

Apoyo para docentes - Ejercicios guiados

Visualización y comunicación con Tableau Public

(Parte I)

La sesión se separa esencialmente en 3 momentos. El inicio y el cierre de la sesión se realizan utilizando la presentación de la clase y corresponden a una introducción a Tableau, sus características e instalación; y el cierre de la clase donde se resumen las ideas fuerza y se presenta el desafío de la semana. Toda la parte intermedia de la sesión se realiza directamente en Tableau, sin utilizar la presentación; los pasos a seguir en esta etapa son los que se describen en este documento.

Luego de realizar la explicación de las características de Tableau e instalarlo, abra el programa y replique los pasos que se describen. Mencione a los estudiantes que los contenidos se encuentran desarrollados en la Guía, si buscan repasar y/o profundizar; la sesión sincrónica debe enfocarse fundamentalmente en la descripción de la forma de uso del programa. Se debe destacar también que existen funciones de Tableau que ya están cubiertas por otros programas, por lo que privilegiaremos que tareas como la limpieza de datos se realice en Python, por ejemplo.

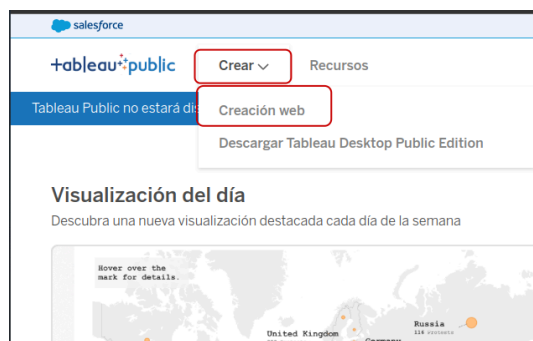
1. Conexión a una fuente de datos

La conexión se hará mostrando que es posible hacerlo tanto a archivos locales almacenados en el computador, como a archivos de Google Drive (incluyendo Google Sheets). Para esto es imprescindible mencionar que los archivos de Google Drive deben encontrarse en modo público para poder realizar la conexión con ellos.

Considere que este trabajo puede realizarse en la versión web como en la versión Desktop y presentan ligeras diferencias; este primer caso conviene realizarlo en ambas plataformas para que los estudiantes puedan familiarizarse. Para esto puede mostrar los siguientes pasos:

En la web:

1. En la página de Tableau Public, hacer clic en "Crear > Creación web" en la pantalla de inicio. Se abrirá la ventana "Conectarse a datos"



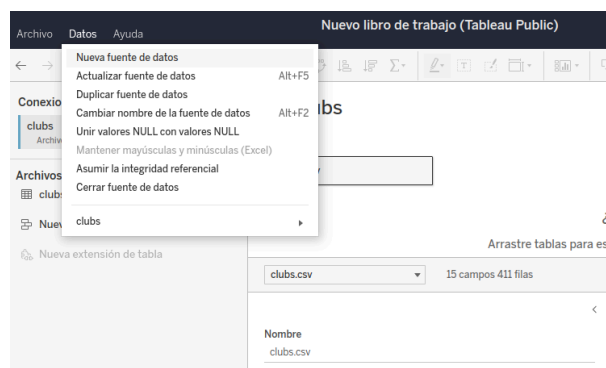
Fuente: <http://public.tableau.com>

2. En la ventana "Conectarse a datos", seleccionar "Cargar desde el equipo" (si se trata de un archivo que tengas almacenado en el computador) o haz clic en la pestaña "Conectores", para seleccionar un archivo de Google Drive (como una hoja de Google Sheets)



Fuente: <http://public.tableau.com>

3. Validar los datos cuando se solicite, y luego se debe navegar hasta la ubicación del archivo u hoja que se desee utilizar. Al seleccionarlo, tras un breve tiempo se cargará.
4. Si fuese necesario cargar más archivos, puede hacer clic en la pestaña "Datos > Nueva fuente de datos"



Fuente: <http://public.tableau.com>

En Desktop:

1. Abra Tableau Desktop. En la parte izquierda verá el Menú “Conectar”



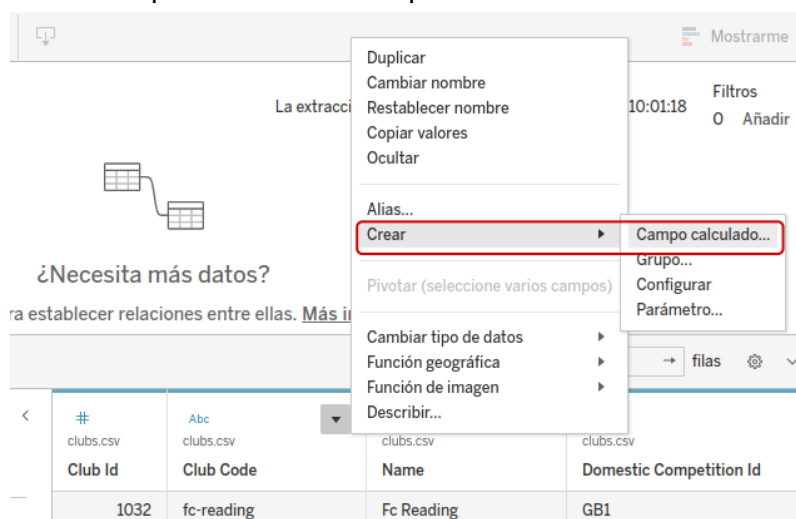
2. Si desea cargar un archivo csv guardado en su computador, haga clic sobre “Archivo de texto” en el submenú “Archivo”. Si va a cargar una hoja de Google Sheets o algún otro archivo que tenga en Google Drive, haga clic sobre “Más...” en el submenú “A un servidor” y escoge “Google Drive”.
3. Seleccione el archivo que desee, ¡y listo!

Mencione a los estudiantes que en adelante mostraremos los procedimientos correspondientes a la versión web. La versión Desktop presenta muy pequeñas diferencias, que indicaremos cuando corresponda.

2. Creación de un campo calculado

El siguiente procedimiento que mostraremos es la creación de campos calculados. Se utilizará para esto el archivo "Notas - Notas.csv". Comience por cargar el archivo en un Libro de trabajo de Tableau Public, y a partir de ello:

1. Muestre que en las columnas es posible hacer clic sobre las celdas de encabezado y observar el tipo de datos que contiene cada campo. Relacione con lo que los estudiantes ya aprendieron de SQL y los Dataframes de Python: numéricos, strings, booleanos, etc.
2. Luego, puede aplicar los siguientes pasos para crear un campo calculado:
 - a. En la pestaña "Fuente de Datos", sitúe el mouse sobre una columna cualquiera, aparecerá una flecha (▼), haga clic y se desplegará un menú. Seleccione la opción "Crear > Campo calculado..." en el menú desplegado.



Fuente: <http://public.tableau.com>

- b. En el cuadro de diálogo "Campo calculado", asigne el nombre al campo y escriba la fórmula que desea utilizar para calcular los valores del campo. En nuestro caso, utilizaremos la fórmula " $[(\text{Notas Promedio Lenguaje}) + (\text{Notas Promedio Matemáticas})]/2$ " y la guardaremos con el nombre "Promedio" (ya que la utilizaremos más adelante). Muestre a los estudiantes que el nombre de los campos se sugerirá cuando los vaya escribiendo.
- c. Muestre que, además, es posible crear campos calculados utilizando funciones, que utilicen como parámetro uno de los campos de nuestra tabla. Así, por ejemplo, si escribimos la fórmula " $\text{LOWER}([\text{Nombre}])$ " obtendremos una nueva columna con todas las celdas del campo "Nombre" en minúsculas.
- d. Haga clic en "Aceptar" para crear el campo calculado y agregarlo a la lista de campos de la fuente de datos.
- e. Arrastre el campo calculado desde la lista de campos de la fuente de datos a la vista de hoja de trabajo para utilizarlo en el posterior análisis y visualización.

Puede mencionar a los estudiantes que hay otras fórmulas que se pueden utilizar; para ello cargue el archivo “Funciones” en Tableau y cree los siguientes campos calculados:

Nombre	Función
Minúsculas	LOWER([NOMBRE])
Años_empresa	=DATEDIFF('year',[Fecha de ingreso],NOW())
Dif_años	=MAX([Edad]-[Edad])

