

Analisis Dataset Netflix

Notebook sencillo basado en el dashboard de Streamlit

1. KPIs: Total de títulos, Mediana de duración películas y duración de Series.
2. Content Health Producción por año y cobertura por rating
3. Detalles & Contexto: Correlación Año de Lanzamiento vs Año agregado vs Duración

Usa el CSV de Kaggle "netflix_titles.csv".

Cargar el dataset

Dataset descargado de Kaggle para práctica

```
In [84]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn

pd.set_option("display.max_colwidth",200)

df = pd.read_csv("netflix_titles.csv")
```

Comandos Dataframe

```
In [2]: df.head()
```

Out[2]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	T M
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	T M
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	T M
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	T M

In [3]: df.head(20)

Out[3]:

	show_id	type	title	director	cast	country	date_added	release_yr
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2020
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2020
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2020
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2020
5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...	NaN	September 24, 2021	2020
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	NaN	September 24, 2021	2020
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso,	September 24, 2021	1990

show_id	type	title	director	cast	country	date_added	release_yr	
					United Kin...			
8	s9	TV Show	The Great British Baking Show	Andy Devonshire	Mel Giedroyc, Sue Perkins, Mary Berry, Paul Ho...	United Kingdom	September 24, 2021	2021
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	September 24, 2021	2020
10	s11	TV Show	Vendetta: Truth, Lies and The Mafia	NaN	NaN	NaN	September 24, 2021	2020
11	s12	TV Show	Bangkok Breaking	Kongkiat Komesiri	Sukollawat Kanarot, Sushar Manaying, Pavarit M...	NaN	September 23, 2021	2020
12	s13	Movie	Je Suis Karl	Christian Schwochow	Luna Wedler, Jannis Niewöhner, Milan Peschel, ...	Germany, Czech Republic	September 23, 2021	2020
13	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Klara Castanho, Lucca Picon, Júlia Gomes, Marc...	NaN	September 22, 2021	2020
14	s15	TV Show	Crime Stories: India Detectives	NaN	NaN	NaN	September 22, 2021	2020
15	s16	TV Show	Dear White People	NaN	Logan Browning, Brandon P. Bell, DeRon Horton,...	United States	September 22, 2021	2020

	show_id	type	title	director	cast	country	date_added	release_yr
16	s17	Movie	Europe's Most Dangerous Man: Otto Skorzeny in ...	Pedro de Echave García, Pablo Azorín Williams	NaN	NaN	September 22, 2021	2021
17	s18	TV Show	Falsa identidad	NaN	Luis Ernesto Franco, Camila Sodi, Sergio Goyri...	Mexico	September 22, 2021	2021
18	s19	Movie	Intrusion	Adam Salky	Freida Pinto, Logan Marshall-Green, Robert Joh...	NaN	September 22, 2021	2021
19	s20	TV Show	Jaguar	NaN	Blanca Suárez, Iván Marcos, Óscar Casas, Adriá...	NaN	September 22, 2021	2021

In [4]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 12 columns):
#   Column          Non-Null Count  Dtype
---  -
0   show_id         8807 non-null   object
1   type            8807 non-null   object
2   title           8807 non-null   object
3   director        6173 non-null   object
4   cast            7982 non-null   object
5   country         7976 non-null   object
6   date_added      8797 non-null   object
7   release_year    8807 non-null   int64
8   rating          8803 non-null   object
9   duration        8804 non-null   object
10  listed_in       8807 non-null   object
11  description      8807 non-null   object
dtypes: int64(1), object(11)
memory usage: 825.8+ KB
```

In [5]: `df.columns`

```
Out[5]: Index(['show_id', 'type', 'title', 'director', 'cast', 'country', 'date_added',  
             'release_year', 'rating', 'duration', 'listed_in', 'description'],  
            dtype='object')
```

```
In [6]: df.describe()
```

```
Out[6]:
```

	release_year
--	--------------

count	8807.000000
-------	-------------

mean	2014.180198
------	-------------

std	8.819312
-----	----------

min	1925.000000
-----	-------------

25%	2013.000000
-----	-------------

50%	2017.000000
-----	-------------

75%	2019.000000
-----	-------------

max	2021.000000
-----	-------------

