The Research on Music Influence Model

In order to quantify the evolution of music, in this article we propose a comprehensive evaluation model to measure the influence of music, and use this model as a tool to review the evolution and revolutionary trends of artists and genres.

First of all, in the directional network of music artist influence relations, we use the PageRank algorithm to initially obtain the PR value to measure the artist's influence, and select multiple indicators that can measure the influence of artists, including the number of influencers and competitiveness level, and establish a grey comprehensive evaluation model through entropy method. A balanced evaluation of multiple artist attributes and a quantitative measure of the influence of influencers on followers. According to this model, we have obtained more than 100 influential master artists, led by Beatles, Bob Dylan, and Rolling Stones.

Second, we use PCA to facilitate visual observation. According to the cross-validation C-H index, we determine that works should be divided into two categories. Through the analysis of the cohesion map, we explained that it is not the genre that determines the essence of the work, but the evolution of aesthetics, history and technology.

Third, through the explanation of the heat map of Pearson correlation coefficient in different situations, we have discovered the preference for the same musical technique within genres and the difference in the use of different musical techniques between genres and we have established a directed graph based on the reference relationship between genres, revealing how each genre can learn from other genres over time.

Fourth, we use the nominal data similarity measurement method to quantitatively characterize the influencer's influence on the music of the followers. In order to measure the degree of influence of the influence on each music feature index, we use the cosine similarity measurement method to measure the similarity of each music feature index of the sub-network and use this as a data column, the gray correlation method is used to analyze the degree of relevance. At the same time, in order to improve the rationality and representativeness, we establish a measure of the change in the style of the followers by imitating the SGD feedback rule of the neural network. It is found that the styles of followers of high-impact artists are more similar to those of influencers and the results are visualized as a double-layer sunburst chart. From the global network and sub-networks, quantitative analysis and visual analysis, it is concluded that music indicators like valence, loudness, tempo, energy and etc. mainly inherit the influence of influencers, and their relevance is above 0.9, and key and mode are the attributes that are least imitated by followers.

Fifth, establish a three-dimensional coordinate map of influence-year-genre to determine the time of change as around 1957. In the genre that has undergone major changes, the actual historical environment is used to analyze the changes in music characteristics within the genre, through the correlation analysis with the level of the entire music field. The correlation analysis of the evaluation indicators between the Pop and the whole music field was carried out. It was found that this genre played an important role in the evolution and development trend of the music field.

Finally, we further analyzed the evolution of the genre and the trend of revolution by evaluating the influence of the model, various musical characteristic indicators and external factors. In order

to better apply to different sizes of global networks and states, we conducted sensitivity analysis and comprehensive optimization. This model has good mobility and application value in the aspects of music evolution and influence measurement.

KEYWORD: PageRank, Grey Comprehensive Evaluation Model, K-Means, Cosine Similarity, Pearson Correlation, Grey Relational Analysis

1 Introduction

1.1 Background

Music is one of the best carriers of popular aesthetic culture based on the self-expression and the joy of inner life. As an important part of cultural heritage, it affects people's values and lifestyles. The series of changes it causes cannot be ignored in any society and culture. Music is also inextricably linked to new music and music artists.

Therefore, this article establishes a method to quantify musical evolution, thus understanding and measuring the influence of previously produced music on new music and musical artists. Our goals are to examine evolutionary and revolutionary trends of artists and genres.

1.2 Our Work

In this article, we will:

- create multiple directed networks of musical influence, then explore a subset of musical influence by creating a sub-network of your directed influencer network.
- develop musical similarity measurement model and use it to measure the similarity of artists within genre and artists between genres.
- compare similarities and influences between and within genres.
- indicate whether the similarity data suggest that the identified influencers in fact influence the respective artists.
- characterize signify revolutions (major leaps) in musical evolution.
- analyze the influence processes of musical evolution that occurred over time in one genre.
- express information about cultural influence of music in time or circumstances and explain how can the effects of social, political or technological changes be identified within the network.

2 Problem analysis

This article aims to examine the evolution and revolutionary trends of artists and genres by establishing a mathematical model to measure the influence of music. To solve this problem, we first carry out preprocessing according to the given conditions and attached data, determine key indicators and discuss modeling.

Problem 1: The core of the problem is to develop certain indicators as parameters to reasonably represent the influence of music. At this time, the question is transformed into selecting indicators to quantitatively represent the influence of music. We use the *Influence_data* set to establish a directional network that represents the relationship between artists. Through the directional network, influencers and followers have many-to-many mapping connections. We develop some potential indicators. We use PR iterative algorithms and establish grey comprehensive evaluation models for these indicators. Characterize musical influence.

Problem 2: The goal of problem 2 is to establish a music similarity measurement model. First, we abstract the musical works into 12 pieces of characteristic information that characterize the works. The problem is transformed into a classification problem of a large data point set. The cluster analysis is performed through the unsupervised learning k-means algorithm, and the similarity of nodes is formed into clusters. Observe and explore the classification that can essentially reflect the characteristics of the work, whether it is divided into genres or clusters. Finally, we plan to successfully explain why genres cannot be used as the essential classification of works by studying the reasons for the clustering of works (evolution of aesthetics, history, and techniques).

Problem 3: From what we found in the second question, we took the preference for skills as a starting point. Musical skills will be reflected in the abnormal performance of information characteristics, and in numerical terms will be manifested as a higher Pearson correlation coefficient. So we plan to make a heat map according to the Pearson correlation coefficient, and answer the connection and difference between the inside and outside of the genre by explaining the very different distribution of the heat map. By establishing a directed graph and demonstrating the evolution of the directed graph over time, it reveals the changes of genres over time, and the extent to which genres can learn from each other.

Problem 4: This problem can be transformed into how to measure the similarity between influencer's influence and follower-influencer music. If influence and similarity are positively correlated, then it can be considered that "influential people" really influence the music created by followers; accordingly, the latter part of the question can be transformed into comparing the similarities between the influencers and the followers. If there is a certain musical characteristic that is much more similar than other musical characteristics, then it can be considered Musical characteristics are more contagious than other characteristics. If the similarities of the various musical characteristics are not very different, then it can be considered that they all play similar roles in influencing the music of a particular artist.

Problem 5: Regarding this issue, we are thinking backwards. Some artists have changed some of the music characteristics, leading the trend and leading to major changes in the overall music. Therefore, we need to determine the time of change first, and determine the music that has changed significantly during that period of time. Then analyzing within the genre that has undergone major changes, the artist with the greatest degree of relevance to changes in music characteristics within the genre is analyzed and determined as the representative changer to be sought.

Problem 6: This question can be transformed into an analysis of the changing process of influencers' influence and the relationship between key music indicators and music development in a music genre over time. Our plan is to map and analyze the changes in the music characteristics of the genre over time and the level of influence over the years, find out the music characteristics that have changed significantly over time in the development process, analyze its influence on influence, the number of new generation artists, popularity, etc., and explain the dynamics in combination with the evaluation model indicators in model1 Influencer indicators and their influence process.

Problem 7: This problem requires us to analyze the actual situation based on the results of the previous problems.

3 Notations and Definitions

symbol	meaning	symbol	meaning
xi	Data column indicators	PR	PageRank evaluation value
Ai	Weights	r	Correlation coefficient
ρ	Distinguishing coefficient	sim	Similarity

4 The Establishment and Solution of Models

4.1 Problem One---Comprehensive Evaluation Model

4.1.1 About Modeling Ideas

The key of the problem is to develop a certain indicator as a parameter to reasonably represent the influence of music. At this time, the problem is transformed into selecting an index to quantitatively represent the influence of music. We use the *influence_data* to establish multiple directional networks that characterize the influence of music. Through the directional network, influencers and followers have multi-to-many mapping connections. However, there are differences in the musical influence between artists, which needs to be reflected in our figure by calculating reasonable weights for nodes and edges.

According to the PageRank[1] network iterative algorithm, the weight of the network nodes is determined according to the specific distribution of the affected people, and the artist's own influence is represented by this. We select multiple indicators that can measure the influence of the artist's music. We use entropy method[2−3] to assign the importance of each indicator and establish a Grey Comprehensive Evaluation Model

4.1.2 Directional network and its sub-networks

According to the influencer-follower relationship given in the *influence_data* data set, the individual musician is represented as each node in the network, and the relationship between influencer (parent node) and follower (child node) is represented as an arrow connecting each node. The link of an influencer node affects its neighbors and then affects the entire network by diffusion, resulting in a global directional network. For the sake of simplicity, a simplified schematic diagram is as following figure 1:

The global network is a multi-to-many mapping analysis of all individual artists. In order to specifically explore the level of influence, arbitrary follower is selected to study the level of influence of all its influencers on this follower. With the follower node as the center, the influencer nodes surround the center to form a one-to-many contact subnet of any research object, as shown in the figure 2.

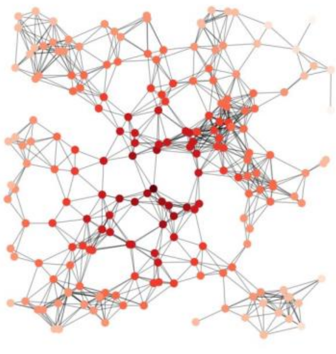


Figure 1. Influencer-follower directed network

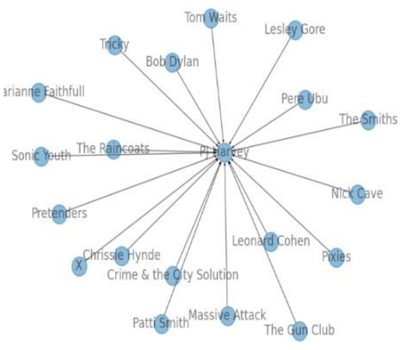


Figure 2. Sub-network

The weight of each node is the PR value, and the weight of the edge is obtained by the gray evaluation model. The calculation principle and method are given below. The sub-network is studied with PJ Harvey as an example. In fact, our model is applicable to any follower.

