

Informe de Análisis Exploratorio: Catálogo de Netflix

Departamento de Análisis de Datos

13 de febrero de 2025

1. Introducción

Este informe analiza el dataset del catálogo de Netflix utilizando técnicas estadísticas y visualizaciones. El conjunto de datos contiene información sobre películas y series, incluyendo metadatos como duración, clasificación, género y país de producción. El análisis se centra en tres aspectos clave: estadísticos descriptivos, relaciones entre variables y distribución geotemporal.

2. Metodología

Se utilizaron las siguientes herramientas:

- **Python** con librerías (pandas, matplotlib, scikit-learn)
- Técnicas de análisis: PCA, test de normalidad (Shapiro-Wilk)
- Métodos estadísticos: Medidas de tendencia central y dispersión

3. Resultados

3.1. Estadísticos Descriptivos

Estadística	Valor
Duración Media	99.528
Duración Mediana	98.000
Moda de la Duración	90.000
Varianza (min^2)	804.816
Desviación Estándar	28.369
Percentil 25	87.000
Percentil 50 (mediana)	98.000
Percentil 75	114.000

Cuadro 1: Duración de Películas

Estadística	Valor
Duración Media	1.765
Duración Mediana	1.000
Moda de la Duración	1.000
Varianza ($\#\text{temporadas}^2$)	2.505
Desviación Estándar	1.583
Percentil 25	1.000
Percentil 50 (mediana)	1.000
Percentil 75	2.000

Cuadro 2: Duración de Series de TV

Estadística	Valor
Mediana	2017
Moda	2018
Varianza	78
Desviación Estándar	9
Percentil 25	2013
Percentil 50 (mediana)	2017
Percentil 75	2019

Cuadro 3: Año de Estreno

Estadística	Valor
Mediana	2019
Moda	2019
Varianza	2
Desviación Estándar	2
Percentil 25	2018
Percentil 50 (mediana)	2019
Percentil 75	2020