```
In [7]: df.shape
```

```
Out[7]: (8807, 12)
```

```
In [8]: df.sample(10)
```

Out[8]:

	show_id	type	title	director	cast	country	date_added	release_y
7213	s7214	Movie	Kiss the Girls	Gary Fleder	Morgan Freeman, Ashley Judd, Cary Elwes, Alex ...	United States	January 1, 2020	1
3378	s3379	TV Show	Nailed It! Spain	NaN	La Terremoto de Alcorcón, Christian Escribà	Spain	October 25, 2019	2
1545	s1546	TV Show	Pup Academy	NaN	Christian Convery, Don Lake, Chance Hurstfield...	United States	December 15, 2020	2
8330	s8331	Movie	The Great Battle	NaN	Zo In-sung, Nam Joo-hyuk, Park Sung-woong, Bae...	South Korea	April 8, 2019	2
7555	s7556	Movie	Nails	Dennis Bartok	Leah McNamara, Shauna Macdonald, Steve Wall, R...	Ireland	March 10, 2018	2
1530	s1531	TV Show	Schulz Saves America	Alexx Media	Andrew Schulz	United States	December 17, 2020	2
4057	s4058	TV Show	Northern Rescue	NaN	William Baldwin, Kathleen Robertson, Spencer M...	Canada	March 1, 2019	2
7532	s7533	Movie	My Brother ... Nikhil	Onir	Juhi Chawla, Sanjay Suri, Victor Banerjee, Lil...	India	October 30, 2018	2

	show_id	type	title	director	cast	country	date_added	release_y
1109	s1110	Movie	Just Say Yes	Appie Boudellah, Aram van de Rest	Yolanthe Cabau, Noortje Herlaar, Kim-Lian van ...	Netherlands	April 2, 2021	2
3969	s3970	TV Show	Traitors	NaN	Emma Appleton, Michael Stuhlbarg, Luke Treadaw...	United Kingdom	March 29, 2019	2

```
In [10]: df.tail()
```


Out[10]:

	show_id	type	title	director	cast	country	date_added	release_year
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015

In [4]: df["country"].value_counts(10)

Out[4]:

country	
United States	0.353310
India	0.121866
United Kingdom	0.052533
Japan	0.030717
South Korea	0.024950
...	
Mexico, United States, Spain, Colombia	0.000125
Canada, Norway	0.000125
Finland, Germany, Belgium	0.000125
Argentina, United States, Mexico	0.000125
United Kingdom, United States, Germany, Denmark, Belgium, Japan	0.000125

Name: proportion, Length: 748, dtype: float64

In [5]: df["country"].value_counts(10).head(10)

```
Out[5]: country
United States    0.353310
India            0.121866
United Kingdom   0.052533
Japan            0.030717
South Korea      0.024950
Canada           0.022693
Spain            0.018180
France           0.015547
Mexico           0.013791
Egypt            0.013290
Name: proportion, dtype: float64
```

In [13]:

Out[13]:

	show_id	type	title	director	cast	country	date_added	release_year	
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	
...
95	s96	TV Show	The Circle	NaN	Michelle Buteau	United States, United Kingdom	September 8, 2021	2021	
96	s97	Movie	If I Leave Here Tomorrow: A Film About Lynyrd ...	Stephen Kijak	Ronnie Van Zandt, Gary Rossington, Allen Colli...	United States	September 7, 2021	2018	

	show_id	type	title	director	cast	country	date_added	release_year	
97	s98	TV Show	Kid Cosmic	NaN	Jack Fisher, Tom Kenny, Amanda C. Miller, Kim ...	United States	September 7, 2021	2021	
98	s99	TV Show	Octonauts: Above & Beyond	NaN	Antonio Aakeel, Chip Chung, Simon Foster, Ter...	United Kingdom	September 7, 2021	2021	
99	s100	TV Show	On the Verge	NaN	Julie Delpy, Elisabeth Shue, Sarah Jones, Alex...	France, United States	September 7, 2021	2021	

100 rows × 12 columns

```
In [14]: df.tail()
```

Out[14]:

	show_id	type	title	director	cast	country	date_added	release_year
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015

Análisis de Datos

Rangos por fecha

```
In [7]: list (df['release_year'].dropna().unique())
```

```
Out[7]: [np.int64(2020),
         np.int64(2021),
         np.int64(1993),
         np.int64(2018),
         np.int64(1996),
         np.int64(1998),
         np.int64(1997),
         np.int64(2010),
         np.int64(2013),
         np.int64(2017),
         np.int64(1975),
         np.int64(1978),
         np.int64(1983),
         np.int64(1987),
         np.int64(2012),
         np.int64(2001),
         np.int64(2014),
         np.int64(2002),
         np.int64(2003),
         np.int64(2004),
         np.int64(2011),
         np.int64(2008),
         np.int64(2009),
         np.int64(2007),
         np.int64(2005),
         np.int64(2006),
         np.int64(1994),
         np.int64(2015),
         np.int64(2019),
         np.int64(2016),
         np.int64(1982),
         np.int64(1989),
         np.int64(1990),
         np.int64(1991),
         np.int64(1999),
         np.int64(1986),
         np.int64(1992),
         np.int64(1984),
         np.int64(1980),
         np.int64(1961),
         np.int64(2000),
         np.int64(1995),
         np.int64(1985),
         np.int64(1976),
         np.int64(1959),
         np.int64(1988),
         np.int64(1981),
         np.int64(1972),
         np.int64(1964),
         np.int64(1945),
         np.int64(1954),
         np.int64(1979),
         np.int64(1958),
         np.int64(1956),
         np.int64(1963),
         np.int64(1970),
```

