotify. The investigation builds upon psychological reactance theory to test the effects of the displacement of different music genre songs as ad background music in the listener's willingness to subscribe to the audio-streaming platform, as well as in their behavioral intention towards the brand providing the service. Along with previously mentioned studies, we propose reactance as the mediator between the exposure to an ad and the plausible negative perception of the advertisement, but unlike already existing works, we apply this concept to a modern audio-streaming platform.

Mediation model



1.5. Methodology

The analysis consists of two sections:

Descriptive Section: The objective of this phase is to gather information to build the experimental design. The first step will be to obtain the most popular music genres in Argentina, meaning those that appear most frequently in "Spotify Viral 50" chart songs, for the country Argentina and for the years 2018, 2019 and 2020. Once we have created the respective dataset and selected the 5 most popular music genres, the following step will be to create another dataset with Spotify queried songs for each genre. This dataset will contain 18.000 songs randomly selected for each genre (80.000 in total), containing information regarding the audio attributes for each song. Using K-Means Clustering, the songs will then be grouped into 5 different clusters (expecting to obtain one cluster representation per music genre). The objective of the clustering model is to gather insight of what centroids/music genres are "opposite" in terms of audio attributes, and once identified, use songs of two "opposite" music genres as input for the experimental step. If the clustering model fails to group songs into clusters that distinctively represent each genre and opposite genres can't be clearly identified in terms of audio attributes, opposition between genres will be identified using a pre-experimental test, in which surveyed people will be asked to select which music genres they perceive as being the most different between each other.

Experimental Section: This part of the thesis will focus on building, running and analyzing the experiment. We will first recruit a sample of Spotify premium and non-premium users and randomly assign them to one of the three experimental groups:

- Control group (CG for brevity): users will be listening to a song of their preferred genre and will be interrupted with a song-free ad (an ad with no background music)

- Treatment group 1 (TG1): users will be listening to a song of their preferred genre and will be interrupted with a congruent ad song (a song of the same music genre)

- Treatment group 2 (TG2): users will be listening to a song of their preferred genre and will be interrupted with an incongruent ad song (a song of an "opposite" music genre)

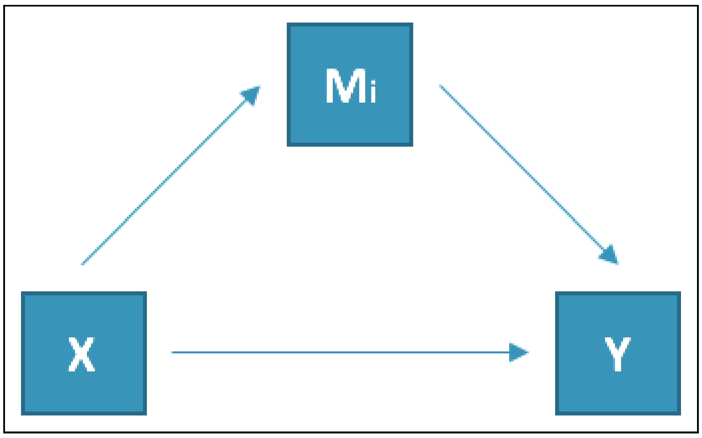
After participants listen to their assigned ad they will be asked to rate the likelihood to subscribe to the premium service, and they will be asked questions regarding their behavioral intention towards the fictional audio-streaming brand created for the experiment. In addition, they will be asked to answer questions designed to measure reactance. The hypotheses to be tested can then be expressed in the following way:

- *Hypothesis 1 a:* The assignment to the experimental groups (I.V.) has an impact on the participant's *Willingness to go Premium* (D.V. 1).

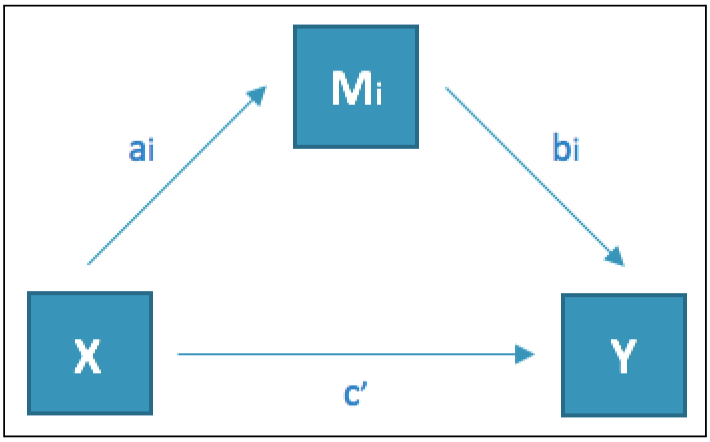
- *Hypothesis 1 b:* The assignment to the experimental groups (I.V.) has an impact on the participant's *Behavioral Intention* towards the brand (D.V. 2).

- *Hypothesis 2*: The effect of the I.V. on the D.V.s is mediated by Psychological *Reactance*.

Mediation model conceptual diagram



Mediation model statistical diagram



Indirect effect of X on Y through Mi = $a_i b_i$
Direct effect of X on Y = c'

We believe that after the experiment, *Willingness to go Premium* and *Behavioral Intention* mean values will be higher in the congruent group (Treatment group 1), and lower in the two incongruent groups (Control & Treatment group 2). On the other hand, we expect *Reactance* mean values to be higher in the incongruent groups (Control & Treatment group 2), and lower in the congruent group (Treatment group 1). In other words, we contemplate that for the D.V.s, experimental groups mean values will be ordered from lower to higher as: Treatment group 2 - Control - Treatment group 1. For Reactance, experimental groups mean values will be ordered from lower to higher as: Treatment group 1 - Control - Treatment group 2.

2. The Data

The data to be used in the research was collected with the Spotify API for Python. The company provides access to some of their data about users, playlists and artists through a Web API.

Specifically, for this analysis, we used the API to collect data and build two core datasets.

a. Viral songs in Argentina[3]: Spotify makes weekly charts with the "viral 50" songs globally and for each country. We used the weekly "viral 50" charts from Argentina for 2018, 2019 and 2020 to collect information about 7750 songs, and built a dataset in which each row corresponds to a song, and each column has information about the song. This dataset has 6 features: *track_name, artist_name, date, region, spotify_id and artist_genre*. Music genres are associated with artists and not songs, so Spotify's viral charts didn't contain this information. This is why we had to make an extra step and track the *artist_genre* associated with the artist using the *artist_name* column.

b. Songs attributes[4]: this dataset was built using Spotify's API to select random songs for the 5 most popular genres identified in the previous phase. Many sub steps were taken when constructing this dataset due to difficulties and particularities of *Spotipy* (Spotify's API), and they are addressed in Section 3.1. The final output of this phase is a dataset containing 80.000 rows. Each row represents a song, and there are 16.000 songs for each music genre. Each column is a feature of the song. The dataset has 15 features: *artist, song_name, song_id, danceability, energy, key, loudness, mode, speechiness, acousticness, instrumentalness, liveness, valence, tempo* and *artist_genre*.

[3] Dataset variables description can be found in Appendix 8.2.1
[4] Dataset variables description can be found in Appendix 8.2.2

Table 1: "Viral songs in Argentina" - 10 first rows

track_name	artist	date	region	spotify_id	artist_genre
Échame La Culpa	Luis Fonsi	2018-01-04--2018-01-04	ar	1zsG4eaZmkA1dvjDDsAGLK	['latin', 'latin pop', 'puerto rican pop', 'tropical']
Cumbia Of Thrones	Cumbia Game	2018-01-04--2018-01-04	ar	5Q6Pqaw1nOeyVZcEqC1S61	[]
Corazón	Maluma	2018-01-04--2018-01-04	ar	4lESS6vuruP6a79KWRaQou	['reggaeton', 'reggaeton colombiano']
Bum Bum Tam Tam	MC Fioti	2018-01-04--2018-01-04	ar	4zWO4gvuFtw6EJZC5FFGlr	['funk carioca', 'funk paulista']
Sensualidad	Bad Bunny	2018-01-04--2018-01-04	ar	5hcisvFMidkMJlElTO9Qmw	['latin', 'reggaeton', 'trap latino']
Déjala Que Vuelva (feat. Manuel Turizo)	Piso 21	2018-01-04--2018-01-04	ar	33bnxcjePIkcmNjEFTJX0l	['colombian pop', 'latin', 'latin pop', 'reggaeton', 'reggaeton colombiano', 'trap latino']
This Is Me	Keala Settle	2018-01-04--2018-01-04	ar	2xGjteMU3E1tkEPVFBO08U	['broadway', 'hollywood', 'show tunes']
Vete	KHEA	2018-01-04--2018-01-04	ar	2rZB6indPhp8C5AZK1PpkS	['argentine hip hop', 'trap argentino', 'trap latino', 'trap triste']
Downtown	Anitta	2018-01-04--2018-01-04	ar	3Ga6eKrUFf12ouh9Yw3v2D	['funk carioca', 'funk pop', 'pagode baiano', 'pop nacional']
The Greatest Show	Hugh Jackman	2018-01-04--2018-01-04	ar	43ay9lQZ5rfNcOOHhRF2cM	['show tunes']

Table 2: "Songs attributes" - 10 first rows

artist	song_name	song_id	danceability	energy	key	loudness	mode	speechiness	acousticness	instrumentalness	liveness	valence	tempo	genre
Arbovirus	Amader Gaan	spotify:track:1YmKMMzVGf7Y3fgWVEDp\	0,324	0,904	10	-7,335	1	0,141	0,000253	0,0112	0,313	0,253	184,786	Rock
I Santoni	Forse un sogno	spotify:track:6kUrMMha9cWe8j1Fl5Oyuj	0,418	0,42	10	-10,658	1	0,0454	0,399	0,00188	0,124	0,69	143,154	Rock
Orion's	Príncipe Cansado	spotify:track:3sQQgu4XbqbwwAukmJofNs	0,216	0,45	9	-7,547	1	0,0276	0,791	0,0000416	0,136	0,125	77,804	Rock
Pavol Hammel & Prudy	Medulienka	spotify:track:7aFSibO8jZ0UOcVK1zSV8m	0,54	0,65	2	-10,394	1	0,0486	0,522	0	0,101	0,357	99,202	Rock
Taylor Hawkins & The Coattail Ride	Crossed The Line	spotify:track:4Sqq4flOBLG8S5DohMkilI	0,417	0,789	0	-5,038	1	0,0515	0,0145	0,00869	0,301	0,38	115,395	Rock
A Giant Dog	(Antichrist Television Blues	spotify:track:5XImJ3ml4UOK3gY0KUtSja	0,433	0,952	7	-5,306	1	0,136	0,00317	0,000778	0,252	0,445	104,668	Rock
Dragon Ash	Jump	spotify:track:6cS4k7wPjeZF7Wu2f3vGow	0,627	0,93	7	-5,445	1	0,0464	0,0183	0,0000791	0,348	0,677	129,98	Rock
No And The Maybes	Hummingbird Bed	spotify:track:12l150jwOaQswnTkuUg7fF	0,572	0,866	5	-4,805	0	0,0399	0,0372	0,0001	0,0781	0,455	80,049	Rock
Bebeto Alves e os Blackbagual	Naval - Ao Vivo	spotify:track:59MpEINIj1zbsvxRLsiCGK	0,505	0,903	4	-6,374	0	0,0625	0,0261	0	0,842	0,779	157,255	Rock
Mägo de Oz	El amor brujo	spotify:track:0nAkkFVCsJ4Rpg4xYxn85M	0,423	0,985	6	-4,491	0	0,0512	0,000634	0,22	0,899	0,435	93,01	Rock

3. Dataset creation & Exploratory Analysis

In this section we will be explaining in greater detail how each dataset was built, and we will be doing an exploratory and descriptive analysis for each one of them. These datasets will then be the input of the K-Means cluster.

3.1. Viral songs dataset

Once we connected with *Spotipy* (Python's API for Spotify), we were able to create a request in which we asked for the 50 weekly viral songs in Argentina during the period that goes from January 2018 to December 2020.

The output of this request was a csv file that contained a list of songs with the attributes *track_name, artist_name, date, region and spotify_id.* Next, we created a function which took as input the csv and, for each song, requested the music genre associated with the artist of the song.

As a result, we obtained a final dataset of 7.750 songs and 6 features: *track_name, artist_name, date, region, spotify_id and artist_genre*.

Once constructed the dataset, we extracted the most popular music genres. This task has its challenges as each artist, and therefore each song, may have no genre, 1 genre, or more than 1 genre assigned. At the same time, a genre can (an almost in every case is) subdivided in many subtypes of music genres. For example, “pop” can appear as “latin pop”, “funk pop”, “colombian pop”, “cumbia pop”, etc. Thereby, we first had to detect all the subgenres available in the dataset, which gave a total of 720 different subgenres. Then, we grouped these subgenres in their respective genre. In doing so, some considerations were taken:

- Many artists were tagged with “latin” genre, and we had to exclude these from the analysis and tag them as “Other”, as latin music encompass a wide range of different genres

- Many artists are tagged with subgenres created by Spotify. Some of them are clear references to existing subgenres of music (such as “deathcore”, which was tagged as “Heavy Metal”), but for many others we had to tag them as “Other”

- When there is more than one genre of music inside a subgenre, the core genre was prioritized (for example, electronic blues was tagged as “Blues”)

After we grouped all of the 720 music subgenres into their core genres, we identified which of them are the most popular, understood as those with the highest frequency count. The results were the following:

- Pop: *6.512 results*

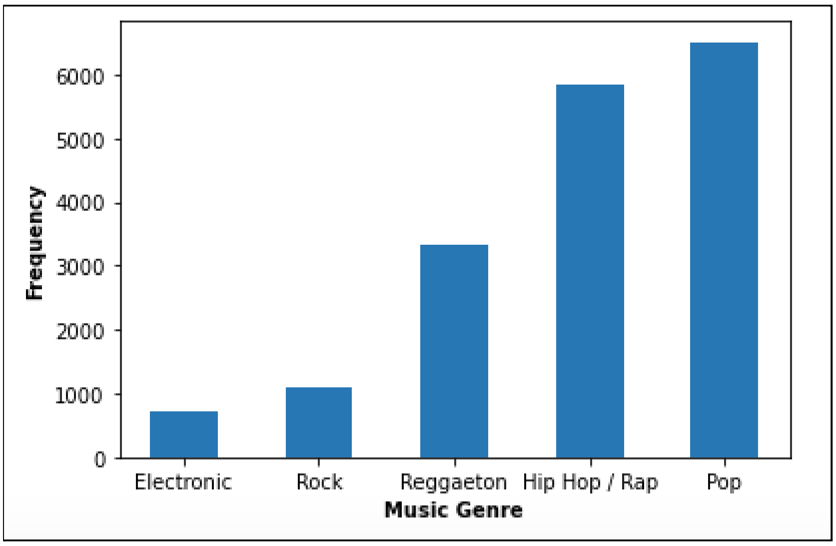
- Hip Hop / Rap: *5.827 results*

- Reggaeton: *3.322 results*

- Rock: *1.109 results*

- Electronic: *730 results*

Figure 1: Music genre's frequency



3.2. Song attributes dataset

In order to construct this dataset, some sub steps had to be taken. Songs for each one of the most popular music genres ("Pop" / "Hip Hop & Rap" / "Reggaeton" / "Rock" / "Electronic") were firstly searched by extracting artists associated to the genre, and then querying songs from the last albums published in Spotify by these artists.

When using the Spotify search query to extract songs, we realized that even when randomizing the search, most of the songs returned were limited to a small number of artists. As a result, when we let the search query run for a considerable amount of time collecting songs in what was supposed to be a “random” search, we found out that of all the songs queried, many of them were repeated. To solve this problem, we decided to divide the dataset creation in two main steps:

1. Randomize the search of artists, until having a considerable number of unique artists for each music genre

2. Search for songs of the artists returned in step 1, by extracting songs from the last albums published by these artists

This approach proved to be much more effective. Not only because songs found were now of a wide variety of artists, but also because, as previously mentioned, music genres are associated with artists and not songs. Therefore, Spotify's search query algorithm runs much faster when we build it to search for artists of a certain genre, and not songs that are associated with artists of a certain genre.

Step 1

Once the search query was built, for each one of the 5 music genres, 500 queries were made, and each one returned 50 results of artists (25 000 artists returned for each genre). Once duplicated values were removed, the final number of artists for each genre was the following:

- Pop: *6.331 artists*

- Rock: *5.688 artists*

- Hip Hop / Rap: *3.168 artists*

- Electronic: *1.255 artists*

- Reggaeton: *870 artists*

The difficulty to obtain *Reggaeton* and *Electronic* music artists from Spotify may be explained by the shorter trajectory of these genres compared to the others analyzed in this section. Although it isn't clear exactly when reggaeton music initiated, its appearance is usually situated in the 1990s (Negrón-Muntaner & Rivera, 2016), while other genres such as *Rock* date from the 1950s. In addition, reggaeton music is a much more geographically limited genre, with almost every artist being from Latin-America. Other attempts to find more reggaeton artists in Spotify were made, such as searching for artists through playlists, or artists that are linked as 'similar' to reggaeton artists by Spotify's listening algorithm, but none of them were successful. Artists found by these means were already queried by the search algorithm built.

