

Universidad de Los Andes
Curso *Análisis no supervisado*- Entrega proyecto

Grupo 19

Carlos Galindo Aguilera, Daniel Enrique Granados Iglesias,
Sebastian Barrera Bernal, Mayra Alejandra Neisa Valero
Septiembre – 2024

Exploración de géneros y recomendación musical basada en Aprendizaje No Supervisado a partir de canciones de Spotify

1. Resumen

En la era digital, la música se ha convertido en un elemento omnipresente, con plataformas como Spotify que ofrecen acceso a millones de canciones. Sin embargo, la vastedad de opciones puede dificultar la personalización de la experiencia del usuario. Este proyecto propone un enfoque basado en Aprendizaje No Supervisado para analizar una base de datos de 30.000 canciones de Spotify, utilizando características como energía, locuacidad, acústica e instrumentalidad. Mediante el uso de Análisis de Componentes Principales (PCA), se reducirá la dimensionalidad del conjunto de datos, permitiendo la identificación de patrones subyacentes. Posteriormente, se aplicarán técnicas de clustering para agrupar las canciones según géneros emergentes y patrones de popularidad. Finalmente, se analizarán sistemas de recomendación para ofrecer sugerencias musicales personalizadas. Este proyecto busca no solo descubrir nuevas tendencias musicales, sino también mejorar la experiencia de usuario mediante recomendaciones más precisas y adaptadas a los gustos individuales.

2. Introducción

En el vasto universo de la música digital, plataformas como Spotify, Amazon Music, YouTube Music y Apple Music compiten por ofrecer a sus usuarios la experiencia más personalizada posible. Sin embargo, en un entorno donde el catálogo de canciones disponible supera los millones, ofrecer recomendaciones musicales precisas y relevantes se ha convertido en un desafío crítico. La diversidad de gustos musicales y la constante aparición de nuevos géneros hacen que las recomendaciones tradicionales basadas en la popularidad o en las preferencias pasadas del usuario resulten insuficientes para satisfacer la demanda de personalización.

El cliente potencial para este proyecto son las plataformas de streaming musical, específicamente los equipos de desarrollo y analítica de datos de empresas como Spotify. Estos equipos enfrentan la necesidad constante de mejorar sus sistemas de recomendación, no solo para aumentar la satisfacción del usuario, sino también para diferenciarse de la competencia en un mercado saturado. En este contexto, el problema a resolver es cómo identificar patrones ocultos y tendencias emergentes en un conjunto masivo de datos musicales, con el fin de mejorar la precisión y relevancia de las recomendaciones.

El uso del aprendizaje no supervisado se presenta como una solución prometedora para este desafío. A diferencia de los métodos supervisados, que requieren datos etiquetados, el aprendizaje no supervisado permite descubrir estructuras subyacentes en los datos sin la necesidad de etiquetas predefinidas. Esto es particularmente relevante en la industria musical, donde los gustos y tendencias cambian rápidamente, y las etiquetas de género tradicionales pueden no capturar la complejidad de la música moderna.

Este proyecto se enmarca dentro del área del aprendizaje no supervisado, con un enfoque en la reducción de dimensionalidad mediante Análisis de Componentes Principales (PCA), la agrupación de datos a través de técnicas de clustering, y el desarrollo de recomendaciones personalizadas. La combinación de estas técnicas permitirá no

solo reducir la complejidad del análisis, sino también identificar nuevos patrones y géneros emergentes que pueden ser utilizados para mejorar las recomendaciones musicales. En última instancia, se espera que este enfoque contribuya a mejorar la experiencia del usuario en plataformas como Spotify, proporcionando recomendaciones más personalizadas y precisas, adaptadas a los gustos individuales y a las tendencias emergentes.

