

ACTIVIDAD 2

Tableau

Autor: Pablo García Lopera

Máster en Big Data y Data Science
Asignatura: 07MBIG - Visualización de Datos

1. Introducción

En esta actividad vamos a utilizar la herramienta de visualización de datos **Tableau online**. Para abrir la herramienta utilizaremos el link: <https://online.tableau.com/>

Vamos a utilizar una fuente de datos de películas que podemos descargar desde el link: <https://www.tableau.com/sites/default/files/pages/movies.xlsx>

Tableau (I)

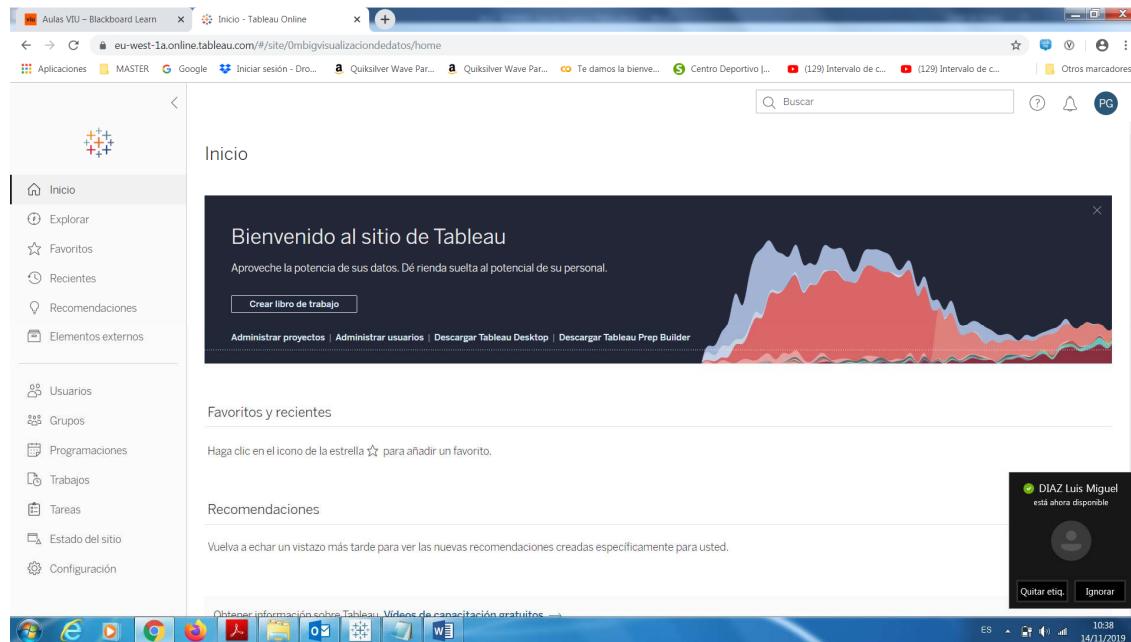
Este es el listado de gráficos que vamos a crear:

1. Presupuesto x Distribuidor/Genero (**Barras Apiladas**)
2. N.º de films x Distribuidor/Genero (**Tree Map**)
3. Evolución en el tiempo: Presupuesto + Recaudación x Distribuidor
4. Contribución del Genero a los Beneficios de cada Distribuidor (**Barras o Áreas apiladas**)
5. Crear un **Dashboard**. Incluir los gráficos en el **Dashboard**.

2. Tableau (II)

En este ejercicio, vamos a cargar los datos de Películas y crear nuestra área de trabajo. Para ello haremos lo siguiente:

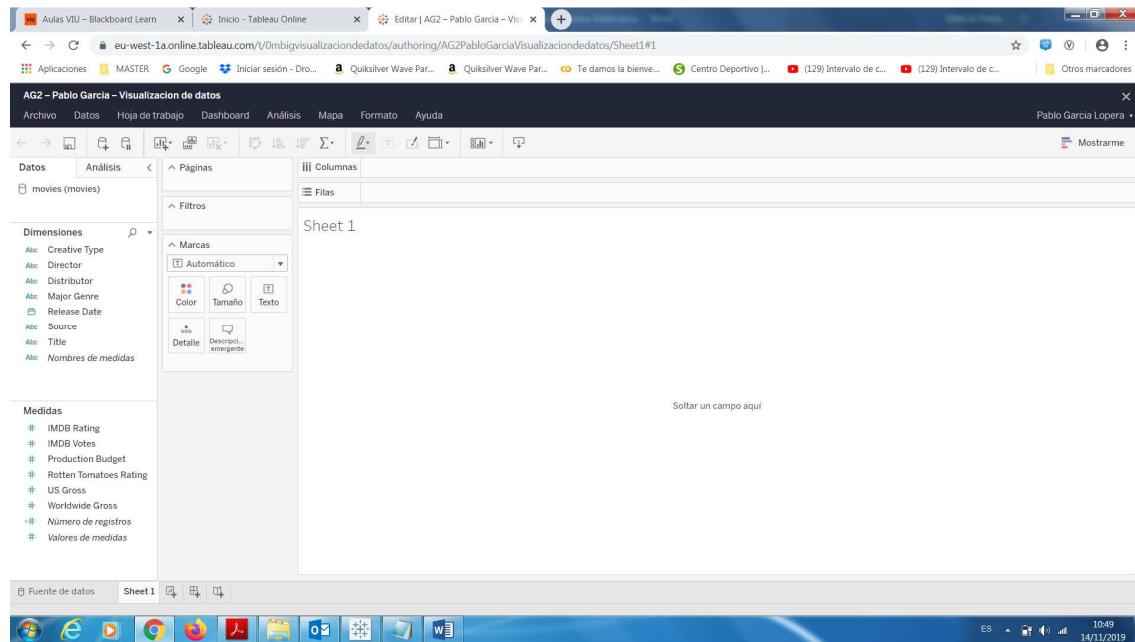
1. Abrimos **Tableau online**. Nos aparecerá la siguiente pantalla:



2. Seleccionar la opción **Crear libro de trabajo**. Nos aparecerá una ventana para cargar la fuente de datos.

3. En nuestro caso, el fichero Excel **movies.xlsx** lo tenemos en el disco duro local, por lo tanto, elegimos la opción **Archivo**, y seleccionamos nuestro fichero. Una vez cargado, ir a **Fuente de datos** (pestaña en la parte inferior izquierda) y podremos ver los datos de las películas.

4. Guardamos nuestro **Libro de trabajo** desde el menú **Archivo**, seleccionando **Guardar como...**. Nos pedirá un nombre. Le vamos a dar el nombre **AG2 – Pablo Garcia – Visualización de datos**.



Cambio de nombre a algunas etiquetas (Dimensiones y Medidas)

Vamos a renombrar algunas etiquetas de nuestra Fuente de datos. Vamos a cambiar el nombre del campo de dimensiones **Major Genre** a **Género**. Para ello haremos lo siguiente:

1. Hacemos clic con el botón derecho sobre el campo a renombrar. En el menú contextual que aparece, seleccionar **Cambiar nombre**



2. Poner **Género** y pulsar **Aceptar**
3. Realizamos los mismos pasos para cambiar los campos: **Distributor**→Distribuidor, **Release Date**→Fecha Estreno, **Title**→Título, **Production Budget**→Presupuesto, **Worldwide Gross**→Recaudación y **Número de registros**→Número de películas.

Tableau (III)

En este apartado vamos a crear un gráfico de Barras apiladas que muestre **Presupuesto x Distribuidor/Género**. Para ello haremos los siguientes pasos:

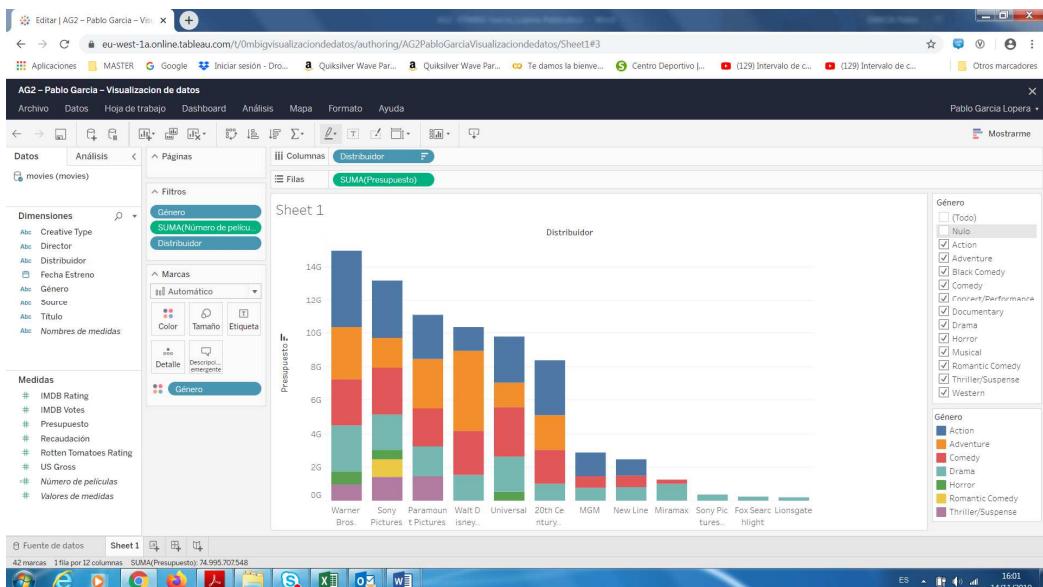
1. Arrastrar el campo **Género** al espacio de trabajo. Aparecerá el campo en Filas. Apuntar al campo y hacer clic con el botón derecho. Aparecerá un menú contextual, seleccionar **Mostrar Filtro**. En la parte derecha de la pantalla, aparecerán los valores de **Género**
2. Para quitar los valores **Nulos**, desactivar el valor **Nulo** del filtro de **Género** de la derecha

The screenshot shows the Tableau interface with the 'movies' data source. In the Rows shelf, the 'Genero' field is selected. A context menu is open on this field, with 'Mostrar Filtro' (Show Filter) highlighted. To the right, the 'Genero' filter pane is open, showing a list of genres: Action, Adventure, Black Comedy, Comedy, Concert/Performance, Documentary, Drama, Horror, Musical, Romantic Comedy, Thriller/Suspense, and Western. The 'Nulo' checkbox is unchecked.

