



Análisis de la plataforma Twitch

Autor: Brian Martin Ortiz

Curso: Data Analytics

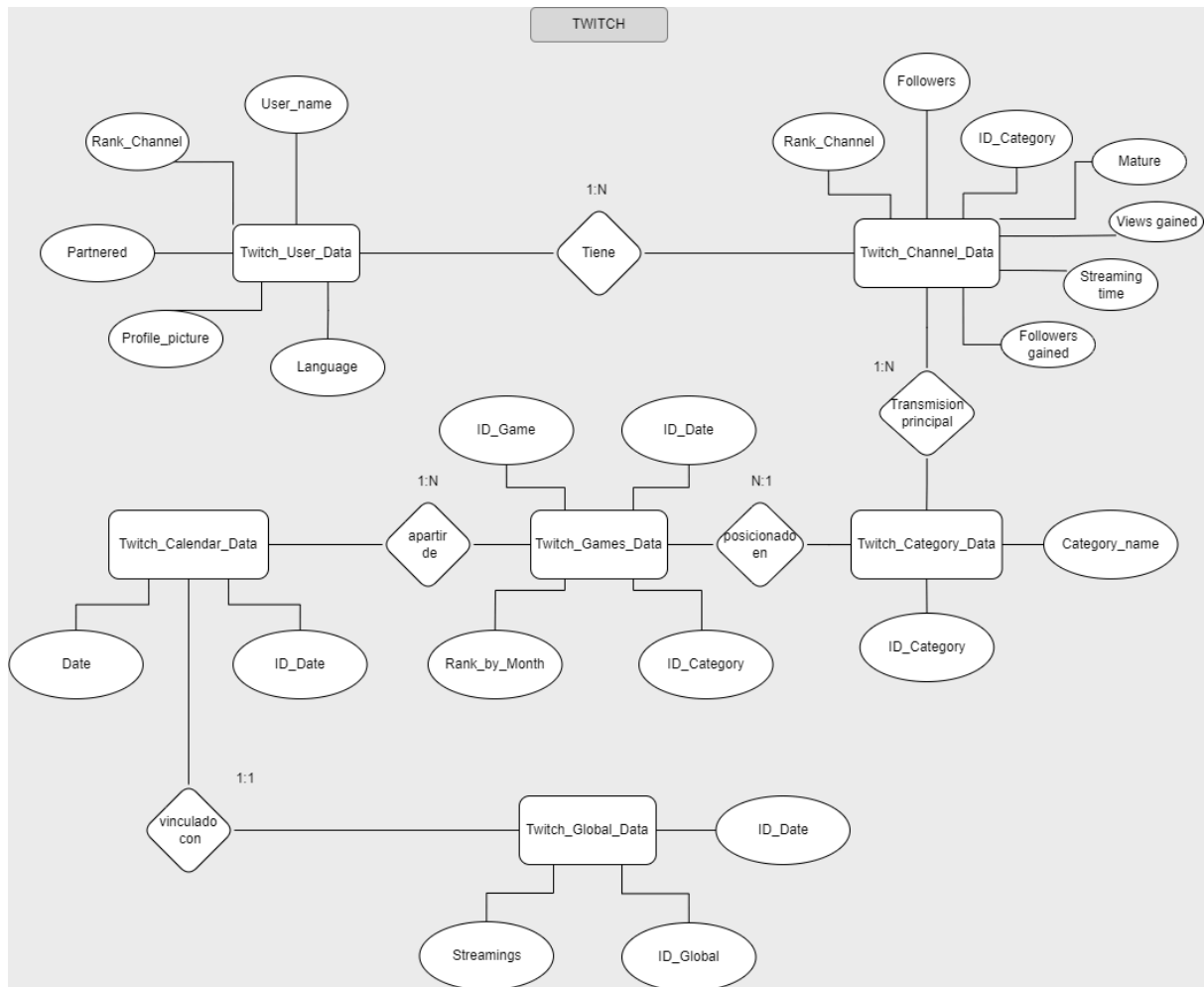
Fecha: Febrero de 2024

Descripción:

En el siguiente trabajo, analizaremos una de las plataformas más populares de la actualidad en cuanto a la transmisión en vivo de contenido digital. Desde su lanzamiento en 2008, hasta la actualidad, la plataforma de Twitch, perteneciente a Amazon, ha experimentado un gran crecimiento de espectadores, que ha revolucionado la forma en que se consumen los contenidos de entretenimiento hoy en día, un campo en donde también tiene como competencia a la famosa plataforma de YouTube.

Estudiaremos en detalle el comportamiento de los streamers y sus espectadores, prestando especial atención al tipo de contenido que retransmiten con mayor frecuencia, así como a la duración de su actividad dentro de la plataforma. Esta información puede resultar muy valiosa para los desarrolladores de videojuegos, ya que les aporta una perspectiva sobre las tendencias predominantes en determinados géneros, así como también para el diseño de estrategias de marketing y retención dentro de la página web. Por último, se propondrán algunas líneas futuras de investigación en aquellas áreas donde sería posible profundizar el análisis para obtener resultados más precisos.

Diagrama entidad - relación



El diagrama también se encuentra dentro del drive haciendo [clic aquí](#).

Listado de tablas

En este apartado, se mencionará cada una de las tablas junto a una breve descripción de las mismas y la definición de la clave primaria y foránea:

Twitch_User_Data: Contiene los nombres de usuarios registrados dentro de la plataforma junto a su lenguaje de preferencia, foto de perfil, ranking del canal y si está asociado o no.

- **PK:** User_name
- **FK:** Rank_channel

Twitch_Channel_Data: Contiene un top de 8800 canales de twitch, junto con sus respectivos datos dentro de la plataforma. (categoría principal de canal, tiempo de

transmisión y visualización total, seguidores y vistas ganadas, promedios, restricciones de edad, etc.)

- **PK:** Rank_channel
- **FK:** ID_Category

Twitch_Category_Data: Contiene la ID y los nombres de cada categoría en la plataforma.

- **PK:** ID_Category

Twitch_Games_Data: Contiene las clasificaciones mensuales, del 1 al 200, de todos los juegos transmitidos desde 2016 en adelante.

- **PK:** ID_Game
- **FK:** ID_Category
- **FK:** ID_Date

Twitch_Calendar_Data: Contiene las fechas de todos los datos recolectados desde el 2016 en adelante.

- **PK:** ID_Date

Twitch_Global_Data: Contiene los datos globales de la plataforma.

- **PK:** ID_Global
- **FK:** ID_Date

Listado de columnas por tablas

A continuación, se mencionará las columnas que posee cada tabla junto con su tipo de campo y clave:

Twitch_User_Data		
Campo	Tipo de campo	Tipo de clave
User_name	VARCHAR (255) NOT NULL	PK
Rank_channel	INT NOT NULL	FK
Partnered	BOOLEAN NOT NULL	-
Language	VARCHAR (100) NOT NULL	-
Profile_picture	VARCHAR (2000) NOT NULL	-

Twitch_Channel_Data		
Campo	Tipo de campo	Tipo de clave
Rank_channel	INT NOT NULL	PK
ID_Category	INT NOT NULL	FK
Wach_time	INT NOT NULL	-
Streaming_time	INT NOT NULL	-
Peak_viewers	INT NOT NULL	-
Avg_viewers	INT NOT NULL	-
Followers	INT NOT NULL	-
Followers_gained	INT NOT NULL	-
Views_gained	INT NOT NULL	-
Mature	BIT NOT NULL	-

Twitch_Category_Data		
Campo	Tipo de campo	Tipo de clave
ID_Category	INT NOT NULL	PK
Category_name	VARCHAR (255) NOT NULL	-

Twitch_Games_Data		
Campo	Tipo de campo	Tipo de clave
ID_Game	INT NOT NULL	PK
Rank_by_month	INT NOT NULL	-
ID_Category	INT NOT NULL	FK
ID_Date	INT NOT NULL	FK
Hours_watched	INT NOT NULL	-
Hours_streamed	INT NOT NULL	-
Peak_viewers	INT NOT NULL	-
Peak_channels	INT NOT NULL	-
Streamers	INT NOT NULL	-
Avg_viewers	INT NOT NULL	-
Avg_channels	INT NOT NULL	-
Avg_viewer_ratio	INT NOT NULL	-

Twitch_Calendar_Data		
Campo	Tipo de campo	Tipo de clave
ID_Date	INT NOT NULL	PK
Date	DATE NOT NULL	-

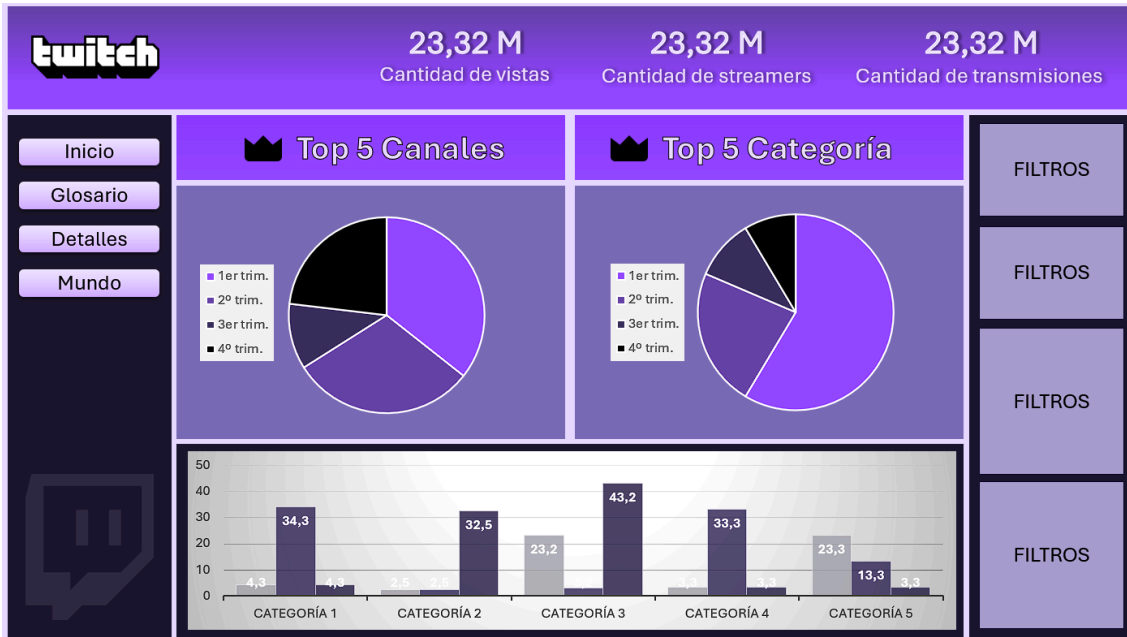
Twitch_Global_Data		
Campo	Tipo de campo	Tipo de clave
ID_Global	INT NOT NULL	PK
ID_Date	INT NOT NULL	FK
Hours_watched	BIGINT NOT NULL	-
Avg_viewers	INT NOT NULL	-
Peak_viewers	INT NOT NULL	-
Streamings	INT NOT NULL	-
Avg_channels	INT NOT NULL	-
Games_Streamed	INT NOT NULL	-
Viewer_ratio	INT NOT NULL	-

Mockup realizado en PowerPoint

Portada



Primera solapa

Medidas calculadas que se crearon

AMedidas Canales	Descripción	Expresión
categoría Preferida	Esta medida fue creada para devolver la categoría preferida de los diferentes canales.	SELECTEDVALUE(Twitch_Category_Data[Category_name], "Sin Datos")
Minutos Transmitido	Esta medida devuelve el total de minutos transmitidos por el usuario en su canal.	SELECTEDVALUE(Twitch_Channel_Data[Streaming_time (in minutes)], "Sin datos")
Minutos Visualizado	Esta medida fue creada para generar un top 5 de canales con más tiempo visto en minutos.	SELECTEDVALUE(Twitch_Channel_Data[Watch_time (in minutes)], 0)
Pico de Vistas	Esta medida fue creada para generar un top 5 de canales con más pico de visualizaciones.	SELECTEDVALUE(Twitch_Channel_Data[Peak_viewers], 0)
Porcentaje de usuarios asociados	Esta medida fue creada para devolver el porcentaje total	VAR a = COUNTROWS(FILTER(Twitch_User_Data, Twitch_User_Data[Partnered] = TRUE())) VAR b = COUNTROWS(Twitch_User_Data)

	de asociados. (Se podrá luego filtrar tanto por canal, como por lenguaje.)	return IF(a = 0, 0, IF(a > 0,DIVIDE(a,b,0)))
Promedio de vistas	Esta medida fue creada para devolver el promedio de vistas de los diferentes canales.	SELECTEDVALUE(Twitch_Channel_Data[Avg_viewers], "Sin datos")
seguidores	Esta medida fue creada para generar un top 5 de canales con más seguidores.	SELECTEDVALUE(Twitch_Channel_Data[Followers], 0)
seguidores ganados	Esta medida fue creada para generar un top 5 de canales con más seguidores ganados.	SELECTEDVALUE(Twitch_Channel_Data[Followers_gained], 0)
Total de categorías Preferidas	Esta medida devuelve el total de categorías preferidas por los streamers.	DISTINCTCOUNT(Twitch_Channel_Data[ID_Category])
Total de Seguidores	Esta medida fue creada para devolver el total de seguidores de los diferentes canales.	SELECTEDVALUE(Twitch_Channel_Data[Followers], "Sin datos")
Total de vistas	Esta medida fue creada para devolver el total de vistas de los diferentes canales.	SELECTEDVALUE(Twitch_Channel_Data[Watch_time (in minutes)], "Sin datos")
Usuarios	Esta medida fue creada para devolver el nombre del canal seleccionado.	SELECTEDVALUE(Twitch_User_Data[User_name],"Seleccione un Canal")

A Medidas Categorías	Descripción	Expresión
Max Pico de vistas	Esta medida fue creada para devolver el Pico Max de vistas logrado por las distintas categorías. Luego se procede a filtrar las mejores 5.	MAX(Twitch_Games_Data[Peak_viewers])
Total de categorías	Esta medida hace un recuento total de todas las categorías existentes.	COUNT(Twitch_Category_Data[ID_Category])
Total de Streamers en Categorías	Esta medida fue creada para devolver la cantidad de transmisores que existen en todas las categorías.	SUM(Twitch_Games_Data[Streamers])
Promedio de vista	Esta medida fue creada para devolver el valor promedio de vista de cada categoría.	SELECTEDVALUE(Twitch_Games_Data[Avg_viewers])
Horas Vista año/mes	Esta medida fue creada para devolver las horas vistas por año/mes de cada categoría.	SUM(Twitch_Games_Data[Hours_watched])
Horas transmitidas año/mes	Esta medida fue creada para devolver las horas transmitidas por año/mes de cada categoría.	SUM(Twitch_Games_Data[Hours_streamed])

A Medidas General	Descripción	Expresión
Total de Canales	Esta medida fue creada para hacer un recuento total de todos los canales existentes. Luego se utiliza para filtrar el total por categoría, lenguaje, asociados y contenido de adulto.	COUNT(Twitch_User_Data[Rank_channel])
Total de horas vistas	Esta medida fue creada para mostrar la cantidad total de horas vistas en la plataforma.	SUM(Twitch_Global_Data[Hours_watched])
Total transmitido en la plataforma	Esta medida fue creada para hacer un recuento total de todas las transmisiones hechas de la plataforma.	SUM(Twitch_Global_Data[Streamings])

Pico de visualizaciones	Esta medida fue creada para mostrar el pico de visibilidad Max que hubo por mes y año.	SUM(Twitch_Global_Data[Peak_viewers])
Pico Max de vistas	Esta medida se creó con el fin de conocer el pico Max de vistas que logro toda la plataforma.	MAX(Twitch_Global_Data[Peak_viewers])
Total juegos únicos transmitidos	Esta medida fue creada con el fin de conocer el total de juegos únicos que se transmitieron en la plataforma.	SUM(Twitch_Global_Data[Games_Streamed])

Parámetros creados

Parámetros	Descripción	Expresión
Parámetro Visual	Fue creada para la página “General” donde se puede interactuar con las diferentes visuales. Contiene 4 gráficos: las transmisiones por mes, Juegos únicos transmitidos por mes, Pico de vistas por mes y horas de vistas por mes.	Parámetro Visual = { ("Total de Transmisiones por Mes", NAMEOF('AMedidas General'[Total transmitido en la plataforma]), 0), ("Juegos únicos transmitidos por Mes", NAMEOF('AMedidas General'[Total de juego únicos transmitidos]), 1), ("Pico de vistas por Mes", NAMEOF('AMedidas General'[Pico de visualizaciones]), 2), ("Horas de vistas por Mes", NAMEOF('AMedidas General'[Total de horas vistas]), 3) }
Parámetro Canales	Fue creada para la página “General” donde se puede interactuar con 2 visuales: Por lenguaje y Por categoría. De esta manera se puede observar la cantidad de canales que existen en cada una por medio de un “Tooltip”.	Parámetro Canales = { ("Por Lenguaje", NAMEOF('Twitch_User_Data'[Language]), 0), ("Por Categoría", NAMEOF('Twitch_Category_Data'[Category_name]), 1) }

Transformaciones de datos

Twitch_Calendar_Data:

```
let
  Origen = Excel.Workbook(File.Contents(TwitchBD),null,true),
  Twitch_Calendar_Data_Sheet = Origen[Item="Twitch_Calendar_Data",Kind="Sheet"][Data],
  #"Encabezados promovidos" = Table.PromoteHeaders(Twitch_Calendar_Data_Sheet,
[PromoteAllScalars=true]),
  #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_Date",
Int64.Type}, {"Date", type date}, {"Column3", type any}}),
  #"Columnas quitadas" = Table.RemoveColumns(#"Tipo cambiado",{"Column3"})
in
  #"Columnas quitadas"
```

Twitch_Games_Data:

```
let
  Origen = Excel.Workbook(File.Contents("C:\Users\brian\Desktop\CODERHOUSE\Carrera - Data
Analytics\Data Analytics\MI_PF\Segunda_Entrega_PF\AnalizandoTwitch_BrianOrtiz.xlsx"), null,
true),
  Twitch_Games_Data_Sheet = Origen[Item="Twitch_Games_Data",Kind="Sheet"][Data],
  #"Encabezados promovidos" = Table.PromoteHeaders(Twitch_Games_Data_Sheet,
[PromoteAllScalars=true]),
  #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_Game",
Int64.Type}, {"Rank_by_month", Int64.Type}, {"ID_Category", Int64.Type}, {"ID_Date",
Int64.Type}, {"Hours_watched", Int64.Type}, {"Hours_streamed", Int64.Type},
{"Peak_viewers", Int64.Type}, {"Peak_channels", Int64.Type}, {"Streamers", Int64.Type},
{"Avg_viewers", Int64.Type}, {"Avg_channels", Int64.Type}, {"Avg_viewer_ratio",
Int64.Type}, {"Column13", type text}}),
  #"Columnas quitadas" = Table.RemoveColumns(#"Tipo cambiado",{"Column13"})
in
  #"Columnas quitadas"
```

Twitch_Global_Data:

```
let
  Origen = Excel.Workbook(File.Contents("C:\Users\brian\Desktop\CODERHOUSE\Carrera - Data
Analytics\Data Analytics\MI_PF\Segunda_Entrega_PF\AnalizandoTwitch_BrianOrtiz.xlsx"), null,
true),
  Twitch_Global_Data_Sheet = Origen[Item="Twitch_Global_Data",Kind="Sheet"][Data],
  #"Encabezados promovidos" = Table.PromoteHeaders(Twitch_Global_Data_Sheet,
[PromoteAllScalars=true]),
  #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_Global",
Int64.Type}, {"ID_Date", Int64.Type}, {"Hours_watched", Int64.Type}, {"Avg_viewers",
Int64.Type}, {"Peak_viewers", Int64.Type}, {"Streamings", Int64.Type}, {"Avg_channels",
Int64.Type}, {"Games_streamed", Int64.Type}, {"Viewer_ratio", Int64.Type}})
in
  #"Tipo cambiado"
```

Twitch_User_Data:

```
let
  Origen = Excel.Workbook(File.Contents("C:\Users\brian\Desktop\CODERHOUSE\Carrera - Data Analytics\Data Analytics\MI_PF\Segunda_Entrega_PF\AnalizandoTwitch_BrianOrtiz.xlsx"), null, true),
  Twitch_User_Data_Sheet = Origen[[Item="Twitch_User_Data",Kind="Sheet"]][Data],
  #"Encabezados promovidos" = Table.PromoteHeaders(Twitch_User_Data_Sheet, [PromoteAllScalars=true]),
  #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"User_name", type text}, {"Rank_channel", Int64.Type}, {"Partnered", type logical}, {"Language", type text}, {"Profile_picture", type text}})
in
  #"Tipo cambiado"
```

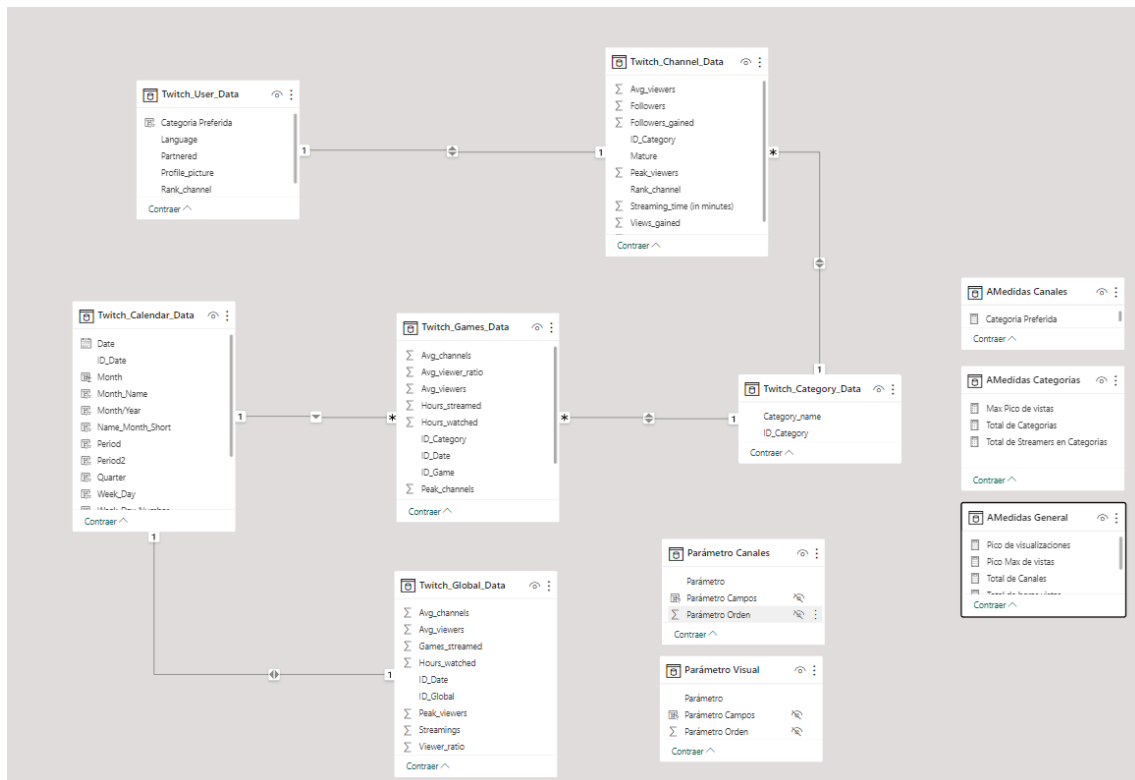
Twitch_Channel_Data:

```
let
  Origen = Excel.Workbook(File.Contents("C:\Users\brian\Desktop\CODERHOUSE\Carrera - Data Analytics\Data Analytics\MI_PF\Segunda_Entrega_PF\AnalizandoTwitch_BrianOrtiz.xlsx"), null, true),
  Twitch_Channel_Data_Sheet = Origen[[Item="Twitch_Channel_Data",Kind="Sheet"]][Data],
  #"Encabezados promovidos" = Table.PromoteHeaders(Twitch_Channel_Data_Sheet, [PromoteAllScalars=true]),
  #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"Rank_channel", Int64.Type}, {"ID_Category", Int64.Type}, {"Watch_time (in minutes)", Int64.Type}, {"Streaming_time (in minutes)", Int64.Type}, {"Peak_viewers", Int64.Type}, {"Avg_viewers", Int64.Type}, {"Followers", Int64.Type}, {"Followers_gained", Int64.Type}, {"Views_gained", Int64.Type}, {"Mature", type logical}})
in
  #"Tipo cambiado"
```

Twitch_Category_Data:

```
let
  Origen = Excel.Workbook(File.Contents("C:\Users\brian\Desktop\CODERHOUSE\Carrera - Data Analytics\Data Analytics\MI_PF\Segunda_Entrega_PF\AnalizandoTwitch_BrianOrtiz.xlsx"), null, true),
  Twitch_Category_Data_Sheet = Origen[[Item="Twitch_Category_Data",Kind="Sheet"]][Data],
  #"Encabezados promovidos" = Table.PromoteHeaders(Twitch_Category_Data_Sheet, [PromoteAllScalars=true]),
  #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_Category", Int64.Type}, {"Category_name", type text}})
in
  #"Tipo cambiado"
```

DER en Power BI



Se agregaron varias columnas calculadas en la tabla "Twitch_Calendar_Data" para hacer más completa la información del calendario.

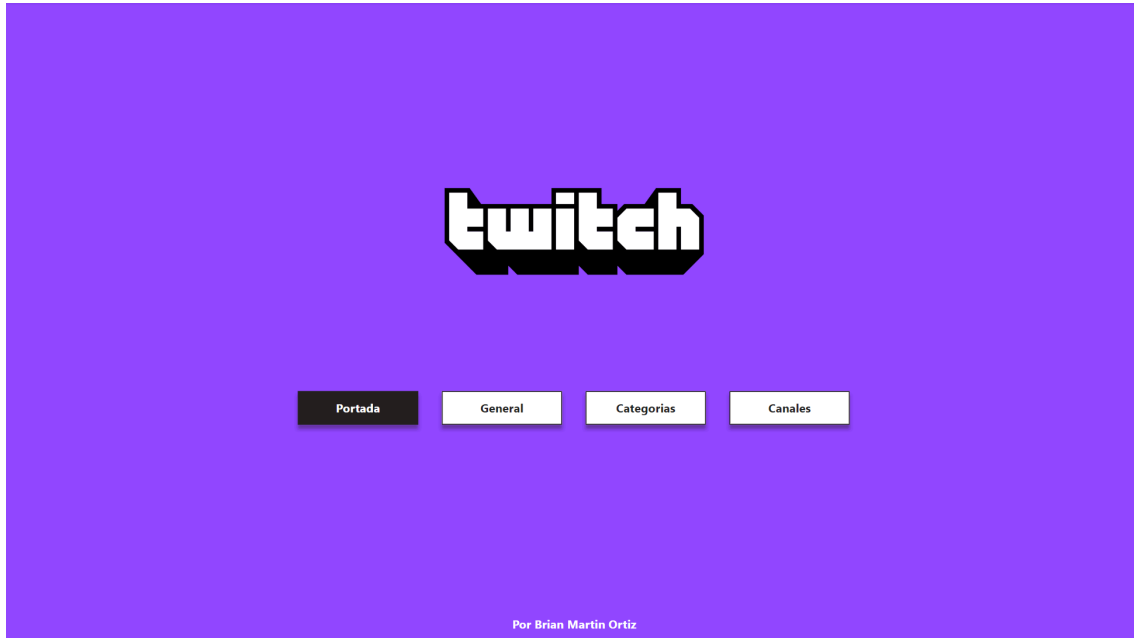
```
Month = MONTH(Twitch_Calendar_Data[Date])
Month_Name = FORMAT (Twitch_Calendar_Data[Date], "mmmm" )
Month/Year = FORMAT(Twitch_Calendar_Data[Date], "mm/yy")
Name_Month_Short = FORMAT (Twitch_Calendar_Data[Date], "mmm" )
Period = FORMAT(Twitch_Calendar_Data[Date], "YYYY/MM" )
Period2 = FORMAT (Twitch_Calendar_Data[Date], "YYYY/mmm" )
Quarter = FORMAT (Twitch_Calendar_Data[Date], "Q" )
Week_Day = FORMAT (Twitch_Calendar_Data[Date], "dddd" )
Week_Day_Number = WEEKDAY (Twitch_Calendar_Data[Date])
Week_Day_Short = FORMAT (Twitch_Calendar_Data[Date], "ddd" )
Year = YEAR(Twitch_Calendar_Data[Date])
Year/Month = FORMAT(Twitch_Calendar_Data[Date], "yy-mm")
Year/Quarter = FORMAT ( Twitch_Calendar_Data[Date], "YYYY" ) & "/Q" & FORMAT (
Twitch_Calendar_Data[Date], "Q" )
```

Se agrego una columna a la tabla "Twitch_User_Data" para identificar mejor la categoría preferida del canal.

```
Categoria Preferida = RELATED(Twitch_Category_Data[Category_name])
```

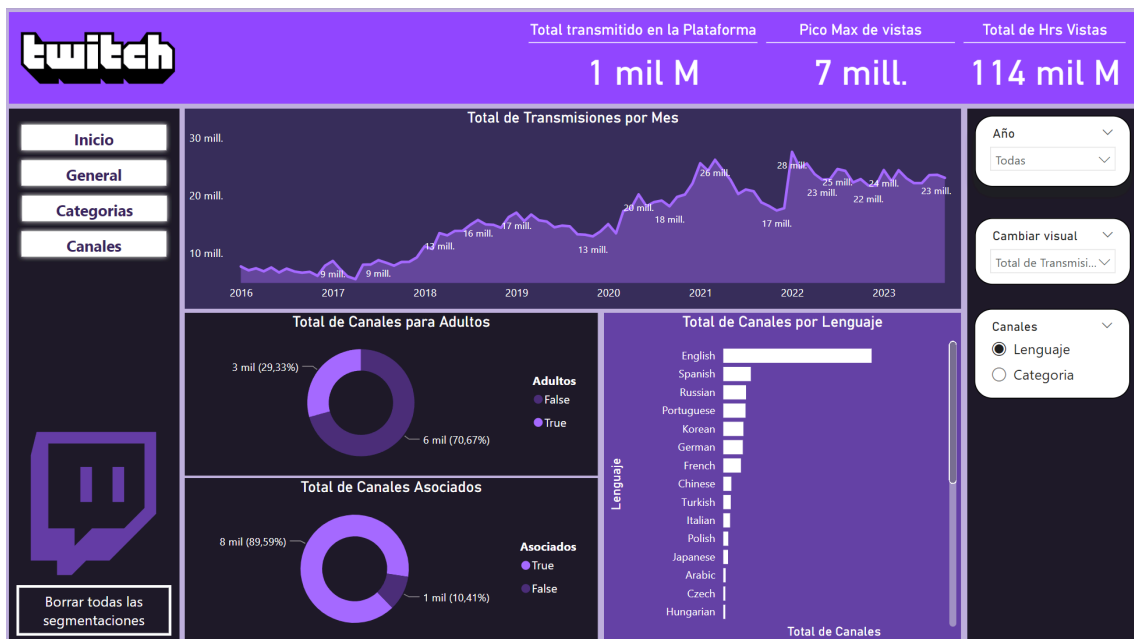
Tablero en Power BI

Portada



Dentro de la portada se crean botones a modo de que el usuario pueda seleccionar lo que desea visualizar de entrada al tablero.

General

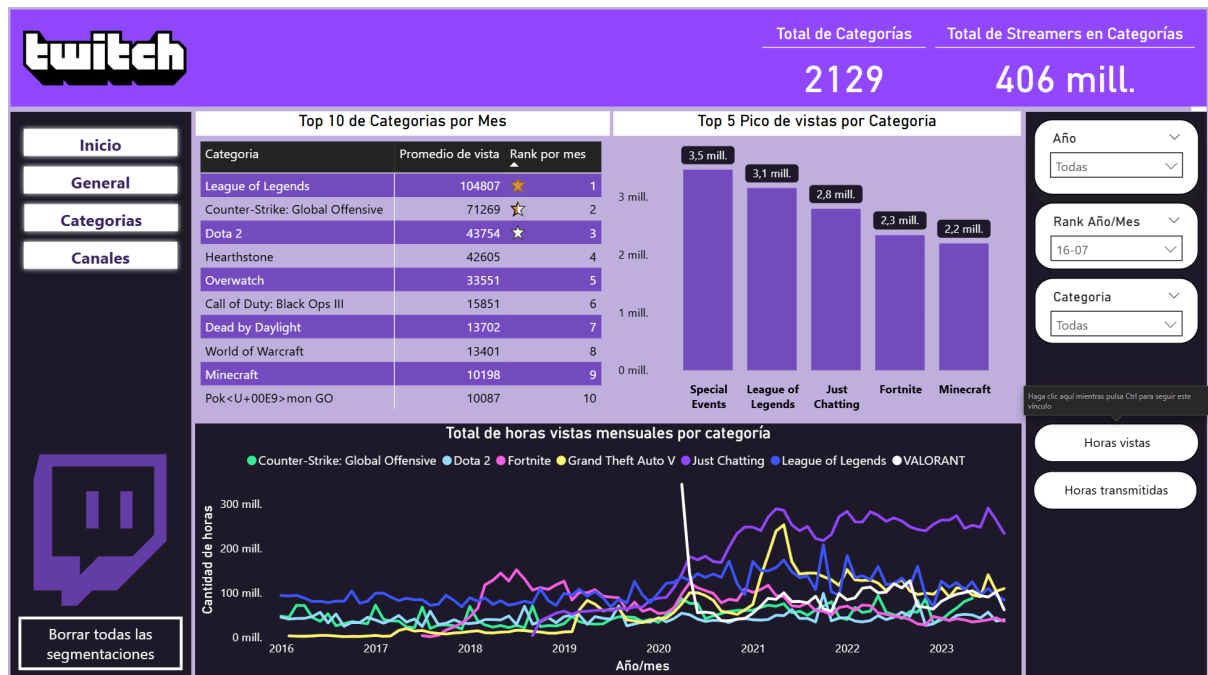


En la primera hoja el objetivo será proporcionar una visión general de los datos acumulados por la plataforma a lo largo del tiempo. Se incluirá un gráfico de área en la parte superior, el

cual permitirá al usuario alternar entre diferentes visualizaciones para mostrar el total de transmisiones, la cantidad de juegos únicos y las horas visualizadas de cada mes del año.

