Vincent Zhang

Professor Mary Ann Smart

Music 128

December 12th, 2020

Personalized Preference

For many music streaming services, there are a lot of methods used to recommend songs. One of which is to use users' streaming history and their other music data as a huge database to calculate and recommend similar songs to the users. It is one of the most effective ways to recommend and predominates in today's era. However, with the advance in technologies and the emergence of much more new ideas, there is a revolution in data analysis for music recommendation. Companies started to expand towards other areas that might relate to one's music tastes such as users' demography, personalities, and new models for calculations. As I get a chance to work with my data, I want to see if there is any correlation between any of these factors to relate to my music tastes.

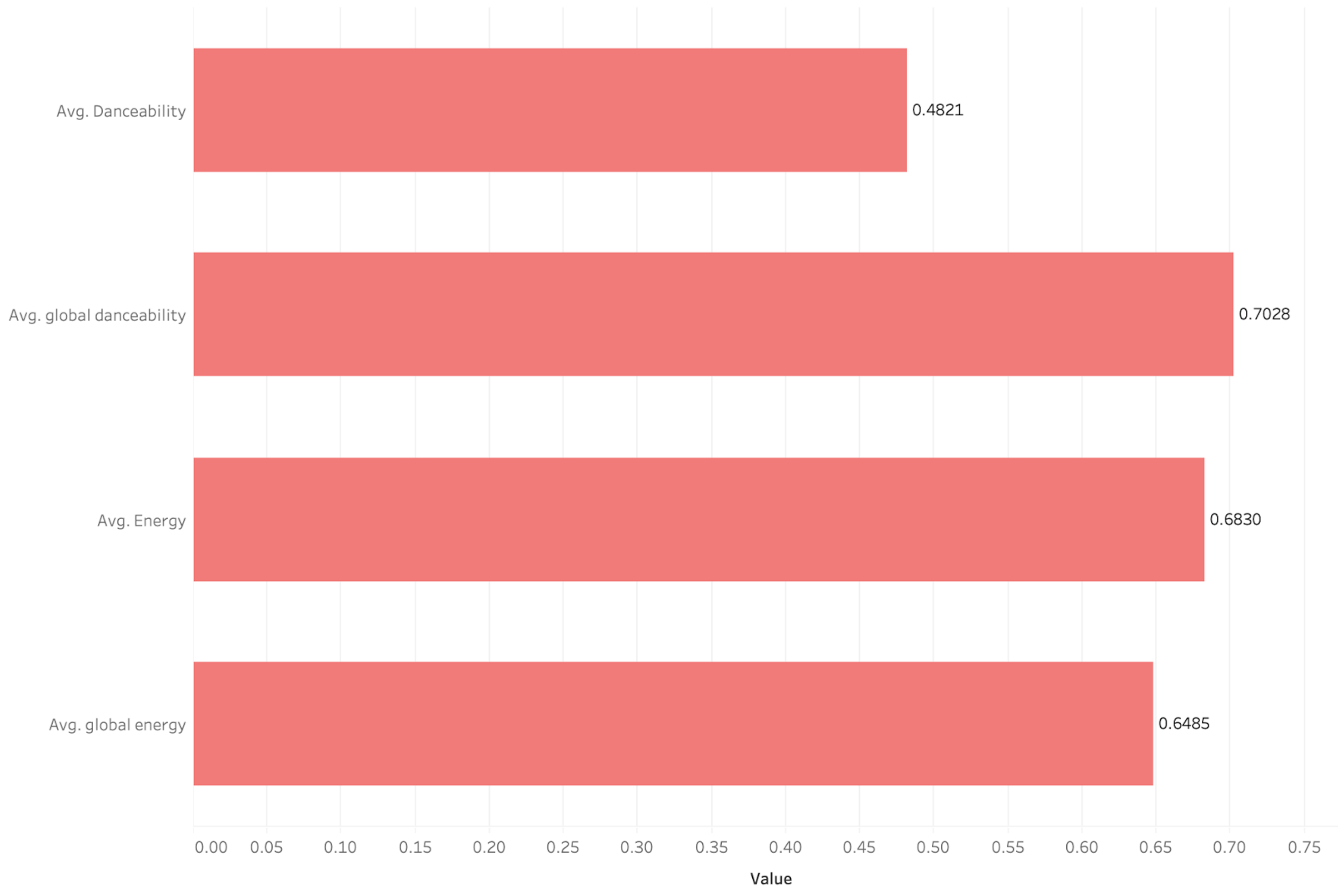
Music data analysis isn't simple; companies have many methods of filtering and calculating people's data to find the optimal recommendation for everyone. Most companies in the 2010s use traditional methods that simply sort through a user's listening, purchase, and rating data and compare them to other music. The calculation is very straightforward, and it does not dive deep into the complexity of the user's music interests. One of the new ways all the companies began to use it in the last couple of years is the use of factor decomposition machines. Xu in his research paper about the factor decomposition machine describes it as "a big data music personalized recommendation method based on big data analysis, which combines user behavior, behavior context, user information, and music work information"(247). This method(FM) dives a lot deeper and focuses on specific user data for the music. In summary of the article, Xu founds out that the new method can find songs and listening attributes from the users to create patterns and models for better calculations. There is more information than ever for music companies to utilize for music recommendations. Some of these attribute information are the elements we have discussed in the course like acoustics, instrumentals, and speechiness of songs. The traditional methods do look at these values more from the ratings and purchase information, but the new factor decomposition method is a big data analysis that brings in multi-layered data that's a lot deeper.

