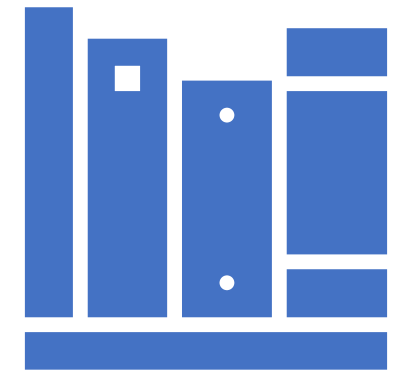


Ranking Spotify



Abstract

El Dataset contiene información sobre distintos aspectos de las canciones que se encuentran en Spotify.

Cuenta con datos cualitativos (tonalidad, ritmo, energía, etc) y cuantitativos (cantidad de reproducciones, comentarios, etc) de 20.718 canciones individuales.

Los datos fueron extraídos al 7 de Febrero de 2023 por lo que la cantidad de reproducciones, visualizaciones, comentarios y cualquier otro dato variable habrá cambiado a esta altura.

Es necesario aclarar que no se están tomando en cuenta canciones que se hayan incorporado a Spotify después de esa fecha.

Fuete: <https://www.kaggle.com/datasets/salvatorerastelli/spotify-and-youtube>



Contexto

Conocer las características distintivas de las canciones que pueden generar interés del público puede ser de gran valor para productores musicales y artistas que quieran tener un marco de referencia.

Aunque creemos que hay factores extramusicales que influyen en el interés por un artista o una canción, el conocimiento de parámetros formales brinda información relevante a la hora de componer.



Resumen de metadata

- 28 columnas
- 20.718 filas
- Variable dependiente: Stream (float64)
- Variables independientes:
 - Album_type (object)
 - Danceability (float64)
 - Energy (float64)
 - Key (float64)
 - Loudness (float64)
 - Speechiness (float64)
 - Acousticness (float64)
 - Instrumentalness (float64)
 - Liveness (float64)
 - Valence (float64)
 - Tempo (float64)
 - Views (float64)
 - Likes (float64)
 - Comments (float64)



Hipótesis y líneas de análisis

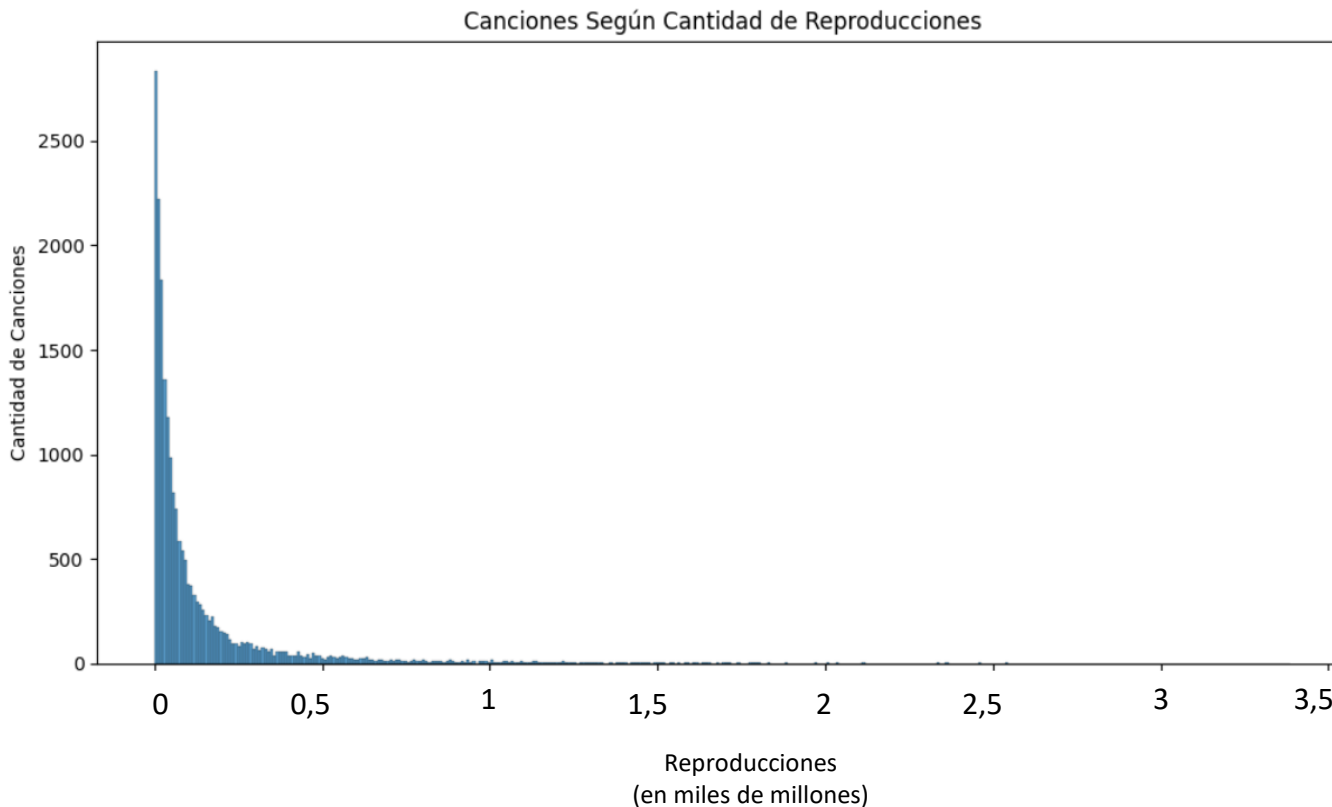
Con el presente dataset trataremos de identificar variables que expliquen por qué una canción o artista son más escuchados que otros.

