

«Talento Tech»

Data Analytics con Python

Clase 08





Clase N° 8 | Conceptos básicos

Temario

- Comenzaremos a utilizar una gran herramienta para la visualización de datos: Tableau.
- Exportaremos archivos de tipo .csv a Tableau para comenzar a crear gráficos interactivos.
- Conoceremos el área de trabajo y cómo utilizarla.
- Empezaremos a crear gráficos para nuestra presentación final.

Introducción a Tableau



En las clases anteriores aprendimos sobre el lenguaje de programación Python. Aprendimos a subir datasets y convertirlos en dataframe con la librería Pandas. Luego aprendimos a graficar nuestros datos con las librerías Matplotlib y Seaborn. En esta clase comenzaremos a crear nuestras presentaciones.

Cuando analizamos datos lo hacemos para obtener una respuesta a una pregunta. En general vamos a querer comunicar nuestros hallazgos. Para esto son útiles los gráficos que realizamos anteriormente. Pero no podríamos presentar nuestros datos solamente mostrando una imagen. Deberíamos juntar nuestros gráficos y poder mostrarlos a la vez. Poner un título a nuestro trabajo. Podríamos también agregar imágenes, jugar con los colores y demás.

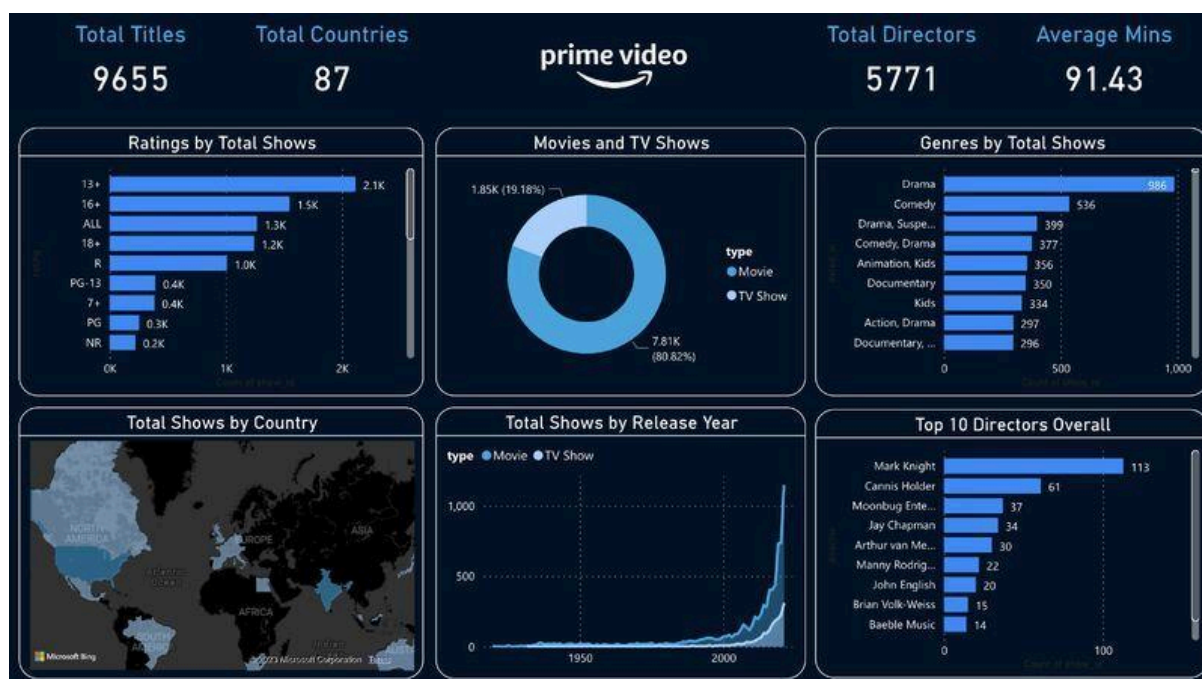
Una gran herramienta que nos permitirá hacer esto de manera más sencilla es Tableau. Tableau es un software que sirve para el análisis y la visualización de datos.

Entre sus características nos permite:

1. Visualizar datos de forma sencilla: Tableau nos permite convertir los datos en gráficos y mapas, pudiendo cambiar los colores, fuentes y agregarle funciones interactivas.
2. Explorar y entender los datos: Gracias a sus funciones interactivas podemos hacer click y observar distintos valores de los datos y aplicarles filtros sin necesidad de realizar cálculos.

3. Crear dashboards (tableros): Podemos crear dashboards o tableros. Los dashboards nos permiten visualizar toda la información y gráficos en una sola vista. Además podemos compartirlos o subirlos a un sitio web.

Algunos ejemplos de dashboards que podemos crear:



Buenos Aires aprende

Agencia de Habilidades para el Futuro












How to read this ?

This visualization intends to analyse the data in a more abstract way than usual, bringing attention to the 3 most important topics, as Family - Painting and Time.

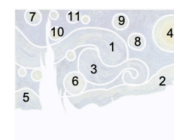
The entire ranking can be found here: [1](#)

The icons symbolize the number of words used on the letters exchanged during the year.

Click to highlight a topic

# Words	Family	Painting	Time
50			
10			
1			

These icons have been grouped into 10 zones to respect Vincent's painting. The most letters on that year and bigger the area's location. The picture below displays the full ranking by size.



