Problem Chosen	2021 MCM/ICM Summary sheet	Team Control Number
D		2126164

The Influence of Music

Summary

In order to explore the influence of music and the development of music over time, we made some models to analysis from the perspective of songs and artists' musical.

First, the data was simplified and classified into 20 kinds of genres, and the neural directed network diagram was made to connect the influencers and followers. Through the analysis, the influence parameters were determined as personal influence, the direction of genre evolution and the speed of evolution. And we took one artist from Avant-Garde as an example, calculated the average transmission speed is 1.7 artists per year, the average annual transmission genres 0.2.

We used the principal component analysis to reduce the dimensions of the indicators and obtain two principal components. It can explained as Characteristics of the music, Type of vocals. Then, we calculated the coefficient of variation of each principal component in each genre. And by comparing the average of the coefficient of variation of each genre, we find that the similarity between Country genres is smaller, and the similarity between Comedy/Spoken genres is larger. We concluded that the average coefficient of variation of Characteristics of the music between different genres is 2.145, and the similarity is small. And the coefficient of variation of vocals is quite similar.What's more,our team believed that influencers really have an impact on the creation of followers by obsercing the images.Then, by calculating the Pearson correlation coefficient and Spearman correlation coefficient of indicators ,we found that there is a strong positive correlation among energy, loundness and popularity, and there is a strong negative correlation between acousticness and popularity. Then we found an artist, the Beatles who is famous, and then we drew energy, and loudness radar charts of his 606 followers , and found that they are indeed similar.

To determine the characteristics that influenced the evolution of music. Based on the "Elbow method" in K-means and the analysis of broken line diagram we combined with the influence of artists, based on the "Elbow method" in K-means and the analysis of broken line diagram, it can find the signify revolution. As a result John Cage was the main agents of change in the classical and main agents of change in the Blue were Jimmy Reed and John Lee Hooker.What's more Boyzone was the main revolutionaries in the Avant-Grade.Then,we first took New age as an example to analyze the evolution of this genres, and found that it was a new genre influenced by Classical, Pop/Rock and Folk. Then,it affected other genres and the International came into being. And we finally found dynamic influencers by looking for specific indicators under the genre.Finally, combining the above analysis with the cultural background of music, we wrote a one-page document to the ICM Society to elaborate our model and consider the impact of other more data on this model.And we also gave some advice about the development of music.

Keywords:Directional network diagram, data visualization, standard coefficient of variation, PCA (principal component analysis)

Contents

1.Introduction

1.1Background

Since ancient times, different artists have chosen different types of music because of age, because of the occurrence of some important historical events or the development of music technology. An artist who influences others can influence multiple followers at the same time. Even if they choose different types of music, however, a large part of artists who follow others in early years have become artists who influence others. Therefore, analyzing the influence of artists on others helps us to better understand the development of music. At the same time, we can observe the characteristics and schools of music to judge the influence of each artist, to provide us with ideas to build a model of measuring influence, and to help people better understand the development of music over time.

1.2 Restatement of problems

In this paper, we are aimed to develop a model that measures musical influence. This problem asks you to examine evolutionary andrevolutionary trends of artists and genres. And we need to solve the following problems:

- For the first poblem, we need to use the Impact_data dataset to create a targeted network of music influence, connect influencers to followers, and establish a connection between influencers and the music category chosen by followers to create a network. At the same time, the parameters of "music influence" need to be determined according to the data, and the parameters need to be applied to a subnetwork of direct influencers for analysis.
- For the second problem, our team needs to find some parameters that can be used to contain the main information of each song. By comparing these parameters, we can get the magnitude of similarity between songs of the same genre and songs of different genres.
- For the third problem, we need to find an indicator to measure whether there are differences between genres. The key lies in how to distinguish different genres and analyze how the market of genres changes in combination with time.
- For the fourth problem, we need to answer whether the similarity data show that the influencers really affected their followers. The key is to find the degree of similarity between the principal components and the two principal components of the followers. The higher the similarity, the greater the influence degree, and the less vice versa.
- For the fifth question, we need to identify characteristics that might indicate a revolution in the evolution of music. And apply those characteristics to our network to find artists who represent revolutionaries.
- For the sixth problem, it requires us to further analyze the process of musical evolution in a genre over time and explain how genres or artists change over time in the light of the previous conclusions.

➢ For the seventh problem, in combination with the historical background, science and technology, we need to analyze the cultural influence of music in time or environment expounded in this paper.

1.3 Overview of our work

Thought the 1.2 show the probolem, we will proceed as follows for the sake of tackling these problems:

- On question one, our team needs to create one or more directed networks of musical influence using Influence_Data data sets or parts of them, where influencers connected to followers. Name of influencer in Influence_Data forminfluencer_name and follower: follower_name. Because the network needs to be drawn in a direction,we decided to connect the connected influencers and followers with a directed edge.This connects them all to construct multiple directional networks. Then through the analysis of the data to develop the impact of "music influence" three parameters, and create a subnetwork of direct influencers. Finally, we apply the formulation parameters to the constructed directional network for analysis

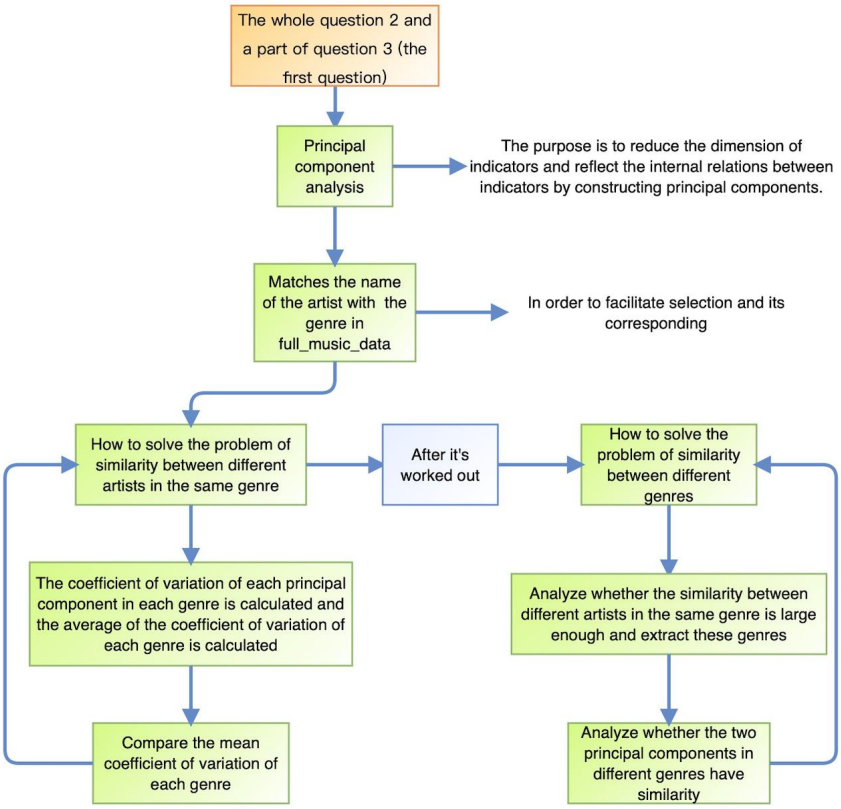


Figure1 summary of our work in task 2&3

- For the third question, first of all, we calculated the mean value of variation coefficient of each genre and compared the difference coefficient of each feature of each genre, and

concluded the comparison ofsimilarity and difference among genres. Then, we took Comedy and Avant-grade as examples and made directional network graphs and analyzed their changes over time. At the same time we analyzed the correlation of different genres. The summary for the second and third questions are shown in the flowchart below.

