



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO

**Modelagem de Aprendizado de Máquina para Compreensão de Fases Musicais: Um Caso de
Estudo com The Beatles**

Paulo Henrique Silva

Belo Horizonte
2023

Paulo Henrique Silva

**Modelagem de Aprendizado de Máquina para Compreensão de Fases Musicais: Um Caso de
Estudo com The Beatles**

Trabalho de Conclusão de Curso apresentado ao curso de Engenharia de Produção da Universidade Federal de Minas Gerais, como requisito à obtenção do grau de bacharel em Engenharia de Produção.

Orientador:

Martín Gomez Ravetti

Banca avaliadora:

Thamara Paula dos Santos Dias

Tiago Alves Schieber de Jesus

Belo Horizonte
2023

AGRADECIMENTOS

Minha sincera gratidão a todos que contribuíram para a construção deste trabalho: minha família, namorada, amigos e professores.

Gratidão ao Martín Gomez Ravetti, por toda orientação e por ter acreditado em mim.

Obrigado aos avaliadores Thamara Paula dos Santos Dias e Tiago Alves Schieber de Jesus, pelas críticas e sugestões.

Meu mundo não estava na sala de aula, mas eu aprendi várias coisas nos últimos cinco anos (e meio). Obrigado a todos os meus heróis nos discos e nos *CDs*.

“Os grandes podem ter partido, mas a grandeza deles vive para sempre” - Stevie Ray Vaughan.

RESUMO

Neste trabalho, é abordada a viabilidade de se traduzir categorizações subjetivas relacionadas à experiência musical em valores objetivos, em meio à complexidade crescente do cenário musical contemporâneo. Este panorama é marcado pela ascensão de plataformas digitais, indexadores musicais e serviços de *streaming*. Em um contexto em que a quantidade de canções cadastradas diariamente na *web* atinge níveis extraordinários, torna-se crucial aprimorar as técnicas de filtragem e personalização na entrega de músicas aos usuários. Para testar as hipóteses, foi desenvolvido um modelo de regressão linear multinomial com o intuito de compreender as distintas fases das obras de uma banda. *The Beatles* foi utilizado como caso de estudo. A implementação prática desse modelo permitiu a aplicação do problema em um contexto real, abordando as implicações técnicas e os desafios envolvidos na utilização da tecnologia para análise e classificação de conteúdo artístico. Por fim, concluiu-se que, apesar dos desafios inerentes à análise, os resultados obtidos sugerem aplicabilidade prática. Os p-valores e a comparação com o modelo nulo indicaram a presença de padrões objetivos relacionados às características musicais subjetivas analisadas.

Palavras-chave: Aprendizado de Máquina. Arte. Estatística. Música.

ABSTRACT

In this work, the feasibility of translating subjective categorizations related to musical experience into objective values is addressed, amid the growing complexity of the contemporary musical scene. This panorama is marked by the rise of digital platforms, music indexers and streaming services. In a context where the number of songs registered daily on the web reaches extraordinary levels, it is crucial to improve filtering and personalization techniques when delivering music to users. To test the hypotheses, a multinomial linear regression model was developed with the aim of understanding the different phases of a band's work. The Beatles was used as a case study. The practical implementation of this model allowed the application of the problem in a real context, addressing the technical implications and challenges involved in using technology to analyze and classify artistic content. Finally, it was concluded that, despite the challenges inherent to the analysis, the results obtained suggest practical applicability. The p-values and the comparison with the null model indicated the presence of objective patterns related to the subjective musical characteristics analyzed.

Keywords: Machine Learning. Art. Statistics. Music.

LISTA DE ABREVIATURAS E SIGLAS

Siglas

MIR	<i>Music Information Retrieval</i>
API	<i>Application Program Interface</i>
AUC	<i>Area Under Curve</i>
ROC	<i>Receiver Operating Characteristic</i>
VIF	<i>Variance Inflation Factor</i>
TP	<i>True Positive</i>
FP	<i>False Positive</i>
TN	<i>True Negative</i>
FN	<i>False Negative</i>
IIA	<i>Independence of Irrelevant Alternatives</i>
PCA	<i>Principal Component Analysis</i>
NLP	<i>Natural Language Processing</i>
OR	<i>Odd Ratio</i>
RLM	<i>Regressão Logística Multinomial</i>
KNN	<i>K-Nearest Neighbors</i>

SUMÁRIO

1 Introdução	6
1.1 Contexto	6
1.2 Justificativa	6
1.3 Objetivos	6
2 Referencial Teórico	7
2.1 Exploring different approaches for music genre classification	7
2.2 Predicting the Geographical Origin of Music	8
2.3 Difficult to define, easy to understand: The use of genre categories while talking about music	9
2.4 Music genre classification using audio features and lyrics sentiment features with multinomial logistic regression	10
2.5 Definition of MIR methodology	10
3 Estudo de Caso	11
3.1 Estudo de Informações Musicais	11
3.2 Escolha do Objeto de Estudo	13
3.2.1 A Evolução das Fases dos Beatles	13
3.2.1.1 Era “Beatlemania”	13
3.2.1.2 Era “Transition”	14
3.2.1.3 Era “Psychedelic”	14
3.2.1.4 Era “Dissolution”	14
3.3 Escolha do Método Computacional	15
3.3.1 Pressupostos para Regressão Logística Multinomial	16
3.3.1.1 Ausência de Multicolinearidade	16
3.3.1.2 Independência das Alternativas Irrelevantes (Teste Hausman-McFadden)	17
3.3.2 Métricas para Regressão Logística Multinomial	18
3.3.2.1 Matriz de Confusão	18
3.3.2.1 ROC e AUC	19
4 Metodologia	20
4.1 Coleta, Limpeza e Transformação de Dados	20
4.1.1 Complexidade dos Acordes	23
4.1.2 Análise Sentimental da Letra	24
4.2 Análise Exploratória	25
4.2.1 Distribuição da Variável Resposta	25
4.2.2 Distribuição das Variáveis Independentes	25
4.2.3 Multicolinearidade	26
4.2.4 Visualização das Variáveis Categóricas	28
4.2.5 Análise de Tendências Temporais	29
4.2.6 Testes de Pressupostos	30
4.2.6.1 Testes Individuais de AUC	30
4.2.6.2 Variável Dependente Nominal e Categorias Mutuamente Exclusivas	30
4.2.6.3 Independência das Observações (Sem Repetições)	30
4.2.6.4 Independência das Alternativas Irrelevantes (Teste Hausman-McFadden)	

5 Resultados	31
5.1 Colunas de Dados	31
5.2 Formulação Matemática	33
5.3 Comparação com o Modelo Nulo	33
5.4 Pseudo- R^2	34
5.5 P-Valores	34
5.6 Matriz de Confusão	35
5.7 Curva ROC e AUC	35
5.8 Validação K-Fold para AUC	36
5.9 Validação Leave-One-Out para Acurácia	37
5.10 Interpretação dos Coeficientes	37
6 Considerações Finais	38
6.1 Discussão dos Resultados	38
6.2 Trabalhos Futuros	40
6.3 Conclusão	40
7 Referências Bibliográficas	40
8 Anexos	46
8.1 Importância Cultural e Histórica dos Beatles	46
8.2 Estudos Sobre Características Musicais	50
8.2.1 Acorde	50
8.2.2 Tonalidade	51
8.2.3 Modo	52
8.2.4 Mudança de Tonalidade	52
8.2.5 Fórmula de Compasso	52
8.2.6 Dançabilidade	53
8.2.7 Energia	53
8.2.8 Acústicidade	54
8.2.9 Valência Emocional	54
8.2.10 Instrumentalidade	55
8.2.11 Vivacidade	55
8.2.12 Volume	56
8.2.13 Fala	59
8.2.14 Tempo	59
8.3 Efeitos Específicos	60

1 INTRODUÇÃO

1.1 CONTEXTO

Em um cenário marcado pela ascensão de plataformas digitais, indexadores musicais e serviços de *streaming*, a quantidade de canções cadastradas diariamente na *web* atinge patamares extraordinários. De acordo com o CEO do Spotify, 40.000 faixas são carregadas em sua plataforma todos os dias (BEAUMONT, 2020). Tais canções possuem uma gigantesca diversidade em gêneros, temas e inspirações. Ao proporcionar uma ampla gama de possibilidades aos ouvintes, nota-se o surgimento de grandes desafios ao segmento.

1.2 JUSTIFICATIVA

Essas tecnologias, fundamentadas na personalização, possuem sistemas de sugestão que dependem da capacidade de se traduzir, de maneira objetiva, as nuances da experiência musical, observando padrões de gostos do usuário e sugerindo canções ajustadas às suas preferências individuais. Nesse sentido, as demandas do mercado por informações musicais objetivas tornam-se cada vez mais presentes.

Nesse contexto, a dualidade entre o subjetivo e o objetivo adquire uma complexidade renovada. A apreciação musical, que é intrinsecamente subjetiva, necessita ser convertida em métricas objetivas para atender aos requisitos dessas plataformas e serviços pois, através de filtros, as plataformas buscam proporcionar a entrega de músicas que se alinhem precisamente com as preferências e expectativas do usuário.

A natureza subjetiva da apreciação e categorização musical, em contraste com a necessidade crescente de objetividade em avaliações musicais contemporâneas, define o contexto em que este trabalho se insere. Construir pontes entre a subjetividade da experiência musical e a exigência de métricas mensuráveis torna-se, então, essencial para atender à integração dessas informações nos sistemas digitais.

1.3 OBJETIVOS

Investigar e testar a possibilidade de traduzir categorizações subjetivas, relacionadas à experiência musical e à classificação subjetiva, em valores objetivos, utilizando técnicas computacionais.

2 REFERENCIAL TEÓRICO

2.1 EXPLORING DIFFERENT APPROACHES FOR MUSIC GENRE CLASSIFICATION

Goulart, Guido e Maciel (2012) realizaram investigações no campo da classificação de gêneros musicais, buscando estabelecer uma abordagem que permitisse uma categorização objetiva.

Foram adotadas cinco características previsoras, extraídas de arquivos digitais. Na primeira fase dos testes, cada música foi dividida em quadros de 1024 amostras (*frames*) e as colunas inseridas no modelo foram compostas por média de entropia, desvio padrão, entropia máxima, entropia mínima e a diferença máxima de entropia entre quadros consecutivos.

Para cada *frame*, foi calculada a entropia por meio da abordagem de energia de Cover e Thomas (2006), utilizando-se de

$$- \sum_{i=0}^{1023} p_i \log_2(p_i) \quad (1),$$

sendo que p_i é a proporção da energia total do sinal, ou seja, a energia do quadro dividida pela energia do sinal completo. A argumentação para se utilizar esse critério foi porque ele mostrou-se mais estável do que abordagens simples de amplitude e frequência.

Para todos os testes, foram usados 90 exemplos de músicas igualmente divididos em três gêneros distintos: *blues*, clássico e *lounge*, e a etapa de classificação foi realizada usando *SVMs* combinados, que se trata de uma técnica de aprendizado de máquina que busca otimizar a classificação por meio da identificação de fronteiras de decisão eficazes. Os autores argumentam que as *SVMs* combinadas foram escolhidas devido à sua capacidade de lidar eficientemente com conjuntos de dados complexos e multidimensionais. Os testes foram conduzidos dividindo as músicas em estágios de treinamento e teste, variando a porcentagem de músicas usadas para treinamento.

Embora o classificador tenha atingido uma precisão máxima de 100% com a utilização de 80% do banco de dados total para treinamento, conjuntos modestos de dados de treinamento (9 músicas) foram suficientes para alcançar precisões variando de 51,8% a 92,6%, demonstrando boa capacidade de generalização.

2.2 PREDICTING THE GEOGRAPHICAL ORIGIN OF MUSIC

Zhou, Claire e King (2014) descrevem um estudo no campo da Etnomusicologia Geográfica, que busca compreender a distribuição geográfica da música ao analisar sua origem em diferentes partes do mundo. Eles também propõem uma abordagem de aprendizado de máquina, onde um programa é treinado para prever a origem geográfica da música com base em seu conteúdo auditivo.

Mantemos a posição filosófica de que não compreendemos completamente um fenômeno a menos que possamos construir uma máquina que o reproduza. A vantagem desta abordagem para a compreensão de uma área temática é que ela é objetiva e operacional. Esta abordagem contrasta notavelmente com a abordagem tradicional, que quase sempre se baseou no julgamento subjetivo dos críticos humanos. Muitas vezes, esta abordagem subjetiva obtém uma grande compreensão, mas também é preciso admitir que existem limitações nos resultados que dependem das peculiaridades do ouvinte (ZHOU; CLAIRE; KING, 2014). Traduzido por mim.

Ao invés de tratar o problema como uma classificação (prever um país ou área), transformou-se em um problema de regressão, buscando prever coordenadas esféricas, latitude e longitude. Isso é feito pela quantidade limitada de exemplos de alguns países, o que levaria a um desequilíbrio em uma classificação.

Os métodos utilizados incluem "*K-Nearest Neighbors*" (KNN) e "*Random Forest Regression*" (RFR) com um conjunto de 1059 faixas de 73 países/áreas, e foram utilizadas características musicais extraídas por meio do programa MARSYAS. Estas características incluem informações como o centróide espectral, que está relacionado à percepção humana de "brilho" do som, e atributos cromáticos, que descrevem as notas da escala utilizada na música. Além disso, são mencionados 68 atributos padrão fornecidos pelo MARSYAS.

O melhor desempenho preditivo alcançado teve uma distância média de 3,113 quilômetros, todavia o algoritmo teve desempenho variado em diferentes países/áreas, mesmo quando o número de faixas disponíveis era o mesmo. Isso sugere que a música de alguns países/áreas é mais distinta do que a de outros. Um problema adicional é o tamanho relativo dos países/áreas. Por exemplo, para Taiwan, a maioria das músicas foi corretamente posicionada, com uma distância máxima de 7.539,853 km e uma média de 1.019,651 km. Já para a Grécia,

houve uma distribuição mais ampla, destacando que a faixa de previsão para um país/área pode variar, com algumas previsões mais precisas e outras mais dispersas.

Em geral, os resultados indicaram que há uma relação preditiva entre os atributos musicais e as origens geográficas. Os autores utilizam o argumento de que os métodos de aprendizado de máquina superam significativamente modelos aleatórios, para demonstrar a eficácia do método. Também dizem que seria interessante comparar os resultados dos programas de aprendizado de máquina com o desempenho humano na previsão da origem musical, afirmando que suspeitam que os métodos de aprendizado de máquina já são bastante competitivos com os humanos para essa identificação.

2.3 DIFFICULT TO DEFINE, EASY TO UNDERSTAND: THE USE OF GENRE CATEGORIES WHILE TALKING ABOUT MUSIC

Em Cerati (2021), há a utilização do termo “compreensão mútua”. Este é apresentado como um elemento fundamental na interação social, especialmente quando se trata do uso de categorias de gênero musical. Essa compreensão é baseada na reciprocidade de perspectivas, o que significa que os participantes assumem que compartilham uma visão semelhante da realidade musical. A ideia é que, ao usar categorias como "jazz" ou "rock", por exemplo, os interlocutores confiam na suposição de que ambos têm uma compreensão compartilhada desses termos.

Entretanto, o texto destaca que essa reciprocidade de perspectivas pode ser desafiada devido à distribuição desigual do conhecimento musical entre os participantes, especialmente quando pertencem a diferentes comunidades de prática. Cada comunidade pode ter suas próprias nuances, entendimentos e preferências em relação aos gêneros musicais, o que pode levar a mal-entendidos ou dificuldades na comunicação.

Essa desigualdade na distribuição do conhecimento musical é apresentada como um fator que pode complicar a compreensão mútua, pois os interlocutores podem não compartilhar o mesmo repertório de conhecimento sobre determinados gêneros musicais. Portanto, a reciprocidade de perspectivas, embora seja uma base para a compreensão, não é garantida devido às diferenças nas experiências e conhecimentos individuais, criando assim desafios na comunicação musical.