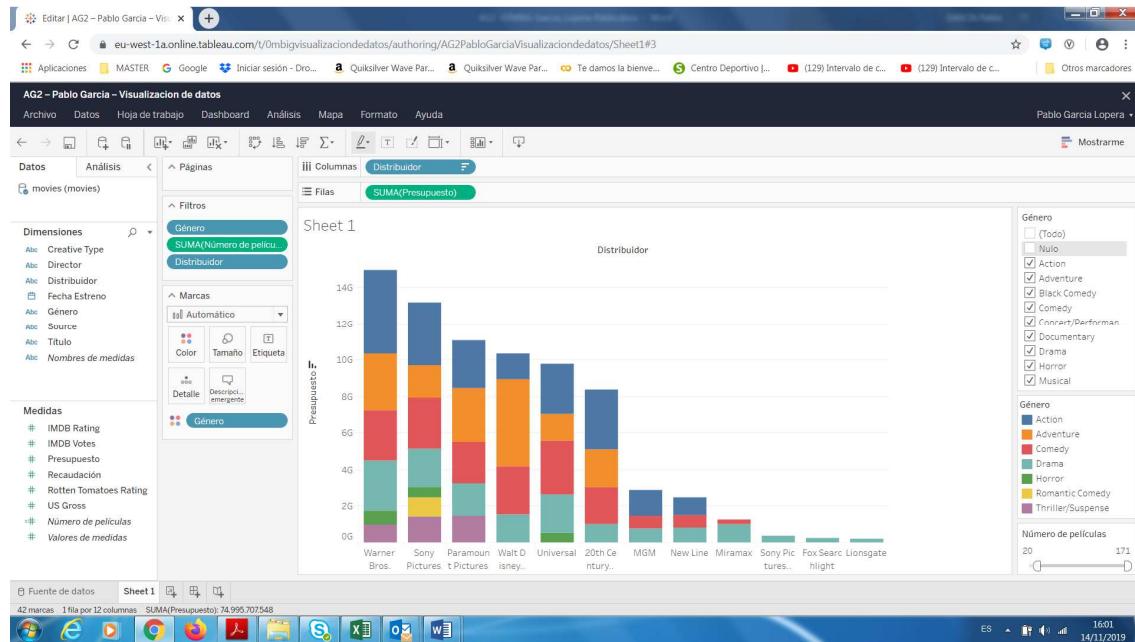
3. Arrastrar el campo **Presupuesto** encima de la columna con valores **Abc** de la Tabla **Género**. Aparecerá el campo en Filas. Al soltar, se mostrarán la suma de presupuestos por Género. Apuntar al campo y hacer clic con el botón derecho. Aparecerá un menú contextual, seleccionar **Mostrar Filtro**. En la parte derecha de la pantalla, aparecerán los valores de **Género**

Creación Gráfico 1 - Presupuesto x Distribuidor/Genero (Barras apiladas)

4. Arrastrar el campo **Distribuidor** a **Columnas** de la parte superior, y ocultar los valores Nulos por Distribuidor. Abrir los tipos de gráficos (Símbolo en la parte superior derecha Muéstrame), y elegir el tipo de gráfico de **Columnas**.



5. Vamos a hacer un filtro para el campo **Número de películas > 20**. Para ello, apuntar al campo y hacer clic con el botón derecho. Aparecerá un menú contextual, seleccionar **Mostrar Filtro**. En la parte derecha de la pantalla, aparecerá una barra de valores entre **1** y **171**. Modificar el campo de inicio de rango a valor **20**.



6. Para renombrar la Hoja de gráfico, hacer doble clic sobre la pestaña, y escribir **Presupuesto x Distribuidor/Genero**. En el título del gráfico se mostrará el mismo texto.

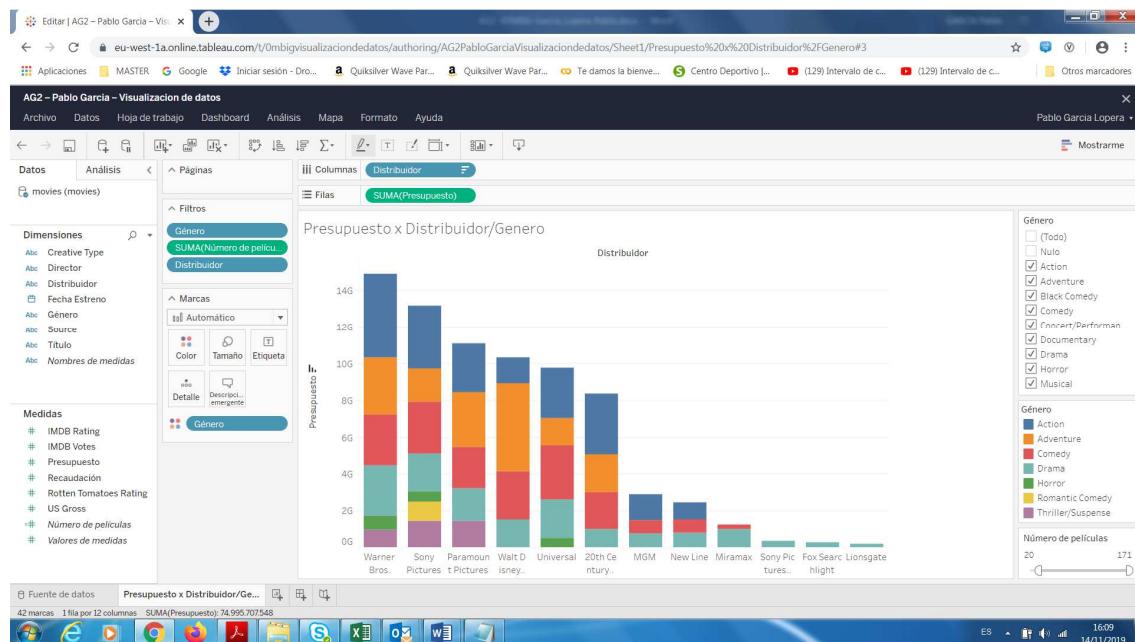
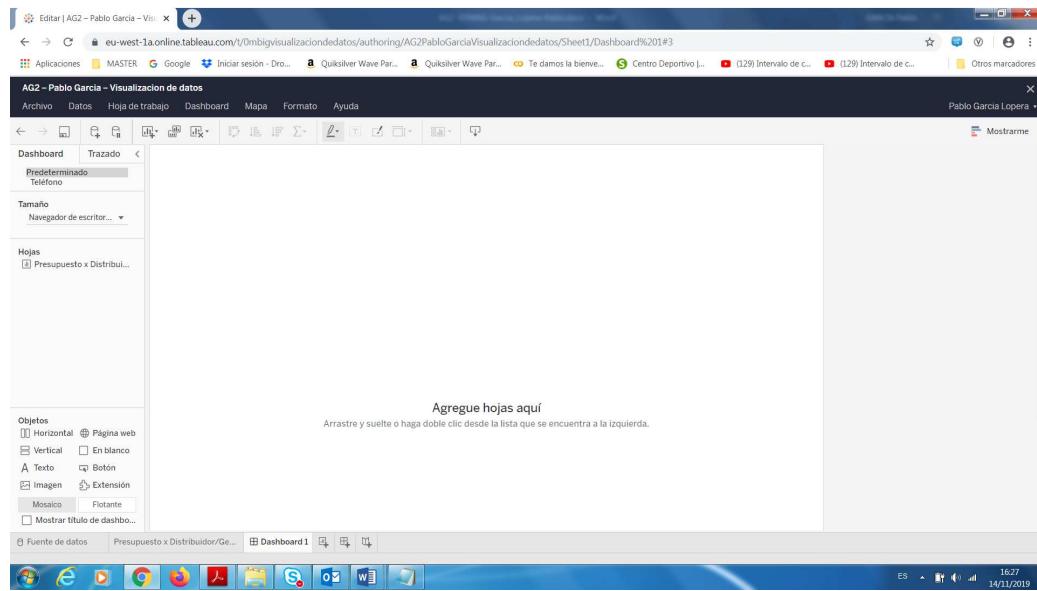


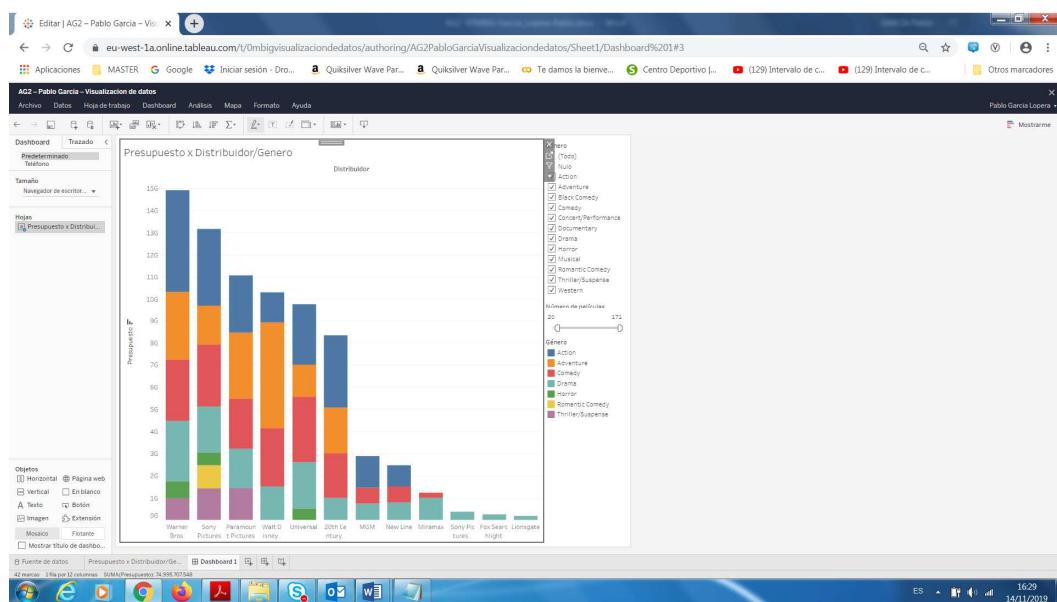
Tableau (IV)

En este apartado, vamos a crear un **Dashboard** e introduciremos en él, el **Gráfico 1**. Los siguientes gráficos creados, también los iremos añadiendo a este Dashboard. Para crear el Dashboard haremos lo siguiente:

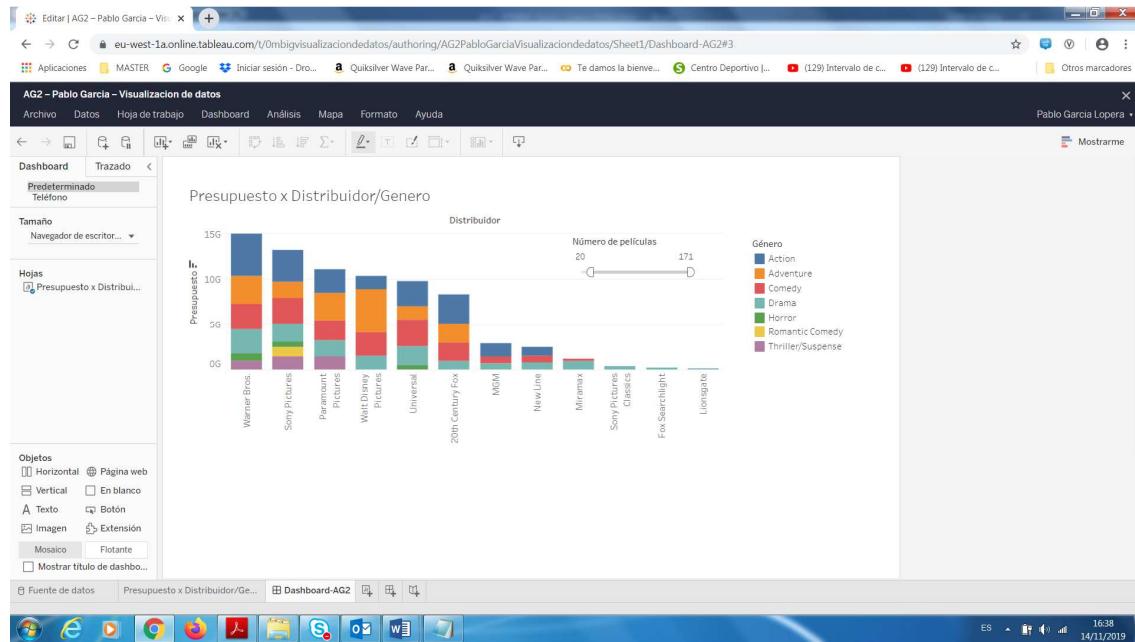
1. A la derecha de las pestañas de Hojas, hay 3 iconos, el primero para crear **Hojas nuevas**, el segundo para crear **Dashboard**, y el tercero para crear **Nuevas historias**. Pulsar sobre la segunda opción. Aparecerá un nuevo Dashboard vacío



2. Arrastrar nuestra Hoja de gráfico (desde Hojas, en la parte izquierda), al cuadro del Dashboard. Se mostrará nuestro gráfico en el Dashboard



3. En el Dashboard hay diferentes objetos: el gráfico, la leyenda del género, la barra del rango de Número de películas, etc.... Si pulsamos sobre cada objeto, en la parte superior derecha aparece una barrita gris con iconos. Pulsamos sobre la flecha y aparecerá un menú. Seleccionamos la opción **Flotante**, y el objeto se podrá trasladar, cambiar su tamaño, etc. Después de realizar estas acciones con los diferentes objetos, este sería el nuevo aspecto:

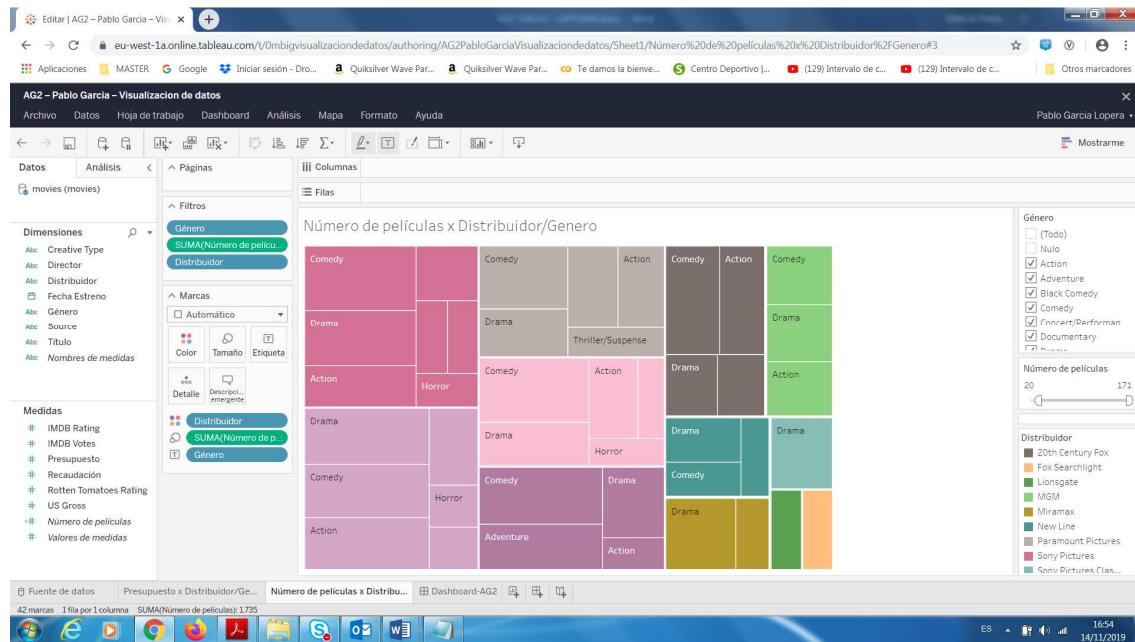


Creación Gráfico 2 - Número de películas x Distribuidor/Genero (Tree Map)

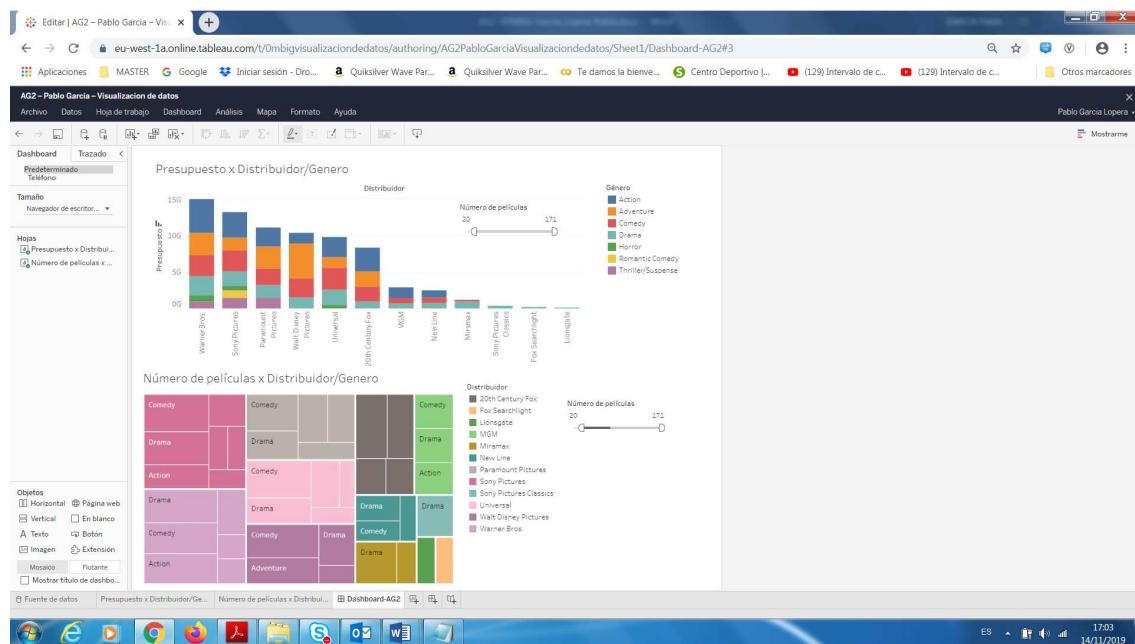
Este nuevo gráfico utiliza casi los mismos datos que el gráfico 1, sólo habrá que cambiar el campo **Presupuesto** por **Número de películas**. Por lo tanto, podemos duplicar el **gráfico 1**, y luego hacer pequeños cambios. Para ello vamos a realizar los siguientes pasos:

1. Hacer clic con el botón derecho encima de la Hoja **Presupuesto x Distribuidor/Genero**. En el menú que aparezca, seleccionar la opción **Duplicar**. Tendremos una nueva Hoja de gráfico, igual que la del gráfico 1.
2. Para cambiar el nombre de la nueva Hoja, hacemos doble clic sobre la pestaña, y escribimos **Número de películas x Distribuidor/Genero**.
3. En gráficos seleccionamos **Tree Map**. 
4. En **Marcas→Color** arrastramos el campo **Distribuidor**

Después de todos estos cambios y acciones, el **Gráfico 2** tendrá el siguiente aspecto:



A continuación, añadimos este gráfico al **Dashboard**, junto con el Gráfico 1



Nota: A partir de ahora, no voy a dar tanto detalle en la creación del gráfico, ya que la mayor parte de las acciones, han sido explicadas en los gráficos anteriores

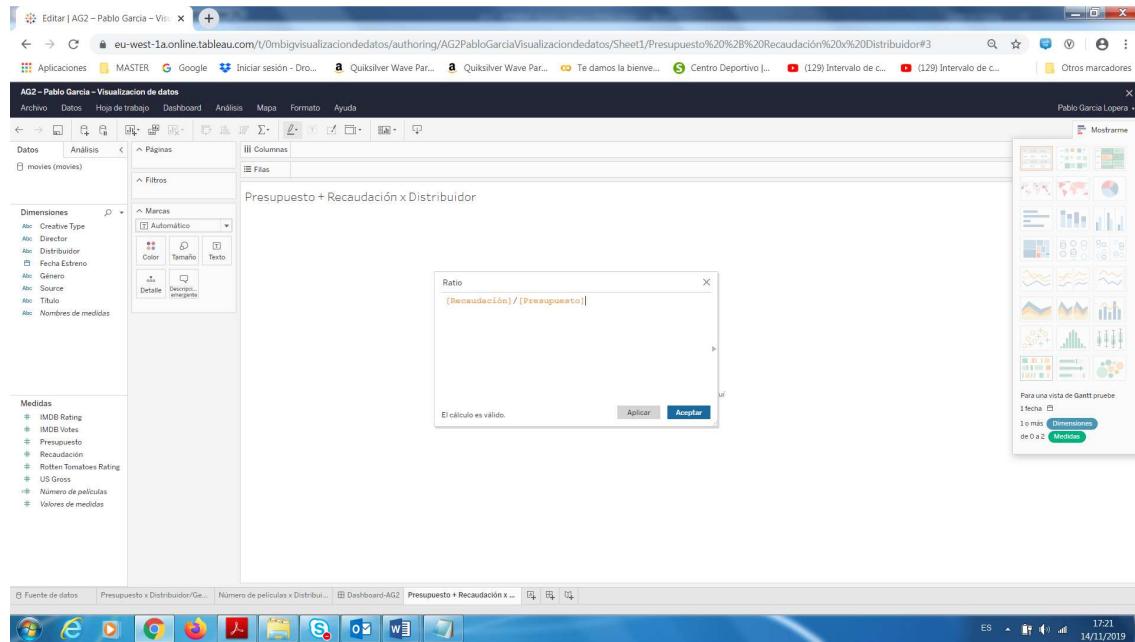
Creación Gráfico 3 - Presupuesto + Recaudación x Distribuidor (Evolución en el tiempo)

Para este gráfico vamos a crear un campo calculado, y lo llamaremos **Ratio**. El cálculo será:

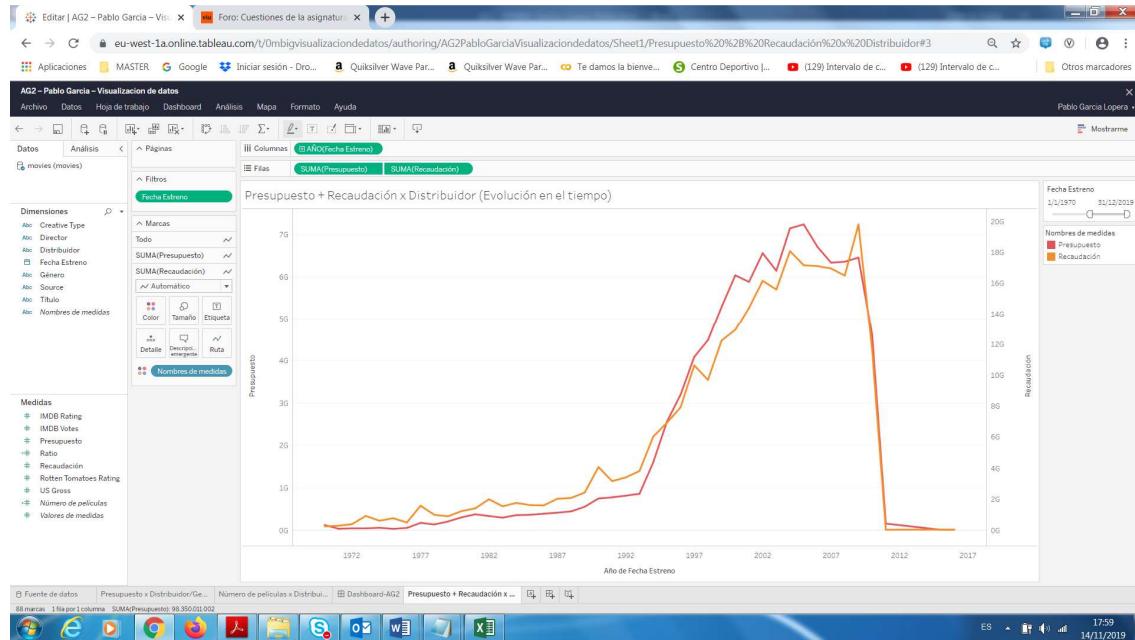
$$\text{Ratio} = [\text{Recaudación}] / [\text{Presupuesto}]$$

Para ello haremos lo siguiente:

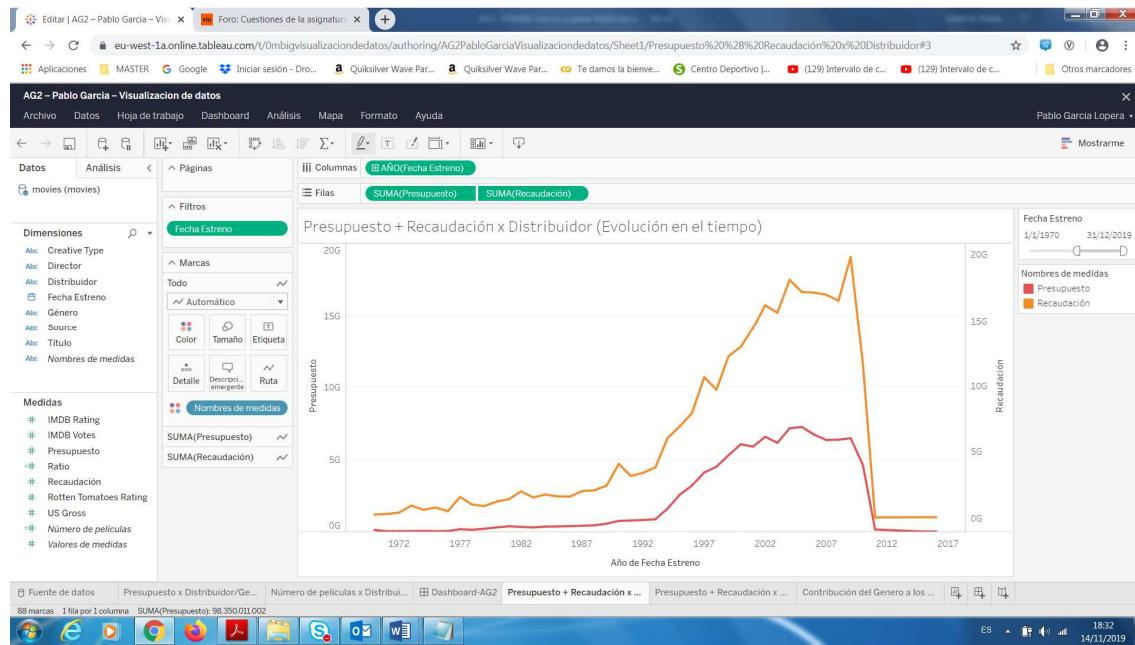
1. Desde el menú **Análisis**, seleccionar **Crear campo calculado**. Aparecerá una ventana que rellenaremos como en la imagen



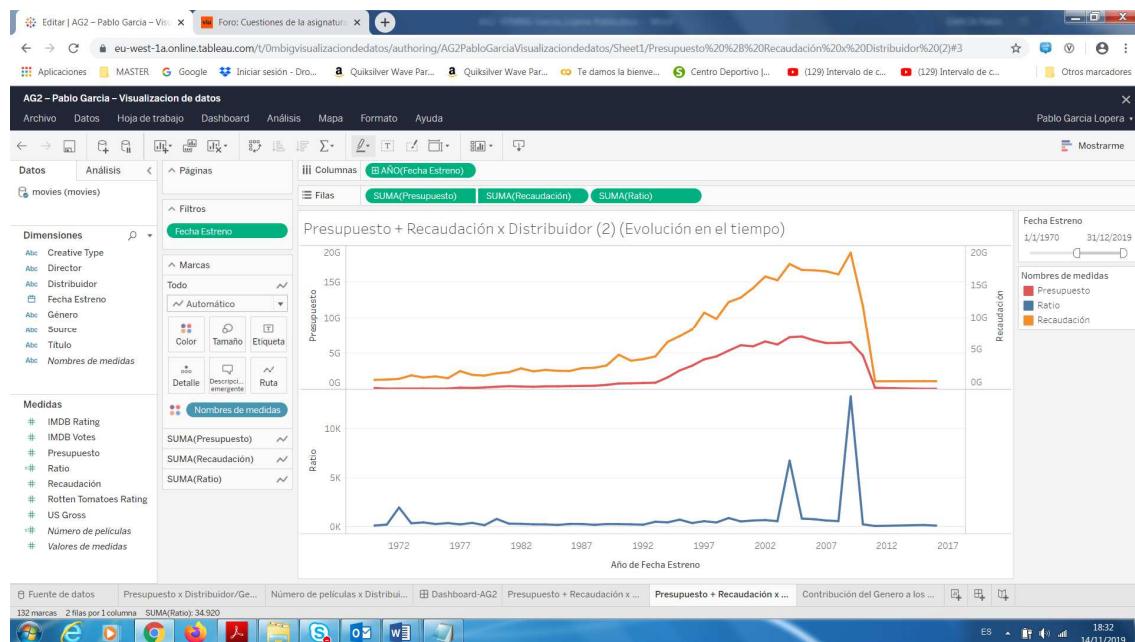
Debemos excluir las películas Fecha de Estreno: **2020, 2025, 2027**, y que empiecen a partir de **1970**. Para ello en la barra de Filtros de **Fecha de estreno**, ponemos como fecha de inicio **01/01/1970** y como Fecha Fin **31/12/2019**. Elegimos **Gráfico de líneas**, y después de añadir los campos necesarios, el gráfico tendrá el siguiente aspecto:



Vamos a poner los dos ejes con el mismo rango. Para ello cambiamos los valores del eje **Presupuesto** de 0G→7G a 0G→20G, igual que el eje **Recaudación**, para que tengan la misma escala. El gráfico quedará así:

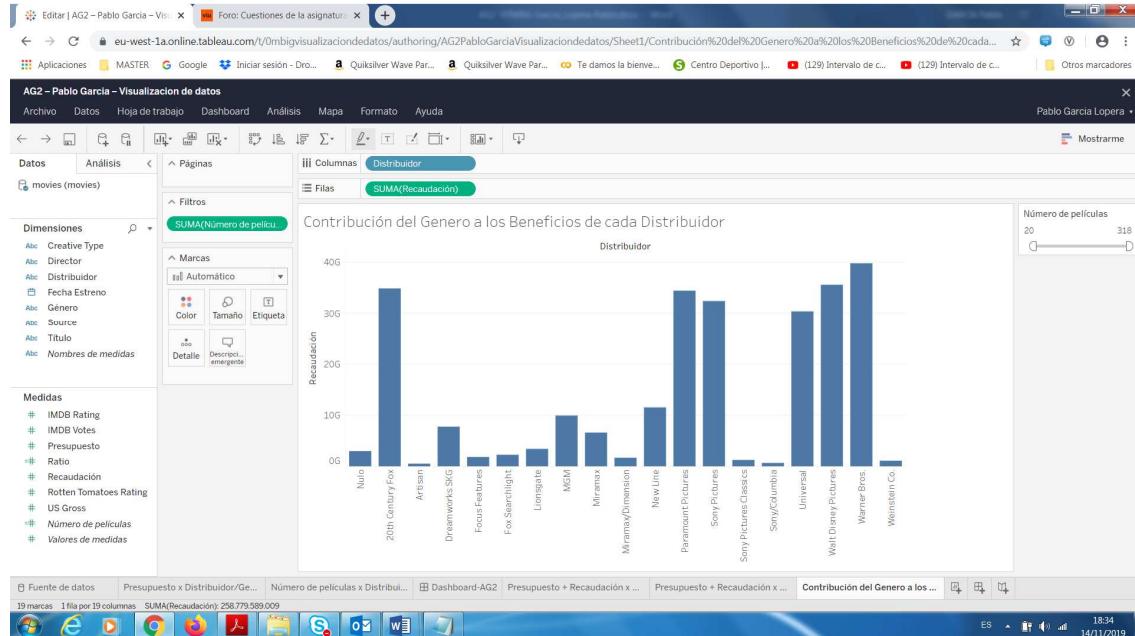


A continuación, vamos a añadir el campo calculado **Ratio** a Filas, y el gráfico tendrá el siguiente aspecto:



Creación Gráfico 4 - Contribución del Genero a los Beneficios de cada Distribuidor (Barras)

Vamos a crear un gráfico de barras poniendo el campo **Distribuidor** en Columnas, y **Recaudación** en Filas. A continuación, filtraremos por distribuidores con más de 20 películas, y eliminaremos distribuidor con valor **Nulo**. El gráfico tendrá el siguiente aspecto:



A continuación, modificaremos el gráfico, arrastrando el campo **Género** a **Color (Marcas)** y ordenaremos de mayor a menor las barras. El gráfico tendrá el siguiente aspecto:

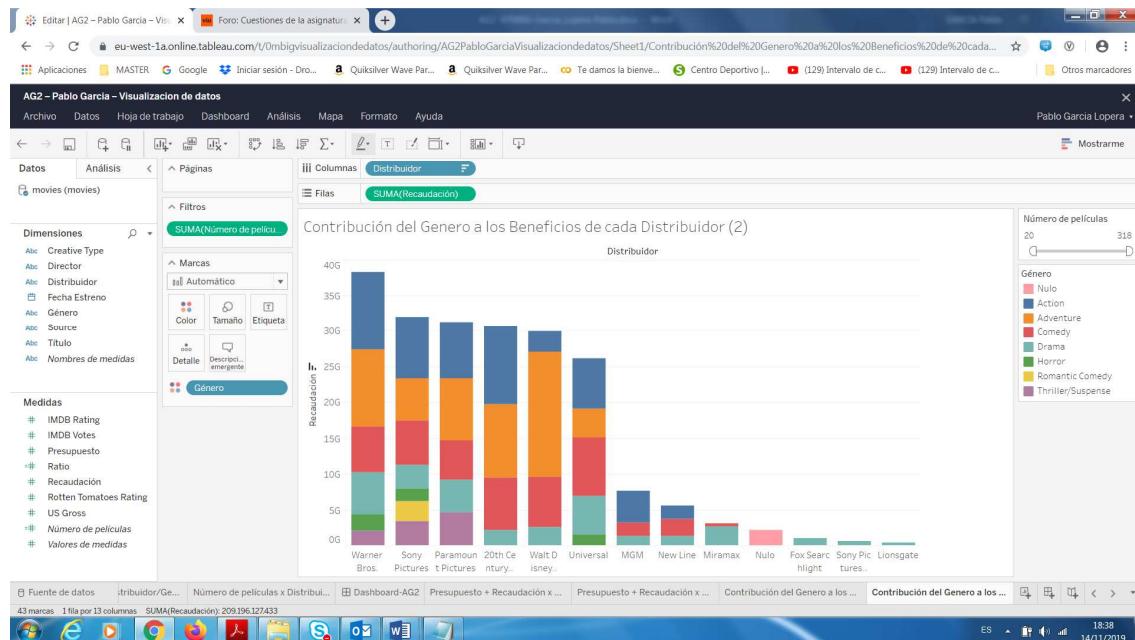
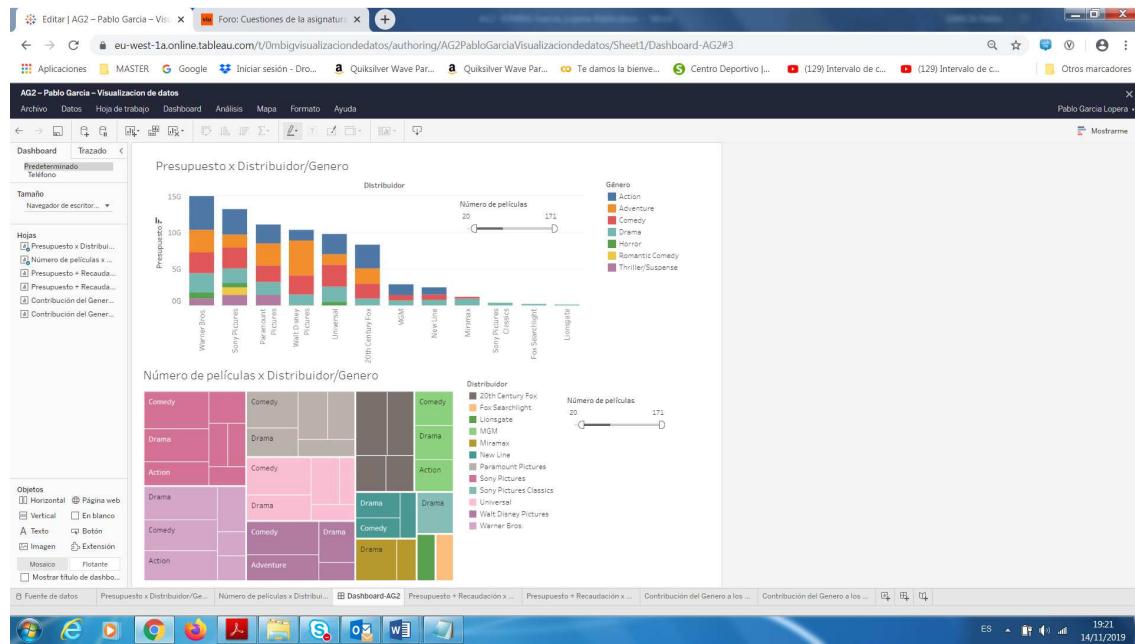
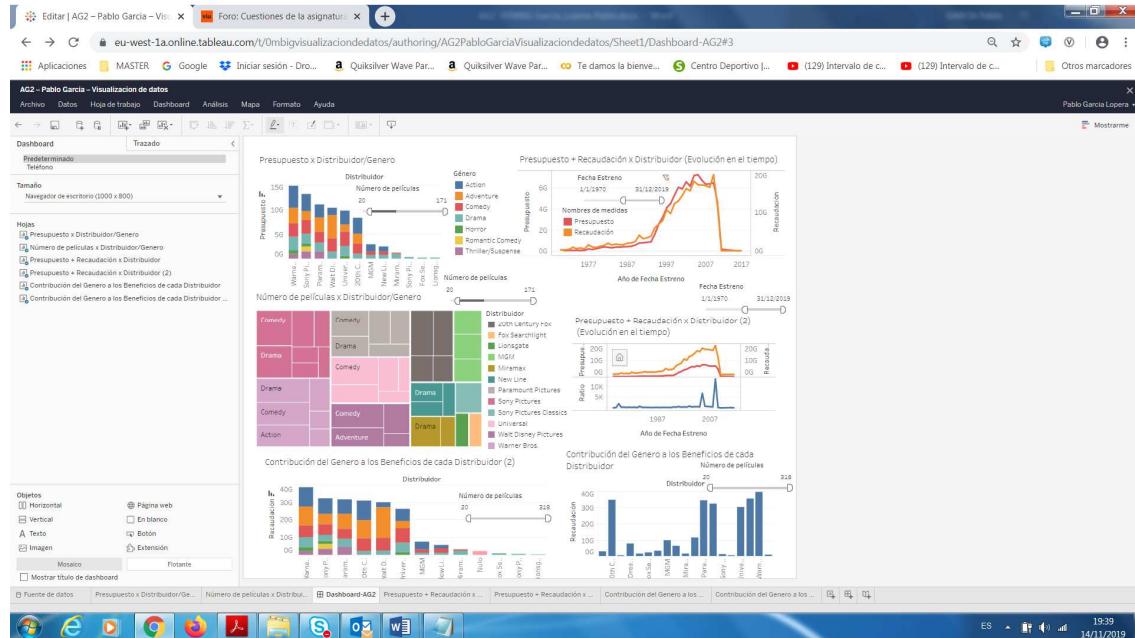


Tableau (V)

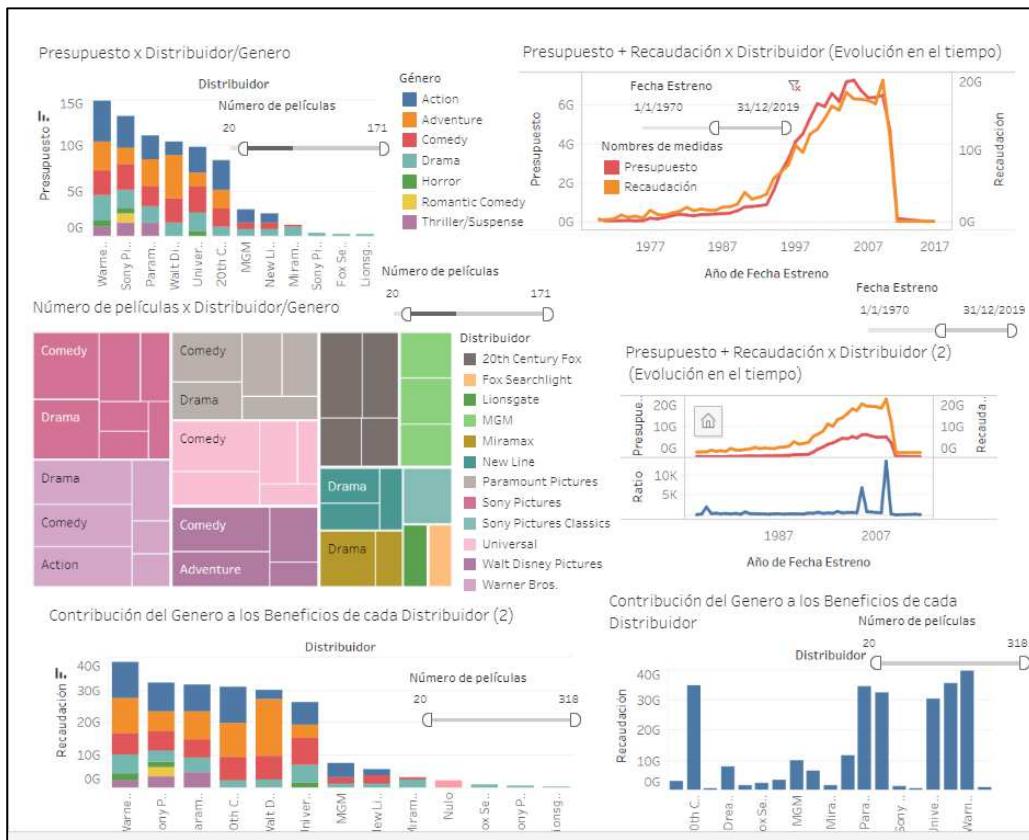
Vamos a añadir todos los gráficos al Dashboard. Hasta aquí, teníamos los dos primeros gráficos



Ahora añadimos los faltantes. **Nota:** del gráfico **Presupuesto + Recaudación x Distribuidor (evolución en el tiempo)**, como se han hecho un primer gráfico, y luego un segundo final, he añadido los dos. Lo mismo para el gráfico **Contribución del Género a los Beneficios de cada Distribuidor**, se añaden, el inicial y el final. El resultado final es:



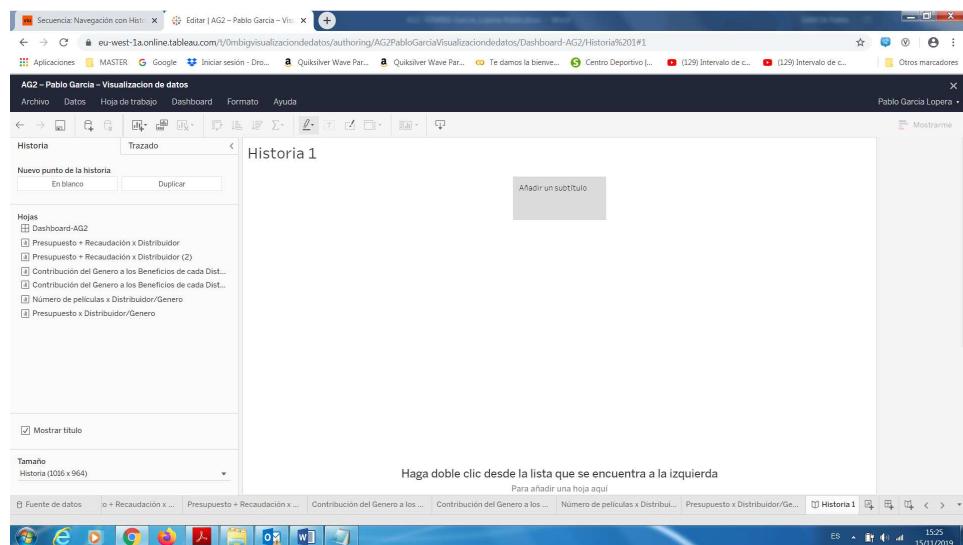
Y si sólo mostramos el Dashboard, este es el aspecto:



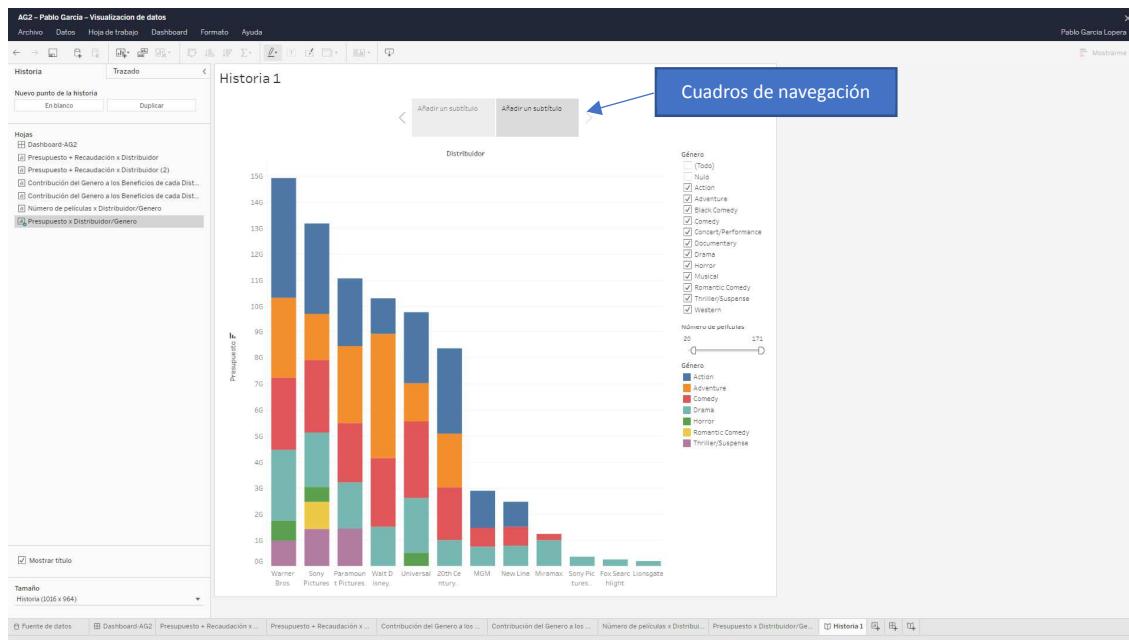
Añadir Navegación con Historias

Vamos a crear una Historia y la llamaremos **Historia AG2**. Esta Historia debe mostrar el gráfico **Presupuesto x Distribuidor/Genero**, y con un panel de navegación que permita, ir de la distribuidora **Sony** a la distribuidora **Warner**. Para ello haremos lo siguiente:

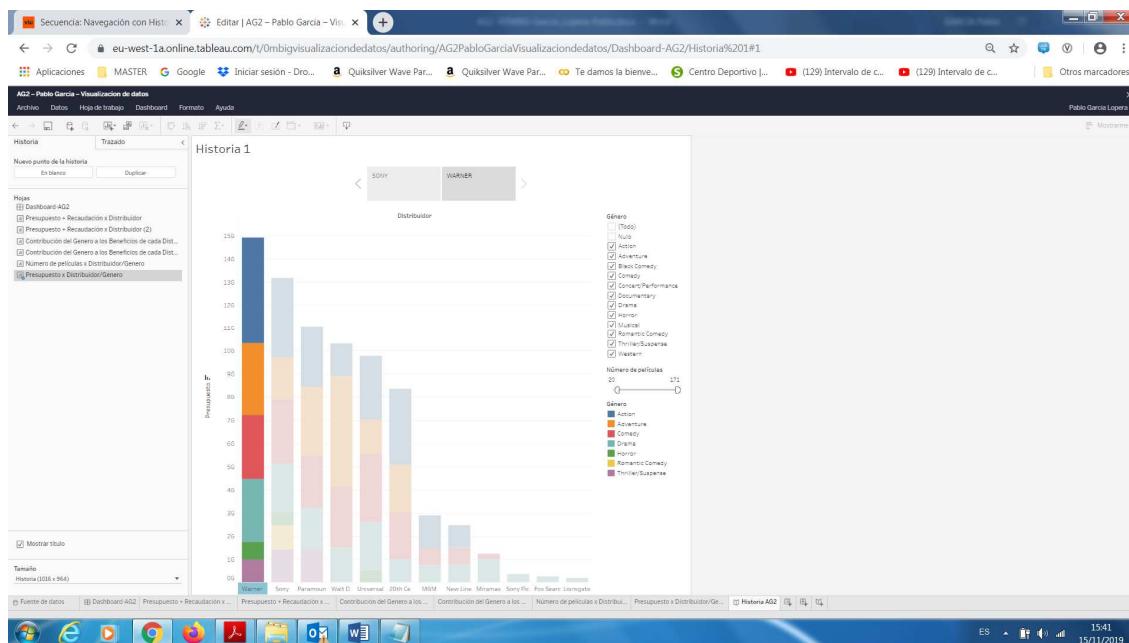
1. Para crear la **Historia**, hacemos clic en el icono de la parte inferior **Nueva historia**. Aparecerá la nueva Historia vacía.



2. En el panel de la izquierda, dónde aparecen nuestras hojas de gráficos, hacemos doble clic en la hoja **Presupuesto x Distribuidor/Genero**. Nos aparecerá el gráfico, y un cuadro de navegación. Volvemos a hacer doble clic en la hoja **Presupuesto x Distribuidor/Genero**, y nos aparecerá un segundo cuadro de navegación.

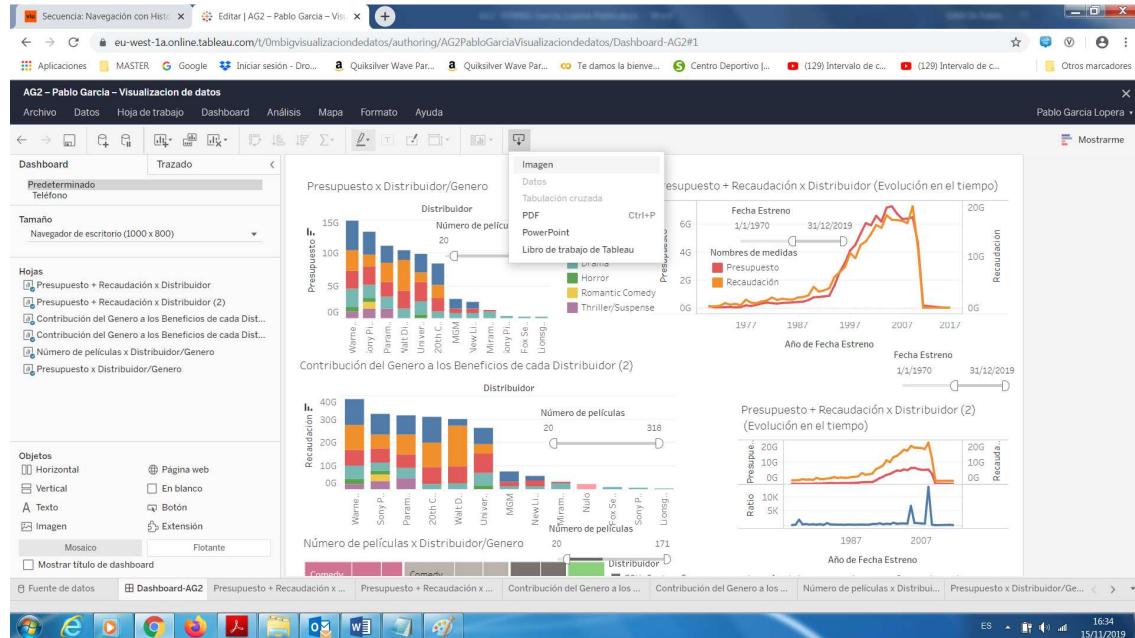


3. Para asociar el primer cuadro de navegación a la distribuidora **Sony**, hacemos clic en él para seleccionarlo, y a continuación, clic en **Sony** del eje X del gráfico. Hacemos lo mismo para el segundo cuadro de navegación, pero seleccionamos **Warner**.
4. Hacemos doble clic en cada cuadro, y ponemos en el primero **Sony**, y en el segundo **Warner**. Ahora si pulsamos en cada cuadro, se mostrará sólo la barra del gráfico de cada distribuidora, alternativamente.

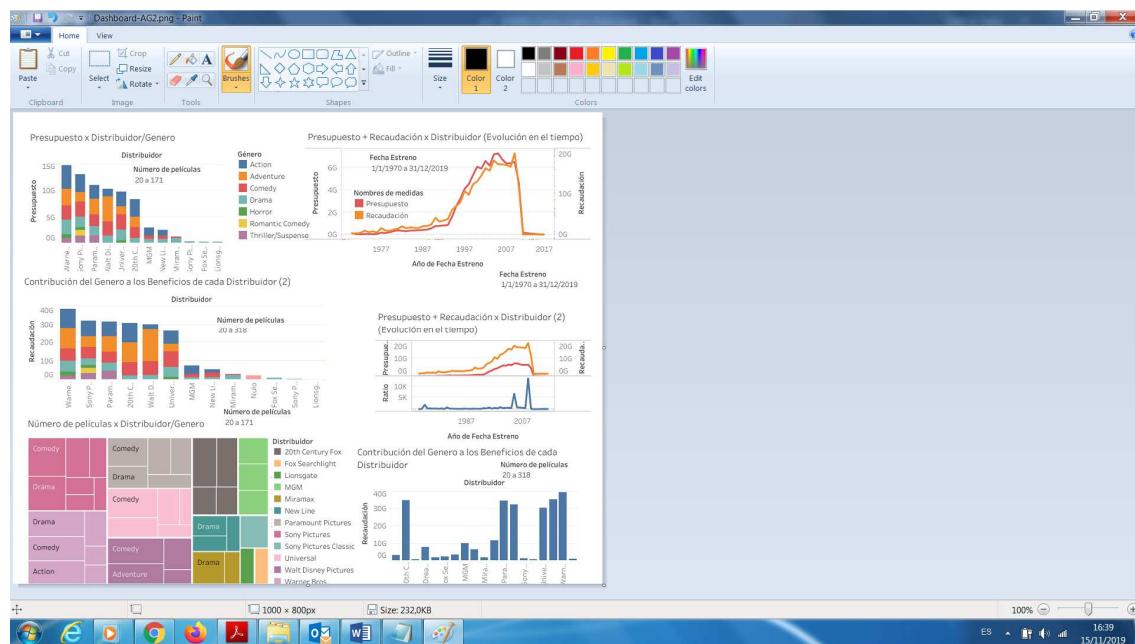


Descargar imagen

Para descargar la imagen de nuestro **Dashboard**, hacemos clic sobre el icono **Descargar** de la barra de herramientas, y seleccionamos **Imagen**. Se nos descargará un fichero de imagen en la carpeta de descargas



Si abrimos este fichero, veremos la imagen guardada en la aplicación **Paint** de Windows



Subir imagen a la web imgbb

Vamos a subir a imagen del **Dashboard** descargada a una web de imágenes <https://imgbb.com/>. El enlace para visualizarla es <https://ibb.co/tBCKgJv>

También podemos obtener de esta web el código **html**:

The screenshot shows a Windows taskbar at the bottom with icons for File Explorer, Internet Explorer, and other applications. Above it, a browser window is open with three tabs: 'Secuencia: Navegación con Hist...', 'Editar | AG2 – Pablo García – Vis...', and 'ImgBB — Subir Imágenes — Al...'. The main content area displays a green circular icon with a white checkmark, indicating the upload is complete. Below it, the text 'Subida completada' is shown. A note says 'Puedes crear un nuevo álbum con el contenido recién subido. Debes crear una cuenta o entrar para guardar este contenido en tu cuenta.' Underneath, there's a preview of the uploaded dashboard image and an 'Insertar' (Insert) button with a dropdown menu set to 'HTML completa enlazada'. The copied HTML code is pasted into the menu:

Añadir el Dashboard al Proyecto de stackblitz de la AG1

Con el código **html** obtenido de la aplicación **imgbb**, lo vamos a pegar en el Proyecto de **stackblitz** de la AG1. También añadiremos algo de código para separar en la página web la parte de AG1 de la AG2. La web tendrá el siguiente aspecto:

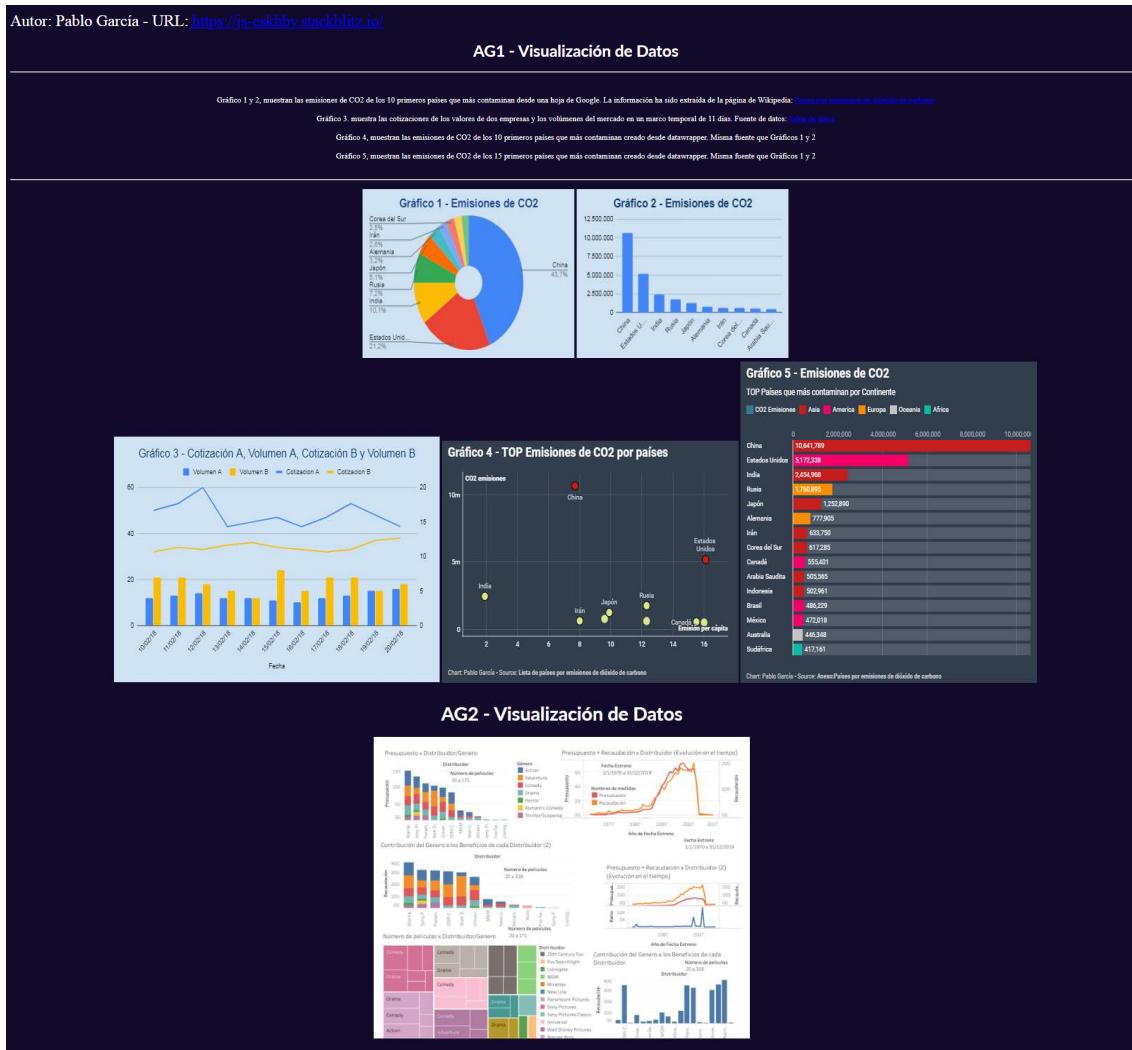


Gráfico - Dispersión con burbujas: Films por MDB Rating + IMDB Votes + Presupuesto (Subir Nota)

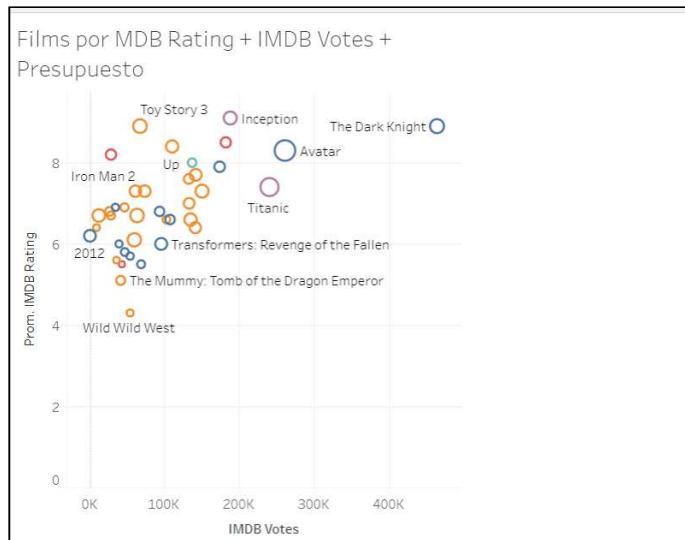
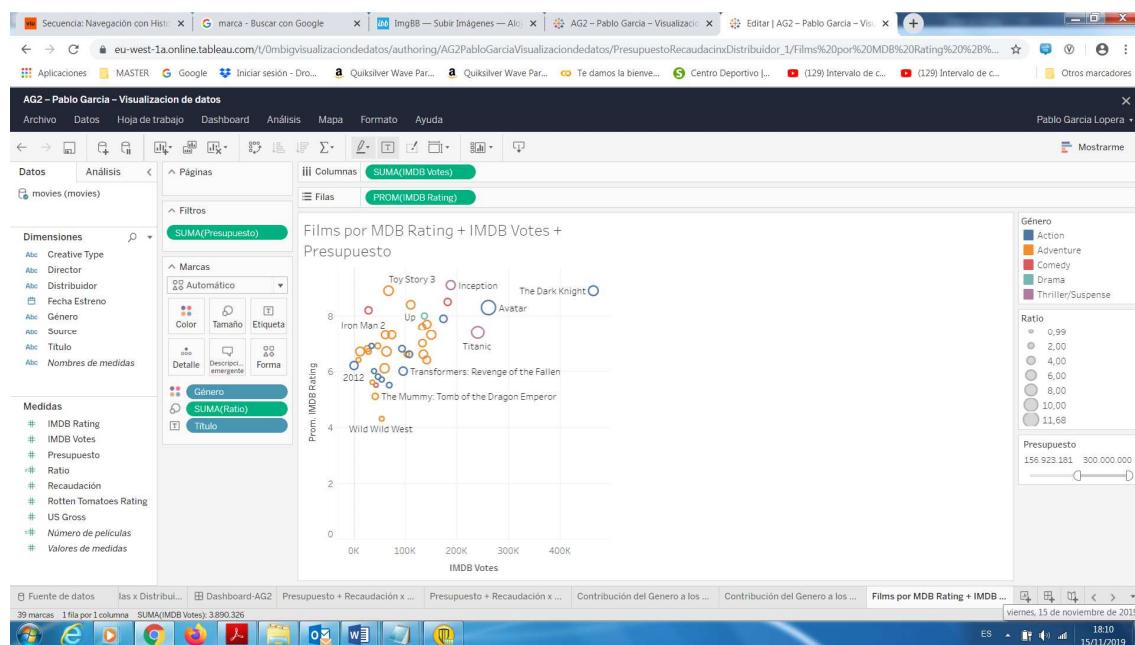
El gráfico debe tener las siguientes características

- 1 variable cualitativa:
 - Género(color)
- 3 variables cuantitativas:
 - Promedio de IMDB Rating (eje vertical)
 - Suma de IMDB Votes (eje horizontal)
 - Presupuesto (tamaño)

Para crear el gráfico hemos hecho las siguientes acciones:

- Tipo de Gráfico: **Burbujas**
- El campo **MDB Rating** en **Columnas** y con valor **Promedio**
- El campo **MDB Votes** en **Filas** y con valor **Suma**
- El campo **Género** en **Marcas (Color)**
- El campo **Ratio** en **Marcas (Tamaño)**
- El campo **Título** en **Marcas (Etiqueta)**
- El campo **Presupuesto** en **Filtros**
- Para que el gráfico salga igual que en el pdf del profesor, el filtro del campo **Presupuesto** debe estar entre **156.923.181** y **300.000.000** (en la imagen del pdf aparecen estos valores)

Este es el gráfico que obtengo (Igual que el del pdf del profesor):



Guardamos la imagen del gráfico en la aplicación **imgbb**, y el código **html** obtenido lo pegamos en el Proyecto de **stackblitz**. El enlace de la imagen subida es: <https://ibb.co/kx0RTJH>, y el código html:

```
<a href="https://ibb.co/kx0RTJH"></a><br /><a target='_blank' href='https://es.imgur.com/'>photo upload</a><br />
```

La web **stackblitz** tendrá este aspecto final:

