

ÉXITO EN SPOTIFY

**PROYECTO FINAL DATA SCIENCE
CODER HOUSE - COMISIÓN 42390**

MELISA CAPUTO

AGENDA

- 1 MOTIVACIÓN Y AUDIENCIA
- 2 RESUMEN DE METADATA
- 3 PREGUNTAS HIPÓTESIS
- 4 VISUALIZACIONES
- 5 INSIGHTS

MOTIVACIÓN Y AUDIENCIA

En la era digital, la música se volvió más accesible que nunca gracias a plataformas como Spotify. Detrás de la aparente simplicidad de escuchar una canción en línea, se esconde un mundo de datos y características musicales que influyen en la cantidad de reproducciones.

El objetivo de este proyecto es realizar un análisis detallado de las tendencias y patrones musicales de los últimos tiempos para elaborar un modelo de Machine Learning que permita predecir el éxito de una canción.

Con este proyecto, se espera ayudar a la industria musical a tomar decisiones más informadas y precisas en cuanto a la producción y promoción de nuevos lanzamientos, lo que en última instancia beneficiará a plataformas de streaming, discográficas, artistas y oyentes.

METADATA

La plataforma Spotify proporciona un dataset que contiene información muy útil para el proyecto. El dataset original contaba con una cantidad muy numerosa de registros ya que contenía el ranking diario de los años 2017 y 2018 para cada región del mundo. Para simplificar el análisis, se trabajó con un nuevo dataset filtrado por la región estadounidense como muestra de la población total. Luego de la limpieza y transformación del dataset, se cuenta con una cantidad de **72692 registros** y **20 columnas**.

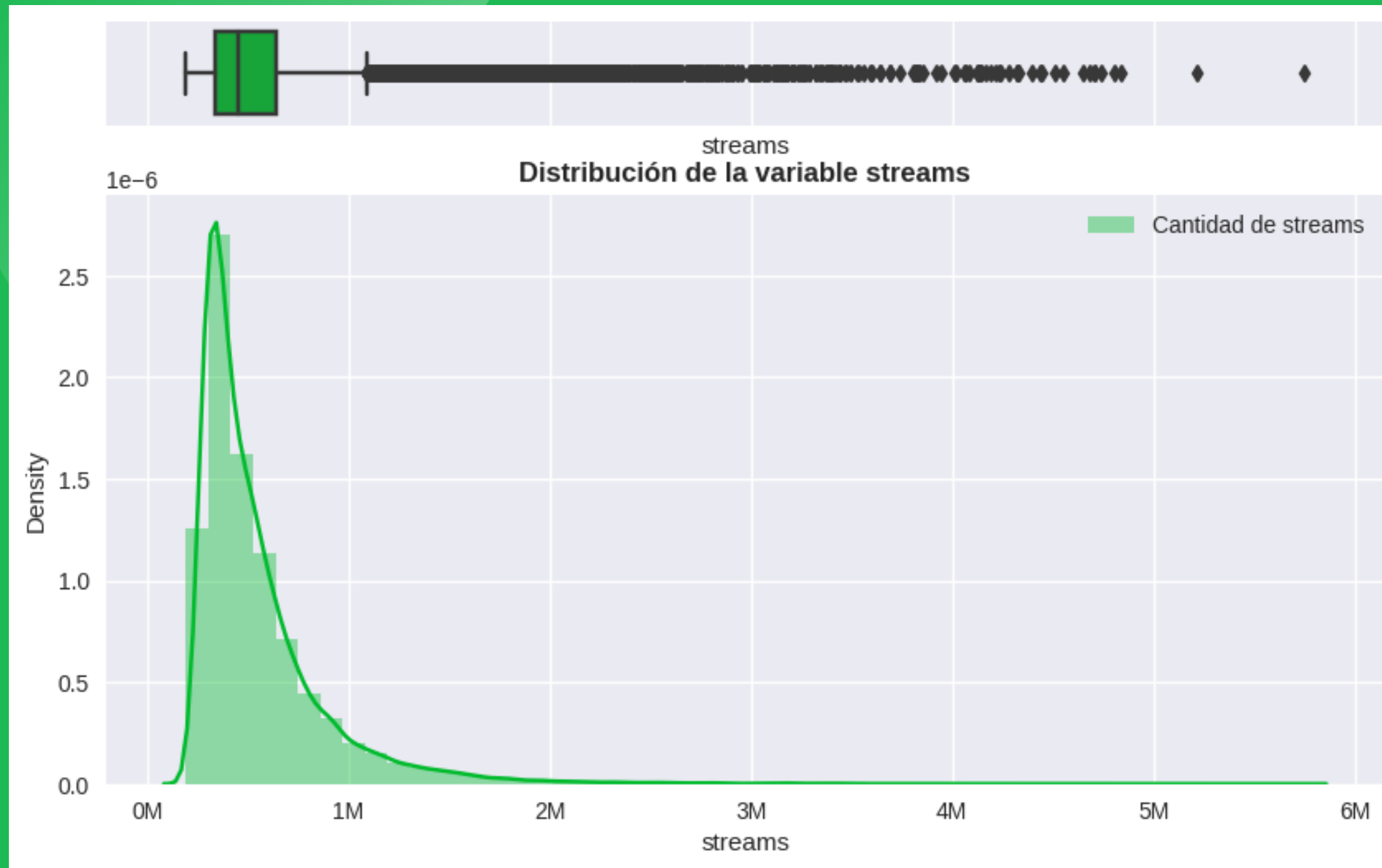
Las variables principales son: posición en el ranking, nombre de la canción, artista, cantidad de reproducciones, año, mes y día del ranking. El resto de las variables corresponden a características musicales medidas en diferentes rangos numéricos: bailabilidad, energía, clave musical, ruido, modo (mayor o menor), presencia de palabras habladas, acústica, volumen de instrumentos, probabilidad de que se haya grabado en vivo, positivismo, tempo, duración, compás medio.

PREGUNTAS DE INTERÉS

- ¿Cómo se distribuye la cantidad de streams en Spotify?
- ¿Cuántas reproducciones suelen tener las canciones más exitosas?
- ¿Se pueden identificar relaciones de dependencia entre las distintas características musicales?
- ¿Quiénes son los artistas que tienen mayor popularidad en la plataforma?
- ¿Cómo influye el género musical en la cantidad de reproducciones?
- ¿Existen géneros que tienden a ser más populares que otros?
- ¿Cuál es el impacto de la duración de una canción en su éxito?
- ¿Suelen ser más exitosas las canciones más cortas?
- ¿Existen patrones en la popularidad de las canciones lanzadas en días específicos de la semana?
- ¿Existen patrones en la popularidad de las canciones lanzadas en ciertos meses del año?

VISUALIZACIONES

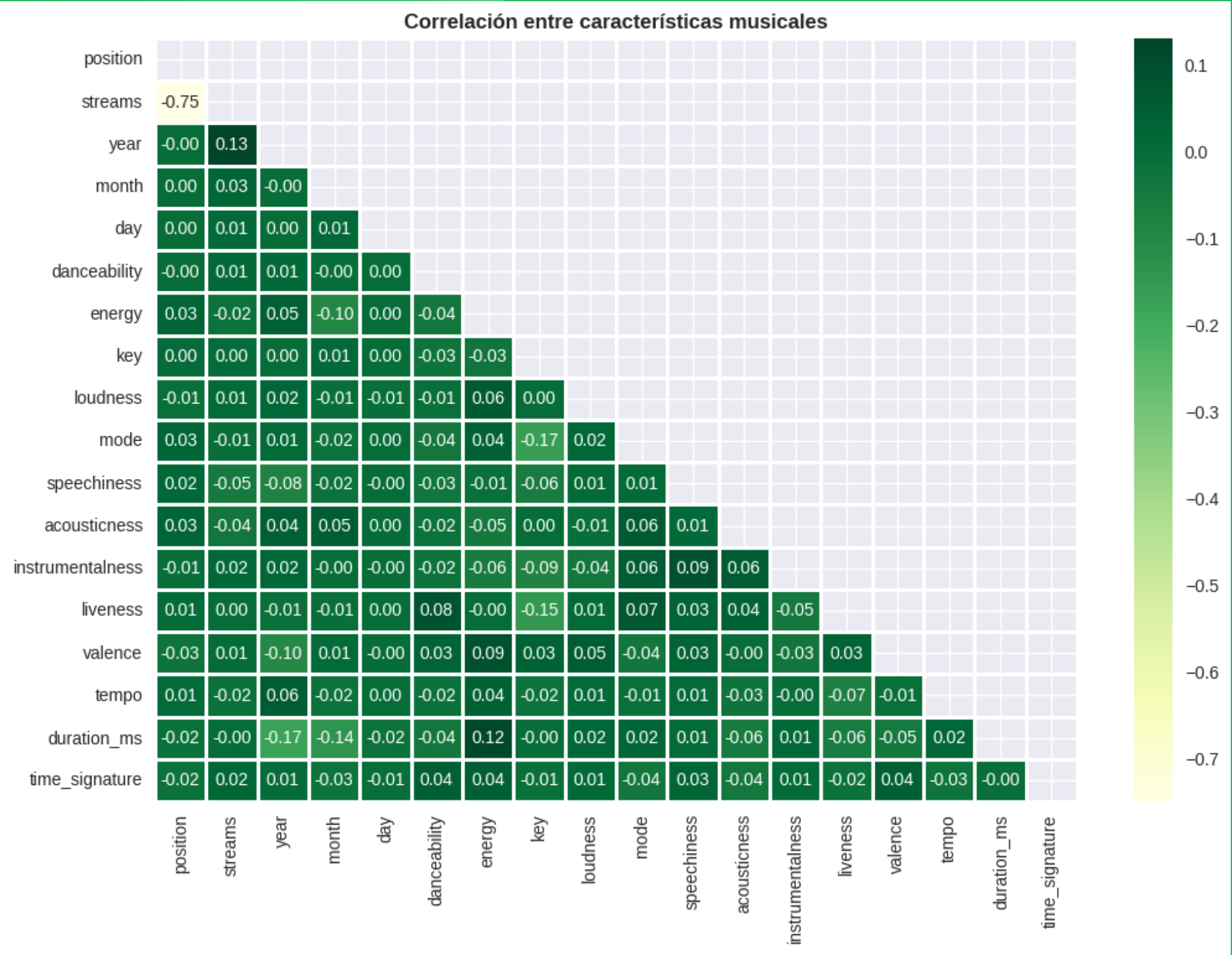
DISTRIBUCIÓN DE LAS REPRODUCCIONES



El promedio diario de reproducciones gira en torno a los 500K, mientras que un porcentaje menor al 25% de las canciones logran superar el millón de reproducciones.

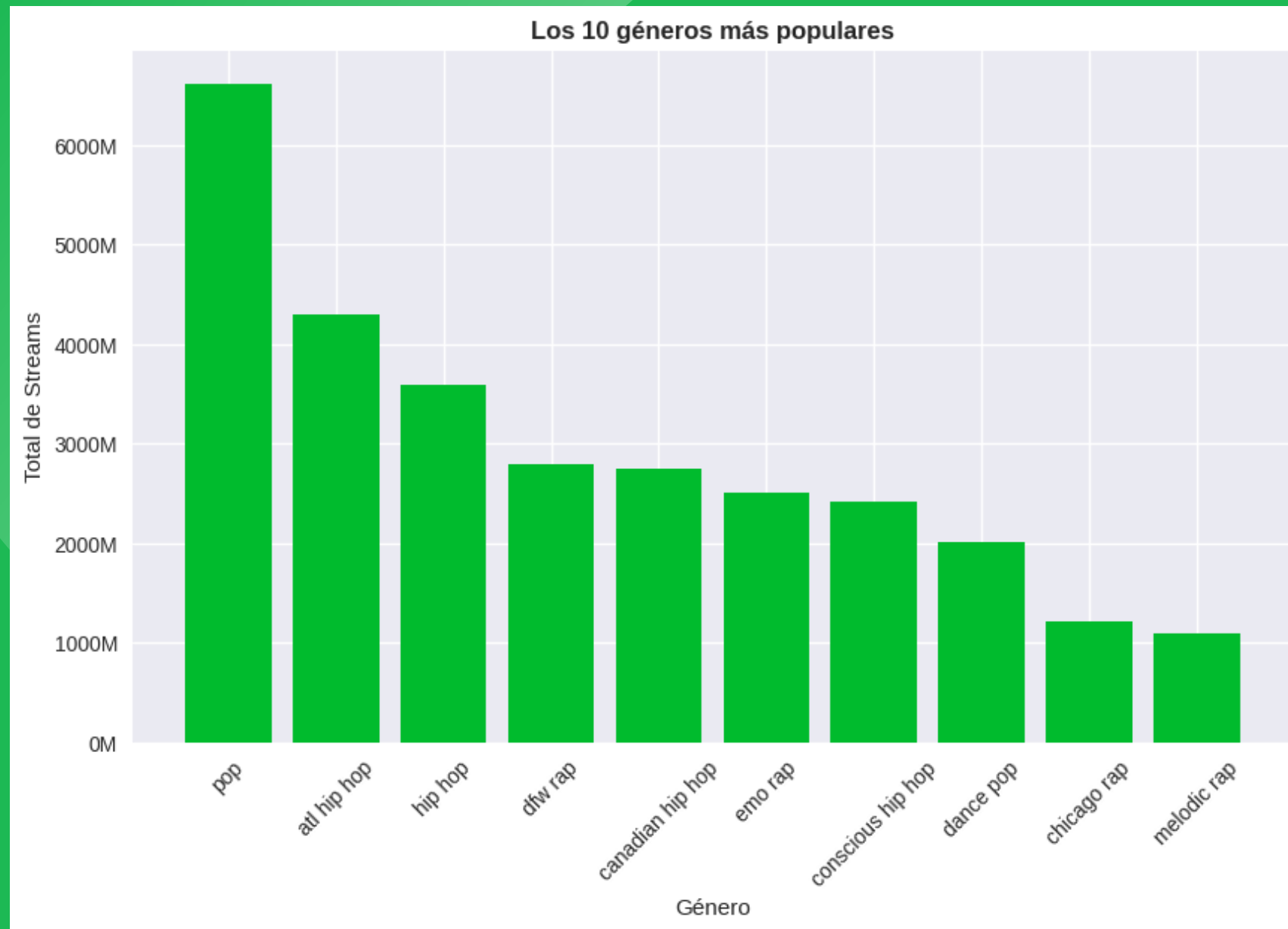
El contraste entre la mayoría de canciones con un bajo número de streams y la presencia de un grupo notable de canciones con un alto número de streams puede deberse a diferentes factores como éxitos internacionales, temas virales o lanzamientos de artistas populares.

