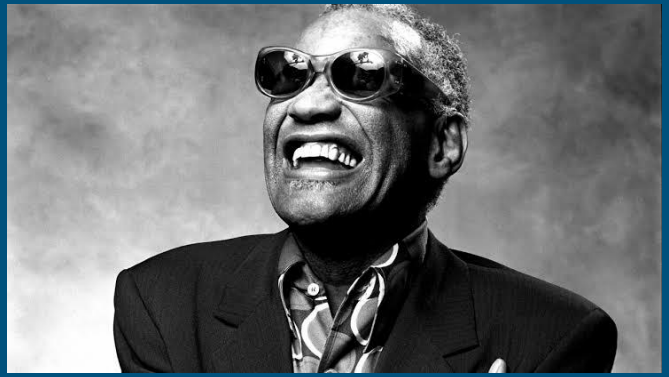
Chord Progression Network Analysis of Ray Charles Songs

Team 09: Gehan Velivitiya & Juanita Ramirez
CNT5805: Final Project Presentation
11/29/2023

Overview & Purpose

In this project, we constructed a chord progression network of 88 songs belonging to Ray Charles (https://www.e-chords.com/ray-charles).
The purpose of this project was to generate complex graphs for songs, identify patterns songs follow, and explore the structural features of Ray Charles's songs.



Basic Concepts

- **Note:** there are 12 unique notes in music denoted by letters, can be represented as numbers from 0 to 11.
- **Chord:** is composed by three or more notes played simultaneously, similar to a set where the number of elements is equal or larger than 3. In this project, one of the hypotheses is that the order of the notes within the chord was not taken into account.
- **Progression of chords:** It is a sequence of chords, can be represented as a vector of sets.

Research Questions

a. Nodes and link attacks could represent failures in the instrument or the person playing it. What is the robustness of the network?
   Can the chord progression network show why some mistakes in songs go unnoticed?
a. Does Ray Charles have a unique music style compared to other artists?
b. In the structure, why do more important links (larger weights) appear?
c. Does Ray Charles have a preference for chords used in his songs?
d. Are there any communities? Is there homophily?

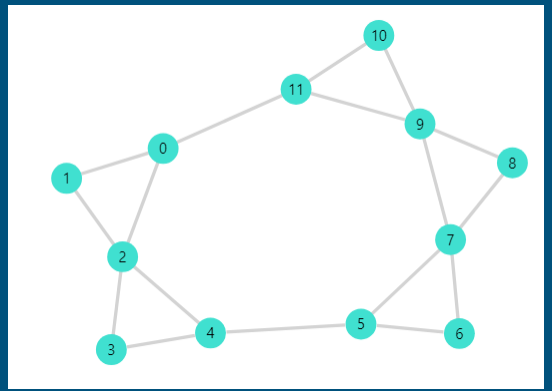
Data

- This is a directed and weighted network
- Nodes: 167
  Edges: 1246
- A node is a chord and a link is a transition between two chords.
- The weight represents in how many songs that transition occurred

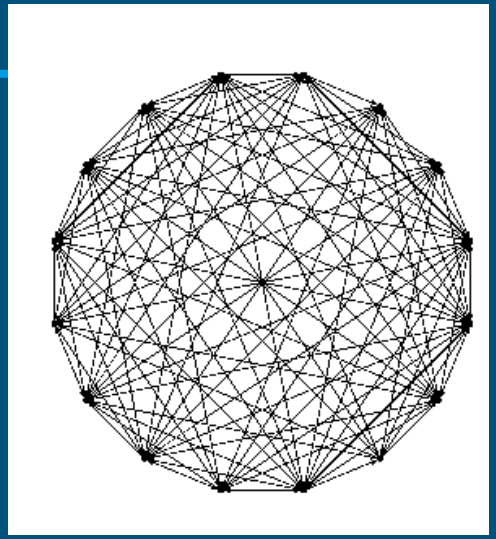
Hypothesis:

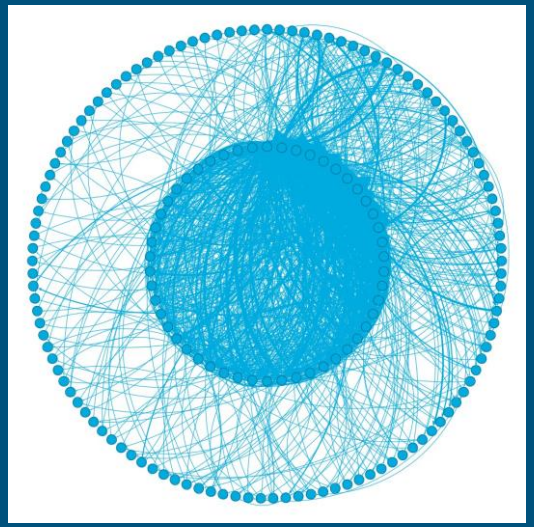
All chords were transposed to the major key “C” or its minor relative “Am” and only unique links were considered for each song.

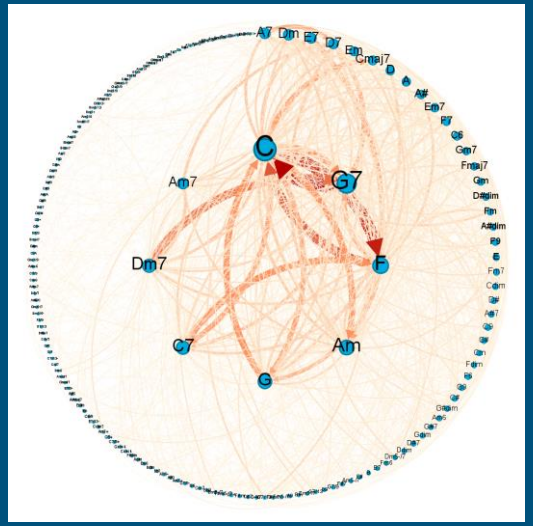
Source	Target
Am	Am7+
Am7+	Am7
Am7	Am6
Am6	Am7
Am7	Am7+
Am7+	Am
Am	E7/9-



Filters and Structural Features

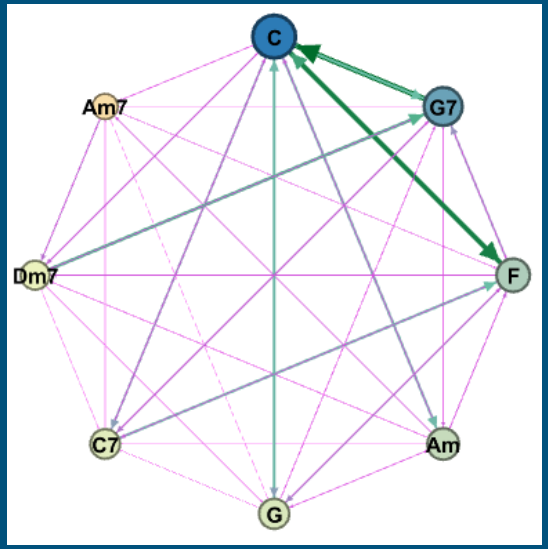


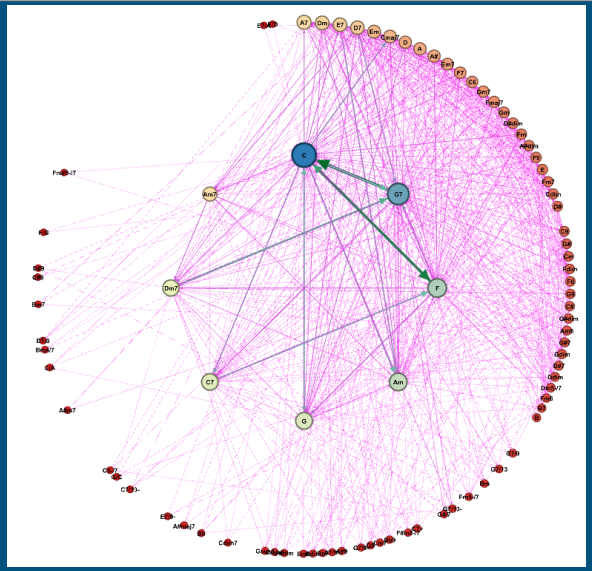




The filters based on the degree returned a subgraph with a density close to 1, this helped to the improvement of graph visualizations in a dual circular layout based on degree. In the network high degree nodes tend to connect with high degree nodes and have heavier weighted links.