```

np.int64(1973),
np.int64(1925),
np.int64(1974),
np.int64(1960),
np.int64(1966),
np.int64(1971),
np.int64(1962),
np.int64(1969),
np.int64(1977),
np.int64(1967),
np.int64(1968),
np.int64(1965),
np.int64(1946),
np.int64(1942),
np.int64(1955),
np.int64(1944),
np.int64(1947),
np.int64(1943)]

```

```
In [85]: df["release_year"].dropna().unique()
```

```

Out[85]: array([2020, 2021, 1993, 2018, 1996, 1998, 1997, 2010, 2013, 2017, 1975,
                1978, 1983, 1987, 2012, 2001, 2014, 2002, 2003, 2004, 2011, 2008,
                2009, 2007, 2005, 2006, 1994, 2015, 2019, 2016, 1982, 1989, 1990,
                1991, 1999, 1986, 1992, 1984, 1980, 1961, 2000, 1995, 1985, 1976,
                1959, 1988, 1981, 1972, 1964, 1945, 1954, 1979, 1958, 1956, 1963,
                1970, 1973, 1925, 1974, 1960, 1966, 1971, 1962, 1969, 1977, 1967,
                1968, 1965, 1946, 1942, 1955, 1944, 1947, 1943])

```

Director por Títulos

```
In [86]: director = df['director'].unique()
print(director)
```

```

['Kirsten Johnson' nan 'Julien Leclercq' ... 'Majid Al Ansari'
 'Peter Hewitt' 'Mozes Singh']

```

```
In [87]: df["director"] = df["director"].fillna("Desconocido")
df[['title', 'director']].head(10)
```

Out[87]:

	title	director
0	Dick Johnson Is Dead	Kirsten Johnson
1	Blood & Water	Desconocido
2	Ganglands	Julien Leclercq
3	Jailbirds New Orleans	Desconocido
4	Kota Factory	Desconocido
5	Midnight Mass	Mike Flanagan
6	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha
7	Sankofa	Haile Gerima
8	The Great British Baking Show	Andy Devonshire
9	The Starling	Theodore Melfi

```
In [68]: data = df[~df["title"].isna()]
len(data)
```

Out[68]: 8807

```
In [69]: data
```


Out[69]:

	show_id	type	title	director	cast	country	date_added	release
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	
1	s2	TV Show	Blood & Water	Desconocido	Ama Qamata, Khosi Ngema, Gail Mabalane, Thabang Molaba, Dillon Windvogel, Natasha Thahane, Arno Greeff, Xolile Tshabalala, Getmore Sithole, Cindy Mahlangu, Ryle De Morny, Greteli Fincham, Sello Ma...	South Africa	September 24, 2021	
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabiha Akkari, Sofia Lesaffre, Salim Kechiouche, Noureddine Farihi, Geert Van Rampelberg,	NaN	September 24, 2021	

show_id	type		title	director	cast	country	date_added	release
3	s4	TV Show	Jailbirds New Orleans	Desconocido	Bakary Diombera	NaN	NaN	September 24, 2021
					NaN			
4	s5	TV Show	Kota Factory	Desconocido	Mayur More, Jitendra Kumar, Ranjan Raj, Alam Khan, Ahsaas Channa, Revathi Pillai, Urvi Singh, Arun Kumar	India		September 24, 2021
...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey Jr., Anthony Edwards, Brian Cox, Elias Koteas, Donal Logue, John Carroll Lynch, Dermot Mulroney, Chloë Sevigny	United States		November 20, 2019
8803	s8804	TV Show	Zombie Dumb	Desconocido	NaN	NaN		July 1, 2019

show_id	type	title	director	cast	country	date_added	release
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, Abigail Breslin, Amber Heard, Bill Murray, Derek Graf	United States	November 1, 2019
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Mara, Ryan Newman, Michael Cassidy, Spencer Breslin, Rip Torn, Kevin Zegers	United States	January 11, 2020
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanana, Manish Chaudhary, Meghna Malik, Malkeet Rauni, Anita Shabdish, Chittaranjan Tripathy	India	March 2, 2019

8807 rows × 12 columns

