

Anexo II. Descripción de las variables

A continuación se muestra una tabla con las variables de nuestro conjunto de datos. Junto con el nombre se indica qué tipo de variable y una breve descripción.

Nombre	Tipo	Descripción
artists	categorical	El artista principal de la canción
artist_ids	id	Identificador único del artista
id	id	Identificador único de la pista
name	categorical	Nombre de la pista
album	categorical	Nombre del álbum de donde proviene la pista
album_id	id	Identificador único del álbum
track_number	categorical	El número de la pista. Si un álbum tiene varios discos, el número de pista es el número del disco especificado.
disc_number	categorical	El número de disco (generalmente uno a menos que el álbum conste de más de un disco)
year	numerical	Año en que se lanzó la pista por primera vez
classical	binary	Valor binario que determina si la pista por género como clásica o no
acousticness	numerical	Una medida de confianza de 0.0 a 1.0 de si la pista es acústica. 1.0 representa una alta confianza en que la pista es acústica
danceability	numerical	La capacidad de baile describe qué tan adecuada es una pista para bailar en función de una combinación de elementos musicales que incluyen el tempo, la estabilidad del ritmo, la fuerza del ritmo y la regularidad general. Un valor de 0 es menosailable y un valor de 1 es másailable
duration_ms	numerical	La duración de la canción en milisegundos
explicit	binary	Si la pista tiene letras explícitas o no, es decir, si su contenido puede ser ofensivo, discriminatorio o violento se considerará explícita su letra

instrumentalness	numerical	Predice si una pista no contiene voces. Los sonidos 'ooh' y 'aah' se tratan como instrumentales en este contexto. Las pistas de rap o de palabra hablada son claramente 'vocales'. Cuanto más cercano esté el valor de instrumentalidad a 0, mayor será 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
key	categorical	La clave en la que se encuentra la pista. Los números enteros se asignan a los tonos utilizando la notación estándar de clase de tono. E.g. 0 = C, 1 = C#/D ♭, 2 = D, and so on.
liveness	numerical	Detecta la presencia de una audiencia en la grabación. Los valores de vivacidad más altos representan una mayor probabilidad de que la pista se haya interpretado en vivo. Un valor superior a 0,8 proporciona una gran probabilidad de que la pista esté activa
loudness	numerical	El volumen general de una pista en decibelios (dB). Los valores de sonoridad se promedian en toda la pista y son útiles para comparar la sonoridad relativa de las pistas. La sonoridad es la calidad de un sonido que es el principal correlato psicológico de la fuerza física (amplitud). Los valores típicos oscilan entre -60 y 0 db
mode	binary	El modo indica la modalidad (mayor o menor) de una pista, el tipo de escala de la que se deriva su contenido melódico. El modo mayor está representado por 1 y el menor por 0

popularity	numerical	<p>La popularidad de la pista. El valor estará entre 0 y 100, siendo 100 el más popular. Se calcula mediante un algoritmo y se basa, en su mayor parte, en el número total de reproducciones que ha tenido la pista y qué tan recientes son esas reproducciones.</p> <p>En términos generales, las canciones que se reproducen mucho ahora tendrán una mayor popularidad que las canciones que se escucharon mucho en el pasado. Las pistas duplicadas (por ejemplo, la misma pista de un sencillo y un álbum) se clasifican de forma independiente. La popularidad del artista y del álbum se deriva de la popularidad de la pista. El valor de popularidad puede retrasarse unos días en la popularidad real: el valor no se actualiza en tiempo real</p>
tempo	numerical	<p>El tempo global estimado de una pista en beats (pulsos) per minute (BPM). En terminología musical, el tempo es la velocidad o el ritmo de una pieza y se deriva directamente de la duración media del tiempo</p>
release_date	date	<p>La fecha en que se lanzó el álbum por primera vez</p>
speechiness	numerical	<p>speechiness detecta la presencia de palabras habladas en una pista. Cuanto más exclusivamente parecida a un discurso sea la grabación (por ejemplo, programa de entrevistas, audio libro, poesía), más cercano a 1.0 será el valor del atributo. Los valores superiores a 0,66 describen pistas que probablemente estén compuestas en su totalidad por palabras habladas. Los valores entre 0,33 y 0,66 describen pistas que pueden contener tanto música como voz, ya sea en secciones o en capas, incluidos casos como la música rap. Los valores por debajo de 0,33 representan música y otras pistas que no se parecen al habla</p>

valence	numerical	Una medida de 0.0 a 1.0 que describe la positividad musical que transmite una pista. Las pistas con valencia alta suenan más positivas (por ejemplo, feliz, alegre, eufórico), mientras que las pistas con valencia baja suenan más negativas (por ejemplo, triste, deprimido, enojado)
time_signature	categorical	Una signatura de tiempo total estimada de una pista. El compás (metro) es una convención de notación para especificar cuántos tiempos hay en cada compás
energy	numerical	La energía es una medida de 0.0 a 1.0 y representa una medida de percepción de intensidad y actividad. Por lo general, las pistas enérgicas se sienten rápidas y ruidosas. Por ejemplo, el death metal tiene mucha energía, mientras que un preludio de Bach tiene una puntuación baja en la escala. Las características de percepción que contribuyen a este atributo incluyen un rango dinámico: volumen percibido, timbre, frecuencia de inicio y entropía general