**Introducción a Tableau**

## **Sobre una de las mejores herramientas de visualización de datos en la actualidad**

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

### **Corta introducción a Tableau,**

por medio de Tableau Public y ejercicio 'abrebocas' *(SneakPeak)* para tener una mínima idea del gran alcance de Tableau como herramienta de visualización de datos por medio de la técnica del 'drag & drop'. Puntualmente **Tableau** sirve mucho para compartir, de la mejor manera visual posible, 'insights' o las conclusiones extraídas de la interpretación y comparación de nuestros datos según un 'dataset' en análisis. Nuestro dataset, el que analizaremos, contiene una lista de transacciones para una determinada tienda. Ahora, el ejercicio consiste en saber cuáles son los Estados más rentables para la tienda en cuestión y mostrar dichos Estados de manera gráfica en un mapa interactivo (todo esto de la manera más eficiente posible).

Adicionalmente, **aprenderá cómo importar libros de Spreadsheets en Tableau (que vendrían siendo el equivalente de un *dataset*) y se dará cuenta que cada pestaña u hoja de cálculo, de un libro de cálculo, sería el equivalente a una tabla en Tableau (es el símil de una entidad para una base de datos relacional). Tanto las bases de datos o libros (datasets) como sus pestañas u hojas (tablas) se importan y se llaman, respectivamente, por medio de una técnica de 'drag & drop'.** Como si fuera poco, como si se tratará de la interfaz gráfica de un manejador de bases de datos relacionales, también puede configurar con facilidad la cantidad de filas (registros) que desea ver de la tabla que seleccione de su dataset.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

#### **Menús de Tableau: Análisis & Formato**

Definitivamente no son los únicos menús de navegación en Tableau, mas, son muy relevantes. Por lo pronto tenga en cuenta que en ***Análisis***se puede especificar cómo queremos nuestros análisis en nuestra hoja de trabajo, también ayuda a diseñar y ajustar la hoja de trabajo según nuestras necesidades; por otro lado, en ***Formato*** puede manipular los datos al definir fuente, alineación, sombreado, borde, línea y demás, con esta herramienta ultimamos detalles como los títulos y las descripciones.

## **Introducción a Dashboard**

### **Manejo & Tipos de campos y sus Jerarquías**

Ahora, luego de arrastrar la tabla de su interés (para nuestro ejemplo será *Orders*), en la esquina inferior izquierda podrá ver una pestaña que se llama: 'Sheet1'.



**'Sheet1', que es una *Worksheet*, nos daría acceso a nuestro panel general (dashboard) de Tableau desde una hoja de trabajo. De hecho, es con *dashboard* que se habilita el acceso a todas nuestras *Worksheet.***

**Entonces he aquí, en 'Sheet1', donde puede manipular directamente los campos, atributos o columnas de la tabla seleccionada en cuestión.** Como se imagina, también se selecciona un campo determinado por medio de la técnica 'drag & drop'.

**Ahora, los campos que usted puede ver en la zona superior son los campos 'categóricos', los que describen nuestros datos (estos campos también se conocen como ‘dimensiones’); y, por otro lado, en la zona inferior, podrá ver los campos 'numéricos' (o también llamados ‘medidas’).** La diferencia entre ambos tipos de campos es que **uno es cualitativo y otro es cuantitativo**, respectivamente. Ahora, Tableau por defecto, bajo su propio criterio, clasifica de forma automática a qué tipo pertenece cada campo; sin embargo, usted puede cambiar esto al poder arrastrar cada campo entre un tipo a otro. **Así la clasificación**:

**Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente con confianza media**

#### **Tipos de Campos a Fondo**

**Los campos cualitativos (dimensiones) son de tipo *variable independiente* y los campos cuantitativos (medidas) son de tipo *variable dependiente*;** es decir, de alguna manera, **las dimensiones definen las medidas (en otras palabras, los campos cualitativos definen a los campos cuantitativos… Esto es así porque las *descripciones, referencias o definiciones, en últimas las cualidades,* le dan contexto a nuestros datos numéricos o cuantitativos: los definen).** Al final, como ya se dijo, queda a nuestra potestad determinar qué campo es de tipo *dependiente* o *independiente*, depende también de nuestras propias necesidades: qué campo queremos que defina y qué campo queremos que sea definido.

