Aidan Draper
CSC401: Final Project
11/28/2017

A Mood Analysis for Detecting Mental Health Conditions in Millennial Spotify Users

Question

Can genres, lyric moods, and lyric topics be analyzed on millennial Spotify users' saved songs to detect oddities, concern, or shifts, in their mental health over time?

Problem Statement

Mental health is a huge issue in millennials. According to the American Foundation for Suicide Prevention, suicide is the second leading cause of death for people aged 44 years and younger and also results in over 1.1 million deaths every year in the United States [10]. It is even more pressing because, many times, it requires self-awareness to find and address. That being said, Statista.com released a report in April of 2015 that stated that, of Spotify's population of over 75 million active listeners at the time, approximately 85% were under the age of 44 years old [11]. This means that if music could show signs of depression or emotional vulnerability through downwards mood trends over time or overall oddities in listening habits, it may show cause for concern and be an accurate predictor of mental health issues.

In a study in 2008, Felicia Baker and William Bor found that there were certain relationships between genres that adolescents listened to and their antisocial behaviors, vulnerability to drug use, and suicide rate. Two genres that showed suicide concern were rock/metal and non-aggressive rap [1]. Essentially, they found that adolescents with mental health had a very different listening habit than that of regular music listeners and were also more subject to higher rates of emotional vulnerability when listening to music [1]. Although their study was inconclusive as to whether the variety of genres were causing the mental health states, they were able to find a strong correlation between specific genres and people's emotionally vulnerability to suggest studying further [1]. Jung et al. also conducted a study in 2017 that analyzed moods from social network data. They were able to use sentiment analyses to conclude depression, and other mental illnesses or behavioral issues, in adolescents based on the text in their posts [4].

Similarly to the case studies, this experiment plans to use the lyrics in songs to conclude the motions being elicited to millennials through music. Jung et al.'s study showed that a sentiment analysis could be a proper tool to conclude deeper levels of sadness in text in a different environment [4]. The hope is that it will be effective for music as well. Nonetheless, various studies have found interesting correlations and patterns as well as a rather strong reason to suggest that music genre, tone, and topics can be a predictor of mental health issues. By using sentiment analysis, preprogrammed Spotify confidence levels, and time-series analyses, shifts in moods and oddities in listening habits can potentially be investigated in the lives of over 140 million people currently using the app to attempt to predict self-harm and suicidal concern [13]. Additionally, topic modeling can attempt to identify the topics being portrayed to listeners, which is important when some listeners might be more emotionally vulnerable to them.

Limitations

Unfortunately, Spofity's API only allows someone access to their own account's saved songs and data that is labeled public. For that reason, only my own account can be analyzed. That being said, if an app were to be programmed, it be implemented on a global scale because it could request access into a user's account, have them sign in, and output their scores and results for this specific test. However, because only my account will be analyzed, I will have to report my own moods during certain periods of time and there is some bias as well as subjectivity to self-diagnosing and reporting. Another limitation is that Spotify's algorithm's score, Valence, is strictly judging the mood of the overall song and nothing else. This algorithm is private, but it essentially rates a song as sad (0) to happy (1) based on a variety of variables from its audio analysis. Juan De Dios Santos conducted a study where the 'boringness' of two Spotify playlists were analyzed that used variables that were available through the Spotify API function 'audio_features'. The formula he developed was boringness = loudness + tempo + (energy*100) + (danceability*100) [7]. He seemed to have some success, however, I've analyzed the correlation between valence and some of these scores in Tableau in my CSC 111 class (as seen below in Figure 1) and they were fairly strong positive correlations, which gives reason to suggest that they were built on confounding variables. This means that there would be a lot of overlapping variables and subjectivity in creating my own formula with scores (that have no public explanation on how they were built) that were likely created with the same variables from the audio analysis data. For this reason, only the valence score will be analyzed.

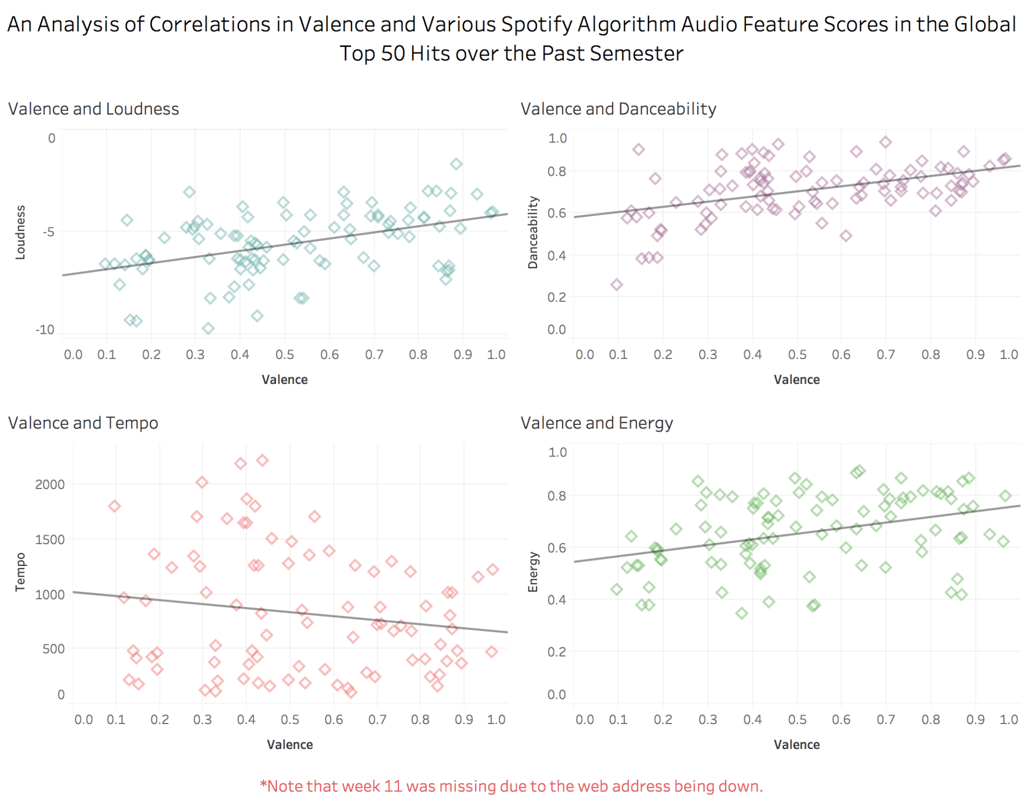

Figure 1. *An analysis of correlations between valence and four other Spotify audio feature variables.* There were strong correlations between valence and loudness, valence and danceability, and valence and energy. Source: Draper, A. *An analysis of mood found in the top 50 hits over the last semester.* Elon CSC 111. 2017.

Tools Used

To conduct this experiment, an open-sourced python library called Spotipy, that takes Spotify's Web API and converts it into an easy-to-use python library, will be used within a python script to collect listening habit data from my own saved song library. From there, song lyrics text from my Spotify's saved songs library will be scraped using a URL found with a search query on Genius's API using the Genius API pyton wrapper, py-genius. The full lyrics will be fed into Vader's sentiment analyzer separately to compare moods. Vader should be able to analyze the overall mood of the lyrics. Based on that, it can be said that the projected lyrics provoke more negative or positive moods in users. The lyrics will also be fed into an LDA model in Gensim to provide an analysis of the topics found in the music. This should show what topics are being elicited to the listener and what moods they also might bring. It is expected that the modeling will pull a lot of words from the chorus, which is believed to be the most influential part of a song. Because of the overall size of the data, only a couple topics will be analyzed, manually labeled and assigned positive or negative moods. Along with that, valence will be analyzed over time as well as average valence scores to see if there have been shifts in mood, or reason for concern. Artist music genre data will be grabbed for each song to develop a count for each genre, which will be used to develop a network analysis of the overall listening habits. Although the overall point is to analyze the count and distribution, it may also show how music genres tend to overlap for artists. The song data, along with their respective scores and lyrics, will be inserted into a MySQL database on phpMyAdmin and pulled into an Excel spreadsheet for the different analyses described later. Topic words will similarly be pulled into an Excel spreadsheet from PhpMyAdmin depending on their genre and then, they will manually be assigned topics. Based on the topic analysis, lyric mood, genre distribution, and valence scores, an attempt at making an overall classification of the mood of my music will be made.

