

# Predicción de *hits* a partir de atributos musicales



Comisión 39960

Curso Data Science

Catalina Miranda

# TABLA DE CONTENIDOS

1. Contexto y audiencia



2. Dataset



3. Hipótesis



4. Análisis exploratorio



5. Insights y recomendaciones

# CONTEXTO Y AUDIENCIA

En un contexto mundial de alta competitividad comercial, una discográfica musical está buscando diferenciarse de sus competidoras. Para ello, necesita saber qué es lo que hace que una canción se convierta en un *hit*.

# DATASET



El conjunto de datos seleccionado agrupa a todas las canciones que aparecieron en el Hot 100 del Billboard Chart desde el año 1958 hasta el 2022. La clasificación no solo está basada en las ventas físicas y digitales de los Estados Unidos, sino también en las reproducciones por radio y por plataformas de streaming a nivel global.

Además, el conjunto de datos incluye información técnica de las canciones presentes en el ranking, como la valencia, el tempo, el modo y la acusticidad, entre otros.

A partir de la información histórica, se identificarán los artistas con mayor trayectoria y las canciones más exitosas. Por otro lado, en base a los datos técnicos de las canciones del ranking, se podrá establecer si existen patrones musicales que hacen que una canción sea más escuchada que otra.

# HIPÓTESIS

La hipótesis fundamental del análisis es que existen relaciones entre las posiciones del ranking y los atributos musicales de las canciones. Se espera que las canciones que más veces alcanzaron el puesto número 1 tengan atributos musicales similares.

# ANÁLISIS EXPLORATORIO

Para probar la hipótesis se plantean dos preguntas fundamentales...

1. ¿Las canciones más exitosas, es decir, aquellas que ocuparon mejores posiciones durante más semanas en el ranking, tienen características en común? De ser así, ¿esas características fueron cambiando con el paso de los años?
2. ¿Existe un género musical predominante dentro del top 10 del ranking? De ser así, ¿existe alguna relación entre el género musical y las características de las canciones?

# ANÁLISIS EXPLORATORIO

Los atributos musicales usados durante el análisis fueron los siguientes:

Sonoridad

Valencia

Tempo

Modo

Energía

Duración  
(ms)

Contenido vocal

Acusticidad

Bailabilidad

Clave

(Los significados y los valores que toman las variables se encuentran en la documentación)

# ANÁLISIS EXPLORATORIO

Pero ... ¿cuáles son las canciones “más exitosas”? Se establecieron tres criterios no mutuamente excluyentes

CRITERIO 1: Canciones con más semanas en el ranking

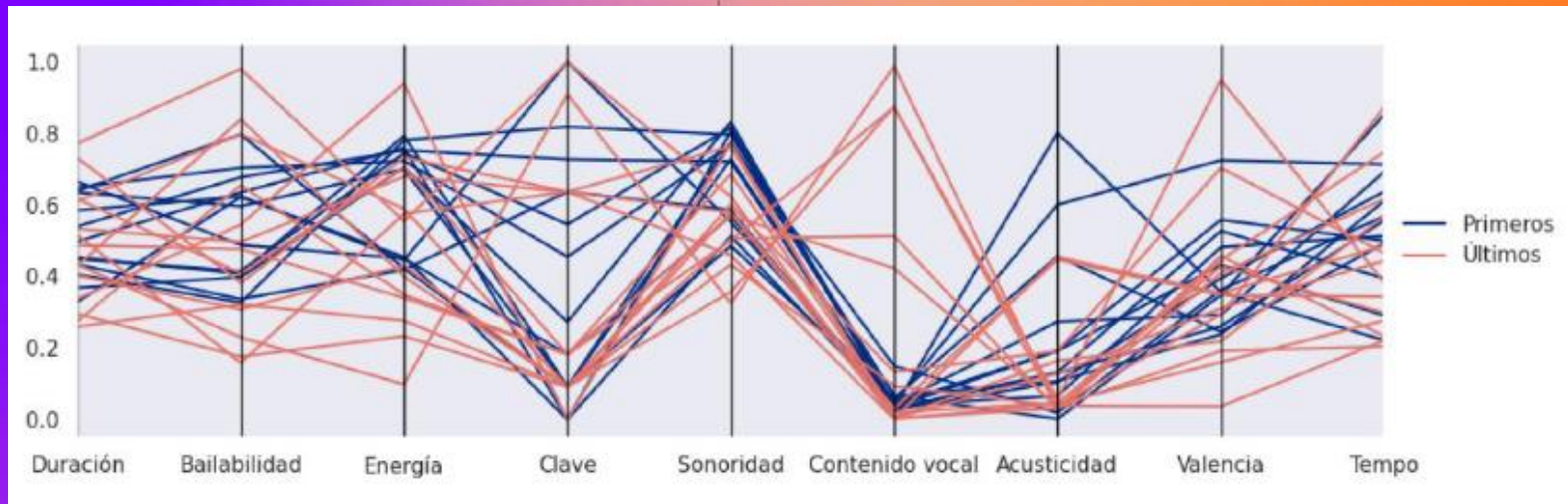
CRITERIO 2: Canciones con más semanas en el top 10

