Problem Chosen	2021 MCM/ICM Summary Sheet	Team Control Number
C		2102422

Influence model of music artists based on Knowledge Map

Since ancient times, music has been an important part of human society. To understand the role music plays in the human community, it is necessary to develop a way to quantify music development. In this article, we build mathematical models from three perspectives: musicians, the process of music and music genres evolving with time, and try to quantitatively and comprehensively analyze the music development process.

In the model for musicians, we first use TOPSIS comprehensive evaluation method and entropy weight method to establish a musician influence measurement model, and establish a knowledge map for all musicians, so that we can intuitively analyze the influence and communication between musicians. In the knowledge map, we mark the nodes of musicians of different schools with different colors, and use the TOPSIS comprehensive evaluation method to evaluate the influence of each musician in the circle, outside the circle and comprehensive influence. Finally, we compare the characteristics of the two musicians who are most closely related to each other, and finally determine that the influencer will actually influence the works created by the followers.

In the model for music, we build a music similarity measurement model based on Pearson correlation coefficient, through the processing and analysis of music indicators. We find that the works of musicians of the same genre are indeed more similar than those of musicians of different genres. At the same time, it is found that musicians do have a significant impact on their followers, and their musical characteristics are highly similar. It is found that the influence of dance ability and value on followers is particularly obvious.

In the model for time series, we extract the features of a specific music collection (such as a singer's works, a genre's works) in each year, and form a line chart that can directly show the evolution process of the current music collection features. It can be found that the music style before the 1960s was a hundred schools of thought, and all kinds of schools had not yet been finalized. However, with the anti war doctrine brought by the Vietnam War and the emergence of a large number of outstanding innovative singers, their music style gradually stabilized.

Key word: Knowledge graph, Correlation analysis, Time series analysis

Content

1 Introduction

1.1 Background

Music is an artistic form and cultural activity whose expression medium is the sound organized on time. Its basic elements include height, strength, duration, timbre and so on. Since ancient times, music has been an important part of human society and is closely related to our daily life. A good music can fully stimulate all kinds of human emotions, and can give people great belief.

For a long time, the study of music has never stopped. Previous work has been done through the analysis of large music data, as long as a specific initial value is input, the machine can automatically compose music that meets the requirements. However, at present, there is no in-depth study on the quantitative analysis of the development and dissemination of music.

1.2 Overview of our work

One of the biggest challenges is the complexity and diversity of music data. In this paper, we develop a series of models to quantitatively analyze the spread and development of music from different perspectives, such as music, musicians, music schools and so on. We have also developed a new assessment framework to quantify the influence of musicians. Finally, we use the time series model to predict the future development of music genres, and focus on the analysis of the deviation, to explore the disturbance of a variety of special circumstances on music indicators.

2 Hypothesis and Symbol

To facilitate the understanding and reading of this article, we will show the basic assumptions and common mathematical symbols involved in this article here. In the article, we will emphasize again when we have related items.

2.1 Hypothesis

In order to clarify the specific scenario of the problem we are studying and eliminate the excessive complexity of the model, we make the following assumptions in this paper:

Hypothesis 1: The data given by the title are true and correct.

Hypothesis 2: The report provided by the artist does not have the influence of civility or politics.

Hypothesis 3: Artists do not create cross-border works frequently

2.2 Symbol

In order to avoid too many repeated explanations of mathematical symbols in this paper, we list and describe the commonly used mathematical symbols in this paper. The rest of the symbols will be explained when used.

Table 1: mathematical symbols commonly used in this paper

Symbol	Definition	Type
S	All artists	aggregate
T_i	Song genre	string
$V_{a,b}$	A artists influence B artists	vector
S_a	Influential artists	aggregate
S_b	Affected artists	aggregate
fet_i	Different characteristics of songs	vector
t_i	Time interval	data

For the mathematical formula appearing in the text, we will mark the number of the formula on the right side of the page so that we can quote it when mentioned below.

3 Establishment and Solution of Model

In this section, we deal with the data, use TOPSIS comprehensive evaluation method and entropy weight method to establish the evaluation model of musicians' influence, so as to achieve the quantitative analysis of musicians' influence ability. Then, a knowledge map of influence relationship between musicians is established to show the subordination and influence degree between influencers and followers. Finally, taking the Beatles and the kinks as examples, it shows that musicians do have an impact on their followers' works.

3.1 Basic concepts

We regard each musician as a node in the relationship network, the influence between musicians as a directed path with weight, and the degree of influence is the weight of the corresponding path.

Each musician has its own genre attribute, and the difference of its influence within and outside its corresponding genre is very significant. Therefore, when we build the influence model of musicians, we use three indicators: in circle influence, out of circle influence and comprehensive influence. These three indicators respectively show the influence degree of musicians in different specific groups, and can explain the "music influence" of musicians from more angles.

3.2 Singer influence scoring model

We start with a quantitative analysis of the data given to the influencers and followers. Here we use Topsis method to build the evaluation model and the entropy weight method to determine the weight.

3.2.1 Selection and Preprocessing of Basic Indicators

Topsis method is a common comprehensive evaluation method, which can make full use of the information of the original data, and the results can accurately reflect the differences between the evaluation schemes.

In the sample with n evaluation indexes and m objects, the original matrix can be obtained: $X = (x_{i,j})_{m\times n}$ .The four initial indicators we selected are: the number of artists affected as $x_{i,1}$ 、 number of people in your field as $x_{i,2}$、 year ranking of influence in your field as $x_{i,3}$ 、 year competitiveness as $x_{i,4}$.

Because the four selected indicators are already very large, the initial matrix is the positive matrix.

$$X' = X \tag{2}$$

Finally, the standardized matrix $X'' = (x''_{i,j})_{m\times n}$ is obtained by the formula $x''_{i,j} = \frac{x'_{i,j}}{\sqrt{\sum_{i=1}^{n} x'_{i,j}}}$.