RELACIONES ENTRE LAS VARIABLES

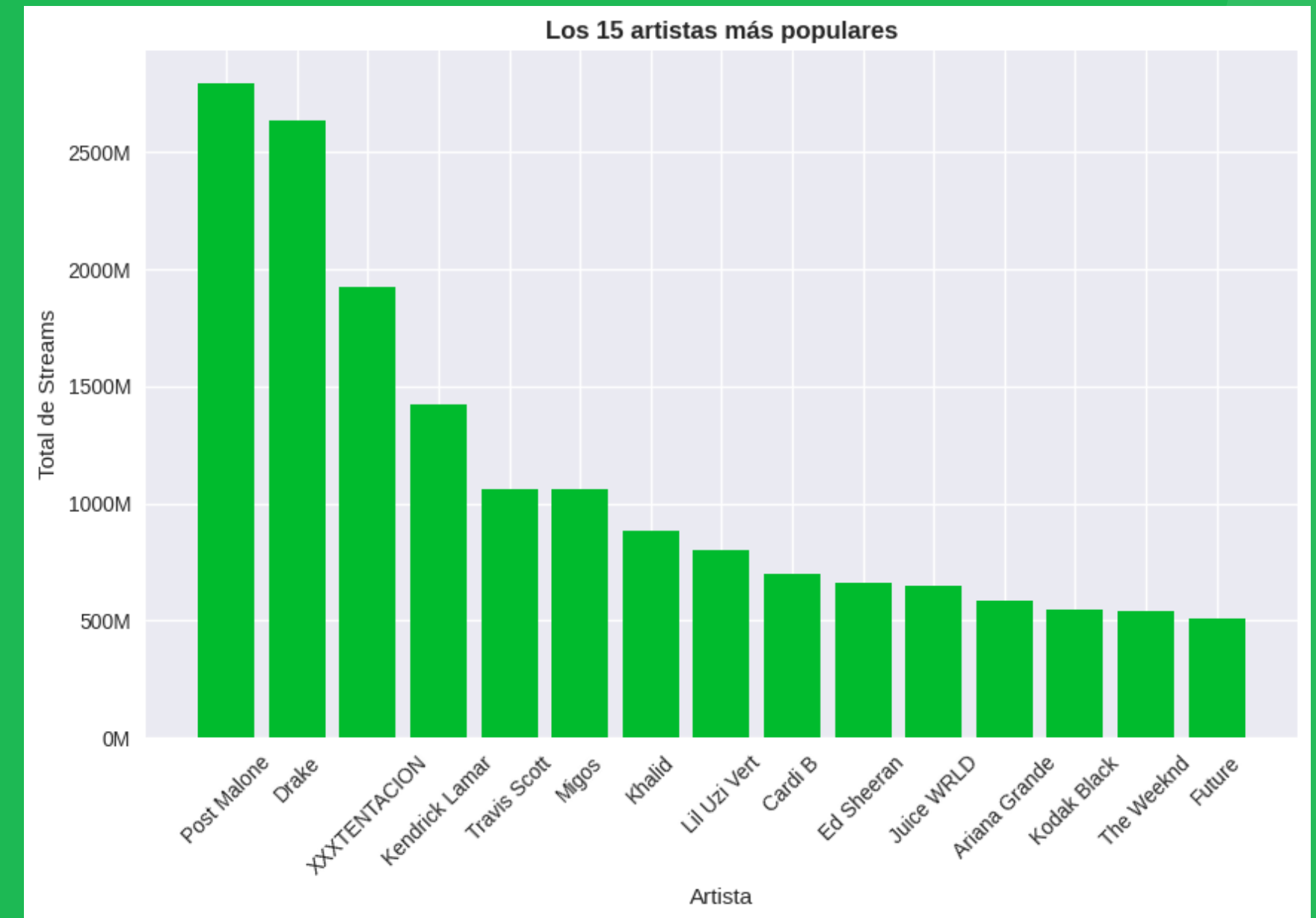


No existen relaciones directas claras entre las variables ni tampoco al compararlas con la cantidad de streams, salvo por la posición del ranking, cosa que tiene sentido ya que es una variable que depende directamente de la cantidad de reproducciones