Además, se presentará información detallada sobre el número de canales registrados, distribuidos según su idioma o categoría preferida, así como el porcentaje de asociados y adultos que se encuentran hasta la fecha considerada.

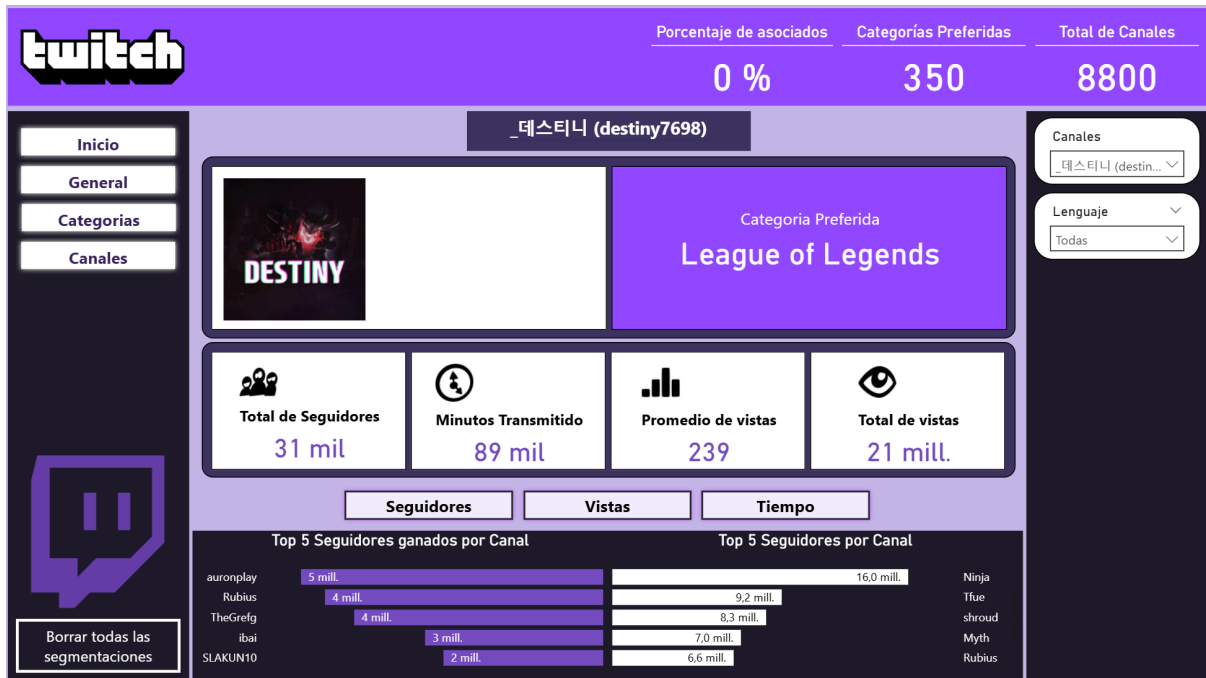
Categorías



En la siguiente hoja, hallarás información relevante sobre las diversas categorías presentes en la plataforma. Los usuarios podrán interactuar con el ‘Top 10’ de categorías de cada mes o elegir una categoría específica para visualizar las horas vistas y transmitidas correspondientes.

También encontrarán visuales que nos muestran el ‘Top 5’ de categorías con el mayor pico de vistas logrado o el total de categorías existentes hasta el momento.

Canales



En la siguiente hoja, se presentará información detallada de cada canal de forma individual. Se analizarán sus números a lo largo del tiempo y se identificará qué contenido logró atraer más al público espectador.

En la parte inferior del tablero podremos interactuar con el 'Top5' de diferentes características, ya sea de seguidores, vistas o tiempos transmitidos dentro de la plataforma.

Futuras Líneas

A partir de los resultados obtenidos, se proponen varias líneas futuras de trabajo. En primer lugar, sería útil incorporar datos más recientes o ampliar la muestra para analizar si las tendencias observadas se mantienen en el tiempo.

Sería especialmente interesante prestar atención a las categorías que destacaron en los últimos años analizados y observar si se mantuvieron vigentes de espectadores hasta la actualidad. Como por ejemplo los géneros de videojuegos como 'League of Legends' o 'Fornite', que por muchos años lograron posicionarse en el top 10 de los juegos más visualizados de la plataforma.

Por otro lado, sería relevante dar seguimiento a la evolución en la cantidad de asociados a Twitch, ya que una disminución sostenida podría indicar una pérdida de interés por parte de los creadores de contenido y, en consecuencia, de los usuarios.

Finalmente, sería valioso comparar estos resultados con los de otras plataformas de streaming, para evaluar si los patrones observados son específicos de Twitch o si pueden generalizarse a todo el ecosistema digital de contenidos en vivo.