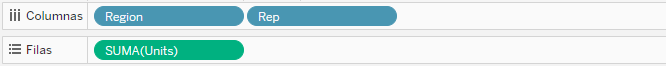
#### **Jerarquía de Campos**

Intente siempre arrastrar primero sus campos cualitativos, y posteriormente, los campos cuantitativos.

Ahora, si hay más de un campo cualitativo (o cuantitativo) en su visualización (gráfico) y hay grados de importancia entre ellos mismos; es decir, si en nuestra visualización tenemos dos campos de tipo *cualitativo,* *campo cualitativo A* vs *campo cualitativo B,* donde el campo cualitativo A puede contener al campo cualitativo B (o lo propio con los cuantitativos); lo ideal es que a la hora de llamar esos campos, dentro de su gráfico, los llame en un orden tal en el que primero se *arrastren* los campos de mayor a menor grado de importancia, en este caso, como el *campo cualitativo A* contiene al *campo cualitativo B,* lo ideal sería arrastrar primero a *A* sobre *B*.

Por ejemplo, si un análisis requerido nos toca llamar el campo de ‘representantes de ventas’ y nos piden evaluar el rendimiento de cada uno de ellos por ‘regiones’, lo que tendría que hacer es llamar primero al campo de ‘regiones’ y, posteriormente, el campo de ‘representantes de venta’; pues, las regiones tendrían un mayor grado de importancia porque contendrían a los representantes.

Ahora, esto es esencial si usted está arrastrando sus campos, para crear una visualización, clasificándolos **por filas y columnas**; en ese caso, si por ejemplo los campos cualitativos están organizados por *columnas*, debe llamar y posicionar primero a ‘regiones’ sobre ‘representantes de ventas’ en la barra de *columnas.* Tal que así:



*No está demás decir que ambas variables, ‘representantes de ventas’ & ‘regiones’, se encuentran representadas en campos con los que puede interactuar dentro de un gráfico o visualización.*

Usted verá un caso de análisis más adelante donde se pondrá en práctica la jerarquía de campos, consistirá en responder qué representante obtiene un bonus anual por región… casualmente. Ese será nuestro siguiente caso práctico: **Visualización de quiénes fueron los representantes con mayores ingresos, por cada región, para obtener un bonus anual.**

### **Caso Práctico 1: Visualización de rentabilidades por cada Estado de los Estados Unidos.**

Continuemos con lo que sería nuestra primera visualización: Saber las rentabilidades por cada Estado de los Estados Unidos. Entonces, como nos interesa ver los datos relacionados a los Estados de los Estados unidos, **arrastramos los campos correspondientes, en principio son: 'Country' & 'State or Province'.** Ahora, el ejercicio es saber cuáles son los Estados de EE. UU más rentables; **dicho eso, nos interesa arrastrar otro campo, uno de tipo numérico que es: 'Profit'***.*

**Ahora, los datos que nos brinda 'Profit' los podemos personalizar de múltiples formas; de tal manera que se vuelva más visual e intuitiva toda la información que podamos extraer de dicho campo, campo 'Profit', al poder arrastrarlo a cuantas opciones disponibles tengamos dentro de la caja de herramientas 'Marcas'.**

*En 'Marcas' podemos determinar por ejemplo para este caso, si queremos que los 'Profit' se diferencien por 'Color'; es decir, que la intensidad del color aumente o disminuye en la medida que el valor del 'Profit' incremente o decremente por Estado. Entonces, tendríamos que arrastrar 'Profit' a la opción 'Color' que está dentro de 'Marcas'. Otra opción de 'Marcas' interesante es 'Etiqueta', en este caso puntual lo que hará 'Etiqueta' es sobreponer sobre cada Estado el valor que le corresponde por 'Profit'; es decir, pasar el valor numérico de cada rentabilidad a cada Estado, la rentabilidad correspondida.*

**Resultado final de nuestra primera *Worksheet*:**

Mapa

Descripción generada automáticamente

Así hay un sin número de opciones más para la caja de herramientas **'Marcas', tales como: 'Tamaño', 'Detalle' & 'Descripción Emergente'. Ahora, sobre cada opción, usted puede configurar aún más los parámetros de dichas opciones.**

Todo lo anterior y mucho más lo irá aprendiendo a fondo en la medida que avance sobre todos los recursos que podrá garantizarle Tableau como uno de los mejores *softwares* de visualización de datos hoy en día.