Table 3: Artist's results for "Rock" genre - 10 first rows

artist	genres	spotify_uri
OXES	['instrumental rock', 'math rock', 'noise rock', 'post-hardcore']	spotify:artist:2Tzy67BolDEM4uqlewpthf
Six Going on Seven	['boston rock']	spotify:artist:49vw6Q4CYxHv0YRC57NI0J
Ojo de Buey	['musica costarricense', 'reggae tico', 'rock tico']	spotify:artist:5SZeAd236vdPV1XXVuET0y
Sultans Of Ping F.C.	['irish indie rock']	spotify:artist:6SdAlCpKHeAXKmNzvhuhSs
Oceano	['panamanian rock']	spotify:artist:3Hz1j47p1SiZ9jlgMfkU01
O Grilo	['brazilian rock', 'rock alternativo brasileiro']	spotify:artist:22KEpOwThQ5q1DGochfayO
I Got You On Tape	['danish alternative rock']	spotify:artist:3HR5AnPFqtwPA8XvCw7VL5
The Brad Pitt Light Orchestra	['irish rock', 'limerick indie']	spotify:artist:7JMRUQ02VgQk4vT9jxS88v
Ombladon	['romanian rap', 'romanian rock']	spotify:artist:6ojVBJkxhgSj5zgXsM3hnF
Obe Dve	['russian alt pop', 'russian indie rock', 'russian rock']	spotify:artist:7pmh8z3Pzz2u68OmucFSZz

Once we created a preliminary dataset for each one of the music genres, the next step taken was to make sure that artists found for each genre weren't also associated with the other music genres of interest. The search query built returns artists labeled with the genre given as a parameter, but not artists whose only label is the genre indicated. Therefore, we had to iterate through all the artists and its genres/genres associated, and remove those artists that, in addition to the music genre by which they were searched for, they had also associated some of the other 4 genres of interest. For example, if the reggaeton query returned an artist associated with the genres 'latin reggaeton' and 'rock', it was removed. Or if the rock query returned an artist whose music genre was 'pop rock', it was also removed. The only exception was made for reggaeton artists whose music genres associated also included 'hip hop' and/or 'pop'. This was a conscious decision as, by definition, reggaeton music is considered 'pop' (a phenomenon that doesn't happen the other way round). Also, reggaeton is very frequently labeled as 'hip hop latino' by Spotify. Therefore, there is almost no reggaeton artist in Spotify whose music genres associated excludes the words 'pop' or 'hip hop'.

After the cleansing process was finished, the final number of artists for each genre was the following:

- Pop: *5.303 unique artists*

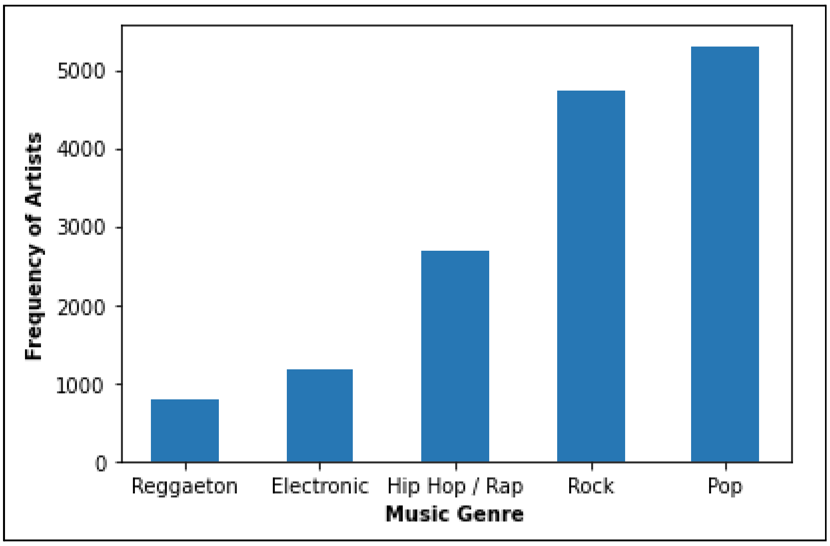
- Rock: *4.734 unique artists*

- Hip Hop / Rap: *2.678 unique artists*

- Electronic: *1.184 unique artists*

- Reggaeton: *776 unique artists*

Figure 2: Artist's frequency



Step 2

Having the clean dataset of artists for each music genre, the following step to complete the dataset creation process was to query songs from these artists. To achieve these, a core function was created, in which the dataset of artists for a particular genre was used as input, and the same dataset containing a new column for each album queried was returned.

Considering that the number of unique artists found in step 1 is different for each music genre, the treatment for each music genre when querying songs was also different:

- For rock and pop genres, only songs from the last album published in Spotify for each artist were queried

- For hip hop genre, songs from the last two albums published for each artist were queried

- For reggaeton and electronic, songs from the last three albums published for each artist were queried

For each music genre, we now have a dataset containing three additional columns, in which the results of the song's query are stored for each row/artist. The unit of analysis is still artists, but what we need is, for each music genre, a dataset where each row is a song and the columns are the audio attributes of the song. Therefore, a function was created to return a dataset for each music genre and each album queried.

For example, the reggaeton genre dataset, that contained for each reggaeton artist information about the artist's last three albums, was transformed into three different datasets for each album, in which the unit of analysis are songs, and the attributes of the dataset are the audio attributes for each song. This gives a result of 10 new datasets (3 for reggaeton, 3 for electronic, 2 for hip hop/rap, 1 for rock and 1 for pop), which were concatenated by each music genre, having a total of 5 song datasets for each genre.

As it was to be expected, each song dataset has a different unique number of songs:

- Rock: *41.428 songs*

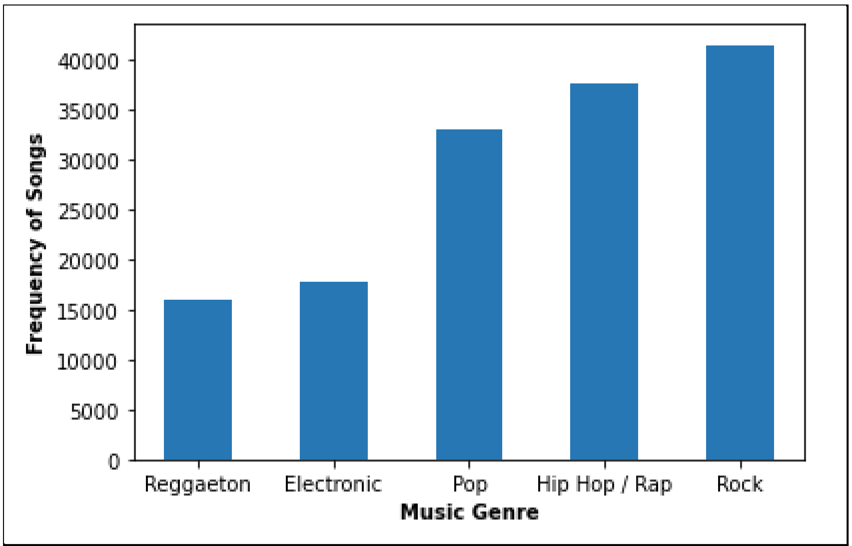
- Hip Hop / Rap: *37.548 songs*

- Pop: *32.967 songs*

- Electronic: *17.720 songs*

- Reggaeton: *16.011 songs*

Figure 3: Song's frequency



Finding reggaeton songs is a difficult task for reasons previously mentioned. For this reason, and to achieve an equilibrium between the 5 music genres, we decided to collect 16.000 random songs for each music genre (as this is approximately the number of songs that we could collect for reggaeton), and join these songs in the same dataset, which will have a total number of 80.000 songs. This dataset of 80.000 songs will be the input to the clustering model described in the next section.

Table 2: "Songs attributes" - 10 first rows

artist	song_name	song_id	danceability	energy	key	loudness	mode	speechiness	acousticness	instrumentalness	liveness	valence	tempo	artist_genre
Arbovirus	Amader Gaan	spotify:track:1YmKMMzVGf7Y3fgWVEDpVG	0,324	0,904	10	-7,335	1	0,141	0,000253	0,0112	0,313	0,253	184,786	Rock
I Santoni	Forse un sogno	spotify:track:6kUrMMha9cWe8j1Fl5Oyuj	0,418	0,42	10	-10,658	1	0,0454	0,399	0,00188	0,124	0,69	143,154	Rock
Orion's	Príncipe Cansado	spotify:track:3sQQgu4XbqbwwAukmJofNs	0,216	0,45	9	-7,547	1	0,0276	0,791	0,0000416	0,136	0,125	77,804	Rock
Pavol Hammel & Prudy	Medulienka	spotify:track:7aFSibO8jZ0UOcVK1zSV8m	0,54	0,65	2	-10,394	1	0,0486	0,522	0	0,101	0,357	99,202	Rock
Taylor Hawkins & The Coattail Riders	Crossed The Line	spotify:track:4Sqq4flOBLG8S5DohMkill	0,417	0,789	0	-5,038	1	0,0515	0,0145	0,00869	0,301	0,38	115,395	Rock
A Giant Dog	(Antichrist Television Blues)	spotify:track:5XlmJ3ml4UOK3gY0KUtSja	0,433	0,952	7	-5,306	1	0,136	0,00317	0,000778	0,252	0,445	104,668	Rock
Dragon Ash	Jump	spotify:track:6cS4k7wPjeZF7Wu2f3vGow	0,627	0,93	7	-5,445	1	0,0464	0,0183	0,0000791	0,348	0,677	129,98	Rock
No And The Maybes	Hummingbird Bed	spotify:track:12l150jwOaQswnTkuUg7fF	0,572	0,866	5	-4,805	0	0,0399	0,0372	0,0001	0,0781	0,455	80,049	Rock
Bebeto Alves e os Blackbagual	Naval - Ao Vivo	spotify:track:59MpElNlj1zbsvxRLsiCGK	0,505	0,903	4	-6,374	0	0,0625	0,0261	0	0,842	0,779	157,255	Rock
Mägo de Oz	El amor brujo	spotify:track:0nAkkFVCsJ4Rpg4xYxn85M	0,423	0,985	6	-4,491	0	0,0512	0,000634	0,22	0,899	0,435	93,01	Rock

4. Clustering: an approach to measure genre similarity

Once we created the dataset containing songs for each one of the music genres of interest, we needed an approach to measure "genre similarity", understood as how similar/different a genre is to another in terms of audio attributes. By using a clustering model, we expect songs to be grouped in different clusters depending on what their audio attributes were. In an ideal world, each final cluster created by the model would contain only songs from a unique music genre, having 5 final clusters representing each music genre in a multidimensional space. An Euclidean measure could then be implemented to detect similarity among music genres.

There are two main reasons why K-Means Clustering was chosen as the best algorithm for this task:

- The dataset attributes are numeric

- Before implementing the algorithm, we already knew the number of clusters to be used

K-Means clustering is a method of vector quantization, which divides the data into a previously determined number of clusters and assigns each observation/song to the nearest centroid in terms of Euclidean distance. The centroid is the mean of the values from all the songs belonging to the cluster. To group songs into clusters, the algorithm initialized with a determined number of k clusters (in this case, 5 clusters, one for each music genre). The centroid of each of the 5 clusters (the mean value of the cluster) starts randomly, and the algorithm calculates the Euclidean distance between each centroid and all the observations/songs. Then, each song is assigned to the nearest cluster, and the algorithm recalculates the centroid's position by averaging all the observations in the cluster and reassigns the songs to the cluster with the nearest centroid. Finally, when the assignment of songs to clusters doesn't change after several iterations, it is considered that the K-Means clustering found the best possible assignment of songs to clusters.

K-Means clustering is a non-supervised model, as it is a model that isn't supervised using training data to predict the value of an attribute in test data. However, in this study the music genre of each song works as the category to be "predicted" for each observation. We want to represent music genres in a multidimensional space. If the clustering algorithm groups songs into clusters composed mostly by a specific music genre, having one cluster for each one of the 5 music genres, this means that audio attributes are a great indicator of genre similarity. If the clustering model fails to group songs into clusters that distinctively represent each genre and opposite genres can't be clearly identified in terms of audio attributes, opposition between genres will be identified using a pre-experimental test, in which surveyed people will be asked to select which music genres they perceive as the most different.

4.1. Data engineering

The clustering model will take as input the dataset containing the 80.000 songs (16.000 songs per music genre).

First, we have to keep only the features of interest[5]. The column containing the music genre cannot be input to the model as it is a categorical feature and hence does not possess an obvious underlying notion of distance. The model is actually trying to find that notion of distance.

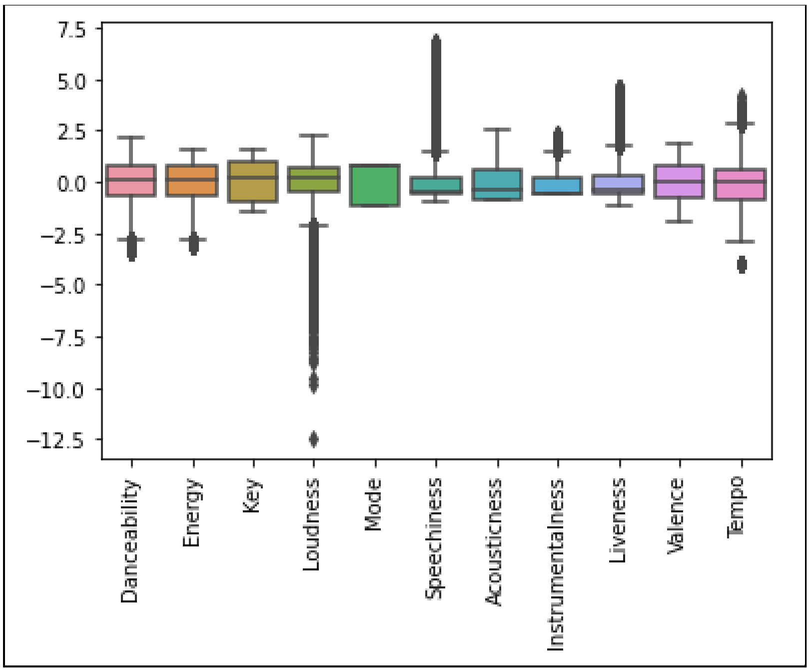
[5] The list of variables used for clustering can be found in Appendix 8.3

Table 4: Clustering input - 5 first rows

	Danceability	Energy	Key	Loudness	Mode	Speechiness	Acousticness	Instrumentalness	Liveness	Valence	Tempo
0	0.324	0.904	10	-7.335	1	0.1410	0.000253	0.011200	0.313	0.253	184.786
1	0.418	0.420	10	-10.658	1	0.0454	0.399000	0.001880	0.124	0.690	143.154
2	0.216	0.450	9	-7.547	1	0.0276	0.791000	0.000042	0.136	0.125	77.804
3	0.540	0.650	2	-10.394	1	0.0486	0.522000	0.000000	0.101	0.357	99.202
4	0.417	0.789	0	-5.038	1	0.0515	0.014500	0.008690	0.301	0.380	115.395

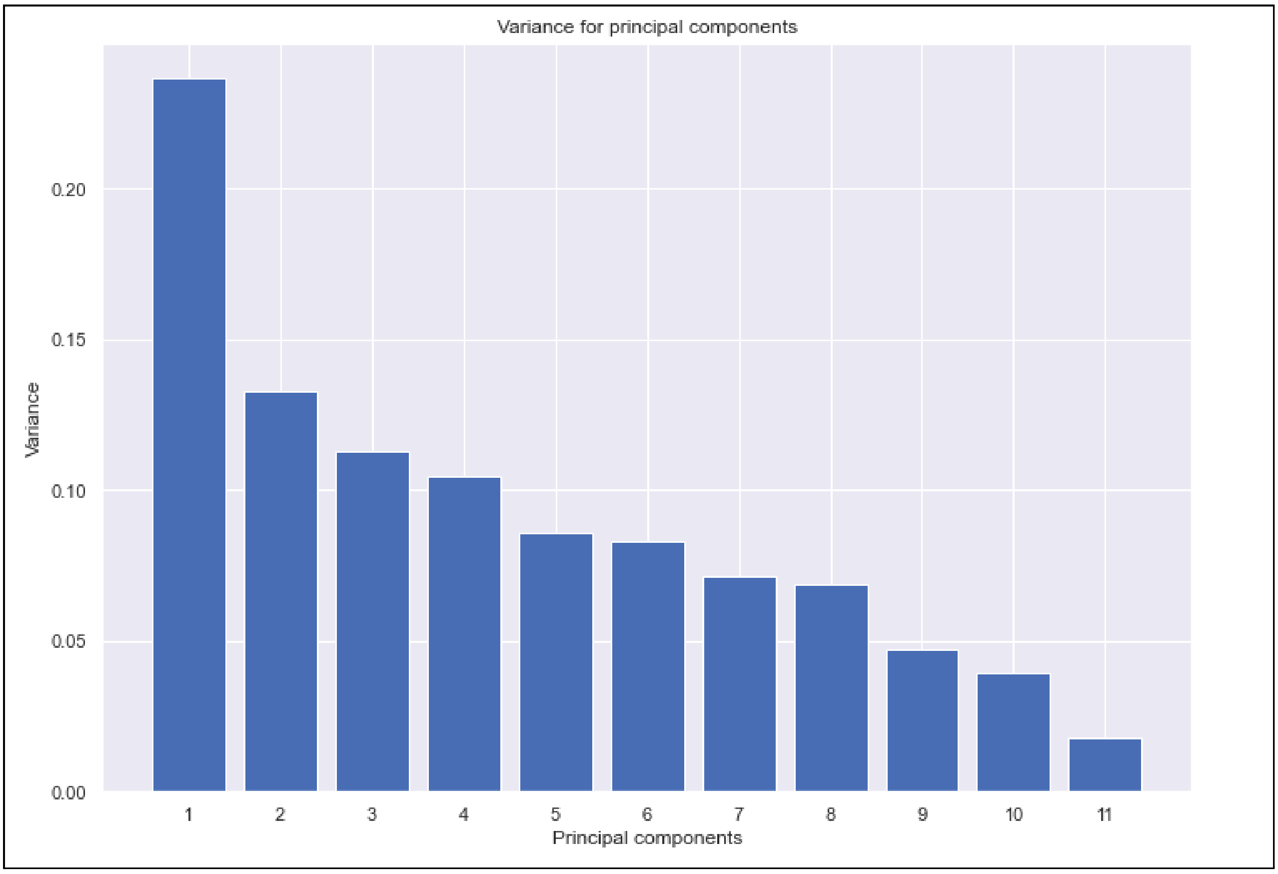
The values of the 11 attributes/audio features are measured using different scales, so the values differ greatly from one attribute to another. As the clustering model uses Euclidean distance to group songs in clusters, we have to normalize the data to convert it into a normal distribution with mean 0 and standard deviation 1.

Figure 4: Audio attributes boxplot



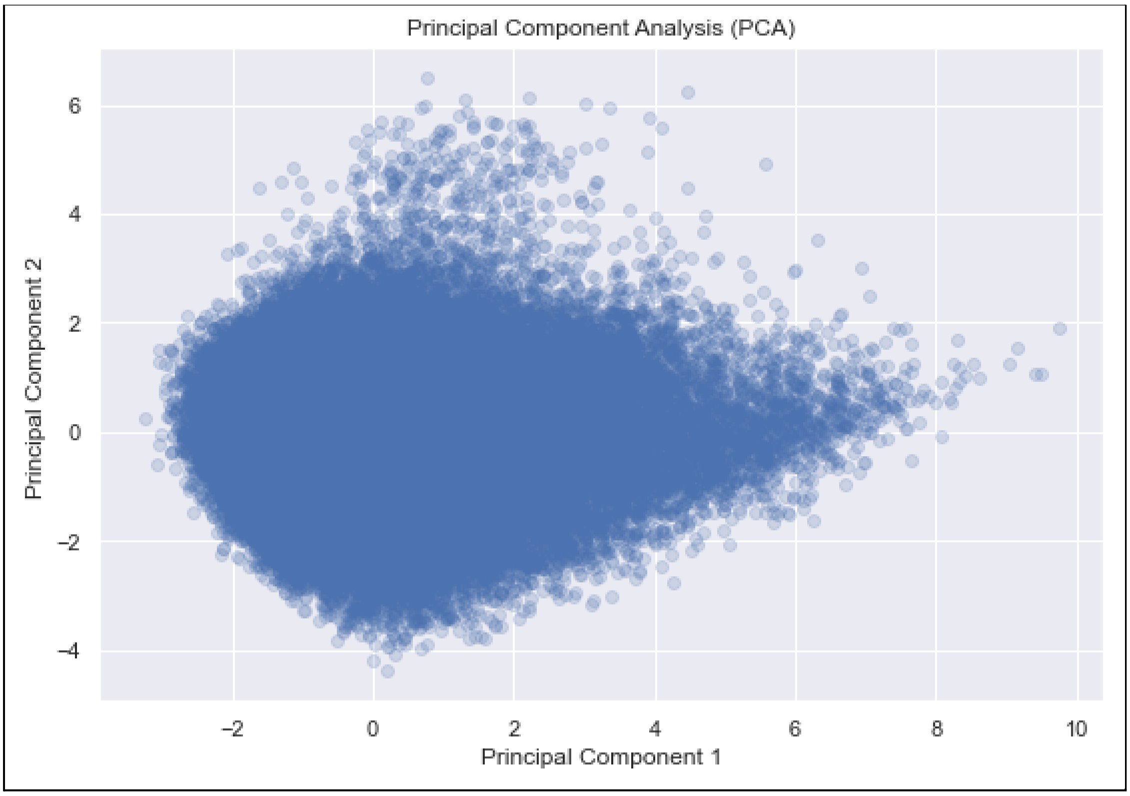
Now that we have the normalized data, we will generate a PCA, using 11 components to explain the variance of the data.

Figure 5: Variance for Principal Components



The *y axis* represents the variance of the data explained by each one of the 11 components generated by the PCA. Now that we know which components explain the most variance, we use the 2 first components to create a scatter plot.

Figure 6: Scatter plot with PCA 1 and 2 components



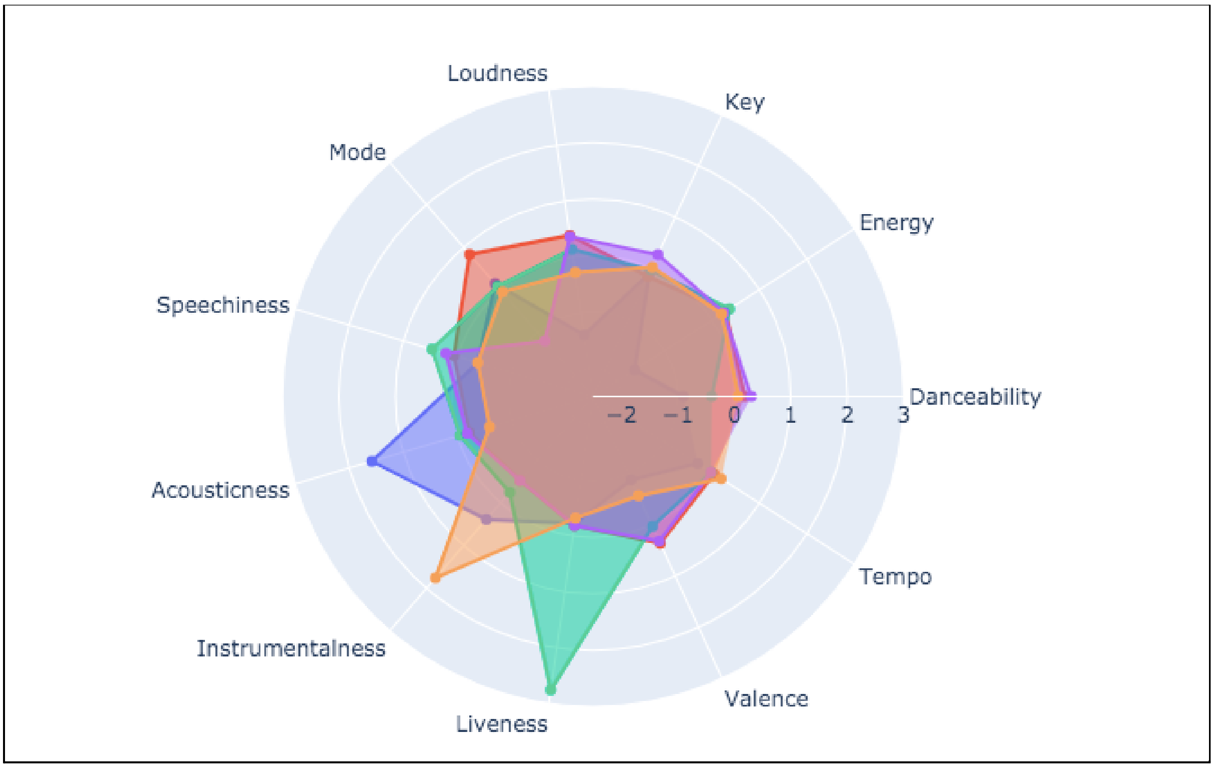
Each dot in the scatter plot is an observation/song, mapped in space by the respective value of *x* (component 1) and *y* (component 2).

4.2. *K-means* implementation

We will initiate a K-Means clustering with 5 centroids, hoping that each one of the 5 final clusters represents a specific music genre. Although we didn't use it to decide on the numbers of centroids, we also runned the elbow method to try to understand what would be the potential number of clusters the model "sees" (see Section 8.4, Figure 20).

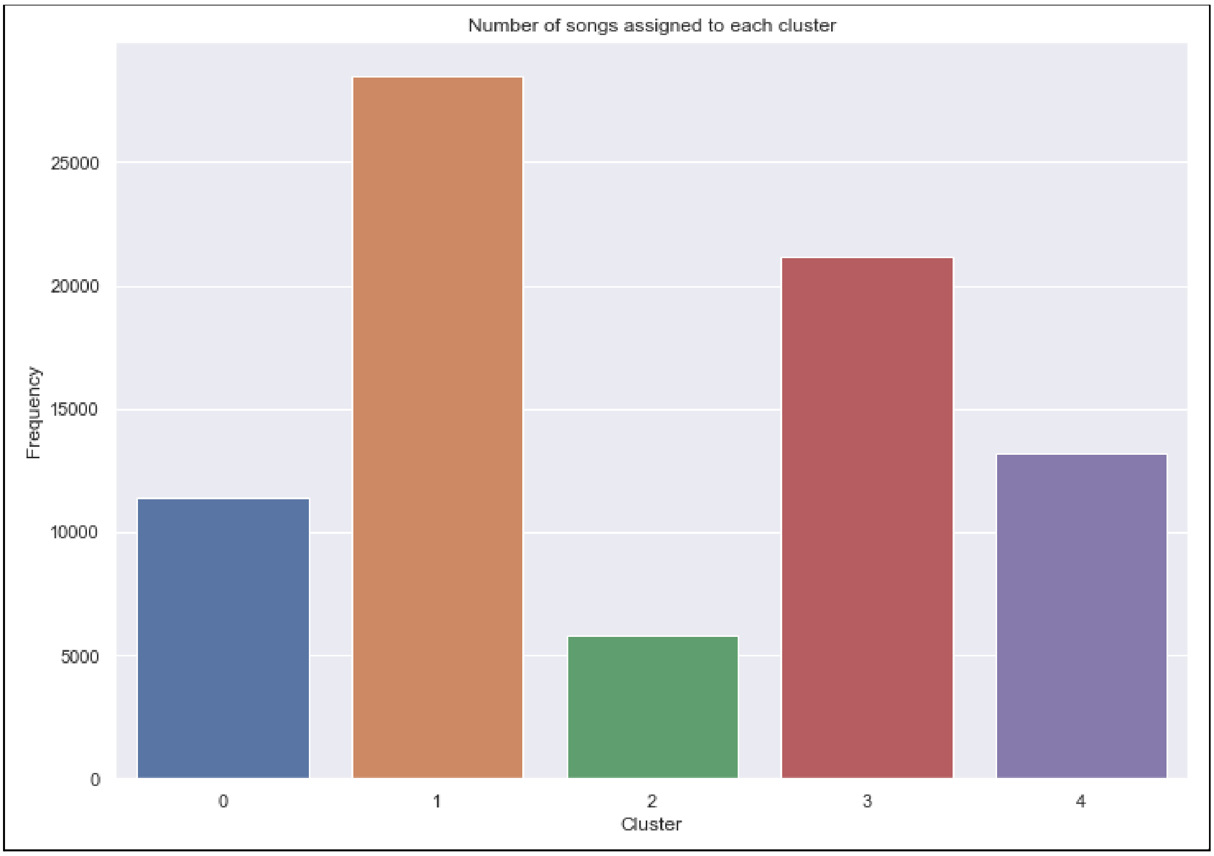
The final centroids are vectors of 11 positions, which represent the 11 audio attributes of the songs. The radar chart displays the values of the 5 cluster centroids for the 11 attributes.

Figure 7: Radar chart displaying centroid values for each audio attribute



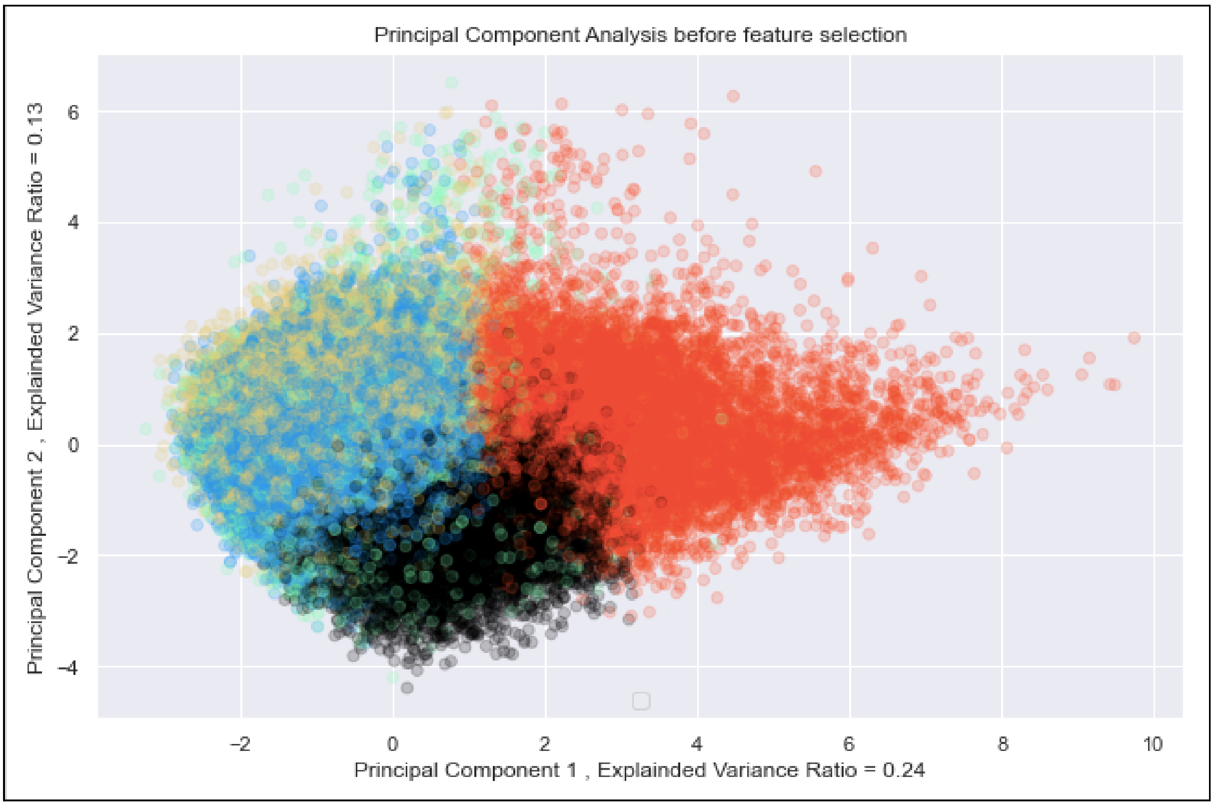
Visualizing the radar chart we can start to sketch some insights. The green colored cluster appears to have songs characterized by high values in the *Liveness* audio attribute, while the orange colored cluster has songs with high values in the *Instrumentalness* audio attribute. The blue colored cluster has some interesting results. It appears to be a cluster containing songs with high values in the *Acousticness* dimension. At the same time, the values in *Energy* and *Loudness* are a lot lower than the other clusters. These results suggest that the blue colored cluster may be a cluster where songs played in acoustic format predominate.

Figure 8: Number of songs in clusters



By looking at the frequency of the songs assigned to each cluster, we can notice that the observations aren't balanced between clusters. While cluster 1 has 28.481 songs assigned, cluster 2 has 5.788. This isn't a good sign, considering that our ideal result would be one in which each cluster contains the greatest percentage of songs possible from only one music genre, which means that each cluster should contain a number near to 16.000 songs.

Figure 9: Colored scatter plot with PCA 1 and 2 components

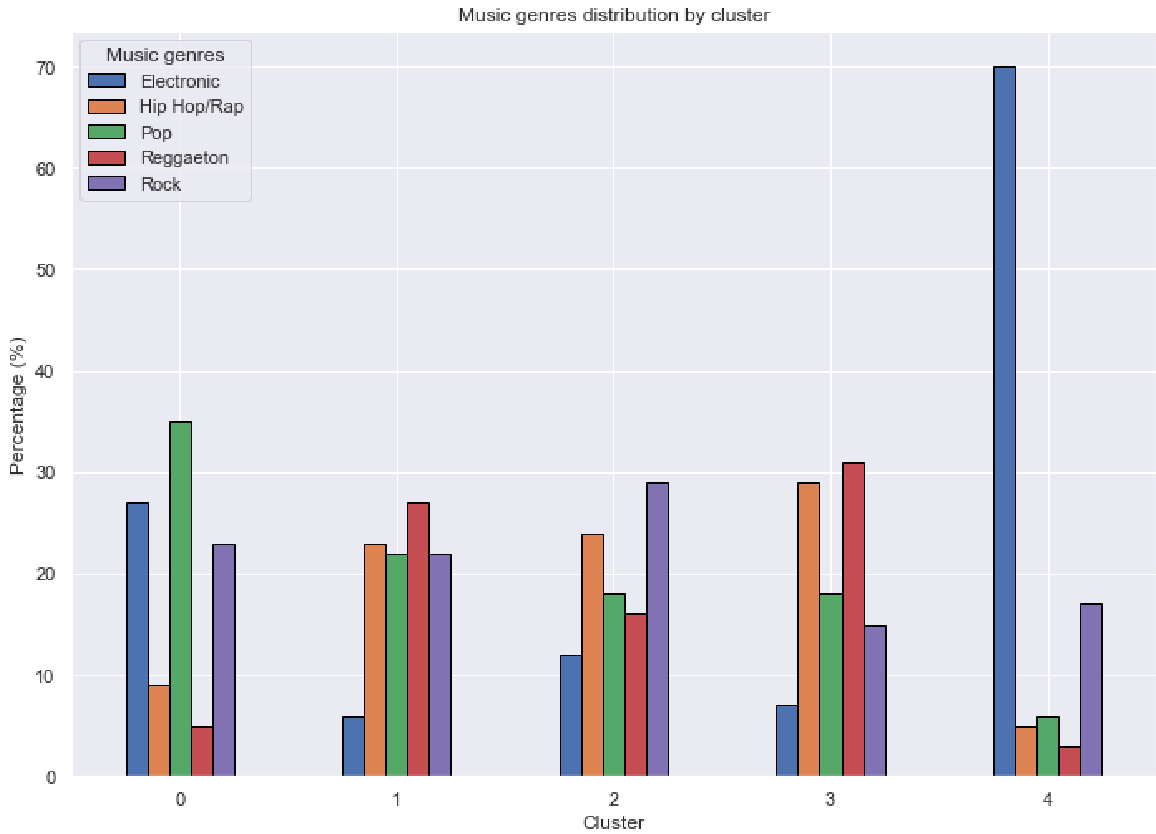


By looking at the scatter plot created with the two first components of the PCA again, and coloring each observation/song with a different color for each one of the 5 clusters, we can also notice that there is a lot of superposition of colors from cluster number 1 (blue), 2 (green), 3 (yellow) and 4 (black). This isn't a good sign either. However, it is interesting how cluster number 0 (red) clearly distinguishes from the others, indicating that the audio attributes from the songs assigned to this cluster have a notable difference compared to the songs from the other clusters.

4.3. Centroid analysis

The results that the graphs from *section 4.1* seemed to announce were finally confirmed when we calculated for each cluster the percentage of songs from each music genre.

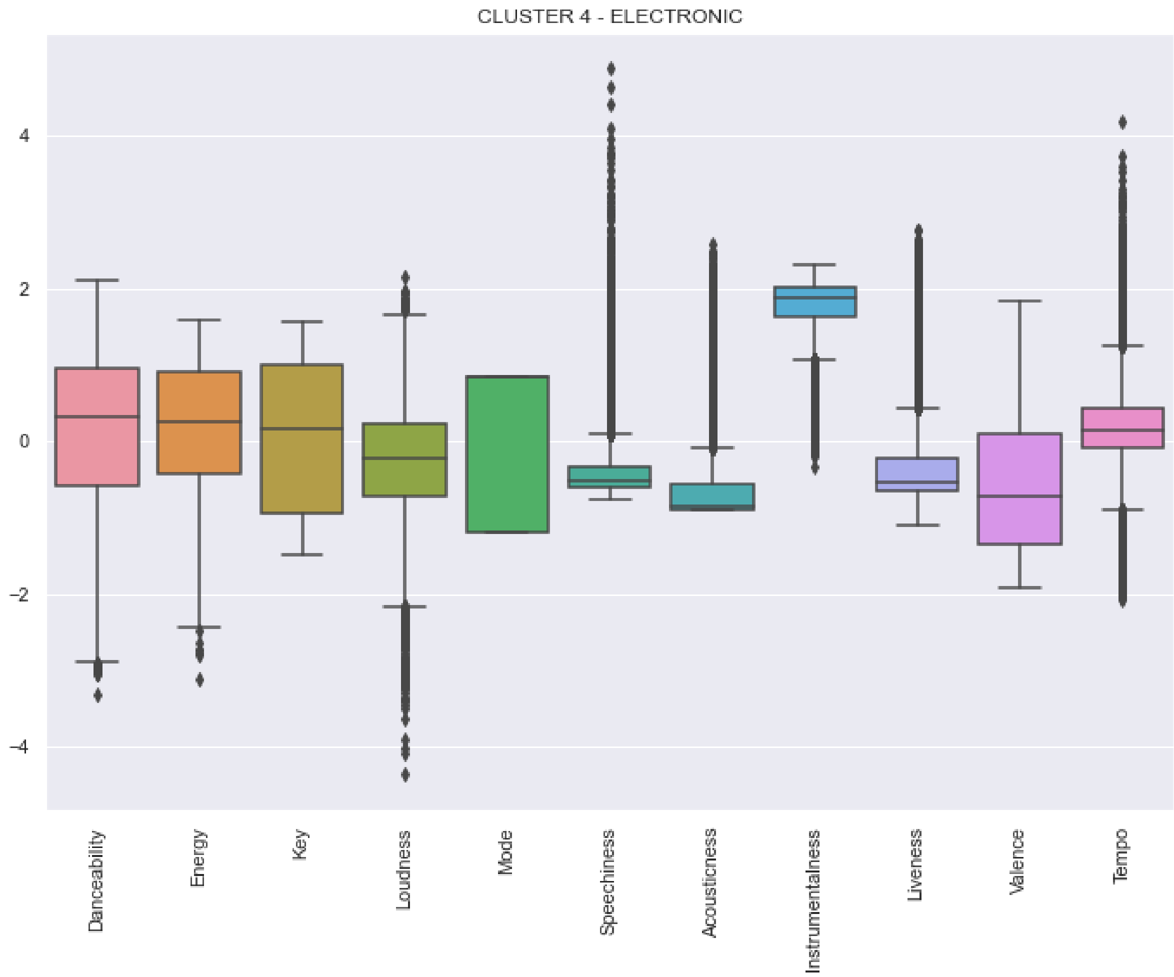
Figure 10: Grouped bar chart with distribution of song's music genres in clusters



Unfortunately, there isn't a clear distinction between music genres regarding the audio attributes of the songs, as the final clusters contain, in most cases, a mix of music genres instead of a predominant one. If music genres can't be clearly identified by the audio attributes of the songs, opposition between genres can't either be recognized by these terms, and a pre-experimental test will have to be done in order to identify which music genres people perceive as being the most different ones between one another.

However, some interesting insights can be taken from the clustering results. From cluster 0, more than 1/3 of the songs are of "Pop" genre, and cluster 4 is clearly the cluster where "Electronic" genre is predominant, as more than 2/3 of the songs assigned to the cluster are from this genre. As we can see in figure 11, it doesn't come as a surprise that the cluster where "Electronic" music genre predominates is the one where Instrumentalness is the audio attribute with the highest values, with half of the observation lying slightly under 2 points in the *y axis*, and *Acousticness* is the audio attribute with the lowest values, with half of the observations lying just above -1 points.

Figure 11: Box plot Cluster 4



Reggaeton and Hip Hop/Rap genres also behave in an interesting way. Except for cluster 2, in all of the other clusters the two genres appear consecutive one to the other when ordering genres by the percentage in which they are in the cluster. When calculating for each possible combination of two music genres what was the average difference of the percentages of appearances for the pair of genres in each cluster, the tuple "Reggaeton – Hip Hop/Rap" returned the lowest result with 0.039%. This means that, on average, these two genres appear in similar percentages when calculating the composition of the clusters, with only 0.039% of difference. For comparison purposes, the tuple with the most percentage difference was "Electronic – Reggaeton".

The similarity between Reggaeton and Hip Hop/Rap genres when analyzing the proportion in which they appear in each cluster isn't a surprise. Both genres started as street music with a great emphasis in percussion. In fact, Reggaeton music was very influenced by American Hip Hop, and many consider Reggaeton as the latin variation of Hip Hop (and as we mentioned earlier, even Spotify labels some reggaeton artists with 'latin hip hop' genre).

4.4. Clustering conclusions

Although the clustering model encountered some interesting insights regarding the behavior of the audio attributes of music genres, clusters weren't a clear representation of each music genre, and this implies that songs from one music genre do not necessarily differentiate from other genres because of the audio attributes the song has. The centroid analysis of *section 4.2* suggests that songs from different genres may have similar audio attributes. Therefore, if songs can't be grouped in their respective music genres because of the numeric values of their audio attributes, opposite genres can't be found using this approach.

A sociological approach will be adopted in a pre-experimental step instead, in which opposition between genres will be found in terms of people's perception of which of the 5 music genres are the most different ones between each other.

5. Experimental step

The experimental step is composed of a pre-experimental step, as well as the main experiment of the study. The pre-step's objective is to detect the tuple of genres that the main experiment will take as input. As we were unable to detect the tuple using the clustering approach, this step will not focus on the audio-attributes of songs, but instead will understand opposition between genres as what people perceive it to be. Once we identify the tuple of opposite music genres, the main experiment will test the hypotheses stated in the *Methodology* section.

5.1 Pretest: Pre – Experimental Step

The pre-experiment was designed to detect opposition between genres. Because of the impossibility to detect opposite music genres in terms of audio attributes in the clustering model, we had to change the approach in which "opposite" genres are understood. The objective from the pre-experimental step was to find a tuple of music genres which people perceived as the most different.

To do so, we built a 5-minute survey in which respondents had to answer 3 main questions. Each question was directed to the 5 music genres that we previously detected as the most popular from the "Viral Songs" dataset (Rock – Pop – Hip Hop/Rap – Reggaeton – Electronic).

1. "Regarding your preferences of music genres, please indicate how much you like the following music genres on a scale of 1 (very little) to 7 (very much)".
2. "Please, indicate from the following music genres which two you find the most different/opposite" (multiple choice with possibility to choose only 2 genres)
3. "Please, indicate from the following music genres which two you find the most similar" (multiple choice with possibility to choose only 2 genres)

The survey was distributed through a link that was posted in different social media, such as Linkedin, Twitter and Whatsapp. Everyone with the link was able to complete the survey, so there weren't any restrictions regarding who could participate in the pre-experiment.

The full outline of the pre-test survey can be found in Appendix 8.5

5.2 Pretest results

Results were analyzed from a total of 128 respondents. The mean age was 27.7 years, with a standard deviation of 10.5. From the total respondents, 52 were men, 75 women and 1 non-binary.

To obtain which tuples of music genres are the most different/opposite, we first analyzed questions number 2 and 3, which asked surveyed people to indicate the 2 music genres they found most similar and the ones they found most different. Results were pretty conclusive:

Table 4: Frequency of music genres declared to be the most different

Tuples of music genres	Frequency count
Rock - Reggaeton	48
Reggaeton – Electronic	25
Rock - Electronic	24
Hip Hop/Rap - Electronic	12
Pop - Electronic	8
Rock – Hip Hop/Rap	6
Pop - Reggaeton	2
Hip Hop/Rap - Reggaeton	2
Rock - Pop	1

Regarding the most different genres, 48 respondents indicated that “Rock – Reggaeton” tuple was the most different, followed by “Reggaeton – Electronic” and “Rock – Electronic”, with 25 and 24 answers respectively.

Table 5: Frequency of music genres declared to be the most similar

Tuples of music genres	Frequency count
Hip Hop/Rap - Reggaeton	34

Rock - Pop	30
Pop - Reggaeton	22
Pop - Electronic	19
Pop – Hip Hop/Rap	13
Rock – Hip Hop/Rap	5
Reggaeton - Electronic	3
Rock - Electronic	2

When asked about the music genres they perceived as the most similar, none of the respondents chose “Rock – Reggaeton” tuple. “Reggaeton – Electronic” and “Rock – Electronic” scored only 3 and 2 answers respectively.

To analyze the results from question 1, which asked respondents to rank from 1 (very little) to 7 (very much) each music genre in order of their likeability, we also performed a frequency count. For this purpose, we first grouped music genres in all the possible combinations of tuples, which gives a total of 10 combinations. For each tuple of music genres, we then calculated the frequency count for the *distance between genres*, understood as the absolute difference between the liking score of the music genres in the tuple. If a respondent likes two genres in the same measure, then the distance between each genre will be 0. On the other hand, the maximum distance between two genres will be 6, which indicates that the respondent likes very little one genre (1), and very much the other (7).

The following table indicates the number of times a tuple of music genres had a score of “6” when calculating the *distance between genres.*

Table 6: Frequency of music genres that scored highest in “distance between genres” measure

Tuple of music genres	**Frequency count**

Rock - Reggaeton	12
Rock - Electronic	11
Pop - Electronic	9
Reggaeton - Electronic	5
Rock – Pop	4
Rock – Hip Hop/Rap	4
Hip Hop/Rap - Electronic	2
Hip Hop/Rap - Reggaeton	1
Pop - Hip Hop/Rap	1
Pop - Reggaeton	0

With a frequency count of 12, “Rock – Reggaeton” appears to be the tuple of music genres people have the most differences in liking, followed by “Rock – Electronic” and “Pop – Electronic”.

Following these results, we conclude that “Rock – Reggaeton” is the tuple of music genres people perceive as being the most different.

5.3 Main Experiment

Now that we have identified *Rock* and *Reggaeton* as the tuples of opposite music genres, the main experiment will use songs of these genres as input for the ad's background music. As we determined, the hypotheses to be tested in the main experiment are:

- *Hypothesis 1 a:* The assignment to the experimental groups (I.V.) has an impact on the participant's *Willingness to go Premium* (D.V. 1).

- *Hypothesis 1 b:* The assignment to the experimental groups (I.V.) has an impact on the participant's *Behavioral Intention* towards the brand (D.V. 2).

- *Hypothesis 2*: The effect of the I.V. on the D.V.s is mediated by *Reactance*.

To carry out the experiment, we have created a fictional audio-streaming platform called *Megatunes*. Each survey completed will be a simulation of *Megatunes* providing its service to users, in which respondents will be listening to a succession of songs which will then be interrupted with an ad, followed by a list of questions designed to test the hypothesis. Previous to completing the survey, all of the participants had to sign an informed consent in which they stated their voluntary participation.

Such as in the pre-experimental step, the survey was distributed through a link that was posted in social media. Everyone with the link was able to complete the survey, so there weren't any restrictions regarding who could participate in the main experiment. Letting anyone with the link complete the survey was a way to enable the most quantity of people to participate, but this kind of distribution also may contribute to an unbalance in the number of people who chose Rock as their preferred genre compared to the ones that chose Reggaeton. In addition, the decision to distribute the link in multiple social media platforms can also contribute to heterogeneity in the respondents (to give an example, users of Linkedin usually have a more academic profile than users of Twitter or Whatsapp). This can cause an unwanted effect if the number of respondents between each social media are unbalanced, as results may be representing a specific population instead of the more general universe of audio streaming platform users.

5.4 Experimental Design

The first question surveyed people had to complete as soon as they started the experiment was to choose between Rock or Reggaeton, depending on their music preferences and their desire to listen to one or other music genres at the moment they answer.

In order to test the hypothesis, respondents were randomly assigned to 3 experimental groups:

- Control group: users will be listening to a 2-minute sequence of three concatenated songs of the music genre they have chosen and will be interrupted with a song-free ad (a 20 second ad promoting “Megatunes”, with no background music)

- Treatment group 1: users will be listening to a 2-minute sequence of three concatenated songs of the music genre they have chosen and will be interrupted with a congruent ad song (e.g. if they are listening to rock songs, the 20 second ad will have as background music a rock song)

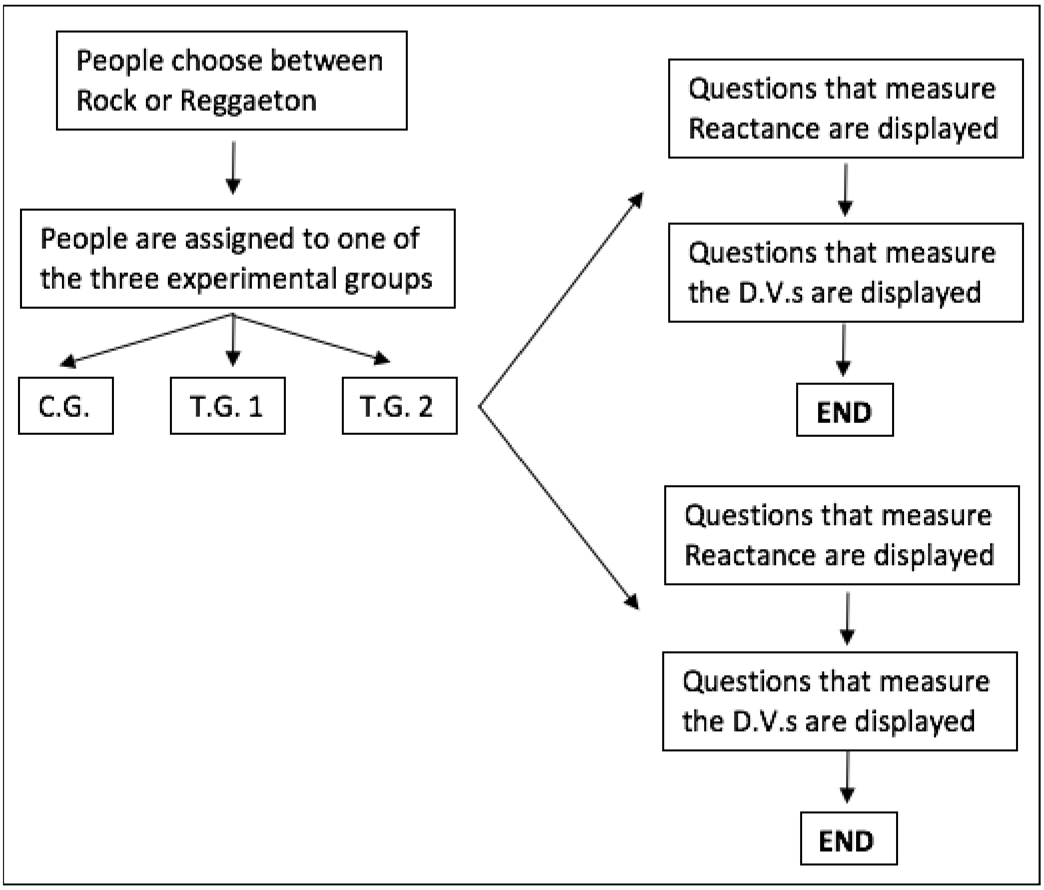
- Treatment group 2: users will be listening to a 2-minute sequence of three concatenated songs of the music genre they have chosen and will be interrupted with a congruent ad song (e.g. if they are listening to rock songs, the 20 second ad will have as background music a reggaeton song)

The songs used in the experiment were taken from the *Song’s attributes dataset.* Although the audio attributes from songs weren’t a good predictor of how songs are classified into their respective music genres, we found this data useful when selecting which songs from Rock and Reggaeton music genres were going to be used as the experiment input. First, we computed the mean of the audio attributes for all the songs of the music genre, until we got a 11-position vector for each one of the two music genres, representing the mean position of the songs in a multidimensional space. Then, we calculated the Euclidean distance of each one of the 16.000 songs vectors with the vector containing the mean values of the audio attributes. After ordering the 16.000 songs of each music genre by the distance to the mean vector, we selected a subset of 1.000 songs for each music genre. From this subset, we then selected the final 4 Rock songs and the final 4 Reggaeton songs that would be played in the experiment. The final songs were chosen taking into account the language in which they were sung (Spanish to avoid introducing a “song language” confound variable into the experiment), and the way in which the songs sounded when played consecutively. We tried different combinations of songs from the 1.000 song subset, until we found a combination that sounded harmonious.

The mediator variable, *Reactance*, was measured by 7 questions displayed after the ad, in which surveyed people had to determine on a scale from 1 (not well) to 5 (extremely well), how good each of the following words described the ad: distracting, disturbing, forced, interfering, intrusive, invasive, annoying.

Regarding the DVs, *Willingness to go Premium* was measured with a single-item question, in which respondents had to decide on a scale from 1 (not likely) to 7 (very likely) how likely is that they would go premium (pay for the Megatunes audio-streaming service). On the other hand, *Behavioral Intention* towards the brand was measured by 4 questions that used the same scale as the former DV. Respondents were asked how likely is it that they would search for more information about the brand, how likely is it that the would consider the brand when thinking on subscribing to a streaming audio service, how likely is it that they would consider the brand's opinions, and how likely is it that they would recommend the brand to a friend. A graphical representation of the experiment's structure is depicted in *Figure 12*.

Figure 12: Main experiment design



After the respondents listened to the ad, a page with the questions measuring Reactance or the D.V.s was displaced. In order to avoid any type of bias, the order in which the two pages appeared was randomized. Some surveyed people answered Reactance questions first, and questions measuring the D.V.s later, and some other surveyed people answered the D.V.s question first, and then questions measuring Reactance.

In order to check whether results from one genre could be compared with the results from the other, a *liking of songs* variable was added to the experiment, in which surveyed people had to answer on a scale from 1 (not at all) to 7 (a lot), how much they liked the songs displayed in the experiment. In addition, respondents had to answer, on a scale from 1 (not at all) to 7 (all of them), if they were familiar with the songs played in the experiment.

The full outline of the pre-test survey can be found in Appendix 8.6

5.5. Pre-processing and Response Analysis

Results were analyzed from a sample of 150 people, formed by 62 men and 88 women. Of the total respondents, 108 chose *Rock* as their prefered music genre, and 42 chose *Reggaeton*.

Composite measures were created for *Behavioral Intention* towards the brand and for negative *Reactance*. Regarding the former, we calculated the mean values of the 4 questions/items measuring *Brand Intention* for each observation. To measure the reliability of the scale, we calculated Cronbach's alpha (Cronbach, 1951)[6], which was sufficiently high (α = .85). To measure *Reactance*, we calculated the mean values of the 7 questions/items measuring negative Reactance. Cronbach's alpha was also calculated and the results were positive (α = .81).

In order to proceed with the analysis of the experiment responses, the first step was to analyze if there were differences between *Rock* and *Reggaeton* groups regarding the liking of songs. We ran a *1-Way Between Subjects ANOVA* and we found statistically significant differences between *Rock* and *Reggaeton* groups, $F(1,144) = 10.65$, $p < 0.01$ (see Appendix 8.7.1). This means that a conjunct analysis for both genres can't be done. The participants who chose Rock systematically expressed a lower liking of the songs displayed in the experiment (average liking = 3.34), compared to the respondents that chose Reggaeton as their prefered genre (average liking = 4.32). When analyzing how familiar respondents were with the songs displayed in the ad, we also found statistically significant differences between *Rock* and *Reggaeton* groups, $F(1,148) = 22.99$, $p < 0.001$ (see Appendix 8.7.1). People who chose *Reggaeton* systematically expressed a higher familiarity (average familiarity = 2.71) than those who chose *Rock* (average familiarity = 1.45). Therefore, a separate analysis had to be done for each one of these groups.

To have a first glance of the data grouped by music choice, we created a table of means and standard-deviations for a 1-Way ANOVA design. By doing so, we could calculate the means of the D.V.s (*Behavioral Intention* & *Willingness to go Premium*) and the mediator (*Reactance*) for each one of the groups in the I.V.:

---

[6] Cronbach's Alpha is a coefficient used to evaluate the reliability of a measure scale. It was devised by Lee Cronbach in 1951, and first appeared in the paper *Coefficient Alpha and the Internal Structure of Tests*.

- Control group: users listening to a song of their preferred genre were interrupted with a song-free ad
- Treatment group 1: users listening to a song of their preferred genre were interrupted with a congruent ad song
- Treatment group 2: users listening to a song of their preferred genre were interrupted with an incongruent ad song

Figure 13: Grouped bar chart with mean values for DVs and Reactance by EG (Reggaeton)

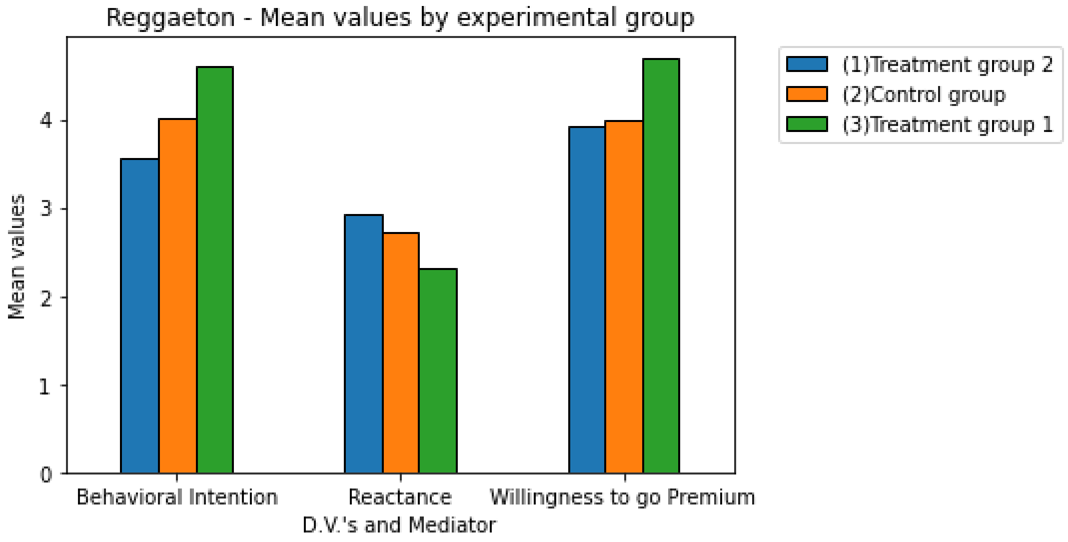
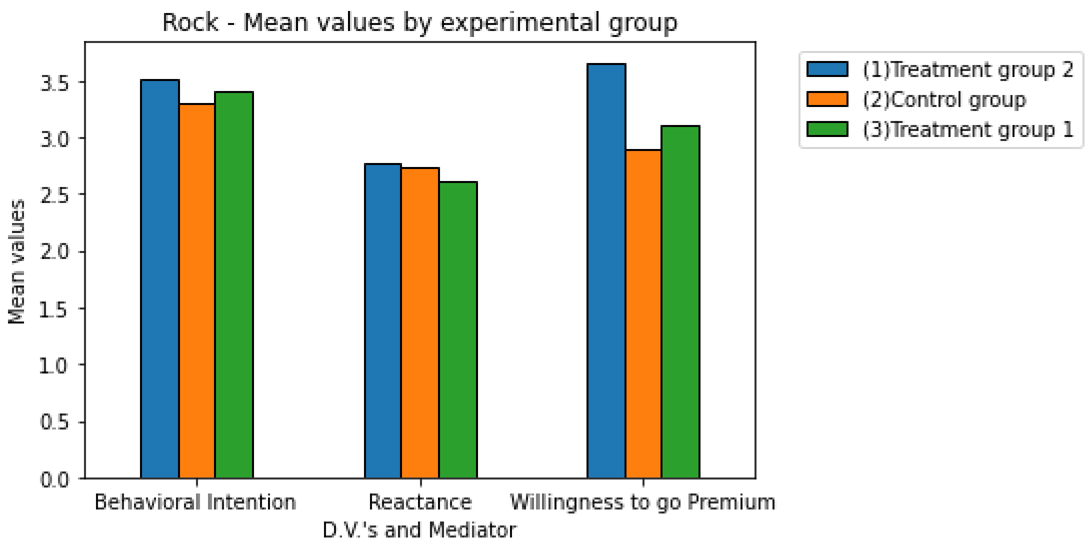


Figure 14: Grouped bar chart with mean values for DVs and Reactance by EG (Rock)



For the Reggaeton sample, the mean of the D.V.s and the Mediator behave as we expected. The order in which the mean values are displayed follow our hypotheses: for *Willingness to go Premium* and *Behavioral Intention*, Treatment group 2 has the lowest mean value, followed by Control group and Treatment group 1. On the other hand, when measuring *Reactance*, Treatment group 2 has the highest mean value, followed by Control group and Treatment group 1. Meaning that being interrupted by an ad with no background music or by an incongruent ad song increases *Reactance* and decreases *Willingness to go Premium* and *Behavioral Intention*, and that changes in mean values are larger when the ad has an incongruent song than when it has no background music.

However, in the Rock sample results weren't as we expected. The D.V.s mean values by group don't behave as hypothesized, as the incongruent experimental group is the one with the highest mean values. Although not as clear as in the Reggaeton sample, *Reactance* follows the order presented in the hypothesis (i.e, Treatment group 2 has the highest mean value, followed by Control group and Treatment group 1).

Post-hoc analyses based on the effect sizes we obtained and on the means of the D.V.s and the Mediator obtained for the Reggaeton and Rock samples, we decided to conduct a post-hoc power analysis to determine which sample size we would have needed to find significance. Post hoc power is the retrospective power of an observed effect based on the sample size and parameter estimates derived from a given data set. It is recommended as a follow-up analysis, especially if a finding is nonsignificant. For this study, we used power analysis in a slightly different way. Instead of trying to find the power of the analysis given the sample size and parameter estimates, what we did was calculate the sample size given the parameter estimates we found in the study, when power *(1 – Type II Error)* = 0.8.

The following results were obtained using the G-power statistical software (Erdfelder, Lang, Buchner, 2007):

Figure 15: Sample sizes for DVs and Reactance (Rock)

Willingness to go Premium	**351**
Behavioral Intention	**2982**
Reactance	**1398**

* Sample sizes when alpha (Type I Error) = 0.05, and power (1 – β) = 0.8

Figure 16: Sample sizes for DVs and Reactance (Reggaeton)

Willingness to go Premium	**264**
Behavioral Intention	**114**
Reactance	**84**

Sample sizes when alpha (Type I Error) = 0.05, and power (1 – β) = 0.8

Table 15 & 16 portray the expected sample sizes when *alpha (Type I Error)* = 0.05, and *power (1 – Type II Error)* = 0.8 for the Rock and Reggaeton samples. Unfortunately, the difference in liking between Rock and Reggaeton reduced our initial sample, and analyses had to be run in samples with low statistical power.

5.6. Results analysis

In order to test the hypotheses with solid empirical methods, we decided to conduct three different follow-up analyses. The first two follow-up analyses that we implemented are part of the bigger category defined as *Planned Contrasts*. These types of analysis are implemented when only a few predicted or *a priori* hypotheses are of interest (Newsom, 2020). This is the case for this study, as the group means for our variables of interest (the D.V.s and the Mediator) were predicted to be different in advance. Although we could use different t-tests to carry out our analysis, this would inflate the familywise error rate. Therefore, a way to contrast different groups without inflating the Type I error rate is to break down the variance accounted for by the model into component parts (Field, 2012). The difference between planned comparisons and post-hoc test can be likened to the difference between one and two-tailed tests in that planned comparisons are done when you have specific hypotheses that you want to test, whereas post hoc tests are done when you have no specific hypotheses (Field, 2012).

The third follow-up analysis we implemented is a mediation model, which seeks to identify and explain the mechanism or process that underlies an observed relationship between an I.V. (in this case, the experimental groups) and a D.V. (in this case, *Willingness to go Premium* and *Behavioral Intention* towards the brand) via the inclusion of a third hypothetical variable, known as the mediator variable (in this case, *Reactance*). Rather than a direct causal relationship between the independent variable and the dependent variable, a mediation model proposes that the independent variable influences the mediator variable, which in turn influences the dependent variable (MacKinnon, 2008).

The full R script of the statistical analyses we ran can be found in Appendix 8.7.

Linear Trend Analysis

The first contrast we implemented is the polynomial contrast, which tests for trends in the data, and in its most basic form it looks for a linear trend (Field, 2012). It represents a simple proportionate change in the value of the dependent variable across ordered categories. A quadratic trend is where there is one change in the direction of the line (e.g., the line is curved in one place). An example of this might be a situation in which a drug enhances performance on a task at first, but then as the dose increases the performance drops again. To find a quadratic trend we need at least three groups, because in the two-group situation there are not enough categories of the independent variable for the means of the dependent variable to change one way and then another (Field, 2012). As happens in this study, polynomial trends should be runned only if it makes sense to order the categories of the independent variable.

After coding the predictor variable groups in the orders we hypothesized (TG2, Control, TG1 for the D.V.s, and TG1, Control, TG2 for the Mediator)[7], we setted the

[7] We expect predictor variables (Willingness to go Premium & Behavioral Intention towards the brand) to be the lowest in Treatment group 2, and increase for Control group and Treatment group 1 (in that

contrast attribute of the predictor variable, creating a contrast matrix for a quadratic polynomial with 3 experimental groups (Field, 2012). Later, we created an ANOVA model to test whether the means of the D.V.s and the Mediator increase across groups in a linear way. This procedure was repeated for the *Reggaeton* and *Rock* samples. For the former, the results exhibited a marginal significant effect of the Experimental Groups on Behavioral Intention towards the brand, $t(39) = 1.69$, $p < 0.1$ (two-tailed), and on Reactance, $t(39) = 1.97$, $p < 0.1$ (two-tailed) - see Appendix 8.7.2. This means that as the experimental groups changed from TG2 to Control group to TG1, *Behavioral Intention* increased proportionately, and that as the experimental groups changed from TG1 to Control group to TG3, *Reactance* also increased proportionately. When we repeated the analyses for the *Rock* sample, no significant effects were found.

Congruent vs Incongruent contrasts

Following the linear trend analysis, we decided to implement a contrast to examine not only trends in the data, but also to check whether congruence among groups has an impact on the D.V.s and the Mediator. Planned comparisons break down the variations due to the experiment in component parts (Field, 2012). Our objective was to contrast the "incongruent" groups (i.e., control and treatment 2) with the "congruent" group (i.e., treatment 1) to reveal more information in the experiment data. The experimental variance is broken down to look at how much variation is created by the incongruent groups compared with the congruent group.

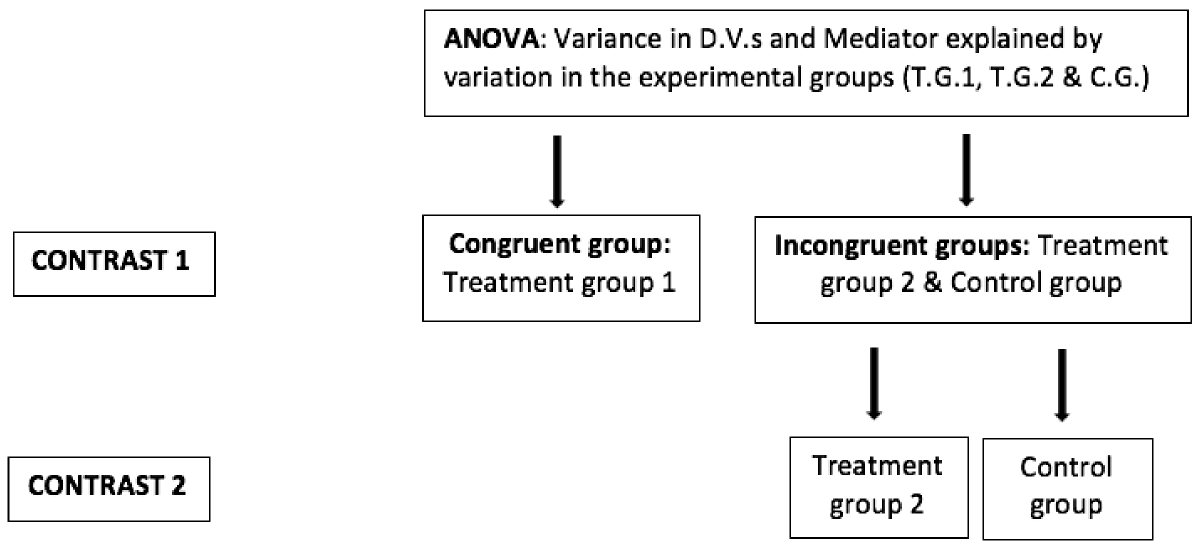
We previously hypothesized that our independent variables are the 3 experimental groups, labeled as "Treatment group 1", "Treatment group 2" & "Control group". As we explained, TG1 participants were listening to a song of their preferred genre to be interrupted with a congruent ad song, while TG2 participants were listening to a song of their preferred genre to be interrupted with an incongruent ad song. Meanwhile, CG surveyed people who were listening to a song of their prefered genre were interrupted with a song-free ad. Therefore, TG1 can be understood as our "congruent group", as the ad displayed in the survey has a song from the same music genre that the participant had been listening to. On the contrary, participants in

order). On the other hand, we expect the mediator (Reactance) to be the lowest in Treatment group 1, and increase for Control group and Treatment group 2 (in that order).

the TG2 and CG were interrupted with an ad that had a song from the opposite music genre, or with an ad that had no music, so they can be assigned as part of the "incongruent group".

To execute the contrast, we decided to run two planned contrasts: first, we wanted to check if the presence of congruency in groups had an impact on the D.V.s and the Mediator. Second, we wanted to know if differences within the incongruent groups (T.G.2 & C.G) were statistically significant.

Figure 17: Planned contrasts squeme

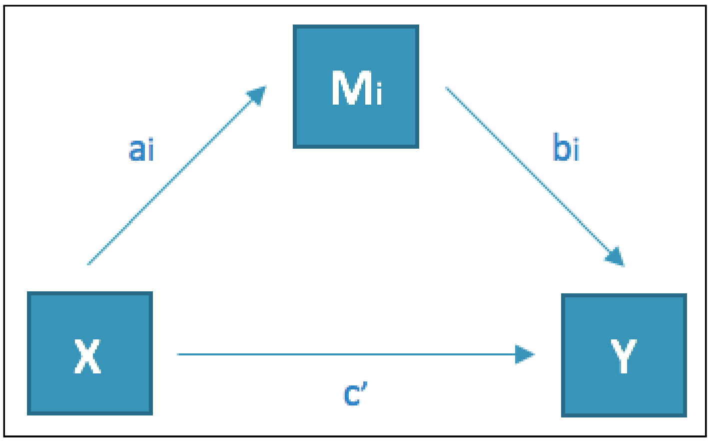


For the reggaeton sample, planned contrasts revealed that incongruence in groups (Treatment group 2 & Control group) was associated with a marginal significant increase in *Reactance* compared to Treatment group 1, $t(39) = 1.87$, $p < 0.1$ (two-tailed), as well as approaching significance in *Behavioral Intention,* $t(39) = -1.15$ (two-tailed) - see Appendix 8.7.2. For the rock sample, planned contrasts revealed that the difference in *Willingness to go Premium* was approaching significance when contrasting the two incongruent groups (Treatment group 2 vs. Control group), $t(105) = 1.6$ (two-tailed) - see Appendix 8.7.2.

Mediation model

In this study, our interest lies not only in the discovery that the assignment to experimental groups affects the respondents *Willingness to go Premium* and their *Behavioral Intention* towards the brand, but we are also interested in the *process* that produces this effect (Preacher & Hayes, 2004). A variable may be called a *mediator* "to the extent that it accounts for the relation between the predictor and the criterion" (Baron & Kenny, 1986). The following figure represents the simplest form of mediation, which occurs when one variable (M) mediates the effect of X on Y.

Mediation model statistical diagram



Indirect effect of X on Y through Mi = a_ib_i
Direct effect of X on Y = c'

We denote c' as the *direct effect* of X on Y after controlling for M. According to Baron and Kenny, variable M is considered a mediator if:

(1) X significantly predicts Y

(2) X significantly predicts M

(3) M significantly predicts Y controlling for X

To run the mediation model, we use model 4 from the macro PROCESS for R developed by Hayes (Hayes, 2013). We run one model for each one of the samples (rock & reggaeton), and one model for each one of the dependent variables. The independent variable and the mediator were kept the same for all the 4 models we ran. When specifying the model using the macro, we asked the program to mean-center all continuous variables and we specified that our independent variable was categorical. Finally, we set the bootstrap resampling to 10.000 iterations. Unfortunately, none of the models we ran gave statistically significant results. Meaning that we couldn't find support for our hypothesis that Psychological Reactance was the process underlying our effect of the assignment of people in experimental groups in the D.V.s. Unfortunately, measuring Reactance is not an easy task, and research still doesn't agree on the appropriate factor structure (i.e., the correlational relationship between a number of variables that are said to measure a particular construct) to achieve this goal. In fact, measurements of the state of reactance are rare, possibly because Brehm conceptualized reactance as “an intervening, hypothetical variable” that cannot be measured directly (Brehm & Brehm, 1981).

6. Conclusion

The present investigation has focused on gathering insights regarding the interruption of voice over ads in audio streaming services, and the effects in consumer behavior when manipulating the background music of the ad. We developed this study through analytical techniques applied to a research question belonging to psychology and behavioral economics disciplines.

Initially, we tried to establish opposition between music genres through a K-means clustering approach, but the algorithm results weren't as we expected, and we were forced to change the method into a sociological one. Therefore, the first lesson we took from the analysis was the importance of consumer's perception, and the impossibility to rely only on numeric interpretation of the data when trying to answer our research question.

Once we succeeded in identifying the tuple of music genres that people perceived as being the most different/opposite, we carried out the main experiment. Its objective was to understand the effects in consumers' *Willingness to go premium* and their *Behavioral Intention* towards the brand providing the service, when manipulating the background music of the ad. We also proposed that the effect of the independent variable on the depent variables was being mediated by *Reactance*. Although results were not statistically significant for the mediation model, we found that the change between experimental groups had significant effects on *Reactance* and *Behavioral Intention* towards the brand in the reggaeton sample. As the experimental groups changed from TG2 to Control group to TG1, *Behavioral Intention* increased proportionately, and as the experimental groups changed from TG1 to Control group to TG3, *Reactance* also increased proportionately. Meaning that when consumers were interrupted by an ad with no background music or by an incongruent ad song, *Reactance* increased and *Behavioral Intention* decreased proportionately (e.g., Reactance increased and Behavioral Intention decreased for consumers listening to Rock music to be interrupted by an ad with no background music, and the increase/decrease was proportionately higher when the ad had Reggaeton background music).

We believe that these results can be very helpful to audio streaming platforms providing freemium services. Managers from these types of companies can find very useful knowing that incongruence in ads may help to attract consumer's attention towards the brand (Dahlén, Lange, Sjodin & Torn, 2012), but this is a risky strategy. Managers from these companies might believe that incongruence in the background music displayed by an ad (e.g., playing Rock when a consumer is listening to Reggaeton) will be more effective in catching consumers' attention when displaying ads, but this can backfire. As we found for the Reggaeton sample collected in this experiment, incongruence in ad's background music increases Psychological Reactance and decreases Behavioral Intention towards the brand providing the service. Therefore, a consumer may be less willing to consider brand's that provide free streaming audio services that display ads with background music that differ from the one the consumer is listening to.

Regarding the limitations of our study, we identified two main constraints:

- There was a difference in liking we did not predict because we did not pre-test for song liking after selecting the songs for the main experiment. This difference in liking prevented us from analyzing the Reggaeton and Rock sample jointly, which substantially reduced the sample size of the two groups. Given the small sample size, and after running a post-hoc power analysis (see Section 5.5), we could confirm that we failed to reach full statistical significance because we were underpowered to do so. Future studies should pretest songs to make sure that participants that choose the genre like them so that analyses can be run on a bigger sample.

- Lack of realism: our scenario tried to mimic the real-life experience of a person using a streaming service. However, experiments are unnatural by definition, so we needed to sacrifice some realism to be able to study this phenomenon in an experimental setting that allowed us to test for causality.

We also understand that there are many opportunities to improve and expand the present study. The improvement we think could be the most valuable for this analysis is to adjust the experiment so that the simulacrum of Megatunes providing it's service gets closer to a real-life situation. For example, the intention by which people listen to music is very important if we want results to be scaled to reality. People usually listen to music because they feel like doing so, so the survey should try to replicate this. Therefore, a good opportunity of improvement we can think of is to enable surveyed people to choose the songs they want to listen to. This would probably avoid the *liking of songs* variable to be different among experimental groups, and a conjunct analysis could be implemented, which unfortunately wasn't our case.

7. Bibliography

Alpert, J. I., & Alpert, M. I. (1990). Music influences on mood and purchase intentions. *Psychology & Marketing*, *7*(2), 109-133.

Baron & Kenny (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182.

Brehm & Brehm (1981). Psychological Reactance: A theory of freedom and control. New York: Academic Press.

Brehm, J.W. (1966). A theory of psychological reactance. New York: Academic Press.

Clee & Wicklund (1980). Consumer Behaviour and Psychological Reactance. Journal of Consumer Research. Vol 6, No. 4, 389-405.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. psychometrika, 16(3), 297-334.

Edwards, Li & Lee (2002). Forced Exposure and Psychological Reactance: Antecedents and Consequences of the Perceived Intrusiveness of Pop-Up Ads. Journal of Advertising, Volume XXXI, Number 3.

Eriksson (2020). The editorial playlist as container technology: On Spotify and the logistical role of digital music packages. *Journal of Cultural Economy,* 13(4), 415-427.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Field, Field, Miles (2012). Discovering Statistics Using R.

Fuentes, Hagberg & Kjellberg (2019). Soundtracking: music listening practices in the digital age. *European Journal of Marketing*, 53(3), 483-503.

Hayes (2013). Introduction to Mediation, Moderation and Conditional.

Hracs & Webster (2021). From selling songs to engineering experiences: Exploring the competitive strategies of music streaming platforms. *Journal of Cultural Economy*, 14(4), 240-257.

Kellaris, J. J., Cox, A. D., & Cox, D. (1993). The effect of background music on ad processing: A contingency explanation. Journal of Marketing, 57(4), 114-125.

McKinnon (2008). Introduction to Statistical Mediation Analysis.

Micael Dahlén Ph.D., Fredrik Lange Ph.D., Henrik Sjödin & Fredrik Törn (2005)
Effects of Ad-Brand Incongruence, Journal of Current Issues & Research in Advertising, 27:2, 1-12, DOI: 10.1080/10641734.2005.10505178.

Morris & Powers (2015). Control, curation and musical experience in streaming music services. *Creative Industries Journal*, 8(2), 106-122.

Negrón - Muntaner & Z. Rivera (2016). Reggaeton Nation. *NACLA Report on the Americas.* 35-39.

Newsom (2020). Univariate Quantitative Methods.

Preacher & Hayes (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. 36(4), 717-731.

Shen & Dillard (2005). On the Nature of Reactance and Its Role in Persuasive Health Communication. 72(2):144-168

Steindl, Christina; Jonas, Eva; Sittenthaler, Sandra; Traut-Mattausch, Eva; Greenberg, Jeff (2015). "Understanding Psychological Reactance". *Zeitschrift für Psychologie*. 223 (4): 205–214.

Webster (2019). Music on-demand: A commentary on the changing relationship between music taste, class and consumption in the streaming age. *Big data Society* 6(2), 1-5

Wilkinson (2000). Chinese History: A New Manual.

Wong (2011). Visualising music: the problems with genre classification.

8. Appendix

8.1. Spotify Audio Features

Audio Features Key	Audio Features Type
danceability: 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.	**Float**
energy: 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,	**Float**

perceived loudness, timbre, onset rate, and general entropy.	
key: The key the track is in. Integers map to pitches using standard Pitch Class notation. E.g. 0 = C, 1 = C♯/D♭, 2 = D, and so on.	**Integer**
loudness: The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing 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.	**Float**
mode: Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.	**Integer**
speechiness: Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, 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.	**Float**
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.	**Float**
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.	**Float**
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 strong likelihood that the track is live.	**Float**

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).	**Float**
tempo: The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.	**Float**

8.2. Datasets variables

8.2.1. "Viral songs" dataset

track_name: String variable containing the title of the track in Spotify.

artist_name: String variable containing the artist name in Spotify.

date: Date variable (format = yyyy-dd-mm) containing the date of "Spotify Viral 50" weekly chart.

region: String variable containing the country code of the "Spotify Viral 50" weekly chart.

spotify_id: String variable containing the ID assigned to songs, formed by a unique combination of numbers and words.

artist_genre: List of strings variable containing the genres/subgenres associated with artists in Spotify.

8.2.2. "Songs attributes" dataset

artist: String variable containing the artist name in Spotify.

song_name: String variable containing the name of the song in Spotify.

song_id: String variable containing the ID assigned to songs, formed by a unique combination of numbers and words.

danceability: Float variable detecting how danceable is a track. A value of 0.0 is least danceable and 1.0 is most danceable.

energy: Float variable representing the level of intensity and activity in a track. It ranges from 0.0 (least energetic) to 1.0 (most energetic).

key: Integer variable containing the key the track is in using standard Pitch Class notation.

loudness: Float variable containing overall loudness of a track in decibels (dB). Values typically range between -60 and 0 dB.

mode: Integer variable containing the modality of a track (Major is 1, Minor is 0).

speechiness: Float variable detecting the presence of spoken words in the track. It ranges from 0.0 (least spoken) to 1.0 (most spoken).

acousticness: Float variable ranging from 0.0 to 1.0, representing the level of confidence a track is acoustic (1.0 is high confidence).

instrumentalness: Float variable that predicts whether a track contains vocals. It ranges from 0.0 to 1.0. The closer to 1.0, the greater likelihood the track contains no vocals.

liveness: Float variable that detects the presence of an audience in the recording. It ranges from 0.0 to 1.0. The closer to 1.0, the greater likelihood the track is live.

valence: Float variable that describes the musical positiveness conveyed by a track. It ranges from 0.0 to 1.0. The closer to 1.0, the more positive a track is.

tempo: Float variable that estimates the tempo of a track in beats per minute (BPM). Values typically range between 110 and 130.

artist_genre: Factor variable containing the music genre of the artist. It has 5 levels: Rock, Pop, Reggaeton, Hip-Hop/Rap, Electronic.

8.3. Final list of variables used for Clustering

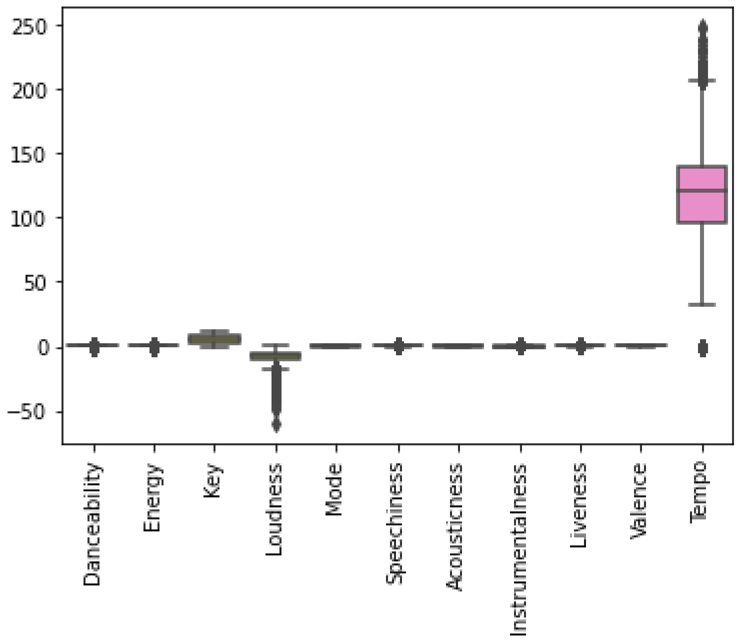
The following list contains the variables from the final dataset (80.000 rows) used for the clustering model (k-means). The 11 variables are the exact same ones listed in *Section 8.1: Spotify Audio Features*, but their values were normalized (mean = 0, standard deviation = 1) to improve the performance of the algorithm.

1. **danceability**
2. **energy**
3. **key**

4. **loudness**
5. **mode**
6. **speechiness**
7. **acousticness**
8. **instrumentalness**
9. **liveness**
10. **valence**
11. **tempo**

8.4. Additional graphs (exploratory analysis)

Figure 18 & 19: Audio features values before and after scaling - boxplot



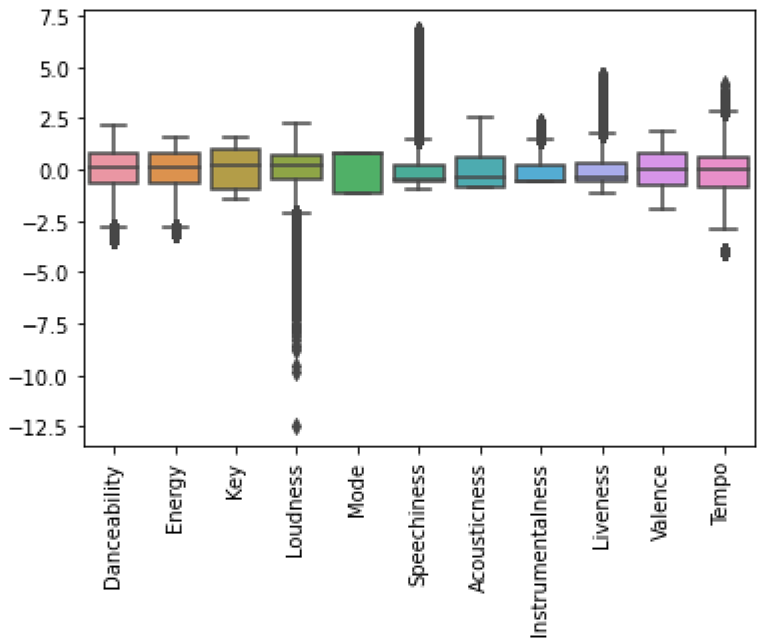


Figure 20: Elbow method (sum of quadratic distances of observations to nearest centroid)

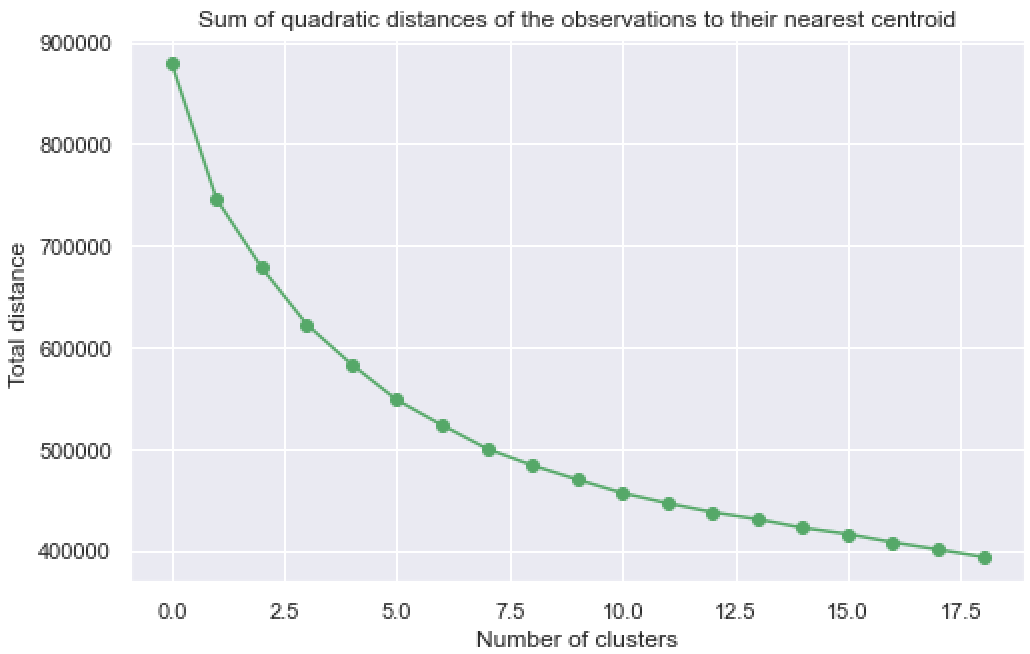
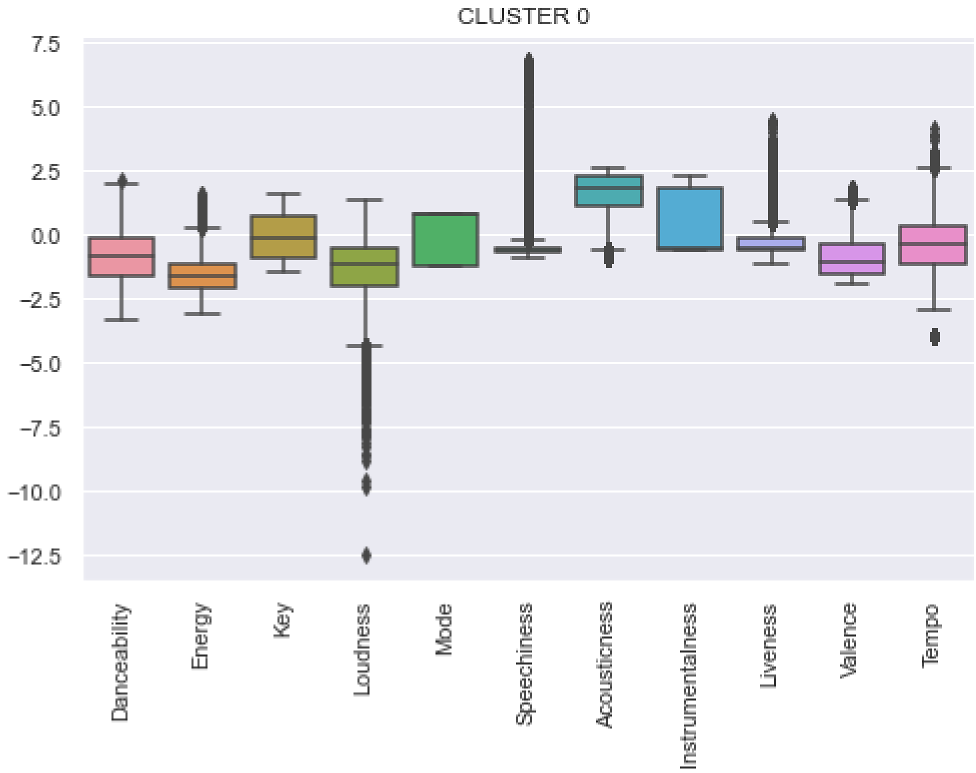
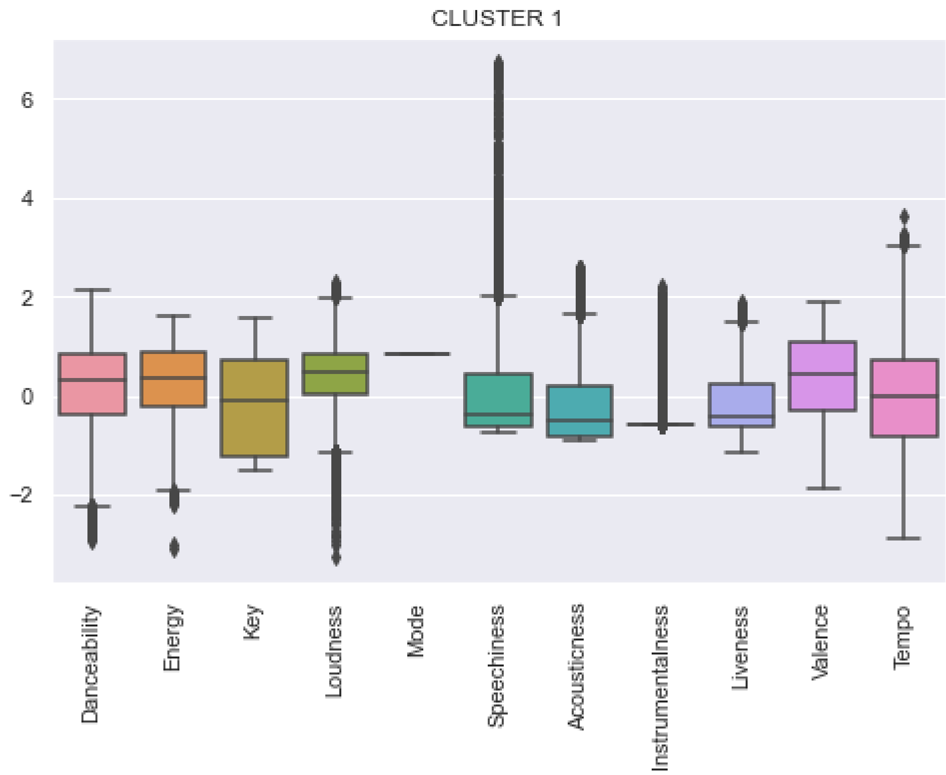
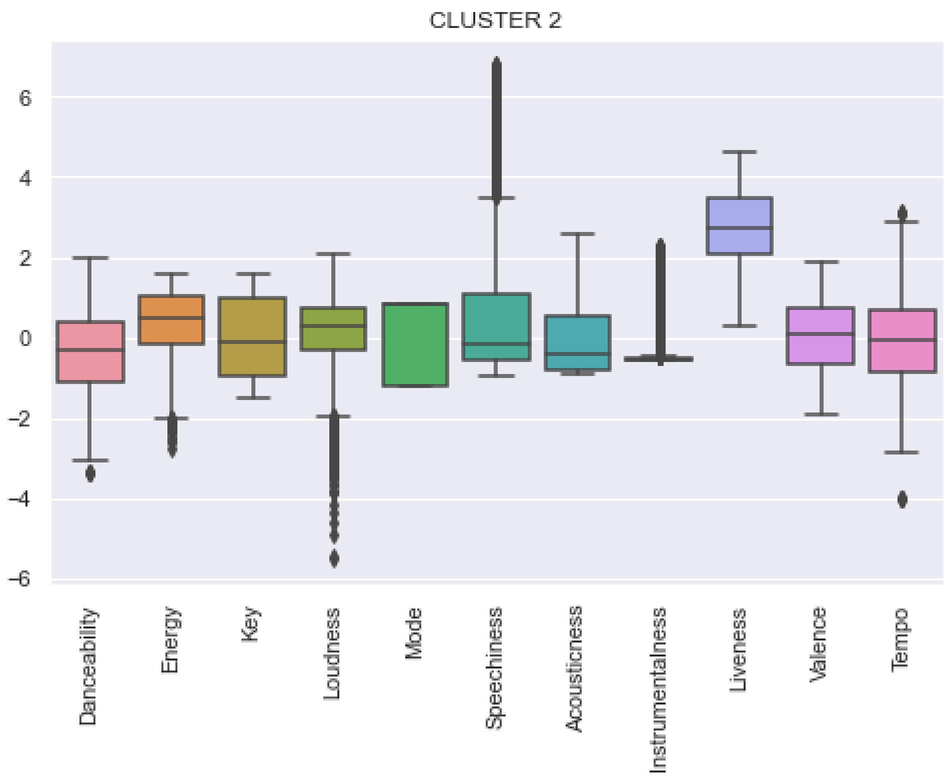
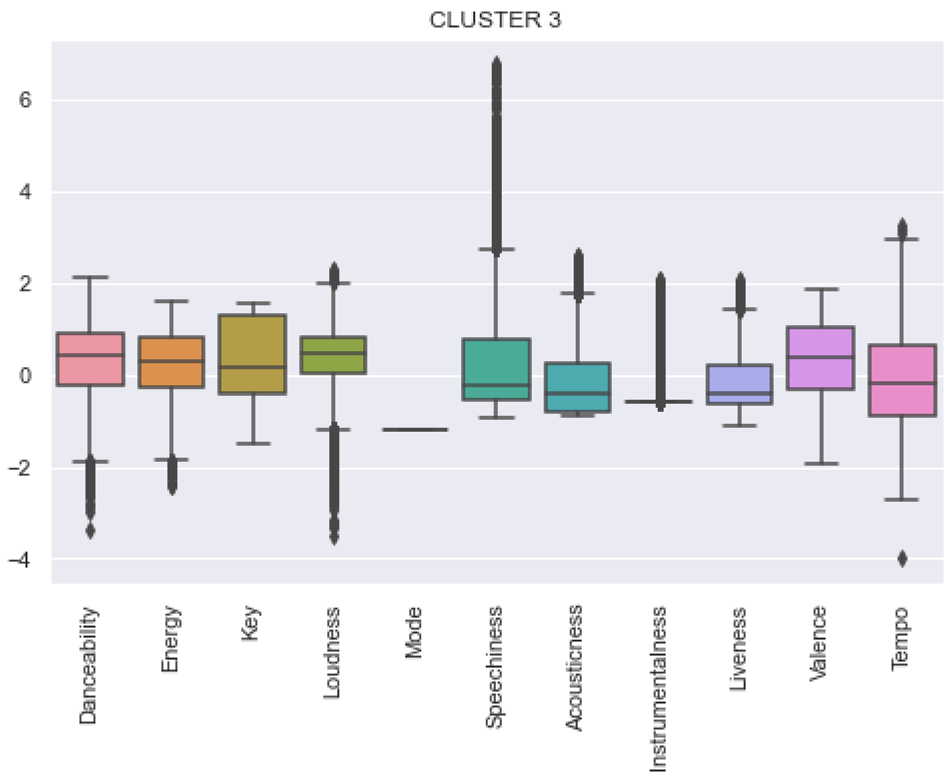


Figure 21: Audio features values after scaling - boxplot for each cluster






CLUSTER 2
6
4
2
0
-2
-4
-6
Danceability
Energy
Key
Loudness
Mode
Speechiness
Acousticness
Instrumentalness
Liveness
Valence
Tempo


CLUSTER 3
6
4
2
0
-2
-4
Danceability
Energy
Key
Loudness
Mode
Speechiness
Acousticness
Instrumentalness
Liveness
Valence
Tempo

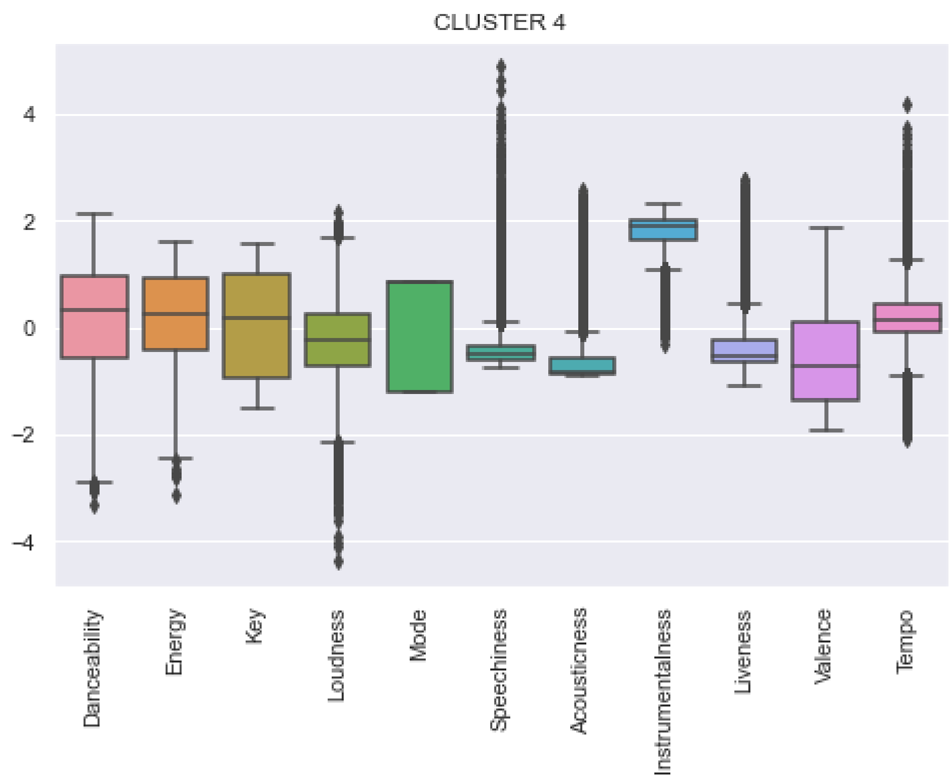


Figure 22: Music genre chosen in experiment - bar chart

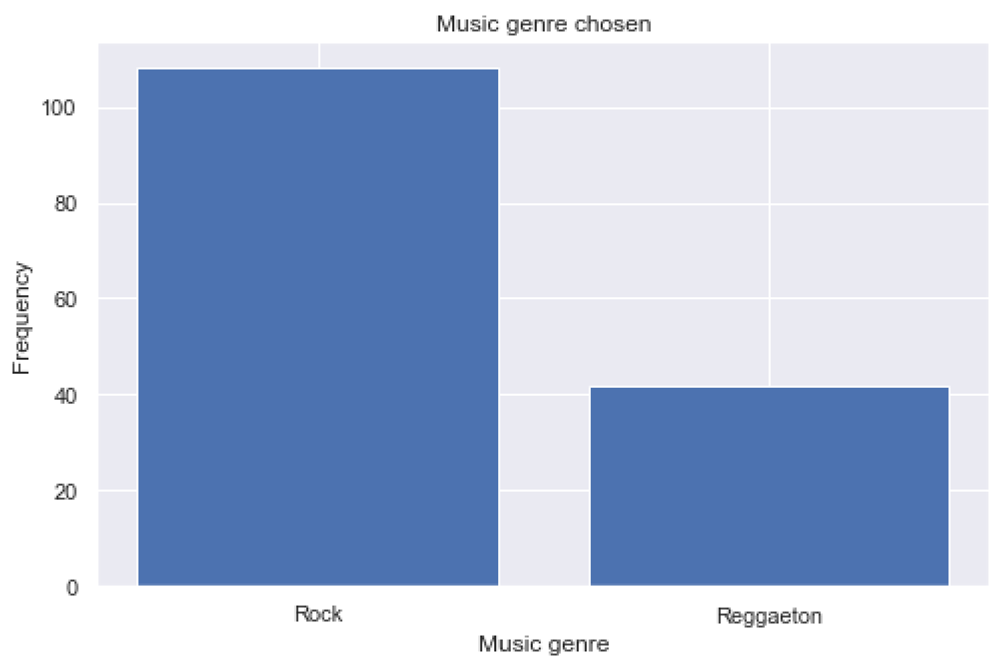
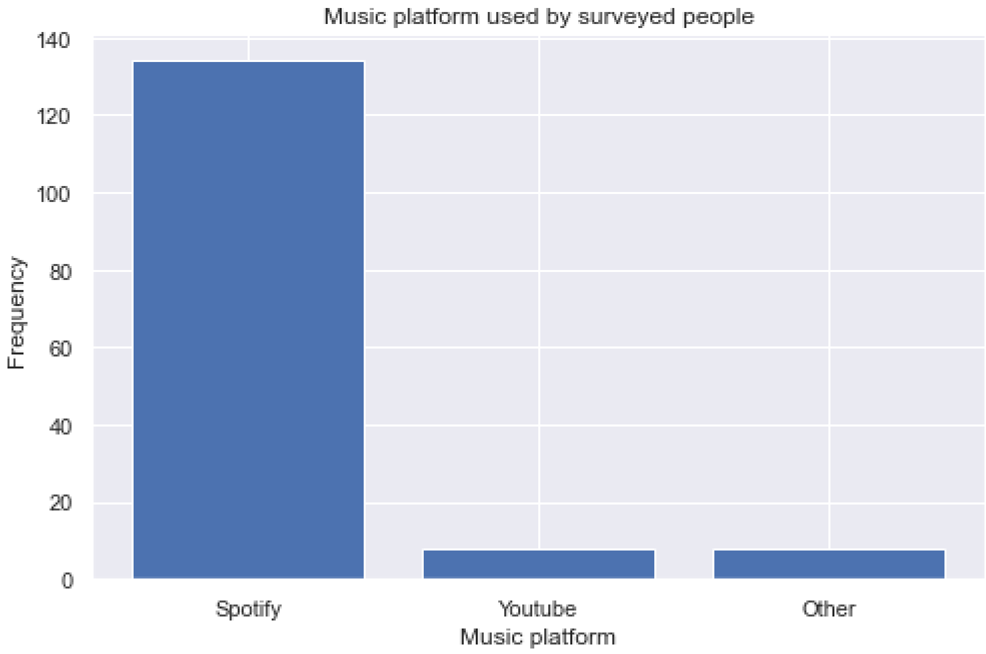


Figure 23: Music platform used by participants - bar chart



8.5. Pretest - questions from the pretest

Estilos/géneros musicales

Respecto a sus preferencias en estilos/géneros musicales, indique cuánto le gustan los siguientes estilos/géneros del 1 (muy poco) al 7 (muchísimo).

- **Rock**
- **Pop**
- **Hip Hop / Rap**
- **Reggaetón**
- **Electrónica**

[page break]

Por favor, indique de los siguientes estilos/géneros musicales cuáles 2 le parecen los más DIFERENTES/OPUESTOS entre sí:

(Pregunta de opción múltiple con posibilidad de seleccionar dos)

- **Rock**
- **Pop**
- **Hip Hop / Rap**
- **Reggaetón**
- **Electrónica**

[page break]

Por favor, indique de los siguientes estilos/géneros musicales cuáles 2 le parecen los más SIMILARES/IGUALES/PARECIDOS entre sí:

(Pregunta de opción múltiple con posibilidad de seleccionar dos)

- **Rock**
- **Pop**
- **Hip Hop / Rap**
- **Reggaetón**
- **Electrónica**

[page break]

Información socio demográfica

- **Edad**
- **Género (Masculino, Femenino, Otro)**

8.6. Main experiment - questions from the main experiment

Nos encontramos frente al lanzamiento de una aplicación de streaming de música llamada MEGATUNES. En este caso estamos buscando simular tu experiencia de usuario, haciéndote escuchar algunas canciones para luego hacerte algunas preguntas. Es importante que te concentres y trates de imaginarte que realmente estás conociendo por primera vez esta aplicación de música.

En primer lugar, te voy a pedir que elijas entre estos dos géneros musicales, de acuerdo a tus preferencias y a las ganas que tengas de escuchar un género u otro en este momento.

- **Rock**
- **Reggaeton**

Te pido por favor que te pongas los auriculares y cuando estés listo para escuchar la música pongas continuar, tocando el siguiente botón.

> ***>> respondents are randomly assigned to one of the three experimental groups***

[page break]

Dale play a la siguiente playlist y disfruta de tu experiencia en Megatunes. Escuchá el audio de manera completa (no dura más de 3 minutos) ya que en las siguientes

páginas te haremos algunas preguntas, y te pediremos que evalúes el servicio de streaming.

> ***>> an audio file is displayed and a sequence of 3 songs of the chosen music genre is played, followed by an ad interruption***

[page break]

Si esta hubiese sido tu primera y única experiencia con Megatunes,

- **¿Qué tan probable es que te pases al servicio Premium? (1 = nada probable, 7 = muy probable)**

Califique las siguientes afirmaciones en orden de probabilidad (1 = nada probable, 7 = muy probable):

- **es probable que busque más información sobre Megatunes**
- **es probable que considere Megatunes la próxima vez que piense en suscribirme a un servicio de streaming**
- **es probable que consulte las opiniones/referencias de Megatunes**
- **es probable que recomiende Megatunes a un amigo/a**

[page break]

Por favor, indique en qué medida cree que cada una de las palabras a continuación describe el anuncio que acaba de escuchar sobre Megatunes premium. Lo que nos interesa es su opinión sobre el anuncio, no sobre la marca o la clase de producto (1 = nada bien, 2 = no muy bien, 3 = bastante bien, 4 = muy bien, 5 = extremadamente bien):

- **intrusivo**
- **forzado**
- **interesante**
- **creíble**
- **interferente**
- **invasivo**
- **perturbador**
- **convincente**
- **distrayente**
- **molesto**
- **informativo**

[page break]

¿Conocías las canciones que escuchaste en la playlist seleccionada? (1 = para nada/ni una, 7 = mucho/todas)

¿Cuánto te gustan las canciones que escuchaste? (1 = para nada, 7 = muchísimo)

¿Qué tan difícil te fue elegir entre los dos géneros que te propusimos (rock vs reggaeton)? (1 = muy fácil, 7 = muy difícil)

¿Qué tan diferentes te resultan los dos géneros que te propusimos (rock vs reggaeton)? (1 = muy similares, 7 = muy diferentes)

Respecto de tus preferencias en estilos/géneros musicales, indique cuánto te gustan los siguientes estilos/géneros (1 = muy poco, 7 = muchísimo):

- **Rock**
- **Reggaeton**
- **Pop**
- **Hip Hop / Rap**
- **Electrónica**

¿Qué tan frecuentemente escucha música?

- **Nunca**
- **A veces**
- **Frecuentemente**
- **Todo el tiempo**

¿Utiliza alguna plataforma digital para escuchar música? (puede seleccionar más de una opción)

- **No, ninguna**
- **Si, Spotify**
- **Si, YouTube**
- **Si, Soundcloud**
- **Si, otra (cuadro de texto)**

¿De la(s) plataforma(s) que mencionaste arriba, de cual(es) tienes la versión premium?

- **Spotify**
- **Youtube**
- **Soundcloud**

- Otra
- Ninguna

[page break]

Información socio demográfica

- Edad
- Género (Masculino, Femenino, Otro)

No dudes en compartir con nosotros cuál crees que puede ser el propósito de este estudio. (opcional)

8.7. Main experiment: Analysis of variance for Willingness to go Premium, Behavioral Intention towards the brand and Reactance

8.7.1. ANOVA

One - way ANOVA: Effect of *musical choice* in *liking of songs*

	DF	Sum Sq	Mean Sq	F value	p
Music choice	1	28.0	27.985	10.65	0.00138 **
Residuals	144	378.5	2.629		

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 150

One - way ANOVA: Effect of *musical choice* in *familiarity of songs*

	DF	Sum Sq	Mean Sq	F value	p
Music choice	1	48.05	48.05	22.99	3.93e-06 ***
Residuals	148	309.34	2.09		

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 150

8.7.2. Contrasts

Linear Trend Contrast: Effect of *experimental groups* in *willingness to go premium (reggaeton sample)*

	Estimate	Std. Error	t	p
(Intercept)	4.2095	0.3020	13.939	<2e-16 ***
EG. DV. L	0.5455	0.5567	0.980	0.333

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 42

Linear Trend Contrast: Effect of *experimental groups* in *behavioral intention (reggaeton sample)*

	Estimate	Std. Error	t	p
(Intercept)	4.06177	0.23309	17.426	<2e-16 ***
EG. DV. L	0.72731	0.42970	1.693	0.0985*

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 42

Linear Trend Contrast: Effect of *experimental groups* in *reactance (reggaeton sample)*

	Estimate	Std. Error	t	p
(Intercept)	2.65083	0.12299	21.553	<2e-16 ***
EG. DV. L	0.44567	0.22674	1.966	0.0565*

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 42

Linear Trend Contrast: Effect of *experimental groups* in *willingness to go premium (rock sample)*

	Estimate	Std. Error	t	p
(Intercept)	3.2212	0.1871	17.216	<2e-16 ***
EG. DV. L	-0.3927	0.3118	-1.260	0.211

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 108

Linear Trend Contrast: Effect of *experimental groups* in *behavioral intention (rock sample)*

	Estimate	Std. Error	t	p
(Intercept)	3.41215	0.15286	22.322	<2e-16 ***
EG. DV. L	-0.07741	0.25470	-0.304	0.762

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 108

Linear Trend Contrast: Effect of *experimental groups* in *reactance (rock sample)*

	Estimate	Std. Error	t	p
(Intercept)	2.71324	0.07987	33.970	<2e-16 ***
EG. DV. L	0.11303	0.13343	0.847	0.399

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 108

Grouped Contrast: Effect of *experimental groups* in *willingness to go premium (reggaeton sample)*

	Estimate	Std. Error	t	p
Intercept	4.20952	0.30199	13.939	<2e-16 ***
EGcontrast1	-0.24524	0.23007	-1.066	0.293
EGcontrast2	-0.03571	0.33882	-0.105	0.917

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 42

Grouped Contrast: Effect of *experimental groups* in *behavioral intention (reggaeton sample)*

	Estimate	Std. Error	t	p
Intercept	4.0618	0.2331	17.426	<2e-16 ***
EGcontrast1	-0.2691	0.1776	-1.515	0.138
EGcontrast2	-0.2212	0.2615	-0.846	0.403

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 42

Grouped Contrast: Effect of *experimental groups* in *reactance (reggaeton sample)*

	Estimate	Std. Error	t	p
Intercept	2.6508	0.1230	21.553	<2e-16 ***
EGcontrast1	0.1754	0.0937	1.872	0.0687
EGcontrast2	0.1040	0.1380	0.754	0.4555*

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 42

Grouped Contrast: Effect of *experimental groups* in *willingness to go premium (rock sample)*

	Estimate	Std. Error	t	p

Intercept	3.22123	0.18711	17.216	<2e-16 ***
EGcontrast1	0.05933	0.12941	0.458	0.648
EGcontrast2	0.37733	0.23407	1.612	0.110

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 108

Grouped Contrast: Effect of *experimental groups* in *behavioral intention (rock sample)*

	Estimate	Std. Error	t	p
Intercept	3.4121483	0.1528596	22.322	<2e-16 ***
EGcontrast1	0.0009459	0.1057224	0.009	0.993
EGcontrast2	0.1066426	0.1912235	0.558	0.578

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 108

Grouped Contrast: Effect of *experimental groups* in *reactance (rock sample)*

	Estimate	Std. Error	t	p
Intercept	2.71324	0.07987	33.970	<2e-16 ***
EGcontrast1	0.04710	0.05513	0.854	0.395
EGcontrast2	0.01856	0.10010	0.185	0.853

* Note: ***, **, * indicate significance of 1%, 5% and 10% respectively. N = 108

8.7.3. Mediation model

Mediation model: *experimental groups* → *reactance* → *behavioral intention (reggaeton sample)*

```
Outcome Variable: Reactance_negative_measure

Model Summary:
          R       R-sq        MSE          F        df1        df2          p
     0.3020     0.0912     0.5998     1.9572     2.0000    39.0000     0.1549

Model:
             coeff         se          t          p       LLCI       ULCI
constant    2.7222     0.1825    14.9128     0.0000     2.3530     3.0915
X1         -0.4222     0.3055    -1.3823     0.1747    -1.0401     0.1956
X2          0.2080     0.2760     0.7539     0.4555    -0.3502     0.7663

**************************************************************************
Outcome Variable: BI_composite_measure

Model Summary:
          R       R-sq        MSE          F        df1        df2          p
     0.2709     0.0734     2.1991     1.0032     3.0000    38.0000     0.4020

Model:
                                coeff         se          t          p       LLCI       ULCI
constant                       4.3893     0.9049     4.8506     0.0000     2.5574     6.2213
X1                             0.5279     0.5990     0.8812     0.3837    -0.6848     1.7406
X2                            -0.4138     0.5323    -0.7773     0.4418    -1.4913     0.6638
Reactance_negative_measure    -0.1379     0.3066    -0.4498     0.6554    -0.7586     0.4828

************************ TOTAL EFFECT MODEL ****************************
Outcome Variable: BI_composite_measure

Model Summary:
          R       R-sq        MSE          F        df1        df2          p
     0.2616     0.0685     2.1542     1.4329     2.0000    39.0000     0.2509

Model:
             coeff         se          t          p       LLCI       ULCI
constant    4.0139     0.3459    11.6028     0.0000     3.3141     4.7136
X1          0.5861     0.5789     1.0125     0.3175    -0.5848     1.7570
X2         -0.4425     0.5230    -0.8460     0.4027    -1.5004     0.6155
```

Mediation model: *experimental groups → reactance → willingness to go premium* (reggaeton sample)

```
Outcome Variable: Reactance_negative_measure

Model Summary:
          R       R-sq        MSE          F        df1        df2          p
     0.3020     0.0912     0.5998     1.9572     2.0000    39.0000     0.1549

Model:
              coeff         se          t          p       LLCI       ULCI
constant     2.7222     0.1825    14.9128     0.0000     2.3530     3.0915
X1          -0.4222     0.3055    -1.3823     0.1747    -1.0401     0.1956
X2           0.2080     0.2760     0.7539     0.4555    -0.3502     0.7663

**************************************************************************
Outcome Variable: premium

Model Summary:
          R       R-sq        MSE          F        df1        df2          p
     0.1948     0.0380     3.6746     0.4997     3.0000    38.0000     0.6847

Model:
                                coeff         se          t          p       LLCI       ULCI
constant                       3.3354     1.1697     2.8514     0.0070     0.9674     5.7034
X1                             0.8031     0.7743     1.0371     0.3062    -0.7645     2.3707
X2                            -0.1222     0.6881    -0.1776     0.8599    -1.5151     1.2707
Reactance_negative_measure     0.2442     0.3963     0.6160     0.5416    -0.5582     1.0465

************************* TOTAL EFFECT MODEL ****************************
Outcome Variable: premium

Model Summary:
          R       R-sq        MSE          F        df1        df2          p
     0.1684     0.0283     3.6161     0.5689     2.0000    39.0000     0.5708

Model:
              coeff         se          t          p       LLCI       ULCI
constant     4.0000     0.4482     8.9243     0.0000     3.0934     4.9066
X1           0.7000     0.7500     0.9333     0.3564    -0.8170     2.2170
X2          -0.0714     0.6776    -0.1054     0.9166    -1.4421     1.2992
```

Mediation model: *experimental groups → reactance → behavioral intention* (rock sample)

```
Outcome Variable: Reactance_negative_measure

Model Summary:
          R       R-sq        MSE          F        df1        df2          p
     0.0865     0.0075     0.6761     0.3916     2.0000   104.0000     0.6769

Model:
              coeff         se          t          p       LLCI       ULCI
constant     2.7418     0.1477    18.5660     0.0000     2.4489     3.0346
X1          -0.1227     0.1978    -0.6203     0.5364    -0.5151     0.2696
X2           0.0371     0.2002     0.1854     0.8533    -0.3599     0.4341

***********************************************************************
Outcome Variable: BI_composite_measure

Model Summary:
          R       R-sq        MSE          F        df1        df2          p
     0.0583     0.0034     2.5378     0.1170     3.0000   103.0000     0.9499

Model:
                                coeff         se          t          p       LLCI       ULCI
constant                       3.4728     0.5943     5.8435     0.0000     2.2941     4.6515
X1                             0.0964     0.3840     0.2509     0.8024    -0.6653     0.8580
X2                             0.1958     0.3879     0.5047     0.6148    -0.5736     0.9652
Reactance_negative_measure    -0.0607     0.1900    -0.3194     0.7501    -0.4375     0.3161

************************ TOTAL EFFECT MODEL ***************************
Outcome Variable: BI_composite_measure

Model Summary:
          R       R-sq        MSE          F        df1        df2          p
     0.0491     0.0024     2.5159     0.1256     2.0000   104.0000     0.8821

Model:
              coeff         se          t          p       LLCI       ULCI
constant     3.3065     0.2849    11.6065     0.0000     2.7415     3.8714
X1           0.1038     0.3817     0.2720     0.7862    -0.6530     0.8607
X2           0.1935     0.3862     0.5012     0.6173    -0.5723     0.9594
```