

Text Mining the Sentiments Behind the Grammys

Gian Carlo Di-Luvi

Business Analytics Coordinator at Pfizer Inc.

E-mail address: GianCarlo.Diluvi@pfizer.com

Music permeates our life as one of the foremost forms of human expression. From our everyday activities to the most important moments in life, music is a part of it all and, thus, there have been many attempts to quantify it in one way or another. Of the many aspects of music, lyrics are of particular relevance because most contemporary artists impose their intended message within them. It is then possible to draw meaningful conclusions from a song by analyzing its lyrics. In this paper, I use text mining and sentiment analysis to look at statistical aspects of lyrics of popular songs from 1960 onwards.

Quantifying music

Objectively studying music is a difficult task at best. Ideally, one would have to analyze all the music released on a given year or years, all around the world. One way to simplify this problem is narrowing down the analysis to only those musical creations that have been held as exceptional. Thankfully, a large number of musical awards are constantly given to artists, producers, composers, and all people involved in the music-making process. These highlight and recognize precisely those musicians that have created outstanding musical pieces. With this in mind, I narrowed down the analysis to winners of one particular award: The Grammy Award for Album of the Year.

The Grammy Awards, named in honor of the gramophone, are arguably one of the most famous and prestigious musical awards. They originated in the late 1950's, when music executives decided to create an award equivalent to the Academy Awards (for film) and the Emmy Awards (for television). Since 1959, Grammy Awards are given every year by The Recording Academy for different categories, such as Record of the Year, Song of the Year, Best New Artist, and Album of the Year. The last one is of special interest because it is awarded based on the quality of the album as a whole, and not on the merit of individual songs.

For this analysis, and with the help of the `geniusR` package in R (Parry, 2018), I obtained from [Genius](#), an online song lyrics repository, the lyrics in English of all the songs of each album that won the Grammy Award for Album of the Year from 1960 to 2018. It should be mentioned that four albums, all from the 60's, were omitted from the analysis: Bob Newhart's 1961 *The Button-Down Mind of Bob Newhart* and Vaughn Meader's 1963 *The First Family* are comedy albums; Judy Garland's 1962 *Judy at Carnegie Hall* is a live album; and Stan Getz and João Gilberto's 1964 *Getz/Gilberto* has lyrics in Portuguese.

After joining the lyrics of the 55 albums—comprising 746 songs—in a single data set, it was possible to start analyzing the data via *text mining*, a set of techniques whose purpose is to process, transform, and study information in the form of text so that statistical conclusions can be derived from it. For instance, the average number of songs per album throughout this period is roughly 14, with the 80s having a low 11.4 and the 2000s having the largest at more than 16 (mainly due to Outkast’s 2003 *Speakerboxxx/The Love Below*, which featured 40 songs).

There are exactly 179,381 words in my data set, averaging around 240 per song. The 60s are by far the decade with the lowest average word count per song, with 160, versus the much higher 302 from 2010 to 2018. Figure 1 shows how there seems to be a slight increase in the number of words per song over time.

Figure 2 shows the ten most common words in all those songs, without taking into account the usual English stop words such as ‘the’, ‘of’, ‘to’, etc. (Stop words can easily be removed from a tidy data set with the `tidytext` R package (Silge and Robinson, 2017).)

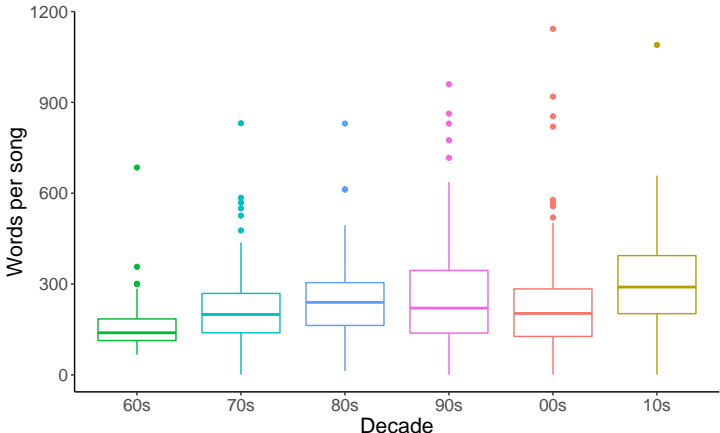


Figure 1: Boxplot of the number of words per song by decade of the Grammy Albums of the Year, 1960–2018.