```
In [70]: df = df[~df["country"].isna()]  
len(df)
```

```
Out[70]: 7976
```

```
In [71]: df
```

Out[71]:

	show_id	type	title	director	cast	country	date_added	release_date
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	
1	s2	TV Show	Blood & Water	Desconocido	Ama Qamata, Khosi Ngema, Gail Mabalane, Thabang Molaba, Dillon Windvogel, Natasha Thahane, Arno Greeff, Xolile Tshabalala, Getmore Sithole, Cindy Mahlangu, Ryle De Morny, Greteli Fincham, Sello Ma...	South Africa	September 24, 2021	
4	s5	TV Show	Kota Factory	Desconocido	Mayur More, Jitendra Kumar, Ranjan Raj, Alam Khan, Ahsaas Channa, Revathi Pillai, Urvi Singh, Arun Kumar	India	September 24, 2021	

	show_id	type	title	director	cast	country	date_added	release_date
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra Duah, Nick Medley, Mutabaruka, Afemo Omilami, Reggie Carter, Mzuri	United States, Ghana, Burkina Faso, United Kingdom, Germany, Ethiopia	September 24, 2021	
8	s9	TV Show	The Great British Baking Show	Andy Devonshire	Mel Giedroyc, Sue Perkins, Mary Berry, Paul Hollywood	United Kingdom	September 24, 2021	
...
8801	s8802	Movie	Zinzana	Majid Al Ansari	Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, Mansoor Alfeeli, Ahd	United Arab Emirates, Jordan	March 9, 2016	
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey Jr., Anthony Edwards, Brian Cox, Elias Koteas, Donal	United States	November 20, 2019	

show_id	type		title	director	cast	country	date_added	release_date
8804	s8805	Movie	Zombieland	Ruben Fleischer	Logue, John Carroll Lynch, Dermot Mulroney, Chloë Sevigny	United States	November 1, 2019	
					Jesse Eisenberg, Woody Harrelson, Emma Stone, Abigail Breslin, Amber Heard, Bill Murray, Derek Graf			
					Tim Allen, Courteney Cox, Chevy Chase, Kate Mara, Ryan Newman, Michael Cassidy, Spencer Breslin, Rip Torn, Kevin Zegers			
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Mara, Ryan Newman, Michael Cassidy, Spencer Breslin, Rip Torn, Kevin Zegers	United States	January 11, 2020	
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanana, Manish Chaudhary, Meghna Malik, Malkeet Rauni, Anita Shabdish, Chittaranjan Tripathy	India	March 2, 2019	

7976 rows × 12 columns

Fechas de agregado "date_added"

```
In [72]: df["date_added"] = pd.to_datetime(df["date_added"], errors="coerce") ; df[["date_ad
```

```
Out[72]:
```

	date_added
0	2021-09-25
1	2021-09-24
4	2021-09-24
7	2021-09-24
8	2021-09-24

Año de lanzamiento "year_added"

```
In [73]: df["year_added"] = df["date_added"].dt.year; df[["date_added", "year_added"]].head(
```

```
Out[73]:
```

	date_added	year_added
0	2021-09-25	2021.0
1	2021-09-24	2021.0
4	2021-09-24	2021.0
7	2021-09-24	2021.0
8	2021-09-24	2021.0

Date added vs Year_added

```
In [74]: df["year_added"] = df["date_added"].dt.year; df[["date_added", "year_added"]].head(
```

```
Out[74]:
```

	date_added	year_added
0	2021-09-25	2021.0
1	2021-09-24	2021.0
4	2021-09-24	2021.0
7	2021-09-24	2021.0
8	2021-09-24	2021.0
9	2021-09-24	2021.0
12	2021-09-23	2021.0
15	2021-09-22	2021.0

KPIs

Promedio de duración películas y series y rating

```
In [75]: titulos = df["title"].value_counts().sum()
series = (df["type"] == "TV Show").sum()
películas = (df["type"] == "TV Show").sum()
print(titulos, series, películas)
```

7976 2285 2285

```
In [88]: df["duration"].str.split(" ")
```

```
Out[88]: 0      [90, min]
1      [2, Seasons]
2      [1, Season]
3      [1, Season]
4      [2, Seasons]
...
8802   [158, min]
8803   [2, Seasons]
8804   [88, min]
8805   [88, min]
8806   [111, min]
Name: duration, Length: 8807, dtype: object
```

```
In [89]: df["duration"] = df["duration"].str.extract("(\\d+)").astype(float)
```

