

Data Analysis and Modeling for Measuring the Impact of Music

Longzhen Yu, Pengfei Zhu and Wandai Li

School of Nanjing, Hohai University, Nanjing 211100, China

Abstract

In order to understand the role of music in human collective experience, we need to develop a method to quantify the evolution of music. To this end, we use the grey comprehensive evaluation method to establish a model to measure the influence of music: First, we connect various artists through the influence of music, draw a knowledge map and trace all the data. We count the direct and indirect influences of influencers and get the masters of the genre. Then we take the total number of influential artists, ranking of influence, the number of people in the genre during the period, and the total number of people in the period as evaluation indicators, and use AHP to determine the weights to establish a grey comprehensive evaluation model. After that, we use PCA and cosine similarity to establish a music similarity measurement model: First we use PCA to reduce the dimensionality of the data, extract three principal components, calculate the cosine similarity between them, and draw a heat map of the correlation relationship. According to the genre of the artist, cluster analysis is carried out to reflect the relationship between artists within the genre and artists of different genres. We further dig into the data and analyze the similarities and influences between genres and genres. Taking Pop/Rock as an example, we analyze the rise and fall of genres and their changes over time by drawing a line chart. Finally, according to the genre of influencers and followers, the relationship between genres is analyzed. In order to find out which music features are more appealing, we establish a correlation analysis model: first, the data is tested for normal distribution, and then the Pearson correlation coefficient between the sample data is calculated. Taking the indicators of some artists and his influencers as an example, a radar chart is drawn, which intuitively reflects the influence of people with musical influence on his followers. To understand the major changes in the development of the genre, we draw a line chart in chronological order of the total number of people affected by the masters of the same genre, and analyze the outstanding changes in the work indicators. Through the above analysis, we submit a music influence measurement model, a music similarity measurement model, and a correlation analysis model to the ICM, thus reflecting the impact of music on the social and cultural environment.

Keywords

Grey fuzzy comprehensive evaluation model; cosine similarity; PCCs.

1. Introduction

1.1. Problem background

Music has been part of human societies since the beginning of time as an essential component of cultural heritage. As part of an effort to understand the role music has played in the collective human experience, we have been asked to develop a method to quantify musical evolution. There are many factors that can influence artists when they create a new piece of music, including their innate ingenuity, current social or political events, access to new instruments or tools, or other personal experiences. Our goal is to understand and measure the influence of previously produced music on new music and musical artists.

Some artists can list a dozen or more other artists who they say influenced their own musical work. It has also been suggested that influence can be measured by the degree of similarity between song characteristics, such as structure, rhythm, or lyrics. There are sometimes revolutionary shifts in music, offering new sounds or tempos, such as when a new genre emerges, or there is a reinvention of an existing genre (e.g. classical, pop/rock, jazz, etc.). This can be due to a sequence of small changes, a cooperative effort of artists, a series of influential artists, or a shift within society.

Many songs have similar sounds, and many artists have contributed to major shifts in a musical genre. Sometimes these shifts are due to one artist influencing another. Sometimes it is a change that emerges in response to external events (such as major world events or technological advances). By considering networks of songs and their musical characteristics, we can begin to capture the influence that musical artists have on each other. And, perhaps, we can also gain a better understanding of how music evolves through societies over time.

My team is willing to explore profound findouts to establish a music influence model based on the dataset provided by ICM.

1.2. Problem analysis

We divide the problem into the following specific analysis, which can be divided into the following five subsections according to requirements:

Find the relationship between influencers and followers, create a knowledge graph, establish a directional network, and establish a comprehensive evaluation model of influence;

Establish a similarity measurement model through the analysis and processing of the data set;

Analyze the degree of similarity between influencers and influenced persons, and use correlation analysis to explore the relationship between various indicators and popularity;

Analyze the rise and fall of genres over time, and major leaps in the evolution of music;

A letter to the ICM Association.

1.3. The concept of the model

Direct influence: musicians directly influenced by influencers.

Indirect influence: the influence of followers of influencers on other musicians as influencers.

2. The Description of the Problem

2.1. Restatement of the problem

Use the influence_data data set or portions of it to create a (multiple) directed network(s) of musical influence, where influencers are connected to followers. Develop parameters that capture 'music influence' in this network. Explore a subset of musical influence by creating a subnetwork of your directed influencer network. Describe this subnetwork. What do your 'music influence' measures reveal in this subnetwork?

Use full_music_data and/or the two summary data sets (with artists and years) of music characteristics, to develop measures of music similarity. Using your measure, are artists within genre more similar than artists between genres?

Compare similarities and influences between and within genres. What distinguishes a genre and how do genres change over time? Are some genres related to others?

Indicate whether the similarity data, as reported in the data_influence data set, suggest that the identified influencers in fact influence the respective artists. Do the 'influencers' actually affect the music created by the followers? Are some music characteristics more 'contagious' than others, or do they all have similar roles in influencing a particular artist's music?

Identify if there are characteristics that might signify revolutions (major leaps) in musical evolution from these data? What artists represent revolutionaries (influencers of major change) in your network?

Analyze the influence processes of musical evolution that occurred over time in one genre. Can your team identify indicators that reveal the dynamic influencers, and explain how the genre(s) or artist(s) changed over time?

How does your work express information about cultural influence of music in time or circumstances? Alternatively, how can the effects of social, political or technological changes (such as the internet) be identified within the network?

2.2. Clarification of the problem

For question 1, we created a directional network based on the data set. The nodes are influencers or followers, and they are connected through the influence of music to draw a knowledge graph. We trace all the data, count the direct and indirect influence of influencers, and calculate the masters of the genre. After analyzing the data, the total number of influencers, the ranking of influence, the number of people in the genre during the period, and the total number of people during the period are used as evaluation indicators. AHP is used to determine the weights and a gray comprehensive evaluation model is established.

For the second question, after dimensionality reduction processing of the data through the PCA method, three principal components are extracted, the cosine similarity between them is calculated, and the correlation heat map is drawn. According to the genre of the artist, cluster analysis is carried out to reflect the relationship between artists within the genre and artists of different genres.

For question three, we further dig into the conclusion of question two, analyze the similarities and influences between and within the genre, the rise and fall of the Pop/Rock genre over time and its changes by drawing a line graph. Finally, according to the genre of influencers and followers, the relationship between genres is analyzed.

For question four, we use the Pearson correlation coefficient to explore the correlation between each indicator and popularity. Higher correlation means that the corresponding music features are more appealing. After that, we selected some artists and his influencers, and drew a radar chart for their indicators to directly reflect the influence of people with musical influence on his followers.

For question five, we understand the inflection point of the genre development that is the major change point based on the line chart drawn at the time of solving the problem, and then screen the works of that period and analyze the outstanding changes in the work indicators. Finally, combining the knowledge graph and the line graph, we can find the artists who lead the change.

For question six, we analyze the influence of the three principal components extracted from the second question to determine the influence of the index on the genre. By sorting all the works of a genre in chronological order and performing linear regression analysis, we can use popularity to measure the development of artists.

For question seven, we first took the Internet as an example to analyze the influence of the Internet on music in the Internet. Then, through the changes of music over time, we analyzed its impact on music by the Renaissance in the 1940s and the Industrial Revolution in the 1960s.