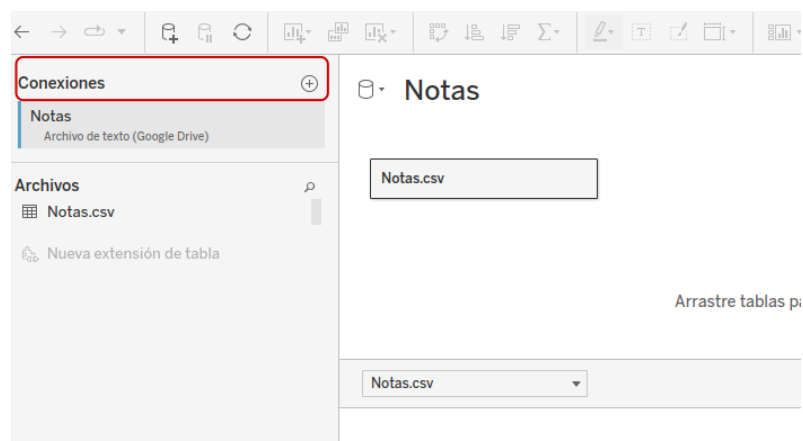
3. Analice junto con los estudiantes qué hace cada función, notando que al crear un campo podemos usar funciones relacionadas con otros campos, con el propio campo o incluso funciones sin parámetros, como NOW(). Preste especial atención a este primer ejemplo, ya que se utilizará en el desafío.
4. NO PASE DEMASIADO TIEMPO EN ESTO, e insista a los estudiantes en que la limpieza y preparación de datos será mejor realizarla siempre con otros programas.

3. Relaciones entre tablas

Para comenzar con este contenido, se puede recordar a los estudiantes la relación entre tablas que se puede realizar en SQL mediante los comandos JOIN. Les puede solicitar que mencionen algunos ejemplos de JOIN, enfatizando en cuál era su utilidad. Puede mencionar tanto su uso en SQL como en la creación de tablas pivote, en Python.

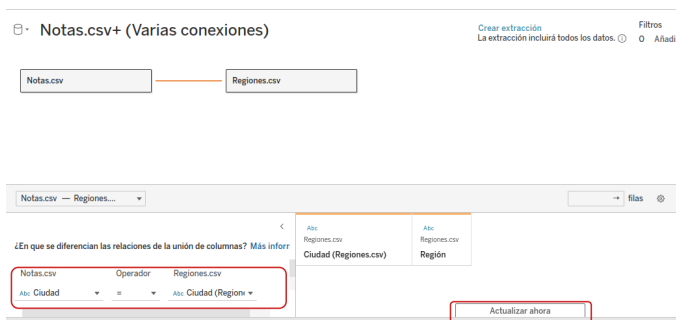
Mencione luego que en Tableau realizaremos una relación entre tablas pero sin unir las, lo que es esencialmente diferente al trabajo realizado en SQL. Para verlo, se sugiere añadir el archivo “Regiones” a nuestro libro (donde ya hemos cargado el archivo “Notas”), para que Tableau inmediatamente detecte que entre los dos archivos hay un campo en común. Para esto aplique los siguientes pasos:

1. Abra Tableau Public, crea un nuevo libro de trabajo y cargue en él “Notas.csv”
2. Haga clic en la pestaña “Conexiones” y cargue un segundo archivo. En nuestro caso será “Regiones.csv”.



Fuente: <http://public.tableau.com>

3. Bajo la pestaña “Conexiones” aparecerán ahora las dos bases de datos. Si hace clic sobre “Regiones” en la pestaña “Archivos” aparecerá “Regiones.csv”. Haga clic sobre él y muévelo hacia la derecha, para ubicarlo al lado de “Notas.csv”



4. Muestre que Tableau detecta de inmediato que hay un campo en común (Ciudad). Presionamos “Actualizar Ahora” y obtendremos una unión entre las dos tablas. Mencione que pueden calcular nuevos campos en esta tabla generada (puede mostrar la creación de uno, que se hace de la misma forma que ya han visto).

4. Nuestra primera visualización - Elementos de Tableau

Por motivos metodológicos, en la Guía se explican los elementos de Tableau para luego construir nuestra primera visualización. Durante la clase, sin embargo, se sugiere comenzar por la construcción de la visualización para que este proceso de pie a conocer los diferentes elementos y explicarlos en la medida en que se vayan necesitando. La explicación, por lo mismo, será más superficial durante la sesión siendo responsabilidad posterior de los estudiantes profundizar en estos tópicos.

Utilizaremos el archivo “Notas - Notas.csv”, que debe ya haber sido cargado y creados los campos de la Actividad Guiada N°1 (y que reproducimos aquí). Explique a los estudiantes que vamos a crear nuestra primera visualización que, en principio, corresponderá a una tabla. Para ello aplique los siguientes pasos:

1. Abra la hoja de trabajo y arrastre el campo “Ciudad” a las Filas y el campo “Genero_V1” a las Columnas. Además, arrastre el promedio recién creado a la Marca de Texto en el estante de Marcas.



