

NETFLIX

#1 1997

#2 California, EEUU



#3 Plataforma Streaming
(2007)



“Disfruta de tus contenidos **donde y cuando quieras**”

DATASET

197.094

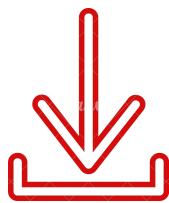
Rows

11

Features

0

Nules



NETFLIX

statista

IMDb

kaggle



OBJETIVO

“ Predecir cuántas semanas va a estar un título top ten en las diferentes regiones, teniendo en cuenta el género, la categoría, la región y el content type”



EDA

Netflix.head(100)

Netflix.info()

Netflix.describe()

Netflix.columns

Netflix['genre'].str

.capitalize()

sorted(Netflix['genre'].
unique())

Netflix.rename

Netflix.dtypes

Netflix.isna().sum()

Netflix['genre'].str.strip()

1

2

3

4

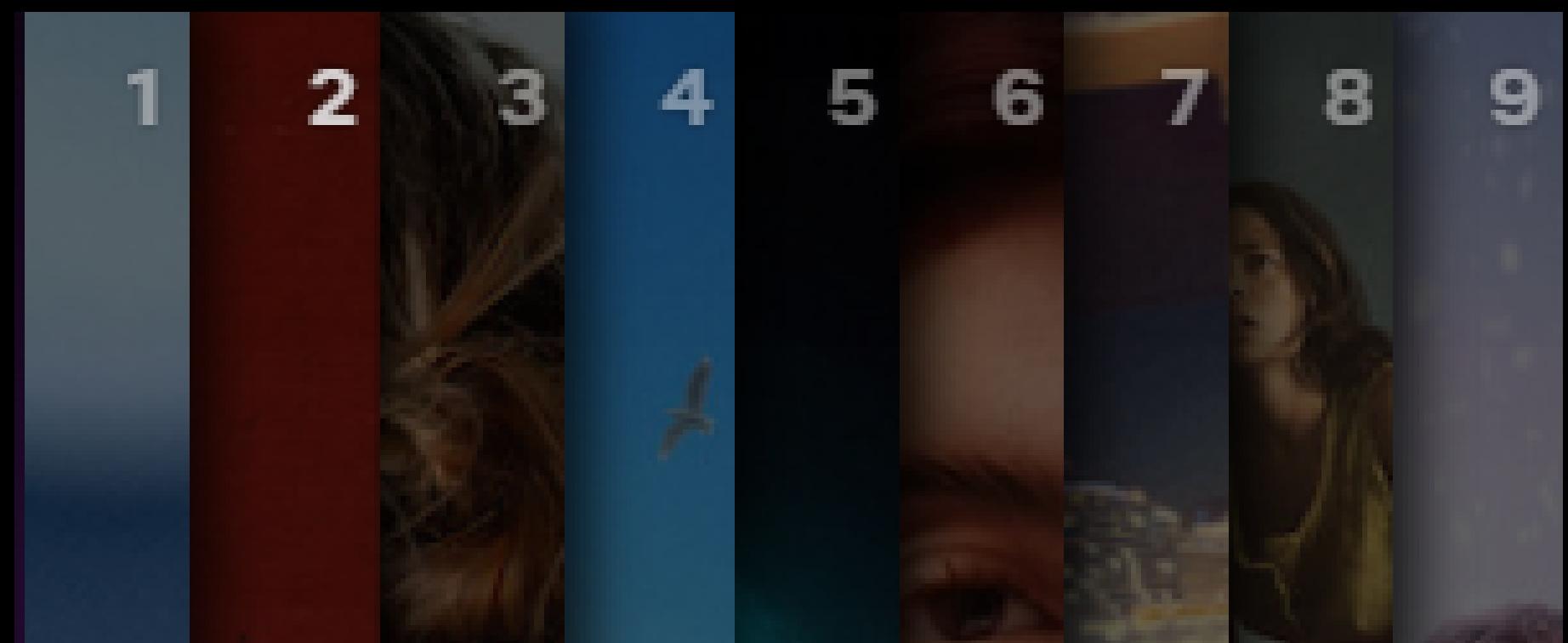
5

6

7

8

9



```

#importa numpy
import numpy as np
#importa pandas
import pandas as pd
#Importa las Librerias matplotlib, seaborn y Scipy
import matplotlib.pyplot as plt
import seaborn as sns
from scipy import stats
%matplotlib inline
import matplotlib.cm as cm

```

#Transformamos los tipos de datos object por strings.

```

Netflix['region'] = Netflix['region'].astype("string")
Netflix['country_name'] = Netflix['country_name'].astype("string")
Netflix['country_iso2'] = Netflix['country_iso2'].astype("string")
Netflix['category'] = Netflix['category'].astype("string")
Netflix['show_title'] = Netflix['show_title'].astype("string")
Netflix['season_title'] = Netflix['season_title'].astype("string")

```

Netflix.dtypes

	region	string
country_name	string	
country_iso2	string	
week	object	
category	string	
weekly_rank	int64	
show_title	string	
season_title	string	
cumulative_weeks_in_top_10	int64	
realease_year	int64	
genre	object	
dtype: object		

Netflix.info()

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 197024 entries, 0 to 197023
Data columns (total 11 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   region          197024 non-null   string  
 1   country_name    197024 non-null   string  
 2   country_iso2    197024 non-null   string  
 3   week            197024 non-null   datetime64[ns]
 4   category        197024 non-null   string  
 5   weekly_rank     197024 non-null   int64  
 6   show_title      197024 non-null   string  
 7   season_title    96884 non-null   string  
 8   cumulative_weeks_in_top_10  197024 non-null   int64  
 9   realease_year   197024 non-null   int64  
 10  genre           197024 non-null   object  
dtypes: datetime64[ns](1), int64(3), object(1), string(6)
memory usage: 16.5+ MB

