

The complexity of measuring music complexity

SFI CSSS 2023

Members: Cheyenne, Virginia, Golnar, Joris, Annie, Yuanmo,
Faith



How does the complexity of an artist's music evolve over time?



Entropy can be measured using image compression

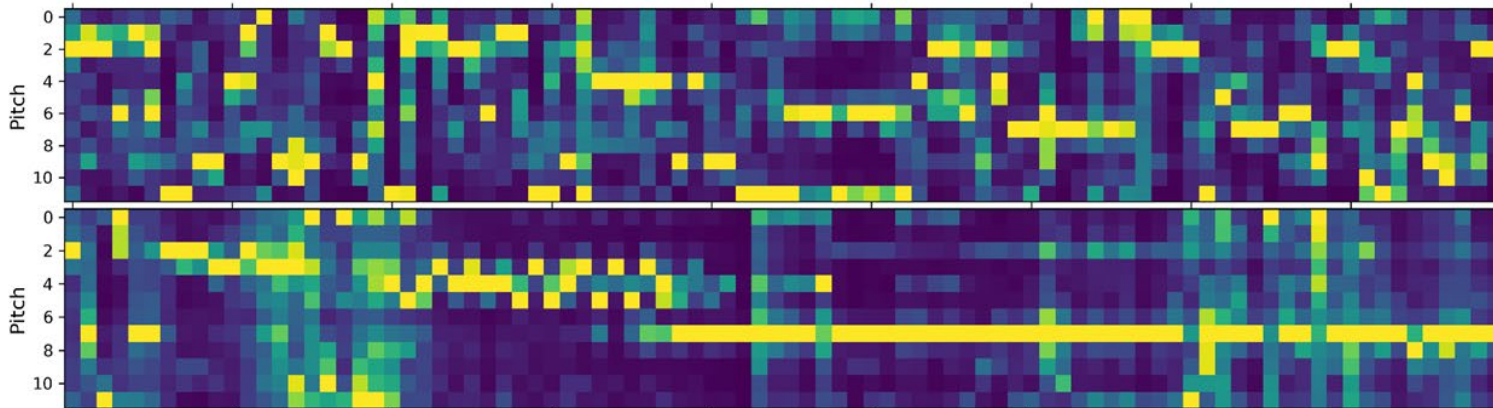
$$\text{Computable Information Density} = \frac{\text{Length}_{\text{compressed}}}{\text{Length}_{\text{raw}}}$$

Where computable information density is a proxy for **entropy**



We convert songs to images Using chromograms from the Spotify API

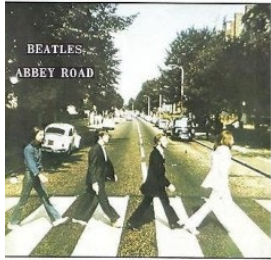
Her Majesty



It's all too much

Time (non-uniform spacing)

