



LA CALIFICACIÓN CRÍTICA DE DISCOS DE DEATH METAL



@chainreaction_84

Jorge Luis Sánchez Ruiz

CONTENIDO

- 01** INTRODUCCIÓN Y OBJETIVOS DEL PROYECTO
- 02** FUENTE DE DATOS|DESCRIPCIÓN DE LOS DATOS
- 03** LIMPIEZA DE DATOS Y EDA| DATA WRANGLING
- 04** MODELOS DE PREDICCIÓN
- 05** DISCUSIÓN Y CONCLUSIONES PRELIMINARES
- 06** DESAFÍOS, LIMITACIONES Y PRÓXIMOS PASOS
- 07** FUTURAS ITERACIONES

INTRODUCCIÓN: EL DEATH METAL Y SU EVOLUCIÓN SONORA

- Originado en los 80s, el death metal es conocido por su intensidad y complejidad técnica.
- Con el tiempo, algunas bandas experimentaron con transiciones hacia otros géneros.
- Estas transiciones son a veces aclamadas y otras veces recibidas con escepticismo.
- Nuestro proyecto busca describir y predecir cómo estos cambios afectan la percepción crítica de los discos en el death metal.

OBJETIVOS

- Analizar el efecto de las transiciones de género en el score crítico.
- Estudiar las relaciones entre el score crítico, la producción musical total de las bandas y el número de reviews recibidas.
- Predecir el efecto de posibles transiciones de género en el score de la crítica.
- Predecir la recepción crítica en términos de la producción musical.

LIMPIEZA DE DATOS Y EDA| DATA WRANGLING

PROCESO DE ANÁLISIS DE DATOS

- Fuentes: Dos datasets de Kaggle.
- Tratamiento de Datos: Uso de expresiones regulares para identificar transiciones de género.
- Variables Derivadas:
- Booleano para clasificar si una banda realizó transiciones.
- Conteo de discos por banda.
- Score medio crítico.
- Conteo de reviews.
- Integración: Añadimos registros adicionales con un dataset complementario.
- Visualización: Raincloud plots para distribuciones de puntuaciones medias.

MODELOS DE PREDICCIÓN

- Exploración de Datos Atípicos: Transformación con Yeo-Johnson debido a colas a la izquierda en distribuciones.
- Intento de Ajuste a Ley de Potencias: Resultó infructuoso debido a baja correlación lineal.
- Evaluación de Características: Cálculo de la Información Mutua entre features del dataset.