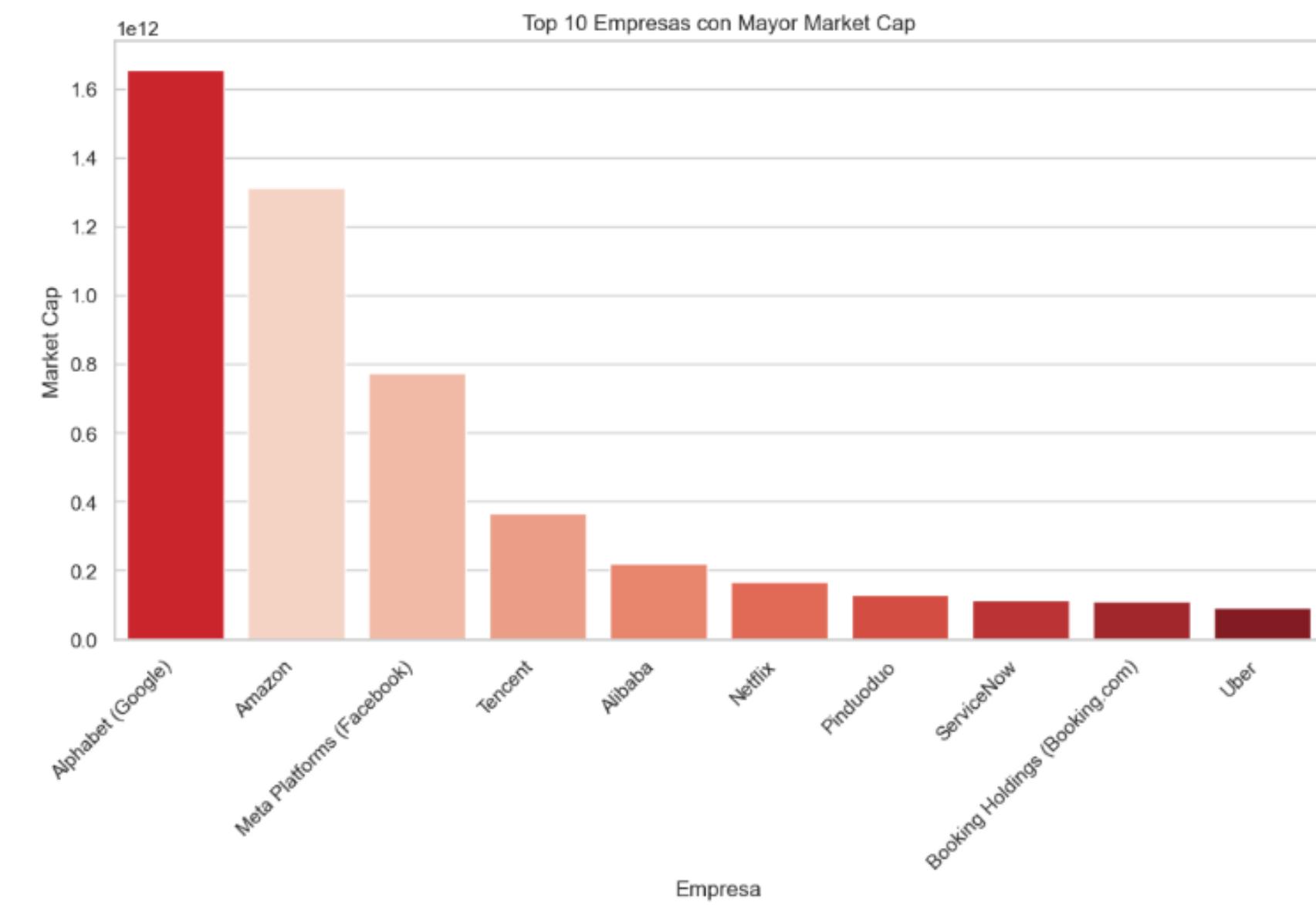
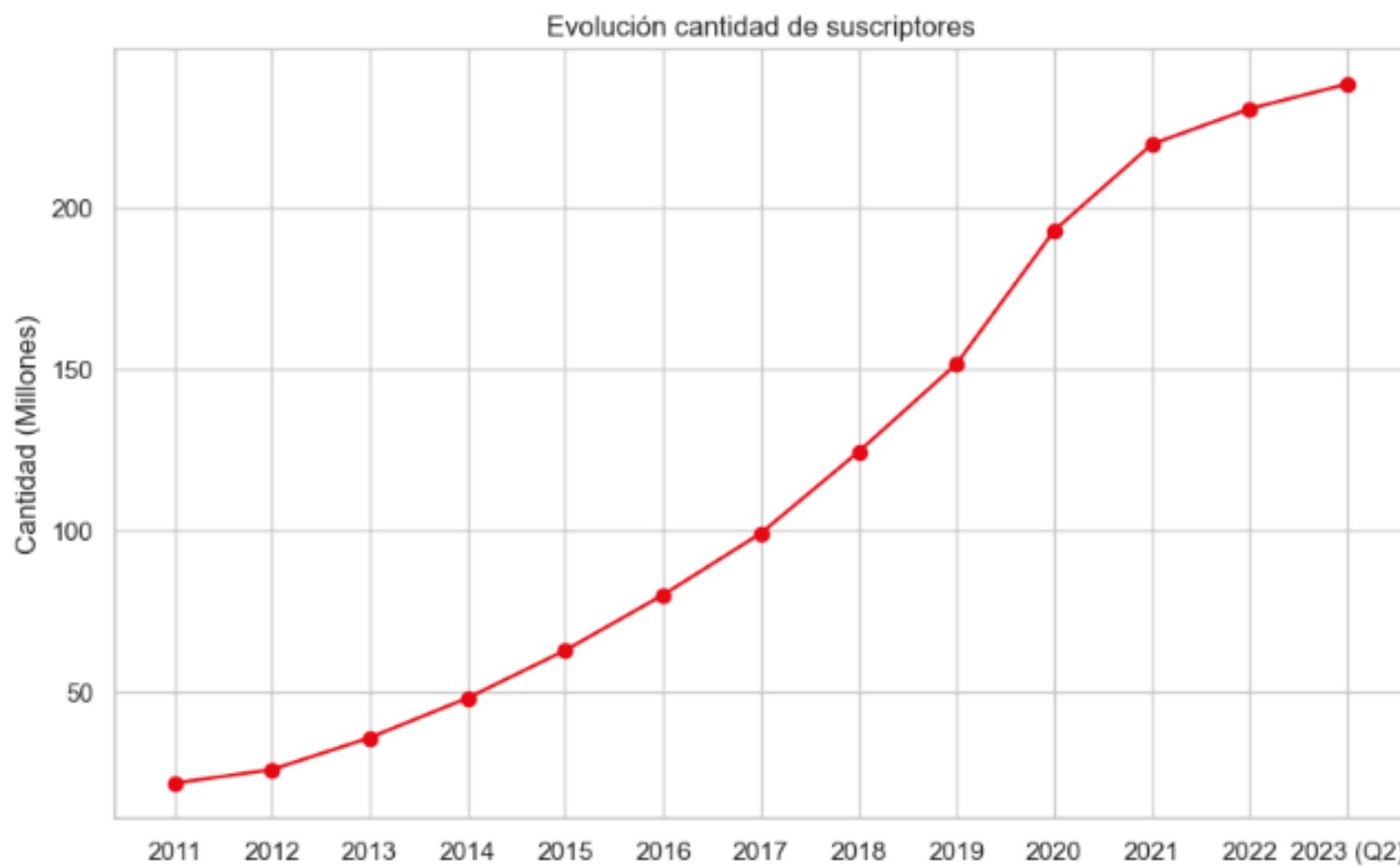
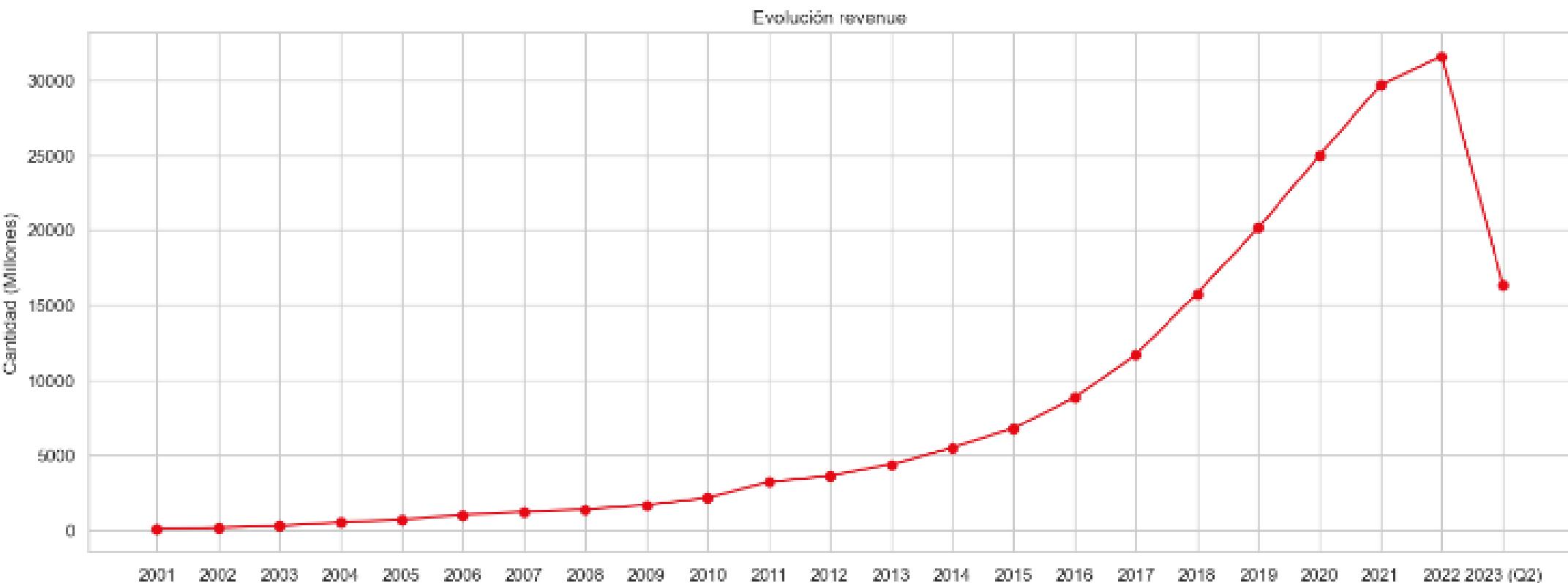
```

Netflix.head(100) #sacar estaciones

	region	country_name	country_iso2	week	category	weekly_rank	show_title	season_title	