3.2.2 Obtaining entropy coefficients

The basic idea of Entropy Weight Method is to determine the objective weight according to the index variability. Generally speaking, the smaller the information entropy of an index, the greater the variation of the index value, the more information provided, the greater the role it can play in the comprehensive evaluation, and the greater its weight. Conversely, the larger the information entropy of an index, the smaller the variation of the index value, the less information provided, the smaller the role played in the comprehensive evaluation, and the smaller the weight.

According to the normalized matrix, the information entropy of the j evaluation index can be obtained as follows: $s_j = -\frac{1}{\ln m}\sum_{i=1}^{m} d_{ij} \ln d_{ij}, \quad j = 1,2,\cdots,n$

Thus, the entropy weight coefficient of item j is obtained: $\omega_j = 1 - s_j \Big/ n - \sum_{j=1}^{n} s_j$

Table 2: Relevant Index Weight

Factor	x_1	x_2	x_3	x_4
Weight (ω_j)	0.191793	0.420922	0.442092	0.171262

3.2.3 Calculate final score

Define the minimum and maximum values first:

$$Z^+ = (\max\{z_{11}, z_{21}, \ldots, z_{n1}\}, \max\{z_{12}, z_{22}, \ldots, z_{n2}\}, \ldots, \max\{z_{1m}, z_{2m}, \ldots, z_{nm}\})$$

$$Z^- = (\min\{z_{11}, z_{21}, \ldots, z_{n1}\}, \min\{z_{12}, z_{22}, \ldots, z_{n2}\}, \ldots, \min\{z_{1m}, z_{2m}, \ldots, z_{nm}\})$$

Define the distance between the i evaluation object and the maximum and minimum value:

$$D_i^+ = \sqrt{\sum_{j=1}^{m} \omega_j (Z_j^+ - z_{ij})^2} \qquad D_i^- = \sqrt{\sum_{j=1}^{m} \omega_j (Z_j^- - z_{ij})^2}$$

Final score is obtained:

$$S_i = \frac{D_i^-}{D_i^+ + D_i^-}$$

Here we show the top five data:

Table 3: Top Five Musicians with Comprehensive Score

Singer's name	*Genre score*	*Out-of-Genre Score*	*Comprehensive score*
The Beatles	1	0.701149425	1
Bob Dylan	0.580470163	0.770114943	0.61952862
The Rolling Stones	0.547920434	0.172413793	0.501683502
David Bowie	0.403254973	0.16091954	0.365319865
Led Zeppelin	0.383363472	0.091954023	0.336700337

3.3 Building Knowledge Map

The data from the above analysis can be sorted out to get the following figure:

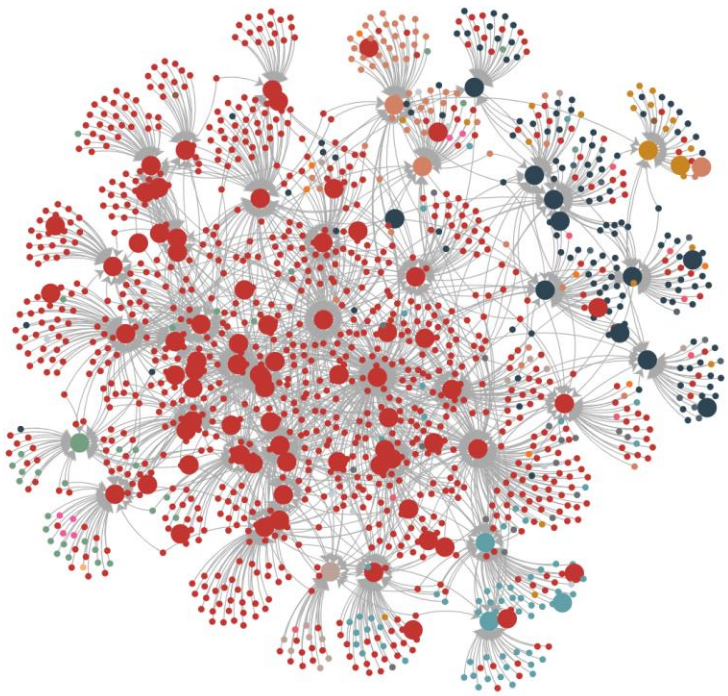
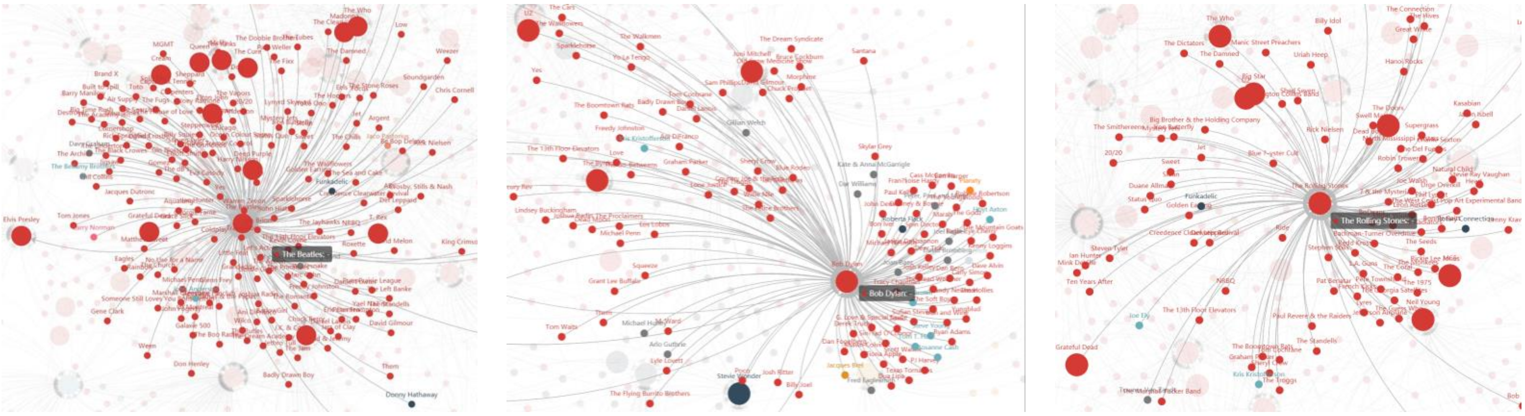