**Nota**: En Tableau Desktop existe también la pestaña **Historia,** ésta es una combinación de *dashboards & worksheets.*

## **Conectando Tableau a Fuentes de datos y vistazo general de Joins**

### **Introducción a Join en Tableau**

Si usted importa cualquier tipo de base de datos en Tableau, al menos estructurada; es decir, con sus datos organizados entre filas y columnas y de manera *limpia*; y, adicionalmente a eso, dicha base de datos importada cuenta con más de una entidad o tabla… se dará cuenta que, al más parecido estilo de SQL, usted podrá hacer *Joins* en Tableau también.

Primero que todo, debe enfocarse en la DB importada de su interés, previamente estructurada (es preciso decir ‘de su interés’ ya que se pueden dejar importadas varias al mismo tiempo) y hacer clic en ella. Desde ella, podrá ver todas sus entidades o tablas (organizadas una debajo de otra). Mire este ejemplo:

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

La fuente de datos del ejemplo es una Base de datos llamada: *‘P1-SuperStoreUS-2015’* que está estructurada, pues, tiene sus datos organizados entre filas y columnas. Dicho eso, podemos revisar si esta fuente de datos cuenta con más de una entidad o tabla (En bases de datos situadas en *Spreadsheets* una hoja de cálculo, pestaña, sería considerada una tabla o entidad dentro de todo el libro de cálculo); en efecto, la base de datos *‘P1-SuperStoreUS-2015’* sí cuenta con más de una entidad o tabla. Las entidades o tablas de ‘P1-SuperStoreUS-2015’ son: *‘Orders’, ‘Returns’ & ‘Users’.*

**¿Qué podría acontecer entonces?**

Con la misma técnica de 'drag & drop' podrá unir dos o más de dos tablas relacionadas y que pertenecen a una misma base de datos (base de datos que debe estar ya estructurada). Esta unión o *Join* se vería así:

Gráfico

Descripción generada automáticamente con confianza media

Manual y fácilmente, por medio de la ayuda de la interfaz gráfica de Tableau y al conectar varias tablas entre sí, usted podrá ir más allá y especificar qué campos o columnas, de estas tablas conectadas, están directamente relacionadas entre sí; esto sería el equivalente en SQL de una tabla con una columna *Primary key* que ha exportado sus llaves a otra columna de otra tabla que las recibe como *Foreign key* (y por eso se pueden unir); es decir, existe la condición de posibilidad del cruce entre ambas tablas: sus llaves.

**Recuerde**: Con *Join* se da el primer puntapié para darle fundamento práctico a la teoría que esconde el concepto de ‘Bases de datos relacionales’; lo que trata sobre, justamente, las relaciones. Relacionar entidades entre sí para extraer información, no datos, para darle un sentido práctico y contextualizado al mundo real a tanta codificación, convertir el código y los datos en información útil y diciente.

### **Conexión a diferentes fuentes de datos en Tableau**

En Tableau se pueden conectar distintas fuentes de datos e, incluso, además de ser diferentes, no tienen que ser necesariamente de la misma naturaleza o formato; es decir, usted podría trabajar con una base de datos relacional (de SQL) de manera conjunta con una base de datos de formato completamente diferente como, por ejemplo, CSV. En consecuencia, usted al final podrá trabajar con muchos ficheros diferentes y, ni siquiera, tener que conservar necesariamente la misma naturaleza o formato entre ellos (En caso de que decida trabajar de manera conjunta con ellos, lo ideal y necesario es que se relacionen por medio de sus llaves, pero pueden ser de formatos diferentes).

