Géneros de Netflix

Abrahám Soto y Odalis Rodriguez

2024-07-11

INTRODUCCIÓN

Netflix es una empresa de entretenimiento, con una plataforma de servicio streaming por suscripción en un dispositivo con coneción a internet, ofrece una cantidad de películas, series, documentales, programas interactivos; sin recibir anuncios publicitarios en tu navegación por la plataforma.

Esto nos lleva a conocer el hecho de que Netflix ofrece una buena experiencia audiovisual para el espectador. Por años el gigante de la pantalla chica se catalogó como la plataforma con más titulos disponibles, a travé de la siguiente base de datos, recolectando 583 titulos desde el año 2014 al año 2021, cada uno con sus características.

Con la finalidad de determinar qué géneros de entretenimiento prefiere la audiencia suscrita a la plataforma.

Análisis de la data

En observación a nuestra data podemos establecer que existen 114 géneros en total, provenientes de la variable "Género", los cuales corresponden a un título caracteristicamente.

table(Netflix.G6\$Genero)

##		
##	Action	Action-adventure
##	7	1
	·	
##	Action-thriller	Action comedy
##	3	5
##	Action thriller	Action/Comedy
##	1	1
##	Action/Science fiction	Adventure
	·	
##	1	2
##	Adventure-romance	Adventure/Comedy
##	1	1
##	Aftershow / Interview	Animated musical comedy
##	6	1
##	Animation	Animation / Comedy
##	5	1
##	Animation / Musicial	Animation / Science Fiction
##	1	1
##	Animation / Short	Animation/Christmas/Comedy/Adventure
##	4	1
##	Animation/Comedy/Adventure	Animation/Musical/Adventure
	•	
##	1	1
##	Animation/Superhero	Anime / Short
##	1	1
##	Anime/Fantasy	Anime/Science fiction
##	1	2
##	_	Biographical/Comedy
	Anthology/Dark comedy	
##	1	1
##	Biopic	Black comedy
##	9	2
##	Christian musical	Christmas comedy
##	1	1
##	Christmas musical	Christmas/Fantasy/Adventure/Comedy
		-
##	1	1
##	Comedy	Comedy-drama
##	49	14
##	Comedy / Musical	Comedy horror
##	2	1
##	Comedy mystery	Comedy/Fantasy/Family
##	1	· · · · · · · · · · · · · · · · · · ·
	_	1
##	Comedy/Horror	Coming-of-age comedy-drama
##	1	1
##	Concert Film	Crime drama
##	6	11
##	Crime thriller	Dance comedy
##	1	1
	-	
##	Dark comedy	Documentary
##	2	159

##	Drama	Drama-Comedy
##	77	1
##	Drama / Short	Drama/Horror
##	1 Famile	1
## ##	Family 2	Family film 2
##	Family/Christmas musical	Family/Comedy-drama
##	1	1
##	- Fantasy	Heist
##	1	1
##	Heist film/Thriller	Hidden-camera prank comedy
##	1	1
##	Historical-epic	Historical drama
##	1	1
##	Horror	Horror-thriller
##	9	11
## ##	Horror comedy 1	Horror thriller
##	Horror/Crime drama	Making-of
##	1	2
##	Mentalism special	Mockumentary
##	1	2
##	Musical	Musical / Short
##	2	1
##	Musical comedy	Musical/Western/Fantasy
##	2	1
##	Mystery 2	One-man show 3
## ##	Political thriller	Psychological horror
##	1	1 sychological norror
##	Psychological thriller	Psychological thriller drama
##	4	1
##	Romance	Romance drama
##	6	1
##	Romantic comedy	Romantic comedy-drama
##	39	1
##	Romantic comedy/Holiday	Romantic drama
##	Demontis to a desma	5
## ##	Romantic teen drama 1	Romantic teenage drama
##	Romantic thriller	Satire
##	1	2
##	Science fiction	Science fiction adventure
##	4	1
##	Science fiction thriller	Science fiction/Action
##	1	1
##	Science fiction/Drama	Science fiction/Mystery
##	3	1
##	Science fiction/Thriller	Sports-drama
##	4	Sny thriller
## ##	Sports film 1	Spy thriller 2
##	Stop Motion	Superhero
##	1	2
	_	

##	Superhero-Comedy	Superhero/Action
##	1	1
##	Supernatural drama	Teen comedy-drama
##	1	1
##	Teen comedy horror	Thriller
##	1	33
##	Urban fantasy	Variety show
##	1	4
##	Variety Show	War
##	1	2
##	War-Comedy	War drama
##	1	2
##	Western	Zombie/Heist
##	3	1

Objetivo 1: Agrupar géneros

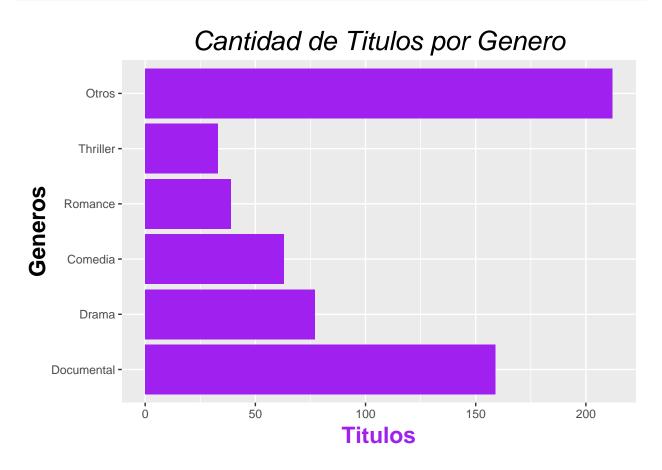
Denotamos que existen similitudes entre los géneros y por ello agrupamos en 6 categorías respectivamente para manejar de mejor manera nuestra información. Contamos con los géneros: Comedia, Thriller, Documentales, Romance, Drama y Otros.

table(Generos)

## Gene	ros					
## Docu	mental	Drama	Comedia	Romance	Thriller	Otros
##	159	77	63	39	33	212

Gracias a nuestra nueva tabla de frecuencias acumuladas, podemos vizualizar a través de un gráfco nuestras cantidades:

GENEROS



Objetivo 2: Puntajes

Evaluaremos la variable imdb_score para analizar los puntajes correspondientes a los títulos, agrupandolos en las 6 categorías de géneros que creamos.

De esta forma creamos una nueva tabla de frecuencias acumuladas:

```
table(Netflix.G6$imdb_score)
```

```
##
## 2.5 2.6 3.2 3.4 3.5 3.7 3.9 4.1 4.2 4.3 4.4 4.5 4.6 4.7 4.8 4.9
                                                                       5 5.1 5.2 5.3
                              1
                                      2
                                          3
                                                                             19
## 5.4 5.5 5.6 5.7 5.8 5.9
                              6 6.1 6.2 6.3 6.4 6.5 6.6 6.7 6.8 6.9
                                                                       7 7.1 7.2 7.3
        19
            15
                20
                    30
                        16
                            13
                                24
                                     18
                                         30
                                             28
                                                 26
                                                     18
                                                         25
                                                              24
                                                                 19
                                                                      19 28
## 7.4 7.5 7.6 7.7 7.8 7.9
                              8 8.1 8.2 8.3 8.4 8.5 8.6
                                                          9
    12
       10 10
                                  3
                                      5
```

Observamos una tabla de frecuencias para los puntajes por lo cual distribuimos en rangos cada puntaje a su respectivo titulo, iniciando con su puntuación más baja 2.5 hasta la más alta 9.

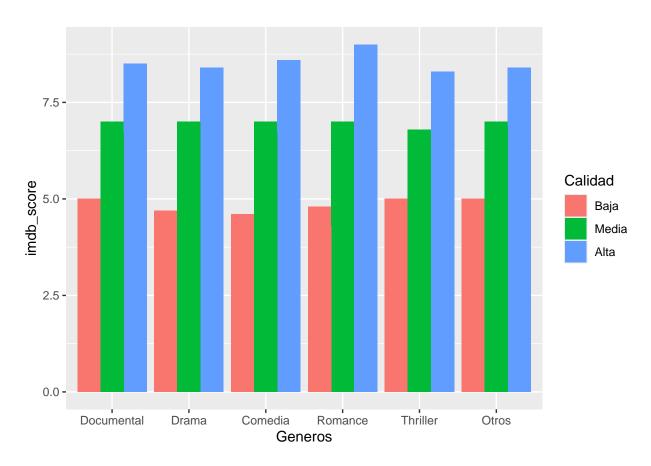
Dando a luz una nueva variable llamada Calidad la cual corresponde a cada titulo, denominada por 3 etiquetas: Buena, Regular y Mala.

table(Netflix.G6\$Calidad)

```
## ## Baja Media Alta
## 58 392 133
```

Podemos viasualizar mejor esta nueva variable a través de un gráfico:

С



Con este gráfico podemos visualizar los rangos de la variable Calidad y agrupando sus respectivas 3 etiquetas en cada género. Observamos que las magnitudes son casi homogeneas.

Objetivo 3: Valores y Procentajes

Generamos una matriz para evaluar las calidades de cada géneros a través de una tabla de frecuencias relativas:

Tabla Acumulada

Datos

```
##
            Comedia Documental Drama Romance Thriller Otros
                 17
                             29
                                             12
                                                       12
## Buena
                                    18
                                                             45
## Regular
                                             23
                                                       17
                 41
                            113
                                    53
                                                            145
## Mala
                             17
                                     6
                                              4
                                                             22
```

Tabla porcentual

prop.table(Datos)

```
## Buena Comedia Documental Drama Romance Thriller Otros ## Buena 0.029159520 0.04974271 0.03087479 0.020583190 0.020583190 0.07718696 ## Regular 0.070325901 0.19382504 0.0909090 0.039451115 0.029159520 0.24871355 ## Mala 0.008576329 0.02915952 0.01029160 0.006861063 0.006861063 0.03773585
```

Ahora conocemos los porcentajes de calidad de cada género, también podemos delimitar los porcentajes marginales para conocer nuevos valores:

```
prop.table(Datos,1)
```

```
## Buena 0.1278195 0.2180451 0.1353383 0.09022556 0.09022556 0.3383459
## Regular 0.1045918 0.2882653 0.1352041 0.05867347 0.04336735 0.3698980
## Mala 0.0862069 0.2931034 0.1034483 0.06896552 0.06896552 0.3793103
```

Distribución de porcentajes de las calidades de titulos.

```
prop.table(Datos, 2)
```

```
## Buena 0.26984127 0.1823899 0.23376623 0.3076923 0.3636364 0.2122642 ## Regular 0.65079365 0.7106918 0.68831169 0.5897436 0.5151515 0.6839623 ## Mala 0.07936508 0.1069182 0.07792208 0.1025641 0.121212 0.1037736
```

Distribución de porcentajes de la cantidad de titulos según sus géneros distribuidos en las 3 categorías de calidad.

Objetivo 4: Correlación.

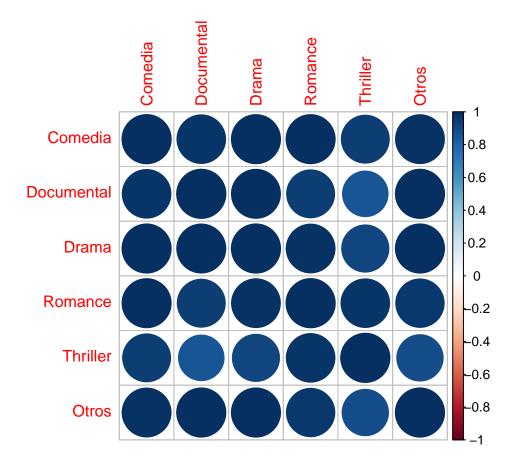
Podemos visualizar la correlación de cada género entre sí.

Correlacion_Generos

```
##
                Comedia Documental
                                       Drama
                                                Romance
                                                        Thriller
                                                                      Otros
              1.0000000
                         0.9762210 0.9963701 0.9950821 0.9484206 0.9877483
## Comedia
## Documental 0.9762210
                         1.0000000 0.9911311 0.9499475 0.8571465 0.9980900
                         0.9911311 1.0000000 0.9830379 0.9179913 0.9974474
## Drama
              0.9963701
              0.9950821
                         0.9499475 0.9830379 1.0000000 0.9751578 0.9674328
## Romance
## Thriller
              0.9484206
                         0.8571465 0.9179913 0.9751578 1.0000000 0.8873290
                         0.9980900 0.9974474 0.9674328 0.8873290 1.0000000
## Otros
              0.9877483
```

Evaluamos nuestros resultados, cada genero se correlaciona con otro, evaluando así cada combinación. Mediante un gráfico podemos visualizarlo:

corrplot(Correlacion_Generos)



Indicamos la correlación de los géneros mediante la línea de rangos, iniciando con -1 como una baja correlación o casi nula, hacia 1 como una alta correlación.

Resultados:

Primer Resultado

Con nuestro primer gráfico trabajando con factores y agrupaciones, observamos que el género "Otros" es con quien más titulos cuenta, siendo casi 250. Mientras que el género con menos titulos sería "Thriller" con menos de 50 titulos.

Segundo Resultado

Definiendo las etiquetas de categorías en la calidad de cada genero observamos que hay en total 133 titulos de buena calidad, 392 titulos de calidad media y 58 titulos de baja calidad.

Tercer Resultado

Ahora podemos encontrar las calidades precisas de cada géneros, pero también evaluarlos entre sí y responder a la interrogante "¿Cuál género es mejor?

El género con la mayor cantidad de titulos de baja calidad es "Otros" con un 3.7% y el género con menor cantidad de titulos de baja calidad es "Thriller" junto con "Romance" con un 0.68%

Además, el género con mayor cantidad de titulos de calidad regular es "Otros" con un 24,8% mientras que el titulo con menor cantidad de titulos de calidad regular es "Thriller" con 2,9%

Por ultimo encontramos que el género con mayor cantidad de titulos de buena calidad es "Otros" con 7.1% y el género con menor cantidad de titulos de buena calidad es "Romance" con 2.05%

Cuarto Resultado

Con nuestras correlaciones determinamos que el género "Otros" y el género "Thriller" tienen una correlación de 0,8, no son iguales pero si muy similares en cuanto calidades.

Mientras que los géneros "Documentales" y "Drama" tienen una correlación de 1 con el género "Otros" siendo practicamente parecidos, pero el género Otros presenta una cantidad significativamente mayor que los otros dos.

Conclusiones

Entendiendo que, el género cinmetográfico es el tema principal de una película o título y este sirve para las clasificaciones de las mismas. Logramos evaluar en criterio de sus puntajes, estipulados por el criterio de los usuarios de esta plataforma de streaming, para gozar de nuestro entretenimiento conociendo así qué género nos resulta convenientes para cada quién.

El género "Otros" abarca una amplia colección de géneros diversos no clasificados en nuestras otras 5 categorías, conociendo así que algún otro tipo de género nos puede resultar más gratificante para nuestro consumo; tal vez el Romance puede ser el género de preferencia de algún usuario, pero nuestras tablas porcentuales nos demuestran que resulta más interesante explorar otros géneros, que pudiesen convertirse en nuestro favorito.