For other attributes of this big data analysis, companies also gather users' listening context and their demography to compare with users' music data in order to formulate better music recommendations. One of which is the context and timing of when the user is playing music. That includes the time, location, occasion, or emotional state of the user. In Pichl and Zangerle's research about the multi-context aware music recommendation, the authors combine *situational clusters* analysis with *archetype clusters* analysis. The first analysis is based on the playlist contexts like its name and played timing during the day, and the second analysis is based on acoustic features(acoustics, danceability) of the tracks in different playlists. By analyzing these variables of different playlists of the user, they were able to provide the most accurate recommendations out of the experiment groups. One thing very notable in their research is that Pichl and Zangerle used the factor decomposition machine from the other paper to conduct recommendation analysis. They believe “FMs allow us to model and exploit the influence of a certain context on the choice of tracks for a given user”(2) and “the use of Factorization Machines allows for easily extending our current approach with further notions of context such as emotion or culture”(20). The use of big data analysis is already prevalent and important for accurate music recommendation. It has extended data analysis to our emotional status, lifestyles, and daily activities, and these calculations actually work. For these big data analyses around users' lifestyles, Drott writes in his article: "Not only does playlist targeting thereby encourage the multiplication of interest-based, lifestyle, and/or psychographic audience segments that can be sold to advertisers; it also expands the range of attributes that may be added to individual users' profiles, increasing the detail—and thus the value—of their data doubles"(255). He analyzes the current systems of music recommendation and praises the efficiency and smart ways of calculations, but that also raises many issues with users' privacy and business ethics. Drott raises the issue that "Spotify and other firms use people's quotidian musical practices as a way of getting them to divulge aspects of their lives that they may not wish to divulge, that they may even be unaware of divulging"(260). These music companies are getting people's lifestyle data without the users knowing it, which they can use for music recommendation, but also as merchandise for customers. Many people are not aware of any of these business practices and continue to supply more data for their music recommendations while their data is also becoming merchandise for the music companies.

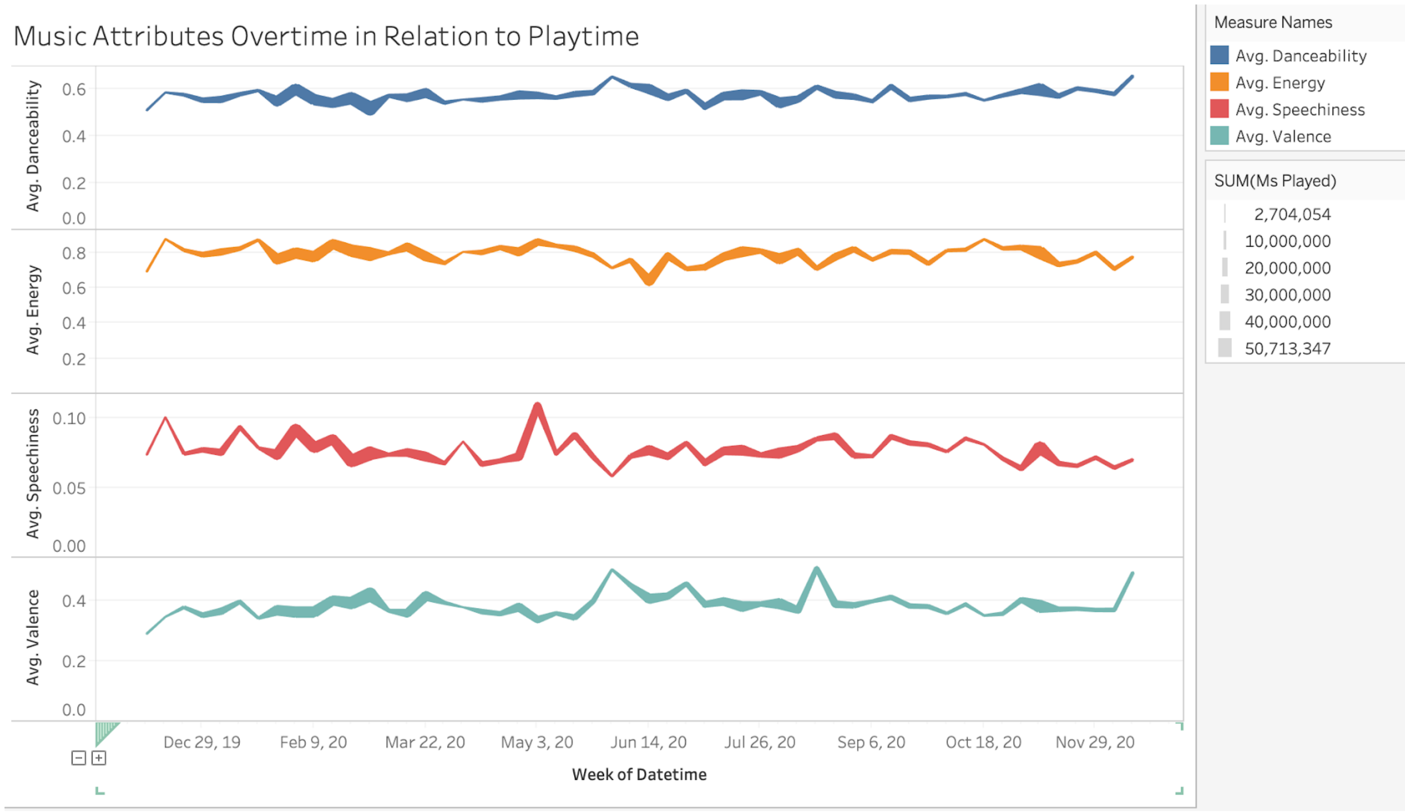
These ways of gathering data are efficient but also invade users' privacy. They take a straightforward approach to analyzing the user's activity around playing music like their demography, music genres, player purpose, etc. Yet, in recent years, there are even more creative ways of getting data from the users concerning their music tastes. Researchers manage to find the patterns of users' humming while listening to music, and many other papers found a correlation between personality and music tastes as well. In Miao, Dong, and Hong's article describing their experiment with Hummings in relation to users' preferences, they conclude "Based on the Recognition of humming songs, we posit that transition preferences are integral to a user's enjoyment of a playlist. We propose an HRRS framework by integrating humming query and reinforcement learning for online learning and adaptation to sequential preferences within a listening session, so as to tailor the playlist to the listener's current scenario" (2162). HRS is Humming-Query and Reinforcement-Learning-based Modeling Approach for Personalized Music Recommendation. It captures users' real-time interactions with the music aka humming, which is surprisingly really helpful in determining if the users like the songs or not. However, such practices also require the recordings of the users while they are listening to music, which again steps into the boundaries of privacy agreement between the users and the companies.

Another practice done by the business for music recommendation is the use of personality. Like our assignments in the course about how personality is correlated to our music tastes, I researched further from the personality quiz and other articles. In one of the articles that is used in the music personality quiz we did for the class, authors Rentfrow and Gosling point out the relationship between people's personalities with their taste in music. In their research, they have concluded many correlations between people's tastes in music to people's personalities in the big five categories: Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness. With their research on all of these perspectives of personalities, I have decided to take an extra step to verify the research results on my own personal Spotify data. They write: "Our analyses suggest that individuals who enjoy listening to upbeat and conventional music are cheerful, socially outgoing, reliable, enjoy helping others, see themselves as physically attractive, and tend to be relatively conventional"(1249). To echo their findings, I found my average danceability and energy data for all the songs I have played on Spotify is shown in the following graph compared to the global top 10 songs' danceability and energy data:

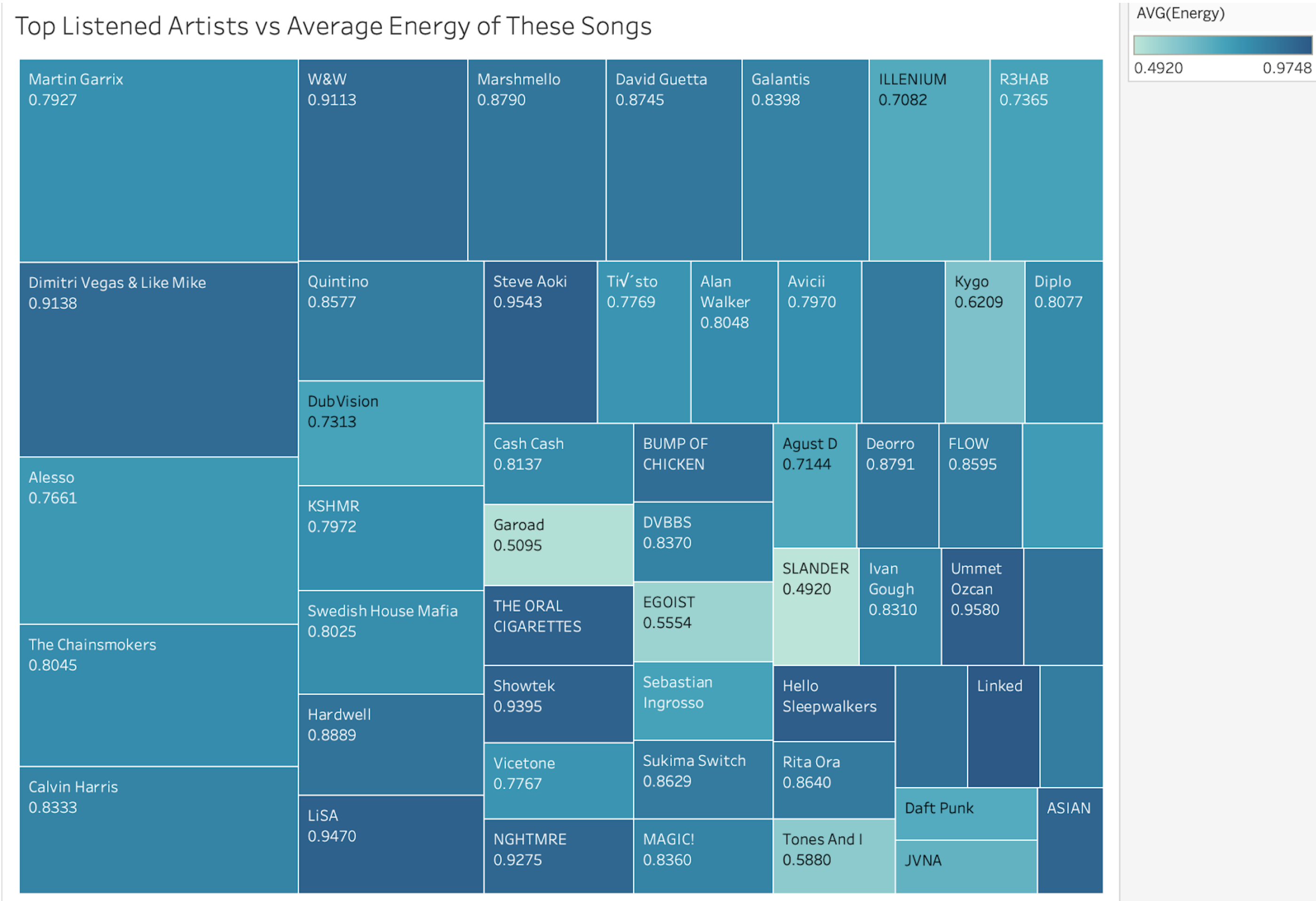
AVG Energy and AVG Danceability for Vince Compare to Global Values



The data shows that my average energy is a lot higher than the global top 10 status, but for danceability, it was lower. I am a socially outgoing person, and many personality tests, including the one I did in class, show that I am genuinely extroverted. I think the data correctly depicts that, but there are errors like the danceability here, which is different from the paper's hypothesis. Here is another graph showing the attributes of the songs I played last year, and the density of the line is based on the minutes of the songs I played over time.



I have included four attributes of the songs and it is based on 2020. The thicker the lines, the more I listened to music at that time. Overall, I think there are no major changes in my listening patterns; the biggest difference is around 0.2 for any attribute values. If my music taste is altered by my personality, I feel like there should be a dramatic change since March due to Covid-19--my living patterns changed greatly and I became a lot more introverted, but the graph doesn't show that. In fact, the Valence seems to have increased a bit. This resonates with one of the points from the paper: "One noteworthy finding was the absence of substantial correlations between the music-preference dimensions and Emotional Stability, depression, and self-esteem, suggesting that chronic emotional states do not have a strong effect on music preferences"(1249). There are some emotional changes and personality shifts that are not shown in our music patterns, and what's more important is that these tests are still incomplete with room to improve. It is impossible to know every aspect of me from my music history, and the models for these recommendation systems are not perfect. At last, I have another graph to show about the topic, displaying my top listened artists:



This graph shows the artists I listened to over 50 times of their songs combined and the average energy value which is based on all of their songs I played. The size is based on how many times I listened to their songs, and the color is based on the average energy, the higher the darker, the lower the brighter. From this chart, I can see the artists I listened to the most and their tracks' average energy. Most of them have surprisingly high energy, much higher than my average, 0.68. I think using my regular routined songs can better reflect my personality compared to using all the songs I play. The recommendation system in the future should consider these factors for the calculation. The model should include more diversity of one user; I might be outgoing, but I do listen to songs that do not match the calculations. The ultimate music recommendation is about people's preferences, not defining who they are.

In conclusion, I want to talk about my overall experience with this project. I took an extended approach from the subject we have learned in this course to further research. I learned so many possible ways that businesses can do to gather our data and use them. These calculations and methods are really smart and effective, but it also comes with drawbacks like business ethics. Privacy is essential coming forward with our reliance on the internet. I can no longer look at Spotify as a simple music app anymore, as I learned its history, practices, and business structure in *Spotify Teardow*n. On the other hand with data, it was also really interesting. I get to play around with my data to see patterns in relation to the research paper. It relates to my career path in the future which is around statistics and data science. Yet, I still didn't know how to use Tableau to its full potential, but I am excited to work with it in the future. I want to express my gratitude for letting me have such an extension for my project. I am very glad that I made this project something I planned originally, and the learning experience was worth my time and effort. Please email me if you need any extra materials like reference articles and Tableau files.

Works Cited

Bauer, C., & Schedl, M. (2019). Global and country-specific mainstreaminess measures: Definitions, analysis, and usage for improving personalized music recommendation systems. *Institute of Computational Perception, Johannes Kepler University Linz, Linz, Austria,* 1-36. doi:https://doi.org/10.1371/journal.pone.0217389

Drott, E. A. (2018). Music as a Technology of Surveillance. *Journal of the Society for American Music, 12*(3), 233-267. doi:10.1017/s1752196318000196

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology, 84*(6), 1236-1256. doi:10.1037/0022-3514.84.6.1236

Miao, D., Lu, X., Dong, Q., & Hong, D. (2020). Humming-Query and Reinforcement-Learning based Modeling Approach for Personalized Music Recommendation. *Procedia Computer Science, 176*, 2154-2163. doi:10.1016/j.procs.2020.09.252

Pichl, M., & Zangerle, E. (2020). User models for multi-context-aware music recommendation. *Multimedia Tools and Applications*. doi:10.1007/s11042-020-09890-7

Xu, D. (2019). Research on music culture personalized recommendation based on factor decomposition machine. *Personal and Ubiquitous Computing, 24*(2), 247-257. doi:10.1007/s00779-019-01343-9