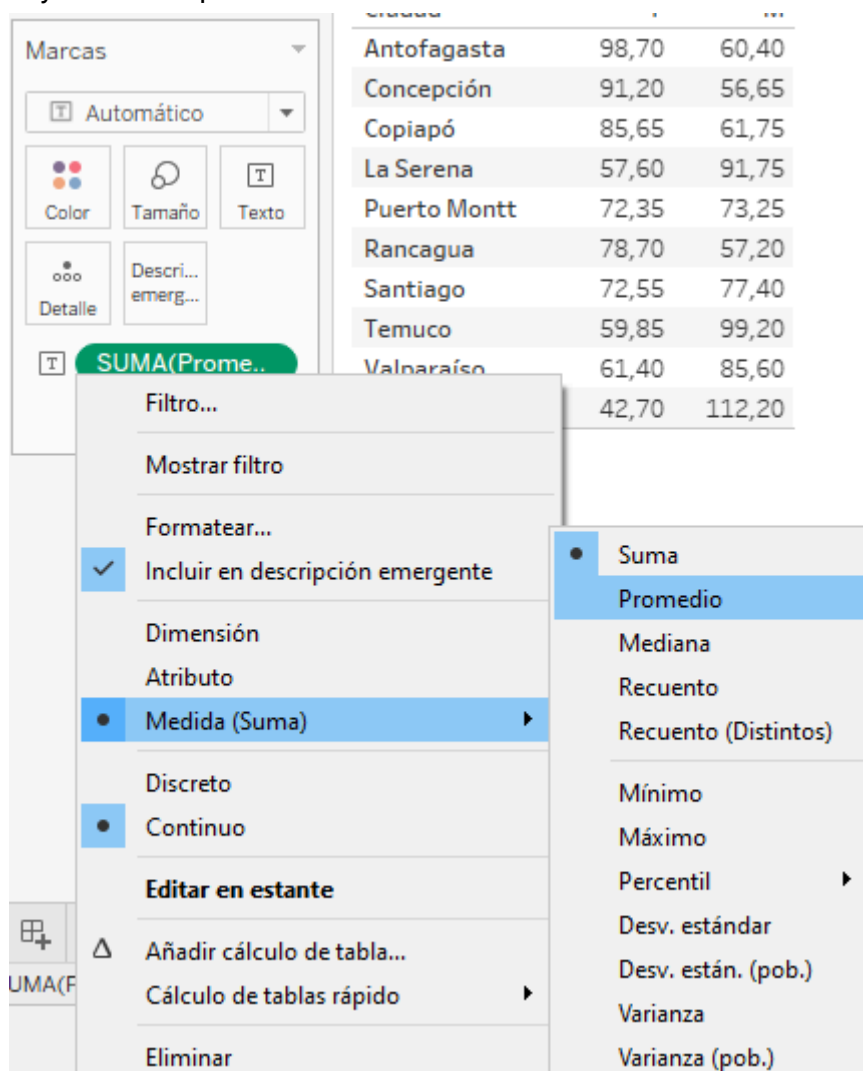
Fuente: <http://public.tableau.com>

Aproveche de mostrar a los estudiantes que:

- Hemos utilizado las filas y columnas. Explique que en Tableau estas son los “estantes” principales de la hoja de trabajo, donde los usuarios pueden arrastrar y soltar los campos de datos para definir las dimensiones y medidas y así definir la estructura básica de la visualización.
- A partir de lo anterior, vendrá la pregunta natural de qué son las dimensiones y medidas. Explique que se trata de dos tipos de campos que se utilizan para analizar y visualizar datos.
 - Las dimensiones son campos que contienen información descriptiva o categorías, es decir, variables cualitativas. Se utilizan para agrupar, filtrar y ordenar los datos, y se colocan generalmente en los estantes de fila o columna en Tableau.
 - Por otro lado, las medidas son campos que contienen valores numéricos que se pueden utilizar para realizar cálculos, como sumas, promedios, máximos o mínimos. Es decir, corresponde a variables cuantitativas. Las medidas se utilizan para construir gráficos y tablas que ilustran patrones o tendencias en los datos. Se colocan generalmente en los estantes de tamaño, color o forma en Tableau.
- Mencione que se puede transformar una dimensión en una medida y viceversa utilizando funciones de agregación que verán en la sesión asincrónica. Por ejemplo, se puede convertir una dimensión en una medida contando el número de registros en

cada categoría de la dimensión. De esta manera, se puede utilizar la dimensión en cálculos numéricos como si fuera una medida. Por otro lado, para convertir una medida en una dimensión se pueden convertir los valores numéricos de la medida en rangos o categorías discretas para así agruparlos. Esto permite que la medida se utilice como una dimensión para agrupar los datos en diferentes categorías.