2.4 MUSIC GENRE CLASSIFICATION USING AUDIO FEATURES AND LYRICS SENTIMENT FEATURES WITH MULTINOMIAL LOGISTIC REGRESSION

O texto de Bal (2018) introduz um estudo sobre a classificação de gêneros musicais com base em características de áudio e sentimentos extraídos das letras das músicas. A pesquisa propõe utilizar um conjunto expandido de dados de sentimento, denominado ANEW, de Bradley (1999) em comparação com estudos anteriores, buscando demonstrar o valor adicionado das características sentimentais das letras na classificação de gênero musical.

O método envolveu a revisão da literatura para identificar as melhores características de áudio, a coleta de dados da *Million Song Dataset* (<http://millionsongdataset.com/>) e de datasets de gênero, letras em formato *bag-of-words* (formato que não leva em conta a ordem ou a estrutura das palavras no texto, apenas a frequência com que aparecem nele) e dados de sentimento. A análise foi realizada por meio de modelos de regressão logística multinomial, que utilizaram as 94,482 linhas e 34 colunas de características como fonte de dados.

O modelo que incorporou tanto características de áudio quanto de sentimentos nas letras alcançou o melhor desempenho, com uma precisão geral de 41,5%. O modelo contendo apenas características de sentimentos nas letras teve um desempenho muito próximo, com uma pontuação de precisão de 39,6%.

Ao analisar a matriz de confusão, observou-se que algumas classes foram previstas com maior precisão do que outras. A análise revelou que todas as classes eram frequentemente erroneamente classificadas como “Pop Rock”, a classe majoritária. Também surgiram resultados interessantes ao examinar as características específicas que levaram à classificação incorreta de determinados gêneros. Por exemplo, “Latin” foi, em alguns casos, classificado como “Rap” devido a semelhanças em sua pontuação de valência emocional.

2.5 DEFINITION OF *MIR* METHODOLOGY

Através de Kaminskas e Ricci (2012), é possível compreender que a Recuperação de Informações Musicais (*MIR*) é uma área de pesquisa que utiliza técnicas de processamento de sinais de áudio e aprendizado de máquina para extrair informações relevantes de conteúdo musical. Essa disciplina faz parte de um campo mais amplo conhecido como Recuperação de Informações Multimídia, que se concentra em extrair informações de diferentes tipos de mídia, como imagens, vídeo e som.

Os pesquisadores se concentram em diferentes níveis de análise do sinal de áudio para extrair características significativas da música. Esses níveis incluem abordagens que envolvem transcrever notas ou acordes individuais na música. Essa abordagem pode ser útil para a detecção de instrumentos em uma música. Também inclui análises que se estendem a sequências de notas para identificar periodicidades, como detecção de tempo, sequenciamento de playlist ou resumo musical. Também aborda a subjetividade na representação mais abstrata da música, mais próxima da percepção do usuário, que pode ser usada para tarefas como detecção de gênero musical ou modelagem de preferências do usuário em sistemas de recomendação de música.

Kaminskas, Ricci (2012) destaca que características comuns usadas na *MIR* incluem o timbre, que refere-se às características perceptuais de um som que distinguem diferentes tipos de produção de som, como vozes ou instrumentos musicais e que muitas vezes envolve a distribuição espectral do sinal de áudio. Também, a melodia e a harmonia, as quais envolvem a análise da distribuição de altura das notas em segmentos do sinal de áudio. Também nota-se o uso do ritmo, que está relacionado à regularidade temporal de uma peça musical. A informação de ritmo é extraída pela análise das periodicidades dos batimentos no sinal de áudio.

Há um interesse crescente em abordagens contextuais na *MIR*, considerando fatores como o estado emocional do usuário, localização e atividade atual para personalizar ainda mais as recomendações musicais. Essas abordagens estão em desenvolvimento, e os desafios incluem a heterogeneidade dos dados musicais e os custos associados à avaliação de sistemas contextuais.

3 ESTUDO DE CASO

Com base na proposta de realizar uma tradução objetiva de critérios subjetivos através de métodos computacionais, optou-se pela realização de uma aplicação prática própria, para servir como estudo de caso.

3.1 ESTUDO DE INFORMAÇÕES MUSICAIS

Visando a aplicação do problema, foram estudadas diversas características de informações musicais abordadas pela *MIR*. Na tabela 1 é apresentado um resumo dos estudos e no anexo 1 estão presentes as informações de forma completa.

Tabela 1 - Características musicais relevantes ao estudo

Característica	Descrição
Análise Sentimental da Letra	Avalia a polaridade emocional da letra da música, geralmente representada por valores de Negatividade, Neutralidade e Positividade.
Tonalidade musical	Indica a escala musical central da música, fornecendo uma referência fundamental para a harmonia.
Modo musical	Define a atmosfera geral da música, indicando se ela tem uma sensação mais alegre (maior) ou melancólica (menor).
Fórmula de Compasso	Representa a estrutura rítmica da música, indicando quantas batidas estão presentes em cada medida (geralmente, 3 ou 4).
Dançabilidade musical	Reflete a adequação da música para a dança, com valores mais altos indicando maior facilidade para dançar.
Energia musical	Mede a intensidade e atividade da música, indicando se ela é mais calma ou mais enérgica.
Acústicidade musical	Indica o quão acústica é a música, se ela é mais voltada para instrumentos acústicos ou se contém elementos eletrônicos.
Valência musical (<i>valence</i>)	Representa a positividade ou negatividade da música, indicando se tem uma atmosfera mais feliz ou triste.
Instrumentalidade musical	Indica a probabilidade de a música ser instrumental, ou seja, sem vocais.
Vivacidade musical (<i>liveness</i>)	Refere-se à presença de elementos ao vivo em uma gravação, indicando se a música foi gravada em um ambiente ao vivo.
Volume	Mede a intensidade do som da música em decibéis, indicando se ela é mais suave ou mais alta.
Falas na música (<i>speechiness</i>)	Indica a probabilidade de haver elementos de fala na música, como diálogo falado ou palavras recitadas.
Tempo (bpm)	Define o ritmo da música, indicando quantas batidas ocorrem por minuto.
Acorde	Refere-se a combinações de notas musicais tocadas simultaneamente, formando a base harmônica da música.
Complexidade dos Acorde	Avalia a sofisticação e diversidade dos acordes na música, indicando o quão intrincada é a estrutura harmônica.

Fonte: Autoria própria.

3.2 ESCOLHA DO OBJETO DE ESTUDO

Para tornar a abordagem aplicável, optou-se por selecionar um contexto musical específico. Devido à grande quantidade de informação disponível na internet e a relevância do grupo, *The Beatles* foi escolhido como objeto de estudo para uma avaliação prática do problema. A partir de estudos, foi possível observar a questão das suas fases musicais como uma boa oportunidade de aplicação, visto que tais fases são marcantes e distintas de forma que os fãs, a partir de uma noção subjetiva, as identifiquem com facilidade.

No subtópico 3.2.1 encontram-se estudos relacionados à banda e características musicais de cada uma delas.

3.2.1 A Evolução das Fases dos Beatles

A trajetória do grupo pode ser dividida em quatro eras distintas que marcaram sua carreira e evolução musical ao longo dos anos.

A música e as letras dos Beatles mudaram drasticamente do primeiro ao último álbum devido à adaptação a uma sociedade em constante mudança, à maturidade musical e ao aumento de sua popularidade. (GUTZMER, 2019). Tradução minha.

3.2.1.1 Era “Beatlemania”

A primeira fase, conhecida como “Beatlemania”, ocorreu no início dos anos 60. Nessa época, os Beatles eram uma sensação global, com suas músicas alegres e cativantes, como “*I Want to Hold Your Hand*” e “*She Loves You*”.

[...] Uma de suas primeiras músicas de sucesso, “*I Want to Hold Your Hand*”, [...] fala sobre um amor simples e inocente, que pode ser percebido nas frases “*Oh yeah, I’ll tell you somethin’ / I think you will understand / When I say that somethin’ / I want to hold your hand*” e “*And when I touch you / I feel happy inside*”. Essas letras são acompanhadas por uma sensação animada e alegre transmitida pela guitarra e pelas harmonias vocais, destacando a inocência e euforia da música. A identidade dos Beatles nessa época era a de um grupo alegre e inocente que qualquer um poderia apreciar. (GUTZMER, 2019). Traduzido por mim.

3.2.1.2 Era “Transition”

Na segunda fase, denominada "Período de Transição", aconteceu no final dos anos 60, a música do grupo começou a se tornar mais complexa e experimental, incorporando influências psicodélicas e introspectivas (GUTZMER, 2019). Álbuns como *Rubber Soul* e *Revolver* marcaram uma mudança significativa em seu som, apresentando letras mais sofisticadas e uma abordagem inovadora na produção musical. À medida que os *Beatles* se tornaram mais populares, eles passaram para um lado de ambiguidade e lirismo que fazia com que o ouvinte tivesse que prestar atenção nas letras para encontrar o significado. (GUTZMER, 2019).

3.2.1.3 Era “Psychedelic”

Na terceira fase, conhecida como "Período Psicodélico", houve um grande salto da banda no que tange a complexidade lírica e experimentação, como vemos em Gallaher (2020):

As diferenças de som, produção, compilação e apresentação entre os álbuns que cercam *Help!* (1965) e *Sgt. Pepper's Lonely Hearts Club Band* (1967), são impressionantes; os dois álbuns, embora separados por apenas dois anos, parecem ter sido criados por bandas completamente diferentes.

Os *Beatles* mergulharam de cabeça na experimentação sonora, utilizando técnicas de estúdio avançadas, letras surrealistas e uma mistura única de estilos musicais (GALLAHER, 2020).

3.2.1.4 Era “Dissolution”

A quarta e última fase, chamada de "Período de Dissolução", ocorreu no final dos anos 60 e início dos anos 70. Durante esse período, surgiram tensões internas na banda, refletidas em seus álbuns, como *The White Album* e *Let It Be*. Esses álbuns mostraram uma variedade de estilos musicais e composições individuais dos membros da banda, indicando a crescente separação criativa entre eles. Blackman (2020), que discute *The White Album* e destaca como as letras representam uma banda altamente individualizada que utilizava as letras como uma forma de contar histórias pessoais e transmitir mensagens ao público:

[...] as letras das músicas deste álbum representam uma banda altamente individualizada que se baseava nas letras como forma de contar histórias pessoais e transmitir mensagens ao público. A

dependência do amor e das mulheres como tema das músicas não é mais evidente neste álbum [...]. Além disso, após minha análise, constatei que, em média, a palavra "amor" é usada 0,97 vezes por música [...]. Tradução própria.

Ainda sobre a individualização da banda, nos últimos álbuns dos *Beatles*, um clima de separação começou a se manifestar. A dinâmica interna da banda estava sofrendo tensões crescentes, refletidas nas composições e nas relações pessoais. *Let It Be* e *Abbey Road* apresentaram um senso de despedida e encerramento, com músicas que expressavam melancolia e reflexão sobre o fim iminente. Esses álbuns capturaram a essência de um período turbulento e marcaram o fim de uma era para os *Beatles*, enquanto cada membro se encaminha para projetos musicais individuais. Fala de John Lennon acerca de *Let It Be*:

“Havia muita tensão durante aquele álbum. Estávamos prestes a nos separar, e isso por si só já era tenso”. (DOGETT, 2009).

Lennon disse que

“a separação dos Beatles pode ser ouvida naquele álbum”. (CARR, 2009).

3.3 ESCOLHA DO MÉTODO COMPUTACIONAL

A partir do conhecimento das características musicais envolvidas no problema e da noção da relevância de tais variáveis nas eras do grupo, foi necessário escolher um método computacional que possibilitasse a aplicação do problema.

Por meio do referencial teórico, foi possível identificar diversos métodos e tecnologias que possibilitam a interpretação de variáveis musicais através de um modelo computacional. Foram citados *Random Forest Classifier*, Redes Neurais, *Support Vector Machine (SVM)*, as Regressões Linear e Logística, entre outros.

Dentre estas, a Regressão Logística Multinomial foi escolhida por sua modelagem matemática, que possibilita a obtenção de valores de relevância estatística de forma facilitada, melhorando a interpretabilidade dos resultados fornecidos pelo modelo. Também, a Regressão Logística é menos propensa a sobreajustar-se em situações de alta dimensionalidade, pelo fato de que seus sobreajustes ocorrem apenas em casos de correlação linear entre as variáveis.

Métodos mais complexos, como Redes Neurais e *KNN*, se sobreajustam, também, com outros tipos de padrões.

Além disso, os resultados fornecidos por esse método são fornecidos em valores de probabilidade, e não em valores absolutos, o que permite gerar gráficos de probabilidade e observar mais intuitivamente a influência de cada atributo nas classificações.

3.3.1 Pressupostos para Regressão Logística Multinomial

A partir das definições de tema de estudo e técnica computacional, iniciaram-se os estudos relacionados aos pressupostos para aplicação da técnica.

3.3.1.1 Ausência de Multicolinearidade

A multicolinearidade é um fenômeno que ocorre quando duas ou mais variáveis independentes em um modelo de regressão estão altamente correlacionadas entre si (SCHAEFER, 1986). Isso significa que uma variável poderia ser prevista com alto grau de precisão a partir de uma combinação linear das outras variáveis independentes, podendo comprometer a interpretabilidade dos coeficientes, a estabilidade das estimativas e a capacidade do modelo de fazer previsões precisas (SCHAEFER, 1986).

Schaefer (1984) propôs o uso do coeficiente de determinação da regressão linear do modelo que tenta prever o valor da *j*-ésima covariável em relação às demais, que ele denomina R_j^2 . Seus resultados pertencem a tal que $0 \leq R_j^2 \leq 1$ e valores próximos de 1 indicam alta multicolinearidade. Dessa forma, se houver alguma relação linear entre quaisquer covariáveis, quando ajustamos um modelo para uma delas em relação às demais, a previsão e o valor observado terão valores muito próximos. Não nos interessa o sinal da associação, logo utiliza-se R_j^2 .

O VIF, estimador de multicolinearidade para a *i*-ésima variável independente, é calculado da seguinte forma:

$$VIF_j = \frac{1}{1 - R_j^2} \quad (3)$$

Na presença de multicolinearidade (bom ajuste no modelo que tenta prever o valor da *j*-ésima covariável em relação às demais), o denominador será muito próximo de zero, o que

inflaciona a variância dos estimadores dos parâmetros do modelo. Montgomery (2001) estabeleceu níveis de multicolinearidade com base no valor do VIF, como demonstrado:

Tabela 2 - Níveis de multicolinearidade

Valor do VIF	Nível de Multicolinearidade
Menor que 1	Baixa
Entre 1 e 5	Moderada
Maior que 5	Alta

3.3.1.2 Independência das Alternativas Irrelevantes (Teste Hausman-McFadden)

O pressuposto de IIA postula que a razão entre as probabilidades de escolher duas categorias em relação a uma categoria de referência não é afetada pela introdução de uma terceira categoria (VIJVERBERG, 2011). Matematicamente, o pressuposto de IIA pode ser expresso da seguinte maneira.

$$\frac{P(Y_i=j|X)}{P(Y_i=k|X)} = \frac{P(Y_i=j|X, Y_i \neq k)}{P(Y_i=k|X, Y_i \neq j)} \quad (2),$$

onde:

- $P(Y_i=j|X)$ (2.1) é a probabilidade de escolher à categoria j , dado um conjunto de variáveis independentes X .
- $P(Y_i=k|X)$ (2.2) é a probabilidade de escolher à categoria k dado o mesmo conjunto de variáveis independentes X .
- Y_i é a variável de escolha do i -ésimo indivíduo.
- As probabilidades condicionais $P(Y_i=j|X, Y_i \neq k)$ (2.3) e $P(Y_i=k|X, Y_i \neq j)$ (2.4) são as probabilidades de escolher as categorias j e k excluindo a categoria de referência, respectivamente.

O Teste *Hausman-McFadden* é utilizado para avaliar a validade da IIA. Ele compara o modelo geral com modelos específicos que excluem uma das categorias. Sua fórmula é dada por

$$H = -2 \cdot \log \left(\frac{L_g}{L_s} \right) \quad (4), \text{ onde}$$

- H é a estatística do teste *Hausman-McFadden*.
- L_g é a verossimilhança do modelo geral com todas as categorias.
- L_s é a verossimilhança dos modelos específicos que excluem uma das categorias.