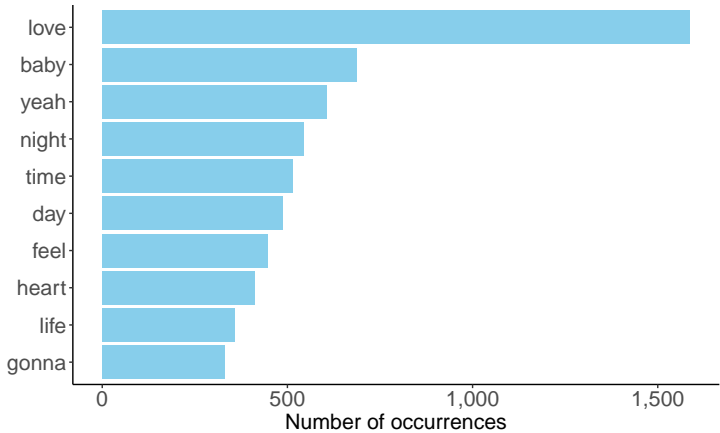


Figure 2: The ten most common words of the songs of all the Grammy Albums of the Year, 1960–2018.

Amusingly, although perhaps not surprisingly, ‘love’ is the most common word across all songs. Moreover, this is also true if one looks at each individual decade, as can be seen in Figure 3. This last

Figure also shows that ‘baby’ has been the second most popular word for the past three decades.

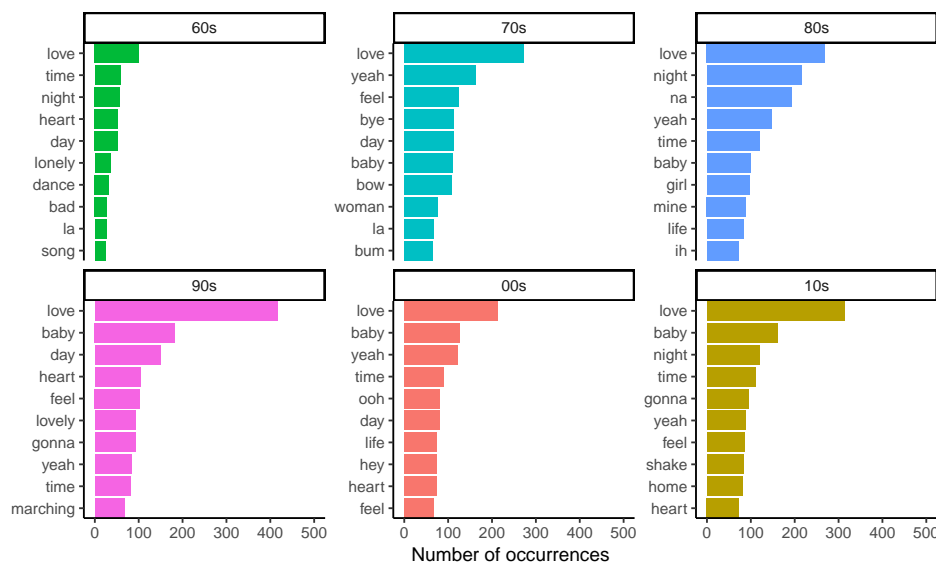


Figure 3: The ten most common words of the songs of all the Grammy Albums of the Year, 1960–2018, by decade.