For a more traditional ranking: [1](#)

The 3 circles below the # letters represent the most discussed topics by Vincent on that year. As other topics than the top 3 may have been discussed in the letters I still wanted to represent those even if in grey.



Buenos Aires aprende

Agencia de Habilidades para el Futuro

DUNE DYNASTY

Dune and Dune part 2 have some of the best ensemble casts in contemporary American cinema. Explore the films of four of DUNE's lead actors.

The network chart shows the main cast members clustered with their films. Actors are connected by the films they've worked on together.



Films that Pair Dune Costars

Aberdeen	Rampling	Skarsgård
Avengers: Infinity War	Bautista	Brolin
Boogie Woogie	Rampling	Skarsgård
Goya's Ghosts	Bardem	Skarsgård
Lady Bird	Chalamet	Henderson
Last Words	Rampling	Skarsgård
Little Women	Chalamet	Pugh
Melancholia	Rampling	Skarsgård
No Country for Old Men	Bardem	Brolin
Signs & Wonders	Rampling	Skarsgård
The Greatest Showman	Ferguson	Zendaya
Thor: Love and Thunder	Bautista	Skarsgård

TIMOTHEE CHALAMET



ZENDAYA



AUSTIN BUTLER



FLORENCE PUGH



Filmographies

(sorted by avg rating; max rating of 5)

HOVER mouse over bars in chart to see character from film.

Timothée Chalamet	Zendaya	Austin Butler	Florence Pugh
Interstellar 4.35	Spider-Man: No Way Home 4.10	Dune 4.00	Oppenheimer 4.15
Dune 4.00	Dune 4.00	Once Upon a Time in Hollywood 3.80	Dune 4.00
Call Me by Your Name 3.90	The Greatest Showman 3.75	Elvis 3.65	Puss in Boots: The Last Wish 3.95
Little Women 3.90	Spider-Man: Far from Home 3.70	The Bikeriders 3.60	Little Women 3.90
Enter Galactic 3.70	Spider-Man: Homecoming 3.70	Aliens in the Attic 2.70	Fighting with My Family 3.55
Lady Bird 3.70	Challengers 3.65	Dude 2.55	Midsommar 3.55
Beautiful Boy 3.65	Malcolm & Marie 3.30	The Dead Don't Die 2.75	Outlaw King 3.45
The King 3.65	Smallfoot 3.30	Aliens in the Attic 2.70	Lady Macbeth 3.40
Don't Look Up 3.60	Duck Duck Goose 2.85	The Intruders 2.25	Black Widow 3.35
Hostiles 3.60	Space Jam: A New Legacy 2.25	Yoga Hosers 2.10	The Wonder 3.30
The French Dispatch 3.55			Don't Worry Darling 3.15
Wonka 3.50			The Commuter 3.15
Bones and All 3.40			The Falling 2.65
Men, Women & Children 3.30			
Miss Stevens 3.30			
A Rainy Day in New York 3.25			
Hot Summer Nights 3.20			
Love the Coopers 2.90			
One and Two 2.60			
The Adderall Diaries 2.60			
Worst Friends 2.25			

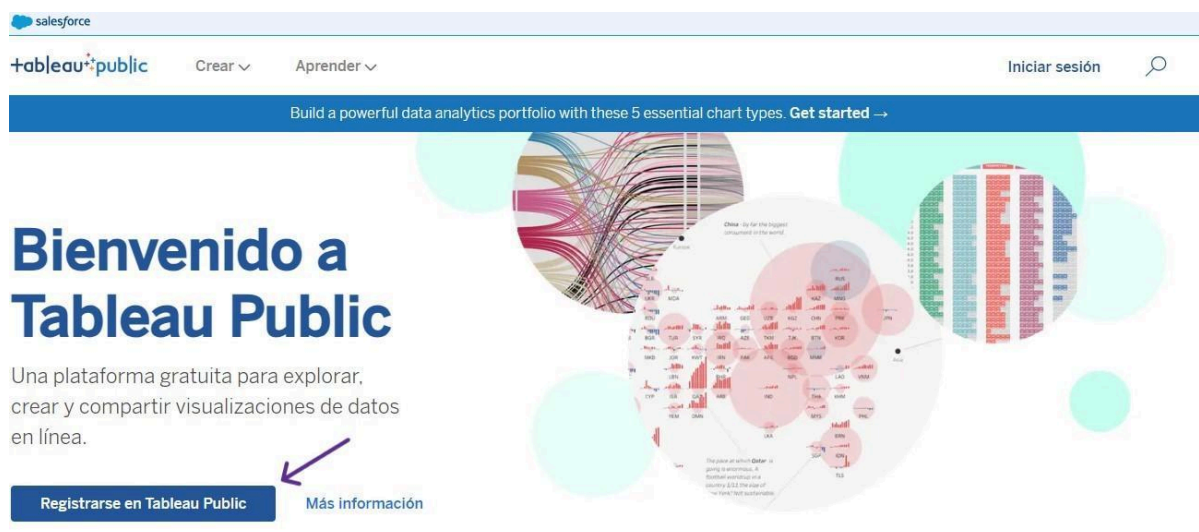
Let the spice flow #DataPlusMovies

Developed by: Josh Preston <https://www.joshpreston.com> Data: IMDB Images: Wikipedia

Conociendo Tableau Public: Crear cuenta

Para crearnos una cuenta y poder utilizar Tableau solo necesitamos un correo electrónico. Ingresaremos al sitio: <https://public.tableau.com/app/discover> para crearnos una cuenta.