cumulative_weeks_in_top_10	realease_year	genre
0	Latin America	Argentina	AR	2023-08-13	Movie	1	Heart of Stone	<NA>		1	2023 Action
1	Latin America	Argentina	AR	2023-08-13	Movie	2	Good Luck to You, Leo Grande	<NA>		2	2022 Comedy
2	Latin America	Argentina	AR	2023-08-13	Movie	3	Shark Bait	<NA>		2	2022 Horror
3	Latin America	Argentina	AR	2023-08-13	Movie	4	Clifford the Big Red Dog	<NA>		2	2019 Adventure
4	Latin America	Argentina	AR	2023-08-13	Movie	5	Paradise	<NA>		3	2023 Action
...

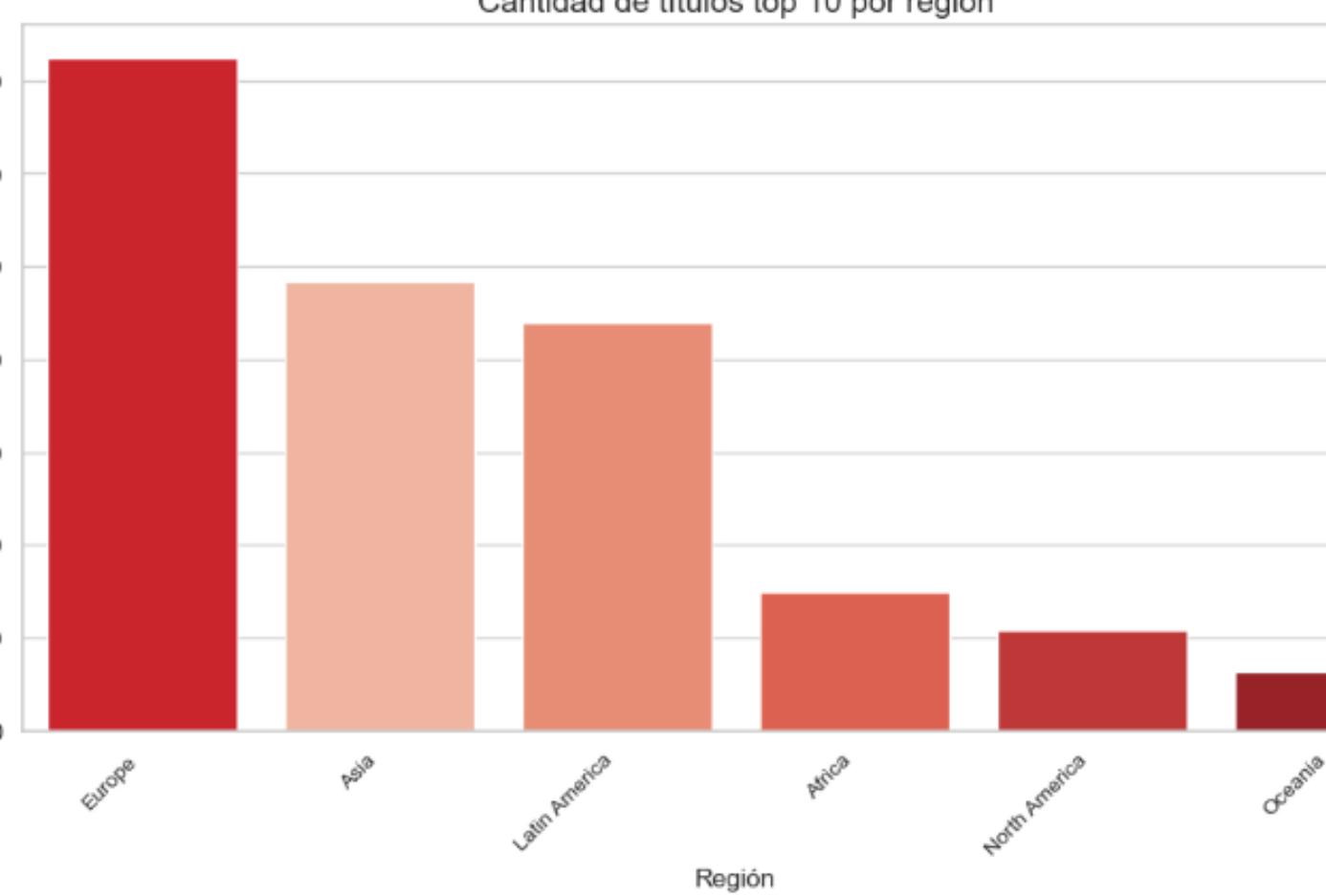


