– Independent Work Report Fall, 2018 –

Visualizing and Predicting Spotify Genre Characteristics

Savannah Du
Adviser: Brian Kernighan

Abstract

This paper details a quantitative approach to analyzing how music genres have changed over time. Using data from the Spotify API, this project addresses the lack of research in the field of music informatics on the evolving traits of traditional genres and the rise of new genres by visualizing and modeling musical characteristics. A bubble chart displaying genre characteristics using color and position allows users to view similarities between genres and correlations between characteristics such as energy and tempo. To supplement user understanding of how genres have changed over time, a motion graph plots musical relationships from 1960 to 2017. To further quantify and concretize musical trends, this project models and classifies genres based on their musical similarities. The resulting work allows users to view the danceability, energy, instrumentalness, speechiness, tempo, and valence of popular Spotify genres and reveals cyclic patterns in many genres' characteristics, providing an example for future work on quantifying musical trends.

1. Introduction

1.1. Motivation

A recent Spotify advertising campaign swamped billboards worldwide with colorful displays of listener statistics, making not-so-subtle references to relevant political and social events and figures. One sample advertisement offered a comforting remark to commuters on the London Tube: *Dear 3,749 people who streamed "It's the End of the World As We Know It" the day of the Brexit vote, hang in there* (Image 1). These advertisements introduce the theory that music consumption is shaped not just by listeners' personal preferences, but also by the current sociopolitical climate. With millions of users worldwide, music streaming services like Spotify and Apple Music have

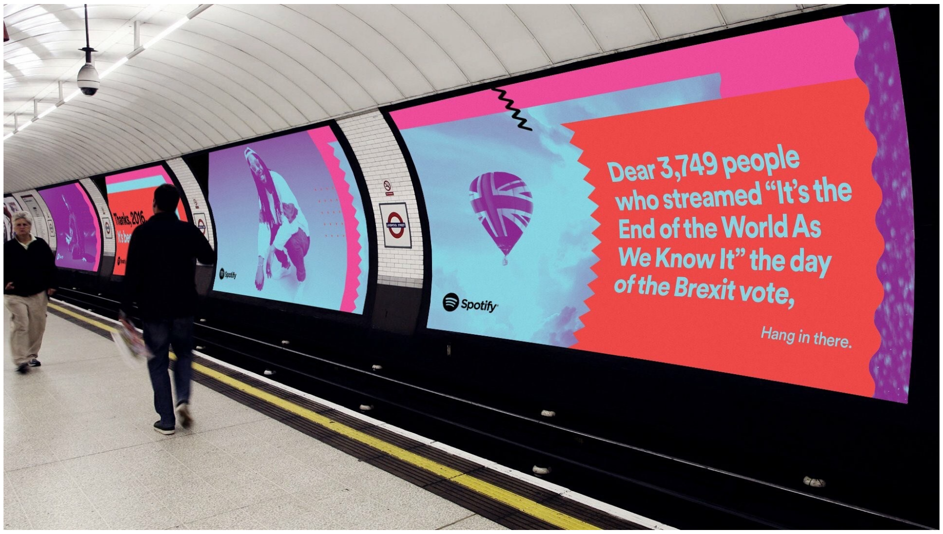

Image 1: Spotify advertising campaign in the London Tube.

access to huge amounts of musician and listener data. While music radio stations tend to offer a limited range of genres, often repeating the top 20 hits in a select genre for months, the new methods of consuming music allow listeners to experiment with a much more diverse set of tracks. Spotify playlists like "lush lofi," which manage to combine hip-hop instrumentals with jazzy vibes, have made it difficult to answer the questions of "what's your favorite genre?" or "what kind of music do you listen to?" Easier access to more diverse styles is not the only thing that has changed; the advance in technologies making the rise of electronic music possible has influenced other, more traditional, genres as well. Switching back and forth between top pop songs from the past few months versus past few years, I have noticed that an increasing number of pop songs have started to include electronic elements, often incorporating much stronger dance beats and more energetic vocals; the line between pop, electronic, and other genres has become blurred.

How have music genres changed over time? How have traditional genres shifted with changing listener preferences, and where on the musical spectrum have new genres emerged? While certain genres, such as blues and rap, are often associated with specific social or historical movements, this project focuses on quantifying the changes in the musical characteristics of popular genres over time. Taking advantage of a relatively recent digitization of musical records which has made music analytics more viable, this project provides an alternate lens through which listeners and musicologists alike can view genre.

1.2. Goal

In order to provide a more quantitative view of how popular genres' musical characteristics have changed over time, this project aims to:

1. Advance user understanding of *the relationships between genres and their musical characteristics* by providing a visualization of the similarities between genres now and in the past, as well as correlations between different musical characteristics.
2. Show users the *changes in musical relationships over time* by providing a visualization of the shifts in musical characteristics of popular genres.
3. Model and classify genres based on their musical characteristics and predict *future trends* based on similarities to genres of the past.

1.3. Problem Background and Previous Work

Recent research in music informatics, an emerging field dealing with the production, distribution, consumption, and analysis of music through technology, has taken advantage of a growing library of digitized audio recordings and technologies available for audio analysis. One study has focused on categorizing modern songs into traditional genre buckets and tracking the quantity of tracks in each bucket over time, constructing a taxonomy of genres styles [12]. Others have focused less on genre classification and more on tracing musical evolution by analyzing a few selected characteristics in popular music of the past few decades and in Western classical music of the past centuries [13, 14].

Studies on song classification are valuable for observing how musician and listener preferences have changed over time, but neglect possible transformations within genre buckets themselves. The technological advances in music production and editing in the recent few decades have changed the way musicians create and listeners consume, resulting in a range of new and quite popular electronic genres. Research focused on more traditional genres, such as country and rock, may not be as adaptable and relevant to projecting future trends as work that incorporates a more fluid set of genres. On the other hand, research focused on tracing overarching musical changes with no regard for song categories is less robust because it does not take into account an increasingly diverse range of genres and styles that may be trending in opposite musical directions and mirror vastly different listener demographics.

This project addresses the focuses of previous studies and offers a more balanced approach by observing a limited selection of musical characteristics within the context of popular genres, tracking changes over time both within genres and over popular music as a whole.

