

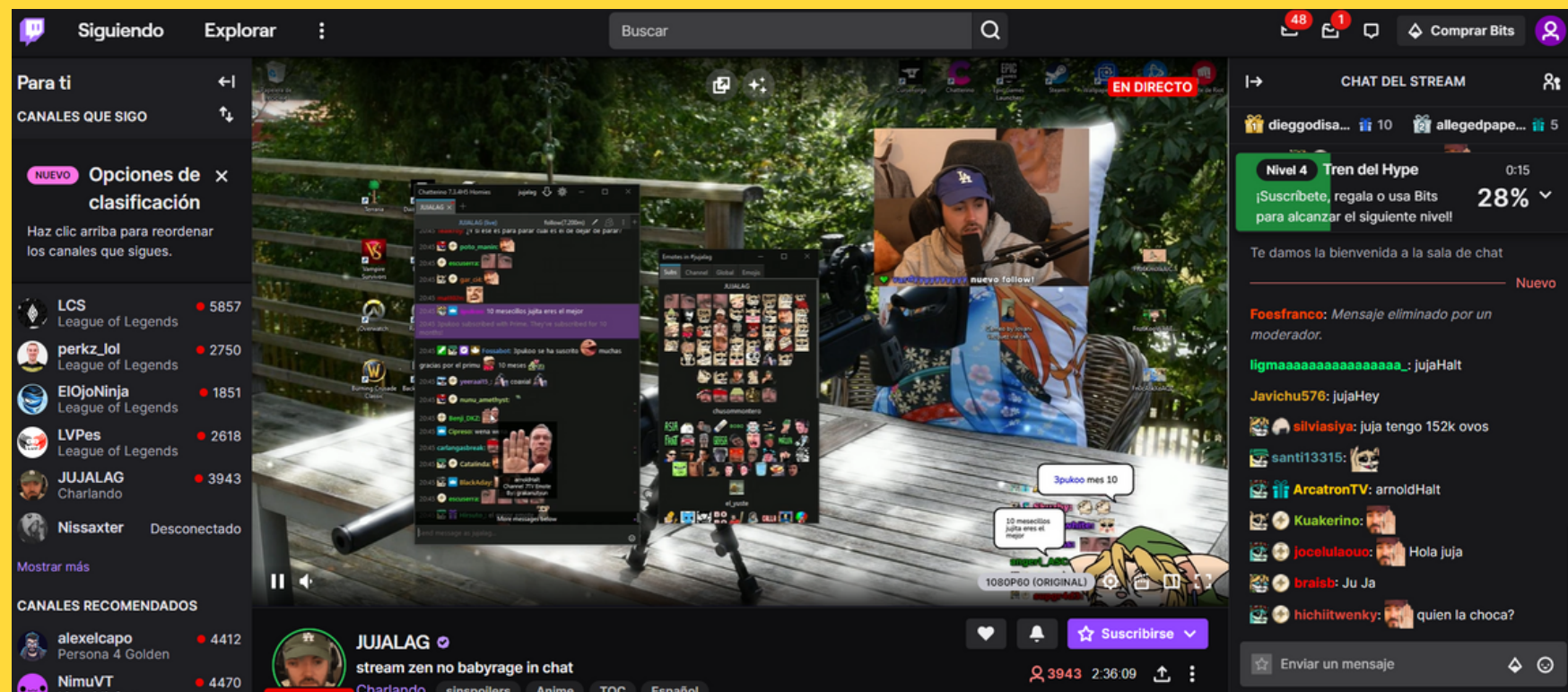
Get your head in the game



DESCUBRIENDO LA RELACION ENTRE LAS ESTADISTICAS DE TWITCH Y LAS GANANCIAS

Un análisis de regresión lineal

Introducción



¿Qué es twitch?

01

Plataforma de streaming en vivo mas popular, tiene chat en vivo, repetición y comerciales cada cierto tiempo.

Proposito

02

Cuantificar el valor en dinero (dólares) de los eventos de streaming en twitch según las visualizaciones promedios del canal.

Problema a solucionar

03

Aunque los números de rating en vivo de los streams en Twitch ya superan a los de la televisión convencional muy pocas marcas buscan campañas con canales de twitch.

Recoleccion de data

Web scraping y consultas al twitch api



Paso 1

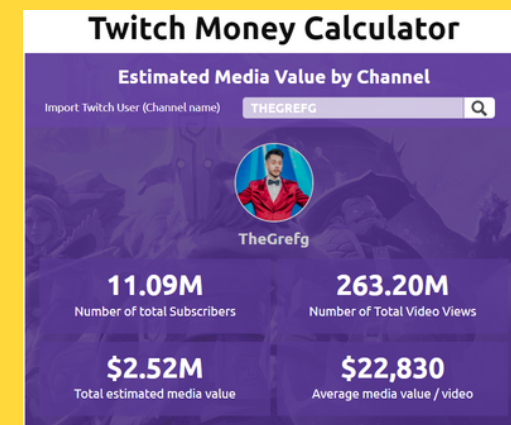
Obtener el top 1000 de los canales con mas total de vistas en los últimos 30 dias con la api de twitch

Channel peak viewership which streamed on Twitch in the past 30 days.

	Channel	Watch time (hours)	Stream time (hours)	Peak viewers	Average viewers
1	TheGrefg	10,442,997 (903.9%) ▲ 91	(33.0%) ▲ 1,745,462 (1,815.1%) ▲ 113,820 (655.0%) ▲		
2	kingsleague	8,334,526 (67.9%) ▲ 35	(35.6%) ▲ 792,033 (34.3%) ▲ 236,440 (23.8%) ▲		
3	Ibai	6,499,088 (3.3%) ▲ 94	(21.2%) ▲ 289,137 (-35.2%) ▼ 68,955 (-14.8%) ▼		

Paso 2

Hacer web scraping para obtener de la pagina “<https://sullygnome.com>” las estadísticas de visualizaciones de los canales de twitch de 30 dias atras.



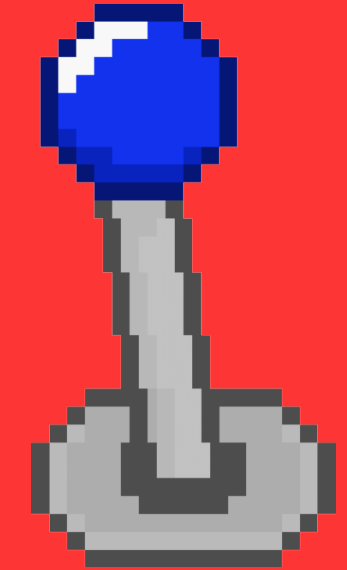
Paso 3

Con la calculadora “Twitch Money Calculator” obtenemos las ganancias de los canales en dólares por stream

4/11

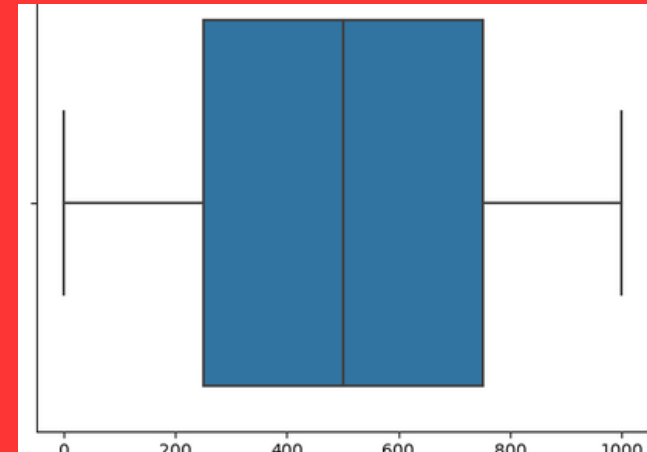
Exploración de los datos

outliers, nulos y mas



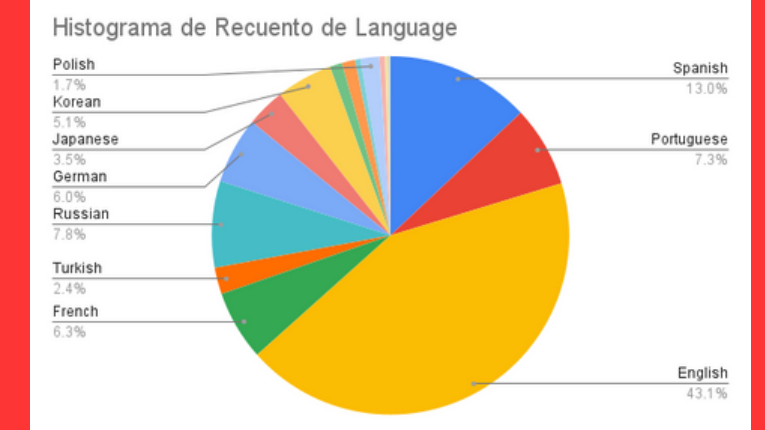
✓ Outliers

No encontramos outliers



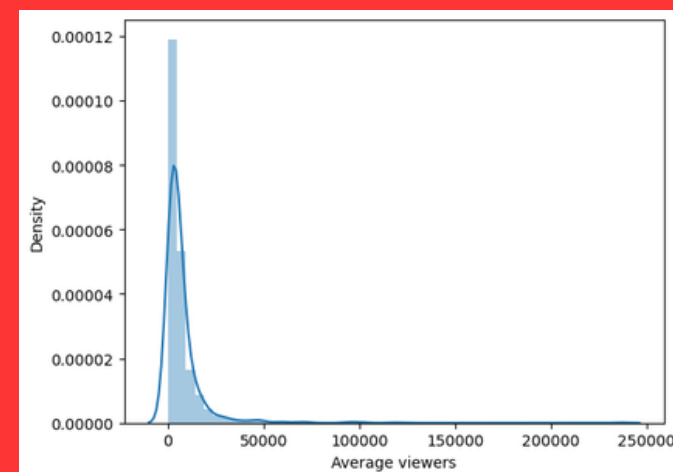
✓ Idiomas

La mayoría de los streamers más populares hablan en inglés



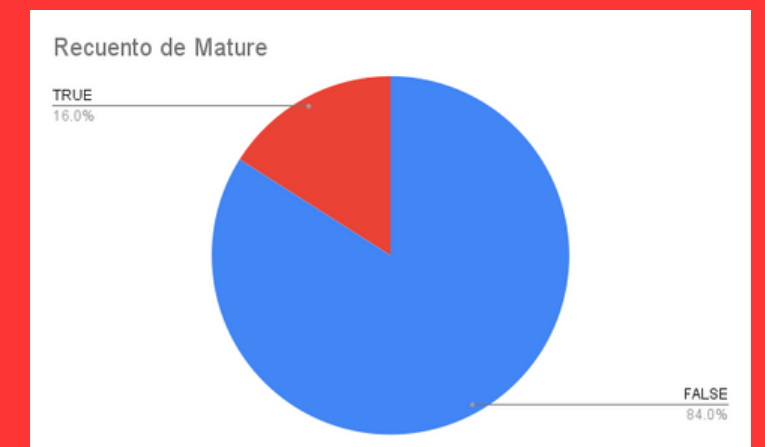
✓ distribuciones

La distribución de las variables numéricas son normales con un ligero sesgo a los valores más bajos



✓ Categóricos

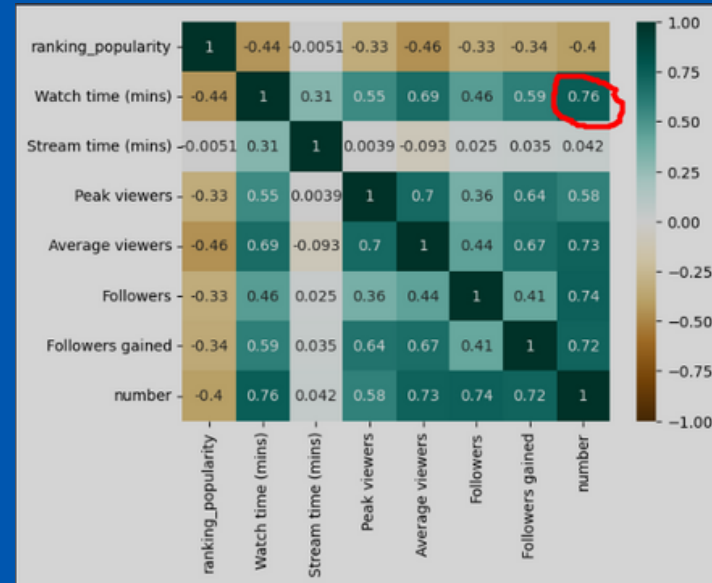
El 16 % del contenido es para mayor de edad. El 12 por ciento de los streamers no está apadronado.



Limpieza y normalización de datos

colinealidad

Eliminamos una columna por alta colinealidad



Labels categóricas

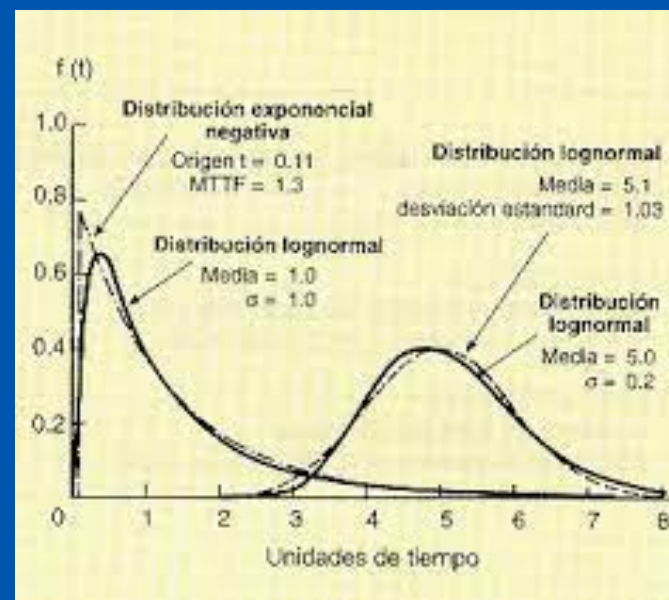
se usó la function LabelEncoder para todos los datos. Se creó un catalogo para no perder el nombre de los idiomas

	ranking_popularity	Partnered	Mature	Language	Esport_channel
0	0	1	0	14	0
1	1	1	0	14	0
2	2	1	0	14	0
3	3	1	0	12	0
4	4	1	0	14	0
...

Dutch, English, French
10, 11, 12, 13, 14, 15, 16

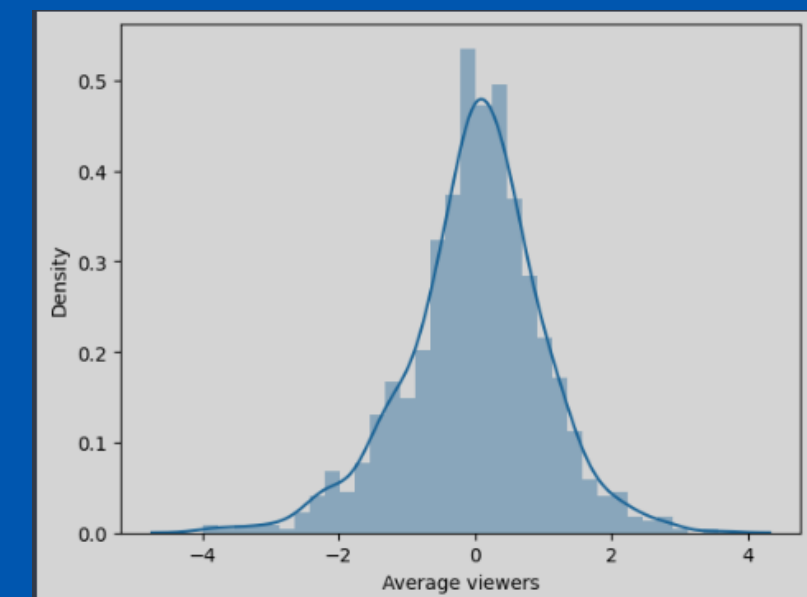
Transformaciones

Se usaron las transformación lognormal y máximos-mínimos en los datos numéricos



Estandarizar

Se estandarizó la data para obtener mejores resultados.



Evaluación del modelo de regresión lineal



R-squared = 0.67

El modelo explica una cantidad de moderada a alta de la varianza en la variable objetivo. Los predictores utilizados en el modelo tienen una fuerte relación con la variable objetivo y pueden considerarse buenos predictores.

Mean Absolute Error (MAE):

Un valor MAE de 850 indica que las predicciones del modelo están, en promedio, a 850 unidades de distancia de los valores reales. El pronóstico está a una considerable distancia de la predicción, lo que hace que el pronóstico tenga un margen de error alto.



♥ ♥ ♥ ♥ ♥
GAME OVER

Análisis de los coeficientes



```
y = 16679x_Followers gained +  
1180x_Peak viewers +  
582x_Followers + 243x_Average  
viewers + 144x_Mature + 24x_Stream  
time (mins) + 3x_ranking_popularity -  
50x_Language - 184x_Esport_channel  
- 32x_Partnered - 2166
```

01

Los followers
coeficientes muy altos

02

Los viewers muy
determinantes

03

El contenido maduro
monetiza bien

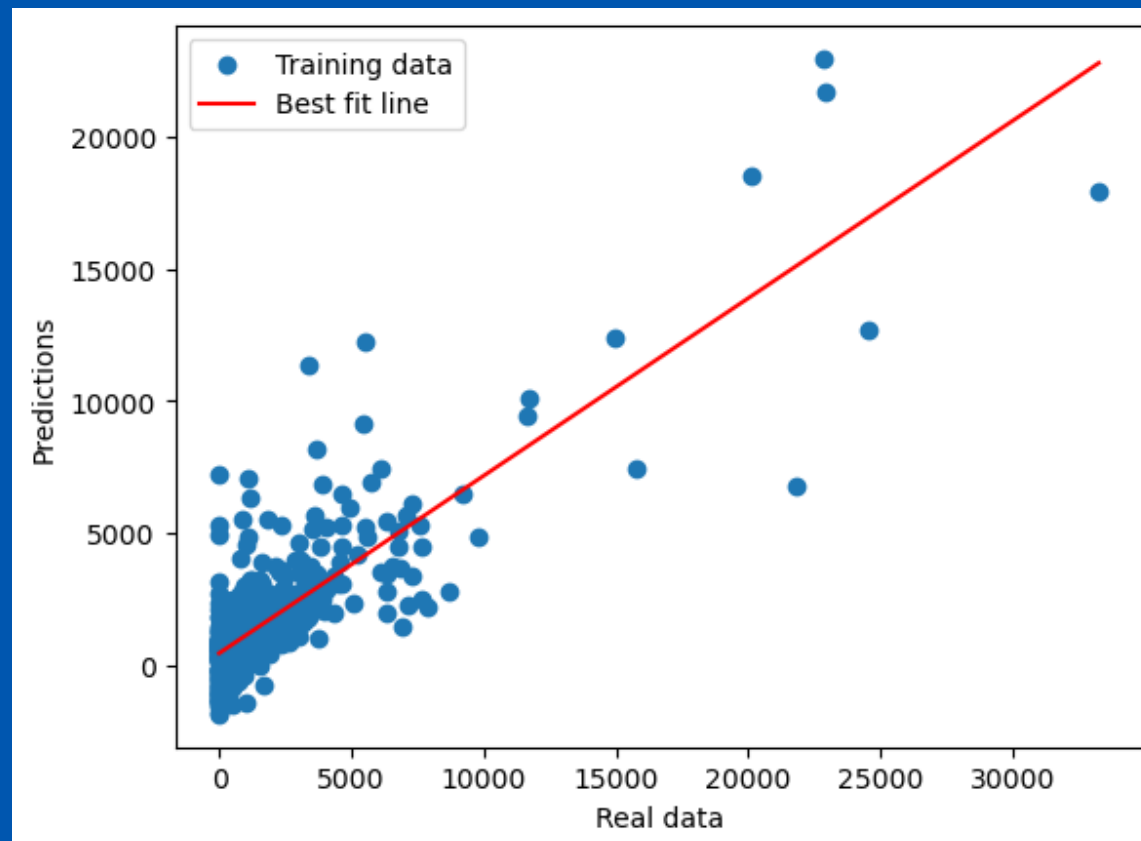
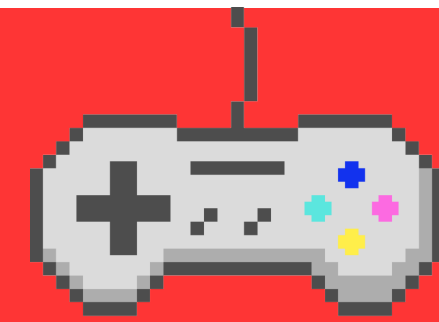
04

Los esports, los partnered
se monetizan mal

05

La constante es alta y
negativa

Pronostico



Si pronosticamos la liga mexicana de League of legends. El resultado fue 123.936 dolares por evento.



Conclusiones



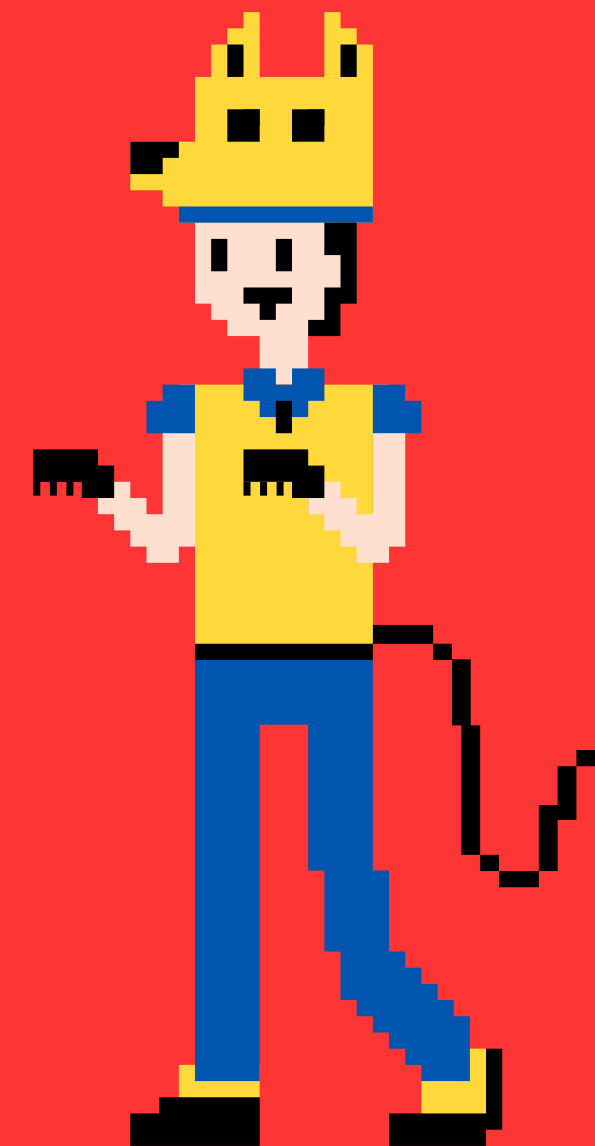
- **Las visualizaciones y los followers son los stats mas importantes para monetizar.**
- **El forecast tiene oportunidades de mejora, debido a la selección de datos.**
- **Los canales que no son de esports, que son para mayores de edad y que no son asociados generan más dinero.**



áreas potenciales para futuras investigaciones y análisis



- agregar métricas diferentes a las vistas en los datos.
- Tener una muestra más representativa de los canales.
- Investigar por qué el esports monetiza poco





THANK YOU!

I hope you learned something new!