Problem Chosen	2021 MCM/ICM Summary Sheet	Team Control Number
D		2103803

The Mystery of Music: Analysis on Musical Influence

Summary

Music is always influenced by various factors. In this report, we construct an evaluation model of music influence, and verify the mutual influence relationship among music, artists and music genres through principal component analysis and cluster analysis. We use time series analysis and time windows to observe the evolution of music. At last we develop a lexical analysis system to explore the cultural influence of music and the effects of social environment.

First, we build a directed network to evaluate the "music influence" of artists. We propose a model which includes three indicators: **Hub Score**, **Authority Score** and **Genre Coefficient**. Consequently, we have obtained a comprehensive assessment of the musical influence of artists.

Then we investigate the similarity and relevance among music, artists and genres. We adopt principal component analysis (PCA) to extract the principal music characteristics, and then we perform cluster analysis to observe similarities and differences between artists and music genres. The conclusion is that music styles are the main factors affecting musical similarity. In addition, we carry out the time series global principal component analysis to explore the variation of genres over time.

On the basis of existing conclusions, we justify the validity of musical influence that we proposed in Task 1. Furthermore, we deeply research in what ways do influencers affect followers.

Next, we analyze the evolution and revolution of music from the perspective of time. We use time windows to search for key points of transformation in the music history. It is believed that the volatility and the growth rate of popularity can be used as the characteristics of innovation. We regard the most popular artists in each period as the revolutionaries. More comprehensively, we identify the trend of music genres by Fisher's optimal timing sequence classification algorithm.

In the end, we develop a lexical analysis system to measure the cultural influence of music and social environmental effects. We can draw a conclusion that popular music has exerted a great influence on traditional culture, and the Second World War has affected all of the music genres.

Keywords: Musical Influence, PCA, Cluster Analysis, Time Windows, Network

Contents

-

1 Introduction

1.1 Problem Background

Music is a treasure in human history. It is of great significance to explore the development of music. Music is not only influenced by music creators, but also may be related to the interaction of artists, and will also change with external social and cultural environment. Under the mutual influence of multiple factors, many music genres keep pace with the times, which leads to constant innovation.

In order to explore music evolution, we need to find a series of approaches to quantify the development of music, such as constructing a directed network to evaluate the influence of music and analyze the reasons.

1.2 Restatement of the Problem

Based on the background information and requirements, we need to complete the following tasks:

- **Task 1**: Build a directed network of music influence, to measure the one-way influence of influencers on followers.
- **Task 2**: Develop music similarity measurements to analyze the similarities and relevance between music, artists and music genres, as well as the development and change of genres
- **Task 3**: Use the similarity measurement to verify whether the influence of influencers on followers is real and effective through the similarity and correlation between musicians and genres, and to analyze the reasons for the influence
- **Task 4**: Research the development and change of music from the perspective of time. Analyze the trend of genres and the key points of musical revolution.
- **Task 5**: Further discuss the cultural influence of music and social environmental effects on music development.

1.3 Our Work

Figure 1 shows the brief outline of our work.

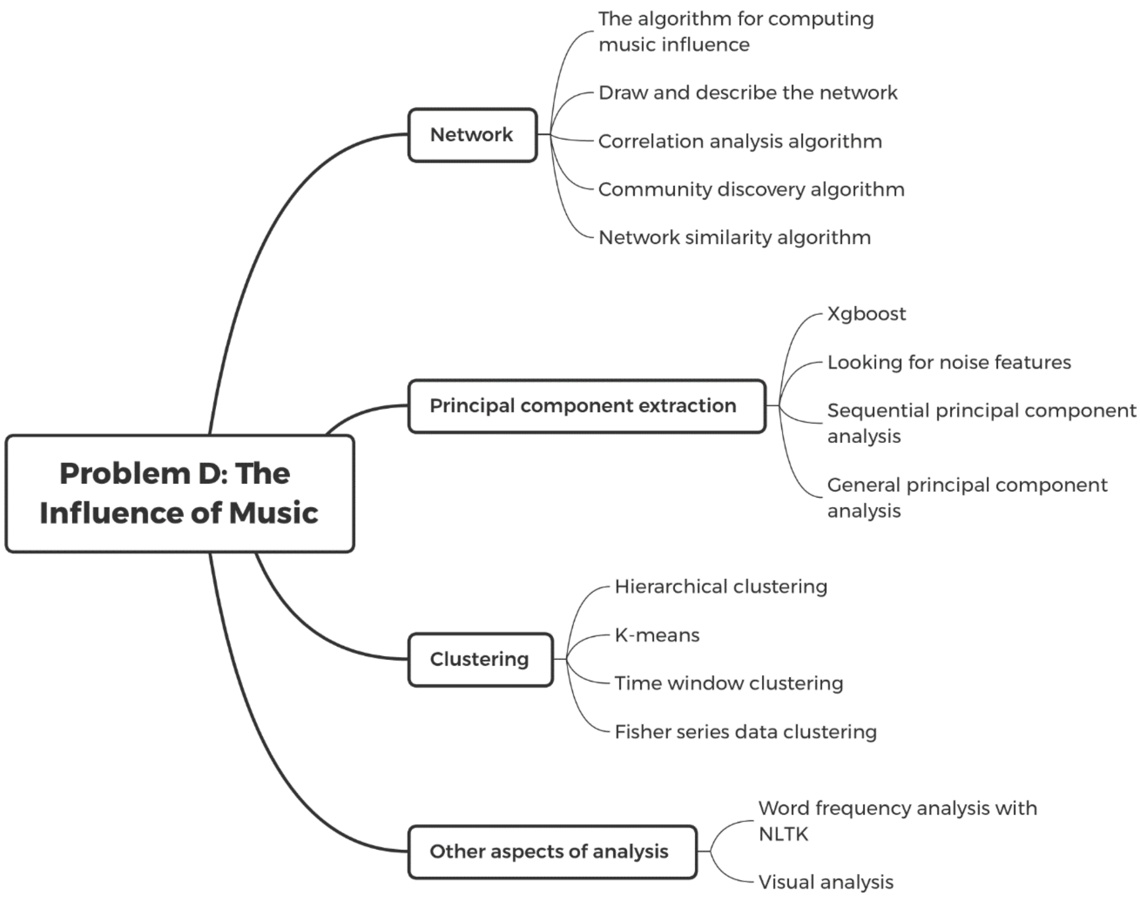

Figure 1: Flow Chart of Our Work

2 Assumptions and Justifications

We make the following assumptions:

1. **the influencer of the same genre and the influencer of the different genre need to be considered separately.** Generally, cross genre communication may bring people great inspiration, but also may have little harvest. Compared with cross genre communication, the harvest of communication with the influencer of the same genre will be more stable, so we need to consider it separately, and we need to try to take the mean value to eliminate the instability of cross genre communication.

2. **The more people you influence, the more influence you have.** It is generally believed that the more people accept an influence, the more influence the influencer should have.

3 Task 1: The Model of Music Influence

In this section, we build a directed network connecting influencers and followers. Based on this network, we comprehensively consider factors such as the number of followers, the quantity of influencers and the genre of music. Then we construct a model to evaluate music influence，and finally we calculate the score of each artist's musical influence from 1930 to 2020.