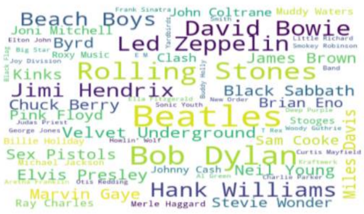


Figure 3. Artist Master Word Cloud

Determination of each model index

a. the PR value[4]

Since each node in the obtained directional network has a certain level of influence, when an artist with a higher influence level is influenced by another artist, then his/her influencer must have a higher influence than the former artist. The artist's directional influence relationship is accumulated layer by layer. We assign weights to the mapping relationship between influencers and followers. First, randomly assign weights to them. According to the PageRank algorithm, the global network can be iterated to obtain the stability of the directional network. For each weight, the weight obtained after iteration is called the PR value. At this time, the PR value indicates the level of the number of people directly or indirectly affected by the affected person, and the number of affected people is converted into a PR value indicator.

We obtain the PR value of each artist in the global network, and take the set of influencers with a higher PR value (top 100) as the master artist. In order to display our results visually, we use the PR value as the basis for the font size to display these master artists in a word cloud[5].

b. Artist influence attribute vector

In order to perform Grey Relational Analysis[6], it is necessary to determine which attributes can measure the influence of the artist. So we believe that the attributes that can be found on the Internet that can measure the influence of music include:

x_1: PR value/node weight, which represents the number of people affected directly or indirectly by each influencer.

x_2: The number of people from the same genre as the researched artist in the directional network, indicating the intensity of competition in the genre of the influencer.

x_3: The number of people from the same genre who debuted in the same year as the researched artist in the directional network, indicating the intensity of competition in the field of influencers that year.

x_4: The total number of people in the directional network who debuted in the same year as the researched artist, indicating the intensity of competition when the influencer debuted.

x_5: Whether the follower inherits the influencer's genre, the same is 1, the difference is 0, which is an important aspect of the influence of the influencer on the follower.

The above constitutes the artist's influence attribute vector:

$$x' = (x_1, x_2, x_3, x_4, x_5)^T$$

In order to reduce the subjectivity of the selection of the above indicators, we adopt the Entropy Weight Method to calculate the entropy weight w of the indicator, and the more objective artist influence attribute vector x is defined as:

$$x = wx'$$

4.1.3 Comprehensive Evaluation Model

According to the above indicators, The Comprehensive Evaluation Model is established to analyze the influence level of each artist in the sub-network. The center of the sub-network is the follower PJ Harvey, and its influence attribute vector is denoted as x_0. The influencers around the center in the sub-network are all influencers who have influenced PJ Harvey. The attribute vector of the i^{th} influencer is denoted as x_i, is shown in the following table.

Table 1. The attribute vector

Artist Name	x_1	x_2	x_3	x_4	x_5
PJ Harvey	0.00024	2808	369	1129	1
Pretenders	0.000195	2808	364	796	1
Bob Dylan	0.000252	2808	372	855	1
Sonic Youth	0.000328	2808	462	951	1
Crime & the City Solution	0.000105	2808	462	951	1
Nick Cave	0.000223	2808	364	796	1
Captain Beefheart	0.000141	2808	372	855	1
Leonard Cohen	0.000116	2808	99	540	1
Pixies	0.000165	2808	462	951	1
X	0.000156	2808	364	796	1
Lesley Gore	0.000111	2808	372	855	1
The Smiths	0.000226	2808	462	951	1
Chrissie Hynde	0.00018	2808	364	796	1
Massive Attack	0.000265	208	31	951	0
Pere Ubu	0.000146	2808	364	796	1
Tom Waits	0.000155	2808	364	796	1
Tricky	0.000165	208	61	1129	0
The Gun Club	0.000142	2808	462	951	1
Patti Smith	0.000165	2808	364	796	1
Marianne Faithfull	9.63E−05	2808	372	855	1
The Raincoats	0.000108	2808	364	796	1

The weight of each index is determined according to the entropy method:

Table 2. The weight of each index

w_1	w_2	w_3	w_4	w_5
0.419255	0.142749	0.167414	0.127833	0.142749

Substitute into the calculation to obtain the influence attribute vector of each artist.

Select x_0 as the reference column, and the attribute vectors $x_1, x_2, \dots, x_n$ of all influencers as the comparison column. According to the data of each indicator, substitute the Gray Correlation Analysis Algorithm to calculate the correlation degree to the reference column, and get the correlation between each influencer and follower as an evaluation index that affects the strength of the relationship. So we get that the edge weight representing the influence level of each artist in the sub-network is:

Table 3. The influence level of each artist

Artist	Edge Weight	Artist	Edge Weight
Pretenders	0.687054	The Smiths	0.698312
Bob Dylan	0.701389	Chrissie Hynde	0.687054
Sonic Youth	0.698312	Massive Attack	0.517105
Crime & the City Solution	0.698312	Pere Ubu	0.687054
Nick Cave	0.687054	Tom Waits	0.687054
Captain Beefheart	0.701389	Tricky	0.532089
Leonard Cohen	0.611421	The Gun Club	0.698312
Pixies	0.698312	Patti Smith	0.687054
X	0.687054	Marianne Faithfull	0.701389
Lesley Gore	0.701389	The Raincoats	0.687054

As what is shown above, the higher the PR value, the stronger the artist’s influence; the higher the weight of the influencer to follower, the stronger the relationship between them. So far, our work has revealed the artist's level of influence from the perspective of the artist's own influence and the strength of the influence relationship.

4.2 Problem Two---Music Similarity Measurement Model

4.2.1 Data Preprocessing

The attached data files *full_music_data and the two summary data sets (with artists and years)* provides us with Characteristics of Music Works by Artists of Different Genres.

Most algorithms are quite sensitive to the relevance, reliability and validity of the data. Therefore, we need to perform data preprocessing by cleaning, selecting and normalizing the data.

Data Screening

The amount of raw data is large, so we should first do data screening according to integrity and usefulness of the information.

- For some obviously wrong data, like Wolfgang Amadeus Mozart's influencer_active_start is 2010 in *influence_data*, we will modify it to the correct age.
- For the cross-border collaborations of multiple artists from different genres, we regard them as belonging to all participating genres simultaneously. E.g. Howlin' Wolf, Eric Clapton, Steve Winwood, Bill Wyman &Charlie Watts done *I Ain't Superstitious* together. It belongs to the genre Blues , Pop/Rock and Jazz.
- For works of unrecorded genres, we ignore them.

Data Normalization

The characteristics of music can be quantified by the value from *danceability* to *duration_ms* in the table. Since we will use indicators with various units, we have to normalize all the indicators to facilitate the further analysis of the data. It can provide an approach for comparison of different kinds of data and reflect the combined results of different factors.

Here we use Min-max normalization, scaling all the values in the range [0,1]. Formula (1) gives the general form of the adopted normalization. Let x' be the normalized value, the formula is :

$$x' = \frac{x - x_{max}}{x - x_{min}} \qquad (1)$$

where x_{max} and x_{min} are respectively the maximum and minimum value of the indicators in the same unit.

Feature selection

Then, we use Pearson Correlation Coefficient, formula (2) to measure the correlation between each feature and popularity.

$$\mathrm{p(x,y)} = \frac{\sum x_i y_i - n\overline{xy}}{(n-1)S_x S_y} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum {x_i}^2 - (\sum x_i)^2}\sqrt{n\sum {y_i}^2 - (\sum y_i)^2}} \qquad (2)$$

The results are shown in Table 4, and then we delete the features whose correlation coefficient is too small, that is, the absolute value is less than 0.15, which are marked in red here.

Table 4. The correlation coefficient of each characteristics

Characteristics	Pearson Correlation Coefficient	Characteristics	Pearson Correlation Coefficient
danceability	0.183133	instrumentalness	−0.216110
energy	0.389819	liveness	−0.064326
valence	0.020748	speechiness	−0.006928
tempo	0.094173	explicit	0.183029
loudness	0.420300	duration_ms	0.057150
mode	−0.036295	year	0.775415
key	0.015031	popularity	1.000000
acousticness	−0.477451		

So far the data set is preprocessed.

4.2.2 Model building and solving

To develop Music Similarity Measurement Model, it is necessary to quantitatively describe the similarity between different music. A work contains many indicators, we can regard each work as a point in space, where n is the number of indicators contained in each work. Next, we describe the similarities between music.

We believe that the works that best represent a genre are those with the highest popularity. Therefore, in each genre, the top 100 most popular works are select to represent their respective genres. Training the model in this way can better initialize the convergence point.

We use PCA to analyze the feature information of the works and color the points according to the genre category, so as to visually observe the clustering of works.

Principal component analysis (PCA)

The two-dimensional visualization of PCA is shown in Figure 4. The features of the work are roughly divided into two clusters. Some genres are quite cohesive (for example, the Pop/Rock genre represented by purple and the electronic music genre represented by dark green), but not all genres are like. The feature information of the genre is not obvious in the PCA dimensionality reduction.

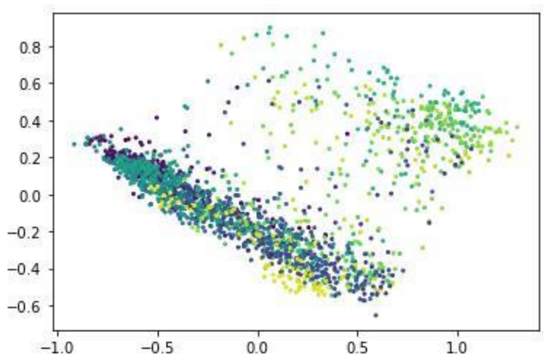


Figure 4. The two-dimensional visualization

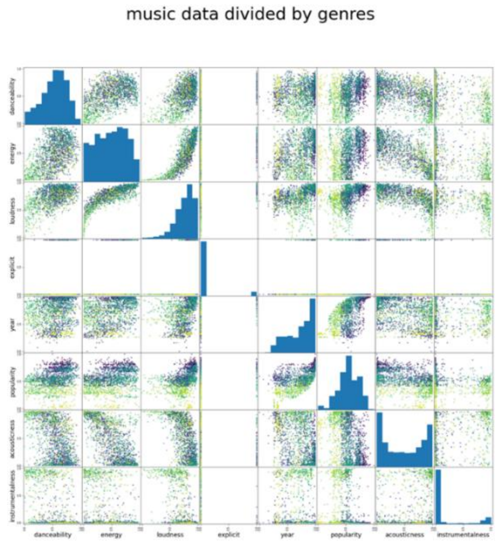


Figure 5. Music data divided by genres

In order to better observe the feature information of the work, we then intercept each two kinds of characteristics, draw it into a two-dimensional space, we view the characteristics of music works within the genre in the scatter matrix. The scatter matrix is shown in Figure 5.