2. Approach

I built my dataset using Spotify's public API, which contains metadata on each of its artists, albums, and tracks. For each of the 30 millions total tracks in its database, Spotify provides information on the track's release year and album genre, as well as values for a number of musical characteristics, such as danceability, timbre, and key, which are calculated based on an audio analysis of the track.

In order to construct a dataset that was focused enough to produce meaningful conclusions, I chose to include only songs that were categorized as one of the most popular Spotify genres within the past century. My reasoning behind this limitation was twofold: first, from a technical standpoint, selecting genres with large numbers of tracks reduces the possibility of outliers significantly affecting the conclusions; second, from a musicology perspective, focusing the analysis on genres that are relevant to a larger subset of listeners means that musical changes can be viewed in the context of larger social movements or historical events.

Since Spotify's records of popularity, either by track or by genre, are not publicly available through the API, I looked into alternate sources to judge popularity, such as Billboard or Shazam top lists. I had difficulty finding lists for earlier decades, so I instead drew from a project created by a Spotify developer that calculates popular genres for each year from 1960 to 2017 by "taking the currently most-popular songs (as of 2017-11-30) from each year back to 1960, finding the genres to which their artists correspond (sometimes in hindsight), and then ranking those genres according to the fraction of that year's artists they each represent," resulting in 160 unique genres [3].

Out of all the musical features that the Spotify API calculates for each track, which range from characteristics of the music itself to how 'live' the track sounds, I limited my analysis to six characteristics that were most quantifiable, easily heard and understood, and likely to be related to or affected by other characteristics in the set. The six characteristics are explained in Table 1.

Characteristic	**Summary of Spotify description**	**Value Range**
1. Danceability	How suitable a track is for dancing based on tempo, rhythm stability, beat strength, and overall regularity. A higher value is more danceable.	[0, 1]
2. Energy	A perceptual measure of intensity and activity based on dynamic range, perceived loudness, timbre, onset rate, and general entropy. A higher value is more energetic.	[0, 1]
3. Instrumentalness	A prediction of whether a track contains no vocals. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.	[0, 1]
4. 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 higher the speechiness.	[0, 1]
5. Tempo	The overall estimated speed of a track in beats per minute (BPM). A higher value means a faster track.	[50, 220]
6. Valence	A measure of the musical positiveness conveyed by a track. A higher value means a happier track.	[0, 1]

Table 1: Selected musical characteristics.

For my visualizations, I first considered comparing the distributions of genres' musical characteristics using histograms and tracking change over time using steamgraphs. However, since these methods of visualizing data are mostly one-dimensional and would make it difficult to compare metrics between genres or among multiple musical characteristics at once, I decided to use a bubble chart to loosely show relationships among genres and characteristics and a motion chart to show change over time. Example bubble charts represent multidimensional data points as bubbles (circles), with position, size, and color mapping to various features [15]. In representing change over time, I wanted to be able to display the six characteristics while effectively adding time as another dimension. I found that motion graphs, which again represent genres as bubbles, would allow me to show time via animation while keeping bubble size, color, and position open for showing musical characteristics [6, 9].

3. Implementation

In order to achieve my three major goals, I began by collecting data from the Spotify database, which I then used to build out two data visualizations and model and classify genres based on their musical features. Overviews of data collection, visualization, and modeling and classification are outlined below.

3.1. Data Collection

I used the Spotify Web API's search functions to collect data for the selected musical characteristics of popular genre tracks, which I then imported into a SQLite database for further analysis. I began by extracting a list of 160 popular genres from the Spotify developer project previously mentioned. Then, searching through Spotify's database, I used the API to extract only the tracks whose album genres were one of the most popular, resulting in 9 million entries. I imported these tracks into a SQL table with columns for genre, trackID (a unique ID provided by the API), and release year, which I used later to count the numbers of tracks in each genre per year. It is important to note that because many albums are listed as more than one genre, using the API to search by genre returned many tracks multiple times. This meant that, out of the 9 million tracks, there were about 4 million unique. For each unique track, I then used the API to search for musical characteristics by trackID, creating a separate SQL table with columns for trackID, release year, danceability, energy, instrumentalness, speechiness, tempo, and valence. I used the two SQL tables, the first containing genre information and the second containing musical characteristics, to calculate average characteristics for each genre per year by cross-referencing the tables based on the trackIDs and release years.

In order to verify that I had accurately imported all the data provided by the API into my SQL tables, I used Python and SQL script to verify that the total number of tracks in each genre from the first SQL table was equal to the total count directly output by the API. For example, if the API returns 84 trackIDs for rock music in 1960, then there should be exactly 84 rows in my SQL table tagged as rock and released in 1960.

3.2. Data Visualization

To accomplish my two visualization goals — first, showing similarities between genres and correlations between musical characteristics, and second, displaying changes in genres and musical relationships over time — I created a bubble chart and a motion graph using D3.js, a Javascript library for data visualization based in HTML, SVG, and CSS.

3.2.1. Bubble chart: similarities between genres and correlations between characteristics

The bubble chart represents each genre as a bubble. The *size* of a bubble scales to the total number of tracks whose albums are tagged as that genre, the *color* is calculated based on the genre's average values for the musical characteristic(s) selected by the user, and the *horizontal position* of the bubble is based on the average value for another selected characteristic. Described below are the different methods that I used to increase the effectiveness of color as a mechanism for visualizing musical relationships, first using color to represent up to three characteristics, and later using color to represent a single characteristic.

First approach: color determined by multiple characteristics

The first approach allowed users to view color as a function of up to three different musical characteristics, giving them a way of viewing similarities and differences between genres along multiple characteristics simultaneously.

This approach was initially based on the RGB coloring system, which defines a color using a set of three 8-bit integers mapping to red, green, and blue values; $rgb(255,0,0)$ is red, $rgb(0,0,0)$ is black, and $rgb(255,255,255)$ is white. In this implementation, the color of each genre bubble was calculated by scaling the values of up to three musical characteristics to the range [0, 255] to represent R, G, or B values, as exemplified in Table 2.