3.1 Description of the Directed Network of Musical Influence

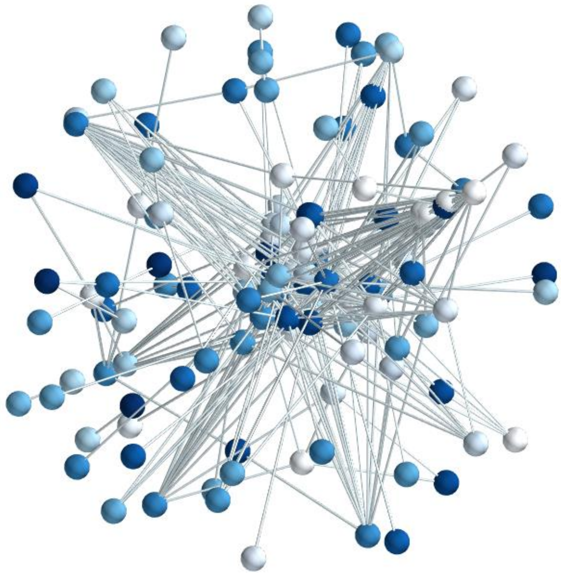

Figure 2: Partial of the Musical Influence Network

We use python-networkx package to analyze the raw data and build our directed network. To improve the 3D visualization effect, we only draw the partial image of our network by Mayavi, as shown in figure 2. In this diagram, each small ball is a node, which represents an artist. If there is a connection between two nodes, it means that there is an influence-follow relationship between them.

3.2 Musical Influence Evaluation Model

3.2.1 Model Establishment process

We divide all artists into three categories:

- **Inf**: Pure influencer. They only affect followers without being influenced by other artists.
- **Inf&Fol**: They are both influencers and followers.
- **Fol**: Pure follower. They follow influencers, but do not affect others.

The number of artists in each category is shown in figure 3.

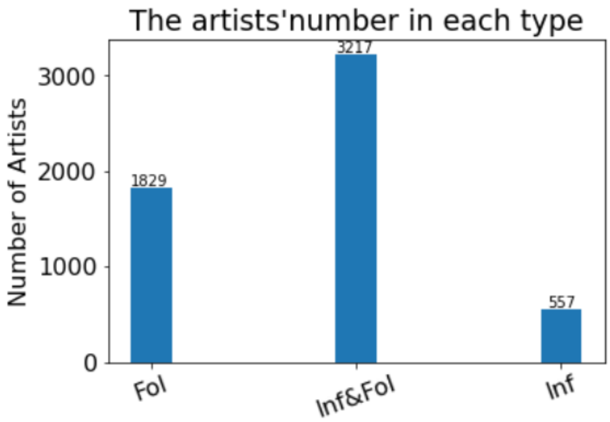

Figure 3: The Number of Artists of Each Type

We calculate the influence of each artist by letting followers rate to their influencers.

We believe that the musical influence of an artist is directly related to the quantity of followers. However, the number of artists is dynamically changing over time, which means a parameter is required as a regulatory factor [1]. Furthermore, a follower may be influenced by many artists, but the degree of music impact from each influencer is different. In order to simplify the problem, we assume that the musical influence is mainly affected by two factors [2]: **Hub Score** and **Authority Score**:

- **Hub Score (HS)**: The score that followers rate to their influencers.
- **Authority Score (AS)**: The weight of the score when followers scoring.

In addition, we need to consider the genre factors. We believe that if the influencer and its followers belong to the same genre, the influence will be greater. Oppositely, an artist will be less influential if he and his followers are of different genres. In view of that, we propose **Genre Coefficient** ρ to indicate the influence of the genres on music impact.

Figure 4 shows the process of model establishment and solution.

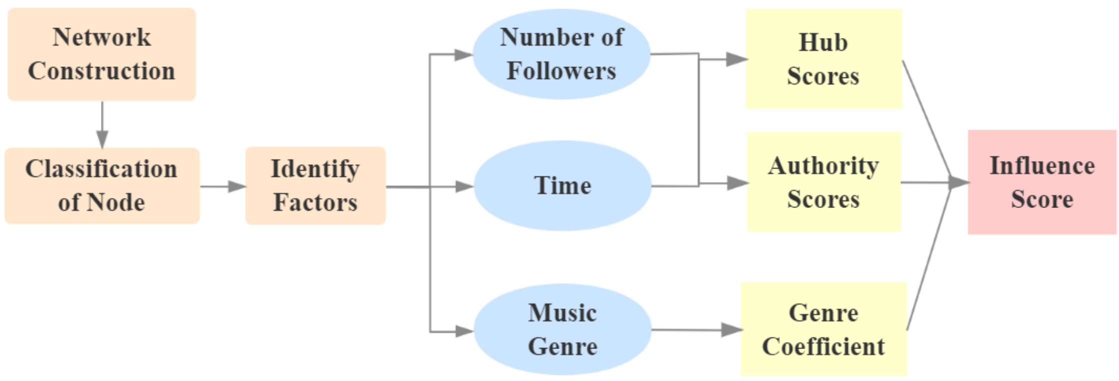

Figure 4: The Flow Chart of Model Establishment

3.2.2 Hub Score

In order to clarify our algorithm process, we take an ideal three-layer network displayed in figure 5 as an example for follow-up elaboration. The symbols and definitions are shown in Table 1.

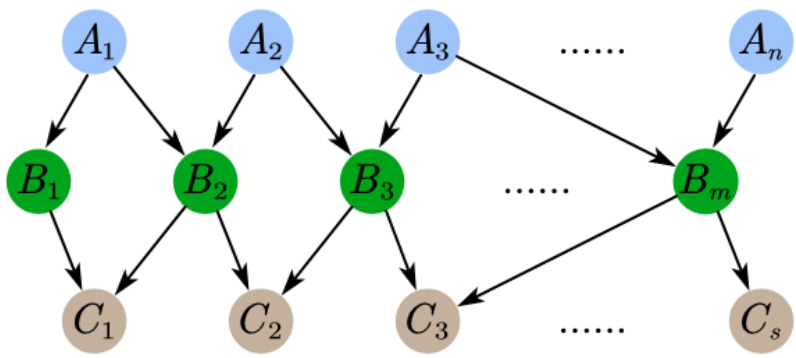

Figure 5: the Ideal Three-layer Network

Table 1: Symbols used in this model

Symbol	Description	Unit
p	Indicate the influence of the genres on music impact	
X	Influencer	
Y	Follower	
A	Pure influencer (an influencer who is not a follower)	
B	a follower who is an influencer	
C	Pure follower (a follower who is not an influencer)	
T	The total number of followers influenced by artists in the same period	
O	The outdegree of B	
HS	Hub Scores of an artist	
AS	Authority Scores of an artist	
S	the covariance matrix	
Z	a collection of related artists	
$p_{k,j}$	the correlation coefficient between genre and genre	
η	the principal components	

Firstly, we only consider the number of followers. We define the outdegree of B_i as $O_{t,b}$, where t represents the start time of the artist's music career. To eliminate the influence of population base growth, we let t be a constant value, and we calculate the total number of followers T_t that influenced by artists in the same period. The calculation formula is as follows:

$$T_t = \sum_{i=1}^{n} O_{t,i} \quad t = 1920, 1930, \ldots, 2020 \tag{1}$$

Then the Hub Scores of an artist should be calculated as:

$$HS_i = \frac{O_{t,i}}{T_t} = \frac{O_{t,i}}{\sum_{i=1}^{n} O_{t,i}} \quad (t = 1920, 1930, \ldots, 2020) \tag{2}$$

But for pure followers, they haven't influenced others, so we describe the Hub Scores of pure followers like:

$$HS_k = \frac{1}{s} (k \in 1..s) \tag{3}$$

3.2.3 Authority Score

A follower is simultaneously affected by multiple influencers, but the degree of impact may be different. Therefore, a weighting factor should be considered. Authority Score indicates the authority of influencers. The more followers, the higher Authority Score. Taking Figure 6 as an example, B1 are influenced by A1 and A2, and A1 affected 2 artists, while A2 affected 1 artist, so A1 is more authoritative.

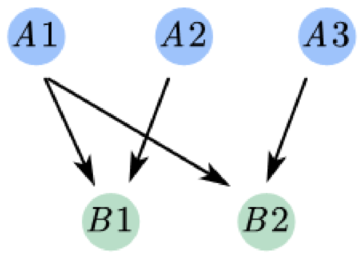

Figure 6: the Simple Network Graph

Suppose a follower X_i is affected by n influencers, the outdegrees of these influencers are $O_{t,i}$ $i \in 1..n$, so the Authority Score of influencers can be expressed as

$$AS_{i,j} = \frac{O_{t,j}}{\sum_{k=1}^{m+s} O_{t,k}}, \ i = 1, 2, ..., m+n \tag{4}$$

3.2.4 Genre Coefficient

When an influencer is of different genres, we can't accurately measure to what extent a follower is affected by the influencer. So the ratio of the scores that a follower rate to his influencers of the same genre is A, which can be obtained as follows:

Step 1: For each followers of the genre Z, calculate the percentage of the influencers of genre Z. Assuming that an follower X_i is affected by l influencers, and e influencers are of genre Z, so the percentage of influencers of this genre is

$$\alpha_i = \frac{e}{n} \tag{5}$$

Step 2: The Genre Coefficient A is defined by **the average of** α_i.

3.3 Results

In conclusion, the scores that a follower rates to an influencer of the same genre is

$$A \cdot HS_i \cdot AS_{i,j} \tag{6}$$

The scores that a follower rates to an influencer of the different genre is

$$(1-A) \cdot HS_i \cdot AS_{i,k} \tag{7}$$

Assuming that each influencer(both Inf and Inf&Fol)will receive j times of ratings. So we can calculate "music influence" by

$$music\ influence_i = \sum_{k=1}^{j} score_k \tag{8}$$

We calculate the "music influence" of each artist and rank them by influence degree, the top 10 are shown in figure 7.

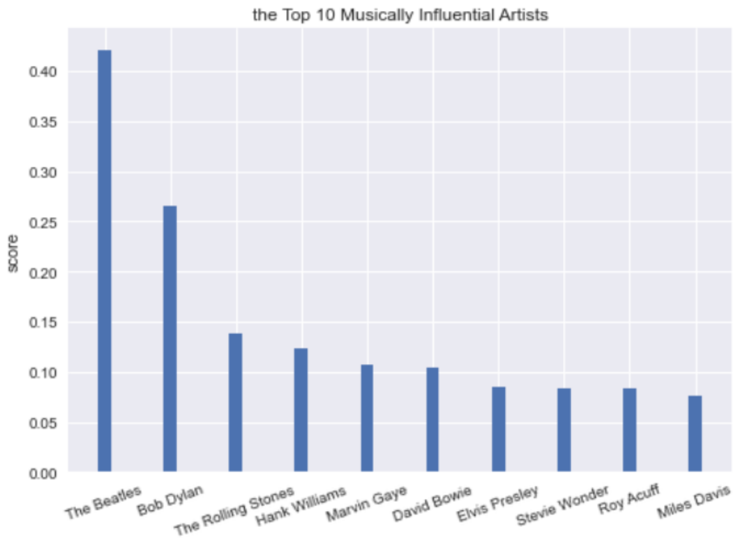

Figure 7: the Top 10 Musically Influential Artists

4 Task 2: Musical Similarity and Relevance

4.1 Data Preprocessing

First of all, we perform data preprocessing, taking their genres as labels, the remaining attributes as characteristics. Then we use the xgboost model for multi-classification, so that we can obtain the scores for each feature, as shown in Figure 8.

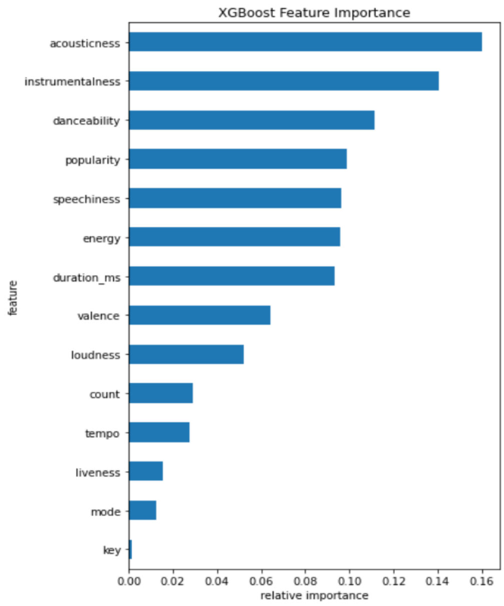

Figure 8：Xgboost Feature Importance

On the basis of thresholds theory, we can determine that there are four noise properties:

'duration_ms', 'count', 'mode' and 'key', adversely affecting the results of clustering. So we delete these attributes and screen out the following indicators, displayed in Table 3. At last, we process the data to eliminate the effect of dimensionality.

Table 2: Music Genre and Numbers

Number	1	2	3	4	5
Genre	Pop/Rock	Country	Classical	Electronic	Comedy/Spoken

Number	6	7	8	9	10
Genre	Easy Listening	Reggae	Jazz	R&B;	Latin

Number	11	12	13	14	15
Genre	Vocal	Folk	International	Avant-Garde	Blues

Number	16	17	18	19	20
Genre	Stage & Screen	New Age	Religious	Children’s	Unknown

Table 3：Property Name and Corresponding Number

Number	1	2	3	4	5
Name	Danceability	Energy	Valence	Temo	Loundness

Number	6	7	8	9	10
Name	Acousticness	Instrumentalness	Speechness	Duration_ms	Popularity

4.2 Similarity of Music and Artists

4.2.1 Measures of music similarity

Since the effective music characteristics have been selected, we use principal component analysis to extract the principal components from the ten characteristics. According to the processed data set, we calculate the covariance matrix S and its corresponding eigenvalues, as well as principal component coefficients. Then we calculate the variance contribution rate according to the eigenvalues, and the number of principal components should be selected to make the cumulative variance contribution rate reach 80%.

According to the cumulative contribution rate in Table 4, two principal components can explain 84.78% of all the information. As shown in Table 5, the coefficient of No. 6 in Component1 is the largest, so Component1 mainly represents 'Acousticness'. While in Component2, we can see that the coefficients of No.2 and No.3 are larger, so Component2 delegates 'energy' and 'valence', which reflects the listener's experience of the music, such as positiveness, vitality and a strong sense of spiritual impact.

Table 4：The Characteristic Root and Contribution Rate of Covariance Matrix

Characteristic Root	Contribution Rate	Cumulative Contribution Rate
4.37	0.7776	0.7776
0.39	0.0702	0.8478
0.31	0.0551	0.9029
0.25	0.0460	0.9489
0.11	0.0209	0.9698
0.07	0.0139	0.9837
0.04	0.0072	0.9909
0.02	0.0044	0.9952
0.02	0.0037	0.9989
0.005	0.0011	1.0000

Table 5：Coefficients of the Two Principal Components（units：10^{-1}）

Number	1	2	3	4	5	6	7	8	9	10
Component1	0.23	-1.40	0.10	-0.18	-0.33	9.86	-0.01	-0.39	-0.04	-0.75
Component2	2.97	5.53	6.30	0.55	1.48	1.03	-1.19	2.70	-0.74	2.86

We extract the principal components of each artist and calculate the Mahalanobis distanc e between them after normalizing the principal components. And use this distance as a measu -re of their similarity. The Formula is as follows:

$$d_{ij}^2(M)=(X_i-X_j)'\Sigma^{-1}(X_i-X_j) \tag{9}$$

4.2.2 Similarity of Artists

Considering the two principal components as the characteristics of each artist, we use Mahalanobis distance to measure the similarity between two artists, then we categorize artists by clustering analysis. Through the contour coefficient diagram (Figure 9) and the graph of intra-cluster square deviation sum (Figure 10), artists can be divided into five groups, as shown in Figure 11.

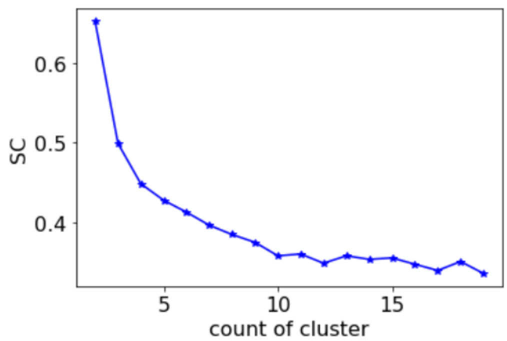

Figure 9: the Contour Coefficient Diagram

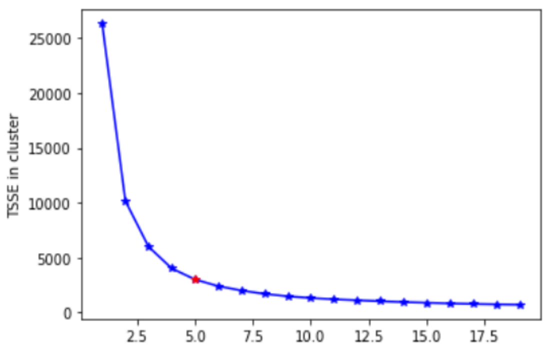

Figure 10: Intra-cluster Square Deviation Sum

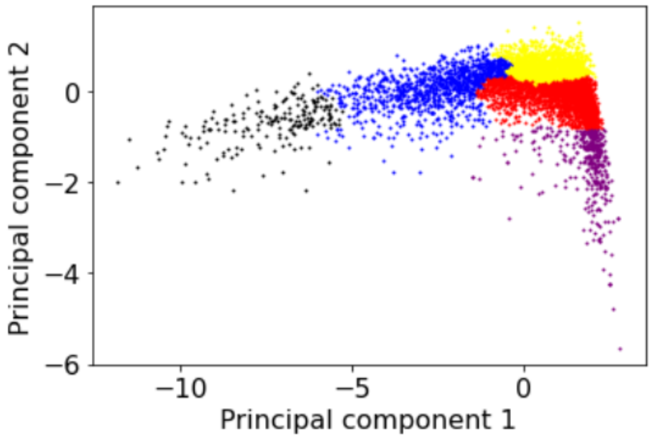

Figure 11： Cluster Analysis Diagram

According to the clustering results, we obtain the proportion of the five types of artists in different genres, which is depicted in Figure 12. Details are shown in dataset 3. The figure reflects that genre is not the main factor that determines the difference of artists. Artists of the same genre may present different styles, while artists of different factions may also show similarities when creating music [3].

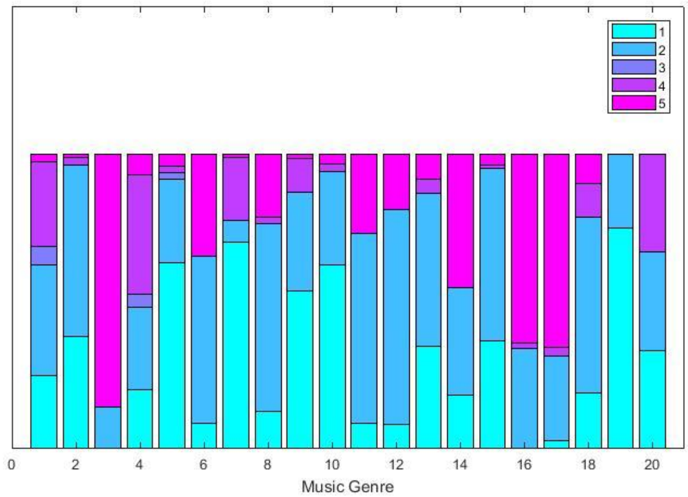

Figure 12： the Distribution of Five Types of Artists In Different Genres

4.3 Relevance and Similarity of Music Genre

4.3.1 Similarity of Genres

Based on the proportion distribution in Figure 12, we identify the type with the largest proportion as the genre's characteristic, calculating the Mahalanobis distance of two genres to perform cluster analysis. The results are shown in figure 13 and figure 14. All the music genres are divided into five categories, as shown in Table 6 and Figure 15.

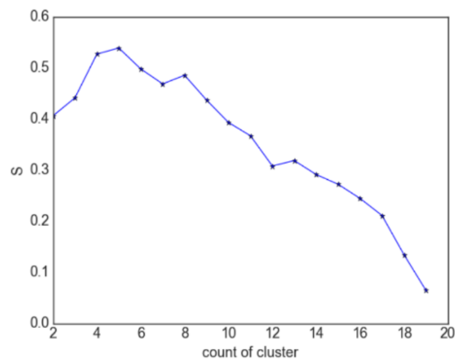

Figure 13: the Contour Coefficient Diagram

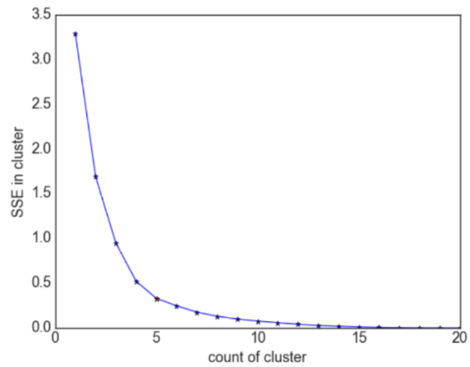

Figure 14: Intra-cluster Square Deviation Sum

Table 6：Clustering results table

Categories	1	2	3	4	5
Genres	1	4	2、6、8、11、12、13、15、18、20	3、14、16、17	5、7、9、10、19
Style	3	4	2	5	1

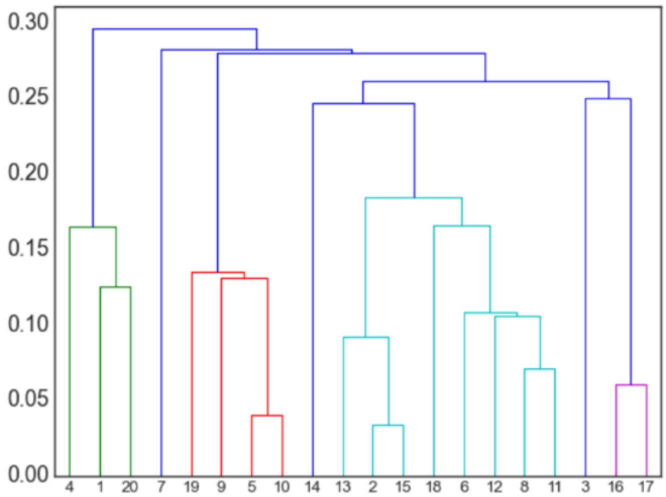

Figure 15: System Clustering Diagram

As can be seen from Table 4, the similarity of music genres is decided by the artists' type with the largest proportion, which indicates that the more similar the music style is, the higher the similarity of the genre.

4.3.2 Relevance of genres

To analyze the correlations between genres, it is necessary to explore the variation of genres over time. Utilizing the dataset, we sort each music by genres, And for each genre, we rank those music by release dates. Songs released in the same year are characterized by the average of each feature, so we analyze the time series of each character that may exists breakpoint from 1920 to 2020.

We adopt the time series global principal component analysis, to obtain the comprehensive evaluation time series of each genre, as shown in figure 16:

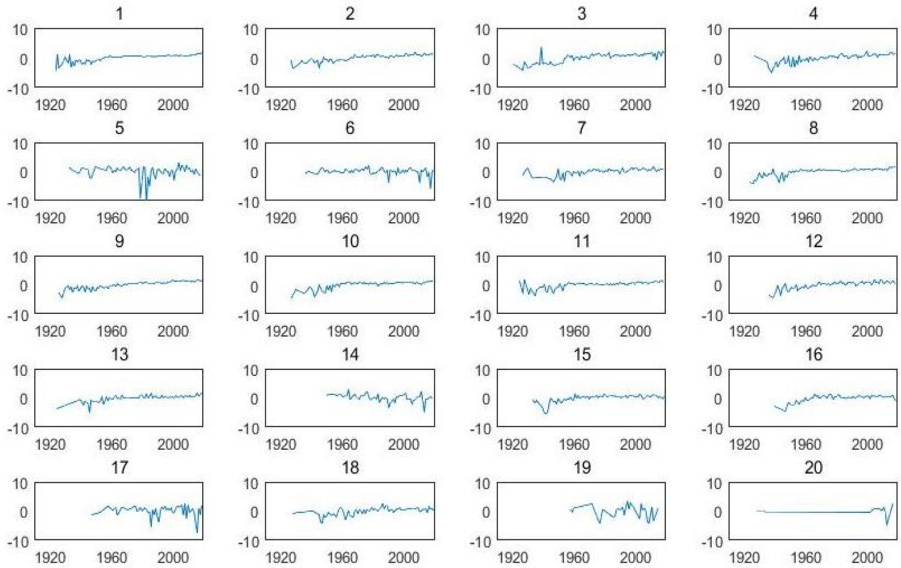

Figure 16：Comprehensive Evaluation Time Series Diagram

Figure 16 shows the trend of each genre's overall score. Due to the relatively small sample size of genre 19 and 20, we don't take them into consideration. It can be seen that genre 5 and 17 have great fluctuations in the middle part of the sequence. Genre 6 and 14 have great ups and downs at the end, while the rest fluctuate on the front end.

Then we need to count the correlation between the two genres. We define a feature Z for each genre, which we call it as "a collection of related artists". Z equals the union of each point and the neighbors of each point in the genre. Once the Z attribute of each genre is calculated, the correlation coefficient between k genre and j genre is defined as [4]

$$\rho_{ij} = \frac{\{Z_i\}_j \cap \{Z_i\}_k}{\{Z_i\}_j \cup \{Z_i\}_k} \tag{10}$$

5 Task 3: Validity of Musical Influence

In this section, we examine whether the influencers actually affect the music created by their followers based on the previous research conclusions. We found the most influential artists according to the 'music influence' obtained in Task 1, and we analyzed the comprehensive evaluation time series of each faction in Task 2. If these influential artists can actually affect the creation of fans, the comprehensive evaluation time series will show different changes at the time nodes when the artist start his career, as shown in figure 17.

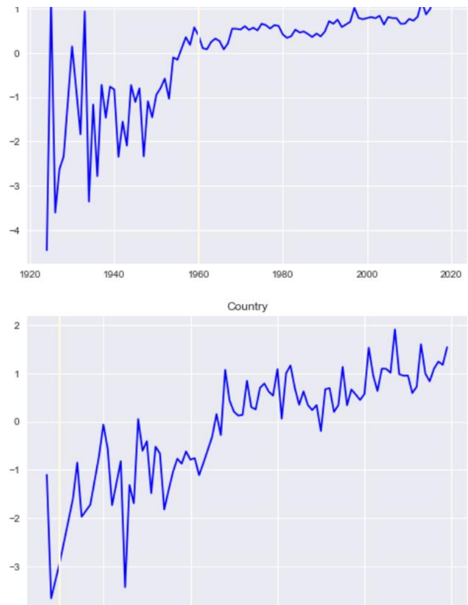

Figure 17：Comprehensive Evaluation Time Series with Nodes

It can be seen from Figure 12 that the comprehensive evaluations of most genres fluctuate greatly before the time node and become stable after the time node. We can believe that the influencer affects the certain creative trend of music genres.

Then we analyze the characteristics which influencers affect their followers. Taking Pop/Rock as an example, use the time when influencer start their career as the separated time point, and the artists before and after the separated time points are represented by scatters in different colors, as shown in figure 18. It can be inferred that these influencers contributed to a general increase in 'acousticness' among artists of this faction.

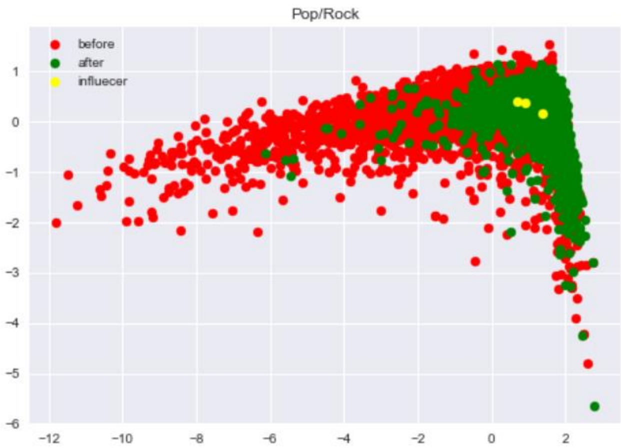

Figure 18：Pop/rock Artist's Principal Components Scatter Diagram

6 Task 4: Development and Change of Music

6.1 Revolutions of Music

We use time windows to analyze key points in the development of music. We extract the principal components of the Pop/Rock faction separately.

The specific steps of time window analysis are as follows:

Step1：After data standardization, take 10 years as a unit to extract time Windows year by year, for example, we regard the year of 1921~1930 as a time window, year 1922~1931 as a time window, and number each time window in order.

Step2：Measuring time windows.

Step3：Clustering the time window.

91 time Windows are obtained, and the clustering results are shown in Table 7. The results divide 1921-1936 into a category, the 1937-1949 into a category, and the 1950-2020 into another.

Table 7：the Clustering Results of Time Windows

Categories	1	2	3	4	5
Time Windows	1	2	3~16	17~20	21~91

Since there is no further differentiation after 1950, we repeat to perform time window clustering to the data from 1953 to 2020. The results are shown in Table 8, which divides the 1954 ~1978 into one category and the 1979~2020 into the other category.

Table 8：the Clustering Results of Time Windows

Categories	1	2	3
Time Windows	1	1~17	17~59

To sum up, the time can be divided into four parts:1921~1936, 1937~1949, 1950~1978 and 1979~2020. According to the time series diagram of each principal component, as shown in Figure 14. According to the corresponding coefficients of each principal component, it can be found that principal component 1 represents 'popularity', principal component 2 represents 'speechiness', and principal component 5 represents 'valence', while the coefficients of principal component 3 and principal component 4 are not significantly prominent.

In combination with the first chart and the time period in Figure 19, it is found that the volatility of popularity is relatively large in the first time periods, and the growth is relatively rapid in the third time period, but much slower in the fourth time period. Therefore, it is believed that **the volatility and the growth rate of popularity can be used as the characteristics of innovation**. Similarly, when the fluctuation or growth rate of popularity, speechiness and valence time series change, it can be regarded as the characteristics of innovation.

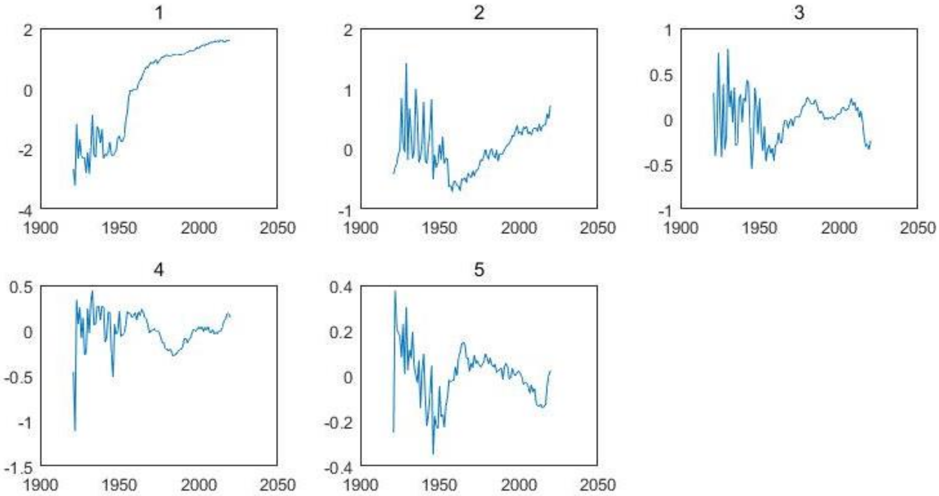

Figure 19： Time Series Diagrams of Each Principal Component

According to the obtained time periods and combined with the artist's popularity ranking, we can find out the artists who rank higher in each time period as the revolutionaries. Figure 20 shows the respective revolutionaries in the time windows.

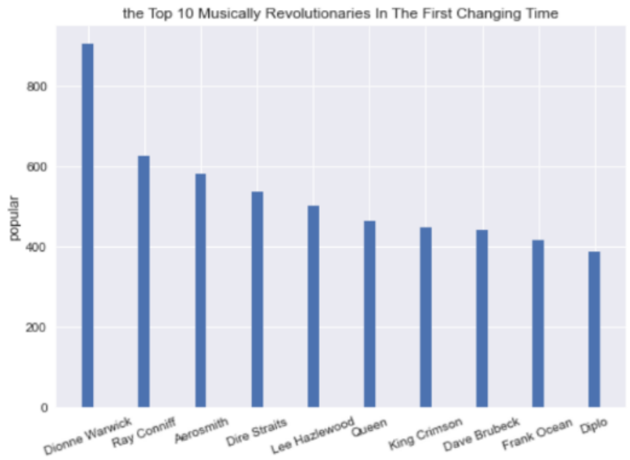

Figure 20(a): Revolutionaries of the First Time Window

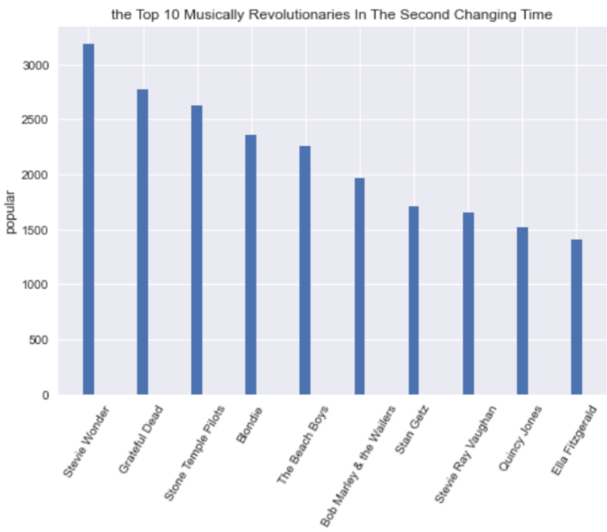

Figure 20(b): Revolutionaries of the Second Time Window

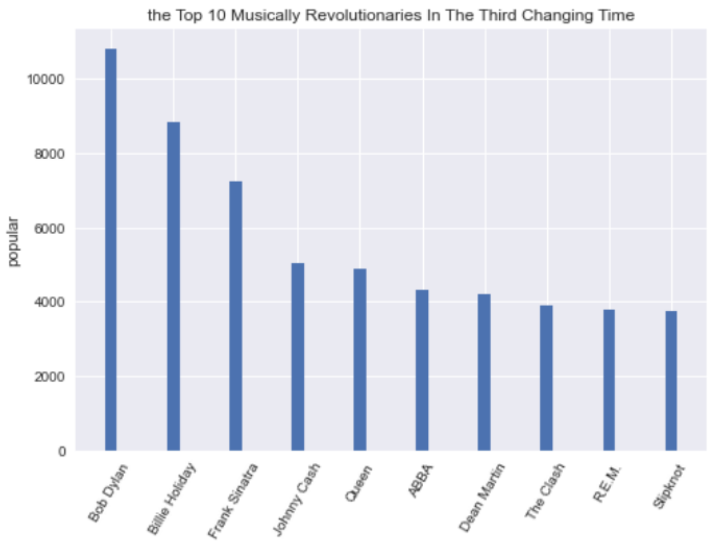

Figure 20(c): Revolutionaries of the Third Time Window

6.2 Development of Music Genres

We choose Pop/Rock as our research genre. We extract the time series principal components, choosing the first three principal components as our coordinates. We use Fisher's optimal timing sequence classification algorithm to classify our timing sequence data, and the conclusion are as follows:

Table 9: Results

Categories	1	2	3
Time Window	1924-1948	1949-1968	1969-2020

Because of the good performance results, we can directly use the principal components to measure the variation of our genres. The equation is as follows:

$$\begin{aligned}\eta = 0.05 \times danceability + 0.32 \times energy + 0.05 \times valence + 0.04 \times tempo \\ + 0.11 \times loudness - 0.42 \times liveness - 0.06 \times instrumentalness \\ - 0.01 \times speechiness + 0.05 \times duration_m s + 0.83 \times popularity\end{aligned} \tag{11}$$

As shown in figure 21, Pop/Rock music is becoming more and more popular, while Pop music is growing more Energy, less liveness, and more dance-oriented.

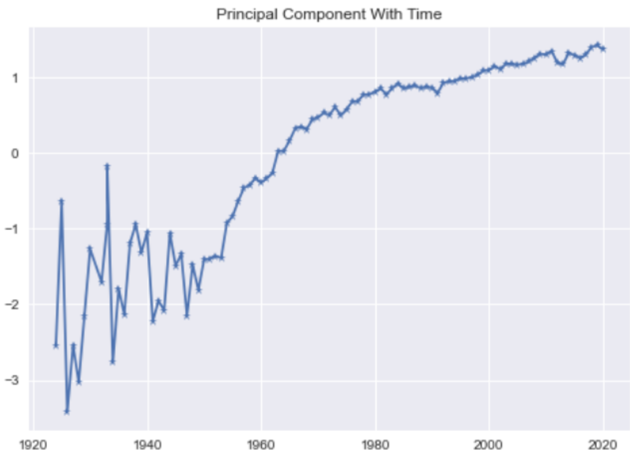

Figure 21: Principal Component with Time

7 Task 5: Cultural Influence of Music and Environmental Effects

We develop a lexical analysis system to measure the impact of music on culture. Changes in each type of music are also visualized, in order to see how cultures are influenced by music over long periods of time. We start our analysis after analyzing the song names.

We use dynamic time windows as the basic unit. Specifically, the closer the time to the present, the shorter the time window will be. On the one hand, the small number of songs in the early years takes a long time to accumulate. On the other hand, considering that with the evolution of times, people's lifestyle and culture are updated faster, so the time window in recent years is naturally relatively small. Starting from the 10-year-span time window, it has gradually evolved to the 5-year-span time window in recent years.

In the current time window, we count the words in the song titles and observe the word frequency distribution of different periods, as shown in Figure 22. In this way, we can infer whether the cultural characteristics of this period are different from those of other periods.

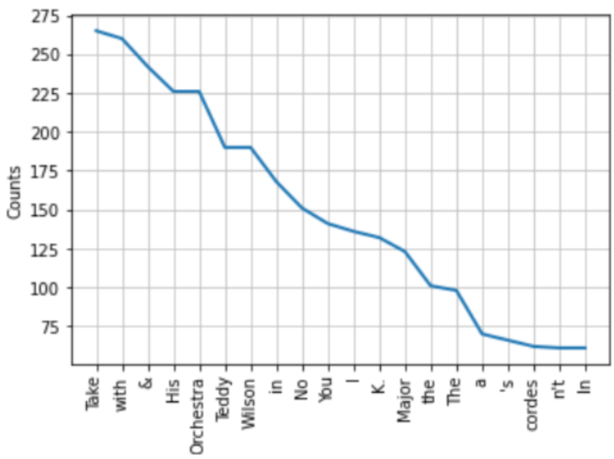

Figure 22： A Graph of Word Frequency Over a Period of Time

By observing the statistics of word frequency in different time periods, we can obviously find the cultural influence caused by music. For example, the word "His" appeared more frequently in the early years, while in recent years, words like "I" and "Yours" appear more frequently. This change of appellation reflects the changes of love and sense of independence.

At the same time, we measure the number of music released in different genres over the years, as shown in the figure 23. It is obvious that classical music is gradually declining in recent decades, which shows that popular music has exerted a great influence on traditional culture.

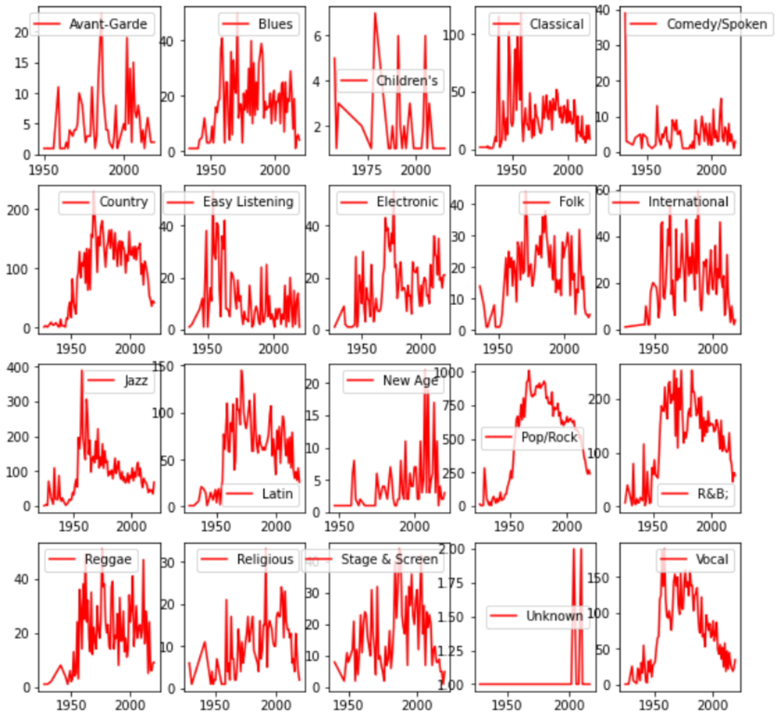

Figure 23: Time Series Diagram of Released Music in Each Genre

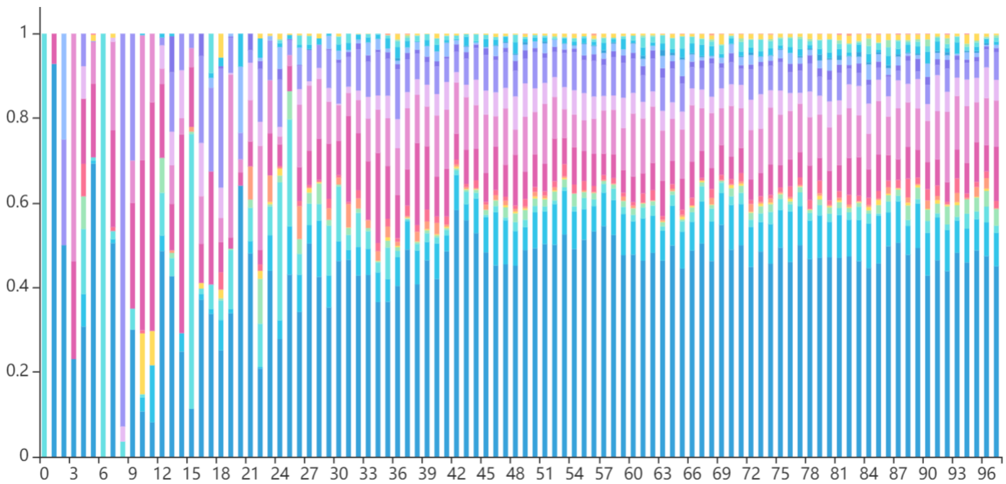

Figure 24: Percentage of the Market Share of Each Genre Per Year

To explore the influence of social, political and technological progress on music, we believe that we should focus on the transformation of some important parameters in the network over time. We calculate more than ten time Windows and draw a network within each time window to obtain a network attribute sequence diagram (as shown in the figure 25).

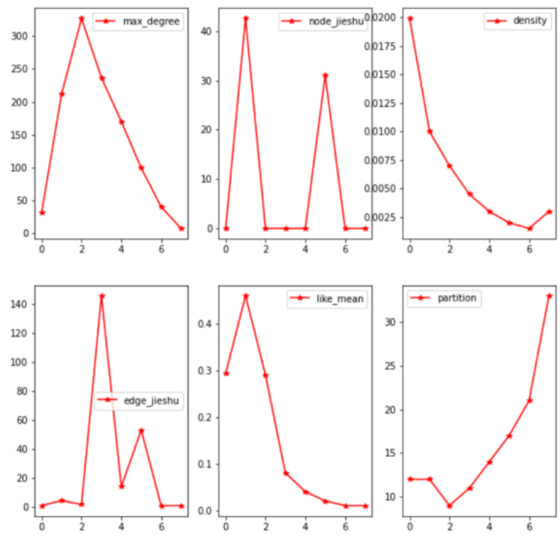

Figure 25: Visualization Result Diagrams of Attributes

Then we calculate the average correlation of the network to reflect the music similarity in a certain period of time, as shown in figure 26: we believe that if the music similarity is high during a period of time, it indicates that there is a common theme of fanaticism at that time. The figure shows that the similarity of each genre is high from 1930 to 1960, which shows that during this period, a major event affecting the national spirit occurred, named the Second World War.

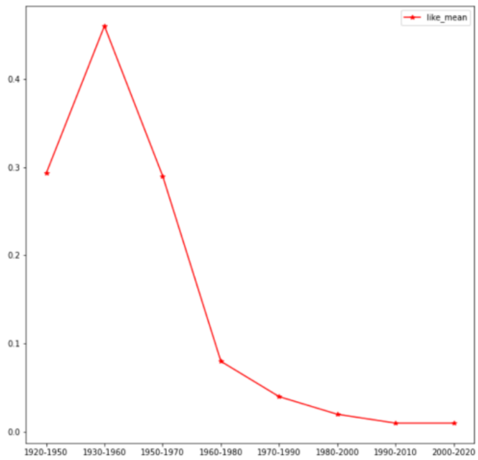

Figure 26: Correlation Mean Time Series Diagram

8 Sensitivity Analysis

In the "music influence" index constructed in the first question.

If we no longer consider the influence of different genre (in our original model, the utility of voting members of this genre is different from that of voting for foreign genres), if we change it here and default that its utility is the same, then we can still have no particularly drastic changes in the top ten, as shown in Figure 27, which means that our model has a good stability.

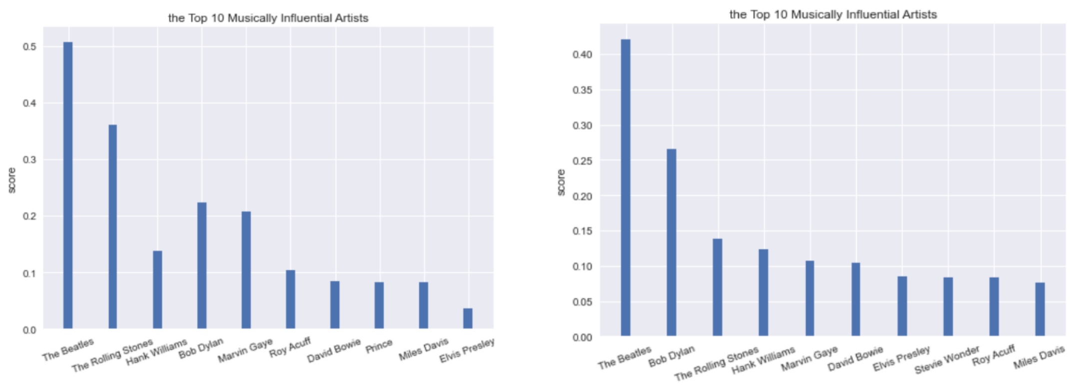

Figure 27: Results after modifying

In the second question, we have four types of data: the original data, data without noisy features, the principal component data, and the principal component data without noisy features. We carry out KNN learning and xgboost multi-classification machine learning respectively. In the two models, both principal component analysis and original data analysis show a better effect after removing noise features.

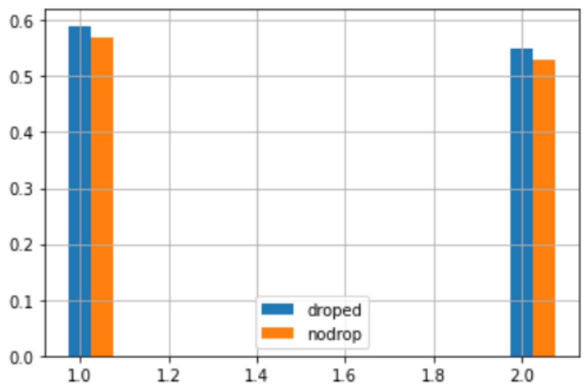

Figure 28: Results after Removing Noise Features

9 Model Evaluation and Further Discussion

9.1 Strengths

1.We use the "music influence" index calculation method, which takes into ac-count the population growth, through the proportion of the number of people affected rather than the direct number of people, and fully considering the differences of influence between different factions, so that our "music influence" index can effectively alleviate the score deposition effect caused by traditional node ranking algorithms (such as Pagerank, etc.), and fully reflect the influence of artists in different times.

2.The "xgboost" algorithm is used to screen features before principal components, fully considering the validity of features. And according to the flexible and changeable characteristics of data, we use different principal component extraction methods (traditional principal component extraction and temporal principal component extraction) for different problems, which makes the pro-cess of principal component extraction more efficient.

3.In the clustering process, a variety of clustering methods are combined (such as K-means, systematic clustering, fisher time series data clustering).

4.We use and create a variety of algorithms. For examples, community discovery algorithm, similarity measurement algorithm, time window clustering algorithm, word frequency analysis algorithm. We analysis the potential information in the data sufficiently.

9.2 Weaknesses

1. The use of principal components makes our data lose some information compared with the original data.
2. In question 6, the time global principal component analysis method selects more variables, and the explanatory power of the model is insufficient.

9.3 Further Discussion

1. If we can get the connection between words and cultural characteristics, or the dictionary of the emotional content of words, then we can get more information from the names of songs.
2. We should consider the specific mechanism of action between each artist, and draw an industry knowledge graph for our music industry with the interpersonal relationship between artists (external data sets).

10 Conclusion

We build a directed network to evaluate the "music influence" of artists. Our model includes three indicators: Hub Score, Authority Score and Genre Coefficient. Then we successfully achieved a comprehensive assessment of the musical influence of artists. We can clearly see that the top 3 musically influential artists are **the Beatles**, **Bob Dylan** and **the Rolling Stones**.

We investigate the similarity and relevance among music, artists and genres. Through principal component analysis (PCA) and cluster analysis, we can draw a conclusion that genre is not the main factor that determines the difference of artists, while music style is the critical factor. At the same time, different music genres are highly relevant.

Our model produces a desirable result which shows that the music influence we calculated is effective, and we can achieve a further discussion about in what aspects that musical influence represents. This enlightens us to explore the influence mechanism of music and apply it in many fields such as music teaching and psychotherapy.

Moreover, we analyze the trends of music over time, looking for key points and exploring the reasons for major shifts at a given moment. We have also discovered the cultural influence of music, changing people's lifestyles and behaviors; At the same time, music will also be affected by the external social and political environment, and then produce the change of style. In view of this, we can further study the social background of music works, which is better for us to understand the influence of music.

References

[1]Xing Su, Research on Community Detection algorithm based on Node similarity in complex Networks [D]. Lanzhou University, 2020.
[2]Liu Jianguo, Ren Zhuoming, Guo Qiang, Wang Binghong. Research progress of node importance ranking in complex networks [J]. Acta Physica Sinica, 2013, 62 (17): 9-18
[3]Fan D J, Ma X M, Wang L J. Hand shape classification and size optimization of Zhejiang young men based on cluster analysis.
[4]Zhao Danfeng, Huang Yanling, Huang Dongmei, Lin Junchen, Song Wei. Research on Mining method of time Series motif Association rules based on AR_TSM [J/OL]. Computer Application Research: 1-7 [2021-02-09]. Https://doi.org/10.19734/j.issn.10013695.2019.12.0664.

The Value of Music Influence Network

Music has been part of human societies since the beginning of time as an essential component of cultural heritage. But if you are a person who has never been exposed to music, it can be challenging for to understand and measure music and its development from an artistic point of view. But it doesn't matter, through the quantitative music development model we developed, you can easily understand and measure music development by some mathematical indicators.

In the next few sections, you can use the model we developed to analyze the complex network of interpersonal worship in the music industry to find out the real music masters. You will know the importance of the different characteristics of a song and understand the similarities between different songs. Understand the genres of music then understand the differences and connections between them. Know the history of music and major points of influence, and get in touch with the great musicians who have changed the music world. It can make you become a “music liker” cause it can have a detailed understanding of music development.

Of course, if you are a gentleman who is interested in music, then our model will contribute to your music career. Not only can you use it to understand and measure music development mentioned above, but you can also use it to learn to measure specific music features. So as to learn faster and develop your music knowledge.

However, the most important thing is that you can understand the real impact of music if you use our model.

You can understand the impact of music on our environment and culture through the model, and you can also understand how music reflects the atmosphere of the surrounding environment.

As our data becomes richer, our model will delve deeper into the impact of music on culture.

What meaning did he have to our culture, and how did it work on our culture?

When our network is gradually enriched, we will construct an unprecedented music-centered knowledge graph to understand the influence of music in culture in an all-round way. So it's time to explore the mystery of music and try a new method!