ARTISTAS Y GÉNEROS

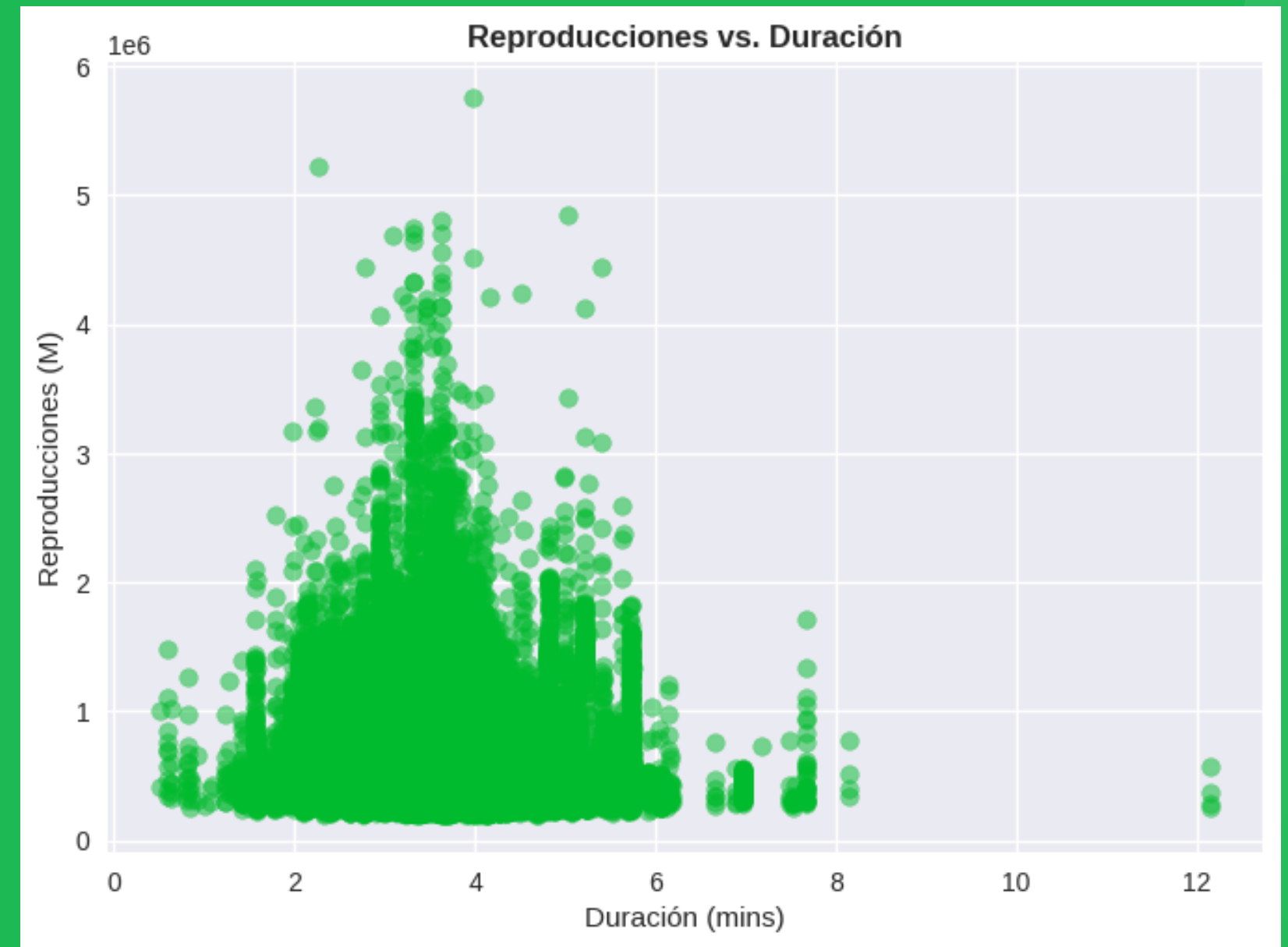
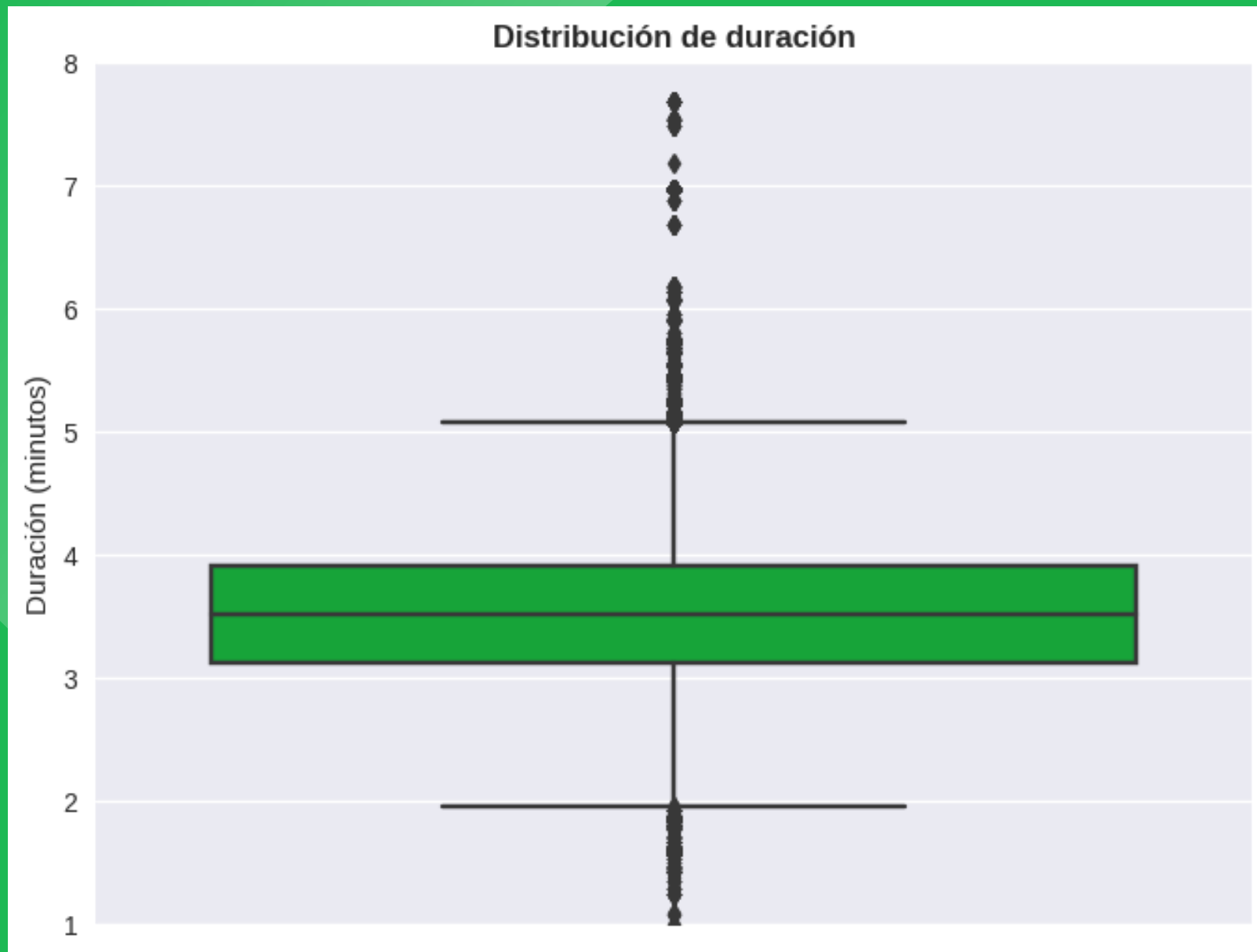