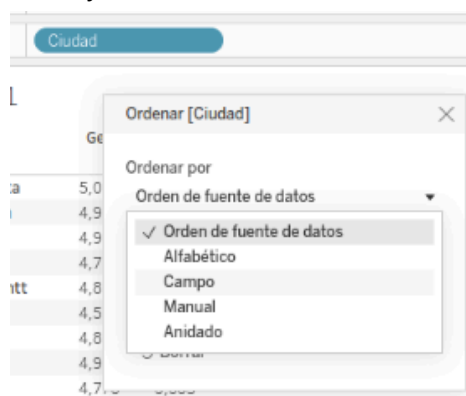
- Puede reforzar la idea anterior con un ejemplo: la edad de una persona es una variable cuantitativa. Sin embargo, si en una muestra agrupamos las diferentes edades y contamos la cantidad de personas que registran cada una, estamos abordándola como variable cualitativa.
- Volvemos a la visualización ahora, en que hemos arrastrado el promedio recién creado a la Marca de Texto en el estante de Marcas. Automáticamente, Tableau creará el campo como SUMA(Promedio); haga clic derecho para ir a la opción Medida y cambiarlo por Promedio.



Fuente: <http://public.tableau.com>

Explique que lo que hemos realizado es permitir que Tableau interprete y muestre correctamente este nuevo campo. Hemos utilizado ahora el estante de marcas, y conviene explicar qué es, señalando que es donde se definen las marcas (puntos, líneas, barras, etc.) que se utilizarán para representar los datos en la visualización. En nuestro caso, ya que nuestra visualización es una tabla lo que queremos es que la marca sea el “texto”, es decir, el valor del promedio (no una marca o símbolo).

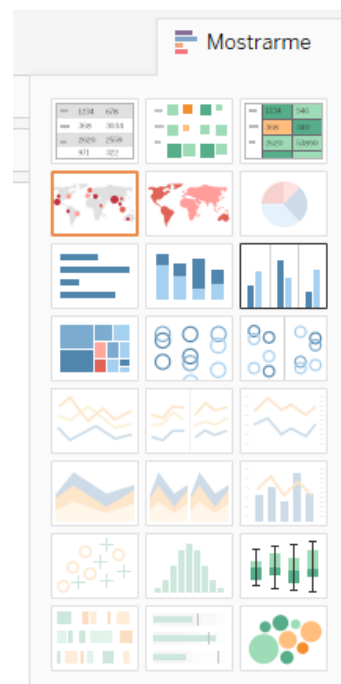
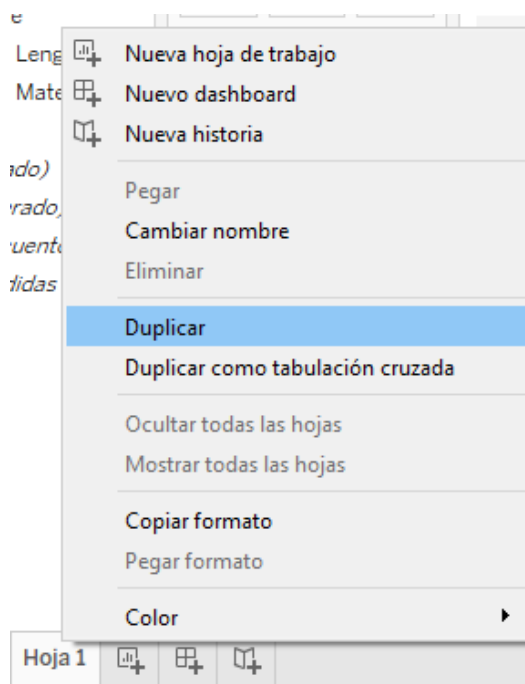
Un aspecto importante es que puede mostrar que, por defecto, los datos aparecen en el orden propio de la fuente de datos de la que provienen. Si considerara que es más cómodo que estos se vean en orden ascendente o descendente, por ejemplo, haga clic en la flecha de “Ciudad” en el estante de Filas, y seleccione “Ordenar”. Obtendremos el siguiente Menú:



La opción “Orden de fuente de datos” es la que aparece por defecto. La opción “Alfabético” es obvia, y podremos seleccionar si queremos que sea ascendente o descendente. “Campo” nos permite ingresar un criterio para el orden; hacemos clic aquí y nos permitirá elegir el campo por el que queremos que se ordene (en nuestro caso, el promedio recién llevado a las marcas). Ahora podemos escoger las opciones que deseemos. Puede relacionar con lo que se vio en tablas dinámicas, en la sesión anterior.