3. Models

3.1. Symbols and definitions

Table 1 symbols and definitions

| SYMBOLS | DEFINITIONS |
|---------|-------------|
|---------|-------------|

| | |
|-----------------|--|
| I_a | Total Music Influence Index of Influencer A |
| I_{ttc} | Follower C's relative influencer A's music influence index |
| M | Target layer |
| C | Criterion layer |
| P | Scheme layer |
| λ_{max} | Maximum eigenvalue |
| ω | Weights |
| α | Best reference data column |
| R | Covariance matrix |

3.2. Assumptions

In order to make the model relatively simple and reasonable enough, we make the following assumptions:

If the followers of an influencer influence those who have its own followers, then the influencer has a certain influence on each follower, but the influence is not necessarily the same.

When a follower is influenced by multiple influencers, we assume that the influence is equally divided.

We only consider the influence of artists of the predecessors on the artists of the later generations (that is, the artists who are older have an influence on the artists of the later age)

Music influence is determined by the total intensity of influence, which depends on the total number of influencers, and has nothing to do with other factors.

3.3. Comprehensive evaluation model of influence based on grey comprehensive evaluation

3.3.1 The basis of the model

As to the definition of musical influence, first of all, we believe that if an influencer directly or indirectly influences many followers, then his musical influence should be relatively high.

In other words, the musical influence of an artist depends on the number of followers it affects and how many people these followers affect. Secondly, we believe that the rise and decline of different genres in different periods will also affect artists in a certain period.

To sum up, we believe that the musical influence of an artist in a genre depends on the total number of people he influences and his age.

The signs and definitions are mostly generated from queuing theory.

3.3.2 The establishment of knowledge graph

When the data set is too large, the image of the knowledge graph is not clear. We use a part of the data set to create a directional network, with nodes as influencers or followers, and connect with music influence to draw knowledge graphs of five typical genres.

Taking into account the large number of music types and the differences between different types of music, we select the more representative five typical areas of Avant-Garde, Classical, Comedy/Spoken, Easy listening and New age in the Influence_Data dataset to draw a knowledge map, where the nodes are followers and influencers in the same music field, and they are connected in the form of musical influence.

