Changes in GDP influence attributes of popular music with lags between one and four years

Heather Shen

Abstract
Song hotness, as measured by the number of sales and/or downloads of a certain record, has been shown to be predicted by the song's attributes. It is observed in this report that the year in which a song is released is almost always a significant variable in estimating the "hotness", or popularity, of the song. This suggests that, aside from the attributes of the song, there are time-related variables at play. Thus, in this investigation, the relationship between the economy and music is analyzed. It is shown that GDP, unemployment, stock prices, and consumer sentiment are positively associated to loudness, tempo, and time signature of popular music between 1961 and 2010. By far the most influential economic factor on music, the effect of changes in GDP occurs mostly at lag one while the effects of unemployment rates and consumer sentiment measurements occur between lag one and lag two. It is also demonstrated that changes in GDP are positively associated with the popularity of certain genres such as Rock and Pop.

Introduction
In a profit-generating industry such as music production, it is useful to analyze what makes a song a "hit". Previous and ongoing research most often look at individual characteristics in a song to determine its potential to become popular (Middleton). Textual analysis of lyrics has become a growing trend (Khan) as data analysis tools for text become more readily available. Easily measured variables of songs are also common to study among scholars (Mauch, MacCallum, Levy, and Leroi). These attributes studied include duration, loudness, danceability, among other metrics. In fact, many research groups currently provide algorithms and prediction models that allow users to examine the probability of creating hit songs. In addition to these services, many music-streaming companies now use machine learning to predict the songs that users will like based on previous history of music preferences.

Though music analysis based only on attributes of the music is useful knowledge, this report posits that there exist variables outside of song attributes that affect a song's popularity. Just as Beyoncé's *Lemonade* (2016) tackled social issues such as police brutality and racism while *The Times They Are a-Changin'* by Bob Dylan (1964) helped popularize the Civil Rights Movement, it is entirely conceivable that there exist external variables that influence the popularity of certain music. This is especially plausible when considering the ubiquity of music in everyday life – an average American listens to four hours of music every day (Stutz).

Thus, this report intends to focus on exploring the effect of economic variables on the music preferences of the population. The music dataset used is compiled by The Echo Nest, a music intelligence company, and includes the attributes of one million songs. These data points were collected through The Echo Nest's company data and among the attributes included by song are Artist Name, Tempo, Loudness, Year Released, Duration, and more (THIERRY). The economic data of the GDP and S&P 500 Adjusted Close were obtained through Yahoo! Finance. Other economic data measured were Unemployment, obtained through the US Bureau of Labor (Bureau of Labor Statistics Data) and Consumer Sentiment, provided by the University of Michigan's Survey of Consumers (Surveys of Consumers).

The analysis performed in this report found that GDP has a leading effect on certain variables that include Loudness, Tempo, and Time Signature of all popular songs. In addition to this, GDP also has a significant leading effect on the overall hotness and single attributes of certain genres.

Following the Introduction, the Methods section will explore the basic trends found in the music data that are of interest. It will also discuss the techniques used to measure data, explain the estimation and prediction models tested, and the use of time series in finding correlation lags between variables. The Results section will follow with an overview of the findings in this investigation and lastly, the Discussion and Conclusions section will interpret the results and expand upon the implications of this research.

Methods

The music dataset was first separated in order to account for missing data. Entries that did not include a Hotness rating or were released outside the year range of 1961-2010 were automatically omitted. The remaining dataset consists of 294,026 points. One subset was then taken to include only songs with a Hotness rating above 0.5 (out of a scale 0-1) so as to analyze popular songs in comparison to all songs released (n = 137071). Figure 1 demonstrates the distribution of Hotness ratings over all available songs in the dataset. As is shown, this distribution appears approximately normal with an average of 0.5. The original dataset was also stratified based on Genre, with a total 42 unique categories (See Appendix 1 for the full list)

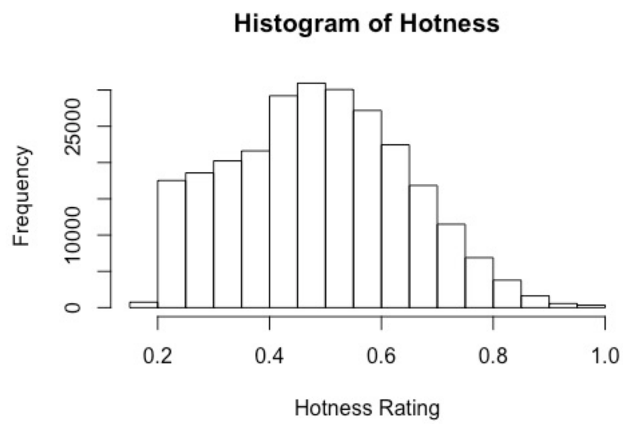


Figure 1: *This histogram shows the distribution of hotness across all songs in the dataset. The average is 0.487 with a standard deviation of 0.157.*

