DS 710 Final Project

Twitter Analysis of Musical Genre Popularity

**Project Summary**

The power of social media has displayed its significance during the digital age. Twitter certainly stands amongst the most prominent social media platforms, especially in terms of identifying trends and popularity. I chose to evaluate the popularity of musical genres based on their perceived popularity on Twitter. Popularity is determined in two ways. First, by overall tweet count then by the number of exclamation marks used within those tweets.

This information is particularly interesting for those in the music industry, whether it be a record company choosing which artist to sign next, or a music store planning their in-store marketing strategies. By identifying popularity of music by genre, the music industry can develop a more inclusive roadmap for the future in order to maximize profit.

**Data Collection**

Tweets were collected using the REST API through Python. The program grabbed (up to 5000) tweets that were categorized into genres. I created five general genres, spilt into sub-genres for gathering tweets. The tweets were collected over a seven-day period which is warranted by the REST API. The genres evaluated were Pop, Country, Rock, Blues, and Dance. I included a series of hashtags that could be related to or signify subsets of each genre (i.e., Rock Music: #rockmusic, #metalmusic, #punkmusic).

Data collection provided two useful pieces of information. First, I counted the number of tweets that included one of the aforementioned hashtags for each genre, up to 5000 tweets, during the past seven days. The overall tweet count can be an indication of popularity (Figure 1). Secondly, I used Python to segregate the text of each tweet, and count for exclamation marks, compiled by genre, to determine how excited people were, while tweeting about a particular genre, also indicating popularity (Figure 2).

**Analysis**

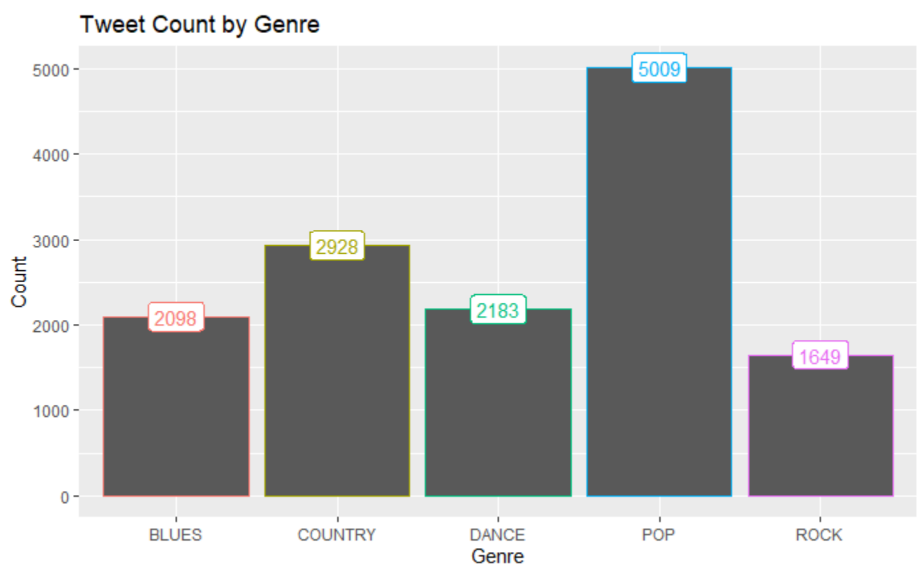
Two methods were used (in R Studio) in order to evaluate the relationship between tweet count and exclamation mark count by genre.

First, I evaluated the potential for a linear relationship between both variables. The variables were plotted against one another in a scatterplot (Figure 3), and a summary of the linear model was created.

Secondly, the variables (Tweet Count & Exclamation Count) were compared in a Chi Squared Test of Independence.

**Conclusion**

Both analyses indicate that there is likely no relationship between the number of tweets mentioning a specific genre and the amount of exclamation marks within those tweets. It is important to mention that popularity is assumed by Tweet count, therefore; popularity cannot be determined by the number of exclamation marks within tweets mentioning specific musical genres.

Figure 1

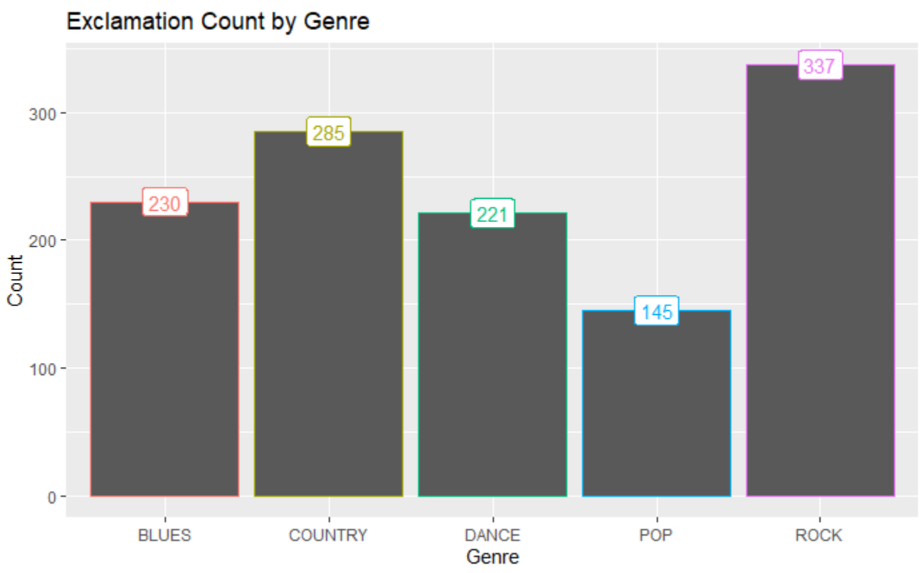
Figure 2

Figure 3

