Greatest Hits Versus Deep Cuts: Exploring Variety in Set-lists Across Artists and Musical Genres

Edward Abel[1,*], Andrew Goddard[2]

[1]University of Southern Denmark, Denmark
[2]Independent Researcher

Abstract

Live music concert analysis provides an opportunity to explore cultural and historical trends. The art of set-list construction, of which songs to play, has many considerations for an artist, and the notion of how much variety different artists play is an interesting topic. Online communities provide rich crowd-sourced encyclopaedic data repositories of live concert set-list data, facilitating the potential for quantitative analysis of live music concerts. In this paper, we explore data acquisition and processing of musical artists' tour histories and propose an approach to analyse and explore the notion of variety, at individual tour level, at artist career level, and for comparisons between a corpus of artists from different musical genres. We propose notions of a shelf and a tail as a means to help explore tour variety and explore how they can be utilised to help define a single metric of variety at tour level, and artist level. Our analysis highlights the wide diversity among artists in terms of their inclinations toward variety, whilst correlation analysis demonstrates how our measure of variety remains robust across differing artist attributes, such as the number of tours and show lengths.

Keywords

computational musicology, statistical music analysis, set-list composition, music information retrieval

1. Introduction

Live music experiences offer a unique glimpse into society and have significant cultural impact [6]. They have also become a crucial source of revenue for artists in the streaming era [17]. Constructing live music set-lists involves several considerations for an artist, such as catering to different types of fans with varying expectations and managing the trade-offs between these expectations [26]. Artists must also consider how performing specific songs or covers could attract more media attention than the concert might otherwise receive [16, 24]. The implications of a set-list now extend beyond just the audience in attendance in the venue, with artists like Bruce Springsteen offering the ability to buy and stream every single show from current and previous tours [15], while bands like Metallica[1] and Pearl Jam[2] provide numerous official live recordings of their performances. In addition, live shows' set-lists have been shown to poten-

CHR 2024: Computational Humanities Research Conference, December 4 – 6, 2024, Aarhus, Denmark.
*Corresponding author.
edabelcs@gmail.com (E. Abel); bossfansheff@gmail.com (A. Goddard)
www.edabel.co.uk (E. Abel)
0000-0002-3694-5116 (E. Abel); 0000-0001-7384-0252 (A. Goddard)

[1]https://www.livemetallica.com/
[2]https://pearljam.com/music/bootlegs

tially impact and influence listening behaviour of an artist's fans regarding non-live material [30].

The process of set-list creation varies, with some artists meticulously planning their sets in advance, while others prefer to make more spontaneous decisions based on the energy and reactions of the crowd during the show. Getting the right balance of set-list variety is crucial. Artists may face backlash if their set-lists are perceived as too formulaic and lacking in variety. For example, during Bruce Springsteen's 2023 tour, fans criticised performances for being too similar night after night, leading band members to defend their choices on social media [25]. Conversely, there is a risk of alienating fans by including too many unexpected songs at the expense of beloved greatest hits, as seen by some fan's reactions to recent Bob Dylan shows [13], or by altering hit songs too much from their album versions, as seen by fan frustrations at recent Arctic Monkeys' shows [7]. Constructing a set list involves navigating diverse expectations, and the degree of variety in an artist's performances is an intriguing topic with broader applications, such as historical live performance recommender systems [1].

In this paper, we explore the data acquisition and processing of musical artists' tour histories, and propose an approach to explore variety, at individual tour level, artist career level, and for comparisons between a corpus of artists. For this, we propose notions of a shelf and a tail, as a means to investigate and explore tours' variety and properties, and explore how they can be used to quantify variety at tour level and artist level. Additionally, our approach explores the impact of cover songs upon variety, and explores variety comparisons across musical genres.

Through analysing many artists, we highlight the wide range of artist variety levels, from those who tend to focus more on greatest hits to those who perform more diverse shows featuring deep cuts, and everything in between. Additional correlation analysis of our concept of variety explores how our measure remains robust across artists with differing characteristics, such as the number of tours or average show lengths. For more information about the project and its data, see the project's GitHub repository,[3] and to interactively explore the data and the proposed approach, see the project's Interactive Web App[4].

The rest of the paper is structured as follows: Section 2 covers background literature, Section 3 details our approach, and Section 4 provides conclusions.

2. Background

Work has explored live music performance exploring performer or audience psychology and touching on how set-lists impact this, such as in terms of presence and representation within shows and set-lists [27], through measuring value of live music as a motivation scale [23], and exploring how the set-list and beyond has a bearing on an artist's impact as part of their stage success [10]. Other work has explored live music performances from the perspective of music theory, such as work defining performance parameters in relation to how performances bring compositions to life through variations in timing and dynamics and its impact on listener perception [19], and work exploring how composers' choices in tones, intervals, and harmonies influence stylistic music changes over time [22].

[3]https://github.com/EdAbel/setlist-variety
[4]www.edabel.co.uk/setlist-variety

Various work exploring live music performance taking a more quantitative approach have focused on an in-depth analysis of a single artist. Work has explored the band The Grateful Dead, examining the band's live recordings over three decades to analyse the performances in relation to cultural trends in music and to investigate how the performances of the band change over time [33]. Others have analysed The Grateful Dead's live concerts from 1972 to 1995 in comparison to listening habits outside of concerts [30], highlighting how there are correlations between live set-lists and home listening. Work focusing on the artist Bruce Springsteen, has explored his live performances and set-lists from the perspective of examining how set-list analysis can provide inclinations about tensions between commercial considerations for playing new album material and playing expected but older hits [2]. Bob Dylan is another artist work has focused upon, such as a study of Bob Dylan's set-lists from the 1960s to the 2020 that investigates his approach to performing, exploring how he curates a show and how this in turn creates a meta-narrative [8].

Limited studies have looked to compare tour and set-lists for multiple artists. The Music & Entertainment publication Consequence set out and discussed the "25 Best Rock Acts with the most Unique set-lists" [31]. Within such articles, although there is an implication of a focus, here on the acts whose shows define the word unique, the methodology involved in compiling such a list is obscured, and so therefore it is difficult to assess.

Some recent discourse has explored how the make up of fan communities of artists such as Taylor Swift and The Grateful Dead may share surprising similarities, despite their different musical styles, and the impact this can have on further similarities between performance set-lists [5]. The notion of a comparison between Taylor Swift and The Grateful Dead performances has been further explored, along with a limited number of other artists, in terms of unique songs [11]. Focusing on considerations of the notion of special songs, analysis of set-lists from individual tours for different artists is carried out, to explore the prominence of unique (and quite unique) song occurrence rates. Comparisons highlight how The Grateful Dead is seen as a very varied artist compared to Taylor Swift, when considering variety only from the perspective of unique special "surprise" songs. In this analysis, only individual tours are considered, which may result in an unrepresentative view of an artist's overall career. Additionally, the methodology favours artists with longer tours and longer sets.

The metric Consecutive Set Similarity (CSS) has been proposed to look to measure variety for an artist [20]. Here, an artist's career of set-lists are arranged in sequence and each set-list is compared to the previous set-list in terms of the amount of different songs, resulting in a value of 1 if the the set-list is identical to the previous set, and -1 if it is completely different. From this, an overall average is derived for each artist, from which comparisons and clustering of different artists can be performed [28]. The measure has a very narrow focus due to only considering two shows at a time, and so will consider oscillating changes of songs that are played frequently but not every night as signalling high variety. Moreover, it does not consider information such as when one tour ends and another begins (likely to signal a significant change in set-lists, which may in turn favour artists with many small tours over a long period).

Some coverage has highlighted how artists themselves explore the art of set-list curation, and the considerations they have for variety. During a tour that contained over 800 shows the band Radiohead made a conscious choice to ensure every show was unique and to never have a repeating set-list. The band would curate the set-lists daily and emphasised the importance of

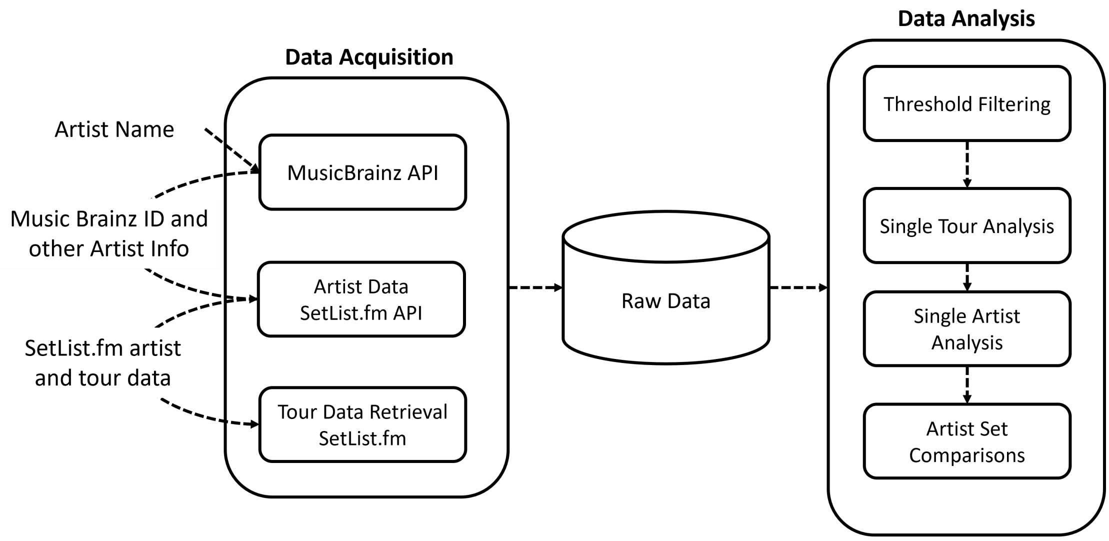


Figure 1: Data Pipeline Stages of our Approach

variety, contrasting their approach with other artist's that play identical sets nightly [21]. More recently, it has been claimed that the band Metallica look to customise concert set-lists based on local Spotify data (and local radio trends), to look to cater to localised fans' preferences, and to help increase show diversity [29].

3. Our Approach

In our approach, raw artist tour data is first collected and processed, then utilized in various stages of analysis, as illustrated in Figure 1. Communities such as MusicBrainz[5] and Setlist.fm[6] provide extensive crowd-sourced encyclopaedic data on musical artists and live concert set-lists. The following sub-section outlines our data acquisition process, detailing how these sources were leveraged to obtain the data used for our analysis.