Let's take Pop/Rock represented by purple as an example. It can be seen that in many cases, the cohesion of Pop/Rock is quite good, which is consistent with what we observe in the PCA dimensionality reduction figure, but in some cases (for example, when it involves instrumentalness) is quite scattered, which shows that Pop/Rock does not show similar characteristics, that is, the music within the genre is quite different in these aspects.

Does it mean that genres cannot absolutely divide the difference between music? For this reason, we use K-Means Cluster Analysis. It is worth mentioning that K-Means is an unsupervised learning algorithm, which ensures that genres will not interfere with classification as prior knowledge.

K-Means Cluster Analysis

Among the classic clustering algorithms, there are K-Means, Special clustering etc., K-Means is still widely used in various fields due to its high efficiency and easy implementation. We use K-Means for cluster analysis, and nodes represented by works with close similarity will form a cluster.

We categorize the works to compare the similarity without taking the genre of the works as prior knowledge.

After inputting new data without labels, we compare each feature of the new data with the corresponding feature of the data in the sample set, calculate the similarity, and then extract the classification label of the most similar data in the feature vector of the sample set. For different sample feature vectors, the value range of the correlation coefficient is [-1,1], and the similarity varies between 0-1.

The K-Means Algorithm is sensitive to the value of k. In fact, in the context of this problem, k represents that music should be divided into several categories according to the characteristic information, which is very important for our analysis. In the dimensionality reduction figure, it can be seen that the points are roughly clustered into two clusters. In order to explore this problem more precisely, we adopt the Cross-validation method and use the Calinski-Harabasz Index, formula (3) to evaluate K. The larger the Calinski-Harabasz Index S, the greater the clustering effect.

$$s(k) = \frac{\mathrm{tr}(B_k)}{\mathrm{tr}(W_k)} \frac{m-k}{k-1} \quad (3)$$

Where *m* is the number of samples in the training set, and *k* is the number of categories. B_k is the covariance matrix between categories, and W_k is the covariance matrix of data within the category. *tr* is the trace of the matrix.

The curve is shown in the figure 6. It can be seen that it is the most reasonable way to divide music into two categories, which also confirms the situation we have observed in the PCA figure.

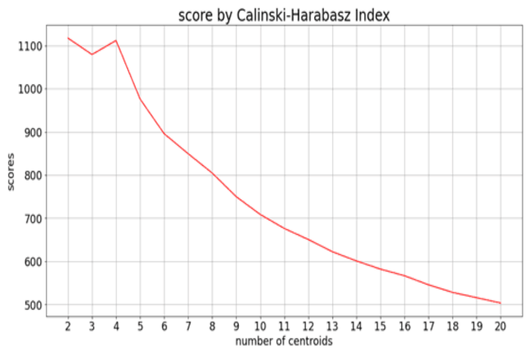


Figure 6. Calinski-Harabasz Index

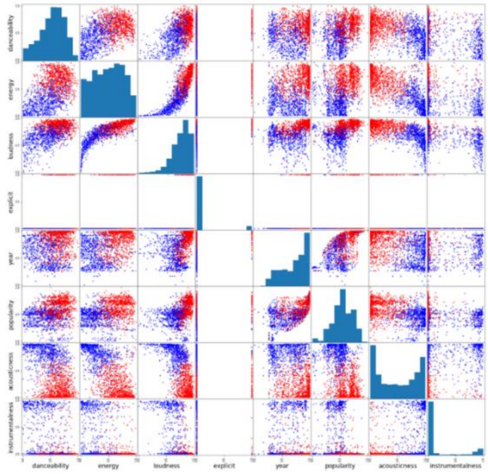


Figure 7. Music data divided into 2 centroids

Next, we color the original image according to the classification indicated by k-means, as shown in Figure 7.

It can be seen that the cohesion of all points is quite good now, which shows that the musical works represented by these points have more similar characteristics. Here, we explain energy row to explain as an example.

In energy-related combinations, the red dots represent one type of music, which is more powerful than another type of music, and another characteristic information naturally explains the reason:

the red have higher value in *danceability*、*loudness*、*year*、*popularity* and have lower value in *acousticness*、*instrumentalness*. This shows that they has strong rhythm and loudness, which is undoubtedly an important reason for the more energy of red. We can know from year and popularity, that the age of this kind of music is closer to us, and its popularity is also high. This tells us from the perspective of music development history that people from modern times have begun to try to create new music. They have a higher popularity and more importantly, our model shows that they have parted ways with the music of the past and formed a new kind of music. We are really exciting to see that their efforts have finally paid off! From the perspective of *acousticness* and *instrumentalness*, this seems to indicate the technical reasons for the changes in new music: the development of vocals and the use of soundtrack processing are quite positively related to the production of new music.

In general, red music is more passionate and unrestrained, while blue music is more classical and introverted, and our model successfully divides music into two more essential categories than genres. It is worth mentioning that the electronic music represented by dark green is almost unwavering red music in almost all situations, which is consistent with our analysis, while other genres have both color at the same time. It doesn't belong to one of them steadily, which shows that the music styles of most genres are quite diverse. Therefore, the division of genres is not that big and different genres of music often show similar characteristics.

4.3 Problem Three---Similarities and influences between and within genres

4.3.1 Relevance of characteristics between genres

According to the music characteristics of the *full_music_data* data set, we further used Pearson Correlation Coefficient to quantitatively analyze the various musical characteristics of works in the genre, and draw a heat map of correlation coefficients.

Works of the same genre are more likely to have a consistent preference for the use of certain characteristics, which can be seen from the Pearson correlation coefficient. For this reason, we divide music works by genre, and make a heat map of characteristic information. First of all, we found that in the case of finding the correlation coefficient of all music features, loudness and energy have a correlation of 0.79, which is quite reasonable, while other features have no obvious correlation. In order to reveal the connections within the genres and the differences between them, we explain two representative pictures among them.

Figure 8 shows the similarity of Avant-Garde's music works in 13 dimensions including danceability, energy, and valence. For Comedy/Spoken genre music, the correlation coefficient between speechiness and liveness is quite high. The reason can be explained as the use of spoken words has brought higher popularity to this type of music, which is consistent with the reality; Features are not clearly related. This shows that the most obvious feature of Stage&Screen music is the unique relationship between the use of spoken words and popularity.

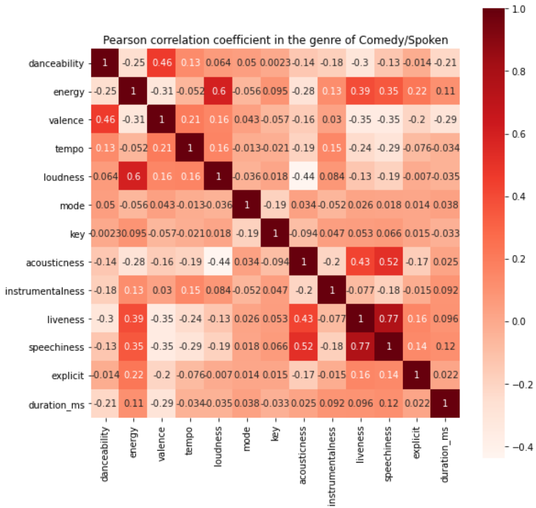


Figure 8. Comedy/Spoken

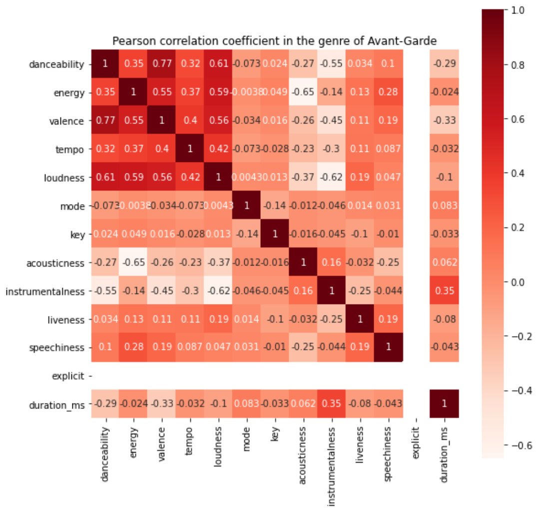


Figure 9. Avant-Garde

Figure 9 shows the similarity of Avant-Garde's music works in 13 dimensions. For Avant-garde music, there are many musical characteristics that are not lowly related, which shows that avant-garde music generally has the courage to try various methods, music is their laboratory, and this is also the passion of Avant-Garde music. The most obvious difference from other genres.

The above two examples show that high correlation coefficients can reveal similarities within the genres, usually this corresponds to the consensus on the use of certain techniques within the genres, and the difference in the heat map between the genres indicates the different priorities between the genres.

4.3.2 Changes Over Time

We extracted the influencers and affected followers in the influence_data.csv file, and got 20 factions after deduplication. As shown in the table, the 20 factions are sorted by the first letter and labeled with 0-19.

Table 5. Music genres

Music genre	Serial number	Music genre	Serial number
Avant–Garde	0	Jazz	10
Blues	1	Latin	11
Children of Bodom	2	New Age	12
Classical	3	Pop/Rock	13
Comedy/Spoken	4	R&B	14
Country	5	Reggae	15
Easy Listening	6	Religious	16
Electronic	7	Stage & Screen	17
Folk	8	Unknown	18
International	9	Vocal	19

Twenty categories correspond to twenty nodes. One edge connects two nodes with a certain category affecting a certain category, so the evolution of music genres over time is transformed into a weighted graph between genres from the source point. The problem of the path to the destination.

Floyd-Warshall algorithm

There are many algorithms for calculating the shortest path, mainly including Dijkstra algorithm, Bellman-Ford algorithm, Floyd algorithm and SPFA algorithm. Considering that the Floyd algorithm adopts a triple loop compact structure and is highly efficient, it is suitable for dense graphs, that is, complex multi-node, multi-Laminar school evolution, we use it to explore the evolution time and evolution path between the two schools.

There are two possibilities for the evolution of any genre i to j:

1. Directly from i to j;
2. Evolve from i to j through several genres.

The path with the shortest weighted path length is called the shortest path, and its path length (the sum of weights) is called the shortest path length or shortest distance.

It is more appropriate to use an average value for one type of time that affects another type. We can approximately assume that each influencer affects another person under the same conditions, and its error obeys a normal distribution, and the average value is closer to the true value. If the test is less than 0, it is actually the fact that the affected faction appears in the influence faction. From a factual point of view, it is impossible to happen. It is considered that it happened at the same time and the propagation time is 0.

The shortest path between the two nodes involved in this problem is not negative and meets the conditions of use.

From what is mentioned above, we get the path of evolution between the two genres. For intuitive reflection, we select the evolution process of five genres, Avant-Garde, Easy Listening, Blues, Country, and Latin, a simplified path diagram is obtained, as shown in Figure 10.

The shortest evolution time is shown in Table 6, which represents the shortest path between two nodes, that is, the shortest time for evolution between two genres.

Table 6 shortest length between genres

	Avant-Garde	Blues	Children's	Classical	Comedy/Spoken	Country	Easy Listening	Electronic	Folk	International
Avant-Garde	0.000	25.167	30.000	6.000	26.810	28.500	23.810	19.412	18.500	28.945
Blues	20.333	0.000	38.333	20.254	27.754	3.333	10.333	36.193	20.957	16.626
Children's	46.364	53.191	0.000	52.364	53.000	56.513	50.000	54.147	51.867	55.135
Classical	14.545	19.167	24.000	0.000	30.000	22.500	19.000	28.500	12.500	27.500
Comedy/Spoken	22.000	16.000	20.000	17.500	0.000	5.000	12.000	34.318	22.000	18.293
Country	17.000	11.000	35.000	23.000	26.892	0.000	7.000	36.412	17.624	13.293
Easy Listening	10.000	29.924	30.000	16.000	27.500	32.500	0.000	29.412	27.000	18.293
Electronic	15.118	20.033	26.754	15.833	21.754	23.367	18.754	0.000	20.622	23.889
Folk	22.000	6.667	33.333	10.000	20.000	10.000	12.000	32.881	0.000	18.293
International	5.000	17.576	25.000	11.000	27.500	20.909	10.000	24.412	10.909	0.000
Jazz	10.000	12.424	20.000	14.000	10.000	15.000	8.333	29.412	17.273	15.909
Latin	12.000	24.316	30.000	18.000	21.892	26.892	2.000	31.412	19.202	8.293
New Age	19.697	11.667	31.333	15.000	25.000	15.000	17.000	21.000	5.000	20.000
Pop/Rock	6.364	13.191	18.000	12.364	13.000	16.513	10.000	14.147	11.867	15.135
R&B;	15.000	9.200	29.000	5.000	12.500	12.533	16.516	21.818	17.500	20.909
Reggae	18.894	19.200	30.530	15.000	22.500	22.533	12.000	24.412	24.398	15.000
Religious	23.391	20.167	27.500	18.250	25.750	23.500	27.027	31.174	23.333	30.227
Stage & Screen	15.000	24.091	10.000	21.000	21.667	20.000	5.000	33.333	20.000	23.293
Unknown	26.364	33.191	38.000	32.364	33.000	36.513	30.000	34.147	31.867	35.135
Vocal	27.727	12.667	20.000	26.875	27.500	16.000	20.000	35.000	16.875	22.727

	Jazz	Latin	New Age	Pop/Rock	R&B;	Reggae	Religious	Stage & Screen	Unknown	Vocal
Avant-Garde	26.000	24.340	13.500	13.810	25.324	30.476	27.785	20.000	31.310	24.000
Blues	20.000	8.333	27.754	22.046	15.254	23.333	18.333	28.333	26.626	21.626
Children's	53.053	50.530	50.811	40.000	51.514	56.667	53.976	48.000	57.500	56.267
Classical	20.000	20.000	7.500	20.833	23.750	34.904	30.000	14.000	37.500	18.000
Comedy/Spoken	21.892	10.000	25.000	20.909	12.500	23.654	20.000	28.909	28.293	23.293
Country	16.892	5.000	29.000	22.496	20.000	30.000	15.000	25.000	23.293	18.293
Easy Listening	17.500	10.000	22.000	23.810	31.250	39.500	20.000	20.000	28.293	17.500
Electronic	21.807	19.285	15.769	8.754	10.833	21.987	22.730	16.754	26.254	23.202
Folk	21.892	10.000	17.500	18.734	11.250	22.404	20.000	24.000	28.293	13.333
International	17.500	20.000	18.500	18.810	22.159	21.818	10.000	20.000	10.000	5.000
Jazz	0.000	12.439	21.500	23.423	21.761	25.455	15.000	18.571	25.909	16.476
Latin	11.892	0.000	24.000	24.369	21.250	30.111	10.000	20.000	18.293	13.293
New Age	16.000	15.000	0.000	13.333	16.250	27.404	25.000	21.333	30.000	18.333
Pop/Rock	13.053	10.530	10.811	0.000	11.514	16.667	13.976	8.000	17.500	16.267
R&B;	16.901	14.516	12.500	16.092	0.000	11.154	24.516	19.000	30.909	12.368
Reggae	21.892	10.000	22.500	12.530	10.000	0.000	20.000	20.530	25.000	20.000
Religious	30.080	27.557	25.750	17.027	13.250	24.404	0.000	22.500	34.527	7.500
Stage & Screen	11.667	15.000	27.000	26.078	31.250	37.121	25.000	0.000	33.293	21.304
Unknown	33.053	30.530	30.811	20.000	31.514	36.667	33.976	28.000	0.000	36.267
Vocal	26.667	21.000	20.000	29.449	22.511	22.000	31.000	15.000	32.727	0.000

Here, we enumerate the evolution of two factions.

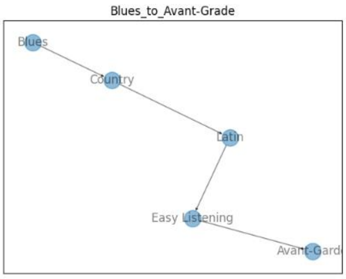


Figure 10. Blues→Country→Latin→Easy Listening→Avant-Garde

The shortest path from Blues to Avant-Garde needs to go through multiple genres, with a weight of 20.33. So such an evolution is difficult, and we think it is almost impossible.

From Electronic to Pop/Rock, the weight is relatively low, which is 14.14, which shows that Electronic is easier to become Pop/Rock after a certain evolution. This is the same as Tangerine Dream (Electronic) in *influence_data* being affected by Pink Floyd, Jimi Hendrix (Pop/Rock) is consistent, in the process of electronic formation, more or less affected by Pop/Rock, so it is more likely to evolve into Pop/Rock later.

The relationship between the two factions, if the evolution time is shorter, the stronger the relationship between the two and the weaker the reproduction. After more middle genres, the weaker the relationship between the two genres.

4.4 The Solution of Problem Four

4.4.1 About Modeling Ideas

First, we analyze the sub-network. Whether the similarity data reported in the *influence_data* data dataset can show that the identified influencers actually affect the respective artists, we plan to use the nominal attribute similarity measurement method to solve the problem. Analysis; in the sub-network relationship between multiple influencers and a follower we proposed in model 1, the follower has a degree of musical similarity to all influencers. In model 2, we are similar to this We have made a measurement model, so that whether influencers have an influence on followers' music can be transformed into characterization by calculating the similarity of these music and whether it is related to the influence of influencers, so we adopt the correlation factor-covariance analysis method Analyze the correlation between influence and similarity; analysis of *full_music_data* and the other two data sets can show that the similarity of a certain follower's music characteristic to that of its influencer, when a certain music characteristic index is similar When it is much higher than other characteristics, it can be considered that some of the musical characteristics are more infectious than others. When the similarity of the various musical characteristics is not obvious, it means that the followers "take all the music" to play similar roles. At this time, we consider using the cosine similarity measurement method to measure the horizontal similarity of each music feature, and use the gray correlation analysis method to characterize the degree of association, so as to solve the problem.

In order to increase the reliability and globality of the model, it is not enough to analyze a specific sub-network alone. For this reason, we select the top 100 highly influential artists selected in model 1 and select 10 samples as the influence In order to compare the difference between the characteristics, we make a double-layer sunburst chart to visually analyze the problem from a global perspective.

4.4.2 Analysis on the Relevance Model of Sub-network

4.4.2.1 Nominal attribute similarity measure

Let the main genre of the follower be O_1, and the main genre of the influencer is O_2. There are a total of n follower-influencer relationship data in the data set. The number of data matched by the two objects O_1 and O_2 is m, then the similarity is the total number of matches in the total number of attributes $\mathrm{sim}(O_1, O_2) = \frac{m}{n}$

We get a similarity of 76.58%, and it can be considered that the similarity data reported in the data set can indicate that the identified influencers actually influenced the respective artists.

4.4.2.2 Cosine similarity measure

First, select nine music characteristic indicators: *Danceability, Energy, Valence, Tempo, Loudness, Acousticness, Duration_ms, Liveness, Speechiness*, and draw the relationship figure between the followers PJ-Harvey and its corresponding influencer indicators in the sub-network. Since it is detailed in the follow-up 5.4.3 that key and mode have little influence on followers, we filter out these two indicators for analysis.