3. Materiales y métodos

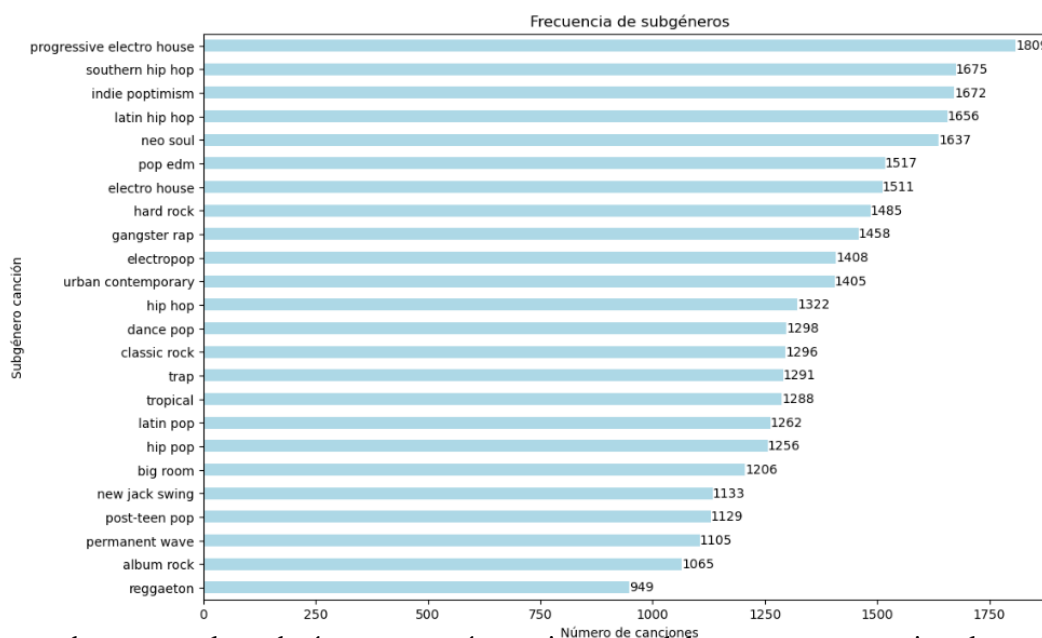
3.1. Fuente de datos

El dataset utilizado para este proyecto, corresponde a '30.000 Spotify Songs' un dataset de Kaggle creado por Joakim Arvidsson que puede ser consultado a través de la siguiente URL: [enlace](#). En este dataset se encuentran casi 30.000 canciones recuperadas desde la API de Spotify, y 23 columnas o variables que las describen.

En el dataset podemos encontrar columnas identificadoras tales como el ID y nombre de la canción, el nombre del artista, el ID y nombre del álbum, y el ID y nombre de la playlist. De igual forma, encontramos la fecha en que fue lanzado el álbum, y variables categóricas como el género y el subgénero al que pertenece la playlist.

