Evaluating Patterns in Critically Acclaimed Music

Paul Bartlett (# *250753008*)
The University Of Western Ontario

PBARTLE7@UWO.CA

Abstract

The purpose of this analysis is to identify relationships between musical genre of critically acclaimed albums and time. The dataset used for this analysis contains over 18,000 reviews from Pitchfork from January 5th, 1999 to January 8th, 2017. It contains important data including release year, artist name, genre, and a score ranging from 0.0-10.0. The findings may be useful for determining what the most successful genre of critically acclaimed music is for each of the last 18 years.

1. Description of Applied Problem

1.1. Existing solutions to similar problems

The trends of popular music can easily be attained through the various Billboard charts that have existed since 1955. A group of scientists from the University of London analysed around 17,000 songs that charted on the U.S. Billboard Hot 100 over the last 50 years and created a visualization of the popularity of musical genres over time (Matthias Mauch and Leroi, 2015a). The problem with getting data from these charts is that popular music generally isn't critically acclaimed, and is therefore not as interesting as data from sources that evaluate music more objectively. Another source that uses visualization of this problem well is musicmap (Crauwels, 2016). The website contains information about hundreds of genres of music and their history. It provides a great overview of all the popular strands of music, but doesn't go into too much depth about specific artists or albums. It does provide a good overview of all genres regardless of popularity, but I'm more interested in evaluating the history of the best albums created by artists.

1.2. Pitchfork solution

Using a dataset that includes over 18,000 reviews from Pitchfork, I will be going through the data to find how critically acclaimed music has changed over time. In addition, I will also be looking at which release from an artist is the most well received. A general pattern I've seen when listening to several albums from an artist is that the first 2-3 albums tend to be the best from their discography. I would like to use the data to confirm or deny this assumption.

2. Description of Available Data

2.1. Pitchfork

The dataset that I will be using is taken from Pitchfork. Pitchfork is an online magazine that focuses on reviewing both popular and independant music. The data set for Pitchfork Reviews from January 5th, 1999 to January 8th, 2017 is available on kaggle (Conaway, 2017). There are 18,393 reviews that include important data including release year, artist name, genre, and a score ranging from 0.0-10.0. Considering that Pitchfork is one of the longest running online review sites, it makes it a primary choice for useful data.

2.2. Best New Music

Pitchfork features a section called "Best New Music" for highlighting recent releases that the staff found to stand out in a positive way. These albums generally have a minimum score of 8.0 and are another useful way for tracking the best music on the site. The dataset includes an identifier for "Best New Music", and would be a useful way to sort through the data.

3. Plan for Analysis and Visualization

3.1. Analysis

To analyse this data properly, we must use a method to take out only the best reviews from the dataset. Fortunately, Pitchfork has a system to distinguish the best albums called "Best New Music". Unfortunately,

this feature launched in 2003 so using it would leave out all the music before it was launched. By looking at the data, we will be able to find the typical rating for an album that gets the "Best New Music" tag and use that rating to take all albums from the data set that are higher than the threshold. From this, we should be able to classify each album that meets the requirement by year and genre so that it can be used for visualization. To analyse what release number is considered to be the best, we will have to query all artists that have multiple releases on the site. A similar analysis and visualization was done on kaggle by the author of the data (**?**).

3.2. Visualization

For visualization, I would like to do something similar to what was done for visualizing the U.S. Billboard Hot 100 over the last 50 years (Matthias Mauch and Leroi, 2015b). The chart features spindles for each genre that run vertically with the width proportional to the frequency of the style. The y-axis contains the year so the viewer can easily compare between each year to see what genre is the most popular or the least popular. Since I do not have much experience with visualization, it is possible that doing something similar will be too difficult to achieve. A simple way to visualize this in a similar way would be to use a line graph, with each line representing a genre, the x-axis covering each year, and the y-axis covering the frequency.

References

Nolan Conaway. 18,393 pitchfork reviews, 2017a. URL https://www.kaggle.com/nolanbconaway/pitchfork-data.

Nolan Conaway. Is the first album the best?, 2017b. URL https://www.kaggle.com/nolanbconaway/is-the-first-album-the-best.

Kwinten Crauwels. musicmap, 2016. URL https://musicmap.info/.

Mark Levy Matthias Mauch, Robert M. MacCallum and Armand M. Leroi. The evolution of popular music: Usa 1960–2010. 2015a. URL http://pubmedcentralcanada.ca/pmcc/articles/PMC4453253/.

Mark Levy Matthias Mauch, Robert M. MacCallum and Armand M. Leroi. Figure 3., 2015b. URL http://pubmedcentralcanada.ca/pmcc/articles/PMC4453253/figure/RSOS150081F3/.