Analysis of Qualities of Hit Music:

Ryan Pascual, Yeehahn Wang-Liu

Abstract:

This study aims to investigate the changes in distinguishing characteristics of popular music – primarily, acousticness, danceability, energy, instrumentalness, liveness, speechiness, and valence – over time through Spotify datasets and Grammy award datasets. The data were preprocessed, merged, and plotted in various ways to answer the following questions:

1. How have qualities of hit music such as danceability, energy, acousticness etc. changed from the 1920s to the present? How do these changes compare to the changes in music in general?
2. What qualities of music makes it popular?
3. What musical qualities do Grammy nominees have in common? Are some more prevalent than others?
4. Is there a bias for certain artists when selecting a Grammy winner? Are some artists favored? Are some artists prevented from winning?

Our analysis reveals clear shifts over time, including a steady increase in danceability and energy, showing the evolution of listener tastes over time. Grammy-winning artists and songs exhibit higher valence and danceability values than general popular music, suggesting that listeners and critics prefer more upbeat songs. A small number of musicians, such as Taylor Swift and Beyoncé, dominate Grammy nominations and awards. Follower-counts and award recognition are weakly positively correlated. The data suggests that Beyoncé has a bias favoring her for winning a Grammy. However, the data is insufficient to support Brian McKnight any claims regarding whether or not there is a bias against him when selecting a Grammy winner.

Methodology

All data is gathered by Spotify or the Grammy awards. All characteristic descriptions can be found here. Values have been adjusted to be from 0 to 100 as opposed to 0 to 1 where possible. This report looks into characteristics of popular music, but it is difficult to qualify popularity. To attempt to adjust for this issue, a large dataset of music in spotify was trimmed such that the only songs remained satisfied the following criteria: $popularity - 0.4(year - 1920) > 10$. The songs that did not meet this criteria were classified as "unpopular". This way older songs are more lenient when it comes to whether or not they're "popular" so that ones that have fallen out of popular media are still accounted for while ensuring a strict criteria for modern music.

Results

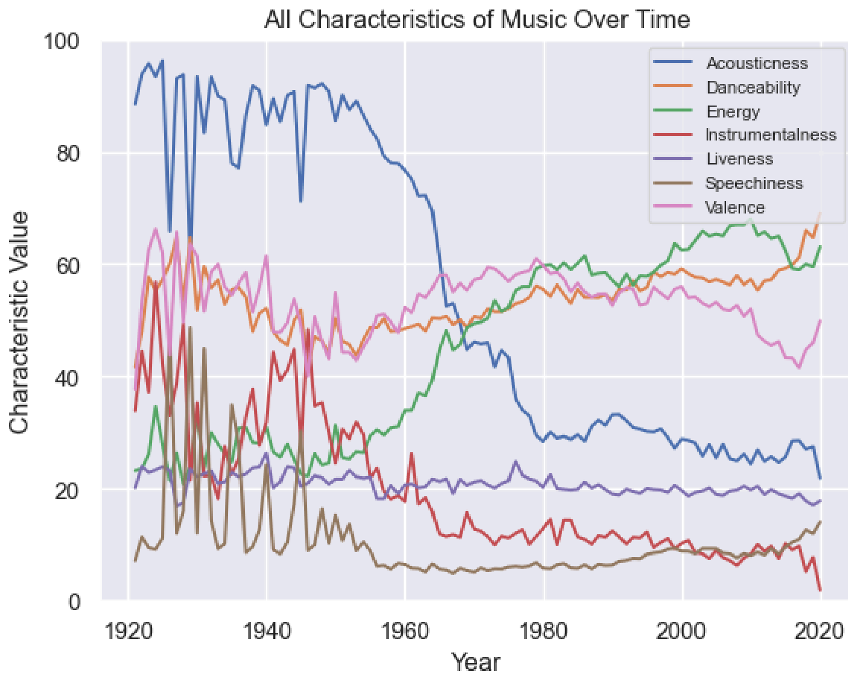


Figure 1: *All Characteristics vs Time*

This plot shows all characteristics within the range of 0-100 and how they changed over time. While very messy from 1920-1950, this can give initial insight as to where to investigate more. In-depth analysis will only occur for characteristics where more insight can be gained by looking further. Assume if the plot is not shown then there is no notable difference seen from this graph. [View all graphs here]. Acousticness sharply declined from approximately 90 to 30 during 1945-1980 where afterwards it stabilized at a relatively low value. Danceability has shown slight increase over time from approximately 50 to 70. Energy has shown a steady increase to around 60. Liveness has remained stable at approximately 25. Instrumentalness has steadily decreased. Speechiness showed sharp fluctuations from the 20's to 40's but then stabilized at around 10. Valence has remained stable at around 50 during the entire time period.

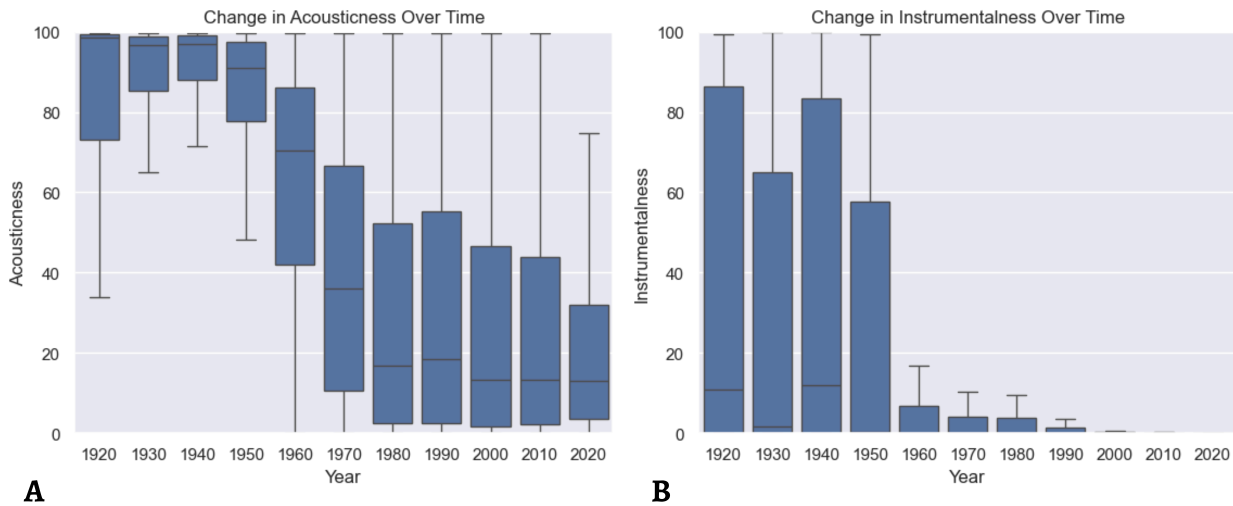


Figure 2.: Change in acousticness and instrumentalness over time box plots.

Acousticness remained varied even as the average decreased. Instrumentalism both decreased on average and in range.

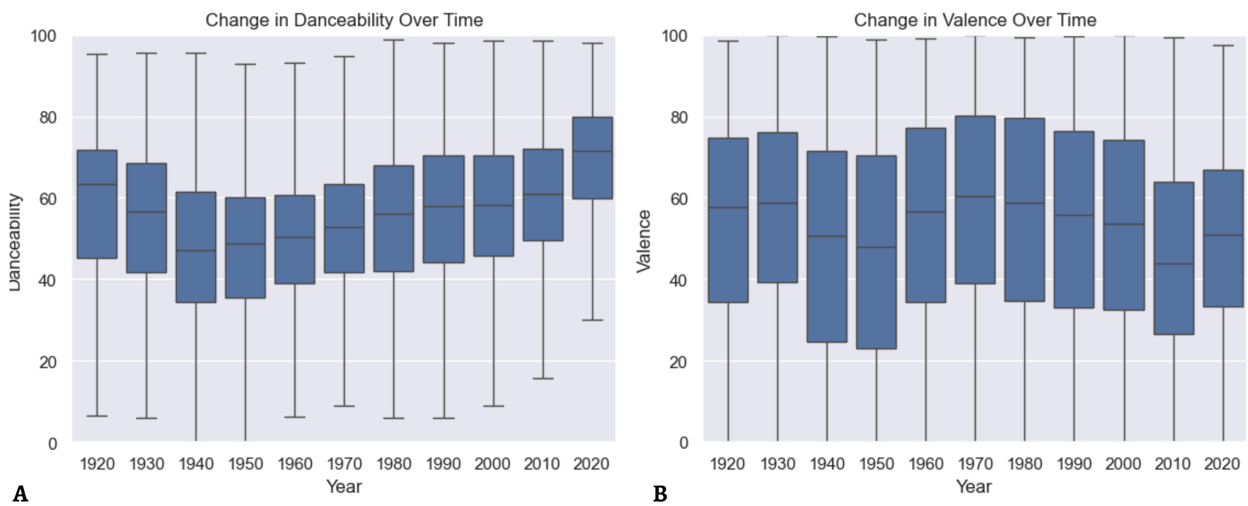


Figure 3: Change in danceability and energy over time box plots.

Danceability and valence showed some variation with danceability increasing but both were relatively stable and showed large ranges.

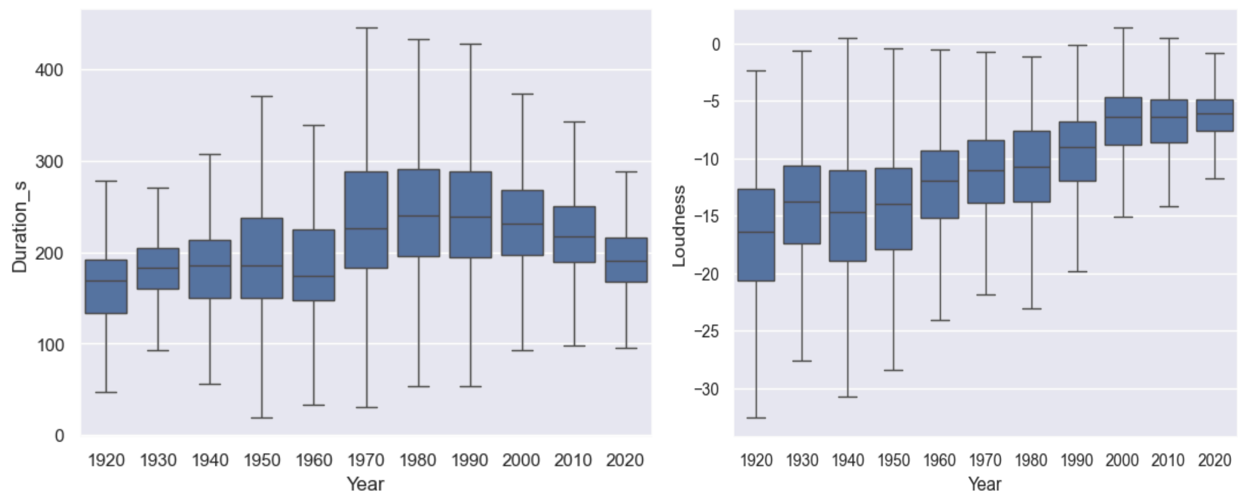
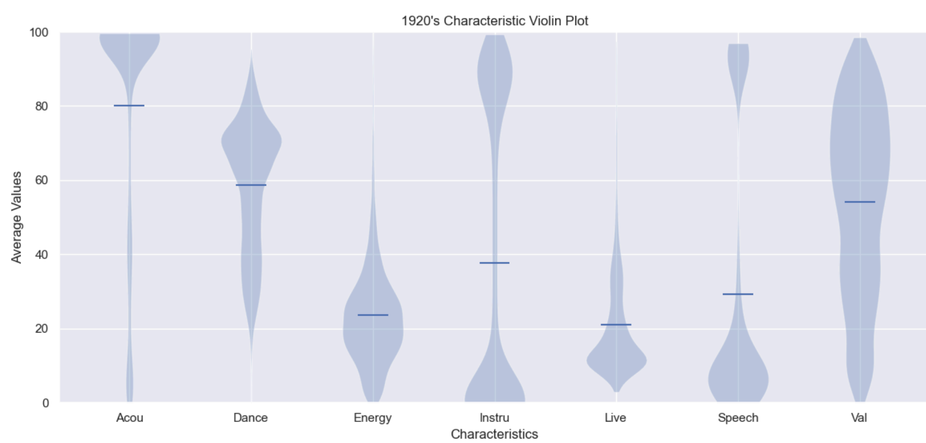


Figure 4: *Change in duration in seconds and loudness over time box plots.*

Once again, due to the large range of these characteristics, the slight fluctuations of the average values are not extremely noteworthy. However, both the range of duration and loudness of music has decreased notably.

A



B

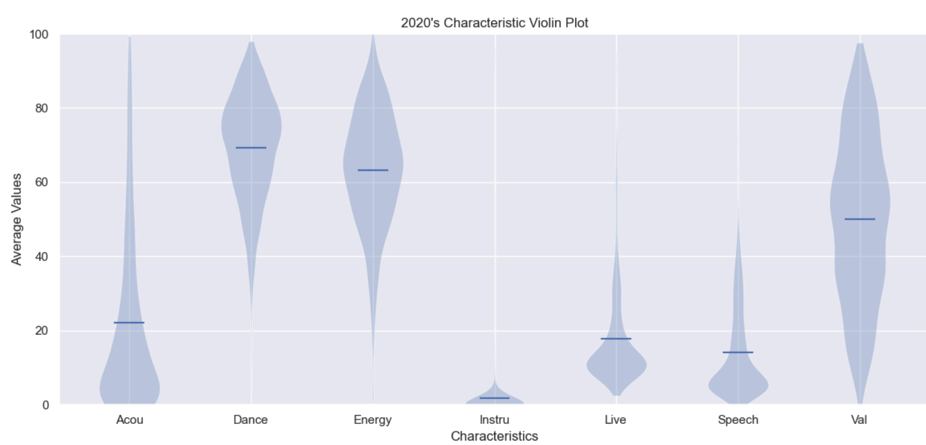


Figure 5: Violin plots of characteristics in 1920 and 2020

Looking at acousticness, it becomes clear that while there are still songs that are high in acousticness, the vast majority are low in acousticness. Danceability and energy have shown increase but are still relatively stable and varied. A new insight in instrumentalness can be made where it is seen that during the 1920's music was either very clearly instrumental or very clearly not instrumental, but by 2020 almost all music was clearly not instrumental. Liveliness and speechiness have not changed notably. Valence's wide range can be clearly seen.

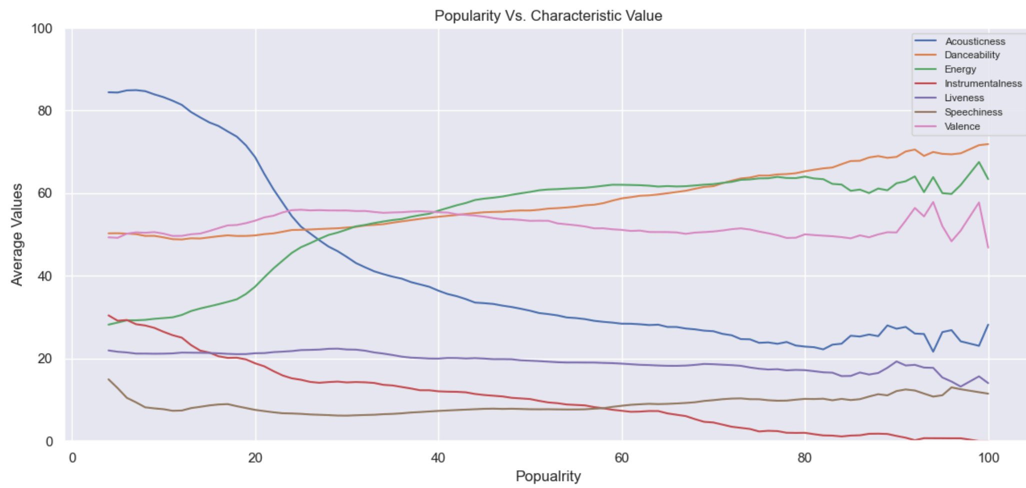


Figure 6: Change in Characteristic vs. Popularity

Danceability and energy are positively associated. Instrumentalness and acousticness are strongly negatively associated. Liveness, speechiness, and valence are negatively associated. To find strength of correlation, r-values were calculated for each characteristic.

Acoustic	Dance	Energy	Instrument	Live	Speech	Valence
-0.573	0.200	0.485	-0.297	-0.076	-0.172	0.014

While there are no characteristics with extremely strong correlation coefficients, acousticness and energy have moderately strong correlations. Suggesting more electric or less acoustic music and more energetic music is typically a differing factor between popular music and unpopular music.

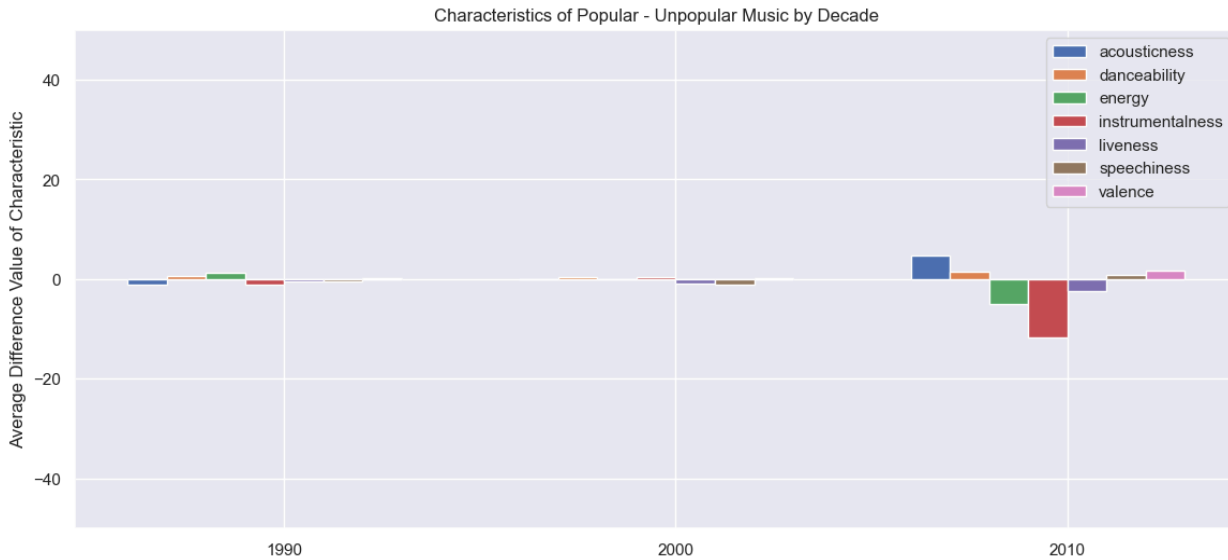


Figure 7: Characteristics of music histogram (popular - unpopular)

There is little difference in characteristics for popular and unpopular music during 1990 and 2000. However, in 2010 differences became more clear and notably acousticness was more prevalent in popular music and instrumentalness was less prevalent. Nonetheless, the differences appear to be minor.

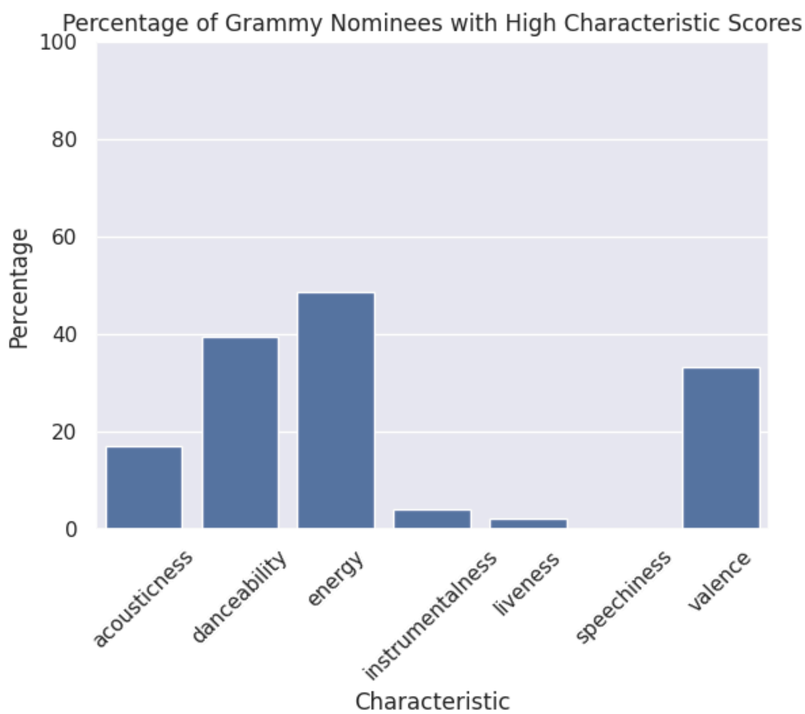


Figure 8: Percentage of Grammy Nominees with High Characteristic Scores Bar Chart.

This plot was generated by finding what percent of grammy nominated songs had a characteristic value greater than 64. This plot shows that Grammy nominees tend to have lots of energy, danceability, and valence, while lacking instrumentalness, liveness, and speechiness.

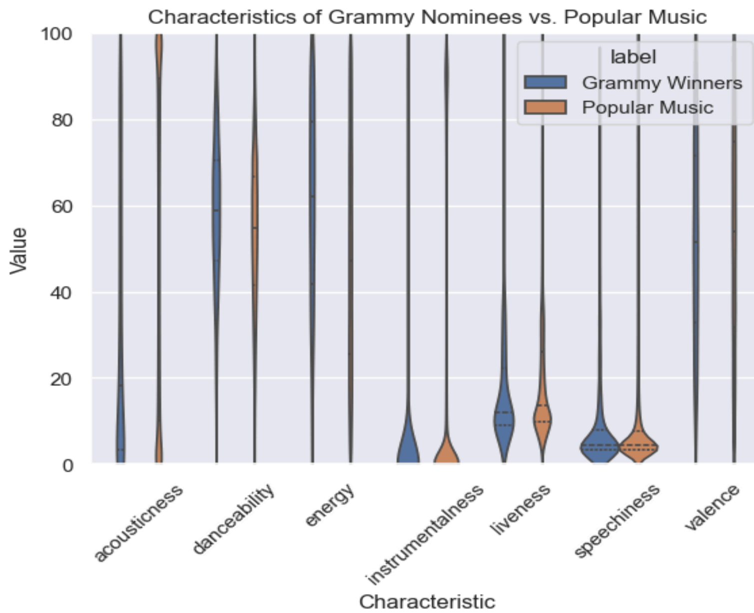


Figure 9: Violin Plot of Characteristics of Grammy Nominees and Popular Music

This plot shows that the musical characteristics of Grammy winners and popular music are similar. Looking closely reveals that Grammy winners tend to have marginally lower instrumentalness, liveness, and speechiness.

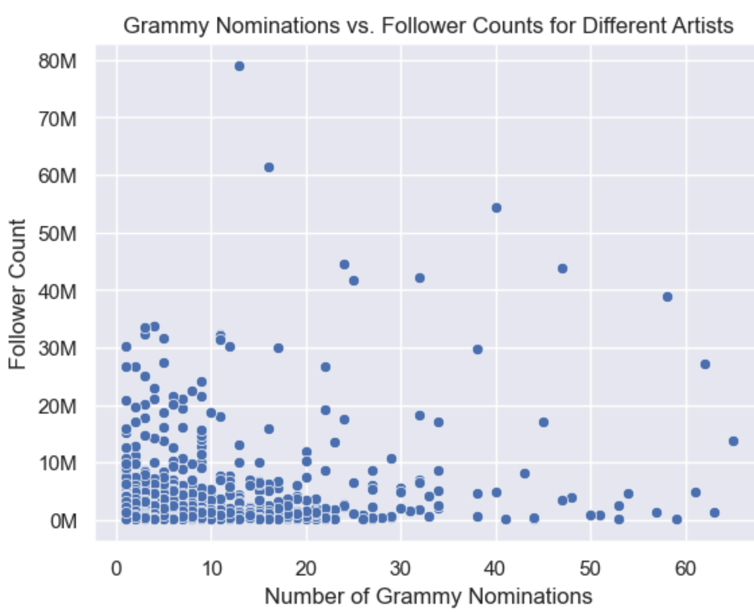


Figure 10: Scatter plot of grammy nominations vs. follower counts

Although there is substantial variation at every level of nominations, there is a weakly positive correlation between number of Grammy nominations and follower count. At lower follower count, there is a higher concentration of artists with a relatively low number of Grammy nominations.
The graphs created to evaluate whether the Grammy's displayed inaccurate information, likely due to the dataset for the Grammy's we used containing several missing entries. Fortunately, the information is easily found online and we continued our statistical analysis with data gathered online. We performed a chi-squared goodness of fit test and found that Beyoncé (99 nominations 35 wins) is likely biased for Grammy winner selection and there was insufficient evidence to suggest that Brian Mcknight (17 nominations 0 wins) was being biased against. A typical grammy award has 5 nominees. Both of the aforementioned conclusions could be drawn regardless if we assumed their music was better than other nominees (25% chance of winning) equal to other nominees (20% chance of winning) or worse than other nominees (15% chance of winning) For a far more in-depth analysis reference the README.

Discussion

The gradual decrease in acousticness and especially instrumentalism could be explained by further development in electronic instruments and microphones. The first electric guitar was made in 1931 [1] and this is when the drop in acousticness is first seen. Furthermore, developing microphones would allow singers to be heard in larger bands, which would allow vocals to be more prevalent in music. The decrease in range of loudness and song duration could be a result of music standardization because of streaming services like Spotify. People do not want to constantly change their volume when listening to different songs and with accurate ways to measure volume, producers have slimmed the volume into a smaller range. For duration, the decreasing range and average value could be caused by the rise in short-form media where popular songs have singular memorable moments that are repeated and the current length is the "ideal" amount. Similarly, the Grammy winners having higher valence and danceability might be due to preference for catchier, more upbeat songs rather than actual listener taste or trends in hit music for the average listener. The lack of a strong correlation between Grammy awards and online popularity could suggest that industry recognition is driven more by connections whereas mass popularity is driven by social media algorithms and artist branding.

Work Flow

Our workflow was rather simple where Yeehahn Wang-Liu would deal with all of the work associated with questions 1 and 2 and Ryan Pascual would deal with all of the work associated with questions 3 and 4. We more or less worked separately from one another besides for occasional check-ins on progress, adjusting and reviewing each other's work, and discussion on findings. The largest challenge we faced was not knowing how to format, clean, and merge our data. With the work flow of the project due to the nature of the assignment (discovery, sketching/planning, and plotting) it was difficult to know exactly what we needed out of the datasets until we started plotting. This forced us to go back and re-preprocess the data on multiple occasions.

Limitations

It is very difficult to precisely define what "popular" means and what qualifies something as "popular". This issue is made worse when considering that the data was culled to be "popular" by using Spotify's popularity statistic which is heavily influenced by current popularity rather than popularity upon release. While attempts were made to adjust for this as outlined in the methodology, ultimately the filter was arbitrary and difficult to ensure that it accurately filtered songs. Furthermore, the idea of qualifying music with characteristics is useful for finding general trends, but this severely over-simplifies music and other micro details (chord progressions, specific instruments, melodic patterns, etc.) could also influence popularity which becomes more evident when viewing Figure 7. Another limitation is the dataset used to analyze the grammy awards had a handful of missing entries and extremely poor and inconsistent formatting which made it nearly impossible to analyze accurately. Fortunately, this issue could be rectified since the relevant data is easily found online and could be used for statistical validation.

Conclusion

People often hear that they must follow the modern American pop style characterized by high energy, high danceability, low acousticness, and low instrumentalism, and certainly this style is popular as seen in Figure 6 and Figure 8 as well as becoming dominant in the modern day as seen in Figure 1. Furthermore, popular music has almost entirely shifted to using electric instruments and components as seen in Figure 3 A and Figure 6. With a moderately strong negative correlation between popularity and acousticness it would seem that popular music is defined by electronic instruments and sounds but it is possible that electronic music is simply easier to produce and produced more often as seen by the wide range in Figure 3 A. Throughout the entirety of the analysis there was a consistent trend of very large ranges of characteristics that fit into popular music. Figure 2 A, Figure 3, and Figure 5 all show that characteristics can vary wildly and still be popular in the music industry. These findings could be beneficial to anyone interested in creating music. While the American pop style is popular, it is by no means the only way to create popular music and perhaps this data can encourage artists to pursue music in a style they enjoy.

The research performed in this project was intended as a starting point for further research into both musical trends and styles, but also culture. We believe this has been a great success in that regard and ideas for further research are listed below.

Further Research

Some questions that can be answered with further research are:

- How does the data change if we use data from other streaming platforms in addition to Spotify?
  - This would reduce bias because in our study, we used data from only one streaming platform
- How would our results change if only data from the United States were included?

    - This would reflect more of our culture and be more compatible with the Grammy dataset, which stems from an American award
- How have smaller details (chords progressions, rhythm, melody, etc.) also changed?
- Are there any official records the Grammy's showing bias when selecting winners?
- Is the change in acousticness and instrumentalism actually a result of technology changes or changes in the taste of listeners?
- Can changes in culture such as the 1960s Counter Culture be reflected in changes in music?

Challenge Goals:

Multiple Datasets:
There were four raw datasets: artists.csv, data_by_genres.csv, grammy_award_data.csv, and spotify_datset.csv. Descriptions of all datasets can be found in the discovery document. There were 6 organized and processed datasets. The majority of the creation of these datasets was done by merging the datasets with spotify_dataset.csv. While simple on paper it proved to be extremely challenging due to the datasets all having different ways of listing artists.
Statistical Validation:
Statistical analysis was almost entirely used to answer question 4 and proved extremely useful for drawing conclusions. While it did require us to ignore some invalidated conditions, given the context of the situation these conditions would not likely largely misconstrue results. However, we nonetheless carried with additional caution which made us unable to claim Brain Mcknight as being biased against.
(A more indepth description of the challenge goals can be found in the README)

References

Spotify_dataset – **dataset from Spotify containing audio features of popular music from 1920 to 2020**

Data_by_genres – **dataset from Spotify that contains audio features of genres**

Artists – **dataset from Spotify that contains characteristics of various artists**

Grammy_award_data – **dataset containing all Grammy awards and nominations from 1966 - 2025**

1. "Invention: Electric Guitar". *www.invention.si.edu*. Lemelson Center for the Study of Invention and Innovation. Archived from the original on 24 August 2018. Retrieved 28 May 2025.