ISSN 2618-1568 Vol. 3, Issue 2: 55-59, DOI: 10.25236/FAR.2021.030211

Research on Music Characteristics and Artist Influence

Shiyi Hao

School of Science, China University of Geosciences, Beijing 100083, China

Abstract: *Nowadays, music has become an inseparable part of our lives, and studying the history of music can also helps us analyze the historical environment of the society at that time. To study the history of music is to analysis the characteristics and vocals of music, as well as data on the ability of influence among artists. Our paper attempts to establish a model that can analyze and evaluate the influence and similarity of musicians and their works to further study the history of music. We use cosine correlation to analyze the correlation of all music and music characteristics. Then we exclude a few sets of data that have little to do with popularity, and perform dimensionality reduction of the data. Then the cosine similarity is used to calculate the similarity of different genres, and the similarity between the internal genres and the artists represented by the genres. Finally, it is concluded that artists are more similar in genres. Moreover, we found through calculations: Within the genre, most of the characteristics have not changed much, but we can see their small changing trends. For example, the changing trend within the rock genre is changing from focusing on lyrics to focusing on musical melody. Another finding is that the artist is not necessarily in the same genre as the artists he follows, but the similarity of the works of the artist and the similarity of the artists he follows are basically relatively high. It can be seen that the artist with a higher influence score is he Generally speaking, the works of "infectiousness" are relatively high. In addition, we also verify the stability by calculating the variance of the data in the similarity model, and the results obtained are between 0.09 and 0.20, which proves that the stability of our model is relatively high.*

Keywords: *Network Science, Music, Cosine Similarity, Influence*

1. Introduction

Music is not only a very important part of human society, but also an important part of human cultural heritage. Music injects vitality into people's life and also gives them the enjoyment of beauty. It combines a variety of emotions that are often more powerful than words. There are interactions between different musicians, between different music genres, and within the same music genres, even external events (such as major world events or technological advances) can influence changes in music. So that we can gain a better understanding of how music evolves through societies over time, examine evolutionary and revolutionary trends of artists and genres, it is significant to build a model to measure the influence of music.

2. Model Establishment and Solution

2.1 Establish a directed network

Using toolkits such as pandas numpy csv textblob nltk in Python software, we counted two statistics for each musician: the number of people he influenced and the number of people who influenced him.Through these data, we can grasp the influence ability of musicians, here's an example by mapping the relationships of four of these musicians (Figure 1). The middle box contains: musician_name, musician_active_start and musician_main_genre.

ISSN 2618-1568 Vol. 3, Issue 2: 55-59, DOI: 10.25236/FAR.2021.030211

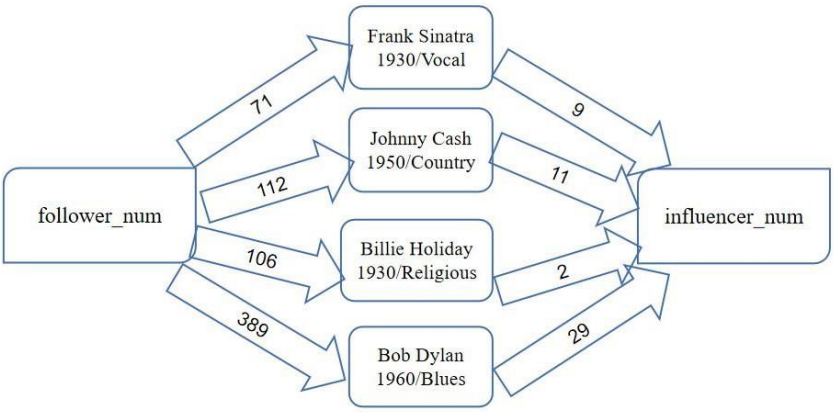

Figure 1: A diagram of four musicians

In the network we've built, we've found that if an artist has a large following, he or she can influence the trends of the day with his or her music style.

At the same time, we've mapped the proportions of musicians from different eras, this is shown in Figure 2.