Hypothesis

I have somewhat of an understanding of my music listening habits based on my mood and, for that, I believe that my music will correlate to times when I've been more happy or sad. I also have diagnosed anxiety and formerly, depression, so it will be interesting to see if that is noticeable in past years more than now. Overall, I believe that the four sources will be able to classify my music as more negative because I believe I am more emotionally vulnerable to lyrics in music and also have a history of mental illness. I believe that the most occurring genres will likely be hip hop/rap and alternative pop. According to a study conducted by Miranda and Claes in 2008, non-violent rap music elicits higher depression scores than even heavy metal [5]. Rap songs also resulted in higher rates of angry responses than any other music genre [5]. For that reason, I believe my overall listening habits will likely be more sad and angry.

Data Collection and Manipulation

To collect the initial saved song and genre data for the study, a python script was built using the libraries spotipy and py-genius. The spotipy python wrapper pulls data, including: song name, song id, artist, artist id, popularity and date added, from the Spotify API when the saved tracks are grabbed. From there, the song id is used to look up the valence score, under audio features. After that, the lyrics are grabbed from the respective genius.com links for each song using a search query within the python program that includes the song name and artist name variables. Of the 379 songs in my saved library, 82 were not found. Therefore, only 78% of the song data could be modeled using the data mining approaches. These eight variables were then inserted into phpMyAdmin in the table, spotify_saved_songs. Finally, with the artistId for each song, the function 'artist' was used to search for a specific artist's genres for each song. Each genre was added to the table, spotify_saved_genres. Unfortunately, Spotify assigns genres just to artists so there are repeats for genres in the table. For that reason, song_id was also added to spotify_saved_genres so that total genre distribution and genre distribution by artist can still be analyzed on their own if need be.

In this study, the analysis is on Spotify listening habits so the Spotify API is the best data source to use, as it is the primary source rather than a secondary source. In Phase 1, MusixMatch was used to find song lyrics. However, due to licensing issues, only the first 30% of the lyrics could be grabbed. Additionally, the sales team and developer team never replied to emails inquiring about grabbing the full lyrics. For this reason alone, the Genius API and the Genius API python wrapper, py-genius, were used to grab song URLs and text scrape with BeautifulSoup. MusixMatch was definitely the better source. Of the 379 saved songs in my library, it found lyrics for all but 24 songs. On the other hand, Genius was not able to grab lyrics for 82 songs. This meant that only approximately 78% of the song's could be analyzed with topic modeling and sentiment analyses.

Once lyrics had been fully collected, the vader scores and topics were found using the scores() method in draper_finalproject.py. Initially, it was anticipated that the lyrics would need cleaning, but they were fairly perfect in the end because of the way they were grabbed off of the site's html. The only two ways in which the data was cleaned was removing any instances of hard brackets that identified a verse or chorus, such as '[Verse 1]', and getting rid of blank lines ('\n') so that the lyrics could be separated line-by-line and analyzed using a common deliminator. That being said, some used verse and chorus identifiers used curly braces so they could not be cleaned without affecting the rest of the lyrics.

To visualize the data in Tableau, the information had to be grabbed from phpMyAdmin and inserted into Excel so that the data was local. Unfortunately, Tableau does not have the ability to connect to phpMyAdmin and, in the future, a local SQL source like MySQL Workbench would provide an easier way to connect the data to Tableau. In the excel file, called draper_visualization_data.xlsx, each sheet represents data collected from numerous queries found in the Appendix of Queries section, which can be found labeled in queries.txt file as well. Queries 1 and 2 were used to collect data for the 'scorebygenre' sheet, query 3 was used to grab data for the 'songs' sheet, and query 4 was used to collect genre data about the population in the 'genre' sheet. There was one calculated column in the 'genre' sheet that found the percentage of the overall population that each genre represented. It is important to note that some artists had no genre while others had multiple though. The 'scorebygenre' sheet had 3 columns that required data manipulation. Using the range of genres found in cell N2, 50 genres, that represent 75% of the total artists' genres found in the saved songs were grabbed and inserted into column J. In column K, a VLOOKUP function was used to grab the respective Valence scores in row B for each of the 50 genres. Similarly, in column L, a VLOOKUP function was used to grab the Vader scores in column G for the songs where lyrics were found. It is also important to note that the valence average is for all 379 songs while the Vader average score was only for 297 songs where lyrics were found so the populations are slightly different due to constraints.

Data Dictionary

Table: **adraper2.spotify_saved_songs**

- **saved_order:** an auto increment primary key that also represents the order in which the songs were saved
- **song:** the name of the saved song
- **song_id:** Spotify's unique key for the song (useful for searches)
- **artist:** the name of the primary artist of the song (artists[0])
- **artist_id:** Spotify's unique key for the specific artist (also useful for searches)
- **lyrics:** the text for the song from scraped from genius.com after grabbing the song's address from a search query (its NULL if the song wasn't found)
- **valence:** Spotify's audio feature algorithm score for sad to happy, from 0 to 1, for a specific song
- **popularity:** Spotify's algorithm score for a song's popularity based on recent user streaming
- **vader_lyrics:** Vader's specific score for the lyrics data on the mood, negative to positive, from -1 to 1
- **date_added:** the date that the song was saved to a user's library in MySQL DATE format

Table: **adraper2.spotify_saved_topics**

- **id:** a synthetic key for topic count (total = saved songs * 3)
- **song_id:** the song id for the lyrics being analyzed
- **topic:** a manually assigned topic based on the five words in the next columns
- **word1:** the first word from the specific topic in the song's lyrics
- **word2:** the second word from the specific topic in the song's lyrics
- **word3:** the third word from the specific topic in the song's lyrics
- **word4:** the fourth word from the specific topic in the song's lyrics
- **word5:** the fifth word from the specific topic in the song's lyrics
- **word6:** the sixth word from the specific topic in the song's lyrics
- **word7:** the seventh word from the specific topic in the song's lyrics
- **vader_words:** Vader's sentiment score for the five words from -1 to 1 for a negative or positive mood

Table: **adraper2.spotify_saved_genres**

- id: the synethic key for the row of the genre
- song_id: the respective song id for the artist's genre
- genre: Spotify's assigned genre for the specific artist

Entity Relationship Diagram

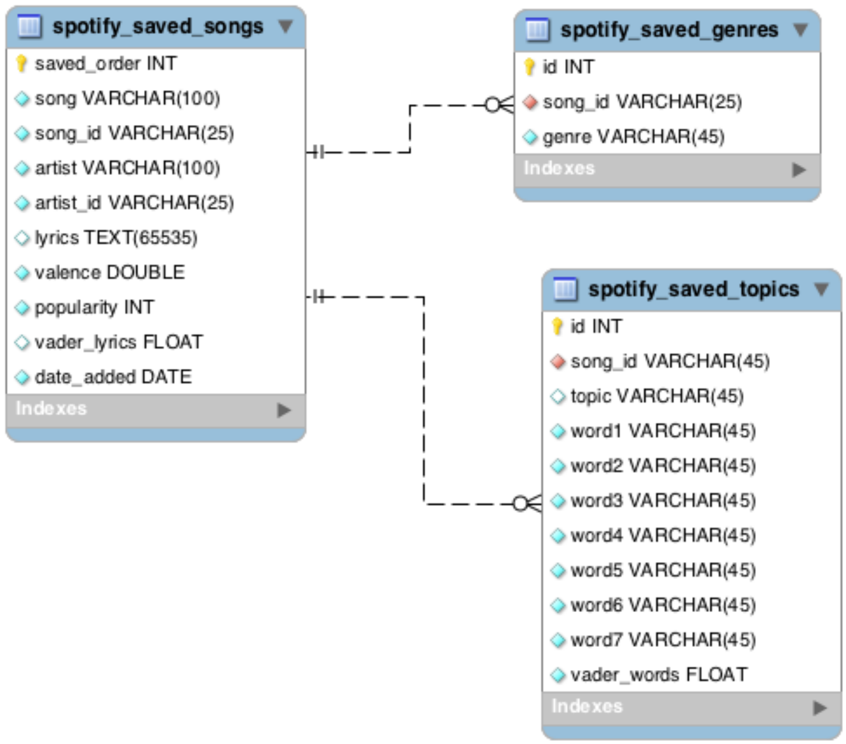

Figure 2. *The Entity Relationship Diagram for the database hierarchy used to store information for this study.* There are a total of three tables. The database is not normalized. However, it could be if song_id became the primary key in spotify_saved_songs.