Nos dirigimos a la opción “Registrarse en Tableau Public”.



Nos dirigirá a la página que vemos a continuación.

Cree su cuenta de Tableau

Nombre

Este campo es obligatorio.

Apellido

Correo electrónico

Contraseña

Confirmar contraseña

Las contraseñas distinguen entre mayúsculas y minúsculas, y deben tener entre 8 y 100 caracteres. Además, deben incluir al menos un número (0-9), una letra mayúscula, una letra minúscula y uno de los siguientes caracteres: ! " \ # \$ % & ' () * + , - . / : ; < = > ? @ [] ^ _ ` { | } ~ .

CREAR MI CUENTA

Completamos nuestros datos y creamos una contraseña para ingresar. Puedes elegir “usar contraseña segura”, eso te creará una contraseña que se usará para ingresar automáticamente si usas un mail de gmail.

Luego te llegará al correo electrónico con el cual te registraste un mail:



Agencia de Habilidades para el Futuro



Gracias por registrarse en Tableau. Estamos muy contentos de que forme parte de la comunidad de Tableau. Esperamos que vuelva a menudo para aprovechar los beneficios de los foros, videos de capacitación, grupos de usuarios y otros recursos.

Para activar la cuenta, haga clic en la siguiente dirección URL:

<https://id.tableau.com/token/38372c7fbd724e6a99502974330b805a>

Para su seguridad, el enlace de arriba vencerá al cabo de 48 horas. Si es necesario, puede solicitar un nuevo enlace aquí: <https://id.tableau.com/resendActivation>

Si usted no realizó esta solicitud, ignore este correo electrónico.

Para asegurarse de recibir futuras comunicaciones, agregue "tableau.com" a su lista de remitentes seguros.

Gracias nuevamente por registrarse en Tableau.

El equipo de Tableau

Activa la cuenta clickeando en el url del mail. Luego debes iniciar sesión.

Al iniciar sesión te pedirá que completes la región geográfica y aceptes los términos y condiciones.



[Iniciar sesión](#) [Registro](#) [Restablecer contraseña](#)

Tableau Public necesita más información

Seleccione su país o región.

Land/Region
Argentina

☒ Acepto los [Términos](#), la [Directiva de datos](#) y el [Código de la comunidad](#).
Al registrarse, confirma que está de acuerdo con que Salesforce trate sus datos personales tal como se describe en la [Declaración de privacidad](#).

FINALIZAR REGISTRO

* Estas preferencias se guardarán en su cuenta. De este modo, ya no se le volverán a pedir cada vez que inicie sesión en el futuro.

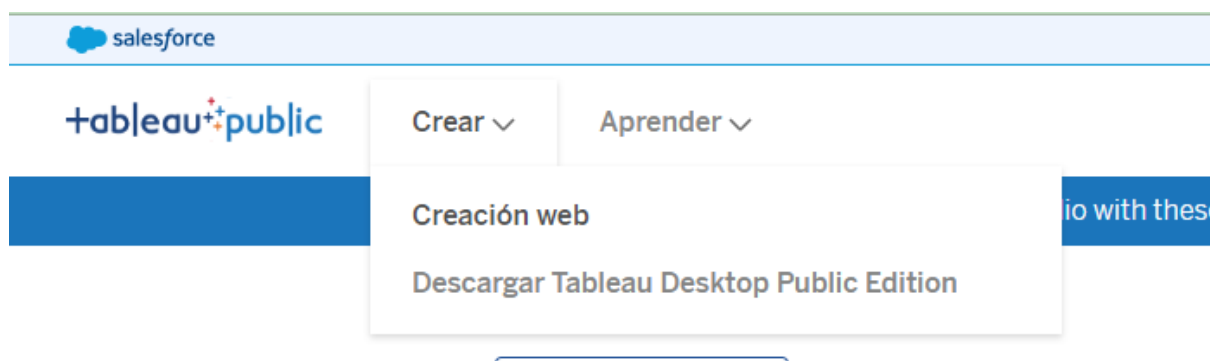
Finalmente clickeamos en "FINALIZAR REGISTRO".



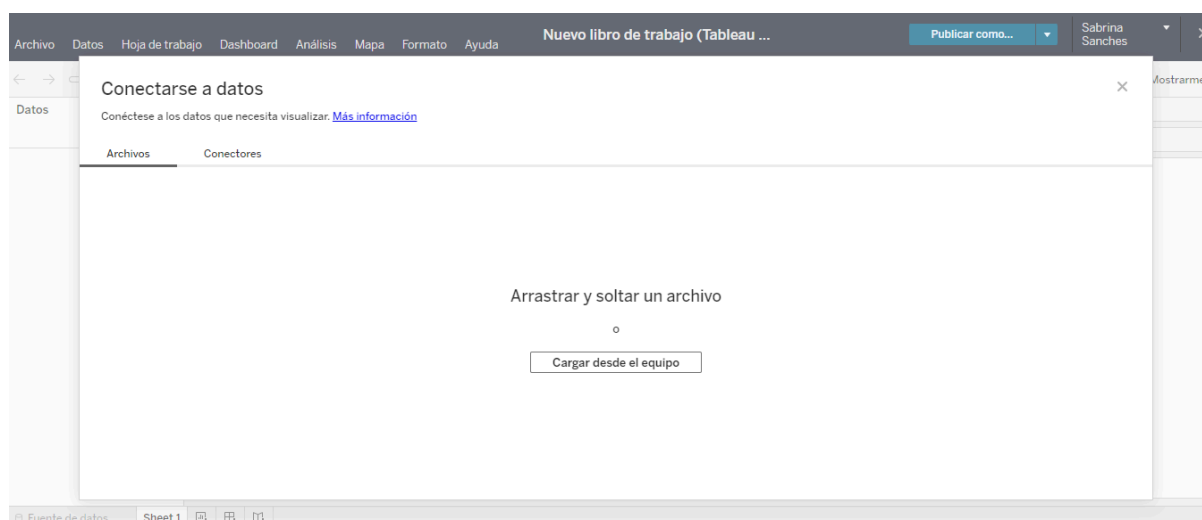
Si seguiste todos los pasos ya tendrás creada tu cuenta.
¡Ahora empecemos a crear!

Subir archivos a Tableau

Ya en nuestro perfil de Tableau nos dirigimos a la opción “Crear”. Como usaremos la versión Web no es necesario que descargues ningún software. En “Crear” seleccionamos “Creación web”.

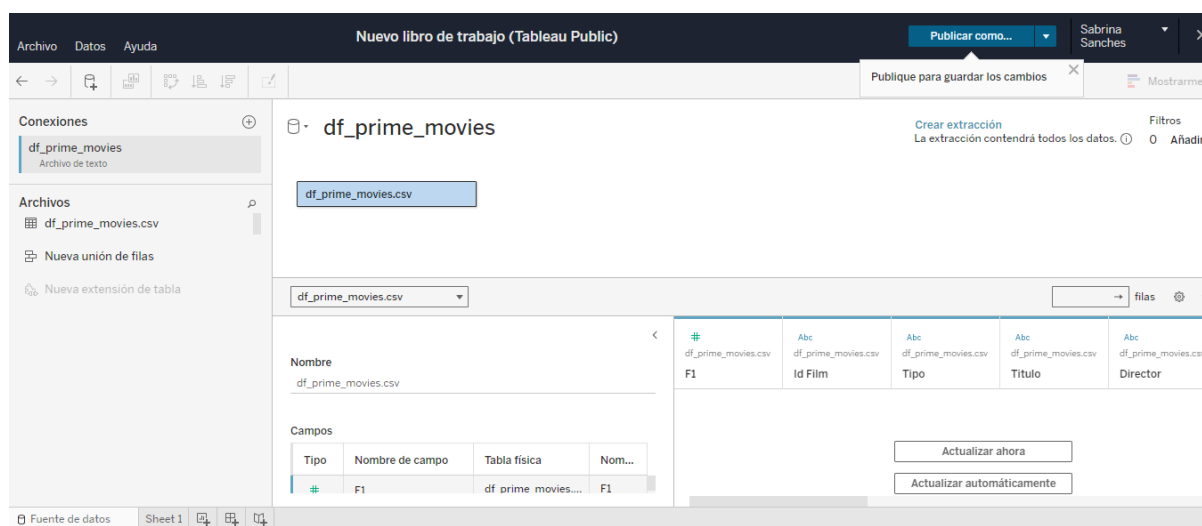


Al seleccionar “Creación web” se nos abrirá nuestro espacio de trabajo con una ventana emergente:



Ya trabajamos anteriormente con nuestro .csv de Títulos de producciones en Amazon Prime. Puedes subir el mismo .cvs o puedes descargar el dataframe como un .csv ya con las columnas modificadas (nombres cambiados) y filtradas (eliminar elementos repetidos o nulos). Lo realizamos en la clase 7.

Usaré el que obtuvimos al modificarlo el la clase anterior.



Nuevo libro de trabajo (Tableau Public)

Publicar como... Sabrina Sanches

Publicar para guardar los cambios

Conexiones

df_prime_movies
Archivo de texto

Archivos

df_prime_movies.csv

Nueva unión de filas

Nueva extensión de tabla

df_prime_movies

Crear extracción
La extracción contendrá todos los datos.