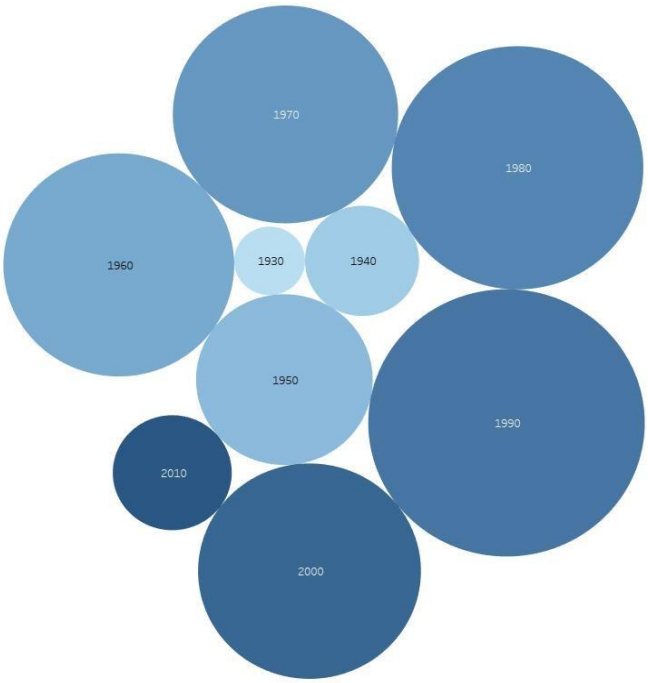

Figure 2: Musicians of different ages appear proportional filling bubble graph

2.2 The parameters of the music influence

We found that the popularity data in the data_by_artist document is a good way to describe the popularity of musicians among the general public, so we looked at the degree of influence between these figures.After consulting relevant materials, we finally believe that the evaluation standard of musicians' influence on musician accounts for 40 percent and musicians' influence on the public accounts for 60 percent is more reasonable.

We didn't know at first how to set a score among musicians, Later, we tried to study the distribution of the number of musicians who influenced each musician and decided to divide their influence on musicians into six levels of 0-5 stars in roughly normal distribution, The specific classification is shown in Table 1, of these, 615 were the most influential musicians in the number of influences.

ISSN 2618-1568 Vol. 3, Issue 2: 55-59, DOI: 10.25236/FAR.2021.030211

Table 1: 0 to 5 star rating

stars	follower_num
0	0—10
1	10—30
2	30—50
3	50—80
4	80—150
5	150—615

We set the total score at 100, the popularity figure is the highest out of 100, the calculation needs to be multiplied by the weight of 0.6. The stars data needs to be multiplied by the total score of 100, multiplied by the weight of 0.4, and then divided by 5 as the final score of this part. That way, we can get a formula to calculate a musician's influence score(Formula 1)

$$s = 0.6x_1 + \frac{0.4*100}{5}x_2 = 0.6x_1 + 8x_2 \tag{1}$$

2.3 Generation pop subnetworks

In order to study the influence of music style on the trend of The Times more carefully, we use Tableau software to visualize the data, this is shown in Figure 3.

Take Bob Dylan as an example. His style is the Blues. When he first started playing (1960), the popularity of country music increased significantly.

In our model, musical influence refers to the ability of an artist's work to influence people, in other words, how many people are influenced by him and how many people like his work. The macro reflection is the popularity of this type of music at that time.

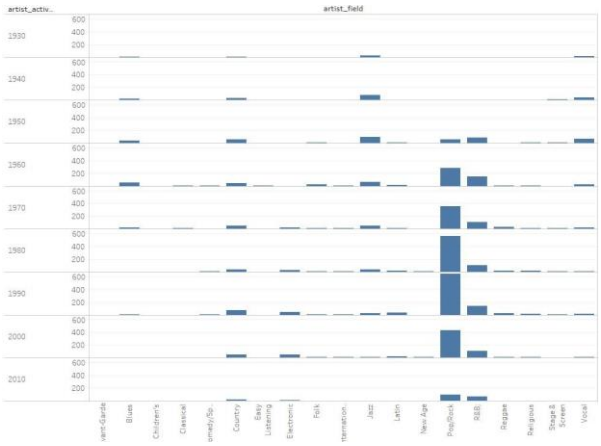

Figure 3: Visual presentation of data

2.4 Cosine similarity

We first analyzed the relationship between these musical characteristics and popularity (data from data_by_artist). Excluding a few values with low correlations, this is a data dimension reduction operation. The correlation analysis is to put these numbers in the same order of magnitude as popularity, then ask for the average after all the data is subtracted.(First we excluded two features: mode and key, because they're hard to quantify.)

And by calculating that we exclude tempo, loudness, instrumentalness, speechiness and duration_ms. Left danceability, energy, valence, acousticness and liveness to study.

We use the cosine similarity method to calculate the similarity of different genres and different ages. In other words, we measure the similarity between two vectors by measuring the cosine of the Angle between them. As shown in Formula 2.

ISSN 2618-1568 Vol. 3, Issue 2: 55-59, DOI: 10.25236/FAR.2021.030211

$$\cos\theta = \frac{\sum_{i=1}^{5} A_i * B_i}{\sqrt{\sum_{i=1}^{5} (A_i)^2} * \sqrt{\sum_{i=1}^{5} (B_i)^2}} \tag{2}$$

Through calculation, we finally get the conclusion: Most artists are similar to their own genres of music, there are certain similarities between some of the genres(Such as Folk and Country, Electronic and Rock).