```
In [78]: movies = df[df["type"] == "Movie"]["duration"].mean()
print("Promedio de duración de películas: " , movies)
```

Promedio de duración de películas: 100.52056962025317

```
In [79]: series = df[df["type"] == "TV Show"] ["duration"].mean()
print("Promedio de duración de las series (temporadas): " , series)
```

Promedio de duración de las series (temporadas): 1.8555798687089715

Rating

```
In [80]: df["rating"].isna().sum()
(df["rating"] == "").sum()
data["rating"] = data["rating"].replace("", pd.NA).fillna("NR") ; data["rating"].va
```

```
Out[80]: rating
TV-MA    3207
TV-14    2160
TV-PG    863
R         799
PG-13     490
TV-Y7     334
TV-Y      307
PG         287
TV-G      220
NR         84
Name: count, dtype: int64
```

```
In [81]: data["rating"].isna().sum()
```

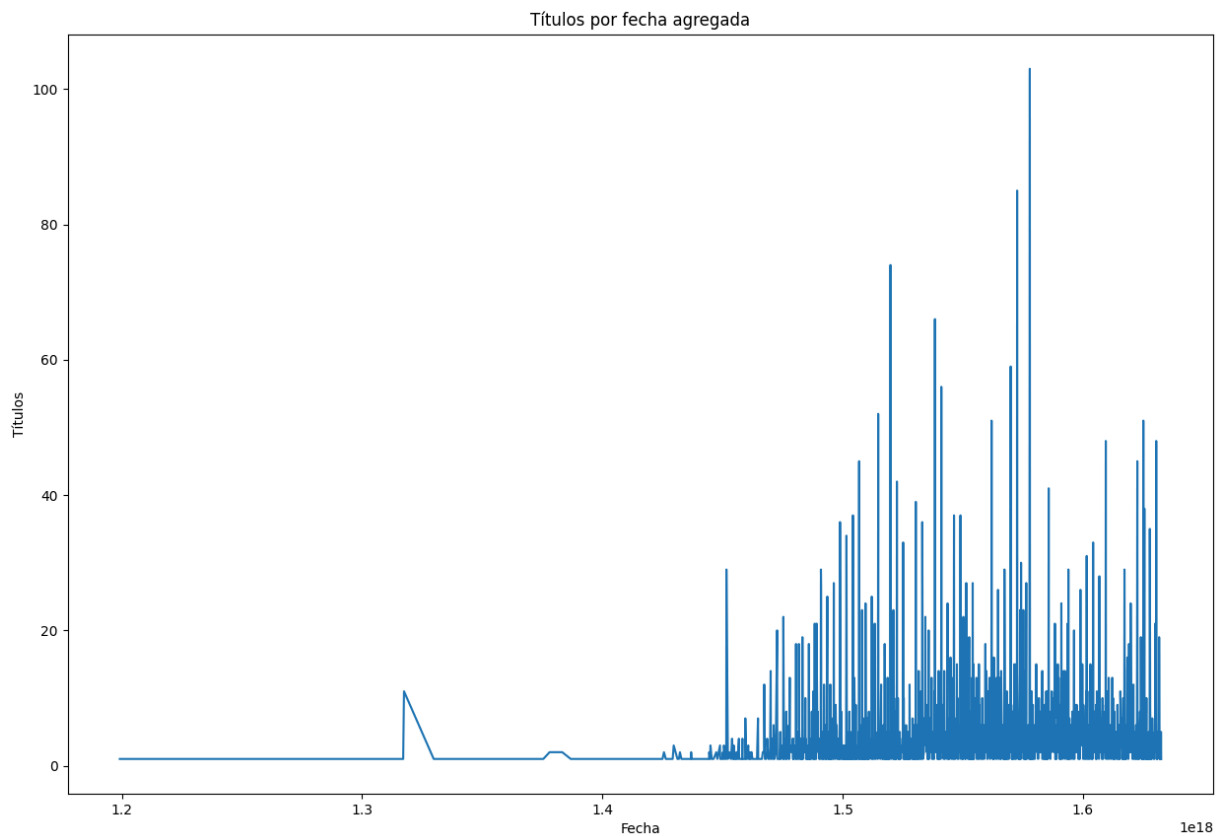
```
Out[81]: np.int64(0)
```

Content Health

Gráfica de títulos por Fecha de Agregado (date_added)