Cuadro 4: Año de Adición a Netflix

3.2. Gráficos Relevantes

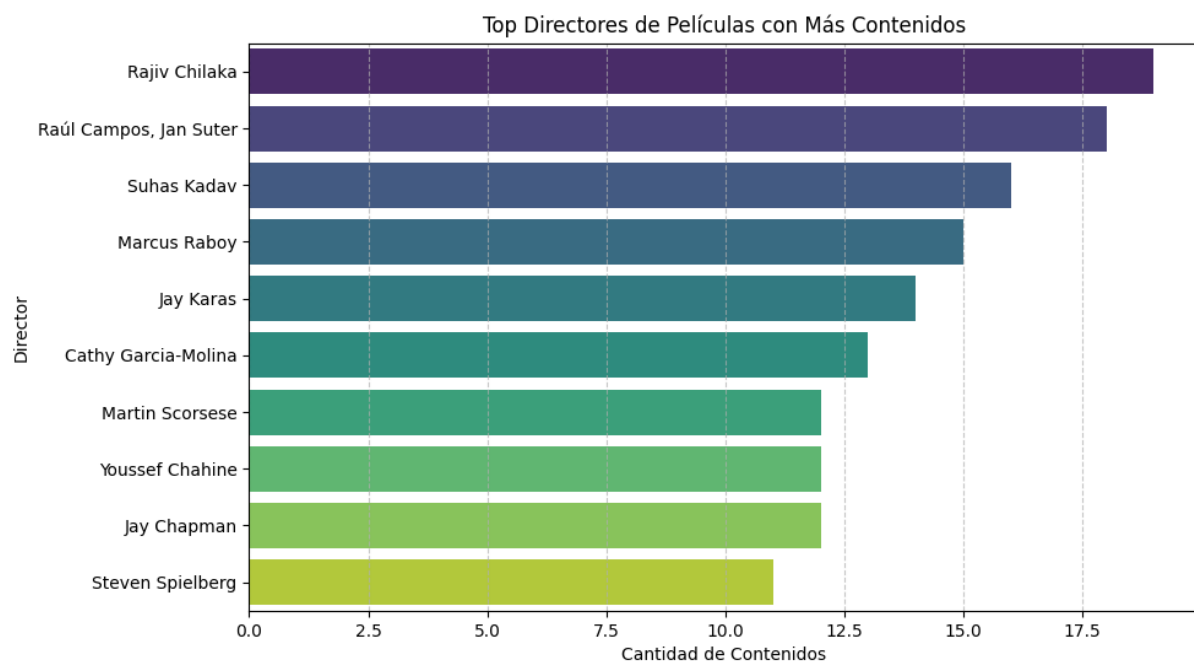


Figura 1: Directores de Películas Más Comunes

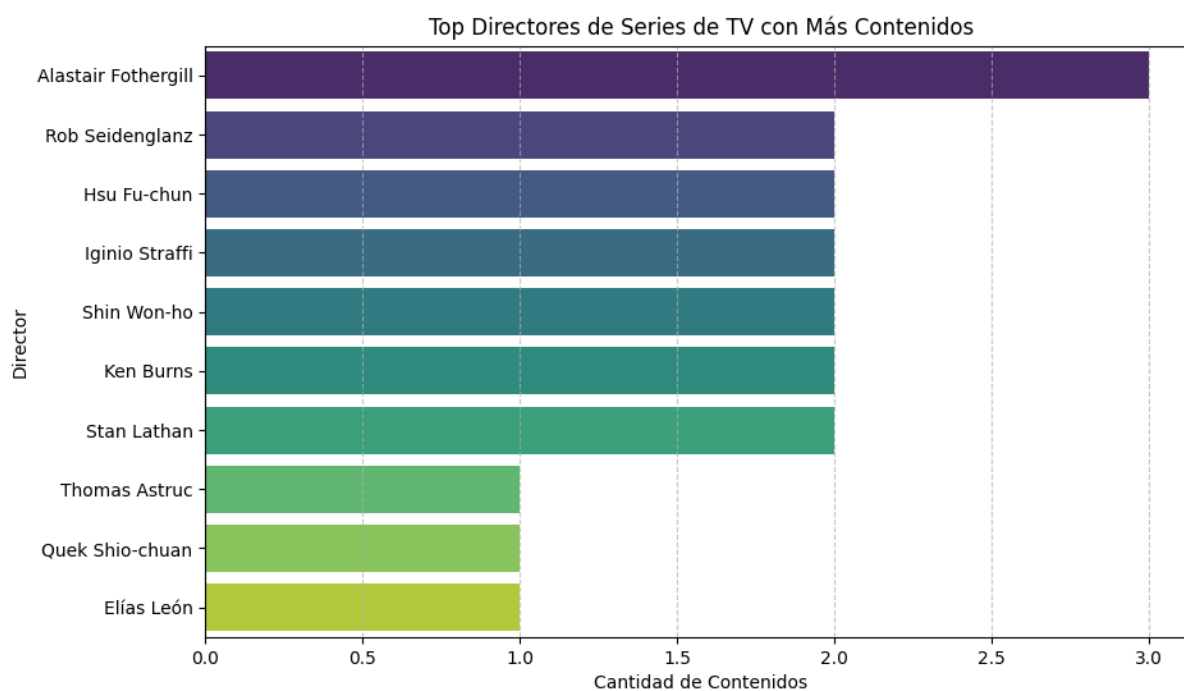


Figura 2: Directores de Series Más Comunes

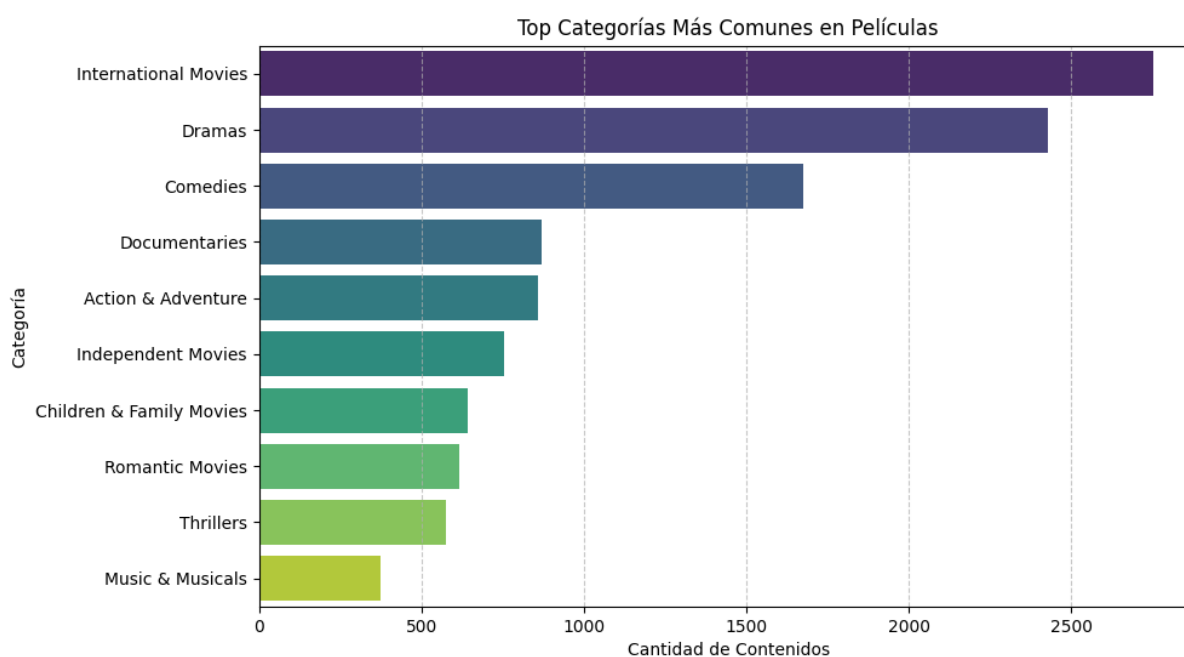