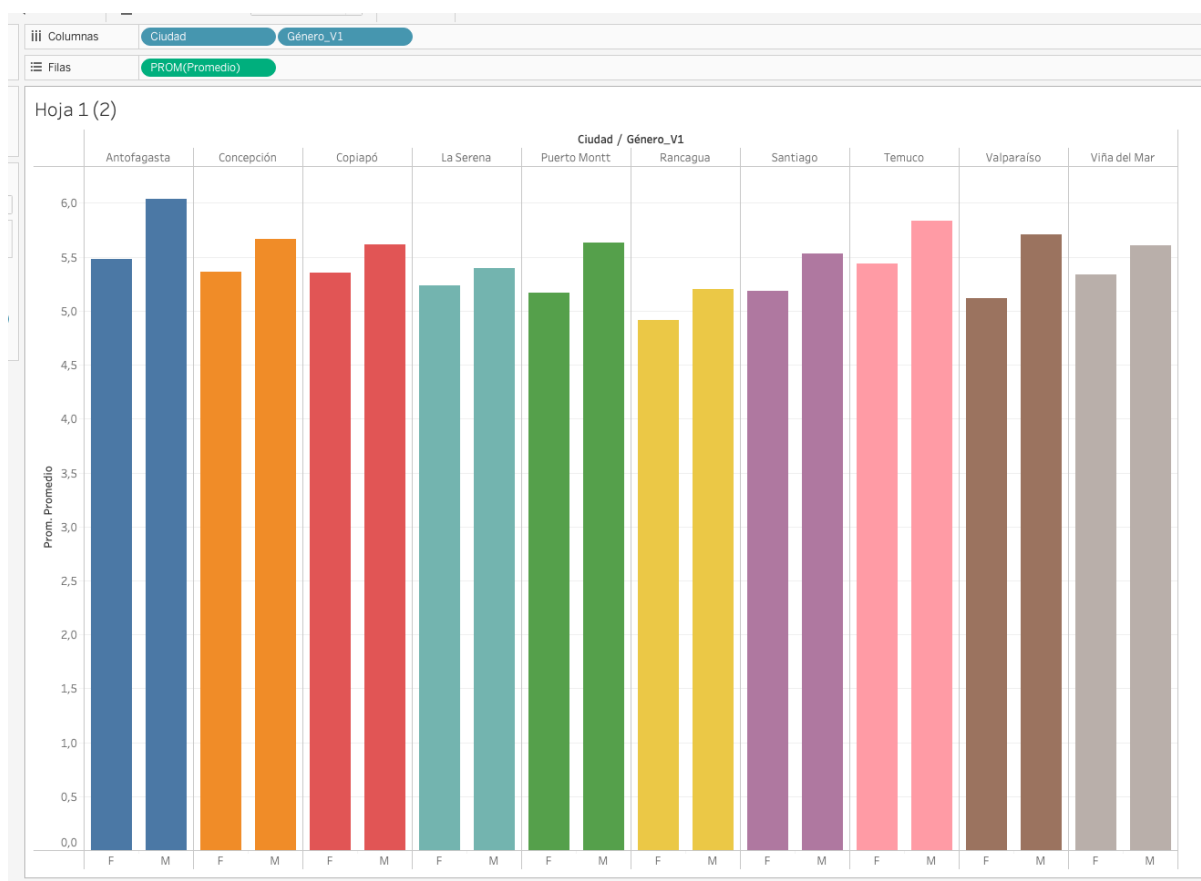
Precise que otros elementos se irán viendo en la medida en que se necesiten, o en la sesión asincrónica.

- El siguiente paso será visualizar los mismos datos en un gráfico. Duplique la hoja actual mostrando el procedimiento para hacerlo. Luego, en la esquina superior derecha, seleccione la opción “Mostrarme” y luego “Barras paralelas”.