- For the fourth question, we need to determine whether influencers actually influence followers to make music, first we screened out the principal components of different followers under the same influencer in the same genre and plotted their scatter plots. In order to address the question of whether some musical traits are more "appealing" than others, we performed a regression analysis of correlations among all indicators and popularity indicators. A higher correlation means it's more contagious. Finally, we randomly selected an artist and analyzed how similar the "more infectious" indicator was to show that it played a similar role in influencing the popularity of a particular musician.
- For the fifth question, we first used full_music_data's data to estimate the characteristics of major revolutions that may affect the evolution of music. Its characteristic indicators include danceability, energy, loundness and popularity, as well as the number of artists and the genres it influences. Then, we used the K-means elbow graph analysis to find the inflection point, and the artist corresponding to this point is the reformer in the genre.
- For the sixth question, we first filtered the data of influence_data and took the New Age genre as an example to analyze its evolution process. The key is to identify the founders of New Age and the genres that influenced those founders by the years given. Using the same principle, the direction of evolution after the New Age can also be deduced. Then, we looked at how popularity and those metrics change similarly within each genre over time, and finally find dynamic influencers by looking for specific metrics within that genre.
- For the seventh question, we first analyzed data_by_year data, take energy, danceability and popularity as indexes, then we explored the change value of these three indexes every year from 1921 to 2020. Finally we analyzed the reasons for the turning point in combination with historical events and social background, and predicted the future trend of music.

2 General Assumptions and Variable Description

2.1 Assumptions

Before modeling, to avoid ambiguity, we will define the following confusing concept.

- It is assumed that the characteristics of music, such as danceability and energy, can fully reflect the creative style of the artist.
- Since there are 40,000 artists' information in the data, some artists' genres are missing, so we assume that this missing has no impact on our data analysis.
- Suppose that the earliest time each genre appears in the data is the time when the genre is generated。

2.2 Variable Description

Symbol	definition
PI	Personal Influence
v	The speed with which genres spread
t	The year of the influence of the influencer
ρ	Standard difference coefficient
σ	Standard deviation of genre
μ	The average

3 Model Establishment and Solutions

3.1 Task1

After reading, we subdivided this question into the following three sub-questions, which are the establishment of the directed network influencers and followers, the establishment of the parameters of the influence of music, and the introduction and description of one of the subnetworks of the established directed network system.The following is a detailed discussion.

3.1.1 a directed network of musical influence

In the first part we need to build a network of music influence. Because the music influence is targeted, the network constructed should be directional. Using some of the data in the Influence_Data data, we randomly selected and took Pop/Rock types of music as an example to make the following diagram.

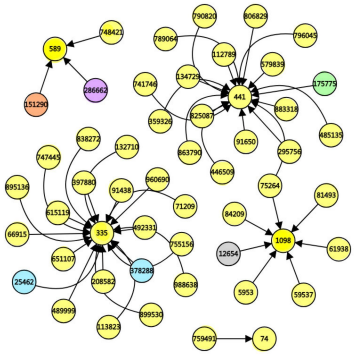


Figure 2: The example of Pop/Rock music's directed network

As shown, for instance the Pop/Rock music, by constructing directional networks had built connections between influencers and followers. Figure out that yellow stands for Pop/Rock, blue stands for Electronic, orange stands for Jazz, Purple represents country, Green representative Reggage and Grey stands for Comedy/Spoken.

From Figure1, through the diagram we can see that for Pop/Rock type followers who ID are 589,441,335,1098 and 74, most of the artists who influence their musical types come from the same field, which is Pop/Rock type. However, there are still a few artists in other fields, such as Electronic, Country and Jazz.But the weight of the impact is small.This applies equally to those who influence and follow others. Therefore, our preliminary analysis shows that the influence of artists in the same field will be more obvious, and based on the Figure 1 inspiration, then the music influence parameters are confirmed.

3.1.2 Explore a subset of musical influence

The question required us to build the establishment of parameters for musical influence. Through the color distribution of ID circles in the Figure 1, we can find out that the main parameters of music influence are the personal influence, the speed of school transmission and the evolution of school.

First of all, we set up two Avant-Garde schools to study the direction, speed and personal influence of Avant-Garde schools. The connection is shown as follows.

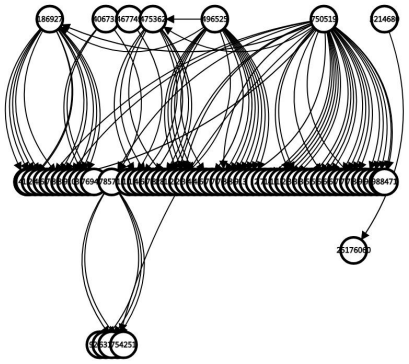


Figure3:the directed network for the Avant-Garde(Overall structure distribution)

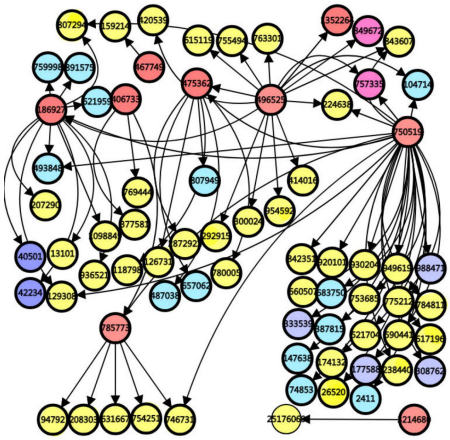


Figure 4 the directed network for the Avant-Garde(For details)

In order to express more clearly that the number and correspond main_genre of each element in the Figure2, we decided to use color in the Figure 3 to distinguish different areas of music, so that achieved the role of supplementary explanation. And in the Figure 3, red stands for Avant-Garde, pale purple for New Age, pink for Classical, yellow stands for Pop/Rock, and blue stands for Electronic.

And then we analyze the speed, direction, and influence of individuals (Taking Avant-Garde as an example)

- With regard to the speed of Avant-Garde transmission, we can see in the Figure2 that the artists of the first class of 7 Avant-Garde schools influence the second level of 78 artists. can be seen that the propagation speed of Avant-Garde is about 11.
- The Figure2 can clearly see that there are two subsets of the influence generated by the Avant-Garde, the second and the third layers in the Figure2. We can also see in figure 3 that Avant-Garde influenced Pop/Rock、New Age、Classical、Electronic four schools. The Pop/Rock was most affected, with 44 artists affected by Avant-Garde artists.
- With regard to personal influence, such as an artist with a id of 750519 in the Figure 3, we take it out separately for analysis. That is, it can be seen in the Figure 4 that the artist influenced 26 other artists of different or field factions between 1930 and 2010, belonging to the four different factions of the Pop/Roc,New Age, Classical, Electronic.

