

**Problem Chosen**

**C**

**2023  
MCM/ICM  
Summary Sheet**

**Team Control Number**

**2307519**

---

# A Glimpse of Music Change through Influence Networks

## Summary

Music is the treasure of humanity. The interplay of artists and genres has also contributed to the continuous development of music. The purpose of this paper is to develop a model for measuring musical influence and to explore the evolutionary and revolutionary trends in examining artists and genres.

In TASK 1, we treat each artist as a node and represent the node-to-node influence with a directed line segment. Then we quantify the direct influence by adapting the TwitterRank model and quantify the indirect influence by constructing a transmission of indirect influence modeled after *Markov chains*. Finally, by studying the influence of individual nodes in a sub-network, we reveal the important people in this sub-network.

**Keywords:** ARIMA   Discrete Normal Distribution   Linear Regression   Hierarchical clustering   SVM

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Problem Background . . . . .	3
1.2	Our work . . . . .	3
<b>2</b>	<b>Assumptions and Notations</b>	<b>4</b>
2.1	Assumptions . . . . .	4
2.2	Notations . . . . .	4
<b>3</b>	<b>Task1: Network Model</b>	<b>5</b>
3.1	Network for influence . . . . .	5
3.2	Music influence . . . . .	6
3.2.1	Direct influence . . . . .	6
3.2.2	Indirect influence . . . . .	7
3.3	Some analysis of subnetworks . . . . .	8
<b>4</b>	<b>Task2: Similarity model and Clustering</b>	<b>9</b>
4.1	Similarity model . . . . .	9
4.1.1	Preparation of the model . . . . .	9
4.1.2	Establishment of the model . . . . .	10
4.2	Hierarchical Clustering . . . . .	10
<b>5</b>	<b>Task3: Similarities and influences</b>	<b>11</b>
5.1	Similarities and influences within the genre . . . . .	11
5.2	Similarities, differences and influences between genres . . . . .	13
5.3	Changes in genre over time . . . . .	13
5.4	The connection between genres . . . . .	14
<b>6</b>	<b>Task4 :The relationship between influencers and followers</b>	<b>15</b>
6.1	The relationship between the works of influencers and followers . . . . .	15
6.2	Influence factor . . . . .	16
<b>7</b>	<b>Task5: Features and Leaders for change</b>	<b>17</b>
7.1	Features that produce change . . . . .	17
7.2	Explore leaders for change . . . . .	18
<b>8</b>	<b>Task6: Genre evolution</b>	<b>19</b>
8.1	Dynamic index of genre evolution . . . . .	19
8.2	Analysis of specific factors . . . . .	20
8.3	An analysis of the influence of artists . . . . .	21
<b>9</b>	<b>Task7: Other factors</b>	<b>22</b>
<b>10</b>	<b>Analysis on Model's Sensitivity</b>	<b>23</b>
<b>11</b>	<b>Strengths and Weaknesses</b>	<b>23</b>
<b>A document to the ICM Society</b>		<b>24</b>
<b>References</b>		<b>25</b>

# 1 Introduction

## 1.1 Problem Background

Wordle is a popular trivia game offered by the New York Times, and many players report their scores on Twitter every day. Our team developed a set of models to investigate what factors influence the number of people participating in Wordle games and how Wordle words affect players' scores. We did the following.

•

## 1.2 Our work

1. We treat each artist as a node and represent the node-to-node influence with a directed line segment. Then we quantify the direct influence by adapting the TwitterRank model and quantify the indirect influence by constructing a transmission of indirect influence modeled after *Markov chains*. Finally, by studying the influence of individual nodes in a sub-network, we reveal the important people in this sub-network.
2. We use the *random forest algorithm* to reduce the dimensionality of the features of artists' works, then vectorize each feature of artists' works, followed by using min-max to normalize the feature values, and then measure the similarity between artists by the *cosine similarity*. Finally, we use *hierarchical clustering method* to analyze the correlation between artists.
3. We analyzed the correlations and influences between and within genres, using the Easy Listening genre as an example, in combination with the relationships of influence, and the degree of similarity. Based on this, we then identified the main characteristics of all genre styles.
4. Using the artist Al Di Meola as an example, we analyzed the relationship between influence and the degree of similarity of the works in relation to the influence model we developed, and developed further analysis based on each of these characteristics.
5. Using the folk genre as an entry point, we analyzed the change curve of each of its characteristic attributes over time, searched for curve inflection points, and identified the influencers of major changes using the similarity metric.
6. We defined dynamic indicators of stylistic evolution based on similarity indicators and selected the Vocal genre to analyze in detail the influential process of its musical evolution over time. The artist Frank Sinatra from the Vocal genre is also used as an example to explain the reasons for the stylistic changes of his works over time.
7. Combined with other sources, we analyzed the specific impact of environmental and temporal factors on music from four perspectives: political, economic, warfare, and technological.

## 2 Assumptions and Notations

### 2.1 Assumptions

- The relationship of influence between artists is related to their works, not their lives.
- The same artist can only belong to one genre.
- The relationship of influence between artists is credible and real.
- The effect between genres is considered only for the genre given in the data and does not include other genres.

### 2.2 Notations

Notations	Definition
$p_{ij}$	the direct influence from $i$ to $j$
$q_{nn+2}$	the indirect influence from $n$ to $n+1$
$\mathcal{P}$	the attenuation degree of the influence
$\eta$	the effective quantity for information
$Inf_i$	the music influence of $i$
$C_{ij}$	the similarity between $i$ and $j$
$\varphi(t)$	Variance of similarity over time
$\psi(t)$	Activity changes with time
$d_{ij}$	distance between $i$ and $j$

Table 1: Notations Table

## 3 Task1: Network Model

### 3.1 Network for influence

By observing the data, we find that each piece of data often contains several parts. The first is the unique ID of the influencer and the ID of the followers. So for these  $n$  authors, we can use an  $n \times n$  matrix to record the influence relationships between them.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad (1)$$

Where the value of  $a_{ij}$  is 1 or 0, indicating whether the author  $i$  has influence on the author  $j$ , respectively.

So from this matrix we can generate a network. We use nodes to represent each writer. Directed line segments are used to indicate that one node has a relationship to another node. We also use different colors to indicate the different schools that writers belong to.

We randomly selected a few authors to generate the following network as a schematic:

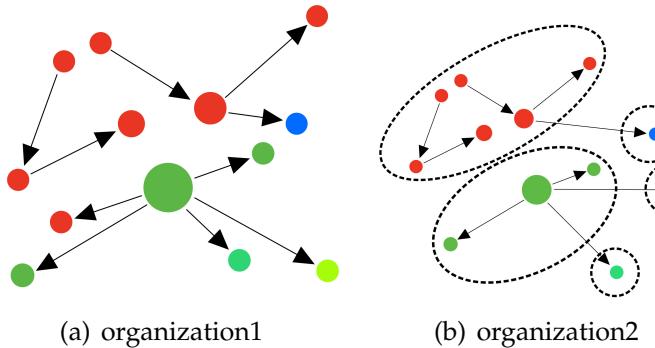


Figure 1: Organization

By combining all the artists' information into a matrix and generating a network, we will have a vast and complex network. In order to better analyze the network, we integrated our network into blocks, imitating the concept of community in complex network theory. Community structure defines a group of nodes in the network that are closely connected internally and sparsely connected externally as a community. These nodes usually have the same attributes or similar functions. Then in our network, we can artificially define a group of nodes belonging to artists of the same genre as a community.

