

1. PV.04 - FAMILIARIZARSE CON TABLEAU

1.1. Objetivo:

Aprender los fundamentos de Tableau mediante la visualización de un conjunto de datos de muestra.

1.2. Resultados del aprendizaje:

- Crear visualizaciones textuales y gráficas en Tableau.
- Combinar datos de múltiples tablas para crear visualizaciones.
- Crear campos calculados para categorizar los datos.
- Crear un panel de control para ver varias visualizaciones a la vez.

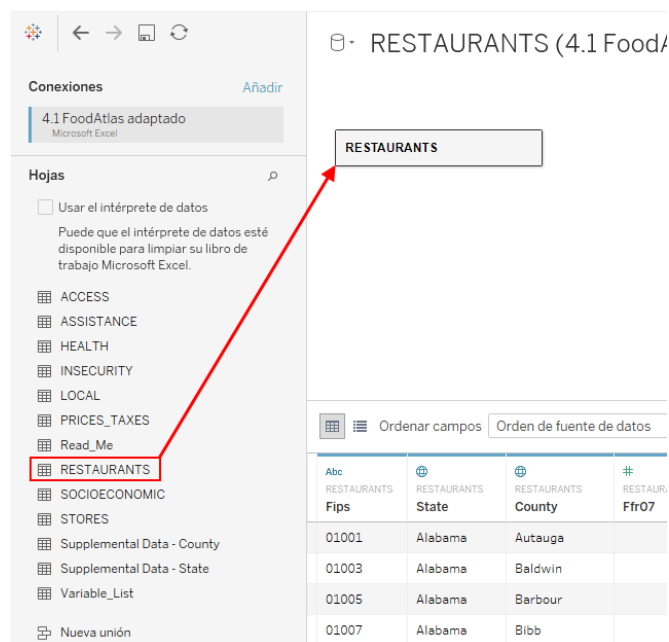
Este tutorial le llevará a través de los fundamentos de Tableau. Trabajaréis con el Atlas de Investigación del Acceso a los Alimentos del Departamento de Agricultura de los Estados Unidos. Se trata de un amplio conjunto de datos, por condado de EE.UU., sobre el acceso a tiendas de comestibles, tiendas 24h, restaurantes y datos socioeconómicos y de salud.

El acceso a este conjunto de datos se realizó a través de data.gov y se alojó en el sitio web del Departamento de Agricultura de EE.UU.

1.3. Parte 1: Iniciar Power BI e importar los primeros datos

1. Abrir el archivo en Excel y echarle un vistazo. Verá que hay una serie de hojas de trabajo. Cada una de ellas es una tabla independiente. El diccionario de datos está en la pestaña "Variable_list".
2. Inicie Tableau.
3. Haga clic en "Microsoft Excel" para conectarse a nuestro archivo de datos.
4. Navegue hasta la ubicación donde está almacenado su archivo de datos (FoodAtlas adaptado.xlsx) y selecciónelo.

- Verá una lista de hojas de cálculo de Excel en la parte izquierda de la pantalla. Estas son todas las hojas contenidas en el libro de trabajo. Arrastre la hoja **RESTAURANTS** al espacio de trabajo:



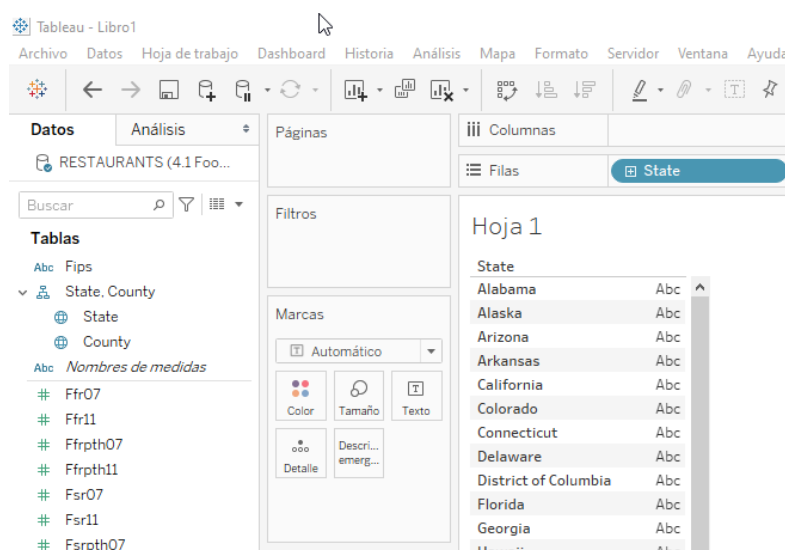
- A continuación, haga clic en ir a la hoja de trabajo en la esquina inferior izquierda.



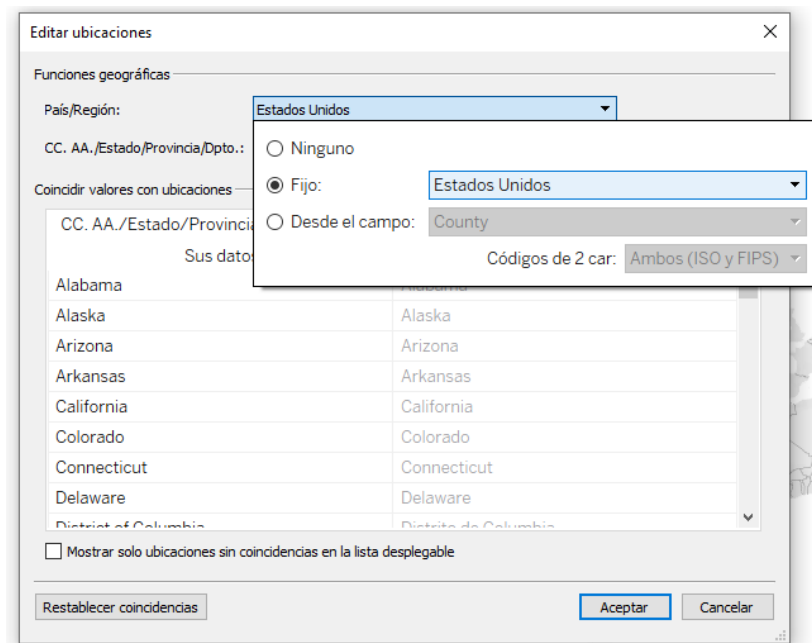
1.4. Parte 2: Trabajar con datos: Filas, Columnas y Marcas

Aquí vamos a crear una tabla con los condados y su información asociada a los restaurantes de comida rápida.

- Arrastre la dimensión **"State"** al estante de Filas. Verá esto:

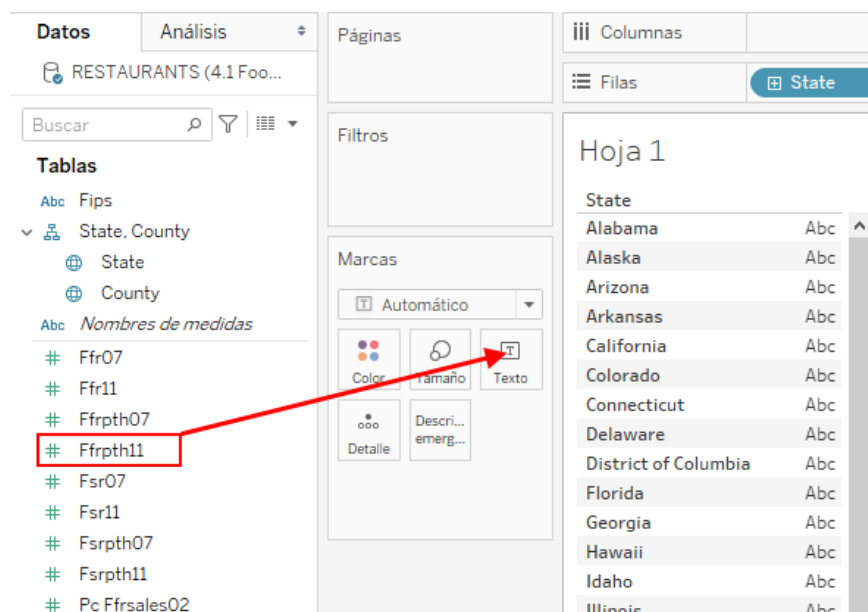


2. Como vamos a trabajar con datos de Estados Unidos y estamos en Europa y con una versión en castellano del Tableau tenemos que hacer algunos cambios. Tenemos que ir al menú **Mapa > Editar ubicaciones...** y se os abrirá la siguiente ventana:

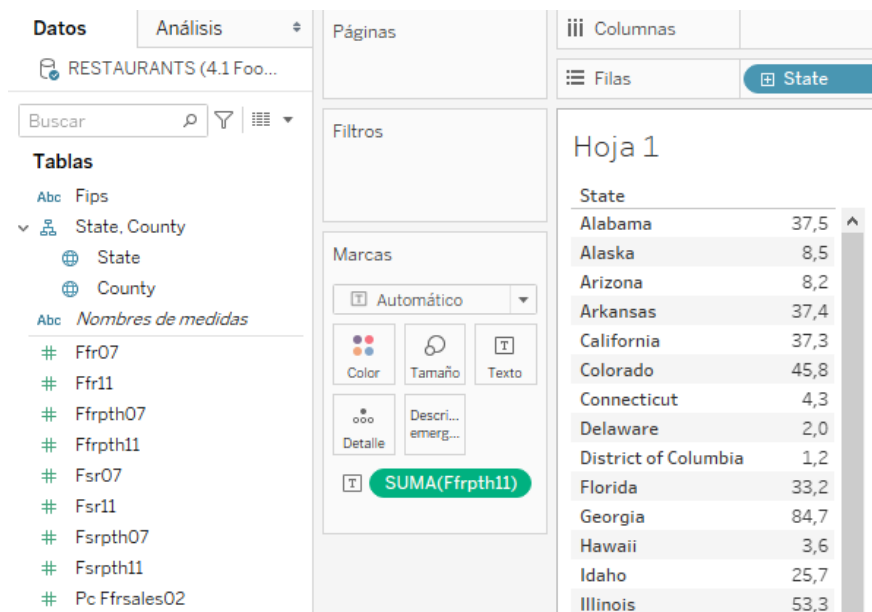


En esta ventana cambiaremos **España** por **Estados Unidos** y pulsamos Aceptar. Observar que las dimensiones **State** y **County** ya tienen un formato jerárquico, por ejemplo, **State > County**.

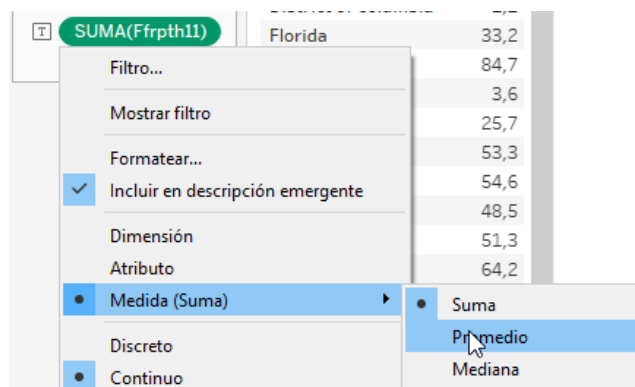
3. Ahora arrastraremos la medida "Ffrpth11" (el número de restaurantes de comida rápida por cada 1000 residentes a partir de 2011) al área de Texto dentro de Marcas (como se mostrarán los valores). Ahora verá esto:



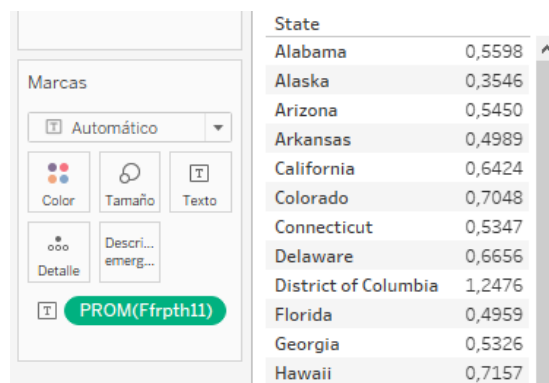
Después de arrastrar:



- Observe que Tableau adivina que quiere sumar todos los valores de cada estado (**SUMA**). Hay muchas entradas en esta tabla para cada estado - una para cada condado. Pero como nuestra medida es una ratio, queremos promediar estos números, no sumarlos. Así que haga clic en **SUMA(Ffrpth11)** y seleccione **Medida>Promedio**. Ahora verá el promedio de todos los valores para ese estado:



Después de cambiar a promedio:



5. Luego arrastre la dimensión **"County"** a Filas, junto a **"State"**. O bien, puede hacer clic en el signo + junto a **"State"** si Tableau ya ha reconocido la jerarquía entre **"State"** y **"County"**.

Ahora verá un desglose por condado. Si ve un diálogo de advertencia, haga clic en "Añadir todos los miembros".

Filtros

Marcas

T Automático

Color

Tamaño

T

Texto

ooo

Detalle

Descri...
emerg...

T PROM(Ffrpth11)

Filas

State

County

Hoja 1

State	County	
Alabama	Autauga	0,615
	Baldwin	0,648
	Barbour	0,701
	Bibb	0,264
	Blount	0,347
	Bullock	0,379
	Butler	0,823
	Calhoun	0,773
	Chambers	0,678
	Cherokee	0,652
	Chilton	0,478

Desplázate hasta "Pennsylvania" y busca el condado de "Philadelphia". Parece que tienen una alta proporción de restaurantes de comida rápida.

También nos encontramos que justo debajo de Pennsylvania aparece un estado **PR** (Puerto Rico) que no tiene datos por lo que vamos a excluirlo. Para eso, hacemos clic derecho sobre **PR** y seleccionamos Excluir y nos aparecerá la etiqueta **State** en Filtros y vemos que desaparece de la tabla.

Filtros

State

Marcas

Automático

Color

Tamaño

Texto

Detalle

Descri...
emerg...

PROM(Ffrpth11)

Hoja 1

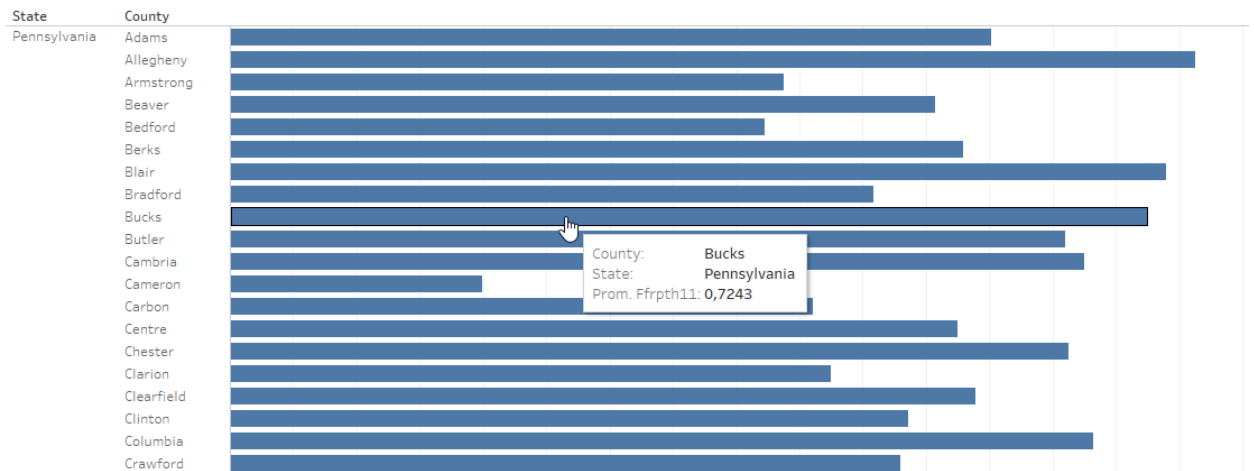
State	County	
Pennsylvania	Perry	0,456
	Philadelphia	0,851
	Pike	0,246
	Potter	0,229
	Schuylkill	0,481
	Snyder	0,578
	Somerset	0,439
	Sullivan	0,772
	Susquehanna	0,324
	Tioga	0,377
	Union	0,424
	Venango	0,384
	Warren	0,386
	Washington	0,586
	Wayne	0,377
	Westmoreland	0,637
	Wyoming	0,387
Rhode Island	York	0,653
	Bristol	0,582
	Kent	0,737
	Newport	0,822

1.5. Parte 3: Creación de gráficos y mapas

Veamos cómo se compara Philadelphia con el resto del estado. Primero, filtramos nuestros datos para ver solo Pennsylvania.