Filtros
0 Añadir

df_prime_movies.csv

filas

#	Id Film	Tipo	Titulo	Director
F1				

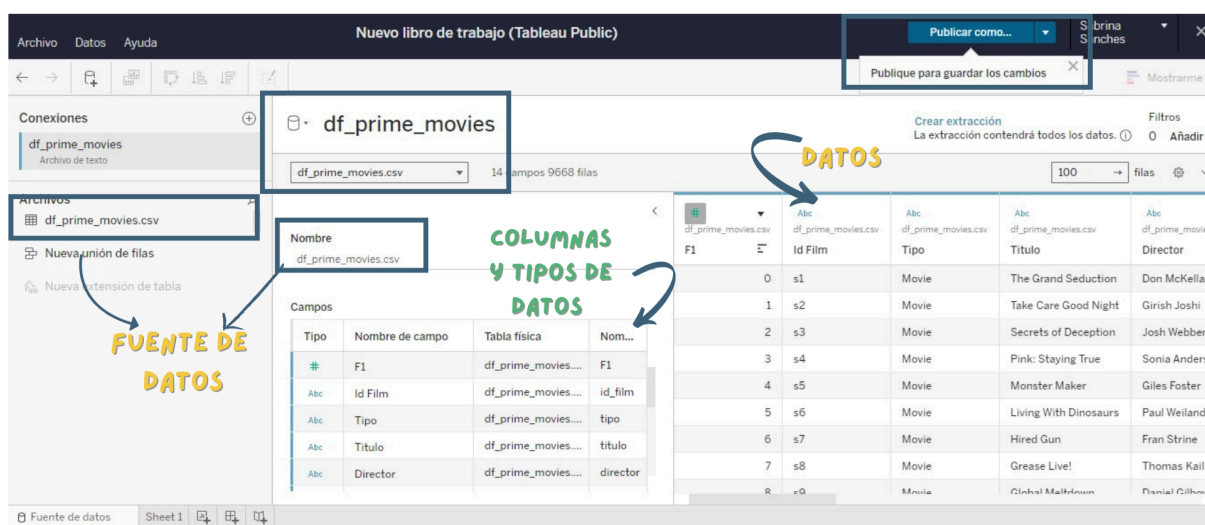
Actualizar ahora

Actualizar automáticamente

Fuente de datos Sheet 1

Nuestro archivo en .csv se llama “df_prime_video”, es lo que vemos en la parte superior. Aún no vemos los datos, para esto presionamos en el botón inferior: “Actualizar ahora”.

Luego de que los datos se actualicen nos mostrará el contenido de nuestro .csv.



The screenshot shows the Tableau Public interface with a new workbook titled "Nuevo libro de trabajo (Tableau Public)". The "Conexiones" pane on the left shows a connection to "df_prime_movies" (Archivo de texto). The "Archivos" pane shows "df_prime_movies.csv" selected. The "Fuentes de datos" pane shows the data source "df_prime_movies.csv" with 14 campos and 9668 filas. The "Campos" pane lists the fields: "Id Film", "Tipo", "Titulo", and "Director". The "Datos" pane shows a preview of the data with columns: "Id Film", "Tipo", "Titulo", and "Director".

Tipo	Nombre de campo	Tabla física	Nom...
#	F1	df_prime_movies....	F1
Abc	Id Film	df_prime_movies....	id_film
Abc	Tipo	df_prime_movies....	tipo
Abc	Titulo	df_prime_movies....	titulo
Abc	Director	df_prime_movies....	director

	Id Film	Tipo	Titulo	Director
0	s1	Movie	The Grand Seduction	Don McKellar
1	s2	Movie	Take Care Good Night	Girish Joshi
2	s3	Movie	Secrets of Deception	Josh Webber
3	s4	Movie	Pink: Staying True	Sonia Anders
4	s5	Movie	Monster Maker	Giles Foster
5	s6	Movie	Living With Dinosaurs	Paul Weiland
6	s7	Movie	Hired Gun	Fran Strine
7	s8	Movie	Grease Live!	Thomas Kail
8	s9	Movie	Global Maltdown	Daniel Gilhe

Veremos que del lado izquierdo aparecerá el nombre de nuestro archivo de datos. También aparece en la parte superior.

Luego tenemos dos secciones. Por un lado vemos los nombres de las columnas como: "Nombres de campo". Nos muestra el tipo de dato, el nombre, la tabla de la cual proviene el campo y su nombre.

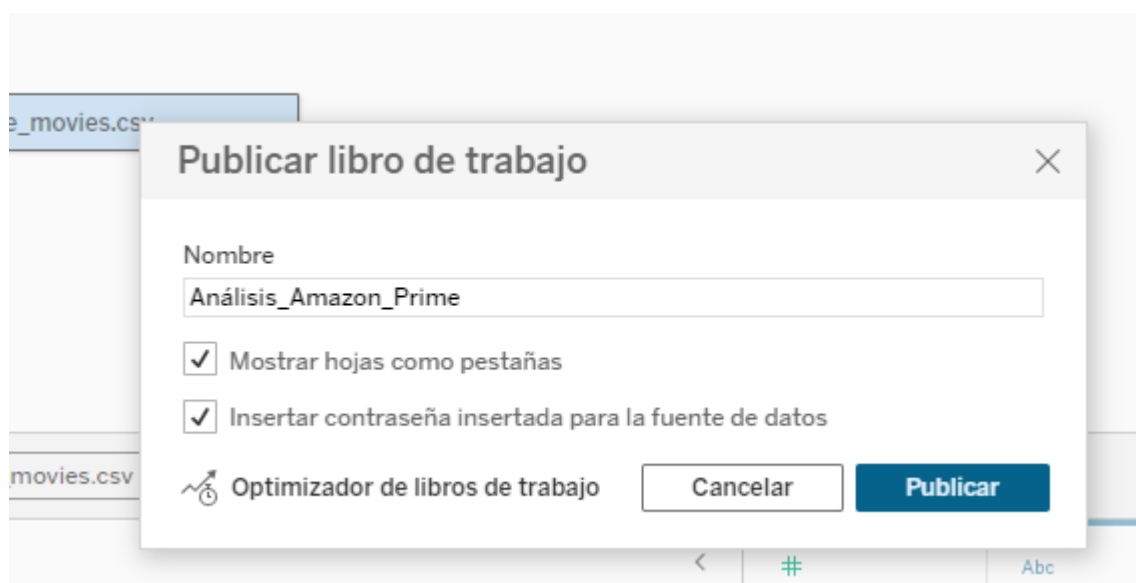
Del lado derecho podemos ver los datos de nuestro .csv. En la parte superior de cada columna vemos el tipo de dato, la fuente y el nombre de la columna.

En la siguiente imagen puedes ver el ícono que representa cada tipo de dato.