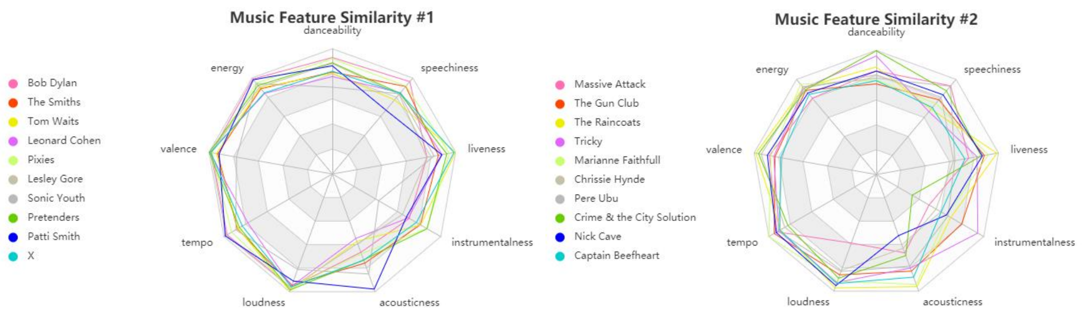


Figure 11. Multi-dimensional music feature similarity

In the following, we will use the cosine similarity measurement method to measure the similarity of each characteristic index of the music of followers PJ-Harvey and its influencers.

For an n-dimensional vector, assuming that A and B are two n-dimensional vectors, A is $[A_1, A_2, \dots \dots, A_n]$ and

B is $[B_1, B_2, \dots\dots, B_n]$, then A and B are The cosine of angle θ is equal to:

$$\cos\theta = \frac{\sum_{i=1}^{n}(A_i \times B_i)}{\sqrt{\sum_{i=1}^{n}(A_i)^2} \times \sqrt{\sum_{i=1}^{n}(B_i)^2}} = \frac{A \cdot B}{|A| \times |B|}$$

The closer the cosine value is to 1, the closer the angle is to 0 degrees, that is, the more similar the two music feature vectors, which means that the two music feature vectors have "cosine similarity".

4.4.2.3 Analysis of Correlation Coefficient and Grey Correlation Degree（GRA）

●Influencer's influence-the correlation coefficient analysis of the similarity of the music of the two parties indicates the degree of correlation

We use the influence of each influencer obtained by model 1 and the music similarity of each influencer-follower obtained by model 2 to perform correlation coefficient-covariance analysis on the two indicators of the sub-network:

First, determine the data column as x={influence of each influencer in the sub-network}, and the comparison sequence as y={the similarity between the music of each influencer and the follower}. The two index data are listed in the following table.

Perform covariance analysis on the above two data column indicators according to the following formula:

$$\mathrm{r} = \frac{\sum_{i=1}^{10}(x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{10}(x_i - \bar{x})^2 \cdot \sum_{i=1}^{10}(y_i - \bar{y})^2}} = 0.3966$$

Since the correlation degree is about 0.4, it shows that there is a correlation between the two indicators of influence and music similarity, but the correlation is weak. In this sub-network, it can be considered that there is a correlation between the influence of the influencer and the similarity of the music of both parties. "Influential people" will influence the music created by followers, and followers maintain their own music creativity and independence to a certain extent.

●Grey Correlation Analysis on the Similarity of Musical Characteristic Indexes of Influencers & Followers

In 5.4.3, we used the cosine similarity measurement method to obtain the similarity of each music feature of the follower and its corresponding multiple influencers in the sub-network, as shown in the appendix, we analyze the degree of gray correlation for each similarity index.

The optimal reference data is selected as y, as shown in the following table, y is the optimal value in each music characteristic index data.

Because the data in each factor column in the system may be different in dimension, it is not easy to compare or it is difficult to get a correct conclusion when comparing. Therefore, the above-mentioned target variables should be non-dimensionalized and the data column should be averaged $x_i(k) = \frac{x_i(k)}{\bar{x_\iota}}$, The meanings of i and k are as shown in the above table. After dimensionless processing, the correlation coefficient is calculated by the following formula. i=1，2，……，20，k=1，2，……，9

$$\varsigma_i(k) = \frac{\underset{i}{min}\,\underset{k}{min}\,|y(k) - x_i(k)| + \rho\,\underset{i}{max}\,\underset{k}{max}\,|y(k) - x_i(k)|}{|y(k) - x_i(k)| + \rho\,\underset{i}{max}\,\underset{k}{max}\,|y(k) - x_i(k)|}$$

Next, calculate the degree of association $\mathrm{r}_k = \frac{\sum_{i=1}^{20}\varsigma_i(k)}{20}$

i=1,2,……,20,k=1,2,……,9

Table 7

artist_name	danceability	energy	valence	tempo	loudness	acousticness	instrumentalness	liveness	speechiness
Bob Dylan	0.929809	0.989175	0.978944	0.970395	0.963929	0.665416	0.709545	0.767035	0.962658
The Smiths	0.814204	0.888569	0.929871	0.875335	0.944179	0.757244	0.811725	0.860545	0.912485
Tom Waits	0.804128	0.897770	0.940211	0.869816	0.993160	0.571916	0.814740	0.998908	0.773014
Leonard Cohen	0.777680	0.840005	0.992770	0.780008	0.954374	0.547787	0.701221	0.890730	0.846677
Pixies	0.919479	0.948344	0.975647	0.929637	0.967293	0.677905	0.872867	0.949560	0.916947
Lesley Gore	0.881011	0.905848	0.988277	0.979591	0.968912	0.589701	0.647318	0.767286	0.828814
Sonic Youth	0.693402	0.943047	0.776727	0.890914	0.810887	0.849811	0.427176	0.825406	0.835686
Pretenders	0.888370	0.922956	0.998592	0.856039	0.984903	0.730113	0.873355	0.928185	0.840644
Patti Smith	0.863749	0.982460	0.918342	0.988088	0.909676	0.977968	0.684335	0.887885	0.665575
X	0.820543	0.845508	0.986224	0.830566	0.958775	0.732825	0.774542	0.987237	0.842984
Massive Attack	0.834075	0.802604	0.874713	0.925344	0.587411	0.671990	0.491168	0.755837	0.932700
The Gun Club	0.729691	0.884466	0.828712	0.950242	0.860511	0.828407	0.795864	0.884144	0.785653
The Raincoats	0.866453	0.913230	0.985302	0.891237	0.972068	0.961954	0.693261	0.987528	0.686486
Tricky	0.955032	0.907314	0.839361	0.906945	0.925810	0.802413	0.944059	0.824869	0.660752
Marianne Faithfull	0.761794	0.948942	0.860715	1.000000	0.915277	0.942203	0.656150	0.917845	0.746120
Chrissie Hynde	0.803422	0.908318	0.942162	0.851960	0.822738	0.633073	0.556037	0.633312	0.812899
Pere Ubu	0.776208	0.923754	0.770744	0.914052	0.831846	0.786634	0.668071	0.643130	0.928177
Crime & the City Solution	0.996820	0.889449	0.963852	0.822299	0.889392	0.695502	0.333999	0.858633	0.879439
Nick Cave	0.833084	0.857537	0.891540	0.932404	0.952582	0.524025	0.652816	0.868552	0.837884
Captain Beefheart	0.756432	0.842166	0.782382	0.897319	0.934723	0.879937	0.604075	0.724790	0.703846

The obtained relevance and the ranking according to the relevance is shown in the figure. Since the relevance of the four indicators of Valence, Liveness, Energy, and Danceability are all above 0.9, which is higher than the relevance of other music feature indicators, we can think that Valence, Liveness, Energy and Danceability are more "infectious" than others.

Table 8

Datasheets									
sequence	6	4	1	3	2	8	9	5	7
index	Danceability	Energy	Valence	Tempo	Loudness	Acousticness	instrumentalnesss	Liveness	Speechiness
Correlation	0.835269	0.902073	0.911254	0.90311	0.907422	0.741341	0.685616	0.848071	0.819972

4.4.3Analysis of the feature difference model of the global network

The characteristics of the artist are given by the *artist_data* dataset, and the top priority is to consider how to measure the influence of the artist on the followers. Inspired by the SGD learning rules in the neural network, influencers are the learning objects of their followers. We imitate the SGD learning rules to establish indicators to measure the influence of influencers on their followers in terms of various characteristics.

In the training and learning of BP neural network, when the output of the network is inconsistent with the target output, this difference needs to be fed back to the network to correct the network weight. The most common feedback method is defined as:

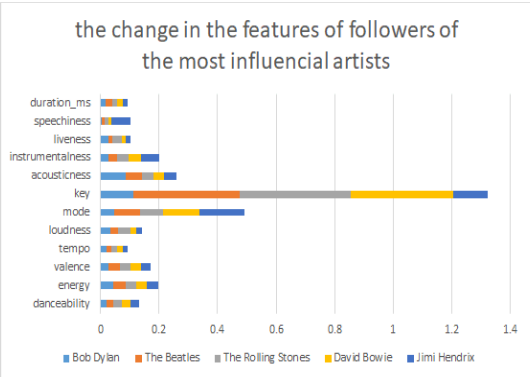
$$\text{Loss} = \frac{1}{2}(output - target)^2$$

The feature vector of the follower is regarded as the learning result of the influencer, and the difference between the feature vector of each follower and the influencer after learning the influencer is quantified in the above way. In the data given by *artist_data*, each artist has 12 feature data. The feature data of influencers and followers are normalized together to obtain the feature vector of each artist. Suppose the eigenvector of the influencer is (x1, x2,..., x12), and the eigenvector of a certain follower is (y1, y2,..., y12), then the change of the i^{th} feature is defined as:

$$Change_i = \frac{1}{2}(x_i - y_i)^2$$

The vector model of the change amount of influencers to followers is obtained above. All the change vector is averaged. The mean change vector reflects the influence ability of the influencer on the follower. At the same time, since the numerical information of each feature is combined, it also reflects the influence of different characteristics on the follower.

