

# An In-depth Risk Analysis of Spotify Listening Habits to Detect Mental Illness

Aidan Draper

CSC 401

Final Project



# MOTIVATION

- Spotify is a global music streaming source that over 140 million people currently use around the world
- I found information on their audio feature scores and believed there was reason to suggest that mood could be analyzed

## THE THOUGHT PROCESS:



# Question to Investigate

- *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?*



DATA COLLECTION AND MANIPULATION

# AUDIO FEATURES SCORES

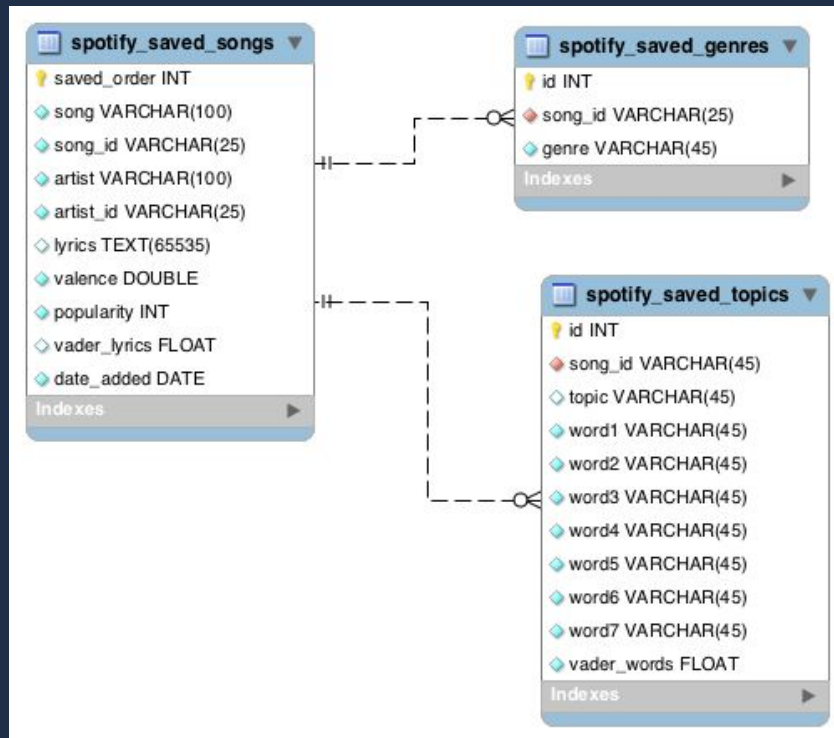


- Spotify's algorithms' scores for songs
- Relevant scores as described by Spotify:
  - Valence
  - Mode (major or minor)
  - Tempo
  - Danceability
  - Loudness
  - Energy
  - Key



# The Entity Relationship Diagram for the Database

6



# Grabbing the Data

- Saved songs, their Valence Score and their Artist Genres were grabbed using spotipy, a python wrapper for Spotify's API
- Using the song name and artist name, a search query was run on py-genius, a python wrapper for Genius's API
  - Only 297 lyrics were found for the 379 songs (~78% of the data)
- Using the returned URL from the genius.com site, BeautifulSoup was used to grab lyrics text off the webpage
- Topic modeling was performed on the lyrics data for individual songs
- NLTK Vader was used on the lyrics data and the topics words

# Constraints and Limitations

- You may only gather your, or public, information on Spotify's API.
- You may only request and grab 50 songs at a time through the API
- The algorithm behind the audio features data (the Valence Score) of Spotify's API are not public.
- Mood, along with emotional vulnerability, are somewhat unmeasurable and some levels of doubt, or deeper investigation, should be considered in a larger study.
- Spotify only assigns genres to artists not songs.
- As mentioned before, only ~78% of the lyrics were found for the songs on Genius's API.