Her Majesty
by JOHN LINDON and PAUL McCARTNEY

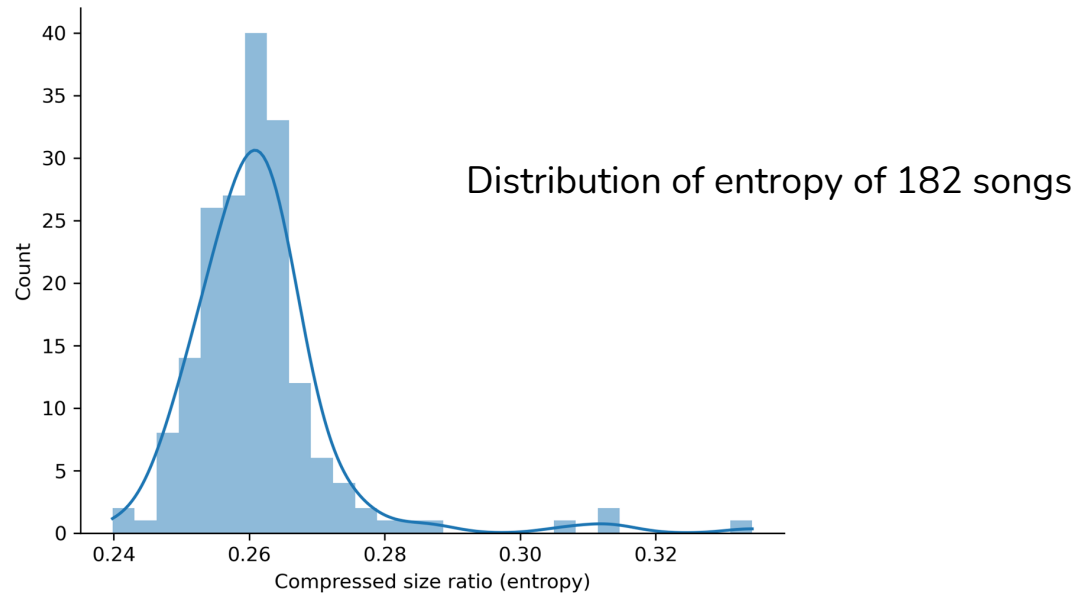


it's all too much
by george barvicka

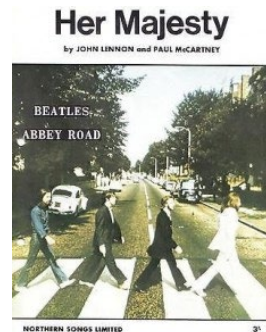
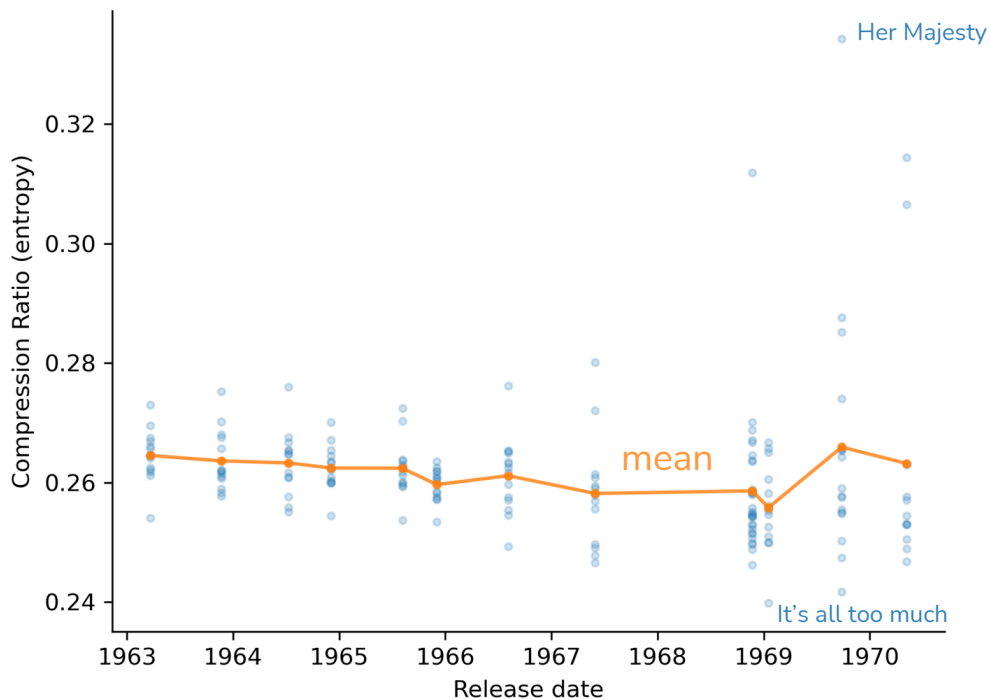