3.1. Data Acquisition

Given a set of music artist names, we acquire tour details, and data for each tour of each artist, and store the data as depicted in the data model in Figure 2. For each artist name, they can be uniquely identified via MusicBrainz Identifiers (MBIDs), which is a universal unambiguous standard artist identification.[7] Through calls to the Music Brainz API,[8] the MBID for each artist name is acquired, along with additional Music Brainz artist information including their Gen-

[5] https://musicbrainz.org/
[6] https://www.setlist.fm
[7] https://musicbrainz.org/doc/MusicBrainz_Identifier
[8] https://musicbrainz.org/doc/MusicBrainz_API (utilising the musicbrainz API wrapper R package - https://github.com/dmi3kno/musicbrainz

Table 1
Genre Mappings

General Genre	(Specific) Genre
Rock	Rock, Hard Rock, Pop Rock, Southern Rock, Gothic Rock, Comedy Rock, Blues Rock, Rap Rock
Alternative	Alternative Rock, Indie Rock, Post-Rock, Dance-Punk, Garage Rock, Indie Pop, Britpop, Emo, Grunge, Alternative Hip Hop
Metal	Heavy Metal, Thrash Metal, Power Metal, Gothic Metal, Melodic Death Metal, Progressive Metal, Nu Metal, Symphonic Metal, Alternative Metal, Groove Metal, Metalcore, Death Metal, Glam Metal
Punk	Pop Punk, Punk Rock, Hardcore Punk, Post-Hardcore, Gypsy Punk
ElectronicAndDance	Synthpop, Electronic, Industrial Metal, Industrial Rock, Dance
Folk	Folk Rock, Indie Folk, Folk Punk
Progressive	Progressive Rock, Progressive Metal, Experimental Rock, Post, Psychedelic Rock
Pop	Pop, Pop Rock

der (one of Male, Female, or Group), and their *Start Date* and *EndDate*.[9] Additionally, a single musical genre considered most representative was curated for each artist using MusicBrainz's genre information. As this data often assigns multiple genre tags to an artist, a manual selection of a single tag was performed where necessary. Next, using each artist's MBIDs, each's corresponding setlist.fm ID (a separate setlist.fm unique identifier for each artist) is obtained via calls to the setlist.fm API.[10] This data is stored as depicted in the Artist table in Figure 2.

The determined single genres for each artist result in a large number of different genre values, many representing similar sub-genres of a more general genre. To make genre analysis more tractable, the set of genres can be mapped onto a smaller set of generalized genres. Such a mapping of dozens of genres onto a set of generalized genres is shown in Table 1. This data is stored as depicted in the Genre table in Figure 2.[11]

Given a set of artist setlist.fm IDs, a list of each artist's tours can be obtained, from which data associated with each tour of each artist can be subsequently acquired. For each tour, overall information of the tour name, the total number of shows on the tour, and date ranges are acquired, and stored as depicted in the Tour table in Figure 2. For each tour's songs, the list of songs played on the tour, along with each song's number of plays on the tour are acquired. Additionally, for each song, whether the song is denoted as a cover song (with respect to the artist the tour is for) is recorded.[12] The song information is stored as depicted in the Song table

[9]For groups this represents their formation date and disband dates, for solo artists it holds just their birth dates and retirement or death dates. For ongoing artists EndDate will be "present".

[10]https://api.setlist.fm/docs/1.0/index.html (utilising the SetListR wrapper R package - https://github.com/fusionet24/SetListR

[11]The assignment of a single genre to each artist, followed by the mapping of these genres to a broader set of generalized categories, has been curated as a proof of concept. However, genre classification is a complex and expansive subject in its own right, with numerous studies addressing the challenges associated with genre categorisation [9] and the phenomenon of genre crossover [32]. Given such complexities, the automation of genre classification represents a compelling area for future exploration.

[12]Additionally, tours that are empty (made up of only empty shows) are removed, as are any artists for which all

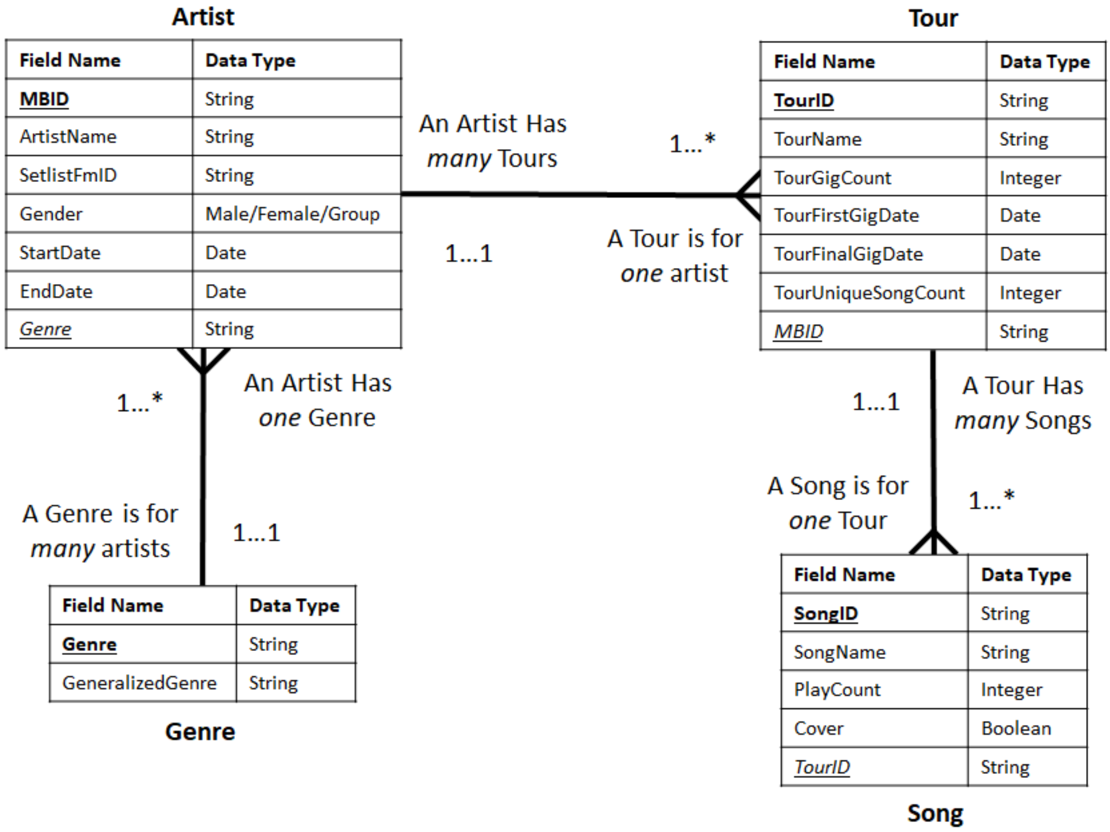


Figure 2: Data Model of our Approach

in Figure 2.

From this, full artist tour history data for over 200 artists was acquired, chosen for their prominence within popular music history and culture. This number is progressively expanding, and up-to-date figures, and access to the data, can be found at the project's GitHub repository.[13] All of the data acquired is stored as raw data, as depicted in Figure 1, to then be utilised within the analysis stage.

3.2. Data Pre-Processing

Before beginning analysis, pre-processing of the raw dataset is performed. Within our analysis, we are interested in artists who have sufficiently substantial touring histories. Therefore, we define thresholds to utilise only artists that have a minimum number of tours and a minimum overall number of shows, and only keep tours that have a minimum number of shows, a minimum number of unique songs played on the tour, and a minimum show length of the

tours are empty. Further, tours identified by their name as a set of Promotional publicity media/private shows, are removed.

[13]https://github.com/EdAbel/setlist-variety

Table 2
Raw Data Threshold Parameters

Variable	Value
Artist Minimum No. of Tours	5
Artist Minimum Total No. of shows	200
Tour Minimum No. of shows	20
Tour Minimum No. of songs	10
Tour Minimum Average Show Length	10

shows from the tour.[14] These threshold values are inherently subjective and context dependent; therefore, we conduct the analysis on the raw data, preserving its integrity so that alternative thresholds could be applied if needed. Within our analysis that follows, the threshold parameter values utilised are shown in Table 2. Following the pre-processing stage, we begin the analysis at individual tour level.

3.3. Tour Analysis

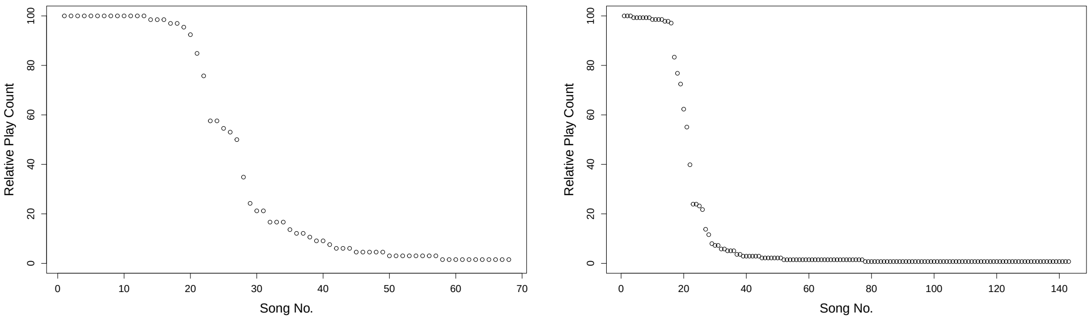
For a tour, each (unique) song has a Play Count (*PC*), denoting how many times it was played on that tour. Tours can vary in terms of how many shows they are made up of, therefore, for comparisons between tours, a tour's (absolute) *PC* values can be normalised with respect to the number of shows in the tour. For a tour, a Relative Play Count (*RPC*) for each song played on the tour can be computed via:

$$RPC_i = (\frac{\mathrm{PC}_i}{tN}) * 100 \tag{1}$$

Where PC_i is the *PC* value of the *i*-th song and tN is the number of shows in the tour. A *RPC* value of 100 represents a song being played every single show of a tour, whilst a value of 50 represents a song being played at exactly half of the tour's shows. We can visualise a tour and its songs in terms of their *RPC* values, where the y-axis denotes *RPC* value, and the x-axis denotes song number where songs are sorted with respect to *RPC* values high to low. For example, Bruce Springsteen (and the E-Street Band's) 2023 tour, is shown in Figure 3a, and Coldplay's Music of the Spheres 2023 Tour is shown in Figure 3b. Such tour visualisations highlight a generality for many tours to map out an s shaped sigmoidal like function shape. From a tour's dataset, notions of *Shelf, Tail, 100%'ers, Uniques and Covers* can be calculated and subsequently highlighted within such visualisations. Each of these notions are defined and explained next.

