Billboard Hot 100 Songs and their Spotify Attributes

Team B.S. (yxing12, tlamlert, jgoshu, rchau1)

Hypothesis

With music trends constantly changing, our team is interested in investigating how measurable traits have changed over time for Billboard Hot 100 songs. Specifically, we would like to test 1. correlation between a song's Danceability score and the duration they remain on the Billboard Hot 100 chart; 2. difference in average Valence score for songs that remain on Billboard for a longer vs shorter time period; 3. difference in average Energy score for songs on the 1970s and 2010s Billboard.

Data

We scraped weekly data from https://www.billboard.com/charts/hot-100/ from Jan 1970 to Apr 2022. This contained information about song name, artist name, weeks on the chart, and date of first appearance on the chart. For every Billboard song that can be found on Spotify, we collected song features including danceability, energy, loudness, valence, tempo, etc. using Spotify API. There are about 200,000 data points. The number of weeks on the chart is skewed toward lower values as depicted in the diagram below. There is also an unusual peak in the number of songs that stayed on the chart for exactly 20 weeks.

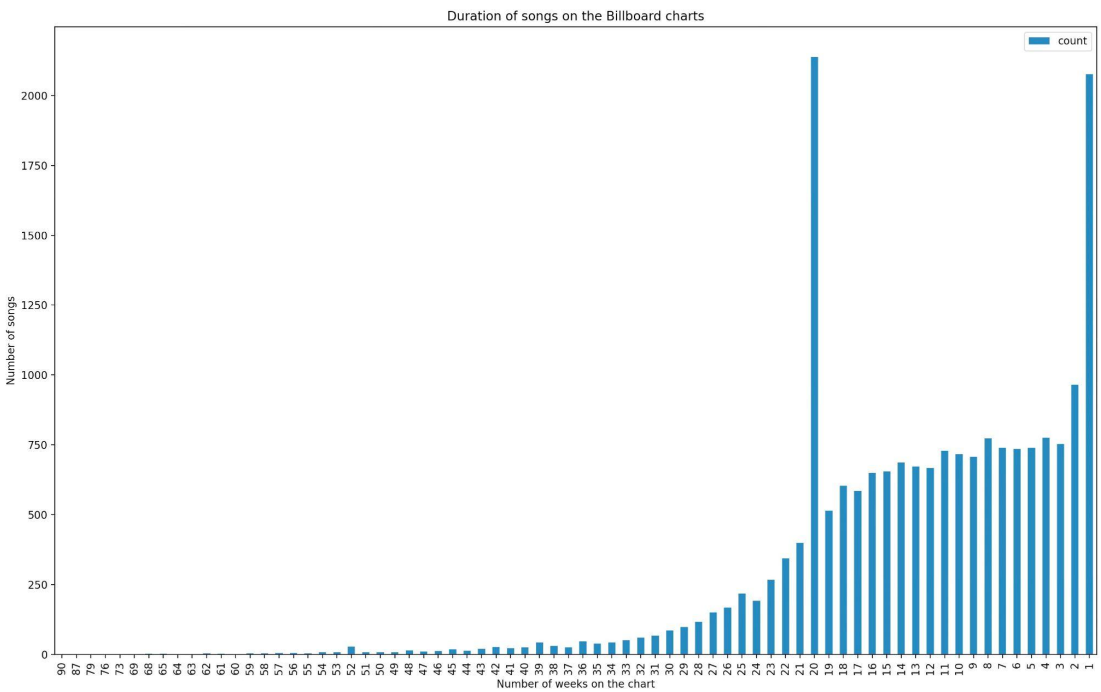

Figure 1. Duration of songs on the Billboard charts

Findings

Claim #1: There is no statistically significant correlation between the song's danceability and the number of weeks it remains on the Billboard Hot 100.

Support: We performed a chi-square independence test to check if danceability is related to a song being on the Billboard chart for a period of 10 weeks. Our test statistic is 51522.998 with a p-value of 1.0. This means that our test is not statistically significant.

Claim #2: There is a statistically significant difference in valence score for songs that stayed longer on the Billboard in the 2010s, 2020s, and overall throughout the decades (1970s - 2020s).
Support: We performed a two-sample t-test to check if the valence (positivity) of songs that are on the Billboard for a long time are statistically different from the songs that are on the Billboard for a shorter period of time. The detailed test statistic and the p-value are shown in the table below. From the p-values, we see that the test is statistically significant for the 2010s, the 2020s, and overall.

Decade	Test Statistic	P-Value
1970s	0.35	0.72
1980s	-0.50	0.62
1990s	-0.42	0.68
2000s	-0.27	0.78
2010s	3.41	0.00066
2020s	3.64	0.00028
All	-4.099	$-4.17E^{-5}$

Figure 2. Table showing test statistics and p-values for two-sample t-test looking at the relationship between average Valence score and number of weeks a song stayed on the chart

Claim #3: There is a significant difference between the energy of Billboard Hot 100 songs in the 1970s versus the 2010s.
Support: We performed a two-sample t-test to check if the energy in the songs from the 1970s decade is significantly different from the energy in the songs from the 2010s. Our test statistic is -17.86 with a p-value of $5.48E^{-70}$. This means that our test is statistically significant. This claim is also supported by the energy bar chart shown below.

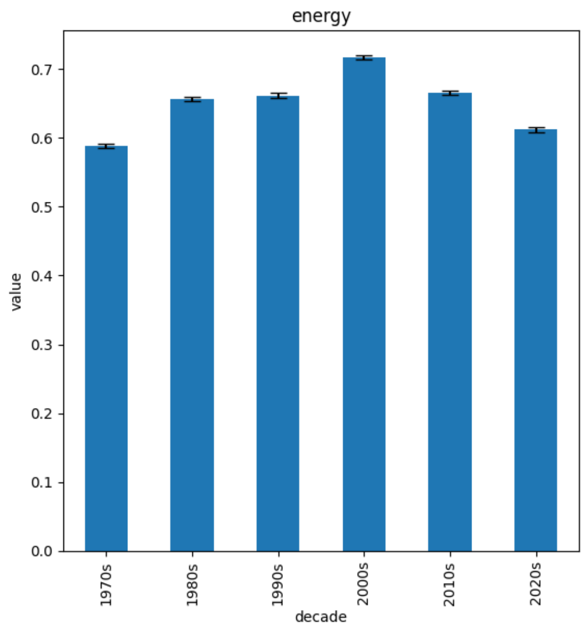

Figure 3. Bar chart displaying average energy levels for songs on the Billboard Hot 100 over the decades, based on songs from the 1970's to 2000's

Predicting Duration of a Billboard Hot 100 Song Lasting on the Chart

Team B.S. (yxing12, tlamlert, jgoshu, rchau1)

Goal

With our analysis of Billboard Hot 100 songs and their traits, we are interested in how they have affected the success of any given song after its release. Presently, influences from digital platforms have shifted the paradigm for mainstream and we would like to investigate which direction these shifts, if any, are taking the music industry. We are interested in the following task setting: given historical Billboard Hot 100 songs from the 1970s to 2020s and the number of weeks they stayed on the chart, predict the number of weeks that any Billboard Hot 100 song can stay on the chart.

Data

We used the dataset described in the above section to train and test our model. Specifically, we used song features given by Spotify API as inputs and the number of weeks each song stayed on the Billboard Hot 100 songs chart as prediction labels. In order to alleviate the skewness of our dataset, we divided prediction labels into 5 discrete ranges: 1-3 wks, 4-8 wks, 9-13 wks, 14-18 wks, and 19+ wks. These are the 20th, 40th, 60th, and 80th percentile intervals, respectively. Due to an unusually high amount of songs that stayed on the chart for exactly 20 weeks, we decided to exclude them from the dataset.

Model+Evaluation Setup

We built a classification machine learning model using a simple feed-forward neural network with 3 dense layers. The model used categorical cross-entropy which was used as the loss function during the training. The model is evaluated based on the accuracy of its prediction since we are interested in predicting how long a song stayed on the chart. We chose a classification model over a regression model due to the uneven distribution of the dataset. By grouping prediction labels into equal groups, we allowed the model to predict just the approximate number of weeks rather than the exact number and only penalized the model when the prediction was far off.

Layer (type)	Output Shape	Param #	Activation function
dense (Dense)	(None, 64)	832	ReLu
dense_1 (Dense)	(None, 32)	2080	ReLu
dense_2 (Dense)	(None, 5)	165	Softmax

Figure 4. Detailed summary of classification model including the dimension and activation function in each layer

Results and Analysis

Claim #1: Our model achieves higher accuracy than a dummy model that always predicts the label with the highest likelihood.

Support for Claim #1: We evaluated our model based on the accuracy of the prediction. Our model achieves 31.21% accuracy on the training data and 30.53% accuracy on the testing data. The distribution of labels across the entire dataset is 21.76% (1-3 wks), 21.57% (4-8 wks), 20.01% (9-13 wks), 18.22% (14-18 wks), and 18.44% (19+ wks). This means the dummy model that predicts a single label with the highest likeliness (1-4 wks) will achieve only 21.76% accuracy.

True labels: [4 4 3 3 0 0 3 0 4 1 3 1 3 2 3 1 1 0 4 1 1 3 2 0 4]
Prediction: [4 1 1 4 0 0 1 0 1 0 2 0 1 4 4 1 1 0 1 1 1 2 2 1 0]
Figure 5. Example target labels and predictions

Claim #2: The model fails to capture the features of songs that stayed for 9-18 weeks.
Support for Claim #2: According to the confusion matrix, our model rarely predicted 9-13 weeks or 14-18 weeks regardless of the input song features.

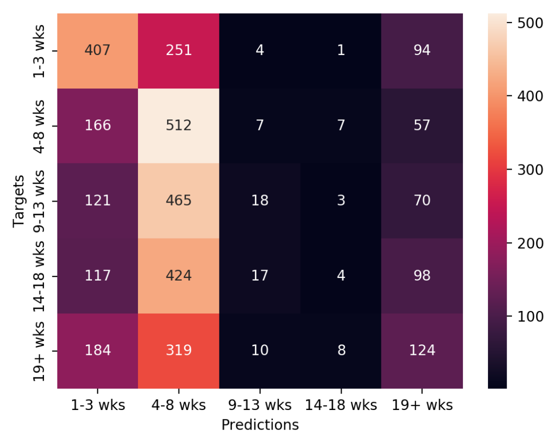

Figure 6. Confusion matrix of the classification model

Socio-Historical context and Impact Report

Socio-historical context

Throughout the decades in the late 20th and early 21st century, the popularity of music at points in time were directly linked to social movements. Evidence shows that music since the 1960s can be categorized into various revolutions. Beginning around 1964 was the introduction of British rock bands (i.e. Rolling Stones, Beatles), which introduced an upbeat and energetic sound that fizzled out the previous slow and smooth music of the blues and jazz era. During the 1980s, there was the introduction of more technology produced music, and in the 1990s came the domination of rap and hip hop (Mauch et. al., 2015). The prevalence of these musical revolutions may have greatly shaped the way our data looked throughout decades. <u>For instance, the introduction and continuation of more upbeat music between the 60s and 80s coincides with the events of the Vietnam War.</u> Soldiers on the battlefield were known to play the upbeat music of the time to keep their morale up while fighting overseas (*Time*, 2017). In our data, we see one of the largest jumps in overall energy from 1970 to 1980, which may very well be a direct consequence of this event. <u>The major stakeholders in our project include artists who are trying to produce more relevant and popular music, as well as record labels who want to see where music trends are heading before they sign new artists.</u> Our data shows the attributes that are most prevalent for America's most popular songs in modern history. Artists from all over could take advantage of the factors that contribute to success in the American music industry and find more success in their music. Unfortunately, some artists who may be harmed are those trying to break into the popular mainstream. The research conducted looks only at songs who have been on the Billboard Top 100 at least once. We observe the factors that contribute to a song remaining on the Billboard for longer. Therefore, this research may contribute to the cycling of the same popular artists to remain in popularity, which will hurt the chances of new artists trying to get their foot in the door. <u>Some relevant research that has been conducted on the popularity of music stems from analyzing the lyrical content of songs and the psychology of its listeners.</u> For instance, a professor at the University of Pennsylvania saw that songs saw more success on the Billboard Top 100 if they used the word "you" more often (*Knowledge at Wharton*, 2021). The societal impacts of this research was that it opens the doors to analyzing the meaning of lyrics songs produce. The messages in music, and especially popular music, may greatly define the culture of its time. Our research does not observe the impact of the lyrics of songs, nor does it look at other important factors such as attributes of the artists (race, age, etc.) that may contribute to the popularity of songs. The societal impacts from a study that observes the meaning of lyrics in popular music, as well as the demographic of artists, may prove itself to be even more meaningful in describing the socio-historical context of the era.

Ethical Considerations

While Billboard is a globally recognized music website, the Hot 100 ranking only feature songs popular in the US. In addition, the song rankings are not solely based on listeners' preferences. A greater emphasis will be given to songs with higher sales volume, paid subscription streams and songs with ad supports. This reflects societal bias contained in our data: the system is in favor of songs produced by larger companies or songs that can be better publicized by their producers. It is not a true reflection of music tastes in a certain era and thereby affects our analysis. Mainstream music companies cooperate with each other to achieve greater success and increased profits for the common good. Therefore, the bias here is relatively hard to mitigate, given the power of capital in the industry. There are also issues we foresee arising from our analysis. Stakeholders that might care about our result include singers, producers and songwriters, as they want their production to be successful and well known across the world. The trends we found might become “formulas” for producers to follow so that they can produce successful pieces. However, diversity and creativity are of great importance in the arts industry and we do not want all songs to be the same just because they are more likely to go viral. As mentioned in the previous section, our analysis might potentially harm those new artists trying to enter the industry, since we only looked at songs that have ever made it onto the Billboard. It might lead to a scenario where the famous become more famous over time, but the newcomers never get a chance to showcase their talent. This could be some possible issues of our project result, as it leads to negative impacts.

We do not foresee any issues related to privacy arising from our project and our analysis. As mentioned earlier, we scraped our data from Billboard Hot 100 ranking and Spotify API, which are both readily available to the public. We performed statistical tests using the data and trained a model using machine learning methods, which are common uses of public datasets. Hence, we are using the data in a manner agreed by the provider. Typically in a setting like this, privacy issues arise when researchers join multiple datasets that inevitably reveal private information about individuals that were not known originally. However, in our case, we choose to join two very publicly available and easily-observed pieces of information about America’s most popular songs. The results we achieve also mainly communicate information about the aggregate music landscape of decades over time, and do not have harmful implications for any individual artist or song.

Works Cited

https://royalsocietypublishing.org/doi/full/10.1098/rsos.150081
https://time.com/4949617/music-vietnam-war/
https://knowledge.wharton.upenn.edu/article/what-makes-songs-popular-its-all-about-you/
https://www.billboard.com/pro/billboard-changes-streaming-weighting-hot-100-billboard-200/