Partimos de la hipótesis que las cualidades intrínsecas de la canción como la tonalidad en la que está compuesta, su ritmo, la posibilidad de ser bailada o su energía son relevantes en este aspecto.



Análisis exploratorio Univariado

¿Cómo está distribuida la cantidad de reproducciones?



Primero revisamos si la cantidad de reproducciones de las canciones nos muestran alguna agrupación a tener en cuenta.

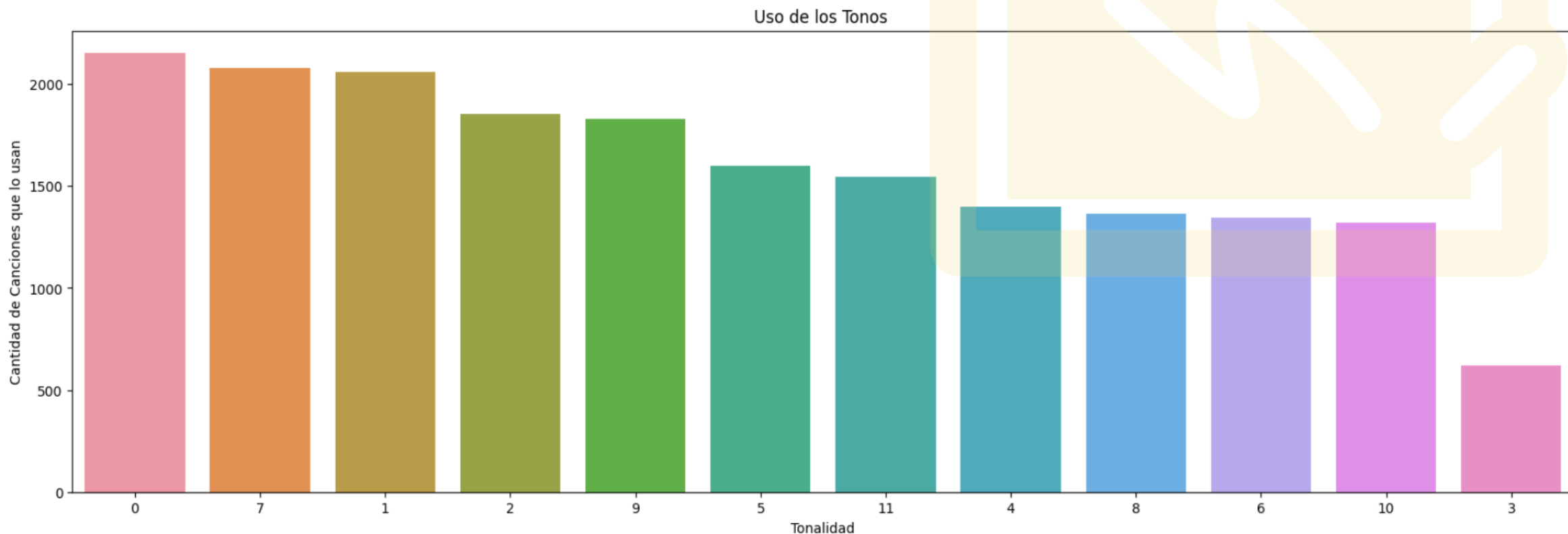
- Observamos que la mayoría de las canciones tienen muy pocas reproducciones.
- No aparecen grupos a primera vista que puedan ser identificados.

Análisis exploratorio Univariado

¿Hay tonalidades más usadas?

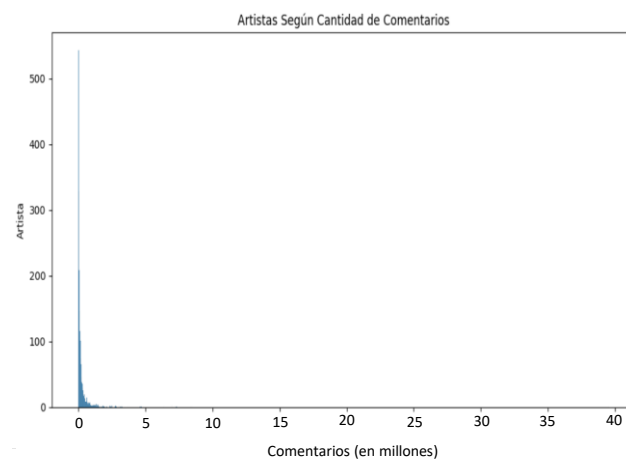
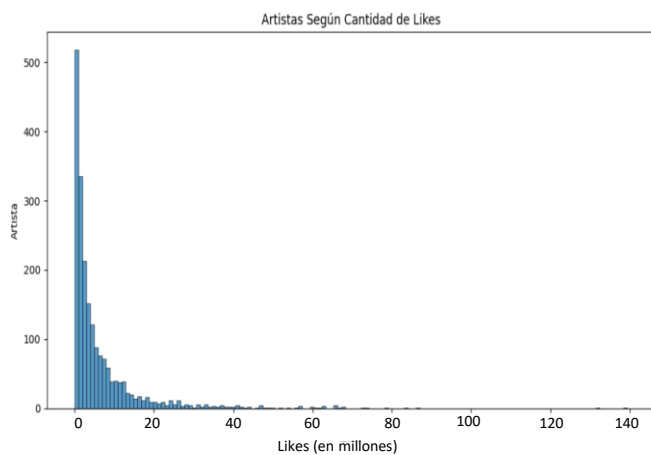
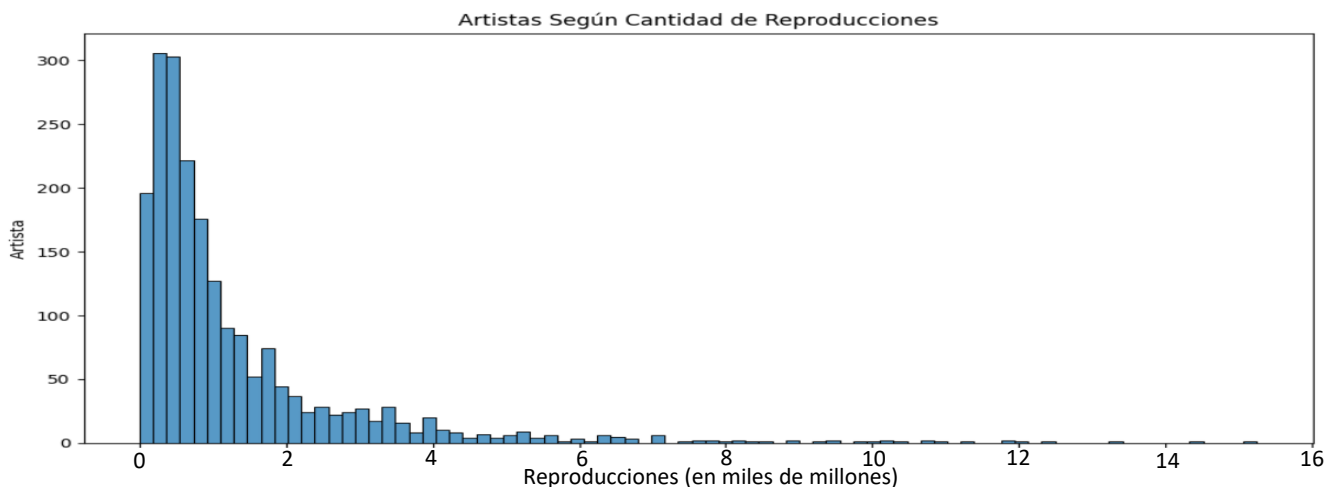
Siguiendo con esta línea, comprobamos si hay alguna tonalidad más utilizada.