Shelf - The notion of a tour's *Shelf* is a measure of the significance of a tour to have a set of of songs that are played at most of the tour's shows, and outline a shelf like shape in the top left of the plots in Figure 3. Given a Shelf Size (*SS*) value, denoting what top percentile of tour songs are to be considered part of the shelf, a Shelf Value *S* can be calculated as the ratio of a

[14] Such thresholds are beneficial for identifying and removing tours with missing data issues, such as tours which only a few of the shows have been added for, or tours that have many shows added but lack song information for many of the shows. Such tours, if left in the data, can unduly impact and bias analysis.



(a) Bruce Springsteen (and The E-Street Band) - 2023 Tour (Just Data Points).

(b) Coldplay - Music of the Spheres 2023 Tour (Just Data Points)

Figure 3: Tours Visualisation Just Data Points

tour's songs that are in its shelf. Given the tour's set of $\mathbf{X}$ songs, of length n, sorted as, x_1 to x_n, from high to low with respect to their *RPC* values:

$$\mathbf{X} = \{x_1, x_2, \dots, x_n\} \quad \text{where} \quad x_1 \geq x_2 \geq \dots \geq x_n \tag{2}$$

The *Shelf Songs* are selected as the top *SS* percentile of $\mathbf{X}$:

$$\textit{Shelf Songs} = \text{Top}_{SS}\text{Percentile} = \{x \in \mathbf{X} | x \geq P_{SS}(\mathbf{X})\} \tag{3}$$

and S is calculated via:

$$S = \frac{|\textit{Shelf Songs}|}{n} \tag{4}$$

An *SS* value of 10% would select all the songs that have been played at 90% or more of a tour's shows, and the corresponding S value represents the ratio of the tour's songs that are played at 90+% of the tour's shows. So, an S value of 0.25 would represent that 25% of the tour's songs are played at 90% or more of its shows.

Tail - The notion of a tour's *Tail* is a measure of the significance of a tour to have a set of songs that are played only rarely on the tour, and outlines a tail like shape in the bottom right of the plots in Figure 3. Given a Tail Size (*TS*) value, denoting what bottom percentile of tour songs are to be considered a part of the tail, a Tail Value T can be calculated as the ratio of a tour's songs that are in its tail. The *Tail Songs* are selected as the bottom *TS* percentile of $\mathbf{X}$:

$$\textit{Tail Songs} = \text{Bottom}_{SS}\text{Percentile} = \{x \in \mathbf{X} | x \leq P_{TS}(\mathbf{X})\} \tag{5}$$

and T is calculated via:

$$T = \frac{|\textit{Tail Songs}|}{n} \tag{6}$$

A TS value of 10% would select the songs that are played at most at 10% of a tour's shows, and the corresponding T value represents the ratio of the tour's songs that are played at 10% or less of the tour's shows. So, a T value of 0.4 would represent that 40% of the tour's songs are played at 10% or less of its shows.

100%'ers - In the set of songs making up the tour's shelf, there exists a subset of 0 or more songs that are played at 100% of the tour's shows. This subset of songs (*100%'er Songs*) can be identified as the set of songs that have $RPC = 100$. A 100%'ers Value H is calculated as the ratio of a tour's songs that are in this subset.

$$H = \frac{|100\%\text{'er Songs}|}{n} \tag{7}$$

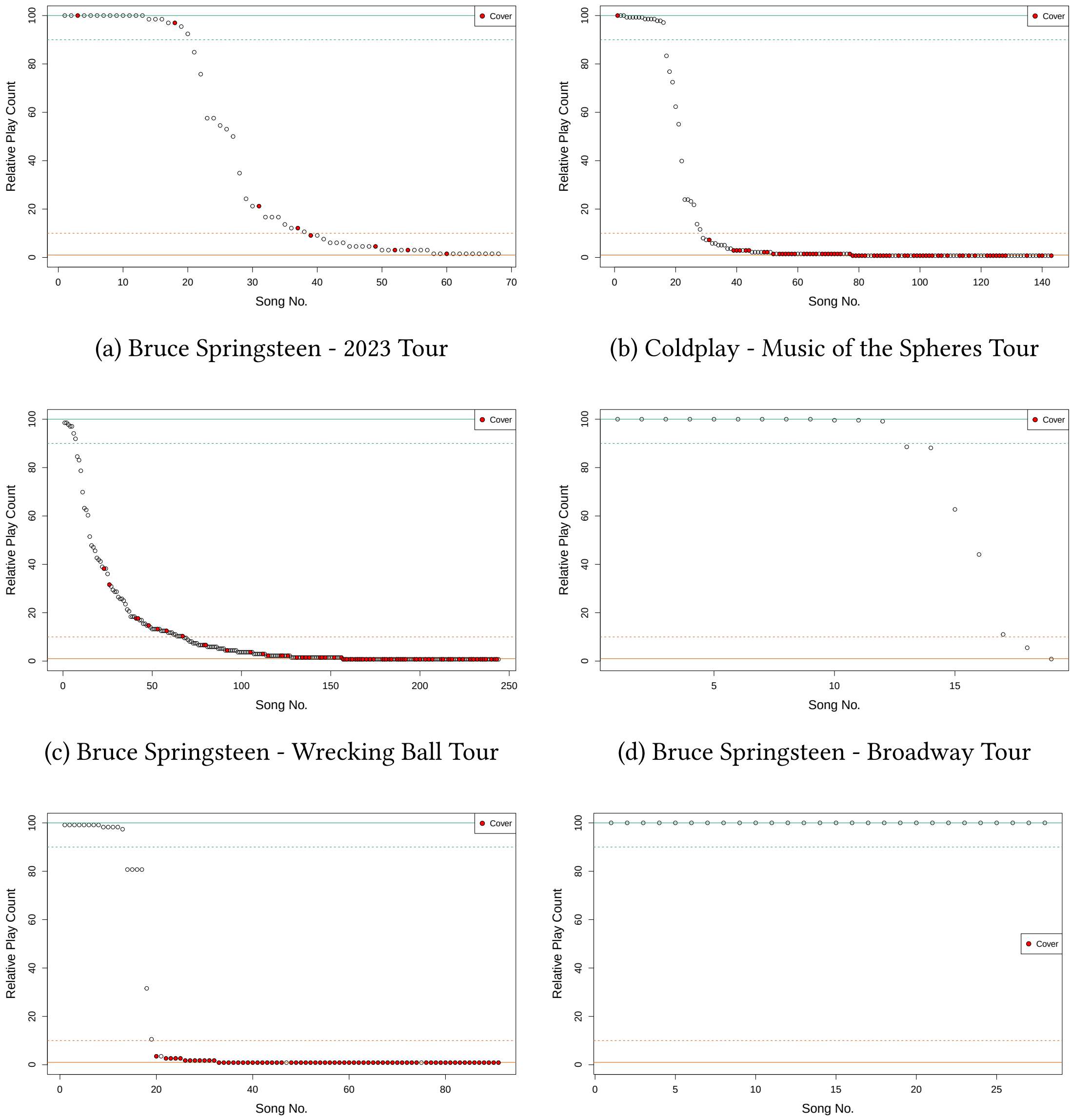
Uniques - In the set of songs making up a tour's tail, there exists a subset of 0 or more songs that are played only once during the whole tour. This subset of songs (*UniqueSongs*) can be identified as the set of songs that have a $PC = 1$. A Uniques Value U is calculated as the ratio of a tour's songs that are in this subset.

$$U = \frac{|\text{Unique Songs}|}{n} \tag{8}$$

Covers - For each song played on the tour we have information denoting which are cover songs (with respect to the artist), from which the shelf songs that are cover songs, and the tail songs that are cover songs, can be determined. From this, the set of shelf songs minus those that are covers can be determined and a Shelf Minus Covers Ratio Value SMC calculated. Similarity, the set of tail songs minus those that are covers can be determined and a Tail Minus Covers Ratio Value TMC calculated.

These calculated notions of Shelf, Tail, 100%'ers, Uniques and Covers can be highlighted visually within our tour plots, as shown for the tours introduced earlier of Bruce Springsteen (and the E-Street Band's) 2023 tour in Figure 4a, and Coldplay's Music of the Spheres 2023 Tour in Figure 4b. In these plots, the shelf lower edge is denoted via a dotted green line and the 100%s'ers are those songs that sit on the solid green line. The tail upper edge is denoted via a dotted orange line and the uniques are those songs that sit on the solid orange line. Cover songs are denoted by filled red data points, and the shelf and tail covers can be identified as those filled data points above the shelf's lower edge and below the tail's upper edge respectively. Here, and within subsequent analysis, SS and TS parameters of 10% are utilised. However, these parameters can be altered to any desired numbers. For experimentation exploring sensitively analysis and impacts of these parameters see Appendix A.

Such analysis can be utilised to explore and compare different tours from the same artist and to compare tours of different artists. For example, regarding other Bruce Springsteen tours, the plot for the Wrecking ball tour is shown in Figure 4c and the plot for the Bruce Springsteen on Broadway tour shown in Figure 4d. These plots highlight how, as in The Wrecking ball tour's case, a tour can have quite a small and sharp shelf, or, as in the Bruce Springsteen on Broadway tour, a tour can alternatively be predominantly made up of a shelf. The tour for Taylor Swift's Speak Now World Tour is shown in Figure 4e, highlighting a tour which has a large long tail that is made up of many uniques that are cover songs. The tour for Pink Floyd's The Wall tour



(a) Bruce Springsteen - 2023 Tour

(b) Coldplay - Music of the Spheres Tour

(c) Bruce Springsteen - Wrecking Ball Tour

(d) Bruce Springsteen - Broadway Tour

(e) Taylor Swift - Speak Now World Tour

(f) Pink Floyd - The Wall Tour

Figure 4: Tour Visualisations with Shelf, Tail, and Covers Identified

is shown in Figure 4f, highlighting a tour where all of the tour's songs are contained in the shelf and in fact are all 100%'ers.

Building on this analysis of single tours, next, we explore analysis of an artist's whole career of tours.

3.4. Artist Career of Tours Analysis

For a single tour, its Shelf Value S and Tail Value T can be calculated. With the calculation of a pair of such values for every tour of an artist's career, their whole career can be visualised, in chronological order, in a bar plot. Such a visualisation, for every tour for Bruce Springsteen is shown in Figure 5a. Here, for each tour, tail values are shown as negative blue bars and each tour's corresponding shelf values are shown as green bars.