We deduced that chords with a high degree are common chords among songs, and in the outer circle rare chords are observed.





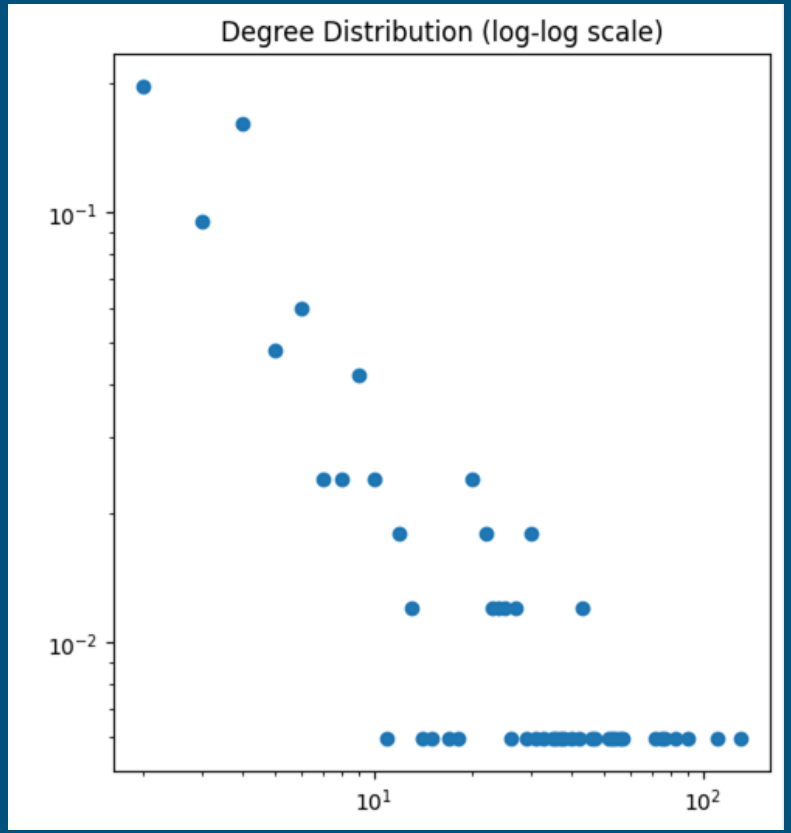
Filter based on the degree of nodes to isolate the inner circle of chord progression network. Density of the directed graph is 0.929 and the undirected graph is 1.

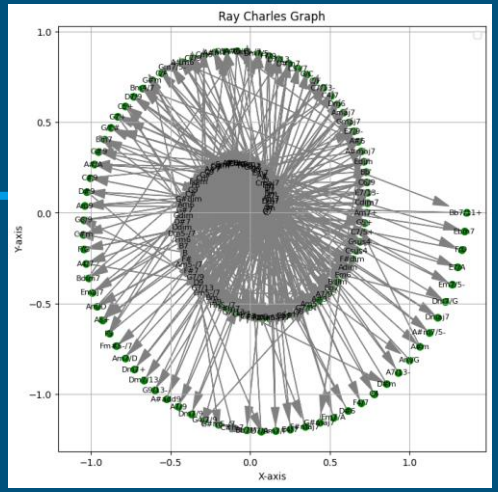
- C, G7, F and Am are hubs in the network.
- The links between C-F-G7-C show one of the most famous chord progressions in music.

