

Nomdedeu Nicolas

Proyecto CoderHouse #46150

Las claves de la musica
comercial

Profesor y Tutor :
Ruben Alfredo Parente Aun,
Juan Carlos Borges Pons

0:10



2:20

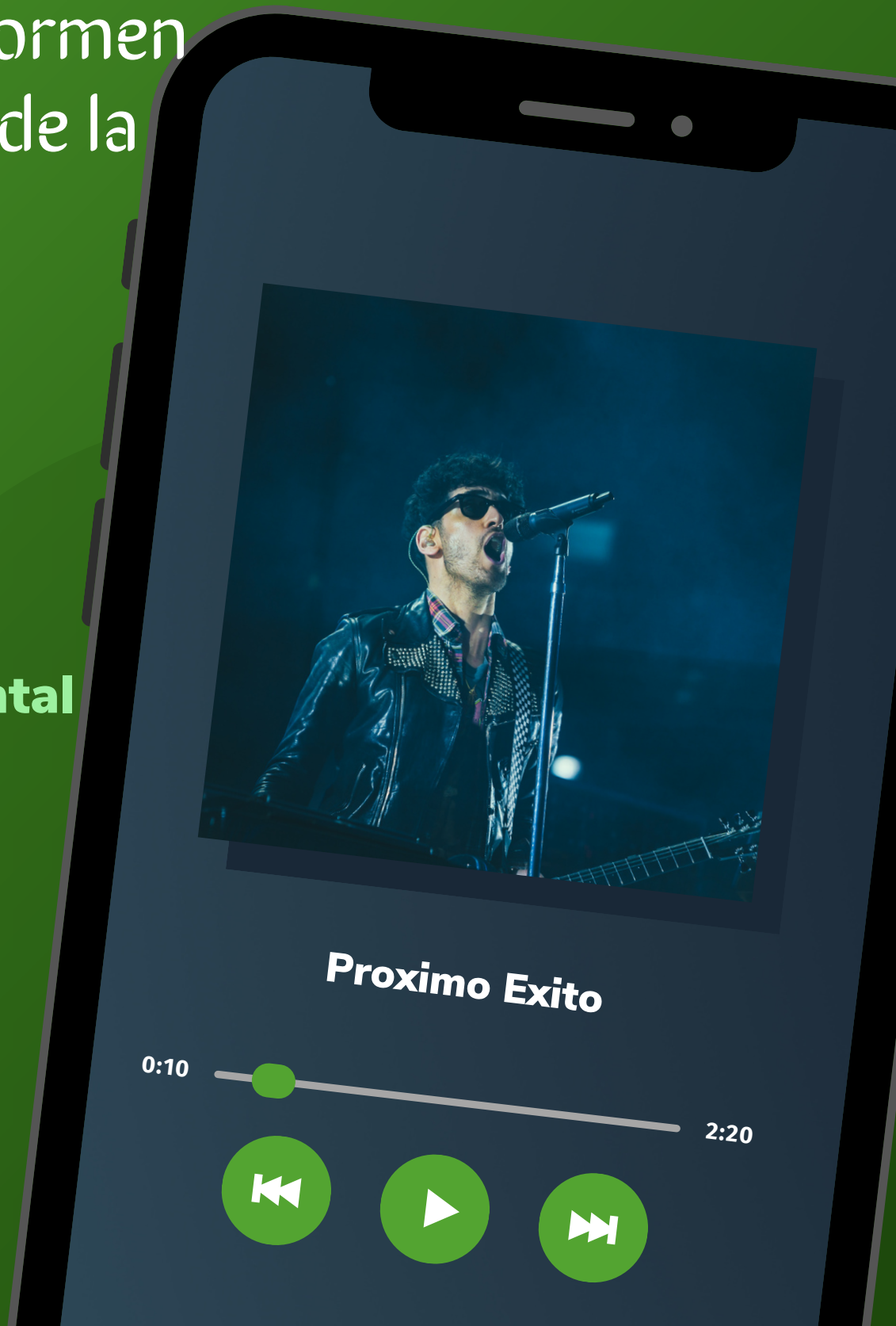


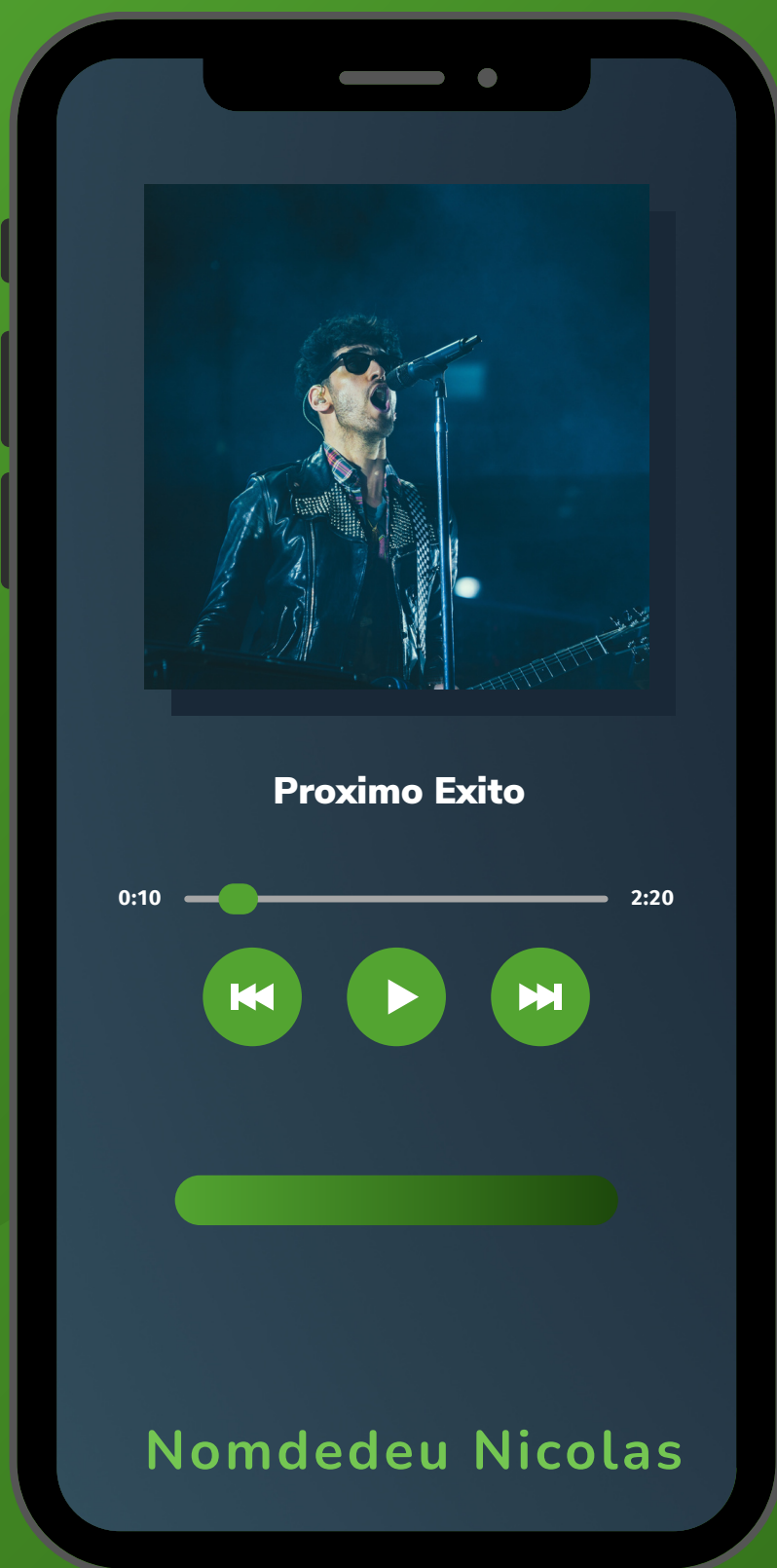
En las proximas semanas
estaremos en busqueda de
caracteristicas que conformen
los top hits y tendencias de la
industria musical



**Bailable, Energico,
Ruidoso o Acustico,
Sintetico, Instrumental**

**Existe una formula
para la musica?**





Esta base de datos contiene informacion sobre miles de canciones en Spotify y su rendimiento en Youtube, midiendo la cantidad de likes, comentarios y streams que reciben. Esto sumado a características musicales como el uso de instrumentos, el ruido, la energía y ritmo, vocalidad, duración, si es performada en vivo, uso de palabras en la descripción/título y positividad. Estas características serán observadas para analizar la siguiente hipótesis:

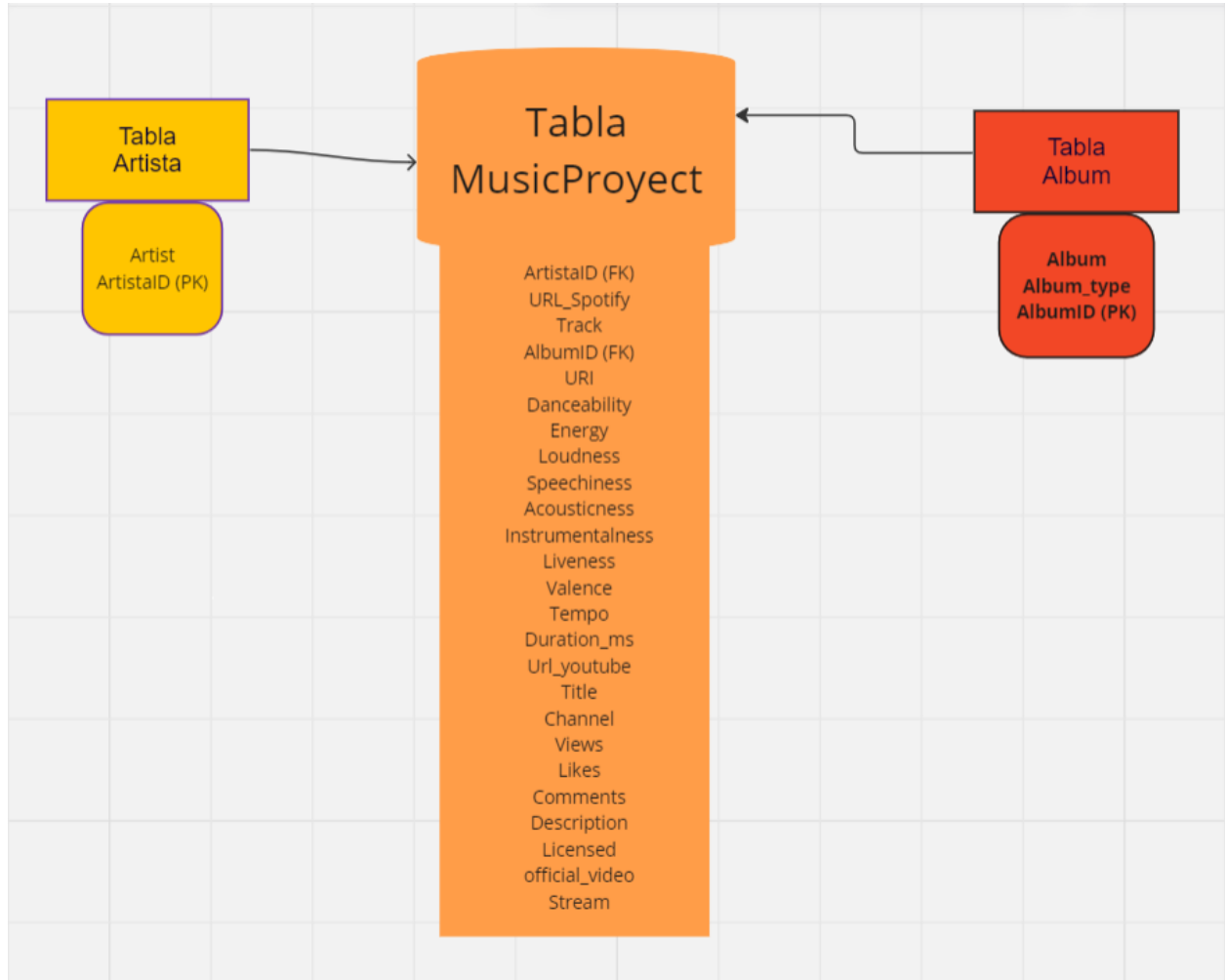
EXISTE UNA FORMULA PARA CREAR MUSICA COMERCIAL Y QUE ESTA REPERCUTA EN EL ENGAGEMENT DE LA BANDA