Figure 1: musicians influence knowledge map

This is the map of knowledge that is constructed in which the nodes of different genres of musicians are labeled with different colors, and the size of the points is the influence of the musician. It can be seen from the observation that the regional division of the map is very obvious. The periphery of the entire network is a number of small path clusters, each of which represents a more influential singer and its affected general singers. Inside the graph is a complex network with the main path cluster dominated by the super-influential singers (The Beatles) and the secondary path cluster dominated by many sub-influential singers.

In this diagram, the red node represents the genre of the musician as Pop/Rock, the dark blue node represents R&B, the light green node represents Electronic, the light pink node represents Blues, the Yellow node represents Vocal, and the blue node represents Country. Looking at the pictures, we can see that most of the popular singers of the late 20th century music genre belong to the Pop/Rock genre.Each genre appears centrally, with distinct zoning on the graph, such as an obvious gap between R&B and Pop/Rock. Compared with other genres, each genre has closer internal connections and closer communication.

According to the previous analysis, the top three most influential singers are: The Beatles, Bob Dylan and the rolling stones. We have extracted their corresponding partial networks, as shown in the following figure:



The Beatles Bob Dylan The Rolling Stones

Figure 2: part of the network of the top three most influential singers.

The Beatles is a British rock band, whose music style originated from the rock music of the 1950s, and developed psychedelic rock, pop rock and other music styles. He won 14 Grammy Awards. They are at the top of the Rolling Stone magazine's "50 greatest pop musicians in history" list and the US Billboard's "most outstanding musicians Hot 100" list. The band made a variety of innovations in the rock music at that time, breaking many narrow formats at that time, and making the rock music more profound in connotation.

In our calculation above, we obtained the relevant parameters of the Beatles:

Table 4: Beatles influence score

Singer's name	*Genre score*	*Out-of-Genre Score*	*Comprehensive score*
The Beatles	1	0.701149425	1

It can be found that the conclusion of our model is really in line with the actual situation, and can really reflect the real situation.

3.4 Musician homogenization evaluation model

In order to measure the impact of musicians' works on the works of the affected people, we use correlation analysis to explore the correlation between various indicators and popularity by using the "full_music_data" data set. The higher the correlation is, the more appealing it is. And take the

Beatles as an example to analyze, according to different indicators, check the similarity degree of indicators between the influencer and the affected.

According to the statistics of the "influencer man" data set, pop / rock is the most popular genre, accounting for almost 90%. Therefore, we choose singers and bands under this genre as the basis of statistics, which can also make the statistics representative and extensive.

We counted the most frequent singers or bands in the "influencer_man" feature, and defined them as influential singers or bands. For these influential singers or bands, we started from "full_music" In the data set, we counted the number of songs produced by each singer or band, and judged their song style and type by looking at their songs. In order to make a comparison of song styles, we chose the band with the most influence, a large number of songs and mutual influence. These two bands are the Beatles and the Kinks. The two bands have outstanding characteristics in personal influence and composition. Their style is very distinct, so they are chosen as representatives of their own style.

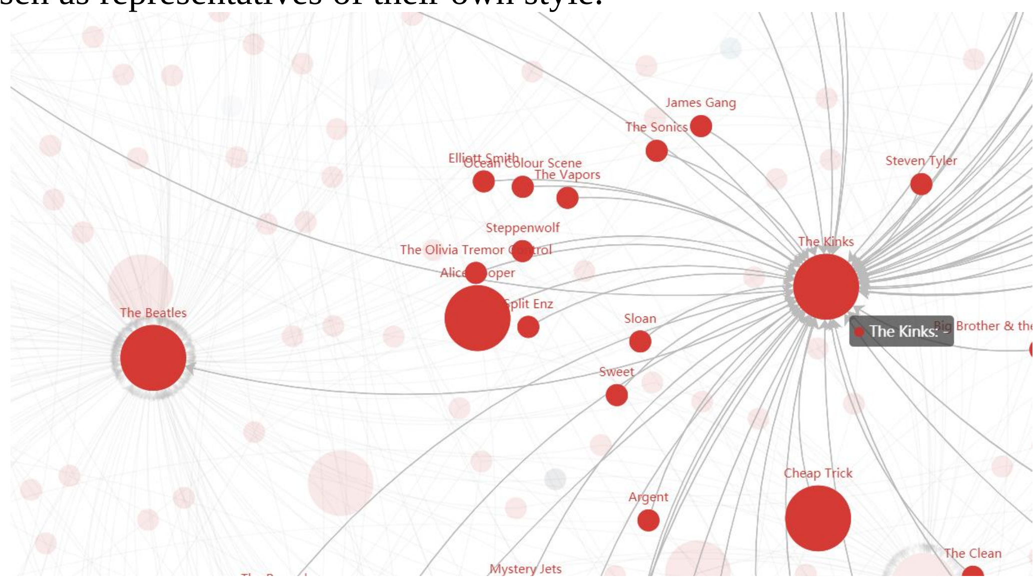


Figure 3: Node diagram of relationship between the Beatles and the Kinks

We can see from the graph that the Beatles really influenced the kinks, and they all have a lot of followers.