DISCUSIÓN Y CONCLUSIONES PRELIMINARES

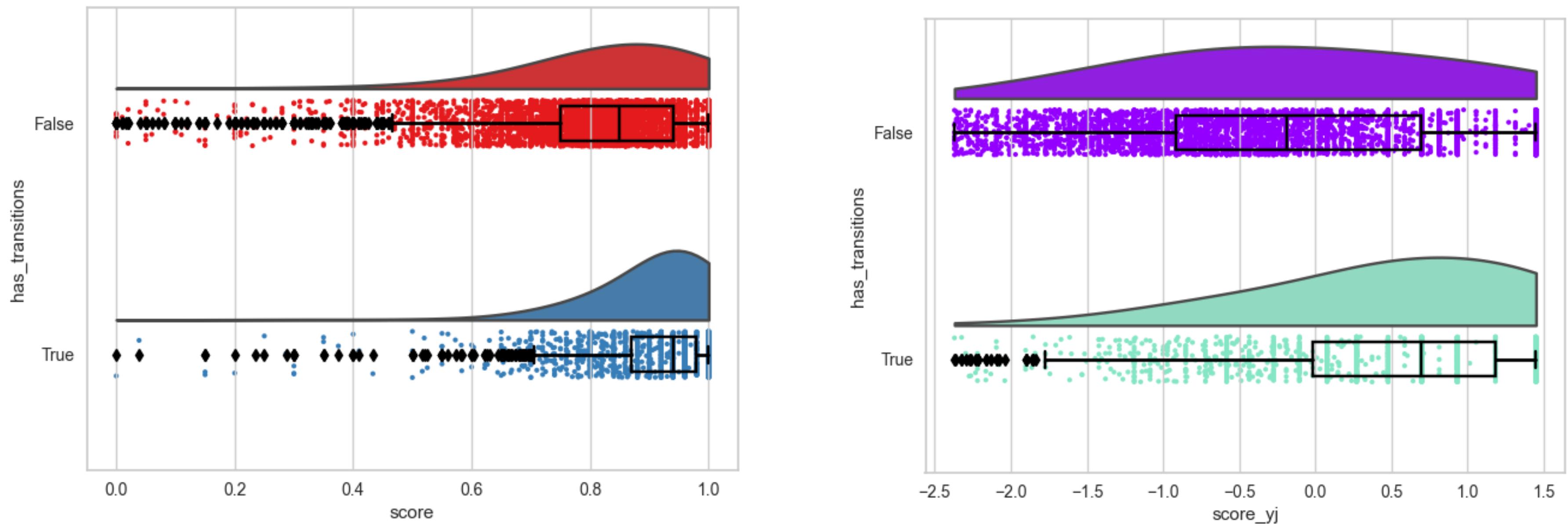


Fig 1.) Raincloudplot distribución original (izquierda) y distribución transformada por Yeo-Johnson (derecha)

DISCUSIÓN Y CONCLUSIONES PRELIMINARES

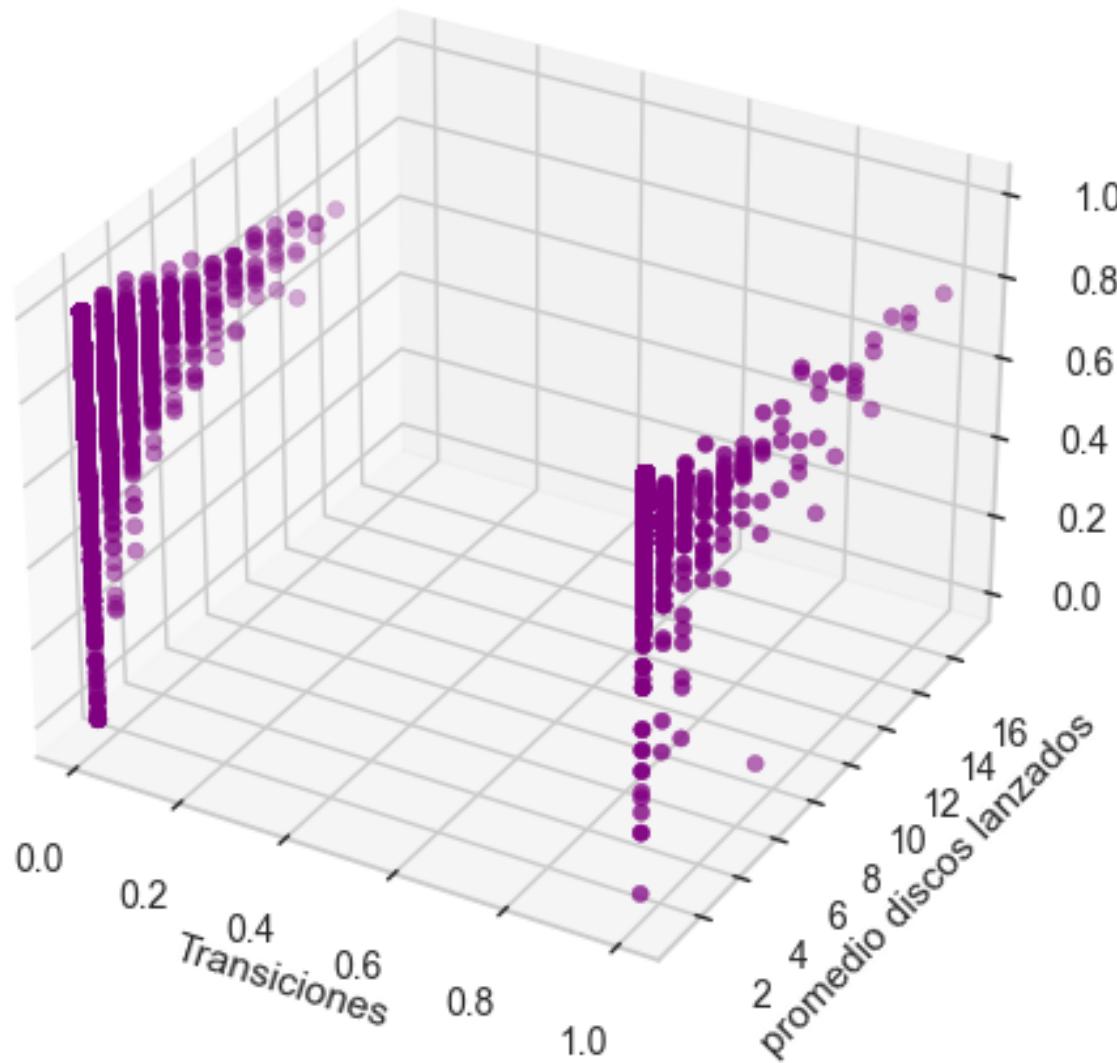
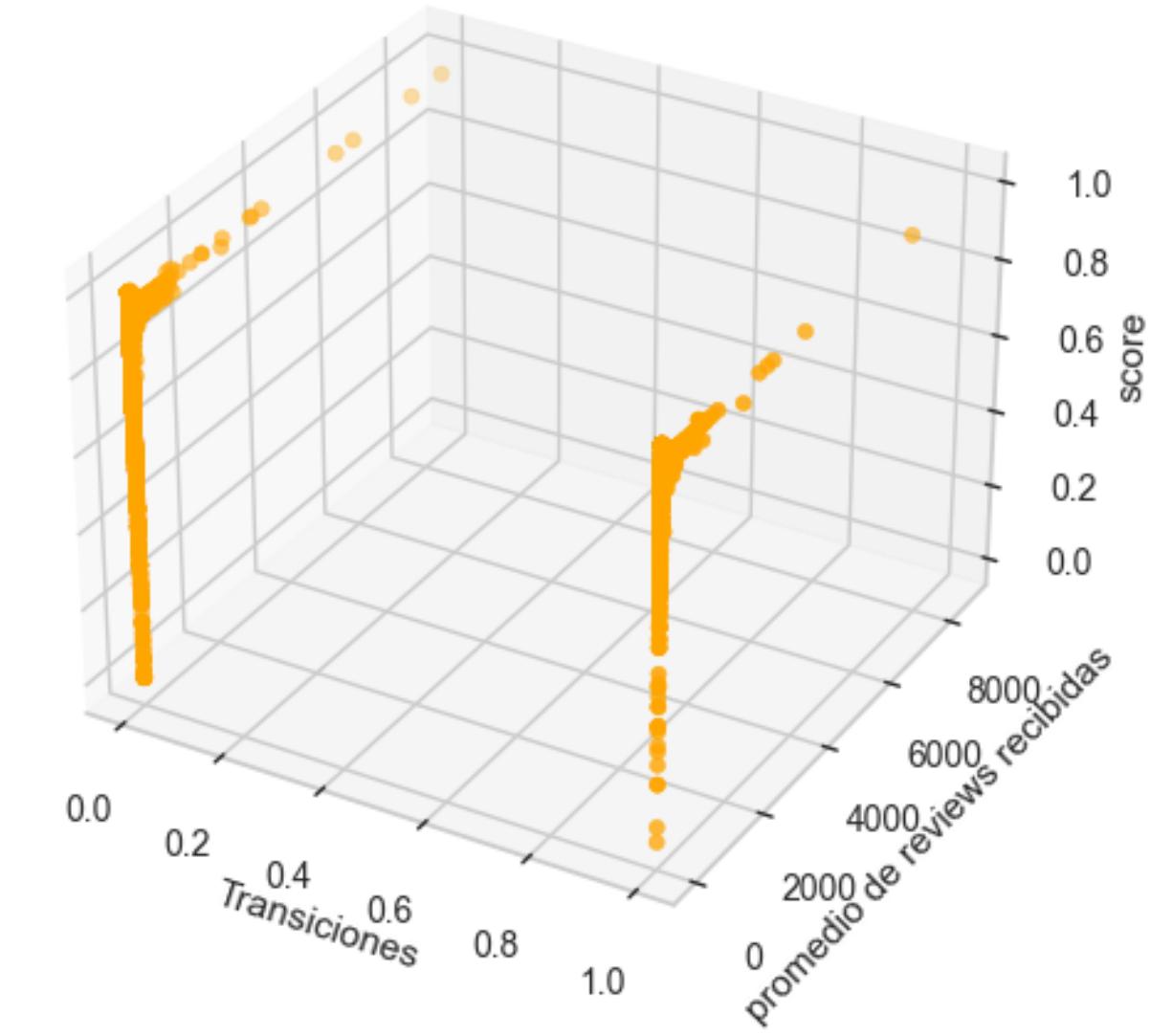


Fig 2.) Score como función del promedio de discos para los escenarios con transiciones y sin transiciones de género

Fig 3.) Score como función del promedio calificaciones para los escenarios con transiciones y sin transiciones de género



DISCUSIÓN Y CONCLUSIONES PRELIMINARES

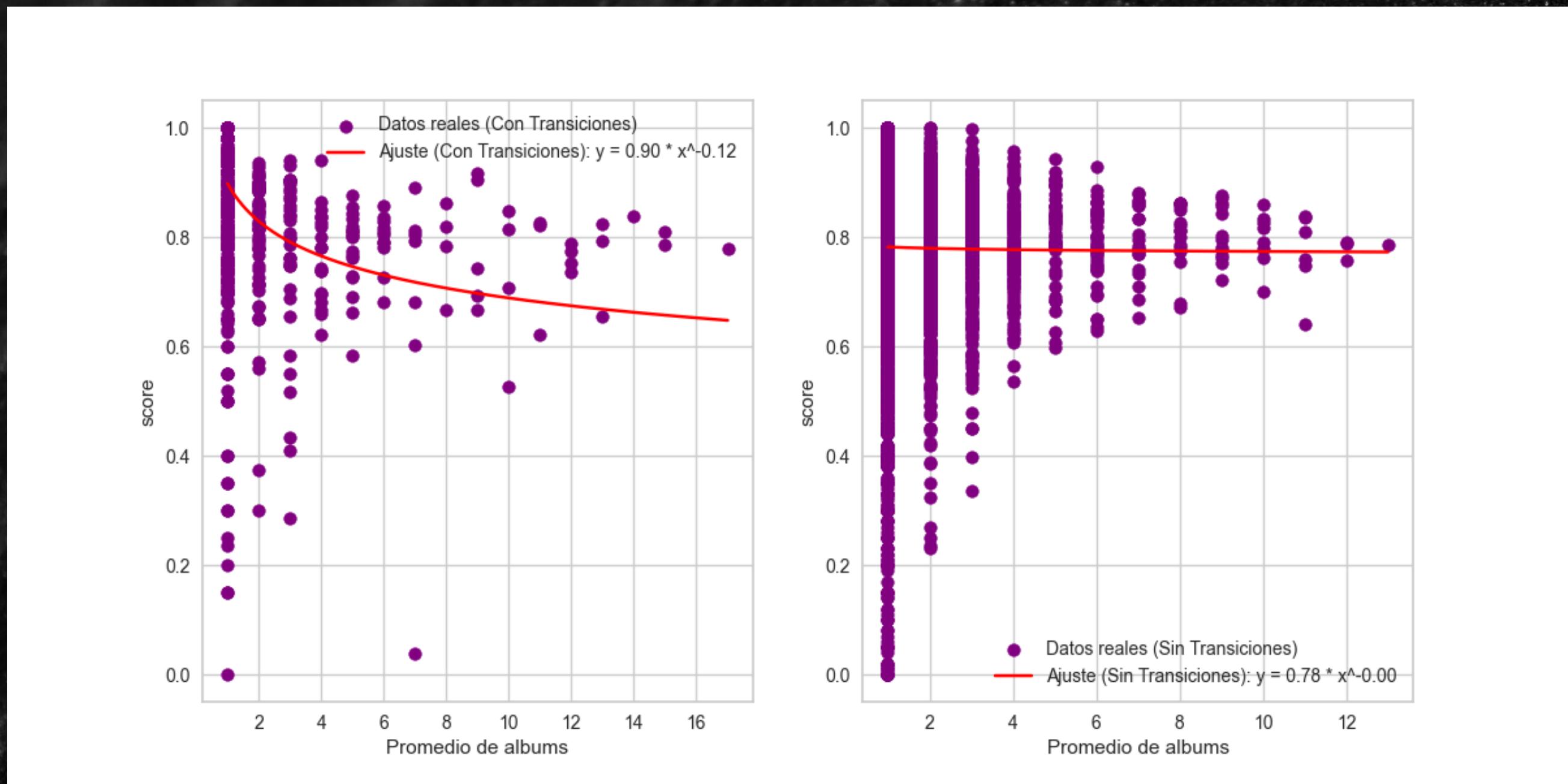
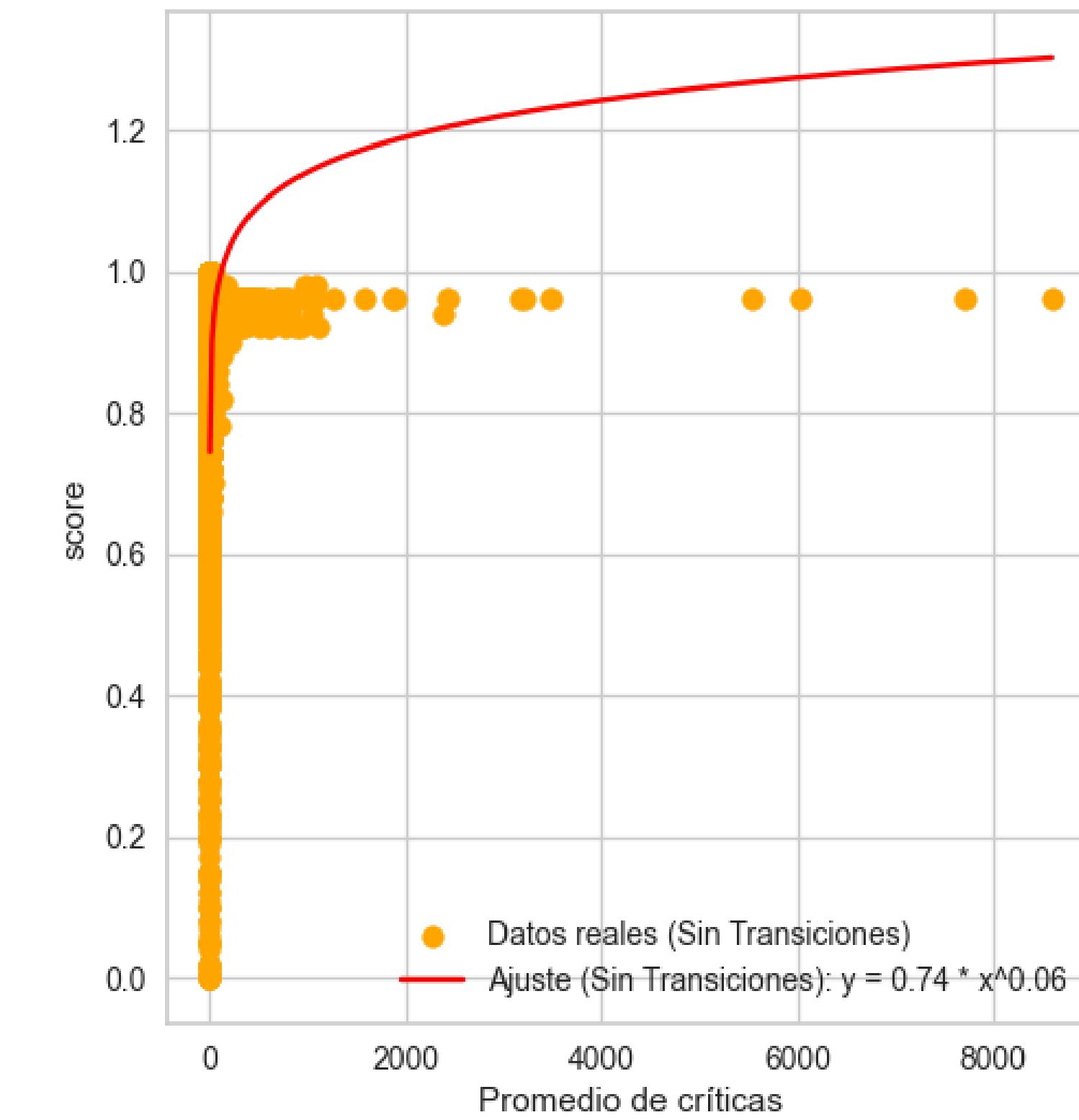
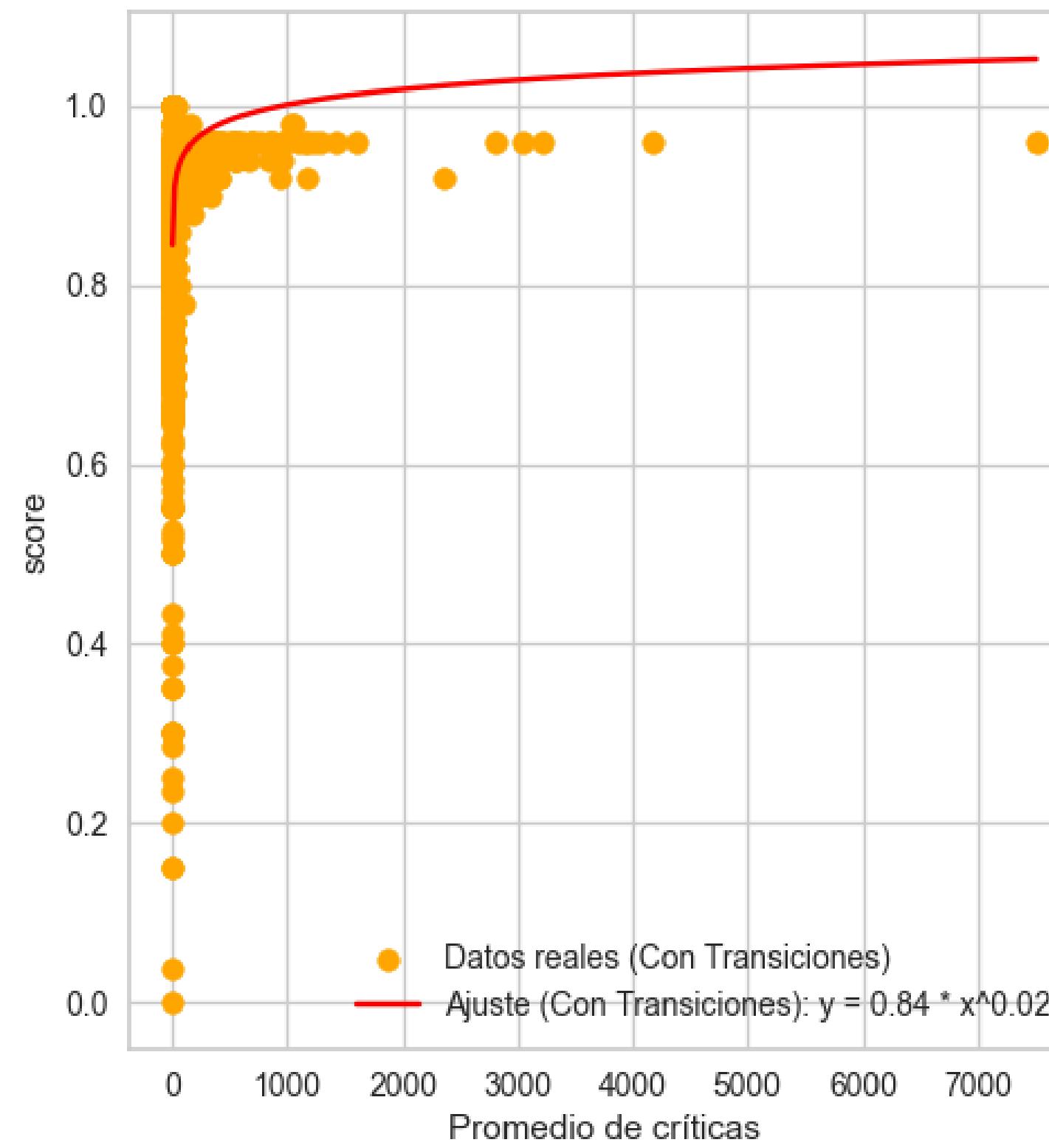


Fig 5.) Ajuste como ley de potencias para el score en función del promedio de discos lanzados el coeficiente de correlación lineal en la representación lineal es de 0.07 lo que pone en evidencia que no es la tendencia de los datos

DISCUSIÓN Y CONCLUSIONES PRELIMINARES



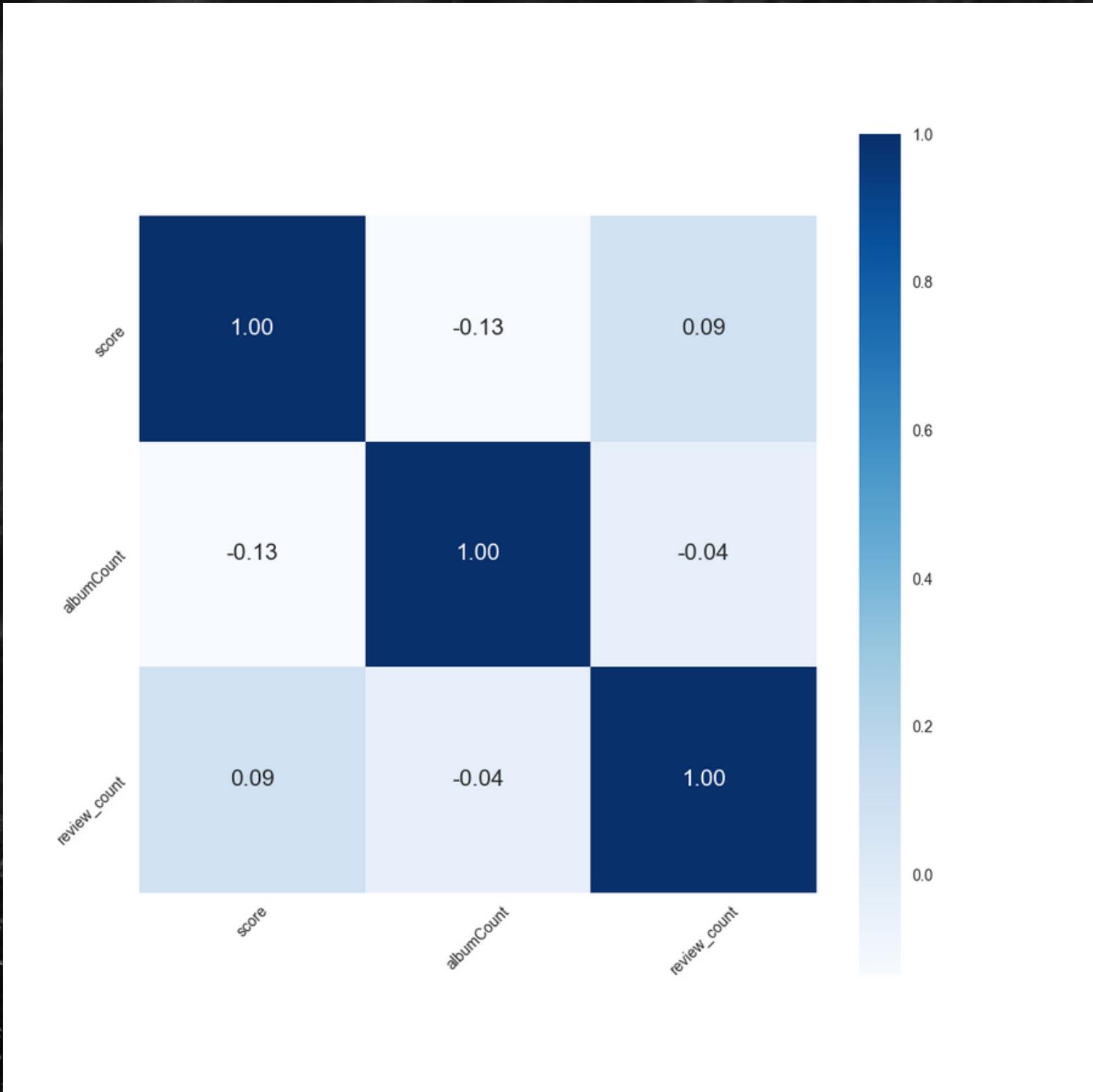
DISCUSIÓN Y CONCLUSIONES PRELIMINARES

ANÁLISIS DE ALBUMCOUNT VS SCORE

- AGRUPACIÓN VERTICAL: VARIABILIDAD DE SCORES EN CUENTAS DE ÁLBUMES BAJAS (1-5).
- VARIABILIDAD DISMINUIDA: MENOS DISPERSIÓN EN SCORES PARA CUENTAS > 5 .
- TENDENCIA LEVE: DESCENSO EN SCORES PARA CUENTAS ENTRE 6 Y 16.
- VALORES ATÍPICOS: NOTABLES EN CUENTAS DE 2-5.
- DENSIDAD: MUCHOS PUNTOS CERCANOS A UN SCORE DE 1.0.

CONCLUSIÓN: TENDENCIAS CLARAS PERO CON VARIABILIDAD SIGNIFICATIVA; BASE PARA FUTUROS ANÁLISIS.

DISCUSIÓN Y CONCLUSIONES PRELIMINARES



- El coeficiente de Pearson muestra ausencia de tendencias lineales entre los features.
- El cálculo de la información mutua entre score y número de discos ($I=0.2$) muestra que si existe dependencia.
- El cálculo de la información mutua entre score y número reviews ($I=0.03$) muestra una dependencia más leve entre estos features.

DESAFÍOS, LIMITACIONES Y PRÓXIMOS PASOS| FUTURAS ITERACIONES

- En las cuentas de álbumes más bajas (aproximadamente de 1 a 5), hay una agrupación vertical de datos que indica una variedad de "scores" para una cantidad específica de álbumes. Esto sugiere que, para las bandas o artistas con pocos álbumes, el número de álbumes lanzados no es un buen predictor del "score".
- Para estudiar las relaciones complejas entre estos features es pertinente explorar otros métodos como modelos basados en árboles o redes neuronales