Additionally, for each tour, its 100%'ers Value H and Uniques Value U can be calculated. These pair of values represent values equal to or less than the tour's S and T values. Therefore, an artist's whole career can be visualised in a bar plot where the amount of each tour's tail that is made up of uniques, and the amount of each tour's shelf that is made up of 100%'ers is highlighted. For Bruce Springsteen, every tour with this information is shown in Figure 5b.

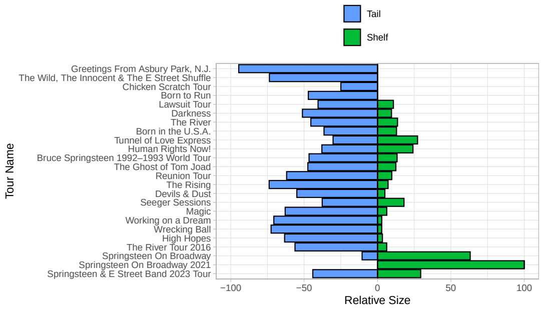
Moreover, for each tour the amount of its shelf that is made up of covers, and the amount of its tail that is made up of covers can be calculated. Then, an artist's whole career can be visualised with the amount of each tour's tail that is made up of covers, and the amount of each tour's shelf that is made up of covers highlighted. Every tour for Bruce Springsteen with this information is shown in Figure 5c. Alternatively, we could visualise the impact of cover songs on each tour's shelf and tail through calculating each tour's Shelf Minus Covers Value SMC and Tail Minus Covers Value TMC. Then, an artist's whole career can be visualised, highlighting how shelves or tails that are made up of a substantial amount of cover songs will become smaller. Every tour for Bruce Springsteen with this information is shown in Figure 5d.

Such analysis can be utilised to explore and compare the careers of different artists. For example, Figure 6a shows the career of Iron Maiden, and Figure 6b shows the career of Slipknot, each showing Shelf and Tail values and how much of them are taken up by 100%s and Uniques. These plots highlight how these artists have a clear leaning towards playing more conformity and greatest hits like sets, and highlight how over their career this has only become more pronounced. The career of Taylor swift, showing each tour's Shelf and Tail values and how much of them are taken up by covers, is shown in Figure 6c. Here, we observe how after her first tour, a similar size of shelves and tails is observed. However, whereas the Speak Now World Tour's tail is made up almost entirely of cover songs, we observe the inverse for the tail of the most recent Eras tour. The range of different shelf and tail values for the tour's of Pink Floyd are shown in Figure 6d. From this plot we observe stark differences between early tours which have little or no shelves, and later tours, that coincide with their commercial peak, having little or no tails and large shelves.

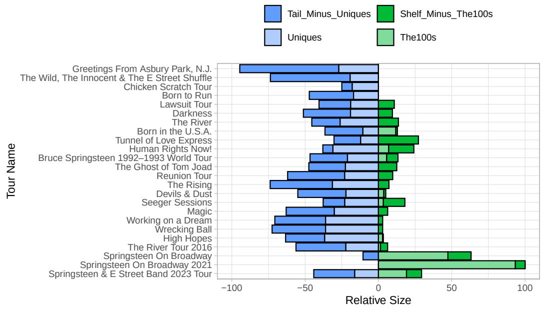
Building on this analysis of the whole touring career of a single artist, next, we explore comparison analysis between a corpus of artists.

3.5. Comparing Artists

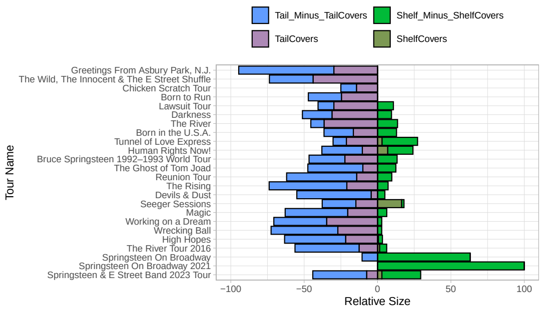
For an artist, Shelf S and Tail T values for each tour can be calculated, denoting the size of shelf and tail for each tour. From the set of shelf values, a single average shelf value $\bar{S}$ can be calculated via:



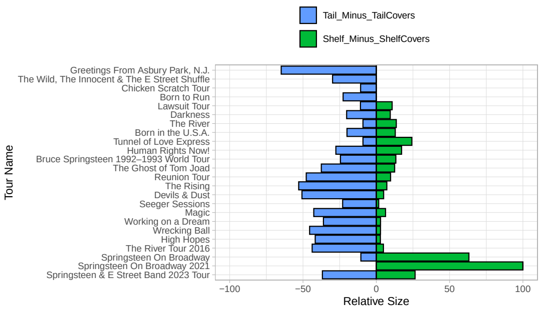
(a) All Tours Shelf and Tail



(b) All Tours Shelf & 100%'ers and Tail & Uniques



(c) All Tours Shelf & Covers and Tail & Covers



(d) All Tours Shelf And Tail Without Covers

Figure 5: All Tours Analysis - Bruce Springsteen

$$\bar{S} = \frac{1}{n}\sum_{i=1}^{n} S_i \tag{9}$$

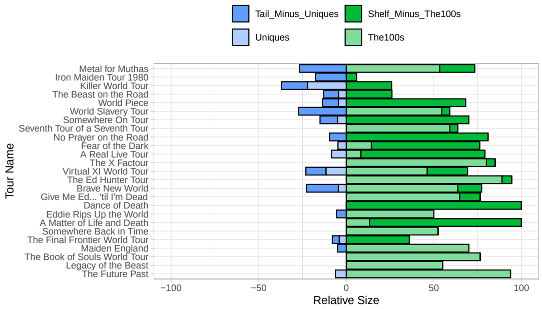
where $\bar{S}$ is the average of the individual Shelf values, S_i represents the Shelf value for Tour i, and n is the number of Tours for the artist.

Similarly, an average tail value $\bar{T}$ for the artist can be calculated via:

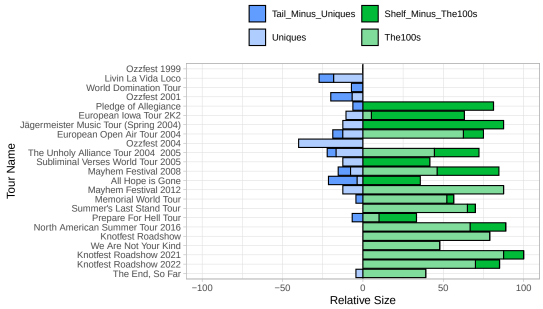
$$\bar{T} = \frac{1}{n}\sum_{i=1}^{n} T_i \tag{10}$$

where $\bar{T}$ is the average of the individual Tail values, T_i represents the Tail value for Tour i, and n is the number of Tours for the artist.

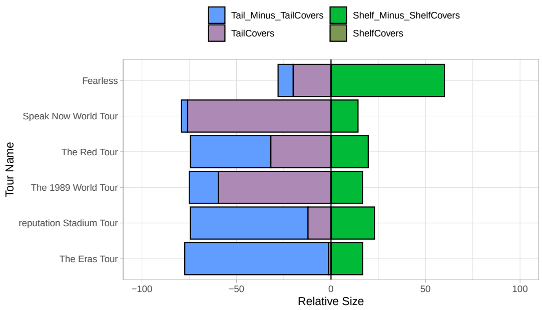
From these calculations, a pair of $\bar{S}$ and $\bar{T}$ values can be calculated for every artist, and the whole set of artists can be shown within a single scatter plot, as shown in Figure 7. Here, the x-axis denotes mean tail values ($\bar{T}$) and the y-axis denotes mean shelf values ($\bar{S}$), with the shelf axis scale inverted to highlight how, in variety terms, the larger the shelf the less variety, and the larger the tail the more variety. Each data point in the plot represents an artist, coloured with respect to their generalized genre value. The solid blue diagonal line, running from the bottom left to the top right, signifies the vector of values where the sum of $\bar{S}$ and $\bar{T}$ is 100, which



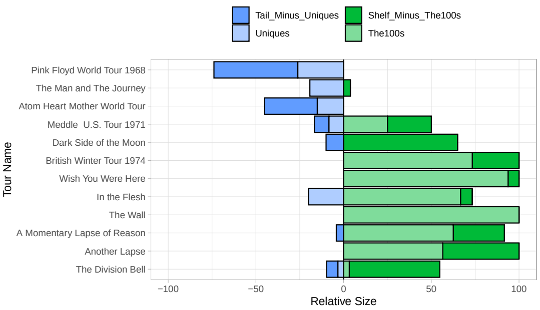
(a) Iron Maiden – All Tours Shelf & 100%'ers and Tail & Uniques



(b) Slipknot – All Tours Shelf & 100%'ers and Tail & Uniques



(c) Taylor Swift – All Tours Shelf & Covers and Tail & Covers



(d) Pink Floyd – All Tours Shelf & 100%'ers and Tail & Uniques

Figure 6: All Tours Analysis - Various Artists

would signify that 100% of songs (in all the artist's tours) are contained within shelves and tails. Therefore, each artist's distance to this line signifies their average combined shelves and tail size. The solid red diagonal line, running from the top left to the the bottom right of the plot, signifies the set of pairs of equal $\bar{S}$ and $\bar{T}$ values. Artists sitting on this line have equally sized average shelf ($\bar{S}$) and tail ($\bar{T}$) values, artists that sit above the line have a greater average tail than shelf suggesting more variety, and artists that sit below the line have a greater average shelf then tail suggesting less variety. The distance each artist is from this line represents the strength of this property. Artists further towards the bottom left of the plot represent those with much larger shelves and smaller tails on average, suggesting they are the artists with the least variety. Artists further towards the top right of the plot represent those with smaller shelves and larger tails on average, suggesting they are the artists with the most variety. Such a visualisation, which preserves the dimensions of the shelf and the tail separately, enables nuanced comparisons within this multi-dimensional space [3]. This facilitates highlighting, for example, differences between artists who are equidistant from the solid red diagonal line but vary in their distance from the solid blue line.