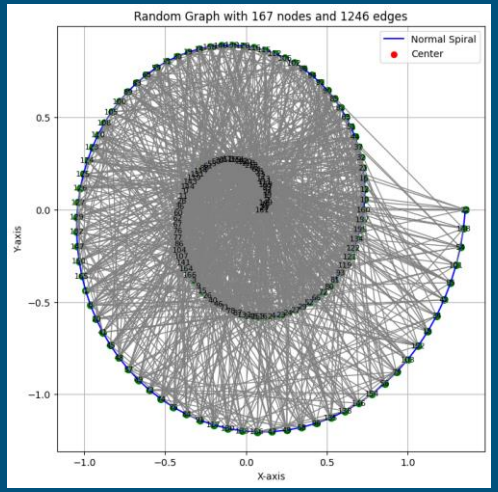
Ego network filter on chord “C”. 51.49% of the nodes are connected to it and 73.43% of edges. Some edges are ones with higher weights. This implies the importance of chord “C” in chord progressions.

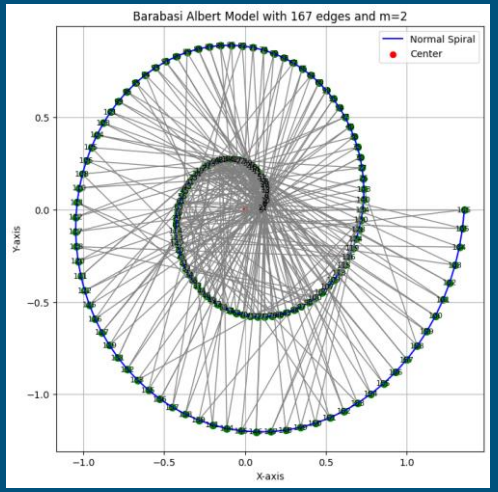
Statistics

- This is a scale free network.
- The density is 0.045 and the clustering is 0.486. While only 4.5% of possible connections appear on average for each node 48.6% of neighbors are connected between them.
- The average path length is 2.686 and the diameter is 6
- The average degree is 7.46 with min=2 and max=130