**Note that the diagram has been slightly altered since phase 1. The one-to-many relationship from spotify_saved_songs to spotify_saved_topics is no longer mandatory because not all of the lyrics were found. Vader variables were changed to floats to store all of the decimals and the number of saved topic words was changed from 5 to 7 to make topic identification easier.*

Both, spotify_saved_genres and spotify_saved_topics, connect to spotify_saved_songs from the column song_id, which is unique to each saved song. They both also have a one-to-many relationship with the table. However, oddly enough, some Spotify artists do not have a listed genre so it is not mandatory for a song_id to be in the saved genres table. There will be three song id matches for every song in the topics table because there should be three topics found for each song. As of right now, the plan is to only collect five words for each topic. Giving them their own unique column seemed fine because they are also ranked in order of relevance to the topic. It also makes assigning topics much easier later down the road. All of the tables have synthetic keys, but the saved songs table actually has meaning, as it also relates to the song's order added.

Data Mining Methodology

As previously mentioned, this study is implementing two data mining algorithms, sentiment analysis and topic modeling, in an attempt to classify my Spotify listening habits in my saved songs library. The two methods are explained in detail below.

A sentiment analysis is basically a judgment of mood, as positive or negative, for text given. Depending on the specific tool used, it might look at unigrams, bigrams and even trigrams to score text as negative and positive, as well as the strength of emotion in the overall text [2]. For example, using Vader in the NLTK python package, the word ‘depressed’ would likely have a score around -1 where as the word ‘happy’ would have much higher score around 1. If they were in a sentence together, like “I felt as if I was becoming depressed, but the sun came out and now I’m feeling happy”, would likely receive a neutral score around 0. Bigrams and trigrams are important because they can detect decrements, increments, and negations of words, like “not bad” or “sort of happy”, which should not be scored as extreme [2]. With all of these word scores, it develops positive, negative, neutral, and compound scores based on a lexicon, or an untrained dictionary of sentiments for the specific text, that the program developed during runtime [2]. In this study, only compound scores were grabbed because the average score seemed to be what would show the best overall mood for each song.

For a discovered topic unknown to the user, topic modeling attempts to pull out a certain number of words based on their probability of relating to the topic. In the case for the Latent Dirichlet Allocation (LDA) model used, the algorithm will grab words from a created bag-of-words and group them around what trained topic they most likely represent [12]. The bag-of-words is essentially an unorganized collection of words from a corpora, or collection of texts or documents, where their order in a sentence is ignored and focus is just put just on frequency and known meaning [12]. If a set of 5 sentences is fed into the program, it will run through them and spit out which sentences describe a certain topic. For example, if two of the sentences were about playing in a lacrosse game and then, a soccer game on TV, but the rest were about breakfast, lunch and dinner, the topic model would spit out two sets of words. They might be [‘goal’, ‘Ronaldo’, ‘attack’, ‘dodge’, ‘scored’, ‘stick’] and [‘eggs’, ‘plate’, ‘Italian’, ‘cook’, ‘chicken’, ‘sandwich’], which would likely mean the topics are sports and food. Unlike Vader, the LDA model is an unsupervised training approach based on the number of ‘passes’, or times, the user tells it to run through the text. It uses about 80% of them to train the set and then, it backtracks to find the most prevalent topics in the text with the remaining 20% [12]. The words are actually scored and organized by their relevance to the topic, but that will not be important in this study. In this specifc study, 7 topic words were grabbed, along with their respective song names and topic vader scores, by genre and then manually assigned topics in an excel workbook, named draper_topic_data.xlsx.

Both of these models are relevant to the study because they can describe moods and topics portrayed in music that might be absorbed by the user. If someone exhibits listening habits that have an overall negative mood and have genre distribution that correlates with negative emotions, then it can be said that they are more likely to be emotionally vulnerable to their music, and may show signs of mental illness. When analyzing the valence scores, shifts in listening habits are also important because they can catch negative trends that may show signs of future self-harm. Using topic modeling, the actually ideas that the user might be subjected to during listening can be picked out. This can be important in identifying depression, catching suicide or self-harm related topics, or even violent acts that are projected in some songs.

Additional models, including text summarization, named entity recognition, and a social network analysis, were considered when conducting this study. However, for the various reasons described below, they were not included in this study. That being said, two models could be useful in future studies in this field. For starters, text summarization would likely not be as effective for this study because it has a higher chance of just pulling out the chorus, which likely contains the most common sentences due to the repetitiveness. Because topics are the ideas the user would be subjected to, topic modeling was much more effective. Named entity recognition was also considered because it can grab all sorts of nouns from the lyric data. This was one of the models that should be considered for future studies because it can grab a variety of proper nouns, from locations, people, and organizations, which could be important for analyzing topics, lead to more analysis into what topics are about on a large scale, and also identifying positive or negative words that were missed from the topic modeling. That being said, for this study, topic modeling was more important for the analysis of overall moods in songs. A social network analysis was considered to see if there were groupings of topics or genres that could identify where the more negative moods were being projected. That being said, a scatterplot depicted in the results and discussion section ended up compensating for the genre distribution idea. However, if models were trained in the future to show mentally ill networks of users versus users who did not show signs of concern, users could be compared to these two models to see how they fit to both models in order to predict concern in overall listening habits.

Visualizations

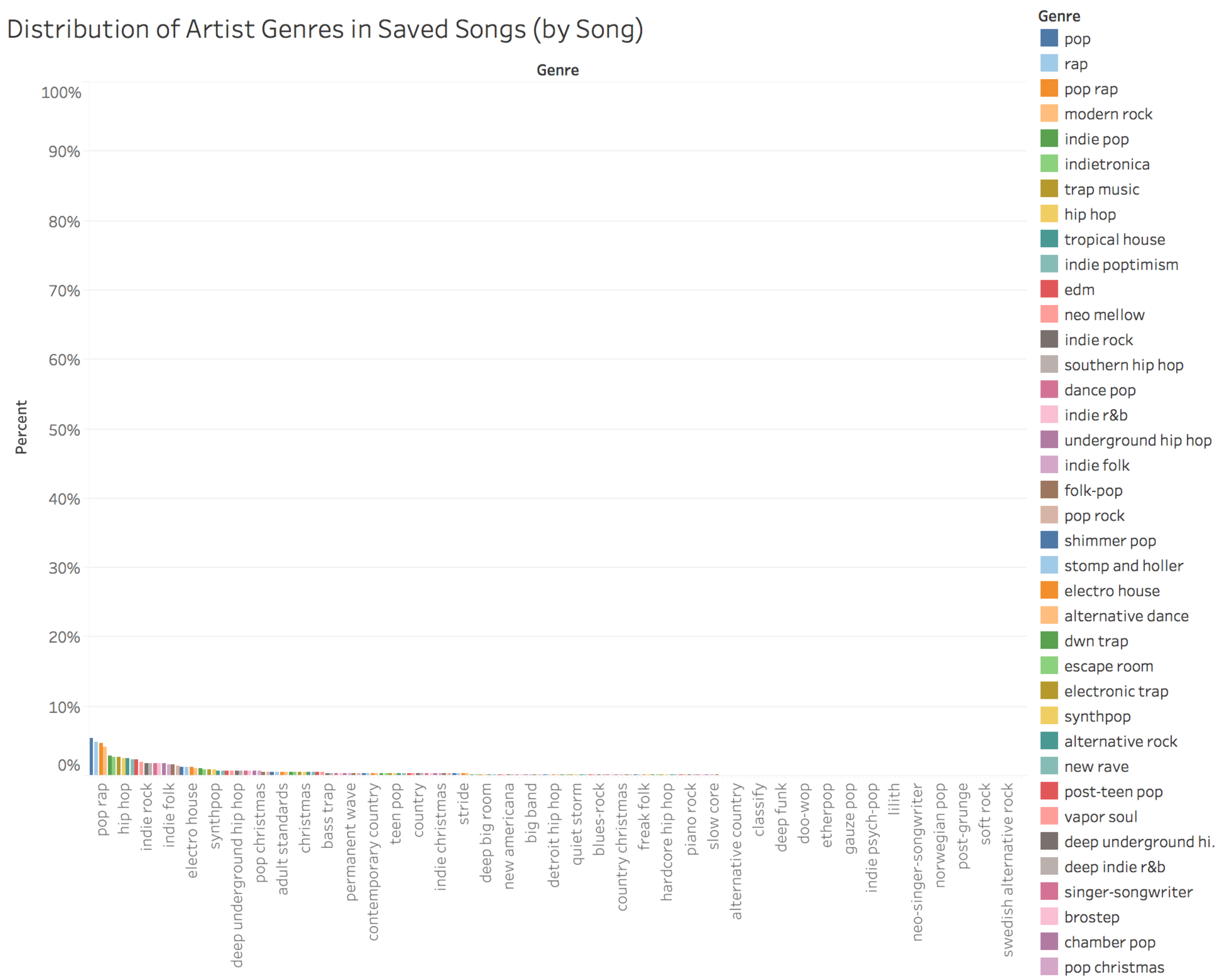

Figure 3. *The distribution of all of the artist genres available (with repeats) in the saved songs library.* This graph is very sparse, which makes it somewhat hard to read. Because it is a distribution graph, it technically has to be out of 100% on the y-axis or the results may be misinterpreted. However, it still successfully shows that the artists of the songs have a variety of genres. Along with that, the genre legend to the right is organized by greatest to least. From this, it can be seen that pop had the most, then rap, then pop rap and son on while (not pictured) Swedish alternative rock had the least.