Something interesting in Figure 3 is the appearance of rather odd words in some decades. For example, the popularity of the word ‘la’ in the 60s is due to Barbra Streisand’s 1964 interpretation of “Who’s Afraid of the Big Bad Wolf” (“...who’s afraid of the big bad wolf / tra la la la la...”), while in the 70s it mostly appears in Simon and Garfunkel’s 1971 “The Boxer.” The word ‘na’ appears more than 100 times in the unreleased version of Paul Simon’s 1987 “Diamonds on the Soles of Her Shoes,” and more than 30 times in the same album’s “You Can Call Me Al,” as well as in Michael Jackson’s 1984 “P.Y.T. (Pretty Young Thing),” making it the third most common word of the 80s. That same decade the tenth most common word was ‘ih’, which refers to the gasping sounds of Paul Simon’s 1987 “Homeless.” Similarly, ‘bum’ appears on the chorus of Stevie Wonder’s 1975 “Smile Please” (“Bum, bum bum ditti / Bum, bum bum ditti...”). Finally, ‘ooh’ appears a large amount of times in both OutKast’s 2004 “Roses” (“Roses really smell like poo-ooh-ooh...”) and in the chorus of U2’s 2006 “City of Blinding Lights.” (It can, of course, be argued that none of these expressions are real words.)

The sentiments of sentimentality

Just by looking at the most common words by decade, it is possible to try to guess which decades were, say, happier than others. For example, ‘lonely’ and ‘bad’ (words with a negative connotation) were prominent words in the 60s, whereas ‘baby’, ‘heart’, ‘lovely’, and ‘yeah’ (words with a more positive connotation) were more common in the 90s. This may lead us to hypothesize that the 90s had a lighter mood than the 60s.

The reasoning behind that example is the basis for *sentiment analysis*, a series of techniques that help quantify emotions or affective states. Human writing is necessarily loaded with emotional intent in its words, and part of successful reading is the correct interpretation of those sentiments. Sentiment analysis aims to reproduce this mental process programmatically, using objective methods (without

forgetting that, in the end, what is being studied is subjective by nature).

There exist a variety of ways of doing this. The most popular ones rely on assigning sentiments to individual words, and then obtaining the net sentiment of a text by adding up the sentiments of the words that make it up. These so-called vocabulary lexicons contain a large number of English words with their respective sentiment (or sentiments) and are usually obtained and verified by hand.

For this analysis I used the Liu et al. (2004) lexicon, which categorizes words as either ‘positive’ or ‘negative’ in a binary fashion. A ‘positive’ and a ‘negative’ score are computed by counting the number of words within each category, and the net sentiment is obtained by subtracting the latter from the former. Liu’s lexicon can be found in the `sentiments` data set from the `tidytext` R package.

It is worth noting some implications of this approach. First, many English words are omitted from Liu’s lexicon because they are relatively neutral, which makes it difficult to assign a sentiment to them. Second, one may also argue that there exist multiple levels of positive and negative connotations, that is, not all positive words are equally positive. Although other lexicons do take this fact into account, it may also be argued that, if Liu’s lexicon is subjective as it is, adding more levels will only increase that subjectivity. Thus, using another lexicon does not bring enough benefits to have a dramatic impact on the quality of the analysis. All in all, lexicon-based sentiment analysis offers an easy approach to analyzing emotions in text, although its limitations must be considered.

With that in mind, I applied Liu’s lexicon to every song in the data set to generate a net sentiment score by album, which is shown in Figure 4. From this graph we can see that the first, second, and fourth albums with the greatest net sentiment belong to the 90s: Whitney Houston’s 1994 *The Bodyguard: Original Soundtrack Album*, Natalie Cole’s 1992 *Unforgettable... with Love*, and Celine Dion’s 1997 *Falling into You*, respectively. Moreover, the 90s are the only decade in which no album has a negative net sentiment score. The 2010s are also an interesting decade in that they have the album with the third highest net sentiment (Bruno Mars’ 2018 *24K Magic*), and also the one with the lowest one (Mumford & Sons’ 2013 *Babel*).

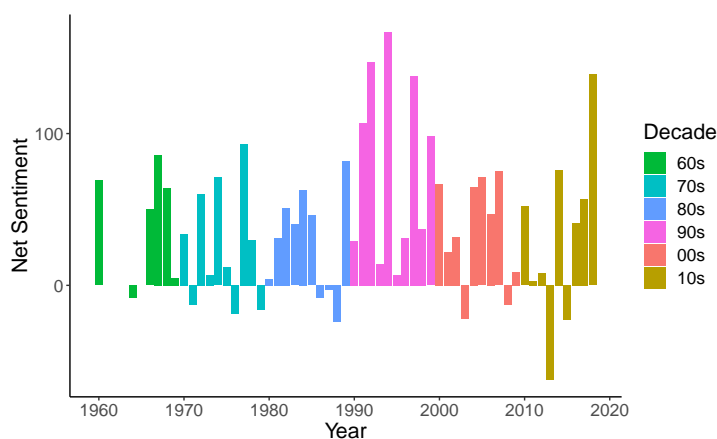


Figure 4: Net sentiment scores of the Grammy Albums of the Year, 1960–2018. Missing values in the 60s correspond to the 4 albums omitted from the analysis.