```
In [82]: count= df["date_added"].dropna().astype(int).value_counts().sort_index()
```

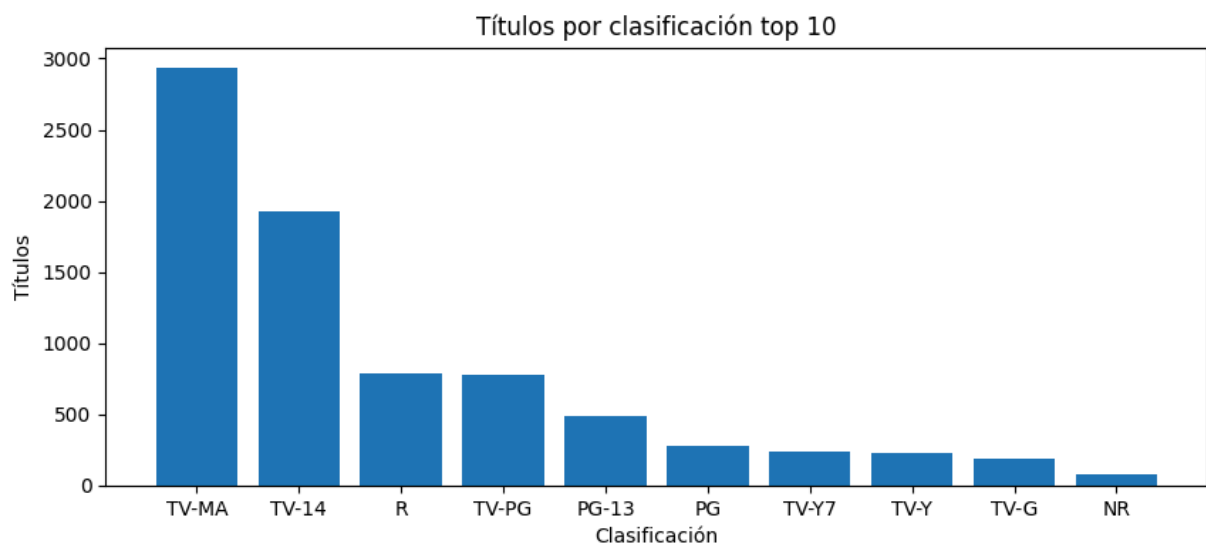
```
In [83]: plt.figure(figsize=(15,10))
plt.plot(count.index, count.values)
plt.xlabel("Fecha")
plt.ylabel("Títulos")
plt.title("Títulos por fecha agregada")
plt.show()
```



Gráfica por Rating

```
In [36]: rating_ct = df["rating"].value_counts().sort_values(ascending=False).head(10)
```

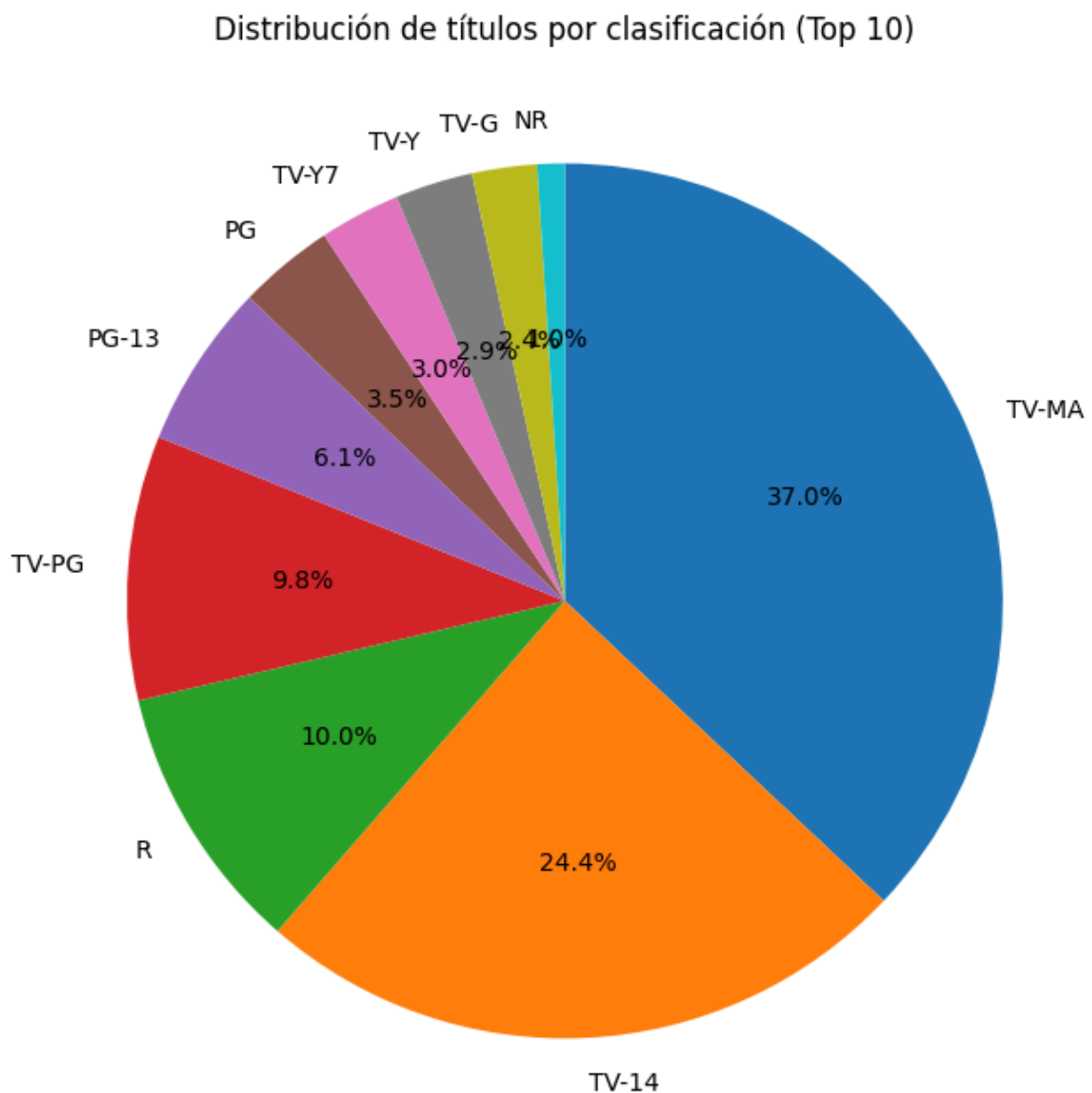
```
In [37]: plt.figure(figsize=(10,4))
plt.bar(rating_ct.index.astype(str), rating_ct.values)
plt.xlabel("Clasificación")
plt.ylabel("Títulos")
plt.title("Títulos por clasificación top 10")
plt.show()
```



Representación Gráfica de Pastel del Top 10