Para esse teste, a hipótese nula afirma que não existe uma diferença sistemática nos coeficientes relativos entre duas estimativas ou modelos em questão. Isso significa que os modelos estão consistentes e não há evidência estatística de que suas relações entre variáveis tenham diferenças significativas. A hipótese alternativa, por outro lado, afirma que há uma diferença sistemática nos coeficientes relativos entre as duas estimativas ou modelos em consideração. Em termos simples, isso significa que os modelos não são consistentes, e há evidência estatística de que as relações entre variáveis têm diferenças significativas. Dessa forma, se o teste de hipóteses levar a um p-valor significativamente baixo, isso fornece evidência estatística de que a hipótese nula de IIA não é válida no modelo (VIJVERBERG, 2011).

3.3.2 Métricas para Regressão Logística Multinomial

3.3.2.1 Matriz de Confusão

A matriz de confusão é uma tabela que é comumente utilizada para avaliar o desempenho de um modelo de classificação. Ela compara as previsões feitas pelo modelo com as classes reais dos dados. Quando aplicada em um problema multiclasse, se dará por uma matriz quadrada de ordem igual ao número de classes, onde, geralmente, cada linha representa a classe real e cada coluna representa a classe prevista pelo modelo.

Figura 1: Exemplo de matriz de confusão para as classes “true”, “false” e “neither”

		predicted		
		true	false	neither
observed	true	257	51	3
	false	32	553	11
	neither	50	76	83

Fonte: DOUVEN, ELQAYAM (2021).

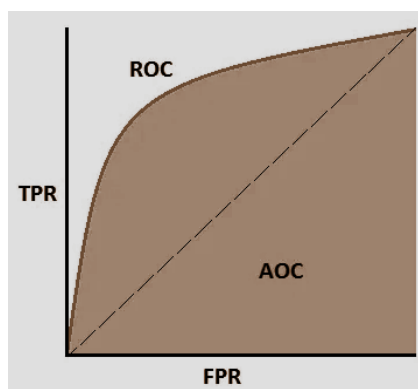
A partir dela, obtém-se alguns valores, tais como *TP* (*True Positives*), que indica o número de observações verdadeiras classificadas corretamente como positivas pelo modelo e geralmente se posicionam na diagonal principal da matriz. Também, o *FP* (*False Positives*), que é o número de observações falsas classificadas erroneamente como positivas pelo modelo. Também, o *TN* (*True Negatives*), significando o número de observações verdadeiras classificadas corretamente como negativas pelo modelo. Ademais, o *FN* (*False Negatives*), que é o número de observações verdadeiras classificadas erroneamente como negativas pelo modelo.

3.3.2.1 ROC e AUC

A Curva *ROC* é uma representação gráfica do desempenho geral de um modelo de regressão (HOSNER; LEMESHOW, 2000). Para sua construção, são obtidas diversas matrizes de confusão, cada uma utilizando um ponto de corte de probabilidade diferente. Os pontos de corte vão de 0 a 1 e, para cada matriz de confusão resultante, obtém-se a sensibilidade (ou

taxa de verdadeiros positivos), que é dada por $\frac{TP}{TP + FN}$ (5). Também, a taxa de falsos positivos, que é dada por $\frac{FP}{TN + FP}$ (6). Dessa forma, ao aplicar tais valores em um gráfico, a ligação dos pontos fornecem a curva *ROC*.

Figura 2 - Representação gráfica da *ROC*



Fonte: NARKHEDE (2018).

A *AUC* representa a área abaixo da curva *ROC*. Ela quantifica a capacidade de um modelo de classificação de distinguir entre as classes corretamente. Na Figura 1, a área total abaixo da linha pontilhada representa a capacidade de previsão do modelo nulo (sem previsores).

Hosmer & Lemeshow (2000) utilizam que uma *AUC* menor que 0,7 demonstra baixa capacidade discriminatória; *AUC* entre 0,7 e 0,8 demonstra discriminação aceitável; entre 0,8 e 0,9 demonstra discriminação excelente e *AUC* maior que 0,9 demonstra discriminação excepcional.

4 METODOLOGIA

4.1 COLETA, LIMPEZA E TRANSFORMAÇÃO DE DADOS

O processo de construção do modelo iniciou-se pela coleta e transformação de dados. A ideia era capturar aspectos diversos de cada música, com identificação de padrões temporais, exploração de complexidades líricas e análise de características sonoras. A ideia que orientou a coleta foi criar um conjunto de dados representativo e balanceado, incluindo variáveis relacionadas à análise sentimental e à dinâmica musical, o que permitiria uma análise abrangente das tendências ao longo das eras. Na tabela 3 são apresentados os processos de coleta, limpeza e transformação de dados, juntamente com as motivações da coleta.

Tabela 3 - Coleta, limpeza e transformação de dados e suas motivações

Coluna	Motivo da Coleta	Tipo de Informação	Limpeza e Transformação no Caso de Exemplo	Fonte da Informação no Caso de Exemplo
Nome da Canção	Servir como <i>id</i> único.	Textual.	NA.	<i>The Beatles Complete Chord Songbook.</i>
Posição de Lançamento do Álbum	Auxilia nas análises de evolução no tempo.	Numérico inteiro.	NA.	<i>The Beatles Complete Chord Songbook.</i>
Álbum	Auxilia nas análises de evolução no tempo, discretizando as evoluções.	Textual.	NA.	<i>The Beatles Complete Chord Songbook.</i>
Letra	É <i>input</i> para a análise sentimental e à quantidade de palavras distintas.	Textual.	NA.	<i>The Beatles Complete Chord Songbook.</i>
Quantidade de Palavras Distintas na Letra	Identificar tendências relacionadas à complexidade lírica no decorrer das eras.	Numérico inteiro de 0 a infinito.	Para cada letra (linha), foi aplicado uma função <i>Python</i> que conta a quantidade de palavras distintas.	Gerado a partir de dados de <i>The Beatles Complete Chord Songbook.</i>

Coluna	Motivo da Coleta	Tipo de Informação	Limpeza e Transformação no Caso de Exemplo	Fonte da Informação no Caso de Exemplo
Análise Sentimental da Letra	Identificar tendências no decorrer das eras.	Numérico flutuante de 0.00 a 1.00, para cada uma das variáveis “Negatividade”, “Neutralidade” e “Positividade”.	Descrito no tópico 4.1.2.	Gerado a partir de dados de <i>The Beatles Complete Chord Songbook</i> e
Cantor Principal	Identificar tendências no decorrer das eras.	Textual transformado em <i>One-Hot-Encoding</i> .	Apenas três canções, dentre todas, possuíam cantor principal igual a “Starr”. Tais linhas foram desconsideradas para a construção do modelo, evitando desequilíbrio nos dados.	<i>The Beatles Complete Chord Songbook</i> .
Criador do Conceito	Identificar tendências no decorrer das eras.	Textual transformado em <i>One-Hot-Encoding</i> .	Apenas três canções, dentre todas, possuíam como criador “Starr” ou “Outros”. Tais linhas foram desconsideradas para a construção do modelo, evitando desequilíbrio nos dados.	<i>The Beatles Complete Chord Songbook</i> .
Presença de Refrão	Identificar tendências no decorrer das eras.	Categórica textual, “Sim” ou “Não”.	NA.	<i>The Beatles Complete Chord Songbook</i> .
Tonalidade Principal	Identificar tendências no decorrer das eras.	Numérico inteiro de 0 a 11 utilizando a <i>Pitch Class Notation</i> . Exemplos: 0 = C, 1 = C#/D b, 2 = D.	Em casos nos quais a mudança de tonalidade ocorre, foi realizada análise pelo autor. Considerou-se a tonalidade presente nas partes mais memoráveis da canção.	<i>Isophonics Website</i> .
Modo	Identificar tendências no decorrer das eras.	Categórica textual, “Menor” ou “Maior”.	O modo das canções fora dos modos menor ou maior foi descrito	<i>Isophonics Website</i> .

Coluna	Motivo da Coleta	Tipo de Informação	Limpeza e Transformação no Caso de Exemplo	Fonte da Informação no Caso de Exemplo
			como o relativo mais parecido.	
Mudança de Tonalidade	Identificar tendências no decorrer das eras.	Categórica textual, “Sim” ou “Não”.	NA.	<i>Isophonics Website.</i>
Fórmula de Compasso	Identificar tendências no decorrer das eras.	Categórica textual, “Three” ou “Four”.	Em casos nos quais a mudança de fórmula de compasso ocorre, foi realizada análise pelo autor. Considerou-se a fórmula de compasso presente nas partes mais memoráveis da canção.	<i>API do Spotify.</i>
Dançabilidade	Identificar tendências no decorrer das eras.	Numérico flutuante de 0.00 (menos dançante) a 1.00 (mais dançante).	NA.	<i>API do Spotify.</i>
Energia	Identificar tendências no decorrer das eras.	Numérico flutuante de 0.00 (menos enérgica) a 1.00 (mais enérgica).	NA.	<i>API do Spotify.</i>
Acústicidade	Identificar tendências no decorrer das eras.	Numérico flutuante de 0.00 (menos acústica) a 1.00 (mais acústica).	NA.	<i>API do Spotify.</i>
Valência	Identificar tendências no decorrer das eras.	Numérico flutuante de 0.00 (menos alegre) a 1.00 (mais alegre).	NA.	<i>API do Spotify.</i>
Instrumentalidade	Identificar tendências no decorrer das eras.	Numérico flutuante de 0.00 (menos instrumental) a 1.00 (mais instrumental).	NA.	<i>API do Spotify.</i>
Vivacidade	Identificar tendências no decorrer das eras.	Numérico flutuante de 0.00 (menos ao vivo) a 1.00 (mais ao vivo).	NA.	<i>API do Spotify.</i>
Volume	Identificar tendências no decorrer das eras.	Numérico flutuante de -60dB (mais baixa) a 0dB (mais	NA.	<i>API do Spotify.</i>

Coluna	Motivo da Coleta	Tipo de Informação	Limpeza e Transformação no Caso de Exemplo	Fonte da Informação no Caso de Exemplo
		alta).		
Fala	Identificar tendências no decorrer das eras.	Numérico flutuante de 0.00 (menos falas) a 1.00 (mais falas).	NA.	API do Spotify.
Tempo	Identificar tendências no decorrer das eras.	Numérico flutuante de 0.00 a infinito.	NA.	API do Spotify.
Duração em Milissegundos	Identificar tendências no decorrer das eras.	Numérico flutuante de 0.00 a infinito.	NA.	API do Spotify.
Quantidade de Acordes	Identificar tendências no decorrer das eras.	Numérico inteiro de 0 a infinito.	NA.	API do Spotify.
Complexidade dos Acordes	Identificar tendências no decorrer das eras.	Numérico inteiro de 0 a infinito.	Descrito no tópico 4.1.1.	Gerado a partir de dados de <i>The Beatles Complete Chord Songbook</i> .
Era	Serve como variável dependente, além de gabarito para previsões.	Categórica.	NA.	Fitzsimmons (2019).

4.1.1 Complexidade dos Acordes

Para cada uma das canções, foi realizada uma tabela contendo todos os seus acordes. Cada acorde foi representado em uma coluna distinta, e as músicas dispostas nas linhas.

Tabela 4 - Acordes

Song	Chord.1	Chord.2	Chord.3	Chord.4	Chord.5	Chord.6
Ask Me Why	E	F#m7	G#m7	G#7	C#m	Am
Baby It's You	A	E7	D7	E	D	A7
Do You Want to Know a Secret	Em	Am	G	F	B7	E

Foi desenvolvido um sistema de classificação, onde cada acorde foi avaliado de acordo com o número de notas que o compõem, tendo sua complexidade classificada em três níveis, sendo a **Complexidade 1** para acordes que consistem em três notas tocadas simultaneamente, como Em (Mi menor). **Complexidade 2** para acordes com quatro notas tocadas simultaneamente, como Em7 (Mi menor com sétima). Por fim, **Complexidade 3** para acordes com cinco notas tocadas simultaneamente, como Em7/9 (Mi menor com sétima e nona).

Dessa forma, a complexidade total dos acordes de cada música consiste na soma das complexidades dividido pela quantidade de acordes. Assim, músicas com mais acordes simples e de Complexidade 1, recebem uma complexidade menor e mais próxima de 1, enquanto músicas com acordes mais complexos recebem uma pontuação mais alta.

4.1.2 Análise Sentimental da Letra

Para quantificar o sentimento presente em cada letra, empregou-se o VADER (*Valence Aware Dictionary and sEntiment Reasoner*) do *Natural Language Toolkit (NLTK)*, que é uma ferramenta de análise de sentimentos projetada para lidar com texto em linguagem natural. O VADER atribui a cada texto analisado três pontuações que refletem diferentes dimensões emocionais, sendo negatividade, neutralidade e positividade, as três variando de 0 (fraca incidência) a 1 (alta incidência).

Tabela 5 - Exemplos de resultados do VADER

<p><i>“Love, love me do You know I love you I’ll always be true So please, love me do Whoa, love me do</i></p> <p><i>Someone to love Somebody new Someone to love Someone like you</i></p> <p><i>Love, love me do You know I love you I’ll always be true So please, love me do Whoa, love me do Yeah, love me do Whoa, oh, love me do”</i></p> <p>Resultado da Análise Sentimental:</p>	<p><i>“Honey Pie Honey Pie</i></p> <p><i>Honey Pie Honey Pie</i></p> <p><i>Honey Pie Honey Pie Honey Pie Honey Pie”</i></p> <p>Resultado da Análise Sentimental:</p> <p>Negatividade: 0,000 Neutralidade: 1,000 Positividade: 0,000</p>
--	---

Negatividade: 0,000 Neutralidade: 0,369 Positividade: 0,631	
---	--

A alta dimensionalidade notada nessa separação em três componentes, a qual poderia resultar em uma maior dificuldade de interpretação das informações e multicolinearidade, motivou a utilização da Análise de Componentes Principais.

Assim, os dados foram normalizados, subtraindo a média e dividindo pelo desvio padrão. Depois, calculou-se as proporções da variância explicada por cada componente gerado pela *PCA*. Isso permitiu avaliar a importância de cada componente na explicação da variância nos dados. Assim, com base na análise das proporções de variância explicada, identificou-se que os dois primeiros componentes principais explicavam aproximadamente 99,9% da variância total. Assim, optou-se por utilizar apenas esses dois componentes no conjunto de dados.

4.2 ANÁLISE EXPLORATÓRIA

4.2.1 Distribuição da Variável Resposta

A contagem de observações em cada uma das quatro categorias de "Era" foi importante para notar que havia relativo desequilíbrio nas quantidades, o que poderia afetar a capacidade do modelo de aprender com igualdade de desempenho em todas as categorias. Todavia, não havia possibilidade de coletar mais linhas, visto que a banda já encerrou. Também, eliminar linhas para alcançar equilíbrio não traria benefícios em termos explicativos.

Figura 3 - Distribuição das classes-resposta

```

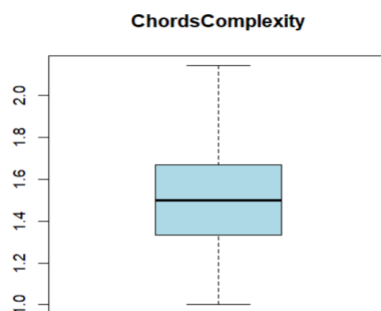
Era
Beatlemania:29
Dissolution:55
Psychedelic:22
Transition :48

```

Fonte: Autoria própria.

4.2.2 Distribuição das Variáveis Independentes

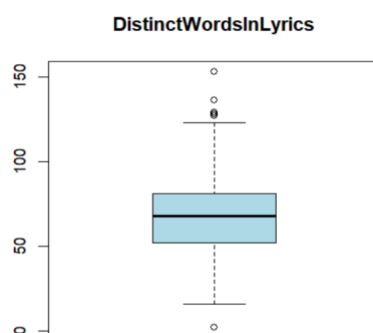
Figura 4 - Boxplot de complexidade dos acordes



Fonte: Autoria própria.

O boxplot da complexidade dos acordes sugere variação considerável. Enquanto a maioria das canções possui uma complexidade moderada, há casos extremos com níveis de complexidade tanto muito mais baixos quanto muito mais altos.