To make our analysis more typical, we choose influencers with a certain number of followers to substitute in the calculation. In the first problem, we gave the word cloud of the top 100 influential artists. We chose the top five as the most influencial artists and the last five as the less influencial artists. Observe the evaluation results respectively.



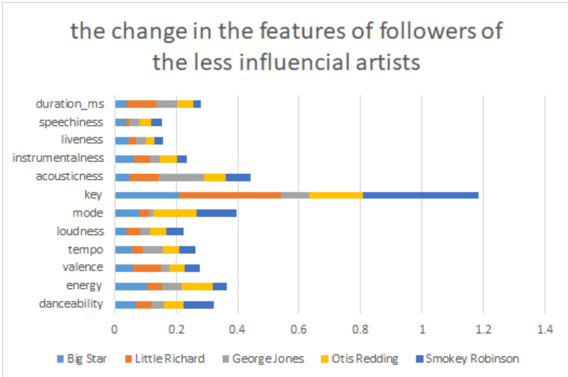


Figure 12 Figure 13

As can be seen in the above figure, more influential artists, such as Bob Dylan, etc., have a smaller follower change vector. It can be concluded that more influential artists will be more strongly Influencing the creation of its followers is manifested as the difference between followers and influencers is smaller.

From the above figure, we can also find that not all features will be imitated by followers. In order to compare the differences between features, we make a double-layer sunburst chart:

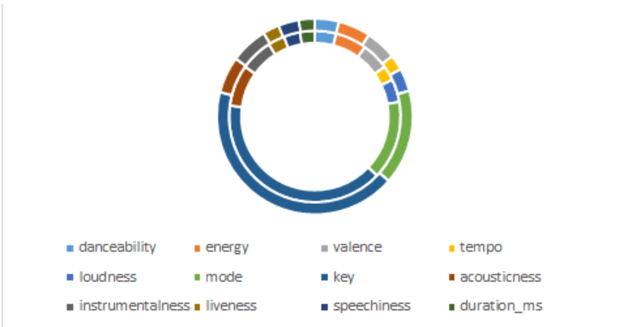


Figure 14. The proportion of the features

In the above figure, different colors represent different characteristics, and the ratio reflects the relative relationship between the changes in characteristics. The outer ring and the inner ring are the most influencial artists and the less influencial artists. The figure shows that keys and mode have a weak influence on followers, which is also in line with music knowledge: most people don't mind if a piece of music is a major or minor key, whether the tonal is C or E flat, the key is absolute but popular The favorite melody is only relative. Except perhaps Sibelius.

The relative strength of the influence of other characteristics is also given in the figure.

4.5 The Solution of Problem Five

4.5.1 About Modeling Ideas

For the problem of finding the major leap characteristics that may exist in the process of music evolution, firstly determine the time of change, and determine the music characteristics that have changed significantly during this period according to the time of change. We plan to draw the three-dimensional influence-year- faction based on the data set according to the coordinate map. According to the time point at which there is a sudden increase or decrease of a certain genre, it can be considered that the moment of these emergencies is the moment when a major historical event occurs. According to the file named by 'data_by_year.csv', draw a genre popularity-time line chart to show the degree of change of various genres over time to determine the degree of change of the genre in this music revolution, and analyze within the genre that has undergone major changes. The artist with the greatest degree of relevance to the change of music characteristics within the genre is determined to be the representative changer to be sought.

4.5.2 Time and faction of major musical changes

We draw a genre-popularity-time line chart according to *data_by_year* to reflect the time of major historical changes, and the visualization results are shown in the figure 15.

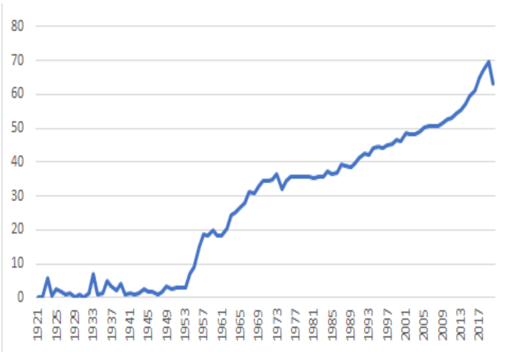


Figure 15. Popularity by year

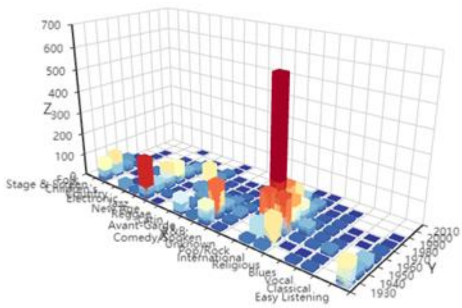


Figure 16. Three-dimensional coordinate map

According to the diagram, we can preliminarily determine that 1957 was a very transformative year in the evolution of music.

In the process of music evolution, music development is affected by technology, economy, culture, environment and other aspects. We exclude other external factors and only conduct research on the field of music, and establish a three-dimensional coordinate map of the degree of influence-year-faction, as shown in the figure 16.

Observing the three-dimensional map, we can intuitively understand that the POP faction developed by leaps and bounds in 1960, and the influence was much higher than that of the factions in each year. By searching for historical events at this point in time,in the mid-1960s, a cultural phenomenon called British Invasion happened , when rock and pop music acts from the United Kingdom[2] and other aspects of British culture became popular in the United States and significant to the rising "counterculture" on both sides of the Atlantic Ocean.[3] Pop and rock groups such as the Beatles, the Rolling Stones, the Who, the Kinks,[4] the Dave Clark Five,[5] Herman's Hermits, the Hollies, the Swinging Blue Jeans, the Animals, Gerry and the Pacemakers, and the Searchers were at the forefront of the "invasion",this result also verifies the judgment of the aforementioned year of change, which is more credible.

4.5.3 Musical characteristics and artists that have undergone major changes in pop sects from 1957 to 1960

According to the data set, we select various music characteristic indicators of pop factions to draw a year-on-year trend chart, as shown in the figure 17.

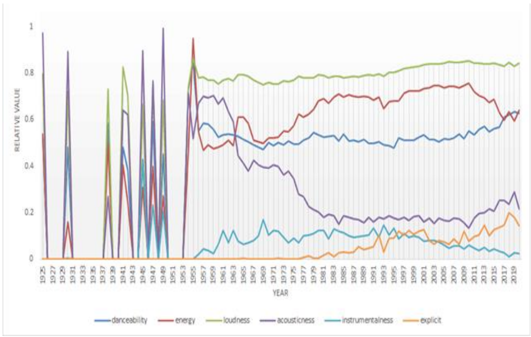


Figure 17. Evolution of Pop/Rock

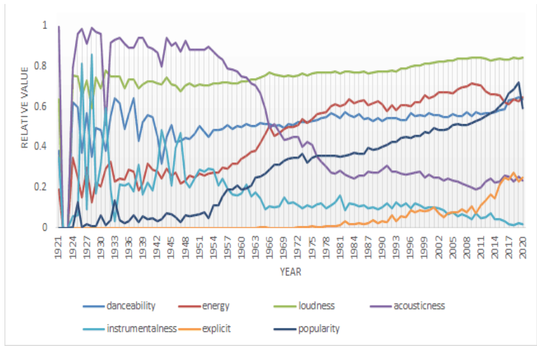


Figure 18. Evolution of the music field

The revolutionary leap of music can be understood in terms of data by reading the data line chart. On the other hand, history tells us that the revolution of music is often closely related to the birth of changers. Therefore, we explore the outstanding musicians at the node of the music revolution to understand the major advances in music.

In the first question, we used the PageRank algorithm to obtain the PR value of each artist on the Internet. The PR value measures the artist's own influence, and changers are more likely to be born from high-impact artists, so we choose Top 3 artists in PR, The Beatles, The Rolling Stones and Bob Dylan, three Pop/Rock artists, coming to explore the relationship between their careers and the progress of Pop/Rock and even the world's music trends.

We took out the characteristics of Pop/Rock and World Music's changes over the years, and obtain interesting information from them. First of all, in the early days of modern music (1921-about 1950), both Pop/Rock music and world music were undergoing rapid changes. Early music did not form a stable style, and various technical directions were being explored. Music features are all being tried. Perhaps this year's popular music skills will be uninterested next year, and their value will be discovered again in another year. Around 1950, that is, shortly after the end of World War II, Pop/Rock music and world music began to stabilize almost at the same time, and the popularity also increased significantly.Mature genres have demonstrated their vigorous vitality after decades of nurturing and baptism of war.

During the 1960s. Pop/Rock music has undergone a major change at this time period, as the popularity of acouticness has fallen sharply at this time, and the emphasis on energy has just emerged. The three most personally influential artists coincidentally debuted almost simultaneously in those few years. Their energy value is unique among contemporary artists, except for Bob Dylan's acousticness, which continues the habits of the old age with a higher value. , The other two have abandoned this technique. The music of world expresses this trend sluggishly. In the following three years, it has undergone major changes similar to Pop/Rock, and was affected by the popular Pop/Rock at the time, and this trend was fixed later.

Table 10

Artist Name	energy	acousticness
Bob Dylan	0.47793	0.562567
The Rolling Stones	0.71992	0.293788
The Beatles	0.54683	0.360356

In the data, we found three talented artists to lead the new trend, which not only happened in their own field but also let the world music absorb their characteristics. The major changes in world music represent the rapid changes in data, and the high-impact artists in our network (such as Beatles, Bob Dylan, Rolling Stones) play the leaders of the moment of change.

4.6 The Solution of Problem Six

4.6.1 About modeling ideas

This problem can be transformed into an analysis of the changing process of influencers' influence and the relationship between key music indicators and music development in a music genre over time. We select a specific music genre (pop/Rock) from the multi-directional network made by model 1 as a sub-network for analysis. First of all, the influence level of each artist in the sub-network can be obtained by the comprehensive evaluation model of model 1. Analyze the changes in the music characteristics of the genre over time, the level of influence, and the changes in popularity over the years, to find out the music characteristics that have changed significantly over time during the development process, and analyze the number and popularity of the new generation of artists. The dynamic influencer index and its influence process are explained in conjunction with the evaluation model index in model1.