ANÁLISIS GLOBAL

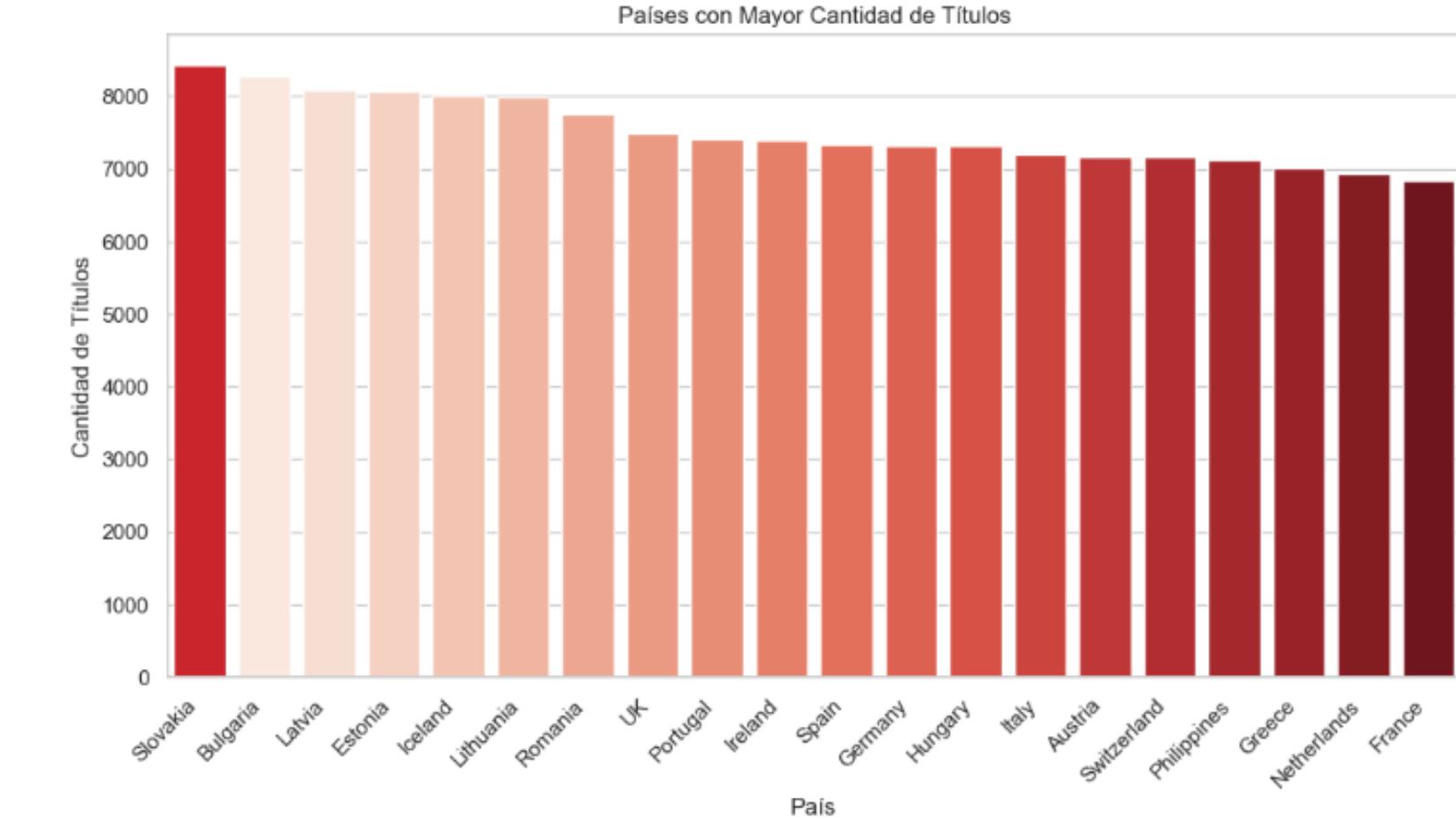


CONTENIDO

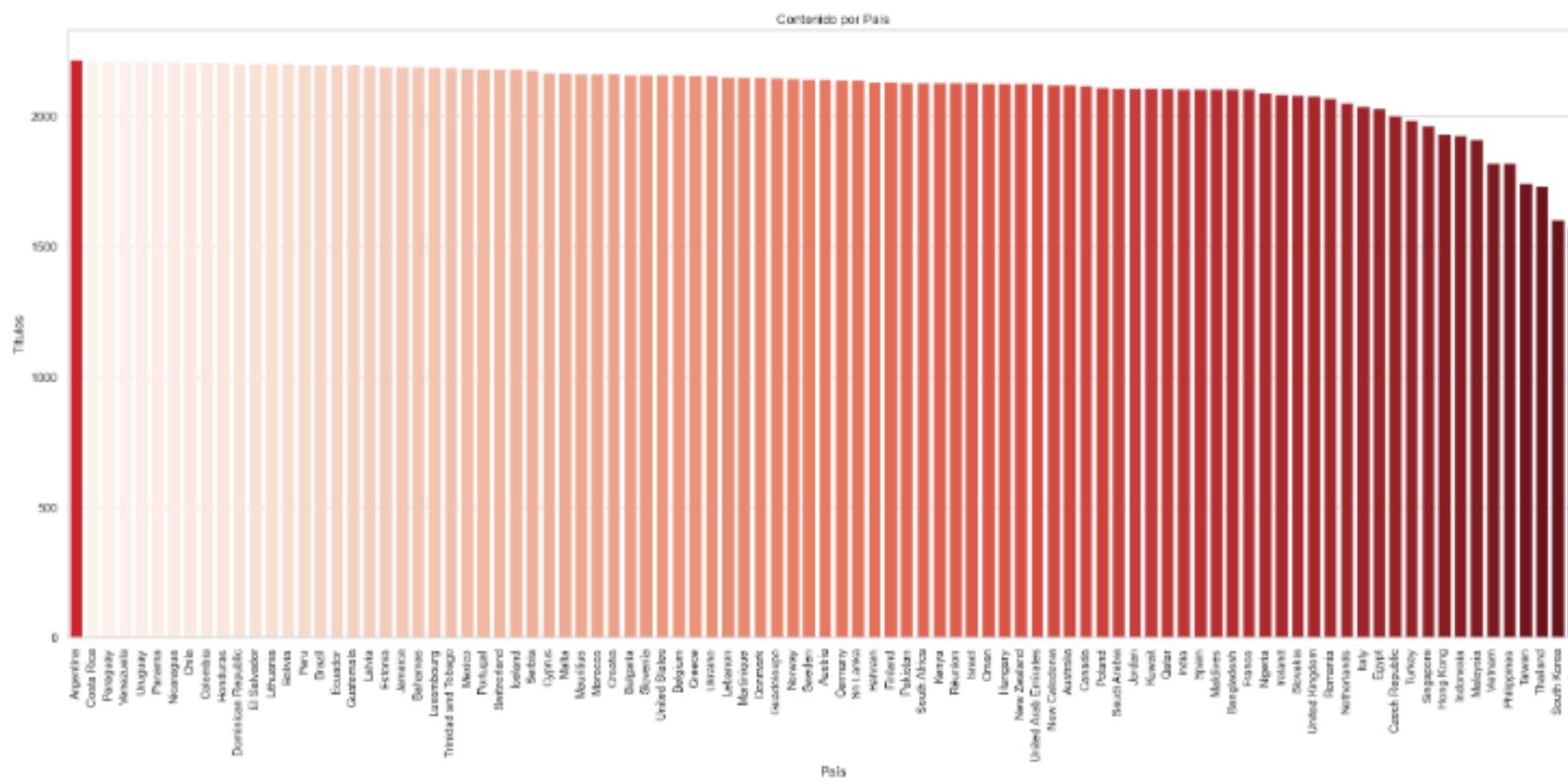
9 10



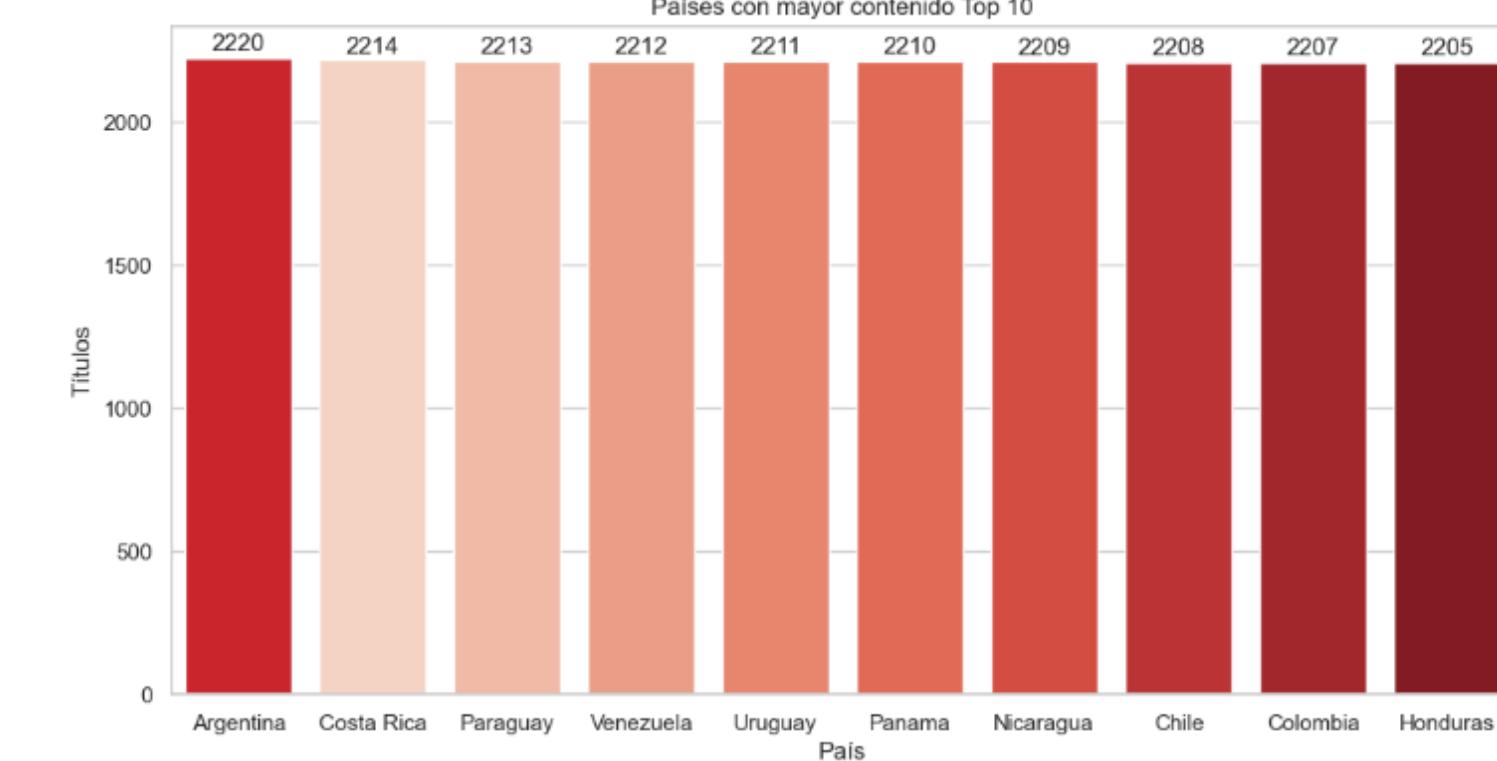
Cantidad de títulos top 10 por región



Países con mayor contenido Top 10