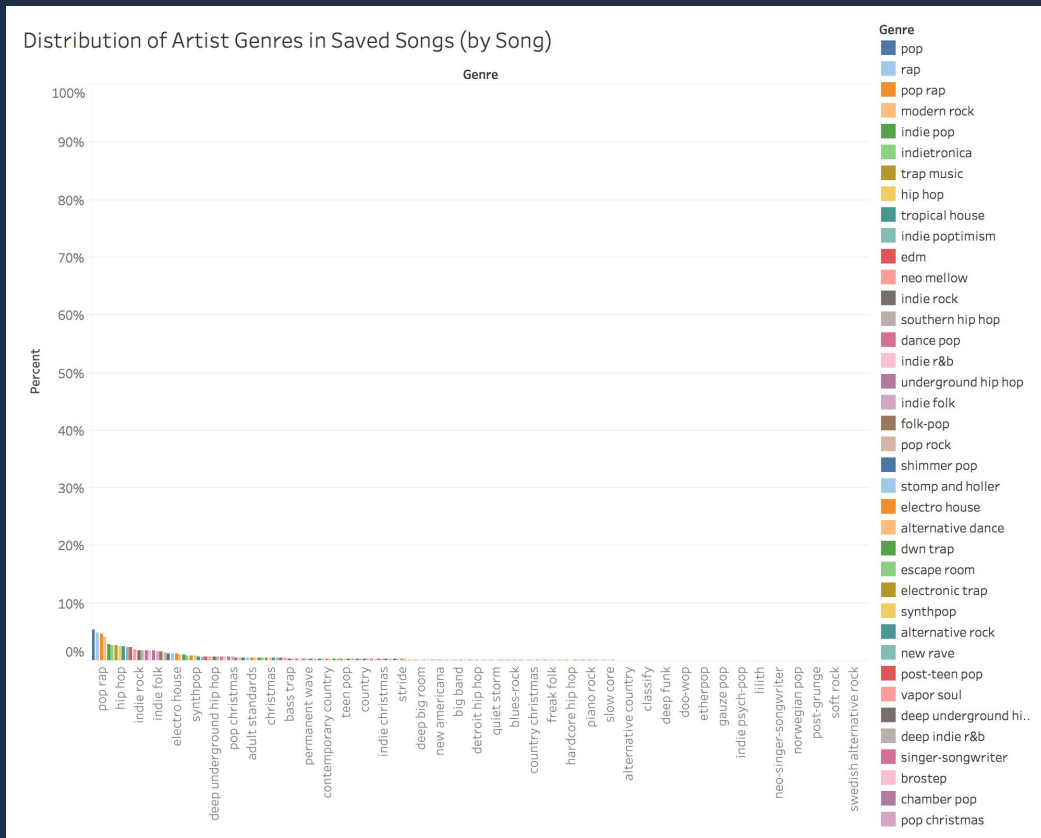


DATA VISUALIZATION

# How are genre's distributed across the saved library?

10

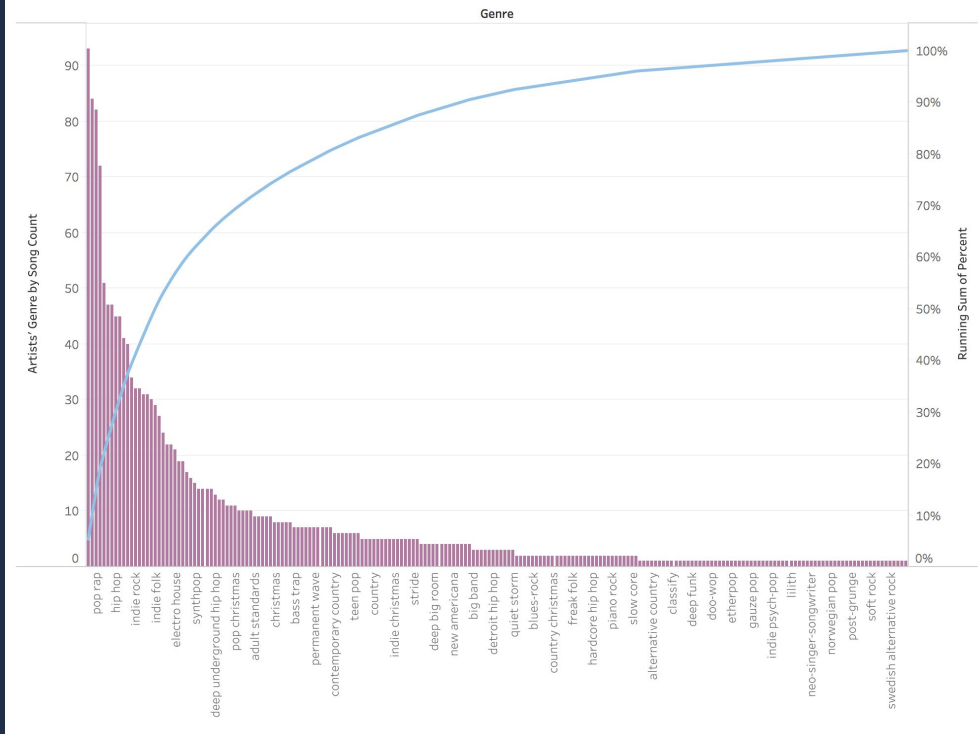
- Top genres were pop, rap, pop rap and modern rock but each made up less than 6% of the total population
- The graph is very sparse!
  - This is good because there is less risk of being exposed to one genre