2.5 The similarity of different genres

We use cosine similarity algorithm to calculate the similarity of different genres. However, due to the large amount of data, we list some genres and contents that have changed a lot.

(1) The similarities between the two genres, blues and country, are declining. According to our calculations, in the 1930s, the similarity between the two genres was as high as 0.99, almost identical. By 2000, the similarity had fallen to 0.90. This decline may not be very large in our opinion, but it reflects some changing trends: The content of country music has become less somber.

(2) Classical and Comedy/Spoken, the similarity between the two is decreasing year by year, but the peak of similarity was in the 1940s, at 0.96, it dropped to 0.66 in the 1990s. We think the Comedy/Spoken genre used a lot of modern technology in its music creation during this period, these modern technologies have led to a change in the genre's creative philosophy, so the similarity between the two gradually decreases(According to the Classical and electric types, the similarity between them is between 0.50 and 0.55)

(3) The similarity of Classical and Pop/Rock reached more than 0.9 in the 1940s and 1950s, this figure is a little beyond our expectation. According to our calculations, the classical type remained basically the same from the 1930s to the 1990s, the similarity is above 0.95, this reflects some changes within Pop/Rock. We notice that the similarity between the two schools changed greatly in the 1960s, the influential musicians in this era include Led Zeppelin, the beach boys and queen and so on, they are all of this type. The similarity between their style and the previous era is relatively low(Between 0.45 and 0.65). It can be said that these people set off a revolution in rock and roll in this era.

2.6 Correlate the network model with the similarity model

We associate the following network model with the similarity model, and find some interesting content. Next, we will explain our findings.

First of all, let's briefly introduce our association method: Modify the previous code to refine the similarity data to individuals and add in the change of year. Finally, we get a huge number of data files, then we try to find out all the following relationships in these data, and find out the trend of this follower.

Almost all artists are similar to the artists they follow (the similarity is more than 0.8). Moreover, most artists are of the same genre as the artists they follow, and only a few artists are different from the artists they follow, for example, Tony Furtado follows three artists, only one of whom is of the same type as him, Pop / Rock, and the other two are country and Jess, with high similarity. The average similarity between country and Pop / Rock is 0.97, and that between Jess and Pop / Rock is 0.98. The variance of these two data is very small, and the similarity between them has always been very high.

Next, we try to find the average score of cumulative influencer in each field, and find that Pop/Rock, country and R&B averaged 19.82, 10.31 and 13.22, respectively. They have a stronger influence on artists than other genres, and the artists influenced by these kinds of artists also follow the art genres of the influencers. We can think that these genres are "more infectious".

3. Conclusions

We build and use two models to analyze all the data: One is the following network model, through which we calculate the influence of different artists and their works. The other is the similarity model, through which we get the similarity of different artists and different music genres. Here are our conclusions:

ISSN 2618-1568 Vol. 3, Issue 2: 55-59, DOI: 10.25236/FAR.2021.030211

1) In our model, there are two main indicators to describe the influence of music: the ability of the musician to inspire others and the ability of the musician to influence himself.

2) We first use cosine correlation of all the characteristics of music and music has carried on the correlation analysis, and then ruled out with epidemic has little to do a few sets of data, the data dimension, and then use the cosine similarity calculation different genre, genre and internal genres on behalf of the artist's similarity, finally it is concluded that the artist in the schools of higher similarity of conclusion.

3) Using the previously established model, we come to the following conclusion: based on the internal similarity of Pop/Rock, most of the features do not change much, but we can see their slight change trend. The change trend within the Rock genre is from focusing on lyrics to focusing more on music melody.

4) According to the previous model, we obtained the artist and oneself follow artists are not the same genre, but the similarity of artists and he followed by the similarity of the artist is essentially higher, visible influence score higher artists of his works in general are "infectious" is quite high.

In general, our model can explain these problems and give a reasonable, appropriate and understandable answer. However, our model also has a lot of room for improvement.

References

[1] Tang, P.; Song, C.; Ding, W.; Ma, J.; Dong, J.; Huang, L. Research on the Node Importance of a Weighted Network Based on the K-Order Propagation Number Algorithm. Entropy 2020, 22, 364.
[2] Lukun Zheng. Using mutual information as a cocitation similarity measure. 2019, 119(3):1695-1713.
[3] Zhou Xuan, Zhang Fengming, Zhou Weiping, Zou Wei, Yang Fan.Evaluation of functional robustness of complex networks using node efficiency [J]. Acta Physica Sinica, 2012, 61(19): 1-7.
[4] Research on Information Extraction Technology for Music Field [D]. Guangdong: South China University of Technology, 2019