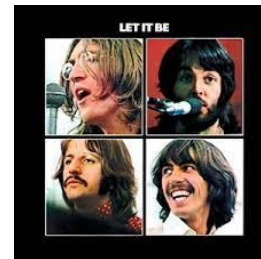
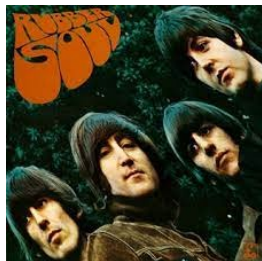
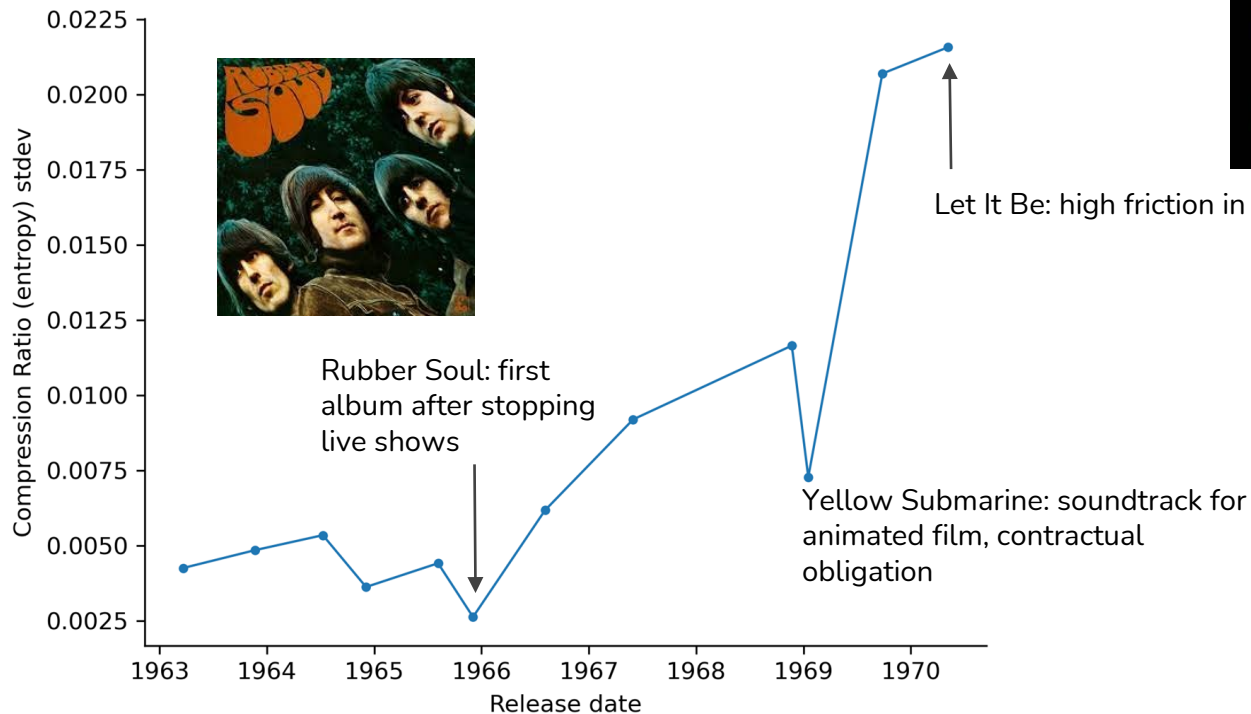
Compression ratio (entropy) of all The Beatles' songs is not uniformly distributed



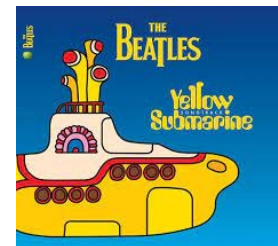
The mean entropy per album stays relatively constant over time



The spread in entropy across songs in each album appears to increase over time



Let It Be: high friction in the group





Entropy can also be used to measure similarity across songs

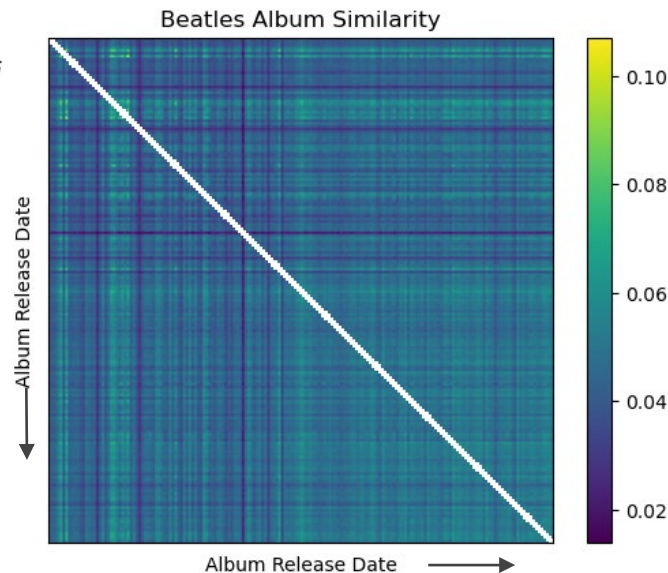
$$\text{Similarity}_{ij} = 1 - \frac{\text{Computable Information Density}_{ij}}{\text{Computable Information Density}_i + \text{Computable Information Density}_j}$$

Where the Computable Information $Density_{ij}$ is the entropy measures on the concatenation of song i and song j