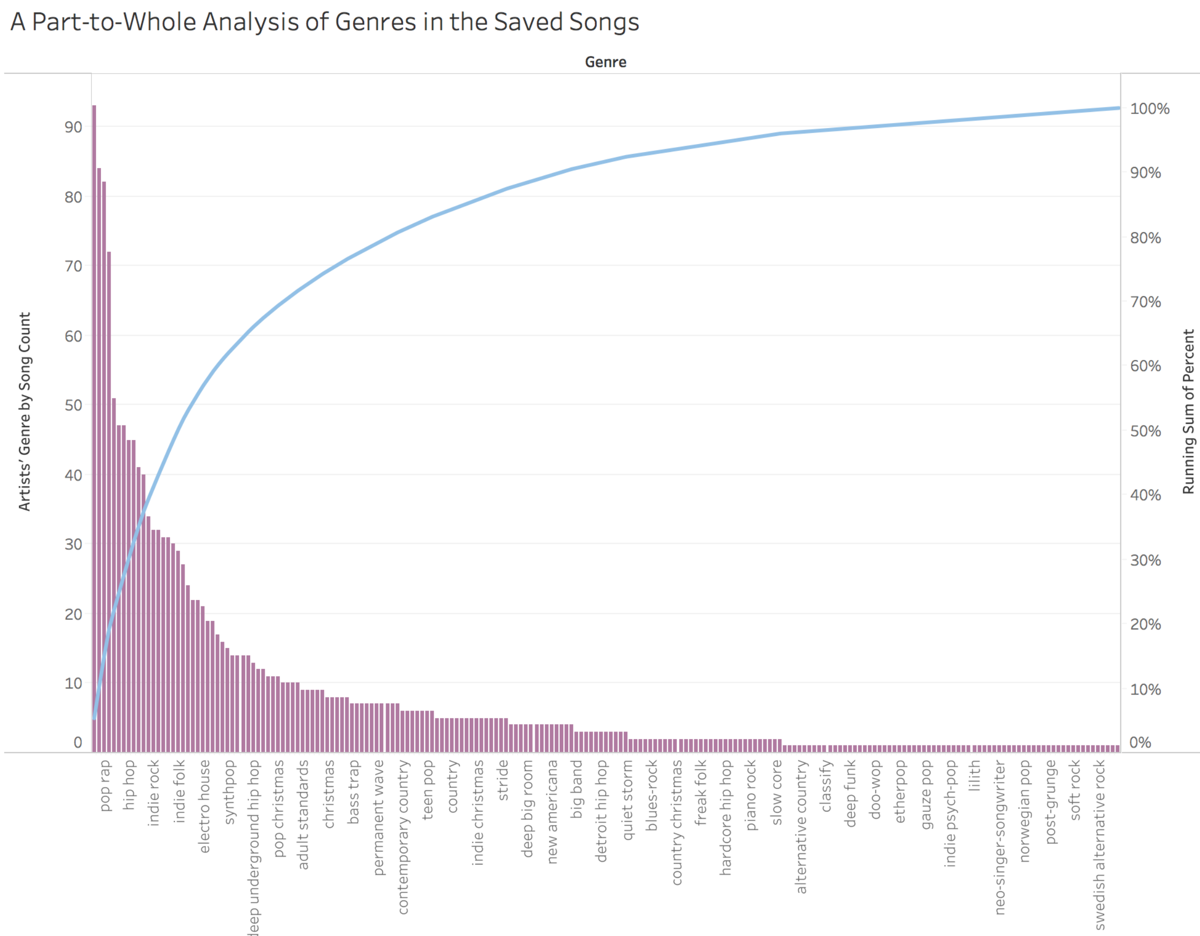

Figure 4. *A ranked part-to-whole pareto chart of the artist genres by song count.* The x-axis shows every third instance of the genres found in the music while the left y-axis shows every count for each genre by song. The purple bars are the count and the running sum is the blue line. This shows how the total population is distributed across the genres. Most of the data, almost 80%, falls within the first 25% of the genres on the x-axis.

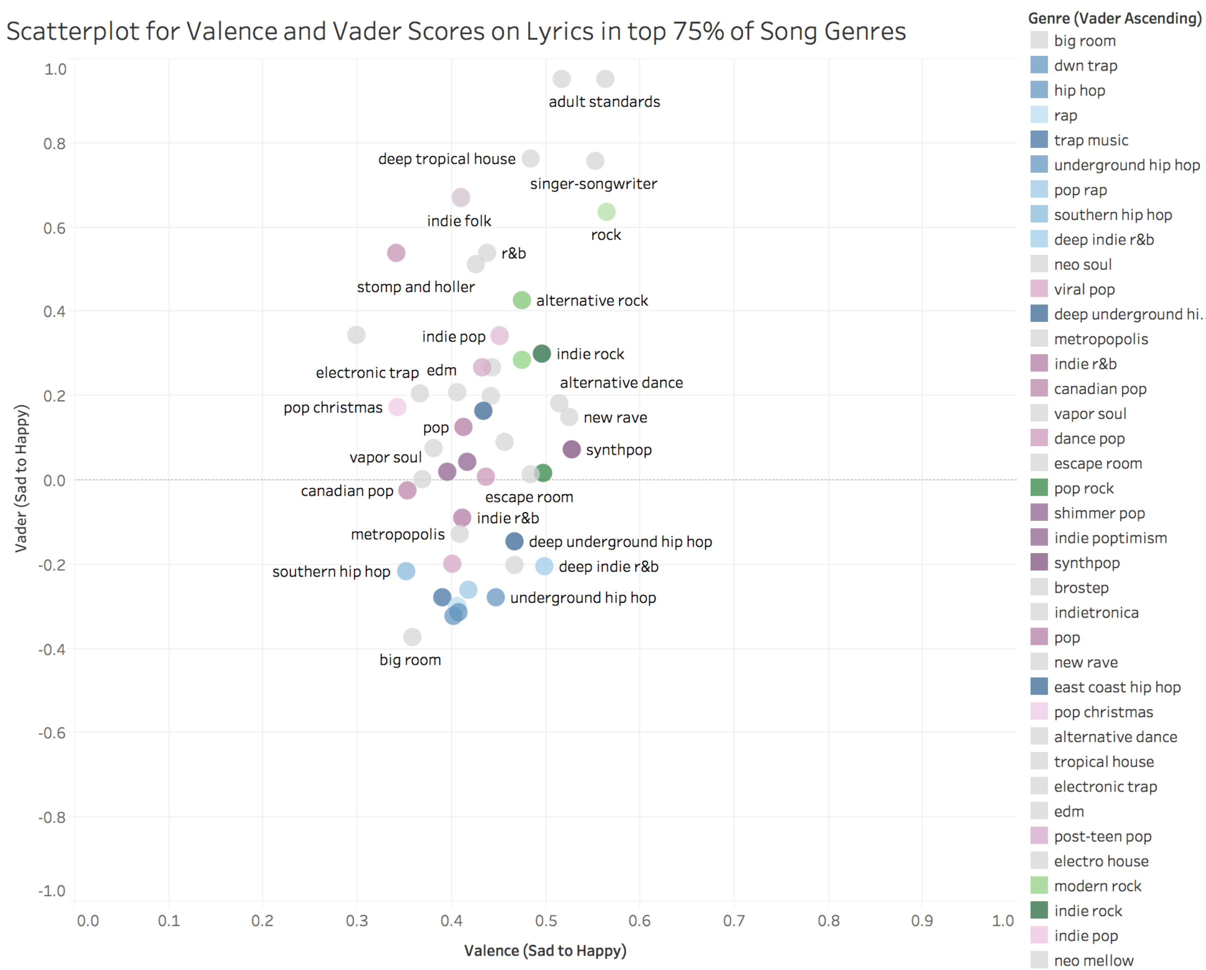

Figure 5. *A scatterplot showing the correlation between Spotify's Valence score on the x-axis and NLTK's Vader score of the song lyrics on the y-axis.* It is important to note that the valence average is based on the entire population of 379 songs while the Vader score is based on the 297 songs where lyrics were found so the results might be slightly skewed. There is no correlation between the two values. The valence score was very clustered around neutral sentiments .3 to .6 while the Vader score had more extremes from -.4 to 1.0. The dots are also color coordinated by genre. Rap and hip hop genres are represented by blue, rock genres are represented by green and pop genres are represented by purple. The legend is organized from negative to positive vader scores.