Fuente: <http://public.tableau.com>

- Mencione a los estudiantes que, por defecto, Tableau separa por Género y luego por Ciudad. Por ello, cambiaremos el orden en las columnas por Ciudad y luego Género, y llevamos "Promedio" a las filas como se ve en la imagen. El resultado debe ser como el siguiente:



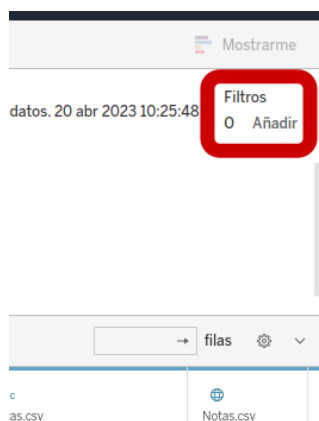
Fuente: <http://public.tableau.com>

Los demás elementos de la Hoja de Trabajo se encuentran explicados en la Guía de la sesión, que debe ser revisada por los estudiantes y consultada en la sesión asincrónica. Sin embargo, en las próximas sesiones se trabajará más en gráficos y se irá comprendiendo mejor el uso de cada elemento

5. Filtros

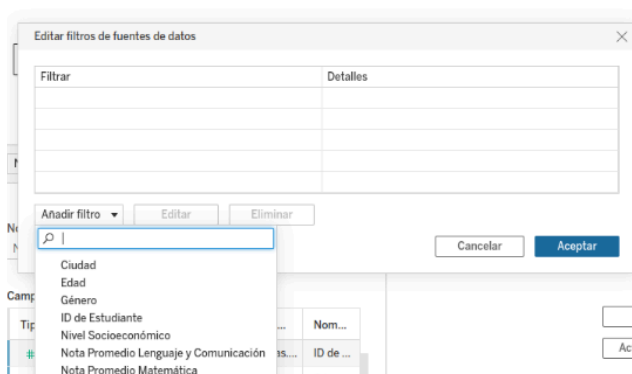
Finalmente, veremos cómo se pueden utilizar filtros en Tableau para lo que utilizaremos el mismo archivo abierto, e iremos aplicando diferentes filtros. Para ello, muestre el proceso de insertar un filtro básico en la Hoja:

- Haga clic sobre "Añadir" junto a "Filtros" en la parte superior derecha de la Hoja de trabajo.



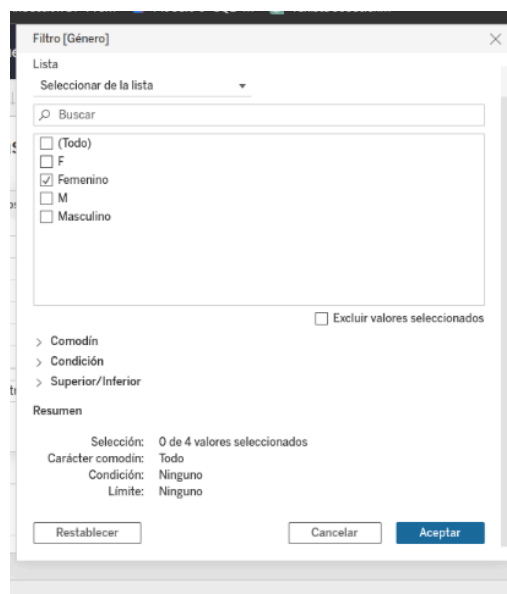
Fuente: <http://public.tableau.com>

- Se abrirá una ventana, y en ella haga clic en “Añadir filtro”. Se desplegará una lista con los campos, y escoja el campo que quiera filtrar. En nuestro caso, escogeremos “Genero_V1”



Fuente: <http://public.tableau.com>

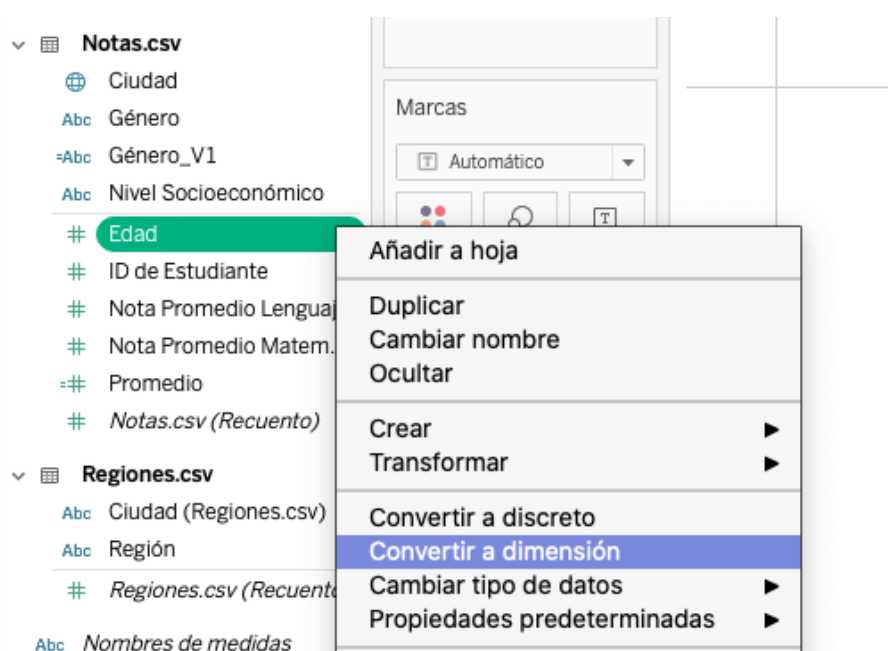
- Muestre que se puede seleccionar los valores que queremos que aparezcan en nuestra tabla, al presionar “Actualizar ahora”. Dado que este filtro se utiliza para limitar el conjunto de datos que se extraen de una fuente de datos, se le conoce como **filtro de extracción**.