Nomdedeu Nicolás

Profesor y Tutor :Ruben Alfredo Parente Aun,Juan Carlos Borges Pons
Comision #46150

Documentación General del Proyecto

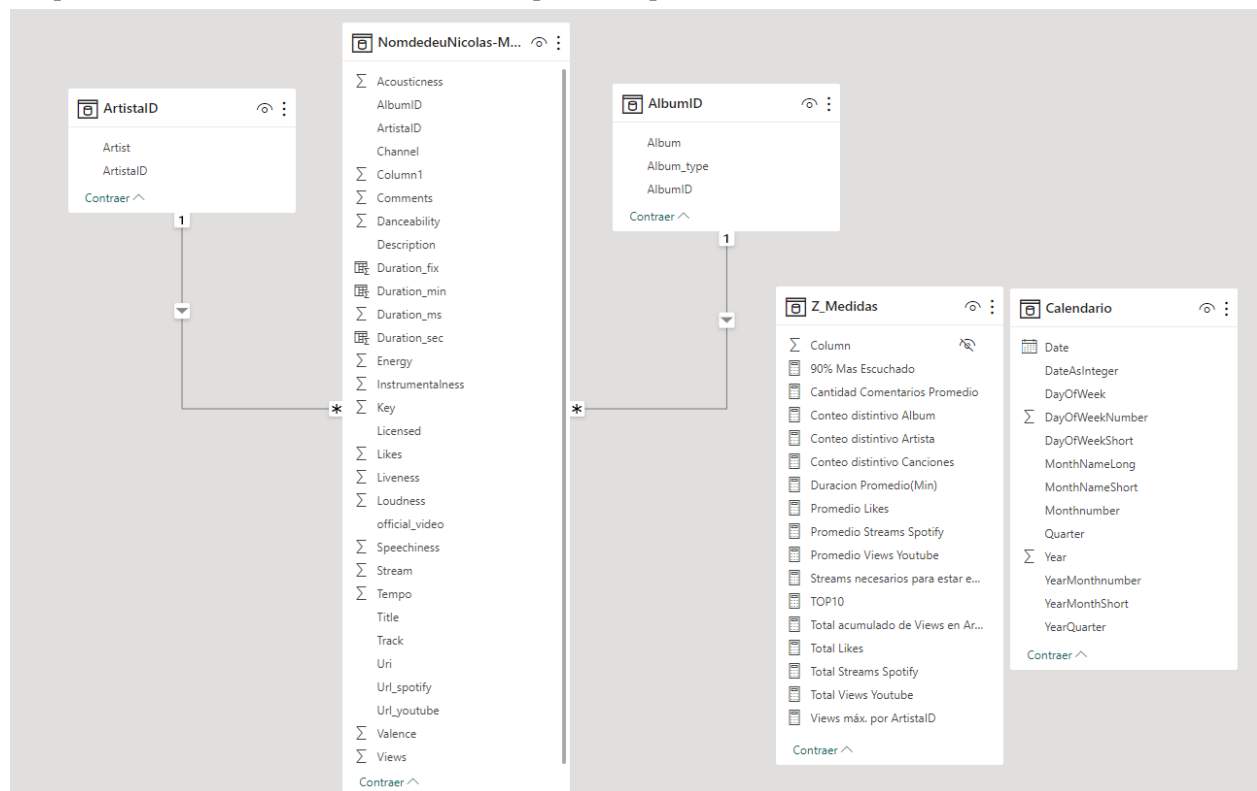
Este diagrama busca representar la relación entre las tres tablas que utilizaré en principio para el proyecto musical. Note que las variables Artista y el Álbum aparecen múltiples veces en la tabla original, por lo cual me pareció lógico normalizarlas y representarlas en otras tablas.



Para generar la relación entre las tablas generamos un ID para cada Artista único: ArtistaID y para cada Álbum único: AlbumID que funcionan como Primary Key de su respectiva tabla.

Luego reemplazamos las columnas de Artista y Álbum que contenían demasiados caracteres por sus números de ID, ahora estas nuevas columnas (nombradas ArtistaID y AlbumID de igual manera que las PK) serían la Foreign Key en la tabla principal y podemos relacionarlas con las tablas Artista y Álbum a través de los ID.