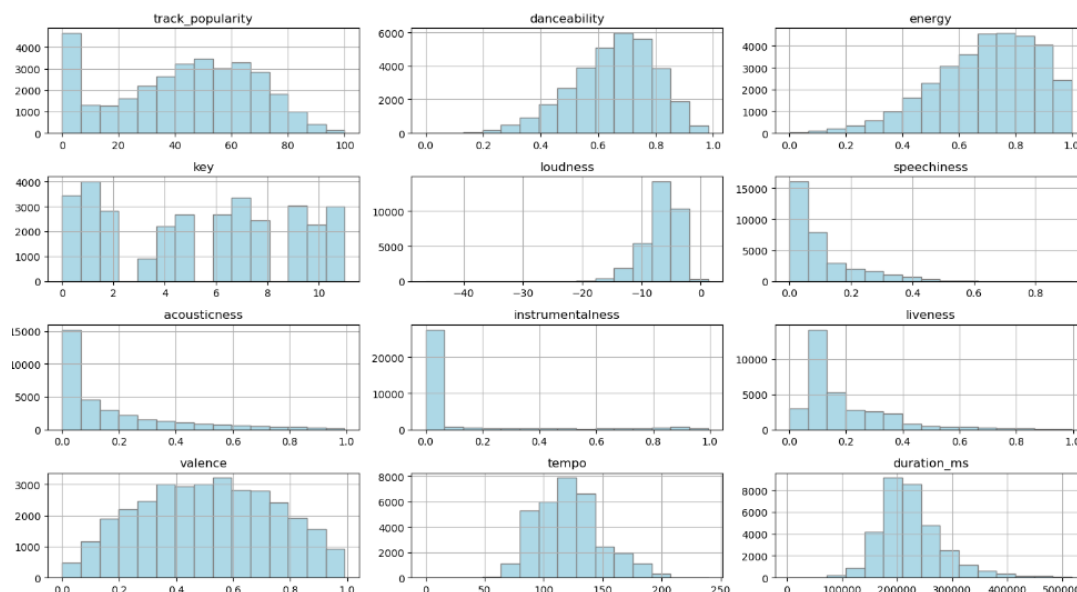
En cuanto a las variables de tipo numéricas, se cuenta con índices que describen diferentes características del sonido de las canciones, como bailabilidad, energía, acústica, entre otros. Además, se cuenta con un índice que representa la popularidad de la canción.

3.2. Análisis de variables categóricas



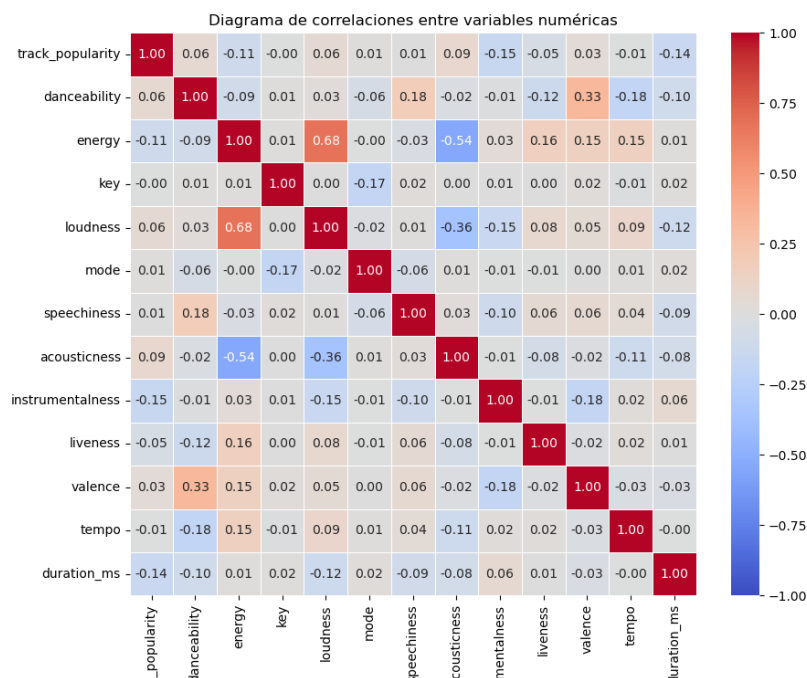
Tal y como observamos, los subgéneros con más canciones en el dataset son progressive electro house, southern hip hop, indie popitism, latin hip hop y neo soul. Mientras que los tres subgéneros con menos canciones son, permanent wave, álbum rock y reggaetón. Sin embargo, las categorías parecen estar lo suficientemente balanceadas, con un rango de 949 canciones a 1.809 canciones entre las 24 categorías.

3.3. Análisis de variables numéricas



En general, tenemos más canciones con baja popularidad que con alta popularidad. La mayor parte de las canciones son relativamente bailables, y parecen reflejar una alta energía. También observamos que la mayoría de las canciones son poco habladas, es decir, están representadas principalmente por la pista. También tenemos canciones con baja acústica y baja instrumentalización, esto último coincide con el grado de habla de la canción.