```
In [43]: rating_ct = df["rating"].value_counts().sort_values(ascending=False).head(10)
```

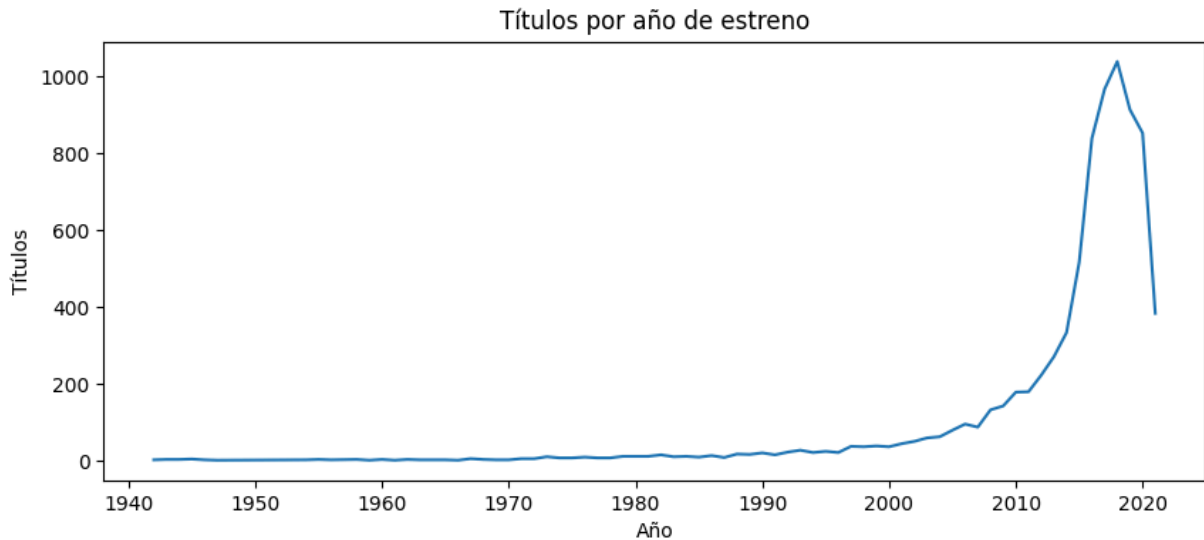
```
In [47]: plt.figure(figsize=(8, 8))
plt.pie(
    rating_ct.values,
    labels=rating_ct.index.astype(str),
    autopct= '%1.1f%%',
    startangle=90,
    counterclock=False
)
plt.title("Distribución de títulos por clasificación (Top 10)")
plt.show()
```



Gráfica por Títulos y Año de Estreno

```
In [38]: count= df["release_year"].dropna().astype(int).value_counts().sort_index()
```