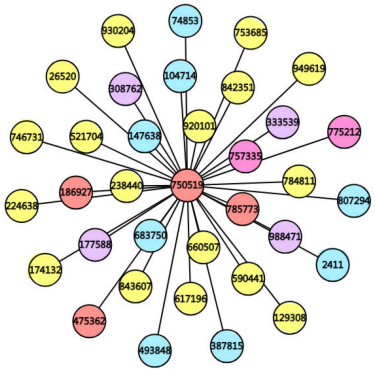


Figure 5 The musical influence by id 750519

3.1.3 Describe this subnetwork

As shown in the Figure 4, the following three parameters are analyzed:

- Personal influence: this artist had influenced 26 other artists for different or similar fields between 1930 and 2010, belonging to four different factions for which are Pop/Rock, New Age, Classical and Electronic.
- Transmission speed: this artist's followers spent 20 years influencing four schools of thought from 1960 to 1980, a total of 34 artists , and 26 of whom were artists of other genres, with an average speed of 1.7 artists per year and an average of 0.2_x0005__x0005_ genres per year
- Direction of transmission: it spread to four different factions including Pop/Rock, New

Age, Classical and Electronic

3.2 Task2

Since the second task has a certain connection with the third task, we will discuss the second and third questions in 3.2. In 3.2.1, we will give a measurement of the musical similarity and discuss the similarities between genres in the third question.

3.2.1 Analysis of musical similarity based on PCA and coefficient of variation

We found that the data provided by full_music_data was too much, and clustering algorithm would lead to close data aggregation, which was not suitable for visual presentation. Therefore, according to the reference literature [1], our group decided to use principal component analysis to reduce the dimension of variables.

The principle of principal component analysis is as follows:

Assume that p indicators are designed for the research of a certain thing and are respectively represented by $X_1, X_2, ..., X_p$. The p-dimensional random vectors composed of these p indicators is $X = (X_1, X_2, ..., X_p)'$. Let's say the mean of our random vector X is μ, and our covariance matrix is $\sum$.

A new comprehensive variable can be formed by a linear transformation of X, which is expressed by Y

$$\begin{cases} Y_1 = u_{11}X_1 + u_{12}X_2 + ... + u_{1p}X_p \\ Y_2 = u_{21}X_1 + u_{22}X_2 + ... + u_{2p}X_p \\ ..... \\ Y_p = u_{p1}X_1 + u_{p2}X_2 + ... + u_{pp}X_p \end{cases}$$

We want the variance of $Y_i = u_i'X$ to be as large as possible and Y_i to be independent of each other. Due to the:

$$\text{var}(Y_i) = \text{var}(u_i'X) = u_i' \sum u_i$$

And for any constant c, there is:

$$\text{var}(cu_i'X) = cu_i' \sum u_i c = c^2 u_i' \sum u_i$$

The process of principal component analysis is shown in the following flow chart:

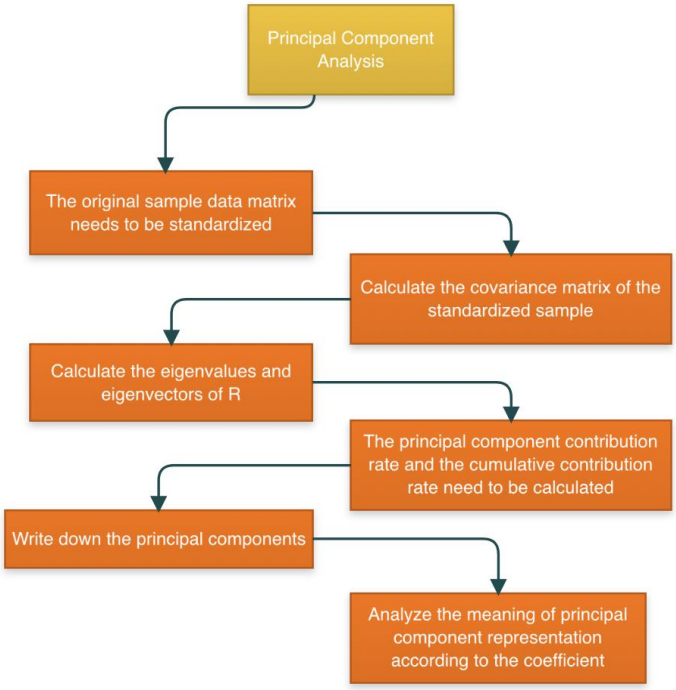


Figure6: The process of principal component analysis

According to the steps of the flow chart, when we standardize the original sample data, we got the matrix:

$$X = \begin{bmatrix} X_{11} & \cdots & X_{1p} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{np} \end{bmatrix}$$

Then we calculated the covariance matrix of standardized samples, and we can get:

$$R = \begin{bmatrix} r_{11} & \cdots & r_{1p} \\ \vdots & \ddots & \vdots \\ r_{p1} & \cdots & r_{pp} \end{bmatrix}$$

Among them, $r_{ij} = \frac{1}{n-1}\sum_{k=1}^{n}(X_{ki} - \overline{X_i})((X_{kj} - \overline{X}_j) = \frac{1}{n-1}\sum_{k=1}^{n} X_{ki}X_{kj}$

$$R = \frac{\sum_{k=1}^{n}(x_{ki} - \overline{x_i})(x_{kj} - \overline{x_j})}{\sqrt{\sum_{k=1}^{n}(x_{ki} - \overline{x_i})^2 \sum_{k=1}^{n}(x_{kj} - \overline{x_j})^2}}$$

Then, the eigenvalue and eigenvector of R, principal component contribution rate and cumulative contribution rate are calculated, and the following results can be obtained:

Eigenvalue of R $\quad \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$

Eigenvector of R $\quad a_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{p1} \end{pmatrix}, \; a_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{p2} \end{pmatrix}, \; \ldots \; , \; a_p = \begin{pmatrix} a_{1p} \\ a_{2p} \\ \cdots \\ a_{pp} \end{pmatrix}$

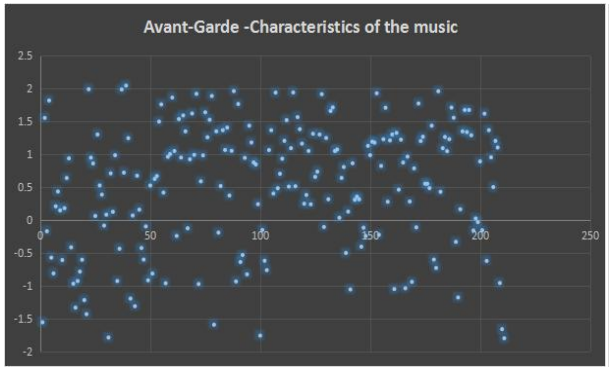
principal component contribution rate $\frac{\lambda_i}{\sum_{k=1}^{p}\lambda_k}(i=1,2,...,p)$

cumulative contribution rate. $\frac{\sum_{k=1}^{i}\lambda_k}{\sum_{k=1}^{p}\lambda_k}(i=1,2,...,p)$

Finally, we wrote the principal component and analyze the meaning of its representative, and we got the following formula:

$$F_i = a_{1i}X_1 + a_{2i}X_2 + ... + a_{Pi}X_P(i=1,2,...,m)$$

We categorize the artists' IDs in FULL_MUSIC_DATA and the influence_DATA after one-to-one correspondence. Then, we calculated the Characteristic of the music &Type of vocals of each genre. We used Avant-Garde as an example to draw the scatter graph. Let's see the following figure5,6.



