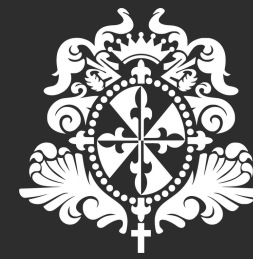


¿Qué hace exitosa a una película?

Un análisis con datos de IMDb

Aldana Gabriela, Arce Valeria, Torres Juan Esteban



Universidad del
Rosario

370 Años
Formando proyectos de vida
al servicio del país

Escuela de ingeniería,
Ciencia y Tecnología

Objetivo

- Identificar los factores clave que influyen en la popularidad y éxito financiero de las películas en IMDb, analizando variables como género, idioma, clasificación por edad, presupuesto, recaudación y calificaciones.

Relevancia

- Comprender qué elementos contribuyen al éxito de una película permite a productores, distribuidores y plataformas tomar mejores decisiones estratégicas. Además, ofrece una aplicación real del análisis estadístico a un tema de interés general.

Variables consideradas:

Categorías: Género, idioma, clasificación adulta.

Númericas: Presupuesto, recaudación, calificación promedio, popularidad.



Metodología

Base de datos: Full IMDb Movies Data (~1 millón de películas).

Análisis descriptivo: Univariado y bivariado.

Pruebas estadísticas:

- Anova
- t-test (una y dos muestras)
- Chi-cuadrado



Análisis de normalidad (Kolmogorov-Smirnov, skewness)

Transformaciones logarítmicas para reducir asimetría

Hallazgos descriptivos

Géneros más comunes: Documentary, Drama, Comedy

Idioma dominante: Inglés (82 %)

Contenido adulto: Solo 9.25 % del total

Calificaciones promedio: entre 5 y 7 (no normal)

Correlaciones importantes:

- Recaudación vs Presupuesto: 0.62
- Recaudación vs Popularidad: 0.67

Pruebas de independencia

Chi-cuadrado entre variables categoricas:

- Género vs Idioma:
- Rechazo de H ($p < 0.001$)
- Ej: Drama sobre-representado en francés/español

Horror y Thriller en inglés

Género vs Contenido adulto:

- Family y Thriller → no adultos
- Crimen, Fantasy → adultos
- Otros géneros → fuertemente adultos

Normalidad y transformaciones

- Revenue:** skew = 8.38 → $\log(\text{skew}) = 0.03$
- Budget:** skew = 4.67 → $\log(\text{skew}) = -0.61$
- Las transformaciones mejoran la simetría pero no logran normalidad perfecta (bimodalidad presente)

Análisis inferencial

Calificaciones promedio:

- $H_0: \mu = 7$ vs $H_1: \mu < 7$
- Se rechaza H_0 ($p < 0.001$)

Ingresos y clasificación adulta:

- Películas no adultas tienen mayores ingresos ($p < 0.001$)
- est t de Welch por desigualdad de varianzas

ANOVA — Popularidad por género:

- Géneros significativos: Drama, Comedy, Action, Horror
- Thriller no significativo ($p = 0.356$)

ANOVA — Presupuesto por idioma:

- Idiomas significativos: inglés, español, francés
- Japonés y ruso no significativos

Bondad de ajuste

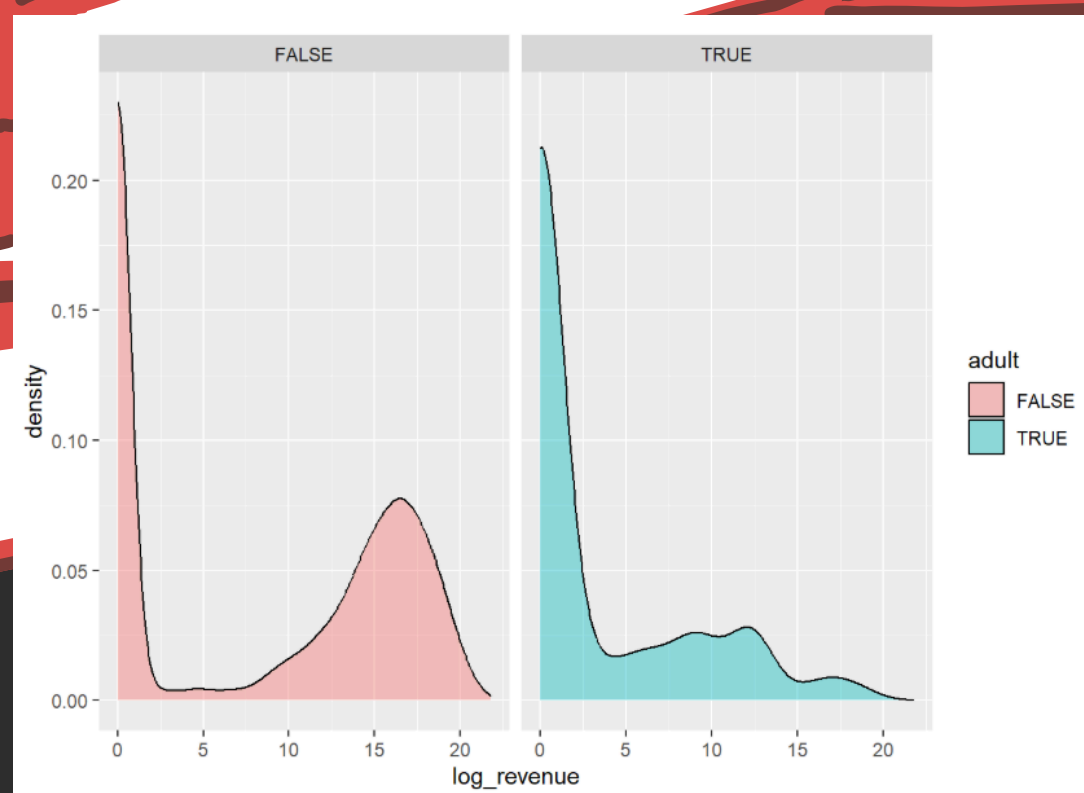
Idiomas: No distribución uniforme ($p < 0.001$)

