



Individual Coursework Submission Form

Specialist Masters Programme

Surname: Lai	First Name: Hoang Duong
MSc in: MSc Business Analytics	Student ID number: 240034730
Module Code: SMM638	
Module Title: Network Analytics	
Lecturer: Dr. Simone Santoni	Submission Date: 10/12/2024
<p>Declaration:</p> <p>By submitting this work, I declare that this work is entirely my own except those parts duly identified and referenced in my submission. It complies with any specified word limits and the requirements and regulations detailed in the coursework instructions and any other relevant programme and module documentation. In submitting this work, I acknowledge that I have read and understood the regulations and code regarding academic misconduct, including that relating to plagiarism, as specified in the Programme Handbook. I also acknowledge that this work will be subject to a variety of checks for academic misconduct.</p> <p>We acknowledge that work submitted late without a granted extension will be subject to penalties, as outlined in the Programme Handbook. Penalties will be applied for a maximum of five days lateness, after which a mark of zero will be awarded.</p>	
Marker's Comments (if not being marked on-line):	

Deduction for Late Submission:

Final Mark:

 %

1. Workflow description:

The objective of this report is to examine the similarities between 84 music genres on the Deezer platform and to analyse how friendship networks influence these similarities.

The analysis begins with the construction of an incidence matrix that captures user behaviour, specifically indicating which users have liked which genres. From this matrix, a Genre-Genre network is derived, where each connection between two genres signifies that at least one user has liked both genres. The weight of each edge in the network is calculated using the following formula:

$$Weight(i, j) = \frac{U_i \cap U_j}{U_i \cup U_j}$$

Where:

- $Weight(i, j)$ = weight of the edge connecting genre i and genre j
- $U_i \cap U_j$ = Number of users like both i and j
- $U_i \cup U_j$ = Total number of users like both i and j

This measurement helps assess how similar the two genres are in terms of their shared listeners.

After obtaining the edge weight data, the next step involves plotting the distribution of edge weights to uncover potential patterns of interest. Additionally, the node degree and the average edge weight for each node are calculated and visualized to provide a more comprehensive understanding of nodes that act as hubs of similarity within the network. The average edge weight reflects the strength of a node's connections within the network, while a high node degree indicates extensive connectivity. When combined, these two measures can highlight key hubs that play a central role in shaping genre similarity across the network.

The Genre-Genre weighted network was analysed using Louvain's community detection algorithm to identify tightly knitted groups of music genres. Further analysis of the detected communities allowed for the refinement of the community structure, resulting in a comprehensive map of music genres on the Deezer platform.

Subsequently, the report examined the friendship network among Deezer users, applying Louvain's algorithm to identify user sub-communities. Key metrics for these sub-communities were extracted and analysed to characterize the central hubs and peripheral satellite communities. This analysis provided insights into how social ties among users influence the similarity and clustering of music genres within the network.

2. Results and Findings

Node Degree and Average Edge Weight of each genre tells us about how connected the genres are with other genres in the network, as well as how strong its ties are with other genres. By analysing the graph below, several key insights can be observed regarding the centrality and influence of specific genres within the network.

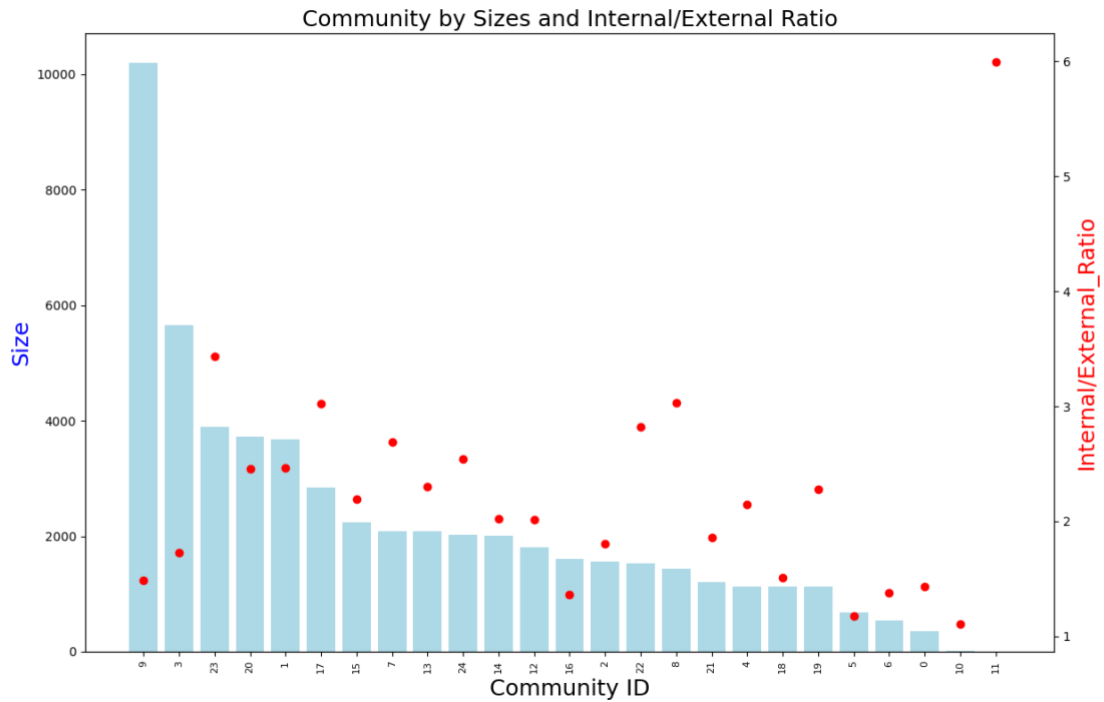


Figure 1. Node Degree and Average Edge Weight for Each Genre Node

The analysis of node degree and average edge weight reveals distinct patterns in the connectivity of music genres within the network. Over half of the genres have low connectivity, acting as peripheral nodes with limited similarity to other genres. These niche genres cater to specific listener communities with minimal overlap in preferences. In contrast, while many genres have high node degrees, only a few exhibit significantly higher average edge weights. These "hub genres" maintain strong, dense connections with other genres, serving as central nodes within the network. Their position allows them to bridge sub-communities and facilitate listener exploration across multiple genres. This dual structure — with influential hubs and isolated peripheral genres — offers key insights for optimizing recommendation systems and understanding user listening behavior on the Deezer platform.

In the process of creating the genre community map for Deezer, several interesting insights emerged.

Sunburst Chart of Global and Local Communities of Music Genres

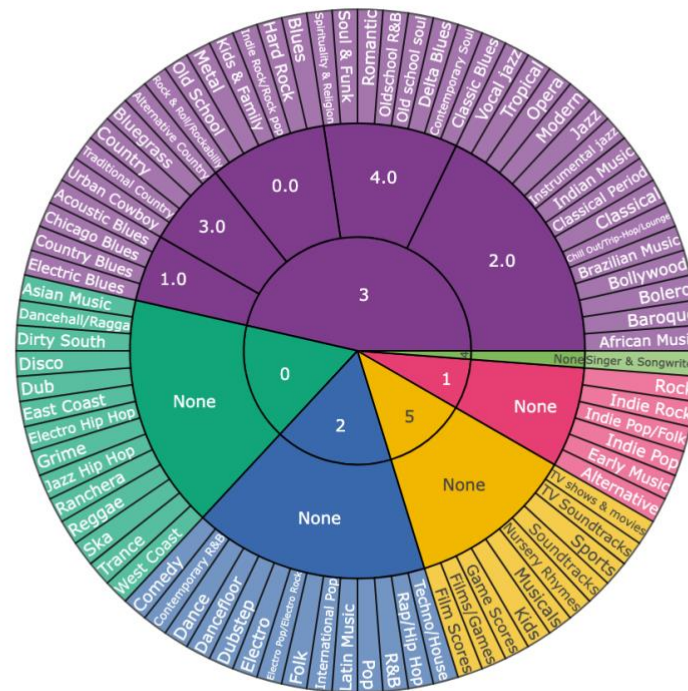


Figure 2. Deezer platform Music Genre Map

Louvain's community detection algorithm, which prioritizes modularity, revealed tightly connected genre clusters. This indicates that listeners of one genre are highly likely to engage with other genres within the same cluster, reflecting shared listener preferences.

Notably, Community 3, which consists of 39 genres, was further subdivided into smaller sub-communities to gain a deeper understanding of the relationships among genres within this large cluster. This approach allowed for a more granular view of genre connectivity, revealing nuanced patterns of listener behavior and genre similarity.

Further investigation of the metrics of user communities, we have found insights into the overall structure of the friendship network on Deezer

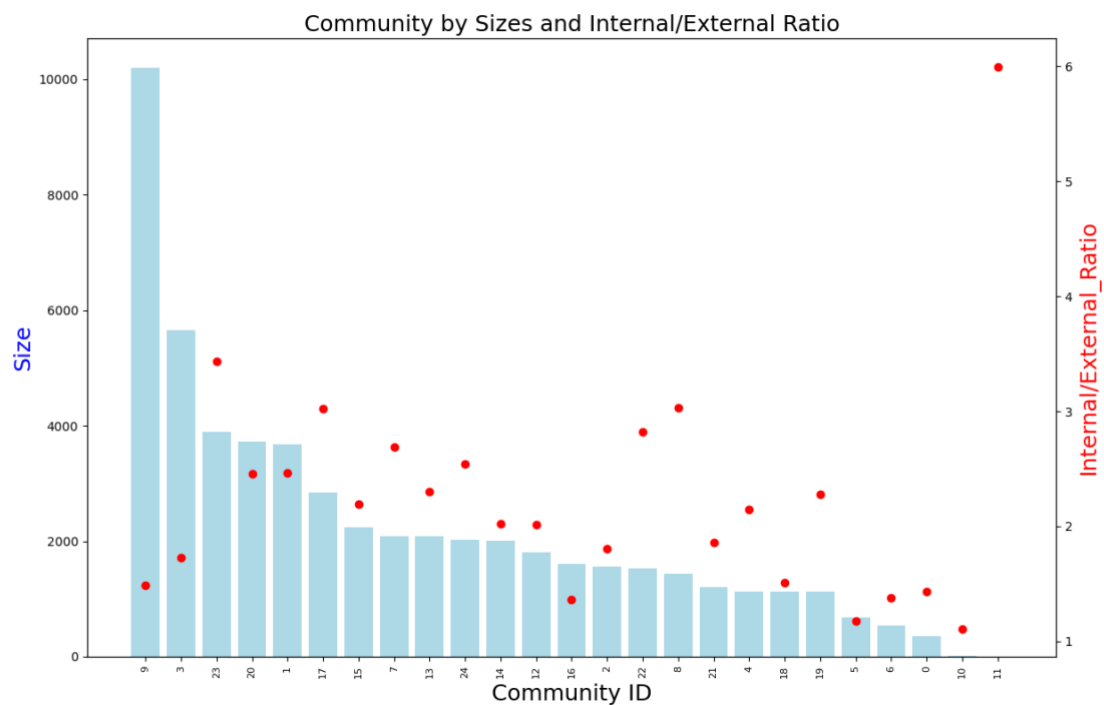


Figure 3. Friendship Community Size and I/E Ratio

Louvain’s algorithm identified 26 friendship communities within Deezer’s network, with key metrics collected for each. The largest sub-communities, containing over 10,000 nodes, are core network groups that act as central hubs within the broader network. Their size and connectivity suggest a significant influence on the overall network structure.

Further analysis of these hubs revealed a low internal-to-external connection ratio, indicating a balanced distribution of connections both within and outside the community. Combined with their large number of total connections, these hubs likely play a central role in shaping the flow of influence and genre similarity across the network.

3. Interpretation of the results:

The "Singer & Songwriter" genre emerged as a standalone community following Louvain's community detection process, raising several interesting questions. Further analysis revealed that this genre maintains connections with nearly every other genre in the network. However, the weights of these connections are relatively weak, preventing the Louvain algorithm from assigning it to a specific community. This phenomenon may be attributed to the inherent nature of the "Singer & Songwriter" genre, which often blends elements of "Pop," "Folk," "Country," "Indie," and even "Rock."

This insight has important implications for Sonic's strategy regarding genre classification. Positioning new releases under the "Singer & Songwriter" label could hinder the effectiveness of recommendation algorithms, as the genre's broad overlap with multiple communities may dilute its distinctiveness. To improve discoverability and enhance promotional efforts for artists, it may be more effective to classify new releases under more

specific genre categories that align more closely with listener preferences and community structures.

The clustering of genres into communities highlights shared listener bases, offering Sonic strategic insights for positioning new releases. Genres within the same community exhibit strong similarities, enabling new songs in one genre to be easily recommended to users who engage with other genres in the same cluster. This approach can enhance discoverability and maximize audience reach for Sonic's artists.

The similarity between music genres is determined by the weight of the ties connecting them. When a user engages with a new genre, its similarity with the existing genres in their listening portfolio increases. Social ties, particularly friendships, play a critical role in this process, as exposure to a friend's music preferences can prompt users to explore and adopt new genres, potentially reshaping their personal music tastes.

An analysis of the friendship network on Deezer reveals the presence of a dominant community that exerts considerable influence on genre similarity across the platform. The structural properties of this community enable even minor shifts, such as the introduction of a new genre, to significantly increase its similarity to popular genres, thereby altering the broader genre landscape. Further analysis of this key community could provide deeper insights into the dynamics of genre connectivity and the role of social influence in shaping music preferences.