Luego incorporamos la Tabla Calendario ante la posibilidad de proveer fechas a nuestra base de datos y analizar la evolución de artistas y variables dinámicas. También sumamos una tabla de medidas, donde nos permite hacer cálculos estadísticos de importancia para el análisis.



De esta manera nuestro diagrama entidad relación queda armado por 5 tablas:

TABLA ARTISTA:

- **Artista:** Nombre de cada artista
- **ArtistaID:** Primary Key de la tabla, con un ID que corresponde individualmente con cada artista

TABLA ALBUM:

- **Album:** Album en el cual aparece la canción en Spotify
- **Album_type:** Especifica si forma parte de un álbum o un single
- **AlbumID:** Primary Key de esta tabla, enlazando cada Álbum/Single con un ID único.

TABLA MUSICPROYECT:

- **Track:** name of the song, as visible on the Spotify platform.
- **ArtistaID:** Foreign Key de la tabla Artista, es una columna con los ArtistaID
- **AlbumID:** Foreign Key de la tabla Album, columna con los AlbumID correspondientes
- **Url_spotify:** el Url del artista en Spotify

- **Uri:** spotify link que puede usarse para encontrar la canción usando una API.
- **Danceability:** Describe qué tanailable es la canción en una combinación de elementos musicales incluyendo tempo, ritmo, fuerza de bajos, y regularidad en general. 0 es menosailable y 1 másailable.
- **Energy:** Medida de 0 a 1 que representa la intensidad de la canción. Por ejemplo, qué tan rápido y ruidoso es. El Death Metal es muy energético, la música clásica puede no serlo.
- **Key:** “the key the track is in. Integers map to pitches using standard Pitch Class notation. E.g. 0 = C, 1 = C#/D ♭, 2 = D, and so on. If no key was detected, the value is -1.”
- **Loudness:** Qué tan ruidoso es el track en decibeles (dB), esto está promediado a través de la canción para generar comparación relativa del uso de sonidos más ruidosos o fuertes.
- **Speechiness:** Detecta la presencia de palabras vocalizadas en la canción. Los formatos que más utilizan la voz, (talk show o un audiolibro) tienden a 1 en esta categoría. Arriba de 0.66 es una canción que utiliza en su mayoría de tiempo palabras, entre 0.33 - 0.66 es una combinación de música y palabras, debajo de 0.33 son canciones con un estilo de no-lyrics.
- **Acousticness:** Una medida de si la canción está grabada en acústico, 1 demuestra gran confianza de que es una versión acústica.
- **Instrumentalness:** Predice la vocalidad de la canción, como los sonidos “Ooh” y “aah” son tomados como instrumentales, mientras que el Rap es claramente vocal. Cercano a 1 indica no vocalidad, arriba de 0.5 representa temas instrumentales pero es más confiable el número si es cercano a 1.
- **Liveness:** Detecta la presencia de una audiencia (en vivo), si es mayor a 0.8 hay mucha confianza de que fue grabada en vivo.
- **Valence:** una medida de 0.0 a 1.0 que describe la positividad musical transmitida por una pista. Las pistas con una valencia alta suenan más positivas (p. ej., felices, alegres, eufóricas), mientras que las pistas con una valencia baja suenan más negativas (p. ej., tristes, deprimidas, enfadadas).
- **Tempo:** el tempo global estimado de una pista en pulsaciones por minuto (BPM). En terminología musical, el tempo es la velocidad o ritmo de una pieza dada y se deriva directamente de la duración promedio del tiempo.
- **Duration_ms:** duración de la canción en milisegundos
- **Stream:** cantidad de streams en Spotify
- **Url_youtube:** url del video en youtube, si lo tiene.
- **Title:** titulo del video en youtube
- **Channel:** canal que publicó el video en youtube
- **Views:** views en youtube
- **Likes:** cantidad de likes en el video
- **Comments:** cantidad de comentarios
- **Description:** description del video en Youtube.
- **Licensed:** Indica si el video representa contenido con licencia, lo que significa que el contenido se subió a un canal vinculado a un socio de contenido de YouTube y luego fue reclamado por ese socio.
- **official_video:** valor en booleano (v/f) de si es el video oficial o no.