0.0 = completely dissimilar

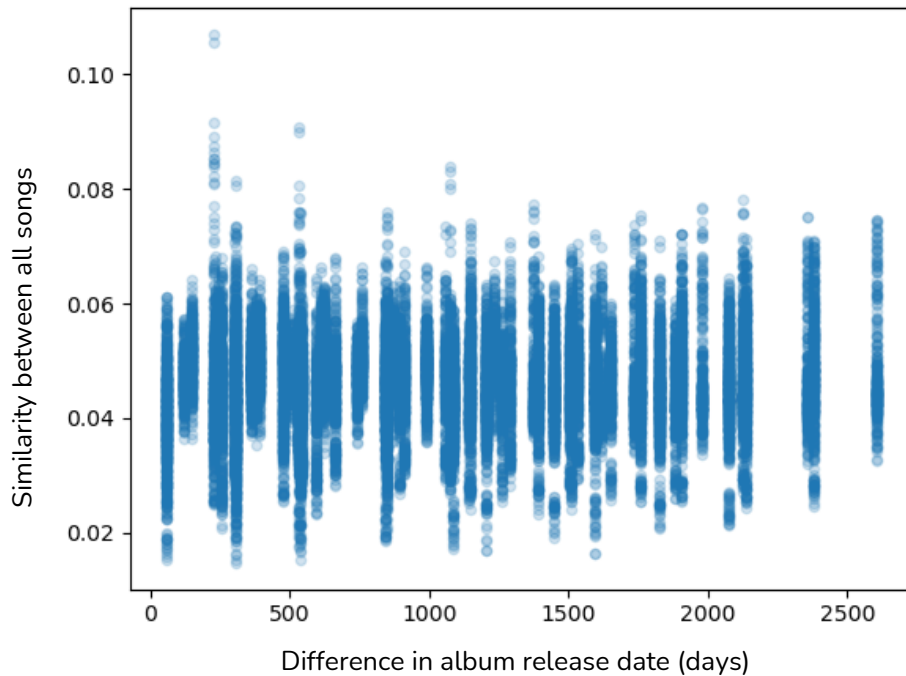
0.5 = same song

More idiosyncratic songs early on





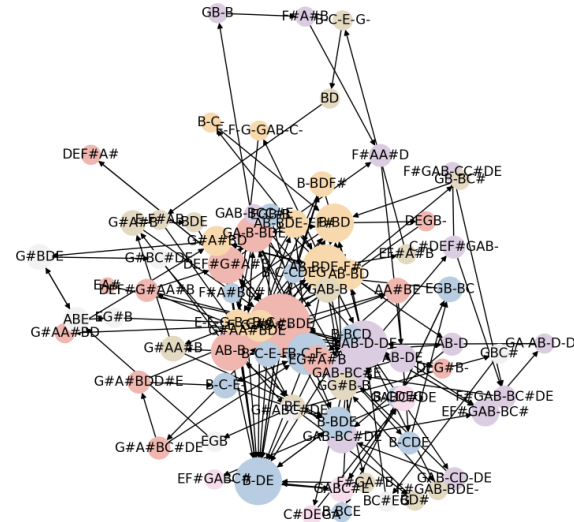
Songs released in close succession do not show a higher similarity score



Entropy can also be measured from song Pitch ...

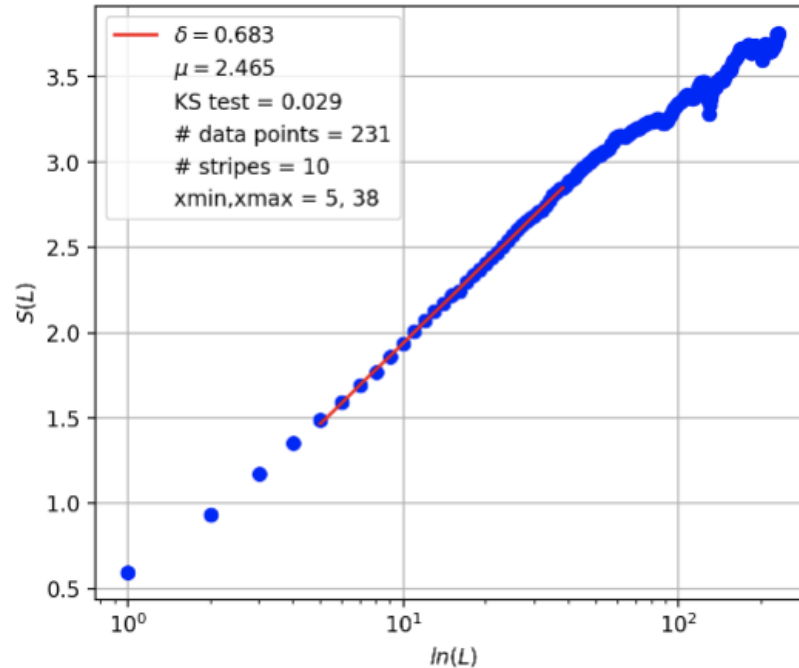
- We use the *musicnetwrk* library to study the complexity of the harmony
- We can represent the song chords as networks

Track 1: Please Please me



Entropy can also be measured from song Pitch ...