En *Conexiones* propiamente en Tableau, podrá todas las fuentes de datos que ha importado, tal que así:

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

En el anterior ejemplo, como se percata, tenemos dos fuentes de datos importadas. La 1ra es un archivo de formato .*xlsx* (Spreadsheets) y la 2da es un archivo de texto, más específicamente de tipo *.csv* (Comma Separated Values**). Entonces, haciendo clic sobre una de ellas, podrá ver internamente sus tablas y, posteriormente, sus campos o columnas (en caso de que se trate de una fuente de datos estructurada u organizada de forma tabular).** En nuestro ejemplo de conexiones, el único fichero organizado de forma tabular es *‘P1-SuperStoreUS-2015’,* contrario a *‘P1-OfficeSupplies’.* **¿Cómo editar las fuentes de datos (Cambiar tabla para analizar en dashboard)?**

Supongamos que ya está situado sobre una base de datos en particular y, en especial, sobre una tabla de dicha base de datos; es decir, ya está viendo el *dashboard* donde puede analizar columna por columna la información de una tabla (de una tabla perteneciente a una base de datos en particular, lo dicho). La imagen es más o menos así:

Interfaz de usuario gráfica, Aplicación, Word

Descripción generada automáticamente

Aquí estamos en el Dashboard con la capacidad de analizar visualmente cada uno de los campos de la tabla *Users* de la base de datos *‘P1-SuperStoreUS-2015’*. Sin embargo, si quisiéramos retroceder y ver el Dashboard de otra tabla, ya sea de la misma base de datos o de otra, tendría que hacer doble clic en la opción que se ve en toda la esquina superior izquierda, en este caso sería ‘Users (Varias conexiones)’:

Interfaz de usuario gráfica, Aplicación, Word

Descripción generada automáticamente

O bien, desplegar el menú despegable de la misma y hacer clic en: “Editar fuente de datos…”

*Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente*

**Aclaración:** Se habla de *varias conexiones* debido a que hay dos o más de dos fuentes de datos, en nuestro caso recuerde que tenemos dos:

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

**Continuemos**. En nuestro ejemplo, luego de hacer clic en “Editar fuente de datos…” verá lo siguiente:

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Ahora, si no haremos ninguna relación con alguna otra tabla (si no haremos ningún *Join*) y sólo nos interesa ver el dashboard de otra tabla, tendríamos que suprimir la tabla actual (haciendo clic derecho en el recuadro ‘Users’ y luego clic en ‘Eliminar’) y arrastrar la otra (la otra tabla de nuestro interés).

Pongamos así: Deseo dejar de ver la tabla ‘Users’ para ver una tabla de la base de datos *‘P1-OfficeSupplies’;* es decir, sería otra base de datos. En ese caso, luego de suprimir ‘Users’, debe hacer clic en *‘P1-OfficeSupplies’* situado en Conexiones; estando ahí se enlistarán todas las tablas relacionadas a *‘P1-OfficeSupplies’.* Se percatará que realmente sólo hay una tabla, está enlistada en la sección de *Archivos,* y ya estará arrastrada en la imagen, tal que así:

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Ya estando seleccionada la única tabla que nos interesa analizar visualmente en un Dashboard, hacemos clic en **Sheets1** (lo que usted ya conoce).

### **Previsualización de una tabla o fichero en Tableau** Ahora, antes de irte al Dashboard, puedes visualizar cada una de las tablas (o también ficheros) de cada base de datos. Tableau te da la opción de tener una previsualización. Por ejemplo, supongamos que quiero ver una previsualización de la única tabla de la base de datos *‘P1-OfficeSupplies’*. Sería así:

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

Se hace clic en el cuadro señalado.

### **Extra: ¿Cómo visualizar el tipo de datos para las columnas de una tabla?**

Veamos por ejemplo una parte de las columnas de una tabla, *Orders*, de la base de datos *‘P1-SuperStoreUS-2015’.*

Interfaz de usuario gráfica

Descripción generada automáticamente

Si se fija en la parte superior izquierda de cada columna podrá ver unos *iconos,* estos iconos representan justamente el tipo de datos de cada columna. Por ejemplo,  o  representan el tipo de dato **‘string’**;  representa el tipo de dato **‘number’** &  representa el tipo de dato **‘date’.**

### **Extra: ¿Cómo seleccionar un tipo de gráfico deseado para nuestra Worksheet?**

Realmente no es difícil, luego de haber arrastrado sus campos, los de su interés para analizar visualmente, usted puede percatarse que en la esquina superior derecha sale una opción que dice: ‘**Muéstrame’,** tal que así:



Si hace clic sobre este botón se desplegará un sin número de gráficos disponibles, a su antojo, para seleccionar con qué tipo de visualización analizar su Worksheet. Generalmente, según la lectura de los datos que tenga Tableau de la Worksheet en cuestión, el mismo software le habilitará unas opciones de visualización disponibles para trabajar y otras no.

## **Creando campos calculados**

### **Intro: ¿Cómo limpiar el *dashboard*?**

Convengamos que, eventualmente, en nuestra Worksheet hay contenido visual y queremos limpiar toda la información ahí contenida. Tenemos actualmente el siguiente gráfico:

Interfaz de usuario gráfica

Descripción generada automáticamente

Nuestro gráfico se trata de una lista y queremos borrarla, ¿cómo lo haríamos? Se trata de hacer un clic en una simple opción y es la siguiente:

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente, hace clic en borrar hoja y listo, automáticamente su gráfico será suprimido de su hoja actual de trabajo para comenzar desde 0 desde esa misma *Worksheet*.

### **Caso Práctico 2: Visualización de quiénes fueron los representantes con mayores ingresos, por cada región, para obtener un bonus anual.**

Para crear un **campo calculado**, que será una herramienta de mucha utilidad para resolver este ejercicio, básicamente tendría que añadirle un campo adicional, de tipo numérico, que haga un cálculo entre campos numéricos ya existentes dentro del fichero o tabla que pretende analizar visualmente; como bien te imaginas, el campo calculado que se pretende crear no existe inicialmente dentro del fichero o tabla en cuestión; es decir, a usted le corresponde crear este *campo calculado* desde Tableau. Pongamos el concepto en práctica mediante nuestro caso actual. Veamos:

Interfaz de usuario gráfica

Descripción generada automáticamente

Si se percata en nuestra tabla no contamos con un campo que haya calculado los ingresos totales (*revenue*) por cada representante de venta, es decir, un campo que resulte de la multiplicación del total de las unidades vendidas (*Units*) por el precio de dichas unidades (*Unit* *Price*); entonces, como no tenemos ese campo, toca calcularlo. Al *campo calculado* lo llamaremos *Revenue*.

Para crear un campo calculado, desde Public Tableau, tiene que hacer clic sobre la pestaña *Análisis* que se encuentra situada en la barra de menús superior; desde ahí podrá ver claramente la opción: **‘Crear campo calculado…’.** Posteriormente verá una ventana emergente donde podrá crear su campo calculado, tal que así:

Interfaz de usuario gráfica, Texto, Aplicación, Word

Descripción generada automáticamente

En *Cálculo1* pásele el nombre a su campo calculado, sería para este caso: *Revenue.* Posteriormente, en la parte inferior de la misma ventana emergente, escriba los campos ya existentes dentro de su fichero con los que pretende calcular su fórmula de ingresos y crear así su nuevo campo, serían: *Units \* Unit* *Price.* Queda así:

Interfaz de usuario gráfica, Texto, Aplicación, Word

Descripción generada automáticamente, *clic en Aceptar.*

Listo, ya estaría creado:

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente

Observe que su único campo cuantitativo con un numeral (#) acompañado del signo igual (=) es *Revenue,* esa simbología nos revela que el campo *Revenue* resulta ser un *Campo calculado.*

Arrastremos ahora al campo *Revenue* a ‘Filas’ para representarlo visualmente en un gráfico de barras (de antemano ya hemos tomado en cuenta la Jerarquía de campos para los campos cualitativos).

Gráfico, Gráfico de barras

Descripción generada automáticamente

***Algo adicional para aclarar***, las barras ya están ordenadas de forma descendente; es decir, representan de mayor a menor los ingresos generados por los vendedores en cada región. Para ordenar datos ya visualizados, especialmente cuando se trata de gráficos de barras, de forma ascendente o descendente (o inclusive intercambiar la información entre filas y columnas), podemos hacerlo desde la siguiente caja de opciones:

