Shruti Sindhi

Griffith College

Email: shrutivijay.sindhi@student.griffith.ie

PART 1

Outline of the selected Domain:

Music Industry is 23.1-billion-dollar business, as per 2020 records. Streaming alone comprises of 56 percent of the total business globally. Listening to mellow music promotes well-being, and calmness of individuals. Music is in the air, and it is one of the largest markets in the world. The world music industry is huge, it is not clear where to exactly to draw the line when sizing it. The detailed explanation of the music industry will include actual customer's expense (it includes concert tickets, streaming subscription and more), B2B Licensing cash flow, ad-revenues of radio and other music related media. Music streaming is the king and more steadfast music consumption medium. There are multiple music consumption structure and none of them is simple, not even streaming.

Name of the Dataset:

Spotify Dataset (1960-2019)

Link of the Dataset:
https://www.kaggle.com/datasets/theoverman/the-spotify-hit-predictor-dataset

Introduction of the dataset:

The dataset that will be utilized is a song track dataset fetched with the aid of Spotify's Web API. Spotify's Web API is an interface that is available for free that is used by programs that can be used to retrieve and handle Spotify data over the internet. And so, it can be concluded it is a **real time data and reliable**. The dataset descends under multiple domains: Internet, Music, Business, and Art & Entertainment. The original dataset consists of **19 attributes**, out of which 10 consist of decimal values. For example, the value of the following 6 is Integer, and that of the last 3 is String. In addition, a new attribute called 'decade' was added to the final sheet as it was made using records for 6 decades from 1960 to 2010.

Total Number of Records: 41006

Attributes with numeric data type: 17

Attributes with string data type:3

Attributes comprising nominal data: 6

Attributes comprising ordinal data: 0 (There is no categorical field whose value needs to have a true and normal order.)

On quest of the dataset, it was learned that no attribute had any missing value.

19 new calculated and grouped features are generated using the available initial features. The uri feature was not used anywhere throughout the analysis. The usage of the chorus hits attribute was minimum.

Description of each attribute in the original dataset:

track: it includes the name of the song
artist: it reveals the artist's name who has sung the song.

(The artist dimension diverges into 2 attributes **leading artist** and **featuring artist** using custom split functionality available in tableau)

uri: it aids in identifying the source of the track.

danceability: this dimension specifies how danceable the track is based on various music elements that include tempo, rhythm, and beat stability. The value ranges between 0.0 and 1.0, where 0 is slightly danceable, and 1 is most danceable. The maximum danceable value found in the dataset is 0.9.

energy: the attribute represents a perceptual measure of the intensity and activeness of the track. Generally, the energetic tracks are fast, boisterous, booming, and noisy. For example, death metal has high energy. Features contributing to the variable include dynamic range, timbre, perceived loudness, onset rate, and general entropy. The track's energy can take any value between 0.0 and 1.0, where 0.0 means no energy and 1.0 indicates high power music.

key: it exhibits the estimated overall key of the track. Integer values are mapped to the pitches using standard pitch class notation. (No track is sung in one keynote. The keynote assigned to a music piece emerges the most in a song.)

loudness: it mentions the loudness of the track. It is measured in decibels (dB). It takes a value between -60 and 0 dB.

mode: a musical track has 2 modes; major or minor. Mode specifies the modality of the track. It is a scale through which a song's melodic content is procured. Major mode is characterized by 1 and minor by 0.

speechiness: this field perceives the presence of spoken words in a song. The closer the recording is to the speech; closer its value will be to 1. The list of speech-like recordings includes talk shows, audiobooks, and poetry. Some artists make speech tracks. Songs with values above 0.66 point out that the tracks are made up of too many spoken words. Records with values between 0.33 and 0.66 suggest that the tracks may encompass both music and speech. Music and speech can be in sections or layered depending upon the artist and the track. The immediate case can be exemplified with rap music. Values below 0.33 mean that the music piece is more musical and contains non-speech-like tracks (for example, piano, a guitar version of the tracks).

acousticness: the feature checks the probability that the track was produced using acoustic instruments, including voice. Its attributes have a confidence measure between 0.0 and 1.0. Value 1 under the acoustics column entails high confidence the track is acoustic.

instrumentalness: the feature foresees if the track consists of vocals, or it is absent. The sounds like 'aaha', 'aah', or 'ooh' are treated as instrumental. If the instrumentalness value is close to or equal to 1.0, then it means the track doesn't comprise any vocal content. The value of the record above or equal to 0.5 intends to represent the track is instrumental.

liveness: this variable detects the existence of the audience in the recording. A higher liveness value indicates that the track was performed live. The value of the attribute range between 0.02 and 0. 99. A record in the dataset with 0.8 and more liveness implies a strong likelihood that the track was played live.

valence: this feature takes any value from 0.0 to 1.0. It illustrates the musical positiveness conveyed by the track. High valence is associated with positive mood, whereas low valence is associated with negative mood. The positive and negative moods list is happy, cheerful, sad, angry, depressed, etc.

tempo: the overall value of the tempo of a track is computed in beats per minute (BPM). The minimum recorded value of the tempo is 46.8. Tempo indicates the speed/pace of the music.

duration_ms: this dimension mentions the duration of the track in milliseconds. The data type of the attribute is an integer.
time_signature: the time signature dimension of the dataset represents an estimated overall time signature. It is a notational convention in western music that specifies how many beats (pulses) are contained in each measure(bar). The maximum beats a measure of the track can hold is 5, and the minimum is 0.
chorus_hit: this feature displays the author's best estimate of when the chorus would start for a particular artist's track. This attribute represents the timestamp value of the third part of the track. The dimension was attained by the API call for Audio Analysis for a particular music piece. The chorus hit consists of decimal values between 0 and 263. It consists of a broad range of numerical, such that the analysis may be complex. The value should be grouped for better analysis.
sections: it represents the number of parts/sections the song is structured into for a better melody. The three main sections of the song are verse, chorus, and bridge. The value for this attribute lies between 1 and 169. So, it is not easy to interpret what these numerical values represent. The numerical matter is acquired from the data received by the API call for Audio Analysis of the track.
target: the value of the target variable can be either '0' or '1'. 1 suggests that the track has arisen in the weekly list of Hot-100 songs. The list gets issued by Billboards. Value 0 implies that the track is a flop.

Reason for selecting the dataset:

Considerable datasets are available for Art, Entertainment, and Music on numerous platforms. **The reason for selecting the Music category over Art and Entertainment is that the data of the other two categories did not have enough attributes for building operational, strategical, analytical, and predictive dashboards**. For example, on the Kaggle platform, there is a particular dataset by year, for instance, Top Track of 2019 or Top Track of 2017. Nevertheless, those datasets can only be used for making good predictions or predictive dashboards. Initially, the exploratory data analysis was started with Spotify – All-Time Top 2000s Mega Hit Dataset available on one of the free dataset websites. However, the Mega Dataset consisted of 18 attributes and a few thousand records, which did not satisfy the condition assigned for the task.

The final selected dataset has a meticulous range of fields to give more information about the track (song), but the database did not consist of field genre from a domain point of view. **Therefore, found a similar dataset consisting of genre fields and linked it to the original dataset in the tableau.** The link of similar dataset is as follows:

https://www.kaggle.com/code/akiboy96/spotify-song-popularity-genre-exploration/data

Problem Statement:

Music is a vocal or instrumental sound or combination of both fused in a way that produces a beautiful form of harmony and expression of emotion. Music is the language of the universe. It brings people together. It not only inspires people but also improves focus and memory and improves cognitive abilities. In addition, it sometimes lowers the blood pressure if listened correctly. In other words, music is not the solution to all the problems, but it can help tackle the issue better. The purpose of labouring with the Spotify dataset is to predict a hit song using different soundtrack features. Also, along the way, find the relationship between other elements in the suggested music dataset from 1960 to 2019. The database comprises American and Latin American artists only. The other sub-objective is to improve the research of **musicologists** employing exploratory and predictive analysis. Simultaneously enhance the creativity of the music virtuoso. The future goal of the study is to discover the hit predictor for each artist from 1960 to 2019 based on their song information available in the existing dataset.

List of Questions intend to investigate using the dataset:

How danceable were the track in each decade?

What was the energy of different artists in 6 decades?

How lively were the tracks based on the musical keys?

Which is the most sung genre by each decade?

Find the relationship between genre and valence dimension.

Which genre emerged the most in the weekly list of Hot-100 tracks (Issued by Billboards)?

Top 2-5 artists in each decade that appeared in the weekly list of Hot-100 tracks.

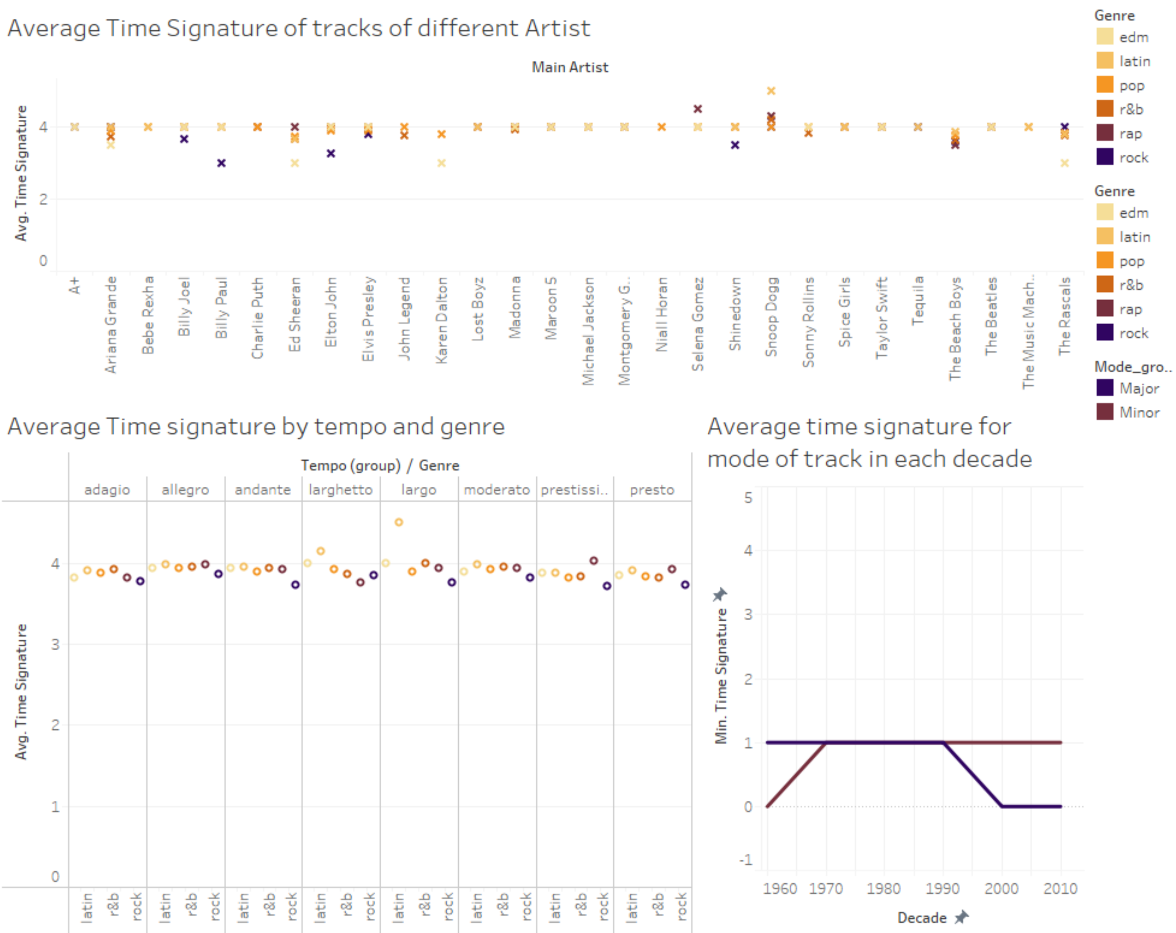
List of dimensions (and its value) that make the track appear in Hot-100 tracks.

To uncover relationships between different musical attributes in each genre.

Locating what the relationship between different musical variables in all 6 decades is?

Note: These questions give a rough idea of which feature will be explored in relation with other features. The dashboard doesn’t answer the above questions straight forwardly. Each dashboard focuses on a particular feature of the dataset and exhibits its association with other components.

1. Time Signature Dashboard:



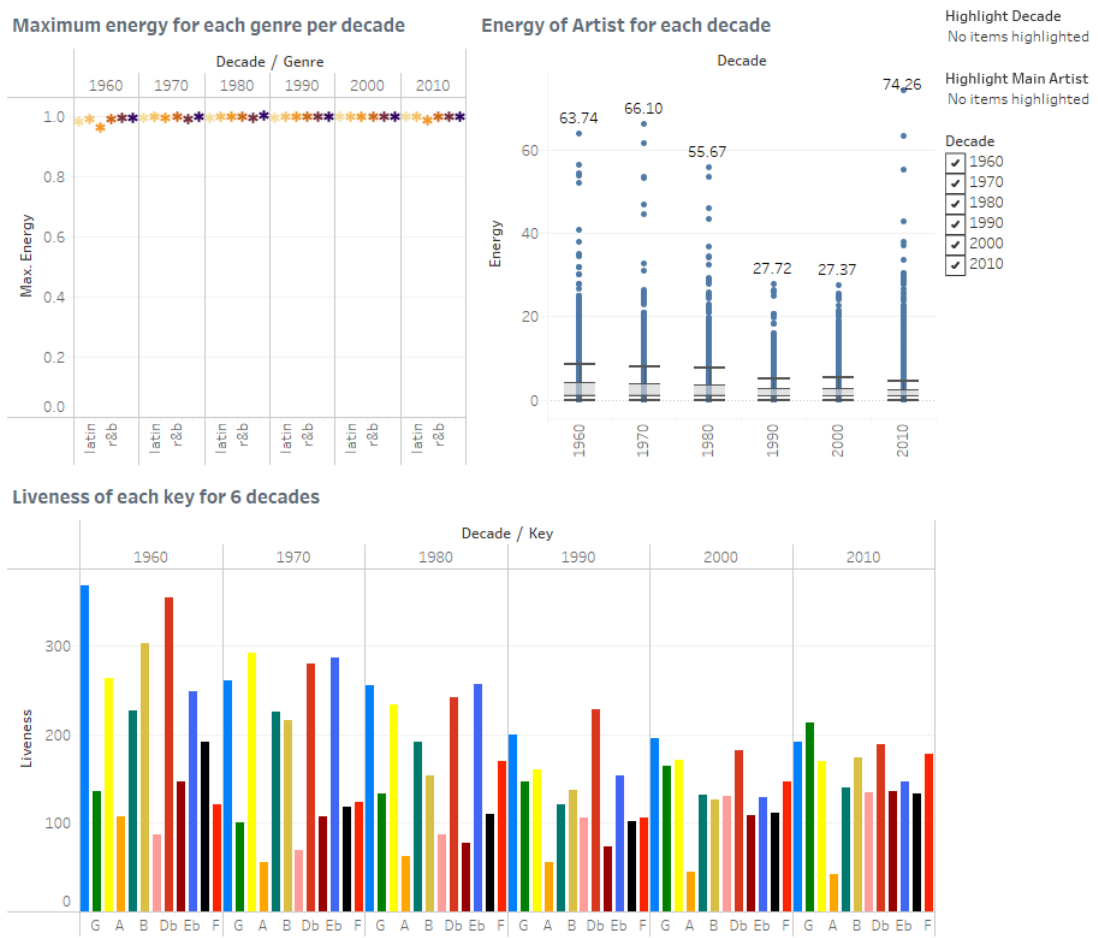
Target Audience and decision-making process: The target audience, in this case, is music producers and music record companies. It will aid a particular record company in uncovering what range of value for different musical features an artist outperforms or prefers singing the song. This way, the record company can make more money. Another beneficiary, the music producer, can make more competitive songs for the market. It will also help guide a better sound along the way to other music experts (**maestro**) and artists.

Purpose: The aim is to discover the average number of beats the leading artist has per bar in each genre's tracks. And the decade and mode the tracks belong to, for different artists. (There are 10569 unique artists in the records. Few artists are randomly filtered out and utilized for better dashboard representation.)

Information it presents: Time Signature Dashboard displays a graph with Time Signature (number of beats in each bar) as the main feature. It presents the average number of beats a singer uses in its track in various genres if they sing or are an expert in more than one genre. In addition, it offers the average TS value for each genre for a specific tempo range. Furthermore, it shows the change in the minimum TS value for numerous artists for different modes of their songs over the years.

Actionable Insights: The dashboard can be filtered by artist along with genre and tempo along with genre. A Latin track with larghetto tempo has a minimum average time signature of 4 for all 6 decades in case of tracks with minor mode. A high average time signature anomaly was seen for tracks sang within the largo tempo range and Latin genre. Artist like Selena Gomez, Taylor Swift, and Ed Sheeran have the ability to sing songs with different tempo. And artist like Elton John, Ariana Grande can sing songs within distinct genre.

2. Energy Dashboard:

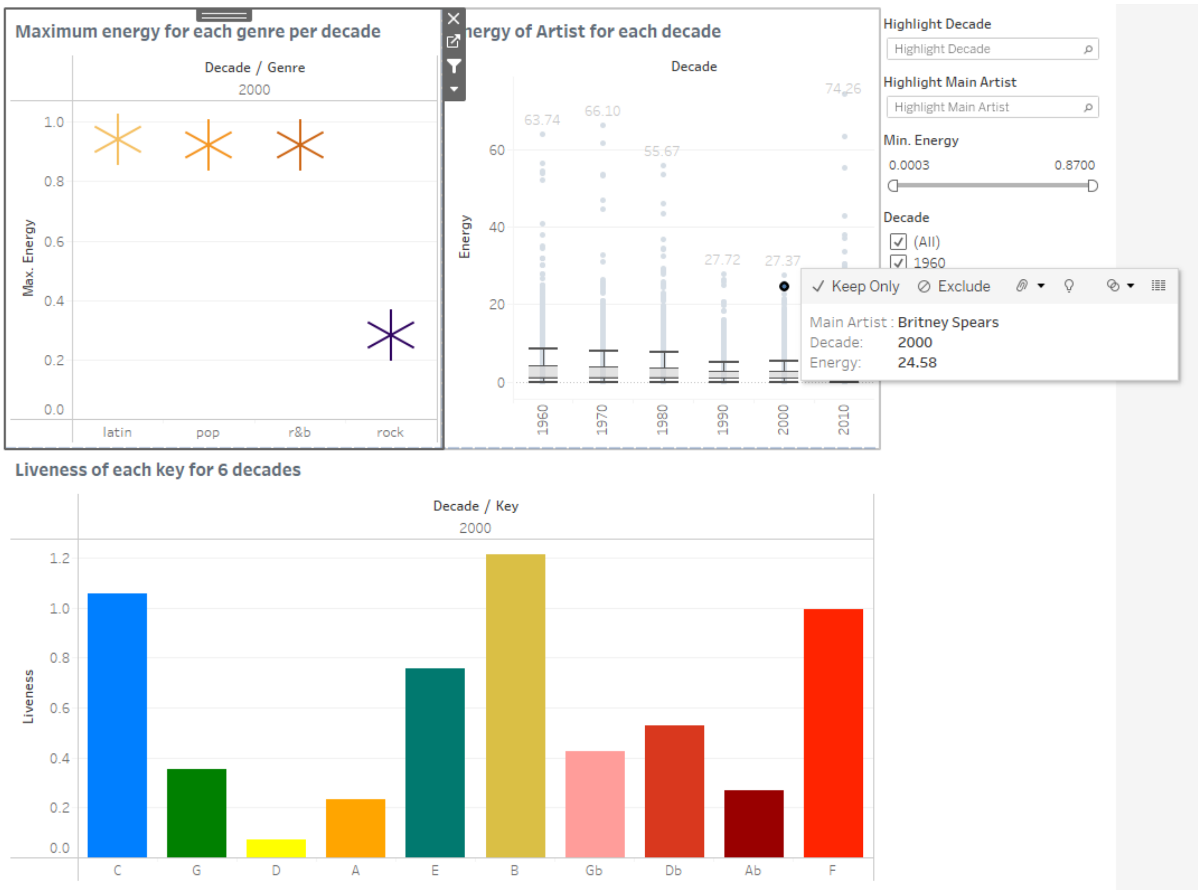


Target Audience and decision-making process: The prime beneficiary of this layout will be the **artists themselves, music composers, and listeners**. First, the performer/singer will recognize the key and genre they are more comfortable with while creating the melody. What is their liveness when they sing a song with a particular key? Finally, they can ascertain their energy range. And work on improving it if required. Listeners can pick songs and artists depending upon the energy range they favour. They can also consider the genre. The layout will work as a blueprint for the music composers can work on developing music after understanding the energy range of the artist. They can also unveil the energy trend in different decades before releasing their work. Since energy is **the sense of forwarding motion in music, whatever keeps the listener engaged and listening.** Composers can **write and arrange music for various media, including film, tv, stage productions, video games, and advertisements, as per the energy requirement in that scenario.**

Purpose: The intent here is to find the liveness of the melody for all the musical keys. The other aspiration is to find the relationship of energy feature with other remaining features of the songs.

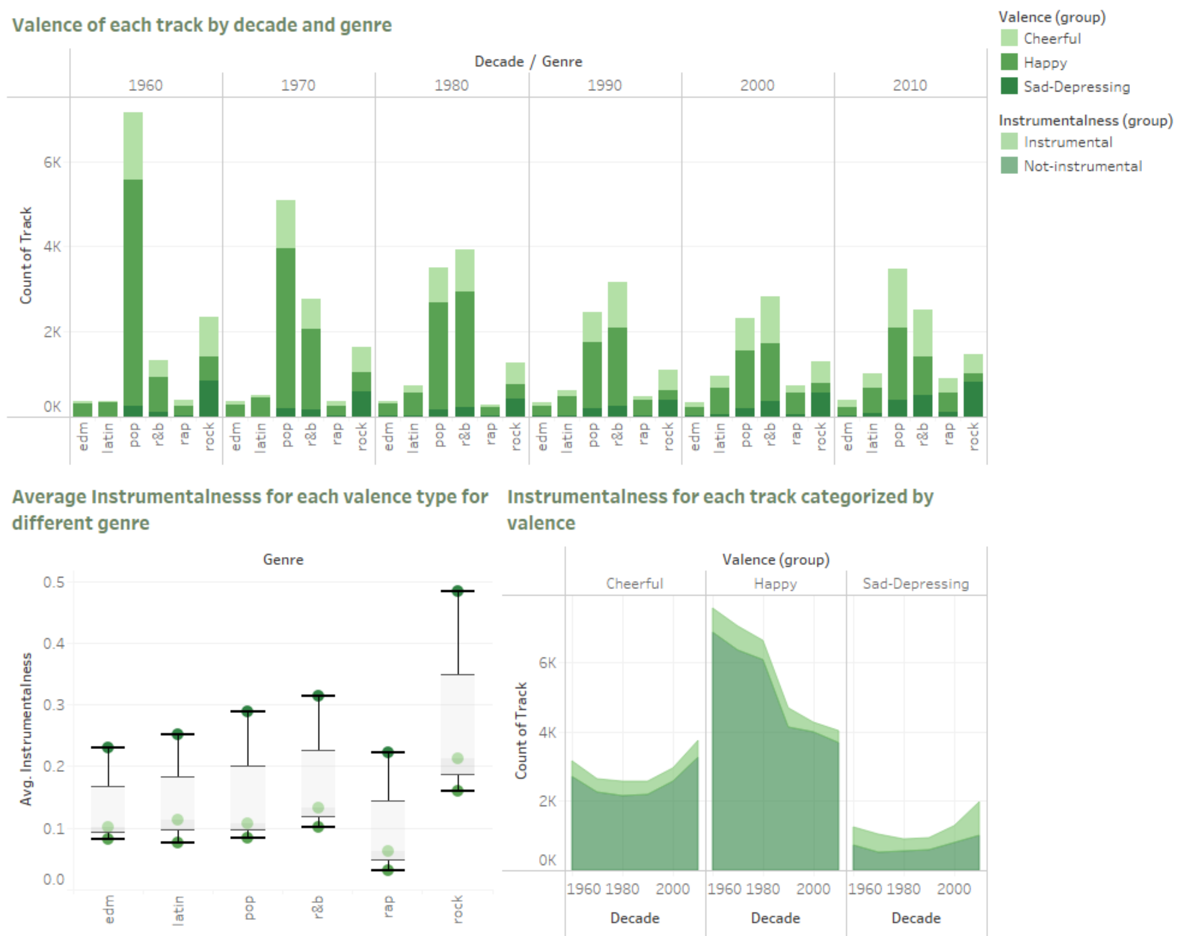
Information it presents: The dashboard exhibits the maximum energy for the different genres for each decade. While the track's energy is found based on the two attributes. At the same time, the summation of energy of all the ways for all the artists can be noted, which is segregated by decade. Both leading and featuring artists are exerted to unveil the total energy here. Suppose only the leading artist attribute is considered. In that case, the energy value of the main artist who also appeared as a featuring artist will get factored out from the closing value. And affect the overall result. Finally, the liveness of tracks by 12 musical keys for each decade can also be seen. Maximum energy is used instead of minimum energy in the first chart. When the graph was formed using the min value, the range was assigned to infer the musical piece with minimum energy. It was noticed that many artists have a high energy range while recording a track. And so, the maximum function was added to set out better results.

Actionable Insights: When the dashboard is filtered by the 'Energy of Artist' graph that is any jitter (each jitter represents a distinct artist from that era), then the maximum energy of the artist for all the genres they have sung the song in is displayed. The liveness of all the keys for different decades for all the tracks sang by a particular artist is exhibited. It can be demonstrated by the example below:



From the first graphical demonstration, only the POP genre had maximum energy equal to 0.96200 in 1960. All other genres had uniformity in the maximum value throughout all six decades. The values were above 0.98.

3. Valence Dashboard:



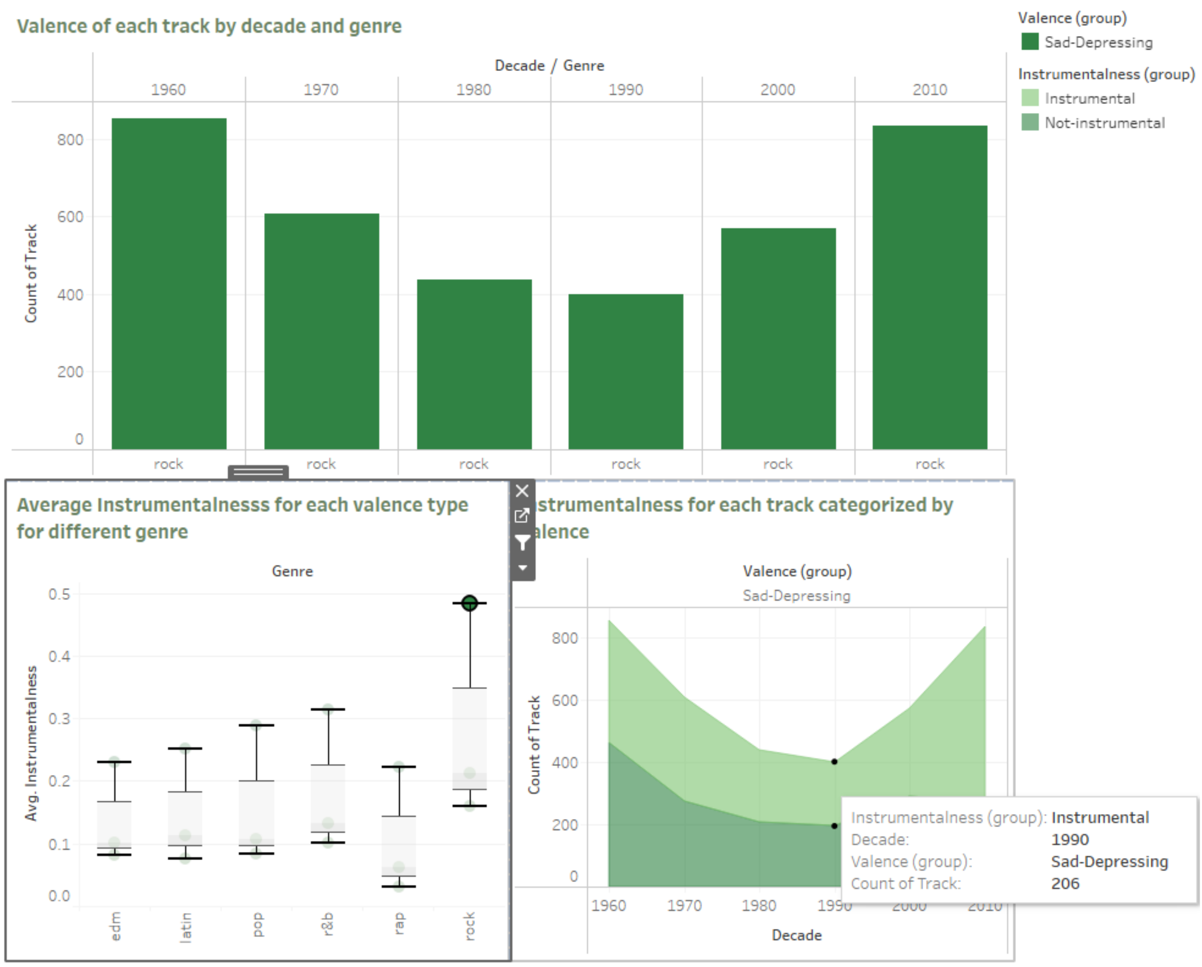
Note: The valence attribute had myriad values between the range of 0.0 and 1.0. For better graphical presentation and analysis. There was categorizing a numerical value to cheerful, happy, and sad depressing.

Target Audience and decision-making process: The target customers to utilize the dashboard are **listeners, music composers, audio engineers, and songwriters.** The layout will help the listeners select a song based on their mood. They don't need to apprehend what instrumental value is; they can opt for the theme just established founded by decade and genre. The information portrayed on the dashboard will suggest songwriters, audio engineers, and music composers select the instrumental value while making the song keeping the genre and the valence in mind. The engineers, composers, and writers must work collectively to settle the song's mood. Otherwise, unexpected, funny music pieces will be created. For instance, a rap song with high instrumental value and not much rap content cannot be classified as a rap song as it will not meet the standard.

Purpose: The layout intends to light the number of cheerful, happy, and sad depressing songs each decade. At the same time, also track down the melody based on genre. And determine the valence by genre. Eventually, the purpose is to find the average instrumental value for each valence type.

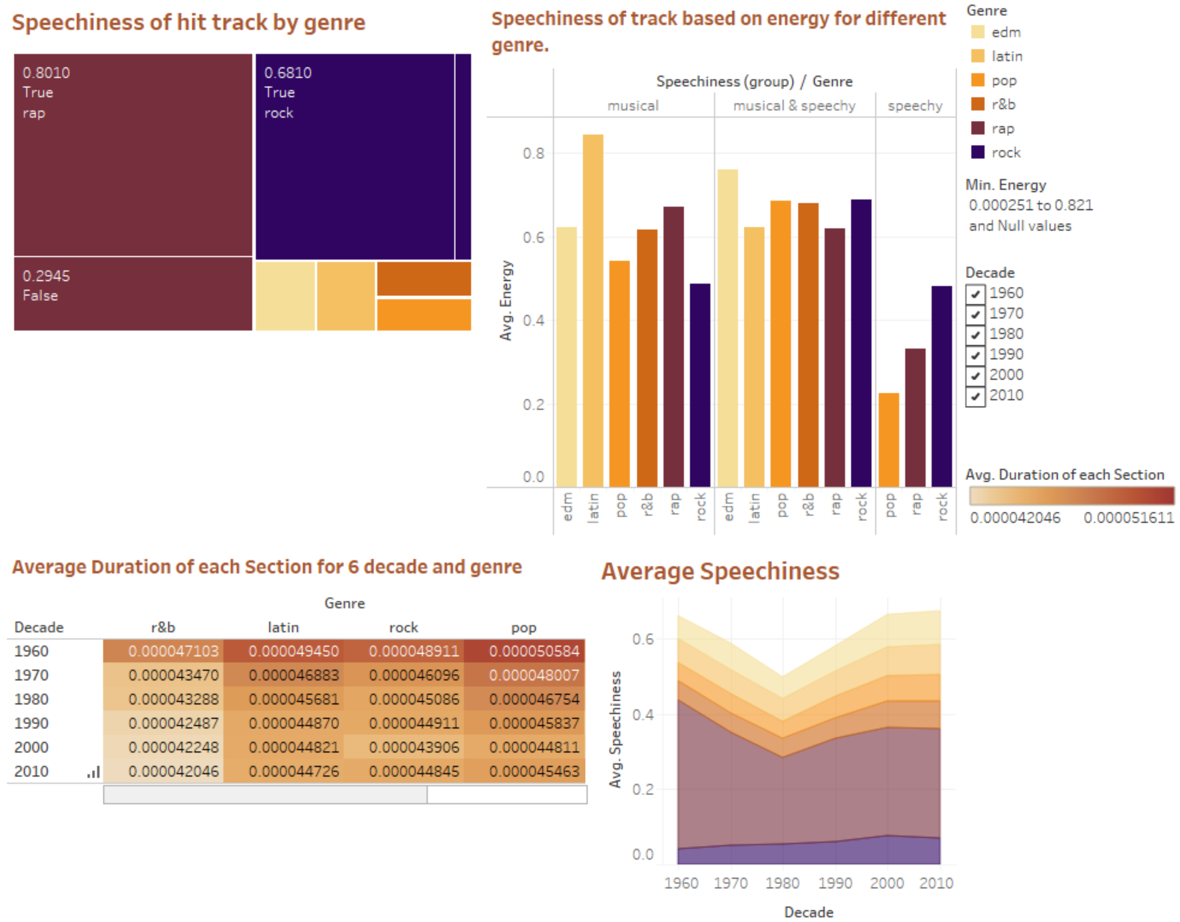
Information it presents and Actionable Insights: It was realized that the maximum number of tracks were sung in 1960 in the POP genre. And most of the songs were happy songs in all six decades. It is generally believed that sad and depressing songs are more instrumental than other types of music. The above exploratory analysis confirmed the belief that the average instrumental value of the sad-depressing piece is more than happy or cheerful songs. It is further followed by cheerful and happy music, respectively. Dazing information revealed from the dashboard was that the average instrumental value for the sad and depressing track was in the rock genre. But there was an expectation that the maximum average weight for sad songs would be revealed in Latin or pop music. The instrumental-valence graph corroborated that the rap songs are more speechy and less instrumental. The happy rap song has the most negligible instrumental value of 0.0320. The cheerful music in the rock genre has almost the same value as sad music in the rap category. The value is close to 0.2.

The was a downfall in the number of pop track compositions till late 2000. But a spike was seen from the beginning and throughout the 2010 decade. On the other hand, the number of EDM tracks created was constantly low over six decades.



The third graph at the bottom right exhibits if the song is instrumental or non-instrumental. There were 206 unique tracks in 1990 that were sad and depressing in the rock genre. Pieces with an instrumental value of 0.5 and above are considered instrumental. All other songs are deemed as non-instrumental.

4. Speechiness Dashboard:



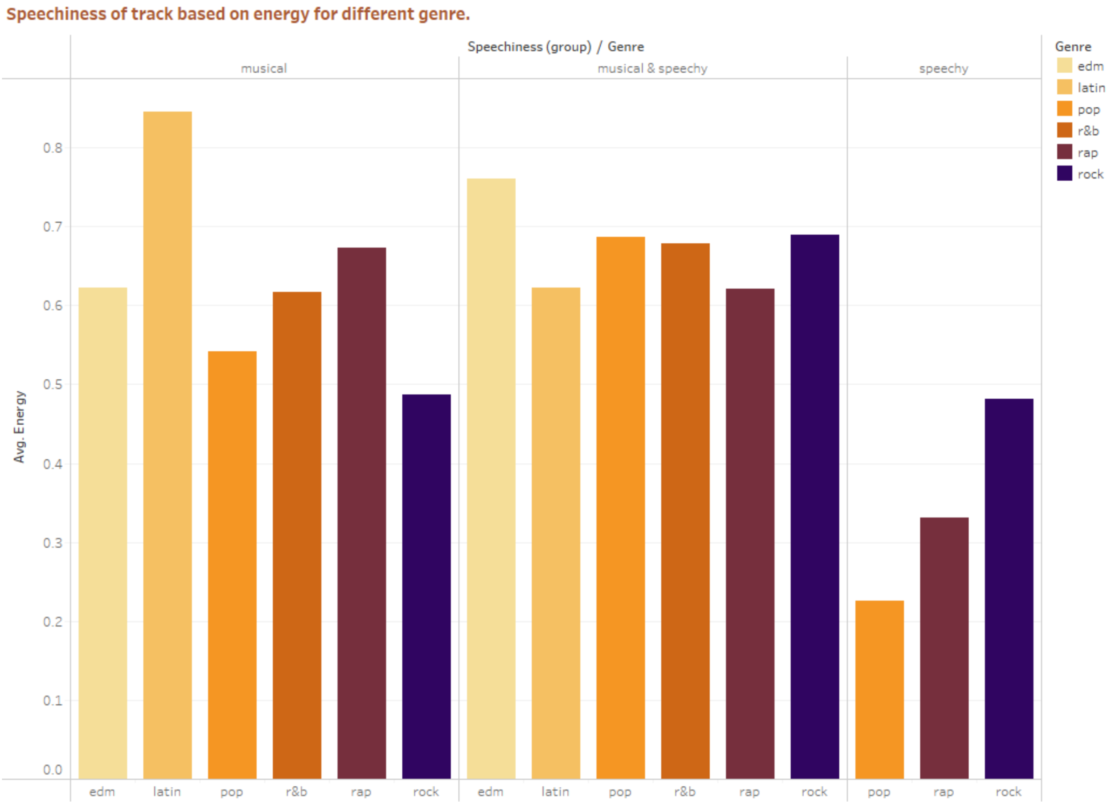
Decade	r&b	latin	rock	pop
1960	0.000047103	0.000049450	0.000048911	0.000050584
1970	0.000043470	0.000046883	0.000046096	0.000048007
1980	0.000043288	0.000045681	0.000045086	0.000046754
1990	0.000042487	0.000044870	0.000044911	0.000045837
2000	0.000042248	0.000044821	0.000043906	0.000044811
2010	0.000042046	0.000044726	0.000044845	0.000045463

Complete representation of half displayed graph:

Average Duration of each Section for 6 decade and genre

Decade	r&b	latin	rock	pop	rap	edm
1960	0.000047103	0.000049450	0.000048911	0.000050584	0.000047078	0.000051611
1970	0.000043470	0.000046883	0.000046096	0.000048007	0.000045365	0.000049046
1980	0.000043288	0.000045681	0.000045086	0.000046754	0.000044917	0.000046214
1990	0.000042487	0.000044870	0.000044911	0.000045837	0.000043770	0.000045698
2000	0.000042248	0.000044821	0.000043906	0.000044811	0.000044259	0.000045058
2010	0.000042046	0.000044726	0.000044845	0.000045463	0.000045893	0.000047313

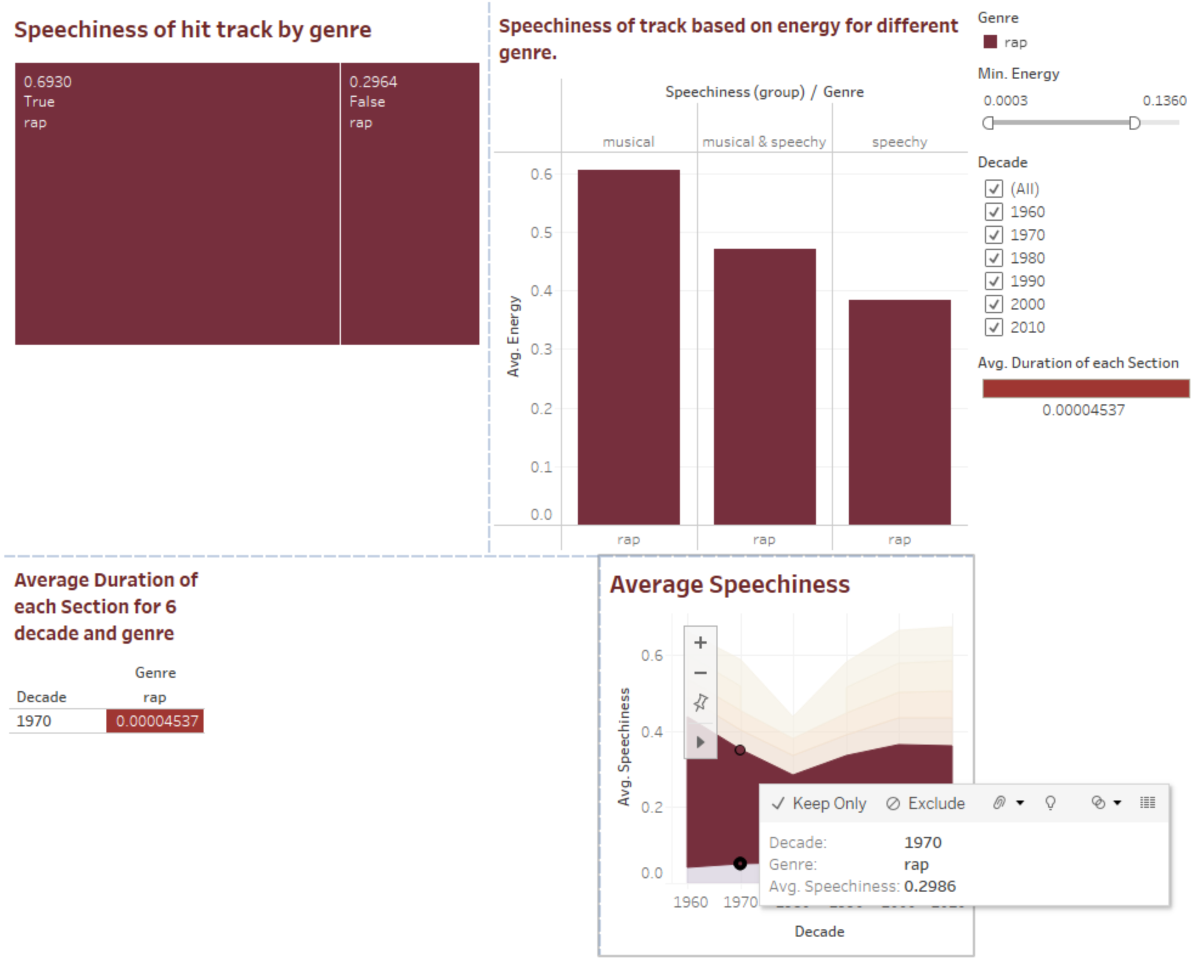
Avg. Duration of each Sect..
0.0000420460.000051611



Target Audience and decision-making process: The primary and sole recipient of this layout is the **lyricist.** The dashboard will guide the pristine lyricist on how big a song they should write depending on the genre theme and singer for whom they are penning the music. The lyricist can also comprehend from the analysis that if the song is too lyrical, then the song's overall energy can decrease. For instance, pop music has the least energy when the track has a high speech value. The **only exception here is Talyor Swift's 'All too Well' 10 minutes version with 52% energy despite the too wordy song. Some can retain it in the musical & loquacious category from the layout above.**

Purpose: The pursuit of this dashboard is to find the relationship between the speechiness of tracks with other components of the music. The diagram in the right bottom corner shows the average speechy value for all six decades. Another purpose is to find the average duration of each section simultaneously. Eventually, the aim is to uncover the hit or flop track depending on the spoken content of the song and further classified by the genre of the music.

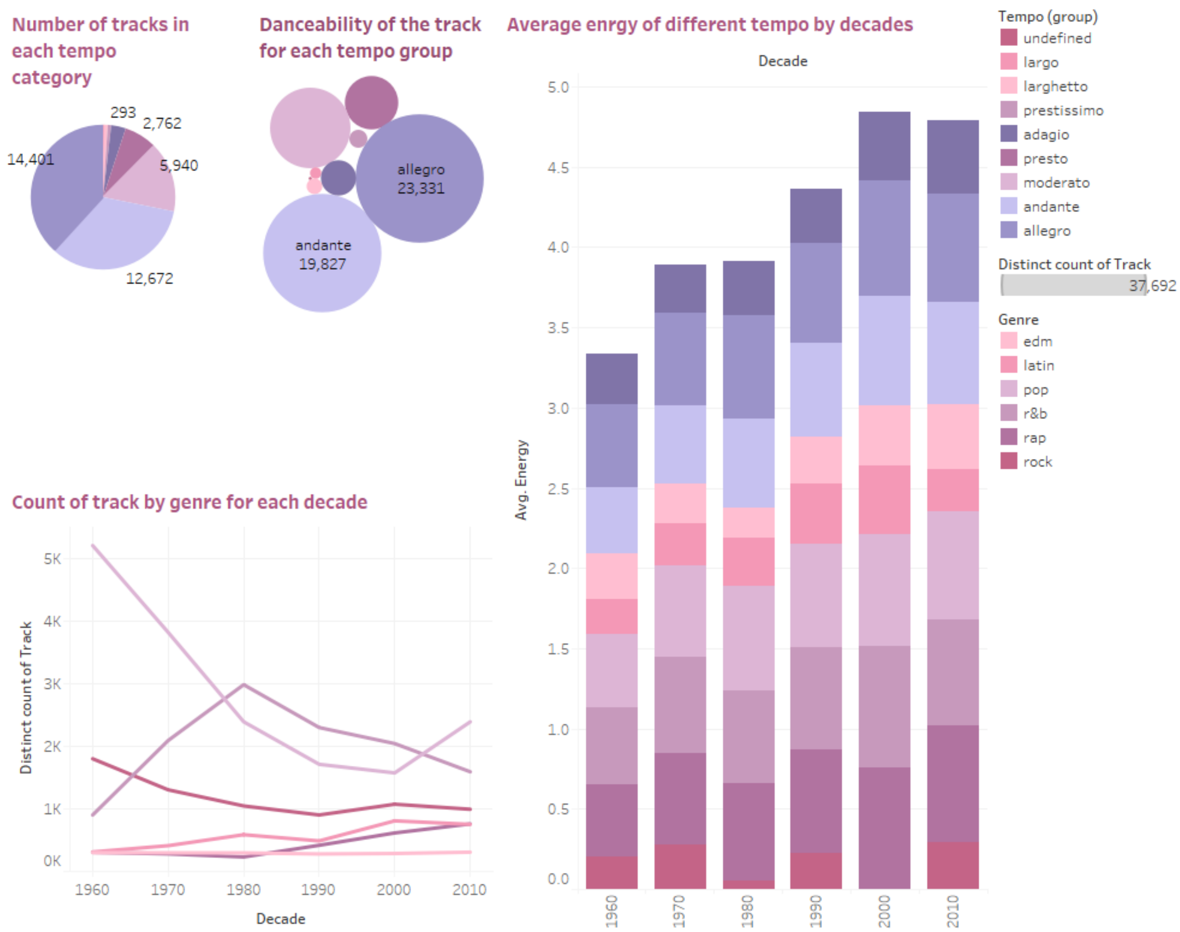
Information it presents and Actionable Insights: It was uncovered that the high average duration of each section in a song was seen between 1960-1970. The most negligible value was recorded in 2010 for all the tracks under the r&b theme. The highest value was recorded in the EDM genre in 1960. The insight from the graph helped get the closure that artists made long songs in the mid-90s, and it decreased with the decade and specifically by each year. From the demonstration, it is glimpsed that between 1970-1990 there was no track with an average speechy value that satisfies the range value available when the minimum energy value was between 0.000251 and 0.1360. The average spoken word value hit the low value of 0.0577 in 1980. A rising curve was seen after 2000, and the average value remained almost the same throughout that decade and the following decade. The size of boxes in the tree maps depends on the number of hit songs, the speech, and the genre theme. The colour is added to the track based on genre and not hit or value indicator. There were no hit songs discovered in EDM, Latin, and R&B genres based on speeches of the musical track. The rock theme has the least flop in all six decades. The average speechy value for both hit and flop pieces of music under rock theme is 0.6810. It was disclosed that the average energy of Latin music was highest if the track was musical and not speechy. EDM, LATIN, and POP theme songs are not lengthy. They include both musical and speech components in it—the pop music energy decrease when the piece is too protracted.



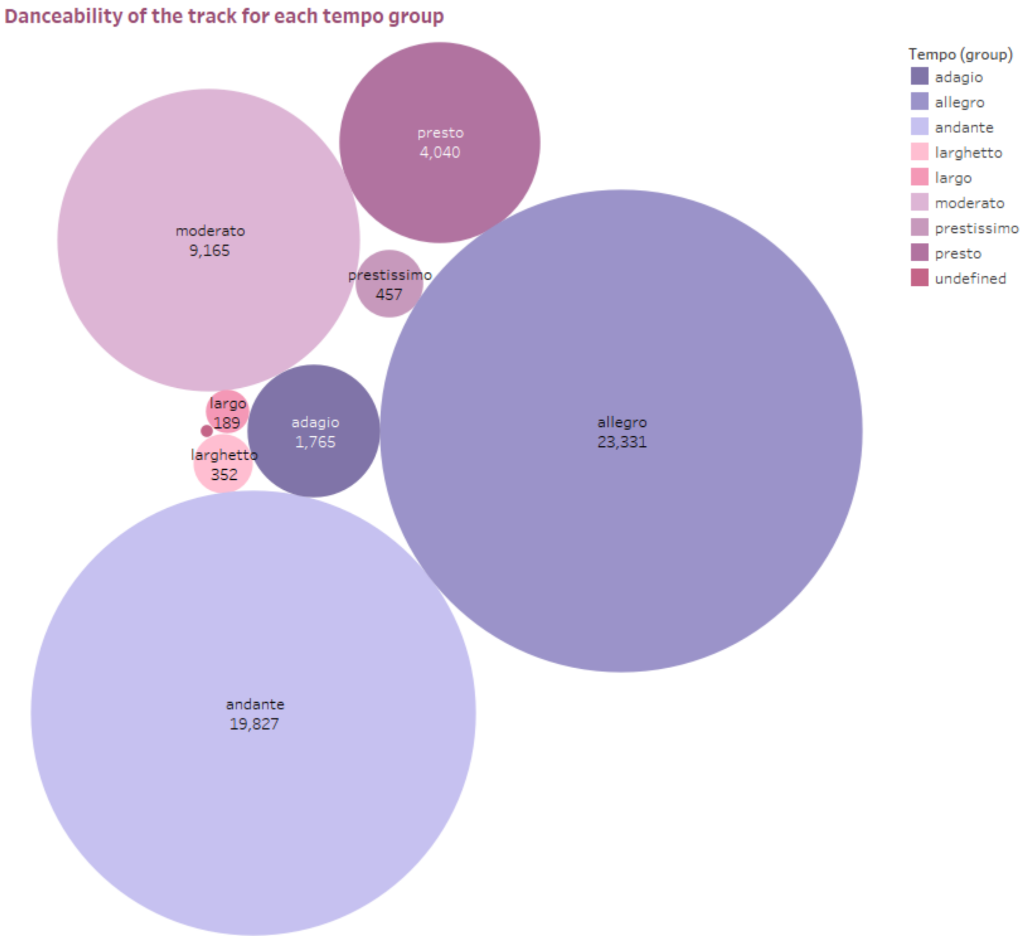
On filtering, it was disclosed that the average energy of Latin music was highest if the track was musical and not speechy. EDM, LATIN, and POP theme songs are not lengthy. They include both musical and speech components in it—the pop music energy decrease when the piece is too protracted.

It is detected that the song's energy decreases for a rap song when the track moves from musical to speechy. As a result, the number of hit songs is more than the number of flop songs when it is a warp on woof on the length of the lyrics. The verboseness of flop rap songs is close to 0.2964. This aids in concluding that song if the speechiness of the rap song is less, then it will be a flop piece.

5. Tempo Dashboard:



Complete representation of half displayed graph:

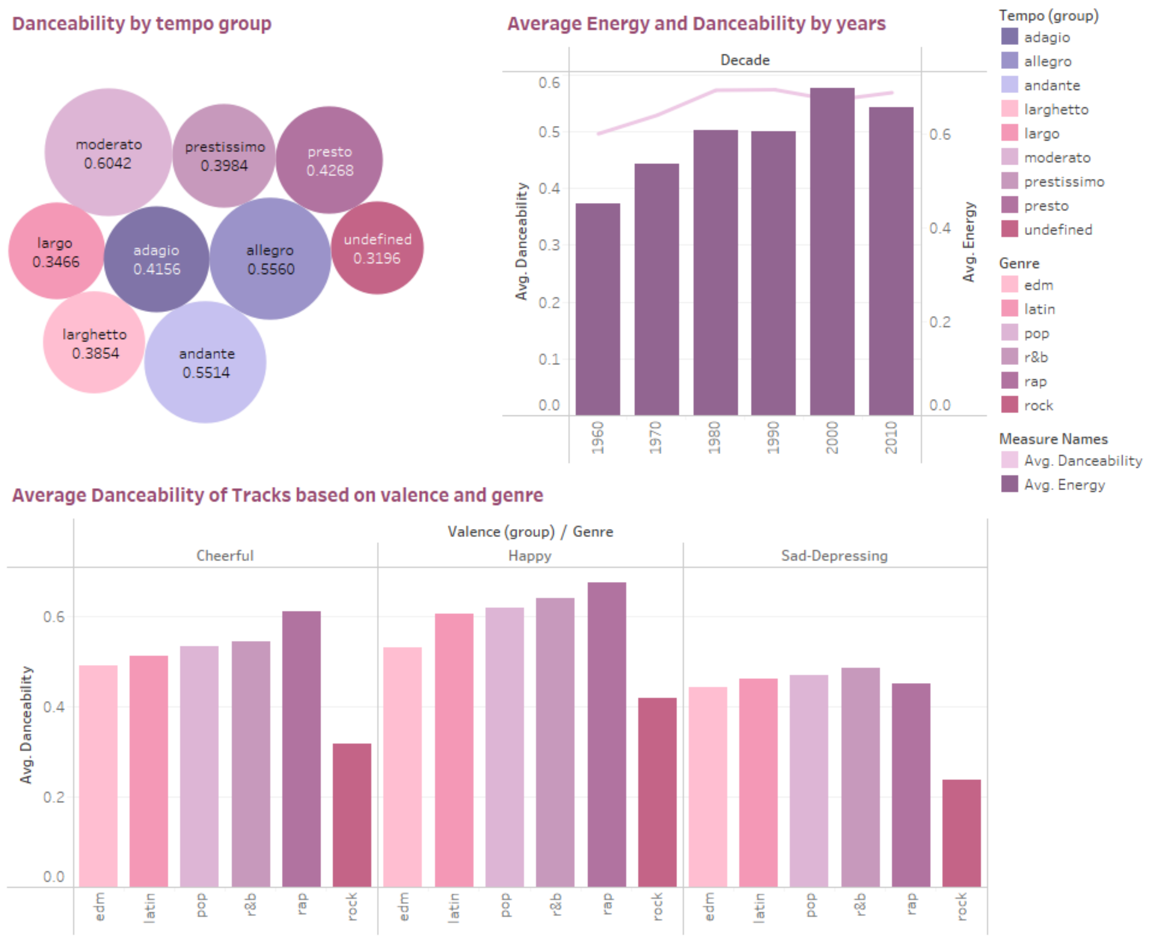


Target Audience and decision-making process: The primary beneficiary of this layout is **musicians,** as they decide the tempo of the music based on the genre theme, danceability, and artist. The dashboard will help the musician make changes in the tempo of the music without affecting the overall quality of the song when working with different artists—or making the cover of an album or the music record. The multigraph layout will also aid the musicians in selecting tempo carefully when they are creating a music sound piece for symphony orchestras, professional bands, and music for films or television.

Purpose: The motivation of the layout is to find a number of tracks in each tempo classification. Simultaneously, the other goal is to find the average energy for each tempo type for 12 lustra. Ultimately, the number of songs is also computed by genre for a period of 6 decades.

Information it presents and Actionable Insights: There were 14,401 distinct tracks for 12 lustra that have allegro-type tempo. The petty count is for larghetto. The allegro only has a high track count, but it is also the most danceable track. The least danceable songs are the ones that have a tempo-type largo. The tally is 189. The cast up of EDM genre theme songs remains approximately the same for all 6-decade. The total was a little above or below close to 300. The Measliest number of tracks was recorded in 1970. The totalization was 6684 records. The highest was in 1960. The value of Latin and rap themes increased throughout the years. The pop genre song had a steep slope from 1960 up till 2000. Alternatively, there was a peak bargain in 1980 for the r&b genre. The average energy of adagio decreased in the early and late 1970s. After that, there was a persistent increase in the tempo value. As a result, the moderato and prestissimo have approximately the same average energy in 1990.

6. Danceability Dashboard:



Target Audience and decision-making process: The target audience for the danceability dashboard is music **directors, choreographers, and everyday listeners.** This will aid the typical frequent listeners select the music based on danceability. They don't need to be a maestro to choose a soundtrack to dance. But the layout will lead to a better choice. Furthermore, it will permit the dance choreographers to pick a music piece from different decades based on the situation**. Finally, the music director is the most essential beneficiary, also called the conductor. The conductor can intensify or make changes to their record based on the danceability requirement. For instance, the condition is a danceable song for a sad-depressing situation or a happy scenario.**

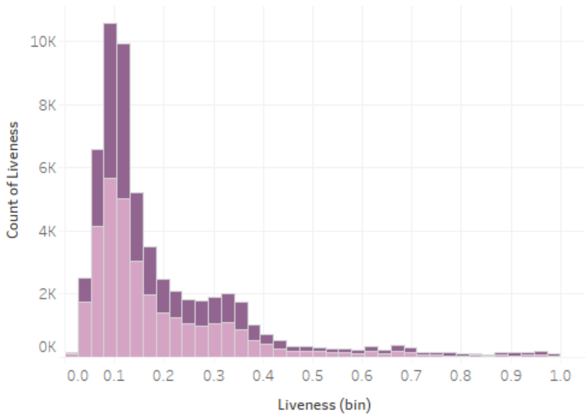
Purpose: The drive of the design was to find the danceability of the track based on other components of the song. The average danceability of the music is witnessed every decade. The affinity between energy and danceability is also found. There is a computation of tracks' average danceability by valence (mood) and genre to better insight if the music is danceable or not.

Information it presents and Actionable Insights: The first insight obtained is that moderato is the most danceable tempo with a value of 0.6042. And the minor danceable tracks are under the undefined category. The average danceability increased gradually from 1960 up till 2000. After that, a fall was seen again in 2010 and years after it. The average energy of the songs boosted as well with the passing years. It indicates there might be a positive affinity between these two attributes.

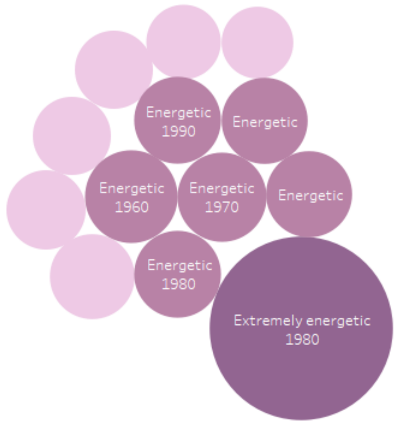
Average danceability by genre and valence is also scanned in the layout. It reveals that the rock songs are the least danceable of all other genres songs, whatever the mood value. The most danceable genre is rap. No strong pattern could be discovered in the remaining three genes. When values are filtered by years, the average danceability for each tempo type changes. A transformation is noticed in the genres for each valence group. The interactive dashboard can also be filtered by genre and valence mood when screened using any one of the ways. Then, there is a modification in the danceability value for each tempo type.

7. Liveness Dashboard:

Hit and Flop Track Count for a particular liveness range



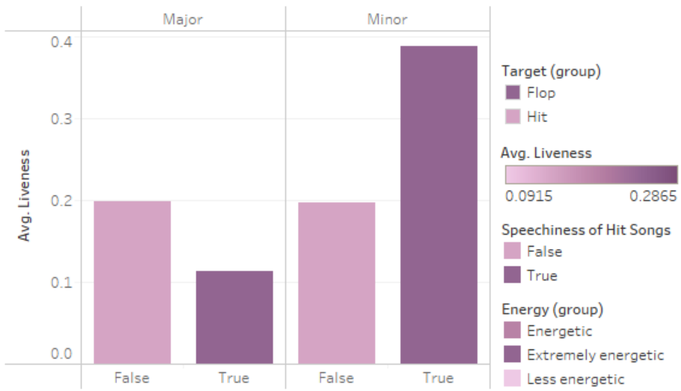
Average Liveness grouped by energy for different decades



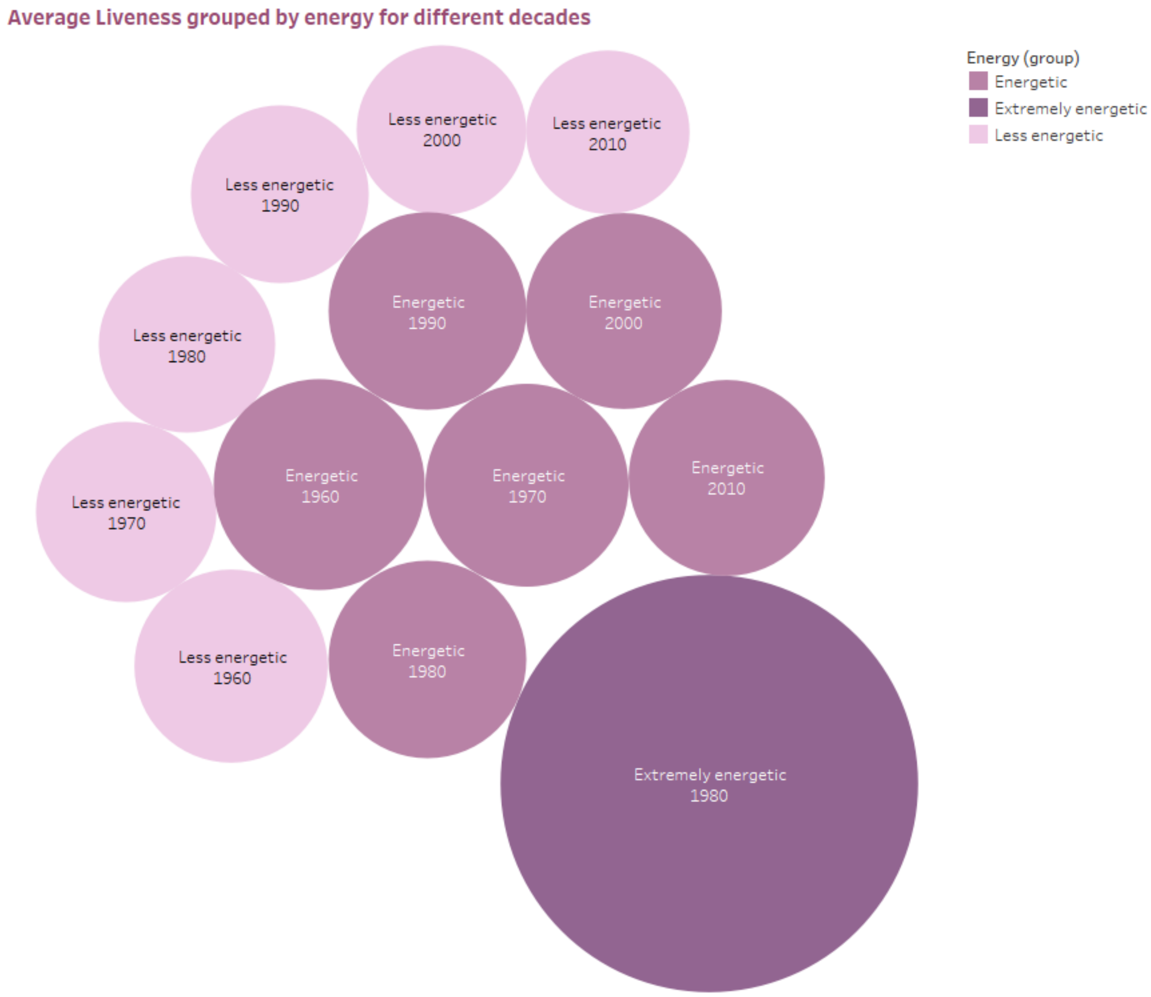
Avrage Liveness by Tempo and Valence

Tempo (gro..	Valence (group) Cheerful	Happy	Sad-Dep..
adagio	0.1789	0.2372	0.1465
allegro	0.1956	0.2046	0.1872
andante	0.2024	0.1973	0.1831
larghetto	0.1762	0.1907	0.1389
largo	0.2608	0.2203	0.1619
moderato	0.2066	0.1954	0.1844
prestissimo	0.2154	0.2137	0.1508
presto	0.2138	0.2002	0.1969
undefined	0.1690	0.0915	0.2865

Average Liveness of the track by mode and speechiness of hit songs



Complete representation of half displayed graph:



Target Audience and decision-making process: The beneficiary of the layout is performers (singers) and music directors. It will let the performers know that they need to work on melodies' energy and not just their energy to increase the overall liveness of the song. Likewise, the music director can improve the overall result of the soundtrack once they detect the affinity of one feature with others. It will ultimately lead to a high probability of top musical records and sweeten the overall performance of their music record company. “Liveness” refers directly to reverberation time. A live room has a long reverberation time and a dead room a short reverberation time.

Purpose: The motivation of the above dashboard is to count hit and flop tracks for a given liveness range. The mean liveness of the music record is split by different decades and energy to explore the affinity of energy with liveness. Its additional purpose is to examine the association of average liveness split by verboseness and tempo of the music record.

Information it presents and Actionable Insights: From the dashboard, it was perceived that the average liveness of the flop track based on speechiness is almost identical for both major and minor modes. The average liveness is 0.3883 when hit songs with high verbose values are true. This is valid for minor mode songs. It was discerned that the most energetic music record was in 1980. The liveness histogram exhibits the count of hit and flop songs for each bin value. Higher bin values have a low count of liveness soundtrack (2-digit number) compared to initial bin values. There are 31 flop songs and 115 hit songs with a liveness bin value of 0. In the final part of the dashboard description, the liveness count has checked against both the valence and tempo of the song. The high liveness value uncovered for sad-depressing music with undefined tempo has surprisingly high values. It does not follow the pattern as the sad-depressing songs should have consistently low liveness value irrespective of the song's bpm. BPM stands for Beat per minute, and it is used to measure the tempo. On filtration by the tempo-valence graph of liveness attribute, it was discovered that there are only flop songs when explored based on verboseness feature of music. The number of hit and flop songs can be checked on the histogram graph of the liveness count of records.

8. Acoustics Dashboard:

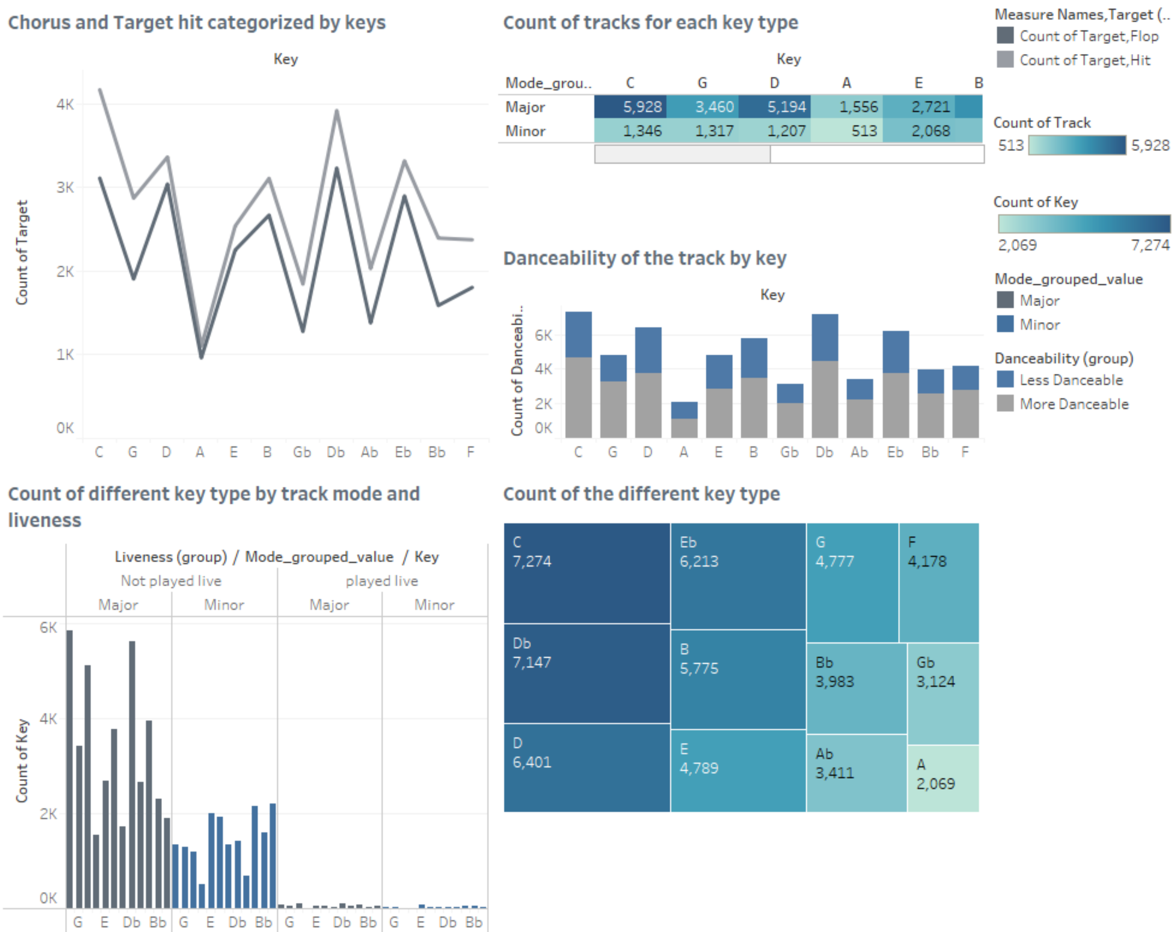
Target Audience and decision-making process: The prime target audience for this multigraph is **audio engineers** and **music producers.** These two music genii are the ones that work on sound, instruments, and electric sounds of the music. Therefore, the layout will aid the maestros in apprehending each genre's pattern and producing the track according. Before making and releasing the final piece, they can keep the time signature, liveness, and mood factor in mind.

Purpose: The dashboard aimed to analyse the acoustic aspect and study its correlation with the instrumental value, genre, liveliness, time signature, tempo, and valence. In addition, there are two comparison graphs to associate the instrumentals with verboseness and the acoustics of soundtracks.

Information it presents and Actionable Insights: The acoustic and instrumental comparison graph demonstrates that the maximum acoustic and instrumental value is close to 1 for all six decades.

However, it doesn't prove that the song with high instrumental value also has a high acoustic value as no condition was applied to the features. For example, if high acoustic and instrumental values display the tracks. A comparison graph between instrumentals and lyric variables is also presented. For example, pop, EDM, and Latin genre themes have no acoustic value in the 5th Time Signature when it is played live. Songs with 0-time signatures have significantly less or no maximum acoustic value. The size of the diamond in the top right graph on the dashboard is determined and established on the maximum acoustic value for that specific genre and the time signature of the soundtrack it belongs to in the data records. The maximum acoustic value for each tempo-valence type is almost the same. The graph represented in the lower-left corner is a bit confusing. The top acoustics is 1 in the analysis process. Each decade, the maximum speech value increases if the soundtrack is in the 4th Time Signature and rock genre and not played live.

9. Key Dashboard:



Complete representation of half displayed graph:

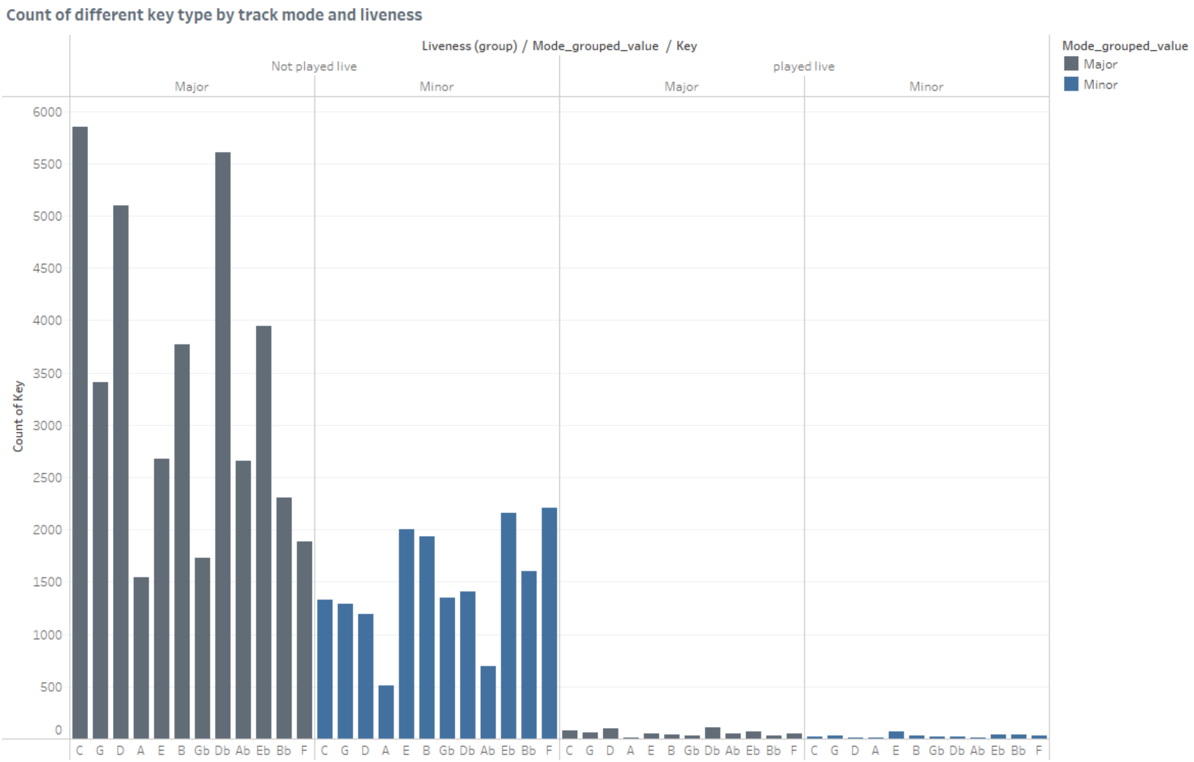
Count of tracks for each key type

Key

Mode_grou..	C	G	D	A	E	B	Gb	Db	Ab	Eb	Bb	F
Major	5,928	3,460	5,194	1,556	2,721	3,815	1,755	5,720	2,705	4,015	2,341	1,934
Minor	1,346	1,317	1,207	513	2,068	1,960	1,369	1,427	706	2,198	1,642	2,244

Count of Track

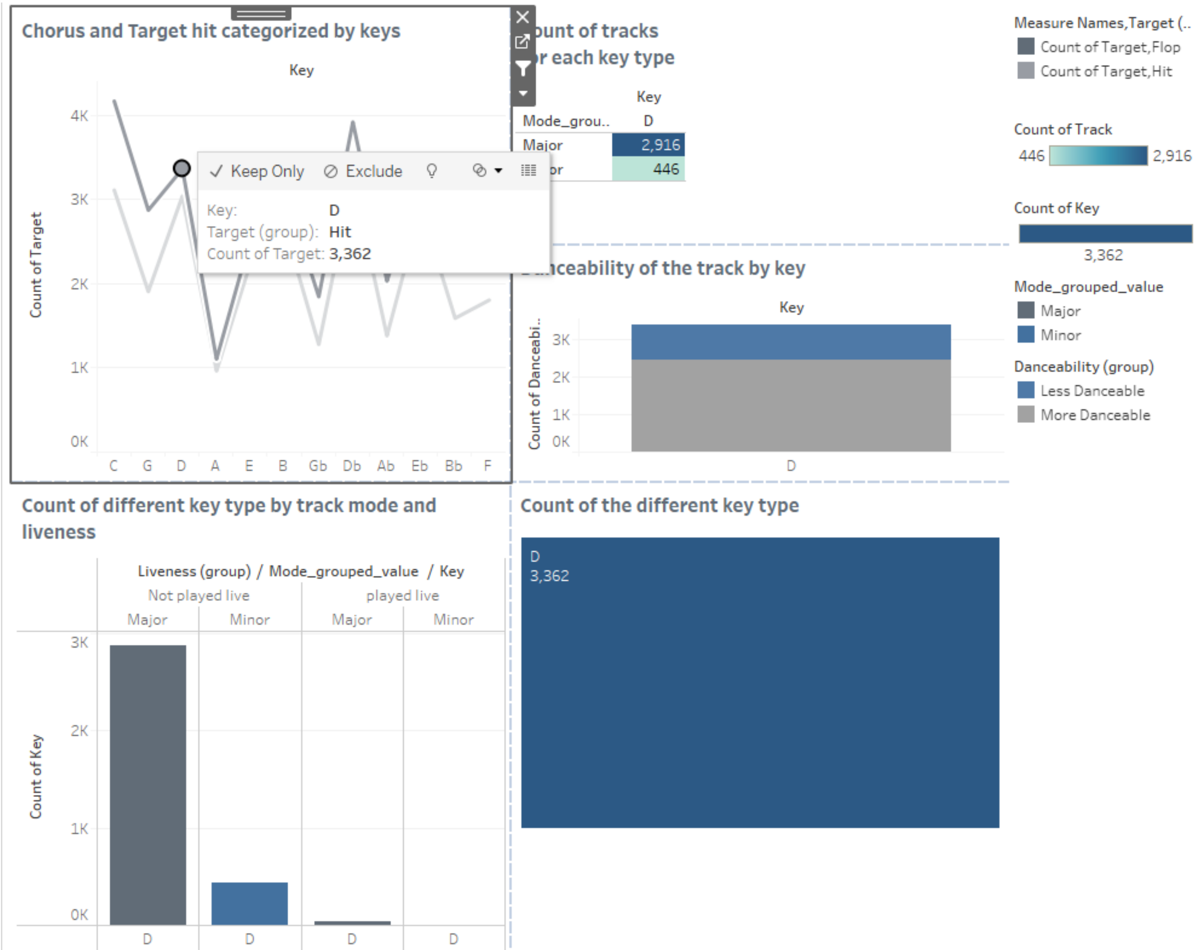
513 5,928



Target Audience and decision-making process: There are numerous target audiences for the above layout. **The list possesses singers, lyricists, music producers, music composers, and audio engineers.** All these music connoisseurs must keep the key component in mind when working on any part of the music. The dashboard will enable the experts to make the correct choice. **In addition, the key dashboard can be advantageous to an individual who is a newbie to the musical domain. It can be a good kickstart for them.**

Purpose: The motive of the polygraph dashboard is to locate the connection between keys and other musical components. The sub-objective is to find the count of music records under each key type. The number of hit and flop tracks for each musical key. The mode the key belongs to and the danceability for each key type. It is also analysed if the key was played live or not.

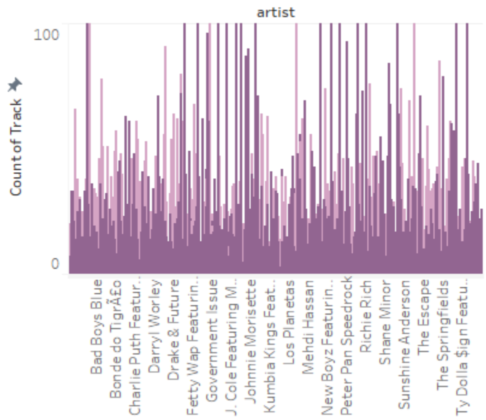
Information it presents and Actionable Insights: The dashboard reveals the total number of records for every key type. Simultaneously it represents the count by mode of the song, which is the major mode or minor mode. Furthermore, the hit and flop count is located. Finally, the danceability is checked for each track and further bifurcated into danceable and less danceable melodies. The number of records with a specific key, whether played live or not, is investigated. Key 'A' has the least danceability of all other keys. The most danceable keys are C and Db. The most common key is C, with a count of 7274. And least used key in the song is 'A', with a total of 2069. The number of data records played live with different keys is relatively less than all the data records that are not played live. The key with the most hit and flops is C, and most of the music records are sung in the major mode. The number of records yodels in minor mode and used less is A in concern with other keys.



The total number of hit (popular) songs sung using key D is 3362. Out of which, 2916 were in major mode and 446 in minor mode. No records were played live in the minor mode in this scenario. The count of the less danceable track is 937, and the more danceable key is 2425. Finally, the number of records played in major mode and not played live is 2883.

10. Hit Flop Dashboard:

Hit and Flp Song by artists



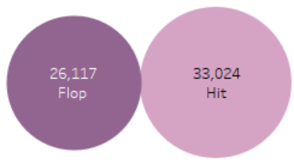
Hit and Flop track by average of chorus hit and valence

Target (gro..	Valence (group) Cheerful	Happy	Sad-Depressing
Flop	41.510	39.600	43.472
Hit	39.956	39.089	40.036

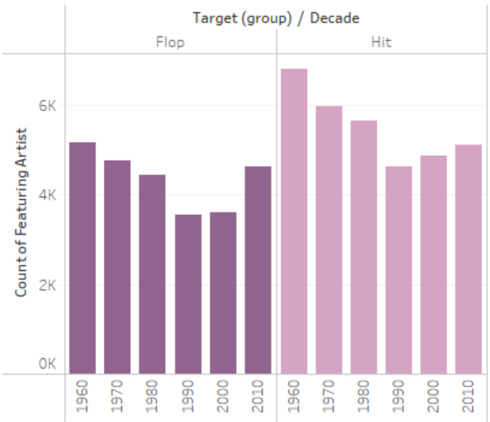
Target (group)
- Flop
- Hit

Genre
- edm
- latin
- pop
- r&b
- rap
- rock

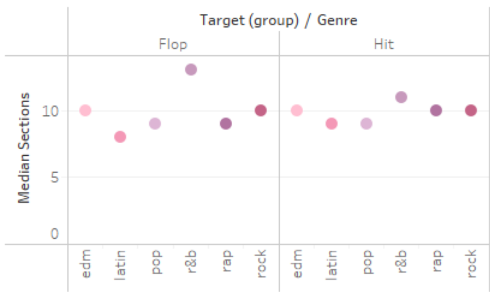
Hit and Flop track count



Hit and Flop Track of featuring Artists

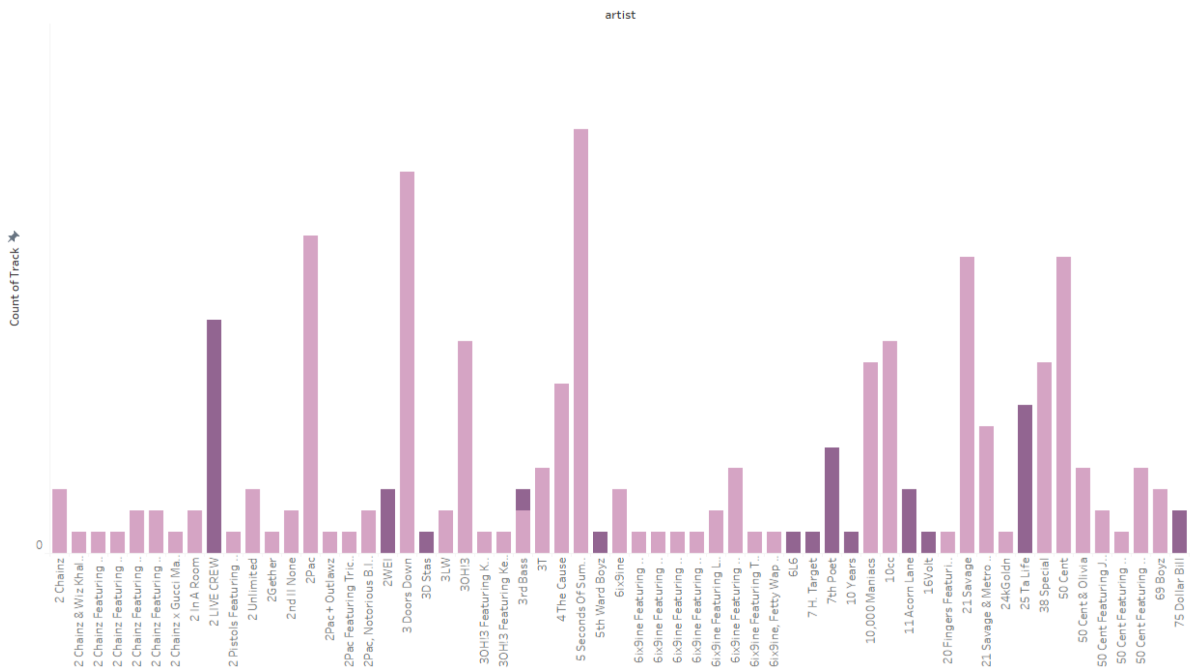


Median of Section of Hit and Flop track, further classfied by genre



Complete representation of half displayed graph:

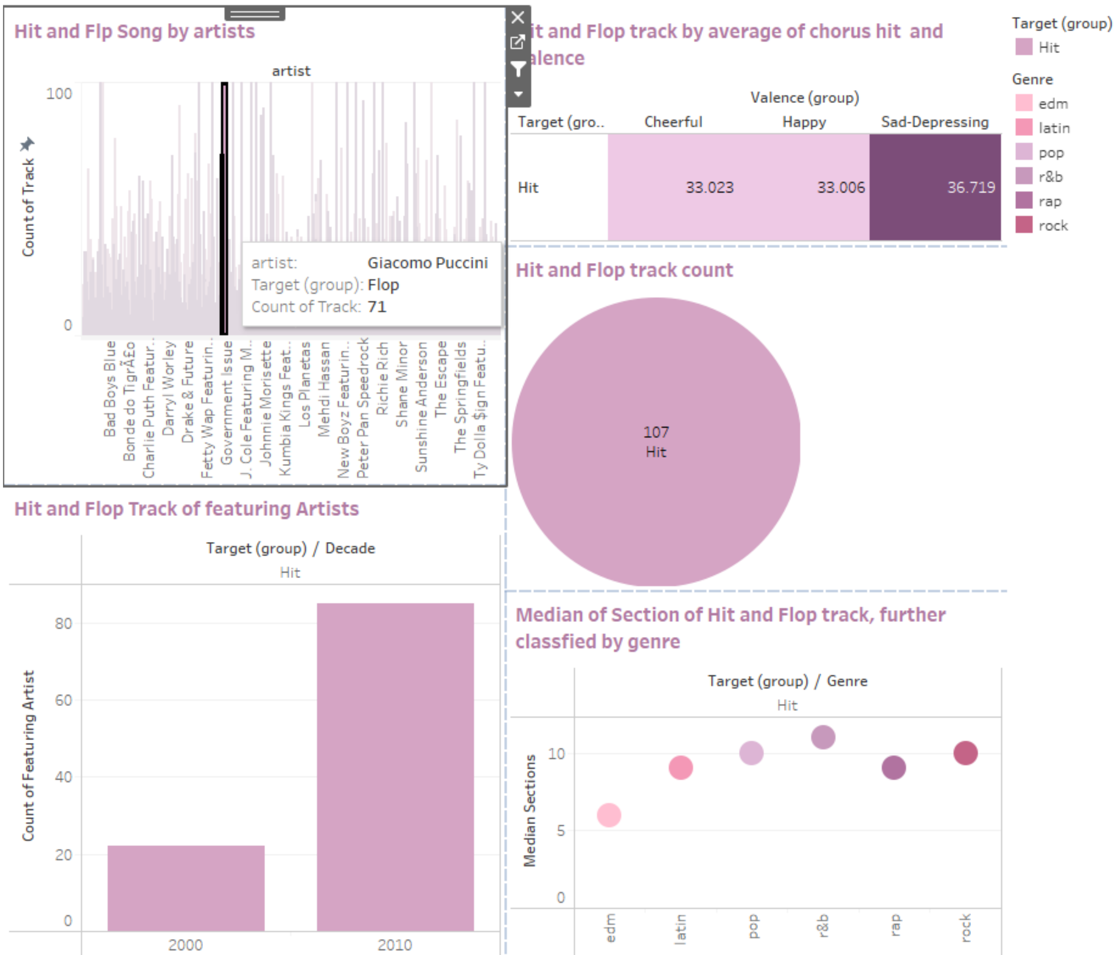
Hit and Flp Song by artists



Target Audience and decision-making process: The primary beneficiary of the above-mentioned target audience is record company and singers. Singers can understand which genre they perform best or worst and opt to be part of the album accordingly. Likewise, record companies will get an insight into which artist and genre to invest in so that they make more money.

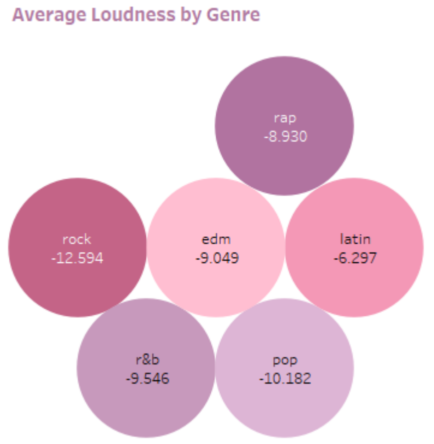
Purpose: The hit flop dashboard exhibits the hit and flop song feature, also known as the target element in the dataset. It illustrates the count of Hit and Flops soundtracks. Its analysis is further broken down to discover if the track is a hit or a flop based on genre, section, valence, and mean of chorus hit. Both leading and featuring artists are filtered to apprehend the hit and flop component of the music.

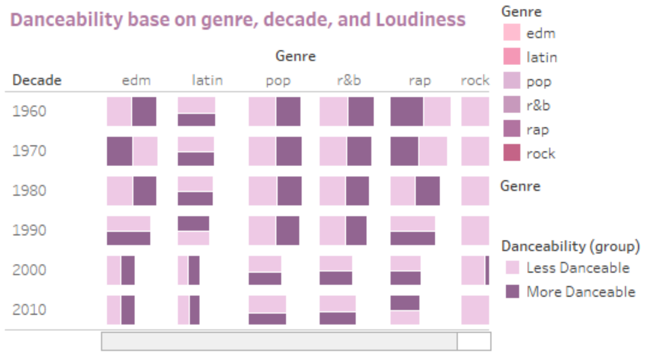
Information it presents and Actionable Insights: Artists can filter the dashboard. All the artists are included here in the dashboard. The total number of hit songs is 33204, whereas the number of flop tracks is 26117. The most flop songs are from r&b, and the least is from Latin. Similarly, the maximum hit was from r&b, and the minimum was pop. Both the most hit and flop featuring artist was from 1960. Minimal flop songs were recorded in 1990 and the following nine years. The number of flop songs, which are sad and depressing, has an average chorus hit of 43.472. It is the maximum value.

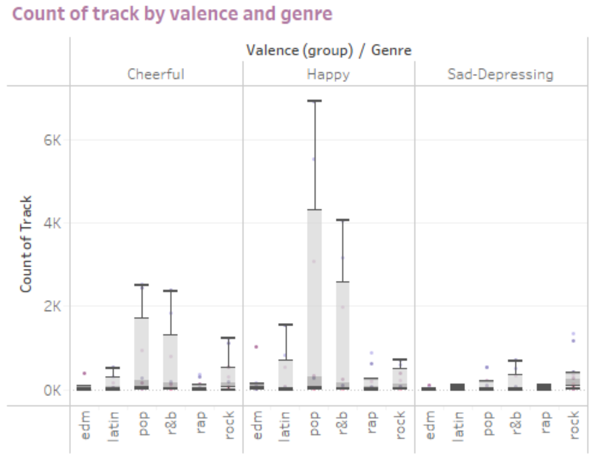


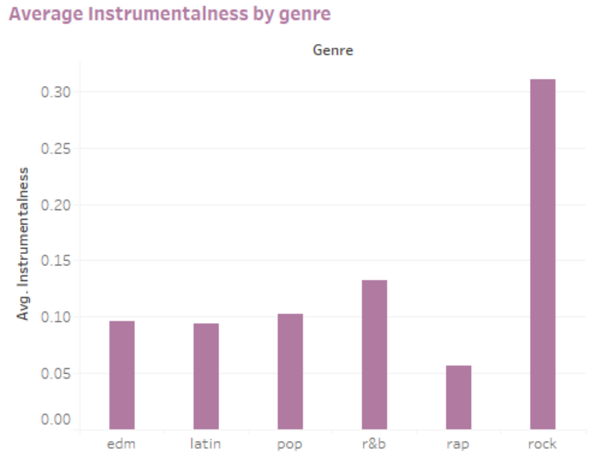
Artist Glacomo Puccini has 71 flop songs and 107 Hit songs. The highest chorus hit was noticed for sad, depressing songs. The artist has sung the music in all 6 genres. So it has a maximum median section in r&b. And the artist was active for 2 decades, that is decade 2000 and 2010. The artist performed with 20 featured artists in 2000 and for the remaining 10 years. In 2010, the leading artist performed with more than 80 featured artists.

11. Genre Dashboard:

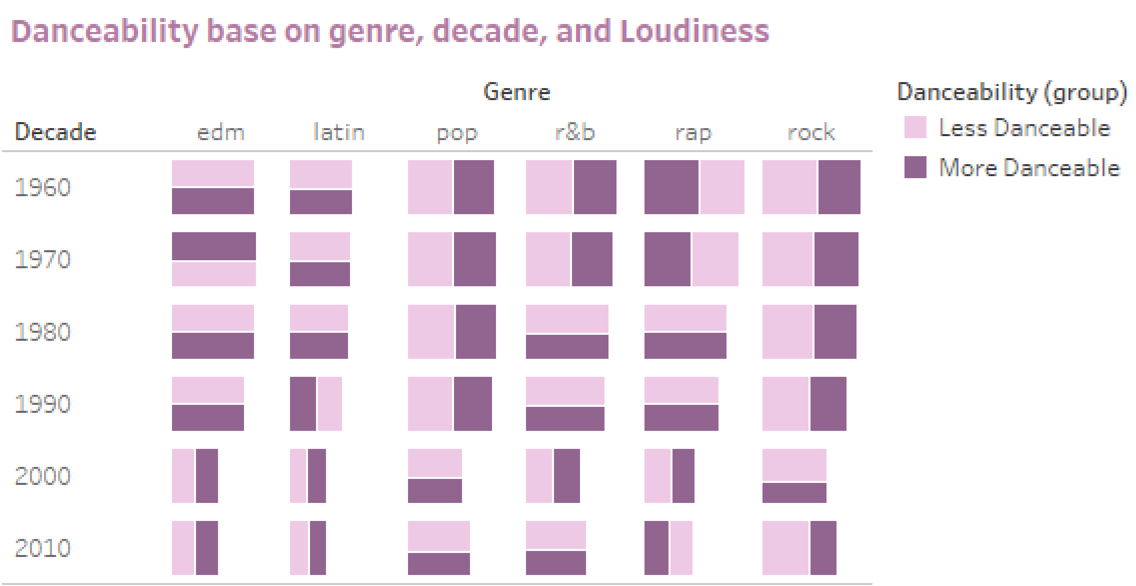


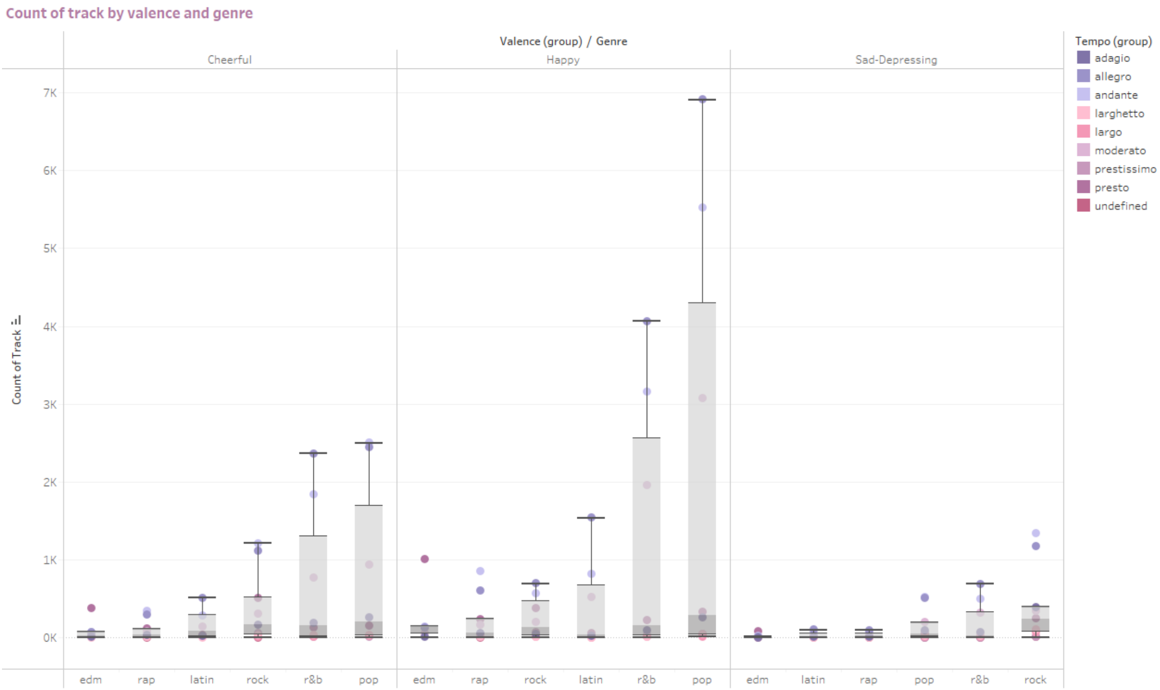






Complete representation of half displayed graph:



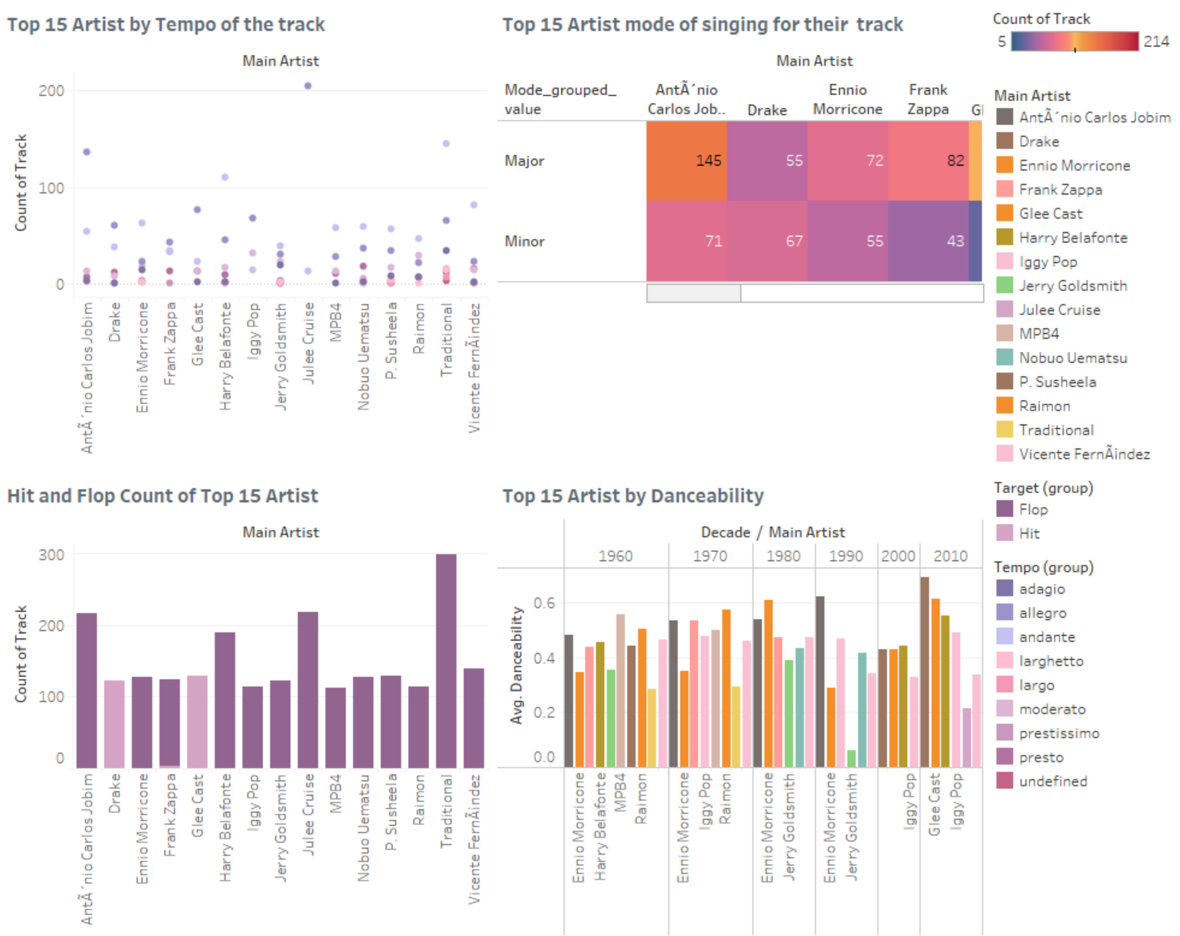


Target Audience and decision-making process: The target audience for genre feature is music listeners, performers, and all the professionals who make an album or record work in terms of creation. It will give the listener some basic information like the genre their favourite song is from and additional necessary information if they are interested in learning. In addition, the performers/artists can improve their performance using the layout. All other music experts can work to develop better musical results by referring to the charts.

Purpose: The layout intends to discuss the genre feature of the music. The sub-objectives desired to be achieved to uncover the loudness, instrumental value, mood of the music record, and danceability value for each genre.

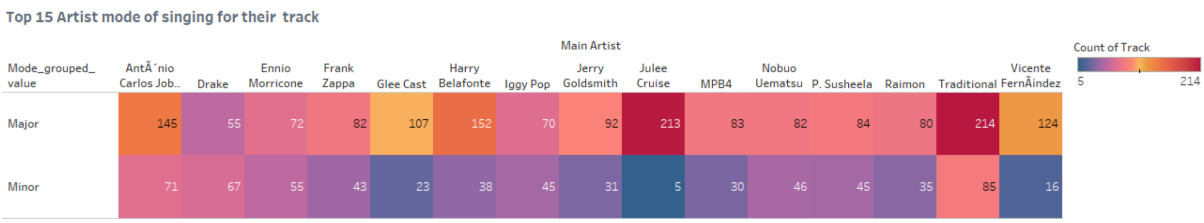
Information it presents and Actionable Insights: The rock genre has the highest value of average instrumental. At the same time, the lowest is in the case of the rap genre theme. The Latin genre has loudness equal to -6.297 bpm. It is closest to 0 and easily audible to the audience. The rock genre has no high danceable value in any decade except for 2000 and the following nine years. The maximum danceability was seen in the rap genre in 1960. The EDM genre has hardly any sad, depressing songs.

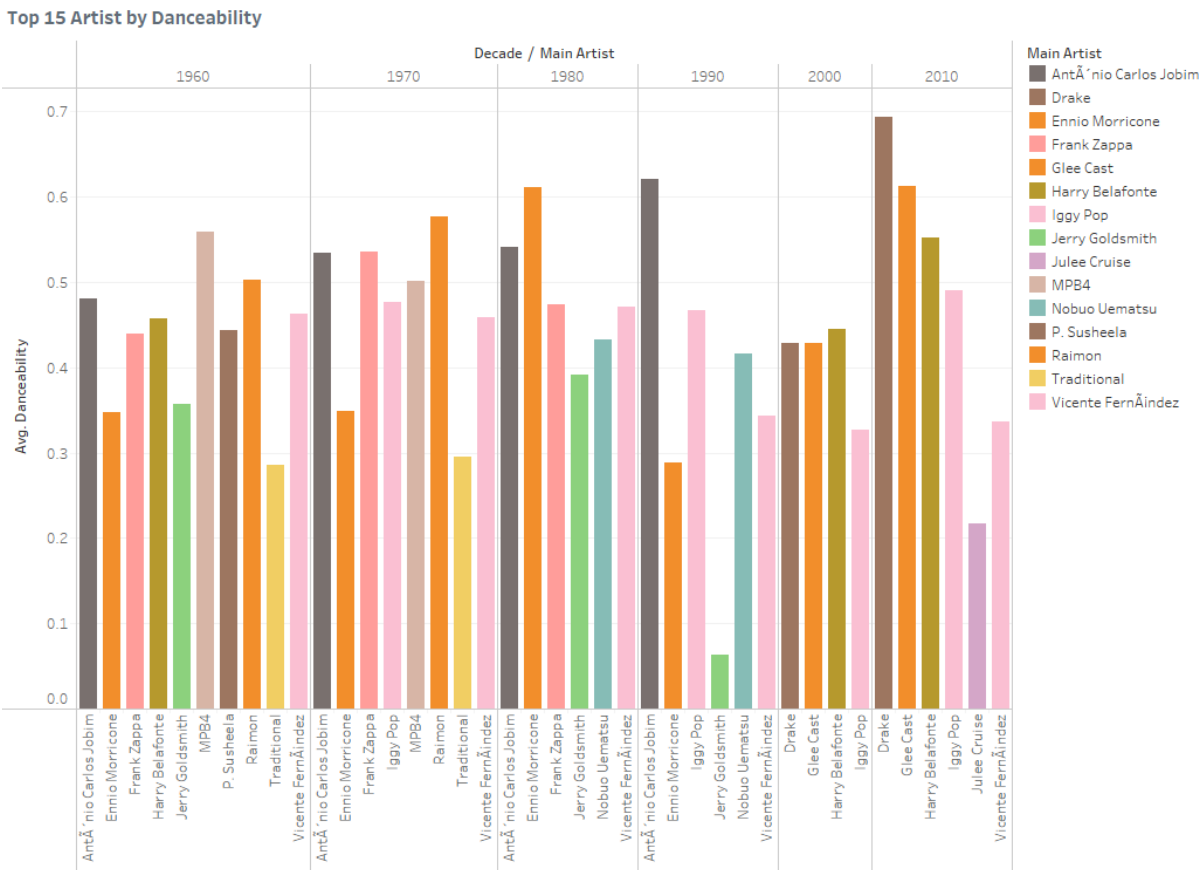
12. Top 15 Artists Dashboard:



Note: Here the top 15 artists are selected based on the artist who has produced a maximum number of songs. A range of other filtering conditions and be added to uncover details about the music artist or so-called maestro.

Complete representation of half displayed graph:



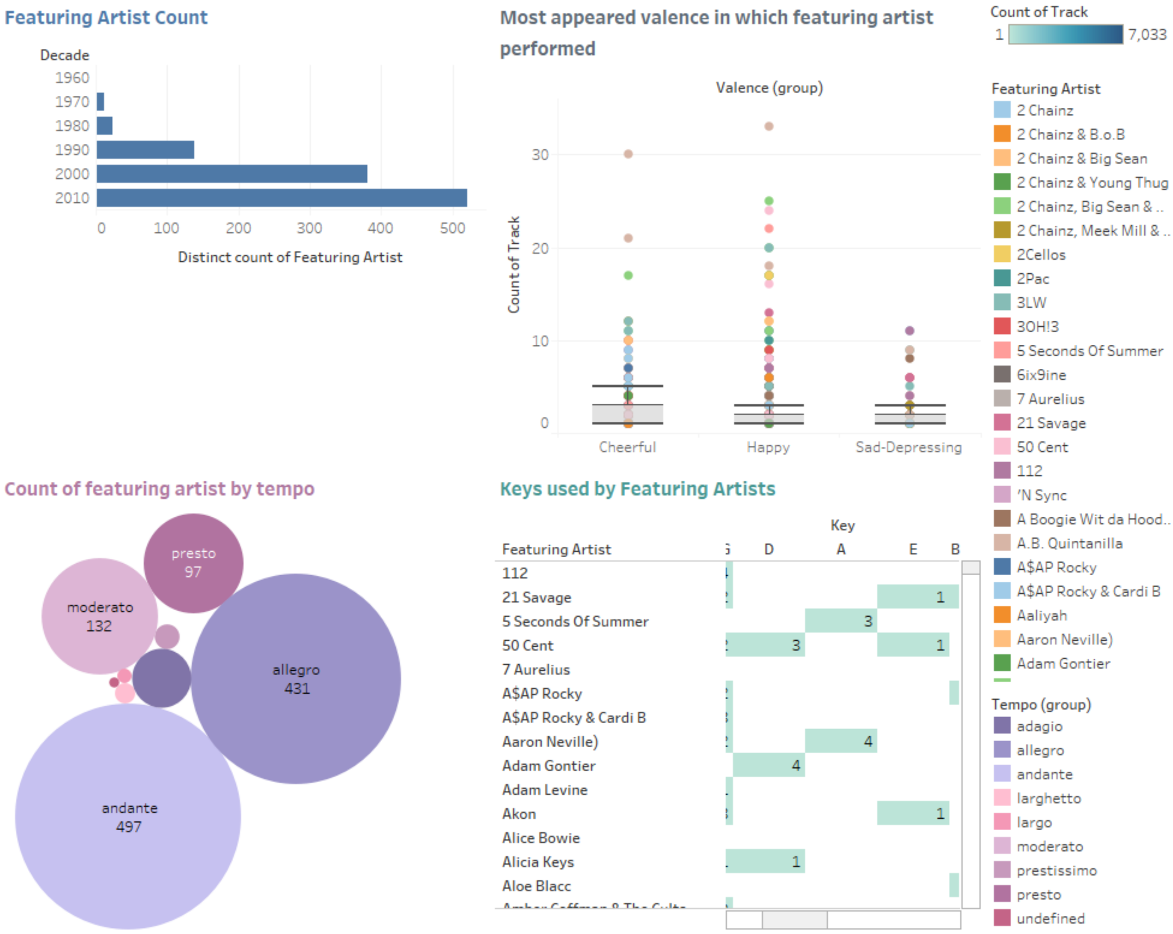


Target Audience and decision-making process: The artists are the beneficiary of the dashboard. They can comprehend their past performance using the available features. It will aid artists in performing better in the future.

Purpose: The goal of the layout was to investigate most of the musical features of the leading artist. The top singer is the focus here, and the sub-goal is to locate a pattern in their work and creation. The prime attribute of artists is examined with track count, a mode they have sung the song in and average danceability. Additionally, it checks the tempo and decade in which most of the songs were recorded. The hit and flop count of the artist was also reviewed.

Information it presents and Actionable Insights: No artist has sung more minor songs than major mode songs. Artist Traditional has song the most major songs from the top 15 Artists. Julie Cruise sings the least songs in Minor Mode. Every singer from the top 15 has sung a song for more than 1 decade. Even though traditional has sung the most number of pieces, his songs are a flop. Drake has produced more than 100 records, and all are hit. Drake can sing in five different tempo ranges. The tempo types are adagio, allegro, andante, moderato, and undefined.

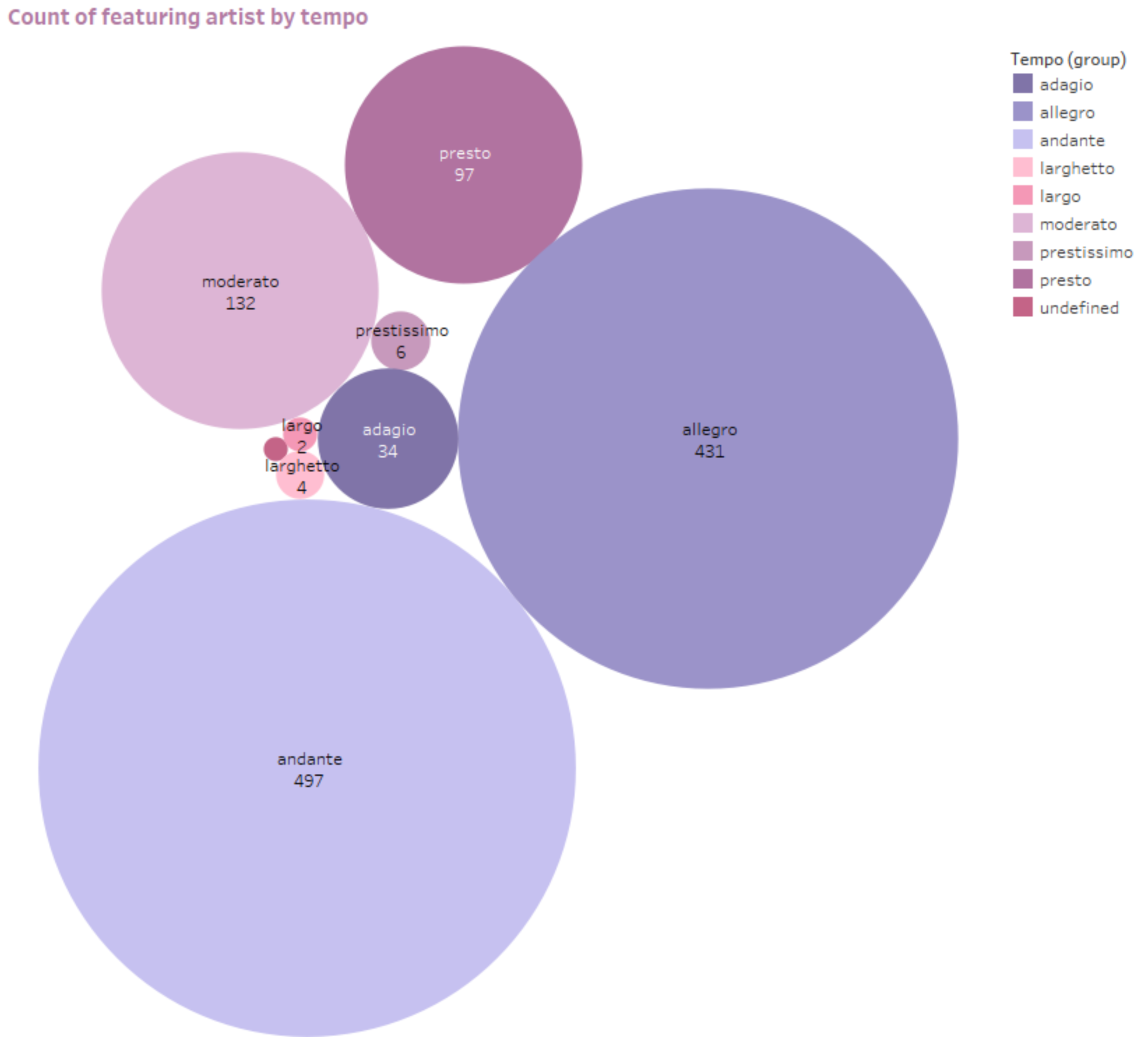
13. Featuring Artist Dashboard:



Complete representation of half displayed graph:

Keys used by Featuring Artists

Featuring Artist	C	G	D	A	E	B	Gb	Db	Ab	Eb	Bb	F
						Key						
112		4						2	1			
21 Savage		2			1	2	1		1		1	2
5 Seconds Of Summer				3								
50 Cent		2	3		1			2	1	1		
7 Aurelius												3
A$AP Rocky		2				1						
A$AP Rocky & Cardi B		3										
Aaron Neville)		2		4								4
Adam Gontier			4									
Adam Levine		1						1		1		
Akon		3			1		1		8	5	4	4
Alice Bowie												2
Alicia Keys		1	1									
Aloe Blacc						2						
Amber Coffman & The Cults		2										
Amy Lee				5								
Andre 3000										2		
Angelina												11
Ariana Grande						2		2	1			
Ashanti		1				1			1	1		
Ashley Monroe			1			1						
Athena Cage										4		
Avant							1				2	
Ayana											2	
Az Yet											2	
B.o.B		6				2						
Baby Bash		4	1						1			
Baby, Turk, Mannie Fresh, Juve..			4									
Bebe Rexha								5				
Bebe Rexha & Digital Farm Ani..						3						
Benny Benassi						3						
Beyonce			1	6								1
Big Sean & Stefflon Don			7									

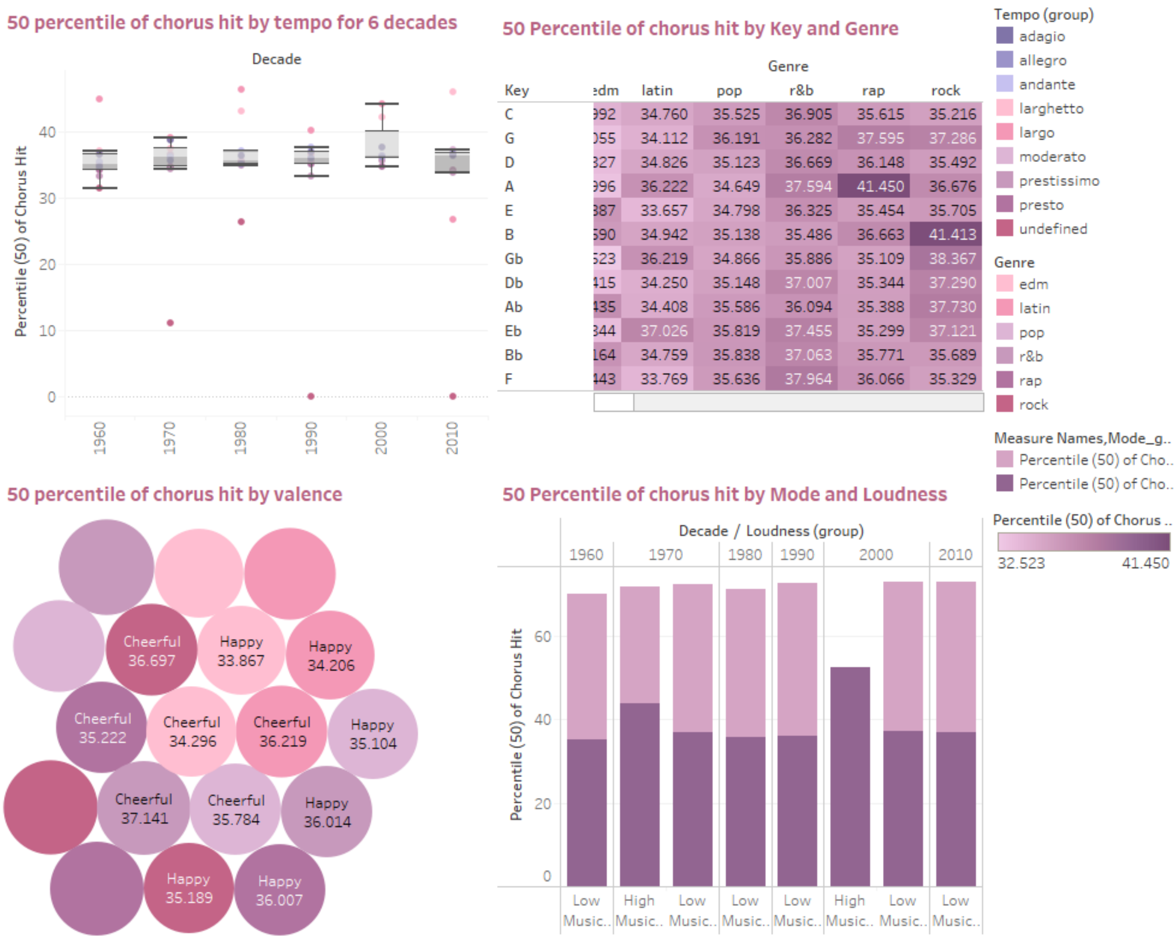


Target Audience and decision-making process: Both artist and the featuring artist will be the target audience for this dashboard. The leading artist will get information about the featuring artist's overall performance and the features affecting their performance. The featuring artists can spot patterns in their performance. As a result, a good set of artists will be able to come together and create good, high-quality, and trending tracks.

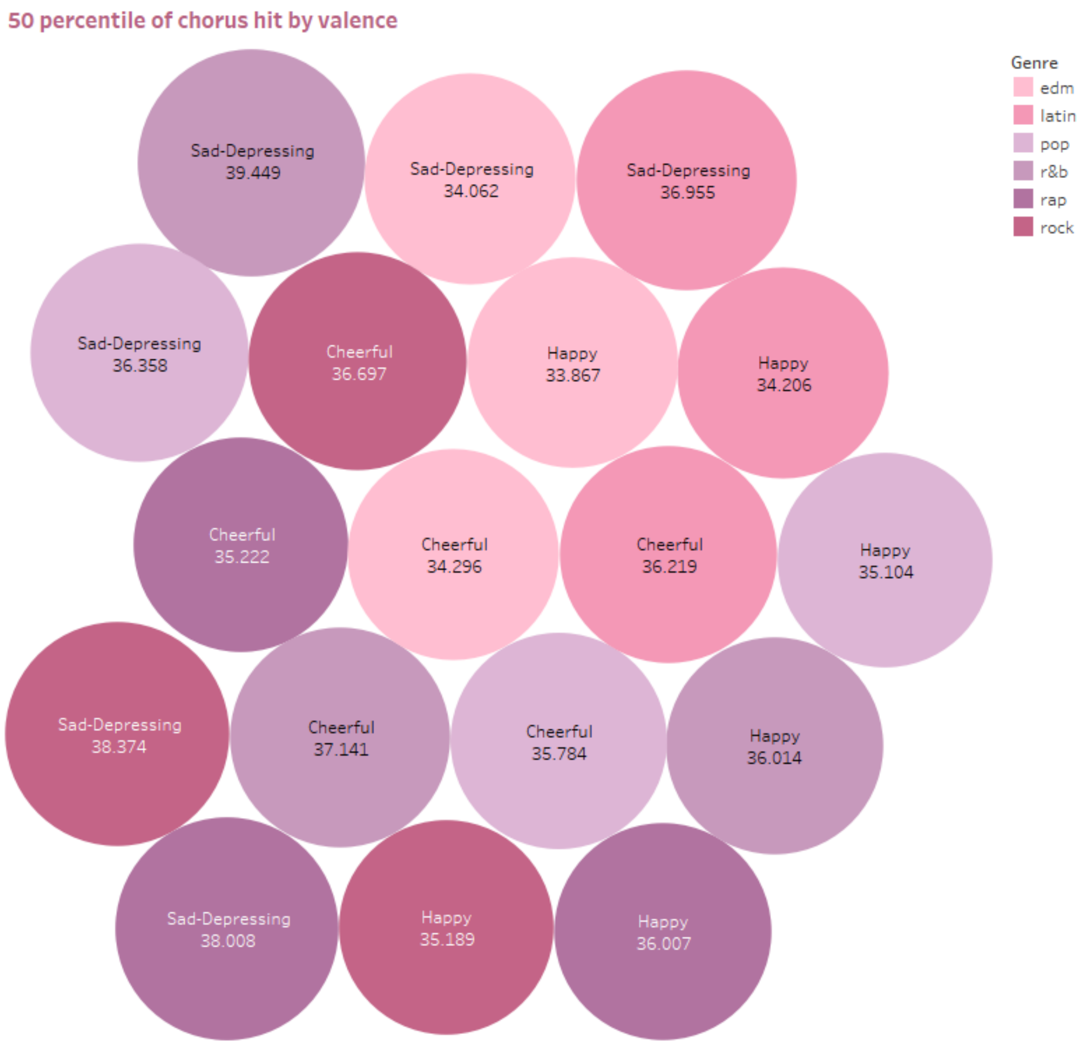
Purpose: The motive for the featuring artist layout is to understand the influence of musical features on featuring artists' work/performance and investigate if there is any pattern featuring artists follow. In addition, the count, tempo employed, valence they appeared the most, and musical key preferred are probed for better interpretation.

Information it presents and Actionable Insights: The count of featuring artists in presto tempo was 97. And it was 497 for andante tempo. The most distinct featuring artist was seen in 2010 and the following nine years. There were no featuring artists in 1960 who were unique. Artist 112 has sung the saddest and most depressing songs.5 seconds of the summer group have used key A in 3 songs. Alicia Keys has used key D in one piece.

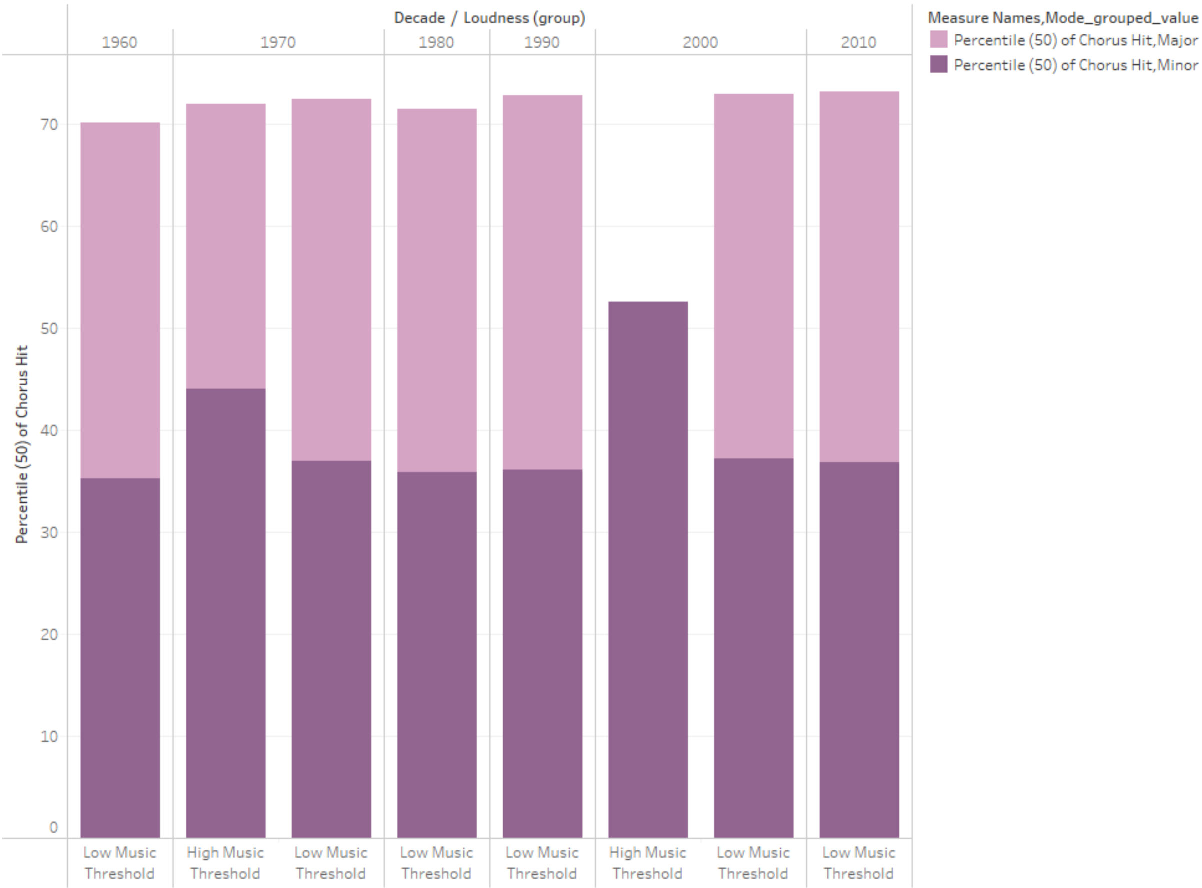
14. **50 Percentile Chorus Hit Dashboard:**



Complete representation of half displayed graph:



50 Percentile of chorus hit by Mode and Loudness



50 Percentile of chorus hit by Key and Genre

	Genre					
Key	edm	latin	pop	r&b	rap	rock
C	33.992	34.760	35.525	36.905	35.615	35.216
G	33.055	34.112	36.191	36.282	37.595	37.286
D	33.327	34.826	35.123	36.669	36.148	35.492
A	32.996	36.222	34.649	37.594	41.450	36.676
E	35.887	33.657	34.798	36.325	35.454	35.705
B	35.590	34.942	35.138	35.486	36.663	41.413
Gb	32.523	36.219	34.866	35.886	35.109	38.367
Db	33.415	34.250	35.148	37.007	35.344	37.290
Ab	36.435	34.408	35.586	36.094	35.388	37.730
Eb	33.344	37.026	35.819	37.455	35.299	37.121
Bb	36.164	34.759	35.838	37.063	35.771	35.689
F	35.443	33.769	35.636	37.964	36.066	35.329

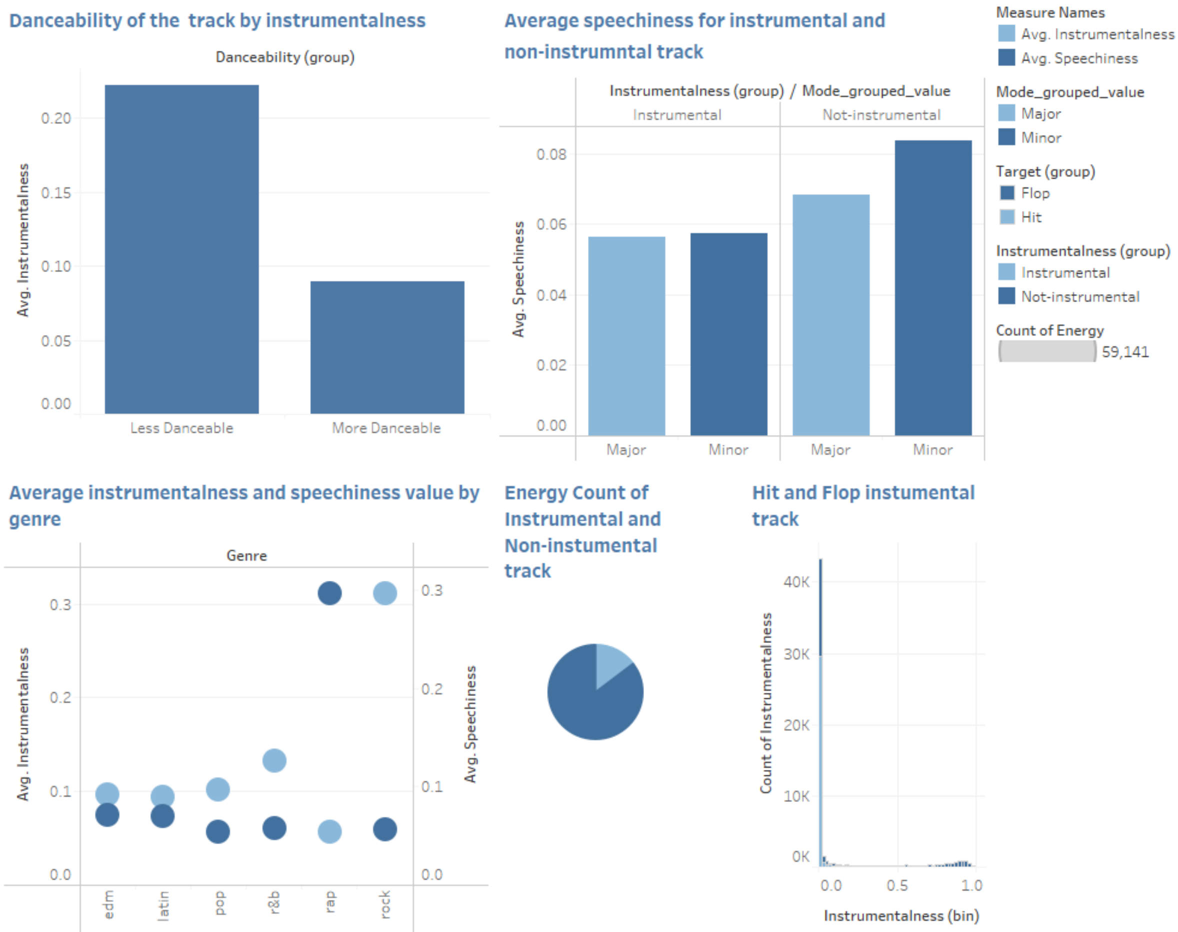
Percentile (50) of Chor..
32.523 41.450

Target Audience and decision-making process: The lyricist and the singers will be the target audience for the dashboard mentioned above. Therefore, the layout will guide these music experts and let them know what they should not do regarding the 50 percentile of chorus hits.

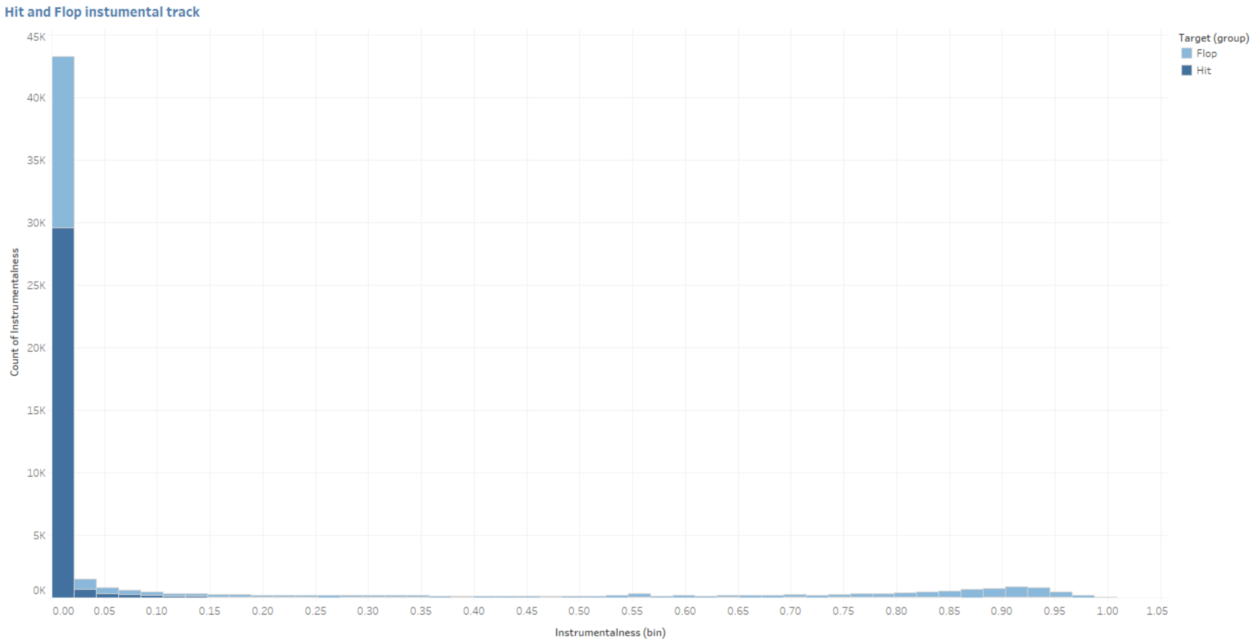
Purpose: The dashboard's purpose was to analyse the affinity between chorus hits of the soundtrack with some of the musical features obtained from the dataset. Here only the middle 50 percentile of the records for each decade is checked to spot if there is any change in the pattern. Simultaneously, the chorus hit is examined by genre, key, valence, loudness, and mode (major-minor mode) of the music record. Ultimately, the track's tempo is explored to learn the feature behaviour.

Information it presents and Actionable Insights: Key A has a maximum 50 percentile value under the rap genre. The value is 41.450. That is followed by key B in the rock genre. In the 2000 decade, there were no 50 percentile chorus hits found in the major mode section and high music threshold. The 50-percentile chorus hit value for each genre and mode has been specified separately.

15. **Instrumental Dashboard:**



Complete representation of half displayed graph:



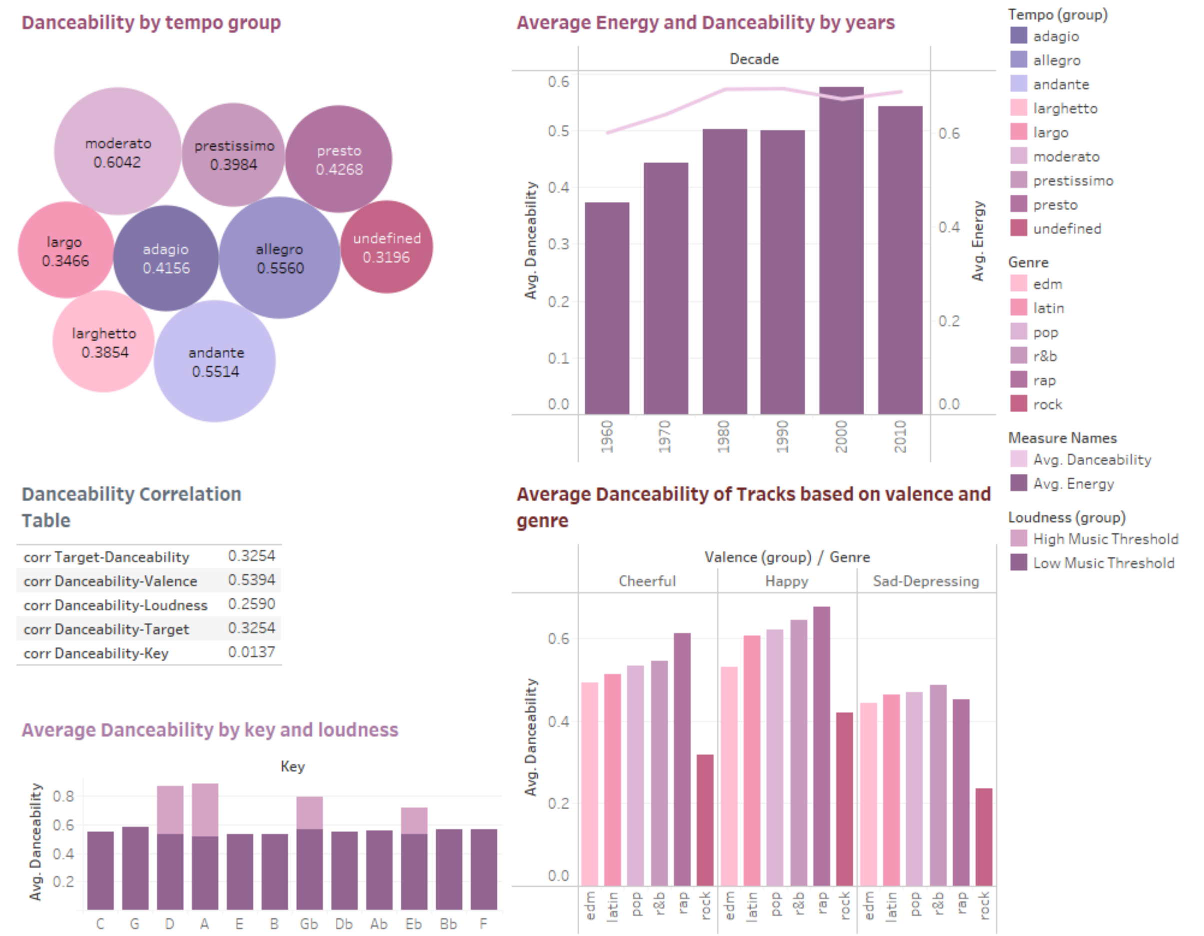
Target Audience and decision-making process: The target audience will be musicians. It will help musicians make more hit songs using all the musical components displayed in the graph.

Purpose: This final dashboard aims to learn about the instrumental feature of soundtracks. It is investigated if the high or low instrumental value affects the music record's danceability, wordiness, or mode. In addition, the hit-flop and energy count of the music are also located based on instrumental value. Finally, the average instrumental value is compared with the song's words.

Information it presents and Actionable Insights: The histogram is made for instrumental features, and different colours further separate the values for hit and flop songs. The danceability group is checked with an average instrumental value. There is a comparison between instrumentals and speeches done. From the comparison graph, it is uncovered that when the instrumental value of the rap genre is high, then the lengthy value is meager. The exact opposite case is noticed with the rock genre. The average speechy and instrumental values are almost identical for EDM songs.

PART 2

1. **Danceability Correlation Dashboard:**

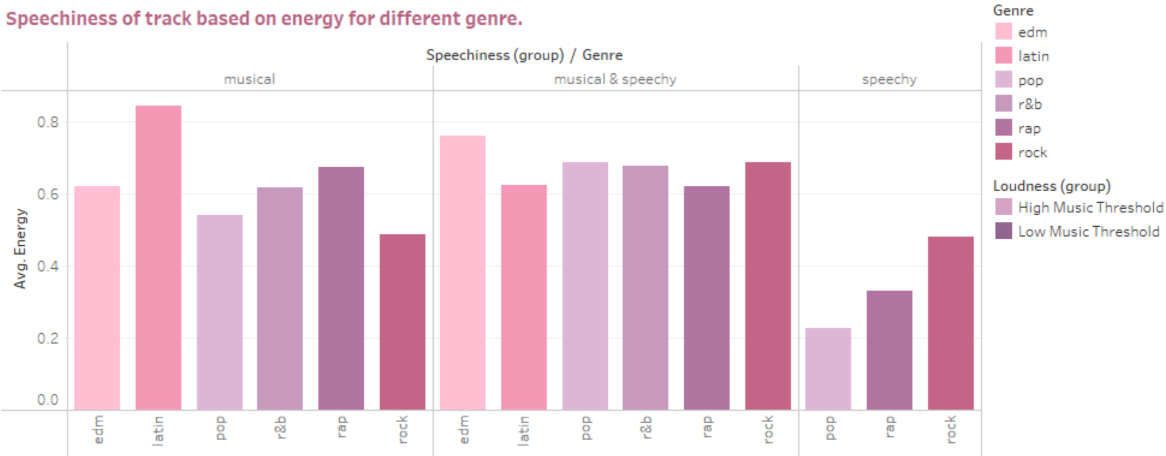


Target Audience and decision-making process: The target audience for the danceability dashboard is music **directors, choreographers, and everyday listeners.** This will aid the typical frequent listeners select the music based on danceability. They don't need to be a maestro to choose a soundtrack to dance. But the layout will lead to a better choice. Furthermore, it will permit the dance choreographers to pick a music piece from different decades based on the situation**. Finally, the music director is the most essential beneficiary, also called the conductor. The conductor can intensify or make changes to their record based on the danceability requirement. For instance, the condition is a danceable song for a sad-depressing situation or a happy scenario.**

Purpose: The purpose of the dashboard is to find the correlation of danceability with other musical components using different graphs and correlation formulas or diagram

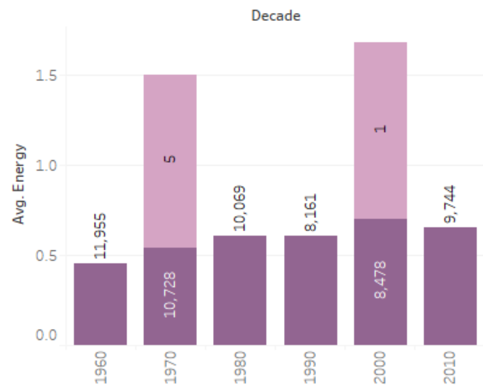
Information it presents and Actionable Insights: Maximum correlation was found between Danceability and valence. The danceability is least in case of rock genre in sad and depressing music category. The danceability of the song increases with the increase in its energy.

2. Energy Correlation Dashboard:



Energy Correlation Graph

corr Energy-Acousticness	-0.710
corr Energy-Liveness	0.156
corr Energy-Speechiness	0.137
corr Energy-Danceability	0.213
corr Energy-Chorus Hit	0.006
corr Energy-Instrumental	-0.202
corr Energy-Valence	0.341
corr Energy-Tempo	0.213
corr Energy-Loudness	0.765
corr Energy- Duration	0.020
corr Energy-Section	-0.041
corr Energy-Mode	-0.049

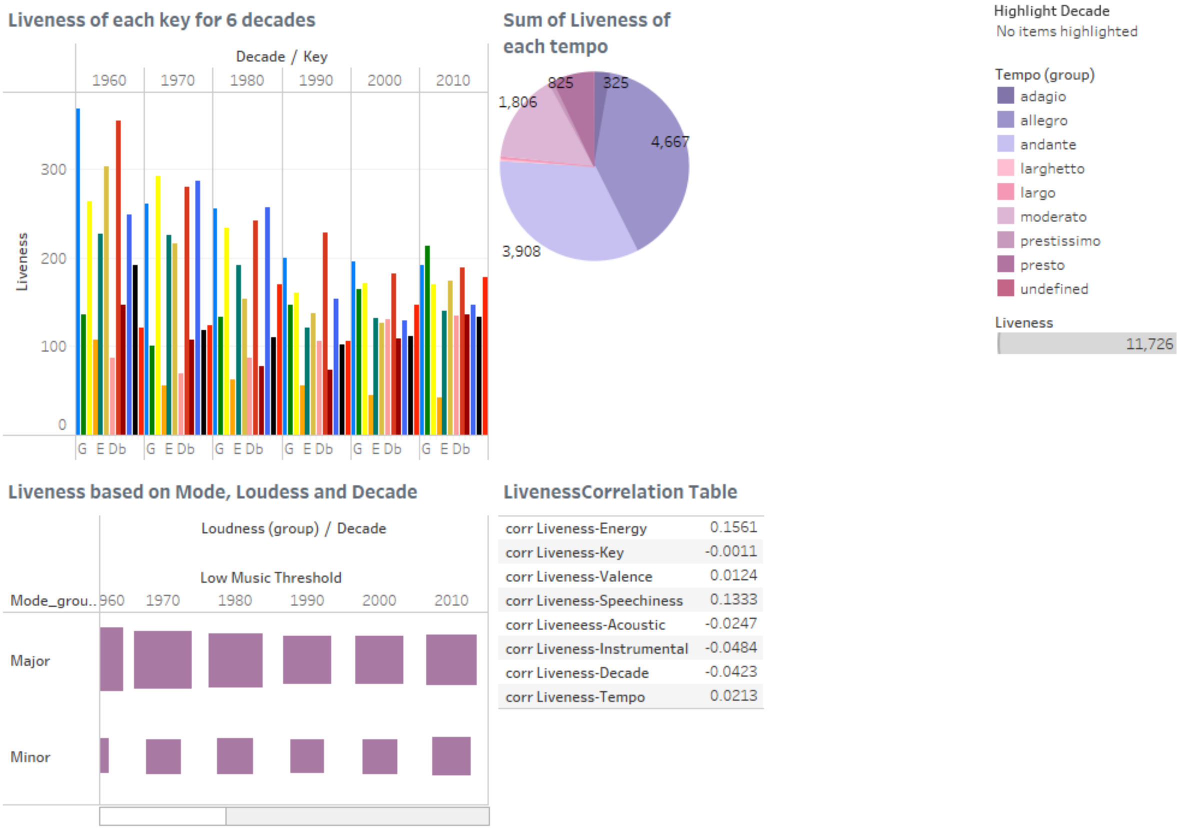


Target Audience and decision-making process: The prime beneficiary of this layout will be the **artists themselves, music composers, and listeners**. First, the performer/singer will recognize the key and genre they are more comfortable with while creating the melody. What is their liveness when they sing a song with a particular key? Finally, they can ascertain their energy range. And work on improving it if required. Listeners can pick songs and artists depending upon the energy range they favour. They can also consider the genre. The layout will work as a blueprint for the music composers can work on developing music after understanding the energy range of the artist. They can also unveil the energy trend in different decades before releasing their work. Since energy is **the sense of forwarding motion in music, whatever keeps the listener engaged and listening.** Composers can **write and arrange music for various media, including film, tv, stage productions, video games, and advertisements, as per the energy requirement in that scenario.**

Purpose: The purpose of the dashboard is to find the correlation of energy with other musical components using different graphs and correlation formulas or diagram

Information it presents and Actionable Insights: Highest correlation was seen between Energy and Loudness. The correlation is 0. 765.It is a clear indication if the music is loud there is definitely a chance it will have high energy. The average energy of pop song is least when the track comprises of too many words.

3. Liveness Corelation Dashboard:

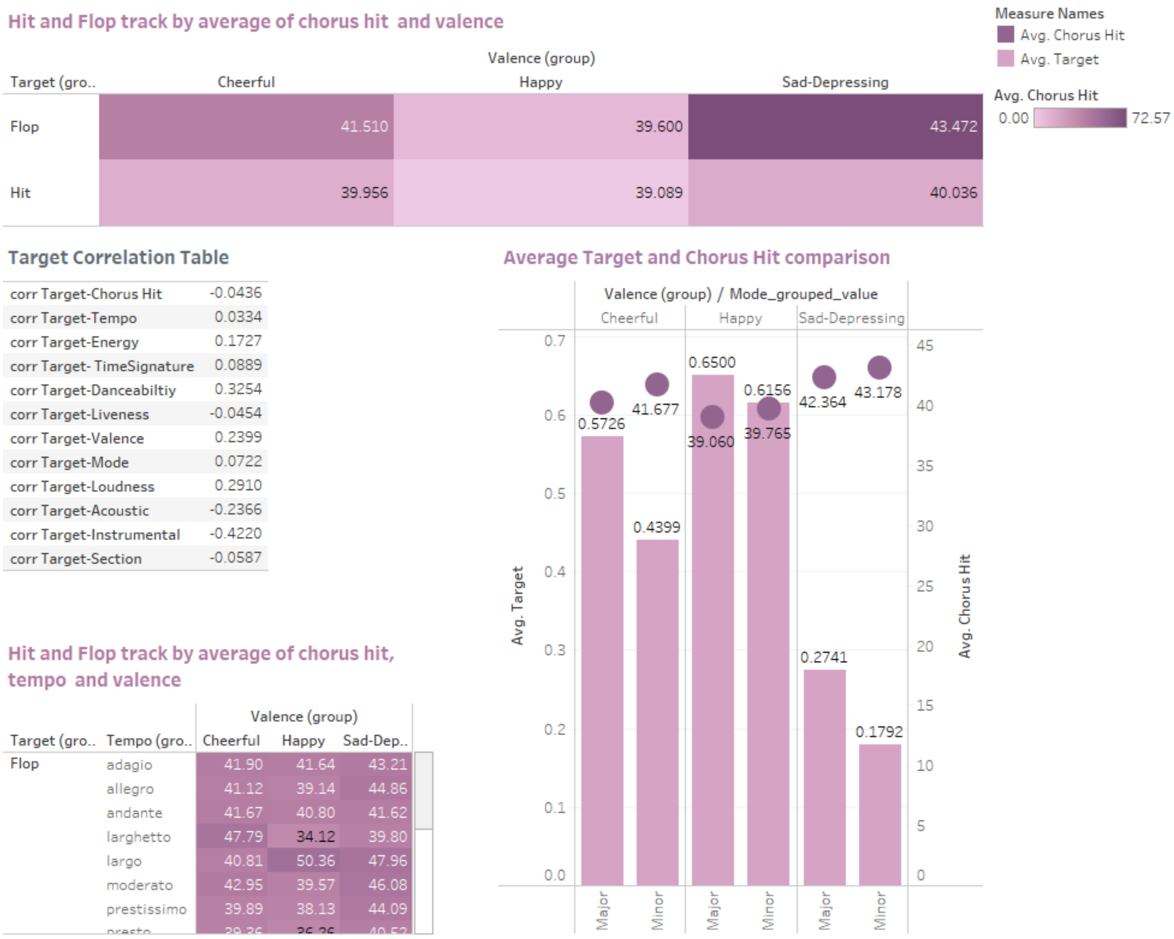


Target Audience and decision-making process: The beneficiary of the layout is performers (singers) and music directors. It will let the performers know that they need to work on melodies' energy and not just their energy to increase the overall liveness of the song. Likewise, the music director can improve the overall result of the soundtrack once they detect the affinity of one feature with others. It will ultimately lead to a high probability of top musical records and sweeten the overall performance of their music record company. "Liveness" refers directly to reverberation time. A live room has a long reverberation time and a dead room a short reverberation time.

Purpose: The purpose of the dashboard is to find the correlation of liveness with other musical components using different graphs and correlation formulas or diagram

Information it presents and Actionable Insights: Most of the features have negative correlation with liveness. Liveness and speechiness have a little high correlation with each other compared to other set of variables. The allegro tempo has the liveliest songs.

4. Target Corelation Dashboard:

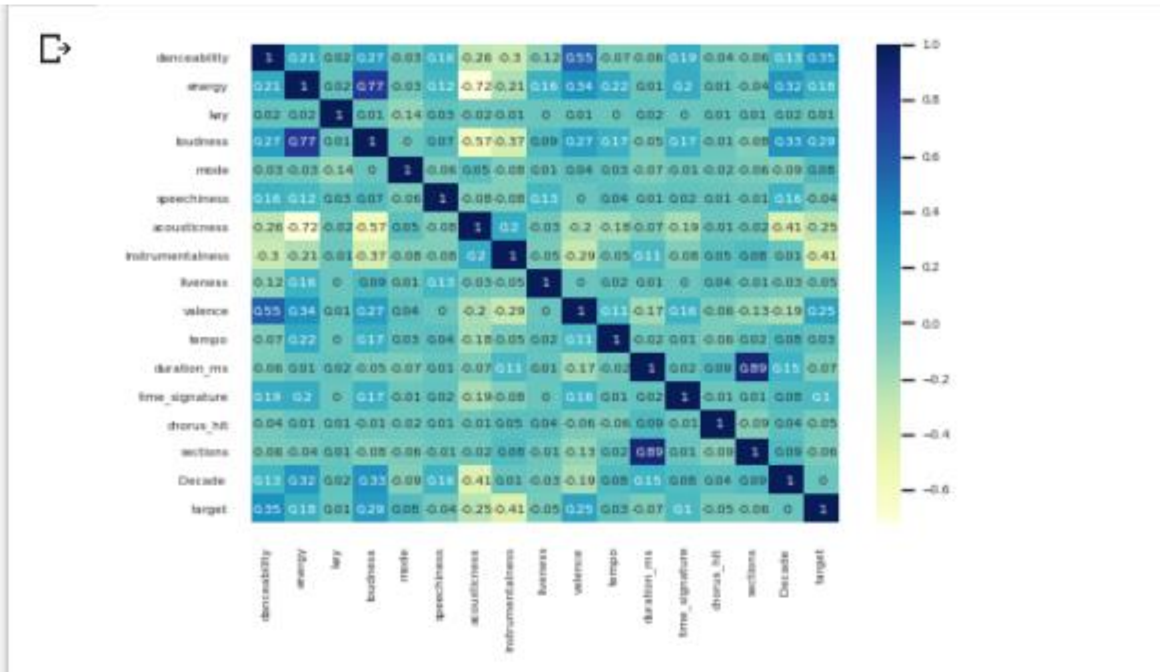


Target Audience and decision-making process: The primary beneficiary of the above-mentioned target audience is record company and singers. Singers can understand which genre they perform best or worst and opt to be part of the album accordingly. Likewise, record companies will get an insight into which artist and genre to invest in so that they make more money.

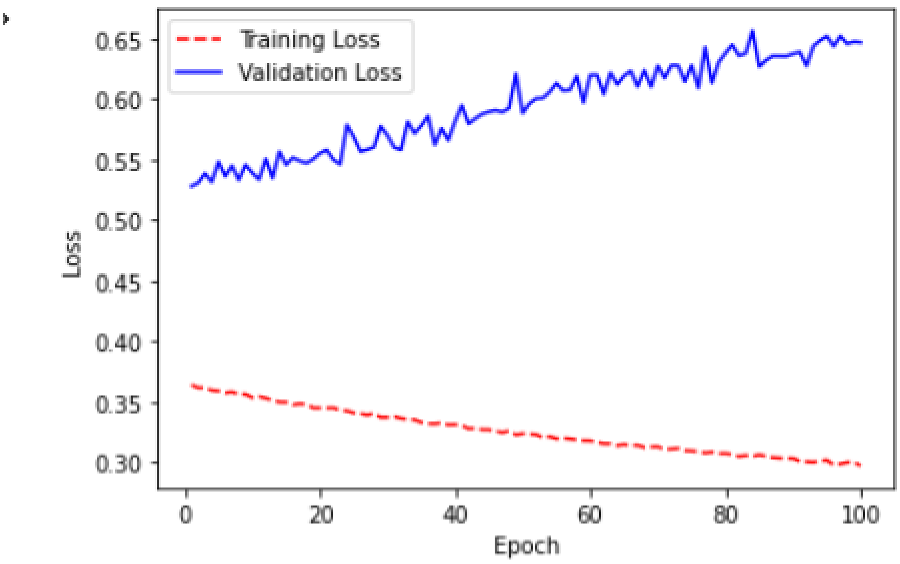
Purpose: The purpose of the dashboard is to find the correlation of target (hit flop feature) with other musical components using different graphs and correlation formulas or diagram

Information it presents and Actionable Insights: All the features had very low correlation with Target feature. Some feature had negative correlation as well. Maximum correlation was seen between target and danceability feature. The chorus hit value is maximum for sad depressing songs.

Correlation Graph for all the features:



Model Related Information: A Deep Learning algorithm is used to analyse the model. Model Analysis Result when a number of epochs are 100 is shown below. The accuracy increases from 75% to 80%, but there is an issue as some test cases are lost. In this scenario, the learning rate is kept at 0.1. And the number of the batch size used is 100.



The number of epochs used finally is ten, and the accuracy attained was 75.15%. The dataset was broken down into training, validation, and testing dataset. The ratio is 80:10:10. Features like Track, URI, Genre, and Artist were not used while training the model.