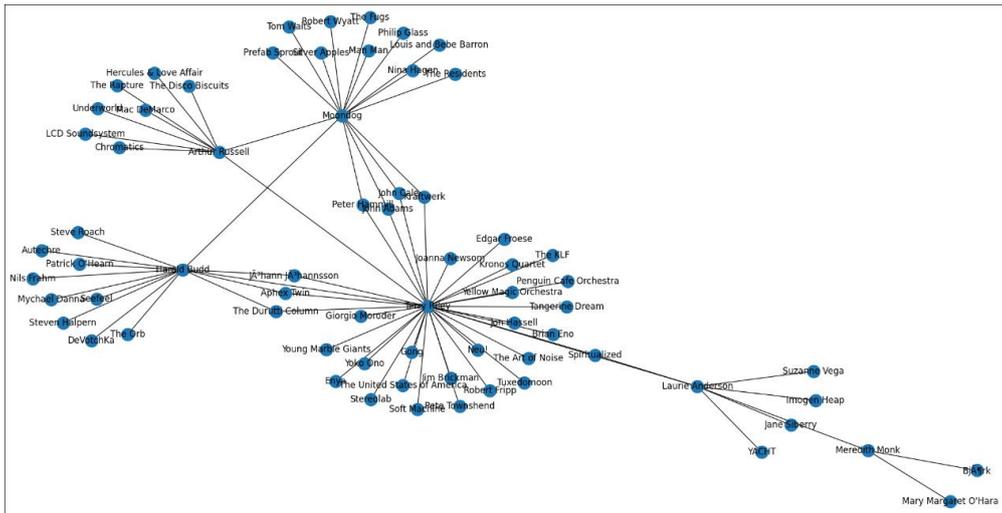


Figure 1 Avant-Garde

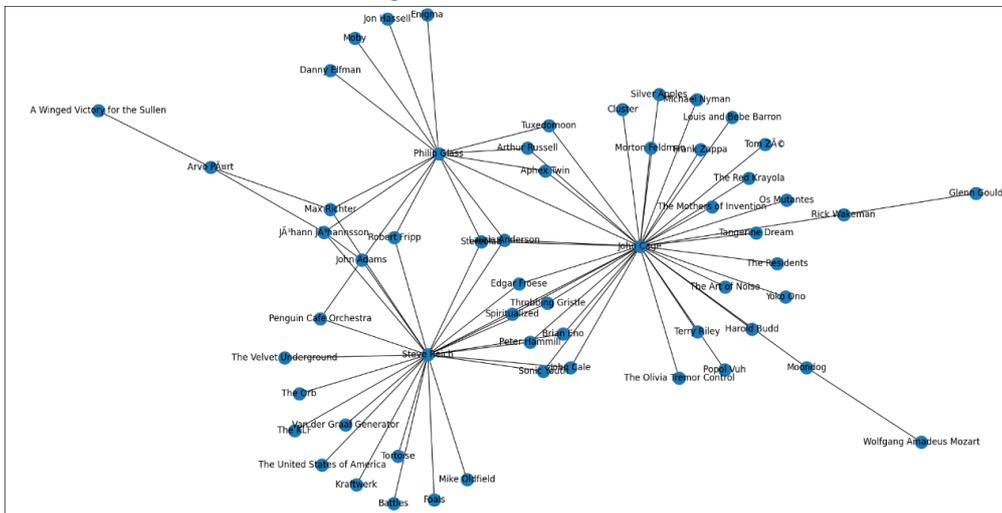


Figure 2 Classical

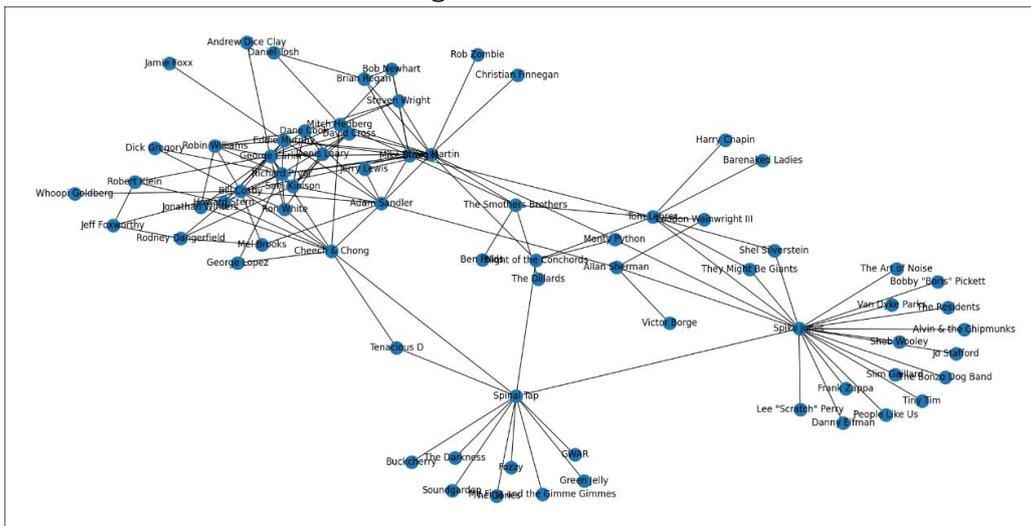


Figure 3 Comedy/Spoken

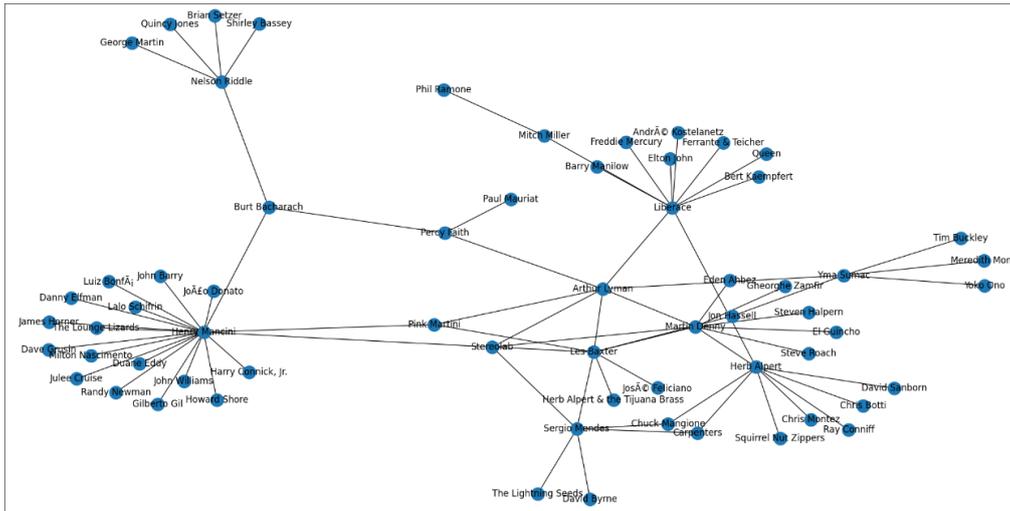


Figure 4 Easy listening

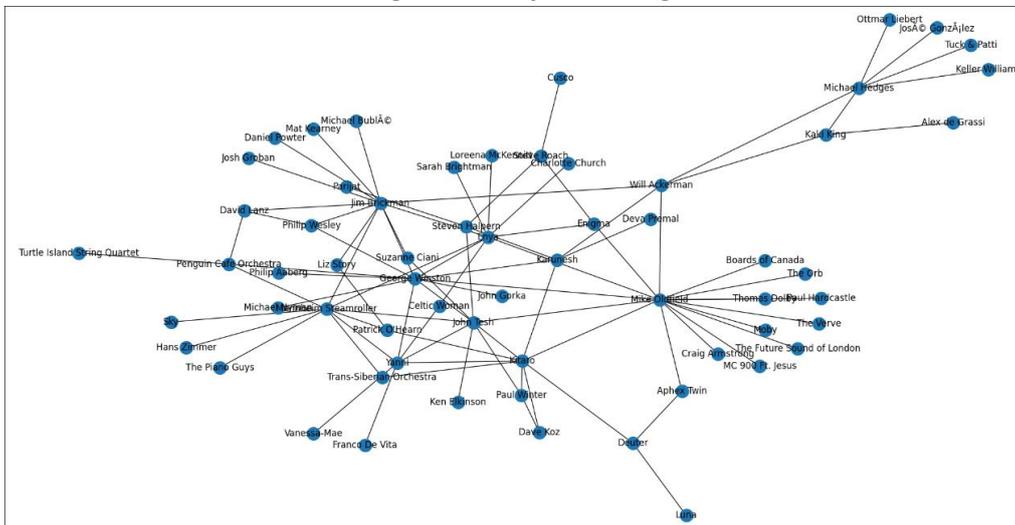


Figure 5 New age

3.3.3 Music influence based on the total number of people affected

As to the definition of musical influence I_a , we give the following formula:

$$I_a = \sum I_{ttc}$$

Artist A is an influencer, and C is a follower who is directly or indirectly influenced by A. Then the influence of A on C can be defined as:

$$I_{ttc} = \frac{1}{n} \rightarrow c$$

Among them, n is the total number of influencers that caused follower C to be affected.

The logical explanation of the formula is: the musical influence of an influencer depends on the number of followers it affects, and how many influencers these followers will be influenced by.

3.3.4 Music influence based on age

After processing the data set, we list the total number of influencers of different genres of art masters and corresponding masters, where the total number of influences is the sum of the direct and indirect influences, and is plotted as a table.

We draw a line graph of the total number of people affected by masters of the same genre in chronological order to visualize the results.