CRITERIO 3: Canciones con más semanas en el puesto Nº1



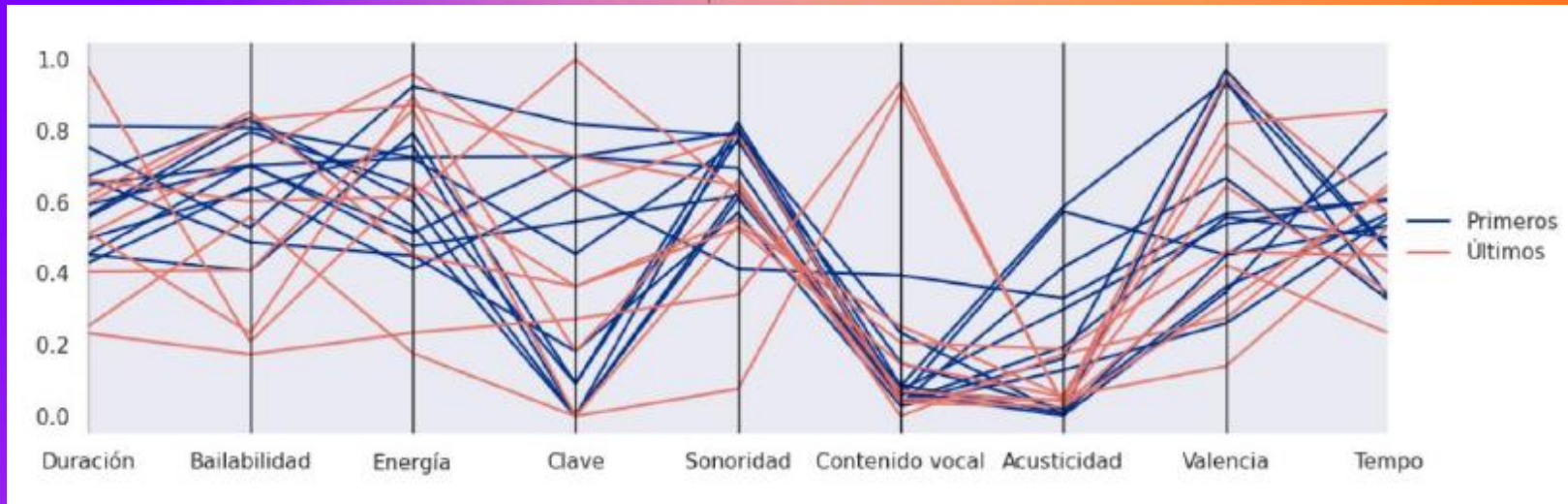
# INSIGHTS

## Criterio 1



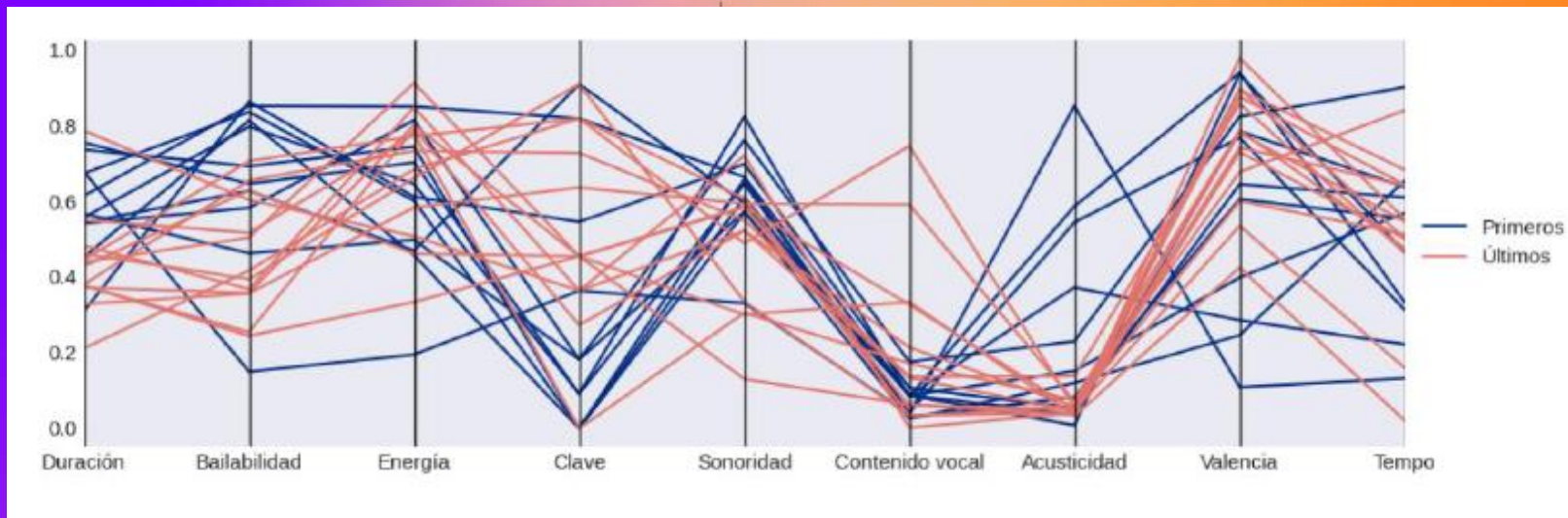
# INSIGHTS

## Criterio 2



# INSIGHTS

## Criterio 3



En base al análisis realizado **NO** puede afirmarse que las canciones exitosas tengan características en común

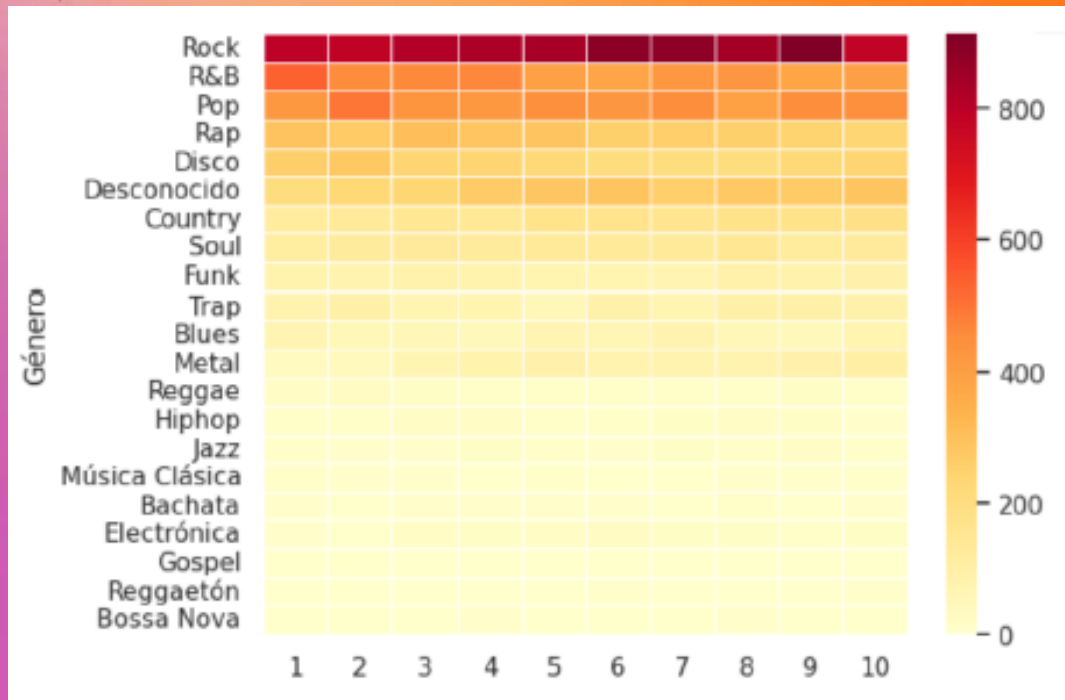
# INSIGHTS

En base al análisis realizado **NO**  
puede afirmarse que las canciones  
exitosas tengan características en  
común

# INSIGHTS

Existe un género predominante en el ranking: **el rock.**

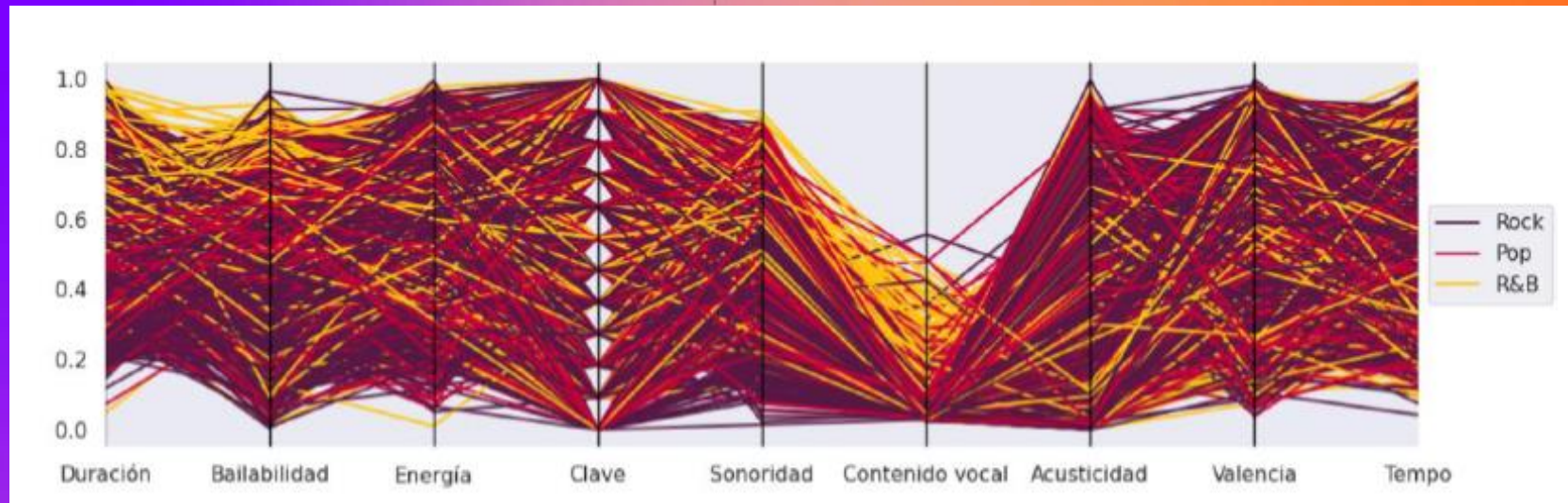
## Géneros predominantes en el top 10.





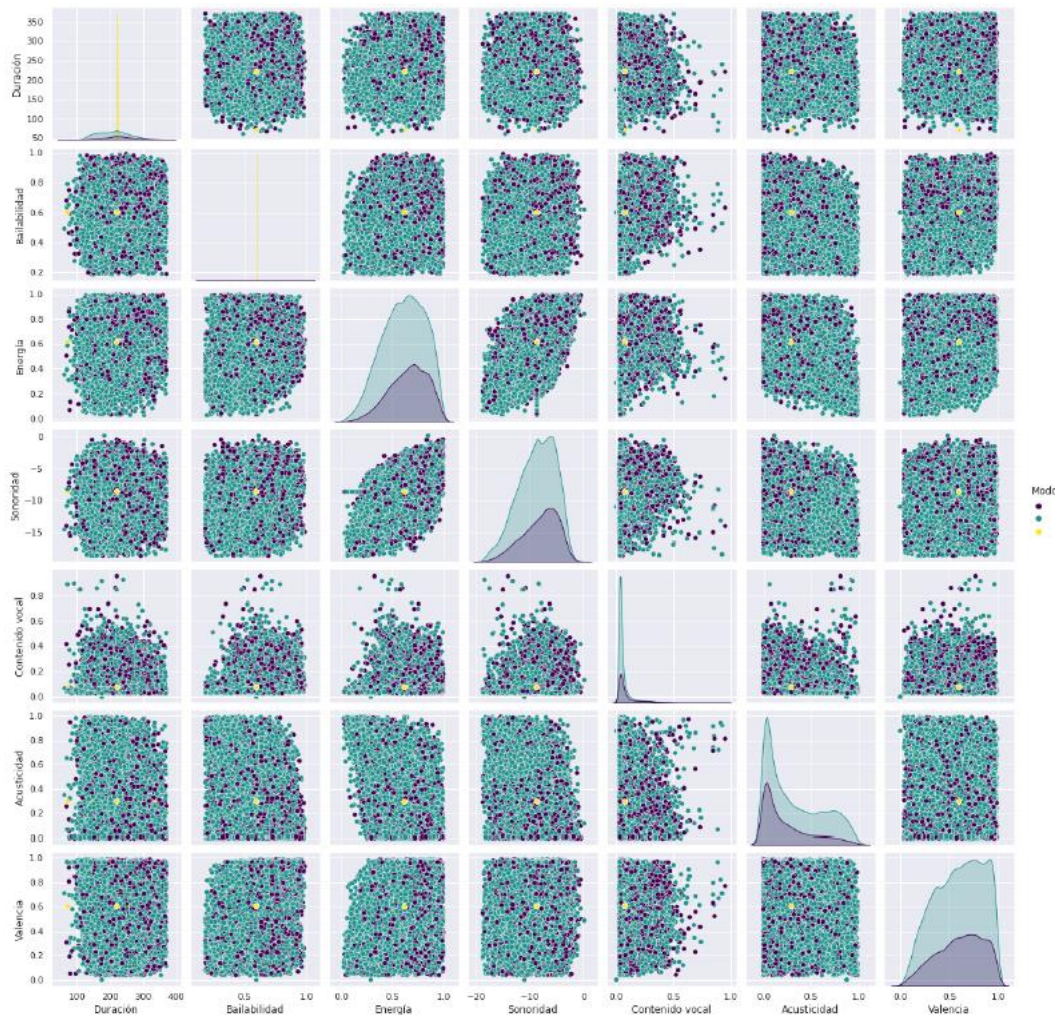
# INSIGHTS

Sin embargo, no se puede afirmar que existan relaciones entre los atributos musicales y el género



# INSIGHTS

Además, se estudió  
la correlación entre  
las variables, pero  
no se encontró  
ninguna.



# RECOMENDACIONES

En base al estudio realizado, *no se puede recomendar* la creación de un modelo de machine learning para predecir el éxito de las canciones