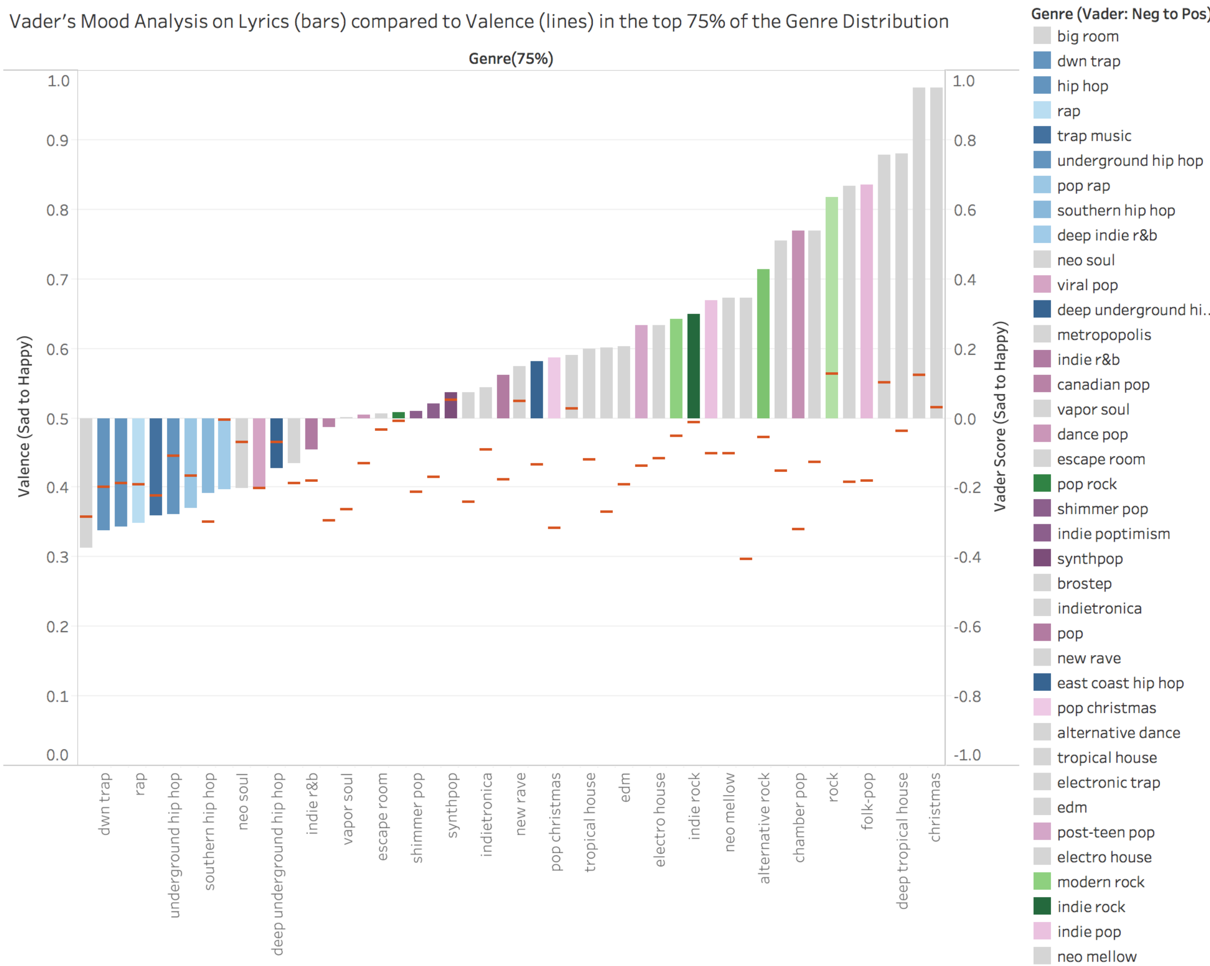

Figure 6. *A ranked comparison of Vader scores on lyrics and Spotify's valence score.* The bars represent the NLTK Vader scores while the small dashes represent the Valence scores. From this diagram, it can be said that there is little correlation between the two variables across each genres, with the exception of a few outliers. The genres were also categorized by color in the same way as figure 5. Rap and hip hop are blue, green represents rock categories, and pop labels are purple. Most of the rap genres fell in the negative range for the Vader lyrics, with the exception of East Coast Rap. Pop, on the other hand, was sparse across the scoring. It fell between -.2 and .65. Rock was only positive and fell between .3 and .63. Refer to the right y-axis for Vader scores and the left y-axis for Valence scores. Both are labeled from sad to happy.

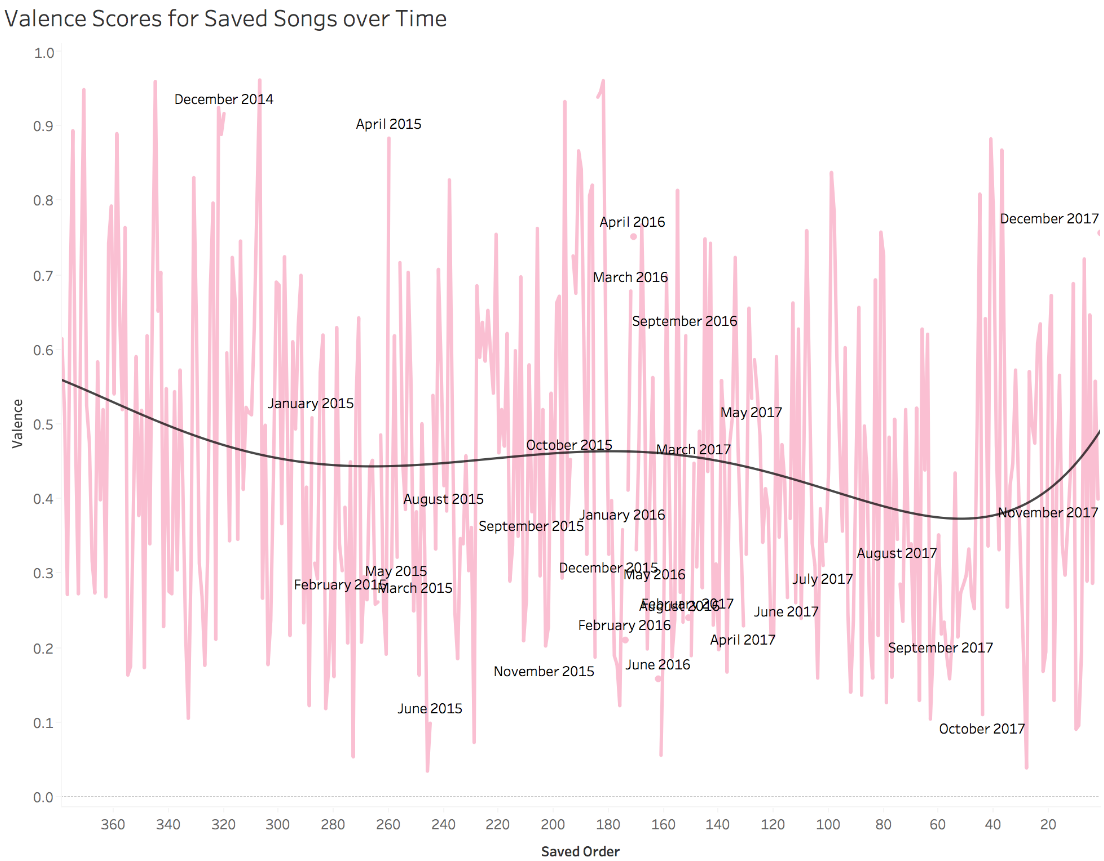

Figure 7. *Valence score from sad to happy over time by order saved in the user's saved songs Spotify library.* The pink lines are the valence scores for each of the saved songs. The black trend line is a five-variable (to the fifth degree), parabolic function to fit the analysis over time. The data is organized in order of the song being saved to a library, but the dates are also labeled by month and year for every gap. There are two troughs around March to May of 2015 and August to October 2017. There is also one peak around March to April of 2016. Recent trends show a very positive trend as well. The dates are slightly skewed because the songs are often saved many times at once over the course of the year so no songs may be saved at all for certain period of time.