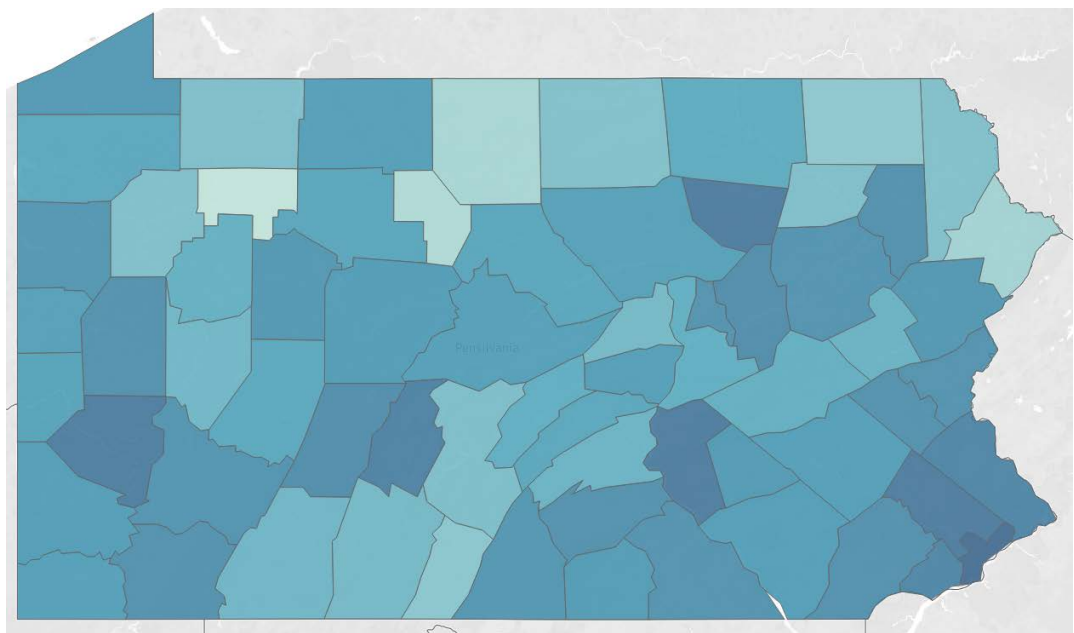
- Haga clic en **"State"** en Filas, y luego en **Editar Filtro...**
- En el cuadro de diálogo Filtro [State], vamos a la pestaña General y hacemos clic en "Seleccionar de la lista" y nos aparecen todos los estados con una casilla para marcar. Hacemos clic en la casilla "Ninguno", luego desplácese hacia abajo y seleccione solo Pennsylvania. Comprobar que al lado derecho no esté marcado Excluir. A continuación, haga clic en "Aceptar".
- No ha cambiado mucho, salvo que ahora sólo vemos los datos de los condados de Pennsylvania. Y todos los datos son difíciles de comparar. Es difícil saber si el condado de Philadelphia realmente tiene el mayor número de restaurantes de comida rápida por cada 1000 residentes en el estado. Así que ahora vamos a probar algunas visualizaciones para que los datos sean más fáciles de ver.
- En la pestaña "Mostrarme" de la esquina superior derecha, haz clic en el icono del gráfico de barras (tercera fila, primera columna).
- Ahora verá un gráfico de barras (también notará que **AVG(Ffrpth11)** se ha movido del área de Marcas a la zona de Columnas.



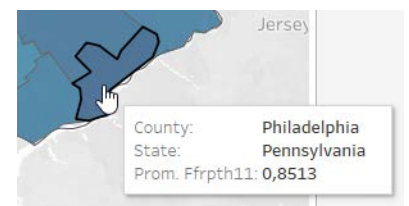


Desplázate hacia abajo en la lista y verás que **Philadelphia** es realmente la campeona cuando se trata de la mayor cantidad de restaurantes de comida rápida por cada 1000 residentes. Pero sigue siendo engorroso navegar por un gráfico de barras tan complejo.

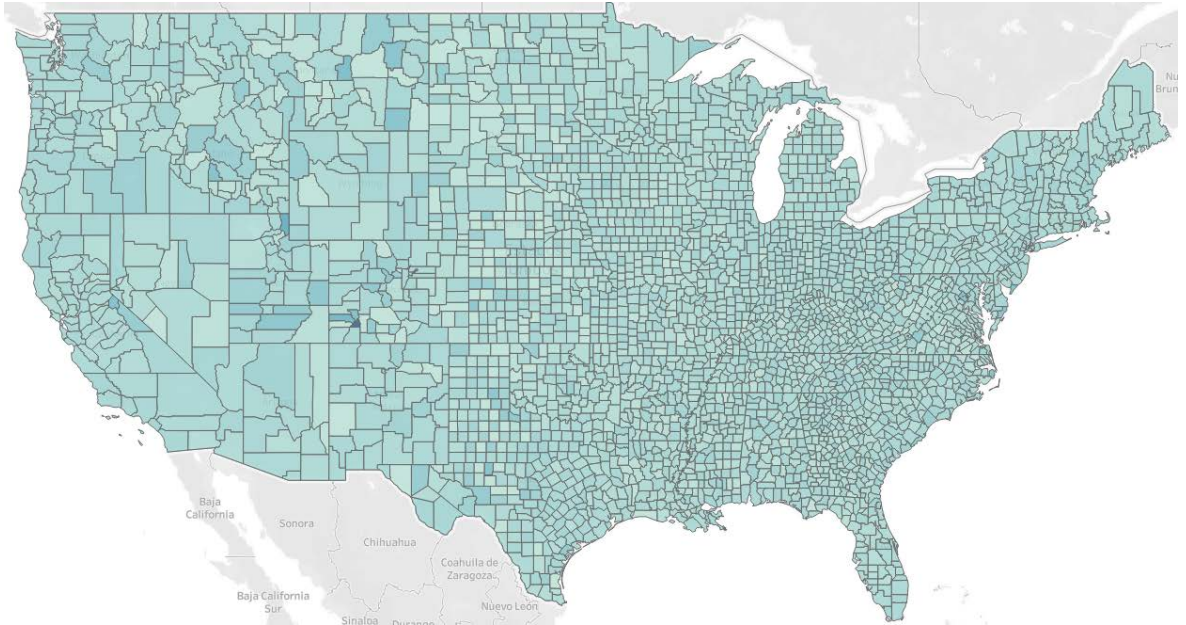
- Así que ahora haz clic en el icono del mapa relleno en la pestaña “Mostrarme” (segunda fila, segunda columna) y verás un mapa sombreado de Pennsylvania por condado, donde las zonas sombreadas significan valores más altos. Observe que también han cambiado los valores en Filas y Columnas a Longitud y Latitud. Además, State, County y AVG(Ffrpth11) están ahora en el área de Marcas.



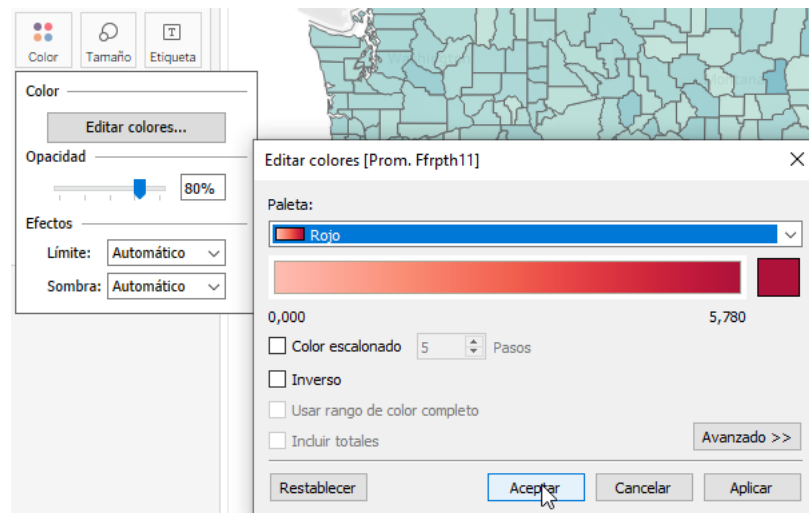
De esto se desprende que **Philadelphia** se encuentra claramente entre los valores más altos de densidad de restaurantes de comida rápida del estado. Para estar seguro, puedes pasar el ratón por encima de los otros condados sombreados de forma similar y ver los valores exactos.



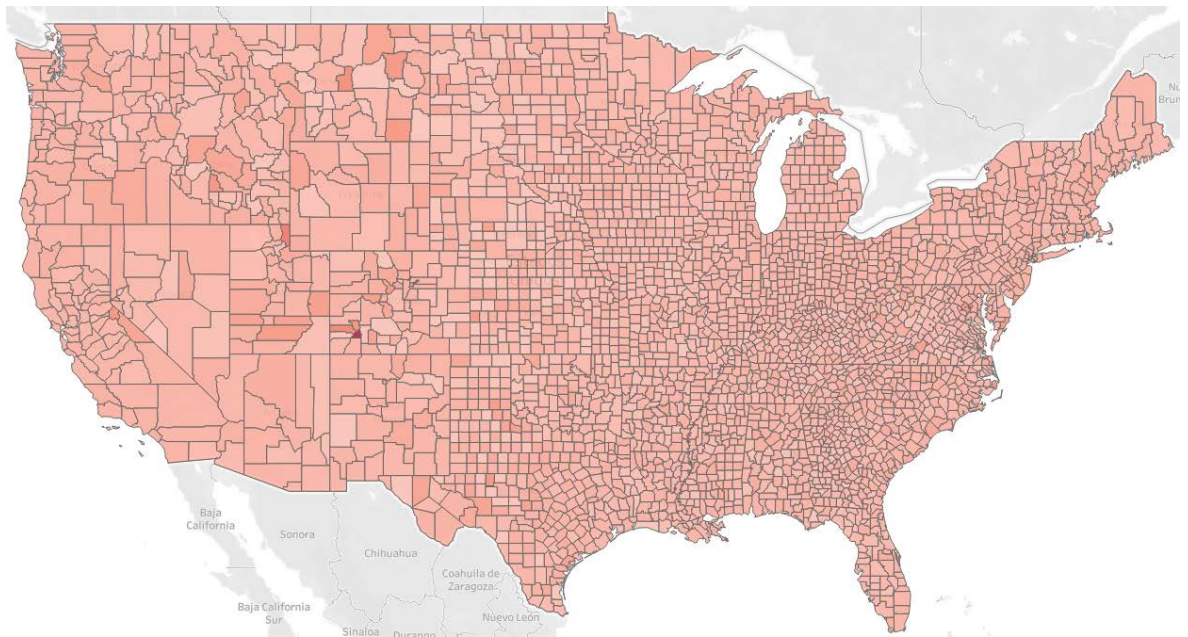
7. Ahora vamos a crear un mapa de todo el territorio continental de Estado Unidos. Haz clic en **"State"** (ahora bajo el área de Marcas) y selecciona **"Editar filtro..."**.
8. Haz clic en **"Todos"** y luego deselecciona Alaska, Hawaii y Puerto Rico (PR) y pulsamos Aceptar. Ahora verá un mapa como este:



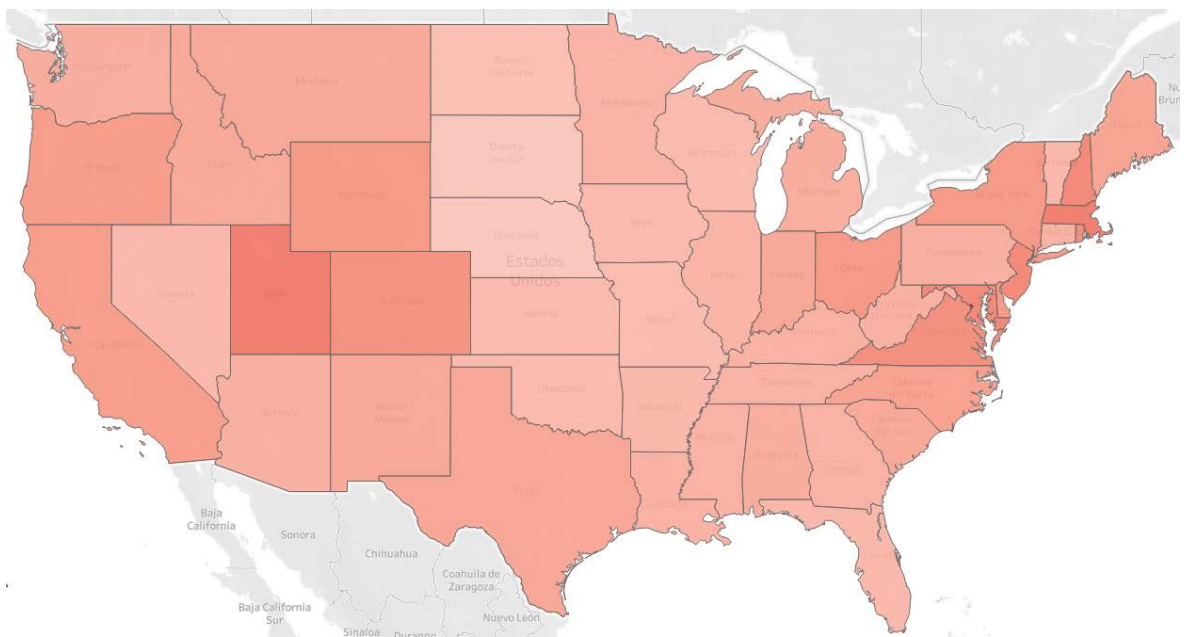
9. Vamos a cambiar el color para que sea más fácil de leer. Haz clic en **"Color"** en Marcas y luego **"Editar colores..."**



10. Escoger en Paleta **"Rojo"** y haz clic en OK. Ahora verás el mapa en tonos de rojo.



11. Ahora vamos a ver el mapa por estados, no por condados. Haz clic en “**County**” dentro del área de Marcas y selecciona “Eliminar”. Ahora verás esto:

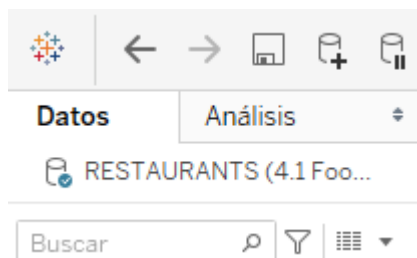


12. Por último, cambie el nombre de la hoja de trabajo que ha creado haciendo doble clic en la pestaña “Hoja 1” en la parte inferior de la ventana de Tableau y escribiendo “Densidad de comida rápida por estado”.
13. Este sería un buen momento para guardar el libro de trabajo de Tableau (Archivo/Guardar). Guarde el archivo en el mismo lugar donde puso la hoja de cálculo de Excel.

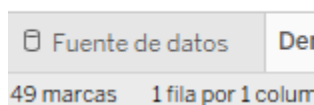
1.6. Parte 4: Análisis de datos en múltiples tablas (hojas de trabajo)

A veces tendrá datos relacionados en más de una tabla (o en el caso de Excel, en más de una hoja de trabajo). Tableau le permitirá combinar esos datos utilizando algo llamado "join".

1. Haga clic con el botón derecho del ratón en la entrada **RESTAURANTS** bajo "Datos" en la parte superior izquierda de la pantalla.



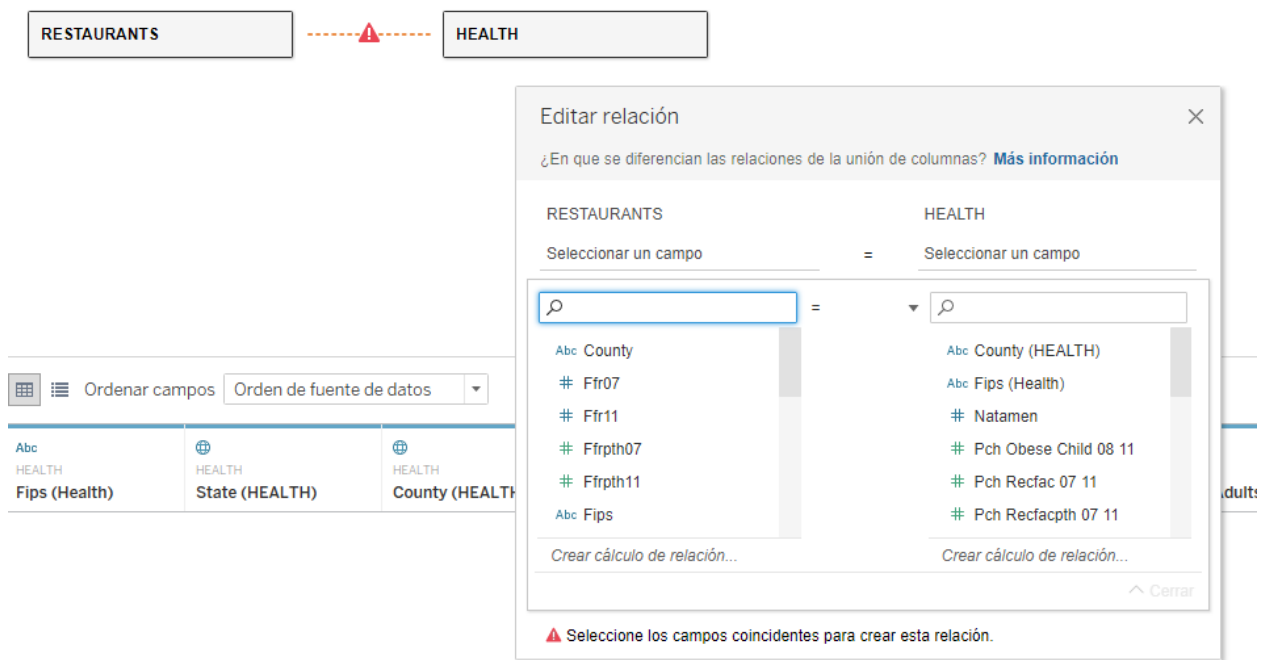
Seleccione **Editar fuente de datos...** O también puede hacer clic en la pestaña "Fuente de datos" en la esquina inferior izquierda.



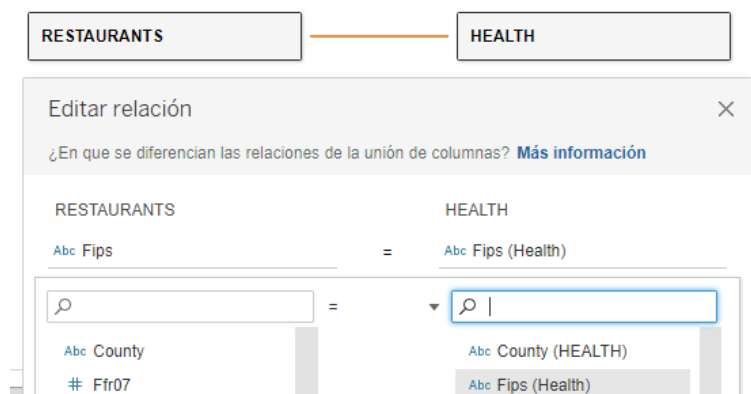
2. Arrastre la hoja de trabajo **HEALTH** al área de espacio en blanco junto a **RESTAURANTS**. La tabla HEALTH tiene datos como las tasas de obesidad y el número de instalaciones recreativas, también registrados por condado.

En Tableau 2020.4, tenemos una nueva forma de fusionar tablas. Se llama vista "Relaciones". Anteriormente utilizábamos la vista "Uniones". La vista de relaciones es una forma inteligente y rápida de fusionar tablas de forma directa. Puede seguir trabajando con la vista "Joins", pero la tabla combinada resultante será una tabla estática (la desventaja de las uniones). Para cambiar a la vista de uniones debe hacer doble clic en el cuadro RESTAURANTS y luego realizar las operaciones de fusión.


Sin embargo, seguiremos la nueva vista "Relaciones". Después de arrastrar la tabla HEALTH a la zona blanca, Tableau intentará establecer la relación entre las filas de estas dos tablas. Debemos especificar la columna que se utilizará para relacionar las filas. En este caso, vamos a emparejaremos los datos de las dos tablas basándonos en el identificador del condado, que es "FIPS". FIPS es un identificador único de condado, por ejemplo, ID de condados. No utilizamos los nombres de los condados para el cotejo porque no es único, es decir, algunos estados tienen condados con el mismo nombre.



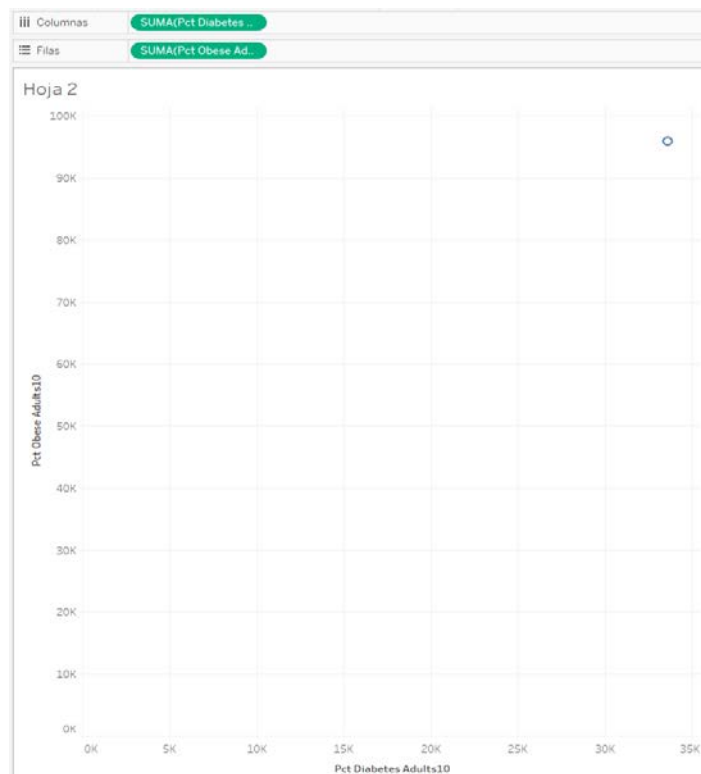
Ahora seleccione **Fips** para la tabla **RESTAURANTS** y luego vuelva a seleccionar **Fips** para la tabla **HEALTH**. El aspecto será el siguiente. La línea entre los rectángulos no debe tener ninguna señal de advertencia.



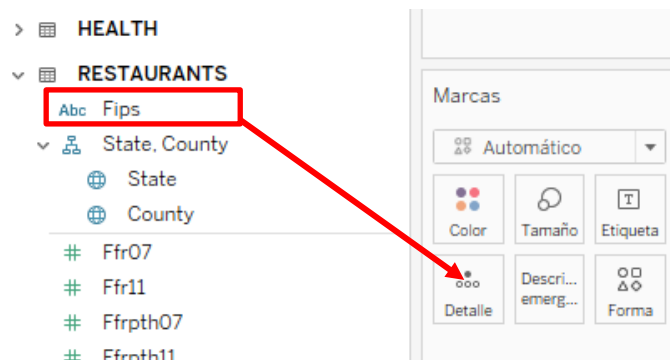
Ya está bien, puede cerrar el cuadro de diálogo Editar relación.

- Ahora haga clic en la hoja de trabajo que acaba de crear (esquina inferior izquierda). Verá nuevas Dimensiones y Medidas agrupadas bajo cada fuente de datos (tabla) como HEALTH y RESTAURANTS.
- Cree una nueva hoja de trabajo de Tableau haciendo clic en  el icono en la parte inferior de la pantalla.
- Busquemos una relación entre las tasas de diabetes y las tasas de obesidad. En "Medidas", arrastre "**Pct Obese Adults10**" (el porcentaje de adultos obesos en 2010, por condado) al estante Filas y arrastre "**Pct Diabetes Adults10**" (el porcentaje de adultos con diabetes en 2010, por condado) al estante Columnas.

6. Tableau agregará automáticamente los datos de todas las observaciones y mostrará sólo un punto de datos en el gráfico.

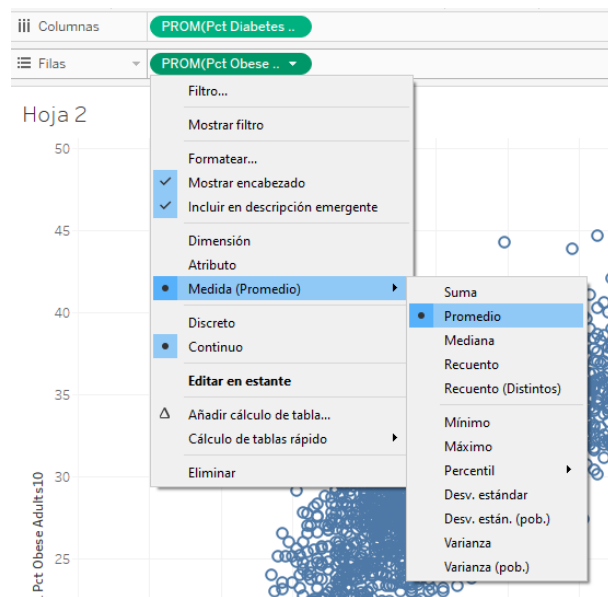


Ahora, necesitamos desagregar nuestros puntos de datos a nivel de condado. Para ello, necesitamos aumentar el detalle de nuestro gráfico. Desde las dimensiones de la tabla RESTAURANTS, arrastre "Fips" al botón de detalle del área de Marcas.

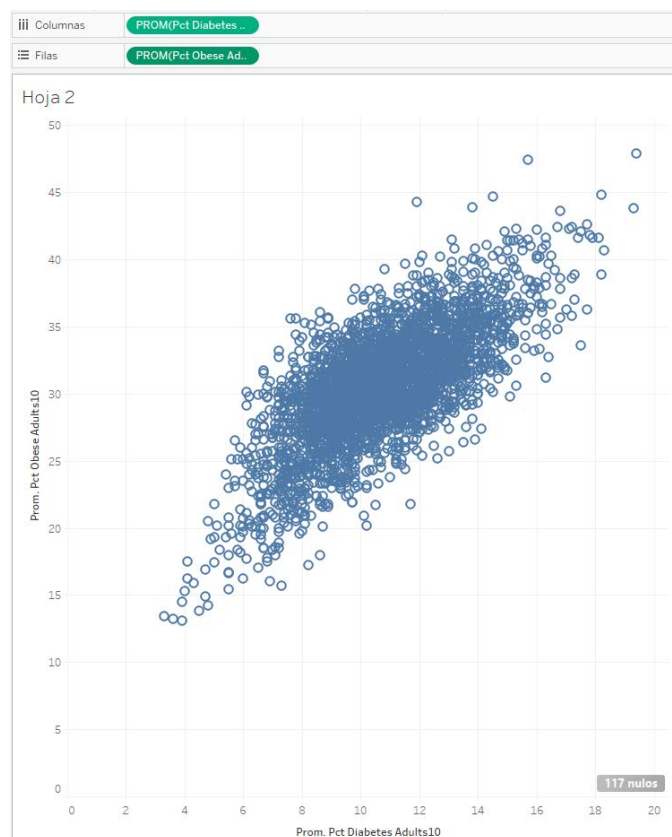


Esto le dice a Tableau que trace cada punto de datos (condado) por separado, y no los sume o promediarlos en un solo número.

Sin embargo, podemos tener datos duplicados o de varios condados, y la suma de esos valores a de nivel de condado no nos dará una buena información. Es mejor que tomemos la media de nuestras medidas a nivel de condado. Para ello, haga clic en la flecha hacia abajo situada junto a la medida **SUMA(Pct Obese Adults)** en el estante de filas y seleccione **Medida(Suma)** y luego seleccione **Promedio**. Esto le dice a Tableau que tome el promedio de todos los valores de un condado. Haga el mismo truco para la medida **SUMA(Pct Diabetes Adults)** en el estante de columnas

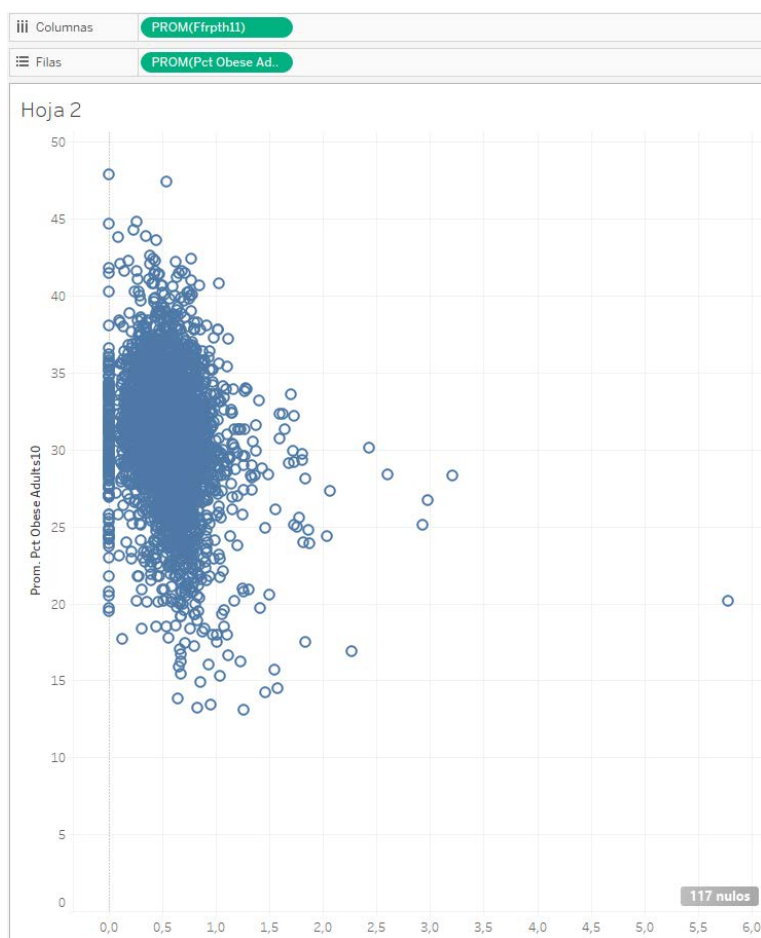


Verás un gráfico de dispersión que muestra claramente una relación positiva entre la obesidad y la diabetes:



7. Ahora vamos a combinar los datos en las tablas. Elimina "**Pct Diabetes Adults10**" de las columnas y sustituirlo por "**Ffrpth11**" - recordemos que es el número de restaurantes de comida rápida por cada 1.000 habitantes de nuestra tabla original de **RESTAURANTS**.

Cambia la agregación por el promedio de nuevo. Ahora se obtiene este gráfico de dispersión:



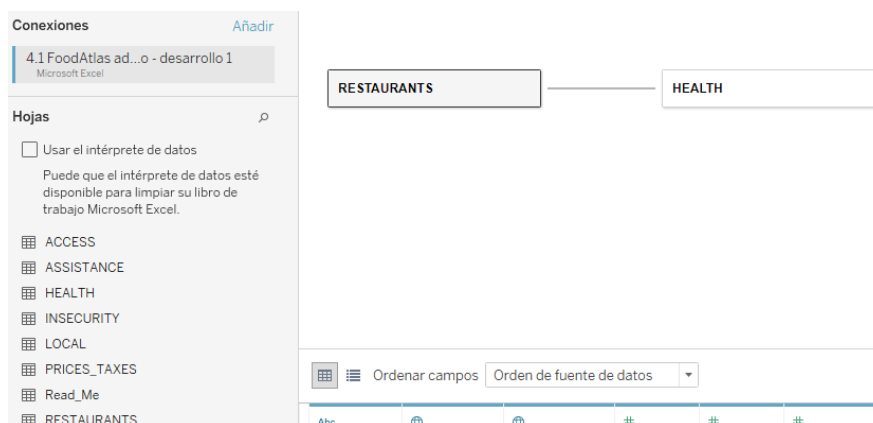
Esto nos da el resultado algo sorprendente de que la densidad de los restaurantes de comida rápida parece no tiene relación con las tasas de obesidad.

8. Cambie el nombre de esta hoja de trabajo por "Densidad de la comida rápida frente a las tasas de obesidad"

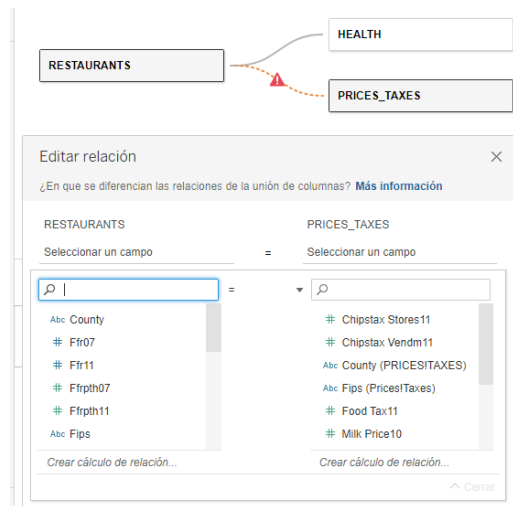
1.7. Parte 5: Crear un campo calculado

Puede crear campos de datos adicionales basados en los que ya existen. Esto puede ser útil si quiere crear categorizaciones adicionales, o si necesita realizar nuevos cálculos.

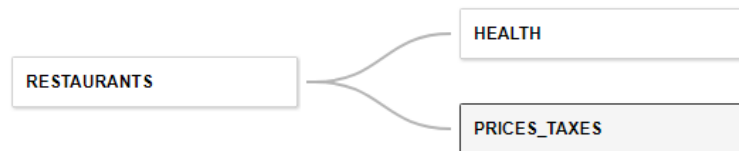
1. Primero, vamos a crear una nueva hoja de trabajo, basada en la que acabamos de crear. Haga clic con el botón derecho del ratón en la pestaña "Densidad de la comida rápida frente a las tasas de obesidad" y seleccione "Duplicar hoja". Se creará "Densidad de Comida Rápida vs Tasas de Obesidad (2)".
2. Ahora elimine **Ffrpth11** del estante de Columnas y reemplácelo por "**Pct Diabetes Adults10**". Cambie la agregación a promedio. Verá el mismo gráfico de dispersión del paso 7 de la última sección.
3. Haga clic con el botón derecho del ratón en la entrada **RESTAURANTS** bajo "Datos" en la parte superior izquierda de la pantalla y elija "Editar fuente de datos..." O también puede hacer clic en la pestaña "Fuente de datos" en la esquina inferior izquierda. esquina inferior izquierda.
4. Arrastre la hoja **PRICES_TAXES** al espacio en blanco debajo de **HEALTH** y **RESTAURANTS**. Esta hoja tiene información sobre los precios relativos de la leche, los refrescos y los tipos impositivos de la comida basura.



Entonces aparecerá la ventana "Editar relación". Seleccione **Fips** para la tabla **RESTAURANTS** y luego seleccione de nuevo **Fips** para la tabla **PRICES_TAXES**. Se verá como la siguiente.



Si el fideo entre los rectángulos no tiene ninguna señal de advertencia, entonces usted es bueno para puede cerrar el cuadro de diálogo.

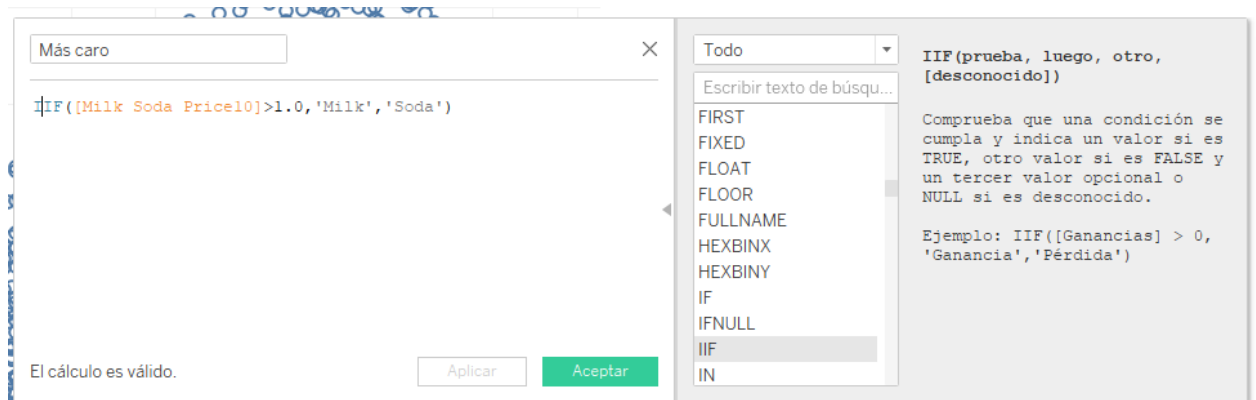


5. Vuelva a la última hoja de cálculo, la que duplicó.
6. Vaya al menú **Análisis** y seleccione "**Crear campo calculado...**".
7. Nos gustaría saber si el precio relativo de la leche con respecto a los refrescos tiene alguna relación con las tasas de obesidad y la diabetes. Tenemos un campo en **PRICES_TAXES** llamado "**Milk Soda Price10**", que es el precio de la leche dividido por el precio de los refrescos, por condado, en 2010.

En el cuadro de diálogo, escriba "**Más caro**" para el nombre, y la siguiente fórmula para la fórmula:

`IIF([Precio de la leche y la soda10]>1.0, 'Milk', 'Soda')`

NOTA: IIF NO es un error tipográfico. Asegúrese de que tiene dos "I" ahí.



8. Si lo has escrito correctamente, verás "El cálculo es válido" bajo el cuadro de texto Fórmula. A continuación, haga clic en "Aceptar". Verá que el nuevo campo aparece como una Dimensión:



9. Arrastre su nueva dimensión "Más caro" al botón "Color" del área de Marcas.
10. Ahora codificará el color de su gráfico de dispersión:



Los puntos naranjas representan los condados donde la leche es más cara, y los puntos rojos representan los condados en los que la soda es más cara. Los puntos azules representan condados donde esta cifra no estaba disponible, así que eliminémosla.

11. Haz clic en "Más caro" bajo el área de Marcas y selecciona "Filtrar..." Deselecciona "Nulo" y haga clic en "Aceptar". También haga clic en "Color" bajo el área de Marcas y seleccione la paleta "Daltónico" en el cuadro desplegable que dice "Seleccionar paleta de colores". Seleccione la paleta de colores y haga clic en "Asignar paleta" y luego Aceptar.

12. El gráfico se verá ahora un poco más limpio:



13. De este gráfico aprendemos algunas cosas interesantes:


- En la mayoría de los países, la leche es más cara que los refrescos.
- Donde la leche es más cara, las tasas de obesidad y diabetes tienden a ser más altas.
- Los puntos azules tienden a estar en la parte superior derecha del gráfico.
- Donde los refrescos son más caros, las tasas de obesidad y diabetes tienden a ser más bajas.
- Los puntos naranjas tienden a estar en la parte inferior izquierda del gráfico.

14. Cambia el nombre de tu hoja de trabajo a "Impacto de los precios de las bebidas"

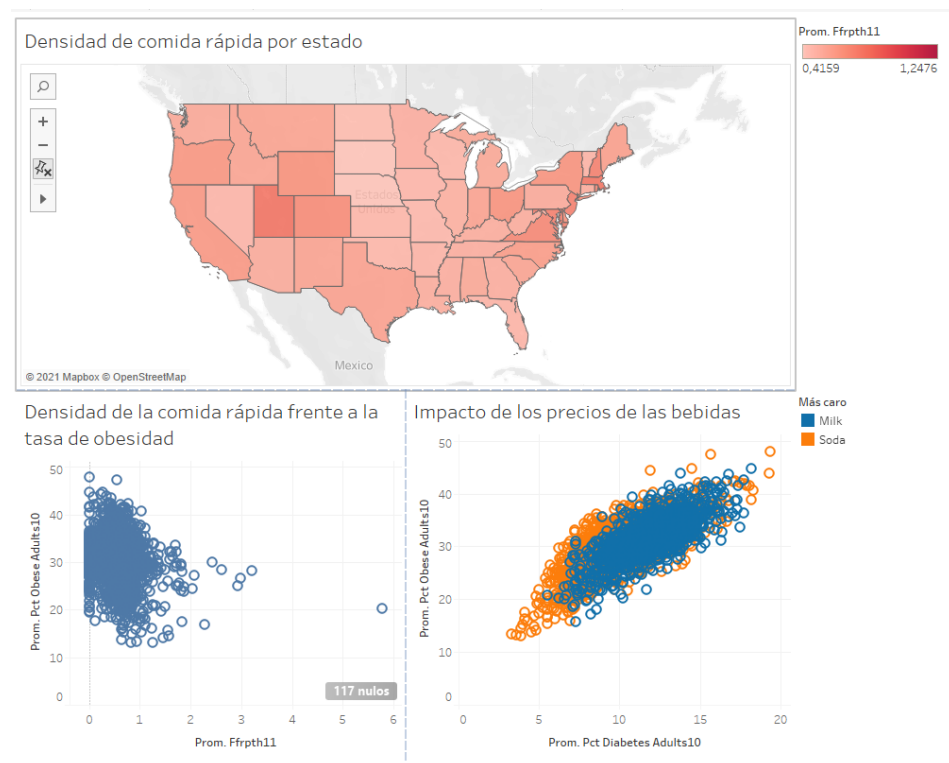
15. Guarda el libro de trabajo de Tableau.

1.8. Parte 6: Crear un cuadro de mando

Un cuadro de mando es una colección de hojas de trabajo, dispuestas de manera que sean fáciles de ver todas a la vez.

1. Haga clic en la  pestaña de la parte inferior de la pantalla para crear un nuevo cuadro de mando.
2. Arrastre la hoja de trabajo "Densidad de la comida rápida por estado" (en el cuadro de mandos) al espacio en blanco "Suelte las hojas aquí".
3. Ahora arrastre la hoja de trabajo "Densidad de la comida rápida frente a las tasas de obesidad" a la mitad inferior del espacio de trabajo (el mapa). espacio de trabajo (el mapa). Cuando esté en el lugar correcto, verá la mitad inferior del mapa sombreada en gris claro.

4. Ahora arrastre la hoja de trabajo "Impacto de los precios de las bebidas" a la esquina inferior derecha del espacio de trabajo (el lado derecho del gráfico de dispersión). Tendrá el siguiente aspecto:



5. Cambia el nombre de la ficha "Dashboard 1" por el que desees y guarda el libro de trabajo.

1.9. Parte 7: ¡Ahora inténtelo por su cuenta!

1. Cree una nueva hoja de trabajo de Tableau. Cree un mapa relleno que muestre las tasas de obesidad en adultos de 2010 por condado. Excluya Alaska (AK), Hawai (HI) y Puerto Rico (PR).

SUGERENCIA: Recuerde promediar los porcentajes, no sumarlos.

2. Cree una nueva hoja de trabajo de Tableau. Cree una tabla que compare el porcentaje de restaurantes de comida rápida por cada 1000 personas (**Ffrpth11**) con el porcentaje de restaurantes de servicio completo por cada 1000 personas (**Fsrpth11**). Limite sus resultados a los condados de New Jersey. La tabla debería tener el siguiente aspecto:

Restaurantes en los condados de Nueva Jersey

State	County	Prom. Ffrpth11	Prom. Fsrpth11
New Jersey	Atlantic	0,747	1,192
	Bergen	0,797	0,894
	Burlington	0,738	0,587
	Camden	0,729	0,563
	Cape May	1,863	2,919
	Cumberland	0,656	0,509
	Essex	0,648	0,629
	Gloucester	0,581	0,498
	Hudson	0,764	0,638
	Hunterdon	0,664	0,929
	Mercer	0,725	0,733
	Middlesex	0,728	0,637
	Monmouth	0,778	1,044
	Morris	0,784	0,871
	Ocean	0,699	0,689
	Passaic	0,681	0,629
	Salem	0,501	0,395
	Somerset	0,776	0,809
	Sussex	0,566	0,734
	Union	0,628	0,745
	Warren	0,729	0,665

SUGERENCIA: Arrastre sus dimensiones a Filas y sus medidas a columnas. Luego cambie el tipo de tabla a "tablas de texto".

3. Modifique la tabla que creó en el punto 2 para codificar por colores el texto de la tabla. Los valores de los condados donde hay más restaurantes de comida rápida que restaurantes de servicio completo que deben ser de un color, y los valores de los condados en los que hay más restaurantes de servicio completo que de comida rápida deben ser de otro color.

SUGERENCIA: Deberá crear un campo calculado para comparar **PROM(Ffrpth11)** con **PROM(Fsrpth11)**.

1.10. Entregable

Esta tarea debe realizarse de forma individual. Complete las instrucciones de este documento y envíe a Moodle:

- Un **PDF** explicando todos los pasos realizados para completar la Parte 7 del guion de prácticas.
- El **fichero de Tableau** (extensión twb) donde se han realizado todas las tareas del guion de prácticas.

1.11. Instrucciones de envío

Envíe el archivo PDF y el Fichero Tableau a través de Moodle **antes de la fecha límite**.

La entrega tardía está permitida, pero habrá una **penalización del 10% por cada 12 horas** hasta un máximo de 48 horas de demora.