Fuente: <http://public.tableau.com>

Mencione a los estudiantes que los filtros pueden aplicarse tanto para dimensiones como para medidas, según el tipo de información que queramos analizar. En la guía podrán encontrar una información más detallada de las características de cada uno. Nos centraremos ahora en mostrar el uso de un filtro de contexto, utilizando el ejercicio descrito en la Guía. Para ello:

1. Cree una nueva Hoja en el archivo Notas que estábamos utilizando. Muestre que tanto la “Edad” como el “Id de Estudiante” aparecen como Medidas ya que son variables numéricas. Hacemos clic derecho en cada una y las transformamos en dimensiones.



Fuente: <http://public.tableau.com>

2. Generamos una tabla, en las filas pondremos el Id de Estudiante y su Edad. En el texto pondremos el promedio de matemáticas.
3. Arrastramos el campo “ID de Estudiante” al estante de Filtros y en las opciones seleccionamos que muestre solo las 10 mejores notas promedio de Matemáticas, tal como se observa en la imagen.

Filtro [ID de Estudiante]

☒ Por campo

Clasificación: Superior ▼

Recuento: Introducir un valor... ▼ 10

Campo: Nota Promedio Matemática ▼

Agregación: Suma ▼

☐ Por fórmula

Clasificación: Superior ▼

Recuento: Introducir un valor... ▼ 10

Fórmula:

Resumen

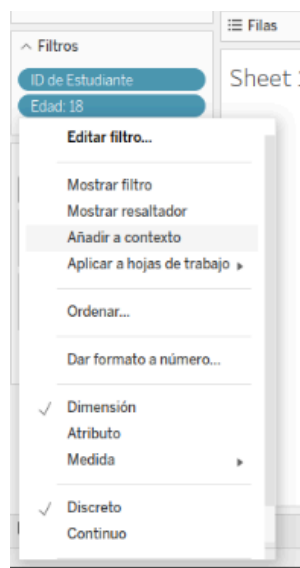
Selección: 275 de 275 valores seleccionados

Carácter comodín: Todo

Fuente: <http://public.tableau.com>

Muestre a los estudiantes que el resultado de esto será obtener el top 10 de estudiantes con mejor promedio en matemáticas, pero necesitamos solamente a los mejores 10 que tengan 18 años. Para esto, aplicaremos un **filtro de contexto**.

4. Arrastramos la Edad al estante de Filtro, y seleccionamos sólo la edad 18 años. ¿Qué sucede? En este caso, Tableau entiende que mostremos a los 10 mejores, y luego filtramos por 18 años seleccionando solo esa casilla. Haga clic en aplicar para mostrar que la lista se vacía: no hay estudiantes de 18 años entre los 10 mejores puntajes.
5. Para que nos muestre a los 10 mejores de 18 años, lo que debemos hacer es forzar que calcule todo con respecto a nuestro filtro, por lo que hacemos clic derecho en el filtro de edad, y lo transformamos en filtro de contexto.



Fuente: <http://public.tableau.com>

Con esto se le indica a Tableau que nos debe mostrar los 10 mejores promedios, pero sólo del universo del contexto. Puede pedir a los estudiantes que mencionen contextos en los que esto podría ser de utilidad.

6. Finalización

Para finalizar, conviene señalar a los estudiantes que en las siguientes sesiones se profundizará en las diferentes opciones que ofrece Tableau (incluyendo colores y tipos de gráficos, etc), pero lo que han visto hasta el momento les sirve para realizar el desafío de la semana.

A partir de esto, es posible regresar a la presentación para revisar el desafío de esta semana, sintetizar las ideas fuerza y terminar la clase.