Genre	Characteristic Values	RGB	HSV
Pop	(0.62, 0.69, 0.02)	(159, 176, 6)	(223°, 0.69, 0.02)
Blues	(0.58, 0.43, 0.12)	(148, 110, 31)	(209°, 0.43, 0.12)
Focus	(0.45, 0.15, 0.85)	(115, 38, 217)	(162°, 0.15, 0.85)

Table 2: RGB and HSV coloring with danceability, energy, and instrumentalness

Professor Adam Finkelstein proposed using an alternate color scale, HSV, which defines colors through hue, saturation, and value (brightness), as shown in Figure 1. HSV is preferable to RGB and often more used in data visualizations and graphics because it is more similar to the way humans perceive color [2]. An artifact of using the HSV color space for my color mapping, however, was that the difference between very dark colors of vastly different hues or saturations was barely distinguishable, as illustrated in Table 2.

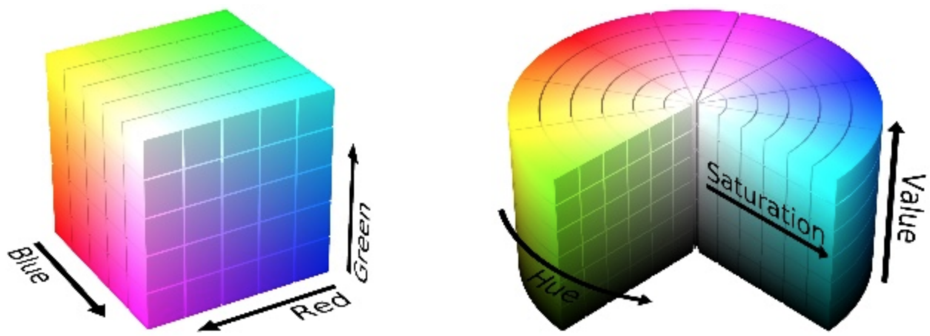

Figure 1: RGB vs. HSV color spaces [5].

Second approach: color determined by one characteristic

A number of users, including more than half of the students in Professor Kernighan's seminar, found the multivariate approach to color difficult to interpret in practice, which led to a second approach which determines color based on the value for a single selected characteristic.

In this approach, the lowest value is represented by one color, the highest value by another, and the values in between map to colors "in between" the two ends. "In-between" colors are determined using color interpolation, a method of filling in color values between two colors for a smooth color gradient.

A simple example of color interpolation is to use white and black for the endpoints, and linearly scale middle values to corresponding values of lighter or darker gray. For example, a genre with a danceability of 0.0 is colored white, a genre with danceability 1.0 is black, and a genre with danceability 0.3 is light gray. However, using white and black as endpoints makes it difficult to distinguish between many of the middle points, since the perceived difference between a light gray and a slightly lighter gray is not that large.

Using different pairs of endpoint colors that are on opposite sides of the color spectrum, such as yellow and blue, offers a range of more easily distinguishable hues that can be used to represent middle values more effectively. D3 provides color interpolations along different color spaces; given two end colors and a float between 0 and 1, it returns a color along a path from one color to the other, where the path is determined by the color space (Figure 2). I found the LAB color space, which moves from yellow through purple to blue, most effective because it offered a greater range of distinguishable hues than the RGB interpolation, but fewer than HSL interpolation, where too many hues could again confuse the user.

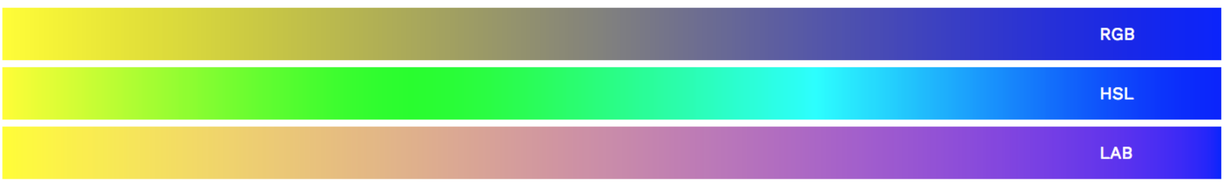

Figure 2: D3 color interpolation between yellow and blue using different color spaces

The effectiveness of the color mapping is also affected by the distribution of the characteristic chosen. For normally distributed data, for example, the majority of points which are colored in the light purple range are again difficult to distinguish, while the extreme yellow and blue ranges are underutilized.

In order to make full use of the color gradient and distinguish the areas where values were concentrated, I tested a nonlinear colormapping, where a small increment in a value in the middle range (the difference between danceability at 0.4 and 0.5) displays a large color change than a small increment in a value in the lower or upper ranges (the difference between 0.0 and 0.1) (Figure 3).

Figure 3: Color interpolation using linear vs. nonlinear colormapping

Since this method was most effective for normal data, I had to manually normalize my datasets by calculating color based on the logs or squares of the data points.

To solve this issue of coloring data with varying distributions, I implemented a final nonlinear colormapping approach based on feedback from Professor Finkelstein, which automatically scales the color for a genre based on the selected characteristic's cumulative distribution (Figure 4). For example, the genre with the median danceability over all genres will be colored as the midpoint color between yellow and blue. This method removes the need for manual adjustment while ensuring that the entire color range is utilized evenly for all characteristics.

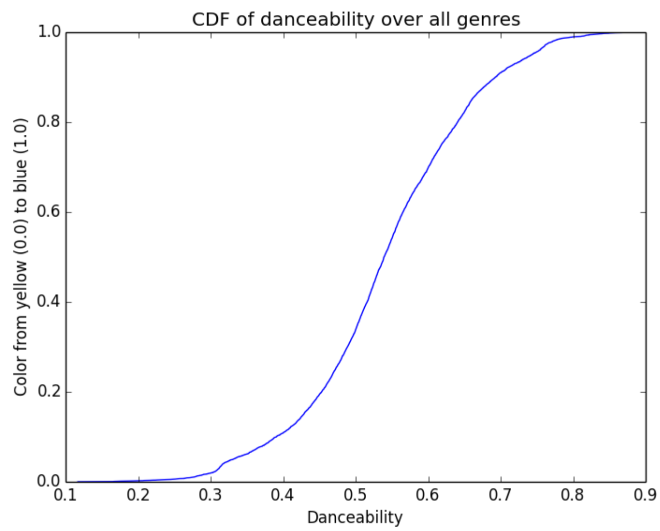

Figure 4: Mapping of danceability values to color

In summary, the multivariate coloring approach has the potential to show relationships among up to four characteristics at once (three with color, one with position), using RGB or HSV color spaces to calculate a single color for each genre. Based on user feedback, I implemented an alternate approach that uses a nonlinear colormapping to calculate color based on a single characteristic; in this approach, users can compare pairs of characteristics using color and position.