# Do most of the genres fall within a subset of the data?

11

A Part-to-Whole Analysis of Genres in the Saved Songs

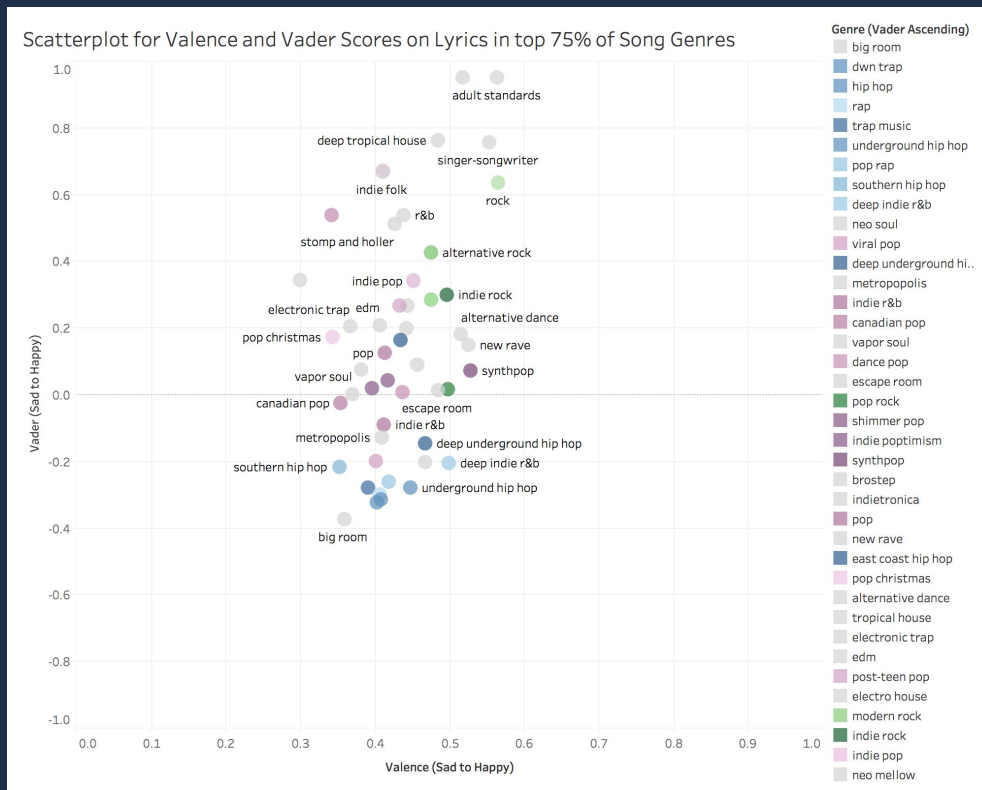


- Yes, so moving forward the top 50 genres (~76%) of the data shall be analyzed
  - This was done to remove outlier genres that contain only a couple songs or less, which might skew the data

