

Influência do clima na escolha musical

Henrique Alvelos^{1,2*}

^{1*}Escola de Engenharia, Universidade do Minho, Rua da Universidade,
City, 4710-057, Braga, Portugal.

Corresponding author(s). E-mail(s): pg50414@uminho.pt;

Abstract

A influência da música na escolha musical é um campo de estudo que tem atraído a atenção de pesquisadores de diversas disciplinas. A psicologia da música, por exemplo, tem sido explorada por North e Hargreaves [1], que destacaram como a música pode influenciar o comportamento e as escolhas das pessoas. Eles argumentam que a música pode afetar as escolhas musicais de várias maneiras, desde a influência das preferências pessoais até o impacto das tendências culturais e sociais.

Além disso, a relação entre as preferências musicais e a personalidade também tem sido um foco de pesquisa. Rentfrow e Gosling [2] investigaram a estrutura e os correlatos de personalidade das preferências musicais. Eles descobriram que as preferências musicais são significativamente correlacionadas com a personalidade, sugerindo que a música que escolhemos pode refletir quem somos como indivíduos. Outro aspecto importante é a relação entre as funções da música e a preferência musical. Schäfer e Sedlmeier [3] exploraram essa relação e descobriram que as funções que atribuímos à música podem desempenhar um papel importante na música que escolhemos ouvir.

Keywords: Koppën-Geiger, Spotify, Weather, Dataset, Python, Pandas, ETL, Parquet, PowerBI

1 Introdução

A escolha musical dos indivíduos é um campo complexo, influenciado por uma vasta gama de fatores, que vão desde preferências pessoais e influências culturais até o estado emocional e o ambiente imediato. Um fator que tem recebido atenção crescente é a meteorologia. Diversos estudos sugerem que as condições climáticas podem afetar o humor e o comportamento humano, levantando a hipótese de que elas também

possam influenciar as preferências musicais. Este estudo visa explorar a relação entre as condições meteorológicas e a escolha de músicas, utilizando dados detalhados de climas, temperaturas e rankings semanais das músicas mais populares. Ao investigar essa conexão, pretendemos fornecer insights valiosos tanto para a indústria musical quanto para a compreensão das interações entre ambiente e comportamento humano.

2 Preparação do dataset

2.1 Definição dos objetivos

Correlação entre Clima e Escolha Musical:

- **Análise de Correlação:** Explorar possíveis correlações entre diferentes condições climáticas e a popularidade de certas propriedades de músicas. Utilizar gráficos e métricas de correlação no Power BI para visualizar e quantificar essas relações;
- **Associações Específicas:** Identificar associações específicas entre certas condições climáticas e propriedades de músicas que frequentemente aparecem nos rankings.

Segmentação por Condições Meteorológicas:

- **Segmentação de Dados:** Dividir o dataset em segmentos baseados em diferentes condições meteorológicas para identificar padrões distintos de preferência musical em cada segmento.
- **Análise Comparativa:** Utilizar funcionalidades de segmentação do Power BI para comparar os rankings musicais em diferentes segmentos climáticos, determinando se há preferências musicais específicas associadas a certas condições meteorológicas.

Inferir a Influência Climática:

- **Quantificação da Influência:** Desenvolver análises no Power BI para quantificar a influência das variáveis meteorológicas nas escolhas musicais. Utilizar gráficos e dashboards interativos para ilustrar essas influências.
- **Identificação de Tendências:** Utilizar as análises para prever quais tipos de músicas têm maior probabilidade de se tornarem populares sob determinadas condições climáticas.

2.2 Acesso aos Dados

Para realizar este estudo sobre a influência da meteorologia na escolha musical, foi necessário aceder e integrar uma vasta quantidade de dados de diferentes fontes.

O processo envolveu a utilização de técnicas de web scraping e APIs especializadas para a coleta dos dados necessários.

2.2.1 Músicas

Inicialmente, utilizou-se um scraper para obter as cidades disponíveis no website <https://charts.spotify.com/>. Esta etapa foi fundamental para definir a meta geográfica de análise, permitindo a identificação de todas as cidades que seriam usadas quer para a

análise da meteorologia, quer para a análise das músicas. Além do nome das cidades, o scraper obteve o país.

Posteriormente, desenvolvi um scraper adicional para extrair semanalmente o Top 100 de músicas de cada cidade previamente identificada. Este scraper foi programado para aceder aos rankings semanais e guardar cada identificador de música (ID) e posição no ranking. Esta informação começou a ser produzida pelo *Spotify* desde 21 de outubro de 2021 e semanalmente tem sido atualizada.

Enquanto o scraper buscava os rankings semanais, utilizei o pacote (Spotipy), uma biblioteca Python para interagir com a API do Spotify, para buscar atributos detalhados de cada música listada. O Spotipy permitiu a obtenção de dados enriquecidos sobre cada faixa, características acústicas como a energia, a valência, capacidade de dança, a instrumentalidade e o tempo (BPM).

2.2.2 Meteorologia

Para complementar a análise da influência da meteorologia na escolha musical, foi necessário aceder dados meteorológicos detalhados e precisos. A obtenção dessas informações envolveu o uso de um website especializado em estações meteorológicas e a aplicação de um pacote Python para a coleta de dados históricos.

Primeiramente, usou-se o <http://mesonet.agron.iastate.edu/geojson/network/AZOS.geojson> que disponibiliza detalhes de várias estações meteorológicas ao redor do mundo. Esse recurso foi essencial para identificar as coordenadas geográficas específicas das estações meteorológicas em cada cidade incluída no estudo.

Com as coordenadas das estações meteorológicas em mãos, utilizei o pacote <https://pypi.org/project/meteostat/>, uma biblioteca Python especializada em dados meteorológicos históricos. O Meteostat permite aceder informações meteorológicas detalhadas de diversas fontes, facilitando a coleta de dados de maneira programática e eficiente.

Através do Meteostat, obtive dados meteorológicos diários desde 15 de outubro de 2021. Visto que o dataset musical começou a ser produzido a partir de 21 de outubro de 2021 e relatava o top 100 de músicas da semana, faz sentido que a análise da influência da meteorologia na escolha musical fosse feita a partir da semana anterior a essa data.

Tendo em conta que existem imensos tipos de climas, independentemente das temperaturas médias, usei a escala de Köppen-Geiger para classificar os climas das cidades. Esta classificação foi feita com base nas coordenadas geográficas das cidades.

2.2.3 Classificação de Climats (Köppen-Geiger)

A classificação de Köppen-Geiger é um sistema de classificação climática amplamente utilizado que divide os climas do mundo em categorias distintas com base em fatores como temperatura, precipitação e sazonalidade. Este sistema é amplamente reconhecido pela sua simplicidade e eficácia na descrição de diferentes tipos de climas, facilitando a compreensão e a comparação de regiões climáticas em todo o mundo. Possíveis valores e significados:

- **A** - Climas Tropicais Chuvosos
 - **Af** - Clima equatorial;

- **Aw** - Clima de Savana (clima tropical com estação seca);
- **Am** - Clima tropical de monção.
- **B** - Climas Secos
 - **BWh** - Clima quente de deserto;
 - **BWk** - Clima frio de deserto;
 - **BSh** - Clima quente de estepe (semiárido);
 - **BSk** - Clima frio de estepe.
- **C** - Climas Temperados Chuvosos e Moderadamente Quentes
 - **Cfa** - Húmido em todas as estações com verão quente.
 - **Cfb** - Húmido em todas as estações com verão moderadamente quente.
 - **Csa** - Chuva de inverno com verão quente.
 - **Cwa** - Chuva de verão com verão quente.
 - **Cwb** - Chuva de verão com verão moderadamente quente.
- **D** - Climas Frios Continentais
 - **Dfa** - Húmido em todas as estações com verão quente.
 - **Dfb** - Húmido em todas as estações com verão frio.

2.3 Limpeza e Transformação de Dados

Esta tarefa foi realizada não só após obter os dados, mas também durante a coleta dos mesmos. Nomeadamente, na obtenção de dados musicais, foi necessário garantir que as músicas obtidas tinham todas as características pretendidas. Caso não tivessem, colocava-se a variável a zero. Para o caso da variável *location*, foi necessário unir o nome da cidade e do país, retirar acentos e caracteres especiais e colocar tudo em minúsculas em todos os datasets, para ficar de uma só forma.

Sendo que haviam variáveis que não eram úteis, como o ID da música ou a pressão atmosférica, estas foram retiradas.

Já após a obtenção dos dados meteorológicos, foi necessário garantir que as datas estavam no formato correto e que as variáveis estavam no tipo de dados correto. Além disso, preenchiam-se os valores nulos com o valor 0.0.

Após a obtenção dos dados climáticos, foi necessário garantir que as variáveis estavam no tipo de dados correto, nomeadamente a coluna *week* e que não haviam valores nulos. Se houvessem, eram atualizados para 0.0.

Por fim, quanto à classificação de Köppen-Geiger, visto que haviam climas que só tinham uma cidade, alterou-se o clima dessa cidade para o clima mais próximo e que já tivesse alguma cidade com esse clima.

2.4 Redução e Integração de Dados

Dado que os datasets obtidos não estavam no formato ideal para a análise, nomeadamente o dataset de músicas estar com 100 músicas por semana por cidade e o dataset de meteorologia estar com uma linha por dia por cidade, foi necessário reduzir os dados.

No dataset da meteorologia, agrupou-se os dados por semana e cidade, calculando a média da temperatura média diária, da temperatura mínima, da temperatura máxima, da precipitação e da velocidade do vento. No dataset das músicas, obteve-se a média da capacidade de dança, da energia, da instrumentalidade, do tempo e da valência do top 15 de músicas de cada cidade por semana. Sendo que fazer a média de 100 músicas não fazia sentido pois seria bastante ambíguo pois pode abordar músicas de diferentes características, decidiu-se fazer a média das 15 músicas mais populares.

Posteriormente, seguiu-se uma normalização dos dados, de forma a que todas as variáveis tivessem a mesma importância na análise e para que os valores fossem mais dispersos. Logo, o novo limite mínimo e máximo de cada variável passou a ser 0 e 30, respetivamente.

Por fim, juntaram-se os datasets de músicas, da meteorologia e dos climas, tendo em conta a localização da cidade e a data da semana. Eis o resultado do dataset final:

- **location:** Nome da cidade e do país;
- **week:** Data de início da semana;
- **danceability:** Capacidade de dança média das 15 músicas mais populares da cidade na semana;
- **energy:** Energia média das 15 músicas mais populares da cidade na semana;
- **valence:** Valência média das 15 músicas mais populares da cidade na semana;
- **instrumentalness:** Instrumentalidade média das 15 músicas mais populares da cidade na semana;
- **tempo:** Tempo médio das 15 músicas mais populares da cidade na semana;
- **tavg:** Temperatura média semanal da cidade;
- **tmin:** Temperatura mínima semanal da cidade;
- **tmax:** Temperatura máxima semanal da cidade;
- **prcp:** Precipitação semanal da cidade;
- **wspd:** Velocidade do vento semanal da cidade;
- **latitude:** Coordenada geográfica da cidade;
- **longitude:** Coordenada geográfica da cidade;
- **kg_zone:** Classificação climática de Köppen-Geiger da cidade.

Para esta integração, usei o Parquet, um formato de arquivo colunar que é otimizado para leitura e gravação de dados em larga escala. O Parquet é especialmente útil para conjuntos de dados grandes e complexos, pois permite a leitura e gravação eficiente de dados em diferentes formatos e estruturas. Assim, optei por duas abordagens: ETL e ELT. Eis uma análise comparativa entre os dois métodos:

ELT: Memória Inicial=78.58 MB - Memória Final=104.42 MB - Diferença=25.84 MB
ETL: Memória Inicial=78.37 MB - Memória Final=85.02 MB - Diferença=6.65 MB

Assim, conclui-se que o método mais eficiente é o ETL.

2.5 Segmentação de Dados e Anonimização

Sendo que era para análise estatística, que não havia identificador da música (era só usado para obter as características dela) e que era importante saber a localização das cidades, não se procedeu nem à segmentação, nem à anonimização dos dados.

2.6 Armazenamento de Dados

Para armazenar os dados, optei pelo mongodb por questões de possuir uma vasta gama de drivers e ferramentas que facilitam a integração com plataformas de análise de dados, como o Power BI. Isto permite extrair, transformar e carregar dados de maneira eficiente para análise subsequente.

3 Resultados

3.1 Página de Análise de Características das Músicas

3.1.1 Capacidade de Dançar

No topo dos gráficos climáticos, destacam-se algumas tendências interessantes nas diferentes classificações climáticas:

	Dançaabilidade Aumenta	Dançaabilidade Mantém	Dançaabilidade Diminui
Temperatura Aumenta	Bsk, Cfa, Cfb, Csa, Dfa, Dfb	Bsh, Bwh, Cwa	
Temperatura Mantém		Cwb, Aw, Am, Af	
Temperatura Diminui			

Com esta tabela, é possível concluir que nos Climas Tropicais Chuvosos (Af, Am, Aw), sendo que a temperatura se mantém, a dançaabilidade também; nos que possuem verão quente (Cfa, Csa, Dfa) ou moderadamente quente e húmido (Cfb) ou húmido em todas as estações com verão frio (Dfb) tende a aumentar quando a temperatura aumenta, enquanto que nos climas com chuva de verão (Cwa, Cwb) a dançaabilidade se mantém; nos climas quentes secos (Bsh, Bwh) a dançaabilidade mantém apesar de aumentar a temperatura.

3.1.2 Energia

	Energia Aumenta	Energia Mantém	Energia Diminui
Temperatura Aumenta	Cfa, Cfb, Dfa, Dfb	Bsk, Bsh, Bwh, Cwa, Csa	
Temperatura Mantém		Cwb, Aw, Am, Af	
Temperatura Diminui			

Com base na tabela da energia, é possível concluir que nos climas com verão quente (Cfa, Dfa) ou moderadamente quente e húmido (Cfb) ou húmido em todas as estações com verão frio (Dfb), a energia tende a aumentar quando a temperatura aumenta. Nos climas quentes secos (Bsk, Bsh, Bwh) e nos climas com chuva de verão (Cwa, Csa), a energia mantém-se apesar do aumento da temperatura. Nos Climas Tropicais Chuvosos (Af, Am, Aw), onde a temperatura se mantém, a energia também se mantém.

	Instrumentalidade Aumenta	Instrumentalidade Mantém	Instrumentalidade Diminui
Temperatura Aumenta	Cfb, Bwh	Bsh, Bsk, Cfa, Csa, Cwa, Dfa, Dfb	
Temperatura Mantém		Cwb, Aw, Am, Af	
Temperatura Diminui			

3.1.3 Instrumental

Com base na tabela da instrumentalidade, é possível concluir que nos climas com verão moderadamente quente e húmido (Cfb) e num dos climas quentes secos (Bwh), a instrumentalidade tende a aumentar quando a temperatura aumenta. Nos restantes climas quentes secos (Bsh, Bsk), nos climas com verão quente (Cfa, Dfa) ou moderadamente quente e húmido (Cfb) ou húmido em todas as estações com verão frio (Dfb) e nos climas com chuva de verão (Cwa), a instrumentalidade mantém-se apesar do aumento da temperatura.

3.1.4 Tempo

	Tempo Aumenta	Tempo Mantém	Tempo Diminui
Temperatura Aumenta		Bsh, Bsk, Bwh, Cfa, Cfb, Csa, Cwa, Dfa, Dfb	
Temperatura Mantém	Af	Cwb, Aw, Am	
Temperatura Diminui			

Quanto ao Tempo, apenas o Clima equatorial é que aumenta, apesar da temperatura não se alterar. Nos restantes, independentemente da temperatura, os valores mantêm-se.

3.1.5 Valência

	Valência Aumenta	Valência Mantém	Valência Diminui
Temperatura Aumenta	Cfa, Dfa	Bsh, Bsk, Bwh, Cfb, Csa, Cwa, Dfb	
Temperatura Mantém		Cwb, Af, Aw, Am	
Temperatura Diminui			

Por último, na característica da valência, os Climas Húmidos em todas as estações com verão quente ou frio é que se destacam com o aumento da valência quando a temperatura aumenta.

3.2 Página de Análise por Hemisférios

Nestas páginas, pretende-se verificar os valores do desvio padrão em relação à característica das músicas entre os climas para determinar os que têm maior dispersão, ou seja, os que potencialmente indicam que há uma relação entre a meteorologia e as características das músicas. Será apresentado uma lista sobre os que mais se destacam.

3.2.1 Hemisfério Norte

- **Capacidade de Dançar:** Af (Clima equatorial)
- **Energia:** Cwa (Clima subtropical húmido com verão quente)
- **Instrumental:** Bwh (Clima desértico quente)
- **Tempo:** Aw (Clima tropical de savana)
- **Valência:** Cwa (Clima subtropical húmido com verão quente)

3.2.2 Hemisfério Sul

- **Capacidade de Dançar:** Aw (Clima tropical de savana)
- **Energia:** Cfa (Clima temperado húmido com verão quente)
- **Instrumental:** Cfa (Clima temperado húmido com verão quente)
- **Tempo:** Cfa (Clima temperado húmido com verão quente)
- **Valência:** Cfa (Clima temperado húmido com verão quente)

3.3 Página de Análise por Clima

Com esta página, conseguimos perceber a relação entre as características das músicas e a meteorologia nas várias cidades de cada clima. Portanto, dentro das "mesmas" condições meteorológicas, poderemos inferir sobre a existência de outros fatores que fazem com que as características das músicas sejam diferentes em cada cidade.

Assim, é possível verificar que as cidades sul-americanas e europeias são mais propensas a terem músicas com mais energia e mais dançáveis, ao contrário das mais orientais, norte-americanas e britânicas.

3.4 Página de Análise por Cidade

Nesta página, é possível perceber a relação entre as características das músicas e a meteorologia numa cidade específica. Assim, fica mais fácil perceber a relação entre a meteorologia e as características das músicas ao longo do ano. É possível inferir que cidades europeias têm tendência a que, no verão, sejam mais "energéticas e dançáveis". Já sobre cidades cuja temperatura não varia muito, os valores das características tendem a manter, o mesmo previsto anteriormente na análise por características.

4 Discussão

Analisando os dados apresentados, podemos observar que a relação entre temperatura e diferentes características climáticas varia conforme o tipo de clima. Climas Tropicais Chuvosos (Af, Am, Aw) demonstram uma estabilidade tanto em temperatura quanto em dançabilidade e energia, indicando uma consistência que não é influenciada por variações térmicas.

Por outro lado, climas com verão quente (Cfa, Csa, Dfa) ou moderadamente quente e húmido (Cfb) ou húmido em todas as estações com verão frio (Dfb) mostram uma tendência de aumento na dançabilidade e energia à medida que a temperatura sobe. Isso sugere que o calor pode ser um fator estimulante para essas características nesses climas específicos.

Interessantemente, nos climas com chuva de verão (Cwa, Cwb), apesar das variações de temperatura, a dançabilidade se mantém, o que pode indicar que outros fatores além da temperatura influenciam essa característica.

Nos climas quentes secos (Bsh, Bwh), a dançabilidade e a energia parecem não ser afetadas pelo aumento da temperatura, o que pode refletir uma adaptação das atividades e do comportamento energético às condições áridas desses climas.

A instrumentalidade mostra um padrão misto, com aumento em climas moderadamente quente e húmido (Cfb) e em um dos climas quentes secos (Bwh), enquanto se mantém estável nos demais. Isso pode indicar que a instrumentalidade é uma característica que pode ser influenciada por uma combinação de fatores climáticos, não apenas pela temperatura.

Finalmente, o Tempo aumenta apenas no Clima equatorial, o que é peculiar, considerando que a temperatura se mantém constante. Isso pode sugerir que o Tempo, como característica, é influenciado por elementos inerentes ao clima equatorial que não estão presentes nos outros climas.

5 Conclusão

Este trabalho abre portas para futuras investigações sobre a relação entre a meteorologia e a música, incentivando a exploração de novas abordagens e metodologias para entender como o ambiente influencia a criação e a apreciação musical. Com uma compreensão mais profunda dessas interações complexas, podemos enriquecer a nossa apreciação da música e aprimorar nossas práticas artísticas para refletir a diversidade e a riqueza do mundo natural.

Contudo, é de notar que este estudo é baseado em correlações observadas nos dados e poderão haver outros fatores, como é possível notar quando se analisa por clima individualmente, onde há cidades com valores das características das músicas diferentes, mesmo tendo o mesmo clima, ou até os valores das características se manterem quando as temperaturas médias dos climas variam. Assim, é importante considerar fatores como a cultura e o gosto pessoal ao interpretar os resultados e ao aplicar essas descobertas na prática musical.

References

- [1] North, A.C., Hargreaves, D.J.: The Social and Applied Psychology of Music. Oxford University Press, ??? (2008)
- [2] Rentfrow, P.J., Gosling, S.D.: The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology* **84**(6), 1236–1256 (2003)
- [3] Schäfer, T., Sedlmeier, P.: From the functions of music to music preference. *Psychology of Music* **37**(3), 279–300 (2009)