正文

摘要

音乐随着社会变迁而演变。为了理解音乐在人类集体经验中所扮演的角色, Our team was asked to 审视艺术家和流派随社会变迁的进化和革命趋势。

in task 1, 我们通过建立有向连接图将影响者与追随者联系起来。图的方向从影响者指向追随者。

in task 2，

in task 3，

Contents

1. Introduction
2. Assumptions and Justification(uncountable)
3. Model developments
4. Sensitivity Analysis
5. Strengths and Weaknesses
6. References
7. Appendices(可有可无？)
8. Intro

1.1Background

**Music** is an [art](https://en.wikipedia.beta.wmflabs.org/w/index.php?title=Art&action=edit&redlink=1) form whose [medium](https://en.wikipedia.beta.wmflabs.org/w/index.php?title=Media_(arts)&action=edit&redlink=1) is [sound](https://en.wikipedia.beta.wmflabs.org/w/index.php?title=Sound&action=edit&redlink=1) and silence.[1]在人类千年历史长河中，音乐是文明发展不可或缺的一部分。当艺术家创作一段新的音乐时，有许多因素会影响他们，包括他们天生的创造力，当前的社会或政治事件，使用新乐器或工具，或其他个人经历。有时，音乐也会发生革命性的转变，提供新的声音或节奏，例如当一种新的流派出现，或现有流派的重新发明。通过分析歌曲网络及其音乐特征，我们可以开始捕捉音乐艺术家对彼此的影响。另外，我们也可以更好地了解音乐是如何随着社会的变迁而演变的。

**Music** is an [art](https://en.wikipedia.beta.wmflabs.org/w/index.php?title=Art&action=edit&redlink=1) form whose [medium](https://en.wikipedia.beta.wmflabs.org/w/index.php?title=Media_(arts)&action=edit&redlink=1) is [sound](https://en.wikipedia.beta.wmflabs.org/w/index.php?title=Sound&action=edit&redlink=1) and silence.[1]在人类千年历史长河中，音乐是文明发展不可或缺的一部分。当艺术家在进行音乐创作时，有许多因素会影响他们，包括他们的音乐创造力、个人经历、当下热点新闻和社会状态。有时，音乐也会发生革命性的转变，创造新的声音或节奏，引申出一种新的流派和潮流，或者对现有流派格局改变。通过分析不同时期歌曲网络及其音乐特征，我们可以开始捕捉音乐艺术家对彼此的影响，理解流派间和流派内部的异同点，从而可以更好地了解音乐是如何随着社会和文化的变迁而演变的。

Music is an art form whose medium is sound and silence.[1] In the thousand-year history of mankind, music is an indispensable part of the development of civilization.When artists are creating music, there are many factors that affect them, including their musical creativity, personal experience, current hot news and social status, and even a trivial voice.Sometimes, music also undergoes revolutionary changes, creating new sounds or rhythms, introducing a new genre and trend, or changing the pattern of existing genres.By analyzing song networks and their musical characteristics in different periods, we can begin to capture the influence of music artists on each other, understand the similarities and differences between and within genres, so as to better understand how music has evolved with social and cultural changes of.