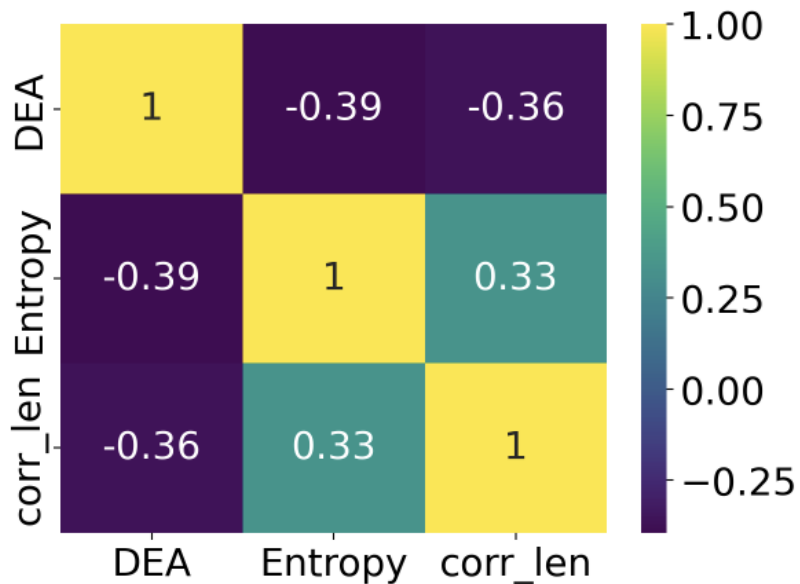
- We use the *musicnetwrk* library to study the complexity of the harmony
- We can represent the song chords as networks
- We obtain the complexity profile of the Pitch over time, and from that the **Diffusion Entropy Analysis (DEA)**





... But it measures different facets of music complexity

- Probably because this analysis does not consider percussion instruments.
- It measures only complexity of Pitch (only for the first album)





Measures of lyrics complexity

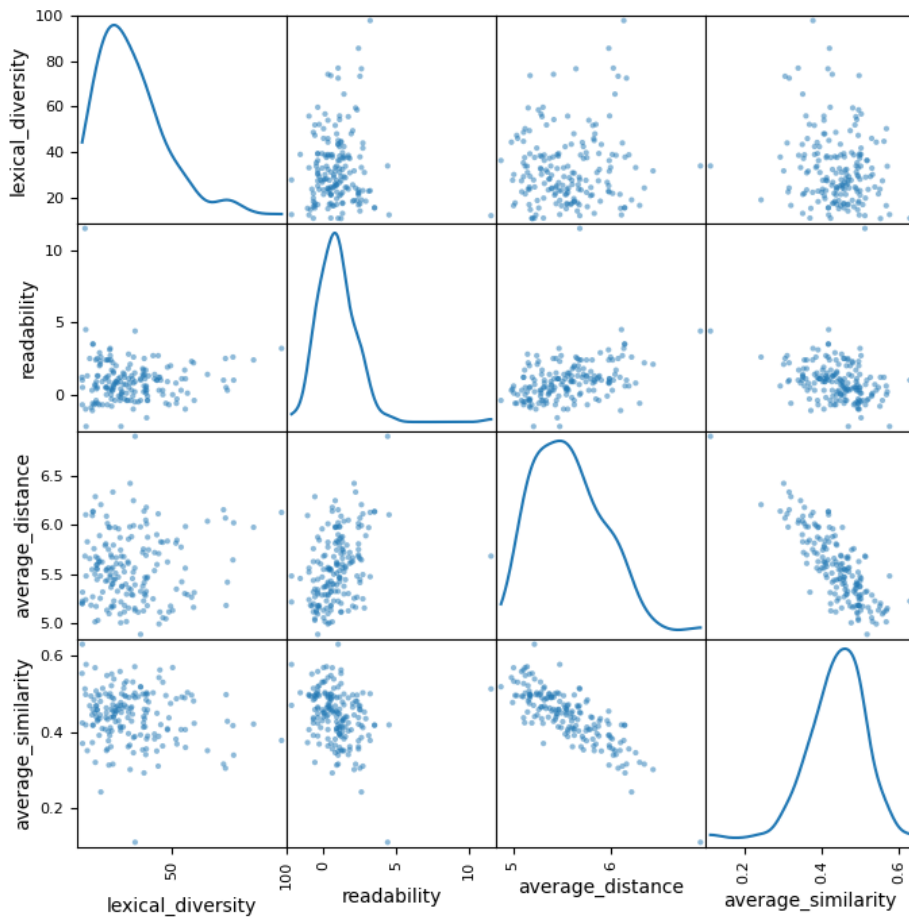
- Lexical Diversity - Measure of Textual Lexical Diversity (MTLD): the mean length of word strings that maintain a criterion level of lexical variation¹
- Readability - Flesch-Kincaid scale (what level of the US educational grade level is required to understand the text)²
- Word embedding³ (pre-trained model: glove, Wikipedia 2014 corpus)
 - average semantic distance between all words in text
 - average cosine similarity between all words in text