Adultos vs No adultos: desequilibrado (9.25 % vs 90.75 %)

Número de géneros (Poisson): No se ajusta, $p < 0.001$

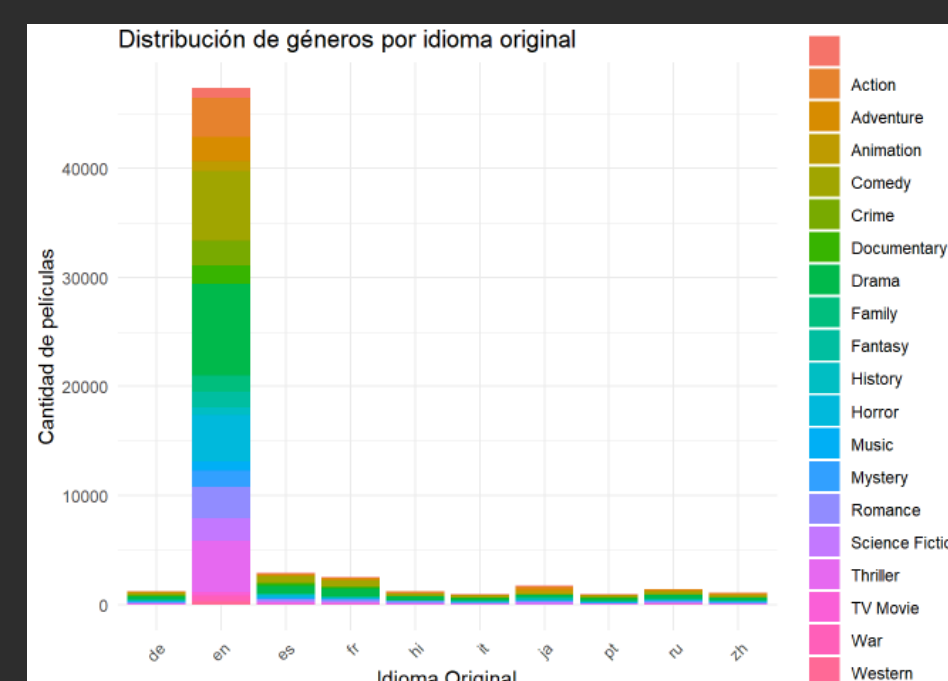
Visualización y Análisis de resultados

Distribución de ingresos por clasificación



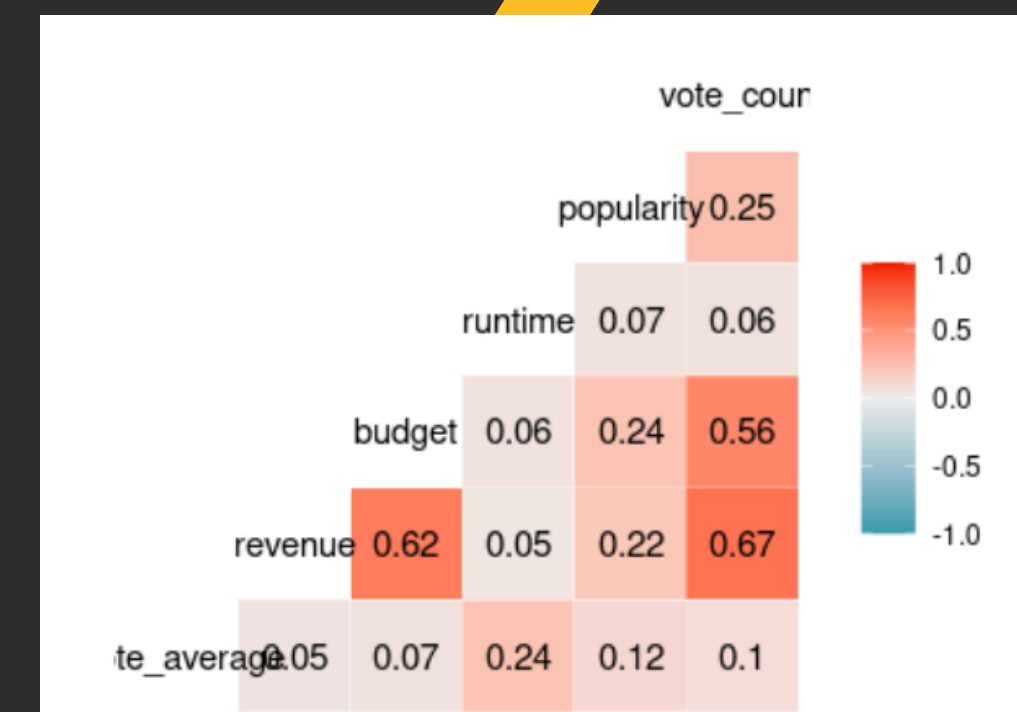
- Las películas no adultas tienen mayor variabilidad en ingresos

Distribucion de generos por idioma



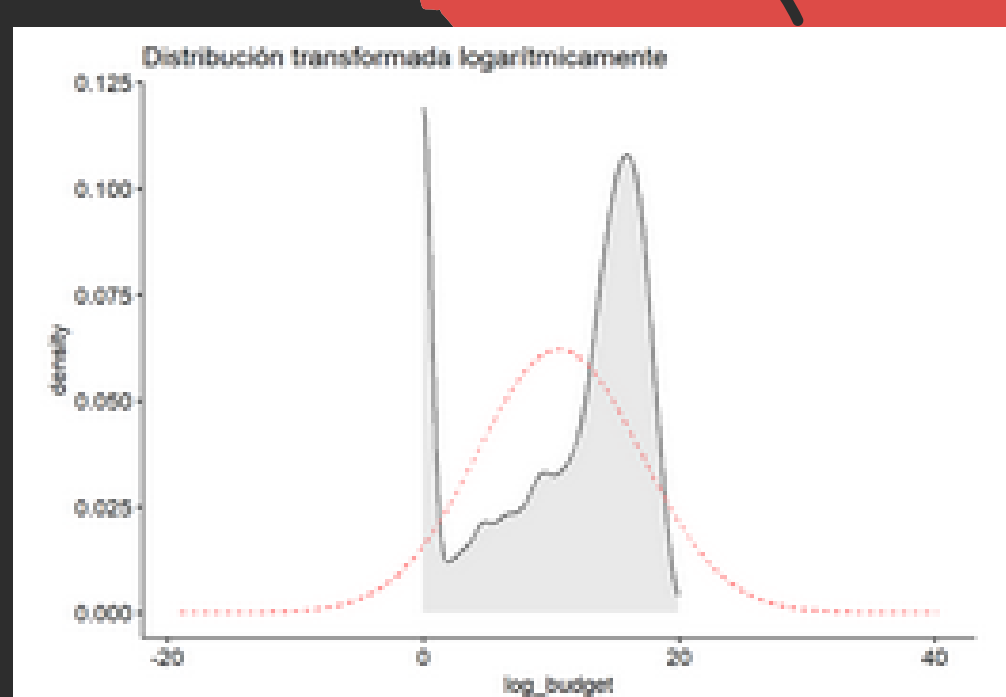
- El inglés domina en cantidad y variedad de géneros.

Matriz de correlación



- La recaudación se relaciona con el presupuesto y la popularidad.

Distribución logarítmica del presupuesto



- Los presupuestos muestran una distribución desigual y no normal

Distribución promedio por idioma



- El idioma inglés concentra los ingresos más altos en promedio.

Conclusiones

- Género, idioma y clasificación adulta si influyen en métricas de éxito
- Asociación fuerte entre género, idioma y tipo de contenido
- Hallazgos clave:**
- Películas no adultas generan mayor recaudación
- Géneros como Action, Drama, Comedy y Horror influyen significativamente en la popularidad.
- Idiomas con mayor presupuesto: inglés, español y francés

Bibliografía

- <https://www.kaggle.com/datasets/anandshaw2001/imdb-data>
- <https://www.r-project.org/other-docs.html>