1.2Problem restatement

1. a

讨论区：

构建音乐家直接影响网络（邻接表）

影响力递归算法

递归算法修正：新生代人才修正、回路断开修正（同年代人才）

1）各个流派内有哪些音乐大师？

TOP5（需要名字）

The Beatles

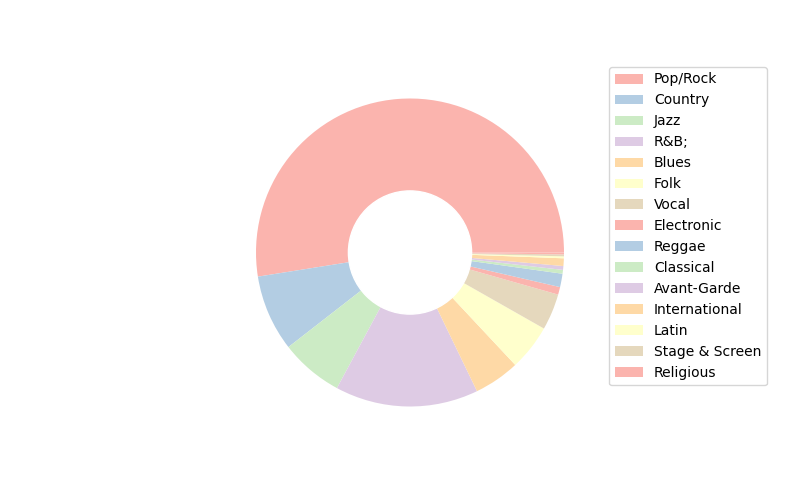
Bob Dylan

Hank Williams

The Rolling Stones

David Bowie

TOP500（只需要各个派别的人数）（一个图，可以是柱状）



~~3）各个派别的巅峰时刻（新生代人才最多，意思就是开始时期在这个年代的人最多）~~

2、相似性度量：

先对14个音乐特征进行标准化处理以消除量纲不同产生的影响，随后使用修正的余弦分析得到余弦相似度，进行归一化至0~1即为相似性（运用为第一题中的a，a从此成为变量）

归一化和标准化的区别：归一化是将样本的特征值转换到同一量纲下把数据映射到[0,1]或者[-1, 1]区间内，仅由变量的极值决定，因区间放缩法是归一化的一种。标准化是依照特征矩阵的列处理数据，其通过求z-score的方法，转换为标准正态分布，和整体样本分布相关，每个样本点都能对标准化产生影响。它们的相同点在于都能取消由于量纲不同引起的误差；都是一种线性变换，都是对向量X按照比例压缩再进行平移。

将同一流派的总平均相似性与全部音乐家总平均相似性作比较（直接做平均）（取Pop/Rock为例）

总体平均相似性：0.49716904539296675

Pop/Rock平均相似性：0.5350414676033829

county平均相似性：0.5866020737544656

Jaz平均相似性：0.6789162860515108

Folk平均相似性：0.7064500139497163

R&B;平均相似性：0.5495801087428339

Blues平均相似性：0.6001108299186905

Vocal平均相似性：0.7425133425349111

计算量过大的备选方案：（分层抽样，按比例，抽取流派内相似性做平均值，抽取流派内与流派外人员之间相似性做平均值）

3、

流派之间的平均相似性作比较：

对每个流派中所有人的相似系数取平均，然后在流派之间作比较得出流派内部相似性的大小关系；

对每个流派中所有人的影响力求和，然后再流派之间作比较得出流派影响力的差别；

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Pop/Rock | Country | Jazz | R&B; | Blues | Folk | Vocal | Electronic | Reggae | Classical | Avant-Garde | International | Latin | Stage & Screen | Religious | Easy Listening | Comedy/Spoken | New Age | Children's | Unknown |
| 平均相似度 | 0.535041 | 0.586602 | 0.678916 | 0.54958 | 0.600111 | 0.70645 | 0.742513 | 0.582089 | 0.673568 | 0.803377 | 0.62834 | 0.550586 | 0.558078 | 0.747444 | 0.557076 | 0.759792 | 0.732481 | 0.782832 | 0.658284 | 0.570205 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Pop/Rock | Country | Jazz | R&B; | Blues | Folk | Vocal | Electronic | Reggae | Classical | Avant-Garde | International | Latin | Stage & Screen | Religious | Easy Listening | Comedy/Spoken | New Age | Children's | Unknown |
| 总影响力 | 35433.96 | 5196.968 | 4238.106 | 8742.44 | 2291.984 | 2207.47 | 2441.471 | 933.946 | 1182.07 | 208.8534 | 149.1958 | 498.9911 | 808.7986 | 273.1309 | 347.2239 | 138.6793 | 205.1443 | 121.2226 | 3 | 2.559976 |

流派的区别：按照流派分类数据取平均的音乐特征，将不同流派的平均音乐特征作比较，得出流派的区别。

流派间的关联：按照流派分类数据取平均的音乐特征，将不同流派的平均音乐特征作相似度分析，得出哪些流派之间有所关联（可以结合史实进行分析，挑选出有希望相似的流派进行比对）

流派随着时间的变化：使用artist+influence数据集算出各流派每年（代）的平均音乐特征，将其与该年平均音乐特征做相似性分析，重点观察，分析哪些音乐特征在其中起到代表性的作用。

4、

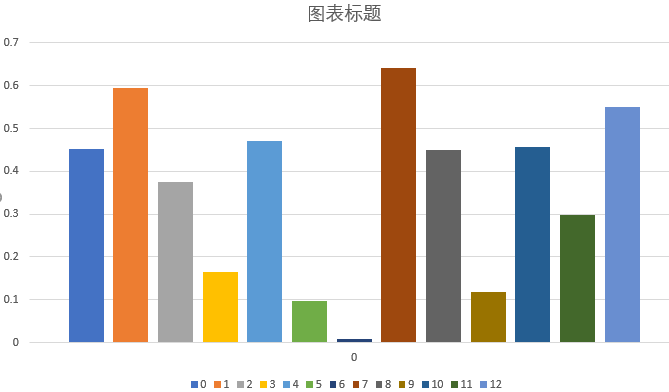
是否表明所识别的影响者实际上影响了各自的艺术家：

求取所有音乐家之间相似度的平均值，求取直接影响音乐家之间相似度的平均值，将二者作比较；

总相似度平均 0.49716904539296675

网络边上相似度平均 0.6635930285869684

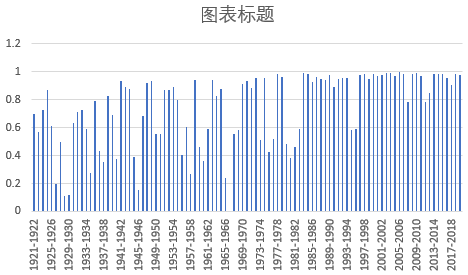
使用皮尔森相关系数，将13个音乐特征分别进行皮尔森比较：将影响者的**某一**音乐特征定为X，而被影响者的**对应**音乐特征定为Y探究这两个变量之间的皮尔森相关性（范围为全部直接影响关系的集合）；做成柱状图来进行比较；

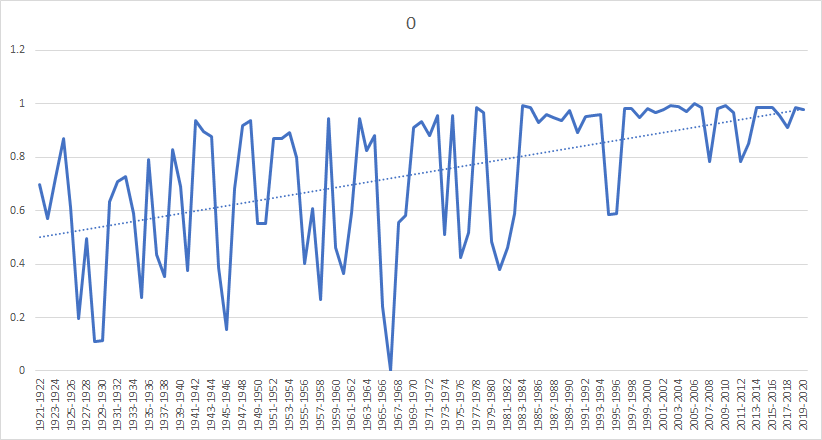


如若全部很像，则它们在影响特定艺术家的音乐方面都有相似的作用

5、

使用year数据集进行逐年的相似性变化的分析，然后对于**重点变化年份**（相似度小）专门进行音乐特征的分析；



****

**1920、1930的音乐家（比对年份为1929）**

艺术家影响力top10

Hank Williams 297.4129563965944 0.0

Woody Guthrie 164.257270916614 0.6938977632319187

Billie Holiday 159.77807700279797 0.6470150554288073

Charlie Parker 129.55722362486398 0.6365481672132528

Howlin' Wolf 128.77688856189425 0.030710959857485554

Nat King Cole 115.55526029177436 0.26254095531472327

Roy Acuff 114.7518054962189 0.48747475655618683

Ella Fitzgerald 111.26396192542714 0.2995538002354844

Pete Seeger 104.67414311953024 1.0

Frank Sinatra 100.27965230552473 0.26829464230415817

**1940的音乐家（1945）**

艺术家影响力top10

Miles Davis 209.70359898269334 0.5259938236552736

Ray Charles 174.78112127978864 1.0

Muddy Waters 169.25386182605288 0.8357802238265349

John Coltrane 149.5183453594907 0.6450407264768107

Bo Diddley 110.27995878747497 0.6767301019607963

John Lee Hooker 99.17312922182342 0.7066826248434311

Percy Mayfield 95.50675348294578 0.8952758785482671

Fats Domino 92.50955264978893 0.0

Bill Monroe 92.21531699341747 0.5901719572322569

Thelonious Monk 88.17061739754786 0.9032470416900962

**1960的音乐家（1967）**

艺术家影响力top10

The Beatles ,693.2233035449233 ,0.9251815952934412

Bob Dylan 445.13402940029334 0.6240716732926987

The Rolling Stones 371.14671532273724 0.8877948506224351

David Bowie 316.2759435748734 0.9157781204926478

Led Zeppelin 279.40931124951885 0.5536449613597928

The Kinks 256.324702817549 0.3309870882272297

Jimi Hendrix 237.8146255282818 0.026814559205774167

The Velvet Underground 236.5774666285356 1.0

Black Sabbath 215.7687956945067 0.0

Pink Floyd 210.9936885530381 0.33477613395609834

**1970、1980的音乐家（1981）**

艺术家影响力top10

Sex Pistols 224.98059892745798 0.7236661727209565

The Clash 204.50777860051926 0.18243745604967984

Ramones 180.59499730106953 0.7443778881016243

Brian Eno 173.5140521152474 0.4245132266963127

Roxy Music 164.5639895359767 0.0

Nirvana 145.66277639742393 0.15974380985204636

Kraftwerk 141.84364244928403 0.27740980317708913

Black Flag 139.2066796086238 0.9840152469916988

Kiss 138.81651509508174 0.2583704889136964

New York Dolls 133.3575636998304 1.0

**就这些了，总共四个重点年份**

对于重点年份寻找附近的影响力较大的音乐家（得出一个集合），使用该年份的音乐特征和这些音乐家进行相似度比较，其中相似度较大的为革命者

6、

分析某一种流派（pop&rock）：

使用full数据集，将每年份POP&ROCK的(每年平均)音乐特征计算出来，以10年为区间比较十年内音乐特征变化的权重（数据的离散程度，使用**熵权法**），最后整合得到能够揭示动态影响者的音乐特征。

重点关注这几个音乐特征，并以此为基础回头分析这个流派的音乐特征随年份变化的特点。（使用第一小问中得到的数据）

7、

音乐在某一时间或某一条件下的文化影响：

在构建好的影响力网络模型中寻找对应年代影响力名列前茅的艺术家

分析他们的流派、音乐风格，结合历史资料就可以得到这一时间下音乐对文化的主要影响。（可以不用举例子，只要说明我们可以使用这样的方法就可以）

识别社会、政治或技术变革的影响：

通过第六问得到的能够揭示动态影响者的音乐特征，将其作为重点分析当前年代音乐特征的变化曲线，结合音乐知识最终识别出外部影响。（必须举出一个例子）

477787！！！！！！！

Music evolves with social changes. In order to understand the role of music in human collective experience, Our team was asked to examine the evolutionary and revolutionary trends of artists and genres following social changes.

For problem 1 and 2, we first use the adjacency list to construct an artist's music influence network. In the network, the relationship between influencers and followers is established in the way of path weighting, and the "music influence" parameter of each artist is obtained through recursive traversal and weighted ratio summation, that is, the "new generation talent influence". In order to determine the specific value of the weighting of the path, we use principal component analysis and cosine similarity algorithm, supplemented by appropriate data processing, to obtain the similarity relationship between each two artists, and use this as the weighting parameter. On this basis, we came to the conclusion that artists of agreed genres are more similar than artists of different genres.

For question three, we use the similarity relationship and the "new generation talent influence" to get the similarity and influence relationship between and within the genres. For the difference of genres, we averaged the data by genre, compared the characteristics of different genres, and found (what difference). On this basis, we analyzed the average data of genres for similarity, and verified the genre relationship with the music history at that time. Finally, we find the years with large changes in similarity for further analysis (what music) to play a representative role.

For question four, we compare the average similarity between every two artists and the average direct impact similarity to get the influence of fan creation. Further, we calculated the Pearson correlation coefficients of each music feature and compared them, and finally made a histogram to show their relationship, and found that it is (how, is it affected by a certain department or both).

For question 5, we calculate the similarity year by year, analyze the music characteristics of the key change years (), and find (). Then, we look for musicians with great musical influence in the key change years, and analyze the similarity between the musicians and the years, and the revolutionaries with greater similarity, for example (who).

摘要

音乐随着社会变迁而演变。为了理解音乐在人类集体经验中所扮演的角色, Our team was asked to 审视艺术家和流派随社会变迁的进化和革命趋势。

对于问题一和问题二，我们首先利用邻接表构造艺术家音乐影响网络。在网络中以有向路径加权方式建立影响者与追随者的关系，并通过递归遍历、加权比例求和得出每一艺术家的”音乐影响“参数，即“新生代人才影响力”。该参数揭示了艺术家对新人才吸引力和对音乐发展的贡献。为了确定该路径加权具体数值，我们利用修正余弦相似度算法，辅以适当的数据处理作为相似度度量，从而得出每两位艺术家间的相似度关系，并以此作为加权参数。在此基础上，我们得出了同一类型艺术家比不同类型艺术家更相似的结论。

对于问题三，我们分别对流派间和流派内部进行分析。我们求出了各流派的平均音乐特征，并对不同流派平均特征进行相似度计算，发现Avant-Garde and New Age 相似度最高，pop/rock和international之间相似度最低。同时我们通过计算流派内相似度得知classic内部艺术家相似度最高，pop/rock内部艺术家相似度最低。此外，我们通过流派内音乐影响求和，发现pop/rock流派的影响力最大，并且通过各流派在各年代音乐特征分析发现pop/rock流派从1980s至今都是最流行的流派。

对于问题四，我们将每两艺术家间平均相似度和直接影响平均相似度比较，得出粉丝创作受影响者影响的结论。进一步的，我们计算各音乐特征的皮尔森相关系数并进行比较，发现影响者的energy和acousticness音乐特征最具“感染力”。

Music evolves with social changes.In order to understand the role of music in human collective experience, Our team was asked to examine the evolutionary and revolutionary trends of artists and genres following social changes.

For questions one and two, we first use the adjacency list to construct an artist music influence network.In the network, the relationship between influencers and followers is established in a directional path weighting method, and the "music influence" parameter of each artist is obtained through recursive traversal and the sum of weighted proportions, that is, the "new generation talent influence".This parameter reveals the artist's attraction to new talents and contribution to the development of music.In order to determine the specific value of the weight of the path, we use the modified cosine similarity algorithm, supplemented by appropriate data processing as the similarity measure, so as to obtain the similarity relationship between each two artists, and use this as the weighting parameter.On this basis, we came to the conclusion that artists of the same type are more similar than artists of different types.

For question three, we separately analyze the inter-genre and the inner-genre.We calculated the average music characteristics of each genre, and calculated the similarity of the average characteristics of different genres, and found that Avant-Garde and New Age had the highest similarity, and pop/rock and international had the lowest similarity.At the same time, by calculating the similarity within the genre, we know that the artist within the classic has the highest similarity, and the artist within pop/rock is the lowest.In addition, we summed up the influence of music within the genre and found that the pop/rock genre has the greatest influence, and through the analysis of the music characteristics of each genre in each era, we found that the pop/rock genre has been the most popular genre since 1980s.

For question four, we compare the average similarity between every two artists with the average similarity of direct influence, and draw a conclusion about the influence of those affected by fan creation.Furthermore, we calculated the Pearson correlation coefficients of various music characteristics and compared them, and found that the energy and acousticness music characteristics of the influencer were the most "infectious".

From the figure, it can be seen that the years of 1929, 1945, 1967, and 1981 have the lowest similarities with the previous year's musical characteristics, indicating that the musical characteristics of some of these years have changed significantly and musical changes may have occurred.Next, in these years of change, we look for artists on the network who have a great influence on music and have a high degree of similarity with the music characteristics of the years of change.We can think that the artist finally found is the music revolutionist.Here we take 1967 as an example. We found that the beatles, the rolling stones and David Bowie were more influential and had a similarity of 0.9 with the music characteristics of that year. Therefore, we believe that the three of them were revolutionaries of the 1967 music revolution.

As can be seen from the figure, the similarity of music characteristics of 1929, 1945, 1967 and 1981 with those of the previous year is the lowest, indicating that some of these years have great changes in music characteristics and may have undergone musical changes.Next, we look for artists whose music has a strong influence on the network and whose music features are highly similar to those of the years of change.We can think of the artists that were eventually found as musical revolutionaries.Taking 1967 as an example, we found that The Beatles, The Rolling Stones and David Bowie had a great influence, and their similarity with that year's music features reached 0.9. Therefore, we believe that they were revolutionaries of the 1967 music revolution

对于问题一和问题二，我们利用邻接表构造艺术家音乐影响网络。通过有向路径加权方式建立影响者与追随者的关系。并通过递归遍历权值得出每一艺术家的”音乐影响“参数，即“新生代人才影响力”。该参数揭示了艺术家对新人才吸引力和对音乐发展的贡献。为了确定该路径加权具体数值，我们采用修正余弦相似度算法，得出每两位艺术家间的相似度关系，并以此作为图的加权参数。在此基础上，我们得出了同一类型艺术家比不同类型艺术家更相似的结论。

对于问题三，我们分别对流派间和流派内部进行分析。我们求出了各流派的平均音乐特征，并对不同流派平均特征进行相似度计算从而得出流派间音乐特征区别。同时我们通过计算流派内每两位艺术家相似度得出流派内艺术家相似排名。此外，我们通过流派内音乐影响求和，发现pop/rock流派的影响力最大，并且通过各流派音乐特征映射到年代分析，得到流行流派变化表。

对于问题四，我们将每两艺术家间平均相似度和直接影响平均相似度比较，得出粉丝创作受影响者影响的结论。进一步的，我们计算各音乐特征的皮尔森相关系数并进行比较，发现影响者的energy和acousticness音乐特征最具“感染力”。

对于问题五和问题六，逐年进行相似度计算，对相似度小的年份进行音乐特征分析。然后，在该年份寻找音乐影响大的音乐家，将音乐家与该年音乐特征进行相似度分析，以1967年为例，the beatles影响力最大并且与变化年份音乐特征相似度高，他是当时的音乐革命者。此外，我们利用熵权法计算十年区间内的音乐特征变化权值，权值越大则特征离散程度越大，最后发现pop/rock流派energy和loudness特征60年代后不断，说明pop/rock流派变得更活跃了。

对于问题七，我们结合引用的相关文献进行了音乐与文化的互相影响关系的分析。在音乐影响文化方面，检索高影响力艺术家的方法得到了使用。在音乐体现文化变化方面，将问题六得到的数据与相关文献进行灵活结合是找到Pop/Rock实例答案的关键。