
Resumen de lo que contiene el dataset

Este dataset contiene datos tanto de películas y tv shows disponibles en Netflix. Cada entrada contiene datos como: tipo (movie, tv show), título, director, país de origen, fecha que fue añadida, año de publicación, clasificación, duración, y género.

Descripción de las columnas

Nombre	Tipo	Descripción
Show_id	Integer	Indica el identificador correspondiente a cada película o TV show.
Type	String	Indica si la entrada es de tipo película o TV show.
Title	String	El título de la película o TV show.
Director	String	El nombre del director de las películas o los TV shows.
Country	String	País de origen dónde la película o TV show fue producida.
Date_added	Date	Fecha en la cual la película o TV show fue añadida a Netflix.
Release_year	Integer	Año de publicación de la película o TV show.
Rating	String	La clasificación a la cual pertenece la película o TV show, incluye TV-MA, TV-14, R, PG-13, TV-PG, TV-Y, PG, TV-G, y TV-Y7.
Duration	Integer	La duración de las películas o TV show. En el caso de las películas se proporciona el mismo en minutos, y en el caso de los TV shows se proporciona en temporadas.
Listed_in	String	Hace referencia al género al que pertenece la película o TV show.

Problemas encontrados en el dataset

Datos sucios / desordenados

- Existencia de 3 entradas con datos duplicados para películas.
- Se formateo toda la columna "date_added" a tipo fecha.
- La columna "duration" tiene un formato de datos incorrecto. Se quitó en todos los elementos "min", dejando solamente el número correspondiente a los minutos.

Datos eliminados del dataset original

- Se eliminó la columna "show_id", debido a que no van a ser objeto de este análisis.
- Se eliminaron los elementos de TV shows, debido a que no van a ser objeto de este análisis.
- Se quitó en todos los elementos "min" de la columna "duration", dejando solamente el número correspondiente a los minutos.

Objetivo

El propósito de este análisis es visualizar a través de un dashboard simple los diferentes KPI (key performance indicator) sólo de las películas, que fueron lanzadas desde 1942 a 2021. Tanto de manera general, como en específico del país Argentina.

Según el set de datos se me ocurrió determinar:

- Total de películas en número.
- Cantidad de directores.
- Cantidad de países.
- Cantidad de categorías.
- Top 5 directores con más películas publicadas.
- Top 5 años con mayor cantidad de películas publicadas.
- Total de películas por clasificación.

Relevamiento en la segunda hoja del power bi orientado a Argentina:

- Total de películas en número de Argentina.
- Cantidad de directores en Argentina.
- Cantidad de películas por categoría en Argentina.
- Top 5 directores en Argentina.
- Top 5 categorías en Argentina.
- Total de películas por clasificación.
- Top 5 años con mayor cantidad de películas publicadas.

Resultado

El Dashboard interactivo, le permite a cualquier usuario utilizar los diferentes filtros y gráficos para obtener información específica y adecuada sobre el dataset en cuestión.

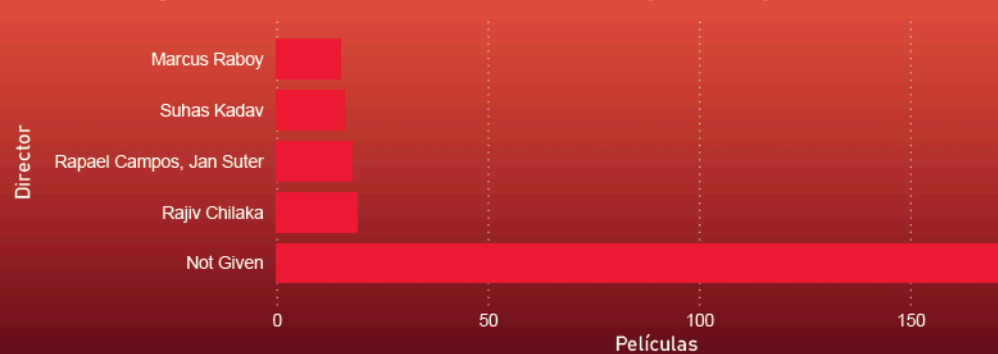
Duración del proyecto

El proceso de análisis, limpieza y transformación de los datos tomó un total de 1 semana. El proceso de elaboración del dashboard interactivo con sus respectivos KPI's un total de 1,5 semanas.



Análisis exploratorio del Dataset de Netflix | Período 1942 a 2021

¿Cuales fueron los 5 directores con más películas publicadas?



6123

Películas



4352

Directores



79

Países



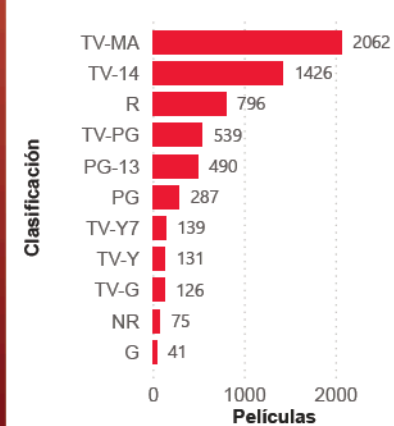
19

Categorías

¿En que años se publicaron más películas?



Total películas por clasificación



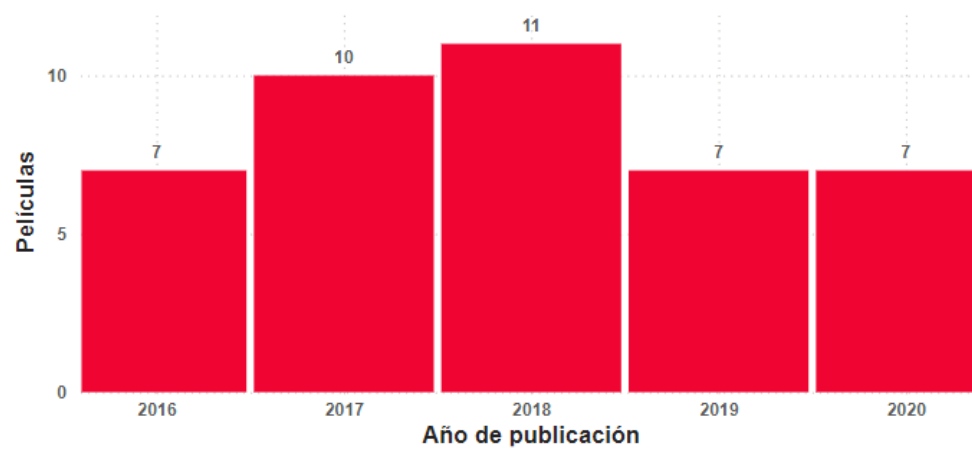


NETFLIX

RESUMEN ARGENTINA



¿En que años se publicaron más películas?



Total películas por clasificación