4.6.2 Pop/Rock genre development analysis

First find out the Pop/Rock genre subnet based on the global directional network. From 5.5.3, we learned that the pop genre was in its infancy before 1950, and its development was not obvious. With the rapid development of 1957, the influence and popularity of the pop genre increased year by year. This is in line with its historical development environment. There is also a great relationship.

Analysis of the line chart Figure 17 of the changes of various music indicators in the Pop/Rock genre over time. We can know that the music feature indicators have played a vital role in the development of the genre. Therefore, we use this as the dynamic influence index of the music characteristics that have changed greatly in the evolution and development process.

According to the average value of the music index parameters of each genre versus time Figure 18, we can get that in the time period of major changes, the trend of music characteristics representing the entire field is affected by the dynamic influence index, that is, the major music index changes in the pop genre have been affected To the entire music and art field, the music indicator style, and then affect the entire social music development.

4.6.3 Music development and artist development influence process

First draw a comparison trend chart of the number of new generation artists in all fields of the genre and the pop field, and the changes in popularity with respect to time. According to the analysis of the relationship between the number of new generations and the popularity of each music feature in different time series, we can get the popularity of the pop genre The degree is related to the trend of the popularity of the entire music field (r=0.96181), and the number of new-generation artists in the pop genre is related to the trend of the number of new-generation artists in the entire music field (r=0.884123), and the two are relatively compared, and the pop genre changes extent Slightly larger than the changes in the entire music field, the correlation analysis quantitatively shows that the development of pop genres has driven the development of the overall music field.

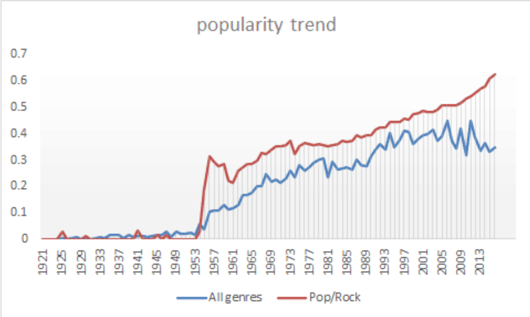


Figure 19. Popularity trend

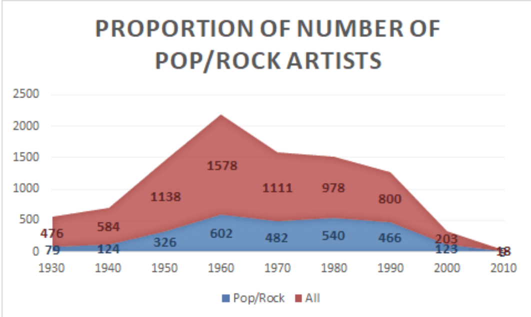


Figure 20. Proprotion of number of Pop/Rock

Combining the foregoing content, we can get a major change in the pop genre, and the major changes in the key music characteristic indicators energy and accousticness map the music characteristics that affect the global music field, leading to the rapid development of the entire music. Moreover, the new generation of artists in both the pop genre and the overall music field reached their peak around 1960, indicating that the pop genre in 1960 caused a large-scale development in various genres and even the entire music field due to representative characters and musical changes. A large number of new-generation artists, and judging the average influence of each year, the results can still support this conclusion. In the following period, the number of new-generation artists has been declining year by year, and the increase in popularity has also continued to decrease (due to popular technology and total number of artists The increase in the number of people, the popularity is still increasing), which also shows that due to the influence of the pop genre, the music field has continued to flourish but the development momentum has declined year by year.

4.7 Music and Society

The background part is supplemented: genres are formed due to commonality, and each composer becomes eternal due to individuality. To understand a series of complex phenomena in the formation of genres, the purpose is to grasp the law of music development from a macro perspective.

The formation of genres is the result of mutual support, mutual influence, and integration of various arts, and its emergence is an objective fact in the development of music. The creative style of musicians in a certain period is bound to be affected by the philosophical concepts, artistic trends, and the influence of social historical phenomena and in the works show a unique style of history.

In Problem 5, we found the major leap characteristics that may exist in the evolution of music. The period 1953-1957 was a very transformative year in the evolution of overall music. World War II ended in 1953-1957 and the economy began to recover.

The plateau period from 1977 to 1983

After the second world economic crisis after the war, the economies of various countries entered a stage of stagflation, the economy rebounded, growth was weak, and unemployment and inflation rates remained high. The capitalist world economic crisis occurred from 1979 to 1982, which is also known as the fifth postwar capitalist world economic crisis. The number of companies closing down in major capitalist countries hit a post-war record, and the unemployment rate reached the highest level since the Great Crisis of the 1930s.

During the crisis, industrial production fell by 11.8% in the United States, 14.8% in the United Kingdom, 7.4% in France, 12.2% in the Federal Republic of Germany, and 41% in Japan. The highest unemployment rate is 10.8% in the United States, 12.5% in the United Kingdom, 8.2% in France, 8.5% in the Federal Republic of Germany, and 2.5% in Japan.

Take the United States and the United Kingdom as examples. The number of unemployed persons in the United

States increased from 6.11 million in 1979 to 12 million, and the unemployment rate was as high as 10.8%, breaking the previous record. Britain fell into crisis in July 1979 and reached its lowest point in May 1981. The industrial and mining production index fell by 12.1%, setting a new record for the previous crisis. Among them, the textile, metallurgy, and construction industries declined the most. Compared with the same period in 1979, the second quarter of 1981 decreased by 29.3%, 29% and 18.6% respectively.

During the crisis, prices in various countries continued to rise, financial conditions deteriorated, interest rates remained high, and world trade shrank severely. In 1980, the US inflation rate was as high as 13.4%. The foreign trade deficit also hit a new record of 36.4 billion U.S. dollars, and the value of exports fell 19.8% from the second quarter of 1981 to the second quarter of 1983.

Trends of thought that emerge in the field of art often spread to other fields soon, and will also have a reaction to time and environment. We are very happy to see that our work can express the influence of music on culture in time or environment.

The smooth rise after the 1990s.

In Problem 5, we also found that after the 1990s, it rose again rapidly and smoothly.

In the late 1980s and early 1990s when culture and economy began to enter the active period, it was also a period when the field of popular music was increasingly emerging. With the development of the Internet media, the Fan economy and Fan culture (which is not very accurate) have also been continuously developed. Fans means fans who are addicted to something. From the beginning, buying albums and watching concerts gave birth to various ways such as buying peripherals (derivative products) for idols, renting advertising spaces for publicity, voting, and charity activities.

Fans have the concept of a group. Fan culture is a conscious search for a sense of collective identity. In his book Understanding Popular Culture, Fisker noticed the cultural group of "fans", thinking that they are excessive readers of popular cultural texts. They are different from the critical and appreciative attitudes of cultural elites towards texts. The input to the text is active and participatory, with strong emotional colors such as enthusiasm and fanaticism.

It is precisely because of Fans' collective identity and active participation that they have unparalleled creativity and productivity. In Question 1 and Question 2, the followers we explored have musical similarity to all influencers, and the appeal and inheritance of the musical characteristics we explored in Question 4 confirm this. New musicians, as fans of old music and old musicians, give us reason to believe that the fan culture will re-create existing genres or lay a new foundation for the emergence of new genres.

We call music art, not only depends on the main characteristics of its art type, but also on its social function. Avant-garde music subverts traditional laws and even breaks people's aesthetic habits, which in turn prompts people to rethink the current social situation.

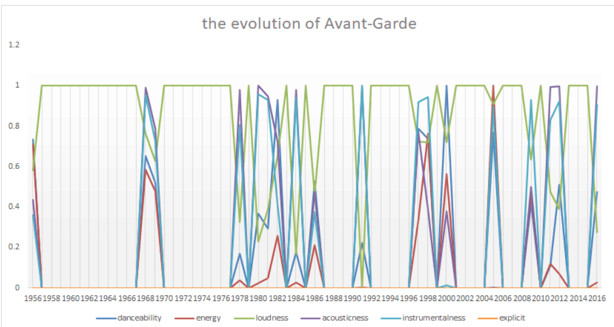


Figure 21. The evolution of Avant-garde

From this figure, we see the evolution of the Avant-garde genre, which is violent and wild.

Avant-gard specifically refers to those creative artists who push the music forward as a whole or at least can create unique ways. The avant-garde does not only refer to a single art school, but generally refers to a variety of art schools that are sequential in time, different in form, and even in opposition. Including avant-garde rock (baiAvant-garde Rock), avant-garde rock (baiAvant-garde Rock) refers to artistic exploration that uses techniques and expressions that are fundamentally opposed to traditional forms.

The formation of the concept of avant-garde music can be traced back to the Renaissance Art Nouveau (Ars Nova). Avant-garde composers generally have the courage to try various methods, and music is their laboratory. Their art, their innovative spirit was not welcomed for a long time at that time, However, it has made a lasting contribution to the development of art in the future, promotes the evolution of art, and is accepted and adopted in the development of human culture in the future.

5. Conclusion

•The directional network created based on the data set is shown in the figure 1. We choose a follower to analyze the multi-influencer sub-network (figure 2), and successfully characterize its own level of influence by calculating the PR value, and in order to characterize the influence of influencers. Regarding the intensity of influence of a specific follower, we consider multiple indicators such as artistic competition and genre inheritance to establish a gray comprehensive evaluation model to measure influence. Among the most influential artists, the top three are Beatles, Bob Dylan, and Rolling Stones. The "features of music influence" here is measured by the following indicators: the level of influencers directly/indirectly influence the number of people, the inheritance of genres, the number of people in each directional network field, the distribution of the number of people in the field by year, and the integration of the various indicators of the intensity of competition in the year and other indicators. The improvement of objectivity through the entropy method reflects the influence of music to a large extent and has practical significance. This sub-network takes arbitrary followers as the research object, so our model also has good scalability.