Tabla Calendario calculada a/t de este DAX: (que no podemos asociar específicamente a estos datos por la ausencia de fechas que hay en la base de datos)

```
Calendario = ADDCOLUMNS ( CALENDAR (DATE (2000, 1, 1), DATE (2015, 12, 31)), "DateAsInteger", FORMAT ( [Date], "YYYYMMDD" ), "Year", YEAR ( [Date] ), "Monthnumber", FORMAT ( [Date], "MM" ), "YearMonthnumber", FORMAT ( [Date], "YYYY/MM" ), "YearMonthShort", FORMAT ( [Date], "YYYY/mmm" ), "MonthNameShort", FORMAT ( [Date], "mmm" ), "MonthNameLong", FORMAT ( [Date], "mmmm" ), "DayOfWeekNumber", WEEKDAY ( [Date] ), "DayOfWeek", FORMAT ( [Date], "dddd" ), "DayOfWeekShort", FORMAT ( [Date], "ddd" ), "Quarter", "Q" & FORMAT ( [Date], "Q" ), "YearQuarter", FORMAT ( [Date], "YYYY" ) & "/" & "Q" & FORMAT ( [Date], "Q" ) )
```

- Date
- DateAsInteger
- DayOfWeek
- DayOfWeekNumber
- DayOfWeekShort
- MonthNameLong
- MonthNameShort
- MonthNumber
- Quarter
- Year
- YearMonthNumber
- YearMonthShort
- YearQuarter

Tabla “Z_ Medidas”:

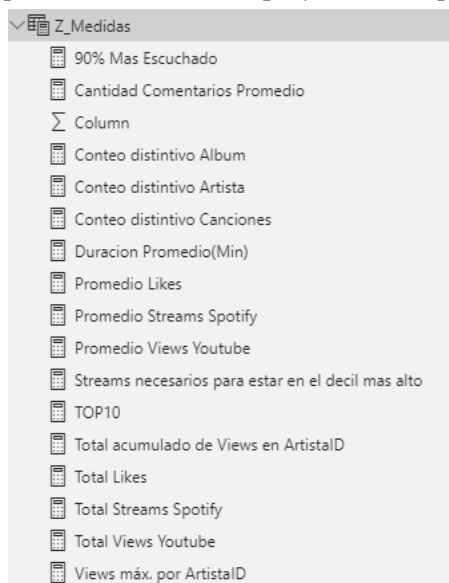
- 90% Más escuchado
- Cantidad Comentarios Promedio
- Conteo distintivo Album
- Conteo distintivo Artista
- Conteo distintivo Canciones
- Duración Promedio (min)
- Promedio Likes
- Promedio Streams Spotify
- Promedio Views Youtube
- Streams necesarios para estar en decil más alto
- TOP10
- Total Acumulado de Views en ArtistaID
- Total Likes
- Total Streams Spotify

- Total Views Youtube
- Views Max por ArtistaID

Dejo en presente algunos de los códigos DAX utilizados para el calculo de estas medidas:

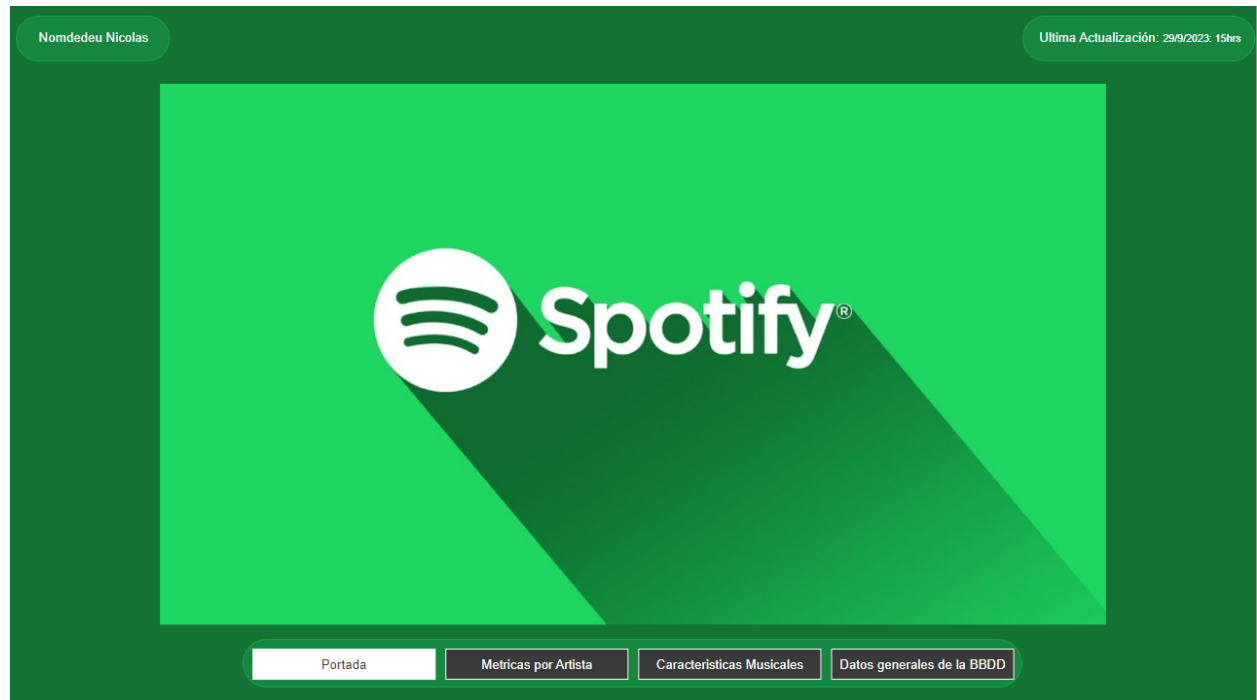
- Conteo distintivo Album = `DISTINCTCOUNT(AlbumID[Album])`
- Conteo distintivo Artista = `DISTINCTCOUNT(ArtistaID[Artist])`
- Conteo distintivo Canciones = `DISTINCTCOUNT('NomdedeuNicolas-MusicProyect'[Track])`
- Duracion Promedio(Min) = `AVERAGE('NomdedeuNicolas-MusicProyect'[Duration_min])`
- Promedio Likes = `AVERAGE('NomdedeuNicolas-MusicProyect'[Likes])`
- Promedio Streams = `AVERAGE('NomdedeuNicolas-MusicProyect'[Stream])`
- Promedio Views Youtube = `AVERAGE('NomdedeuNicolas-MusicProyect'[Views])`
- Total Likes = `SUM('NomdedeuNicolas-MusicProyect'[Likes])`
- Total Streams Spotify = `SUM('NomdedeuNicolas-MusicProyect'[Stream])`
- Total Views Youtube = `SUM('NomdedeuNicolas-MusicProyect'[Views])`
- Views máx. por ArtistaID
= `MAXX(KEEPFILTERS(VALUES('NomdedeuNicolas-MusicProyect'[ArtistaID])),
CALCULATE(SUM('NomdedeuNicolas-MusicProyect'[Views])))`
- 90% Mas Escuchado = `PERCENTILE.EXC('NomdedeuNicolas-MusicProyect'[Stream],0.1)`
- Streams necesarios para estar en el decil mas alto =
`PERCENTILE.EXC('NomdedeuNicolas-MusicProyect'[Stream],0.9)`

Estas Medidas son utilizadas como datos para ayudar a contar una historia y fortalecer argumentos planteados en nuestra proyección. Aquí un pantallazo de estas medidas:



Procedo a dejar imagenes de como luce el progresó de creación de los dashboards:

Portada:



Primera solapa informativa del Dashboard: Métricas y Estadísticas Generales (antes)



1. Solapa de Métricas y Estadísticas con la adición de Tooltips para poder hacer una comparativa entre la plataforma de Youtube y Spotify:



2. Características Musicales, buscamos poder estudiar la naturaleza de los Artistas



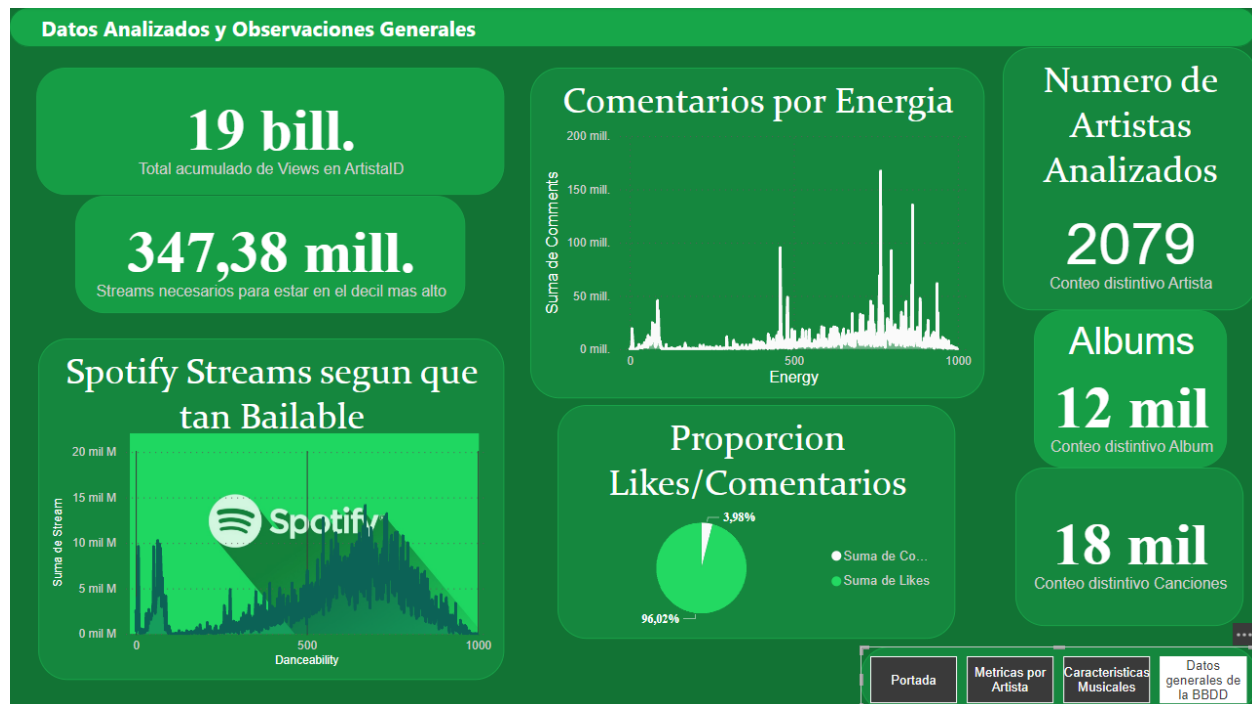
2. Lo transformamos en la solapa para conocer musicalmente, las características que forman la identidad del artista, a través de números. La idea es entender si el Artista busca ser bailable, energético, instrumental, o vocal, y cuales resultados generan ese estilo:



3. La tercer solapa buscaba ser un estudio del mercado musical, con estadísticas generales de que busca el mundo con la industria musical:



3. Luego de ciertas conclusiones mixtas, determiné que mi objetivo es más que ambicioso, uno no podría generalizar a más de 2000 artistas distintos con audiencias tan variadas, por lo cual tome ciertas observaciones de interés y use este espacio para contar de qué se trata tanto la base de datos como el proyecto. Esta es la solapa final:



De esta manera obtuve un Dashboard informativo, que puede ser utilizado para la observación con datos duros de los más grandes Artistas mundiales, que analiza resultados entre las dos plataformas mas masivas de reproduccion de musica Youtube y Spotify, con variables algorítmicas de gran importancia como los Likes y los Comentarios, que pueden verse reflejado como una medida de éxito en la industria, o al menos de cuántos individuos disfrutaron de dicho contenido. Me gustaría concluir con esta idea que resonó en mi cabeza durante la creación de este proyecto, la música no tiene una sola fórmula. Se trata de estilos, que dentro de estos pueden existir formas de hacer música más apetecible y destacada. Uno de los gráficos más llamativos de este informe es el que muestra la relación entre Bailable y los Streams en Spotify (última solapa en la esquina izquierda). Es claro como hay una tendencia que creíamos lógica, la música se baila por ende mientras más bailable más se escucha. Ahora también existe una gran concentración de los Streams de Spotify en canciones con valores bailables cercanos a 0.

Esto demuestra la esencia de la música, se adapta a los gustos y sensaciones de millones de personas, es, y siempre será un grave error tratar de generalizarlo a una simple fórmula.