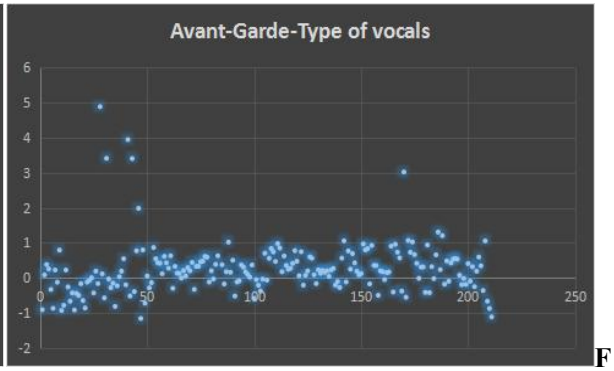


Figure7: Avant-Garde -Characteristics of the music **Figure8:** Avant-Garde-Type of vocals

Figure 5-6 shows that artists in Avant-grade genre have a high degree of dispersion in terms of Characteristics of the music, while Type of vocals is relatively low. In other words, types of vocals in Avant-grade genre are similar.

Our group made statistics according to the genres these artists belong to, and calculated the average values of different genres under the two principal components, as well as their standard deviations, and then measured the dispersion degree of the same genre and different genres by means of standard deviation coefficient variation.

standard deviation coefficient variation $\rho = \frac{\sigma}{\mu}$

The value of each **coefficient of variation** can be seen in the following table:

Table 1 Coefficients of variation of the two principal components of each genre

Type	Characteristics of the music	Type of vocals
Avant-Garde	1.8257811	2.967493995
Blues	1.650764234	2.778943711
Children's	1.196184192	1.223537604
Classical	-1.900593548	10.48242259

Comedy/Spoken	-1.026070878	0.522835473
Country	1.63250846	-47.4230303
Easy Listening	-3.884024145	-2.666754617
Electronic	-2.76550786	-4.113045356
.....	.....	.....

3.2.2 Using measure to determine whether artists within genre more similar than artists between genres

Next, we calculated the average of the coefficients of variation for each genre and plotted a histogram based on the average.See Figure 9 below:

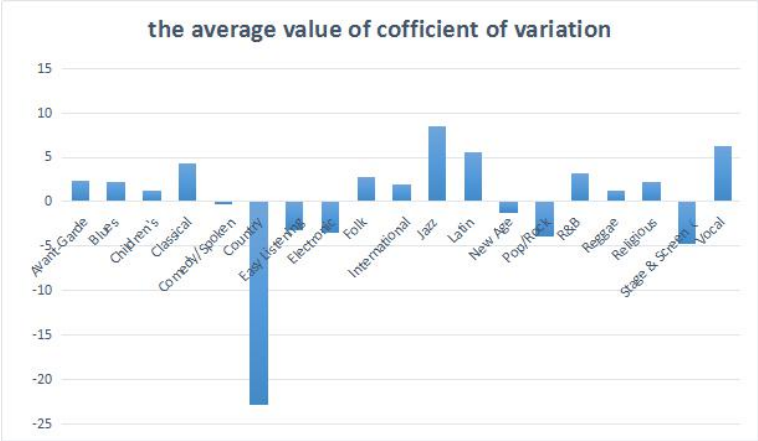


Figure9:the average value of cofficient of variation

From Figure 9, we can see that the coefficient of variation of Country is much larger than that of other genres. Therefore, our team can draw the following conclusions:

- There are great differences in the works created by different artists in the same genre of Country.
- The works created by different artists in the same genre of Comedy/Spoken are relatively similar.

Table 2 the coefficient of each variation

	Characteristics of the music	Type of vocals
the coefficient of variation	2.1447552	0.210293794

With the above conclusions, it is possible to compare whether there are similarities between artists of different genres. Since the dispersion coefficient of other genres except Country genres is small, we first excluded Country genres and calculated the average Characteristics of the music and Type of vocals of artists of other genres, and take the average value of these indicators as the music creation information of a certain genre. Then, we can find out whether different genres are similar to each other by comparing whether different genres are similar to each other in various indicators. The results of our analysis are as following Table2 shown below:

As it can be seen from Table 2, the Characteristics of the music of different genres are

greatly different, while the types of vocals are similar.

3.2.2 Using ourmeasure to determine whether artists within genre more similar than artists between genres

3.3 Task3

3.3.1 Compare similarities and influences between and within genres.

According to 3.2.1 Firgue8, we calculated the average coefficient of variation of each genre and drew a bar chart of them. We found that the coefficient of variation of Comedy/Spoken was much larger than that of other genres. Therefore, we drew the following conclusion: the works created by different artists in the same genre of Comedy/Spoken are greatly different, and the works created by different artists in the same genre of Easy Listening are relatively similar.

3.3.2 How genres change over time

Take comedy as an example to illustrate the genre's change over time

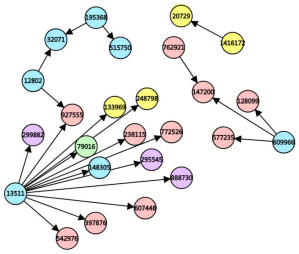


Figure10 the spread of comedy from1930 to-1950

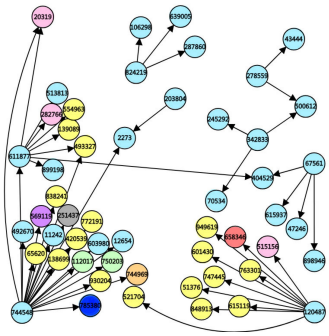


Figure11 the spread of comedy from1950 to-1970

In 1930 to 1950,blue stands for Comedy, pink stands for R&B and green stands for Country, orange stands for Folk

In 1950-1970 bule represents Comedy, pink represents folk, green represents Vocal, grey represents Electronic, purple represents stage&screen, orange represent Children,yellow represnet Pop&Rock, dark bule represents reggage and red represents R&B.