3.4. Análisis de correlaciones de variables numéricas



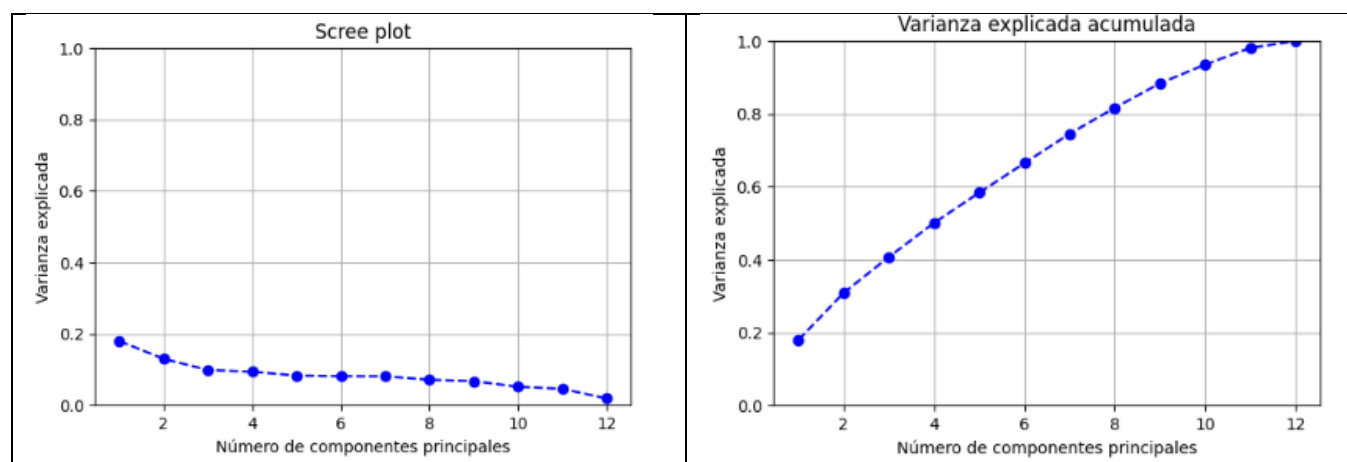
De manera general, observamos correlaciones no muy altas entre las variables numéricas, lo cual pudiera sugerir la no viabilidad de emplear técnicas como Análisis de Componentes Principales al set de datos para reducir su dimensión, sin embargo, esto se determinará al momento de aplicar en detalle esta técnica. Las correlaciones más altas se evidencian entre la variable 'energy' y 'loudness' (0.68), y 'energy' y 'acousticness' (0.54).

3.5. Preparación de datos

Para la preparación de los datos, se hace una revisión del contenido, no se encuentran valores faltantes por lo que se considera el set completo. En particular para los distintos procesamiento se considerarán las variables numéricas por lo que se genera un subconjunto solo con estas variables y se estandarizan todos los valores. Este set está conformado por 13 variables. Se hace una limpieza eliminando duplicados en la columna `track_id` que representa el identificador único de cada canción.

4. Resultados y discusión

PCA (Análisis de Componentes Principales), para reducir la dimensionalidad de las 30.000 canciones de Spotify al transformar las características musicales como acústica, instrumentalidad y energía. Se esperaría simplificar el análisis al retener la mayor variación posible en un espacio de menor dimensión, facilitando la identificación de patrones y relaciones en el dataset.



En la búsqueda del número de componentes a retener se recurre al método de la gráfica del codo, con la varianza explicada, sin embargo, no se identifica un punto claro que indique el corte y si fuese por la varianza explicada acumulada, para tener un valor entre 70% y 80% se deberían conservar alrededor de 8 dimensiones, lo cual no parece una gran reducción, por lo que se mantendrán todas las variables originales.