Figura 3: Categorías Más Comunes en Películas

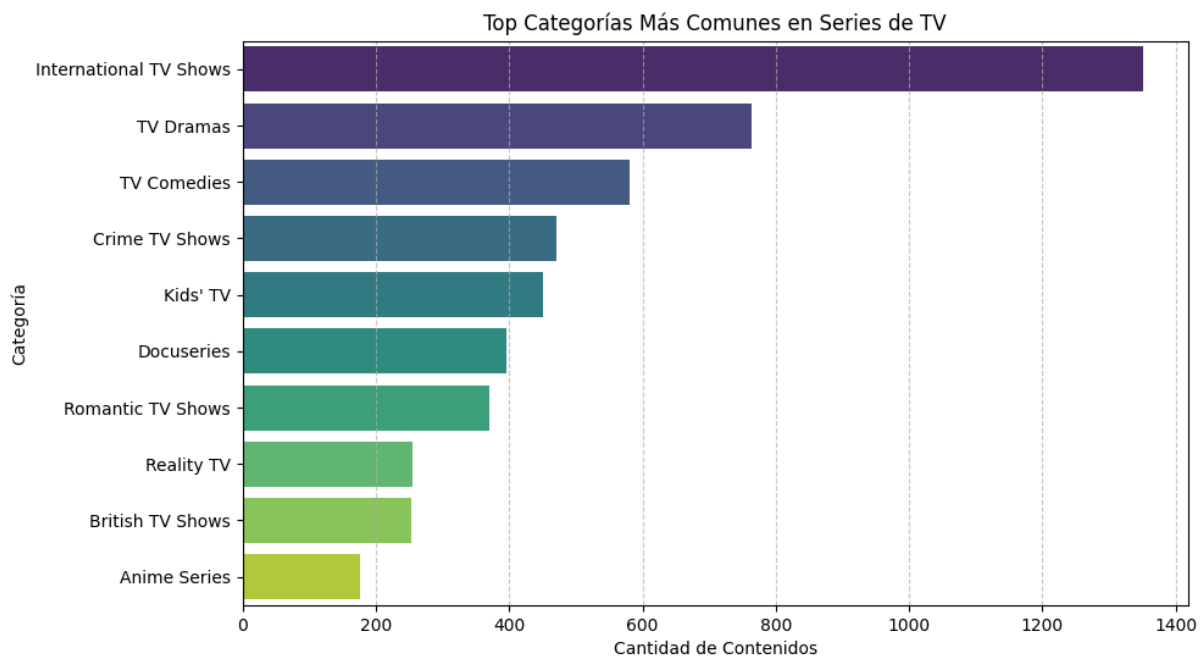


Figura 4: Categorías Más Comunes en Series de TV

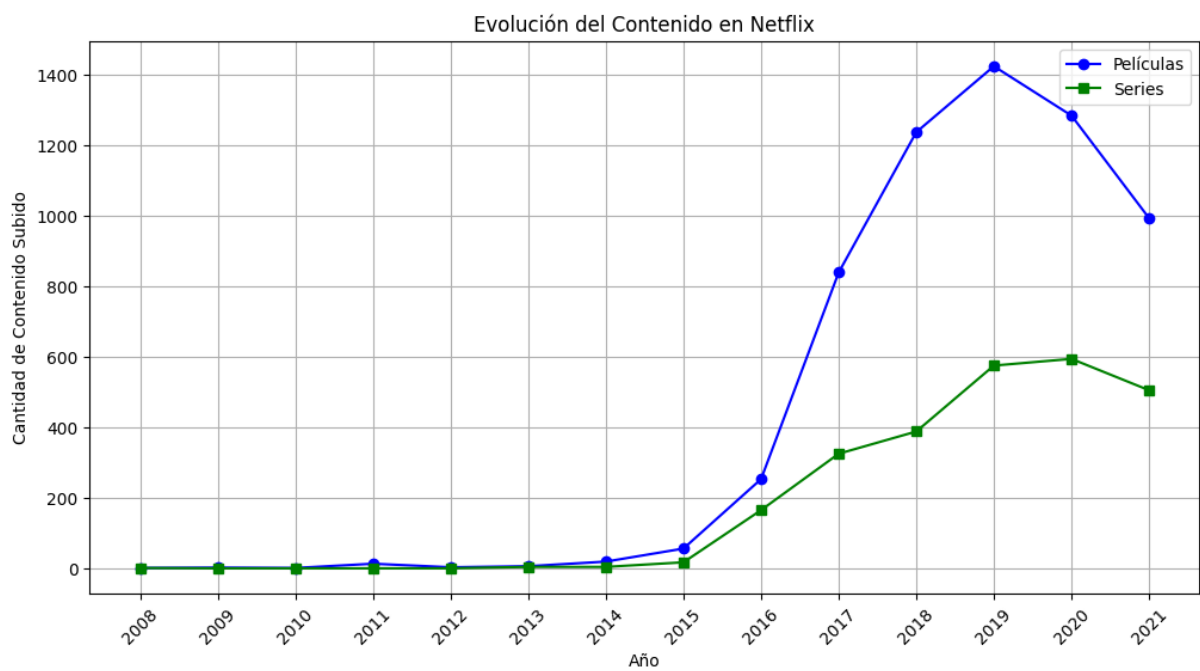


Figura 5: Evolución del Contenido

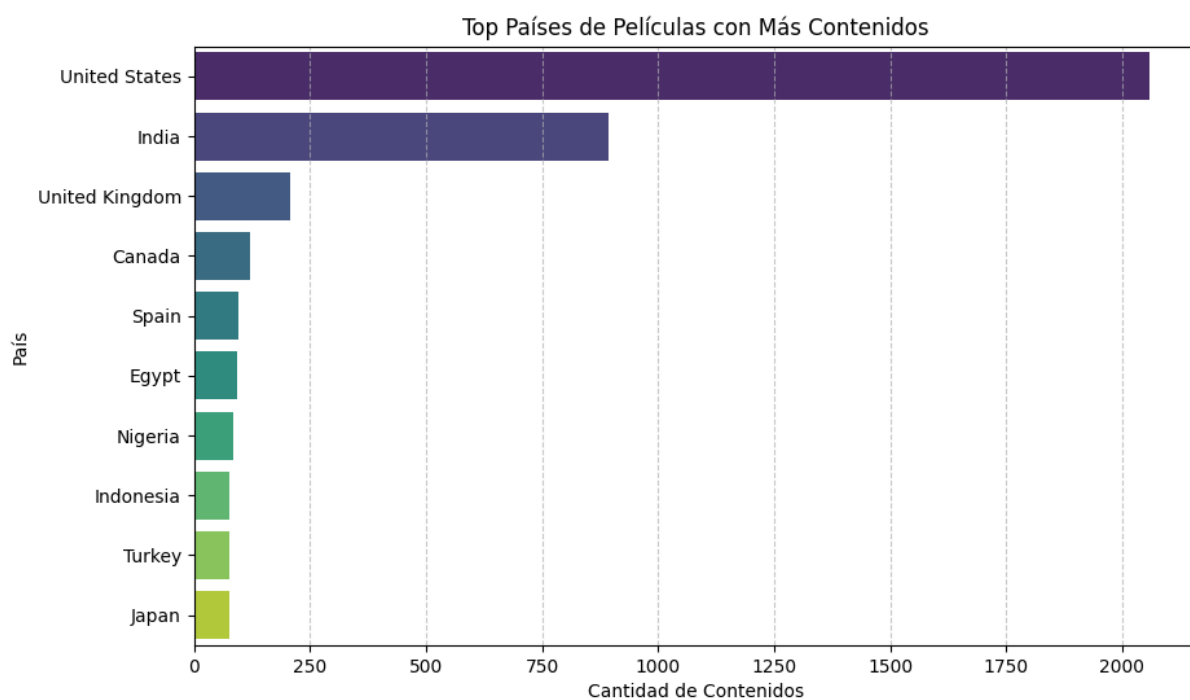


Figura 6: Países con más películas