3.2.2. Motion graph: changes in genres and musical relationships over time

The motion graph, which complements the bubble chart by adding time as a variable, shows users changes in genres and musical relationships from 1960 to 2017. This second visualization again represents each genre as a bubble. The *size* of a bubble scales to the total number of tracks whose albums are tagged as that genre in the selected year, the *color* of a bubble functions as the genre identifier and remains consistent over changes in year, and the *position* of the bubble on the (x, y)

coordinate system is based on the average values for two user-selected musical characteristic. A slider allows the user to view the average characteristics in a selected year.

3.3. Modeling and Classifying Genre Characteristics

To provide a more concrete and quantitative analysis of genre changes, particularly for those less visually inclined, I worked towards modeling and classifying trends using scikit-learn, a tool for machine learning in Python [7].

Each genre is modeled using six different regressions, one for each musical characteristic. Due to high variation in the characteristics of the individual tracks within genres, I instead used the average characteristics per year as inputs for my models, which yielded more accurate regressions and more apparent downward or upward trends.

The resulting six datasets per genre were split into training sets (a randomly selected 80% subset to prevent overfitting) used for modeling genre behavior and training a classifier, and validation sets (the remaining 20% of the data) used to gage the accuracy of the modeling and classification.

3.3.1. Modeling genre behavior

Characteristic for each genre are modeled using separate polynomial regressions. To accurately reflect the nonlinear genre trends, I used polynomial interpolation to approximate polynomial functions of up to degree 6 by using ridge regression, a model which addresses some of the problems of linear regression by imposing a penalty on the size of coefficients.

To check for overfitting, I used K-Fold cross-validation to evaluate the accuracy of the models. Each dataset is partitioned into k subsets of equal size. The model is trained and tested k times, each time using a different $k-1$ subsets for training and the remaining subset for testing. The average error is calculated and used to evaluate the model.

3.3.2. Classification

For classification, I used scikit-learn's K-Nearest-Neighbors algorithm classifier, which stores the genres of all training points, and labels each test point as the most common genre of the closest k points [10]. For each set of classifications, I set $k = \sqrt{n}$, where n is the total number of points in the training sets used [8]. I tested the predictions for the years not in the training set (output by the models from the previous part), as well as predicted values for 2018 - 2030.

4. Results

Interactive versions of the bubble chart (using nonlinear mapping) and motion graph can be viewed at https://savdu.github.io/spotify-genres/.

4.1. Bubble chart

The bubble chart aims to advance user understanding of the relationships between genres and their musical characteristics.

4.1.1. Multivariate color approach

Figure 5 shows the bubble chart with color determined by mapping multiple variables to RGB values. Danceability / energy / valence are mapped to Red / Green / Blue, and position is determined by danceability, with genres of lower danceability to the left and higher to the right.

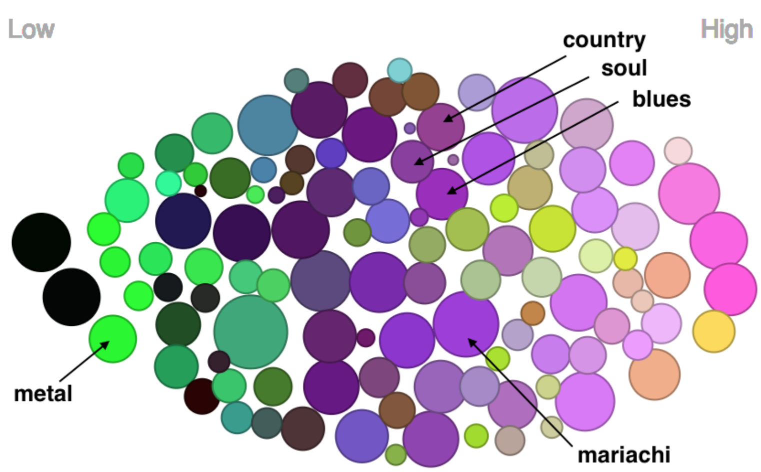

Figure 5: RGB color mapping.

In this implementation, users made a few successful observations relating different musical characteristics. The high concentration of pink and red bubbles on the right (genres of higher danceability) and cluster of green bubbles on the left (genres of high energy, low danceability) suggest that genres with high energy are not necessarily the most danceable. The green bubbles include genres such as metal, screamo, and emo, which are indeed high energy but tend not to have very steady and danceable beats. Further, genres which have high valence values (blue, turquoise, and purple hues) seem to be concentrated in the horizontal center, showing that genres of high happiness tend not to be on either low or high extreme of danceability.

This implementation was less successful in distinguishing genres with similar dark colors, particularly in the blue and purple ranges. The cluster of deep purple bubbles in the center correspond to a range of genres from mariachi to soul, country to blues — genres that sound quite different in reality. The deep purple means that these genres have relatively low danceability and valence values, but due to the similar shades, it is difficult to make any more concrete or specific conclusions on how the genres compare to each other.

The RGB colormapping can be more successful depending on the variables selected. Figure 6 shows Red / Green / Blue as determined by danceability / instrumentalness / danceability, and position determined by energy. In this example, genres of high danceability are colored in shades of purple. Users can discern that genres of high danceability tend to be of medium energy, which is consistent with the previous 3-variable visualization, whereas genres of high instrumentalness tend to be either very low energy (focus music) or very high energy (metal).