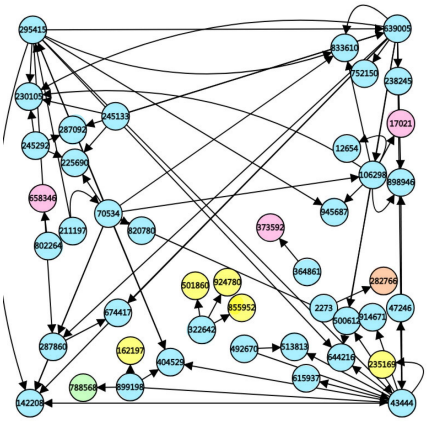


Figure12the spread of comedy from1970 to-1990

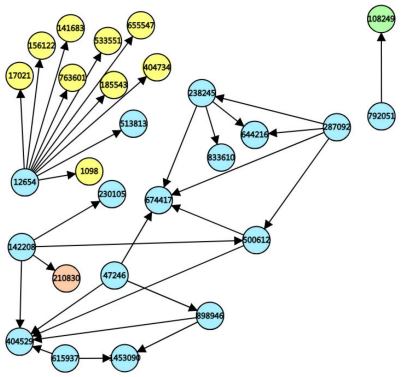


Figure13the spread of comedy from1990 to-2010

In 1970 to 1990,blue represents Comedy, pink represents R&B, green represents Country and orange represents folk

In 1990 to 2010,orange represents R&B, blue represents Comedy, green represents vocal and yellow represents Pop/Rock

Table3 the changing of comedy over years

Years	1930-1950	1950-1970	1970-1990	1990-2010
The number of influencers	4	6	3	3
The number of genres affected	5	7	4	3
The number of followers	21	44	44	25

By **Figure 10-13** and **Table 3** it can be obtained, for an interval in every 20 years, the evolution of the Comedy direction and speed. It spread the slowest in the first 20 years (from 1930 to 1950), and only affected 21 artists, but the direction of the genre's evolution was diversified. From 1950 to 1970 and from 1970 to 1990, the number of artists subjected to the Comedy genre increased, both being 44. As time went by, the types and amounts of genres that were affected began to decrease, and the genres that were affected began to stabilize. As it can be seen from **the Figure13** that most artists influenced by comedy belong to the genres of comedy and POP/Rock.

The chart below shows the changes of Avant-Garde from 1930 to 2010. It can be seen from the graph that there are 5 genres that are influenced by Avant-Garde, and POP/Rock is the most popular. This is also consistent with the results of our analysis in the second part. In the second question, POP/Rock is obtained by principal component analysis (pca) analysis and Avant - Grade about the Characteristics of the music has the very high similarity, which can be concluded that the POP/Rock _x0005_ is associated with Avant - Grade.

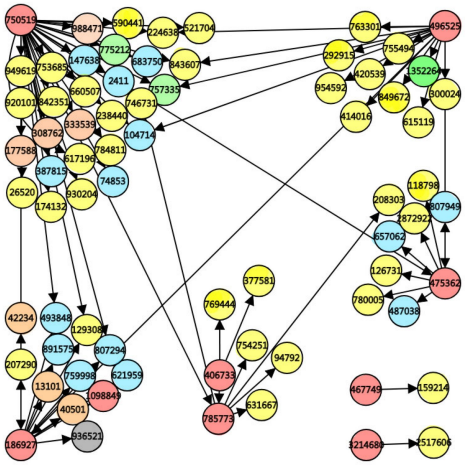


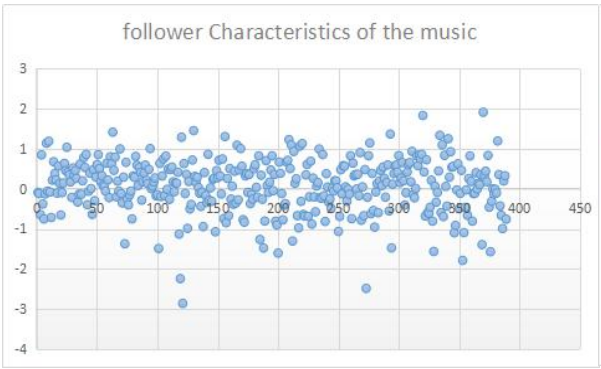
Figure14: Avant-Garde1930-2010

In 1930 to 2010 red represents avant-garde, blue represents electric , green represent classical , gray represents stage-screen, and orange represents New age

3.4 Task4

3.4.1 Difference analysis of Characteristics of the music and Type of Vocals

In this question, we need to analyze whether influencers actually affect the music created by their followers. Whether the music created by the followers is actually affected by the influencers depends on how similar the two principal components of the influencer are to the two principal components of the follower.The higher the degree of similarity, the greater the influence of followers, and vice versa. Therefore, our team made full use of the data of two main components: Characteristics of the music and Type of vocals. We took Pop/Rock as an example, and selected the influencer_id: 66915 to analyze the discrete distribution of the music index data of his followers, and draw a scatter chart. Figure 15、16are as follows:



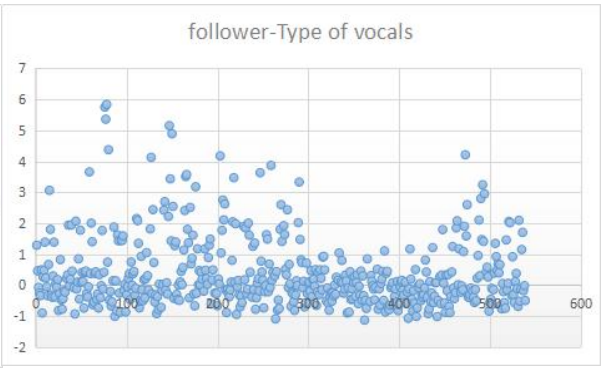


Figure15:follower-Characteristic of the music **Figure16:follower-Type of vocals**

From Figure 15 and 16 above, we can find that under the influence of the same influencer, the distribution of Characteristics of the music and types of vocals of different followers is relatively concentrated, that is, it can be considered that the Characteristics of the music and types of vocals of different followers have a high degree of similarity. So it's not hard to conclude that influencers do have some influence on followers.

3.4.2 Analyze the correlation between various indicators and popularity

This question needs to analyze whether each musical feature is more appealing, so it is necessary to analyze the correlation between these musical characteristics and popularity. The higher the correlation, the more appealing the musical feature.Through the literature, Pearson correlation coefficient and Spearman correlation coefficient can be used as the criteria of correlation evaluation. We analyzed the data in full_music_data to get the results of the following table:

table4： Other indicators of popularity Pearson correlation coefficient and Spearman correlation coefficient of color scale plot

	Pearson	spearman
danceability	0.16	0.166
energy	0.429	0.405
valence	-0.044	-0.076
tempo	0.137	0.146
loudness	0.446	0.486
key	0.049	0.053
acousticness	-0.56	-0.485
instrumentalness	-0.19	-0.197
liveness	-0.098	-0.074
speechiness	-0.057	-0.002

As it can be seen from the above table 4, energy, loudness have a strong positive correlation with popularity, while acousticness and popularity have a strong negative correlation. Therefore, energy and loudness are more "infectious".