The schematic diagram is as figure1(b):

Obviously, the community structure of this network is non overlapping, that is, each node can only belong to one community. This is a very important point. Through this point, we can use a larger node to replace the community, so that we can more intuitively observe the whole network. At the same time, we make the thickness of the directed line segment from node A to node B proportional to the number of small nodes in node A that affect the small nodes in node B.

In this way, we get a network graph based on all the data. Through this network graph, we can intuitively see the relationship between different communities and the important position of a community in the whole music field.

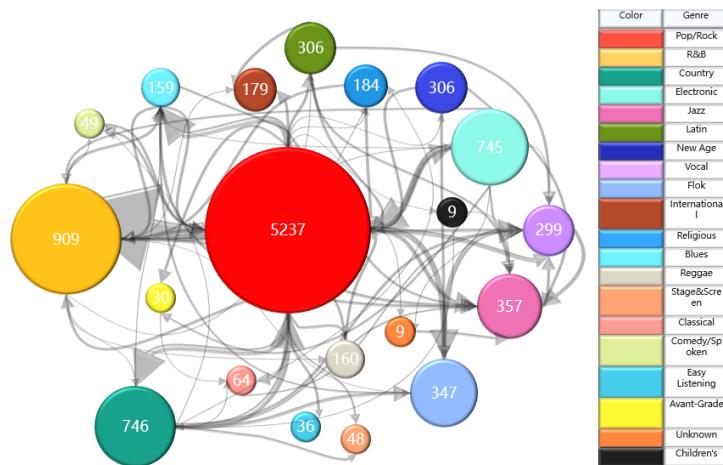


Figure 2: network for all the nodes at the community level

From this, we can roughly observe that pop, R&B, country and electronic play an

important role in the whole history of music development.

## 3.2 Music influence

### 3.2.1 Direct influence

The influence of nodes in the network has been studied for a long time. Many scholars such as singla and Goyal<sup>[1]</sup> define the influence as the similarity and correlation between nodes. However, based on the data given in influence date, we can not understand the similarity between artists. The only similarity is whether these artists belong to the same school, that is, whether these nodes belong to the same community. Therefore, we introduce the twitterrank model proposed by Jianshu Weng:

$$p_{ij} = \frac{|T_j|}{\sum_{\alpha:i \text{ follows } \alpha} |T_j|} \cdot sim_{ij} \quad (2)$$

It's the influence of user i on user j. It should be noted that this model is proposed to study the influence among users in social networks. So we only need to understand its general idea, in which *sim* represents the similarity between users. The first item on the right of the equation expresses the degree of difficulty that the message sent by the influencer is accepted by the follower. Based on the communication theory of community network structure, a piece of information is easier to spread within a community. At the same time, for nodes in the same community, they should have higher similarity. In order to determine the similarity between the two, we use entropy weight method to make a simple comparison: The similarity of same organization is 0.6337, The similarity of different organization is 0.3662.

According to the information obtained, it is determined that the *sim* between two nodes belonging to the same community is 0.63, while the *sim* between two nodes in different communities is 0.36.

Then we determine the first term on the right side of the equation, because a piece of information is more easily spread within a community, so it has a greater weight in the influence on the followers. Meanwhile, we consider that the difficulty of information transmission between two nodes increases with the increase of time. So let's write this quantity in terms of  $\eta$ . We call it the effective quantity of information.

$$\eta_{ij} = \begin{cases} \frac{\rho_1}{\log_{time} time} & \text{If they belong to the same community} \\ \frac{\rho_2}{\log_{time} time} & \text{If they belong to different communities} \end{cases} \quad (3)$$

Among them,  $\rho_1$  and  $\rho_2$  are the transmission difficulty coefficients of the same music genre and different music genres respectively. By consulting the music research materials, we set them as 1 and 0.65. We use log function to represent the decay of effective value of information over time, *time* is the time difference between two nodes.

So we get the following formula:

$$\begin{cases} \frac{\sum_{\gamma: \text{who follows the node } i} \eta_{i\gamma}}{\sum_{\alpha: \text{who the node } j \text{ follows; } s: \text{who follows } \alpha} \eta_{\alpha s}} \cdot 0.63 & \text{If they belong to the same community} \\ \frac{\sum_{\gamma: \text{who follows the node } i} \eta_{i\gamma}}{\sum_{\alpha: \text{who the node } j \text{ follows; } s: \text{who follows } \alpha} \eta_{\alpha s}} \cdot 0.36 & \text{If they belong to different communities} \end{cases} \quad (4)$$

Consistent with twi model,  $p_{ij}$  represents the influence of node i on node j, and the first term on the right of the equation represents the difficulty degree of the influence of node i on node j, which is reflected by the ratio of  $\sum_{\gamma: \text{who follows the node } i} \eta_{i\gamma}$  and

$\sum_{\alpha: \text{who the node } j \text{ follows}; s: \text{who follows } \alpha} \eta_{\alpha s}$ . In short, in a network, the larger the degree of a node, the easier it is to affect other nodes.

### 3.2.2 Indirect influence

At the same time, we take into account the indirect influence factor, an influential artists often influence on his following the often with what he has to follow the influence of the artist. For this kind of node between the process of passing from one generation to another, we simulate markov chain, the state and the state transition probability between replace influence of transmission attenuation degree between nodes:

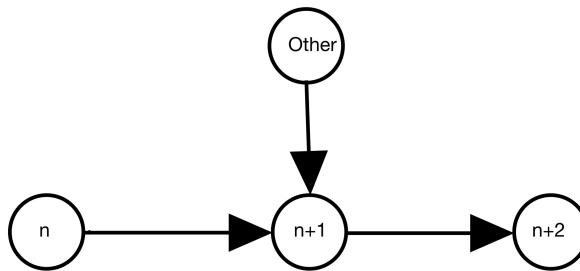


Figure 3: Indirect influence transmission chains

In this process, we first assume that the node to which the follower belongs is inactive, and the influence of node  $n$  on node  $n+1$  is  $\eta_{nn+1}$ , then the indirect influence of node  $n+1$  on the node of the  $n+2$  generation belonging to the first generation is:

$$q_{nn+2} = \frac{p_{nn+1} q_{nn+1}}{\sum p_{n+1}} \mathcal{P} \quad (5)$$

Where  $\sum p_{n+1}$  represents the sum of the direct influence of all other nodes on node  $n+1$ , and  $\mathcal{P}$  represents the attenuation degree of the indirect influence of node  $n$  on node  $n+2$ .

Obviously,  $0 \leq \mathcal{P} \leq 1$ . So in this network, starting from node  $n$ , any node that can be reached through a directed segment will be indirectly influenced by node  $n$  so the indirect influence of node  $n$  should be the sum of these values, which is obviously of the form  $\sum \alpha_n \mathcal{P}^n$ . So even though the whole network is huge, this value should converge. However, in order to prevent some artists from making this value too large to be reasonable due to the large number of indirect influences, we set a threshold  $Q$ . When the indirect influence on a certain transmission chain is less than this threshold  $Q$ , the transmission ends.

Here we set  $\mathcal{P}$  as 0.5, and we will test the sensitivity at the end of the paper. So we get a quantitative measure of the total amount of influence that a node generates:

$$Inf_i = \sum p_i + \sum q_i \quad (6)$$

Where  $\sum p_i$  represents the sum of the quantized values of the direct influence of all nodes directly pointed to by node  $i$ , and  $\sum q_i$  represents the sum of the quantized values of the indirect influence of all nodes having an indirect influence relationship with node  $i$ .