Sistema de recomendación base: Utilizando las características de sonido de cada canción, se desarrollan tres sistemas de recomendación. Estos sistemas sugieren el top cinco de canciones más similares a la que se busca, basándose en la similitud de coseno aplicada a las variables mencionadas. Al utilizar el `track_id` de la canción *Call You Mine - Keanu Silva Remix* de The Chainsmokers, se obtienen los siguientes resultados en cada uno de los sistemas de recomendación.

Para el primer sistema, se recomienda cualquier canción que sea similar a la buscada, independientemente de su género. Los resultados obtenidos fueron los siguientes:

	track_id	track_name	track_artist	playlist_genre
31439	6XtOpnP7gitKdibHXylwsC	I Don't Like You - Nicky Romero Remix	Eva Simons	edm
27757	79Uq59YH1T7EJ16uWxK8Oc	Bathing Suit Body	Badpoken	edm
2129	6YZdkObH88npeKrrkb8Ggf	DUELE EL CORAZON	Enrique Iglesias	pop
27333	5NwWGVjLvbxJZrJLhcRqm1	Music Has the Power - Radio Edit	Ideladap	edm
3014	48QmG1dfvMuYLxMPT7KSRA	No Lie	Sean Paul	pop

Se observa que las canciones recomendadas pertenecen mayoritariamente a dos géneros: EDM y pop, ambos comúnmente asociados con el estilo de este artista. Sin embargo, resulta interesante notar que, aunque la canción original pertenece al género pop, las dos primeras recomendaciones no pertenecen a este género. Por ello, el segundo sistema de recomendación se centra en sugerir únicamente canciones del mismo género que la canción buscada, arrojando los siguientes resultados:

	track_id	track_name	track_artist	playlist_genre
2129	6YZdkObH88npeKrrkb8Ggf	DUELE EL CORAZON	Enrique Iglesias	pop
3014	48QmG1dfvMuYLxMPT7KSRA	No Lie	Sean Paul	pop
3825	5m5TRf7ggoguXtcXQGprN4	Written in the Stars (feat. Eric Turner)	Tinie Tempah	pop
2497	3cFsmM51N5zMAuu1nNY4Tw	No Lie	Sean Paul	pop
3755	1jLsirPDKUS2g4gnkYua58	Ignite	Alan Walker	pop

Como se muestra en la imagen, las dos primeras recomendaciones coinciden con las del modelo anterior, que no consideraba el género. Además, se añaden nuevas canciones a la lista, mostrando aquellas que pertenecen al mismo género y son más afines a la canción que estamos utilizando como referencia. En este caso se repite la canción No Lie de Sean Paul dos veces ya que contiene identificadores distintos posiblemente refiriendo a otra versión de la canción. Finalmente, desarrollamos un tercer sistema de recomendación que sugiere canciones del mismo artista que la canción buscada. Los resultados obtenidos fueron los siguientes:

	track_id	track_name	track_artist	playlist_genre
29718	4B3NahKokb7ILkvaLFZFzn	Sick Boy - Owen Norton Remix	The Chainsmokers	edm
10915	33ecyF2OHCK4HFUSp1hItE	Closer - T-Mass Remix	The Chainsmokers	rap
30142	4Sdp10fc5XgHxL1VaiMVYN	Inside Out - DubVision Remix	The Chainsmokers	edm
24	4sqABRRGU7CzcHXCyUzFw	Do You Mean (feat. Ty Dolla \$ign & bülöw) - My...	The Chainsmokers	pop
30110	1JI70I1IE5IF2tgJm5TnMD	Don't Let Me Down - Hardwell & Sephyx Remix	The Chainsmokers	edm

Lo más interesante de este caso es que, a pesar de centrarnos en el mismo artista, los géneros recomendados coinciden con los que vimos en el primer sistema de recomendación. Esto demuestra cómo la similitud de coseno y los parámetros utilizados capturan eficazmente las características musicales del artista, permitiendo recomendar canciones similares sin depender directamente de variables como el género o el propio artista.