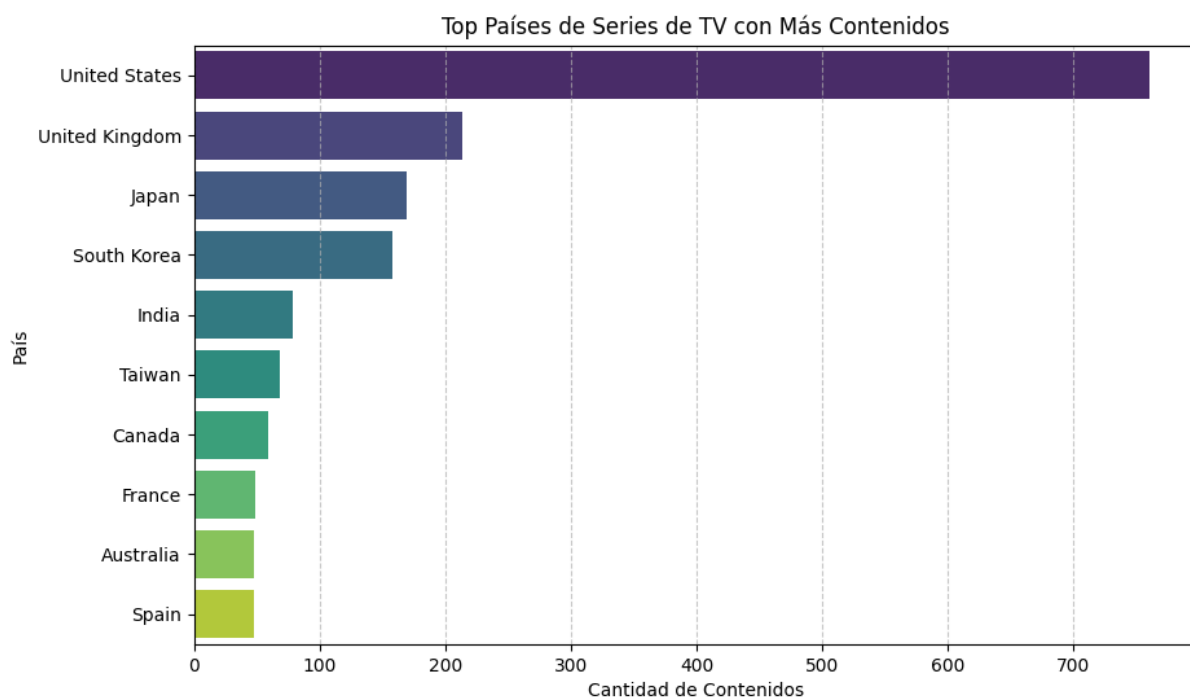


Figura 7: Países con más series

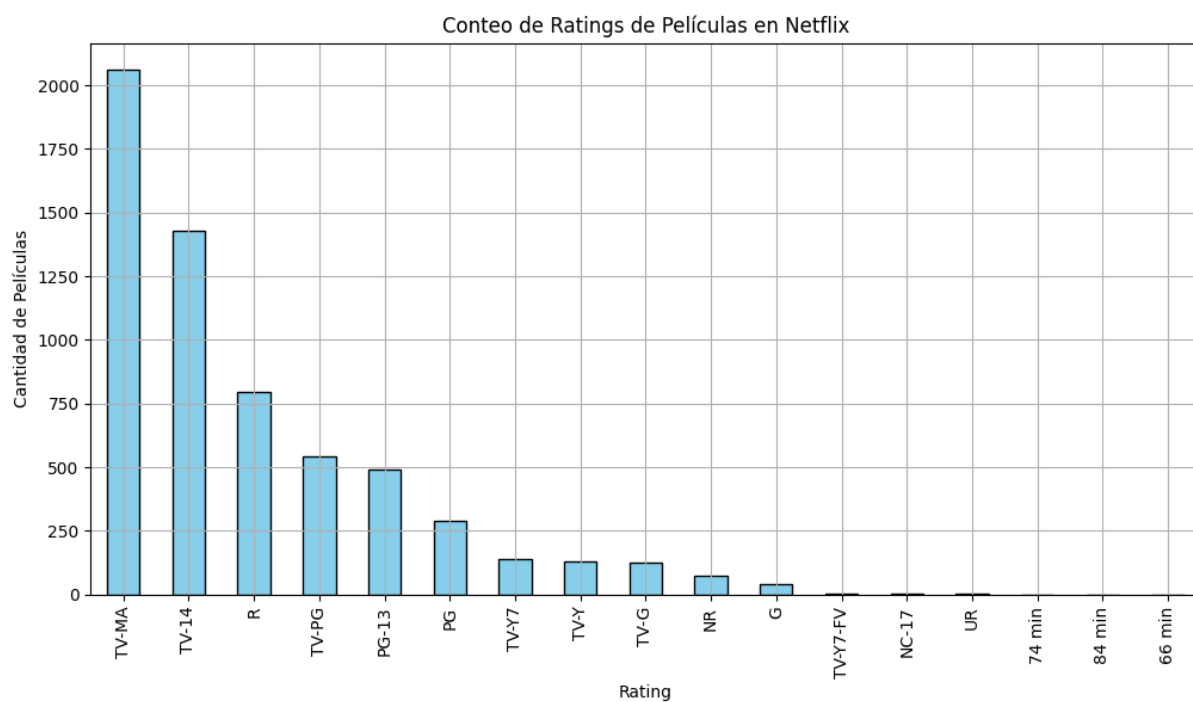


Figura 8: Conteo de Ratings para Películas

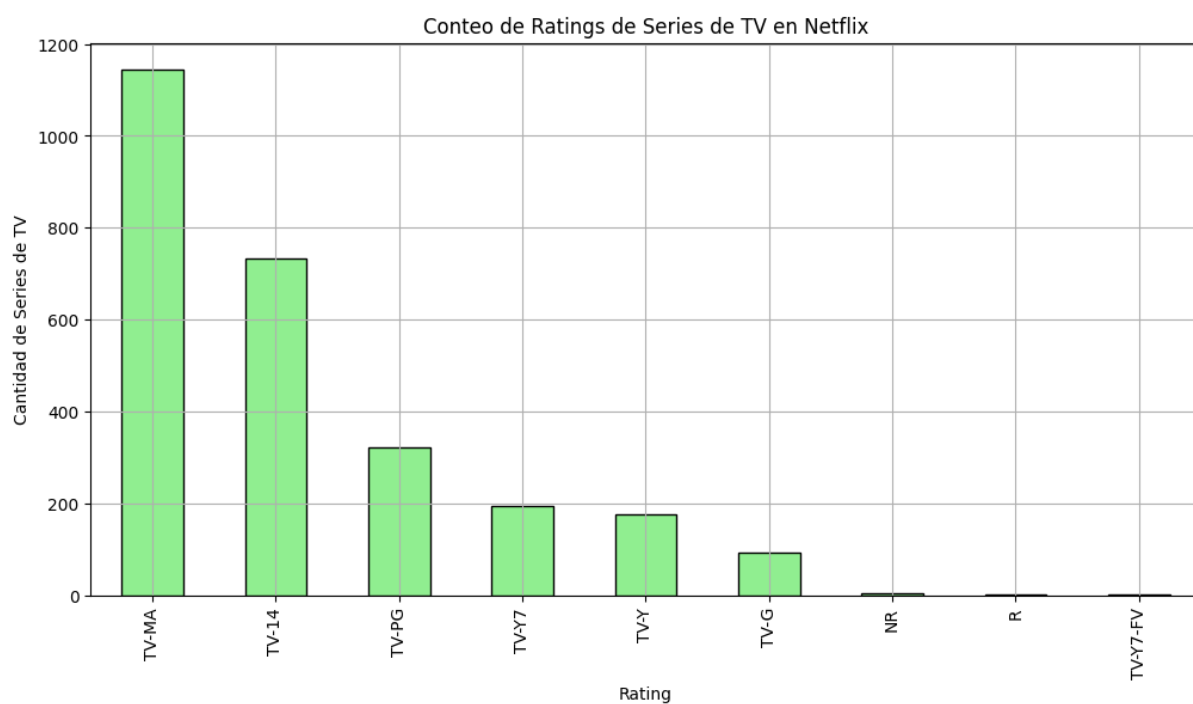


Figura 9: Conteo de Ratings para Series de TV

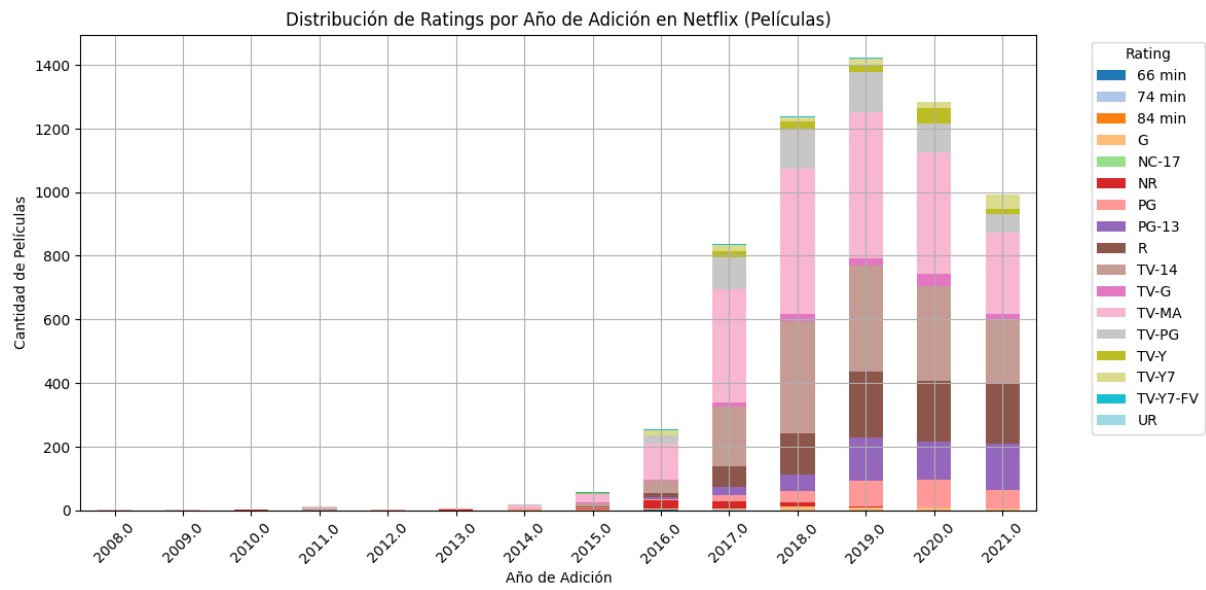