To analyze the economic data, Functional Data Analysis was used to smooth the data and also calculate the first derivatives of each variable. The assumption of smoothness over the same time grid for all points of each economic dataset is fulfilled by the nature of the data, as each set is continuous throughout the years 1961-2010 (Wang). The first derivatives were calculated because, in addition to looking at the trends in economic data in relation to music, it was posited that large changes in the economic data – for instance, a large increase or decrease in GDP – would be a better influencer of music preferences. These were then converted to time series to demonstrate the changes over time of the data and to compare the effects of these variables on the music (Figure 2).

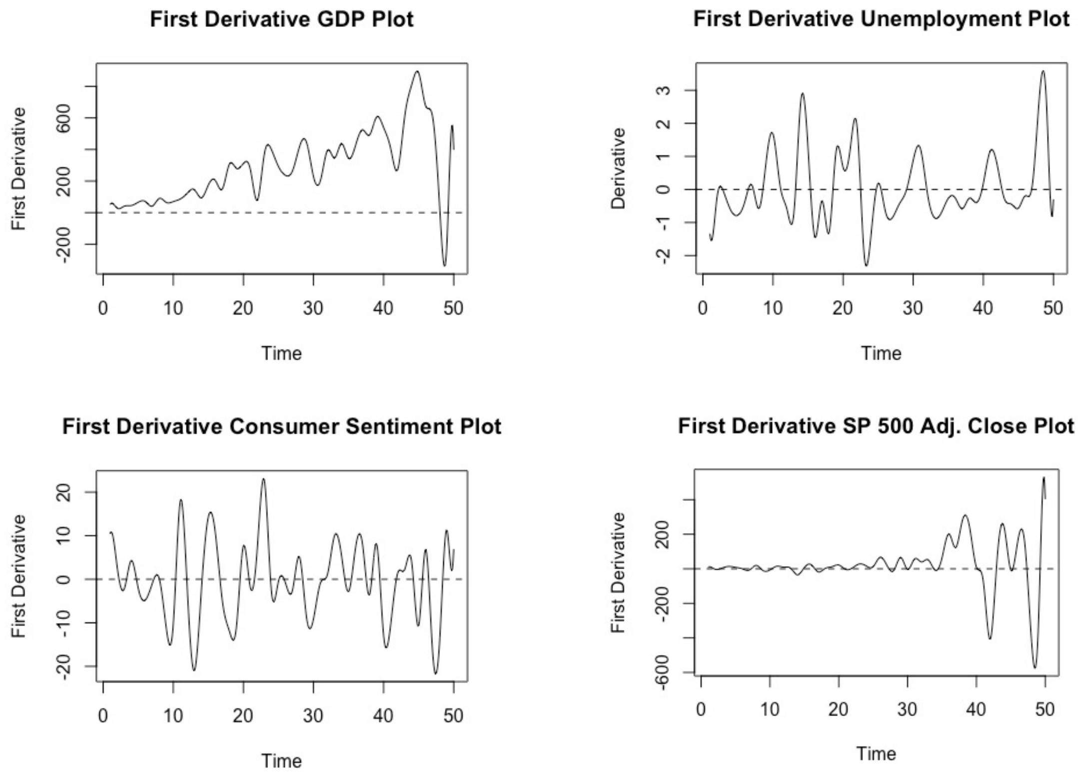


Figure 2: *These plots show the estimated derivative from functional data analysis for GDP (top left), Unemployment (top right), Consumer Sentiment (bottom left), and S&P 500 Adjusted Close (bottom right)*

Looking at overall popular songs (all songs with hotness rating greater than 0.5), there are certain trends by attribute that are noteworthy. For instance, as seen in Appendix 2, average tempo of popular music starts with some instability but levels out around 126 beats per minute. This is consistent with previous research on tempo (Ulhôa). Loudness has been trending upwards since the early 1990's, with little indication of stopping (Appendix 3). The increasing trend has been noted by a wide range of researchers and has been labeled the Loudness War, in which artists and music production companies make their music stand out by increasing the volume ("The Loudness Wars"). Similar to Loudness, Time Signature has been increasing since the 1980's (Appendix 4). This trend shows that popular music has tended towards 4 beats per measure, known as "common time" for being the most popular time signature in western music, from 3 beats per measure.