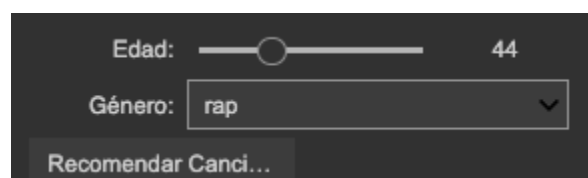
Sistema de recomendación basado en contenido: Para presentar un sistema de recomendación de música que utilizara formas más innovadoras en la implementación del modelo, se decidió crear la manera de interactuar con un usuario y, con base en sus gustos, hacer los análisis pertinentes para recomendar las canciones que pudieran ser de su agrado.

Según algunos [estudios](#) (ver enlaces [http](#)) las personas adquieren sus gustos musicales entre los 13 y 16 años, este rango de edad parece ser definitivo en la selección de preferencias musicales. Con base en estos datos se implementó un modelo donde:

1. Se le pregunta al usuario la edad: Esto permite calcular en qué momento en la línea de tiempo cumplió los 13 y 16 años.
2. Se le pregunta al usuario el género: Esto permite filtrar la base de canciones existente y afinar la recomendación final.

Para la implementación del modelo se usó el cálculo de similitud entre canciones se realiza usando la similitud de coseno sobre las características seleccionadas, lo que permite encontrar canciones que son similares en cuanto a sus propiedades musicales. Cabe mencionar que se filtran algunas de las características musicalmente relevantes para dicho cálculo *danceability*, *energy*, *valence*, entre otras, para comparar y recomendar canciones similares. Estas mismas características fueron también escaladas para que contribuyan de forma equilibrada en el cálculo de la similitud del coseno.

De manera intencional se probó el modelo bajo el supuesto de que la persona tendría 44 años (es decir que nació en 1980) y que su género favorito fuese el rap. Estas condiciones indicarían que sus preferencias musicales fueron adquiridas entre el año 1.993 y 1.996, época en la que se encontraba en *boom* el rap en Los Angeles (West Coast) y New York (East Coast).



Edad: 44

Género:

Recomendar Canci...

Después de ingresada la información el sistema de recomendación genera la siguiente salida:

	track_name	track_artist	album_year	playlist_genre
9191	Straight Outta Compton - Extended Mix / Edit	N.W.A.	1996	rap
7794	Let Me Clear My Throat - Old School Reunion Re...	DJ Kool	1996	rap
7068	Down For My Niggaz	C-Murder	2005	rap
7374	Let Me Clear My Throat - Old School Reunion Re...	DJ Kool	1996	rap
9645	Natural Born Killaz	Dr. Dre	2001	rap

Como primera recomendación observamos una canción del grupo N.W.A. precisamente del año 1.996 similar las otras 2 recomendaciones de la misma lista. Esto indica que el modelo está teniendo en cuenta el input del usuario para realizar la recomendación.

De forma interesante se observa también una canción de Dr. Dre, esta vez con un álbum de lanzamiento fuera de rango, pero curiosamente consistente, pues este artista perteneció a N.W.A. durante la década de los 90's.

Esto demuestra que el modelo fue exitoso porque combina personalización (a través de la edad y el género musical) con un análisis profundo de las características musicales de las canciones. La normalización y el uso de la similitud de coseno ayudaron a hacer recomendaciones precisas y consistentes, asegurando que las canciones seleccionadas fueran adecuadas y alineadas con las expectativas del usuario.

5. Conclusiones

5.1 Personalización con información explícita del usuario vs. Similitud por características musicales:

- Modelo interactivo basado en preferencias del usuario (edad y género): Este enfoque utiliza la edad del usuario para filtrar canciones de una época en la que es probable que hayan desarrollado sus gustos musicales, y también toma en cuenta el género musical preferido. Esto introduce una capa adicional de contexto personal.
- Modelo basado en características musicales (sin información del usuario):
 - El segundo modelo se basa exclusivamente en características musicales como danceability, energy, key, loudness, entre otras. Esto significa que las recomendaciones están impulsadas únicamente por las propiedades intrínsecas de las canciones.
 - Las recomendaciones aquí tienden a ser más objetivas, ya que se basan en qué tan similares son las canciones en términos de características musicales, sin depender directamente de la información proporcionada por el usuario.