- El siguiente gráfico de barras nos muestra que no hay diferencias significativas en el uso de los tonos.
- La tonalidad RE# (3) es la menos utilizada y se corresponde con el análisis anterior que también es la menos escuchada. Tiene sentido porque si hay menos canciones en esa tonalidad, va a tener menos reproducciones.



Análisis exploratorio Univariado

¿Hay alguna relación entre la cantidad de Reproducciones, Likes y Comentarios por Artista?

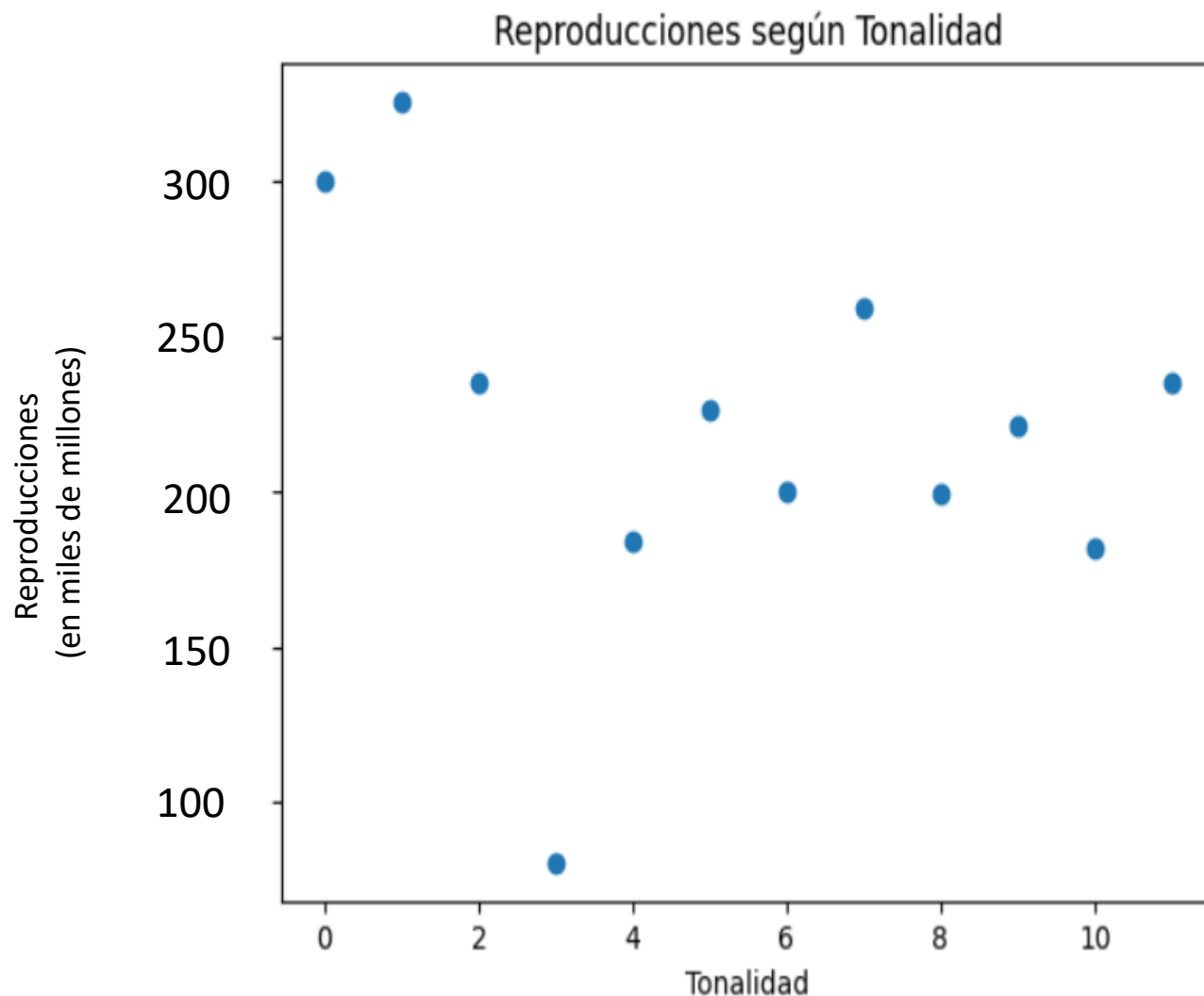


Aquí podemos ver las reproducciones, los likes y los comentarios totales por artista.

- Los artistas más escuchados no coinciden con los que tienen más likes.
- Este comportamiento puede explicarse si se tiene en cuenta que una canción puede reproducirse cuantas veces se quiera y el like va a contar solo una vez.
- También podría explicarse por el grado de participación de la audiencia. Quizás los oyentes de un artista o grupo tiendan a ser más participativos que otros.

Análisis exploratorio Bivariado

¿Hay tonalidades más escuchadas que otras?

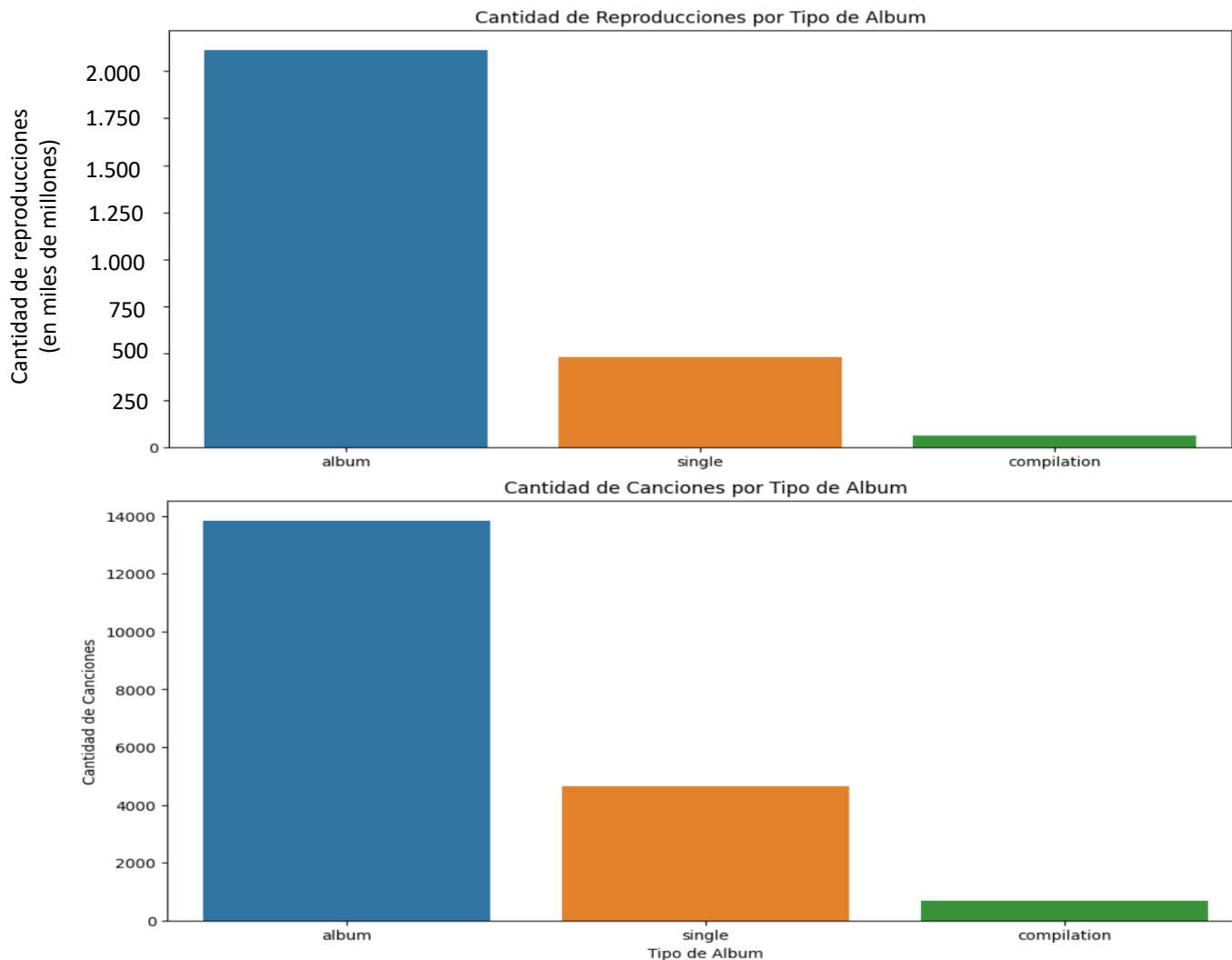


Partimos de la hipótesis que cualidades intrínsecas a la canción, como el tono en el que están compuestas, pueden tener relación con la cantidad de reproducciones.

- La mayoría de las tonalidades están entre 190 mil millones y 260 mil millones de reproducciones totales.
- Solo la tonalidad RE# (3) está por debajo de los 100 mil millones de reproducciones.
- Mientras que DO (0) y DO# (1) son claramente las tonalidades más escuchadas.

Análisis exploratorio Bivariado

¿Son más escuchas las canciones que son parte de un disco o las que son lanzadas como sencillos?



Revisamos si la manera de publicar canciones (álbum, single o compilación) influye en la cantidad de reproducciones.

- Observamos que la mayoría de las escuchas están dentro de álbumes.
- Esas canciones representan el 75% del dataset por lo que es de esperarse que tengan mayor cantidad de reproducciones.

Insights y recomendaciones

Insights

- La mayoría de las canciones tienen muy pocas reproducciones.
- No aparecen grupos a primera vista.
- No hay diferencias significativas en el uso de los tonos.
- La tonalidad RE# (3) es la menos utilizada y la menos escuchada.
- Los artistas más escuchados no son los que tienen más likes.
- Hay que tener en cuenta la participación de la audiencia.
- Solo la tonalidad RE# (3) está por debajo de los 100 mil millones de reproducciones.
- DO (0) y DO# (1) son claramente las tonalidades más escuchadas.
- La mayoría de las escuchas están dentro de álbumes.



Recomendaciones

- Proponemos incentivar la escucha de canciones similares sugiriendo canciones con cualidades (tono, ritmo, etc) similares a la de canción reproducida.
- Sugerimos recomendar canciones del mismo artista, pero que tengan menor cantidad de reproducciones que el promedio para el músico.
- Proponemos aprovechar la participación del público ludificando la escucha activa.
- Sugerimos crear listas de reproducción que mezclen las canciones de acuerdo a las características de bailable o no, más enérgicas o menos enérgicas, usando nombres como "Para bailar toda la noche" o "Para escuchar en tranquilidad" con canciones en tonalidades distintas a DO y DO#.



¡Muchas Gracias!