In gauging the most popular genres, a Pareto chart was created based on songs with a hotness rating over 0.5, as shown in Figure 3. Using the Pareto Principle, subsets of the music data were analyzed by genres Rock, Dance, Hip Hop/Rap, Pop, Singer/Songwriter, Jazz, and Alternative, which make up the top genres in 80% of popular music. This chart was also re-created by year to find changes in popular genres; however, the top genres did not change much – rather, they generally differed in order presented but always remained the same few genres. One can see that Rock is by far the most frequently popular genre. The wide reach of the genre and its sub-genres totaling over one hundred can explain the high frequency of Rock in the data. Stratifying the data by genres allows the investigation to pinpoint certain attributes that make a particular genre "hot", while it can also give insight into how economic factors impact genres differently as well as across all music.

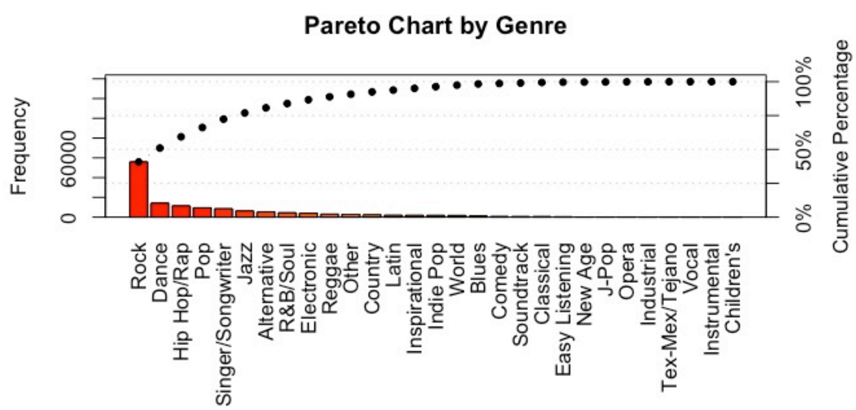


Figure 3: *The Pareto Chart shows the genres that make up the top 80% of hot music between 1961 and 2010.*

After stratifying by genre, generalized linear models were computed to find significant variables when estimating Hotness for each genre. Having found the significant variables, cross validation was then used to estimate the best polynomial combination of variables in predicting Hotness. To test the model, a training and test subset of the data were created and a generalized linear model based on the training set was used to predict possible Hotness values for the test set. Mean squared error (MSE) was then calculated to estimate the performance of these models. In general, the MSE fell in the range of 0.01 to 0.03.

An analysis of classification was used in the report as well to predict classes based on Hotness. Linear and quadratic discriminant analysis was performed to predict the class of hotness rating that a song would achieve based three classes of songs (Bad, Average, Good). Random forest classification was attempted; however, this could not be completed due to the size of the dataset and the computing power necessary for this analysis to be computed.

Lastly, each variable was converted into a time series for years 1961-2010 in order to investigate the effect of economic trends on musical preferences. With continuous variables, related pairs, absence of outliers (outside +/- 3.29 standard deviations), normality of variables, linearity, and homoscedasticity, Pearson correlation can be used in measuring the correlation between variables. Using cross-correlation of time series, the effects of GDP, unemployment, consumer sentiment, and the S&P 500 were measured against each variable that was found to be significant in the generalized linear models. Those found to have significant correlations across time generally included GDP as the economic variable in relation to another music attribute.

Granger-causality tests were used to determine a causal relationship between the economic variables and the attributes of music. The assumption that the future does not predict the past was tested through reversing the relationship in the test (Appendix 5).

Results

In first analyzing the data from popular music (songs with hotness ratings over 0.5), linear regression models identified four significant variables: Genre, Year, Time Signature, and Loudness. At first glance, plotting average Loudness and average Time Signature by year seems to reveal multicollinearity, as shown in Figure 4. However, the Pearson's product moment correlation coefficient reveals that the correlation coefficient of Loudness and Time Signature is 0.08 with a p-value less than 2.2e-16.

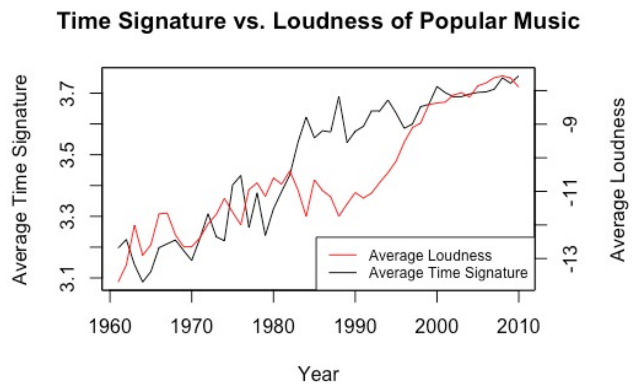


Figure 4: *This graph demonstrates the trends of Time Signature and Loudness of popular music between 1961 and 2010.*

Using nested models, an ANOVA test was performed to find the best model. The results from the ANOVA test are shown in Table 1. Though both the second and third model are significant, the second model is much more significant than the third and demonstrates that this is the best model to use when estimating Hotness for popular music.