Results and Discussion

The genre distribution can be seen in Figure 3 above. This graph is slightly hard to read because of how little each genre represents of the total population of genres found, but it shows a characteristic and a variable important for this study. One characteristic of my listening habits that can be observed was that the genre distribution was very sparse. This is important for identifying my listening habits because it shows that I do not listen to just one or two genres. This reflects positively on my overall mood because it means I am less likely to put all my eggs in one basket and be subjected to just one genre, which can either be very positive or very negative, as mentioned in Miranda and Claes's study. However, what it does show is that my most common genres were pop, rap, pop rap and modern rock, which can be perceived as a concern because, according to Miranda and Claes's study, rock and rap can show signs of depression and self-harm depending on the topics portrayed. Luckily, pop correlates with more positive moods, which slightly balances out the moods commonly projected from the other two.

Because of the sparseness of the data in Figure 3, it seemed necessary to create a part-to-whole analysis to see if it was evenly distributed across a bunch of genres or if most of the songs still fell within a few genres. Figure 4 shows this part-to-whole analysis as a pareto chart. The blue line represents the running average of percentage that is represented by the genres. Without even traveling across ¼ of the genres, more than 75% of the genre distribution can be collected. This successfully showed that most of the genres were found in a just a few genres. As mentioned in the data collection section, cell N2 was used to calculate the range need for 75% of that data. 50 genres were found, ranked by highest percentage of the total count of song genres found.

These genres were then analyzed to see if the average vader score for the 297 lyrics found correlated with the entire populations of songs Spotify valence scores. In retrospect, it may have been more correlated if only the 297 songs were analyzed for both groups, but, by including the entire population, the valence score is a better representation for the overall listening habits. The graph shows that the valence scores were fairly clustered around neutral ratings, but the vader lyrics scores showed much more variability across genres. By color-coordinating the sub-genres under the larger genre categories rap and hip hop, rock and pop, initial observations into trends within these genres could be made. Rap and hip hop was the only overarching genre that was super clustered. It actually correlated with Miranda and Claes's later work in 2012 too. Every sub-genre's lyrics scored negatively, with the exception of east coast rap, and it was by far the most negative genre. This is important because the rap and hip hop genre was said to convey feelings of anger, but also, when the lyrics were non-aggressive, feelings of depression and even self-harm [6]. Using topic modeling, an attempt at categorizing where these saved songs fall within the two moods, angry and sad, can be made. This will hopefully reveal whether there is concern for self-harm or aggressive behavior within my listening habits.

Figure 6 also shows the correlations between the Vader and valence scores. However, it does so in a different way that makes it easier to see how the average Vader scores lined up within the three overarching genres, rap and hip hop, pop and rock. As figure 5 also showed, the variables are far from correlated. However, from this graph, it can be seen that the average valence score falls around 0.4, which is slightly negative. This signifies that my overall listening habits across all genres are somewhat negative. The lyrics on the other hand, suggest that the issue is within the rap and hip hop genre. Unlike Miranda and Claes's studies, rock scored only positively and pop scored both positively and negatively. That being said, they included metal within the rock category, which I rarely listen to, and that may be why the rock data is skewed left. Because of this graph, it can be said that the most concerning genre lyrics is rap and hip hop. Although rock and pop are fairly positive, they will still be analyzed and shown below for a comparison in the topic modeling figures.

The topic modeling tables are listed from 1 to 3. For the following tables, only the first resultant topic was grabbed for each song. The tables were separated by the genres: rap and hip hop, pop and rock. The topics are ordered by their cumulative Vader score from least to greatest in order to compare the range of most negative to most positive topics for each genre. The subgenre is just the first genre available for each artist so the topics are not necessarily definitive, or unique, to each subgenre. It was decided that the best way to analyze specific topics was by separating topics by song. There is far too much symbolism in music for the words to necessarily imply the same thing and artists reference many of the same topics differently. It might be interesting to group genres as a whole in a future study though. Anyway, because Baker and Bor found higher risk in nonaggressive rap songs than aggressive ones, rap and hip hop songs were also separated by aggressive and non-aggressive. Query 6 was used to grab 105 rap and hip hop genre topics, Query 7 was used to grab 90 rock genre topics, and Query 8 was used to grab 159 pop genre topics. Initially, the plan was to assign topics for every result grabbed. However, due to symbolism and deeper meanings in lyrics, it was very difficult to identify many of the topics. In order to avoid identifying many incorrectly, I chose 10 from each genre, including aggressive and non-aggressive rap and hip hop as their own genres. They were then assigned topics based on their words and relevance in their respective song lyrics and meanings. The most interesting 10 for each genre are shown below.

The Rock Genre

Song	Subgenre	Vader	Topic	Words
We Come Ru…	indie rock..	-0.802	Optimism - future	soon, dawn, echoes, dead, gold, face, dirty
Mr. Brightside	alt rock	-0.743	Suspicious of lies	calling, sick, alibis, choking, swimming, …
Entertainment	indie rock	-0.4588	Symbolism of ex	low, want, refused, volume, turn, loud, heard
Hunger of the…	indie rock	-0.25	Pain of a breakup	pine, hunger, heart, female, armour, wears, …
Fred Astaire	alt rock	-0.1531	Still loves his ex	breakfast, baby, lately, anymore, miss, think…
Pumpin Blood	alt rock	0.0258	Excitement	alive, heart, feeling, best, worst, wrong, pumpin
Girlfriend	indie rock	0.4404	A failing relation	miracle, tears, whispering, fortune, bought, …
Last Nite	alt rock	0.6003	Ignoring a GF	turned, walking, care, little, alright, door, miles
Gravity	pop rock	0.6369	Suicide, but why?	light, come, gravity, working, heart, love, …
San Francisco	indie rock	0.8885	Power of love	love, feel, beats, systematic, heart, smiling, …

Table 1. *Noteworthy topics for rock songs found from query 7.* There are a total of 10 songs and 5 fields for each row. The rows are organized first by the sum Vader score for the 7 respective topic words. Some names and topic words were cut short due to length. The songs were selected from the 90 rows that were returned from query 7 and exported into draper_topic_data.xlsx under the sheet, 'rock'.

The Pop Genre

Song	Subgenre	Vader	Topic	Words
Needed Me	dance pop	-0.7976	Hard to date her	shit, little, fix, bitch, inner, trynna, issues
Daddy Issues	indie-pop	-0.7294	Parental issues	thinking, girl, little, matters, ask, forget, dead
Dirty Paws	folk-pop	-0.4404	Fantasy	Dirty, paws, story, dragonfly, furry, birds, beast
Greek Tragedy	indie-pop	-0.4215	Stuck in love	ecstasy, hits, hate, feeling, falling, tired, steps
Riptide	folk-pop	0	Entranced	dark, taken, away, come, unstuck, pfeiffer, closest
Lisztomania	indie pop	0.1027	stuck in the past	time, fortunate, lonely, end, burn, instead, pict…
Compass	gauze pop	0.4019	Love guides us	love, east, calling, way, needed, fallen, leaves
Different Skies	indie pop	0.7579	Jealous of ex	aisle, walking, good, came, heartache, nostalgia...
Skinny Love	folk-pop	0.8555	One-sided relation	told, love, come, skinny, fine, patient, kind
Hooked on a F..	pop	0.8779	entranced in lips	love, girl, good, turn, candy, lips, sweet

Table 2. *Noteworthy topics for pop songs found from query 8.* There are a total of 10 songs and 5 fields for each row. The rows are organized first by the sum Vader score for the 7 respective topic words. Some names and topic words were cut short due to length. The songs were selected from the 159 rows that were returned from query 8 and exported into draper_topic_data.xlsx under the sheet, 'pop'.