•By comparing the scatter matrix diagram classified by genre and the scatter matrix diagram classified by cluster, we deny that genre is the essential classification of musical works. Then, in the analysis and explanation of the clustering situation, we reveal that musical works are classified more essentially based on their cohesion in their feature space, and the decisive factor influencing music classification is the evolution of human society in terms of aesthetics, trends, and techniques. The similarity measurement model has been detailed in 5.2. Artists within genres are not more similar than artists between genres. Only electronic music has better similarity.

•Through the calculation and analysis of the heat map, we found that there is often a consistent preference for the use of skills within genres, but not among genres. We explained why speechiness and liveness are so highly correlated in the heat map of the comedy/spoken genre, and found a causal relationship in it. We explained why the heat map of the avant-garde is clearly different from other schools, and lamented the avant-garde's bold attempts at various techniques. In order to understand the changes of genres over time, we have established a directed graph between genres and let them evolve in time. From this, we understand the mutual reference and transformation between genres, and discover that the reference relationship among certain genres is closer than other factions.

•According to the nominal attribute similarity measurement method, its similarity reached 76.58%. It can be considered that the reported similarity data can show that the identified influencers actually influenced the respective artists; according to the correlation analysis of influence-music similarity, the correlation between the two indicators is reached 0.3966. It can be considered that “influential people” will really affect the music created by followers; gray correlation analysis can be done based on the cosine similarity of the various music characteristic indicators of the influencer-follower It is obtained that the correlation between the four music characteristics of Valence, Loudness, Tempo and Energy and the optimal parameter index is above 0.9, which is higher than the correlation between other music characteristics and the optimal index, and the correlation between Key and mode is very small, so it can be considered that Valence , Loudness, Tempo, Energy are more "infectious" than other features.

•A major leap has taken place in the evolution of music. Around 1957, the pop genre has developed by leaps and bounds. The analysis shows that the key music characteristic index that affects its rapid development is energy and acousticness. In the directional network, the top three most representative changers are named Bob Dylan, The Rolling Stones and The Beatles.

•For detailed conclusions, see 5.6 & 5.7

6 Sensitivity Test

- In previous section, we assume that the distinguishing coefficient ρ = 0.2 to reflect the ranking scores of the musicians more differently. Therefore, we need to change the value of ρ to see if the value of ρ will exert significant influence on our model.

Figure 22 show the rank of music influence when ρ is equal to 0.1, 0.2, 0.3 respectively.

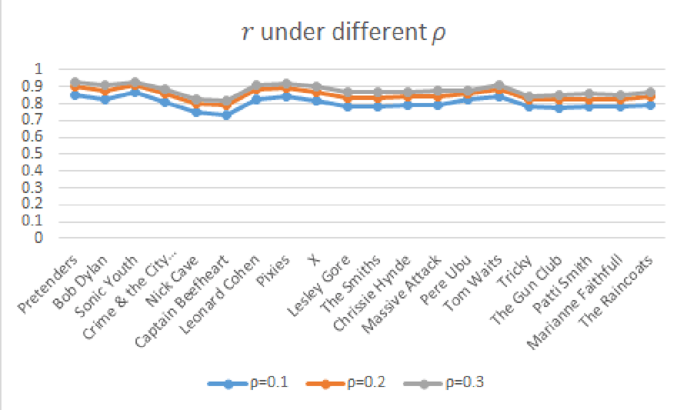


Figure 22. Rank of music influence (ρ =0.1,0.2, 0.3)

By comparing the above three tables, we find that the rank of music influence are identical, which means that our model is insensitive to the change of ρ .

- In the previous research of music similarity based on K-means clustering, we determined K=2 as the optimal value according to the Calinski-Harabasz criterion, and considering that K=4 is the sub-optimal choice, we try to change the value of K to see if the value of K will exert significant influence on our model.

Figure 23 shows the clustering of music data when K = 2, while Figure 24 show the clustering of music data when K is equal to 4.

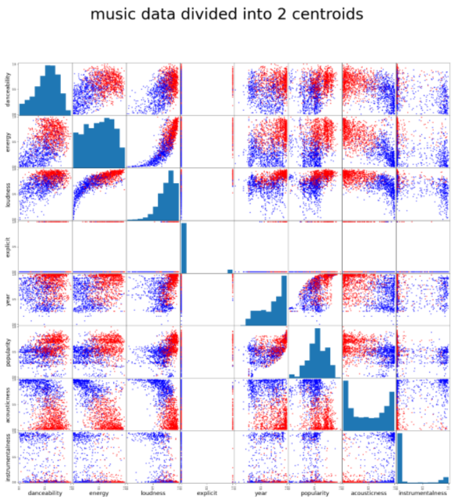


Figure 23. Clustering of music data (K = 2)

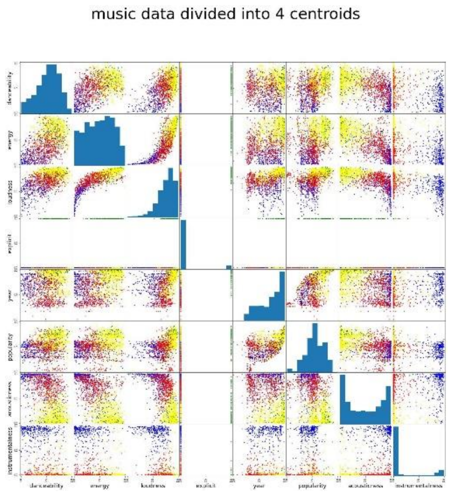


Figure 24 Clustering of music data (K = 4)

Comparing the two figures, it can be seen that if they are clustered into four categories, one of them is quite rare, while the other three are the mainstream; and the newly emerging category is in the middle of the original two categories, which shows that there is a transitional period for change, and the same before and after the transitional period It is divided into different categories, which is consistent with the previous analysis and further supplements the details of the music revolution.

7 Strengths and Weaknesses

7.1 Strengths

- It is convenient and broadly applicable to use our model to select influential people in various fields. Utilizing the grey correlation analysis method, our model has high precision and is convenient for researches to assess the influence of an industry figure. The five factors we select can be universally adopted when analyzing different similar fields.
- Our model has good stability. Our outcomes are nearly identical when different method of analyzing stability and sensitivity is adopted.
- We use an easy-to-understand algorithm to explain the connection between influencers and followers and conduct a comprehensive analysis of the data so that the changes in music characteristics can be observed more intuitively.

7.2 Weaknesses

- Using correlation coefficient to analyze correlation is not comprehensive enough
- The analysis of discrete music features is not accurate enough

8 A Letter to the ICM Society

Dear ICM Society:

Our team has developed a method to quantify music evolution, thus understanding and measuring the impact of previously produced music on new music and music artists. Through the establishment of a Gray Comprehensive Evaluation Model to measure the music influence, thereby investigating the evolution and revolutionary trends of artists and genres. We measure the degree of influence of each music indicator according to the relevance degree.

Through the quantification of time series, we explore the evolution and revolutionary trend of music. On the basis of the model, we also considered the influence of external factors such as environment and politics on the evolution of music, and analyzed both the artist sub-network and the global network to adapt to more music development situations and satisfy the model's mobility and value.

First, Let us give you some explanations about the basic model. In the music artist's influence relationship orientation network, we selected multiple indicators that can measure the artist's music influence, including the number of influencers and the level of competitiveness. We establish a Gray Comprehensive Evaluation Model through Entropy Method, and quantitatively measure the influence of influencers on followers through the balanced evaluation of multiple artist attributes. According to this model, we obtain 100 influential master artists, including Beatles, Bob Dylan, and Rolling Stones.

In the artist influencer-follower directional network that we created, each mapping connection is measured by influence. Simultaneously, the inheritance of music is a process from influencer to follower. In this process, followers with great influence increase, and their inheritance ability increases correspondingly. The inheritance with music as the carrier will inevitably cause some similarity of some musical characteristic indicators to a certain degree, leading to the similarity of artist music. At this time, the similarity is measured by the similarity model. The significance of the influence is measured by the correlation model. Music spreads between different genres through the network. When a certain faction develops by leaps and bounds, the entire network will have a chain reaction to drive the continuous development of the overall music field.

Since the content of the data set provided in the problem solved is limited to certain genres, and there are restrictions on shared artists, we improve the local network and the overall network, quantitative characterization and visual analysis in the model so that with the introduction of more or more abundant data, our model can also play a role in measuring impact. At this time, we improve the Comprehensive Evaluation Model according to the type of data introduced, add data-related indicators to the measurement indicators of the model, and use the entropy method to assign weights and model to obtain the corresponding influence. The main improvement is reflected in the model evaluation system. The introduction and weighting of new indicators.

In the article, we not only solve the problem quantitatively from the perspective of mathematics, but also analyze and connect the relevant results in combination with historical environment, economic and political factors, and social background. For details, please refer to section 5.6 & 5.7 of the article. With the development of music, social culture has continuously developed and enriched. We warmly welcome your suggestions for our model and analysis.

References

[1]Page L, Brin S, Motwani R, et al. The PageRank citation ranking: Bringing order to the web[R]. Stanford InfoLab, 1999.[pagerank]

[2]Wu N. The maximum entropy method[M]. Springer Science & Business Media, 2012.

[3]Gull S F, Skilling J. Maximum entropy method in image processing[C]//IEE Proceedings F (Communications, Radar and Signal Proces sing). IET Digital Library, 1984, 131(6): 646-659.

[4]Sullivan, Danny,"What Is Google PageRank? A Guide For Searchers & Webmasters". Search Engine Land. Archived from the original on 2016-07-03.

[5]Heimerl F, Lohmann S, Lange S, et al. Word cloud explorer: Text analytics based on word clouds[C]//2014 47th Hawaii International Conference on System Sciences. IEEE, 2014: 1833-1842.

[6] Kuo Y, Yang T, Huang G W. The use of grey relational analysis in solving multiple attribute decision-making problems[J]. Computers & industrial engineering, 2008, 55(1): 80-93.