Icono	Tipo de datos
Abc	Valores de texto (cadena)
📅	Valores de fecha
🕒	Valores de fecha y hora
#	Valores numéricos
T/F	Valores booleanos (solo relacionales)
🌐	Valores geográficos (se usan en mapas)
🖼️	Rol de imagen (usado con URL de enlace de imagen)
🗺️	Grupo de clústeres (se usa con Buscar clústeres en datos)

Notarás que no aparece un botón de guardado. Para ir guardando nuestro trabajo debemos ir al botón superior “Publicar como...”

Puede que al presionar “Publicar como...” tarde un poco, dependerá del tamaño de tu dataset.



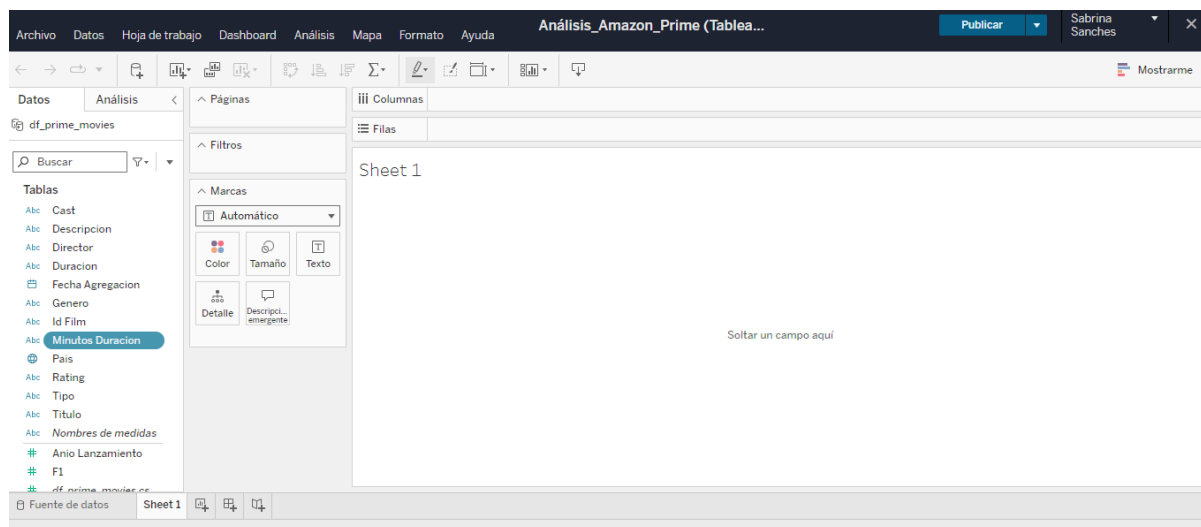


Le ponemos un nombre a nuestro libro de trabajo y tildamos las opciones “Mostrar hojas como pestañas” e “Insertar contraseña insertada para la fuente de datos”. Luego presionamos en “Publicar”.

Ya tienes tu primer proyecto publicado en Tableau. Por el momento solamente hemos cargado el dataset. Examinemos ahora nuestro espacio de trabajo.

Conociendo el espacio de trabajo

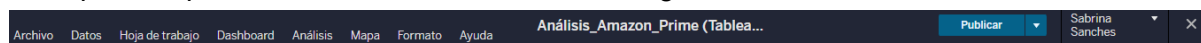
Luego de publicar nuestro proyecto nos dirigirá a una vista como la siguiente:



Analicemos qué tenemos en esta vista.

Barra menú

En la parte superior encontramos la barra de navegación.



Archivo: Nos permite publicar o cerrar el archivo.

Datos: Nos permite agregar una nueva fuente de datos o modificar la que ya tenemos.



Hoja de trabajo: desde aquí podemos seleccionar entre las diferentes hojas de trabajo dentro del libro. Cada hoja puede contener diferentes gráficos, tablas o visualizaciones. También puedes añadir nuevas hojas desde este menú.

Dashboard: Permite seleccionar y trabajar con dashboards dentro del libro. Los dashboards combinan múltiples hojas de trabajo y elementos interactivos en una sola vista. También puedes añadir nuevos dashboards o ajustar los existentes desde aquí.

Análisis: Ofrece herramientas y funciones para realizar análisis avanzados dentro de tu visualización. Incluye opciones para aplicar filtros, agregar cálculos, y usar otras herramientas de análisis.

Mapa: Si estás trabajando con datos geospaciales, este menú te permite ajustar y configurar opciones relacionadas con los mapas en tu visualización, como la proyección, capas, y más.

Formato: Permite ajustar el formato de los elementos en tus hojas y dashboards, como fuentes, colores, bordes y alineaciones. Aquí puedes cambiar cómo se presenta visualmente tu contenido.

Ayuda: Nos lleva a la documentación de Tableau.