At the same time, through the reference of music history. We find that in the study of music history, people regard the connection between the three generations as a kind

of inheritance. So we make  $Q$  a little less than the indirect influence of the third generation artists. This can greatly simplify the calculation of the program.

### 3.3 Some analysis of subnetworks

We take the internal network of easy listening as the initial network, select all the artists associated with it from the whole complex network and join it to form a sub network.

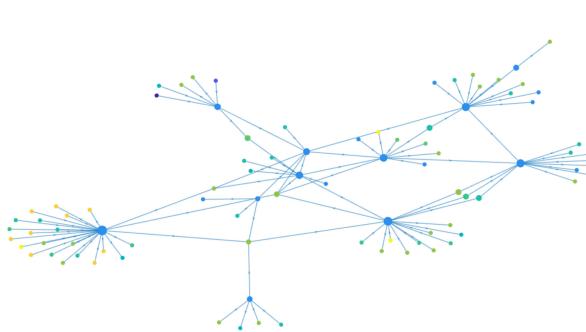


Figure 4: the subnetworks (different genres with different colors)

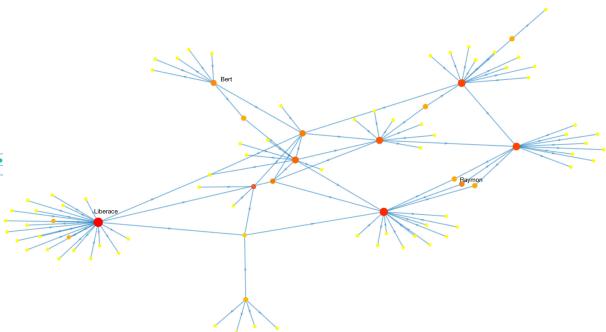


Figure 5: color depth with influence

Through the comparison of the two pictures, we can clearly see that the artists who have a great influence in this sub network are the founders of this genre and some influential figures in this genre. By looking up the history of music, we find that Andr and Liberace, which correspond to these points, have a relatively important influence in this genre. At the same time, we found that some artists outside the genre, such as Raymond, also had a greater impact, because in the whole network, they had enough influence to influence other genres. This also verifies that the influence index inf established by us can well reflect the specific influence of an artist.

## 4 Task2: Similarity model and Clustering

### 4.1 Similarity model

#### 4.1.1 Preparation of the model

For the quantification model of music similarity, we first extract the eigenvalues from the data set. Fortunately, the data set has classified the features of the song for us, and quantified each feature. But there are a lot of these features.

Therefore, we need to reduce the dimensionality of the data before building the similarity model.

We use the random forest algorithm to reduce its dimensionality, First we define the information entropy:

$$H(X) = - \sum_{i=1}^n m_i \log(m_i) \quad (7)$$

where  $m_i$  is the probability of occurrence of the event. After that we define the conditional entropy:

$$H(Y | X) = H(X, Y) - H(X) \quad (8)$$

The information gain:

$$g(D, A) = H(D) - H(D | A) = I(D, A) \quad (9)$$

The decision tree uses the Bootstrap Aggregating method to obtain the information gain of the features, and if the information gain is greater, the feature importance is greater.

We calculated all the eigenvalues and obtained their importance:



Figure 6: feature importance

We set the valve field to 0.5, and we can see that the two quantities, Key and Mode, are obviously smaller than other quantities and are smaller than the valve field, so we eliminate these two eigenvalues.

#### 4.1.2 Establishment of the model

Next, we take a set of eigenvalues of an artist as a vector:

$$\vec{X}_i = (x_{i1}, x_{i2}, \dots, x_{i12}) \quad (10)$$

They correspond to a, b, c, d of the artist  $i$ .

Then, for the similarity between any two artists, we can express it by the cosine of the Angle between two vectors.

$$C_{ij} = \frac{\vec{X}_i \cdot \vec{X}_j}{\sqrt{\sum_{n=1}^{12} x_{in}^2 \cdot \sum_{n=1}^{12} x_{jn}^2}} \quad (11)$$

Obviously, the closer the absolute value of  $C_{ij}$  is to 1, the higher the correlation. For the similarity, when  $C_{ij}$  is low, it means that they are quite different.

Now, of course, up until this point, because each of these data classes have different dimensions, we normalized the data and we used what is called min-max Normalization. The transformation formula is as follows:

$$x^* = \frac{x - \min}{\max - \min} \quad (12)$$

To understand the relationship between artists and genres, we randomly selected ten artists from different genres, calculated their similarity to each other, and produced the following heat map:

	Jazz1	Blues1	Classical	Blues2	Latin	Country	Jazz2	Vocal	Reggae	R&B
Jazz1	1.00	0.91	0.84	0.89	0.92	0.91	0.93	0.82	0.81	0.88
Blues1	0.91	1.00	0.93	0.99	0.98	1.00	0.99	0.94	0.91	0.98
Classical	0.84	0.93	1.00	0.97	0.96	0.94	0.93	0.99	0.99	0.95
Blues2	0.89	0.99	0.97	1.00	0.99	0.99	0.98	0.98	0.96	0.99
Latin	0.92	0.98	0.96	0.99	1.00	0.99	0.98	0.97	0.95	0.99
Country	0.91	1.00	0.94	0.99	0.99	1.00	0.99	0.95	0.92	0.99
Jazz2	0.93	0.99	0.93	0.98	0.98	0.99	1.00	0.93	0.91	0.98
Vocal	0.82	0.94	0.99	0.98	0.97	0.95	0.93	1.00	0.99	0.96
Reggae	0.81	0.91	0.99	0.96	0.95	0.92	0.91	0.99	1.00	0.94
R&B	0.88	0.98	0.95	0.99	0.99	0.99	0.98	0.96	0.94	1.00

Figure 7: similarity

## 4.2 Hierarchical Clustering

Then we conduct hierarchical clustering of these artists according to their similarity, so that we can get a classification map of them. By comparing this classification with their genre, we can know whether the similarity between artists is related to their genre.

Hierarchical clustering is a common way to categorize the data, here we stated briefly about this method First of all, we use the reciprocal of the similarity between artists to represent the distance  $d_{ij}$  between  $i$  and  $j$ :

$$d_{ij} = \frac{1}{c_{ij}} \quad (13)$$

But for those nodes with negative similarity, we assume that the distance between them is greater than the distance between any set of nodes with positive similarity. Therefore, we define the distance between them as:

$$d_{ij} = 1 + \frac{1}{c_{ij}} \quad (14)$$

And then all of the minimum distance  $d_{pq}$  consolidation of the artists who are  $p$  and  $q$  involved into a class  $r$ , and that he and other artists is defined as the distance between them both average distance with other artists.

That is:

$$d_{rs} = \frac{1}{2}(d_{ps} + d_{qs}) \quad (15)$$

And so on until everyone is in the same category.

The result is shown in the figure:

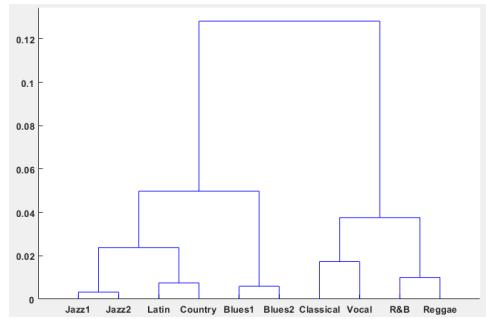


Figure 8: Hierarchical clustering result

We can see that in the same genre blues1 and blues2, jazz1 and jazz2 indeed have a higher similarity, that is to say, the artist has a higher similarity of the same genre. But some artists may be in a border of two schools, the birth of at the same time they also bring some changes to the genre and development.

## 5 Task3: Similarities and influences

### 5.1 Similarities and influences within the genre

In order to compare the degree of similarity and influence within genres, we need to compare the similarity of the styles of some artists within a specific genre.

We have selected some of the artists in the Easy Listening genre as examples for our analysis, and their influence relationships are shown in the figure below:

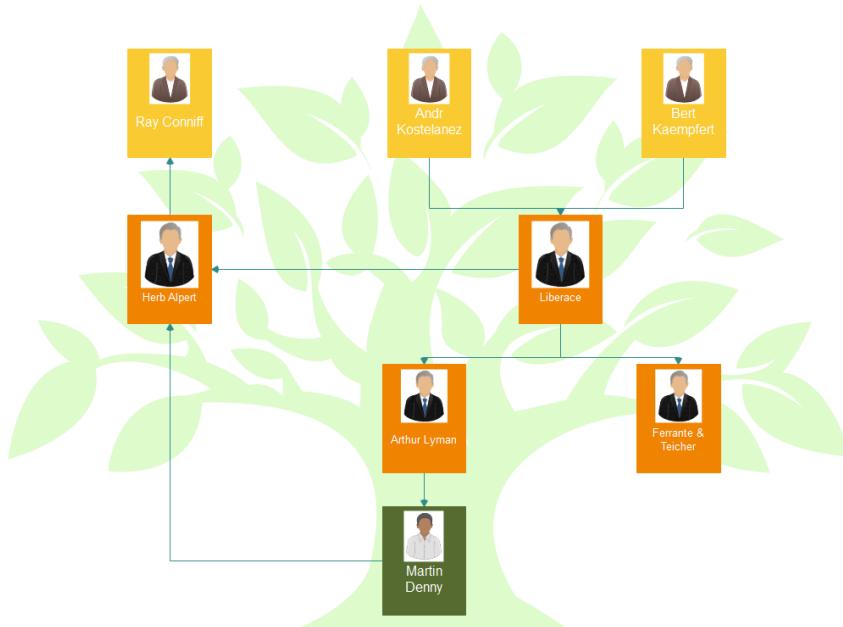
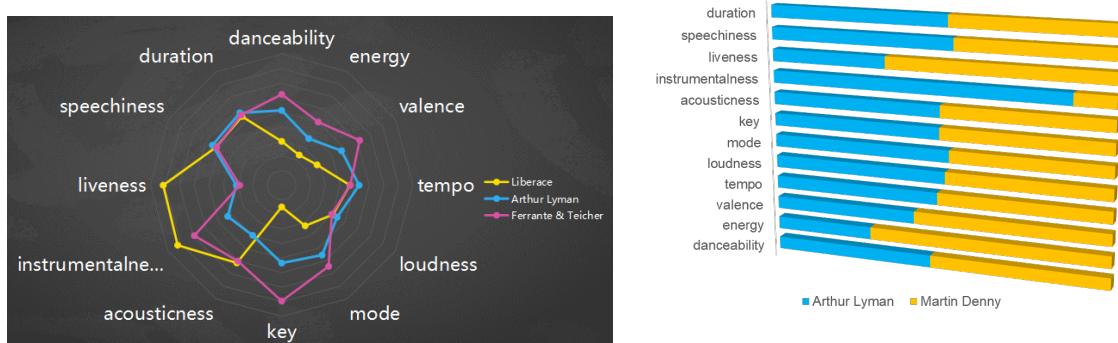


Figure 9: Selected artist-influenced relationships in the Easy Listening genre

We first examine the three artists Liberace, Arthur Lyman, and Ferrante & Teicher, according to the figure above, Liberace influenced Arthur Lyman and Ferrante & Teicher, and we select the average of all three of their works based on their influence. The results are shown below. The results are shown in the following figure:



(a) Comparison of the styles of Liberace, Arthur Lyman, Ferrante & Teicher's works  
(b) Comparison of the styles of Arthur Lyman and Martin Denny's works

Figure 10: Comparison of artists

From the figure, we can see that the three artists remain highly similar in duration, speechiness, tempo, and loudness, verifying that Arthur Lyman and Ferrante & Teicher are influenced by Easy Listening in several distinctive aspects. Liberace, but the differences in liveness are significant, suggesting that as the network evolves, the later members of the genre perform less live. Moreover, danceability, energy, and validity have all improved compared to Liberace in the 1940s, suggesting that the music of Easy Listening was gradually developed to be more danceable, energetic, and active.

After that, we then analyzed the characteristics of the works of Arthur Lyman and the follower Martin Denny, and the results are shown in Figure 10(b)

From the diagram, we can find that the characteristics of Arthur Lyman and Martin Denny's works are mostly similar, except for the three characteristics of energy, liveness, and instrumentalness, which may be influenced by the environment. This shows that Arthur Lyman has a great influence on the style of Martin Denny's works.

## 5.2 Similarities, differences and influences between genres

In order to explore the similarities, differences and influences between genres, we expressed the characteristic quantity of each genre in terms of the average characteristic quantity of all works in that genre based on the data in the full\_music\_data and influence\_data tables. And we visualize it as shown in the figure below:



Figure 11: Comparison of the styles of each genre

### 5.3 Changes in genre over time

In order to explore the changes that occur in a genre over time, we analyze the classical genre as an example.

We counted the mean values of each attribute at each interval from 1930-2010, after which we calculated its standard deviation, where the four attributes of valence, instrumentalness, popularity, and key fluctuate significantly, and we made a graph of them as follows:

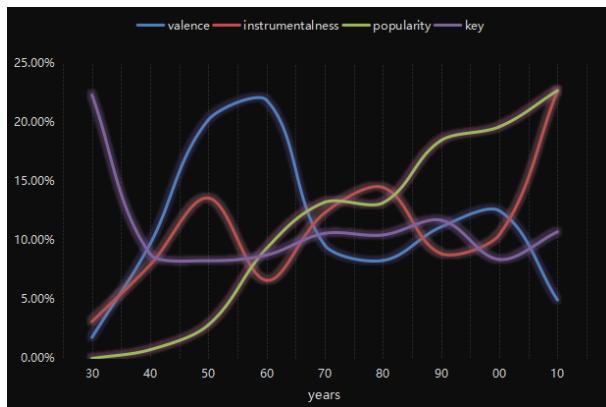


Figure 12: Changes in classical genres over time

We can see a general upward trend in the popularity of classical music, with the 1940s being the turning point for key, compared to the 1930s, when the tone of classical music slowed down from the 1940s and was inherited all the way through. at the same time, valence increased sharply in the 1940s, which we presume was influenced by the Second World War.

The 1950s were the turning point for instrumentalness turned a corner, with a sharp decrease in instrumentalness values, accompanied by a further increase in valence values to a peak, indicating the inclusion of more vocals in the music and a high level of emotion. In the 60s, the values of valence decreased sharply, accompanied by an increase in instrumentalness, indicating that the effect of certain events that occurred in the 50s In the 1970s, the popularity of classical music declined for the first time, while other characteristics changed little, probably due to the influence of the explosion of other musical styles.

### 5.4 The connection between genres

In order to explore the degree of association between genres, we first analyze them from the perspective of the originator. We define the originator of a genre as the artist who first developed the style and influenced the largest number of people within the genre.