As for the two characteristics of energy and loudness play similar roles in whether they influence the music of a particular artist,we chose a famouse artist (influencer): The Beatles, who had 606 followers. We draw radar charts of popularity, energy and loudness of his followers, as shown in Figure 17-19below:

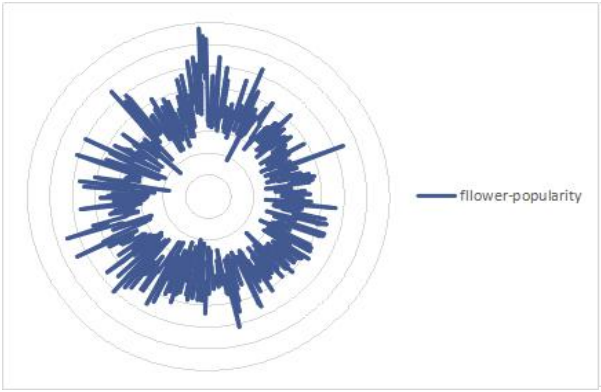


Figure17 followers' radar chart of popularity

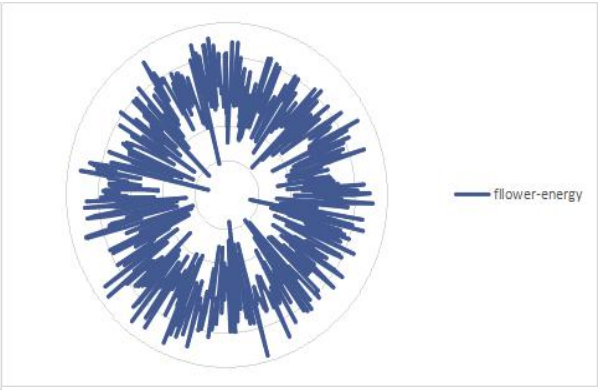


Figure18 followers' radar chart of energy

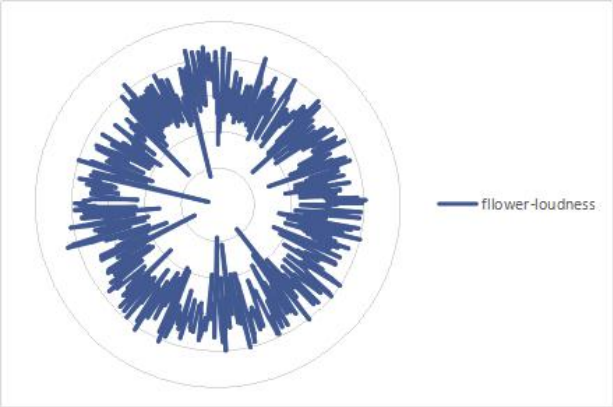


Figure19 followers' radar chart of loudness

From the above figures 17, 18 and 19, we found that the figures are similar. So we could see that more infectious indicators play a similar role in influencing the popularity of a particular musician.

3.5 Task5

3.5.1 The characteristics that might signify revolutions in musical evolution.

In this passage, we need to identify the characteristics of major revolutions that may have influenced the evolution of music. Its characteristic indicators include danceability, energy, loundness and popularity. Through the processing of full_music_data, we gave 50% weight to danceability and energy respectively, then we calculated the comprehensive index, and the one with the highest score is the important reformer of the genre. For example, the index Boyzone is obtained in Avant-grade. The results are shown in the figure below.

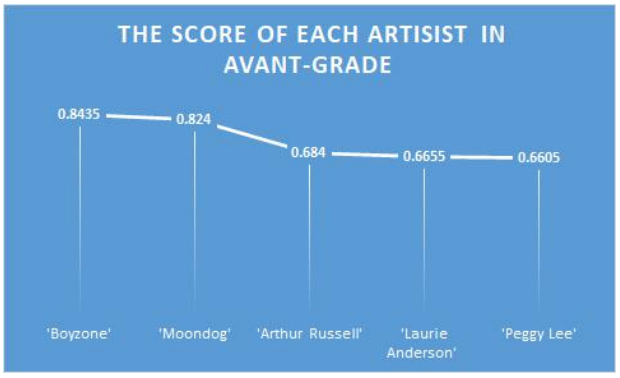


Figure20 The score of each artisist in Avant-Grade

3.5.2 Looking for turning point based on K-means method -- revolutionaries

In order to make a more systematic and scientific judgment, our group established the K-mean model with Matlab to find out the revolutionaries.
Principle and method

Its core indicators are SSE(Sum of Squared Errors)，and The formula for the sum of squared errors is

$$SSE = \sum_{i=1}^{k} \sum_{x \in C_i} ||x - \mu_i||_2^2$$

Ci is the ith cluster, x is a sample point in Ci, μi is the center of mass of Ci. (The mean of all the samples in Ci) Intuitively, this formula is the sum of squares of Euclidean distances from each sample point in the cluster corresponding to the current K value to the center point.

In order to make a more systematic and scientific judgment, our group established the K-mean model with Matlab to find out the revolutionaries.

This value can reflect the density of sample points in each cluster after clustering, that is, the smaller the SSE value is, the denser the sample points in each cluster.

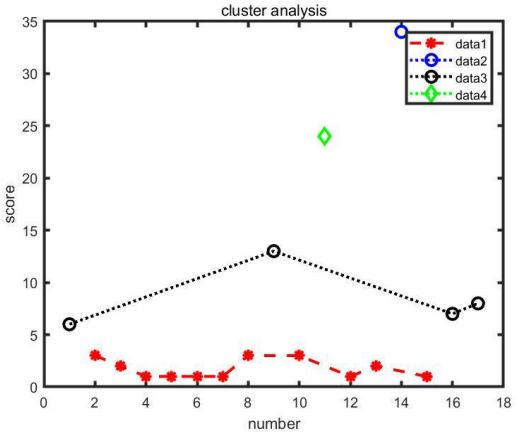


Figure21 The Classical cluster analysis

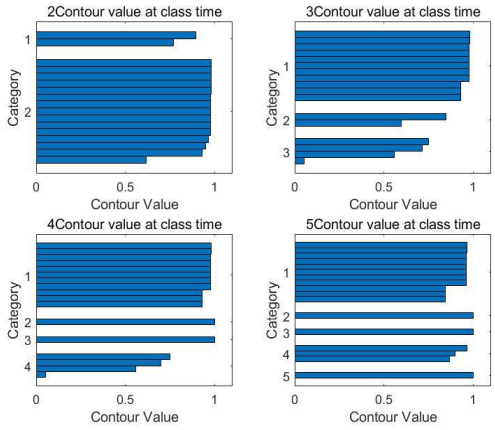


Figure22 Number of clusters in Classical

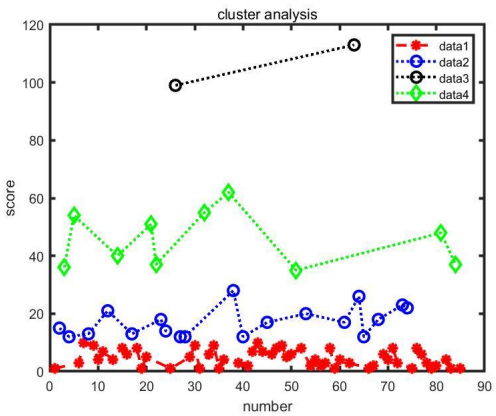


Figure23 The cluster analysis in Blue

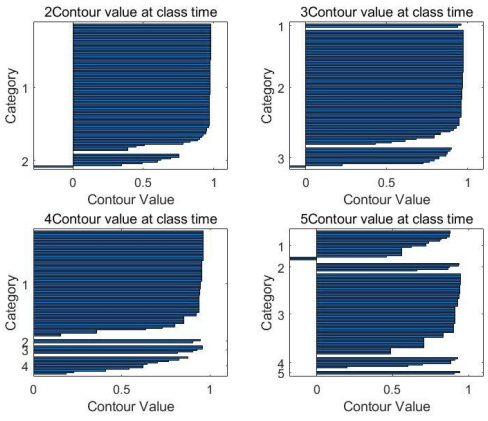


Figure24 The number of clusters in Blue

We used the clustering to determine the K-elbow method and made the image with MATLAB to find the main turning point and the serial number of the corresponding data. The main innovator in classical was John Cage. The major leaps in Blue were Jimmy Reed and John Lee Hooker. The main revolutionary in Avant-grade is Boyzone. This is also consistent with the major change agents corresponding to the characteristics analyzed in 3.5.1.