Las diferentes variantes de Hip Hop y Rap se encuentran entre los géneros más populares junto con el pop.



La mayoría de artistas ubicados en el top 15 pertenecen al género Hip Hop / Rap, confirmando su popularidad que veíamos en el gráfico anterior.

DURACIÓN DE LAS CANCIONES

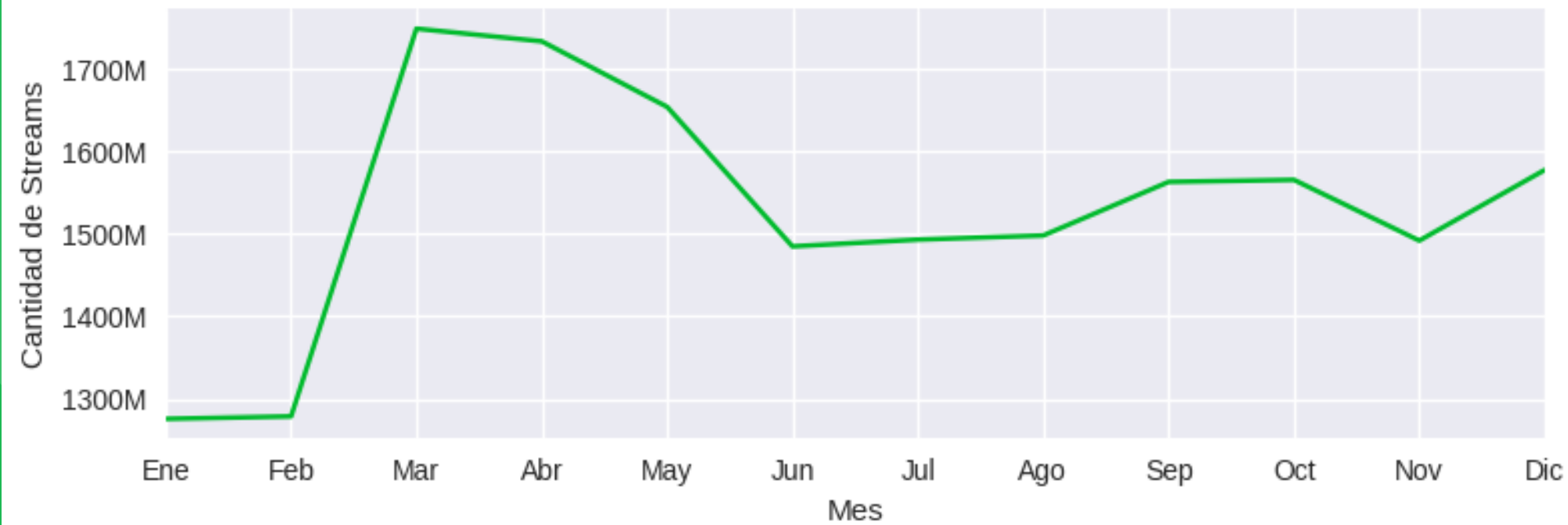


El promedio de duración de las canciones es de entre 3 a 4 minutos.

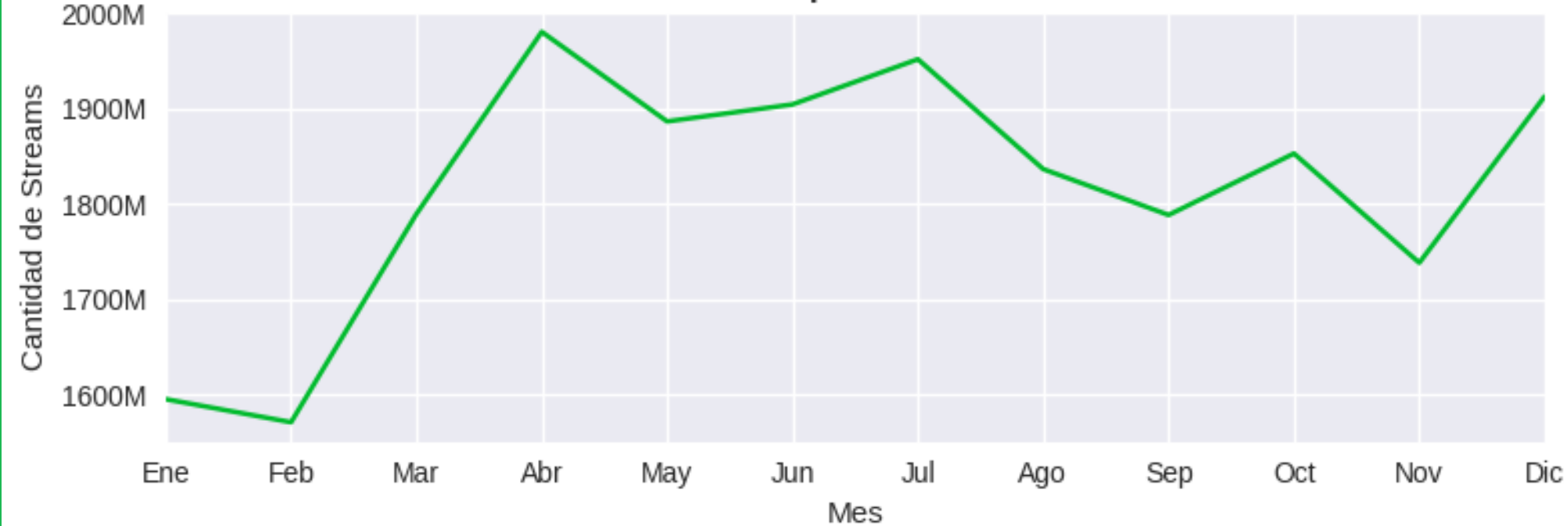
Las canciones con una duración promedio tienen más posibilidades de ser exitosas, mientras que las canciones con corta o larga duración no consiguen más de 2M reproducciones.