Formula	F Score	P-value
Hotness ~ Genre (*Base model*)		
Hotness ~ Genre + Year + Loudness + TimeSignature	255.7562	< 2.2e-16
Hotness ~ Duration + KeySignature + Loudness + Mode + StartofFadeOut + Tempo + TimeSignature + Year + Genre	9.4327	5.28e-09

Table 1: *The above table outlines the models tested along with their respective F-Scores and p-values.*

To examine the prediction capabilities of this model, a training set was created consisting of 75% of the dataset and the remainder was set aside for the test set. The resulting Mean Squared Error is 0.00844, an extremely small error term.

A linear discriminant analysis classification was also performed on this data, with hotness categorized into 'Bad', 'Average', and 'Good'. This process, however, did not perform well. The error rate for the LDA is 0.5499. Though better than randomly assigning a song into one category, a process where the error term is expected to be 66%, 54.99% is not an ideal error rate and thus, LDA classification is not successful in this analysis. Similar to LDA, the quadratic discriminant analysis classification does not perform well. The error rate for this process is 0.5527.

For each genre, a generalized linear model and cross validation was performed to identify significant variables and the best combination of the polynomial to use. The significant variables for each genre are outlined in Table 2 along with the mean squared error rates of the prediction model based on the generalized linear model fit.

Genre	**Significant Variables**	**Mean Squared Error Rate**
Rock	Loudness, Year, Time Signature	0.0243
Dance	Loudness, Year, Key Signature	0.0229
Hip Hop/Rap	Loudness, Year	0.0245
Pop	Loudness, Year, Tempo	0.0258
Singer/Songwriter	Loudness, Year, Tempo	0.0224
Jazz	Loudness, Year, Time Signature	0.0199
Alternative	Loudness, Year, Time Signature	0.0196

Table 2: *The above table shows each Genre along with the variables that were shown to be significant in predicting hotness in the Genre. The Mean Squared Error of the prediction model using the significant variables is outlined in the right-most column*

As observed from Table 2, attributes Loudness and Year are significant variables in all genres when fitting for Hotness ratings, and Time Signature and Tempo appear frequently between the variables.

From the time series analysis performed on the significant variables of each genre and also the overall dataset of popular music versus economic data, it is shown that the most influential economic variable is GDP. First, with regards to the analysis with overall dataset of popular music, the first derivative estimates of GDP from performing functional data analysis are highly correlated with Loudness, Time Signature, and Tempo. Figure 5 shows the time series on the same graph, highlighting the correlation between attributes Loudness, Tempo, and Time Signature versus changes in GDP. The

Pearson's product moment correlation coefficient between average Loudness of popular music and changes in the GDP for lag one is 0.761 (p-value = 2.066e-10). The correlation coefficient between average Tempo and changes in the GDP for lag one is weaker than Loudness, at 0.312 (p-value = 0.02). Lastly, the correlation coefficient between average Time Signature and changes in the GDP for lag four is 0.742 (p-value = 3.465e-9). Thus, of these significant variables, Loudness and Time Signature have the strongest correlation with changes in the GDP.