3.6 Task6

3.6.1 Analyze the influence processes of musical evolution that occurred over time in one genre.

This question involves analyzing the process of musical evolution over time within a genre. The process of musical evolution, that is to say, the emergence of a new genre, is the gradual evolution of a follower under the influence of an influencer. After the evolution of this genre, a certain number of followers will be generated accordingly. Then the identity of the artists of this genre will correspondingly change from followers to influencers, making another new genre come into being. Therefore, our group chose the evolution process of New age, which is relatively representative, as an example.

First, we used the influence_data to set the date back to the earliest point, 1930. We found all the genres produced in 1930 and assumed that subsequent genres were directly or indirectly influenced by these genres. The genres found are shown in Table 5 below.

Table5 The firstly genres

influencer_main_genre	
Blues	Classical
Comedy/Spoken	Country
Easy Listening	Folk
International	Jazz
Latin	Pop/Rock
R&B;	Religious
Stage & Screen	Vocal

Next, we screened out the artists who chose the New Age genre through Follower_Main_Gener and set the time to the earliest. Our team discovered that two artists, Mike Oldfield (489520), Penguin Cafe Orchestra (308762) meet the requirements. So, our team identified these two as the founders of the New Age genre, and then we found out how and when they were influenced by whom in which genre.

Table 6 the information of the firstly genres

influencer_name	influencer_main_genre	influencer_active_start	follower_name	follower_main_genre	follower_active_start
Steve Reich	Classical	1960	Mike Oldfield	New Age	1960
Kevin Ayers	Pop/Rock	1960	Mike Oldfield	New Age	1960
John Renbourn	Folk	1960	Mike Oldfield	New Age	1960

.....	.....	.....	.....	.....	.....

Table7 Information about New Age genre influencers

influencer_name	influencer_main_genre	influencer_active_start	follower_name	follower_main_genre	follower_active_start
Steve Reich	Classical	1960	Penguin Cafe Orchestra	New Age	1970
Terry Riley	Avant-Garde	1960	Penguin Cafe Orchestra	New Age	1970
John Adams	Classical	1960	Penguin Cafe Orchestra	New Age	1970

Finally, our team found the schools that influenced the generation of New Age: Classical, Pop/Rock and Folk.

So what genre did New Age influence as time went by? In order to solve this puzzle, we selected New age from the influencer_mian_genre and add follower_active_start, influencer_id, follower_id, follower_main_genre to filter the genre.The specific process is shown in the **Figure25** below:

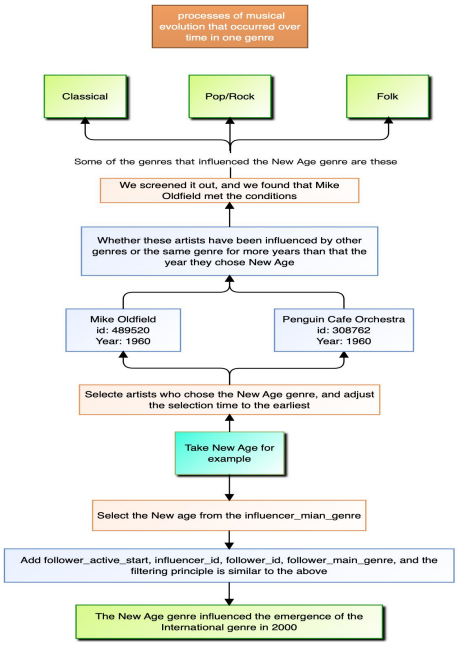


Figure25

3.6.2 Identify the indicators that reveal the dynamic influencers

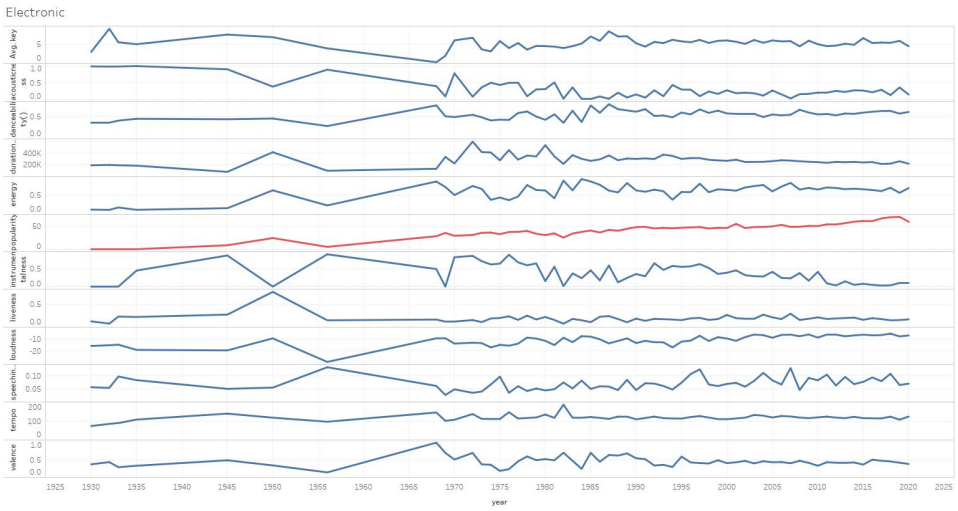


Figure26：Changes of indicators in Electronic schools over time

According to the graph shown above, we can find which index has a higher correlation with the popularity over time, so as to a more "infectious" index belonging to our own school. The results are as follows:

Table 8:Indicators similar to popularity indicators under various schools

gener	Similarity index1	Similarity index2
Comedy/Spoken	danceability	tempo
Country	loundness	tempo
Easy Listening	energy	dancebility
Electronic	energy	duration
Folk	dancebility	energy
International	duration	dancebility
Pop/Rock	dancebility	tempo
R&B;	tempo	dancebility
.....	.....	.....

From the above table, we can find the dynamic influence of different schools, the results are shown in the following table.

Table9:Dynamic Influencers under Various genres

Dynamic influencer1	Dynamic influencer2
Luke Kaufman	Sarah Ross
The Paul Butterfield Blues Band	Albert King
Alvin & the Chipmunks	-

Teyana Taylor	Carnage
Monty Python	Gossip
Jake Owen	Hank Williams III
.....	.....

3.7 Task7

In order to illustrate the cultural influence of music in time and environment, our group first analyzed the data from data_by_year. We used valence, danceability and popularity, which are the main characteristics of music that have been verified above, as indicators. And we explored change value of these three indicators every year from 1921 to 2020. After that, we combined historical events and social background to analyze the reasons for the turning point and predicted the future trend.

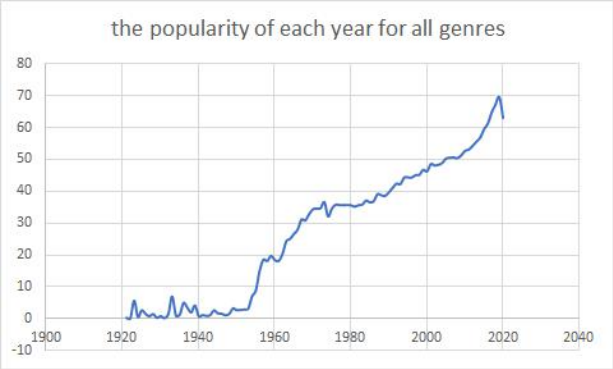


Figure27:the popluarity of each year for all genres

As we can see from the the trend chart in Figure27, music's popularity has increased over time, indicating that as more and more people like music, music has a higher and higher influence on people. It is worth noting that 1960 is the most significant turning point in popularity. In order to find out the reason, we put it in the historical context and found out that in the 1960s, there was a group of British bands visited the United States and they brought the Blues back to its homeland. The Blues played by white bands from Britain, such as The Rolling Stone, The Yardbirds, and The Animals, not only sparked an interest in The Blues among white Americans, but also started a Blues Revival.

And in 1962, thanks to two Germans, a group of older blues artists from the United States began to cross the Atlantic and tour Europe. Originally, they called the event the American Negro Blues Festival, but it was renamed the American Folk Blues Festival and was held every year until 1971. The tour featured videos of some of the most famous Blues artists in the country at the time, including Muddy Waters, T-Bone Walker, John Lee Hooker, Willie Dixon, Sonny Boy Williamson, , Lightnin， Hopkins, Roosevelt Sykes, Howlin' Wolf, Buddy Guy, Son House, Big Mama Thornton etc.

As a result, the decade between 1960 and 1970 saw a rapid development of music and a higher level of participation.

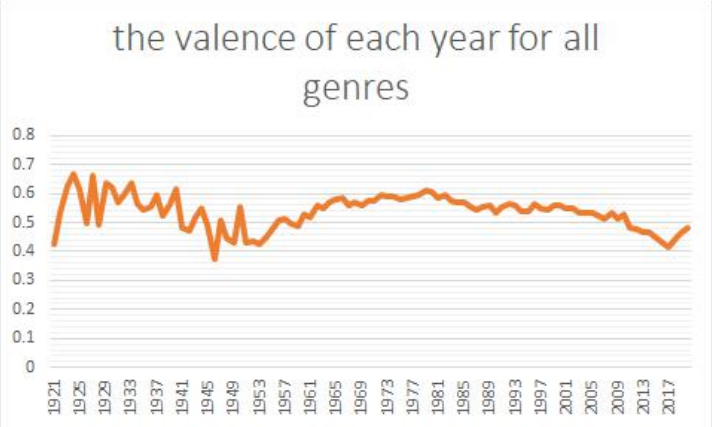


Figure28:the valence of each year for all genres

By analyzing the changes in the music style from 1921 to 2020, we could see that in the early stage of development, the overall music value of valence was very high, and its music style was mainly brisk and optimistic. However, by 1955 the music had become less cheerful and more depressing. It was also related to the Blues style of the time, which was composed by African American blues musicians.

In Task3, we analyzed the trend of genre over time. (Table in Fig. 3.3.2) The types and directions of genres are diversified, and their stability is strong. As the development of science and technology, especially the development of computers, have had a significant impact on digital music, which is why the popularity of Figure28 continues to rise.

4 Evaluation of the model

4.1 Strengths

- In the first question of design parameters, the factors are divided into 3 parts, including personal influence, school propagation speed and genres propagation direction, and the parameter results are visualized and mapped through directional network.
- In the second task, because we don't have the data of each frame of each music, in order to make up for this defect, we used principal component analysis and concluded that the coefficient of variation is used to deal with the index.
- In the third question, we used data visualization to add time to the local graph of directional network graph.
- When seeking the characteristics of the major revolution affecting the evolution of music, we divided the data at first, and then the model was simplified by the formula, and the "elbow map" was obtained by using the MATLAB of the K-means method to find out the inflection point, that is, the major leap.

4.2 Weaknesses

1、In the first question, our team did not use all the data, but made a targeted network by extracting the data.
2、In the second question, the group did not use the data of frame number when calculating the appropriate amount of long-duration features due to limited conditions, but replaced it with other factors. Therefore, the model has some defects.

5 Reference

[1] Liu Shanshan. A music recommendation system [D].] combining audio features with social labels Hubei: Huazhong University of Science and Technology ,2011.20
[2] Yuan Bin. A Study on the Classification of Music Schools Based on Deep Learning [D].]; and North University of Technology ,2019.
[3] Dan. A Study on Similarity Detection Algorithm for Music Style [D].] Dalian University of Technology ,2013.

6 A document to the ICM Society

From: Team 2126164, ICM 2021
To: The ICM Society
Date: February 2/8
Subject: Goals for measuring musical influence

In this paper,we will introduce the following information.

1、Models that measure the influence of music:

- Principal component analysis (PCA) : After 11 index dimension reduction, we will get two principal components, and 11 index shows the internal connection. Through the analysis of the coefficient matrix, we will name for the two main components: Characteristics of the music and the Type of vocals. Then, by analyzing whether this index of each genre is different, we can explain what is created between the same genre.
- Pearson correlation coefficient and Spearman correlation coefficient were used as the criteria for evaluating the correlation relationship. And we used them to determine whether there are differences in the music that is made between different genres. We used these two correlations to figure out which index is more associated with popularity.
- K-means elbow method can find out which revolutionaries are important and which genres represent the analysis of each.

2、The value of our model's impact on the web's understanding of music

- We built a neural network across genres and found parameters to measure the impact of music, including the influence of individual artists, the speed at which a genre spreads, and the direction in which a genre spreads. We can analyze the development process and speed of music through these three indicators.
- With the increase of data, our model changes. For example, when we estimate the similarity of songs, if the information of each frame in the music is obtained, octave spectrum contrast is used to roughly reflect the harmonic distribution of the music signal. Gaussian model can be added to more accurately judge the difference of each frame and then accurately judge the difference of songs on the basis of the model we made, which is an optimization on our model.

3、The significance of our evaluation of music is that the combination of music and culture contributes to the development of society and economy

Music can not only relieve the pressure caused by the fast pace of life and work, but also enrich People's Daily activities. In addition, pop music plays a certain role in promoting economic development. It can not only improve the living standard of the public, but also play a good role in promoting the improvement of public literacy. Today, the popularity of new media has provided a good opportunity for the development of pop music. In order to further promote the development of pop music, it must continue to integrate with other cultural elements, so as to expand the rendering ability of pop music. Under the mutual integration of pop culture and music culture, pop music is guided to develop together with The Times.