

Music Genre Sentiment Analysis

Julie Bazalewski

Summary

In the 2019 Netflix documentary *Garth Brooks: The Road I'm On*, Brooks states “country music I don’t think has ever been treated with the same respect as pop, rock...” and that “in country music... you’re competing against the other forms of music”.

A sentiment analysis of each genre was performed to determine if country music is viewed more negatively by Twitter users compared to rock and pop music, and how sentiment differs among the three genres. Knowing this information could help labels or artists determine how their genre compares to the others and could use this information to become more profitable and/or appealing to consumers.

Data Collection

Approximately 1000 tweets for each genre were collected by using the REST API from the Tweepy python package at one time. The hashtags #countrymusic, #rockmusic, and #popmusic were used to represent each genre. After collecting the data, python was used to clean the tweets, and the TextBlob package was used to determine a sentiment score for each tweet. The scores range from -1 to +1, with -1 being the most negative. Finally, the tweets were exported to a .csv file for further analysis in R. The full tweet text, usernames, user locations, and sentiment scores were saved.

Analysis

After importing the .csv data into R, a χ^2 test for independence was performed to determine whether each genre is equally likely to have negative sentiment (H_0) or if at least one genre of music has a different possibility of having negative sentiment compared to the other music types (H_1). The sentiment scores were grouped into positive, negative, and neutral sentiments for the test. Two-sample one-tail T-tests were also performed on each combination of genres using the sentiment scores rather than the categorical sentiments to determine which groups differed.

Findings and Conclusions

A p-value of 0.01124 from the χ^2 test for independence suggests there is enough evidence to claim that there is an association between genre and sentiment at a significance level of 0.05. As seen in the box plot below (Figure 1), most of the tweets have neutral to slightly positive sentiment. The t-tests indicate that there is evidence at a significance level of $\alpha = 0.01$ that the true sentiment of the pop tweets is greater than the sentiment of the rock tweets. I also found that there is evidence at a significance level of $\alpha = 0.05$ that the true sentiment of the country tweets is greater than the sentiment of the rock tweets. There is not sufficient evidence that the mean of the pop and country tweets differ significantly. From this analysis, it can be concluded that the sentiment of country tweets is not more negative than the other genres. Rock music has the most negative sentiment of the three categories.

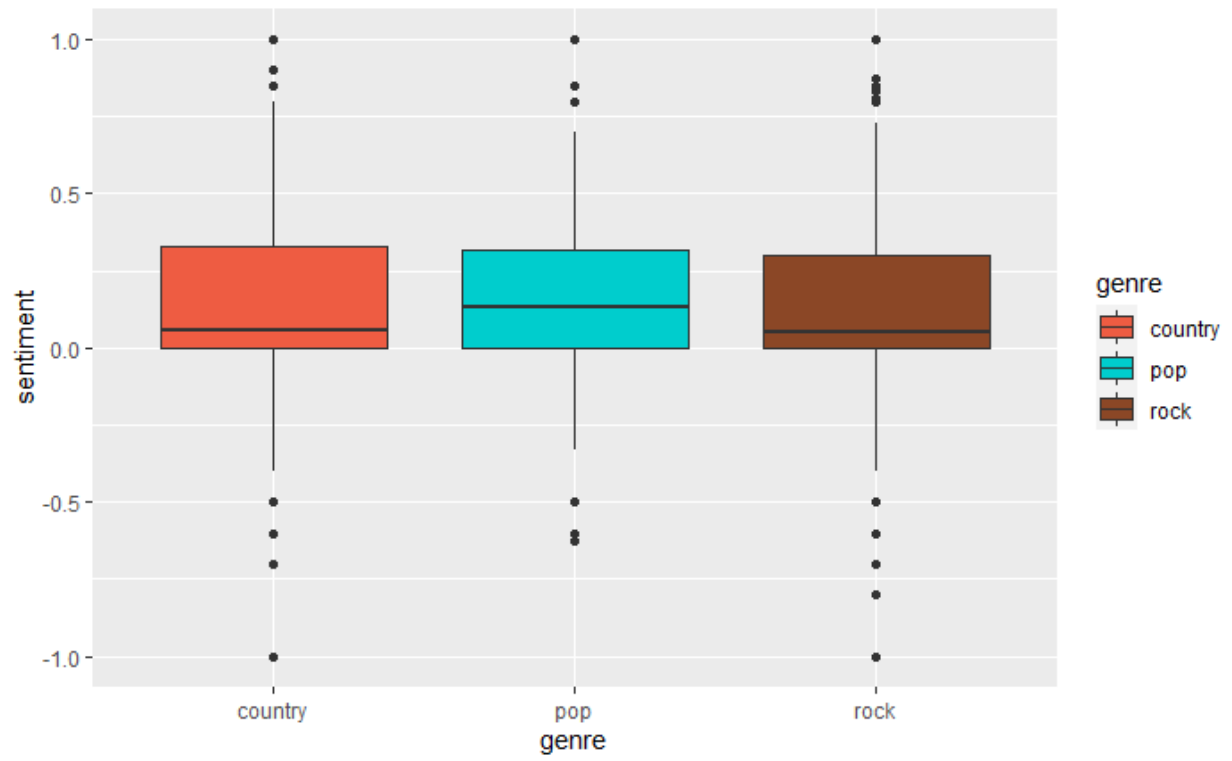


Figure 1: Boxplot of Sentiment Scores by Genre