Figura 5 - Boxplot da quantidade de palavras distintas



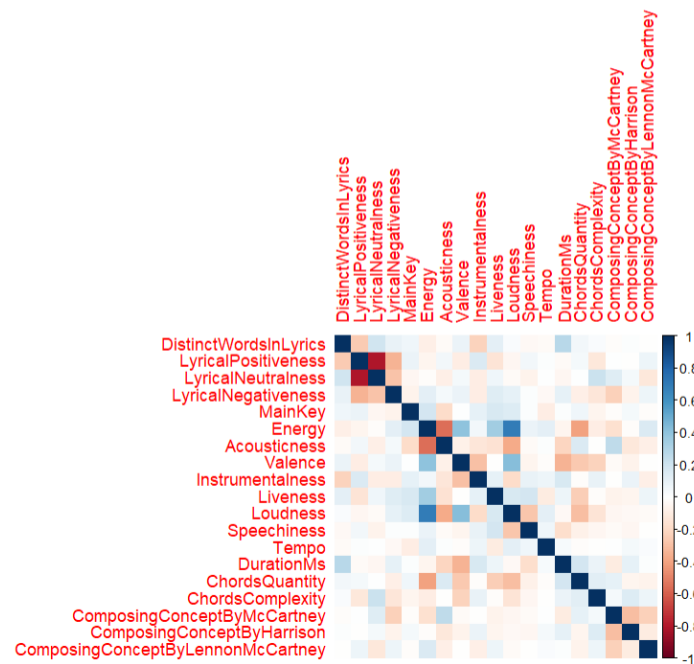
Fonte: Autoria própria.

Ao observar a Figura 5, nota-se que muitas canções estão na faixa intermediária, com cerca de 68 palavras distintas. Porém há variações significativas, com algumas músicas contendo muito menos ou muito mais palavras distintas. Essa variabilidade poderia ser um fator importante a ser considerado ao analisar como a riqueza vocabular nas letras das músicas pode estar relacionada à era musical.

4.2.3 Multicolinearidade

Foram separadas as variáveis não-fator para a observação da multicolinearidade através de um gráfico. As que são fatores foram observadas posteriormente através do *VIF*. Assim, para as não-fator, obteve-se o *plot*:

Figura 7 - Plotagem de multicolinearidade entre variáveis não-fator



Fonte: Autoria própria.

Também calculou-se estatisticamente o *VIF* para cada variável independente, obtendo os resultados abaixo.

Figura 8 - Resultados de *VIF* para todas as variáveis

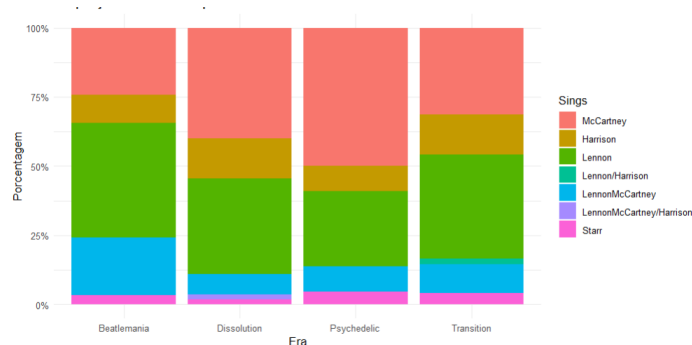
	GVIF
DistinctWordsInLyrics	1.521565
LyricalPositiveness	79520.159623
LyricalNeutrality	76455.414036
LyricalNegativity	30545.663405
Sings	424.995669
ChorusPresence	1.343792
MainKey	1.512349
Mode	1.587638
ChangesKey	1.634810
TimeSignatureIn	1.466048
Energy	5.076379
Acousticness	2.289139
Valence	1.944031
Instrumentalness	1.752357
Liveness	1.550383
Loudness	3.730540
Speechiness	1.597050
Tempo	1.490714
DurationMs	1.945458
ChordsQuantity	2.466069
ChordsComplexity	2.001278
HasChordA	2.165414
HasChordG	2.038830
HasChordD	2.235603
HasChordC	1.787659
HasChordF	2.056950
HasChordD7	1.816715
HasChordA7	1.498990
HasChordAm	1.518892
HasChordBm	1.671944
ComposingConceptByMcCartney	10.394847
ComposingConceptByHarrison	10.859289
ComposingConceptByLennonMcCartney	2.612984

Fonte: Autoria própria.

A partir dos resultados, foi possível identificar a multicolinearidade presente entre as variáveis de análise lírica, tanto quanto as variáveis *Sings* e *ComposingConceptBy*, situação tratada posteriormente.

4.2.4 Visualização das Variáveis Categóricas

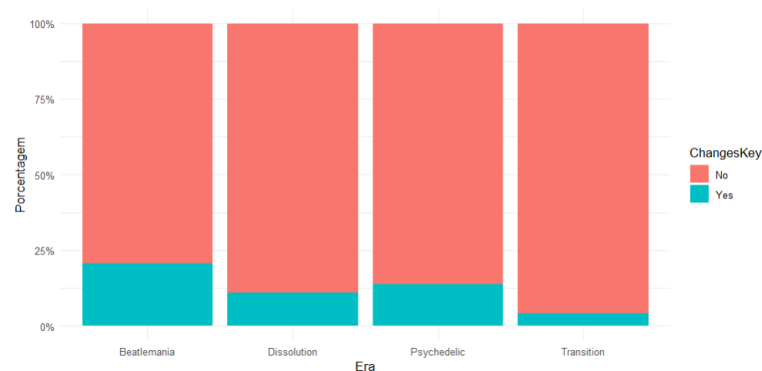
Figura 9 - Proporção de cantores por era



Fonte: Autoria própria.

A figura 8 demonstrou um aumento na participação de McCartney como cantor nas canções, em contraposição ao espaço de Lennon, que diminui tanto nas participações como dupla com McCartney quanto sozinho.

Figura 10 - Proporção de canções com mudanças de tom a cada era

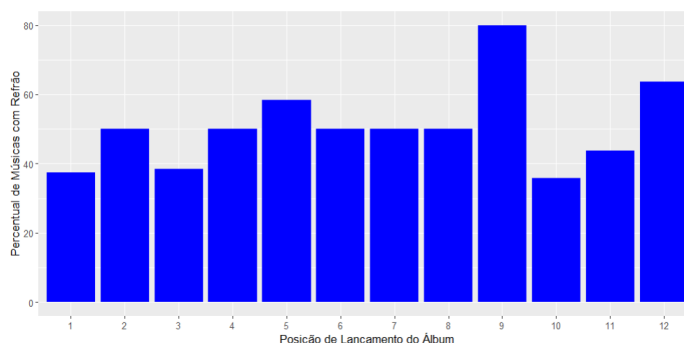


Fonte: Autoria própria.

Foi possível notar uma diminuição sólida na proporção de canções com mudança de tom. O observado no gráfico se demonstrou como o oposto do esperado, visto que quanto mais avançada a era, maior a complexidade harmônica notada auditivamente. Dessa forma, a expectativa era que houvesse um aumento na proporção de canções com mudanças de tonalidade.

4.2.5 Análise de Tendências Temporais

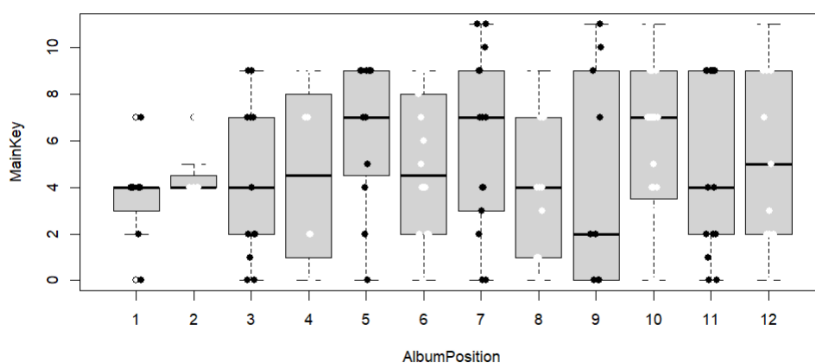
Figura 11 - Evolução da presença de refrão



Fonte: Autoria própria.

Através da figura 10, que demonstra a evolução do percentual de canções com presença de refrão, não foi possível notar uma tendência muito forte. Todavia, foi possível notar que um álbum apenas, o nono lançado, possui um percentual de canções com refrões muito mais alto que os outros de sua era (*Psychedelic*).

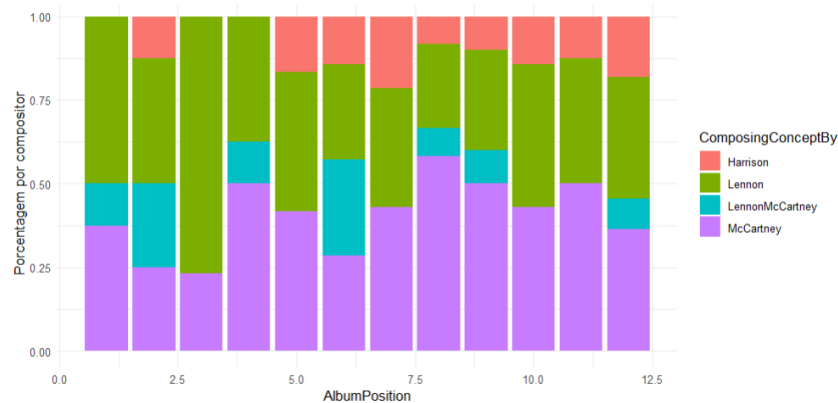
Figura 12 - Boxplots de tonalidade por posição de lançamento do álbum



Fonte: Autoria própria.

É possível notar uma dispersão cada vez maior das tonalidades no decorrer do lançamento dos álbuns, visto que os primeiros álbuns possuíam mediana e quartis muito próximos, diferentemente dos últimos.

Figura 13 - Distribuição de compositores por posição de lançamento do álbum



Fonte: Autoria própria.

Aqui, foi possível notar uma participação cada vez mais sólida de George Harrison nas composições da banda. Também, um afunilamento das participações de Lennon, dando espaço ao compositor citado anteriormente.

4.2.6 Testes de Pressupostos

4.2.6.1 Testes Individuais de AUC

Foi utilizado um modelo de teste contendo como fonte de dados apenas uma variável independente e a variável dependente, no qual foi aplicada uma iteração para aplicar uma variável independente de cada vez e obter a *AUC* individual de cada uma delas. Através desse teste, foi possível conferir se havia alguma variável que possuía participação desleal na predição da Era, além de obter uma estimativa da capacidade de predição de cada uma. Os maiores potenciais individuais de predição se deram através da duração da canção (*DurationMs*), instrumentalidade (*Instrumentalness*) e a complexidade dos acordes (*ChordsComplexity*), respectivamente com 0,6944, 0,6557 e 0,6017 de *AUC*.

4.2.6.2 Variável Dependente Nominal e Categorias Mutuamente Exclusivas

Foi assegurado que as categorias da Era eram distintas e não se sobrepunham, garantindo que cada observação fosse corretamente classificada em uma categoria exclusiva. Como no caso em questão as eras são separadas por álbuns distintos, não existe sobreposição de valores.

4.2.6.3 Independência das Observações (Sem Repetições)

Foi garantido que todas as observações contidas no conjunto de dados eram únicas, sem qualquer forma de repetição. Esse formato de dados com repetição é nomeado *Long*, todavia

os dados se encontravam em formato *Wide*. Desse modo, cada linha representou uma canção e todas as informações de uma mesma canção estavam em uma só linha, logo o teste passou.

4.2.6.4 Independência das Alternativas Irrelevantes (*Teste Hausman-McFadden*)

O teste avaliou se a probabilidade de escolha entre duas categorias é influenciada pela introdução de uma terceira categoria. Para isso, desenvolveram-se quatro modelos auxiliares distintos, cada um contando com apenas três das quatro classes de resposta, com o objetivo de analisar se as classes resposta são independentes entre si.

No caso do *Teste Hausman-McFadden*, um p-valor de 1 sugere que não há evidências estatísticas para rejeitar a hipótese nula de que as categorias têm um efeito significativo nas escolhas entre si. Os testes obtiveram p-valor igual a 1, indicando fortemente que a hipótese nula de independência das alternativas irrelevantes não foi rejeitada.

5 RESULTADOS

5.1 COLUNAS DE DADOS

Tabela 6 - Variáveis independentes e dependentes no modelo

Coluna	Tipo de Dado
Quantidade de Palavras Distintas na Letra	Numérico inteiro de 0 a infinito.
PCA da Análise Sentimental da Letra 1	Numérico flutuante de 0 a infinito.
PCA da Análise Sentimental da Letra 2	Numérico flutuante de 0 a infinito.
Criador do Conceito: McCartney	Categórica textual, “Sim” ou “Não”.
Criador do Conceito: Harrison	Categórica textual, “Sim” ou “Não”.
Criador do Conceito: LennonMcCartney	Categórica textual, “Sim” ou “Não”.
Presença de Refrão	Categórica textual, “Sim” ou “Não”.
Tonalidade Principal	Numérico inteiro de 0 a 11 utilizando a Pitch Class Notation. Exemplos: 0 = C, 1 = C#/D ♭, 2 = D.

Modo	Categórica textual, “Menor” ou “Maior”.
Mudança de Tonalidade	Categórica textual, “Sim” ou “Não”.
Fórmula de Compasso	Categórica textual, “Three” ou “Four”.
Energia	Numérico flutuante de 0.00 (menos enérgica) a 1.00 (mais enérgica).
Acústicidade	Numérico flutuante de 0.00 (menos acústica) a 1.00 (mais acústica).
Valência	Numérico flutuante de 0.00 (menos alegre) a 1.00 (mais alegre).
Instrumentalidade	Numérico flutuante de 0.00 (menos instrumental) a 1.00 (mais instrumental).
Vivacidade	Numérico flutuante de 0.00 (menos ao vivo) a 1.00 (mais ao vivo).
Volume	Numérico flutuante de -60dB (mais baixa) a 0dB (mais alta).
Fala	Numérico flutuante de 0.00 (menos falas) a 1.00 (mais falas).
Tempo	Numérico flutuante de 0.00 a infinito.
Duração em Milissegundos	Numérico flutuante de 0.00 a infinito.
Quantidade de Acordes	Numérico inteiro de 0 a infinito.
Complexidade dos Acordes	Numérico flutuante de 1 a infinito.
Era (variável dependente)	Categórica, sendo as opções textuais “Beatlemania”, “Transition”, “Psychedelic” e “Dissolution”.

5.2 FORMULAÇÃO MATEMÁTICA

A regressão logística multinomial é uma extensão da regressão logística binomial, adaptada para problemas de classificação com mais de duas categorias. A formulação de um modelo

genérico é representada da forma

$$\log\left(\frac{P(Y=k|X)}{P(Y=K|X)}\right) = \beta_{0k} + \beta_{1k}X_1 + \beta_{2k}X_2 + \dots + \beta_{pk}X_p$$

(7), onde:

- Y é o resultado da variável resposta.
- k é uma das opções das categorias-resposta.
- X são as variáveis preditoras.
- $\beta_{0k}, \beta_{1k}, \dots, \beta_{pk}$ são os coeficientes do modelo para a categoria k .
- P é a probabilidade de se resultar no que está entre parênteses.
- K é a categoria escolhida como referência.