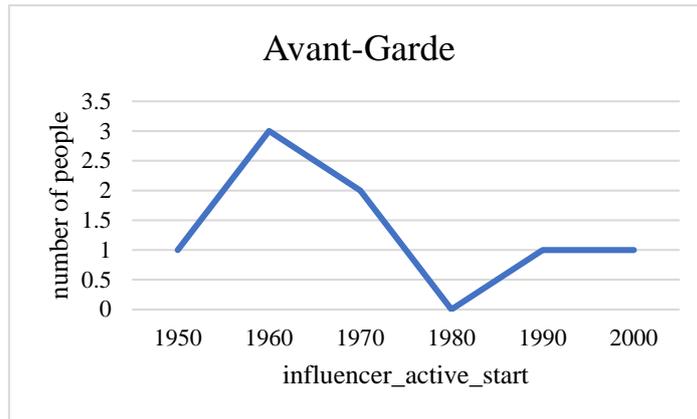


Figure 6 Avant-Grade

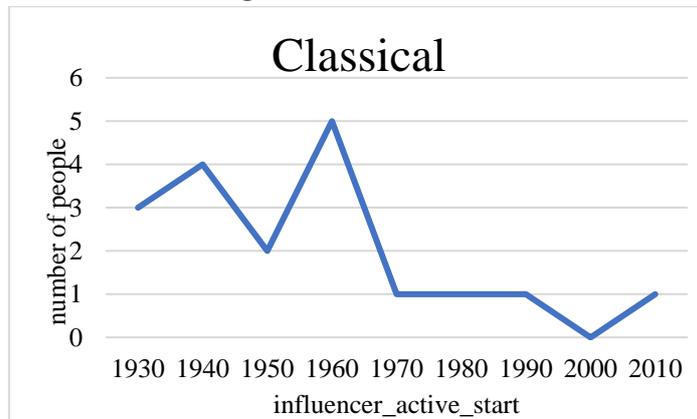


Figure 7 Classical

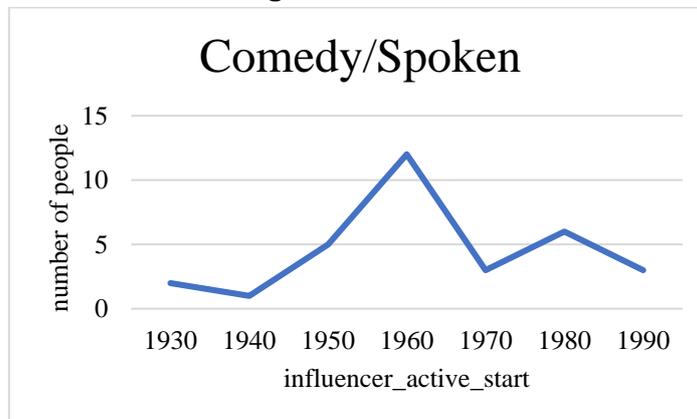


Figure 8 Comedy/Spoken

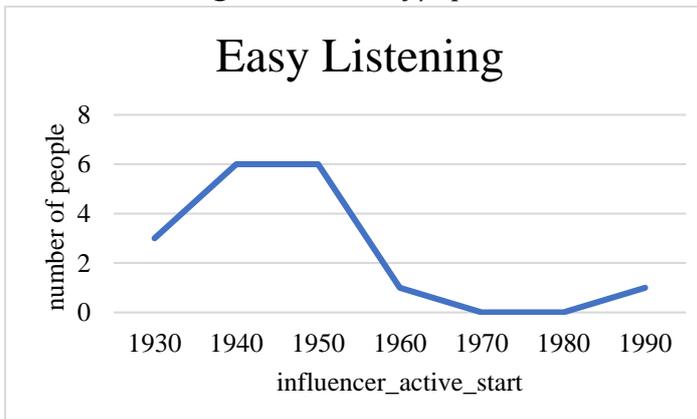


Figure 9 Easy Listening

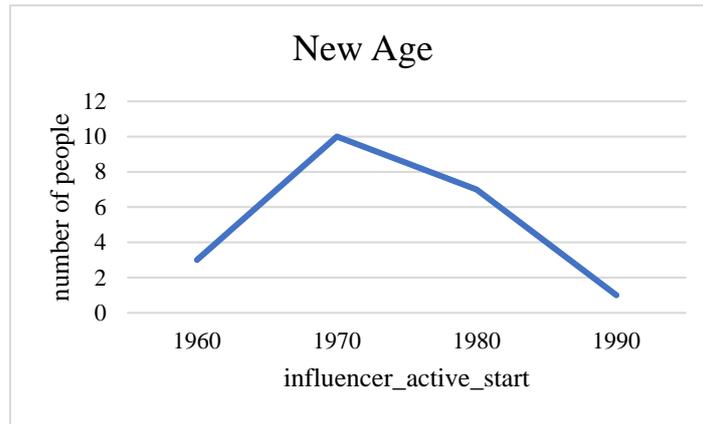


Figure 10 New Age

Weight matrix calculation based on AHP

Through the extraction and analysis of data, we select the total number of artists influenced by the artist, the ranking of the influence of the artist's genre in his period, the number of artists in the genre in that period, and the total number of all genres in the artist's period as evaluation indicators.

Firstly, we use the analytic hierarchy process to select the weight, the target layer and the criterion layer are as shown in the figure below.

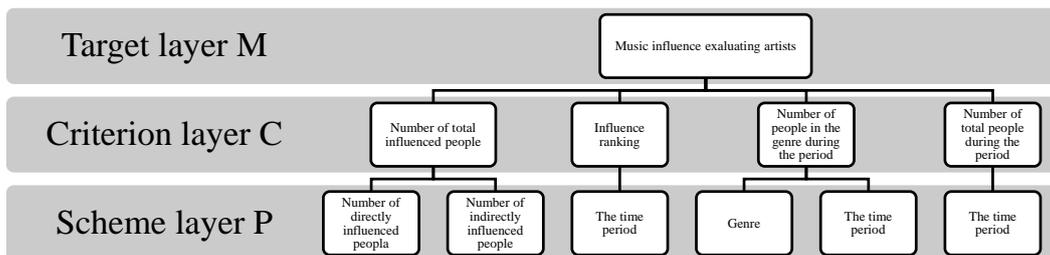


Figure 11 the target layer and the criterion layer

By reading the literature and consulting experts, we construct a judgment matrix M-C, and compare the elements in the criterion layer C pair by pair to obtain a comparison matrix:

Table 2 matrix M-C

| M | C1 | C2 | C3 | C4 |
|----|------|-----|----|----|
| C1 | 1 | 2 | 4 | 4 |
| C2 | 0.5 | 1 | 2 | 2 |
| C3 | 0.25 | 0.5 | 1 | 1 |
| C4 | 0.25 | 0.5 | 1 | 1 |

Solve the eigenvalues of M-C, we can find that $\lambda_{max} = 4$

And the weight vector $\omega = (0.5, 0.25, 0.125, 0.125)^T$

So the importance coefficient of each index is determined, as shown in the following table:

Table 3 the weight of each index

| INDEX | TOTAL NUMBER OF PEOPLE AFFECTED | INFLUENCE RANKING | NUMBER OF PEOPLE IN THE GENRE DURING THE PERIOD | TOTAL NUMBER OF PEOPLE IN THE PERIOD |
|-------|---------------------------------|-------------------|---|--------------------------------------|
| | | | | |

| | | | | |
|---------|-----|------|-------|-------|
| WEIGHTS | 0.5 | 0.25 | 0.125 | 0.125 |
|---------|-----|------|-------|-------|

Establishment of Grey Comprehensive Evaluation Model

This paper takes Comedy/Spoken as an example to establish a grey comprehensive evaluation model.

First, determine the optimal reference data is $\alpha = \{68,1,13,1129\}$

To determine the important coefficients of each index, see 3.3.5.

Calculate the relevance of each artist index data column in Comedy/Spoken to the optimal reference data column, and draw the table in descending order of relevance as follows.

Table 4 correlation

| ARTIST_NAME | CORRELATION | ARTIST_NAME | CORRELATION | ARTIST_NAME | CORRELATION |
|-----------------------|-------------|--------------------|-------------|------------------|-------------|
| ROBIN WILLIAMS | 1.302 | Steven Wright | 1.202 | Tom Lehrer | 1.066 |
| VICTOR BORGE | 1.300 | Rodney Dangerfield | 1.174 | Brian Regan | 1.064 |
| JACK KEROUAC | 1.277 | Bill Cosby | 1.146 | Jello Biafra | 1.058 |
| MONTY PYTHON | 1.273 | Cheech & Chong | 1.133 | Mel Brooks | 1.049 |
| BOB NEWHART | 1.256 | Allan Sherman | 1.125 | Sam Kinison | 1.047 |
| EDDIE MURPHY | 1.255 | Robert Klein | 1.120 | Rudy Ray Moore | 1.037 |
| CASSIUS CLAY | 1.252 | George Carlin | 1.092 | Mitch Hedberg | 1.030 |
| SPINAL TAP | 1.251 | Jerry Lewis | 1.079 | Jonathan Winters | 1.001 |
| ADAM SANDLER | 1.231 | Dick Gregory | 1.079 | Spike Jones | 0.922 |
| THE SMOTHERS BROTHERS | 1.210 | Denis Leary | 1.074 | Bebe Neuwirth | 0.881 |
| STEVE MARTIN | 1.208 | Richard Pryor | 1.068 | | |

3.4. Music similarity measurement model based on cosine similarity

3.4.1 PCA dimensionality reduction processing

As there are too many indicators, we organize the data_by_year data set and use the principal component analysis (PCA) to reduce the data.

In order to facilitate the processing, we do not consider the influence of the Boolean variables mode and key.

First, we standardize the sample matrix and calculate the covariance matrix of the standardized matrix, that is, the correlation coefficient matrix:

$$R = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n ((x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2)}}$$

Calculate the eigenvalues and eigenvectors of the standardized matrix through MATLAB, and calculate the principal component contribution rate and cumulative contribution rate, and finally get the principal components we need.

The eigenvalues, corresponding eigenvectors and contribution rate of the final correlation coefficient matrix are as follows.

Table 5 PCA

| FEATURE VECTOR | A1 | A2 | A3 | A4 | A5 | A6 | A7 | A8 | A9 | A10 | A11 |
|----------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| X1 | 0.25 | -0.54 | 0.24 | 0.11 | -0.19 | -0.35 | -0.30 | -0.09 | -0.54 | -0.20 | 0.00 |
| X2 | 0.38 | 0.00 | -0.07 | -0.10 | 0.09 | -0.14 | 0.34 | -0.13 | 0.17 | -0.42 | -0.69 |
| X3 | 0.02 | -0.50 | -0.63 | 0.29 | -0.27 | 0.21 | 0.11 | -0.27 | 0.18 | 0.17 | 0.00 |
| X4 | 0.36 | -0.15 | -0.21 | 0.09 | 0.07 | -0.06 | -0.02 | 0.88 | 0.07 | 0.10 | 0.01 |
| X5 | 0.37 | -0.05 | 0.07 | -0.01 | 0.25 | -0.31 | -0.41 | -0.25 | 0.62 | 0.02 | 0.29 |
| X6 | -0.37 | 0.02 | 0.00 | 0.23 | 0.03 | -0.08 | -0.58 | 0.08 | 0.11 | 0.23 | -0.63 |
| X7 | -0.35 | 0.00 | 0.00 | 0.11 | -0.23 | -0.78 | 0.40 | 0.07 | 0.16 | 0.14 | 0.06 |
| X8 | -0.23 | -0.18 | -0.48 | -0.51 | 0.55 | -0.22 | -0.08 | -0.04 | -0.25 | 0.00 | 0.02 |
| X9 | -0.16 | -0.50 | 0.36 | -0.60 | -0.22 | 0.18 | 0.06 | 0.12 | 0.32 | 0.15 | -0.11 |
| X10 | 0.22 | 0.39 | -0.33 | -0.45 | -0.61 | -0.14 | -0.28 | -0.03 | -0.08 | 0.08 | -0.02 |
| X11 | 0.38 | 0.04 | 0.14 | -0.03 | 0.18 | -0.05 | 0.20 | -0.20 | -0.22 | 0.80 | -0.18 |
| EIGENVALUES | 6.59 | 1.56 | 1.25 | 0.69 | 0.38 | 0.25 | 0.10 | 0.09 | 0.07 | 0.02 | 0.00 |

| | | | | | | | | | | | |
|------------------------------|------|------|------|------|------|------|------|------|------|------|------|
| CONTRIBUTION RATE | 0.60 | 0.14 | 0.11 | 0.06 | 0.03 | 0.02 | 0.01 | 0.01 | 0.01 | 0.00 | 0.00 |
| CUMULATIVE CONTRIBUTION RATE | 0.60 | 0.74 | 0.85 | 0.92 | 0.95 | 0.97 | 0.98 | 0.99 | 1.00 | 1.00 | 1.00 |

It can be seen from the table that the cumulative contribution rate of the first three principal components is 85%, so we can consider only taking the first three principal components, which can well summarize the original variables.

$$F1 = 0.25x1 + 0.38x2 + 0.02x3 + 0.36x4 + 0.37x5 - 0.37x6 - 0.35x7 - 0.23x8 - 0.16x9 + 0.22x10 + 0.38x11$$

$$F2 = -0.54x1 + 0.00x2 - 0.50x3 - 0.15x4 - 0.05x5 + 0.02x6 + 0.00x7 - 0.18x8 - 0.50x9 + 0.39x10 + 0.04x11$$

$$F3 = 0.24x1 - 0.07x2 - 0.63x3 - 0.21x4 + 0.07x5 + 0.00x6 + 0.00x7 - 0.48x8 + 0.36x9 - 0.33x10 + 0.14x11$$

Based on linear algebra theory, we can think that these 3 principal components can represent all indicators.

3.4.2 Additional Assumptions

Cosine similarity measures the similarity between two vectors by measuring the cosine of the angle between them. When two vectors have the same pointing direction, the cosine similarity value is 1; when the angle between the two vectors is 90°, the cosine similarity value is 0; when the two vectors point to completely opposite directions, the cosine similarity value is -1.

Given two attribute vectors, A and B, the cosine similarity is given by the dot product and the vector length, as shown below:

$$similarity = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|}$$

Due to the large amount of data, we randomly selected 6 songs as samples, used the three principal components obtained after PCA dimensionality reduction as indicators, and calculated the cosine similarity by comparing them in pairs.

Table 6 the heat map

| song_title (censored) | Out Getting Ribs | If I Could Only Flag Her Down | Give It Up | The New Symphony Sid | Tomorrow Is Another Day | Don't Worry About Me |
|-------------------------------|------------------|-------------------------------|-------------|----------------------|-------------------------|----------------------|
| Out Getting Ribs | 1 | 0.999999986 | 0.999999995 | 0.999999994 | 0.999999992 | 0.999999995 |
| If I Could Only Flag Her Down | 0.999999986 | 1 | 0.999999998 | 0.999999973 | 0.999999965 | 0.999999998 |
| Give It Up | 0.999999995 | 0.999999998 | 1 | 0.999999985 | 0.999999979 | 0.999999989 |
| The New Symphony Sid | 0.999999994 | 0.999999973 | 0.999999985 | 1 | 0.999999999 | 0.999999999 |
| Tomorrow Is Another Day | 0.999999992 | 0.999999965 | 0.999999979 | 0.999999999 | 1 | 0.999999998 |
| Don't Worry About Me | 0.999999995 | 0.999999998 | 0.999999989 | 0.999999999 | 0.999999998 | 1 |

3.4.3 Comparison of similarity among artists

Based on the five genres in the first question, for musicians corresponding to the five genres in data_by_artist, the similarity between musicians in the same genre is calculated by cosine similarity, and the mean value of cosine similarity in the same genre is calculated. The results are as follows table.

Table 7 sim of same genre

| GENRE | AVANT-GARDE | CLASSICAL | COMEDY/SPOKEN | EASY LISTENING | NEW AGE |
|-------|-------------|-----------|---------------|----------------|---------|
| SIM | -0.0913 | -0.0264 | -0.0191 | -0.0309 | -0.0169 |

Finally, we selected two music genres, easy listening and classical, and also compared the music similarity between these two fields through cosine similarity, and calculated the mean cosine similarity $sim = 0.002291$

By comparing the degree of similarity between the artists in the selected genres and among the genres, we can conclude that artists in different genres are more similar than artists in the same genre.

3.4.4 Cluster analysis

Cluster analysis refers to the analysis process of grouping a collection of physical or abstract objects into multiple classes composed of similar objects. It is an important human behavior.

The goal of cluster analysis is to collect data to classify on the basis of similarity. Clustering comes from many fields, including mathematics, computer science, statistics, biology, and economics. In different application fields, many clustering techniques have been developed. These technical methods are used to describe data, measure the similarity between different data sources, and classify data sources into different clusters.

The cluster pedigree diagram after cluster analysis of some artists is as follows:

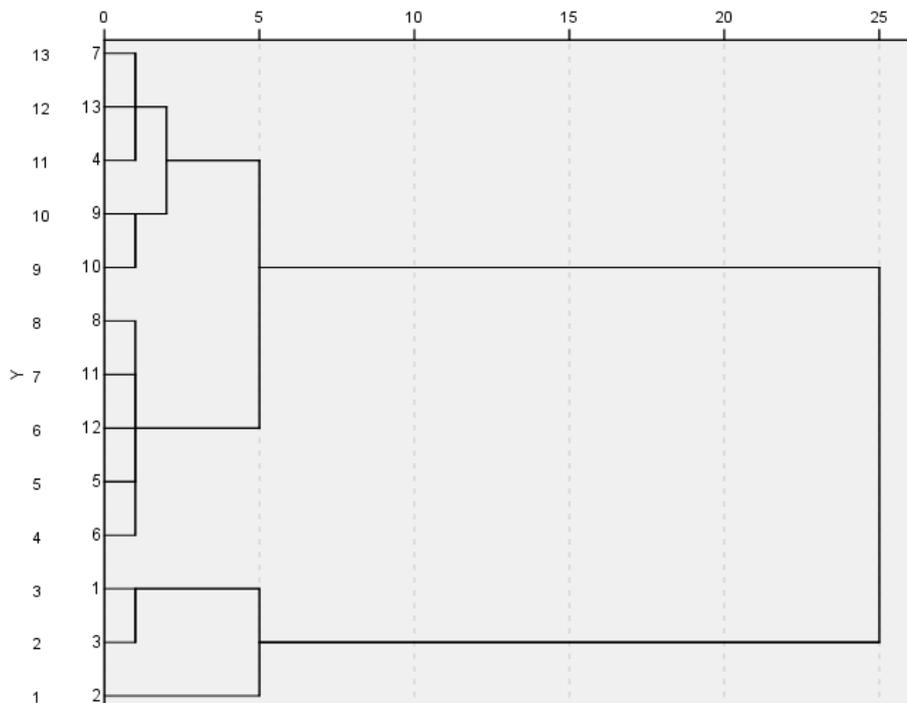


Figure 12 The cluster pedigree diagram

The horizontal axis represents the distance between various types. We can intuitively see the correlation between artists within a genre and artists between genres, and verify that artists between different genres are more similar than artists within the same genre.

3.4.5 The relationship between influence and musical similarity

In order to explore whether people with greater musical influence will affect the music of his followers, we selected Moondog, an artist in the Avant-Garde genre, and Harold, who is also an Avant-Garde genre and a follower of Moondog. Budd and Arthur Russell compare and draw a radar chart as follows.

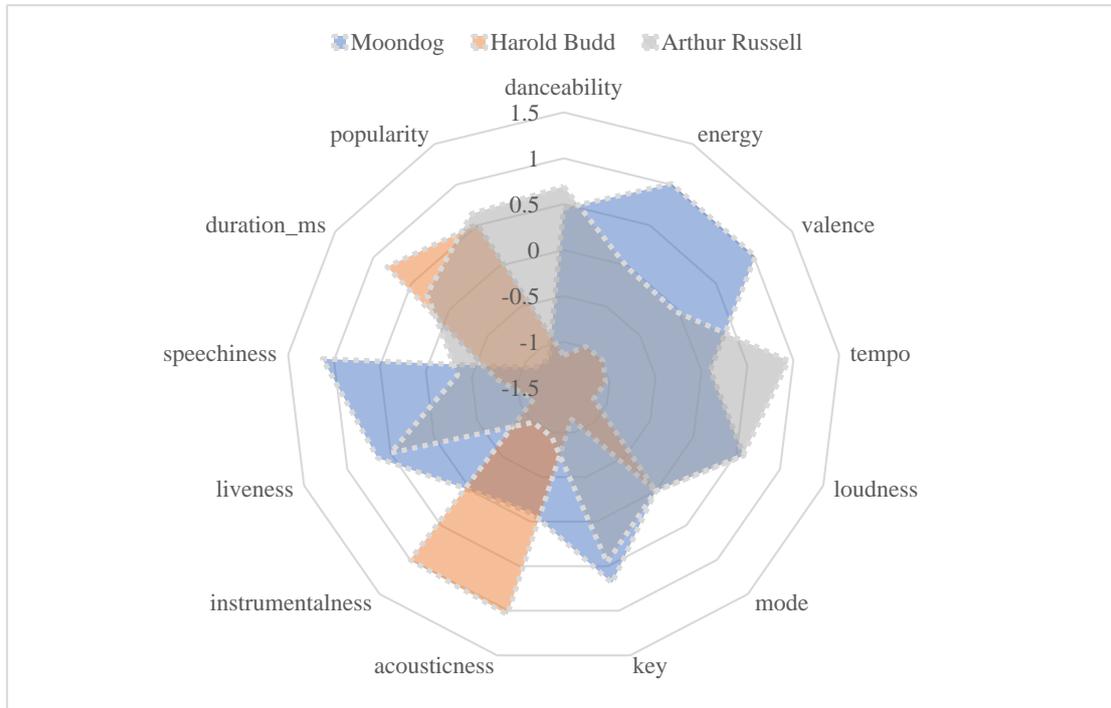


Figure 13 radar chart

It can be seen from the figure that the indicators of Harold Budd and Arthur Russell overlap with Moondog's indicators in many places, which are obviously affected by Moondog.

3.5. Correlation analysis model

3.5.1 Normal distribution test

The Pearson correlation coefficient requires the sample data to meet the requirements of normal distribution. Therefore, we use the Jarque-Bera test to test the normal distribution of the samples.

For a random variable $\{X_i\}$, Assuming its skewness is S and kurtosis is K, then we can construct the JB statistic:

$$JB = \frac{n}{6} \left[S^2 + \frac{(K - 3)^2}{4} \right]$$

It can be proved that if $\{X_i\}$ is a normal distribution, then in the case of a large sample, $JB \sim \chi^2(2)$. Therefore, we integrate the data in full_music_data JB*, calculate the skewness and kurtosis of the sample data, obtain the test value, and calculate the p value of the sample data to be $0.001 < 0.05$, so the sample is considered to meet the requirements of the normal distribution.

3.5.2 Pearson correlation coefficient model

The Pearson correlation coefficient is used to measure whether two data sets are on a line, that is, to measure the linear relationship between distance variables. When the two variables are both normal continuous variables and there is a linear relationship between the two, the Pearson correlation coefficient is often used to describe the degree of correlation between the two. The specific calculation formula is as follows:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \cdot \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

Where r represents the correlation coefficient, n is the number of samples, and X_i and Y_i respectively represent the two sets of attribute values of the i -th sample. When $r = 1$, it is said that X and Y are completely correlated. At this time, there is a linear functional relationship between X and Y ; when $r > 0.8$, it is called high correlation, when $r < 0.3$, it is called low correlation, and other times it is moderate. Related.

Calculate the correlation coefficient with MATLAB

To facilitate visualization, only the pearson correlation coefficient between each indicator and popularity is shown here:

Table 8 pearson correlation

| | DANCEABILITY | ENERGY | VALENCE | TEMPO |
|------------|------------------|----------|-------------|--------------|
| POPULARITY | 0.1831 | 0.3898 | 0.0207 | 0.0942 |
| | loudness | mode | key | acousticness |
| POPULARITY | 0.4203 | -0.0363 | 0.0150 | -0.4775 |
| | instrumentalness | liveness | speechiness | explicit |
| POPULARITY | -0.2161 | -0.0643 | -0.0069 | 0.1830 |
| | duration_ms | year | | |
| | 0.0572 | 0.7754 | | |

It can be seen from the above table that the popularity of music is moderately related to energy, loudness, and year.

3.6. Pros and cons of the model

Only five genres are considered in our model and paper,so partial factors can't be avoided.

During PCA link,we ignore two indexes,mode and key.Although the two indexes are not important in a high degree,there exists slight impact on our findouts.

For the reasom of data processing,some data can't be analysed accurately.

4. Conclusion

4.1. Conclusions of the problem

Conclusion of Question 1

In 3.3, we use the total number of artists influenced by the artist, the ranking of the influence of the artist's genre in his period, the number of the artist's genre in that period, and the total number of all genres in the artist's period as the evaluation index, and the AHP is used to determine the weight. Matrix to establish a grey comprehensive evaluation model.

After that, we take the artists in Comedy/Spoken as an example to calculate their gray relevance. Among them, Robin Williams has the highest relevance, that is, the greatest influence.

In this model, musical influence represents the influence of a certain artist on the creation of other artists.

Conclusion of Question 2

In 3.4, we use PCA dimensionality reduction to process each index to obtain three principal components, and use these three principal components to represent all indexes, and calculate the cosine similarity between musicians in pairs, and establish a music similarity measurement model.

In our measurement standards, artists in different genres are more similar than artists in the same genre.

Conclusion of Question 3

Similarities and influences between and within genres

Between genres: For artists of different genres, their music has certain similarities. The analysis can be obtained because the influencers and followers are in different genres, so there are certain similarities in their music between different genres.

Within the genre: The music within the genre is in the same genre, so when the music is spread within the genre, the similarity will have a certain degree of transitivity.

Influence: The similarity between genres and within genres will affect the reform of genres. When artists of some genres are influenced by artists of other genres, their music works will also have certain similarities with their influencers, so genre music will Certain development and reforms have taken place.

What is the difference between genres?

Different genres are mainly embodied in music. The specific manifestations are energy, valence, tempo and other related factors of music. These are all influencing factors of music. We define a genre mainly through these factors, and analysis can quantify these factors. And comprehensively evaluating these factors, we can draw the division of music genres.

How does the genre change over time?

We selected the Pop/rock genre with a large amount of data in the full data data set, and used the year as the abscissa and the song volume as the ordinate to draw a line chart of the music of this genre over time.

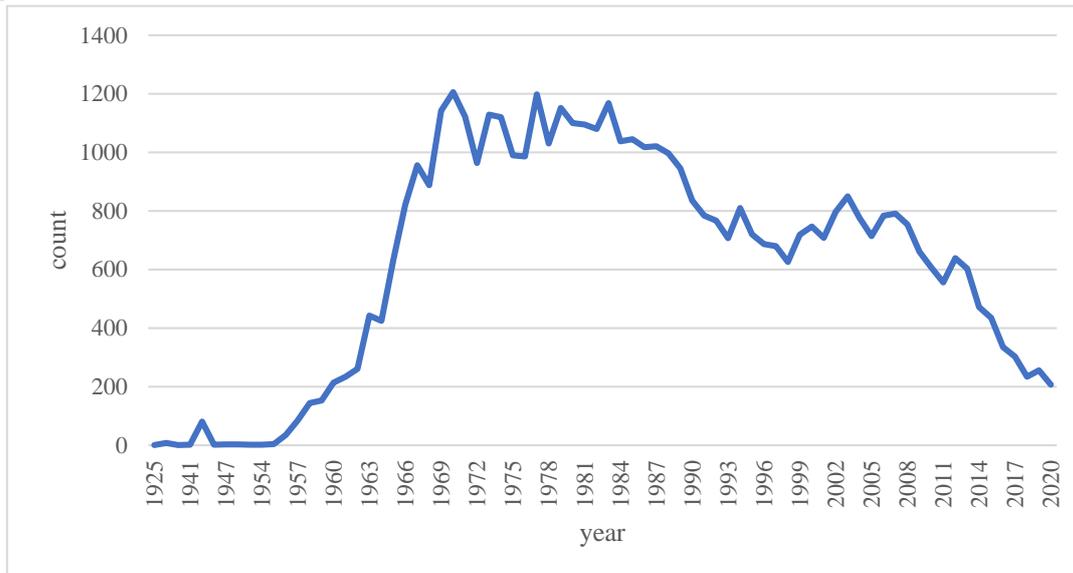


Figure 14 The music of Pop/Rock over time

Through the analysis of the line chart, we can find that the development of this genre has been on the rise from 1925 to 1970, and reached its peak in the 1970s. Our analysis can conclude that the influence of some historical events caused the music During this period of time, it was in a rising period, such as the Renaissance and the Industrial Revolution; after 1970, the music of this genre gradually declined. By comparing the development of other genres in the same period of time and the analysis of the social environment at that time, it can be concluded Two reasons: the influence of war (external reasons), and the continuous development of new genres with the development of various types of music, which led to the decline of the music competitiveness of the genre (internal reasons)

Are some genres related to other genres?

By analyzing the similarity of music of different genres, we can conclude that the music of different genres has a certain similarity. Among them, we can get through the analysis of the

Influence_Data data set that there is a mutual influence relationship between different artists, and there are many influencers and followers in different genres.

Therefore, we can see from the picture that there is a certain correlation between the genres, and the previous exploration of the similarity of artists between the genres also demonstrates this point of view: there are connections between different genres.

Conclusion of Question 4

According to the relevant analysis model in 3.5, we conclude that the energy, loudness, and year characteristics of music are more appealing than other characteristics. In 3.4.5, we compared Moondog, an artist in the Avant-Garde genre, with Harold Budd and Arthur Russell, who are also Avant-Garde genres and followers of Moondog, and drew a radar chart. Intuitively, people with musical influence will obviously influence the music created by followers.

Conclusion of Question 5

A major leap refers to the sudden rise or decline of a sudden change in time in the development process of a genre. When our team analyzed the five genres in question 1, we drew the corresponding line chart, so observe the inflection point of the line chart Or nodes with high or low growth rates can know that the corresponding genre has undergone major changes at a certain point in time.

Immediately afterwards, the reasons for major changes can be analyzed by screening the artists and works of the period, and analyzing the outstanding changes of the work indicators, that is, comparing with the average value of the indicators in all works. Such indicators are used to describe the major changes in the title. The characteristics of change. Then, by analyzing the knowledge graph and the line graph of the first question, we can know which artists have caused major changes.

Take the Classical genre as an example for analysis. Through the question 1 line chart, it is not difficult to see that the classical genre reached its peak in the 1960s and experienced a major decline in the 1960s and 1970s. Take Wolfgang Amadeus Mozart who was active in the 1960s and 1970s as an example. For example, after investigating his data, we found that the average popularity of songs he composed in the 60s and 70s was 23.25, which was lower than the historical average of 26.42 for the genre. We all know that Wolfgang Amadeus Mozart is a famous music master. The popularity of his music in the 1960s and 1970s has been smaller than the historical average (it should have been far above the historical average), indicating that the development of the Classical genre in the 60s and 70s was not very smooth, so the trend reflected by the line chart is reasonable. At the same time, we found that the line chart shows that the New Age genre arose and reached its peak in the 1970s. According to survey data, the music of the New Age represented by the New Age genre just emerged in the 1970s. This shows that The results of our data processing are accurate and consistent with the facts. In addition, because New Age music is in a competitive relationship with Classical music, there is also a certain connection between the rise and fall of the two.

From the knowledge graph shown in question 1, we select the graph of Classical genre for observation, and combine the analysis of the music influence model proposed in question 1, we can conclude that the change makers in the classical field are at the center of the graph Artists such as John Cage and Steve Reich.

Conclusion of Question 6

Take a certain genre as an example, which changes over time, manifested as changes in various indicators in the full music data data set. In the second question, we have calculated the contribution rate of each indicator through PCA, and extracted 3 linearly independent principal components to represent all indicators (except mode and key), and analyzed the average of the three principal components to the genre popularity index The impact of the indicator on the development of the genre can be determined.

Regarding the changes of genres over time, conclusions can be drawn by analyzing the line chart in question 1.

Therefore, we perform linear regression analysis on the data of the Avant-Garde genre:

Table 9 linear regression analysis

| MODEL | R | R2 | ADJUSTED R2 | STANDARD ESTIMATE ERROR |
|-------|--------|-------|-------------|-------------------------|
| 1 | 0.792a | 0.627 | 0.625 | 9.30345 |

The closer R is to 1, the more linear it is.

We found that the most obvious linear relationship with time is the popularity indicator.

Conclusion of Question 7

For society and politics, the characteristics of songs in the era of peace and stability are different from those in the war years. The Cold War, World War II, the Industrial Revolution, and the rapid development of the Internet may have a great impact on music, and music also affects these historical events. Have an impact.

Take the Internet as an example. Nowadays, major music websites have been popularized. These websites have also signed contracts with some well-known record companies so that people can download many original music works. The combination of these music and the Internet has become popular with the Internet. It also brought a revolution in the way of music dissemination, appreciation, and creation.

Through the analysis of the data in the data by year data set, we process data in different fields, analyze and obtain the line graphs of music at different time periods, and analyze the data of these line graphs. We can get that with the popularity of the Internet, the development of music The better it comes, especially as the Internet, as a new medium of communication, breaks the geographical boundaries of communication and speeds up the spread of music.

Among them, we can find that with the development of the Renaissance in the 1400s, a major revolution in music has taken place. The works of the Renaissance period embody humanistic ideas: advocate individual liberation, oppose medieval asceticism and religious views; advocate scientific culture, oppose obscurantism, get rid of the shackles of the church on people's thoughts; affirm human rights, oppose divine power, and reject deeds All authority and traditional dogmas on the basis of theology and scholastic philosophy; supporting centralization and opposing feudal separatism are the main ideas of humanism. The art of the Renaissance extolled the beauty of the human body, advocating that the proportion of the human body is the most harmonious proportion in the world, and applying it to architecture. A series of paintings and sculptures that still take religious stories as the theme, but all express The scene of ordinary people dragged God to the ground. Humanists began to study the Bible by studying classical literature, and translated the Bible into the language of the nation, which led to the rise of the Reformation Movement. Humanism praises the secular and despise heaven, flaunts reason to replace divine enlightenment, affirms that "man" is the creator and enjoyer of life in this world, requires literature and art to express people's thoughts and feelings, science for the welfare of people, and education must develop human personality. Free people's thoughts, feelings and wisdom from the shackles of theology. Advocating the freedom of individuality, it has played a very important role in historical development.

In the 1960s, with the progress of the Industrial Revolution, we can find the trend of a hundred flowers blooming in all kinds of music through the line chart. Since the Industrial Revolution, we can find that the development trend of classical music has declined. On the contrary, some new music has begun to develop. This is because music has gradually become diversified and widespread with the progress of the Industrial Revolution. The Industrial Revolution promoted social progress and scientific development, and therefore accelerated the improvement of

musical instruments. More and more new musical instruments were developed, and the diversification of music was gradually realized.

The industrial revolution's development of art is also reflected in the cultural exchanges at that time. With the advancement of science and technology, social information communication methods are gradually diversified and efficient. Therefore, more and more music influences each other, which is why the similarities between works in different music fields are also increasing. The impact of the Industrial Revolution on music has also changed people's ideology and values. Social changes have allowed classical music to make greater progress in culture and other aspects, which has also promoted the spread of music.

5. Document

Memo recipient: Recipient of the memo: ICM Association

From: Team#2127288

Date: February 8, 2021

Topic: The value of musical influence

Dear ICM Association:

Our team established a comprehensive evaluation model of music influence through analysis of a large amount of data. As we all know, music has a great impact on the social and cultural environment, so this is a very meaningful thing.

In our model, the period, genre, and the total number of people he influences are all indicators of musical influence. We found that when artists of some genres are influenced by artists of other genres, his music The works will also have certain similarities with their influencers, so certain development and reform of genre music will occur.

Through linear regression analysis, we compared the changes of musicians' composing styles over time, and finally we found that the most important indicator of changes over time is popularity. It can be seen that as musicians' composing styles gradually change When he matures, his works will become more and more infectious.

In addition to the above, we also did a lot of data analysis, but only limited to the data given in the title. If more data can be obtained in the future, our analysis of the influence of music may be more accurate, which will also help us to study the influence of music on culture.

References

- [1]M. Kaminskas, F. Ricci. Contextual music information retrieval and recommendation: State of the art and challenges[J]. Computer Science Review, 2012(6): 89-119.
- [2]H. Ezzaidi, J. Rouat. Automatic musical genre classification using divergence and average information measures[R]. Research Report of the World Academy of Science Engineering and Technology, 2006.
- [3]B. Logan, A. Salomom. A music similarity function based on signal analysis[C]. IEEE International Conference on Multimedia and Expo (ICME), Tokyo, Japan, 2001: 745-748.
- [4]Dan Shi. Research on music genre similarity detection algorithm[D]. Dalian University of Technology, 2013.
- [5]Gang Xu. Research on audio similarity and their applications in music retrieval[D]. University of Electronic Science and Technology of China, 2015.
- [6]P. Knees, M. Schedl. Music similarity and retrieval[C]. SIGIR, Dublin, 2013, 4-14.
- [7]B. Logan, A. Salomon. A music similarity function based on signal analysis[C]. IEEE ICME, 2001,17-18.