One could then try to study the evolution of the Grammy Albums of the Year’s net sentiments over time. One way to do this is by explaining the relationship between the year and the net sentiment with a *linear regression*. This model associates the net sentiment of an album that won the Grammy Award with the year it was released via a linear equation, that is,

$$\text{Net sentiment} = a + b \times \text{year},$$

where both a and b are unknown parameters to be estimated with the information of the albums that have won in previous years. The parameter b quantifies how much the net sentiment changes from one year to the next, and it has to be statistically different from zero for the model to be significant. However, the estimated value obtained from the data ($\hat{b} = 0.125$) could not be determined to be different from zero. Figure 5 shows the data points with the estimated linear equation in red, where it is clear that there does not seem to be any meaningful relationship between the net sentiment and the year. This means that the net sentiment of Grammy Award winners, although certainly higher or lower some years, has remained relatively stable overall.

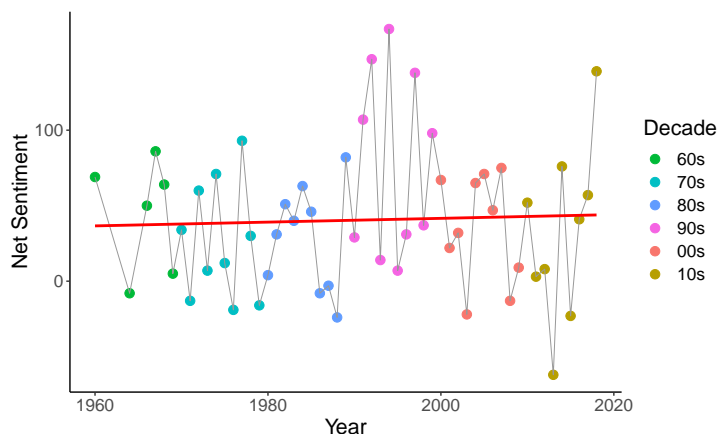


Figure 5: Net sentiment of the Grammy Albums of the Year, 1960–2018, with a linear model fit in red.

It should also be noted that, in order for this type of model to be valid, some assumptions need to occur: net sentiments across different years should have equal variance, follow a Normal distribution, and have no correlation between them. The validity of these assumptions was ascertained using common graphical methods. (Particularly using a residual vs fitted plot, a Q-Q plot, and an autocorrelation plot, respectively.)

Notably, the 90s do appear to have albums with a greater net sentiment than the 60s’—evidence that appears to favor our previous hypothesis. (However, the 60s do not seem to be a decade with particularly low net sentiments, unlike the 2000s, for example.) In order to further confirm this, I obtained the average net sentiment of all the albums by decade, which is shown in Figure 6.

By doing this it becomes evident that the 90s was the decade with the highest average net sentiment, with the 60s coming in second. Interestingly, average net sentiment in other decades seems to vary around 30, making the 60s and 90s much higher than average, with net sentiments of 44 and 78, respectively.

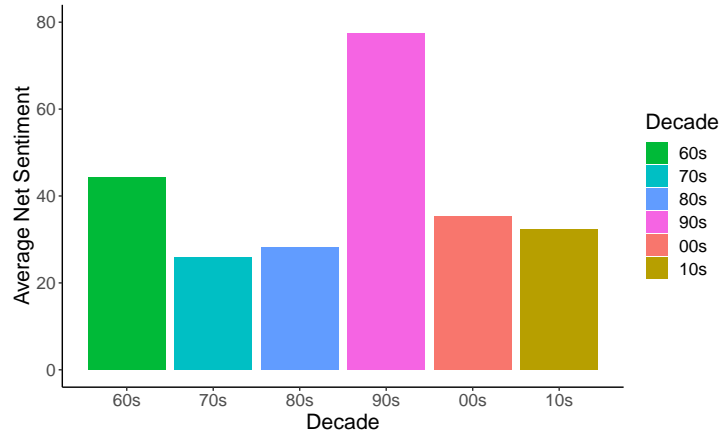


Figure 6: Average net sentiment scores of the Grammy Albums of the Year, 1960–2018, by decade.

And what about the losers?

Albums that win the Grammy Award for Album of the Year are clearly outstanding, but it can be argued that merely being nominated is an honor in and of itself: albums that were nominated to the Grammy but did not win are also incredibly relevant. Thus, for every album in my original data set, I obtained the net sentiment of the albums that were also nominated that year but did not win (excluding live or comedy albums, albums whose lyrics were not in English, and classical music albums). This way I had the net sentiments of 55 albums that won and 209 albums that were only nominated.

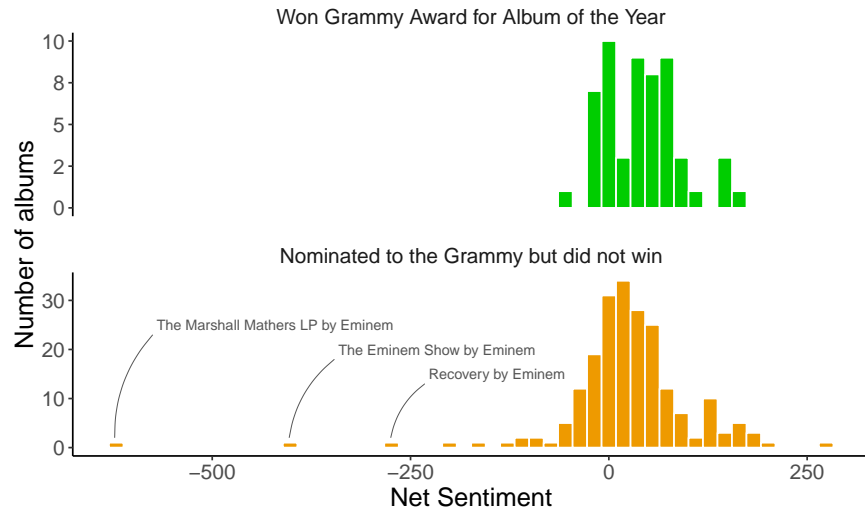


Figure 7: Histograms of the net sentiments of the albums that won the Grammy Award for Album of the Year (green) and those that were nominated but did not win (orange), 1960–2018.

Winners and nominees tend to lie in the same range: the former’s first and third quartiles are 6 and 68, respectively, versus -5 and 58 for the latter. Interestingly, both winners and nominees have almost the same *interquartile range* (IQR, the difference between the third and first quartiles), although the nominees’ IQR is shifted ten units to the left. This means that the nominees’ net sentiments tend to be lower than the winners’, which is corroborated in Figure 7. This graph also shows that there are no

winners with extreme sentiments, unlike some nominees. Of special interest are the three nominated albums with the lowest net sentiment. All correspond to the same artist, Eminem, one of the most renowned rappers of all time, whose music often deals with controversial and dark themes.

All in all, it seems that the only constant that can be observed is that Grammy Albums of the Year always seem to talk about love, whether the general tone is happier, as in the 1990s, or darker, as in the 2010s. This need not be particular to the Grammys: music in general often serves as an escape from reality in an age where, due to technology, we are constantly under a barrage of news and information. It offers a space in which artists express themselves and people from all around the globe can liken that music to their own actuality. Because of this, we will most likely never be able to completely grasp the complexities behind this subjective interchange. More so, maybe that is precisely the point of music. However, just by taking a data-driven glance at it, we can find all sorts of patterns and curiosities that, although falling short of giving us a full understanding, do offer a deeper look at it than what meets the eye.

References

- Bing Liu et al. Opinion lexicon: A list of english positive and negative opinion words or sentiment words., 2004. URL <https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html>.
- Josiah Parry. *Quickstart: geniusR*. URL github.com/JosiahParry/geniusR.
- Julia Silge and David Robinson. *Text Mining with R: a Tidy Approach*. O'Reilly Media, 1 edition, 2017.