# Are the Vader Scores and Valence Scores correlated?

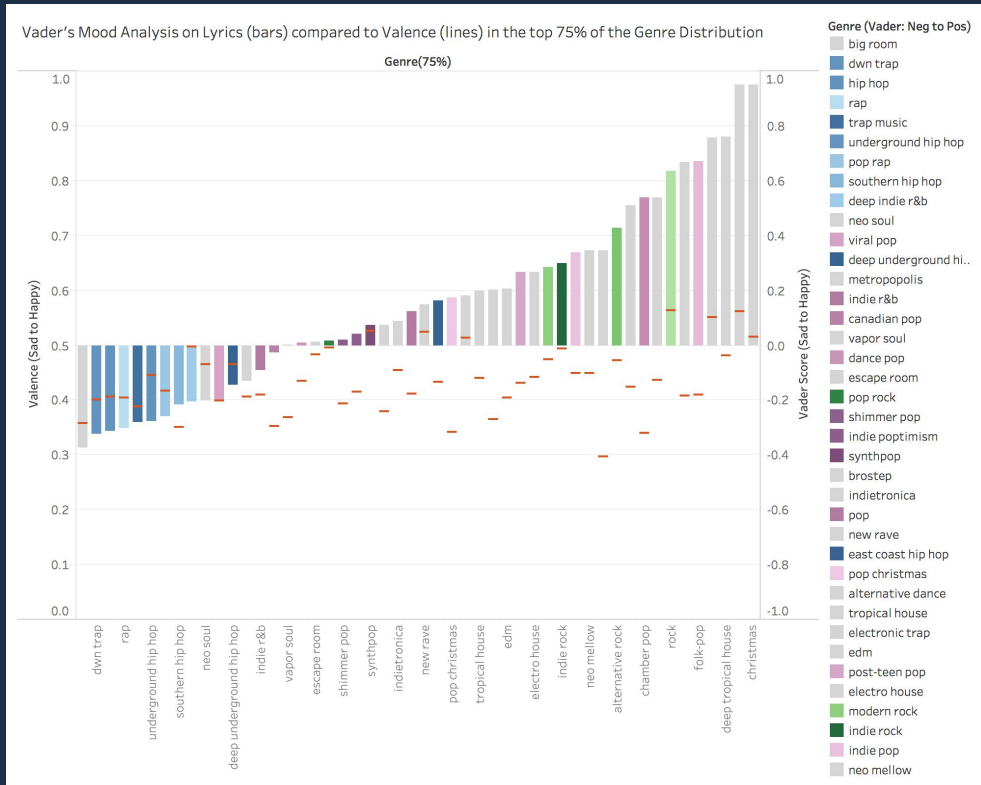
12

- Not really, the lyrics Vader score has a wider range from -0.4 to ~1 while Valence is clustered around .35 and .55
- Coloring by genre reveals some interesting clustering for the Vader scores



# How do the different genres (rap and hip hop, pop and rock) score?

13



- Almost all of the Valence bars fall below 0.5
  - The valence average is around 0.4 = Slightly Negative
- Every rap and hip hop subgenre (except East Coast Hip Hop) fall below 0 for Vader Lyric

What does this mean for the topics  
within certain genres?



## Topic Modeling Results Table for the Rock genre

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'.

## Topic Modeling Results Table for the Pop genre

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'.

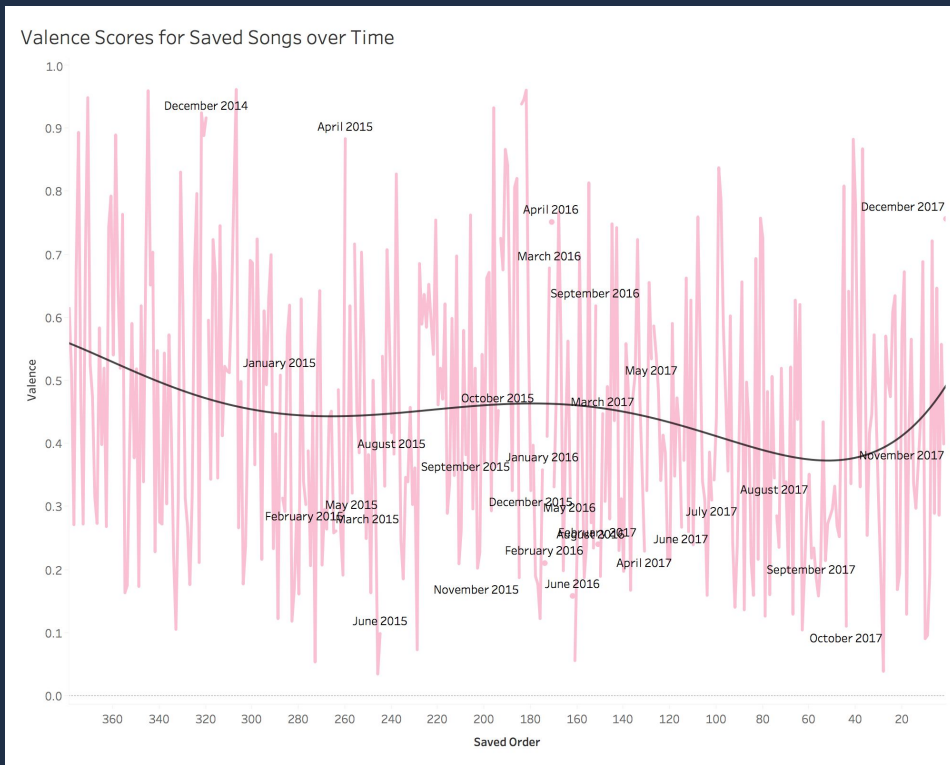
# Topic Modeling Results Table for the Rap and Hip Hop genre

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,...
Non-aggressive	Jumpman	pop rap	0.6124	Bragging about their careers	way, heard, jordan, trust, shoot, metro, ...
	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'.

Am I potentially vulnerable to these topics and are my listening habits correlated to my mood?

# A Time-Series Analysis of the Valence scores for songs in the Saved Library



- Yes, the peaks and troughs in the trend line correlate to events in my life and my overall mood during those times
- This suggests that I may be emotionally vulnerable to my music during troughs





# CONCLUSIONS



# Conclusion

- As the single case specimen in this study, I would consider my listening habits as somewhat risky and overall, slightly negative.
  - This makes sense because I have had mental illness in the past.
- Non-aggressive rap and hip hop as well as Indie Rock are somewhat concerning subgenres to be found in user's libraries because of their a good portion of their topics revolve around depression and self-criticism.
- Based on my correlated mood and listening habits, it suggests that Valence can track peaks and troughs in moods for users who may be more emotionally vulnerable than others
  - That being said, further research is needed on a larger sample population than 1 (lol)

# References

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*, April 2015, www.statista.com/statistics/475821/spotify-users-age-usa/.