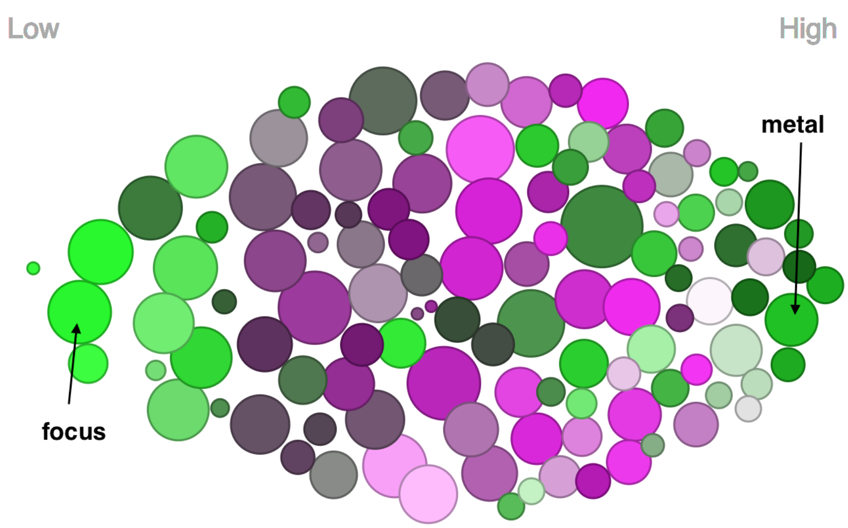

Figure 6: RGB colormapping with two variables

Figure 7 shows the bubble chart with color determined by mapping variables to HSV values. Danceability / energy / and valence are mapped to Hue / Saturation / Value, and position is again determined by danceability.

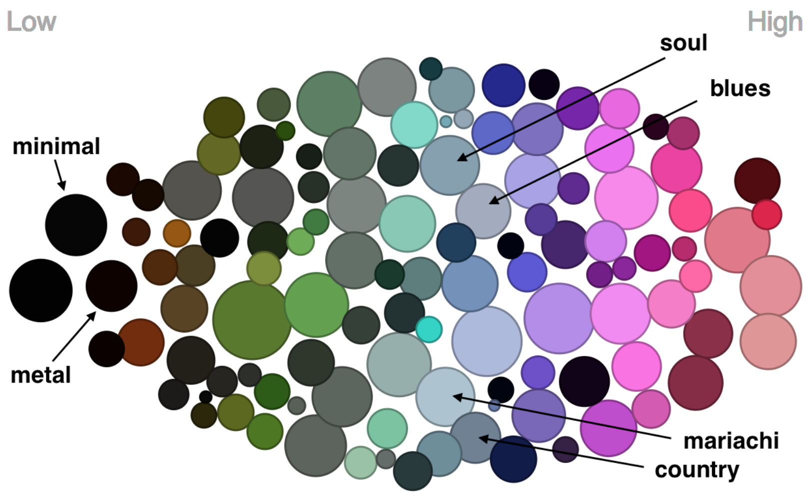

Figure 7: HSV color mapping.

The HSV implementation has different strengths and weaknesses in comparison to RGB. As noted in the implementation section, it is difficult to differentiate between genres with dark values and vastly different hues and saturations; the black bubble on the far left looks very similar to the dark red bubble on the far right, and the metal and minimal bubbles look identical because of their low danceability values, even though their energy and valence values are very different. In this implementation, however, genres of similar hues are easier to differentiate because of the way the color system is implemented. Here, even though blues and country are both light blue-purple, the blues bubble is noticeably lighter in shade and saturation than country, showing that it has a higher energy and valence while being similar in danceability.

4.1.2. Color as single variable, linear vs. nonlinear colormapping

The second implementation of the bubble chart allows the user to position genre bubbles based on one selected characteristic, and color based on another. Compared to the multivariate color approach, this implementation allows users to focus on the correlations between pairs of characteristics by comparing color and position.

Figures 8 and 9 show bubbles colored based on danceability values, with low danceability corresponding to yellow and high to blue, and position based on valence. In this comparison of danceability to valence, users more easily concluded that there was a positive correlation between danceability and valence in the visualization using nonlinear color mapping, since the nonlinear approach accentuates the middle values that appear murky purple in the linear mapping.

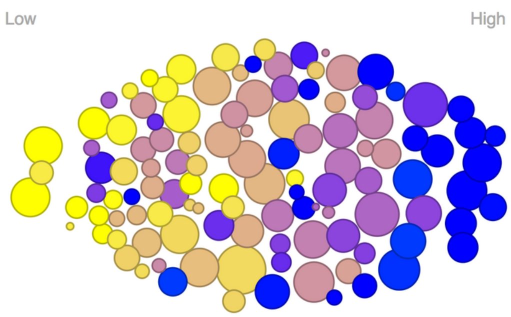

Figure 8: Danceability vs valence (linear).

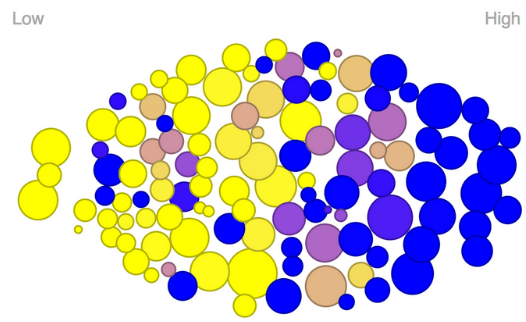

Figure 9: Danceability vs valence (nonlinear).

Figures 10 and 11 calculate color and position as danceability and energy, respectively. Users again found the nonlinear approach more effective, concluding that genres of high danceability are generally of medium energy, which is again consistent with the conclusions drawn from the 3-variable visualizations. The nonlinear approach also makes it easier to distinguish individual bubbles. The distinction between the rock and electronic, for example, is more pronounced in Figure 11 than in Figure 10

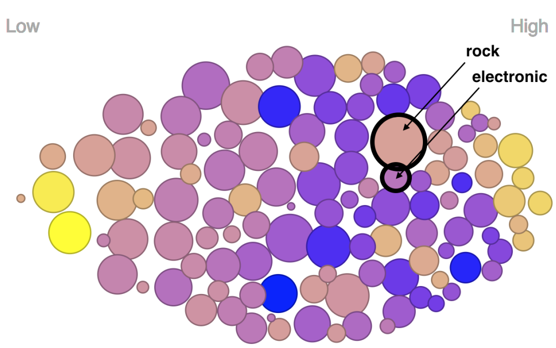

Figure 10: Danceability vs energy (linear).