The rock genre songs in Table 1 did not seem very reflective of rock music at all. It was heavily weighed down by alternative and indie rock, which are far different from the categories from what Baker and Bor were analyzing. Unfortunately, it seems that I did not listen to enough heavy, metal, or classic rock to analyze. Instead, songs like Respect by Aretha Franklin, Ordinary Human by One Republic and Wagon Wheel by Darius Rucker began showing up. This is due to Spotify's artist tags though. These artists might have published rock songs before, but these songs were certainly not that genre. Interestingly enough, the word 'heart' appeared in the table three times and 'love' appeared twice. Most topics were about women, exes, girlfriends, and love in general. Gravity by John Mayer is a concerning exception. The lyrics question how the author could even contemplate suicide and have these dark thoughts given the loving relationship he is in. This song, along with a few others, like Hunger of the Pine by 'Portugal. the Man', have underlying meanings that are encompassed in sadness and depression, which was very unexpected. As seen in Figure 5 and 6, rock placed much higher on the Valence and Vader scores than the other two categories, yet the song topics are very negative.

The pop songs in Table 2 strictly revolved around relationships, with the exception of Dirty Paws, which was a fantasy that referenced events in World War II. Unfortunately, it seems that, for both tables, the topic words did not reflect whether the topic was positive or negative. Skinny love had one of the saddest topics where the author is worried that they will wake up and their relationship with their partner will just be over all of a sudden yet the topic's Vader score was 0.8555. It's hard to detect relevant topics in these lyrics though because they are usually not straightforward. With the exception of Skinny Love, the pop songs were actually fairly positive in topics, which differed from rock. They definitely do not show any reason for concern. Most are love songs or where the author is aware of, and reflecting on, a feeling their partner has, such as in Daddy Issues by The Neighborhood and and Compass by Zella Day. The lead singer in The Neighborhood empathizes with his partner because they both had issues with their fathers. In Compass, the author is analyzing why her partner is obsessed with her. Nonetheless, these topics are less sad overall and should not be worried about as much as some of the other genres.

Before the next table is introduced, it is important to note that Vader seemed to skew some of the topics based on cursing, which does not always reflect negative moods. This is especially the case in rap and hip hop, which can be seen in Table 3 below, but certain songs, like Needed Me by Rihanna in Table 2, likely should have scored higher than they did because of swearing.

The Rap and Hip Hop Genre					
	Song	**Subgenre**	**Vader**	**Topic**	**Words**
Aggressive	Boss	dwn trap	-0.8481	Having sex on Xanax	bitch, pop, xans, fucked, school, aye, …
	LVL	hip hop	-0.8481	His success vs. other rappers	niggas, fuck, new, round, pistol, blew, …
	Bring Dem Thi…	pop rap	-0.7964	Drug dealer life (symbolism)	rings, looking, goons, kill, shoot, canary …
	I Don't Fuck …	hip hop	-0.6808	Comparing ex to girlfriend	bitch, bad, guess, new, lose, thank, hope
	We Dem Boyz	pop rap	-0.5423	He's sick of hoes and lazies	boyz, man, acting, money, hoes, low, …
	Fashion Killa	hip hop	0	A girl's fashion	away, shirt, newest, alexander, wang, …
	Backseat Freest...	hip hop	0.128	Respect him or watch out	mind, respect, die, lead, pray, tower, …
	Ignorance is Bl…	hip hop	0.296	His hard upbringing in LA	ignorance, compton, discharge, bliss, …
	Phone Jumpin	south rap	0.4939	Profiting of his music	hand, residue, feel, kitchen, buy, selling,…
	Jumpman	pop rap	0.6124	Bragging about their careers	way, heard, jordan, trust, shoot, metro, …
Non-aggressive	PRBLMS	dwn trap	-0.8591	Frustration about a girl	hate, worried, tried, remember, explain, …
	She Belongs to…	deep trap	-0.8126	Wife cheating on him	knew, hoes, bitch, shit, tryna, right, …
	911 / Mr. Lonely	hip hop	-0.8074	Anger? about being lonely	line, loneliest, bang, bitches, blues, …
	Nights	pop rap	-0.5574	Nirvana w/ drugs (not death)	want, body, right, family, die, money, …
	No Role Modelz	pop rap	-0.2023	Being fooled by shallow girl	fool, time, love, peace, shame, nigga, fuck
	1-800-273-8255	pop rap	-0.0314	Keep striving (suicide prev.)	little, right, hard, light, shed, snow, …
	All The Way	pop rap	0.4404	Asking a girl to 'choose' him	alright, okay, walk, watch, afraid, long …
	MICHIGAN	hip hop	0.6369	Falling in love in Michigan	brain, stay, fell, think, love, wild, Michigan
	Ex Calling	dwn trap	0.8126	Consider getting back w/ ex	want, talk, best, bird, pick, good, long
	Under Ground…	hip hop	0.9062	Love for his city, Toronto	city, nigga, love, greatest, best, story, …

Table 3. *A comparison of aggressive and non-aggressive rap songs found from query 6.* There are a total of 20 songs and 5 fields for each row. The rows are organized first by somberness and lyrics topics (aggressive and non-aggressive) and then, by the sum Vader score for the 7 respective topic words. Some names and topic words were cut short due to length. The songs were selected from the 105 rows that were returned from query 6 and exported into draper_topic_data.xlsx under the sheet, 'rap_hiphop'.

The rap and hip hop songs in Table 3 showed almost exactly what was expected. In the non-aggressive category, most of the topics revolve around relationships and former partners. There are feelings of grief, sadness, anxiety and even subtle signs of depression, especially in 1-800-273-8255, which is all about Logic's former struggles with depression. Even with the large range of Vader scores, most of the non-aggressive songs were far more somber and it seemed the artists were either extremely critical of themselves or someone else in their lyrics. The aggressive rap and hip hop category was much different though. Most of the artists spent their verses talking about their success and bragging about material things they now had in these songs. In Jumpman, Drake calls himself the Michael Jordan of hip hop because of the success of his career. Dave East talks almost as if he is a drug dealer in Phone Jumpin, but on further analysis, it becomes apparent that he's just referencing to how 'hot' his music is becoming and how money is just rolling in. Compared to the more somber, non-aggressive songs, there is not much to retain from these topics. Most lines are just pop culture references and showing off where they are compared to where they had come from. The non-aggressive lyric topics definitely show more signs for concern. Along with that, the non-aggressive topics songs were very similar to topics found in the indie rock songs of Table 1, which is odd because the subgenre as a whole had very neutral scores with an average Vader of 0.2 and a Valence of 0.5. Perhaps indie rock is a hidden subgenre of concern within rock though.

After the genres had been analyzed, it was still important to attempt to rationalize whether I had emotional vulnerability to these topics. In Figure 7, although respective songs and genres were not included, the Valence scores of every song from the saved library (from old to new) were mapped over time based on the order in which they were saved. Along with that, tags were added to show what period of time related to the saved song. A 5^{th} degree polynomial trend line was added to fit the data and show the overall trend shifts overtime. There ended up being two troughs, between March and May of 2015 and between August and October of 2017, as well as a peak and then a general upwards trend for recent listening habits. Without context, this graph means nothing, but the actual time periods in which these troughs occurred are fairly important times in my own life.

My mother suffered from breast cancer most of my high school career, but it had been re-diagnosed as Stage IV cancer around early January 2015. By analyzing the graph, you can see there is a sharp dip in song selection right around January. Along with that, the trend line continue gradually traveling downwards until about when she passed away in the middle of May of 2015. That being said, things began to look up. I started college, met a girl, and made a few really good friends right off the bat. Losing my mother made it hard to stay positive at times, but I was happy for the most part. Following the minimum of the trough in May of 2015, the trend line begins to rise again. However, that was a difficult Spring. My girlfriend and I parted ways and that was a tough loss. This lead to a decline in my mood, which is visible in my listening habits for that Spring. I will reframe from all of the details of my sophomore year, but, overall, it was a very negative time in my life. I had many conflicts I was dealing with internally and externally and this was even reflected in my grades, which were the lowest they had ever been. That summer, I moved to California, which was very unwanted on my end because it meant leaving some of my best friends. This continuing downward spiral in mood can be seen throughout Figure 7, from saved song 150 to about saved song 70 where school finally started back up. I have had outside confirmation from many friends as well, but this has been one of my best and most positive semesters in terms of my mood. The figure shows this too as my listening habits show a sharp spike in overall positivity. Nonetheless, my listening habits show that I am most likely emotionally vulnerable and my mood correlates to my listening habits. What this means is that I may me more vulnerable to the topics displayed during certain troughs. This does signify that my listening habits might be a risk to my overall mental health.