```
In [39]: plt.figure(figsize=(10,4))
plt.plot(count.index, count.values)
plt.xlabel("Año")
plt.ylabel("Títulos")
plt.title("Títulos por año de estreno")
plt.show()
```

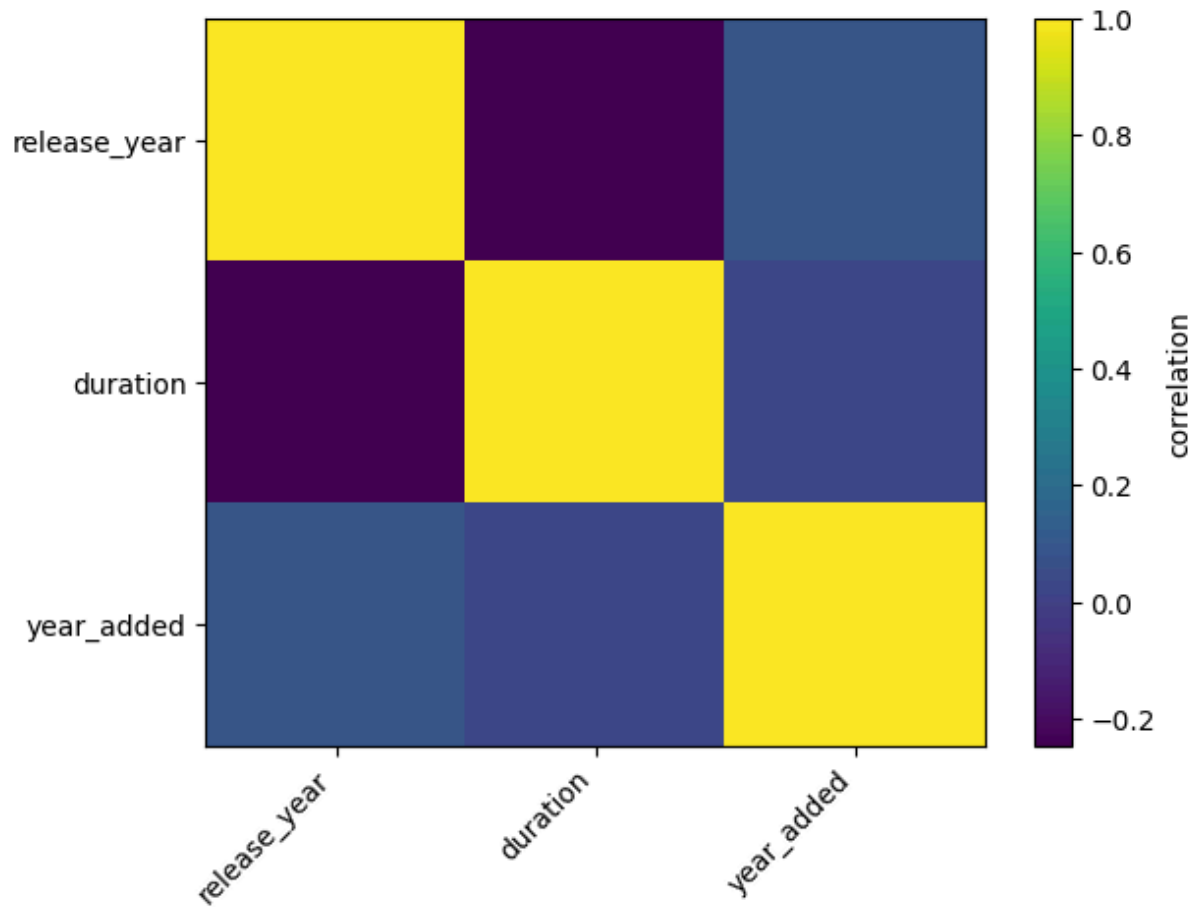


Correlación entre columnas "release_year" vs "year_added" vs "duration"

```
In [48]: num= df.select_dtypes(include="number")
```

```
In [49]: corr = num.corr(numeric_only=True)
plt.imshow(corr, aspect="auto")
plt.colorbar(label="correlation")
plt.xticks(range(len(corr.columns)), corr.columns, rotation=45, ha="right")
plt.yticks(range(len(corr.columns)), corr.columns)
```

```
Out[49]: ([<matplotlib.axis.YTick at 0x7ef25d2c7a10>,
<matplotlib.axis.YTick at 0x7ef25a3e5d50>,
<matplotlib.axis.YTick at 0x7ef25a48c950>],
[Text(0, 0, 'release_year'),
Text(0, 1, 'duration'),
Text(0, 2, 'year_added')])
```



In []:

In []:

In []:

In []: