



VENTAS Y RATING DE VIDEOJUEGOS ENTRE 1990 Y 2016

Para la inclusion en una plataforma de videojuegos online

Alumnos: Néstor Diaz, Nahuel Facello, Miguel Silguero

COMISION 24425



Contenido

Breve introduccion	2
Objetivo	2
Descripcion de la tematica	2
Diagrama entidad-relacion	3
Listado de tablas	4
Listado de columnas y tipo de datos	4
Modelo entidad relacion final	6
MockUp	7
Workflow del proyecto	8
Fuente de base de datos	8
Definicion de tablas	8
Base de datos del proyecto final	9
Edición de la base de datos	9
Carga de la base de datos en Power BI	10
Transformación de datos	10
Análisis funcional del tablero	10
Botones	10
Tooltip y bookmark	11
Presentacion final.....	11
Solapa de Dashboard:.....	12
Solapa Ventas y rating de videojuegos:	13
Solapa Publisher y desarrolladores:	16
Solapa consumo de videojuegos:.....	18
Solapa analisis de ventas:.....	20
Tecnologia utilizada:	22
Conclusion	23

Breve introduccion

Como equipo nos propusimos analizar qué tipos de plataformas son las más utilizadas para jugar videojuegos y ver cuáles son los géneros más valorados y vendidos. Para eso utilizamos el siguiente [dataset](#) que proviene de Kaggle, el cual se basa en obtener información de páginas como Vgchartz y de Metacritic las cuales son muy conocidas y especializadas en este rubro.

Objetivo

Determinar juegos más vendidos y mejores rankeados en cada región para ofrecer una plataforma de videojuegos en línea. Ver la evolución de las ventas y los años de los mismos para así tambien poder determinar el rango etario al que iría dirigido.

Descripcion de la tematica

Descriptivo

Analisis de la base de datos de juegos lanzados entre 1980 y 2016 para identificar aquellos juegos más vendidos y mejores rankeados en el periodo de tiempo descripto.

Diagnostico

Implementación de dashboards dinámicos realizados en Power BI extrayendo los datos de un dataset de videojuegos con el objetivo de observar por año, genero, ranking y ventas.

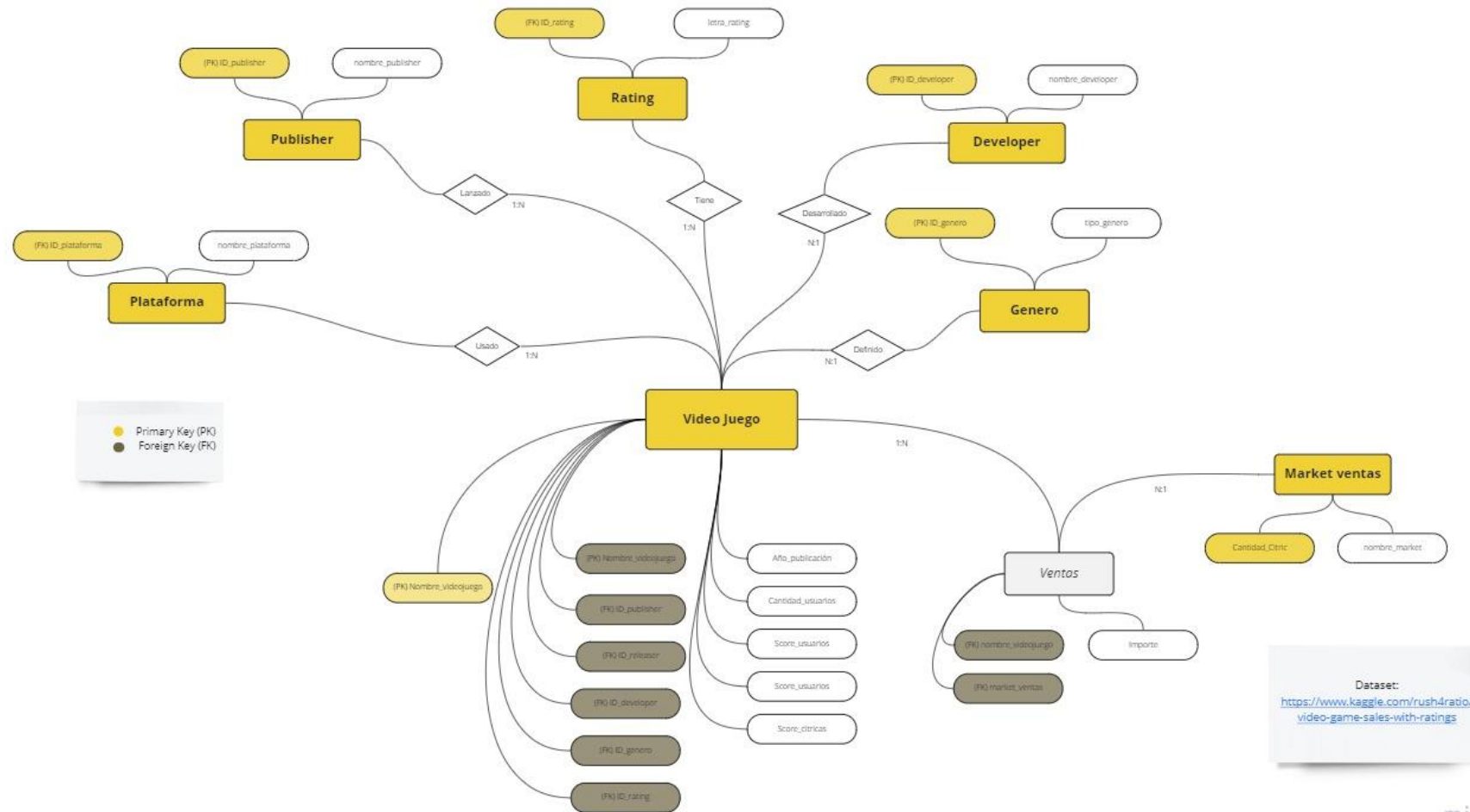
Predictivo

Modelo predictivo de aquellos juegos mejor valorados en tiempo real en funcion de los gustos de los jugadores más frecuentes. Buscamos saber cuál es la tendencia de los jugadores al elegir plataformas y géneros.

Prescriptivo

Inclusión en la plataforma de aquellos juegos mejor valorados que basado en el modelo predictivo, sean utilizados por los jugadores en línea.

Diagrama entidad-relacion



miro

[Link al diagrama](#)

Listado de tablas

Clave/ Tabla	Tabla 1 <i>VideoJuegos</i>	Tabla 2 <i>Plataforma</i>	Tabla 3 <i>Publisher</i>	Tabla 4 <i>Rating</i>	Tabla 5 <i>Developer</i>	Tabla 6 <i>Genero</i>	Tabla 7 <i>Market ventas</i>	Tabla 8 <i>Ventas</i>
Clave primaria	ID_Videojuego	ID_Plataforma	ID_Publisher	ID_Rating	ID_Developer	ID_Genero	ID_Market_ventas	
Clave foranea	ID_Plataforma; ID_Publisher; ID_Developer; ID_Genero; ID_Rating							ID_Videojuego; ID_market_ventas

Listado de columnas y tipo de datos

Tabla 1: Video Juegos

TIPO DE CLAVES	CAMPO	TIPO DE CAMPO
PK	ID_Videojuego	Int (n)
-	Nombre_videojuego	Varchar (n)
FK	ID_Plataforma	Int (n)
FK	ID_Publisher	Int (n)
FK	ID_Developer	Int (n)
FK	ID_Genero	Int (n)
FK	ID_Rating	Int (n)
-	Año publicacion	Int (n)
-	Cantidad_usuarios	Int (n)
-	Cantidad_criticas	Int (n)
-	Score_usuarios	Decimal
-	Score_criticas	Decimal

Tabla 2: Plataforma

TIPO DE CLAVES	CAMPO	TIPO DE CAMPO
PK	ID_Plataforma	Int (n)
-	Nombre_plataforma	Varchar (n)

Tabla 3: Publisher

TIPO DE CLAVES	CAMPO	TIPO DE CAMPO
PK	ID_Publisher	Int (n)
-	Nombre_publisher	Varchar (n)

Tabla 4: Rating

TIPO DE CLAVES	CAMPO	TIPO DE CAMPO
PK	ID_Rating	Int (n)
-	Letra_Rating	Varchar (n)
-	Descripcion_rating	Varchar (n)

Tabla 5: Developer

TIPO DE CLAVES	CAMPO	TIPO DE CAMPO
PK	ID_Developer	Int (n)
-	Nombre_Developer	Varchar (n)

Tabla 6: Genero

TIPO DE CLAVES	CAMPO	TIPO DE CAMPO
PK	ID_Genero	Int (n)
-	Tipo_Genero	Varchar (n)

Tabla 7: Market Ventas

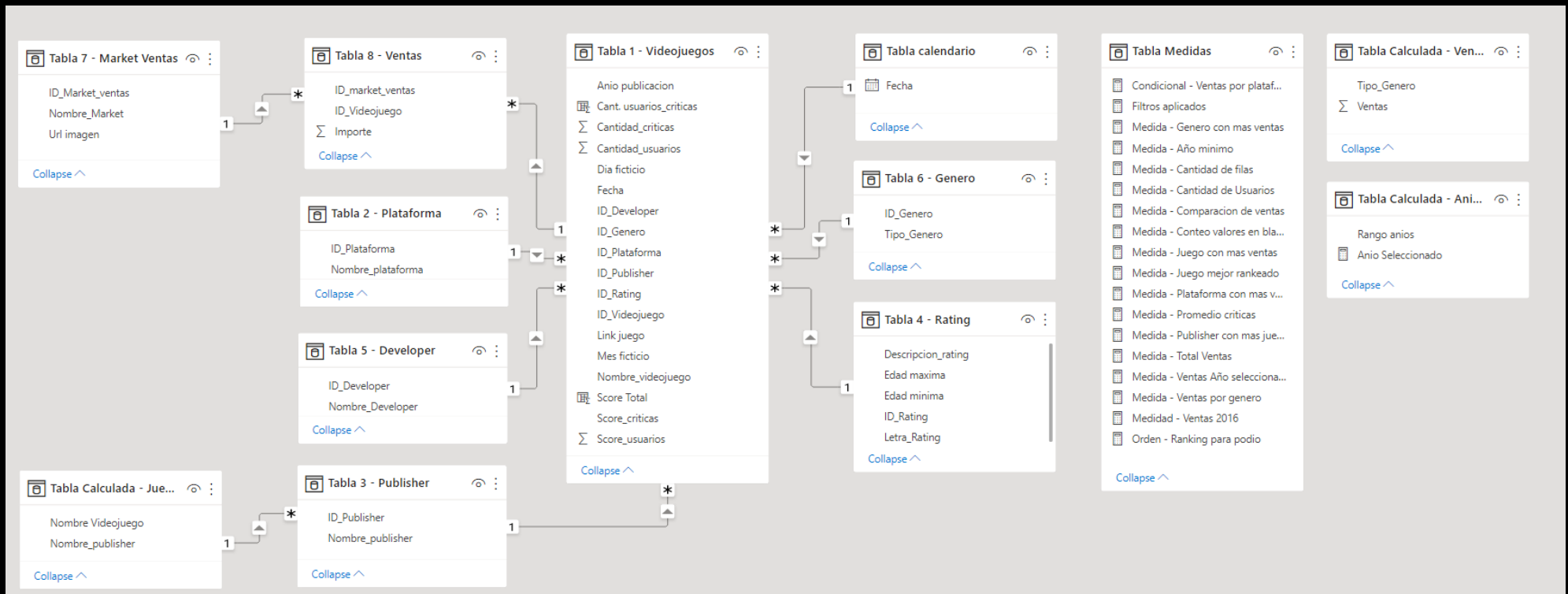
TIPO DE CLAVES	CAMPO	TIPO DE CAMPO
PK	ID_Market_ventas	Int (n)
-	Nombre_Market	Varchar (n)

Tabla 8: Ventas

TIPO DE CLAVES	CAMPO	TIPO DE CAMPO
FK	ID_Videojuegos	Int (n)
FK	ID_Market	Int (n)
-	Importe	Decimal

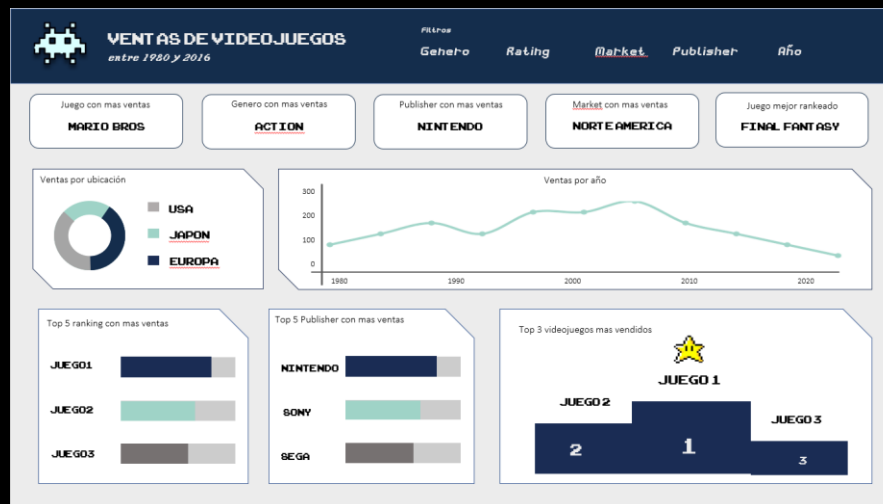
Al modelo final se le han agregado columnas relacionadas a las URL de las imágenes utilizadas en los tooltip quedando de la siguiente forma:

Modelo entidad relacion final

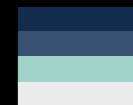


Todas las relaciones pudieron ser creadas automáticamente por el inteligencia de power BI. En ella se incluyen algunas tablas creadas y la tabla de medidas que contiene todas las utilizadas para las distintas tarjetas e indicadores.

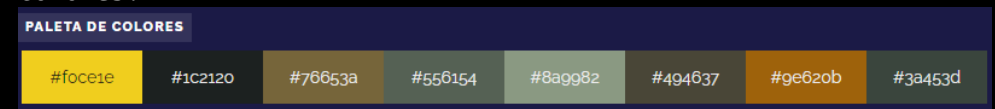
MockUp



La paleta de colores originalmente elegida para el mockup fue la siguiente:



Avanzando en la cursada ya con informacion en el dashboard de Power BI se eligio una foto como portada de la cual se tomaron los colores utilizando la siguiente página web: <https://pinetools.com/es/obtener-colores-imagen>. De la imagen se obtuvo la siguiente paleta de 8 colores:



Workflow del proyecto

Diciembre 13, 2021:

Fuente de base de datos

Como parte del proyecto se decidió seleccionar un [dataset](#) de ventas de videojuegos entre los años 1980 y 2016 distinguidos entre 4 categorías relacionadas a la ubicación: Norte América, Japón, Europa y otros. El Dataset posee una sola tabla en formato csv.

Diciembre 20, 2021:

Definicion de tablas

Utilizando Excel a traves del proceso texto a columnas, utilizando la coma como separador, se generaron las columnas correspondientes a la base. Sobre cada columna realizaron del dataset se realizaron cambios para que este sea más ameno a la lectura y el uso. Las actualizaciones se pueden ver detalladas en la columna “Actualización” debajo:

Columna	Actualización
Name	Renombrada como: Nombre_videojuego
Platform	Renombrada como: Nombre_plataforma
Year_of_Release	Renombrada como: Año publicacion
Genre	Renombrada como: Tipo_Genero
Publisher	Renombrada como: Nombre_publisher
NA_Sales	Los nombres de las 4 columnas se utilizaron para generar un campo nuevo llamado market_ventas.
EU_Sales	
JP_Sales	
Other_Sales	
Global_Sales	Esta columna se elimino siendo que es un campo generado por la sumatoria de: NA_Sales, EU_Sales, JP_Sales y Other:Sales
Critic_Score	Renombrada como: Score_criticas
Critic_Count	Renombrada como: Cantidad_criticas
User_Score	Renombrada como: Score_usuarios
User_Count	Renombrada como: Cantidad_usuarios
Developer	Renombrada como: Nombre_Developer
Rating	Renombrada como: Letra_Rating

Diciembre 27, 2021

Base de datos del proyecto final

Con fines didácticos se generaron 8 tablas distintas. Cada tabla contiene los valores atomizados de las siguientes columnas: Nombre_videojuego, Nombre_plataforma, Tipo_Genero, Nombre_publisher, Nombre_Developer, Letra_Rating.

[dataset](#)

Siendo que la relacion entre el importe de ventas por market era de muchos a muchos, se creó una tabla llamada Ventas que contiene dos foreign keys relacionadas a nombre_videojuego y market_ventas con el importe de ventas por market y videojuegos.

Se generaron primary keys autoincrementales para cada valor único de las columnas descriptas arriba: Tabla Nombre Primary key creadas Columnas que incluye del master Foreign keys que incluye:

Tabla	Nombre	Primary key creadas	Columnas que incluye del master	Foreign keys que incluye
Tabla1	VideoJuegos	ID_Videojuego	Nombre_videojuego; Anio publicacion; Score_criticas; Cantidad_criticas; Score_usuarios; Cantidad_usuarios	ID_Plataforma; ID_Publisher; ID_Rating; ID_Developer; ID_Genero
Tabla2	Plataforma	ID_Plataforma	Nombre_plataforma	
Tabla3	Publisher	ID_Publisher	Nombre_publisher	
Tabla4	Rating	ID_Rating	Letra_Rating	
Tabla5	Developer	ID_Developer	Nombre_Developer	
Tabla6	Genero	ID_Genero	Tipo_Genero	
Tabla7	Market Ventas	ID_Market_ventas	NA_Sales; EU_Sales; JP_Sales; Other_Sales	
Tabla8	Ventas			ID_Videojuego; ID_Market_ventas

Enero 4, 2022

Edición de la base de datos

En la tabla Rating se agregó una columna llamada Descripcion_rating para entender mejor el significado de cada letra. Se identificaron fechas de videojuegos que no correspondían y se actualizaron tomando como fuente de datos páginas del desarrollador del videojuego.

Enero 19,2022

Carga de la base de datos en Power BI

Al momento de cargar la base de datos en Power Bi se debieron eliminar filas en blanco para que el intelligence pueda generar las relaciones entre las tablas por medio de las foreign y primary key.

Enero 24,2022

Transformación de datos

Siendo que nuestro dataset no incluía fechas, solo años. A partir de una copia de la tabla videojuegos que contenía el año se crearon dos variables random entre 1 y 28 para una columna de días y entre 1 y 12 para una columna de mes. El resto de las columnas fue eliminado y se conservaron los datos únicos para conseguir una tabla con jerarquía de fechas. Adicional, se incorporó un parametro con el URL de la base de datos.

Febrero 7,2022

Análisis funcional del tablero

Se agrego la solapa de portada con un fondo que hace alusión al motivo de nuestro análisis, se agregaron el titulo y subtítulo y además se incluyó 4 botones para navegar entre las solapas siguientes.

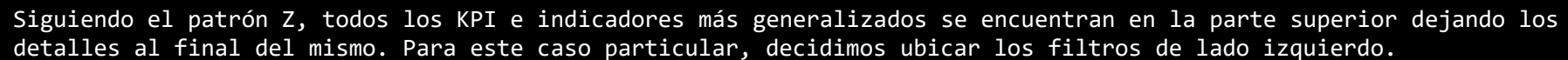
Febrero 9,2022

Botones

Se agregaron botones de Inicio, avanzar y retroceder pagina.