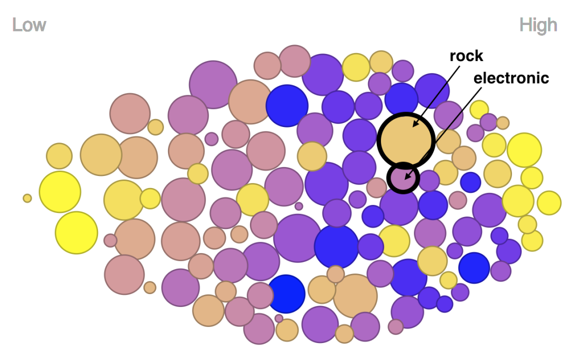

Figure 11: Danceability vs energy (nonlinear).

4.2. Motion graph

The motion graph addresses my second goal of showing changes in musical relationships over time by providing a visualization of the shifts in musical characteristics from 1960 to 2017. Users can select two characteristics that determine genre bubbles' positions for a chosen year.

Figure 12 positions bubbles with energy on the *x*-axis and valence on the *y*-axis. In 1960, high energy tracks were often associated with high valence as well, while less energetic tracks generally scored low on happiness. This highly positive correlation begins to break down by 1980, when there arose a handful of genres that were highly energetic but scored extremely low in valence, such as metal, screamo, and emo. These genres, which were built on the concept of controversy and counterculture, clearly break from the mold in their musical features as well [11]. The increasing number of genres that diverge from the direct energy-valence relationship can be seen especially well in 2000, when more genres arise with consistent valence around 0.5, but energy ranging from 0.4 to 0.7 — genres from experimental to worship to Chinese pop.

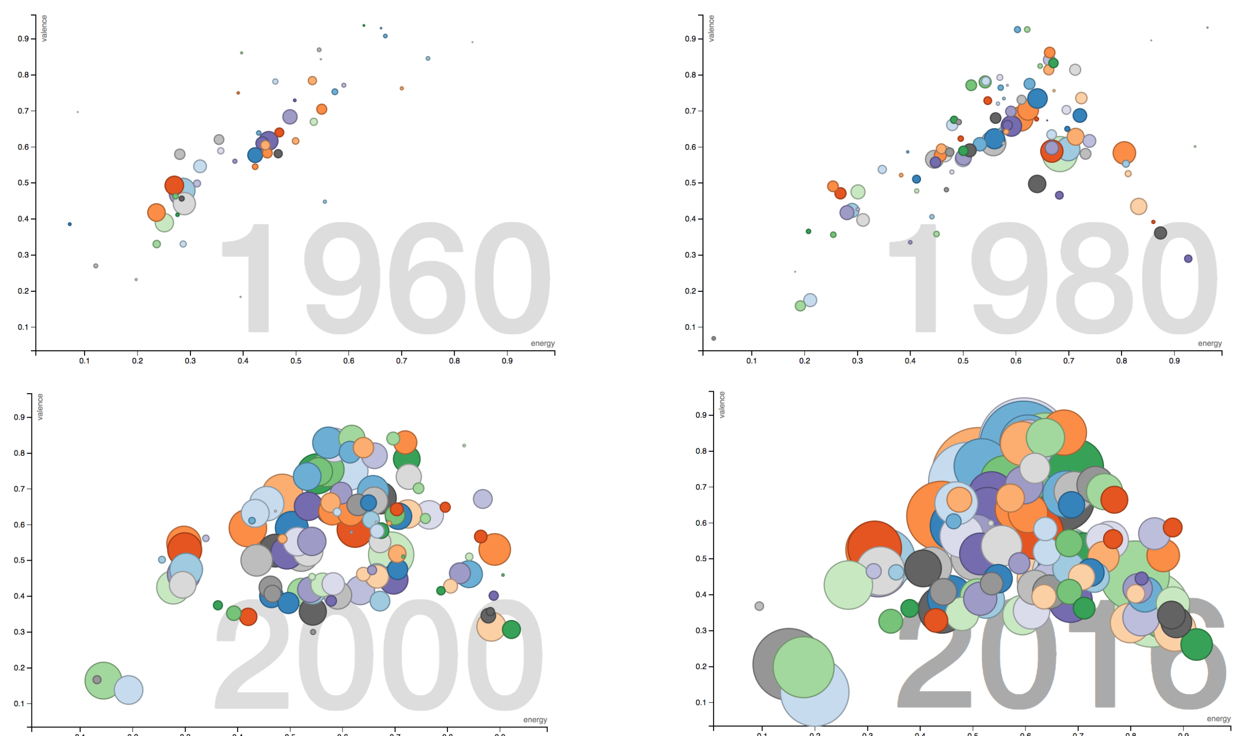

Figure 12: Energy vs. Valence in 1960, 1980, 2000, 2016.

While older genres such as blues, folk, rock, and country all stay consistently close to the central line from start to present, the same cannot be said for new genres (electronic is high energy and low valence), suggesting that music now is indeed becoming more diverse, even if just from an energy and happiness perspective. Some genres have significantly shifted over time, as well; experimental music, perhaps true to its name, has scored increasingly low in both energy and valence from the 80s to the present.

However, not all musical relationships have showed significant change. A view of energy versus tempo shows that the relationship between the two has steadied over time, with faster speeds signifying higher energy (Figure 13). As expected, classical performance, minimal, and focus music remain low energy as well as low speed, and metal and screamo top both the energy and tempo data.

Though tempo is not listed as one of the factors that directly influences Spotify's energy calculations, it may affect other metrics that are tied to energy. In the same way that major keys have come to signify happiness in Western culture, tempo may be one metric that musicians take for granted as having a direct connection to energy.