1. McCarthy, P. M.; Jarvis, S. MTLD, Vocd-D, and HD-D: A Validation Study of Sophisticated Approaches to Lexical Diversity Assessment. *Behavior Research Methods* **2010**, *42* (2), 381–392. <https://doi.org/10.3758/BRM.42.2.381>.(1)
2. Kincaid, J.; Fishburne, R.; Rogers, R.; Chissom, B. Derivation Of New Readability Formulas (Automated Readability Index, Fog Count And Flesch Reading Ease Formula) For Navy Enlisted Personnel. *Institute for Simulation and Training* **1975**.
3. Tandon, A.; Fortunato, S.; Evans, J. Breadth Not Specialization Characterizes How Modern Science Advances. *Working Paper* **2023**

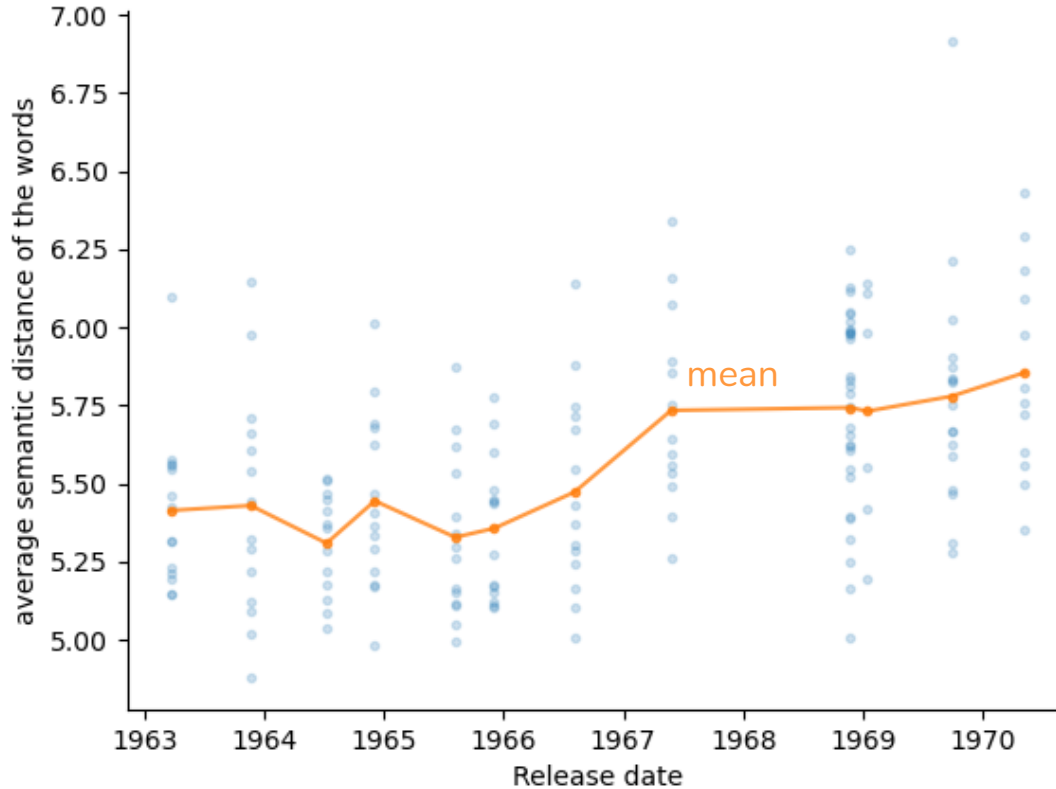


Lyrics of 175 Beatles songs

- Similar distributions
- Correlated apart from lexical diversity

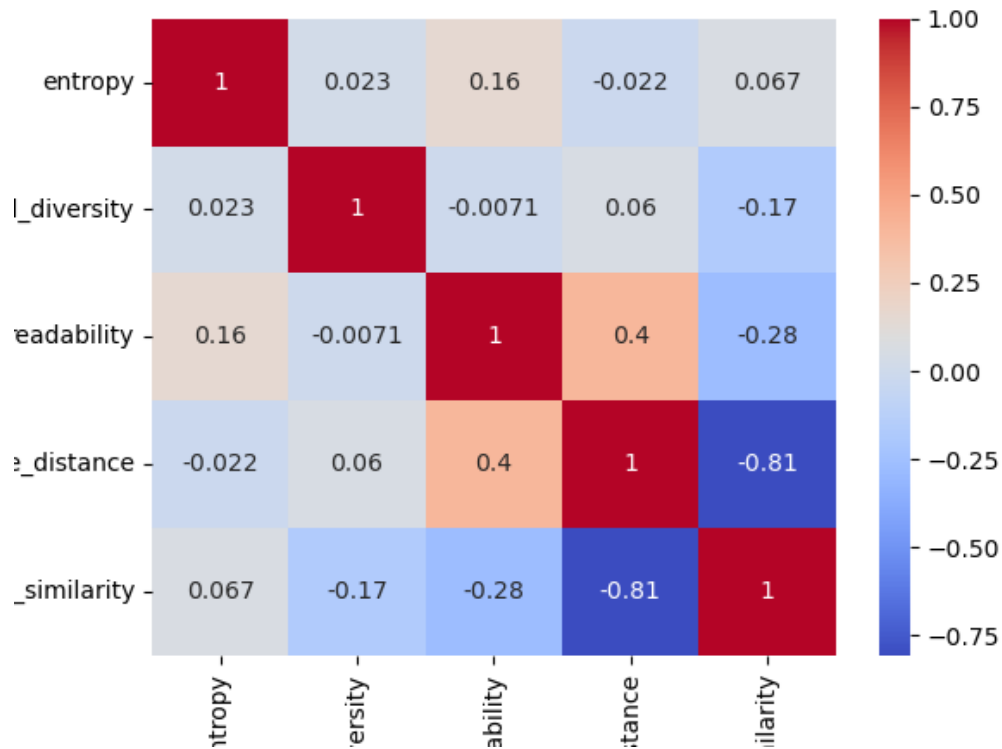


Mean average distance in each album appears to increase over time



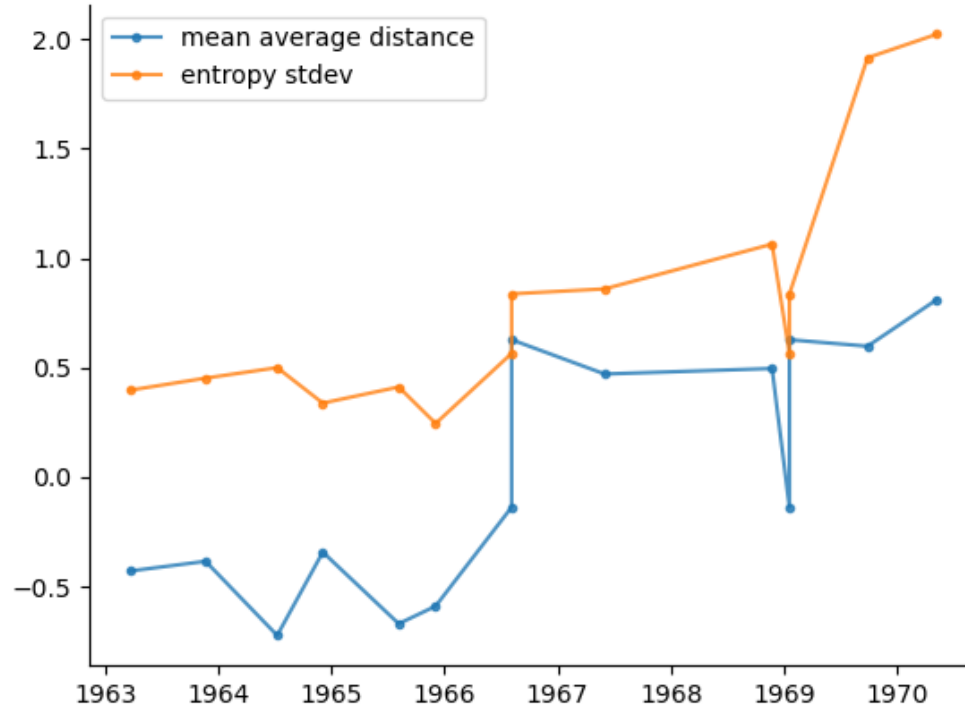


Audio complexity is not really correlated with lyric complexity





Standardised entropy spread and mean average distance (hacking?)





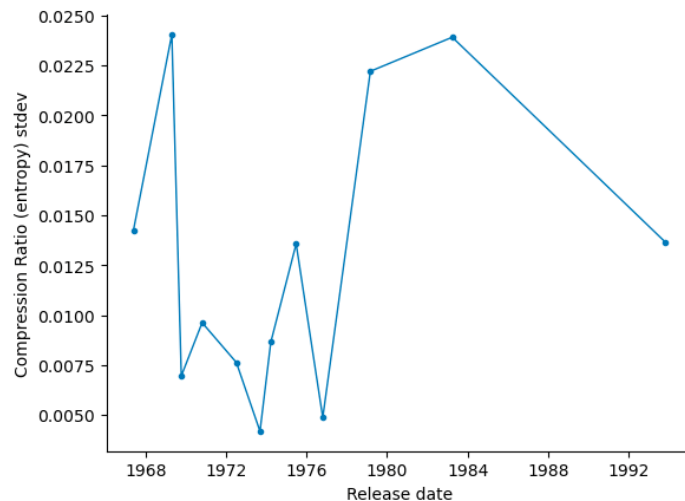
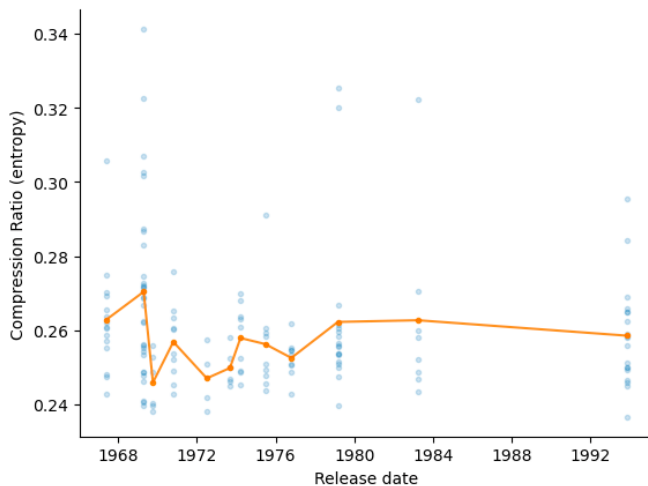
Conclusion and discussion

- **Audio**
 - Mean entropy across songs per album did not vary significantly
 - Standard deviation increased over time
 - Similar trend of standard deviation for within-album song similarity
 - **Songs that were released in close succession are not more similar!**
 - **Lyrics**
 - The mean text complexity, especially measured as the average semantic distance between words, increase over time
- **Audio and lyrics complexity are not directly correlated with each other, but there are similar trends on average**
- We find that we can relate the **compression-based complexity measures** and **semantic lyrics complexity** for *the Beatles* to their **musical progression**, which indicates a promising route to find music complexity more generally



Next steps

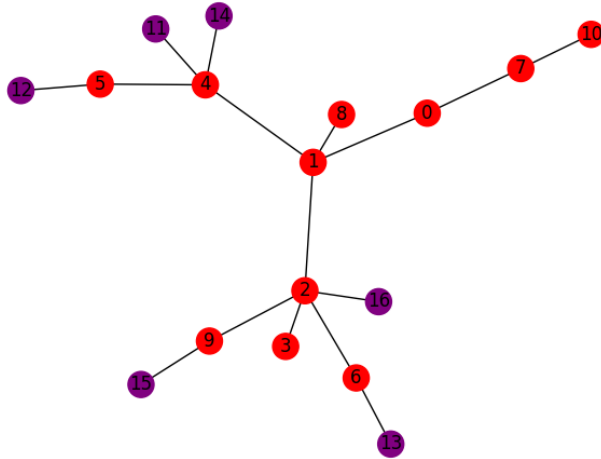
- Song network based on pitches - measure KL divergence (Nardelli et al. 2022)
- Calculate compression ratio for other artists within and across genres



Frank Zappa



Our music genre preferences represented using networks!



Above: Tree network generated according to model

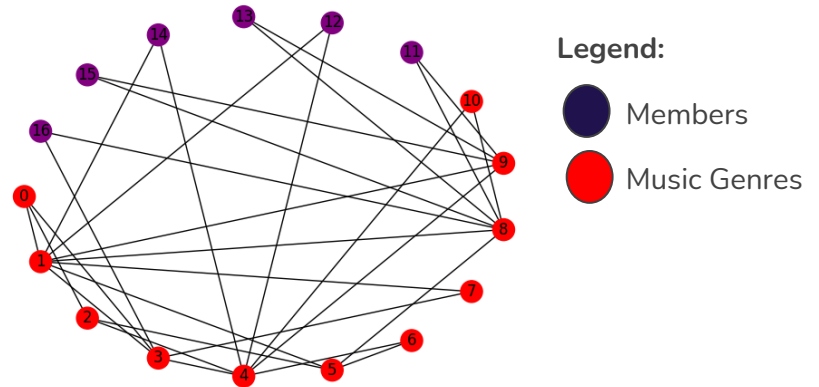
Right: Shell network generated according to model

Random Network Graph Parameters:



Model: Barabasi Albert preferential attachment

n: number of nodes (int)

m: number of edges to attach from new node to existing nodes(int)



Legend:

-  Members
-  Music Genres