Contenido por



Países con mayor contenido Top 10

TOP GÉNEROS



1

Drama

6

Sci-Fi

2

Acción

7

Romance

3

Comedia

8

Documental

4

Thriller

9

Reality

5

Crimen

10

Kids

TOP 10 GLOBAL



1



6



2



7



3



4



5

Acción

Drama

Drama

Acción

Crimen

Crimen

Drama

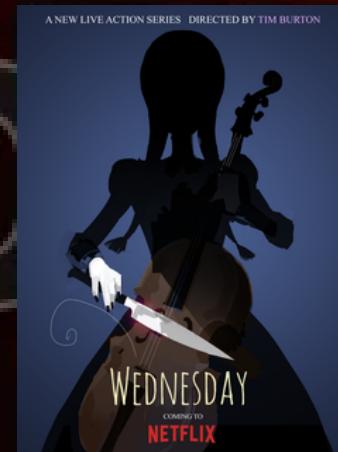
Comedia



8

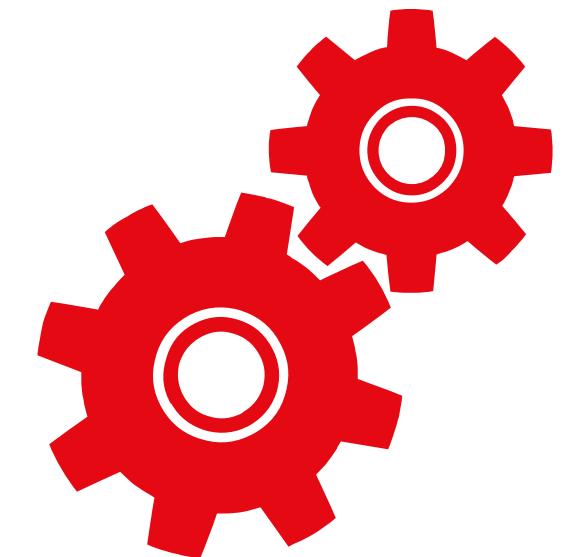


Drama



10

MODELOS APRENDIZAJE SUPERVISADO



Regresión
Lineal

Árboles de
decisión

Random
Forest



Regresión Lineal