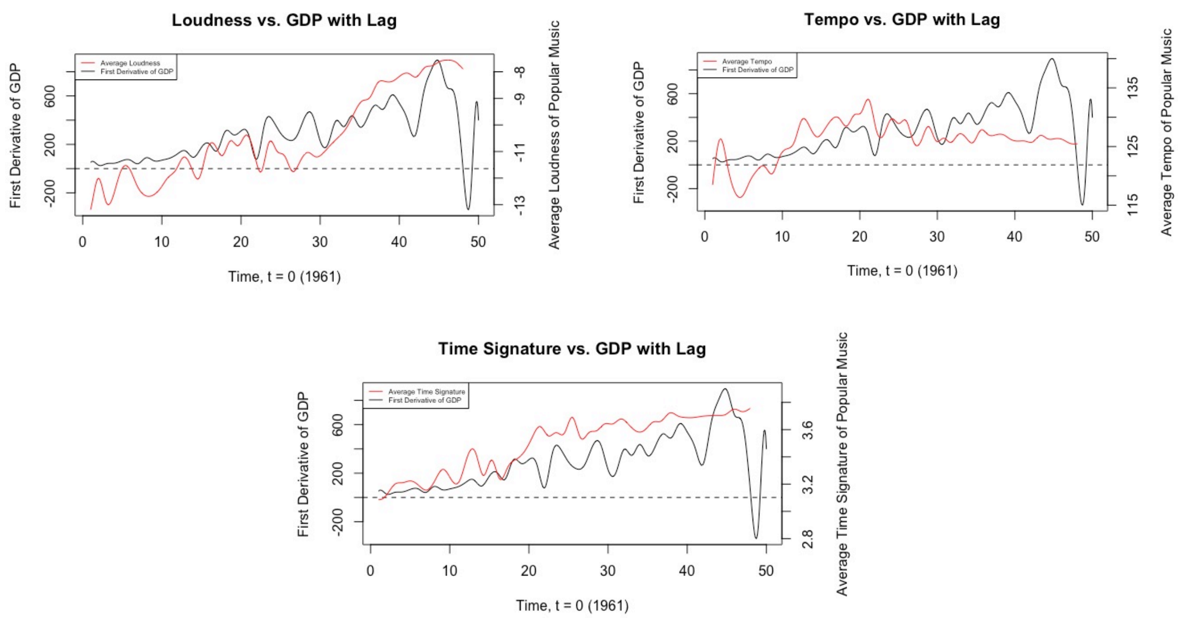


Figure 5: *The graphs above show the changes in GDP and trends of average Loudness (top left), Tempo (top right), and Time Signature (bottom) throughout time. The black line is the first derivatives of GDP while the red is the attribute of the music.*

Next, the correlation coefficients between attributes of each genre and economic variables were calculated. For the Rock genre, changes in the GDP are significantly correlated with the average hotness of the whole genre, its time signature, and its loudness. Figure 6 demonstrates these relationships by plotting, similar to Figure 4, the time series together with the lag variables included. The Pearson's product moment correlation coefficients for Hotness, Loudness, and Time Signature in relation to changes in the GDP are 0.594 with lag four (p-value = 1.31e-5), 0.726 with lag one (p-value = 3.623e-9), and 0.623 with lag one (p-value = 1.772e-6), respectively.

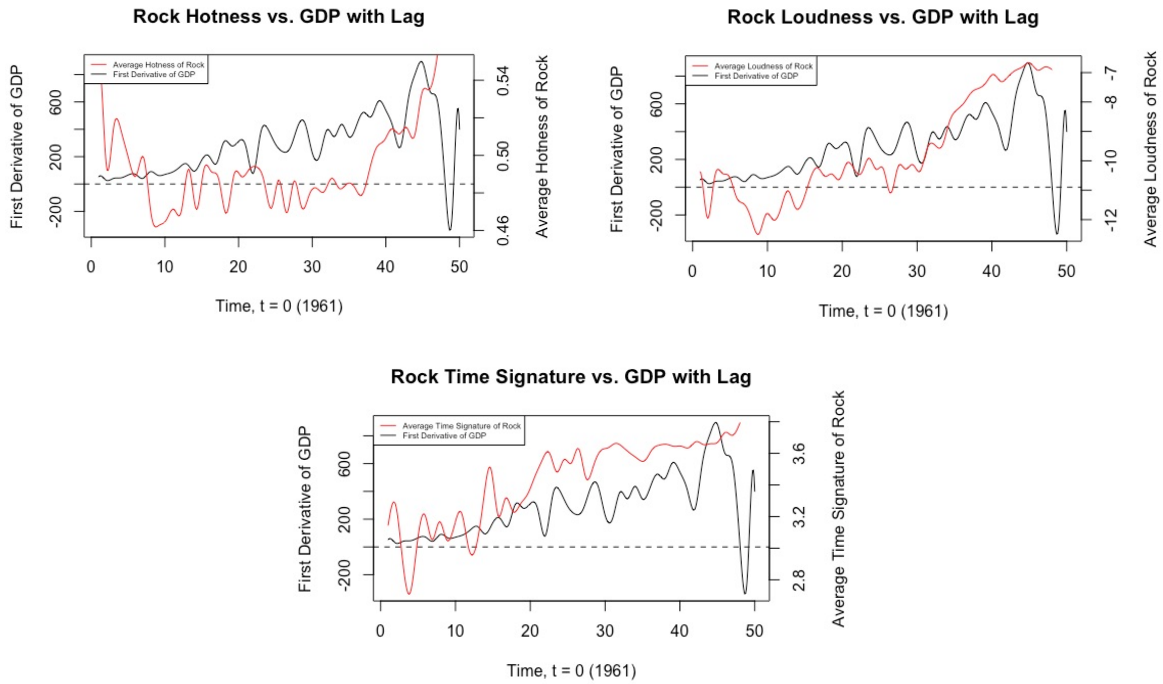


Figure 6: *These graphs show the changes in GDP and the Hotness of Rock (top left), Loudness of Rock (top right), and Time Signature of Rock (bottom) plotted over time. The black line is first derivatives of GDP while the red is property of Rock.*

Seeing that Loudness of the overall popular songs correlates so strongly with changes in the GDP, it is logical that the loudness of certain genres would also correlate strongly with this economic variable. The correlation coefficient between the average Loudness of Pop music and changes in the GDP is 0.53 with lag three (p-value = 0.0001). Between changes in the GDP and the average Loudness of Singer/Songwriter music, the coefficient is 0.664 with lag three (p-value = 8.642e05).

Moving to different variables, changes in GDP continue to have strong correlation with different attributes of popular music. For the average Hotness of the Singer/Songwriter genre, changes in GDP have a 0.484 (p-value = 0.0005) correlation coefficient with lag four. With the average Time Signature of Jazz music, changes in GDP have a 0.426 (p-value = 0.002) correlation coefficient with lag four. These two relationships with changes in GDP are depicted in Figure 7.

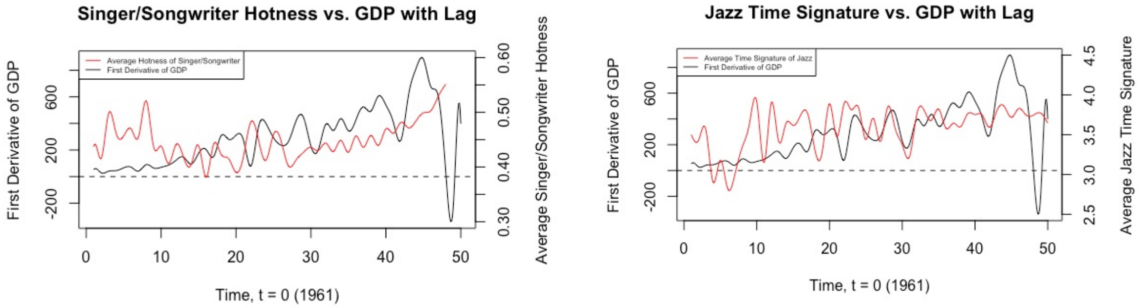


Figure 7: *The graphs above show the changes in GDP in relation to the Hotness of Singer/Songwriter genre (left) and Time Signature of Jazz (right) over time. The black line represents the first derivatives of GDP while the red represents the song attributes.*

Aside from GDP being highly correlated with attributes of popular music, the other economic variables also have significant correlations with certain aspects of this music. Though not as common, these relationships are still important to note. First, the changes in adjusted close of the S&P 500 and the average tempo of the Alternative genre have a correlation coefficient of 0.526 with lead two (p-value = 0.0003). Next, the tempo of Hip Hop/Rap has a correlation coefficient of 0.5 with lead one (p-value = 0.0049) in relation to changes in the unemployment (Figure 8).

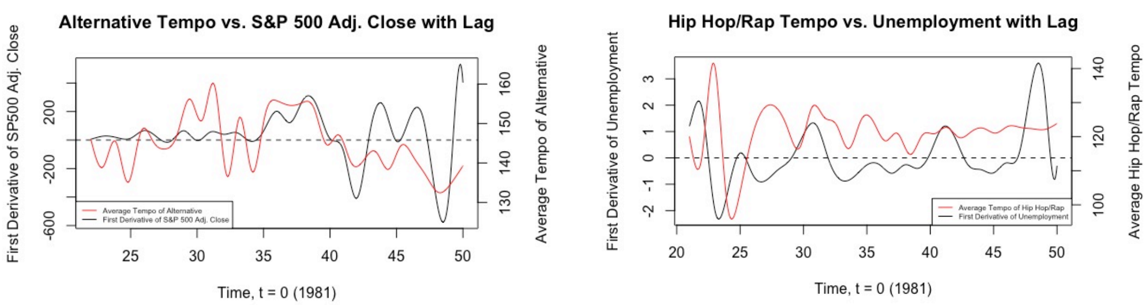


Figure 8: *The left plot shows the changes in the S&P 500 Adjusted Close and average Tempo of Alternative music across time. The right plot shows the changes in the Unemployment rate and average Tempo of Hip Hop/Rap across time.*

Lastly, rates the time signature of the Alternative genre has a correlation coefficient of -0.414 with lag two (p-value = 0.02) in relation to the changes in Consumer Sentiment (Figure 9).

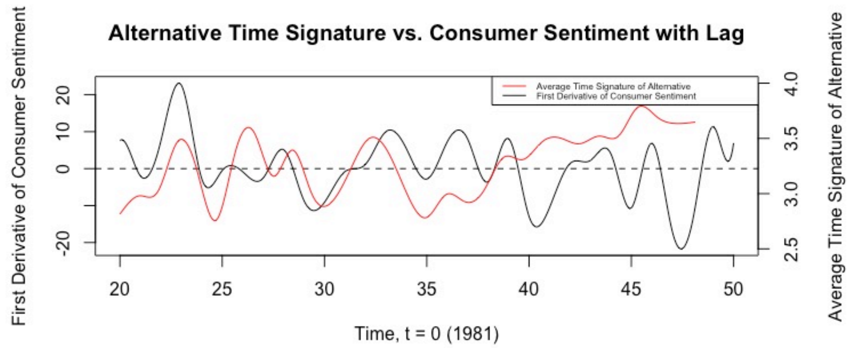


Figure 9: *This plot shows the changes in Consumer Sentiment and the Time Signature of Alternative genre across time.*

Granger causality tests were performed on each set of these variables where the correlation coefficient was significant between music attributes and the economic data. At the usual level of statistical significance (p-value < 0.05), results of the Granger-Causality test indicate that changes in the GDP with lag of one or few years help to predict the Loudness of popular music, Time Signature of popular Rock music, among a few other attributes of genres (Appendix 5). Each Granger Causality test was reversed also to test the assumption that the future does not predict the past (Appendix 5). Only the relationships between changes in GDP and Time Signature of Rock and the Loudness of Singer/Songwriter did not hold up to the assumption.

Discussion and Conclusions

There are limitations to this study and its results. First, this report does not delve into the influence of a combination of economic variables. A combination of explanatory economic variables may show to be a stronger predictor of music preferences. In addition, this study looked only at GDP, Unemployment, Consumer Sentiment, and the S&P 500. Different stock choices such as the NASDAQ or Dow Jones along with other variables like Consumer Spending may turn out to be better predictors than the variables studied here. Secondly, this investigation did not look into important cultural and historical events such as the Civil Rights Movement, the Vietnam War, September 11th, and many others. As mentioned in the introduction, such events do influence the lyrics of songs written and potentially the popularity of these songs. Thus, these variables should be studied in addition to the results reported here. Lastly, this study is based on annual data rather than monthly and the songs represented in this data are not followed through time. Though this allows the data to capture the overall hotness of the songs released in each year, it is more difficult explore the effects of the economy on musical preferences in detail. Therefore, a more encompassing dataset that follows each song through its popularity over time as well as monthly, or even weekly, data would provide for a more robust exploration of the relations studied here.

The conclusions of this study are presented in the hopes of expanding music research beyond the scopes of just music. Thus, this study explored the impacts of the economy on musical preferences and the results of the investigation suggest that the economy significantly impacts the population's preference in music. Changes in GDP are shown to be the strongest influencer, with a positive association in relation to attributes of music popular music as well as the popularity of certain genres (Rock, Singer/Songwriter) with lags between one and four. This means that positive changes in GDP are generally followed by increases in Loudness and Time Signature of popular music as well as an increase in popularity of Rock and Singer/Songwriter genres after one of a few years, and vice versa. This has widespread implications throughout music research and production. Music producers can use economic data to predict certain levels of attributes in the music in order to improve the hotness of their new music. In addition, this study's motivation is to expand music research beyond the limited view of musical attributes of a song. The results of this study show that external factors do influence music preferences and thus, demonstrates that it is valuable for the music industry to better develop prediction models that take new explanatory variables into consideration. In brief, based on the results of the analysis, this study finds that changes in the economy are good predictors of the attributes in popular music, with a lag of one or a few years.

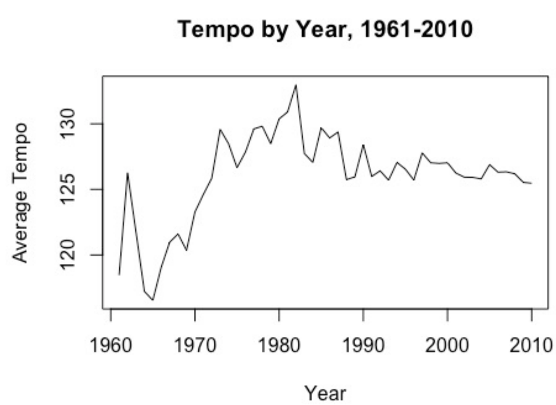
Acknowledgements
I thank The Echo Nest and Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere for providing the One Million Song dataset. I also thank Professor Edward Rothman from the Department of Statistics at the University of Michigan for his extensive support and comments.

Appendix

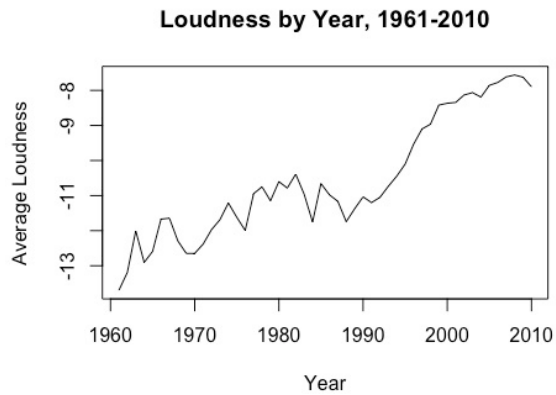
1. This is a comprehensive list of all unique genres from the dataset used in analysis.

Alternative	Anime
Blues	Children's
Classical	Comedy
Commercial	Country
Dance	Disney
Easy Listening	Electronic
Enka	French Pop
German Folk	German Pop
Fitness + Workout	Hip Hop/Rap
Holiday	Indie Pop
Industrial	Inspirational
Instrumental	J-Pop
Jazz	K-Pop
Karaoke	Kayokyoku
Latin	New Age
Opera	Pop
R&B/Soul	Reggae
Rock	Singer/Songwriter
Soundtrack	Spoken Word
Tex-Mex/Tejano	Vocal
World	Other

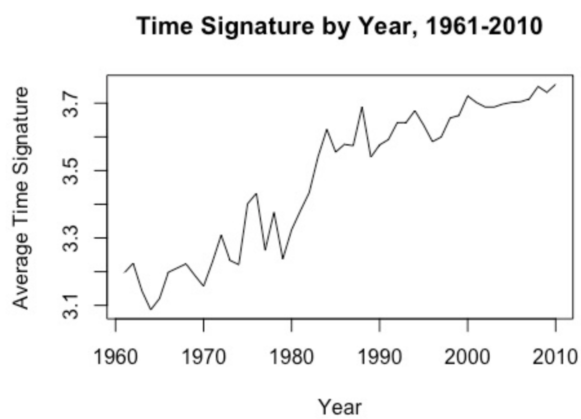
2. This figure shows the average Tempo of popular songs between 1961 and 2010



3. This figure shows the average Loudness of popular songs between 1961 and 2010



4. This figure shows the average Time Signature of popular songs between 1961 and 2010.



5. This table provides the variables tested with the Granger-Causality test that returned significant p-values.

Variables tested	P-value	Reversed relationship p-value
GDP, Loudness	0.0038	0.7
GDP, Hotness of Rock	0.0030	0.39
GDP, Time Signature of Rock	0.0061	0.004*
GDP, Loudness of Pop	0.0373	0.07
GDP, Time Signature of Jazz	0.0180	0.339
GDP, Hotness of Singer/Songwriter	0.0053	0.425
GDP, Loudness of Singer/Songwriter	0.0220	0.001*
Consumer Sentiment, Time Signature of Alternative	0.0396	0.849

References

"Bureau of Labor Statistics Data." U.S. Bureau of Labor Statistics. U.S. Bureau of Labor Statistics, 2017. Web. 17 Apr. 2017.

Khan, Amina. "The Science behind What Makes a Hit Single." Los Angeles Times. Los Angeles Times, 03 Feb. 2012. Web. 17 Apr. 2017.

Mauch, Matthias, Robert M. MacCallum, Mark Levy, and Armand M. Leroi. "The Evolution of Popular Music: USA 1960–2010." (2015): 1-10. Web. 12 Apr. 2017.

Middleton, Richard. Reading Pop: Approaches to Textual Analysis in Popular Music. New York: Oxford UP, 2003. Print.

Stutz, Colin. "The Average American Listens to Four Hours of Music Each Day." Spin. N.p., 19 June 2014. Web. 17 Apr. 2017.

Surveys of Consumers. University of Michigan, 2017. Web. 17 Apr. 2017.

"The Loudness Wars: Why Music Sounds Worse." NPR. NPR, 31 Dec. 2009. Web. 17 Apr. 2017.

Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere. The Million Song Dataset. In Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011), 2011.

Ulhôa, M., Azevedo, C., & Trotta, F. (2015). Made in Brazil: studies in popular music. New York, NY: Routledge.

Wang, J., Chiou, J., & Muller, H. (2015). Review of functional data analysis. Annual Reviews Statistics, . doi:10.1146