Taking this as the starting point, we collected the songs of the two bands in the ‘full_music_data’ data set, and made a statistical analysis of all their songs. First, we sorted all their songs according to the year from small to large. After sorting, we made a statistical analysis of all the indicators of all their songs, because these indicators are specific descriptions Songs, so there may be a sudden increase or decrease in a certain index, which will affect the final statistical data, so we made a median statistical analysis of these data

Table 5: The median of each characteristic of the two bands (part)

	year	loudness	energy	explicit	danceability	tempo	duration_ms	valence
The Beatles	0.5355	0.5495	0.661	121.899	-9.1745	1	4	0.3055
The Kinks	0.561	0.759	0.7165	119.501	-7.4975	1	5.5	0.231

Based on the above data, we can draw the radar chart of the songs of the two bands

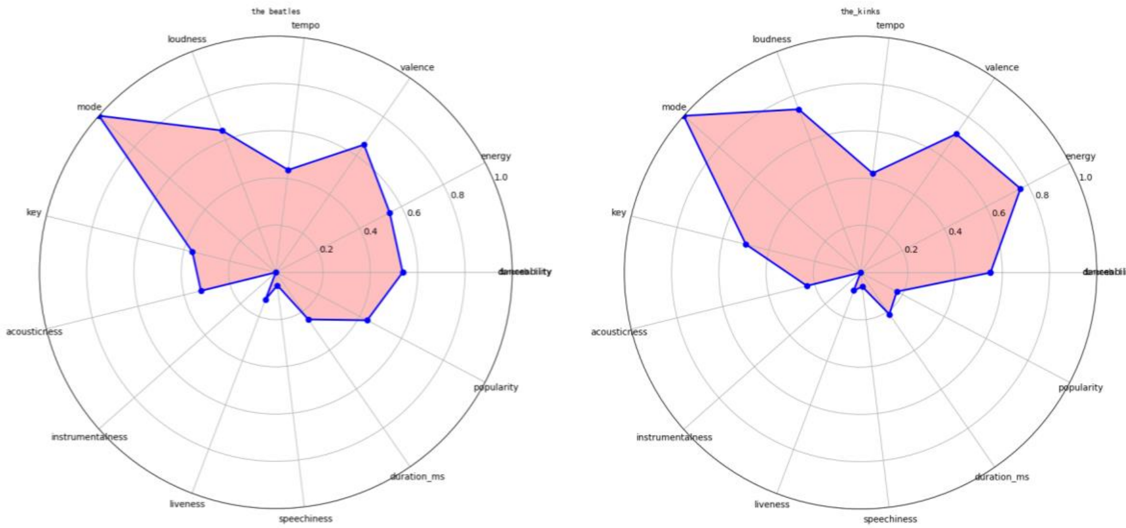


Figure 4: Radar chart of two bands' music styles

From the legend, we can see that the kinks influenced by the Beatles are almost the same as the Beatles in terms of style, and they are very close to each other in terms of indicators. After being influenced by the Beatles, the kinks gradually approached the former in terms of composition.

So we come to the conclusion that influencers do actually influence the music that their followers make

4 Music based measurement model

In this section, we build a music similarity model. First, using the principal component analysis method to explore the correlation between the indicators and popularity, take the main related variables, song characteristics include not limited to structure, rhythm or lyrics. Then, according to the main related variables, Pearson correlation coefficient is used to calculate the degree of similarity between different music, and draw the thermal graph of correlation relationship, which directly reflects the similarity between music. Then the k-means algorithm is used to cluster the artists, and the point graph of correlation is drawn to intuitively view and analyze the correlation between the artists in the genre and the artists in the genre.

4.1 Music similarity measurement model

4.1.1 Principal component analysis

In this section, we use principal component analysis to process the given music data.

Principal component analysis is essentially a multivariate statistical analysis method which transforms several indexes into several comprehensive indexes by orthogonal rotation transformation on the premise of little information loss. Using dimension reduction method to better grasp the main contradictions and improve the efficiency of analysis. The steps of using this method to solve the weight of each factor are as follows:

There are j evaluation indexes, the number of samples is T, $X_{T\times k}$ is formed according to the original sample data, and $R_{k\times k}$ is the covariance matrix of k index sequence, where $\lambda_i(i=1,\cdots,k)$ represents the i characteristic value of matrix $R_{k\times k}$, and $\alpha'_{k\times 1}$ represents the i characteristic vector of matrix $R_{k\times k}$. $PC_i = X\alpha'$ represents the i principal component and $\lambda_i = Var(PC_i)$. According to the principal component analysis, the weight of the j index is determined as: $\omega_j = \sum_{i=1}^{k}\lambda_i\alpha'_j / \sum_{i=1}^{k}\lambda_i$

The results are as follows:

Table 6: Beatles influence score

	year	loudness	energy	explicit	danceability	tempo	duration_ms	valence
weight	0.773853	0.422792	0.38665	0.186826	0.175577	0.090769	0.056057	0.020778

The popularity of the topic is greatly influenced by time in terms of computation. Therefore, the term "year" should be excluded.

Then all the features of music are screened. Firstly, we select the pop / rock music style data, take the "influencer man" in the "influence data" data set as a reference, find out the pop / rock songs in the "full music data" data set, make a summary of them, and then reduce the dimension of PCA. The data processed by PCA are shown in the following table

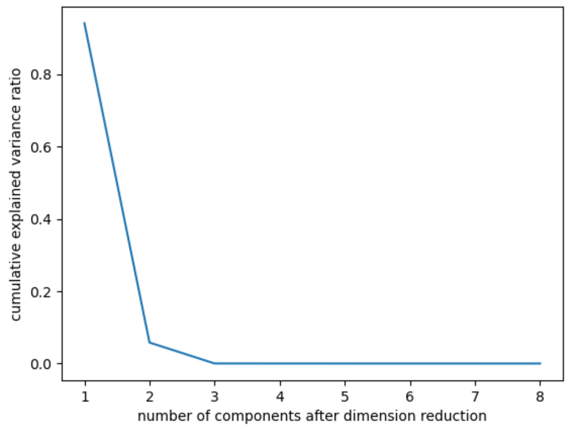


Figure 5: Analysis results of PCA dimension reduction results

It can be seen that the first two features can already cover the characteristics of the whole data, so we select the first two data features as the main data sources. In order to intuitively observe the data distribution of these two features, we use scatter plot to draw, and convert them into thermal graph for analysis.

Figure 6: Eigenvector scatterplot

Observation pictures can confirm that these two data are indeed the main feature sources.

4.2.2 Pearson correlation coefficient

On the basis of the previous paper, we analyzed the correlation of the other eight characteristics except the prevalence, so as to determine which factors are influencing the prevalence. Here we use Pearson correlation coefficient for quantitative analysis of correlation.

Pearson correlation coefficient is obtained by dividing covariance by the standard deviation of two variables. Although covariance can reflect the correlation degree of two random variables (when covariance is greater than 0, it means positive correlation and when covariance is less than 0, it means negative correlation), its value is greatly affected by dimension, so it is not easy to judge the correlation degree of variables from the value of covariance. In order to eliminate the influence of this dimension, Pearson correlation coefficient is used. The calculation formula is as follows.

$$p(x,y)=\frac{\sum x_i y_i - n\overline{xy}}{(n-1)_{SxSy}}=\frac{n\sum x_i y_i-\sum x_i \sum y_i}{\sqrt{n\sum x_i^{\ 2}-(\sum x_i)^2}\sqrt{n\sum y_i^2-(\sum y_i)^2}}$$

According to Pearson correlation coefficient, the correlation matrix is established to study the correlation between the other eight features except popularity, and the feature vector scatter diagram and thermal diagram are drawn.

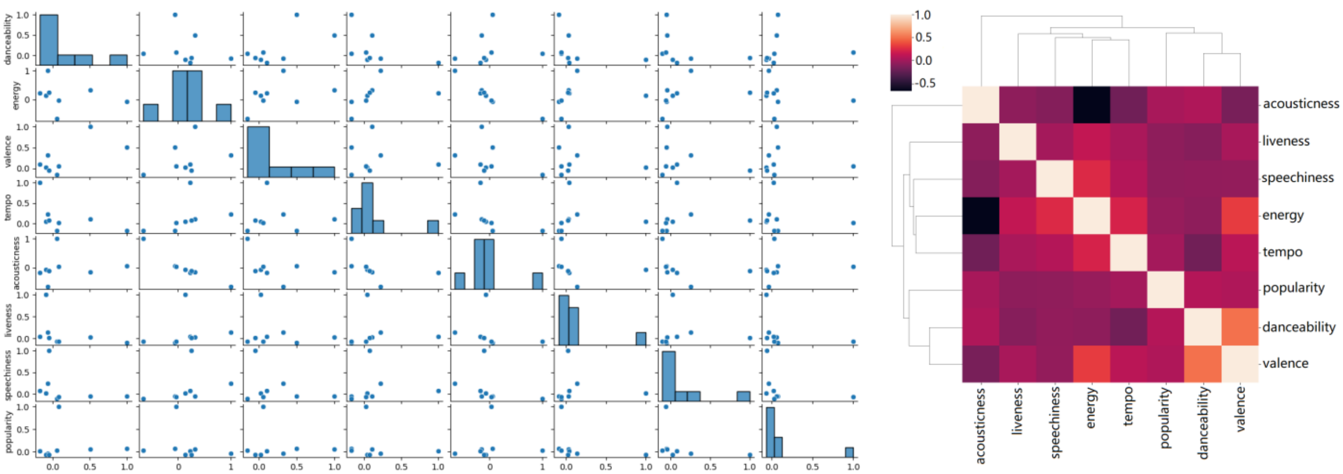


Figure 7: Correlation analysis diagram and its thermodynamic diagram

By observing the above images, we can find that only dance ability, energy, loudness, acuity and explicit are positively correlated with the epidemic degree,

4.2 Similarity between genres

Based on the above research, we extracted the music data of each genre, pretreated it, and got multiple music indicators for each genre. As shown in the following table:

Table 7: Musical Traits of Four Genres (Part)

	danceability	energy	valence	tempo	loudness	key	acousticness
pop/rock	0.5394	0.6667	0.5643	0.5001	0.7525	0.4545	0.1466
R&B	0.627	0.548	0.658	114.44	-9.454	6	0.325
country	0.5955	0.45	0.6285	114.15	-11.13	5	0.518
Jazz	0.6002	0.2934	0.5118	0.5379	0.7148	0.4545	0.8342

Based on these four sets of data, we can draw the following radar charts:

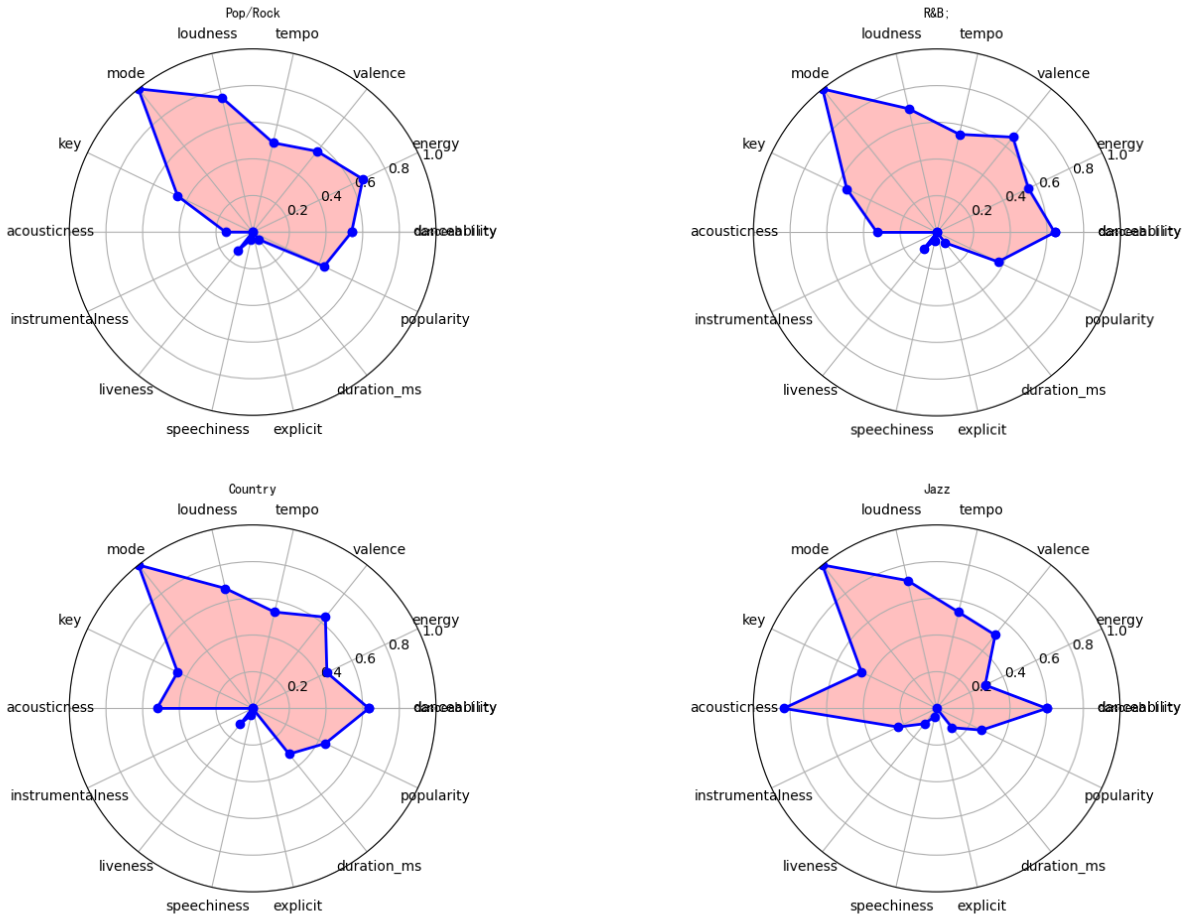


Figure 8: Radar Group of Musical Features of Four Genres

By observing the images, we can find that for R & B, county and jazz, if compared with pop / rock, we can find that dance ability is larger and energy is smaller. But they all have a high loudness, and they tend to be consistent graphically, which indicates that music is gradually integrating with each other, but it retains its own characteristics

5 Model based on time series

5.1 Analyzing the style of songs in time series

Our modeling of time series is a comprehensive time series analysis of songs of different years, songs of different styles and songs of different singers.

For the macro, we first choose to use the data in "data by year". First, we have some macro judgment on the wind transformation. For the data with different characteristics, we first choose to use the normalization method to scale the range. The features with a large range of data, such as' speech 'and' explicit ', or the features with negative numbers, or the features with 0 and 1, are scaled or deleted. Finally, the appropriate data set is obtained, and the centralized processing is done on this data set. The statistical analysis is as follows

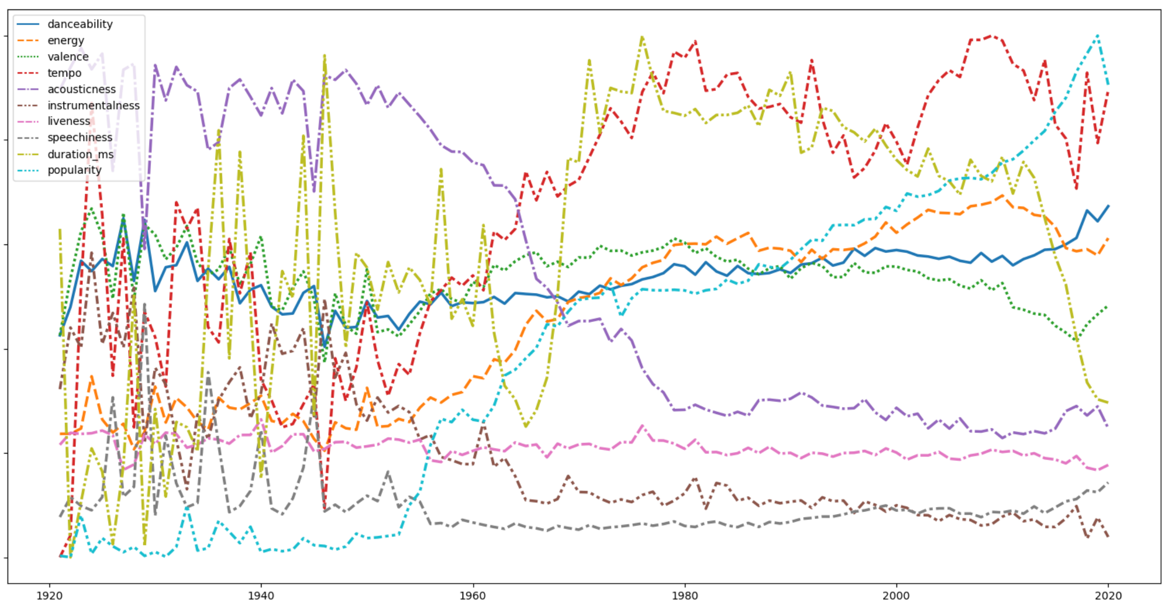


Figure 9: Line charts of music characteristics evolving over time

We can see from the statistical analysis that the music style was still erratic before 1960, when the music circles were looking for their own styles, so all kinds of songs with different characteristics prevailed. With the rise of anti-war sentiment in the Vietnam War around 1960, people have a deeper yearning for world peace, the cultural thoughts of the "collapsed generation", and the change of people's view of the world after World War II, which made many bands in the music world have a large degree of style unity. At that time, anti-war songs headed by Bob Dylon were prevalent. Most of the songs at that time were also pop/rock style. People expressed their yearning for world peace here. This was an unprecedented trend at that time, so that the singers or bands who came out in the following decades were deeply influenced by them, so they also established the pop/rock style mode.

5.2 Analysis of Music Traits of Different Genres with Years

On the macro basis, we can see that pop/rock style songs have a large market share, so we make a detailed quantitative analysis of the songs under Pop/Rock style, 'Full_Music_Data' data is a very large dataset that contains all songs. We sampled 20,000 samples evenly from it, and filtered out pop/rock songs, then analyzed them according to time series.

We analyzed these data, as follows, we counted the trend of music in different fields in the same two characteristics, and the results are as follows:

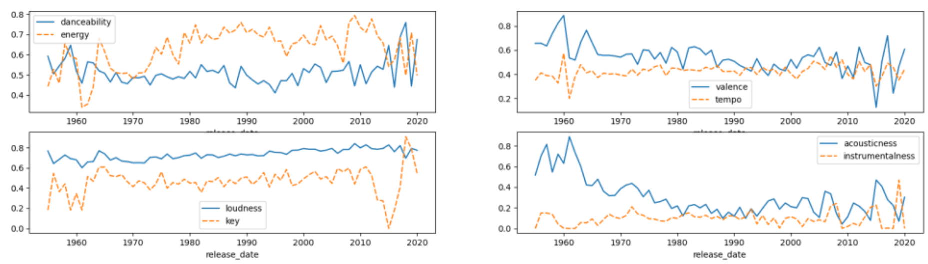


Figure 10: Music in different fields has the same trend.

We can see that the styles of these songs were different around 1960, but in modern times, the styles of songs began not to be limited to themselves, and all kinds of characteristics began to merge.

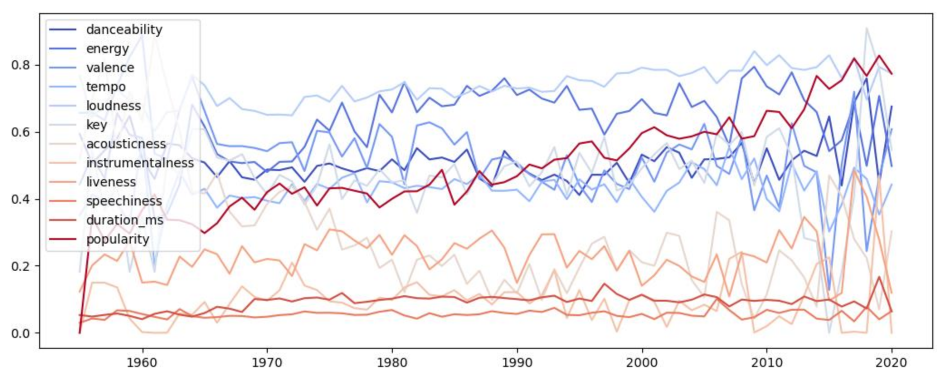


Figure 11: The characteristics of all the songs of the same genre over the years

As shown in the figure above, we made a statistical analysis of all songs of the same type, and found that songs of the same type will also become diversified under the influence of modern music style, and the characteristics will change greatly. In the same field, it also indicates that songs in the field are also trying to match different characteristics

From a statistical point of view, the future music style will tend to be diversified, whether it is in agreement with the field or between different fields, will become diversified

6 Evaluation and extension of models

6.1 Advantage

- The model integrates various factors to ensure the stability and rationality of the results.
- In this paper, for different aspects of the problem to be considered, combined with the characteristics of each aspect, the mathematical model is established, which has relative rationality and good generalization.

6.2 weakness

- Data limitation: the limitation of this report is that it only uses the data provided by the stem.
- The model in this paper is more based on the analysis of the results of mathematical model, less considering the political, economic, cultural and other factors in reality.
- Musicians themselves are not a stable and changeable thing. In reality, there are unpredictable changes brought about by their personal accidents or decisions.

6.3 Ideas for improvement and promotion

Generally speaking, the model is universal and can be further promoted. It can be used not only to explore the influence of music system and musicians and their schools, but also to evaluate other cultural activities with strong communication and long history.

In the specific implementation, the model can be modified according to different actual background, the parameter setting can also change at any time according to the actual situation, and the data can be processed reasonably according to the specific analysis. According to the core method of the model, the problem can be solved more conveniently.

References

[1] Fengguohe, Zhu Siming. Cluster-based large sample support vector machine research [J]. Computer Science, 2006 (04): 145-147.
[2] Fengguo and. SVM classification kernel function and parameter selection comparison [J]. Computer Engineering and Application, 2011,47(03): 123-124+128.
[3] Wang Jiangpan, Guo Qiang, Liu Jiangguo. Author contribution allocation algorithm based on PageRank's co-author paper [J]. Journal of University of Electronic Science and Technology, 2020,49(06): 918-923.
[4] Li Yangtao, Li Chuan, Xu Chao, Lei Xiao, Xu Hongyu, Tang Changjie,Yang Ning. Spatial Scoring: Node Importance Measurement in Information Network Visualization Based on PageRank [A]. Chinese Computer Federation. Thesis Collection (B) (NDBC2012) [C] of 29th Chinese Academic Conference on Database. Chinese Computer Federation: Database Professional Committee of Chinese Computer Society, 2012:8.

Report

Music has always played an important role in human history, so the quantitative analysis of music is also very important. In order to quantitatively explore the value of music from multiple perspectives, we first established a directed weighted network with musicians as nodes. Then it makes a comprehensive analysis from the time series and the degree of influence of transmission, and finally, it analyses the impact of music on political and technological changes. We have got some meaningful results, which we hope will contribute to the development and dissemination of music.

From the network we built, it is intuitive to see that the mainstream music has been Pop/Rock for the past 90 years, followed by R&B. These two genres together accounted for more than 90% of the total number of musicians and music works, which also shows that the most influential music at that time was Pop/Rock. Quantitative analysis of the musician's influence has calculated that the most influential musician in the past 90 years was the Beatles. The band influenced 615 singers, only 41 of whom were not Pop/Rock, suggesting the Beatles 'indelible position in the rise and spread of Pop/Rock. Combining with the historical facts at that time, the Beatles had different degrees of innovation in the form and connotation of songs, making a great contribution to the development of Pop/Rock. The Beatles also ranked first in our musician impact assessment model, indicating that our model really worked well.

Historically, our model reflects the impact of the anti-war boom caused by the Vietnam War in the 1960s on all major schools.And since the 1960s, the music world has been full of talent, becoming the leader in their respective fields, and the music style of each genre has stabilized.

Appendix

Some Important Codes

```
def guiyi(df):
    df = (df - df.min()) / (df.max() - df.min())
    return df


# %%
influence_data = pd.read_csv('./influence_data.csv', encoding='gbk')  # Pop/Rock
# %%

# Pop/Rock          24141
# R&B;              5530
# Country           3301
# Jazz              2716
# Vocal
gener_name = 'Pop/Rock'
pop_man = influence_data[influence_data["influencer_main_genre"] ==
gener_name].loc[:,
          ['influencer_id', 'influencer_name', 'influencer_main_genre',
'influencer_active_start']]
pop_man = pop_man.drop_duplicates("influencer_name")

# %%
RB_man = influence_data[influence_data["influencer_main_genre"] == "R&B;"].loc[:,
         ['influencer_id', 'influencer_name', 'influencer_main_genre',
'influencer_active_start']]
RB_man = RB_man.drop_duplicates("influencer_name")
RB_man.index = np.arange(len(RB_man))
# %%
country_man = influence_data[influence_data["influencer_main_genre"] ==
"Country"].loc[:,
              ['influencer_id', 'influencer_name', 'influencer_main_genre',
'influencer_active_start']]
country_man = country_man.drop_duplicates("influencer_name")
country_man.index = np.arange(len(country_man))
# %%

full_music_sample = pd.read_csv('./full_music_data_sample.csv')
full_music_sample["artist_names"] = full_music_sample["artist_names"].apply(eval)


def get_x(x):
    return x[0]
```

```
full_music_sample["artist_names"] = full_music_sample["artist_names"].apply(get_x)

# %%
#
full_music_sample[full_music_sample["artist_names"].isin(pop_man["influencer_name"
])]

pop_rock_song = pd.read_csv('./pop_rock_song.csv')

# %%

songs =
full_music_sample[full_music_sample["artist_names"].isin(pop_man["influencer_name"
])]
songs_feature = guiyi(songs.loc[:, ['danceability', 'energy', 'valence',
                          'tempo', 'loudness', 'mode', 'key', 'acousticness', 'instrumentalness',
                          'liveness', 'speechiness', 'explicit', 'duration_ms', 'popularity']])
songs_feature = songs_feature.median()
# %%

labels = songs_feature.index.tolist()

dataLenth = songs_feature.shape[0]

data = songs_feature.values.tolist()

angles = np.linspace(0, 2 * np.pi, dataLenth, endpoint=False)
labels = np.concatenate((labels, [labels[0]]))
data = np.concatenate((data, [data[0]]))
angles = np.concatenate((angles, [angles[0]]))

fig = plt.figure()
ax = fig.add_subplot(111, polar=True)
ax.plot(angles, data, 'bo-', linewidth=2)
ax.fill(angles, data, facecolor='r', alpha=0.25)
ax.set_thetagrids(angles * 180 / np.pi, labels, fontproperties="SimHei")
ax.set_title(gener_name, va='bottom', fontproperties="SimHei")
ax.set_rlim(0, 1)
ax.grid(True)
# %%
country_song =
full_music_sample[full_music_sample["artist_names"].isin(country_man["influencer_na
me"])]
RB_song =
```

```
full_music_sample[full_music_sample["artist_names"].isin(RB_man["influencer_name"]
)]
# %%
country_song_feature = guiyi(country_song.loc[:, ['danceability', 'energy', 'valence',
                                            'tempo', 'loudness', 'mode', 'key', 'acousticness',
                                            'instrumentalness',
                                            'liveness', 'speechiness', 'explicit', 'duration_ms',
'popularity']])
RB_song_feature = guiyi(RB_song.loc[:, ['danceability', 'energy', 'valence',
                                  'tempo', 'loudness', 'mode', 'key', 'acousticness',
'instrumentalness',
                                  'liveness', 'speechiness', 'explicit', 'duration_ms', 'popularity']])
# %%
country_song_feature = country_song_feature.median()
RB_song_feature = RB_song_feature.median()

# %%

labels = RB_song_feature.index.tolist()

dataLenth = RB_song_feature.shape[0]

data = RB_song_feature.values.tolist()

angles = np.linspace(0, 2 * np.pi, dataLenth, endpoint=False)


ax.set_thetagrids(angles * 180 / np.pi, labels, fontproperties="SimHei")
ax.set_title("RB_song_feature", va='bottom', fontproperties="SimHei")
ax.set_rlim(0, 1)
ax.grid(True)
```