We counted the originators of all genres and the relationships of influence between them, as shown in the following figure:

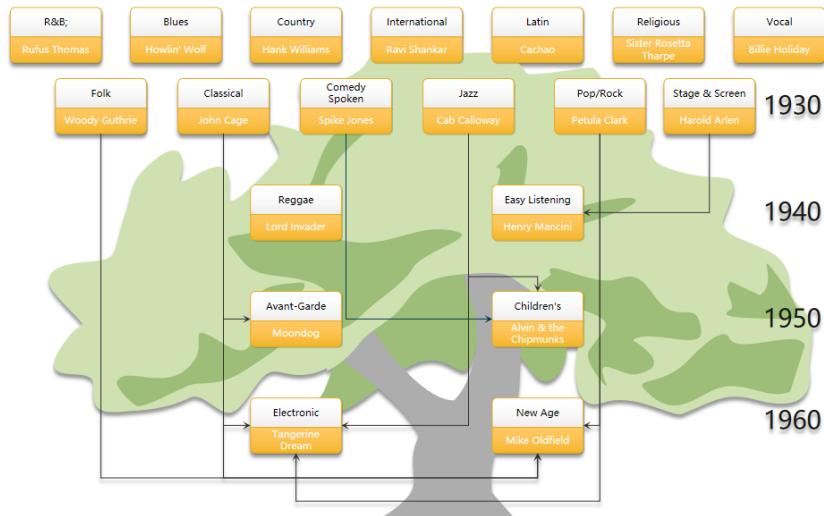


Figure 13: The relationship of influence between the originators of each school

In the figure, we find 15 genres whose originators were formed in the 1830s, and in the 1840s, the Reggae and Easy Listening genres emerged, where Henry Mancini, the originator of Easy Listening, was influenced by Harold Arlen, the originator of Stage&Screen. Arlen's influence led to the creation of Easy Listening.

In the same way that the Avant-Garde genre was influenced by the Classical genre in the mid-1850s, the Children's genre was influenced by the Jazz genre and the Comedy Spoken genre together. By the 1860s, the Electronic genre emerged from the combined influence of the Jazz and Classical genres and the Pop/Rock genre, and the New Age genre was born from the combined influence of the Classical and Pop/Rock genres.

Based on the data of works in each genre, we can dig out the salient features of each genre, as shown in the following table:

Genre	Characteristic	
Pop/Rock	energy	valence
Vocal	danceability	acousticness
International	valence	acousticness
R&B	valence	key
Jazz	acousticness	instrumentalness
Comedy/Spoken	speechiness	explicit
Latin	energy	valence
Country	valence	mode
Blues	energy	valence
Children's	danceability	instrumentalness
Electronic	instrumentalness	duration
Classical	instrumentalness	explicit
Stage & Screen	acousticness	instrumentalness
Folk	key	acousticness
Avant-Garde	loudness	instrumentalness
Religious	energy	tempo
Reggae	mode	instrumentalness
New Age	instrumentalness	duration
Easy Listening	loudness	instrumentalness

Table 2: Characteristics of each genre

From the table we can get the same characteristics of these genres, such as Pop/Rock, International, R&B, Latin, Country, and Blues all have significant values of valence, which means that their music is full of positive emotions.

## 6 Task4 :The relationship between influencers and followers

### 6.1 The relationship between the works of influencers and followers

To explore whether followers' musical styles are influenced by influencers, we need to consider the following online scenario, where one artist follows multiple influencers at the same time.

Taking Al Di Meola in The 1970s as an example, according to The data of influence-data, he was influenced by six artists including Julian Bream, John McLaughlin, Jim Hall, Jimi Hendrix, Miles Davis, and The Beatles. The network structure is shown as follows:

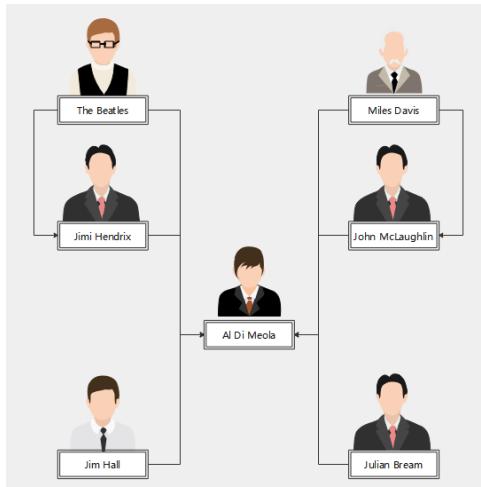


Figure 14: Network structure of Al Di Meola

We use the statistical data in `data_artists` table to formulate corresponding eigenvalue phasor for each of their works, and then calculate the similarity between artists as followers and influencers.

Meanwhile, according to the influence model in Task1, we calculate the influence of six influencers, among which John McLaughlin, Jim Hall, Miles Davis and followers Al Di Meola belong to the same Jazz genre, while the others belong to different genres.

The results are as follows:

	influence	simarity
<b>The Beatles</b>	0.7864	0.8429
<b>John McLaughlin</b>	0.6794	0.7627
<b>Miles Davis</b>	0.4376	-0.3557
<b>Jimi Hendrix</b>	0.3467	0.2738
<b>Jim Hall</b>	0.2467	-0.1849
<b>Julian Bream</b>	0.1349	0.0705

Table 3: the table of influence and simarity

It can be found from the table that Miles Davis and Jim Hall had a negative influence on Al Di Meola, indicating that Al Di Meola innovated the style of Miles Davis and Jim Hall's works, while he basically inherited the style of others' works.

If we only consider the numerical size, we find that the influence of the influencer is related to the similarity between the influencer and the follower's works, that is, the more influential the person is, the more likely the follower will imitate him. But at the same time, some artists will be influenced by some influential artists and make some big changes to the style of the influencer to form their own style.

## 6.2 Influence factor

In order to further explore which features of influentials' works influenced their followers, we selected two of them, The Beatles and John McLaughlin, who had The greatest influence on The style of Al Di Meola's works, and analyzed The detailed data of all aspects of their works, as shown in The figure below

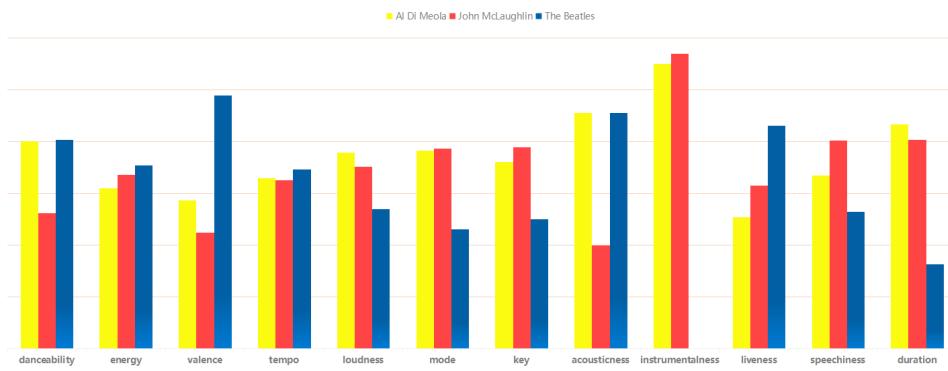


Figure 15: eigenvalue comparison

We can see from The table, The AI DI Meola in tempo, mode, key, instrumentalness, mainly inherited The style of The John McLaughlin duration, in danceability acousticness aspects mainly inherited The style of The Beatles.