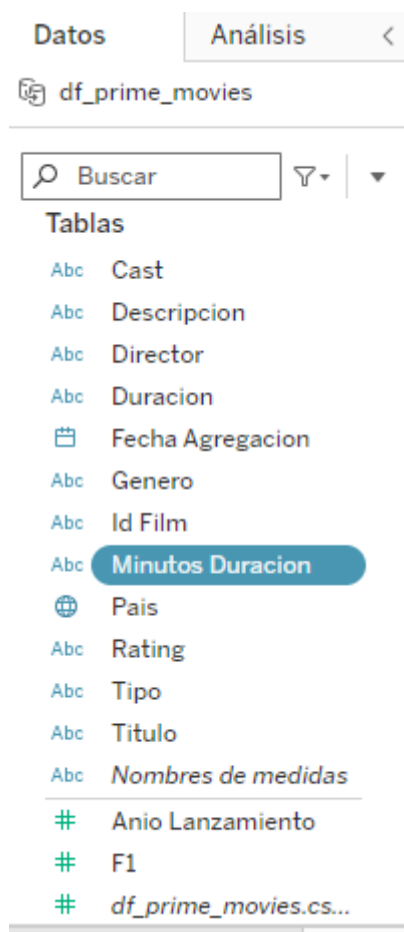
Barra de herramientas



Proporciona acceso rápido a herramientas y comandos frecuentemente usados, como *Guardar*, *Deshacer*, *Rehacer*, *Nueva hoja de trabajo*, *Ejecutar*, *Filtrar*, *Agrupar*, entre otros. También puede incluir botones para cambiar el tipo de visualización.

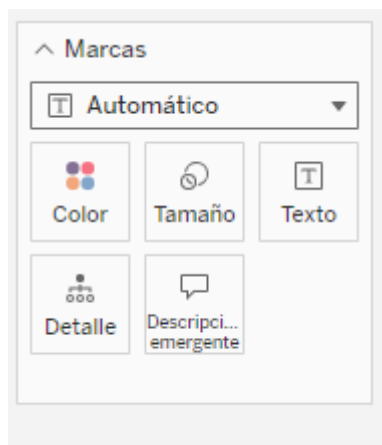
Panel de Datos (Data Pane)

Muestra las fuentes de datos conectadas y sus campos. Aquí puedes arrastrar y soltar campos para crear visualizaciones. Está dividido en Campos de Datos, Campos Medidos y Campos Dimensionales. Verás una línea dividiendo los campos categóricos de los numéricos.



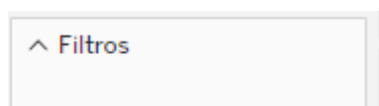
Panel de Marcas (Marks Card)

Ofrece opciones para personalizar la apariencia de los datos en la visualización. Aquí puedes ajustar aspectos como color, tamaño, forma, etiqueta y detalle de los datos. Es donde se configuran las propiedades visuales de los elementos en la hoja.



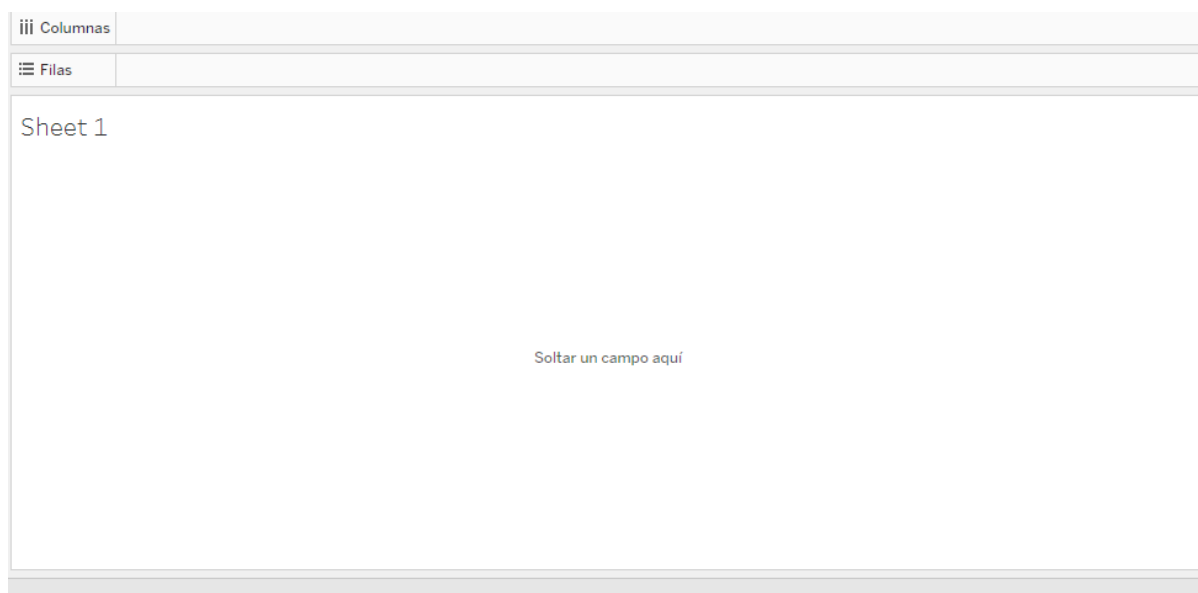
Panel de Filtros (Filters Pane)

Permite aplicar filtros a los datos para limitar la información mostrada en las visualizaciones. Puedes arrastrar campos aquí para establecer filtros específicos.



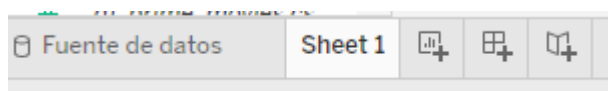
Área de Diseño (Canvas)

Es el lugar donde se construyen y visualizan las hojas de trabajo y dashboards. Aquí puedes arrastrar los campos desde el Panel de Datos y organizarlo en diferentes tipos de gráficos y visualizaciones.



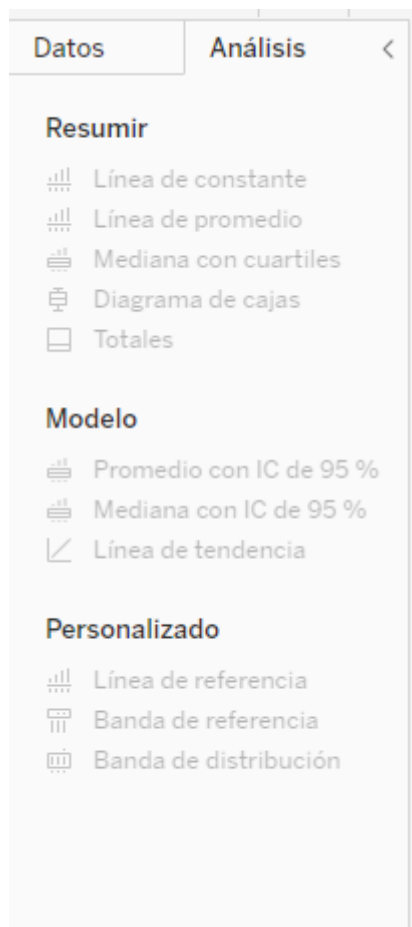
Panel de Hojas de Trabajo y Dashboards (Sheet Tabs)

Muestra todas las hojas de trabajo, dashboards e historias en el libro de trabajo. Puedes hacer click en las pestañas para cambiar entre diferentes hojas y dashboards, o agregar nuevos desde aquí.



Panel de Análisis (Analysis Pane)

Ofrece herramientas y opciones para realizar análisis avanzados, como líneas de referencia, agrupaciones, y calculadoras.



Generar gráficos

En nuestro anterior análisis del dataset de producciones realizamos algunos gráficos. Intentemos replicar esto usando Tableau.

Cada gráfico lo realizaremos en una hoja distinta y le pondremos su nombre correspondiente.

Para generar los gráficos arrastraremos los campos desde datos (al arrastrar el campo tiene una píldora) y lo posicionamos donde necesitamos.



Gráfico de barras

Para generar un gráfico de barras debemos arrastrar las cápsulas de datos a filas y columnas, según queramos que nuestro gráfico sea horizontal o vertical.

Queremos graficar cuantas veces aparece cada director. Para eso arrastramos “Director” al campo *Filas*. Nota que “Director” no es una variable numérica. Ahora necesitamos una medida que nos calcule cuántas veces aparece cada director, es decir, un campo numérico. Tomamos la píldora “Director” y la arrastramos al campo *Columnas*. Director sin embargo no es una variable numérica, así que debemos convertirla de alguna manera. Lo que queremos es la cantidad de producciones según el director, es decir la cantidad de veces que aparece un director. Nos posicionamos sobre la píldora que acabamos de arrastrar y desplegamos el menú. En “Medida” seleccionamos la opción “Recuento”.

Una vez generado el gráfico podemos excluir valores. Excluimos los nulos que son mayoritarios.

Además puedes modificar el nombre de la hoja y el título.

[poner video gráfico de barras simple](#)

Editamos ahora los colores, marcas y apliquemos un filtro.

Para modificar los colores debemos pensar en función de que queremos que nuestros colores sean diferentes, si es un valor numérico o una categoría, luego debemos arrastrar esa píldora a la Marca “Color”. Cambiemos el color según el director.

Luego agreguemos algunas etiquetas sobre el gráfico. Las etiquetas son valores que se escriben en el gráfico. Podríamos elegir “Director” o el valor de recuento. Como los directores ya se ven en los ejes del gráfico usemos el valor numérico como etiqueta. Arrastramos el dataset a “Etiqueta”.

Queremos además tener un filtro sobre lo que nos muestra el gráfico, que nos muestre un intervalo de datos. En este caso filtraremos por la cantidad de producciones, así que arrastramos la píldora con nuestro dataset a filtro.

[poner video grafico_barras_filtro_colores](#)

Gráfico de torta

Creemos una nueva hoja. En esta hoja realizaremos un gráfico de torta. Graficamos la proporción de cada tipo de producción: *TV Show* o *Movie*. Primero debemos arrastrar la



píldora con el nombre de nuestra fuente de datos al espacio de trabajo. Como queremos un gráfico de torta en “Marca” elegimos la opción “Circular”. Luego queremos que se diferencie por el “Tipo”, entonces arrastramos la píldora “Tipo” a “Colores”. También debemos indicar que el ángulo debe ser respecto a la píldora con el nombre de nuestra fuente de datos.

[video](#)

Mapa

[video](#)

Tablas

A menudo ciertos datos vamos a querer mostrarlos en tablas en lugar de gráficos. Veamos cómo podemos generar tablas a partir de los datos.

[video](#)

Extras

Te dejo algunos gráficos extras que puedes realizar con Tableau.

[video](#)



Desafío N° 8:

Crea una cuenta en [Tableau](#). Inicia sesión. Exporta el dataset de producciones de Amazon que trabajamos en las clases anteriores. Puedes exportar el dataframe que obtuviste luego de cambiar los nombres de las columnas y hacer el análisis (debes descargarlo como .csv). Crea en una hoja un gráfico de torta, distinguiendo tipos de producciones (TV Show o Movie), agrega un título y modifica los colores. Modifica el nombre de tu proyecto. Recuerda que para que tu trabajo quede guardado debes publicarlo.



Buenos Aires
aprende
Agencia de Actividades para el Futuro

BA Buenos
Aires
Ciudad