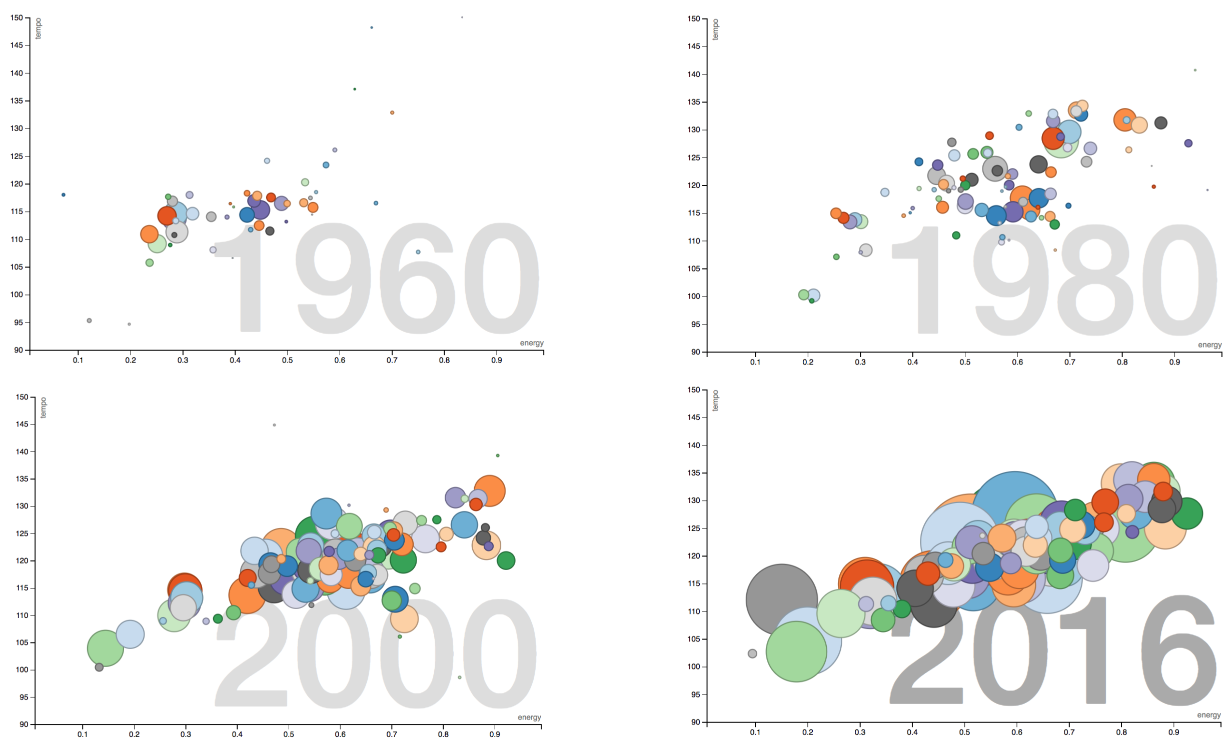

Figure 13: Energy vs. Tempo in 1960, 1980, 2000, 2016.

4.3. Modeling and classification

In addressing my final goal of modeling and classifying genre trends, I found that characteristics for a number of genres showed highly cyclic patterns, with predictable peaks and rebounds, while some genre characteristics remained relatively consistent, and others showed steady increases or decreases over time. Overall, there was high variation in the accuracy of the models for each genre characteristic.

Figures 14 and 15 show rock danceability and classical energy with polynomial regressions of degree 6 fit to 80% of their respective datasets.

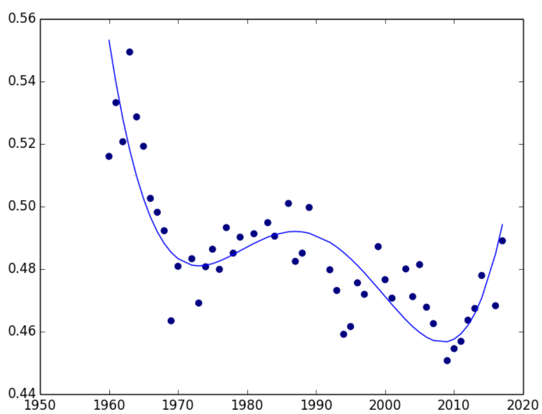

Figure 14: Rock danceability

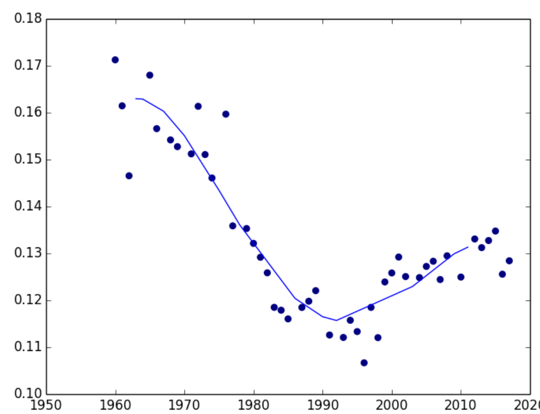

Figure 15: Classical energy

Rock danceability shows a clear nonlinear path, falling to a low in the early 1970s, rising somewhat in the 1980s, and falling again in the 2000s. These patterns seem to reflect rock history; the emergence of hard rock, progressive rock, and heavy metal in the 1970s could explain the decline in danceability, and rock's assimilation into the mainstream music industry throughout the 1980s could explain its increase in danceability, given the increasing popularity in dance-pop during the same time period [4]. Rock danceability has K-Fold cross-validation scores of 0.56 ± 0.31.

Classical music shows a dip in energy until the 1990s, and a steady rise since, which is interesting because Spotify's classical genre refers to recordings of existing Western Classical era pieces as well as new compositions. The changes in classical energy could reflect both the characteristics of newly composed pieces, as well as changing musical interpretations of existing classical pieces, or even shifts in popularity of lower- or higher-energy pieces. Classical energy has higher K-Fold cross-validation scores of 0.82 ± 0.17.

Using the validation points predicted by the regression models for both the original time range and for the next decade, I tested classification on a number of genre combinations, the results of which are displayed in Tables 3. The validation set predicted points for 20% of the original time range of 1960 - 2017, resulting in about 11 or 12 points per genre, so I chose the 12-year window of 2018 - 2029 to test the model's behavior on future characteristics.

Genres	**Accuracy**	
	1960 - 2017	2018 - 2029
Classical / Rock	1.0	1.0
Pop / Electronic	0.71	1.0
Pop / K-pop	0.59	0.46
Pop / Europop	0.77	0.88

Table 3: Classification accuracy.