So we can conclude that influencers are contagious in certain ways.

## 7 Task5: Features and Leaders for change

### 7.1 Features that produce change

In order to identify revolutionary features in the evolution of music, we can take a decade as a time unit and plot the change curve of each feature of the genre over time, and if there is a large fluctuation, the feature belongs to a major leap.

We counted folk genre from the 1950 s to 2010 s work characteristics change, we calculate the variance of each characteristic value, after selecting the obvious fluctuations in the eigenvalue instrumentalness, speechiness and popularity made their variation, the result is shown in the following figure:

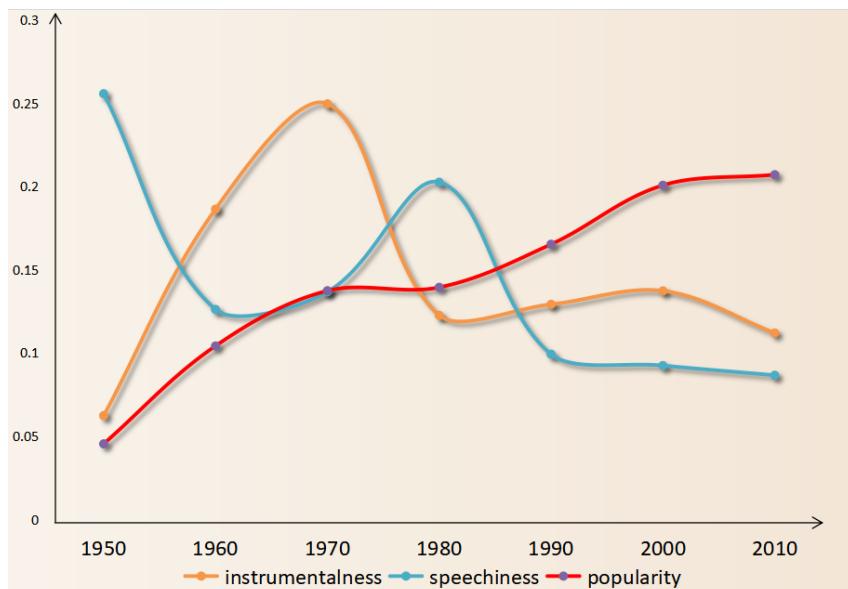


Figure 16: Some change in eigenvalues

From the graph we can see that in the 20 years from 1950 to 1970, there was a sharp decline in speechiness, accompanied by a sharp rise in instrumentalness, which, combined with a dramatic increase in popularity, leads us to speculate that perhaps folk emerged as the star artists who drove the genre forward.

## 7.2 Explore leaders for change

To explore who caused this change, we extracted the characteristics of all artists within the genre in that time period, and used the cosine similarity in task2 to calculate the degree of similarity between the characteristics of their artists' works and the characteristics of the genre between this point in time. If the degree of similarity is greater and the artist has some influence, then he is the change agent of the genre.

We have listed some of the most influential artists in the 1950s and Folk genre. The results are shown below:

Name	similarity	influence
The Kingston Trio	0.948	8.96
Ramblin' Jack Elliott	0.903	10.2
John Fahey	0.877	9.31
Joan Baez	0.702	15.5
Dave Van Ronk	0.852	9.74

Table 4: Similarity and influence of some Folk genre artists

From the table, we can find that Joan Baez is the artist who has the most influence and whose style is most inconsistent with the previous style of the genre. Therefore, we can infer that Joan Baez is a reformer of this era.

We select the average data of his works in this period and the average data of all Folk works in this period to calculate the distance between them, and measure his contribution to the style of the genre. The farther away from the average, the greater the contribution is. The result is shown in the figure below:

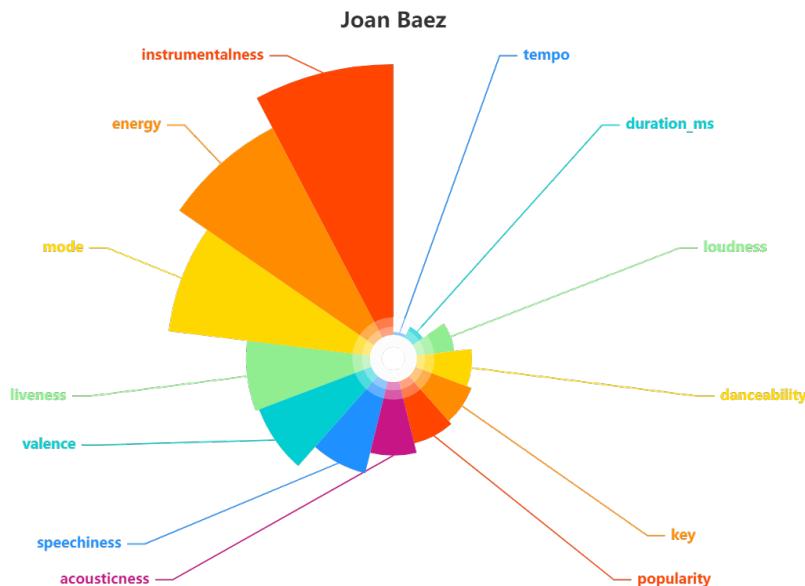


Figure 17: Joan Baez

From the graph, we can find that the contribution of Joan Baez's instrumentalness feature far exceeds the other features, so we can be sure that Joan Baez is the driving force that caused the stylistic change in the Folk genre.

The history of the development of the Folk genre<sup>[2]</sup>. shows that in the 1960s, Joan Baez combined the Folk genre with the Rock genre, introducing rock instruments such as the guitar to the Folk genre, thus driving a significant increase in the Folk genre's instrumentalness characteristics, further supporting our speculation.

We used the same method to filter out other genre changers and the results are shown in the following table:

Genre	Avant-Garde	Blues	Children's	Classical
Name	Terry Riley	Muddy Waters	Alvin	John Cage
Genre	Latin	New Age	Pop/Rock	R&B;
Name	Antnio Carlos Jobim	Mike Oldfield	The Beatles	Marvin Gaye
Genre	Comedy/Spoken	Country	Easy Listening	Electronic
Name	Spike Jones	Johnny Cash	Henry Mancini	Kraftwerk
Genre	Reggae	Religious	Stage & Screen	Vocal
Name	Toots & the Maytals	James Cleveland	Ennio Morricone	Billie Holiday
Genre	International	Easy Listening	Jazz	
Name	Ravi Shankar	loudness	Miles Davis	

Table 5: Change makers in every genre

## 8 Task6: Genre evolution

### 8.1 Dynamic index of genre evolution

According to the community model of Task1, we can regard a genre in the influence network as a community. When a genre changes greatly, the variance of the similarity of songs belonging to this genre in this period will expand. When the songs of this genre tend to be unified, their variance will decrease. We define the dynamic variance is " $\varphi$ ".

And then, we ues 10 years as a period to get a change chart of  $\varphi(t)$ :

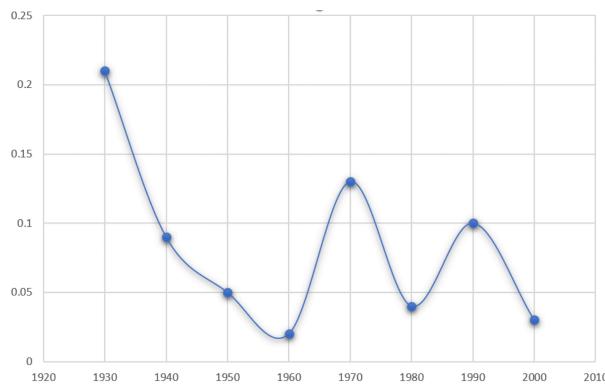


Figure 18: Variance of similarity over time

We can intuitively see that the vocal school was in the era of change in 1930s and

1970s. At other times, the style area is stable.