Further analysis can consider shelves and tails not including the songs in them that are cover songs, through computing average shelf and tail values for each artist in relation to this, and

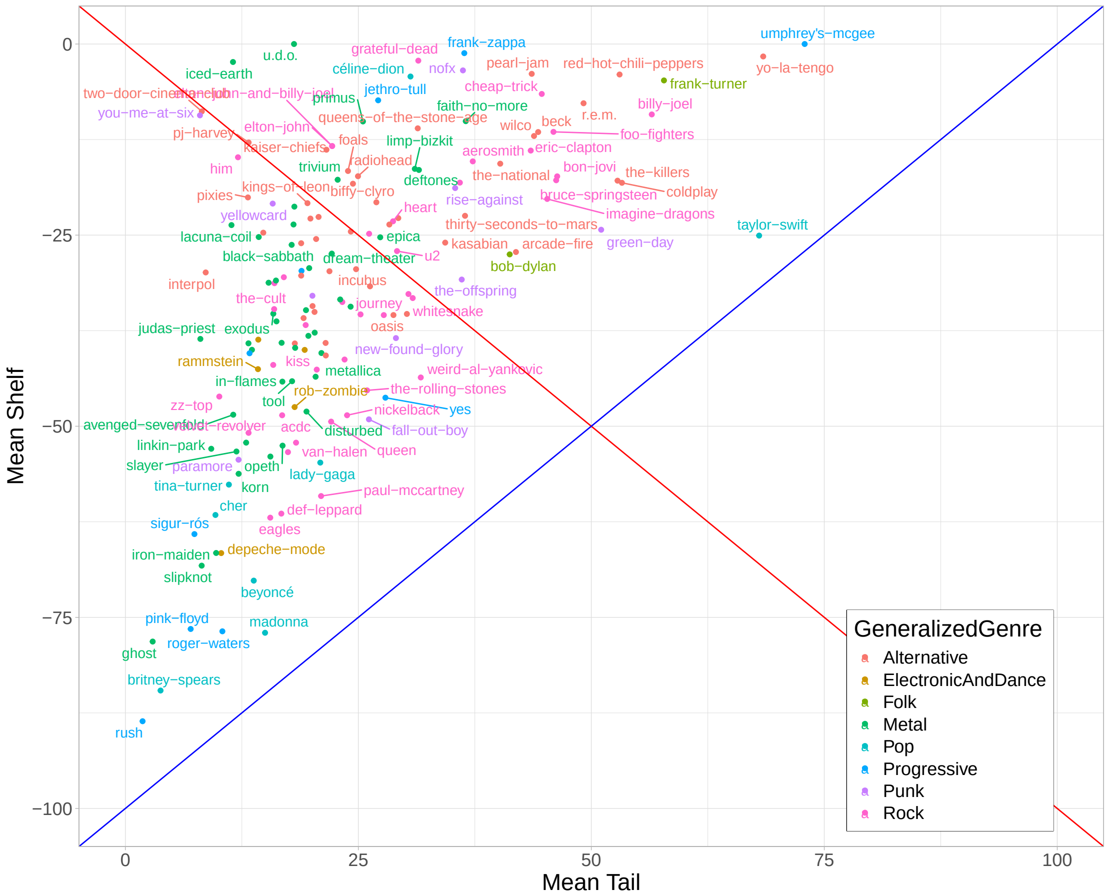


Figure 7: All Artists Comparisons - Average Tail Vs Average Shelf

creating a scatter plot of these results, as shown in Figure 8. In this plot, we see how some artists, that play a lot of covers within their tail, move further away from the top right of the plot, highlighting the importance, for some artists, of playing cover songs as part of their attainment of variety.[15]

From Figures 7 and 8's data, additional analysis can compute overall averages for each generalized genre. Calculated genre averages, for Tail and Shelf values are shown in Figure 9a. The plot highlights how genres, such as Electronic and Dance, and Pop, on average exhibit less variety that other genres, such as Folk, Alternative, and Punk. Calculated genre averages, when shelf and tail cover songs are not considered are shown in Figure 9b. Here, we observe the impacts removing covers has on the genre averages, and how the impacts are greater for some genres, such Rock and Alternative, than others, such as Punk.

Further analysis from Figure 7 and 8's data can, through division of the 2-dimensional plot

[15]Similarity we could explore utilising data pertaining to the amount of uniques and 100%'ers within shelves and tails to, for example, use a weighting system to give these songs more impact.

Figure 8: All Artists Comparisons - Average Tail Minus Tail Covers Vs Average Shelf Minus Shelf Covers

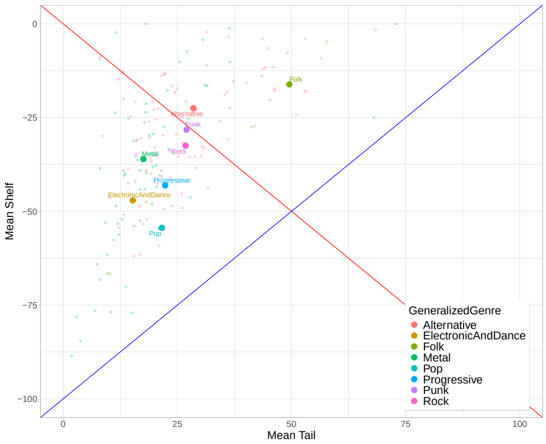
space, cluster the artists into a set of ordinal clusters from very high variety, to very low variety. For discussions and results from such clustering analysis see Appendix B.

Finally, in the pursuit of a single measure of variety for each artist, the shelf and tail values of a tour are combined, to derive a single measure of Variety V for each tour. For tour i, its Variety measure V_i can be calculated via:

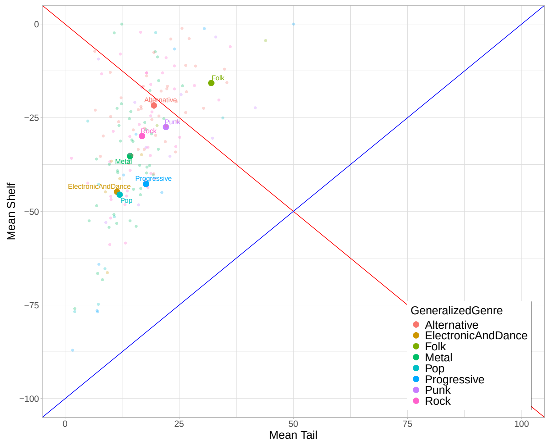
$$V_i = T_i - S_i \tag{11}$$

where T_i is the Tail value of tour i and S_i is the Shelf value of tour i. A positive value represents a tour with a tail larger than its shelf, suggesting more variety, and a negative value represents a tour with a tail smaller than its shelf, suggesting less variety. From this, an average overall variety $\bar{V}$ value of an artist's tours can be calculated via:

$$\bar{V} = \frac{1}{n}\sum_{i=1}^{n} V_i \tag{12}$$



(a) All Artists - Tail and Shelf analysis with Genre Averages



(b) All Artists - Tail without covers and Shelf without covers analysis with Genre Averages

Figure 9: Genre Averages Analysis

where $\bar{V}$ is the average of the individual tour V values, V_i represents the Shelf value for Tour i, and n is the number of Tours for the artist. The set of all artists and their $\bar{V}$ values, ordered with respect to $\bar{V}$, and coloured with respect to generalized genre, is shown in Figure 10, providing an overall visualisation of a corpus of artists with respect to variety.

3.6. Correlation Analysis of Variety with Other Features

To examine the robustness of our notion of variety for comparing different artists, despite their varying characteristics, such as the number of tours, the length of performances, and activity during different time periods, we conducted a correlation analysis between our $\bar{V}$ Variety measure and such properties. Table 3 shows the correlation results for seven artist properties, along with definitions of each property. The Correlation values are the correlation levels found between each of the seven properties and our $\bar{V}$ measure. Here, correlation is calculated with respect to Pearson Correlation Coefficient. For fuller descriptions and discussions of each of these properties see Appendix C, which also contains visualisation scatter plots of each property against our $\bar{V}$ measure. Table 3 highlights how our measure has only very weak correlation to these properties,[16] suggesting our measure is robust for analysis between artists, and will not be unduly bias by, for example, different artists having more tours or longer shows.

[16]Where the semantics of correlation strength can be classified as - Very Weak Correlation: $|r| < 0.2$, Weak Correlation: $0.2 \leq |r| < 0.4$, Moderate Correlation: $0.4 \leq |r| < 0.6$, Strong Correlation: $0.6 \leq |r| < 0.8$, Very Strong Correlation: $|r| \geq 0.8$ [12]