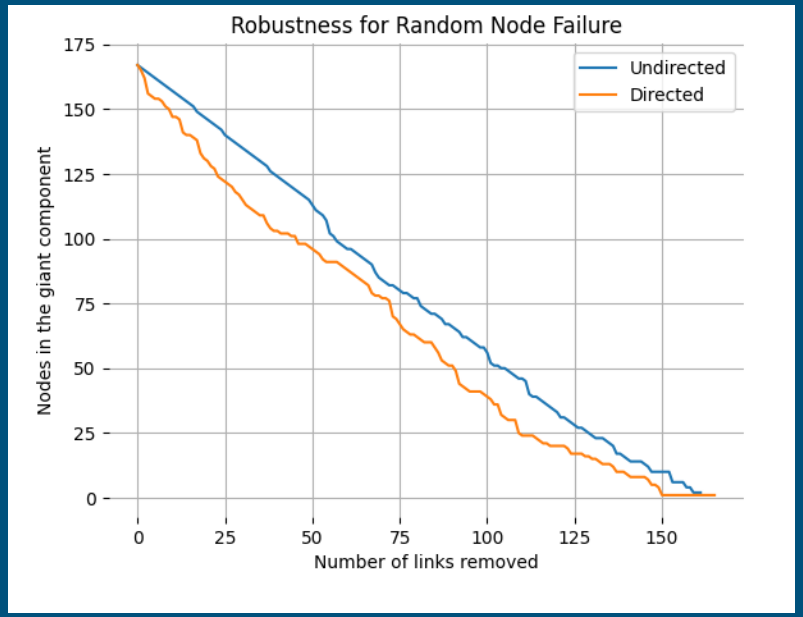


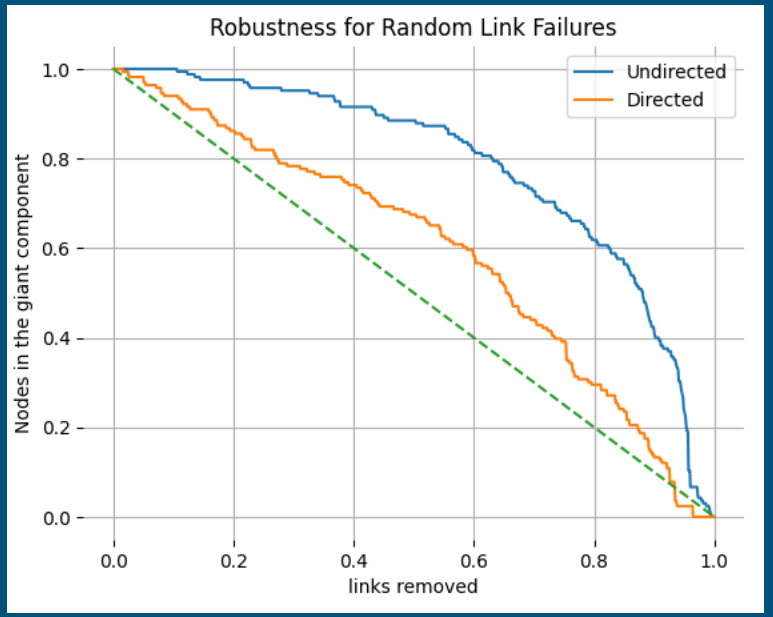


Graph	N and L	Density	Avg. Clustering Coefficient	Avg. Degree
Ray Charles	N= 167, E= 1246	0.0449	0.486	7.461
Random	N= 167, E= 1246	0.0898	0.0904	14.9221
Albert Barabasi Model	N-167, E=330	0.02380	0.0508	3.9520

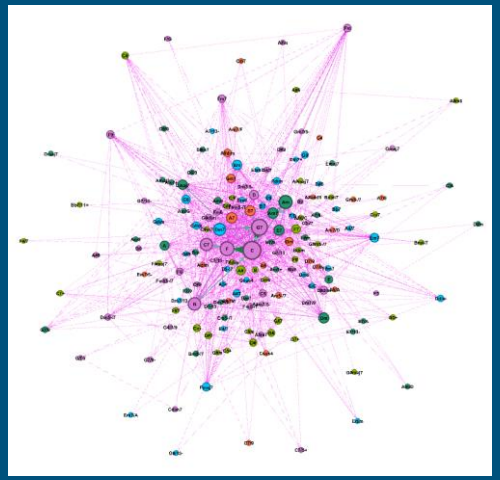
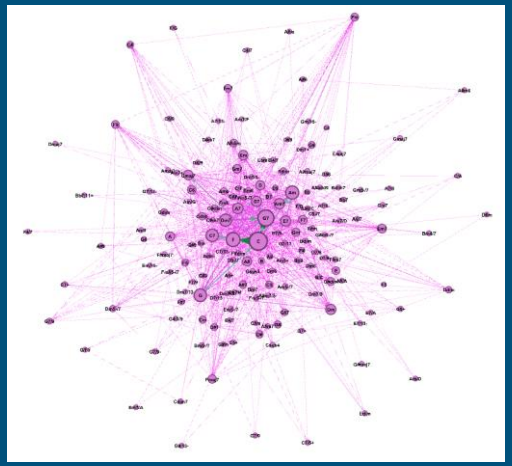
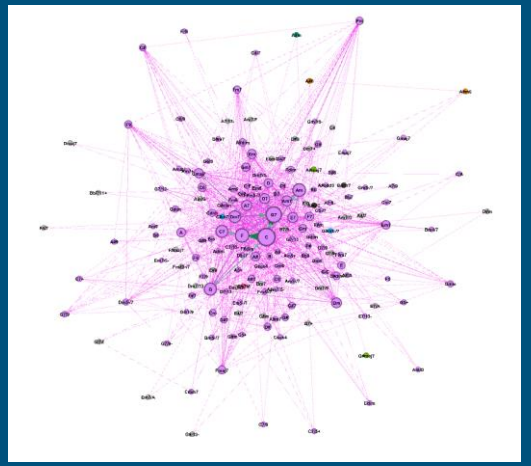
Robustness

In order to analyze failures in music the Robustness of the network was analyzed considering the network directed and undirected, the thresholds were 0.97 and 0.96 respectively for node failures. The directed network is less robust because when rare chords are attacked the flow of the progression is damaged.





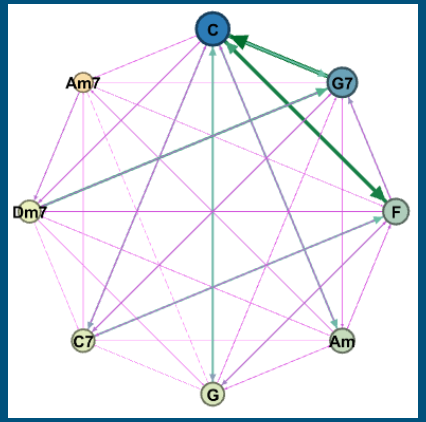
Communities

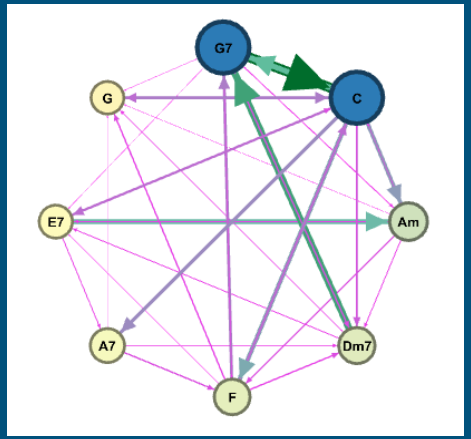
		
Louvain algorithm with resolution 1.0. 5 communities are identified	Louvain algorithm with resolution 2.0. 1 community is identified	37 clusters, however there is a single cluster with 76.05% of the nodes and the rest have 1.2% or less. Therefore can be considered as one giant cluster.

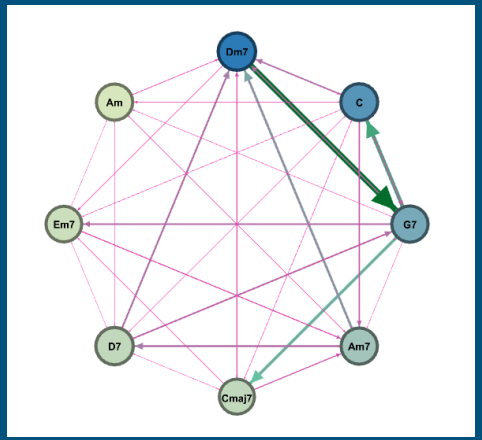
Comparisons Between Networks

- (Ray Charles graph edges ∩ General Jazz edges) / General Jazz edges = 0.182 (281 edges belong to the intersection).
- (Ray Charles Jazz graph edges ∩ General Jazz edges) / General Jazz edges = 0.111 (172 edges belong to the intersection) .
- (Ray Charles graph nodes ∩ General Jazz nodes)/ General Jazz nodes = 0.255 (61 nodes belong to the intersection).
- (Ray Charles Jazz graph nodes ∩ General Jazz nodes)/ General Jazz nodes = 0.213 (51 nodes belong to the intersection).

- The average jaccard similarity score between Ray Charles network and General Jazz Network was 0.149.
- The average jaccard similarity score between Ray Charles Jazz Genre network and General Jazz network was 0.135.







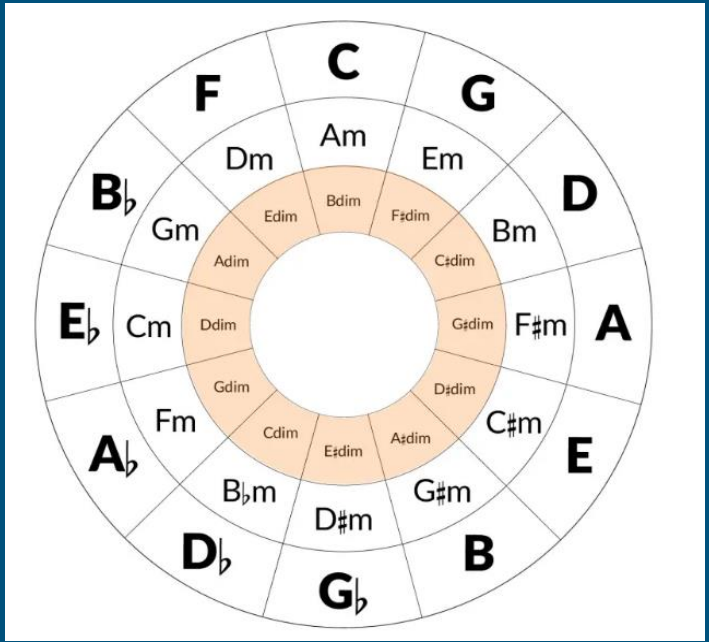
Ray Charles general
General Jazz

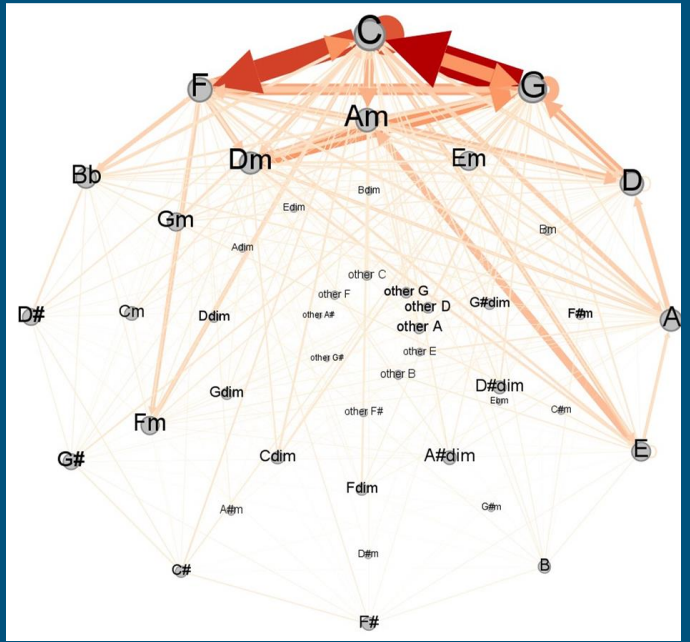
Ray Charles Jazz

Significant observation from these three inner circle graphs is that the majority of chords in the inner circle of the Ray Charles network are major chords whereas in the general Jazz they are minor chords.

Categorical Structure

The chords were categorized based on the three first notes of each chord. The categories were the 36 chords of the circle of fifths, heavier weights and more links appear when categories approach C.





Results

a. The high robustness of the graph explains why it is difficult to hear failures in music. The node failure is similar as when a musician plays a chord wrong, while link failures represent when a musician is playing a song and doesn't make a transition between two chords.
b. The small Jaccard similarity scores, and the low number of nodes and edges in the intersection, shows that the two networks are different. Also the comparison of inner circles of graphs shows significant differences. This confirms Ray Charles has unique style.
c. Since chords have been transposed to keys C and Am, it is not possible to say which exact chords are used commonly in his songs. Despite that, the Circle of Fifths analysis offers insights into a clear pattern for chord and chord progression choices based on the chosen key.
d. As it was observed with the Louvain Algorithm and the Girvan Newman Clustering, there is only one community in the network, this is because chords were transposed. The concept of similarity in chords is subjective.
e. Homophily in chords is difficult to define as there are many chords but chords are constructed in 12 notes, then different chords could share a high amount of notes.
   If we base homophily in distance, the order of the notes and the difference in number of notes is an impediment to define an accurate distance between chords.

Contemplations

- We faced challenges in data preparation including, chord extraction from lyric files, transposition of chords and standardizing duplicate chords names.
- Applying music theory concepts in chord progression network and drawing conclusions involved tedious effort.
- Some additional data and ways that would help extend our study are:
  - More information about chords used.
  - Time interval information for chord progressions.
  - A method to measure distance between songs.
  - More songs and a balanced dataset for the genres.