But at the same time, the vitality of a school is also an important indicator to measure the change of a school. For a community in a network, we introduce the view of dynamic network to record the snapshot of network structure in each period. We use the size of the node to represent the degree of activity within the network, and use the links with other schools to represent the degree of communication activity with other schools. Finally, we use the sum  $\psi$  of the size and degree of the node and the node to represent the prosperity and decline of a community.

$$\psi(t) = \text{size}(t) + \text{indegree}(t) + \text{outdegree}(t) \quad (16)$$

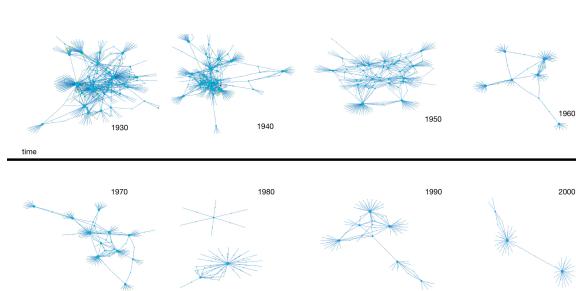


Figure 19: vocal-organization

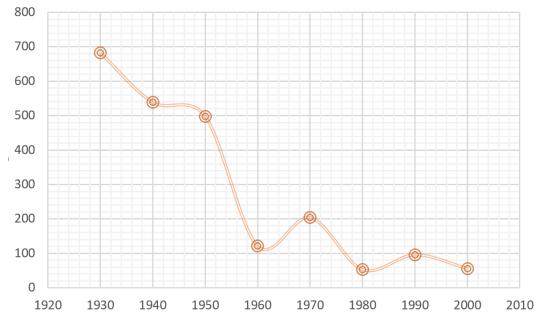


Figure 20:  $\psi(t)$  of every year

Obviously, we can see that after experiencing the glory of the 1930s, vocal gradually declined. In the 1970s and 1990s, vocal actively absorbed some characteristics of other schools, such as jazz and Pop. However, it can not avoid the gradual decline of this community.

## 8.2 Analysis of specific factors

Then, in order to explore the specific changes of the style of vocal school over time, we make statistics of the main changes in the data of vocal.

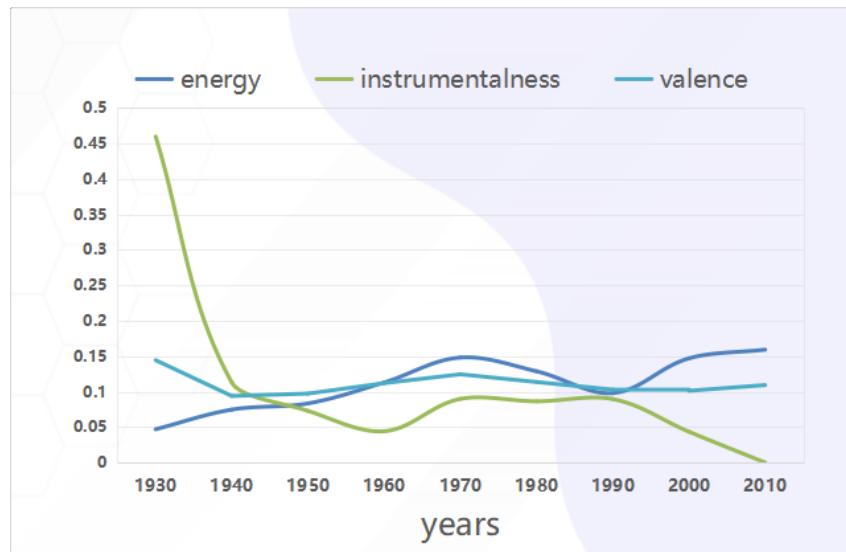


Figure 21: The main factors of style change

In Task3, we summarized the main characteristics of each genre, where the Jazz genre is mainly characterized by the use of instruments, the Pop/Rock genre is mainly characterized by high energy attribute value, and the R&B genre is mainly characterized by high valence attribute, which is more aggressive.

In the 1970s and 1980s, the Vocal genre was briefly influenced by the R&B genre, but its valence values did not fluctuate significantly, so we can assume that the R&B genre did not have much influence on the Vocal genre in terms of valence.

The influence of the Pop/Rock genre on the Vocal genre is obvious, as the Pop/Rock genre began to influence the Vocal genre from the 1950s onwards, and the value of the energy attribute of Vocal increased rapidly. By the 1980s, the Pop/Rock genre no longer influenced the Vocal genre, and Vocal's energy attribute values fell. In the 1990s, the Pop/Rock genre once again entered the ranks of influencers, and Vocal's energy attribute values began to climb again. Therefore, we can speculate that the characteristics of Vocal's energy were mainly influenced by the Pop/Rock genre.

### 8.3 An analysis of the influence of artists

In order to explore how artists are influenced with the passage of time, we select Frank Sinatra, the main character of vocal school, for specific analysis.

As a veteran artist who became famous in the 1920s, Frank Sinatra did not have many influences during his career. According to statistics, he was only influenced by the Vocal genre five times in the 1930s, the Jazz genre twice, and the Vocal genre once in the 1940s.

Based on the data of his career releases, the characteristics of his works were counted every ten years, and the fluctuating attributes of instrumentalness, liveness, and acousticness were plotted over time, as shown in the following graph:

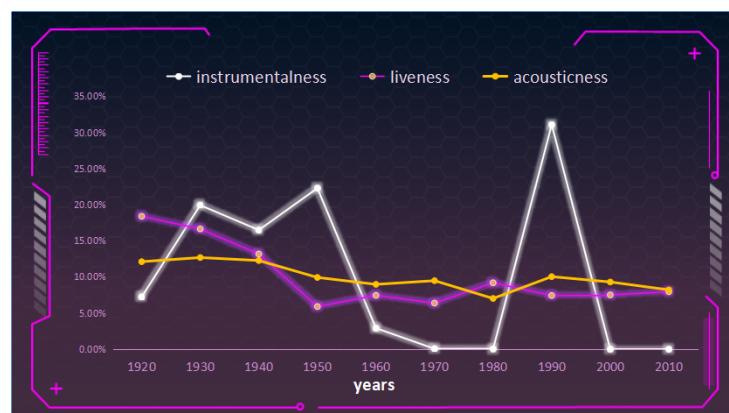


Figure 22: The main changes of Frank

It is clear from the graph that Frank Sinatra's instrumental use improves significantly in the 1930s as he is influenced by two composers of the Jazz genre. He was then influenced by the Vocal genre in the 1940s, and the value of the instrumentalness attribute then slipped. The year-to-year decline in the liveness attribute is most likely due to the popularity of electronic playback technology, to the extent that his performances were more often transmitted via television. The spike in his instrumentalness attribute in the 1990s was likely influenced by the explosion of electronic music at the time.

## 9 Task7: Other factors

We make a curve of eigenvalues over time based on the data\_by\_year data and selected a few eigenvalues with large variance:

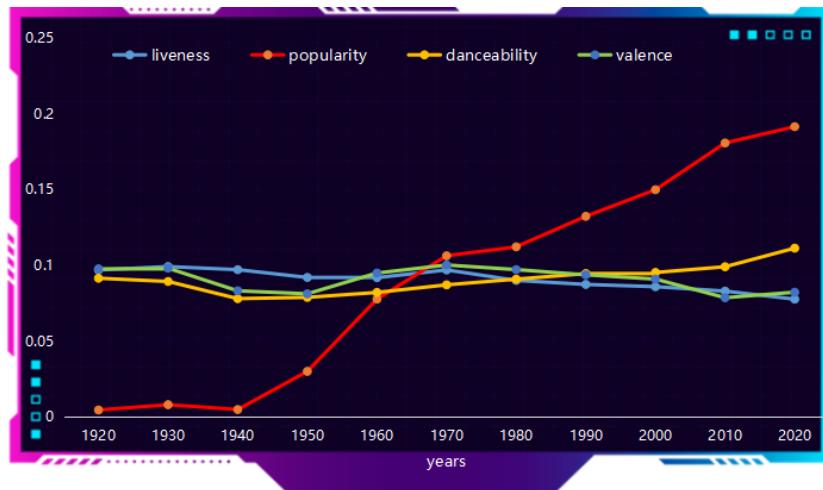


Figure 23: The main changes of Frank

We can explore the impact of fluctuations in the curve in the following ways:

- **Impact of the World War II**

In the 1940s, World War II broke out, and because the war had a melancholy color, the works created in the 1940s were also more negative (less valence) and less suitable for dance, resulting in less danceability. At the same time, Europe was caught in the middle of the war, and a large number of European artists immigrated to the United States for refuge. After World War II, the music artists who immigrated from Europe and the native American music artists influenced each other, which made the music develop well. Coupled with the relatively peaceful international environment after World War II, people had more energy and leisure time to enjoy music, and the popularity of music had a great development after World War II<sup>[3]</sup>.

- **The impact of information technology**

The emergence of the Internet in the 1970s and its subsequent development diversified the vehicles and media for disseminating musical works and made it cheaper and more convenient, and the popularity of music continued to increase. At the same time, however, offline music was more costly and therefore accounted for less than the ease and affordability of distributing music online, thus reducing the liveness characteristics of music.

- **The impact of politics**

During the anti-war and civil rights movements in the United States in the 1960s, rock and roll made a huge impact on mainstream culture as a powerful weapon for youth subculture groups. During these movements, rock music was loved by many people. The danceability of the music increased due to the passionate rhythm of rock and roll, which made people dance to the music involuntarily<sup>[4]</sup>.

- **The impact of the economy**

In the 1930s, the economic crisis broke out across the United States, and the unemployment rate rose sharply in the face of closures and bankruptcies brought about by the Great Depression. The record industry at the time was hit hard. The music industry was hit hard and the popularity of music declined from the 1930s to the 1940s<sup>[5]</sup>

## 10 Analysis on Model's Sensitivity

In Task 1, we introduce the attenuation coefficient  $\mathcal{P}$  and assign it a value of 0.5 in order to measure the effect of indirect influence in the influence network, and now we do a sensitivity test on the attenuation coefficient  $\mathcal{P}$ .

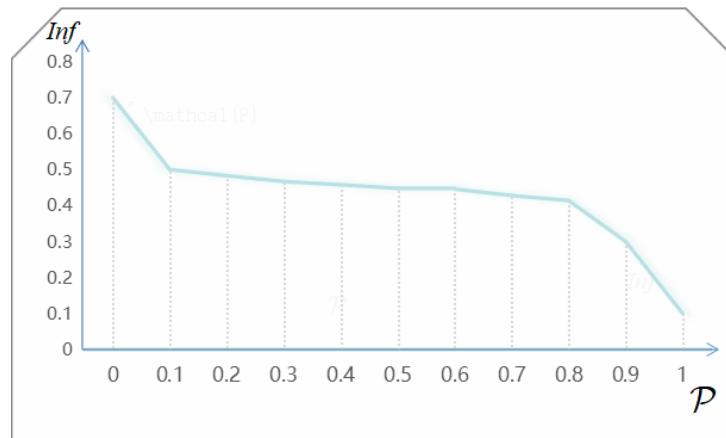


Figure 24: The main changes of Frank

From the figure we can find that there is no large change in the influence when  $\mathcal{P}$  is taken at [0.1,0.8]. Therefore, our model is insensitive to  $\mathcal{P}$  in most cases.

## 11 Strengths and Weaknesses

- Strengths:

1. It is innovative to construct influence models through networks.

2. The selection of the network parameters of the influence network is scientific and reasonable.

- Weaknessess:

1. Because of the large amount of data, our model cannot analyze each genre thoroughly.

2. The indicators of the popularization model cannot be quantified, resulting in a low persuasibility of the model.

3. As time changes, artists in one genre may create music in other genre styles, and we cannot consider this situation.

4. We did not consider the impact that a collaboration between artists would have on the work.

## A document to the ICM Society

To whom it may concern,

Music, as an essential component of cultural heritage, contains tremendous value. In order to measure musical influence and to explore the evolutionary and revolutionary trends in examining artists and genres, we build a network model.

Through our network model, we can identify the change-makers in the genre, such as the artist Joan Baez, who pushed the Folk genre to merge with the Jazz genre, and the artist The Beatles, who influenced not only the genre, but all other genres.

Unfortunately, however, due to the limited amount of data available, we only know the main genre of an artist and some of his works, while in reality, an artist does not belong to only one genre, but his works often dabble in multiple genres. And, in reality, there are fusions and divisions between genres. For example, the country genre can be subdivided into Bluegrass, Country Pop, Countrypolitan, etc. These factors also interfere with the accuracy of our model.

In addition, due to the lack of data, our model is unable to quantify the impact of environmental and major events on music. For example, we found that the Second World War around 1940s saw relatively large ups and downs in the characteristics of all music genres, but it is difficult to accurately reflect this factor in our model.

With this data, we can revise our network model, add influence factors in the network, and thus more accurately analyze the development and change of music genres.

Other factors will have an impact on the music, and likewise, the music will have an impact on other factors. For example, the rise of rock music in the United States in the 1950s elevated white Americans' identification with blacks, eased the problem of racial tension in the United States, and made the country more culturally inclusive.

Music as a part of culture, while having a profound impact on the rest of culture. Music can not only be enjoyed, but it can also permeate social life. And this is where the value of music lies.

Yours Sincerely,  
Team #2107542

## References

- [1] Jianshu Weng, Ee-Peng Lim, Jing Jiang. TwitterRank: Finding Topic-sensitive Influential Twitterers
- [2] [http://www.360doc.com/content/14/0505/10/15883912\\_374732455.shtml](http://www.360doc.com/content/14/0505/10/15883912_374732455.shtml).
- [3] Wang Lei. The Post-World War II American Art Patronage System and the Rise of American Contemporary Art [D]. China Academy of Art, 2015.
- [4] Zhao Dongliang. The relationship between rock music and social movements in the United States in the 1960s[J]. Literary and Educational Materials,2018(6):94-95.
- [5] Ding Zhu. The history of popular music (serialized five) [J]. Music Life,2011(07):28-29.
- [6] YUNJING DAI. A Research of Quantification of Pairwise Influence in Online Social Network.
- [7] Yu Xin. Tradition in innovation: a glimpse of western music in the twentieth century.2006sk079
- [8] Grout, A History of Western Music. People's Music Publishing House,1996, p. 720.
- [9] Shen Xuan. A Brief History of Western Music. Shanghai Music Publishing House,1999, p. 319.