No caso de estudo em questão, as variáveis de resposta são as eras musicais e as variáveis preditoras são as características musicais coletadas. O modelo atribui probabilidades a cada categoria de era musical e, em seguida, classifica a música na categoria com a probabilidade mais alta. Dessa forma, ao realizar a substituição dos valores genéricos (X , k , K) descritos acima pelos valores específicos ao caso de estudo em questão e isolando a parte referente à probabilidade de se obter um valor de igual à opção, obtemos equações como

$$P(Era \sim Transition) =$$

$$P(Era \sim Beatlemania) \cdot e^{\beta_{intercepto} + \beta_{DistinctWords} \cdot DistinctWords + \dots + \beta_{Danceability} \cdot Danceability} \quad (8),$$

$$P(Era \sim Psychedelic) =$$

$$P(Era \sim Beatlemania) \cdot e^{\beta_{intercepto} + \beta_{DistinctWords} \cdot DistinctWords + \dots + \beta_{Danceability} \cdot Danceability} \quad (9),$$

$$P(Era \sim Dissolution) =$$

$$P(Era \sim Beatlemania) \cdot e^{\beta_{intercepto} + \beta_{DistinctWords} \cdot DistinctWords + \dots + \beta_{Danceability} \cdot Danceability} \quad (10).$$

5.3 COMPARAÇÃO COM O MODELO NULO

Tabela 7 - Comparação ao modelo sem previsoires

Modelo	Graus de Liberdade Residuais (Resid. df)	Diferença nos Graus de Liberdade (Df)	Desvio Residual (Resid. Dev)	Diferença de Desvio Residual (LR stat)	Pr(Chi)
Nulo	459		407.6304		
Com previsoires	390	69	175.8181	231.8123	0

Os graus de liberdade associados aos erros residuais do modelo (Resid. df) favoreceram o modelo com previsores, com 69 graus de liberdade residuais a menos. Para o desvio residual (Resid. Dev), que é uma medida de ajuste do modelo aos dados, quanto menor o valor, significa que melhor é o ajuste. O modelo com previsores tem um desvio residual menor em 231.8123 unidades do que o modelo nulo. O p-valor (Pr(Chi)), que indica a probabilidade de observar a estatística de teste (LR stat.) assumindo que a hipótese nula (de que não há diferença relevante entre os modelos) é verdadeira, ficou abaixo de 0.05, demonstrando que, de fato, o modelo com previsores é estatisticamente significativamente melhor que o modelo nulo.

Em resumo, todos os resultados apresentados na Tabela 7 sugerem que o modelo com previsores oferece um ajuste significativamente superior e é capaz de explicar as eras de forma mais eficaz.

5.4 PSEUDO- R^2

O valor do Pseudo R^2 de Nagelkerke obtido foi de 0,8373. Isso é um indício de que o modelo é capaz de capturar uma grande parte da variação na era das músicas com base nas variáveis consideradas. Todavia quaisquer Pseudo- R^2 possuem baixa interpretabilidade, necessitando fortemente de outros testes para demonstrar como cada variável contribui para a classificação (SMITH; MCKENNA, 2013).

5.5 P-VALORES

No contexto do caso de estudo, o p-valor associado a uma variável é uma medida estatística que indica a probabilidade de observar uma associação entre essa variável e a variável de resposta no modelo, assumindo que não há efeito real (hipótese nula). Para o modelo em questão, foi definido no início do trabalho que quando o p-valor fosse igual ou inferior a 0,05 serviria como evidência estatística de que há uma associação significativa entre a variável em questão e a variável de resposta no modelo, enquanto um p-valor acima de 0,05 indicaria que a variável pode não ser tão relevante. Algo que deve ser notado é que todas as variáveis presentes no modelo, independente da classe, possuíram p-valores muito próximos de 0, demonstrando relevância estatística em todas elas na previsão da era. A tabela contendo todos os resultados está presente no Anexo 3, relacionado aos efeitos específicos.

5.6 MATRIZ DE CONFUSÃO

Figura 14 - Matriz de confusão utilizando o mesmo conjunto para treino e teste

Prediction	Reference			
	Beatlemania	Dissolution	Psychedelic	Transition
Beatlemania	24	2	0	4
Dissolution	2	42	5	5
Psychedelic	0	4	14	2
Transition	3	7	3	37

Fonte: Autoria própria.

É possível observar a maior incidência numérica na diagonal principal da matriz, indicando que a maior parte das previsões foram corretas. Ao realizar a divisão da diagonal principal em relação à soma total, obtém-se uma acurácia de 75,97%. O intervalo de confiança de 95% para essa precisão está entre 68,44% e 82,48%.

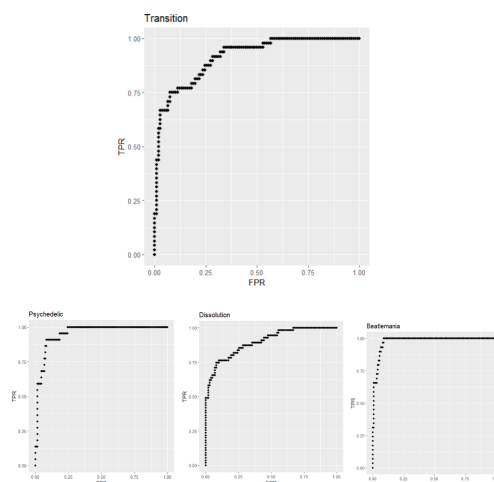
A taxa de acerto esperada se fizéssemos previsões sem nenhum conhecimento (No Information Rate) é de 35,71%, com um p-valor para a comparação entre a precisão do modelo e a taxa de acerto do modelo aleatório muito próximo de 0, indicando que o modelo é estatisticamente significativamente melhor do que fazer previsões aleatórias.

5.7 CURVA ROC E AUC

Foi calculada a *AUC* através da utilização de todos os dados coletados como conjunto de treino. Isso foi feito para se obter um parâmetro da capacidade de generalização do modelo.

Foi obtido o resultado individual da classe Beatlemania de 0.9787; Dissolution de 0.8991; Psychedelic de 0.9528 e Transition de 0.9137. Isso resulta na média para todas as classes resposta de 0.9361.

Figura 15 - Gráficos da curva ROC para cada classe-resposta



Fonte: Autoria própria.

A média dos melhores *thresholds*, que indicam o valor de probabilidade que melhor separa uma classe das outras em termos da relação $TPR \times FPR$, foi de 0.52. Isso significa que, para se obter melhores resultados, o pertencimento a uma classe só deve ser atribuído se sua probabilidade for maior que o *threshold* de 0.52.

5.8 VALIDAÇÃO *K-FOLD* PARA *AUC*

Foi realizada a validação *K-Fold*, que consiste em dividir o conjunto de dados em K partes (ou *folds*) de aproximadamente igual quantidade de linhas. O valor de K é escolhido com base na quantidade desejada de divisões no conjunto de dados. Cada uma das K partes é utilizada como conjunto de teste uma vez, enquanto as $K-1$ partes restantes formam o conjunto de treinamento. Esse processo é repetido K vezes, garantindo que cada parte seja utilizada como conjunto de teste exatamente uma vez, permitindo uma análise abrangente do desempenho em comparação com uma única divisão entre conjunto de treinamento e teste.

O cálculo da *AUC* através das bibliotecas utilizadas depende da existência de pelo menos uma evidência de cada categoria-resposta. Assim, haja visto que o modelo em questão possui quantidade de linhas muito limitada, um valor grande de K diminuiria muito a probabilidade de tal suposição ser atendida. Dessa forma, o valor de K escolhido para os testes foi de 7, pois sendo que um valor menor que este poderia acarretar em um conjunto de treino muito curto, dificultando a identificação de padrões de forma clara.

Através desse teste, obteve-se uma média de *AUC* de 0.7184. É um resultado interessante e demonstra certa capacidade de generalização do modelo, todavia a queda com relação ao teste de *AUC* utilizando o mesmo conjunto para treino e teste foi de 30%.

5.9 VALIDAÇÃO LEAVE-ONE-OUT PARA ACURÁCIA

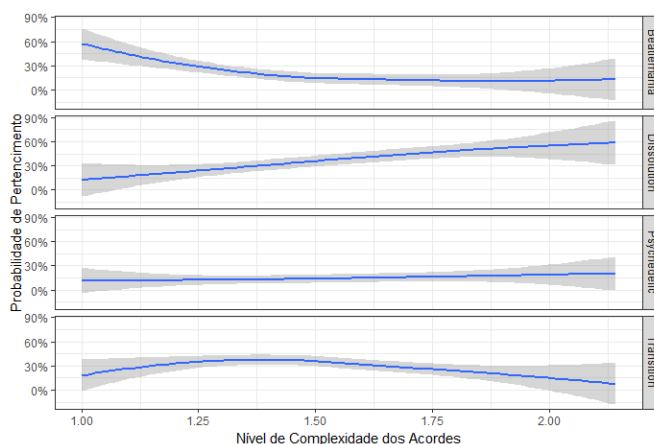
O método *leave-one-out* é uma técnica de validação cruzada. Ele envolve treinar um modelo usando todos os dados, exceto uma única linha, que é retida como conjunto de teste. Esse processo é repetido, removendo cada vez uma linha de dados, permitindo avaliar a média do desempenho do modelo para cada observação individual. Essa abordagem é útil para avaliar a capacidade de generalização de um modelo e reduzir o viés na avaliação do desempenho.

A partir dessa técnica e com a utilização do *threshold* de 0.52 defendido no capítulo 5.6, a quantidade de previsões corretas em relação ao total foi de aproximadamente 71.5%.

5.10 INTERPRETAÇÃO DOS COEFICIENTES

Através dos gráficos dos resultados obtidos para as probabilidades de pertencimento a cada uma das eras, foi possível realizar diversas interpretações.

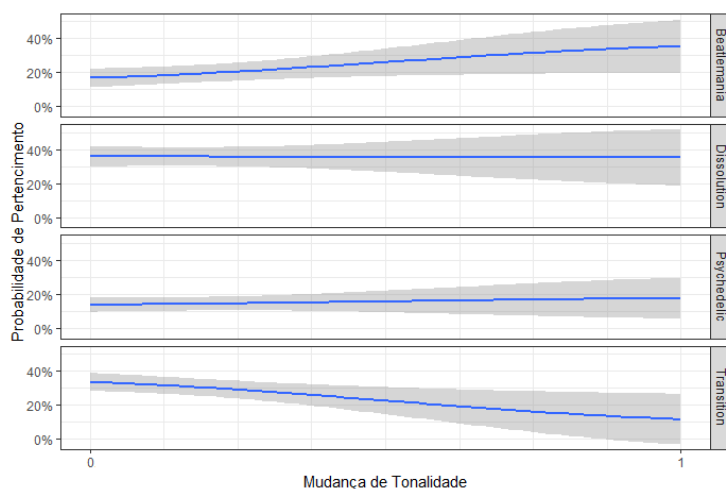
Figura 16 - Probabilidades de pertencimento das classes em relação à complexidade dos acordes



Fonte: Autoria própria

Nota-se que o nível de complexidade dos acordes possui um peso significativo para a probabilidade de pertencimento das eras Beatlemania, Dissolution e Transition. Quanto maior o nível de complexidade, menor a probabilidade de pertencimento a Beatlemania, enquanto para Dissolution, a relação se torna inversa com uma relação que se assemelha à linearidade. Para Transition, notamos uma relação de que quanto mais próximos dos extremos, menor a probabilidade, enquanto os valores próximos a 1.35 de complexidade aumentam a probabilidade de pertencimento a esta classe. Já para a classe Psychedelic, independente do nível, sua probabilidade se mantém em cerca de 14%.

Figura 17 - Probabilidades de pertencimento das classes em relação à presença de mudança de tonalidade



Fonte: Autoria própria.

A presença de mudanças de tonalidade afeta significativamente as probabilidades de pertencimento a Transition de forma negativa. Se a canção possui essa técnica em sua harmonia, muito provavelmente não pertence à era Transition. Já para Beatlemania, a relação é inversa. Se não possuir mudança de tonalidade, há uma probabilidade menor de pertencimento.

A tabela contendo todos os coeficientes está presente no Anexo 3, de efeitos específicos. Sob posse deles, foi possível aplicar a função exponencial e obter a razão de chances (*OR*) de escolha da categoria em questão em detrimento da referência. Da mesma forma, aplicando a função exponencial ao intervalo de confiança, obteve-se o intervalo de confiança relacionado à razão de chances. Tais valores de *Odds Ratio* e seus intervalos de confiança também estão no Anexo 1.

6 CONSIDERAÇÕES FINAIS

6.1 DISCUSSÃO DOS RESULTADOS

A partir dos resultados notou-se que, a todo momento, é necessário realizar *tradeoffs*. Por exemplo, optar pelo uso do PCA na análise lírica foi uma escolha visando minimizar a perda de informações, em contraste com a alternativa de simplesmente remover colunas. Todavia, apesar de conseguir manter mais de 99% de sua variância através da técnica do PCA, a interpretabilidade final dos resultados da análise lírica foi muito comprometida. Essa

necessidade constante de *tradeoffs* presente no universo do aprendizado de máquina dificultou a obtenção de uma resposta para o problema proposto, visto que há sempre uma quantidade muito grande de possibilidades de decisão, cada uma com suas vantagens e desvantagens e, dentro de um estudo com tempo limitado, torna-se inviável testar todas elas.

Além de trocas, foi necessário realizar reduções. Tais reduções não foram exclusivas deste estudo de caso, mas estão presentes em todo o referencial teórico estudado. Isso se deu tanto para valores de entrada (dançabilidade e complexidade dos acordes, por exemplo), os quais precisaram ser processados a partir de critérios objetivos para a separação em níveis, quanto nos valores subjetivos, os quais procurava-se compreender. Através de Cerati (2021) e o termo “compreensão mútua”, pode-se compreender as reduções como a remoção de tudo que não participa da “compreensão mútua”. Em outras palavras, a remoção do que é, de fato, subjetivo. Por exemplo, uma canção dos Beatles não necessariamente é notada como pertencente a uma única era. A sonoridade e os timbres podem fazer com que, de fato, os apreciadores observem alguma canção como um meio-termo entre duas ou mais eras. Todavia, para que seja viável a implementação do problema, há a necessidade de atribuí-la para alguma das opções a partir de um critério objetivo. Tais reduções se vêm como problemáticas pois a intenção foi realizar traduções objetivas para critérios subjetivos. No entanto, o melhor que se parece conseguir é a tradução objetiva de outros critérios objetivos, os quais sofreram reduções de significado.

Zhou, Claire e King (2014) utilizam o argumento da suspeita de que os resultados do aprendizado de máquina superariam o desempenho humano na previsão da localização geográfica musical. Tal argumento da aplicabilidade foi utilizado para demonstrar a competitividade dos modelos. De forma análoga, os resultados e métricas obtidas no estudo de caso dos *Beatles* já demonstram aplicabilidade. Os p-valores observados para cada uma das variáveis fornecem indícios de que, sim, uma característica musical subjetiva (no caso de estudo, a Era) possui rastros objetivos, possibilitando vantagens no desenvolvimento de filtros e recomendações. Também, a comparação entre o modelo nulo e o modelo desenvolvido no estudo de caso demonstra ampla vantagem em utilizar esta metodologia como guia no desenvolvimento de recomendações ao invés de simplesmente escolher uma canção aleatória. Em um caso hipotético, no qual uma pessoa é fã da era Beatlemania, há respaldo, através do observado na Figura 17, para afirmar que a classe Beatlemania pode, dentre todo o resto, ser distinguida das outras eras através da mudança de tonalidade. Isso significa que, caso um serviço de streaming tenha a intenção de indicar uma próxima música dentro de uma playlist

contendo apenas Beatles, caso uma canção conte com uma mudança de tonalidade, aumenta-se a probabilidade de que a canção tenha semelhanças às da era Beatlemania, satisfazendo o gosto do usuário.

6.2 TRABALHOS FUTUROS

A metodologia realizada se deu através da noção prévia da existência das fases musicais dos *Beatles*. Todavia, seria interessante a construção de uma metodologia que possibilita a identificação de eras musicais sem saber previamente das mesmas. Isso agregaria informação útil à discussão subjetiva, trazendo à tona mudanças relevantes que eram despercebidas até então.

Os resultados foram observados analisando a viabilidade das conversões, todavia há a carência de testes relacionados à experiência do usuário. Dessa forma, nota-se que é necessário evoluir a implementação do sistema de sugestão musical através do método realizado, aplicando-o em uma plataforma real de *streaming* e colhendo resultados relacionados à satisfação do ouvinte.

6.3 CONCLUSÃO

Em síntese, a análise e metodologia realizados neste estudo revelam a complexidade e os desafios inerentes à aplicação de técnicas computacionais na interpretação de características musicais subjetivas. A constante necessidade de *tradeoffs* e reduções, tanto na escolha de métodos como no recolhimento de variáveis, põe “em xeque” o pressuposto da tradução sem perdas de significado e destaca o balanço entre preservar informações e garantir interpretabilidade.