Figura 10: Distribución de Ratings por Año de Adición en Películas

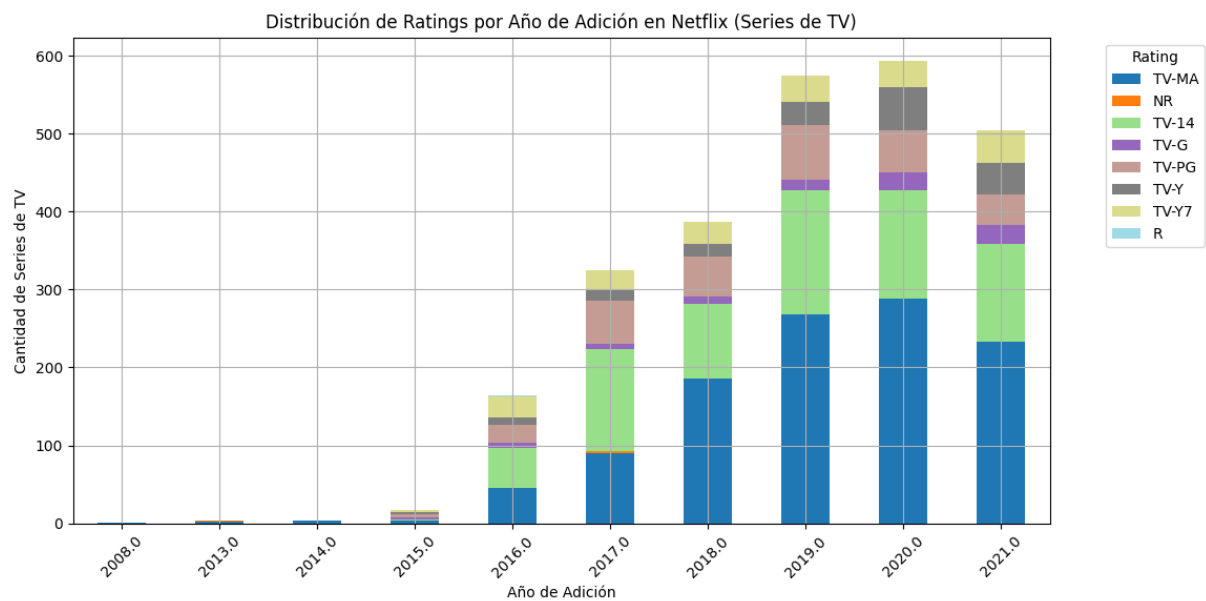


Figura 11: Distribución de Ratings por Año de Adición en Series de TV

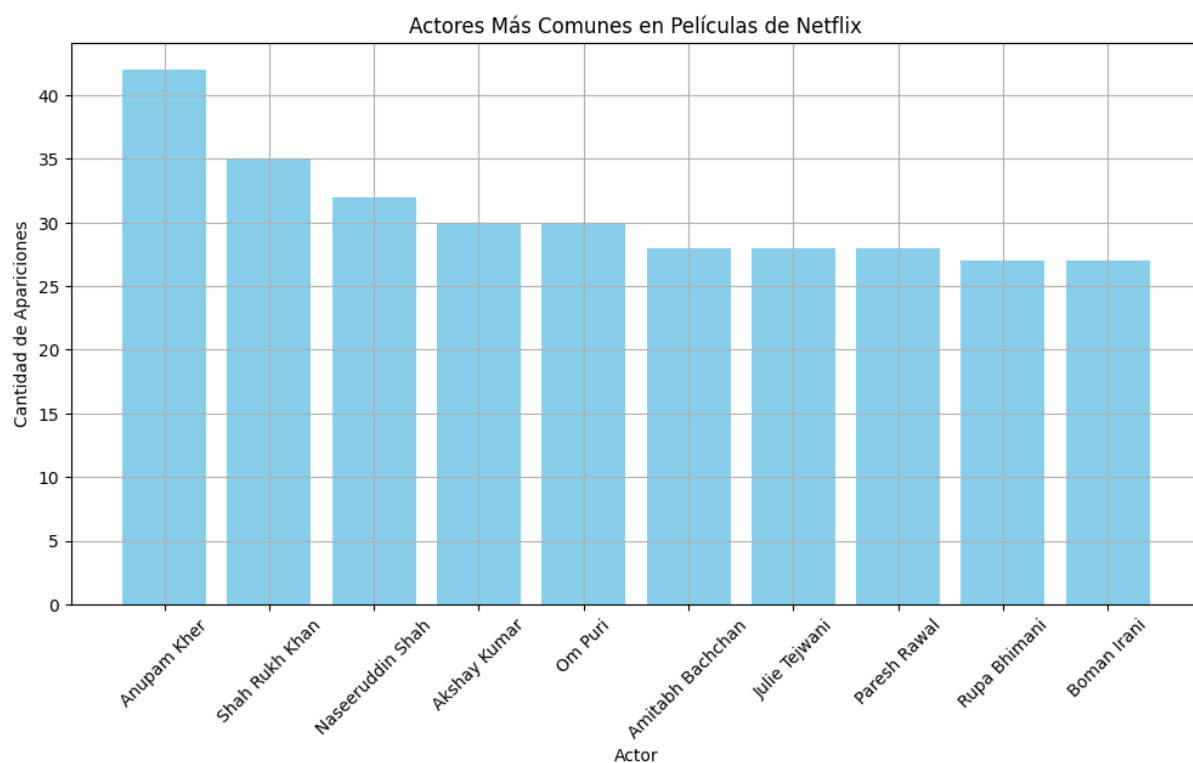


Figura 12: Actores Más Comunes en Películas