Classification of genres that seem intuitively very different, such as classical and rock, is predictably high. Classification of pop and electronic is less accurate, as I had predicted with modern pop tracks infusing more electronic elements. A comparison of 'pop' music from different continents suggests that there is higher overlap between American pop and Korean pop music than American and Europop.

Classification of future points was still accurate for classical and rock, as expected because the two genres occupy very different musical spaces. Surprisingly, classification of pop and electronic became more accurate, possibly reflecting that the genres are headed in different musical directions. Classification of pop versus Korean pop dropped in accuracy to less than 50%, suggesting the musical styles of the two pop genres are converging, while Europop is diverging.

4.4. Additional tests for similarities between genres

Are the predicted pop characteristics in 2000 most similar to the actual characteristics of pop music in 2000, or are they closer to a different genre in a different year? To understand the similarities between predicted points and their actual values, I calculated which genre and year each predicted point was closest to by summing for each characteristic the differences between the predicted point and each actual value in my dataset (pop, rock, classical, etc. from 1960 to 2017).

Table 4 shows the resulting statistics for pop, electronic, classical, and rock music, each genre testing the approximately 11 or 12 points in the validation set. The genre accuracy is the percentage of the validation points that were categorized as the correct genres, and the year accuracy is the percentage of predictions that were within a decade of the real year. Pop scores predictably low in both genre and time accuracy, with most similarities to other types of pop music. Classical music is predictably close to Romantic music, since both refer to Western classical eras. It seems that in my years of listening to pop radio stations, I have actually been listening to worship music and salsa this entire time.

Genre	**Accuracy**		**Most similar genres**
	Genre	Year	
Pop	0.27	0.54	c-pop, cantopop, europop, worship, salsa
Electronic	0.80	0.70	indietronica
Classical	0.66	0.75	romantic
Rock	0.50	0.75	country, protopunk, freakbeat

Table 4: Predicted versus actual genre and year

5. Conclusion

In summary, this project has introduced a quantitative approach to analyzing how music genres have changed over time, providing an example of how musicologists can take advantage of a growing collection of digitized audio and technology for audio analysis to better understand and quantify musical trends in both specific genres and in popular music overall. This project provides two visualizations, a bubble chart and a motion graph, which allow users to view similarities between genres and correlations between characteristics. In viewing changing relationships over time, users can map musical changes to specific points in history, resulting in a more robust understanding of if and how traditional genres have changed and where new genres have emerged. Modeling of genre characteristics has illustrated significant changes in certain characteristics, which can be tentatively explained by historical events and the rise of subgenres. Classification of predicted values has shown that while certain genre pairs are easily distinguishable, pop music from around the world has similar features, and some genres are unexpectedly similar based their musical characteristics alone. This project has provided an initial view of what directions popular genres in the coming few years might move toward.

Though this project uses only data from popular songs available on Spotify, the initial results illustrate the potential of taking a quantitative approach to viewing musical change. Finding alternate sources of song data, particularly in areas like Asia where Spotify is less popular, could provide further evidence of musical patterns.

On the visualization front, color still has the potential to be more informative, particularly when paired with other factors, such as position. Experimenting with 3-dimensional visualizations to help users process more characteristics at once could increase the number of observations made using visualizations alone.

This project attempts to both trace musical patterns of the past and predict musical trends of the future. While the results have shown the possibility of quantifying past patterns, I would like to go further back by mapping them to historical events in even more detail. Experimenting with different regression models, particularly those for cyclic data, could possibly improve the accuracy of predictions of future trends. Since music is often shaped by its political and social environment and even a single iconic musician can drive a new musical movement, it is difficult to confidently predict what music will sound like twenty years down the line. Nevertheless, future work may have the potential to reasonably predict where pop music will head in the coming new year.

6. Acknowledgements

I would like to thank Professor Kernighan for his invaluable guidance throughout the semester, Meagan Wilson for her constant support and enthusiasm, Professor Finkelstein for his advice on colormapping, Jessica Shi for bouncing ideas around with me, and the many users who offered feedback through the process, without whom this project would have (literally) looked very different.

References

[1] "Cross validation," https://www.cs.cmu.edu/~schneide/tut5/node42.html.

[2] "Hsv (hue, saturation and value)," http://www.tech-faq.com/hsv.html.

[3] "A retromatic history of music (or love), 2017 - 1960," http://everynoise.com/retromatic.html.

[4] "Rock music timeline," http://www.rockmusictimeline.com/.

[5] M. Alves, "About perception and hue histograms in hsv space," https://www.slideshare.net/michelalves/about-perception-and-hue-histograms-in-hsv-space.

[6] M. Bostock, "The wealth & health of nations," https://bost.ocks.org/mike/nations/.

[7] L. Buitinck *et al.*, "API design for machine learning software: experiences from the scikit-learn project," in *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, 2013, pp. 108–122.

[8] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.

[9] edent, "motion-chart," https://github.com/edent/motion-chart.

[10] C. Elkan, "Nearest neighbor classification," *elkan@cs.ucsd.edu, January*, vol. 11, p. 3, 2011.

[11] T. Hjelm, K. Kahn-Harris, and M. Levine, "Heavy metal as controversy and counterculture," *Popular Music History*, vol. 6, no. 1, 2012. Available: https://journals.equinoxpub.com/index.php/PMH/article/view/14426/11937

[12] M. Mauch *et al.*, "The evolution of popular music: Usa 1960–2010," *Open Science*, vol. 2, no. 5, 2015. Available: http://rsos.royalsocietypublishing.org/content/2/5/150081

[13] P. H. Rodriguez Zivic, F. Shifres, and G. A. Cecchi, "Perceptual basis of evolving western musical styles," *Proceedings of the National Academy of Sciences*, vol. 110, no. 24, pp. 10 034–10 038, 2013. Available: http://www.pnas.org/content/110/24/10034.abstract

[14] J. Serrà *et al.*, "Measuring the evolution of contemporary western popular music," *Scientific Reports*, vol. 2, pp. 521 EP –, 07 2012. Available: http://dx.doi.org/10.1038/srep00521

[15] J. Vallandingham, "Creating animated bubble charts in d3," http://vallandingham.me/bubble_charts_in_d3.html.