STREAMINGS POR MES

Streams por mes - 2017

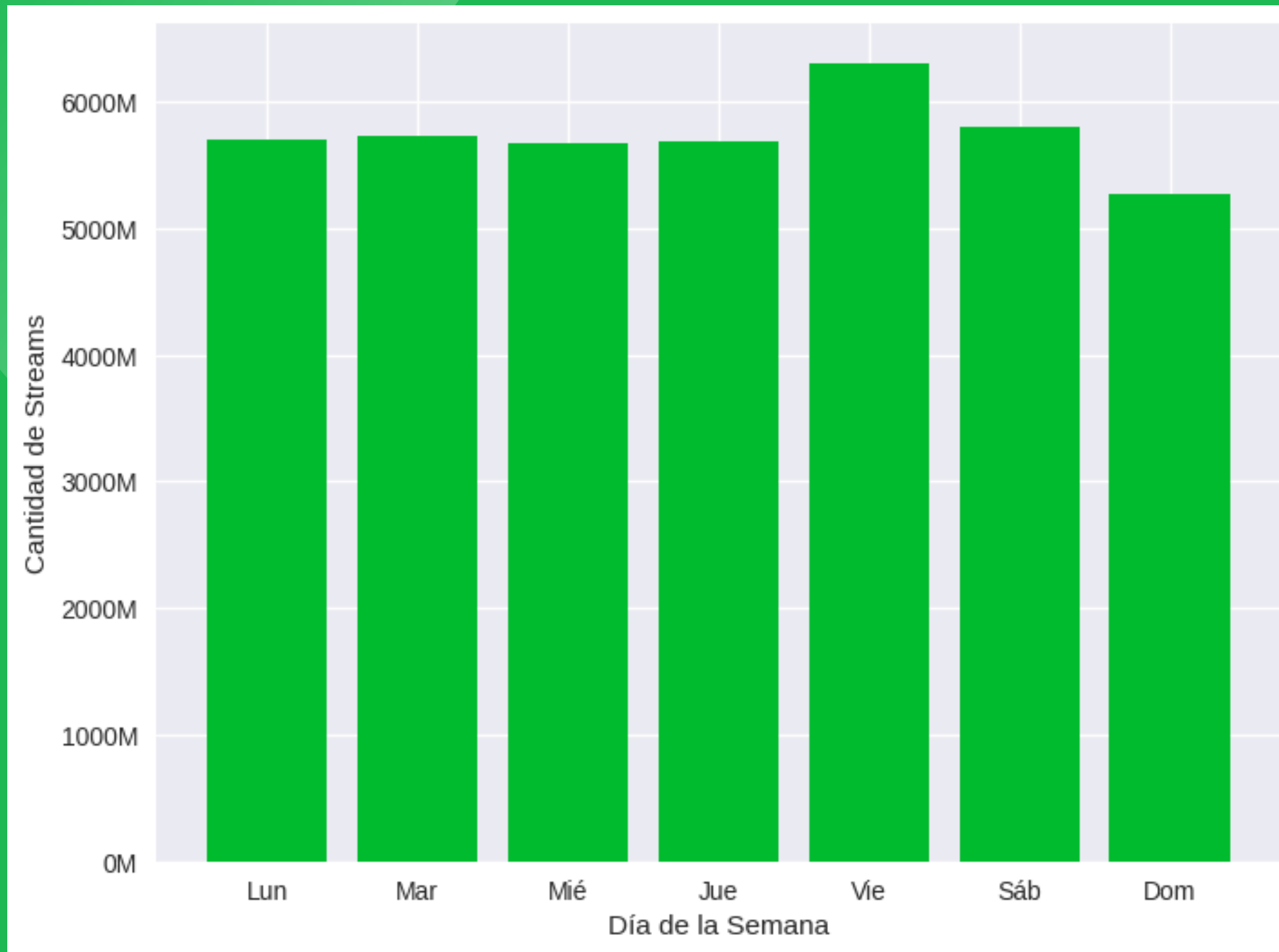


Streams por mes - 2018



Los dos años analizados muestran un patrón similar en la evolución de la cantidad de streams durante los primeros meses del año. Los primeros dos meses tienen la menor cantidad de streams, lo que podría deberse a las tendencias de consumo durante las vacaciones o días festivos. Luego, se observa un aumento exponencial en ambos años, alcanzando su punto máximo en marzo para el año 2017 y en abril para el año 2018.

STREAMINGS POR DÍA



Si bien la cantidad de reproducciones por día de la semana parece ser bastante equilibrada, los viernes se destacan como los días con mayor cantidad de reproducciones.

INSIGHTS

Se recomienda tener en cuenta las siguientes conclusiones a la hora de realizar un lanzamiento musical:

- Una canción puede considerarse más exitosa que el promedio si supera las 500k reproducciones diarias
- Las características musicales no tienen relaciones directas fuertes entre sí, existe mucha diversidad
- Los artistas más exitosos pertenecen a los géneros más exitosos
- Una colaboración entre dos o más artistas de Hip Hop / Rap puede resultar en una canción muy exitosa
- Las canciones con una duración promedio de entre 3 a 4 minutos son las más exitosas
- Los primeros dos meses del año no suelen tener mucho éxito en reproducciones
- Abril parece ser el mejor mes para lanzamientos musicales
- Los viernes son los mejores días para lanzamientos musicales

En el notebook complementario a esta presentación se desarrolla un modelo de regresión lineal que tiene en cuenta todas estas cuestiones a un nivel mucho más detallado.

¡GRACIAS!