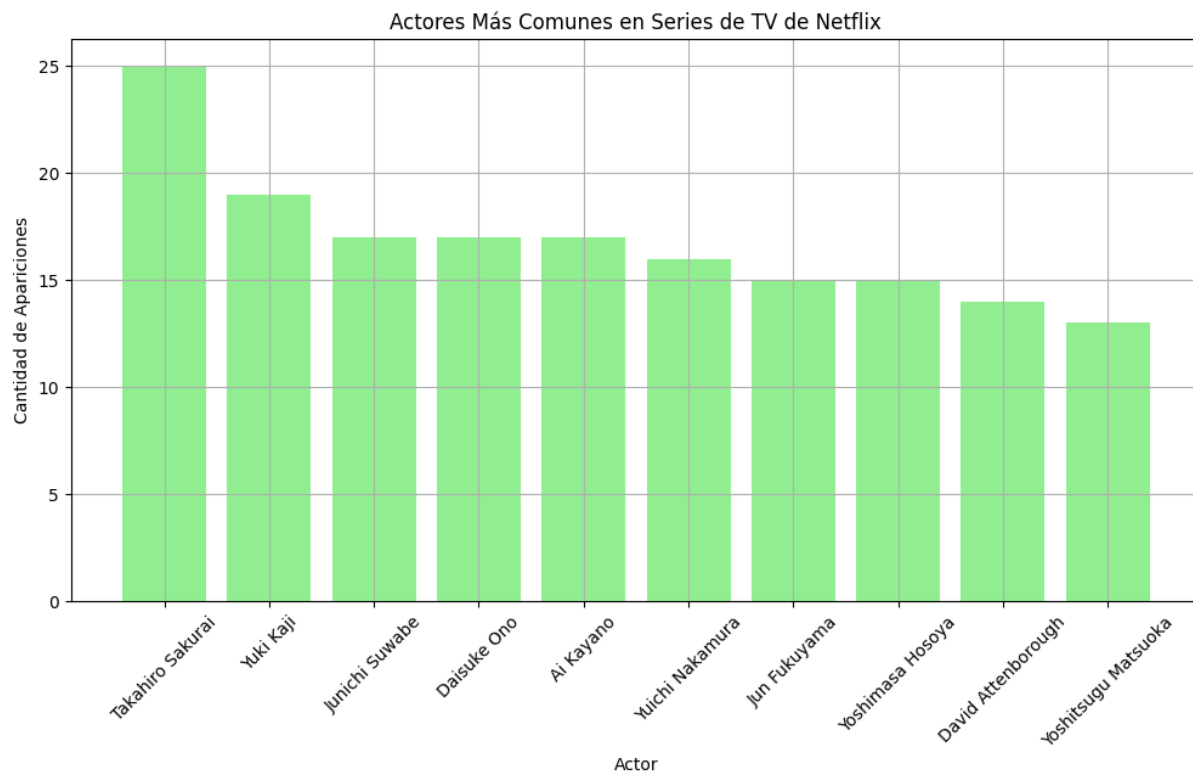


Figura 13: Actores Más Comunes en Series de TV

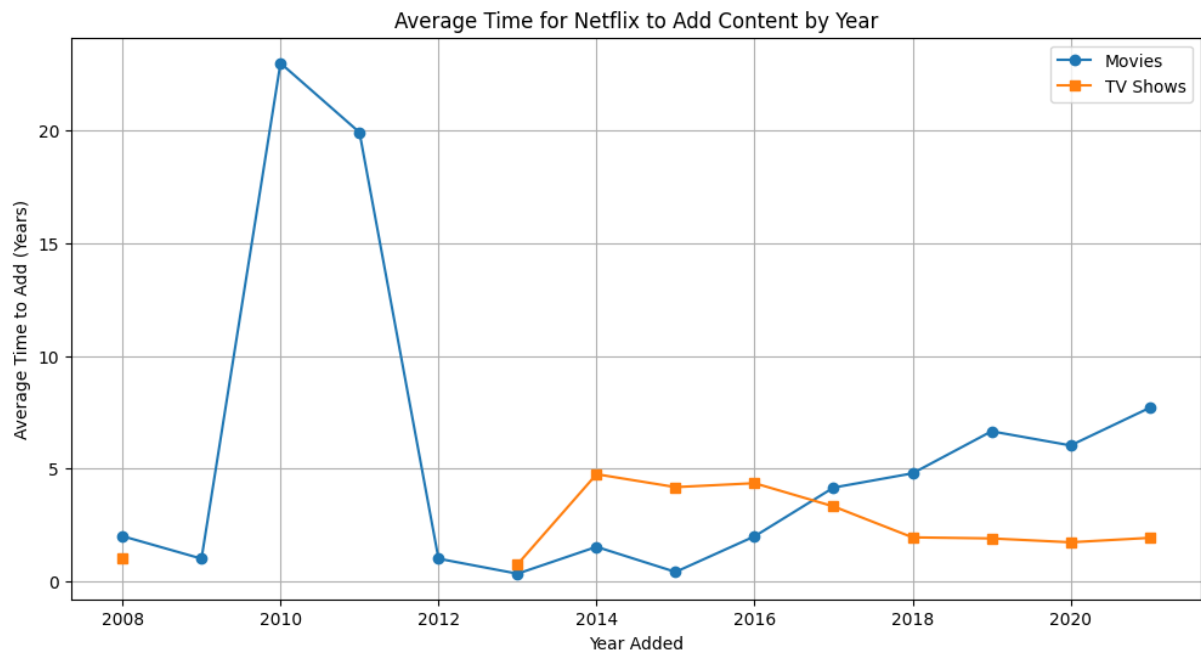


Figura 14: Tiempo promedio anual en añadir el contenido a la plataforma desde su estreno

3.3. Análisis de Componentes Principales (PCA)

3.4. Distribución Temporal

3.5. Relaciones entre Variables

4. Conclusiones

Los principales hallazgos incluyen:

- 68 % del contenido son películas con duración promedio de 99.5 minutos
- Estados Unidos produce el 35 % del contenido total
- Existe correlación positiva ($r=0.62$) entre año de lanzamiento y fecha de incorporación
- El primer componente principal explica el 45 % de la varianza en características técnicas

5. Recomendaciones

- Profundizar análisis de preferencias regionales
- Incorporar datos de valoración de usuarios
- Estudiar correlación entre géneros y éxito comercial