Conclusion

The study was ultimately a success and justifies further research. The hypothesis was correct in saying that the overall listening habits were negative. Although, they were just fairly negative according to the average trend of valence scores seen in Figure 6. The sentiment analysis was able to identify the rap and hip hop group as having the most negative lyrics, which follows Miranda and Claes's work. The sparseness of the genres revealed in Figure 3 is less risky for mental illness concerns and what a user might consistently be subjected to, but the pareto chart in Figure 4 showed that most of the population still fell within just 25% of the genres. Because of the distribution, these became the main genres of investigation in order to see specific average scores and correlations in genres. Figure 5 and 6 revealed significant concern for an overarching genre, rap and hip hop. Because of this, the main investigation of topics was conducted within the rap and hip hop genre.

Topic modeling was somewhat difficult because of the symbolism used in many of the songs. In order to properly choose topics, I had to lookup song lyrics and meanings. Finding better ways to streamline this process would be beneficial for future research and implementation on larger scales. However, the modeling still revealed that aggressive rap and hip hop music differed greatly from non-aggressive rap and hip hop music. The non-aggressive songs conveyed many saddening and even depressing topics usually surround girls. The artists also were very critical on themselves or a significant other. On the other hand, the aggressive songs varied in topic, but usually revolved around overcoming something, bragging, or violence. Being able to classify what songs are aggressive or non-aggressive in further research would be huge for correctly identifying what songs should be investigated. That being said, the songs selected as non-aggressive showed topics of frustration, anxiety and depression, which is concerning.

When valence was analyzed over time in Figure 7, something really cool happened. The trend line, although being slightly under-fit, successfully correlated to periods in my life where I had a significant shift in my mood. The two troughs lined up almost perfectly with stressful and intense events in my own life. This may be subjective because I would say that I am likely more emotionally vulnerable to song lyrics because of my past and present mental illnesses, but this data may still correlate for the average user. Although this data is only for one user, it still suggests further studying, and with larger sample sizes, because the correlation in moods could help benefit suicide prevention. Correlating users' moods with song moods over time shows that a user more likely has emotional vulnerability to music and that's when topics seem to become concerning. Analyzing what genres people are listening to and the topics in the respective songs during troughs in a time-series graph could be a good way to judge risk of Spotify listeners' current or past mental health.

In a future study, it would also be interesting to color-coordinate peeks and troughs in listening habits by genre to see if that can be a telltale sign for trends in listening habits. A possible conclusion that could be develop is that, when a user starts to change the main genres they listen to, the moods they are subjected to are likely to significantly change.

In my own continued research on this study, I plan on using MusixMatch's API if I can get full access to song lyrics because it seems to grab song lyrics much more efficiently and the library is much larger. Along with that change, I found that I had actually listened to 961 songs in total during just the past year. By grabbing a larger dataset of songs listened to across saved and unsaved songs, a better time-series analysis could be developed. In terms of categorizing concern in self-harm and mental illness, grabbing social media data may be effective as well according to Jung et al.'s study. Nonetheless, there is still much to explore and even more reason to do so because of the correlations and success found in this study.

Appendix of Queries

Query 1:

```
SELECT sg.genre, AVG(ss.valence)
FROM`spotify_saved_songs` as ss
INNER JOIN spotify_saved_genres as sg
ON ss.song_id = sg.song_id
GROUP BY 1
ORDER BY 2 DESC
```

Query 2:

```
SELECT sg.genre, AVG(ss.vader_lyrics)
FROM`spotify_saved_songs` as ss
INNER JOIN spotify_saved_genres as sg
ON ss.song_id = sg.song_id
WHERE ss.vader_lyrics IS NOT NULL
GROUP BY 1
ORDER BY 2 DESC
```

Query 3:

```
SELECT song, date_added, saved_order, valence, popularity, vader_lyrics
FROM spotify_saved_songs
ORDER BY 3 ASC
```

Query 4:

```
SELECT genre, count(id)
FROM spotify_saved_genres
GROUP BY genre
ORDER BY 2 DESC
```

Query 5:

```
SELECT songs
FROM spotify_saved_songs
WHERE lyrics IS NOT NULL
```

Query 6:

```
SELECT ss.song, sg.genre, st.topic, st.word1, st.word2, st.word3, st.word4, st.word5,
st.word6, st.word7, st.vader_words
FROM `spotify_saved_genres` as sg
INNER JOIN spotify_saved_topics as st
ON sg.song_id = st.song_id
INNER JOIN spotify_saved_songs as ss
ON sg.song_id = ss.song_id
WHERE sg.genre LIKE '%rap'
OR sg.genre LIKE '%hip hop'
OR sg.genre = 'dwn trap'
```

```sql
GROUP BY sg.song_id
ORDER BY vader_words ASC
```

Query 7:

```sql
SELECT ss.song, sg.genre, st.topic, st.word1, st.word2, st.word3, st.word4, st.word5,
st.word6, st.word7, st.vader_words
FROM `spotify_saved_genres` as sg
INNER JOIN spotify_saved_topics as st
ON sg.song_id = st.song_id
INNER JOIN spotify_saved_songs as ss
ON sg.song_id = ss.song_id
WHERE sg.genre LIKE '%rock%'
GROUP BY sg.song_id
ORDER BY vader_words ASC
```

Query 8:

```sql
SELECT ss.song, sg.genre, st.topic, st.word1, st.word2, st.word3, st.word4, st.word5,
st.word6, st.word7, st.vader_words
FROM `spotify_saved_genres` as sg
INNER JOIN spotify_saved_topics as st
ON sg.song_id = st.song_id
INNER JOIN spotify_saved_songs as ss
ON sg.song_id = ss.song_id
WHERE sg.genre LIKE '%pop%' AND sg.genre NOT LIKE '%pop rap'
GROUP BY sg.song_id
ORDER BY vader_words ASC
```

Resources

1. Baker, Felicity, and William Bor. "Can Music Preference Indicate Mental Health Status in Young People?" *Australasian Psychiatry*, vol. 16, no. 4, Aug. 2008, pp. 284–288., doi:10.1080/10398560701879589.
2. Calderon, Gab. "VADER Sentiment Analysis Explained." *Data Meets Media*, 3 July 2017, datameetsmedia.com/vader-sentiment-analysis-explained/.
3. Charlie, R. "FitteR HappieR." *RCharlie*, 1 Mar. 2017, www.rcharlie.com/post/fitter-happier/.
4. Jung, Hyesil, et al. "Ontology-Based Approach to Social Data Sentiment Analysis: Detection of Adolescent Depression Signals." *Journal of Medical Internet Research*, JMIR Publications, July 2017, www.ncbi.nlm.nih.gov/pmc/articles/PMC5547245/.
5. Miranda, Dave, and Michel Claes. "Personality Traits, Music Preferences and Depression in Adolescence." *Taylor & Francis Online*, International Journal of Adolescence and Youth, 2008, www.tandfonline.com/doi/pdf/10.1080/02673843.2008.9748008?needAccess=true.
6. Miranda, Dave, and Michel Claes. "Musical Preferences and Depression in Adolescence." *Taylor & Francis Online*, 27 Mar. 2012, www.tandfonline.com/doi/abs/10.1080/02673843.2007.9747981?src=recsys.
7. Santos, Juan De Dios. "Is My Spotify Music Boring? An Analysis Involving Music, Data, and Machine Learning." *Towards Data Science*, Towards Data Science, 28 May 2017, towardsdatascience.com/is-my-spotify-music-boring-an-analysis-involving-music-data-and-machine-learning-47550ae931de.
8. Supalov, Vladislav. "Analyzing My Spotify Music Library With Jupyter And a Bit of Pandas." *Vsupalov.com*, 5 July 2017, vsupalov.com/analyze-spotify-music-library-with-jupyter-pandas/.
9. Neo, Ting. "Audio Music Mood Classification." *Tinkering with Data*, 7 Jan. 2017, neokt.github.io/projects/audio-music-mood-classification/
10. "2016 National Facts and Figures." *Afsp.org*, American Foundation for Suicide Prevention, Dec. 2015, afsp.org/wp-content/uploads/2016/06/2016-National-Facts-Figures.pdf
11. "Monthly Active Spotify Users Worldwide 2017 | Statistic." *Statista - The Statistics Portal*, June 2017, www.statista.com/statistics/367739/spotify-global-mau/.
12. "Topic Modeling Explained: LDA to Bayesian Inference." *TDK Technologies*, TDK Technologies, 2017, www.tdktech.com/tech-talks/topic-modeling-explained-lda-to-bayesian-inference.
13. "U.S. Spotify Users by Age 2015 | Statistic." *Statista - The Statistics Portal*, Apirl 2015, www.statista.com/statistics/475821/spotify-users-age-usa/.