Conclui-se que, apesar das limitações e dificuldades encontradas, os resultados obtidos sugerem aplicabilidade prática, evidenciada pelos p-valores e os valores de comparação ao modelo nulo, os quais indicam a presença de padrões objetivos relacionados às características musicais analisadas. Tais resultados sugerem que características subjetivas podem ter fundamentos objetivos, proporcionando oportunidades para o desenvolvimento de filtros e sistemas de sugestão eficazes em plataformas de *streaming*.

7 REFERÊNCIAS BIBLIOGRÁFICAS

- [1] GUTZMER, Jared. The Beatles and the crisis of self-identify. **Sosland Journal**, Kansas, agosto de 2019. Disponível em:

- https://info.umkc.edu/sosland_journal/wp-content/uploads/2019/08/Gutzmer_Sosland_2019.pdf. Acesso em: 2 de julho de 2023.
- [2] STURM, Bob. Machine Learning Research that Matters for Music Creation: A Case Study. **Journal of New Music Research**, 2019. Disponível em: <https://hal.science/hal-03278041/document>. Acesso em: 2 de julho de 2023.
- [3] DREYFUS, Hubert L. **On the Internet**. 2. ed. London and New York: Routledge, 2009.
- [4] SOUZA, José Geraldo de: Renovação Pedagógica no Ensino de Engenharia. **ASEE**, 2002. Disponível em: <http://www.asee.org/international/INTERTECH2002/533.pdf>. Acesso em: 2 de julho de 2023.
- [5] HUYGUE, René. **A Arte E A Alma**. Editora Bertrand, 1960.
- [6] Ploin, A., Eynon, R., Hjorth I. e Osborne, M.A. (2022). AI and the Arts: How Machine Learning is Changing Artistic Work. Paper do projeto **Creative Algorithmic Intelligence Research Project**. Oxford Internet Institute, Universidade de Oxford, UK.
- [7] GRIMES, William: Volumes to Go Before You Die. **New York Times**, 23 de maio de 2008.
- [8] MACDONALD, Ian. **Revolution in the Head: The Beatles' Records and the Sixties**. Prague: Volvox Globator, 1997.
- [9] FRIEDLANDER, Paul. **Rock and Roll: Uma História Social**. Tradução de A. Costa. 4º ed, RJ: Record, 2006. 485 páginas.
- [10] FRITH, Simon. **A Sociologia do Rock**. Milan Feltrinelli. 1999.
- [11] GALLAHER, Lisette F. **Think of what you're saying: mapping the evolution and impact of the Beatles' lyrical approaches to gender in 1965-1966**. MRes thesis, University of Glasgow.
- [12] GUERRERO, Rodrigo. The Role of The Beatles in Popularizing Indian Music and Culture in the West. **The Owl Florida State University Undergraduate Research Journal**, 2015.
- [13] WICKE, Peter, ZIEGENRÜCKER, Kai-Erik. **Handbook of Popular Music**. Alemanha: Atlantis Musikbuch-Verlag, 2001.

- [14] DORUŽKA, Lubomír. **Music Overview**. Praga: Mladá fronta, 1981.
- [15] GAMMOND, Peter. **The Oxford Companion to Popular Music**. **Oxford: Oxford University Press**, 1993.
- [16] CRAWFORD, Richard. **America's Musical Life: A History**. Nova York: Norton, W. W. & Company, Inc., 2005.
- [17] HECL, Rudolf. **The Beatles and Their Influence on Culture**. Brno, 2006. Masaryk University, Brno Faculty of Arts Department of English and American Studies.
- [18] **GUINNESS World Records**. 2005. Disponível em: <<http://www.guinnessworldrecords.com/>>. Acesso em: 2 de julho de 2023.
- [19] BLACKMAN, Marisa. **The Development of The Beatles' Programmatic Music. The Journal Of The Student Scholarship**. Volume XXII. Edição de abril de 2020. Disponível em: <<https://joss.tcnj.edu/wp-content/uploads/sites/176/2020/04/2020-Blackman-submission.pdf>>. Acesso em: 2 de julho de 2023.
- [20] DOGGETT, Peter. **The Art And Music Of John Lennon**. Omnibus Press, 2009. 368 págs.
- [21] CARR, Paul, KENNETH, Womack, **The Cambridge Companion to the Beatles**. Cambridge University Press, 2009. Twentieth-Century Music, 2011, páginas. 256–260.
- [22] HOWLIN, Carl., ROONEY, Bradley. **Patients choose music with high energy, danceability, and lyrics in analgesic music listening interventions**. 2021. Psychology of Music. Páginas 931–944.
- [23] LIEW, Uchida. (2022). **Energetic music is used for anger downregulation: A cross-cultural differentiation of intensity from rhythmic arousal**. Journal of Applied Social Psychology. Disponível em: <<https://doi.org/10.1111/jasp.12951>>. Acesso em: 2 de julho de 2023.
- [24] WAITE, Simon. **Liveness and Interactivity in Popular Music**. Innovation in Music Performance, Production, Technology, and Business, 1st Edition. 2019.

- [25] FELS, Sidney., GADD, Ashley. MULDER, Axel. **Mapping Transparency Through Metaphor: Towards More Expressive Musical Instruments.** 2002. Organised Sound. Páginas 109–126.
- [26] CROFT, John. **Theses on Liveness.** 2007. Organised Sound. Páginas. 59–66.
- [27] AUSLANDER, Philip. **Liveness: Performance in a Mediatized Culture.** 2nd edition. 2008. Editora Routledge.
- [28] SANDEN, Paul. **Liveness in Modern Music: Musicians, Technology, and the Perception of Performance.** 2013. Editora Routledge.
- [29] VUVAN, Dominique, PODOLACK, Olivia, SCHMUCKLER, Mark. **Memory for musical tones: the impact of tonality and the creation of false memories.** 2014. Sec. Auditory Cognitive Neuroscience Disponível em: <<https://doi.org/10.3389/fpsyg.2014.00582>>. Acesso em 02 de julho de 2023.
- [30] HARASIM, Daniel, MOSS, Fabian, RAMIREZ, Matthias, ROHRMEIER. **Exploring the foundations of tonality: statistical cognitive modeling of modes in the history of Western classical music.** 2021. Disponível em: <<https://www.nature.com/articles/s41599-020-00678-6>>. Acesso em 2 de julho de 2023.
- [31] SUNDARARAJAN, Han Se, PURVES, Lake J. **Co-Variation of Tonality in the Music and Speech of Different Cultures.** 2011. Disponível em: <<https://doi.org/10.1371/journal.pone.0020160>>. Acesso em 2 de julho de 2023.
- [32] DOĞANTAN-DACK, Mine. **Tonality: The Shape of Affect.** 2013. Vol 8, Nos 3-4 (2013): Empirical Musicology Review.
- [33] LONG, L. Hu, JIN, F. **Analysis of Main Characteristics of Music Genre Based on PCA Algorithm.** 2021. IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE).
- [34] FRAISSE, Paul. **Rhythm and tempo in The Psychology of Music,** 1982. Academic Press. Páginas 149–180
- [35] BELLA, Dalla. **A developmental study of the affective value of tempo and mode in music.** 2001.

- [36] JOHANSSON, Kg. **The Harmonic Language of The Beatles**. 1999. Luleå University of Technology. School of Music in Piteå.
- [37] UHLE, Christian, HERRE, Juergen. **Estimation of tempo, micro time and time signature from percussive music**. 2003.
- [38] DOWNIE, J. Stephen. **Music Information Retrieval**. Annual Review of Information Science and Technology. 2003. Páginas 295–340.
- [39] SCHEDL, Markus; Gómez, Emilia; Urbano, Julián. **Music Information Retrieval: Recent Developments and Applications**. 2014.
- [40] MOSCHETTA, Paulo, VIEIRA, José. **Música na era do streaming: curadoria e descoberta musical no Spotify**. Sociologias [Internet]. 2018. Disponível em: <<https://doi.org/10.1590/15174522-02004911>>. Acesso em: 2 de Julho de 2023.
- [41] STURM, Bob L, BEN-TAL Oded, MONAGHAN, Úna, COLLINS, Nick, Herremans, Dorien. **Machine Learning Research that Matters for Music Creation: A Case Study**. Journal of New Music Research. 2019. Páginas 36 a 55.
- [42] KOELLREUTTER, Hans-Joachim. **Harmonia Funcional**: Introdução à teoria das funções harmônicas. 3ª edição. São Paulo: Ricordi, Janeiro de 2008. 74 p.
- [43] FARNUM, Theo. **Trippy Sounds**: Recording Studio Effects of Psychedelic Rock, 1960s and Present. 2020. Tese do curso de Bacharel em Artes. Universidade do Estado da Califórnia. Disponível em https://digitalcommons.csumb.edu/caps_thes_all/786/?utm_source=digitalcommons.csumb.edu%2Fcaps_thes_all%2F786&utm_medium=PDF&utm_campaign=PDFCoverPages. Acesso em 16/11/2023.
- [44] PAVAN, G., BUDNEY, G., KLINCK, H., GLOTIN, H., CLINK, D.J., THOMAS, J.A. History of Sound Recording and Analysis Equipment. **Exploring Animal Behavior Through Sound**. Suíça, v.1, 2020. Disponível em https://doi.org/10.1007/978-3-030-97540-1_1. Acesso em 16/11/2023.
- [45] CLARK, Christopher. The Loudness Wars: Why Music Sounds Worse. **The Decade In Music: '00s**. Estados Unidos da América, 2009. Disponível em <https://www.npr.org/2009/12/31/122114058/the-loudness-wars-why-music-sounds-worse>.
- [46] DANTAS, J. D.; CRUZ, S. DA S.. Um olhar físico sobre a teoria musical. **Revista Brasileira de Ensino de Física**, v. 41, n. 1, 2019.
- [47] Benevenuto, F., Ribeiro, F., and Araújo, M. (2015). Métodos para análise de sentimentos em mídias sociais. **Simpósio Brasileiro de Multimídia na Web**, Manaus, Brasil.
- [48] SCHAEFER, Robert L. Alternative estimators in logistic regression when the data are collinear. **Journal of Statistical Computation and Simulation**, páginas 25:75-91. 20 de março de 2007. Disponível em

- <https://www.tandfonline.com/doi/abs/10.1080/00949658608810925>. Acesso em 16/11/2023.
- [49] Souza, Renato Rocha; CAFÉ, Lígia Maria Arruda. Análise de Sentimento aplicada ao Estudo de Letras de Música. **INFORMAÇÃO & SOCIEDADE (UFPB. ONLINE)**, v. 28, p. 275-286, 2018.
- [50] MONTGOMERY, Douglas, PECK, Elizabeth, VINING, Geoffrey. Introduction to Linear Regression Analysis. **Wiley Series in Probability and Statistics**. Editora Wiley, 2012. Disponível em <https://books.google.com.br/books?id=0yR4KUL4VDkC>.
- [51] VIJVERBERG, Wim. **Testing for IIA with the Hausman-Mcfadden Test**. SSRN Electronic Journal. 2011. Disponível em https://www.researchgate.net/publication/228251943_Testing_for_IIA_with_the_Hausman-Mcfadden_Test. Acesso em 16/11/2023.
- [52] NARKHEDE, Sarang. Understanding AUC-ROC Curve. **Towards Data Science**. Postado em 26 de junho de 2018. Disponível em <https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5>. Acesso em 16/11/2023.
- [53] HOSMER, David. W., LEMESHOW, Stanley. **Applied logistic regression**. Vol 1. 2000. Editora John Wiley Sons.
- [54] HAGHBAYAN, Hourmazd. COOMES, Eric. CURRAN, David. Temporal Trends in the Loudness of Popular Music over Six Decades. **Journal of general internal medicine**. 2020. Páginas 394–395. Disponível em <https://doi.org/10.1007/s11606-019-05210-4>. Acesso em 16/11/2023.
- [55] BRADLEY, Margaret. LANG, Peter. **Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings**. 1999. Reporte Técnico do Centro de Pesquisa em Sociopsicologia, Universidade da Flórida.
- [56] BEAUMONT, Rea. **AI as DJ: Artificial Intelligence as the Arbiter of Musical Taste in Digital Media Platforms and the Cultural Continuum**. 2020. Faculdade de Música da Universidade de Toronto.
- [57] SMITH, Thomas. MCKENNA, Cornelius. **A Comparison of Logistic Regression Pseudo R² Indices**. 2013. Multiple Linear Regression Viewpoints, Volume 39.

8 ANEXOS

8.1 IMPORTÂNCIA CULTURAL E HISTÓRICA DOS BEATLES

A importância histórica e cultural dos Beatles é indiscutível. A banda composta por John Lennon, Paul McCartney, George Harrison e Ringo Starr emergiu na década de 1960 e conquistou o mundo, deixando um legado duradouro que transcende gerações. O compositor americano Aaron Copland havia dito que quando as pessoas pedem para recriar o clima dos anos 60, elas tocam músicas dos *Beatles* (MACDONALD, 1997). Estudar os Beatles é

fundamental para compreendermos o impacto profundo que eles tiveram e ainda têm na história e cultura contemporânea, tal como apreendemos de FRIEDLANDER (2003, p. 147-148):

Os Beatles exploraram a criatividade, foram além dos limites, se tornando o maior catalisador do amadurecimento da música rock. Eles fizeram inovações musicais importantes e ajudaram a transformar a indústria musical. Os temas e as ideias contidas em suas canções refletiam a consciência não só dos membros da banda, mas também da crescente contracultura da época. Tradução minha.

Historicamente, os *Beatles* surgiram em um momento de efervescência social e cultural. Eles se tornaram a trilha sonora de uma época marcada por mudanças significativas, como o movimento pelos direitos civis, a Guerra do Vietnã e a revolução sexual. Como citado em (Frith, 1999, p. 22-25):

O rock era algo mais que o pop, mais que o rock'n'roll. Os músicos de rock combinavam uma ênfase de virtuosismo e técnica com a concepção romântica da arte como expressão individual, original e sincera. Eles alegavam não ser comerciais – a lógica de sua música não era a de fazer dinheiro ou de encontrar um mercado; eles estavam, na verdade, desdenhando o pop comercial. Entretanto, a música feita naqueles princípios aparentemente não-comerciais veio a ter seu próprio público. Os fãs da música pop também estavam mudando, preferindo os álbuns aos compactos, investindo em som, usando música como fonte de instrução pessoal e experiência, como também para dançar e divertir-se. O rock abasteceu o lado rebelde da juventude do final do anos 60: o movimento contra a guerra do Vietnã – dentro e fora do exército -, as revoltas nos campi das universidades, as celebrações hedonísticas de drogas e sexo. Tradução minha.

Suas músicas refletiam e influenciavam as atitudes e os sentimentos da juventude da época, tornando-se um símbolo de liberdade, expressão e contestação.

Desde o início, os Beatles eram contraculturais, seu trabalho contrariando os rígidos estilos de vida ditados pelo consumismo dos

anos 1950. Sua música (não necessariamente suas letras, mas sua música real) rompeu barreiras de raça e classe e evocou respostas emocionais de seu público. (GALLAHER, 2021). Tradução minha.

Culturalmente, os *Beatles* transformaram a música popular. Eles introduziram inovações musicais, explorando novos estilos, harmonias e arranjos -- , além de incorporar elementos de outras culturas, como a música indiana. É possível encontrar diversos artigos acadêmicos que demonstram a relevância de tais fatos, citado Guerrero (2015):

A década de 1960 foi um período de experimentação na música popular. Os Beatles, a banda de rock britânica que dominou a cena da música popular durante essa década, foi responsável por inovações duradouras que moldaram a história e o estilo da música pop e da cultura ocidental. Uma das contribuições mais significativas, porém negligenciadas, dos Beatles para a cultura ocidental foi a incorporação de elementos musicais indianos em suas canções. Esta análise discutirá o papel desempenhado pelos Beatles na popularização de certos elementos da música e cultura indiana no hemisfério ocidental.

Sgt. Pepper's Lonely Hearts Club Band, um grande exemplo e, provavelmente, o mais importante álbum da banda, foi lançado em 1967. Este álbum é amplamente reconhecido como o melhor deles - por exemplo, por DORŮŽKA (1981) -, e desempenha um papel crucial no desenvolvimento da música popular por diversos fatores. GAMMOND (1993) descreveu o álbum como uma mistura de surrealismo, misticismo, vaudeville e rock, que elevou a música pop a novos patamares de inventividade. Trata-se do primeiro álbum conceitual (WICKE E ZIEGENRÜCKER, 2001): um álbum que não é composto por diferentes músicas individuais juntas, mas planejado como uma obra completa. Seu som distintivo e revolucionário influenciou não apenas outras bandas da época, mas também moldou o futuro da música, abrindo caminho para o rock progressivo, o rock psicodélico e o rock alternativo (CRAWFORD, 2005).

Em 1967, o álbum conceitual encontrou seu lugar na música pop e rock quando George Martin produziu o aclamado álbum de estúdio *Sgt. Pepper's Lonely Hearts Club Band*. Embora John Lennon tenha negado que suas músicas tenham sido escritas com um tema geral em mente, o álbum ainda se apresenta como um álbum conceitual. Isso se

deve à faixa inicial, "*Sgt. Pepper's Lonely Hearts Club Band*", à faixa relacionada "*With a Little Help from My Friends*" e à eventual reprise da faixa-título perto do final do lado B. Outros fatores contribuintes incluem a capa do álbum, que apresenta os membros dos Beatles vestidos como a persona da banda, *Sgt. Pepper's Lonely Hearts Club Band*, a multitude de críticas positivas que compararam o álbum a obras-primas como "*Guernica*" de Picasso, e a inspiração que ele proporcionou a outras bandas de rock no que diz respeito ao álbum conceitual. Em última análise, a ideia de um conceito unificado apresentado em *Sgt. Pepper's* permitiu que o álbum se tornasse uma obra de arte elevada, algo inédito na música pop e rock na época. (KLEIN, 2019). Tradução minha.

Além disso, os Beatles abordavam questões sociais e políticas em suas letras. Trouxeram à tona questões como paz, amor - Eles eram seguidores de, entre outros, Chuck Berry, Little Richard e especialmente Elvis Presley, e, portanto, não é surpresa para ninguém que a maioria de suas músicas fosse sobre amor. No primeiro período de sua carreira, eles escreviam letras quase exclusivamente sobre amor. De um total de 100 músicas (76 músicas originais e 24 versões de covers) gravadas entre 1962 e 1965, apenas 14 não são explicitamente sobre amor, e mesmo nessas, há algumas alusões ao amor, garotas, mulheres ou relacionamentos (HECL, 2009), igualdade racial e críticas à guerra, estimulando o pensamento crítico e a reflexão em seus ouvintes.

Mais tarde, os Beatles usariam esse gênero para criticar e ridicularizar a sociedade ou tipos específicos de pessoas. "Eleanor Rigby", do álbum *Revolver* de 1966, que não é uma música cômica, mas que definitivamente pode ser considerada uma música narrativa, é uma música sobre uma solteirona que frequentemente vai à igreja e é muito solitária. McCartney canta que ela usa "o rosto que ela mantém em um pote perto da porta" e pergunta: "Para quem é?" No final, Eleanor Rigby morre e ninguém comparece ao seu funeral. A música é um ataque óbvio aos moralismos reprimidos da geração de pais e avós, que são entediantes e inúteis aos olhos dos jovens. "*The Fool on the Hill*" (do álbum *Magical Mystery Tour*, 1967) é uma crítica à cegueira e ao pensamento estreito das pessoas, e à tendência delas de

considerar qualquer um que seja diferente de um tolo. (HECL, 2009).
Tradução minha.

O impacto do fenômeno do “*fab-four*” e seu legado continuam a influenciar artistas, músicos e admiradores em todo o mundo. Isso se torna algo notável ao observar que, de acordo com o *Guinness Book*, dentre as 10 músicas que mais foram reinterpretadas na história da música, 3 pertencem aos Beatles, sendo essa a única banda (ou artista) a figurar mais de uma vez no ranking (GUINNESS, 2005).

8.2 ESTUDOS SOBRE CARACTERÍSTICAS MUSICAIS

8.2.1 Acorde

Um grupo de no mínimo três notas diferentes, quando tocadas simultaneamente, forma uma estrutura harmônica denominada de acorde. (DANTAS; CRUZ, 2019).

Essas notas são geralmente selecionadas a partir de uma escala musical e são combinadas para criar harmonia na música (KOELLREUTTER, 2008). Os acordes são uma parte essencial da teoria musical e têm um grande impacto na sonoridade e na expressão musical. Consistem em elementos essenciais que contribuem para a sua identidade e qualidade sonora. Dois desses elementos fundamentais são as notas fundamentais e os intervalos.

Os acordes são frequentemente nomeados com base na nota mais baixa, que é chamada de "nota fundamental" ou "tônica" (DANTAS; CRUZ, 2019). Essa nota dá o nome ao acorde. Por exemplo, em um acorde de C maior, a nota fundamental é o C (KOELLREUTTER, 2008).

Os intervalos entre as notas que compõem um acorde são determinantes para a qualidade e o tipo do acorde. Intervalos comuns incluem terças (3ª), quintas (5ª) e sétimas (7ª) (DANTAS; CRUZ, 2019).

Acordes podem ser classificados de várias maneiras com base na estrutura das notas que os compõem (KOELLREUTTER, 2008). Acordes maiores são construídos a partir de uma tônica, uma terça maior e uma quinta justa. Por exemplo, o acorde de C maior consiste nas notas C, E e G (DANTAS; CRUZ, 2019). São geralmente associados a emoções mais alegres, luminosas e otimistas.

Acordes menores são semelhantes aos acordes maiores, mas com uma terça menor em vez de uma terça maior (DANTAS; CRUZ, 2019). Por exemplo, o acorde de A menor consiste nas notas A, C e E. Acordes menores tendem a criar uma atmosfera mais melancólica e reflexiva. Sua sonoridade dissonante sugere um estado de tensão emocional.

Em resumo, compositores e músicos utilizam uma variedade de acordes para moldar a experiência emocional da música. A escolha entre acordes maiores, menores e outros tipos desempenha um papel fundamental na criação de uma paisagem sonora que ressoa com o público, influenciando como a música é percebida e sentida (KOELLREUTTER, 2008).

8.2.2 Tonalidade

Um conceito fundamental na música ocidental é a tonalidade, que engloba as relações entre vários elementos musicais, como tons, acordes e tonalidades. Esses elementos são organizados em uma estrutura hierárquica centrada em um tom de referência, estabelecendo uma sensação de estabilidade e ordem (VUVAN, 2014).

A tonalidade na música ocidental refere-se à interação de elementos musicais dentro de uma estrutura hierárquica centrada em torno de um tom de referência. As tonalidades maiores e menores são duas formas comumente usadas, cada uma com sua própria hierarquia tonal. Embora ambas as formas sigam uma organização hierárquica, elas diferem na colocação das notas dentro dessa hierarquia. A tonalidade maior e a tonalidade menor têm, cada uma, suas hierarquias de notas distintas (SCHMUCKLER, 2014).

A tonalidade é um conceito fundamental para analisar a música clássica ocidental e influencia nosso reconhecimento de padrões de movimento e estabilidade percebida. O conceito de tonalidade fornece uma estrutura para a compreensão de suas progressões harmônicas, relações de acordes e modulações de tonalidade que moldam as composições musicais.

Vale notar que as diferenças tonais entre a música tradicional oriental e ocidental são facilmente perceptíveis, e seu contraste é evidente através de seus distintos sistemas tonais e escalas (DOĞANTAN-DACK, 2013). As tradições musicais orientais geralmente têm sistemas tonais e escalas distintas que diferem do sistema ocidental. Essas diferenças contribuem para os sons e “sabores” únicos associados à música oriental. (SUNDARARAJAN, 2021).

8.2.3 Modo

Em complemento ao conceito acima, Harasim (2021) define *Mode* como uma série de sete escalas musicais, cada uma derivada da escala “diatônica” – o que significa que elas têm sete notas e incluem dois intervalos que são semitons (meios tons) e cinco intervalos que são tons (tons inteiros).

Uma explicação mais facilmente compreensível é que uma canção em modo maior, por exemplo, segue de uma sequência específica de acordes que é construída com base nas notas da escala maior correspondente (DANTAS; CRUZ, 2019). Por exemplo, se fixarmos um acorde como tonalidade, a sequência de acordes comuns nesse modo incluirá acordes como tônica (como sendo um acorde maior), segunda maior, terça maior, quarta justa, quinta justa, sexta maior e sétima maior. Já uma música em modo menor segue uma sequência de acordes que é construída com base nas notas da escala menor correspondente, logo os acordes comuns incluirão tônica (como sendo um acorde menor), segunda maior, terça menor, quarta justa, quinta justa, sexta menor e sétima menor (DANTAS; CRUZ, 2019).

8.2.4 Mudança de Tonalidade

A capacidade de variar as tonalidades dentro de uma música é uma técnica musical sofisticada que permite aos compositores explorar emoções e atmosferas. Em vez de aderir a uma única tonalidade ao longo de toda a composição, se introduz em alguns momentos mudanças sutis ou drásticas em seções das músicas (JOHANSSON, 1999).

8.2.5 Fórmula de Compasso

Fórmula de compasso, ou *time signature*, é um recurso rítmico básico para uma ampla variedade de músicas (UHLE, HERRE, 2003). *Time signature* é uma notação musical que indica a organização rítmica de uma música. Consiste em dois valores em sobreposição colocados no início de uma partitura. A parte superior indica o número de pulsos por medida, enquanto a cifra inferior representa a figura de tempo que recebe um pulso. Por exemplo, um compasso 4/4 indica que há quatro pulsos por medida e que cada pulso é uma figura de tempo igual à denominada “semínima”. Já um compasso 3/4 indica que há três pulsos por medida e que cada pulso é uma figura de tempo igual a uma semínima. A *time signature* é essencial para determinar o ritmo e a subdivisão dos compassos em uma composição musical (UHLE, HERRE, 2003).

8.2.6 Dançabilidade

A dançabilidade de uma música refere-se à sua capacidade de inspirar o movimento e a dança nas pessoas. É uma medida subjetiva que avalia a facilidade com que uma música pode ser dançada e o seu potencial de animar e envolver o ouvinte.

A *danceability* descreve quão adequada uma faixa é para dançar com base em uma combinação de elementos musicais, incluindo andamento, estabilidade do ritmo, força da batida e regularidade geral. Um valor menor é menos dançável e um valor maior é mais dançável. (HOWLIN, ROONEY, 2021). Tradução minha.

Liew (2022) descreve a *danceability* como uma medida de excitação rítmica.

8.2.7 Energia

Howlin e Rooney (2021) descrevem que energia, no contexto musical, é uma medida que representa uma medida perceptiva de intensidade e atividade. Normalmente, as faixas energéticas parecem rápidas, altas e barulhentas:

Por exemplo, death metal tem alta energia, enquanto um prelúdio de Bach pontua baixo na escala. Os recursos perceptivos que contribuem para esse atributo incluem faixa dinâmica, sonoridade percebida, timbre, taxa de início e entropia geral. Tradução minha.

Culturas do Leste Asiático geralmente preferem músicas com níveis mais altos de energia, e canções enérgicas estão fortemente associadas à regulação da raiva, em detrimento da regulação da tristeza e da ansiedade (HOWLIN, ROONEY, 2021). Especula-se que a energia, como um elemento de intensidade musical, possa representar uma forma de regulação interna (controle) que as pessoas utilizam para "abafar" a raiva, o que seria mais prevalente nas culturas do Leste Asiático devido às normas socioculturais de expressão emocional. Por outro lado, isso difere da regulação externa associada à intensidade rítmica da música (ou seja, a capacidade de dançar) (HOWLIN, ROONEY, 2021).

No contexto da música ocidental, as preferências podem variar amplamente, com diferentes estilos e gêneros musicais apreciados por diferentes pessoas. Embora seja difícil generalizar, é possível encontrar semelhanças com o que foi mencionado sobre as culturas do Leste Asiático. Por exemplo, certos gêneros musicais ocidentais, como o rock ou o heavy metal, são

conhecidos por sua energia e intensidade sonora, e muitas vezes são apreciados por sua capacidade de liberar emoções intensas, como a raiva, de forma catártica (HOWLIN, ROONEY, 2021). No entanto, também existem gêneros musicais ocidentais que valorizam a expressão emocional mais suave, como o pop ou a música clássica, que podem enfatizar a melodia e a contemplação.

8.2.8 Acústicidade

Howlin e Rooney (2021) exploram o atributo *acousticness*, que é uma estimativa de quão acústica é uma determinada música. As músicas com alta acústicidade consistirão principalmente de sons acústicos naturais (como violão acústico, piano, orquestra e voz humana não processada), enquanto as músicas com baixa *acousticness* serão compostas principalmente por sons elétricos (como guitarras elétricas, sintetizadores, baterias eletrônicas, vocais *autotunados*, entre outros). Howlin e Rooney (2021) ressaltam a diferença entre os dois extremos e enfatizam que o nível de *acousticness* de uma música pode influenciar seu estilo e sonoridade geral.

No trecho a seguir, temos noção de como foi treinado um modelo para obtenção do *feature* de uma base de dados pelo *Echo Nest* API, de 2013 e que fornece bases de dados para *MIR*):

Derivamos esse novo parâmetro diretamente do áudio. Construímos e treinamos modelos de máquinas que previram a acústica com base no sinal de áudio. Isso significa que podemos aplicar facilmente esse modelo aos nossos muitos milhões de músicas, permitindo-nos obter uma estimativa da acústica de praticamente qualquer música que já tenha sido gravada. *EchoNest* (2013).

8.2.9 Valência Emocional

Howlin e Rooney (2021) descrevem *Valence* como

[...] uma medida que descreve a positividade musical transmitida por uma faixa. Faixas com alta valência soam mais positivas (por exemplo, feliz, alegre, eufórica), enquanto faixas com baixa valência soam mais negativas (por exemplo, triste, deprimido, zangado). Tradução minha.

Long (2021) realizou um estudo que demonstra que podemos pensar que entre os artistas ou a música, quanto mais altos forem os índices de energia, valência, andamento, contagem e vivacidade, mais a música se enquadra nas características do gênero *Pop Rock* e mais popular ela é entre os fãs fiéis do estilo.

8.2.10 Instrumentalidade

Howlin e Rooney (2021) dizem que a *instrumentalness* identifica se uma faixa não contém vocais. Os sons “Ooh” e “aah” são tratados como instrumentais neste contexto. Faixas de rap ou um áudio de piadas são claramente “vocais”. Na *API* do Spotify, quanto mais próximo de 1 estiver o valor da instrumentalidade, maior a probabilidade de a faixa não conter nenhum conteúdo vocal. Valores pequenos destinam-se a representar faixas instrumentais, mas a confiança é maior à medida que o valor se aproxima.

8.2.11 Vivacidade

Para compreender o conceito de *liveness*, foi consultado Waite (2019), que define o termo a partir de diversas outras referências. O texto discute a evolução do conceito de “*liveness*” (vivacidade ou presença ao vivo) nas performances artísticas, especialmente no contexto da tecnologia e da mídia. Anteriormente, a distinção binária entre uma performance ser “ao vivo” ou “não ao vivo” era predominante, mas agora essa distinção foi abandonada em favor de um contínuo.

Diferentes teorias e perspectivas foram propostas para entender o conceito de *liveness*. Sidney Fels e seus colegas apresentaram o conceito de “transparência”, que está relacionado à efetividade da tradução da entrada gestual para a saída sonora. Em outras palavras, é a qualidade da correspondência entre as ações do *performer* e os resultados sonoros produzidos (Fels, 2002).

Auslander (2008) discute como o uso de materiais midiáticos nas performances ao vivo reduz a *liveness*, ou seja, diminui os valores tradicionais da performance. Isso ocorre porque a incorporação de elementos pré-gravados ou mediatizados pode afetar a autenticidade e a imprevisibilidade associadas às performances ao vivo.

Croft (2007) faz uma distinção entre “*liveness* procedural”, em que os eventos são gerados em tempo real, e “*liveness* estético”, em que as ações perceptíveis e significativas do performer são mapeadas para resultados perceptíveis e significativos (Croft, 2007). Essa distinção está

relacionada à ideia do "sublime tecnológico", em que as ferramentas tecnológicas em si se tornam objetos estéticos.

Sanden (2013) sugere que a percepção geral de *liveness* surge de uma rede de diferentes aspectos de vivacidade. Ele divide a *liveness* em três componentes, sendo

Liveness espacial-temporal: refere-se à presença de um agente (performer) no tempo e no espaço físico da performance.

Liveness de fidelidade: está relacionada a quão fiel a performance é em relação à sua versão inicial. Isso envolve questões de reprodução precisa e repetibilidade.

Liveness de espontaneidade: indica o grau em que a performance é fixa ou improvisada. Quanto mais espontânea e imprevisível for a performance, maior será sua *liveness* nesse aspecto.

Essas diferentes perspectivas e componentes destacam a complexidade do conceito de *liveness* nas performances contemporâneas, levando em consideração fatores como tecnologia, autenticidade, imprevisibilidade e interação entre o performer e o público.

Apesar de todas essas possíveis abstrações do termo, a *API* utilizada para a construção deste trabalho utiliza-se da detecção da

“[...] presença de público na gravação. Valores mais altos de vivacidade representam uma probabilidade aumentada de que a faixa foi tocada ao vivo”. Howlin e Rooney (2021). Tradução minha.

Dessa forma, o conceito é reduzido apenas ao que Croft (2007) identifica como “*liveness* estético”, pois até mesmo características que Auslander (2008) rotula como materiais midiáticos nas performances (por exemplo, utilização de falsa plateia na gravação) são identificadas pelo modelo como “ao vivo”.

8.2.12 Volume

A *loudness*, ou volume, é um dos aspectos fundamentais da qualidade sonora em uma faixa musical. Ela desempenha um papel crucial na percepção da música e pode influenciar significativamente a experiência do ouvinte.

Howlin e Rooney (2021) definem a *loudness* como a intensidade global do som em uma faixa musical, medida em decibéis (dB). A *loudness* é uma característica fundamental da música

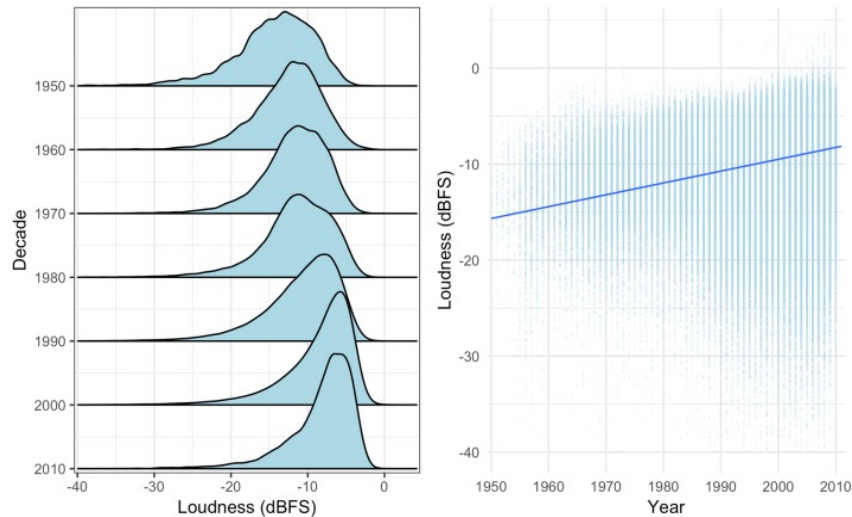
que reflete a percepção subjetiva do volume. Em outras palavras, é a qualidade do som que está diretamente associada à força física da onda sonora, ou seja, sua amplitude. Quando falamos sobre *loudness*, estamos nos referindo à sensação de quão alto ou suave um som parece para o ouvinte.

Theo Farnum (2020) explica que durante a transição da década de 1960 para a década de 1970, houve avanços significativos nos equipamentos de áudio e técnicas de gravação que possibilitaram um maior controle da *loudness* sem comprometer a qualidade do som. Nesse período, os estúdios de gravação começaram a adotar tecnologias mais avançadas (ERBE; THOMAS, 2014), como mixers de áudio mais sofisticados, gravadores de fita de alta qualidade e processadores de sinal mais precisos. Esses avanços permitiram aos produtores e engenheiros de áudio uma manipulação mais refinada da *loudness* durante a mixagem e a masterização, resultando em faixas musicais mais equilibradas e dinâmicas. Além disso, a introdução de técnicas como a compressão e a equalização paramétrica possibilitou um controle mais preciso da *loudness*, permitindo que os artistas alcançassem o impacto desejado em suas músicas sem comprometer a fidelidade do som.

Todavia, em contrapartida,

[...] nos últimos trinta anos, e até mesmo antes disso, a música que ouvimos tem se tornado mais alta. E eu não quero dizer apenas mais agressiva e irritada. Ela está ficando fisicamente mais alta. Isso ocorre porque a música mais alta cria um efeito imediatamente mais agradável no ouvinte, e os executivos de gravadoras têm aumentado o volume por pelo menos três décadas. (CLARK, 2009). Traduzido por mim.

Figura 18 - *Loudness War* no decorrer dos anos




Fonte: HAGHBAYAN; COOMES; CURRAN (2019).

Na Figura 18, os autores abordam a *Loudness War*, fenômeno com o qual a música ao longo dos últimos anos tem se tornado mais alta, não apenas em termos de agressividade e intensidade emocional, mas também em relação ao volume físico. Isso significa que a música está sendo reproduzida a um volume mais alto. A razão para isso é que músicas mais altas tendem a criar um efeito imediatamente mais agradável no ouvinte. Isso ocorre porque o aumento do volume pode dar a impressão de que a música é mais envolvente e emocionante. A imagem abaixo introduz o conceito de dinâmica musical e descreve seu padrão de comportamento durante as décadas.


Figura 19 - Explicação da dinâmica musical

Understanding Dynamic


Dynamic is the distance between any part of a song.




Here is what a visual diagram of what a song might look like around 1979. Notice that the highest points in the sound still only peak at about -4 or -5. This is like a beautiful painting with a generous mat around it.



Here is what a popular song might look like in the early nineties. Notice how the peak is regularly hitting zero? This would be like if I took my painting and expanded it all the way to the edges.



By the early 00's, new digital technologies allowed producers to push the lower levels of the song up even louder without distorting the peaks. This would be like me blowing up my painting outside of the frame.



It is not uncommon, in more extreme situations, to see songs today that look like this. It would be the equivalent of blowing up our image so that only a small portion is visible.

Fonte: CLARK (2009).

8.2.13 Fala

Howlin e Rooney (2021) discutem a métrica *Speechiness* (em português, “fala”), que é capaz de detectar a presença de palavras faladas em uma faixa musical, sendo medida em uma escala de 0.0 a 1.0. Quanto mais a gravação se assemelha ao discurso (como em programas de entrevistas, audiolivros e poesia), mais próximo de 1.0 será o valor atribuído à *Speechiness*. No entanto, destaca-se que as músicas populares atuais têm letras, mas consistem principalmente de elementos musicais. Howlin e Rooney (2021) sugerem que um nível moderado de *Speechiness* - ou seja, músicas que contêm tanto música quanto letras - está provavelmente relacionado de forma positiva a um maior número de reproduções em serviços de *streaming*.

8.2.14 Tempo

Howlin e Rooney (2021) definem como o andamento geral estimado de uma faixa em batidas por minuto (bpm). Na terminologia musical, tempo é a velocidade ou ritmo de uma determinada peça e deriva diretamente da duração média da batida.

É a velocidade na qual as batidas são repetidas, a periodicidade das batidas da música. Geralmente é medido em batidas por minuto (bpm). Tal parâmetro é importante pois influencia diretamente a sensação de movimento, energia e emoção transmitidas por uma determinada composição. O tempo é um importante fator que contribui para a percepção do ritmo (FRAISSE, 1982).

Ele pode variar amplamente de uma música para outra, e até mesmo dentro da mesma música, criando diferentes atmosferas e estilos musicais. Também está intimamente ligado à expressão musical. O mesmo trecho musical pode transmitir emoções diferentes dependendo do andamento escolhido pelo intérprete. Fortes associações foram observadas entre andamento rápido e lento na música e classificações de felicidade e tristeza (BELLA, 2001), um aumento no andamento pode intensificar a emoção e criar uma sensação de urgência, enquanto uma diminuição no andamento pode transmitir uma atmosfera mais serena ou melancólica.

Movimentos espontâneos podem estar intimamente relacionados com as previsões de rajadas locais de energia no fluxo de áudio musical, em particular para a batida e os padrões rítmicos. Essa utilização do corpo é o núcleo conceito de cognição incorporada, que afirma que o

corpo é envolvido ou mesmo necessário para processos cognitivos. (BELLA, 2001). Tradução minha.

8.3 EFEITOS ESPECÍFICOS

Characteristic	log(OR)	95% CI	p-value	Characteristic	log(OR)	95% CI	p-value	Characteristic	log(OR)	95% CI	p-value
Dissolution				Psychedelic				Transition			
DistinctWordsInLyrics	194	192, 196	<0.001	DistinctWordsInLyrics	495	494, 496	<0.001	DistinctWordsInLyrics	212	211, 213	<0.001
ChorusPresence				ChorusPresence				ChorusPresence			
No	—	—		No	—	—		No	—	—	
Yes	25.028	25.028, 25.028	<0.001	Yes	48.173	48.173, 48.173	<0.001	Yes	24.512	24.512, 24.512	<0.001
MainKey	-458	-458, -458	<0.001	MainKey	-3.343	-3.343, -3.343	<0.001	MainKey	-475	-475, -475	<0.001
Mode				Mode				Mode			
Major	—	—		Major	—	—		Major	—	—	
Minor	-6.564	-6.564, -6.564	<0.001	Minor	-11.532	-11.532, -11.532	<0.001	Minor	-7.460	-7.460, -7.460	<0.001
ChangesKey				ChangesKey				ChangesKey			
No	—	—		No	—	—		No	—	—	
Yes	-21.632	-21.632, -21.632	<0.001	Yes	-53.428	-53.428, -53.428	<0.001	Yes	-17.922	-17.922, -17.922	<0.001
TimeSignatureIn				TimeSignatureIn				TimeSignatureIn			
Four	—	—		Four	—	—		Four	—	—	
Three	31.064	31.064, 31.064	<0.001	Three	59.239	59.239, 59.239	<0.001	Three	33.159	33.159, 33.159	<0.001
Energy	-93.094	-93.094, -93.094	<0.001	Energy	-183.832	-183.832, -183.832	<0.001	Energy	-96.971	-96.971, -96.971	<0.001
Acousticness	-40.317	-40.317, -40.317	<0.001	Acousticness	-90.595	-90.595, -90.595	<0.001	Acousticness	-40.603	-40.603, -40.603	<0.001
Valence	-15.368	-15.368, -15.368	<0.001	Valence	-9.826	-9.826, -9.826	<0.001	Valence	-10.019	-10.019, -10.019	<0.001
Instrumentalness	713.837	713.837, 713.837	<0.001	Instrumentalness	-397.955	-397.955, -397.955	<0.001	Instrumentalness	711.602	711.602, 711.602	<0.001
Liveness	-9.387	-9.387, -9.387	<0.001	Liveness	-19.026	-19.026, -19.026	<0.001	Liveness	-17.444	-17.444, -17.444	<0.001
Loudness	-6.994	-6.994, -6.994	<0.001	Loudness	1.576	1.576, 1.576	<0.001	Loudness	-5.488	-5.488, -5.488	<0.001
Speechiness	269.373	269.373, 269.373	<0.001	Speechiness	1.034.659	1.034.659, 1.034.659	<0.001	Speechiness	278.734	278.734, 278.734	<0.001
Tempo	-229	-230, -228	<0.001	Tempo	-538	-538, -538	<0.001	Tempo	-233	-234, -232	<0.001
DurationMs	0.41	0.35, 0.47	<0.001	DurationMs	0.92	0.86, 0.99	<0.001	DurationMs	0.40	0.33, 0.46	<0.001
ChordsQuantity	-3.365	-3.365, -3.365	<0.001	ChordsQuantity	2.951	2.951, 2.951	<0.001	ChordsQuantity	-3.394	-3.394, -3.394	<0.001
ChordsComplexity	50.808	50.808, 50.808	<0.001	ChordsComplexity	23.920	23.920, 23.920	<0.001	ChordsComplexity	46.951	46.951, 46.951	<0.001
HasChordA				HasChordA				HasChordA			
No	—	—		No	—	—		No	—	—	
Yes	16.574	16.574, 16.574	<0.001	Yes	37.285	37.285, 37.285	<0.001	Yes	14.682	14.682, 14.682	<0.001
HasChordG				HasChordG				HasChordG			
No	—	—		No	—	—		No	—	—	
Yes	18.242	18.242, 18.242	<0.001	Yes	135	135, 135	<0.001	Yes	17.948	17.948, 17.948	<0.001
HasChordD				HasChordD				HasChordD			
No	—	—		No	—	—		No	—	—	
Yes	7.037	7.037, 7.037	<0.001	Yes	-15.718	-15.718, -15.718	<0.001	Yes	5.048	5.048, 5.048	<0.001
HasChordC				HasChordC				HasChordC			
No	—	—		No	—	—		No	—	—	
Yes	14.571	14.571, 14.571	<0.001	Yes	31.488	31.488, 31.488	<0.001	Yes	16.324	16.324, 16.324	<0.001
HasChordF				HasChordF				HasChordF			
No	—	—		No	—	—		No	—	—	
Yes	27.828	27.828, 27.828	<0.001	Yes	42.964	42.964, 42.964	<0.001	Yes	25.694	25.694, 25.694	<0.001
HasChordD7				HasChordD7				HasChordD7			
No	—	—		No	—	—		No	—	—	
Yes	18.329	18.329, 18.329	<0.001	Yes	-3.489	-3.489, -3.489	<0.001	Yes	16.273	16.273, 16.273	<0.001
HasChordA7				HasChordA7				HasChordA7			
No	—	—		No	—	—		No	—	—	
Yes	-19.932	-19.932, -19.932	<0.001	Yes	47.235	47.235, 47.235	<0.001	Yes	-20.160	-20.160, -20.160	<0.001
HasChordAm				HasChordAm				HasChordAm			
No	—	—		No	—	—		No	—	—	
Yes	-3.468	-3.469, -3.468	<0.001	Yes	-49.879	-49.879, -49.879	<0.001	Yes	-4.607	-4.607, -4.607	<0.001
HasChordBm				HasChordBm				HasChordBm			
No	—	—		No	—	—		No	—	—	
Yes	-2.923	-2.923, -2.923	<0.001	Yes	-58.558	-58.558, -58.558	<0.001	Yes	880	880, 880	<0.001
ComposingConceptByMcCartney	29.261	29.261, 29.261	<0.001	ComposingConceptByMcCartney	54.243	54.243, 54.243	<0.001	ComposingConceptByMcCartney	29.356	29.356, 29.356	<0.001
ComposingConceptByHarrison	52.891	52.891, 52.891	<0.001	ComposingConceptByHarrison	67.160	67.160, 67.160	<0.001	ComposingConceptByHarrison	52.006	52.006, 52.006	<0.001
ComposingConceptByLennonMcCartney	-3.745	-3.745, -3.745	<0.001	ComposingConceptByLennonMcCartney	15.481	15.481, 15.481	<0.001	ComposingConceptByLennonMcCartney	4.018	4.018, 4.018	<0.001
PCALyrcalAnalysis1	9.739	9.739, 9.739	<0.001	PCALyrcalAnalysis1	22.403	22.403, 22.403	<0.001	PCALyrcalAnalysis1	9.649	9.649, 9.649	<0.001
PCALyrcalAnalysis2	-5.140	-5.140, -5.140	<0.001	PCALyrcalAnalysis2	14.886	14.886, 14.886	<0.001	PCALyrcalAnalysis2	-5.821	-5.821, -5.821	<0.001