5.2 Exploración de nuevas canciones vs. afinidad con el contexto del usuario:

- Modelo interactivo con preferencias de usuario: Este modelo es más efectivo para adaptar las recomendaciones al **contexto personal** del usuario, lo que es especialmente útil cuando queremos que el sistema sugiera música nostálgica o relevante a una época o un género específico.
- Modelo basado en características musicales: Este enfoque puede ser mejor para **explorar** canciones que comparten similitudes musicales, sin estar limitado por las preferencias de género o el contexto de tiempo del usuario. Permite descubrir canciones que son musicalmente similares, incluso si el usuario no las ha considerado antes.

5.3 Flexibilidad de personalización:

- Interacción basada en la edad y género: El modelo basado en las preferencias del usuario es más personalizable en función de las entradas explícitas del usuario. Esto lo hace más flexible a la hora de adaptar las recomendaciones basadas en información subjetiva, como los gustos de una época específica.
- Similitud por características musicales: Aunque este modelo es menos dependiente de entradas directas del usuario, se enfoca más en la música como tal, lo que puede ser útil para descubrir nueva música **independientemente del género o el periodo de tiempo**. La personalización en este caso es implícita, ya que depende de las propiedades de las canciones.

5.4 Aplicaciones en distintos escenarios:

- Modelo interactivo (edad/género): Este modelo es ideal cuando el usuario busca recomendaciones más cercanas a su historia musical y su contexto personal. Puede ser útil en aplicaciones que buscan despertar emociones basadas en la nostalgia o afinidad de género.

- Modelo de características musicales: Este modelo es excelente para el descubrimiento **de música** cuando el usuario no proporciona información personal directa. Puede ofrecer sugerencias basadas únicamente en los atributos musicales, lo que puede generar descubrimientos interesantes.

5.5 Resultados consistentes en ambos enfoques:

- Interacción con el usuario: El modelo que utiliza la edad y el género permite hacer recomendaciones muy alineadas con las preferencias personales.
- Similitud musical: El modelo basado en características musicales permite encontrar canciones similares en términos de su estructura y composición, lo que también genera recomendaciones de calidad.

6. Bibliografía

- Cilibrasi, R., Vitanyi, P., & de Wolf, R. (2004). Algorithmic clustering of music. In *Proceedings of the Fourth International Conference on Web Delivering of Music (EDELMUSIC 2004)*. IEEE. Recuperado de: [enlace](#).
- Jondya, A. G., & Iswanto, B. H. (2017). Indonesian's traditional music clustering based on audio features. *Procedia Computer Science*, 116, 174–181. ScienceDirect. Recuperado de: [enlace](#).
- Li, T., Ogihara, M., Peng, W., Shao, B., & Zhu, S. (2009). Music clustering with features from different information sources. *IEEE Transactions on Multimedia*, 11(3), 423-434. IEEE. Recuperado de: [enlace](#).
- Schedl, M. (2019). Deep learning in music recommendation systems. *Frontiers in Applied Mathematics and Statistics*. Institute of Computational Perception, Johannes Kepler University, Linz, Austria. Recuperado de: [enlace](#).
- Song, Y., Dixon, S., & Pearce, M. (2012). A survey of music recommendation systems and future perspectives. In *Proceedings of the 9th International Symposium on Computer Music Modelling and Retrieval (CMMR 2012)* (pp. 19-22). Queen Mary University of London. Recuperado de: [enlace](#).
- <https://scholarlykitchen.sspnet.org/2024/05/03/music-paralysis-as-we-age/#:~:text=Daniel%20Parris%20was%20inspired%20to,formative%20for%20our%20musical%20tastes>.
- <https://www.duna.cl/noticias/2018/02/13/a-que-edad-definimos-nuestro-gusto-musical-estudio-lo-revelo/#:~:text=Gracias%20a%20bases%20de%20informaci%C3%B3n,a%20los%2016%20en%20promedio>.