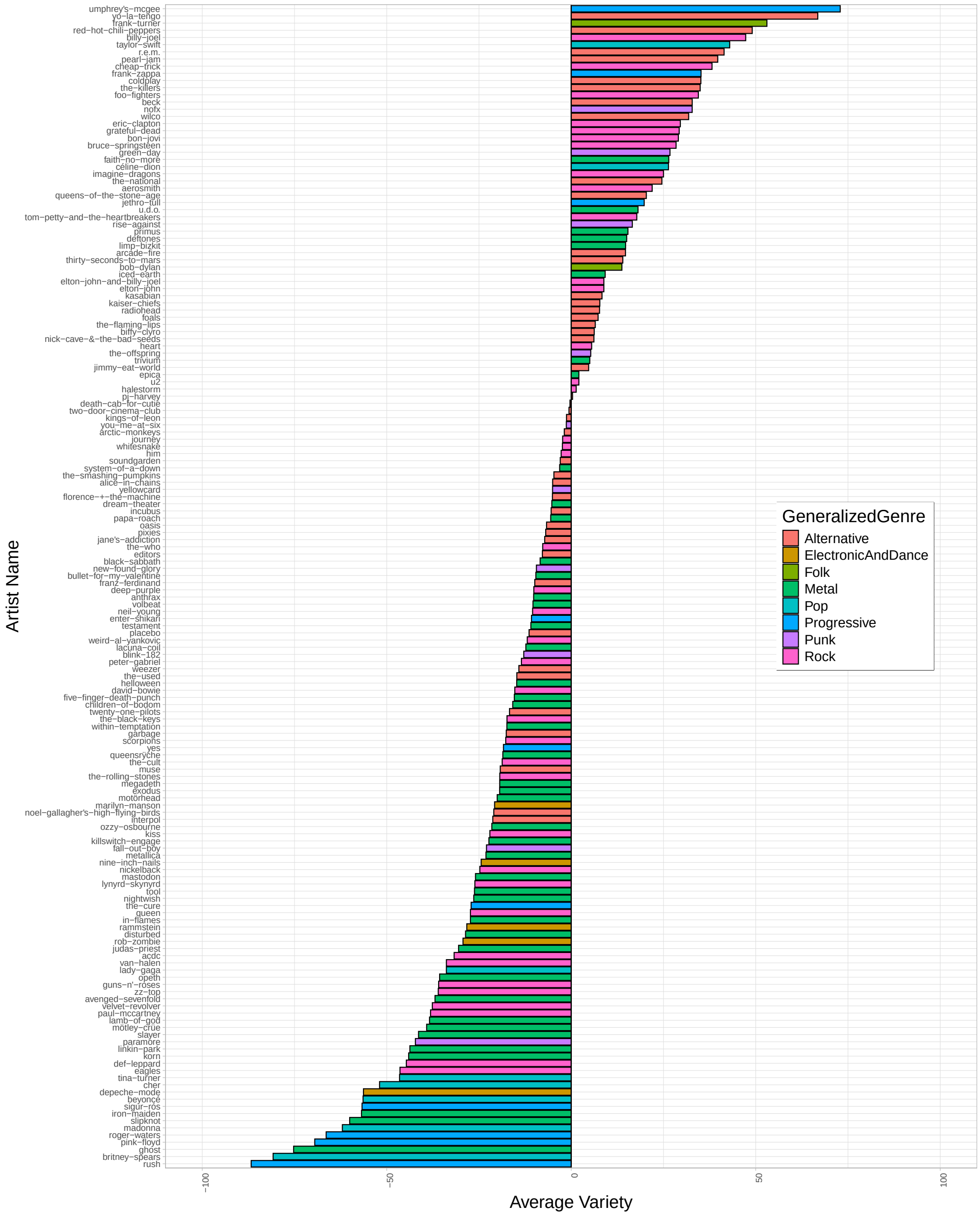


Figure 10: All Artists, Variety analysis, coloured by Generalized Genre

4. Conclusions

In this paper, we explored data acquisition and processing of musical artists' touring histories, and proposed an approach to explore set-list variety, at tour level, artist career level, and for

Table 3
Correlation Analysis

Property Name	**Description**	**Correlation**
Number of Tours	The total number of tours	-0.1786
Total Number of Shows	The total number of shows from all tours	-0.0725
Length of Tours	The average number of shows per tour	0.0908
Average Show Length:	The average show legnth in terms of number of songs	0.0952
H-Index	The careear H-Index, where an artist has a h-index of h if they have played h songs at least h times each	0.0630
Artist Start Date (Groups Only)	The formation incarnation date of the artist (for groups only)	0.0933
Amount Time Period (Groups Only)	The amount time active in terms of years (for groups only)	-0.0952

comparisons between artists. Our approach proposed the notions of a shelf and a tail, to aid explorations of, and to quantify, variety at tour level and artist level. Furthermore, the approach explores the impact of cover songs on these notions of variety, and explores variety comparisons between different musical genres. The analysis of variety highlighted the diversity among artists, in terms of a prevalence to lean towards playing more conformative or more diverse shows. Additional correlation analysis explored the robustness of the proposed notion of variety, with respect to differing artist properties, such as the number of tours or the average lengths of shows.

From our data processing, some data quality issues were uncovered, such as incomplete or empty data, for which such instances can be flagged, and filtering thresholds utilised. Generally, we found more setlist.fm data issues for older tours and shows, and for less popular artists. Additionally, setlist.fm provides set-list data without details on other potential set-list semantics, such as variations in how a song is performed or the inclusion of special elements like artist monologues or other forms of communication during a show. Therefore,future work will explore integrating additional data sources, such as artist fan community databases, to enrich our dataset and model, offering potential for incremental improvements. Additionally, future work will investigate incorporating our analysis into live music recommender systems, which suggest items based on user preferences [4]. Given that factors such as variety and diversity have become increasingly important in this field [18], our analysis may provide valuable insights.

References

[1] E. Abel and A. Goddard. “A Live Concert Performance Recommender System Utilizing User Ideal and Antithesis Ideal Setlist Preferences”. In: *14th International Conference on Smart Computing and Artificial Intelligence.* IEEE Computer Society Press, 2023, pp. 330–335.

[2] E. Abel and A. Goddard. *The Art Behind Bruce Springsteen’s Setlist Composition as Part of His Stagecraft.* 2024.

[3] E. Abel, L. Mikhailov, and J. Keane. “Inconsistency Reduction in decision making via Multi-objective Optimisation”. In: *European Journal of Operational Research* (2017). DOI: 10.1016/j.ejor.2017.11.044. URL: http://linkinghub.elsevier.com/retrieve/pii/S037722171731055X.

[4] M. Aljukhadar, S. Senecal, and C.-E. Daoust. “Using Recommendation Agents to Cope with Information Overload”. In: *International Journal of Electronic Commerce* 17.2 (2012), pp. 41–70. URL: http://www.jstor.org/stable/41739511.

[5] S. Ante. *Why Taylor Swift is the new Grateful Dead.* 2023. URL: https://www.fastcompany.com/90901513/taylor-swift-grateful-dead-cult-brands.

[6] N. Baxter-Moore and T. M. Kitts. “The Live Concert Experience: An Introduction”. In: *Rock Music Studies* 3.1 (2016), pp. 1–4. URL: https://doi.org/10.1080/19401159.2015.1131923.

[7] A. Bullard. *Arctic Monkeys blasted for ‘changing lyrics and rhythm of best known songs’ leaving fans unable to singalong - MyLondon.* 2023. URL: https://www.mylondon.news/whats-on/music-nightlife-news/arctic-monkeys-blasted-changing-lyrics-27104114.

[8] E. C. Callahan and C. Carney. *The politics and power of Bob Dylan’s live performances : play a song for me.* 2024, p. 229. URL: https://www.routledge.com/The-Politics-and-Power-of-Bob-Dylans-Live-Performances-Play-a-Song-for-Me/Callahan-Carney/p/book/9781032315416.

[9] J. R. Castillo and M. J. Flores. “Web-Based Music Genre Classification for Timeline Song Visualization and Analysis”. In: *IEEE Access* 9 (2021), pp. 18801–18816. DOI: 10.1109/access.2021.3053864.

[10] P. Chianca. “Springsteen’s stage success”. In: *Bruce Springsteen and Popular Music.* Routledge, 2018, pp. 178–188. DOI: 10.4324/9781315672144-15/springsteen-stage-success-peter-chianca.

[11] C. Dalla Riva. *Swifties vs. Deadheads: A Meditation on Live Music.* 2023. URL: https://chrisdallariva.substack.com/p/swifties-vs-deadheads-a-meditation?utm%5C%5Fsource=publication-search.

[12] J. D. Evans. *Straightforward Statistics for the Behavioral Sciences.* Brooks/Cole Publishing Company, 1996.

[13] E. Gleadow. *Bob Dylan divides fans by ’doing whatever he wants’ and snubbing hit songs at shows - Mirror Online.* 2024. URL: https://www.mirror.co.uk/3am/celebrity-news/bob-dylan-divides-fans-doing-33095766.

[14] J. E. Hirsch. “An index to quantify an individual’s scientific research output”. In: *Proceedings of the National Academy of Sciences of the United States of America* 102.46 (2005), p. 16569. DOI: 10.1073/pnas.0507655102. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1283832/.

[15] R. Johnston. *How to Stream Bruce Springsteen 2024 Tour Online.* 2024. URL: https://www.billboard.com/culture/product-recommendations/watch-bruce-springsteen-tour-online-streaming-1235669883/.

[16] D. Kreps. *Bruce Springsteen's Poignant Cover of Prince's 'Purple Rain'*. 2016. URL: https://www.rollingstone.com/music/music-news/see-bruce-springsteens-poignant-cover-of-princes-purple-rain-171985/.

[17] A. B. Krueger. *Land of Hope and Dreams: Rock and Roll, Economics and Rebuilding the Middle Class*. Tech. rep. obama whitehouse archives, 2013.

[18] M. Kunaver and T. Požrl. "Diversity in recommender systems – A survey". In: *Knowledge-Based Systems* 123 (2017), pp. 154–162. DOI: 10.1016/j.knosys.2017.02.009.

[19] A. Lerch, C. Arthur, A. Pati, and S. Gururani. "An Interdisciplinary Review of Music Performance Analysis". In: *Transactions of the International Society for Music Information Retrieval* 3.1 (2020), pp. 221–245. DOI: 10.5334/tismir.53. URL: https://transactions.ismir.net/articles/10.5334/tismir.53.

[20] C. Love. *On Repeat. Are artists trotting out the same old set lists gig after gig?* Tech. rep. Medium, 2018. URL: https://databeats.medium.com/on-repeat-70aba1cdc5f8.

[21] B. Mathis-Lilley. "Secrets of the Radiohead Set List". In: *New York Magazine* (2006). URL: https://nymag.com/arts/all/process/17306/.

[22] F. C. Moss, R. Lieck, and M. Rohrmeier. "Computational modeling of interval distributions in tonal space reveals paradigmatic stylistic changes in Western music history". In: *Humanities and Social Sciences Communications* 11.1 (2024), p. 684. URL: https://doi.org/10.1057/s41599-024-03168-1.

[23] M. Mulder and E. Hitters. "Visiting pop concerts and festivals: measuring the value of an integrated live music motivation scale". In: *Cultural Trends* 30.4 (2021), pp. 355–375. URL: https://doi.org/10.1080/09548963.2021.1916738.

[24] T. Murray. *Taylor Swift unexpectedly covers Calvin Harris, Rihanna hit during Liverpool show*. 2024. URL: https://www.independent.co.uk/arts-entertainment/music/news/taylor-swift-this-is-what-you-came-for-eras-tour-b2563075.html.

[25] MusicNews. *Steve Van Zandt Defends Static Bruce Springsteen Setlists*. 2023. URL: https://vermilioncountyfirst.com/2023/03/29/steve-van-zandt-defends-static-bruce-springsteen-setlists/.

[26] M. Pandey. *Crowd-pleasers: The art of choosing the perfect setlist*. 2024. URL: https://www.bbc.com/news/articles/c4nn9expp04o.

[27] D. Pattie. *Rock Music in performance*. Palgrave Macmillan, 2007, pp. 1–188. DOI: 10.1057/9780230593305/cover.

[28] R. Radburn and C. Love. *Digging into concert setlist data: Which artists play the same songs over and over?* Tech. rep. Tableau, 2018. URL: https://www.tableau.com/blog/data-music-which-artists-use-same-old-setlists-gig-after-gig.

[29] A. Rodriguez. *Metallica bases its setlist on what fans listen to on Spotify*. 2018. URL: https://qz.com/1340887/metallica-bases-its-setlist-on-what-fans-listen-to-on-spotify.

[30] M. Rodriguez, V. Gintautas, and A. Pepe. "A Grateful Dead Analysis: The Relationship Between Concert and Listening Behavior". In: *First Monday* 14 (2008). DOI: 10.5210/fm.v14i1.2273.

[31] M. Roffman. *The 25 Best Rock Acts with Unique Setlists.* 2015. URL: https://consequence.net/2016/08/the-25-best-rock-acts-with-unique-setlists/.

[32] D. Silver, M. Lee, and C. C. Childress. "Genre Complexes in Popular Music". In: *Plos One* 11.5 (2016), e0155471. URL: https://doi.org/10.1371/journal.pone.0155471.

[33] F. Thalmann, E. Nakamura, and K. Yoshii. "Tracking The Evolution Of A Band's Live Performances Over Decades". In: *Proc. of the 23rd Int. Society for Music Information Retrieval Conf.* 2022, pp. 850–857. URL: https://madmom.readthedocs.io.

A. Exploration of Shelf and Tail Parameter Values

To aid selection of pertinent *SS* and *TS* parameters, and to aid understanding our of dataset, sensitively analysis experimentation was performed exploring the average percentage of tour songs that are contained within different combined shelf and tail size parameter values. The impact of experimentation with different sized shelf and tail values is shown in Figure 11. The x-axis denotes different combined sizes of shelf and tail values (so 20 represents where the shelf and tail values are both 10) and the y-axis denotes the overall average percentage of songs that are contained within the combined shelf and tail.

Figure 11 highlights how, due to the general trend observed for tours to map out an s shaped sigmoidal like function, there is a pattern that the percentage of songs contained within the shelf and tail is greater than the percentile values denoting the shelf and tail size. From example, a shelf and tail percentile size both of 10% (20% combined value) results in over 50% on average of tours' songs being contained within this 20% percentile space. Moreover, these values denote a point in the plot where the decrease of the gradient of the line is levelling off, suggesting their suitability as shelf and tail percentile sizes.

Further analysis could explore additional experimentation such as, breaking the results down to see the separate contributions of the shelf and tail to the total, using unequal shelf and tail size values, and exploring differences within the results when the data is subsetted for features such as genre.

B. Ordinal Clustering Analysis

From calculations of a pair of $\bar{S}$ and $\bar{T}$ values for every artist, the whole set of artists can be visualised within a single scatter plot, as shown in Figure 7, and from calculations of shelves and tails not including cover songs, the set of artists can be visualised within a single scatter plot, as shown in Figure 8. Further analysis of the data in Figures 7 and 8 can cluster artists into ordinal groups by dividing the 2-dimensional plot space. Variety can be viewed as a combination of levels of shelves and tails, with the plot space divided accordingly, as shown in Figure 12a, here, for average shelf and tail values for each artist. In this plot, the dotted red lines divide the space into 8 levels of variety, representing 8 ordinal clusters. Each artist belongs to only one cluster, as shown by the data point colours in Figure 12a. The membership constraints of each of the 8 clusters can be defined in terms of mean shelf ($\bar{S}$) and mean tail ($\bar{T}$) value ranges, and assigned semantic ordinal names such as:

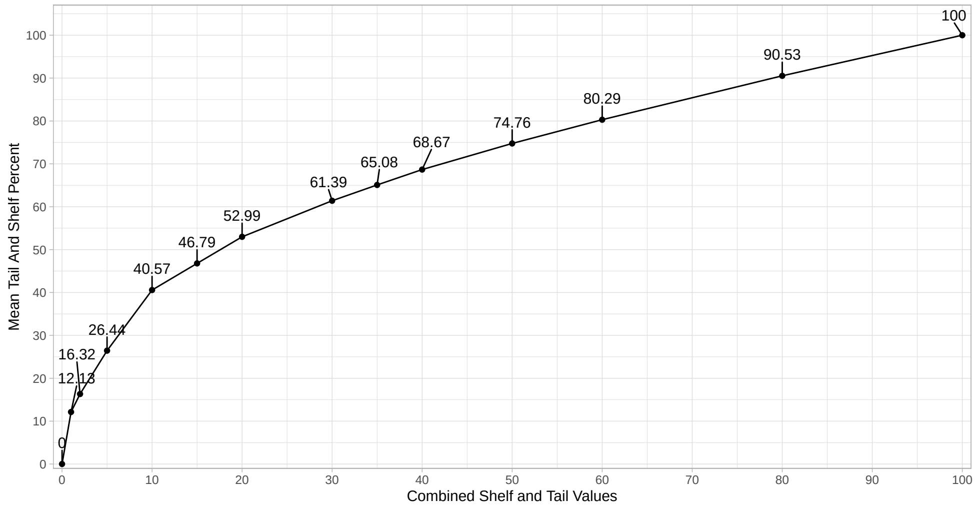
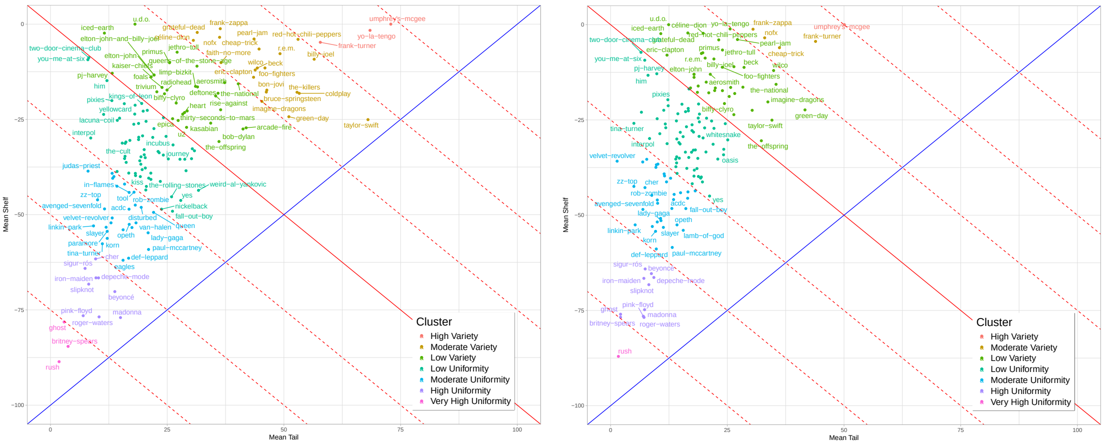


Figure 11: Shelf and Tail Parameter Size Impact on Overall Percentage of Songs Included

1. **Very High Variety**: Where the difference between the mean tail and mean shelf value is greater than or equal to 75.
2. **High Variety**: Where the difference between the mean tail and mean shelf value is greater than or equal to 50 and less than 75.
3. **Moderate Variety**: Where the difference between the mean tail and mean shelf value is greater than or equal to 25 and less than 50.
4. **Low Variety**: Where the difference between the mean tail and mean shelf value is greater than or equal to 0 and less than 25.
5. **Low Uniformity**: Where the difference between the mean tail and mean shelf value is greater than or equal to -25 and less than 0.
6. **Moderate Uniformity**: Where the difference between the mean tail and mean shelf value is greater than or equal to -50 and less than -25.
7. **High Uniformity**: Where the difference between the mean tail and mean shelf value is greater than or equal to -75 and less than -50.
8. **Very High Uniformity**: Where the difference between the mean tail and mean shelf value is less than -75.

Similar analysis can be conducted for data calculations of shelves and tails not including any songs that are cover songs, as shown in Figure 12b. Here, we observe that when covers are excluded, some artists shift in the plot to the extent that they belong to a different cluster, invariably one with less variety. Within such cluster analysis, the % of artists in each cluster can be computed. The breakdown of each cluster artist %, for both data including cover songs, and data not considering cover songs, is shown in Table 4. The Table highlights how when cover songs are not considered, cluster memberships exhibit a general trend for the distribution to move from variety to uniformity.



(a) All Artists Average Tail Vs Average Shelf Cluster Membership

(b) All Artists Average Tail Without Covers Vs Average Shelf Without Covers Cluster Membership

Figure 12: Cluster Membership Analysis

Table 4
Cluster Membership Breakdown Percentages

Cluster Name	Artist % (Including Covers)	Artist % (Not Including Covers)
Very High Variety	0.00	0.00
High Variety	1.85	0.62
Moderate Variety	12.96	2.47
Low Variety	19.14	25.31
Low Uniformity	40.74	43.21
Moderate Uniformity	17.90	21.60
High Uniformity	5.56	6.17
Very High Uniformity	1.85	0.62

C. Correlation Investigation and Discussions of Variety

Our analysis of artist variety explores comparisons of shelves and tails, and our single artist level measure of variety, for comparisons between artists. The robustness of our notion of variety can be explored for its capability to compare different tours and different artists, even though different artists have different career characteristics, in terms of properties such as the number of tours, tour show count sizes, show lengths, being active and touring within different time periods and more. We explore the presence of correlation between our $\bar{V}$ variety measure and such properties, with the Pearson Correlation Coefficient calculated via:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \tag{13}$$

The explored properties and results are shown in Table 5, and visualisation scatter plots of each property against our $\bar{V}$ measure are shown in Figures 13 – 16. Next, each property is outlined and discussed.

Number of Tours: Our dataset includes artists with varying numbers of tours. Figure 13a compares each artist's tour count with their $\bar{V}$, showing no strong relationship between the variables. The Pearson Correlation Coefficient value for these two variables is -0.1786.

Total Number of Shows: For the artists in our dataset, the overall number of shows varies. Comparisons between artists' total show count and their $\bar{V}$ is shown in Figure 13b. Here, the plot highlights there is no strong relationship between these variables; the Pearson Correlation Coefficient value for these two variables is -0.0725.

Length of Tours: Within our dataset, the number of shows within our artists' tours differs. Comparisons between each artist's average tour show count and their $\bar{V}$ is shown in Figure 14a, and the correlation value between these variables is 0.0908. The outlier in the plot is due to the fact Bob Dylan's "Never Ending Tour" is officially billed as a single continuous tour spanning from 1988 to present day and contains around 4000 shows. A plot with Bob Dylan excluded to aid readability is shown in Figure 14b.

Average Show Length: For our set of artists, their average show length, in terms of the number of songs in the set-lists, varies. Comparisons of each artist's average show Length and their $\bar{V}$ is shown in Figure 15a, and the correlation value between these variables is 0.0952.

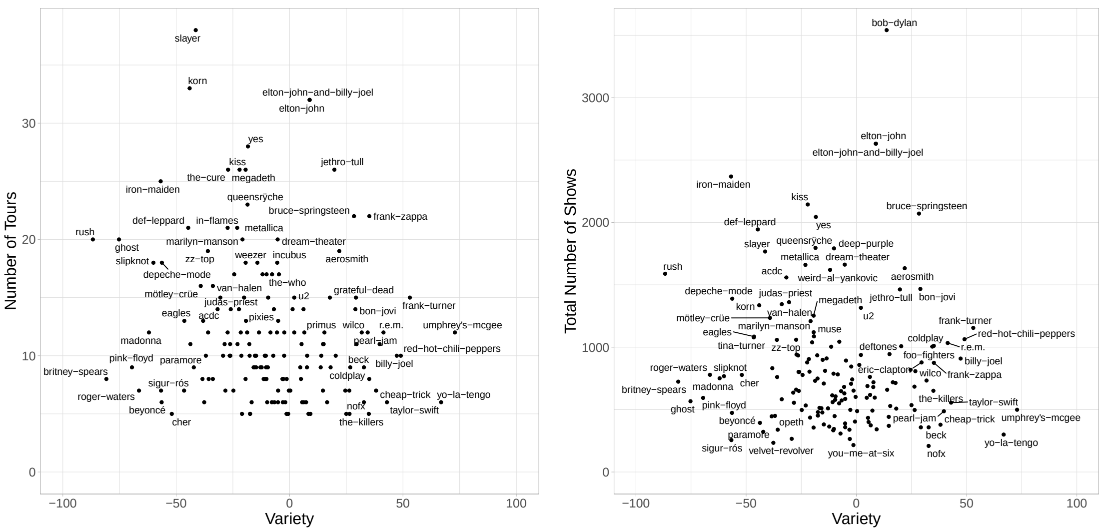
H-Index: The H-index is a scientific output metric that quantifies both the productivity and citation impact of a researcher's publications, where a scholar has an index of h if h of their papers have been cited at least h times each [14]. We could consider this notion within the set-list domain as a measure for an artist (instead of a researcher) and songs (instead of papers), where an artist has a h-index of h if they have played h songs at least h times each. To calculate this for each artist, the set of tours for an artist is combined and overall songs counts calculated from which the H-Index can then be derived from. The plot comparing artists' H-Index against $\bar{V}$ is shown in Figure 15b, and the correlation value between these variables is 0.0630.

Artist Start Dates and Amount of Active Years (Groups Only): For the groups within our artist dataset, information pertaining to the groups formation start date and, if applicable, disbandment date (denoted as present day for still going groups) is retrieved from MusicBrainz. Comparisons of each group's start date and their $\bar{V}$ is shown in Figure 16a, and the correlation value between these variables is 0.0933. From this timeline data, the length of time each group were/have been active for can be calculated (with groups that are still active measured up to present day). Comparisons of each group's Years Active and their $\bar{V}$ is shown in Figure 16b, and the correlation value between these variables is -0.0952.

The set of correlation values for these properties against $\bar{V}$ is shown in Table 5. Considering correlation strength classified as, Very Weak Correlation: $|r| < 0.2$, Weak Correlation: $0.2 \leq |r| < 0.4$, Moderate Correlation: $0.4 \leq |r| < 0.6$, Strong Correlation: $0.6 \leq |r| < 0.8$, Very Strong Correlation: $|r| \geq 0.8$ [12], Table 5 highlights how our $\bar{V}$ variety measure has only very weak correlation to all these properties. This suggests the measure is robust across the varying properties for our artists, and that comparisons between artists are not unduly impacted by these variations.

Table 5
Correlation Data Analysis

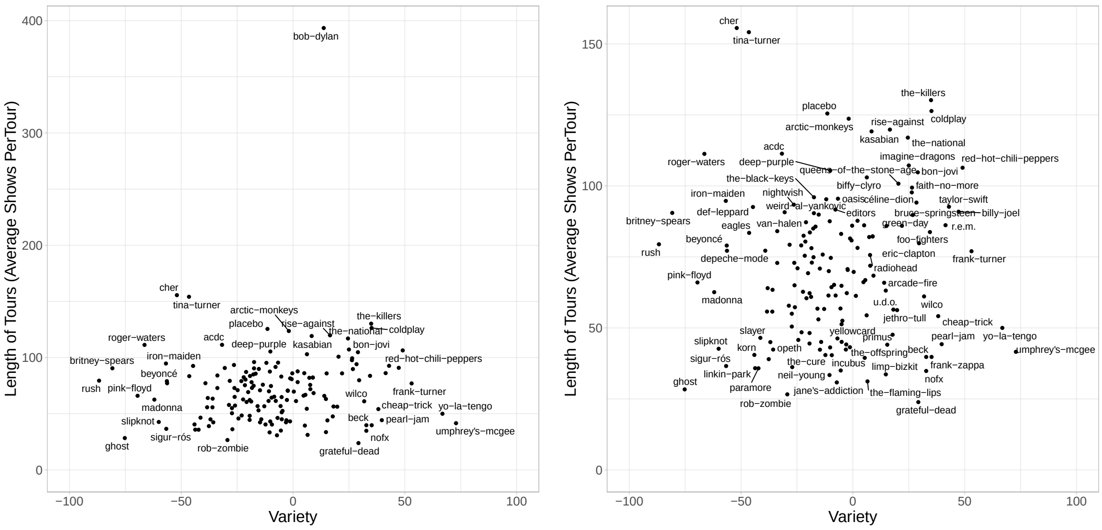
Property Name	Correlation	Min Value	Max Value	Standard Deviation
Number of Tours	-0.1786	5	38	6.2613
Total Number of Shows	-0.0725	209	3541	518.9224
Length of Tours	0.0908	23.87	393.4	35.7931
Average Show Length	0.0952	12.87	57.52	5.7834
H-Index	0.0630	28	126	15.6733
Artist Start Date (Groups Only)	0.0933	1962-01-01	2011-01-01	–
Active Time in years (Groups Only)	-0.0952	5.999	62.52	11.6033



(a) Number of Tours Vs $\bar{V}$

(b) Total Number of Shows Vs $\bar{V}$

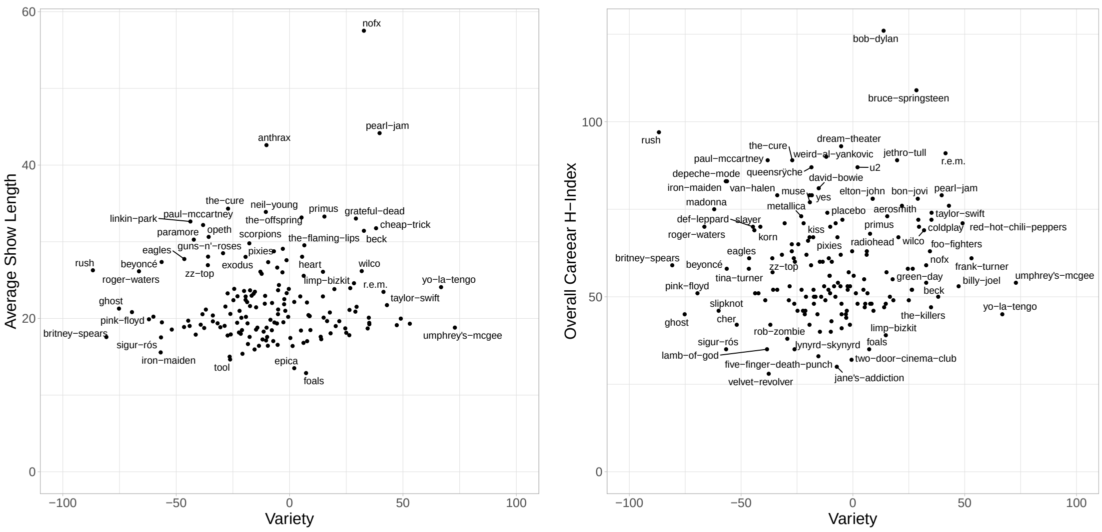
Figure 13: Number of Tours and Total Number of Shows Correlation Analysis



(a) Length of Tours Vs $\bar{V}$

(b) Length of Tours Vs $\bar{V}$ (sans Bob Dylan)

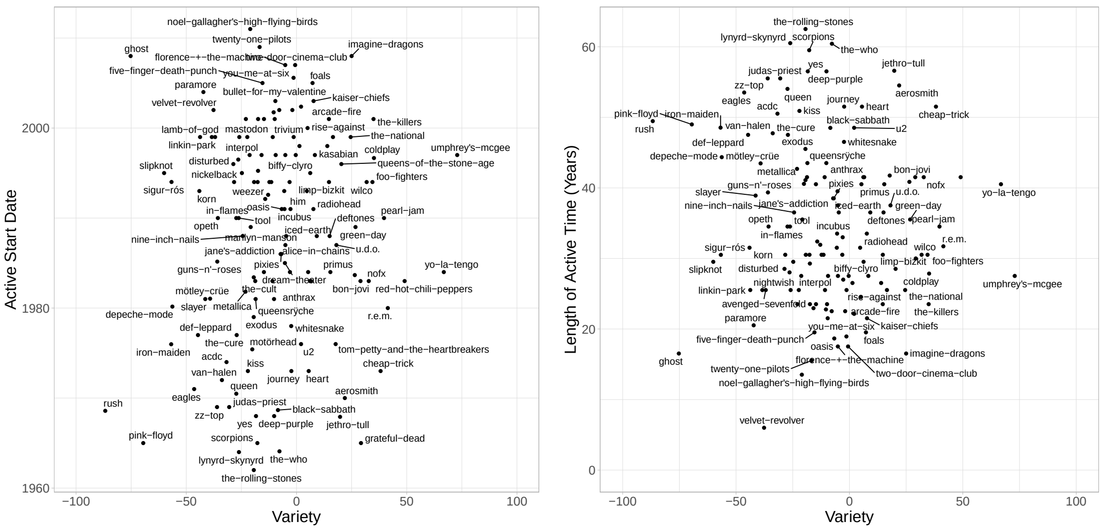
Figure 14: Length of Tours Correlation Analysis



(a) Average Show Length Vs $\bar{V}$

(b) H-Index Vs $\bar{V}$

Figure 15: Average Show Length and H-Index Correlation Analysis



(a) Artist Start Date (Groups Only) Vs $\bar{V}$

(b) Artist Active Years (Groups Only) Vs $\bar{V}$

Figure 16: Start Date and Active Years Correlation Analysis (Groups Only)