Tooltip y bookmark

Presentacion final

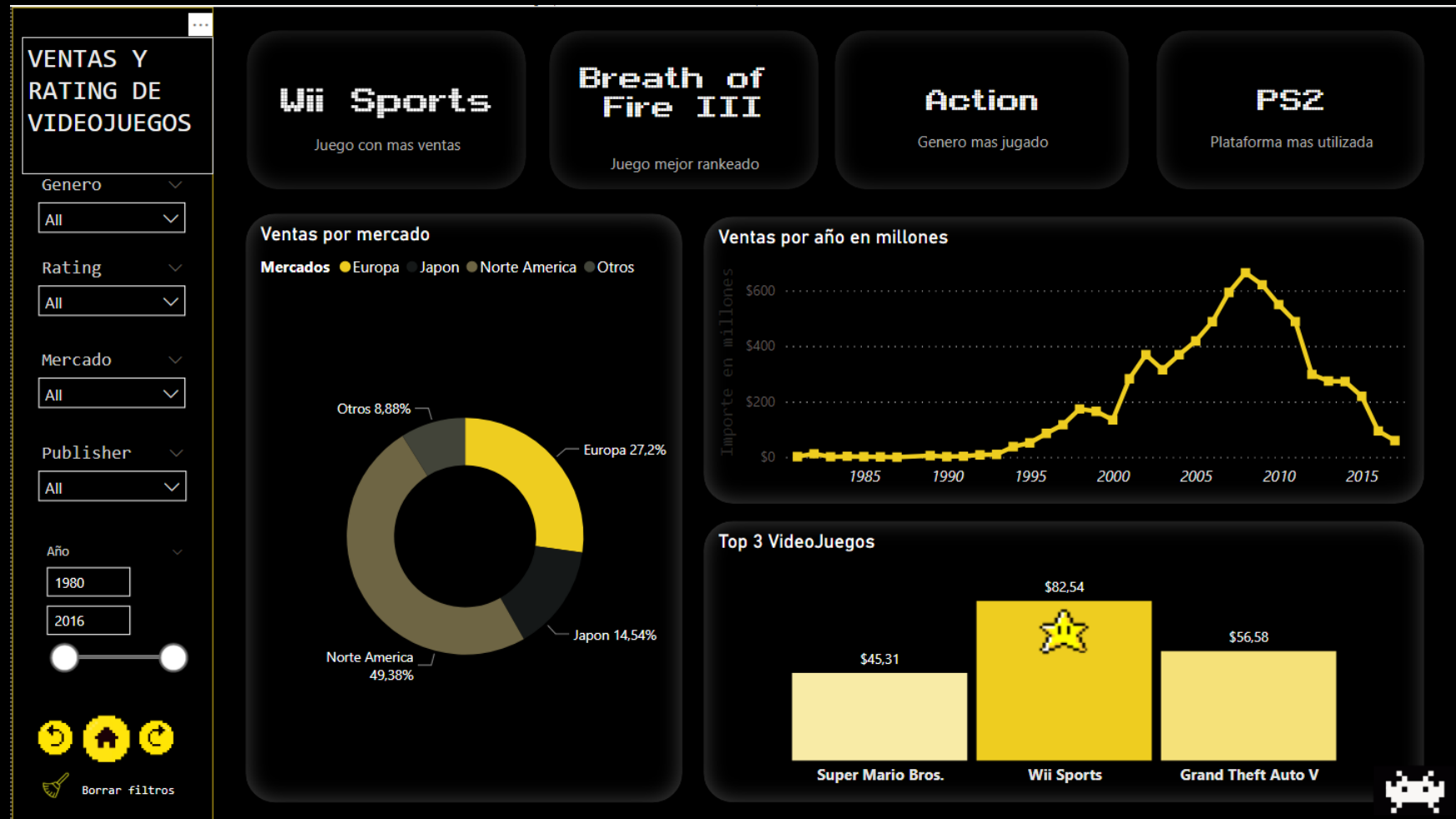


Solapa de Dashboard:



La solapa tiene una imagen de fondo del juego "space invaders" en el cual nos basamos para tomar la paleta de colores. En el se encuentras 4 botones que llevan a cada una de las solapas en orden de importancia.

Solapa Ventas y rating de videojuegos:



Basándonos en el Mockup incluimos los filtros que consideramos importantes en la parte superior siendo estos: Juego con más ventas, juego mejor rankeado, genero más jugado y plataforma más utilizada. En la parte inferior izquierda incluimos un gráfico de anillos para las ventas por mercado. En la parte derecha incluimos un gráfico de ventas por año usando un gráfico de líneas y por último utilizamos un gráfico de columnas apiladas para mostrar un Top 3 de videojuegos. Las medidas utilizadas en esta solapa fueron las siguientes:

Wii Sports

Juego con mas ventas

Juego con más ventas: Medida estática - Utilizando la funcion CALCULATE filtrada con la funcion DAX TOPN filtramos el juego mas vendido combinando la tabla videojuegos con la tabla ventas. La funcion DAX es la siguiente:

```
Medida - Juego con mas ventas = CALCULATE(SELECTEDVALUE('Tabla 1 - Videojuegos'[Nombre_videojuego]),  
TOPN(1,ALL('Tabla 8 - Ventas'),'Tabla 8 - Ventas'[Importe],DESC))
```

Breath of Fire III

Juego mejor rankeado

Juego mejor rankeado: Medida estática -Utilizando la funcion CALCULATE filtrada con la funcion DAX TOPN filtramos el juego mejor rankeado por los usuarios. La funcion DAX es la siguiente:

```
Medida - Juego mejor rankeado = CALCULATE(SELECTEDVALUE('Tabla 1 - Videojuegos'[Nombre_videojuego]),  
FILTER(TOPN(1,ALL('Tabla 1 - Videojuegos'),'Tabla 1 - Videojuegos'[Score_usuarios],DESC),TRUE()))
```

Action

Genero mas jugado

Genero más jugado: Medida variable - Utilizando la funcion CALCULATE, TOPN y SUMMARIZE creamos una tabla de la suma de ventas por genero para finalmente filtrarla mediante el TOPN dentro del CALCULATE para obtener el género más vendido. La funcion DAX es la siguiente:

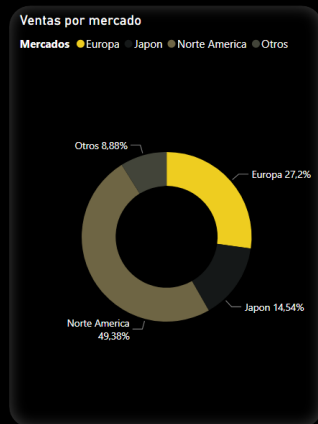
```
Medida - Ventas por genero = CALCULATE(SELECTEDVALUE('Tabla 6 - Genero'[Tipo_Genero]),TOPN(1,SUMMARIZE  
( 'Tabla 6 - Genero','Tabla 6 - Genero'[Tipo_Genero],"Ingresos", SUM ( 'Tabla 8 - Ventas'[Importe] )),  
[Ingresos],DESC))
```

PS2

Plataforma mas utilizada

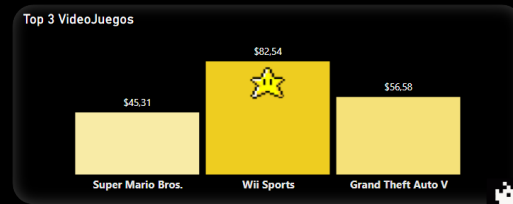
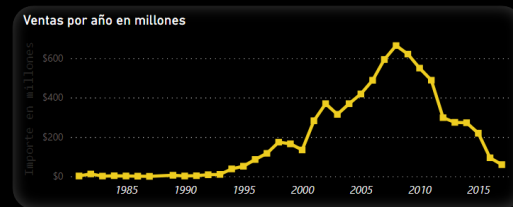
Plataforma más utilizada: Media variable -Para el cálculo de la plataforma, se realizó la misma funcion que en genero mas jugado para crear la medida. La funcion DAX es la siguiente:

```
Medida - Plataforma con mas ventas = CALCULATE(SELECTEDVALUE('Tabla 2 -  
Plataforma'[Nombre_plataforma]),TOPN(1,SUMMARIZE ( 'Tabla 2 - Plataforma','Tabla 2 -  
Plataforma'[Nombre_plataforma],"Ingresos", SUM ( 'Tabla 8 - Ventas'[Importe] )), [Ingresos],DESC))
```



Los graficos agregados en esta solapa fueron un grafico de dona, grafico de linea y grafico de barras:

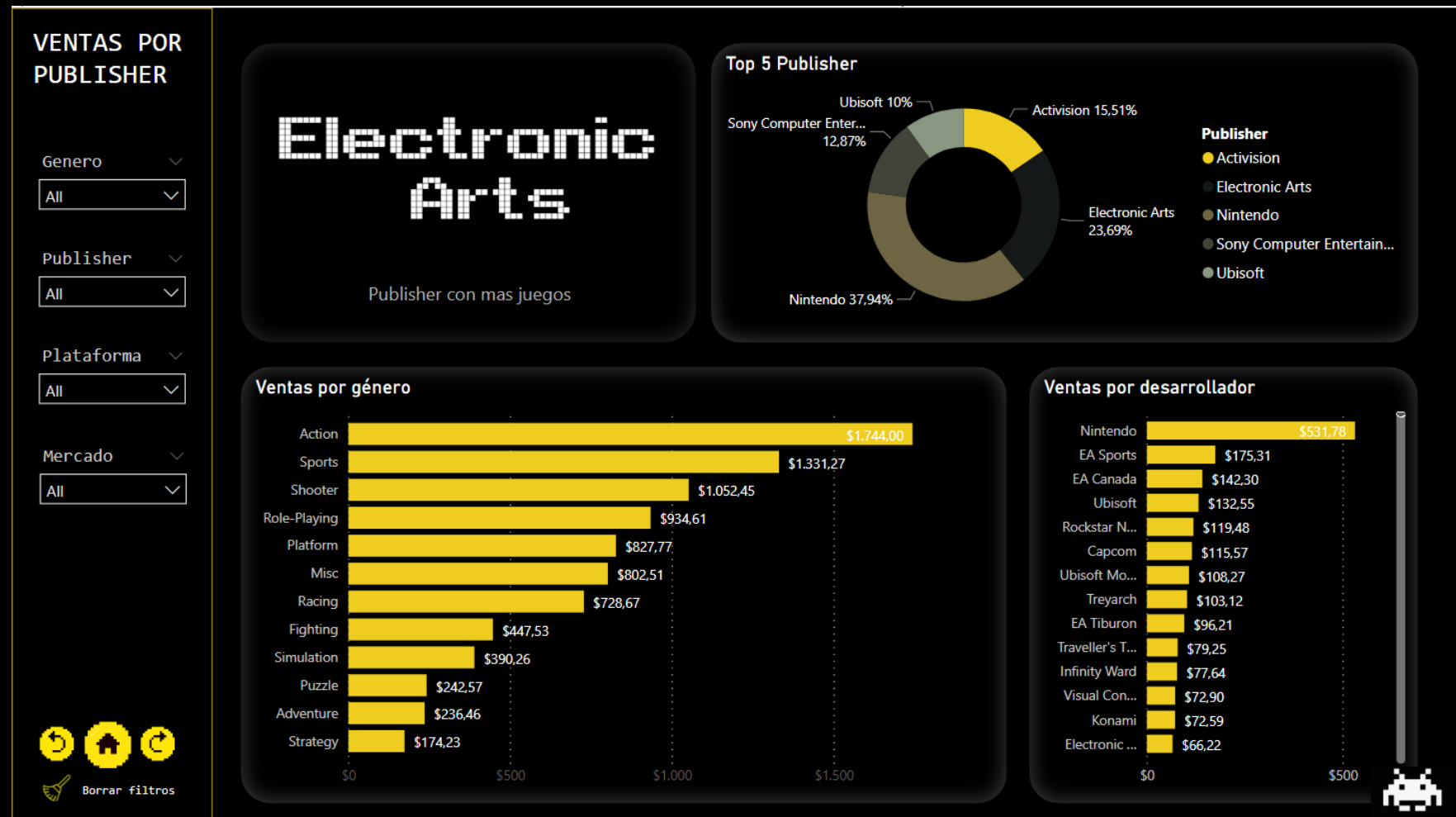
- Ventas por mercado: Muestra de forma porcentual las ventas por mercado de ventas.
- Ventas por año en miles: Grafico lineal de las ventas de videojuegos por año.
- Top 3 videojuegos: Indica los 3 juegos más vendidos usando el filtro TOP N. Para que el grafico muestre el TOP 1 en el centro se creo una medida la cual creando una variable rankPodio que contiene el numero de ranqueo por videojuego basado en la medida total ventas (juego con mayor venta) retorne a traves de un switch, un cambio en el orden para luego poder ordenar el grafico. Por ejemplo, si el ranking es 1, le va a sumar 1 para que la posicion sea 2. La formula DAX es la siguiente:



Orden - Ranking para podio =

```
VAR rankPodio = RANKX(ALL('Tabla 1 - Videojuegos'[Nombre_videojuego]),[Medida - Total Ventas])
RETURN
SWITCH(
    TRUE(),
    rankPodio = 1, rankPodio + 1,
    rankPodio = 2, rankPodio - 1,
    rankPodio = 3, rankPodio,
    4
)
```


Solapa Publisher y desarrolladores:

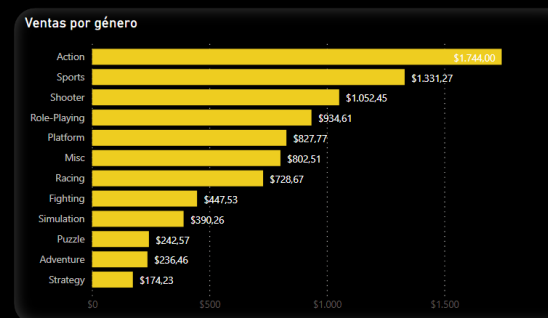
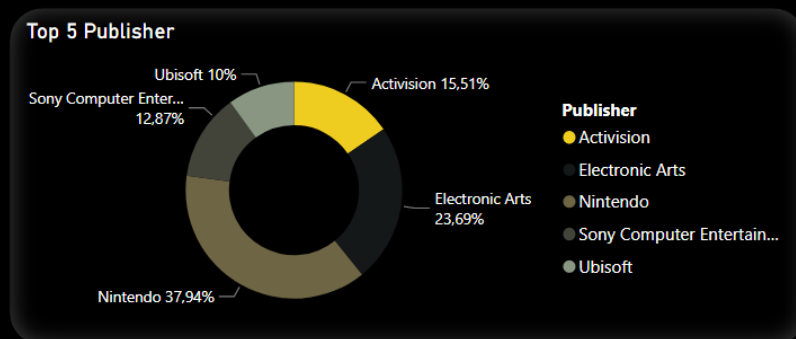


Esta solapa ayuda a identificar fácilmente cual es el Publisher con mas juegos, más ventas y el desarrollador que más ventas ha tenido. Esto lo consideramos practico ya que, al momento de elegir algún juego de la plataforma, el género, el desarrollador y el Publisher son elementos clave a tener en consideración.



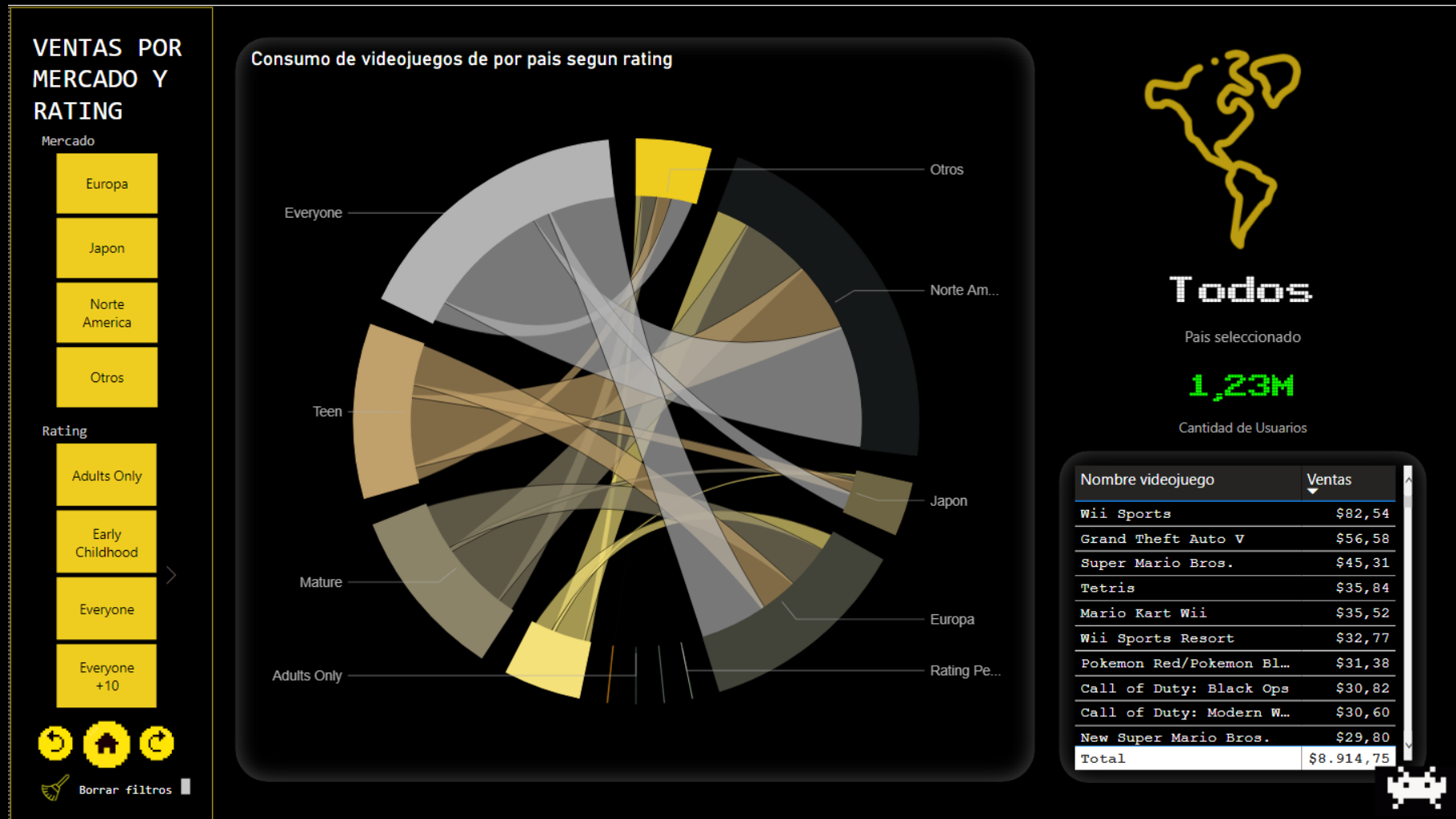
Publisher con más juegos: Medida variable - Para la medida se utilizaron CALCULATE, TOPN, SUMMARIZE y COUNTA. Siendo que la cantidad de juegos es texto, para contar estos valores se debió usar la funcion COUNTA para crear una table por cantidad de juegos por Publisher con la funcion SUMMARIZE, para luego filtrarla dentro del CALCULATE con un TOPN. Quedando la siguiente formula:

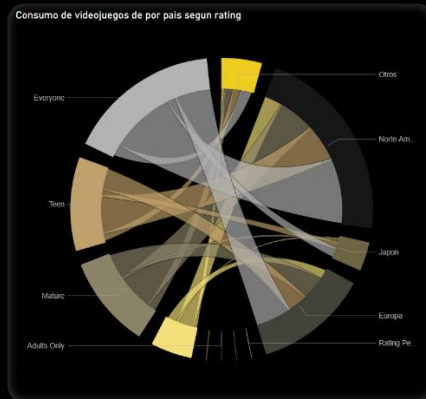
```
Medida - Publisher con mas juegos = CALCULATE(SELECTEDVALUE('Tabla 3
Publisher'[Nombre_publisher]),TOPN(1,SUMMARIZE ('Tabla 3 - Publisher','Tabla 3 -
Publisher'[Nombre_publisher],"Cantidad", COUNTA( 'Tabla 1 - Videojuegos'[Nombre_videojuego] )),
[Cantidad],DESC))
```



Para el resto de la grafica se seleccionaron un grafico de dona, que mostraría los 5 Publisher con mas ventas utilizando como filtro un TOP5 y dos graficos de barra simples que mostrarían en detalle las ventas por genero y las ventas por desarrollador.

Solapa consumo de videojuegos!





Consumo de videojuegos por país según rating: Este grafico de tipo chord descargado desde el store de Power BI muestra el hacia donde se dirige el consumo de rating a mercado según sus ventas. Es un grafico interesante y si bien parece complejo de leer a simple vista notamos que los juegos mas consumos son los juegos para todos en el mercado Norte América. De esta manera una decisión se puede tomar una decisión táctica sobre que tipo de rating ofrecer para cada mercado.

Todos

País seleccionado

1,23M

Cantidad de Usuarios

Filtro de país seleccionado: Medida variable - Para mostrar de una forma más dinámica que país se seleccionó para el filtro, se realizó una medida DAX que retorna el filtro aplicado. En caso que no haya ninguno, a traves del IFERROR, retorna "Todos"

```
Filtros aplicados = IFERROR(FILTERS('Tabla 7 - Market Ventas'[Nombre_Market]),"Todos")
```

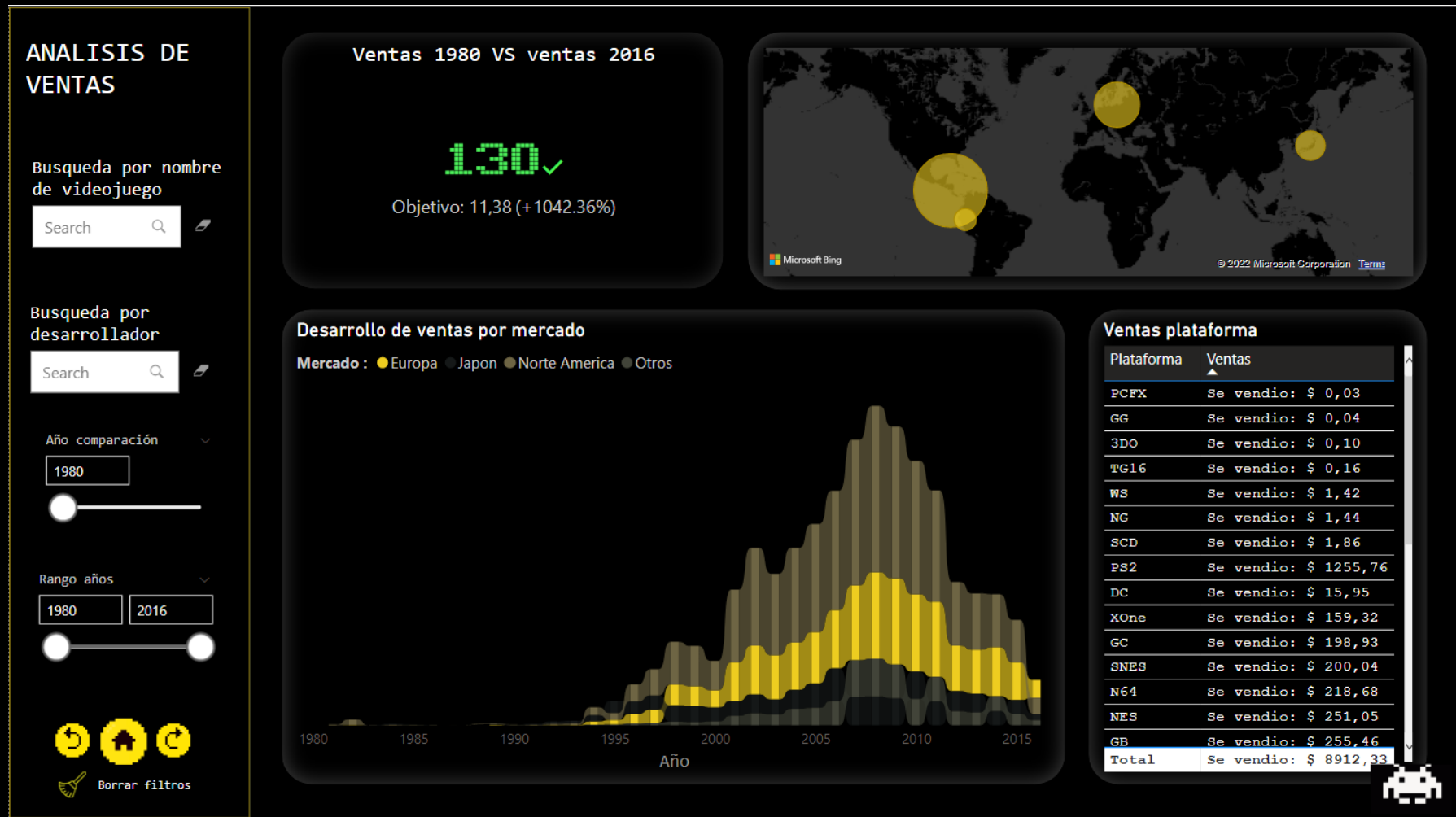
Cantidad de usuarios: Medida estática - Esta medida es la suma total de la cantidad de usuarios.

```
Medida - Cantidad de Usuarios = SUM('Tabla 1 - Videojuegos'[Cantidad_usuarios])
```

Nombre videojuego	Ventas
Wii Sports	\$82,54
Grand Theft Auto V	\$56,58
Super Mario Bros.	\$45,31
Tetris	\$35,84
Mario Kart Wii	\$35,52
Wii Sports Resort	\$32,77
Pokemon Red/Pokemon Bl...	\$31,38
Call of Duty: Black Ops	\$30,82
Call of Duty: Modern W...	\$30,60
New Super Mario Bros.	\$29,80
Total	\$8.914,75

Ventas por videojuego: Se agrego una tabla para ver en detalle las ventas por videojuegos. Esta tabla es filtrada por los segmentadores de dato de la solapa dando mas conocimiento sobre los juegos mas vendidos y menos.