7. Anexos

Variable	Descripción
Track popularity	Popularidad de la canción entre 0 y 100, donde entre más grande el número, mayor es la popularidad.
Danceability	Describe cuán adecuada es la pista para bailar basada en una combinación de elementos musicales que incluyen tempo, estabilidad de ritmo, fuerza de ritmo y regularidad general. Un valor de 0.0 es menos bailable y 1.0 es el más bailable.
Energy	La energía es una medida de 0.0 a 1.0 y representa una medida perceptiva de intensidad y actividad. Por lo general, las pistas enérgicas se sienten rápidas, fuertes y ruidosas. Por ejemplo, el death metal tiene alta energía, mientras que un preludio de Bach obtiene un puntaje bajo en la escala. Las características perceptivas que contribuyen a este atributo incluyen el rango dinámico, el volumen percibido, el timbre, la tasa de inicio y la entropía general.
Key	La clave general estimada de la pista. Se asignan enteros a lanzamientos utilizando notación de clase de tono estándar. P.ej. 0 = C, 1 = C#/D b, 2 = D, y así sucesivamente. Si no se detectó ninguna clave, el valor es -1.
Loudness	El volumen general de una pista en Decibels (DB). Los valores de volumen se promedian en toda la pista y son útiles para comparar el volumen relativo de las pistas. El volumen es la calidad de un sonido que es el correlato psicológico primario de la fuerza física (amplitud). Los valores rango típico se encuentran entre -60 y 0 dB.
Mode	El modo indica la modalidad (mayor o menor) de una pista, el tipo de escala de la que se deriva su contenido melódico. Mayor está representado por 1 y menor es 0.
Speechiness	El habla detecta la presencia de palabras habladas en una pista. Cuanto más exclusivamente segmera la grabación (por ejemplo, programa de entrevistas, audiolibro, poesía), más cerca de 1.0 el valor del atributo. Los valores superiores a 0.66 describen pistas que probablemente se hagan completamente de palabras habladas. Los valores entre 0.33 y 0.66 describen pistas que pueden contener música y habla, ya sea en secciones o en capas, incluidos casos como la música rap. Los valores por debajo de 0.33 probablemente representan la música y otras pistas sin voz.
Acousticness	Una medida de confianza de 0.0 a 1.0 de si la pista es acústica. 1.0 representa una alta confianza, la pista es acústica.
Instrumentalness	Predice si una pista no contiene voces. Los sonidos "OOH" y "AAH" se tratan como instrumentales en este contexto. Las pistas de rap o palabras habladas son claramente "vocales". Cuanto más cerca sea el valor de instrumentalidad a 1.0, mayor es la probabilidad de que la pista no contenga contenido vocal. Los valores superiores a 0.5 están destinados a representar pistas instrumentales, pero la confianza es mayor a medida que el valor se acerca a 1.0.
Liveness	Detecta la presencia de una audiencia en la grabación. Los valores de liveness más altos representan una mayor probabilidad de que la pista se haya realizado en vivo. Un valor superior a 0.8 proporciona una fuerte probabilidad de que la pista esté en vivo.
Valence	Una medida de 0.0 a 1.0 que describe la positividad musical transmitida por una pista. Las pistas con alta valencia suenan más positivas (por ejemplo, felices, alegres, eufóricas), mientras que las pistas con baja valencia suenan más negativas (por ejemplo, triste, deprimida, enojada).
Tempo	El tempo estimado general de una pista en ritmos por minuto (BPM). En la terminología musical, el tempo es la velocidad o el ritmo de una pieza dada y se deriva directamente de la duración promedio del latido.
Duration	Duración de la canción en milisegundos.