Solapa analisis de ventas:



En esta ultima solapa se puede ver en detalle la tendencia de las ventas por mercado, el target de ventas del 2016 en comparación a un año seleccionado como parametro, las ventas por locación en el mapa y si hubo ventas de juegos para determinadas plataformas en los distintos años. Para que esta solapa sea más dinámica, se incluyó un gráfico del tipo Text Filter, que filtra por palabra siendo que hay 16mil líneas de videojuegos.

Ventas 1980 VS ventas 2016

130✓

Objetivo: 11,38 (+1042.36%)

Ventas año seleccionado VS ventas 2016: KPI – Este indicador que muestra la comparación de ventas del 2016 (último año del dataset) contra un año seleccionado por parametro. Para generar este fueron necesarias distintas medidas:

- *Ventas 2016:* Retorna la suma de ventas para el año 2016 utilizando un CALCULATE

```
Medida - Ventas 2016 = CALCULATE(SUM('Tabla 8 - Ventas'[Importe]), 'Tabla 1 - Videojuegos'[Anio publicacion]= "2016")
```

- *Ventas año seleccionado:* Retorna la suma de las ventas de un valor seleccionado en un parametro:

```
Medida - Ventas Año seleccionado =  
VAR anioSeleccionado = CONVERT(SELECTEDVALUE('Tabla Calculada - Anio comparacion'[Rango anios]),STRING)  
VAR retorno = CONVERT(VALUE(CALCULATE(SUM('Tabla 8 - Ventas'[Importe]), 'Tabla 1 - Videojuegos'[Anio publicacion] = anioSeleccionado)),DOUBLE)
```

Parametro año comparación: Este parametro fue creado y modificado para que sea automatizado tomando el año mas grande y el año mas chico a traves de la funcion MAX y MIN respectivamente:

```
Tabla Calculada - Anio comparacion =  
VAR anioMinimo = VALUE(MIN('Tabla 1 - Videojuegos'[Anio publicacion]))  
VAR anioMaximo = VALUE(MAX('Tabla 1 - Videojuegos'[Anio publicacion]))  
RETURN GENERATESERIES(anioMinimo,anioMaximo,1)
```

Año comparación

1980



Ventas plataforma	
Plataforma	Ventas
PCTX	Se vendió: \$ 0,03
GG	Se vendió: \$ 0,04
3DO	Se vendió: \$ 0,10
TG16	Se vendió: \$ 0,16
WS	Se vendió: \$ 1,42
NG	Se vendió: \$ 1,44
SCD	Se vendió: \$ 1,86
P82	Se vendió: \$ 1255,76
DC	Se vendió: \$ 15,95
XOne	Se vendió: \$ 159,32
GC	Se vendió: \$ 198,93
SNES	Se vendió: \$ 200,04
N64	Se vendió: \$ 218,68
NES	Se vendió: \$ 251,05
GB	Se vendió: \$ 255,46
Total	Se vendió: \$ 8912,33

Ventas por plataforma: Esta tabla incluye un condicional que retorna “Se vendió” o “Sin ventas” respectivamente si la plataforma filtrada tuvo o no tuvo ventas. Se utilizo la siguiente formula DAX:

```
Condicional - Ventas por plataforma =  
VAR TotalSales = SUM('Tabla 8 - Ventas'[Importe])  
VAR SalesTexto = CONVERT(FORMAT(SUM('Tabla 8 - Ventas'[Importe]),"$ ##0.00"),STRING)  
VAR Text1 = "Se vendio"  
VAR Text2 = "Sin ventas"  
RETURN IF(TotalSales> 0, CONCATENATE(CONCATENATE(Text1, ": "), SalesTexto),Text2)
```

Busqueda por nombre
de videojuego



Busqueda por
desarrollador



Text Filter: Nos pareció una propuesta interesante y dinámica, agregar del store de Power BI un grafico Text Filter que permite filtrar por palabra todos los datos coincidentes de la columna seleccionada. Por ejemplo, si quisiera ver todos los juegos de Mario sin tener que ir uno por uno, simplemente escribo Mario y pobre ver en analisis de ventas de sus juegos. Este filtro lo agregamos para nombre de videojuegos y desarrollador.

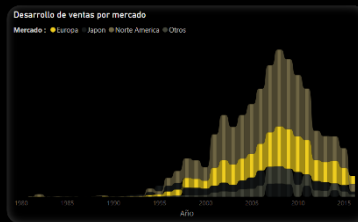


Gráfico Ribbon: Este grafico muestra el desarrollo de ventas de los videojuegos, pero a diferencia del de la primer solapa, este incorpora los valores por mercado para lo cual resulta, mas que evidente que en casi todos los años el mayor consumidor de videojuegos es Norte América excepto en el ultimo año del dataset que Europa parece haber superado el consumo de videojuegos de Norte América.

Tecnologia utilizada:

- <https://github.com> - Para poner en desarrollo las imágenes del tooltip y los iconos de los países
- <https://pinetools.com/es/obtener-colores-imagen> - Obtener los colores de la imagen de la portada
- <https://miro.com> - Crear el modelo relacional
- <https://canva.com> - Edición de imágenes
- <https://google.com> - Descarga de imágenes y contenido

Conclusion

En conclusión, podemos identificar que los juegos a integrar en una plataforma de videojuegos tienen que ser Wii Sports y Breath of Fire III principalmente. Luego, tomando los top 3 sabiendo que a futuro un desarrollador y publisher que siempre está en el top siendo Nintendo seleccionando juegos de acción principalmente para todos. El mercado más importante y que más consume videojuegos es Norte América lo cual lo hace ideal para incorporar la plataforma a ese mercado. Como futuras líneas creemos conveniente actualizar las ventas por país, para tener un conocimiento más certero sobre cada población y así ofrecer juegos de forma más personalizada. Creemos también conveniente a futuro que el gráfico debería contar con los jugadores activos por país, una vez en producción la plataforma de videojuegos, ofreciendo este mismo en función a la selección de juegos del jugador distintas sugerencias de juegos en función a sus gustos.

